Lectures on Symplectic Field Theory

Chris Wendl

Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

Email address: wendl@math.hu-berlin.de

Contents

Preface

This book is an expanded version of the lecture notes I produced for a two-semester course taught at University College London in 2015–16, for Ph.D. students with a background in basic symplectic geometry and interest in symplectic topology and/or geometric analysis. For the most part, each chapter corresponds to a two-hour lecture in the original course, though the reader will quickly notice that in this "expanded" version, most individual chapters contain far more material than can reasonably fit into one lecture (or even two). In reality, much of that material was only sketched or mentioned in passing during lectures, and I ended up using the notes to discuss everything that I would like to have explained if I'd had unlimited time. This includes relatively detailed discussions of several important technical points (e.g. the definition of spectral flow, generic transversality in symplectizations, the punctured Riemann-Roch formula, finite energy and asymptotics with arbitrary stable Hamiltonian structures) which are either incompletely covered by the existing literature or, in my opinion, simply more difficult to learn from other sources than they should be. For topics that are, on the other hand, well covered elsewhere, I have usually not felt obliged to explain every detail, but have tried always to provide adequate references.

One of the interesting features of SFT is that its foundations are—at the time of this writing—not yet complete. When the original "propaganda paper" [**EGH00**] appeared in 2000, it was widely believed that the technical details would be filled in within a few years, and several papers introducing important applications of SFT to contact topology were written under this assumption. Since then, a certain realization has set in that the results in those papers cannot truly be regarded as "theorems" in the sense of mathematics, and it has become less socially acceptable to preface statements of results with caveats of the form, "this theorem is dependent on the foundations of SFT". At the same time, the need for a robust perturbation scheme to achieve transversality in SFT spawned the development of a whole new approach to infinite-dimensional differential geometry, the *polyfold* project [**Hof06**], which is intended for much more general applications but is not yet finished. Opinions vary among symplectic topologists as to how unsatisfied we should all be with this state of affairs, and what could be done about it—among other things, one could make an entire course out of the discussion of such issues, but I have not chosen to do that. My approach is instead to develop the *classical*[1] analysis of pseudoholomorphic curves in symplectizations and symplectic cobordisms, to explain how this would lead to a theory of algebraic contact invariants if transversality for multiple

[1]For the purposes of this discussion, the word "classical" may be defined as "not involving the words *polyfold*, *virtual* or *Kuranishi*".

covers were not an issue, and then to use the tools and insights gained from this discussion to prove *rigorous mathematical theorems* about contact manifolds. Typically, such theorems can be regarded informally as consequences of computations in a (not yet well-defined) theory called SFT, but in a rigorous sense, they are actually consequences of the methods used in those computations. Examples covered in these notes include distinguishing tight contact structures on the 3-torus that are homotopic but not isomorphic (Chapter 11), and the nonexistence of symplectic fillings or symplectic cobordisms between certain pairs of contact manifolds (Chapter 17). The choice of applications is of course biased somewhat toward my own research interests.

Prerequisites. The stated target audience for the original lecture course was "advanced Masters and Ph.D. students in differential geometry or related fields who are not afraid of analysis". More precisely, the notes assume some knowledge of the following topics:

- Differential geometry: manifolds and vector bundles, differential forms and Stokes' theorem, connections, basic familiarity with symplectic manifolds;
- Functional analysis: linear operators on Banach spaces, basics of Sobolev spaces, Fredholm operators;
- Differential topology: smooth mapping degree, intersection numbers, Sard's theorem;
- Algebraic topology: fundamental group, homology and cohomology of manifolds, Poincaré duality, first Chern class, homological intersection numbers.

The following topics are not considered formal prerequisites, but some knowledge of them is likely in any case to be helpful to the reader, who may want to have a good reference for them (as suggested below) within arm's reach:

- Contact manifolds (e.g. Geiges [**Gei08**]);
- Differential calculus on Banach spaces and Banach manifolds (e.g. these two books by Lang: [**Lan93**] and [**Lan99**]);
- Closed pseudoholomorphic curves (e.g. McDuff-Salamon [**MS12**] or my other book in preparation [**Wenb**]);
- Floer homology (e.g. Salamon [**Sal99**] or Audin-Damian [**AD14**]).

Acknowledgements. I would like to thank the students who have sat through various iterations of the course that gave rise to this book, notably Alexandru Cioba and Agustín Moreno for their assistance in editing the first several lectures, as well as Adrian Dawid, Milica Đukić, Shah Faisal, Solveig Hepp, Catalina Jurja, and Michael Rothgang for useful comments. My understanding of Taubes's approach to the Riemann-Roch formula (explained in Chapter 5) and its generalization to the punctured case emerged in part from discussions with Chris Gerig, and I am grateful also to Tim Perutz for helpful hints about Weitzenböck formulas, and Patrick Massot for patient discussions of singular integral operators and elliptic regularity. Thanks also to Michael Hutchings and Janko Latschev for helping me understand the combinatorial factors in Chapter 13, to Jo Nelson for helpful comments on coefficients and orbifold singularities, and to Sam Lisi and Barney Bramham for advice on the

Floer C_ϵ space. And also to Klaus Niederkrüger and Helmut Hofer for enlightening discussions on all manner of things.

About the current version

The version you see in front of you is being revised and updated regularly to accompany a Masters-level special topics course on symplectic field theory at the Humboldt-Universität zu Berlin in the 2026 summer semester.

I have tried to produce a manuscript that is relatively well polished, but I have not tried quite as diligently for that as I do with most of my research papers. As of the beginning of the 2026 summer semester, some of the later chapters that have been in planning for over a decade are not yet complete, and one or two additional chapters exist only as vague plans in my head, so if those chapters exist by the end of the semester, they are unlikely to be error-free. I apologize for any sloppiness that I may have failed so far to expunge. All comments and corrections are welcome,[2] and may be sent to `wendl@math.hu-berlin.de`. Updates on the publication of the book will be posted periodically on my website at

https://www.mathematik.hu-berlin.de/~wendl/publications.html#notes

Most recent update: **May 18, 2026**

[2]especially if those corrections are received before the book goes to press

CHAPTER 1

Introduction

Contents

Symplectic field theory is a general framework for defining invariants of contact manifolds and symplectic cobordisms between them via counts of "asymptotically cylindrical" pseudoholomorphic curves. In this first chapter, we'll summarize some of the historical background of the subject, and then sketch the basic algebraic formalism of SFT.

1.1. In the beginning, Gromov wrote a paper

Pseudoholomorphic curves first appeared in symplectic geometry in a 1985 paper of Gromov [Gro85]. The development was revolutionary for the field of symplectic topology, but it was not unprecedented: a few years before this, Donaldson had demonstrated the power of using elliptic PDEs in geometric contexts to define invariants of smooth 4-manifolds (see [DK90]). The PDE that Gromov used was a slight generalization of one that was already familiar from complex geometry.

Recall that if M is a smooth $2n$-dimensional manifold, an **almost complex structure** on M is a smooth linear bundle map $J : TM \to TM$ such that $J^2 = -\mathbb{1}$. This makes the tangent spaces of M into complex vector spaces and thus induces an orientation on M; the pair (M, J) is called an **almost complex manifold**. In this context, a **Riemann surface** is an almost complex manifold of real dimension 2 (hence complex dimension 1), and a **pseudoholomorphic curve** (also called J-**holomorphic**) is a smooth map

$$u : \Sigma \to M$$

satisfying the **nonlinear Cauchy-Riemann equation**

$$Tu \circ j = J \circ Tu, \tag{1.1}$$

where (Σ, j) is a Riemann surface and (M, J) is an almost complex manifold (of arbitrary dimension). The almost complex structure J is called **integrable** if M admits the structure of a complex manifold such that J is multiplication by i in holomorphic coordinate charts. By a basic theorem due to Gauss, every almost complex structure in real dimension two is integrable, hence one can always find local coordinates (s,t) on neighorhoods in Σ such that

$$j\partial_s = \partial_t, \qquad j\partial_t = -\partial_s.$$

In these coordinates, (1.1) takes the form

$$\partial_s u + J(u)\partial_t u = 0.$$

The fundamental insight of [Gro85] was that solutions to the equation (1.1) capture information about symplectic structures on M whenever they are related to J in the following way.

DEFINITION 1.1.1. Suppose (M, ω) is a symplectic manifold. An almost complex structure J on M is said to be **tamed** by ω if

$$\omega(X, JX) > 0 \quad \text{for all } X \in TM \text{ with } X \neq 0.$$

Additionally, J is **compatible** with ω if the pairing

$$g(X, Y) := \omega(X, JY)$$

defines a Riemannian metric on M.

EXERCISE 1.1.2. Show that an almost complex structure J is compatible with a symplectic form ω if and only if it is tame and ω is J-invariant.

We shall denote by $\mathcal{J}(M)$ the space of all smooth almost complex structures on M, with the C^∞_{loc}-topology, and if ω is a symplectic form on M, let

$$\mathcal{J}_\tau(M, \omega), \mathcal{J}(M, \omega) \subset \mathcal{J}(M)$$

denote the subsets consisting of almost complex structures that are tamed by or compatible with ω respectively. Notice that $\mathcal{J}_\tau(M, \omega)$ is an open subset of $\mathcal{J}(M)$, but $\mathcal{J}(M, \omega)$ is not. Proofs of the following may be found in [MS17, §2.5] or [Wenb, §2.2], among other places.

PROPOSITION 1.1.3. *On any symplectic manifold (M, ω), the spaces $\mathcal{J}_\tau(M, \omega)$ and $\mathcal{J}(M, \omega)$ are each nonempty and contractible.* □

Tameness implies that the **energy** of a J-holomorphic curve $u : \Sigma \to M$,

$$E(u) := \int_\Sigma u^*\omega,$$

is always nonnegative, and it is strictly positive unless u is constant. Notice moreover that if the domain Σ is closed, then $E(u)$ depends only on the cohomology class $[\omega] \in H^2_{\text{dR}}(M)$ and the homology class

$$[u] := u_*[\Sigma] \in H_2(M),$$

so in particular, any family of J-holomorphic curves in a fixed homology class satisfies a uniform energy bound. This basic observation is one of the key facts behind

Gromov's compactness theorem, which states that moduli spaces of closed curves in a fixed homology class are compact up to "nodal" degenerations.

The most famous application of pseudoholomorphic curves presented in [Gro85] is Gromov's *nonsqueezing theorem*, which was the first known example of an obstruction for embedding symplectic domains that is subtler than the obvious obstruction defined by volume. The technology introduced in [Gro85] also led directly to the development of the *Gromov-Witten invariants* (see [MS12, RT95, RT97]), which follow the same pattern as Donaldson's earlier smooth 4-manifold invariants: they use counts of J-holomorphic curves to define invariants of symplectic manifolds up to symplectic deformation equivalence.

Here is another sample application from [Gro85]. We denote by

$$A \cdot B \in \mathbb{Z}$$

the intersection number between two homology classes $A, B \in H_2(M)$ in a closed oriented 4-manifold M.

THEOREM 1.1.4. *Suppose (M, ω) is a closed and connected symplectic 4-manifold with the following properties:*

(i) (M, ω) does not contain any symplectic submanifold $S \subset M$ that is diffeomorphic to S^2 and satisfies $[S] \cdot [S] = -1$.

(ii) (M, ω) contains two symplectic submanifolds $S_1, S_2 \subset M$ which are both diffeomorphic to S^2, satisfy

$$[S_1] \cdot [S_1] = [S_2] \cdot [S_2] = 0,$$

and have exactly one intersection point with each other, which is transverse and positive.

Then (M, ω) is symplectomorphic to $(S^2 \times S^2, \sigma_1 \oplus \sigma_2)$, where for $i = 1, 2$, the σ_i are area forms on S^2 satisfying

$$\int_{S^2} \sigma_i = \langle [\omega], [S_i] \rangle.$$

SKETCH OF THE PROOF. Since S_1 and S_2 are both symplectic submanifolds, one can choose a compatible almost complex structure J on M for which both of them are the images of embedded J-holomorphic curves. One then considers the moduli spaces $\mathcal{M}_1(J)$ and $\mathcal{M}_2(J)$ of equivalence classes of J-holomorphic spheres homologous to S_1 and S_2 respectively, where any two such curves are considered equivalent if one is a reparametrization of the other (in the present setting this just means they have the same image). These spaces are both manifestly nonempty, and one can argue via Gromov's compactness theorem for J-holomorphic curves that both are compact. Moreover, an infinte-dimensional version of the implicit function theorem implies that both are smooth 2-dimensional manifolds, carrying canonical orientations, hence both are diffeomorphic to closed surfaces. Finally, one uses *positivity of intersections* to show that every curve in $\mathcal{M}_1(J)$ intersects every curve in $\mathcal{M}_2(J)$ exactly once, and this intersection is always transverse and positive; moreover, any two curves in the same space $\mathcal{M}_1(J)$ or $\mathcal{M}_2(J)$ are either identical or disjoint. It follows that both moduli spaces are diffeomorphic to S^2, and both consist of smooth families of J-holomorphic spheres that foliate M, hence defining a diffeomorphism

$$\mathcal{M}_1(J) \times \mathcal{M}_2(J) \to M$$

that sends (u_1, u_2) to the unique point in the intersection $\operatorname{im} u_1 \cap \operatorname{im} u_2$. This identifies M with $S^2 \times S^2$ such that each of the submanifolds $S^2 \times \{*\}$ and $\{*\} \times S^2$ are symplectic. The latter observation can be used to determine the symplectic form up to deformation, so that by the Moser stability theorem, ω is determined up to isotopy by its cohomology class $[\omega] \in H^2_{\mathrm{dR}}(S^2 \times S^2)$, which depends only on the evaluation of ω on $[S^2 \times \{*\}]$ and $[\{*\} \times S^2] \in H_2(S^2 \times S^2)$. □

For a detailed exposition of the above proof of Theorem 1.1.4, see [**Wen18**, Theorem E].

1.2. Hamiltonian Floer homology

Throughout the following, we write

$$S^1 := \mathbb{R}/\mathbb{Z},$$

so maps on S^1 are the same as 1-periodic maps on $\mathbb{R}$. One popular version of the *Arnol′d conjecture* on symplectic fixed points can be stated as follows. Suppose (M, ω) is a closed symplectic manifold and $H : S^1 \times M \to \mathbb{R}$ is a smooth function. Writing $H_t := H(t, \cdot) : M \to \mathbb{R}$, H determines a 1-periodic time-dependent Hamiltonian vector field X_t via the relation[1]

$$\omega(X_t, \cdot) = -dH_t. \tag{1.2}$$

CONJECTURE 1.2.1 (Arnol′d conjecture). *If all 1-periodic orbits of X_t are nondegenerate, then the number of these orbits is at least the sum of the Betti numbers of M.*

Here a 1-periodic orbit $\gamma : S^1 \to M$ of X_t is called **nondegenerate** if, denoting the flow of X_t by φ^t, the linearized time 1 flow

$$d\varphi^1(\gamma(0)) : T_{\gamma(0)}M \to T_{\gamma(0)}M$$

does not have 1 as an eigenvalue. This can be thought of as a Morse condition for an action functional on the loop space whose critical points are periodic orbits; like Morse critical points, nondegenerate periodic orbits occur in isolation. To simplify our lives, let's restrict attention to *contractible* orbits and also assume that (M, ω) is **symplectically aspherical**, which means

$$[\omega]|_{\pi_2(M)} = 0, \quad \text{i.e.} \quad \langle [\omega], [u] \rangle = 0 \text{ for all continuous maps } u : S^2 \to M.$$

Then if $C^\infty_{\mathrm{contr}}(S^1, M)$ denotes the space of all smoothly contractible smooth loops in M, the **symplectic action functional** can be defined by

$$\mathcal{A}_H : C^\infty_{\mathrm{contr}}(S^1, M) \to \mathbb{R} : \gamma \mapsto -\int_{\mathbb{D}} \bar{\gamma}^*\omega + \int_{S^1} H_t(\gamma(t))\, dt,$$

[1] Elsewhere in the literature, you will sometimes see (1.2) without the minus sign on the right hand side. If you want to know why I strongly believe that the minus sign belongs there, see [**Wena**], but to some extent this is just a personal opinion.

where $\bar{\gamma} : \mathbb{D} \to M$ is any smooth map on the closed unit disk $\mathbb{D} \subset \mathbb{C}$ satisfying

$$\bar{\gamma}(e^{2\pi i t}) = \gamma(t),$$

and the symplectic asphericity condition guarantees that $\mathcal{A}_H(\gamma)$ does not depend on the choice of $\bar{\gamma}$.

EXERCISE 1.2.2. The **first variation** of a functional such as $\mathcal{A}_H : C^\infty_{\mathrm{contr}}(S^1, M)$ at $\gamma \in C^\infty_{\mathrm{contr}}(S^1, M)$ is by definition the unique linear map $d\mathcal{A}_H(\gamma) : \Gamma(\gamma^*TM) \to \mathbb{R}$ such that for any smooth 1-parameter family $\{\gamma_s \in C^\infty_{\mathrm{contr}}(S^1, M)\}_{s\in(-\epsilon,\epsilon)}$ with $\gamma_0 = \gamma$ and $\partial_s \gamma_s|_{s=0} = \eta$, one has

$$\left.\frac{d}{ds}\mathcal{A}_H(\gamma_s)\right|_{s=0} = d\mathcal{A}_H(\gamma)\eta.$$

In other words, if we think of $C^\infty_{\mathrm{contr}}(S^1, M)$ as an infinite-dimensional manifold[2] with tangent spaces $T_\gamma C^\infty_{\mathrm{contr}}(S^1, M) = \Gamma(\gamma^*TM)$, then $d\mathcal{A}_H(\gamma)$ is simply the differential of $\mathcal{A}_H$ at γ. Prove the formula

$$d\mathcal{A}_H(\gamma)\eta = \int_{S^1} [\omega(\dot{\gamma}, \eta) + dH_t(\eta)]\, dt = \int_{S^1} \omega(\dot{\gamma} - X_t(\gamma), \eta)\, dt.$$

Using the nondegeneracy of ω, this shows that the critical points of $\mathcal{A}_H$ are precisely the contractible 1-periodic orbits of X_t.

A few years after Gromov's introduction of pseudoholomorphic curves, Floer proved the most important cases of the Arnol′d conjecture by developing a novel version of infinite-dimensional Morse theory for the functional $\mathcal{A}_H$. This approach mimicked the homological approach to Morse theory which has since been popularized in books such as [AD14, Sch93], but was apparently only known to experts at the time. In *Morse homology*, one considers a smooth Riemannian manifold (M, g) with a Morse function $f : M \to \mathbb{R}$, and defines a chain complex whose generators are the critical points of f, graded according to their Morse index. If we denote the generator corresponding to a given critical point $x \in \mathrm{Crit}(f)$ by $\langle x\rangle$, the boundary map on this complex is defined by

$$\partial\langle x\rangle = \sum_{\mathrm{Morse}(y)=\mathrm{Morse}(x)-1} \#\left(\mathcal{M}(x, y)/\mathbb{R}\right)\langle y\rangle,$$

where $\mathcal{M}(x, y)$ denotes the moduli space of negative gradient flow lines $u : \mathbb{R} \to M$, satisfying $\partial_s u = -\nabla f(u(s))$, $\lim_{s\to-\infty} u(s) = x$ and $\lim_{s\to+\infty} u(s) = y$. This space admits a natural $\mathbb{R}$-action by shifting the variable in the domain, and one can show that for generic choices of f and the metric g, $\mathcal{M}(x, y)/\mathbb{R}$ is a finite set whenever $\mathrm{Morse}(x) - \mathrm{Morse}(y) = 1$. The real magic, however, is contained in the following statement about the case $\mathrm{Morse}(x) - \mathrm{Morse}(y) = 2$:

[2]At this stage, it is best not to worry so much over exactly what kind of infinite-dimensional manifold $C^\infty_{\mathrm{contr}}(S^1, M)$ is. It is locally modeled on open subsets of the space of smooth sections $\Gamma(\gamma^*TM)$, which is a Fréchet space, but the notion of a "Fréchet manifold" is not so straightforward to define precisely, and one can easily define the term "first variation" without worrying about it.

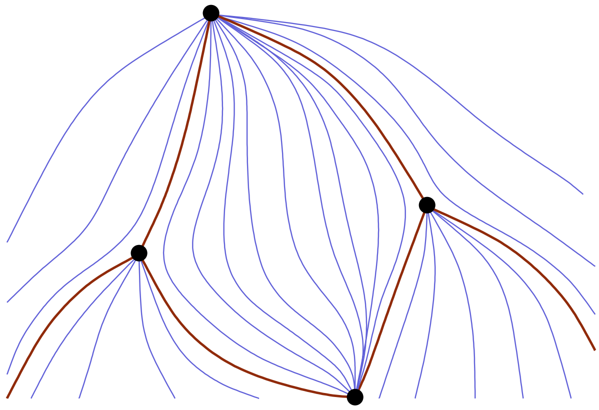

FIGURE 1.1. One-parameter families of gradient flow lines on a Riemannian manifold degenerate to broken flow lines.

PROPOSITION 1.2.3. *For generic choices of f and g and any two critical points $x, y \in \mathrm{Crit}(f)$ with* $\mathrm{Morse}(x) - \mathrm{Morse}(y) = 2$*, $\mathcal{M}(x,y)/\mathbb{R}$ is homeomorphic to a finite collection of circles and open intervals whose end points are canonically identified with the finite set*

$$\partial\overline{\mathcal{M}}(x,y) := \bigcup_{\mathrm{Morse}(z)=\mathrm{Morse}(x)-1} \mathcal{M}(x,z) \times \mathcal{M}(z,y).$$

We say that $\mathcal{M}(x,y)$ has a natural **compactification** $\overline{\mathcal{M}}(x,y)$, which has the topology of a compact 1-manifold with boundary, and its boundary is the set of all **broken flow lines** from x to y, cf. Figure 1.1. This set of broken flow lines is precisely what is counted if one computes the $\langle y \rangle$ coefficient of $\partial^2\langle x \rangle$, hence we deduce

$$\partial^2 = 0$$

as a consequence of the fact that compact 1-manifolds always have zero boundary points when counted with appropriate signs.[3] The homology of the resulting chain complex can be denoted by $HM_*(M\,;\,g,f)$ and is called the **Morse homology** of M. The well-known Morse inequalities can then be deduced from a fundamental theorem stating that $HM_*(M\,;\,g,f)$ is, for generic f and g, isomorphic to the singular homology of M.

With the above notion of Morse homology understood, Floer's approach to the Arnol′d conjecture can now be summarized as follows:

Step 1: Under suitable technical assumptions, construct a homology theory

$$HF_*(M,\omega\,;\,H,\{J_t\}),$$

depending *a priori* on the choices of a Hamiltonian $H : S^1 \times M \to \mathbb{R}$ with all 1-periodic orbits nondegenerate, and a generic S^1-parametrized family of ω-compatible almost complex structures $\{J_t\}_{t\in S^1}$. The generators of the

[3]Counting with signs presumes that we have chosen suitable orientations for the moduli spaces $\mathcal{M}(x,y)$, and this can always be done. Alternatively, one can avoid this issue by counting modulo 2, and thus define a homology theory with $\mathbb{Z}_2$ coefficients.

chain complex are the critical points of the symplectic action functional $\mathcal{A}_H$, i.e. 1-periodic orbits of the Hamiltonian flow, and the boundary map is defined by counting a suitable notion of gradient flow lines connecting pairs of orbits (more on this below).

Step 2: Prove that $HF_*(M, \omega) := HF_*(M, \omega\,;\, H, \{J_t\})$ is a *symplectic invariant*, i.e. it depends on ω, but not on the auxiliary choices H and $\{J_t\}$.

Step 3: Show that if H and $\{J_t\}$ are chosen to be time-independent and H is also C^2-small, then the chain complex for $HF_*(M, \omega\,;\, H, \{J_t\})$ is isomorphic (with a suitable grading shift) to the chain complex for Morse homology $HM_*(M\,;\, g, H)$ with $g := \omega(\cdot, J_t \cdot)$. The isomorphism between $HM_*(M\,;\, g, H)$ and singular homology thus implies that the Floer complex must have at least as many generators (i.e. periodic orbits) as there are generators of $H_*(M)$, proving the Arnol′d conjecture.

The implementation of Floer's idea required a different type of analysis than what is needed for Morse homology. The moduli space $\mathcal{M}(x, y)$ in Morse homology is simple to understand as the (generically transverse) intersection between the unstable manifold of x and the stable manifold of y with respect to the negative gradient flow. Conveniently, both of those are finite-dimensional manifolds, with their dimensions determined by the Morse indices of x and y. We will see in Chapter 3 that no such thing is true for the symplectic action functional: to the extent that $\mathcal{A}_H$ can be thought of as a Morse function on an infinite-dimensional manifold, its Morse index and its Morse "co-index" at every critical point are both infinite, hence the stable and unstable manifolds are not nearly as nice as finite-dimensional manifolds, providing no reason to expect that their intersection should be. There are additional problems since $C^\infty_{\text{contr}}(S^1, M)$ does not have a Banach space topology: in order to view the negative gradient flow of $\mathcal{A}_H$ as an ODE and make use of the usual local existence/uniqueness theorems (as in [**Lan99**, Chapter IV]), one would have to extend $\mathcal{A}_H$ to a smooth function on a suitable Hilbert manifold with a Riemannian metric. There is a very limited range of situations in which one can do this and obtain a reasonable formula for $\nabla \mathcal{A}_H$, e.g. [**HZ94**, §6.2] explains the case $M = \mathbb{T}^{2n}$, in which $\mathcal{A}_H$ can be defined on the Sobolev space $H^{1/2}(S^1, \mathbb{R}^{2n})$ and then studied using Fourier series. This approach is very dependent on the fact that the torus $\mathbb{T}^{2n}$ is a quotient of $\mathbb{R}^{2n}$. For general symplectic manifolds (M, ω), one cannot even define $H^{1/2}(S^1, M)$, since functions of class $H^{1/2}$ on S^1 need not be continuous ($H^{1/2}$ is a "Sobolev borderline case" in dimension one).

One of the novelties in Floer's approach was to refrain from viewing the gradient flow as an ODE in a Banach space setting, but instead to write down a formal version of the gradient flow equation and regard it as an elliptic PDE. To this end, let us regard $C^\infty_{\text{contr}}(S^1, M)$ formally as a manifold with tangent spaces

$$T_\gamma C^\infty_{\text{contr}}(S^1, M) := \Gamma(\gamma^* TM),$$

choose a formal Riemannian metric on this manifold (i.e. a smoothly varying family of L^2-inner products on the spaces $\Gamma(\gamma^* TM)$) and write down the resulting equation for the negative gradient flow. A suitable Riemannian metric can be defined by

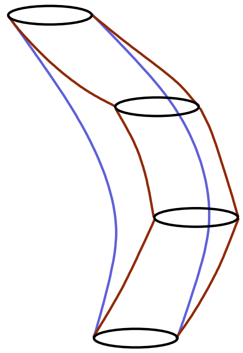

FIGURE 1.2. A family of smooth Floer trajectories can degenerate into a broken Floer trajectory.

choosing a smooth S^1-parametrized family of compatible almost complex structures

$$\{J_t \in \mathcal{J}(M,\omega)\}_{t\in S^1},$$

abbreviated in the following as $\{J_t\}$, and setting

$$\langle \xi, \eta \rangle_{L^2} := \int_{S^1} \omega(\xi(t), J_t\eta(t))\, dt$$

for $\xi, \eta \in \Gamma(\gamma^*TM)$. Exercise 1.2.2 then yields the formula

$$d\mathcal{A}_H(\gamma)\eta = \langle J_t(\dot{\gamma} - X_t(\gamma)), \eta \rangle_{L^2},$$

so that it seems reasonable to define the so-called *unregularized* gradient of $\mathcal{A}_H$ by

$$\nabla \mathcal{A}_H(\gamma) := J_t(\dot{\gamma} - X_t(\gamma)) \in \Gamma(\gamma^*TM). \tag{1.3}$$

Let us also think of a path $u : \mathbb{R} \to C^\infty_{\text{contr}}(S^1, M)$ as a map $u : \mathbb{R} \times S^1 \to M$, writing $u(s,t) := u(s)(t)$. The negative gradient flow equation $\partial_s u + \nabla\mathcal{A}_H(u(s)) = 0$ then becomes the elliptic PDE

$$\partial_s u + J_t(u)\,(\partial_t u - X_t(u)) = 0. \tag{1.4}$$

This is called the **Floer equation**, and its solutions are often called **Floer trajectories**. The relevance of Floer homology to our previous discussion of pseudoholomorphic curves should now be obvious. Indeed, the resemblance of the Floer equation to the nonlinear Cauchy-Riemann equation is not merely superficial—we will see in Chapter 6 that the former can always be viewed as a special case of the latter. In any case, one can use the same set of analytical techniques for both: elliptic regularity theory implies that Floer trajectories are always smooth, Fredholm theory and the implicit function theorem imply that (under appropriate assumptions) they form smooth finite-dimensional moduli spaces. Most importantly, the same "bubbling off" analysis that underlies Gromov's compactness theorem can be used to prove that spaces of Floer trajectories are compact up to "breaking", just as in Morse homology (see Figure 1.2)—this is the main reason for the relation $\partial^2 = 0$ in Floer homology.

We should mention one complication that does not arise either in the study of closed holomorphic curves or in finite-dimensional Morse theory. Since the gradient flow in Morse homology takes place on a closed manifold, it is obvious that every gradient flow line asymptotically approaches critical points at both $-\infty$ and $+\infty$. The following example shows that in the infinite-dimensional setting of Floer theory, this is no longer true.

EXAMPLE 1.2.4. Consider the Floer equation on $M := S^2 = \mathbb{C} \cup \{\infty\}$ with $H := 0$ and J_t defined as the standard complex structure i for every t. Then the orbits of X_t are all constant, and a map $u : \mathbb{R} \times S^1 \to S^2$ satisfies the Floer equation if and only if it is holomorphic. Identifying $\mathbb{R} \times S^1$ with $\mathbb{C}^* := \mathbb{C} \backslash \{0\}$ via the biholomorphic map $(s,t) \mapsto e^{2\pi(s+it)}$, a solution u approaches periodic orbits as $s \to \pm\infty$ if and only if the corresponding holomorphic map $\mathbb{C}^* \to S^2$ extends continuously (and therefore holomorphically) over 0 and ∞. But this is not true for every holomorphic map $\mathbb{C}^* \to S^2$, e.g. take any entire function $\mathbb{C} \to \mathbb{C}$ that has an essential singularity at ∞.

EXERCISE 1.2.5. Show that in the above example with an essential singularity at ∞, the symplectic action $\mathcal{A}_H(u(s,\cdot))$ is unbounded as $s \to \infty$.

EXERCISE 1.2.6. Suppose $u : \mathbb{R} \times S^1 \to M$ is a solution to the Floer equation with $\lim_{s\to\pm\infty} u(s,\cdot) = \gamma_\pm$ uniformly for a pair of 1-periodic orbits $\gamma_\pm \in \mathrm{Crit}(\mathcal{A}_H)$. Show that

(1.5)
$$\mathcal{A}(\gamma_-) - \mathcal{A}(\gamma_+) = \int_{\mathbb{R}\times S^1} \omega(\partial_s u, \partial_t u - X_t(u))\, ds\, dt = \int_{\mathbb{R}\times S^1} \omega(\partial_s u, J_t(u)\partial_s u)\, ds\, dt.$$

The right hand side of (1.5) is manifestly nonnegative since J_t is compatible with ω, and it is strictly positive unless $\gamma_- = \gamma_+$. It is therefore sensible to call this expression the **energy** $E(u)$ of a Floer trajectory. The following converse of Exercise 1.2.6 plays a crucial role in the compactness theory for Floer trajectories, as it guarantees that all the "levels" in a broken Floer trajectory are asymptotically well behaved. We will prove a variant of this result in the SFT context (see Prop. 1.3.12 below) in Chapter 7.

PROPOSITION 1.2.7. *If $u : \mathbb{R} \times S^1 \to M$ is a Floer trajectory with $E(u) < \infty$ and all 1-periodic orbits of X_t are nondegenerate, then there exist orbits $\gamma_-, \gamma_+ \in$ $\mathrm{Crit}(\mathcal{A}_H)$ such that $\lim_{s\to\pm\infty} u(s,\cdot) = \gamma_\pm$ uniformly.*

REMARK 1.2.8. It should be emphasized again that we have assumed $[\omega]|_{\pi_2(M)} = 0$ throughout this discussion. Floer homology can also be defined under more general assumptions, but several details become more complicated.

For nice comprehensive treatments of Hamiltonian Floer homology—unfortunately not always with the same sign conventions as used here—see [**Sal99**, **AD14**]. Note that this is only one of a few "Floer homologies" that were introduced by Floer in the late 80's: the others include *Lagrangian intersection Floer homology* [**Flo88a**] (which has since evolved into the *Fukaya category*, see [**Sei08**, **FOOO09**]), and *instanton homology* [**Flo88c**], an extension of Donaldson's gauge-theoretic smooth 4-manifold invariants to dimension three. The development of new Floer-type theories has since become a major industry; see [**AS19**] for a survey.

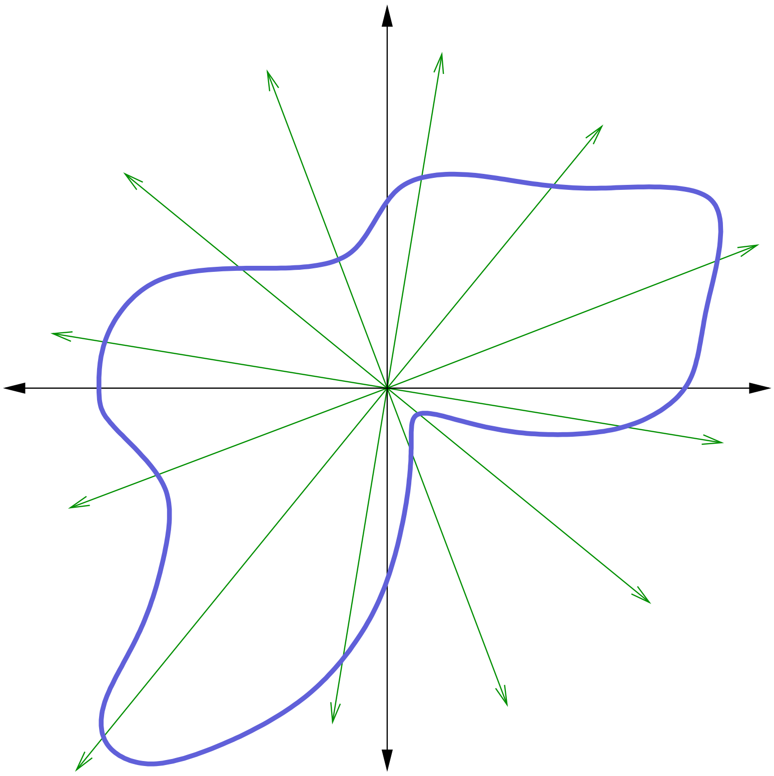

FIGURE 1.3. A star-shaped hypersurface in Euclidean space

1.3. Contact manifolds and the Weinstein conjecture

A Hamiltonian system on a symplectic manifold (W, ω) is called **autonomous** if the Hamiltonian $H : W \to \mathbb{R}$ does not depend on time. In this case, the Hamiltonian vector field X_H defined by

$$\omega(X_H, \cdot) = -dH$$

is time-independent and its orbits are confined to level sets of H. The images of these orbits on a given regular level set $H^{-1}(c)$ depend on the geometry of $H^{-1}(c)$, but not on H itself, as they are the integral curves (also known as **characteristics**) of the **characteristic line field** on $H^{-1}(c)$, defined as the unique direction spanned by a vector X such that $\omega(X, Y) = 0$ for all Y tangent to $H^{-1}(c)$. In 1978, Weinstein [**Wei78**] and Rabinowitz [**Rab78**] proved that certain kinds of regular level sets in symplectic manifolds are guaranteed to admit closed characteristics, hence implying the existence of periodic Hamiltonian orbits. In particular, this is true whenever $H^{-1}(c)$ is a *star-shaped* hypersurface in the standard symplectic $\mathbb{R}^{2n}$ (see Figure 1.3).

The following symplectic interpretation of the star-shaped condition provides both an intuitive reason to believe Rabinowitz's existence result and motivation for the more general conjecture of Weinstein. In any symplectic manifold (W, ω), a **Liouville vector field** is a smooth vector field V that satisfies

$$\mathcal{L}_V \omega = \omega.$$

By Cartan's formula for the Lie derivative, the 1-form λ defined by $\lambda := \omega(V, \cdot)$ satisfies $d\lambda = \omega$ if and only if V is a Liouville vector field; moreover, λ then also satisfies $\mathcal{L}_V \lambda = \lambda$, and it is referred to as a **Liouville form**. We sometimes say in this situation that the Liouville form λ and Liouville vector field V are ω-**dual** to each other. A hypersurface $M \subset (W, \omega)$ is said to be of **contact type** if it is transverse to a Liouville vector field defined on a neighborhood of M.

EXAMPLE 1.3.1. Using coordinates $(q_1, p_1, \ldots, q_n, p_n)$ on $\mathbb{R}^{2n}$, the standard symplectic form is written as

$$\omega_{\text{std}} := \sum_{j=1}^{n} dp_j \wedge dq_j,$$

and the Liouville form $\lambda_{\text{std}} := \frac{1}{2} \sum_{j=1}^{n} (p_j \, dq_j - q_j \, dp_j)$ is dual to the radial Liouville vector field

$$V_{\text{std}} := \frac{1}{2} \sum_{j=1}^{n} \left(p_j \frac{\partial}{\partial p_j} + q_j \frac{\partial}{\partial q_j} \right).$$

Any star-shaped hypersurface is therefore of contact type.

EXERCISE 1.3.2. Suppose (W, ω) is a symplectic manifold of dimension $2n$, $M \subset W$ is a smoothly embedded and oriented hypersurface, V is a Liouville vector field defined near M and $\lambda := \omega(V, \cdot)$ is the dual Liouville form. Define a 1-form on M by $\alpha := \lambda|_{TM}$.

(a) Show that V is positively transverse to M if and only if α satisfies

$$\alpha \wedge (d\alpha)^{n-1} > 0. \tag{1.6}$$

(b) If V is positively transverse to M, choose $\epsilon > 0$ sufficiently small and consider the embedding

$$\Phi : (-\epsilon, \epsilon) \times M \hookrightarrow W : (r, x) \mapsto \varphi_V^r(x),$$

where φ_V^t denotes the time t flow of V. Show that

$$\Phi^* \lambda = e^r \alpha,$$

where we are abusing notation on the right hand side by identifying $\alpha \in \Omega^1(M)$ with its pullback under the projection $(-\epsilon, \epsilon) \times M \to M$. In particular, we obtain from this the formula $\Phi^* \omega = d(e^r \alpha)$.

The above exercise presents any contact-type hypersurface $M \subset (W, \omega)$ as one member of a smooth 1-parameter family of contact-type hypersurfaces $M_r := \varphi_V^r(M) \subset W$, each canonically identified with M such that $\omega|_{TM_r} = e^r \, d\alpha$. In particular, the characteristic line fields on M_r are the same for all r, thus the existence of a closed characteristic on any of these implies that there also exists one on M. This observation has sometimes been used to prove such existence theorems, e.g. it is used in [HZ94, Chapter 4] to reduce Rabinowitz's result to an "almost existence" theorem based on symplectic capacities. This discussion hopefully makes the following conjecture seem believable.

CONJECTURE 1.3.3 (Weinstein conjecture, symplectic version). *Any closed contact-type hypersurface in a symplectic manifold admits a closed characteristic.*

Weinstein's conjecture admits a natural rephrasing in the language of contact geometry. A 1-form α on an oriented $(2n-1)$-dimensional manifold M is called a (positive) **contact form** if it satisfies (1.6), and the resulting co-oriented hyperplane field

$$\xi := \ker \alpha \subset TM$$

is then called a (positive and co-oriented) **contact structure**.[4] We call the pair (M, ξ) a **contact manifold**, and refer to a diffeomorphism $\varphi : M \to M'$ as a **contactomorphism** from (M, ξ) to (M', ξ') if φ_* maps ξ to ξ' and also preserves the respective co-orientations. Equivalently, if ξ and ξ' are defined via contact forms α and α' respectively, this means

$$\varphi^* \alpha' = f\alpha \quad \text{for some } f \in C^\infty(M, (0, \infty)).$$

Contact topology studies the category of contact manifolds (M, ξ) up to contactomorphism. The following basic result provides one good reason to regard ξ rather than α as the geometrically meaningful data, as the result holds for contact *structures*, but not for contact *forms*.

THEOREM 1.3.4 (Gray's stability theorem). *If M is a closed $(2n-1)$-dimensional manifold and $\{\xi_t\}_{t \in [0,1]}$ is a smooth 1-parameter family of contact structures on M, then there exists a smooth 1-parameter family of diffeomorphisms $\{\varphi_t\}_{t \in [0,1]}$ such that $\varphi_0 = \mathrm{Id}$ and $(\varphi_t)_* \xi_0 = \xi_t$.*

PROOF. See [Gei08, §2.2] or [Wenb, Theorem 1.6.12]. □

A corollary is that while the contact form α induced on a contact-type hypersurface $M \subset (W, \omega)$ via Exercise 1.3.2 is not unique, its induced contact structure is unique up to isotopy. Indeed, the space of all Liouville vector fields transverse to M is very large (e.g. one can add to V any sufficiently small Hamiltonian vector field), but it is *convex*, hence any two choices of the induced contact form α on M are connected by a smooth 1-parameter family of contact forms, implying an isotopy of contact structures via Gray's theorem.

EXERCISE 1.3.5. If α is a nowhere zero 1-form on M and $\xi = \ker \alpha$, show that α is contact if and only if $d\alpha|_\xi$ defines a symplectic vector bundle structure on $\xi \to M$. Moreover, the orientation of ξ determined by this symplectic bundle structure is compatible with the co-orientation determined by α and the orientation of M for which $\alpha \wedge (d\alpha)^{n-1} > 0$.

The following definition is based on the fact that since $d\alpha|_\xi$ is nondegenerate when α is contact, $\ker d\alpha \subset TM$ is always 1-dimensional and transverse to ξ.

DEFINITION 1.3.6. Given a contact form α on M, the **Reeb vector field** is the unique vector field R_α that satisfies

$$d\alpha(R_\alpha, \cdot) \equiv 0, \quad \text{and} \quad \alpha(R_\alpha) \equiv 1.$$

[4] The adjective "positive" refers to the fact that the orientation of M agrees with the one determined by the volume form $\alpha \wedge (d\alpha)^{n-1}$; we call α a *negative* contact form if these two orientations disagree. It is also possible in general to define contact structures without co-orientations, but contact structures of this type will never appear in this book; for our purposes, the co-orientation is *always* considered to be part of the data of a contact structure.

EXERCISE 1.3.7. Show that the flow of any Reeb vector field R_α preserves both $\xi = \ker \alpha$ and the symplectic vector bundle structure $d\alpha|_\xi$.

CONJECTURE 1.3.8 (Weinstein conjecture, contact version). *On any closed contact manifold (M, ξ) with contact form α, the Reeb vector field R_α admits a periodic orbit.*

To see that this is equivalent to the symplectic version of the conjecture, observe that any contact manifold $(M, \xi = \ker \alpha)$ can be viewed as the contact-type hypersurface $\{0\} \times M$ in the open symplectic manifold

$$(\mathbb{R} \times M, d(e^r \alpha)),$$

called the **symplectization** of (M, ξ). Here, as in Exercise 1.3.2, we are abusing notation and identifying $\alpha \in \Omega^1(M)$ with its pullback to $\mathbb{R} \times M$ via the projection $\mathbb{R} \times M \to M$.

EXERCISE 1.3.9. Recall that for any smooth manifold M, the cotangent bundle T^*M carries a tautological 1-form $\lambda_{\text{std}} \in \Omega^1(T^*M)$ that locally takes the form $\lambda_{\text{std}} = \sum_{j=1}^n p_j \, dq_j$ in any choice of local coordinates $(q_1, \ldots, q_n)$ on a neighbood $\mathcal{U} \subset M$, with $(p_1, \ldots, p_n)$ denoting the induced coordinates on the cotangent fibers over $\mathcal{U}$. (We will discuss cotangent bundles in somewhat more detail in §3.11.) This defines a Liouville form, with $d\lambda_{\text{std}}$ defining the canonical symplectic structure of T^*M. Now if $\xi \subset TM$ is a co-oriented hyperplane field on M, consider the submanifold

$$S_\xi M := \left\{ p \in T^*M \;\middle|\; \ker p = \xi \text{ and } p(X) > 0 \; \forall X \in TM \text{ pos. transverse to } \xi \right\}.$$

Show that ξ is contact if and only if $S_\xi M$ is a symplectic submanifold of $(T^*M, d\lambda_{\text{std}})$, and the Liouville vector field on T^*M dual to λ_{std} is tangent to $S_\xi M$. Moreover, if ξ is contact, then any choice of contact form for ξ determines a diffeomorphism of $S_\xi M$ to $\mathbb{R} \times M$ identifying the Liouville form λ_{std} along $S_\xi M$ with $e^r \alpha$.

REMARK 1.3.10. Exercise 1.3.9 shows that up to symplectomorphism, our definition of the symplectization of (M, ξ) above actually depends only on ξ and not on α.

In 1993, Hofer [**Hof93**] introduced a new approach to the Weinstein conjecture that was based in part on ideas of Gromov and Floer. Fix a contact manifold (M, ξ) with contact form α, and let

$$\mathcal{J}(\alpha) \subset \mathcal{J}(\mathbb{R} \times M)$$

denote the nonempty and contractible space of all almost complex structures J on $\mathbb{R} \times M$ satisfying the following conditions:

(1) The natural translation action on $\mathbb{R} \times M$ preserves J;
(2) $J\partial_r = R_\alpha$ and $JR_\alpha = -\partial_r$, where r denotes the canonical coordinate on the $\mathbb{R}$-factor in $\mathbb{R} \times M$;
(3) $J\xi = \xi$ and $d\alpha(\cdot, J\cdot)|_\xi$ defines a bundle metric on ξ.

It is easy to check that any $J \in \mathcal{J}(\alpha)$ is compatible with the symplectic structure $d(e^r \alpha)$ on $\mathbb{R} \times M$. Moreover, if $\gamma : \mathbb{R} \to M$ is any periodic orbit of R_α with period

$T > 0$, then for any $J \in \mathcal{J}(\alpha)$, the so-called **trivial cylinder**

$$u : \mathbb{R} \times S^1 \to \mathbb{R} \times M : (s,t) \mapsto (Ts, \gamma(Tt))$$

is a J-holomorphic curve. Following Floer, one version of Hofer's idea would be to look for J-holomorphic cylinders that satisfy a finite energy condition as in Prop. 1.2.7, forcing them to approach trivial cylinders asymptotically—the existence of such a cylinder would then imply the existence of a closed Reeb orbit, and thus prove the Weinstein conjecture. The first hindrance is that the "obvious" definition of energy in this context,

$$\int_{\mathbb{R}\times S^1} u^* d(e^r \alpha),$$

is not very useful: this integral is infinite if u is a trivial cylinder. To circumvent this, notice that every $J \in \mathcal{J}(\alpha)$ is also compatible with any symplectic structure of the form

$$\omega_\varphi := d(e^{\varphi(r)}\alpha),$$

where φ is a function chosen freely from the set

$$\mathcal{T} := \left\{ \varphi \in C^\infty(\mathbb{R}, (-1,1)) \;\middle|\; \varphi' > 0 \right\}. \tag{1.7}$$

Essentially, choosing ω_φ means identifying $\mathbb{R} \times M$ with a subset of the bounded region $(-1,1) \times M$, in which trivial cylinders have finite symplectic area. Since there is no preferred choice for the function φ, we define the **Hofer energy**[5] of a J-holomorphic curve $u : \Sigma \to \mathbb{R} \times M$ by

$$E(u) := \sup_{\varphi \in \mathcal{T}} \int_\Sigma u^* \omega_\varphi. \tag{1.8}$$

This has the desired property of being finite for trivial cylinders, and it is also nonnegative, with strict positivity whenever u is not constant.

Another useful observation from [Hof93] was that if the goal is to find periodic orbits, then we need not restrict our attention to J-holomorphic *cylinders* in particular. One can more generally consider curves defined on an arbitrary *punctured* Riemann surface

$$\dot{\Sigma} := \Sigma \backslash \Gamma,$$

where (Σ, j) is a closed connected Riemann surface and $\Gamma \subset \Sigma$ is a finite set of punctures. For any $\zeta \in \Gamma$, one can find coordinates identifying some punctured neighborhood of ζ biholomorphically with the closed punctured disk

$$\dot{\mathbb{D}} := \mathbb{D} \backslash \{0\} \subset \mathbb{C},$$

and then identify this with either the positive or negative half-cylinder

$$Z_+ := [0,\infty) \times S^1, \qquad Z_- := (-\infty, 0] \times S^1$$

via the biholomorphic maps

$$Z_+ \to \dot{\mathbb{D}} : (s,t) \mapsto e^{-2\pi(s+it)}, \qquad Z_- \to \dot{\mathbb{D}} : (s,t) \mapsto e^{2\pi(s+it)}.$$

[5]Strictly speaking, the energy defined in (1.8) is not identical to the notion introduced in [Hof93] and used in many of Hofer's papers, but it is equivalent to it in the sense that uniform bounds on either notion of energy imply uniform bounds on the other.

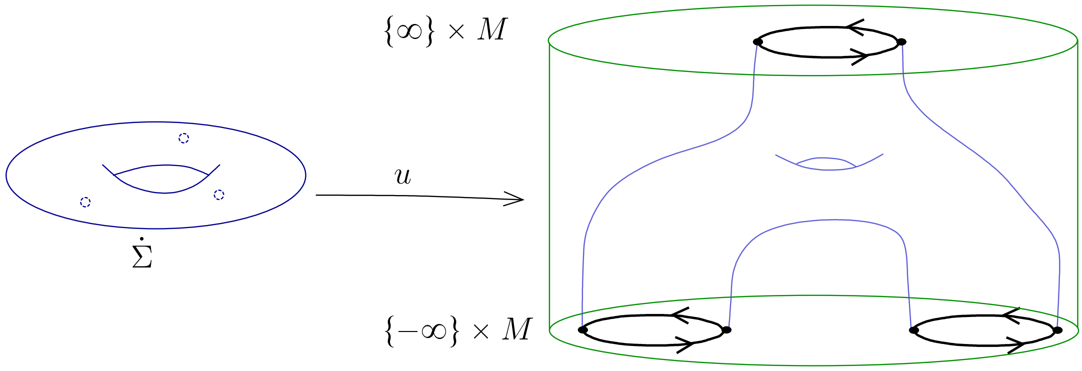

FIGURE 1.4. An asymptotically cylindrical holomorphic curve in a symplectization, with genus 1, one positive puncture and two negative punctures.

We will refer to such a choice as a (positive or negative) **holomorphic cylindrical coordinate** system near ζ, and in this way, we can present $(\dot{\Sigma}, j)$ as a *Riemann surface with cylindrical ends*, i.e. the union of some compact Riemann surface with boundary with a finite collection of half-cylinders $Z_\pm$ on which j takes the standard form $j\partial_s = \partial_t$. Note that the standard cylinder $\mathbb{R} \times S^1$ is a special case of this, as it can be identified biholomorphically with $S^2 \backslash \{0, \infty\}$. Another important special case is the plane, $\mathbb{C} = S^2 \backslash \{\infty\}$.

If $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is a J-holomorphic curve and $\zeta \in \Gamma$ is one of its punctures, we will say that u is **positively/negatively asymptotic** to a T-periodic Reeb orbit $\gamma : \mathbb{R} \to M$ at ζ if one can choose holomorphic cylindrical coordinates $(s, t) \in Z_\pm$ near ζ such that

$$u(s, t) = \exp_{(Ts, \gamma(Tt))} h(s, t) \qquad \text{for } |s| \text{ sufficiently large,}$$

where $h(s, t)$ is a vector field along the trivial cylinder satisfying $h(s, \cdot) \to 0$ uniformly as $|s| \to \infty$, and the exponential map is defined with respect to any $\mathbb{R}$-invariant choice of Riemannian metric on $\mathbb{R} \times M$. We say that $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is **asymptotically cylindrical** if it is (positively or negatively) asymptotic to some closed Reeb orbit at each of its punctures. Note that this partitions the finite set of punctures $\Gamma \subset \Sigma$ into two subsets,

$$\Gamma = \Gamma^+ \cup \Gamma^-,$$

the *positive* and *negative* punctures respectively, see Figure 1.4.

EXERCISE 1.3.11. Suppose $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is an asymptotically cylindrical J-holomorphic curve, with the asymptotic orbit at each puncture $\zeta \in \Gamma^\pm$ denoted by γ_ζ, having period $T_\zeta > 0$. Show that

$$\sum_{\zeta \in \Gamma^+} T_\zeta - \sum_{\zeta \in \Gamma^-} T_\zeta = \int_{\dot{\Sigma}} u^* d\alpha \geqslant 0,$$

with equality if and only if the image of u is contained in that of a trivial cylinder. In particular, u must have at least one positive puncture unless it is constant. Show also that $E(u)$ is finite and satisfies an upper bound determined only by the periods of the positive asymptotic orbits.

The following analogue of Prop. 1.2.7 will be proved in Chapter 7. For simplicity, we shall state a weakened version of what Hofer proved in [**Hof93**], which did not require any nondegeneracy assumption. A T-periodic Reeb orbit $\gamma : \mathbb{R} \to M$ is called **nondegenerate** if the Reeb flow φ_α^t has the property that its linearization along the contact bundle (cf. Exercise 1.3.7),

$$d\varphi_\alpha^T(\gamma(0))|_{\xi_{\gamma(0)}} : \xi_{\gamma(0)} \to \xi_{\gamma(0)}$$

does not have 1 as an eigenvalue. Note that since R_α is not time-dependent, closed Reeb orbits are never completely isolated—they always exist in S^1-parametrized families—but these families are isolated in the nondegenerate case. A **nondegenerate contact form** is one for which every closed Reeb orbit is nondegenerate. One can show that this condition is generic, meaning for instance that on any closed manifold, the nondegenerate contact forms constitute a C^∞-dense subset of the space of all contact forms (see Remark 1.3.13 below). The following result is the contact analogue of Proposition 1.2.7.

PROPOSITION 1.3.12. *Suppose (M, ξ) is a closed contact manifold with a nondegenerate contact form α. If $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is a J-holomorphic curve with $E(u) < \infty$ on a punctured Riemann surface such that none of the punctures are removable, then u is asymptotically cylindrical.*

The main results in [**Hof93**] state that under certain assumptions on a closed contact 3-manifold (M, ξ), namely if either ξ is *overtwisted* (as defined in [**Eli89**]) or $\pi_2(M) \neq 0$, one can find for any contact form α on (M, ξ) and any $J \in \mathcal{J}(\alpha)$ a finite-energy J-holomorphic plane. By Proposition 1.3.12, this implies the existence of a contractible periodic Reeb orbit and thus proves the Weinstein conjecture in these settings.

REMARK 1.3.13. The standard genericity result mentioned above for nondegenerate contact forms can be proved in various ways, e.g. it follows from a slightly more general result about generic regular level sets in Hamiltonian systems proved in [**Rob70**]. A more direct proof via the Sard-Smale theorem that is similar in spirit to the transversality arguments in Chapter 9 may be found in the appendix of [**ABW10**].

1.4. Symplectic cobordisms and their completions

After the developments described in the previous three sections, it seemed natural that one might define invariants of contact manifolds via a Floer-type theory generated by closed Reeb orbits and counting asymptotically cylindrical holomorphic curves in symplectizations. This theory is what is now called SFT, and its basic structure was outlined in a paper by Eliashberg, Givental and Hofer [**EGH00**] in 2000, though some of its analytical foundations remain unfinished as of 2026. The term "field theory" is an allusion to "topological quantum field theories," which associate vector spaces to certain geometric objects and morphisms to cobordisms between those objects. Thus in order to place SFT in its proper setting, we need to introduce symplectic cobordisms between contact manifolds.

Recall that if M_+ and M_- are smooth oriented closed manifolds of the same dimension, an oriented cobordism from M_- to M_+ is a compact smooth oriented manifold W with oriented boundary

$$\partial W \cong -M_- \coprod M_+,$$

where the symbol "$\cong$" in this setting means orientation-preserving diffeomorphism, and $-M_-$ denotes M_- with its orientation reversed. Given positive contact structures $\xi_\pm$ on $M_\pm$, we say that a symplectic manifold (W, ω) is a **symplectic cobordism from** (M_-, ξ_-) **to** (M_+, ξ_+) if W is an oriented cobordism[6] from M_- to M_+ such that both components of ∂W are contact-type hypersurfaces with induced contact structures isotopic to $\xi_\pm$. Note that our chosen orientation conventions imply that the Liouville vector field chosen near ∂W must point *outward* at M_+ and *inward* at M_-; we say in this case that M_+ is a symplectically **convex** boundary component, while M_- is symplectically **concave**. As important special cases, (W, ω) is a **symplectic filling** of (M_+, ξ_+) if $M_- = \varnothing$, and it is a **symplectic cap** of (M_-, ξ_-) if $M_+ = \varnothing$. In the literature, fillings and caps are sometimes also referred to as *convex fillings* or *concave fillings* respectively.

The contact-type condition implies the existence of a Liouville form λ near ∂W with $d\lambda = \omega$, such that by Exercise 1.3.2, neighborhoods of M_+ and M_- in W can be identified with the collars (see Figure 1.5)

$$(-\epsilon, 0] \times M_+ \quad \text{or} \quad [0, \epsilon) \times M_-$$

respectively for sufficiently small $\epsilon > 0$, with λ taking the form

$$\lambda = e^r \alpha_\pm,$$

where $\alpha_\pm := \lambda|_{TM_\pm}$ are contact forms for $\xi_\pm$, and r as usual denotes the canonical coordinate on the first factor in $\mathbb{R} \times M$. The **symplectic completion** of (W, ω) is the noncompact symplectic manifold $(\widehat{W}, \hat{\omega})$ defined by attaching cylindrical ends to these collar neighborhoods (Figure 1.6):

$$(\widehat{W}, \hat{\omega}) = ((-\infty, 0] \times M_-, d(e^r \alpha_-)) \cup_{M_-} (W, \omega) \cup_{M_+} ([0, \infty) \times M_+, d(e^r \alpha_+)) . \tag{1.9}$$

In this context, the symplectization $(\mathbb{R} \times M, d(e^r \alpha))$ is symplectomorphic to the completion of the **trivial symplectic cobordism** $([0, 1] \times M, d(e^r \alpha))$ from $(M, \xi = \ker \alpha)$ to itself. More generally, the object in the following easy exercise can also sensibly be called a trivial symplectic cobordism:

EXERCISE 1.4.1. Suppose (M, ξ) is a closed contact manifold with contact form α, and $f_\pm : M \to \mathbb{R}$ is a pair of functions with $f_- < f_+$ everywhere. Show that the domain

$$\left\{ (r, x) \in \mathbb{R} \times M \;\middle|\; f_-(x) \leqslant r \leqslant f_+(x) \right\} \subset \mathbb{R} \times M$$

[6] We assume of course that W is assigned the orientation determined by its symplectic form.

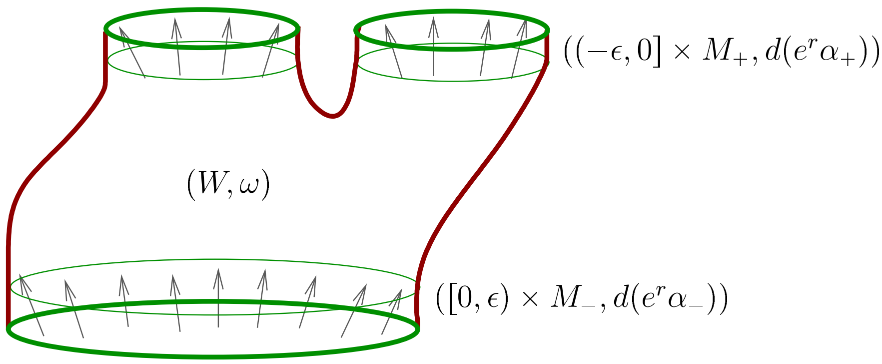

FIGURE 1.5. A symplectic cobordism with concave boundary (M_-, ξ_-) and convex boundary (M_+, ξ_+), with symplectic collar neighborhoods defined by flowing along Liouville vector fields near the boundary.

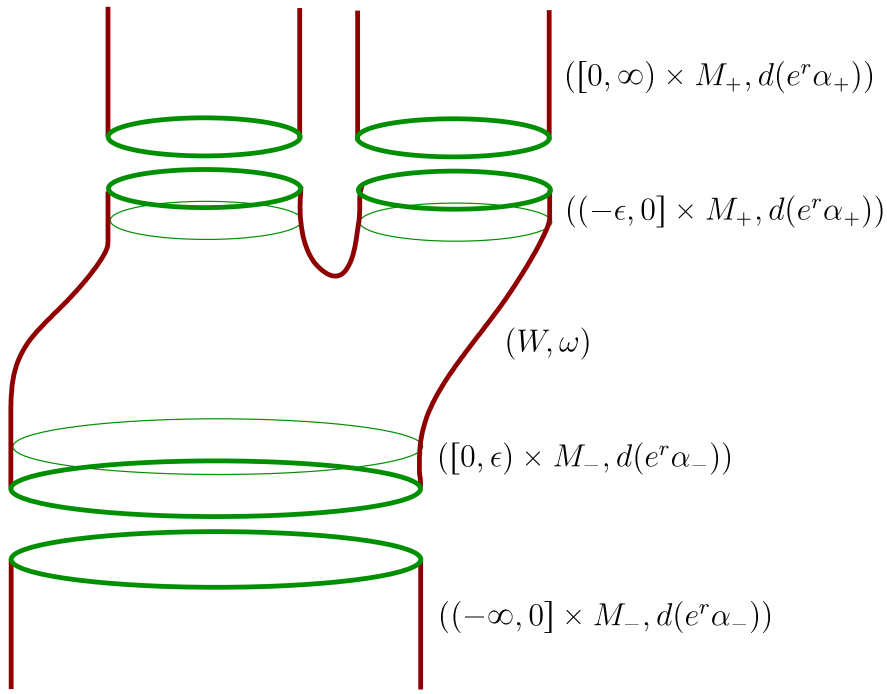

FIGURE 1.6. The completion of a symplectic cobordism

defines a symplectic cobordism from (M, ξ) to itself, with a global Liouville form $\lambda = e^r\alpha$ inducing contact forms $e^{f_-}\alpha$ and $e^{f_+}\alpha$ on its concave and convex boundaries respectively.

We say that (W, ω) is an **exact symplectic cobordism** or **Liouville cobordism** if the Liouville form λ can be extended from a neighborhood of ∂W to define a global primitive of ω on W. Equivalently, this means that ω admits a global Liouville vector field that points inward at M_- and outward at M_+. An **exact filling** of (M_+, ξ_+) is an exact cobordism whose concave boundary is empty. Observe that if (W, ω) is exact, then its completion $(\widehat{W}, \hat{\omega})$ also inherits a global Liouville form.

EXERCISE 1.4.2. Use Stokes' theorem to show that there is no such thing as an exact symplectic cap.

The above exercise hints at an important difference between cobordisms in the *symplectic* as opposed to the *oriented smooth* category: symplectic cobordisms are not generally reversible. If W is an oriented cobordism from M_- to M_+, then reversing the orientation of W produces an oriented cobordism from M_+ to M_-. But one cannot simply reverse orientations in the symplectic category, since the orientation is determined by the symplectic form. For example, many obstructions to the existence of symplectic fillings of given contact manifolds are known—some of them defined in terms of SFT—but there are no obstructions at all to symplectic caps, in fact it is known that all closed contact manifolds admit them (see [**EH02**, **CE20**, **Laz20**]).

The definitions for holomorphic curves in symplectizations in the previous section generalize to completions of symplectic cobordisms in a fairly straightforward way, since these completions look exactly like symplectizations outside of a compact subset. Define

$$\mathcal{J}(W, \omega, \alpha_+, \alpha_-) \subset \mathcal{J}(\widehat{W})$$

as the space of all almost complex structures J on $\widehat{W}$ such that

$$J|_W \in \mathcal{J}(W, \omega), \qquad J|_{[0,\infty)\times M_+} \in \mathcal{J}(\alpha_+) \quad \text{and} \quad J|_{(-\infty,0]\times M_-} \in \mathcal{J}(\alpha_-).$$

Occasionally it is useful to relax the compatibility condition on W to tameness,[7] i.e. $J|_W \in \mathcal{J}_\tau(W, \omega)$, producing a space that we shall denote by

$$\mathcal{J}_\tau(W, \omega, \alpha_+, \alpha_-) \subset \mathcal{J}(\widehat{W}).$$

As in Prop. 1.1.3, both of these spaces are nonempty and contractible. We can then consider asymptotically cylindrical J-holomorphic curves

$$u : (\dot{\Sigma} = \Sigma \backslash (\Gamma^+ \cup \Gamma^-), j) \to (\widehat{W}, J),$$

which are proper maps asymptotic to closed orbits of $R_{\alpha_\pm}$ in $M_\pm$ at punctures in $\Gamma^\pm$, see Figure 1.7.

One must again tinker with the symplectic form on $\widehat{W}$ in order to define a notion of energy that is finite when we need it to be. We generalize (1.7) as

$$\mathcal{T} := \left\{ \varphi \in C^\infty(\mathbb{R}, (-1,1)) \mid \varphi' > 0 \text{ and } \varphi(r) = r \text{ near } r = 0 \right\},$$

[7]It seems natural to wonder whether one could not also relax the conditions on the cylindrical ends and require $J|_{\xi_\pm}$ to be tamed by $d\alpha_\pm|_{\xi_\pm}$ instead of compatible with it. I do not currently know whether this works, but in later chapters we will see some reasons to worry that it might not (see §6.9.2).

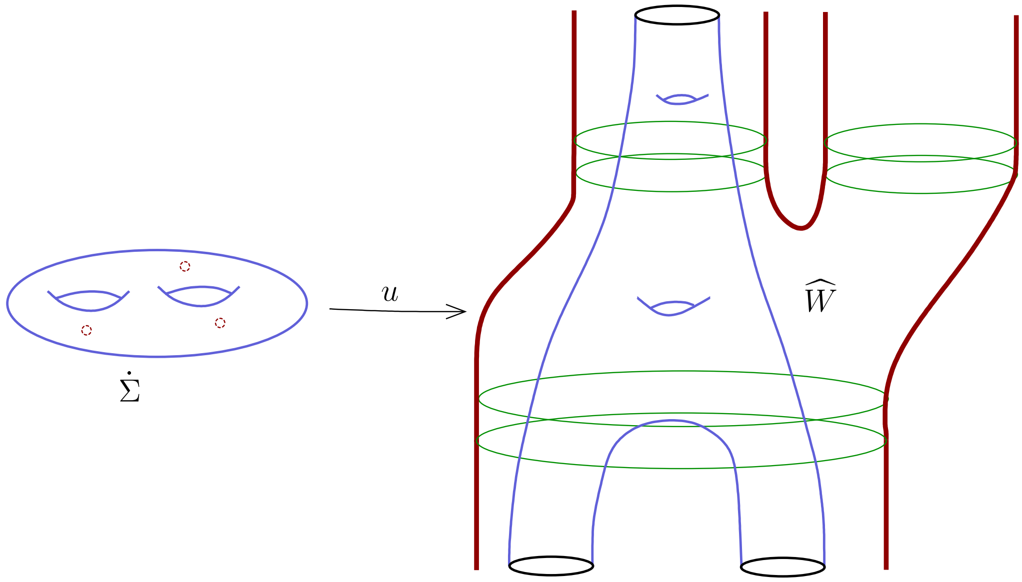

FIGURE 1.7. An asymptotically cylindrical holomorphic curve in a completed symplectic cobordism, with genus 2, one positive puncture and two negative punctures.

and associate to each $\varphi \in \mathcal{T}$ a symplectic form $\hat{\omega}_\varphi$ on $\widehat{W}$ defined by

$$\hat{\omega}_\varphi := \begin{cases} d(e^{\varphi(r)}\alpha_+) & \text{on } [0,\infty) \times M_+, \\ \omega & \text{on } W, \\ d(e^{\varphi(r)}\alpha_-) & \text{on } (-\infty, 0] \times M_-. \end{cases}$$

One can again check that every $J \in \mathcal{J}(W, \omega, \alpha_+, \alpha_-)$ or $\mathcal{J}_\tau(W, \omega, \alpha_+, \alpha_-)$ is tamed by $\hat{\omega}_\varphi$ for every $\varphi \in \mathcal{T}$. Thus it makes sense to define the **energy** of $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ by

$$E(u) := \sup_{\varphi \in \mathcal{T}} \int_{\dot{\Sigma}} u^* \hat{\omega}_\varphi.$$

It will be a straightforward matter to generalize Proposition 1.3.12 and show that finite energy implies asymptotically cylindrical behavior in completed cobordisms.

EXERCISE 1.4.3. Show that if (W, ω) is an exact cobordism, then every asymptotically cylindrical J-holomorphic curve in $\widehat{W}$ has at least one positive puncture.

1.5. Contact homology and SFT

We can now sketch the algebraic structure of SFT. We shall ignore or suppress several pesky details that are best dealt with later, some of them algebraic, others analytical. Due to analytical problems, some of the "theorems" that we shall (often imprecisely) state in this section are not yet provable at the current level of technology, though we expect that they will be in the foreseeable future. We shall use quotation marks to indicate this caveat wherever appropriate.

The standard versions of SFT all define homology theories with varying levels of algebraic structure which are meant to be invariants of a contact manifold (M, ξ). The chain complexes always depend on certain auxiliary choices, including a nondegenerate contact form α and a generic $J \in \mathcal{J}(\alpha)$. The generators consist of formal variables q_γ, one for each[8] closed Reeb orbit γ. In the most straightforward generalization of Hamiltonian Floer homology, the chain complex is simply a graded $\mathbb{Q}$-vector space generated by the variables q_γ, and the boundary map is defined by

$$\partial_{\mathrm{CCH}} q_\gamma = \sum_{\gamma'} \# \left(\mathcal{M}(\gamma, \gamma')/\mathbb{R}\right) q_{\gamma'},$$

where $\mathcal{M}(\gamma, \gamma')$ is the moduli space of J-holomorphic cylinders in $\mathbb{R} \times M$ with a positive puncture asymptotic to γ and a negative puncture asymptotic to γ', and the sum ranges over all orbits γ' for which this moduli space is 1-dimensional. The count $\# \left(\mathcal{M}(\gamma, \gamma')/\mathbb{R}\right)$ is rational, as it includes rational weighting factors that depend on combinatorial information and are best not discussed right now.[9]

"Theorem" 1.5.1. *If α admits no contractible Reeb orbits, then $\partial_{\mathrm{CCH}}^2 = 0$, and the resulting homology is independent of the choices of α with this property and generic $J \in \mathcal{J}(\alpha)$.*

The invariant arising from this result is known as **cylindrical contact homology**, and it is sometimes quite easy to work with when it is well defined, though it has the disadvantage of not always being defined. Namely, the relation $\partial_{\mathrm{CCH}}^2 = 0$ can fail if α admits contractible Reeb orbits, because unlike in Floer homology, the compactification of the space of cylinders $\mathcal{M}(\gamma, \gamma')$ generally includes objects that are not broken cylinders. In fact, the objects arising in the "SFT compactification" of moduli spaces of finite-energy curves in completed cobordisms can be quite elaborate, see Figure 1.8. The combinatorics of the situation are not so bad however if the cobordism is exact, as is the case for a symplectization: Exercise 1.4.3 then prevents curves without positive ends from appearing. The only possible degenerations for cylinders then consist of broken configurations whose levels each have *exactly one positive puncture* and arbitrary negative punctures; moreover, all but one of the negative punctures must eventually be capped off by planes, which is why "Theorem" 1.5.1 holds in the absence of planes.

If planes do exist, then one can account for them by defining the chain complex as an *algebra* rather than a vector space, producing the theory known as **contact homology**. For this, the chain complex is taken to be a graded unital algebra over $\mathbb{Q}$, and we define

$$\partial_{\mathrm{CH}} q_\gamma = \sum_{(\gamma_1, \dots, \gamma_m)} \# \left(\mathcal{M}(\gamma; \gamma_1, \dots, \gamma_m)/\mathbb{R}\right) q_{\gamma_1} \dots q_{\gamma_m},$$

with $\mathcal{M}(\gamma; \gamma_1, \dots, \gamma_m)$ denoting the moduli space of punctured J-holomorphic spheres in $\mathbb{R} \times M$ with a positive puncture at γ and m negative punctures at the orbits

[8] Actually, I should be making a distinction here between "good" and "bad" Reeb orbits, but let's discuss that later; see Chapter 12.

[9] Similar combinatorial factors are hidden behind the symbol "#" in our definitions of ∂_{CH} and $\mathbf{H}$, and will be discussed in earnest in Chapter 13.

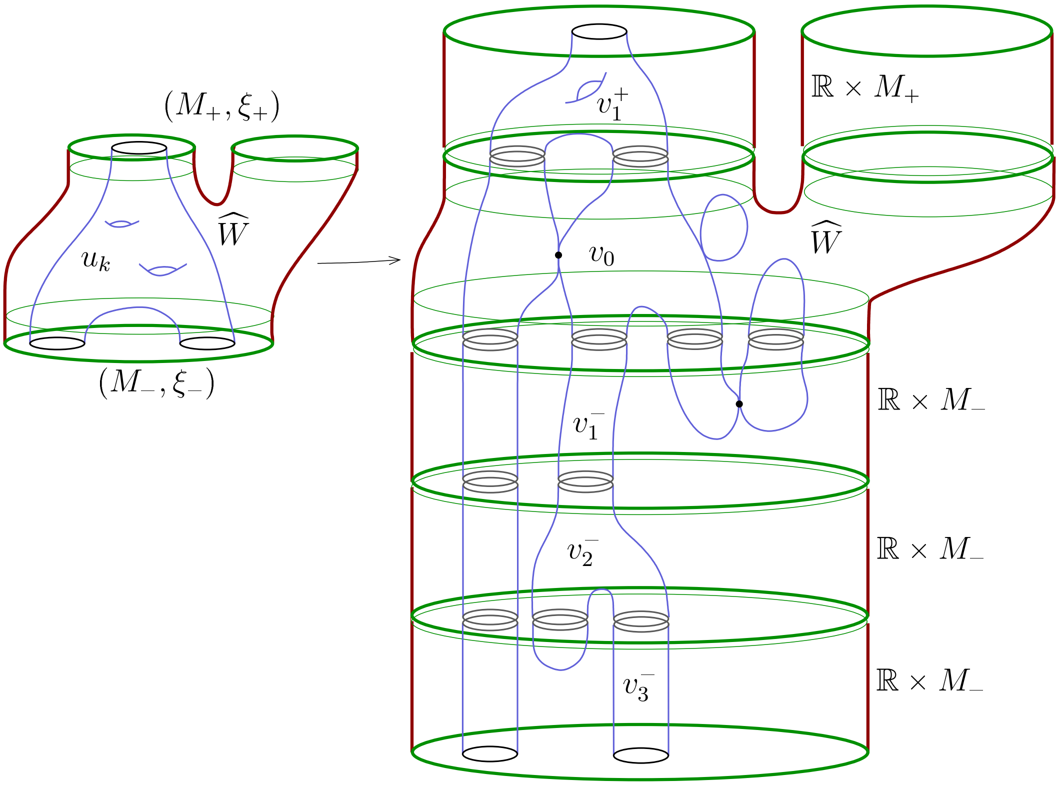

FIGURE 1.8. Degeneration of a sequence u_k of finite energy punctured holomorphic curves with genus 2, one positive puncture and two negative punctures in a symplectic cobordism. The limiting holomorphic building $(v_1^+, v_0, v_1^-, v_2^-, v_3^-)$ in this example has one upper level living in the symplectization $\mathbb{R} \times M_+$, a main level living in $\widehat{W}$, and three lower levels, each of which is a (possibly disconnected) finite-energy punctured nodal holomorphic curve in $\mathbb{R} \times M_-$. The building has arithmetic genus 2 and the same numbers of positive and negative punctures as u_k.

$\gamma_1, \ldots, \gamma_m$, and the sum ranges over all integers $m \geqslant 0$ and all m-tuples of orbits for which the moduli space is 1-dimensional. The action of ∂_{CH} is then extended to the whole algebra via a graded Leibniz rule

$$\partial_{\mathrm{CH}}(q_\gamma q_{\gamma'}) := (\partial_{\mathrm{CH}} q_\gamma)\, q_{\gamma'} + (-1)^{|\gamma|} q_\gamma\, (\partial_{\mathrm{CH}} q_{\gamma'})\,.$$

The general compactness and gluing theory for genus zero curves with one positive puncture now implies:

"THEOREM" 1.5.2. $\partial_{\mathrm{CH}}^2 = 0$, *and the resulting homology is (as a graded unital $\mathbb{Q}$-algebra) independent of the choices α and J.*

Maybe you've noticed the pattern: in order to accommodate more general classes of holomorphic curves, we need to add more algebraic structure. The **full SFT** algebra counts all rigid holomorphic curves in $\mathbb{R} \times M$, including all combinations of positive and negative punctures and all genera. Here is a brief picture of what it looks like. Counting all the 1-dimensional moduli spaces of J-holomorphic curves modulo $\mathbb{R}$-translation in $\mathbb{R} \times M$ produces a formal power series

$$\mathbf{H} := \sum \# \left(\mathcal{M}_g(\gamma_1^+, \dots, \gamma_{m_+}^+ ; \gamma_1^-, \dots, \gamma_{m_-}^-) \big/ \mathbb{R} \right) q_{\gamma_1^-} \dots q_{\gamma_{m_-}^-} p_{\gamma_1^+} \dots p_{\gamma_{m_+}^+} \hbar^{g-1},$$

where the sum ranges over all integers $g, m_+, m_- \geqslant 0$ and tuples of orbits, $\hbar$ and p_γ (one for each orbit γ) are additional formal variables, and

$$\mathcal{M}_g(\gamma_1^+, \dots, \gamma_{m_+}^+ ; \gamma_1^-, \dots, \gamma_{m_-}^-)$$

denotes the moduli space of J-holomorphic curves in $\mathbb{R} \times M$ with genus g, m_+ positive punctures at the orbits $\gamma_1^+, \dots, \gamma_{m_+}^+$, and m_- negative punctures at the orbits $\gamma_1^-, \dots, \gamma_{m_-}^+$. We can regard $\mathbf{H}$ as an operator on a graded algebra $\mathfrak{W}$ of formal power series in the variables $\{p_\gamma\}$, $\{q_\gamma\}$ and $\hbar$, equipped with a graded bracket operation that satisfies the quantum mechanical commutation relation

$$[p_\gamma, q_\gamma] = \kappa_\gamma \hbar,$$

where κ_γ is a combinatorial factor that is best ignored for now. Note that due to the signs that accompany the grading, odd elements $\mathbf{F} \in \mathfrak{W}$ need not satisfy $[\mathbf{F}, \mathbf{F}] = 0$, and $\mathbf{H}$ itself is an odd element, thus the following statement is nontrivial; in fact, it is the algebraic manifestation of the general compactness and gluing theory for punctured holomorphic curves in symplectizations.

"THEOREM" 1.5.3. $[\mathbf{H}, \mathbf{H}] = 0$, *hence by the graded Jacobi identity,* $\mathbf{H}$ *determines an operator*

$$D_{\mathrm{SFT}} : \mathfrak{W} \to \mathfrak{W} : \mathbf{F} \mapsto [\mathbf{H}, \mathbf{F}]$$

satisfying $D_{\mathrm{SFT}}^2 = 0$. *The resulting homology depends on* (M, ξ) *but not on the auxiliary choices* α *and* J.

It takes some time to understand how pictures such as Figure 1.8 translate into algebraic relations like $[\mathbf{H}, \mathbf{H}] = 0$, but this is a subject we'll come back to. There is also an intermediate theory between contact homology and full SFT, called **rational SFT**, which counts only genus zero curves with arbitrary positive and negative punctures. Algebraically, it is obtained from the full SFT algebra as a "semiclassical approximation" by discarding higher-order factors of $\hbar$, so that the commutation bracket in $\mathfrak{W}$ becomes a graded Poisson bracket. We will discuss all of this in Chapter 13.

1.6. Two applications

We briefly mention two applications that we will be able to establish rigorously using the methods developed in this book. Since SFT itself is not yet well defined in full generality, this sometimes means using SFT for *inspiration*, while proving corollaries via more direct methods.

1.6.1. Tight contact structures on $\mathbb{T}^3$. The 3-torus $\mathbb{T}^3 = S^1 \times S^1 \times S^1$ with coordinates (t, θ, ϕ) admits a sequence of contact structures

$$\xi_k := \ker \left(\cos(2\pi k t)\, d\theta + \sin(2\pi k t)\, d\phi \right),$$

one for each $k \in \mathbb{N}$. These cannot be distinguished from each other by any classical invariants, e.g. they all have the same Euler class, in fact they are all homotopic as co-oriented 2-plane fields. Nonetheless:

THEOREM 1.6.1. *For $k \neq \ell$, $(\mathbb{T}^3, \xi_k)$ and $(\mathbb{T}^3, \xi_\ell)$ are not contactomorphic.*

We will be able to prove this in Chapter 11 by rigorously defining and computing cylindrical contact homology for a suitable choice of contact forms on $(\mathbb{T}^3, \xi_k)$.

1.6.2. Filling and cobordism obstructions. Consider a closed connected and oriented surface Σ presented as $\Sigma_+ \cup_\Gamma \Sigma_-$, where $\Sigma_\pm \subset \Sigma$ are each (not necessarily connected) compact surfaces with a common boundary Γ. By an old result of Lutz [**Lut77**], the 3-manifold $S^1 \times \Sigma$ admits a unique isotopy class of S^1-invariant contact structures ξ_Γ such that the loops $S^1 \times \{z\}$ are positively/negatively transverse to ξ_Γ for $z \in \mathring{\Sigma}_\pm$ and tangent to ξ_Γ for $z \in \Gamma$. Now for each $k \in \mathbb{N}$, define

$$(V_k, \xi_k) := (S^1 \times \Sigma, \xi_\Gamma)$$

where $\Sigma = \Sigma_+ \cup_\Gamma \Sigma_-$ is chosen such that Γ has k connected components, Σ_- is connected with genus zero, and Σ_+ is connected with positive genus (see Figure 1.9).

THEOREM 1.6.2. *The contact manifolds (V_k, ξ_k) do not admit any symplectic fillings. Moreover, if $k > \ell$, then there exists no exact symplectic cobordism from (V_k, ξ_k) to (V_ℓ, ξ_ℓ).*

For these examples, one can use explicit constructions from [**Wen13**, **Avd21**] to show that non-exact cobordisms from (V_k, ξ_k) to (V_ℓ, ξ_ℓ) do exist, and so do exact cobordisms from (V_ℓ, ξ_ℓ) to (V_k, ξ_k), thus both the directionality of the cobordism relation and the distinction between exact and non-exact are crucial. The proof of the theorem, due to the author with Latschev and Hutchings [**LW11**], uses a numerical contact invariant based on the full SFT algebra—in particular, the curves that cause this phenomenon have multiple positive ends and are thus not seen by contact homology. We will introduce the relevant numerical invariant in Chapter 14 and compute it for these examples in Chapter 17.

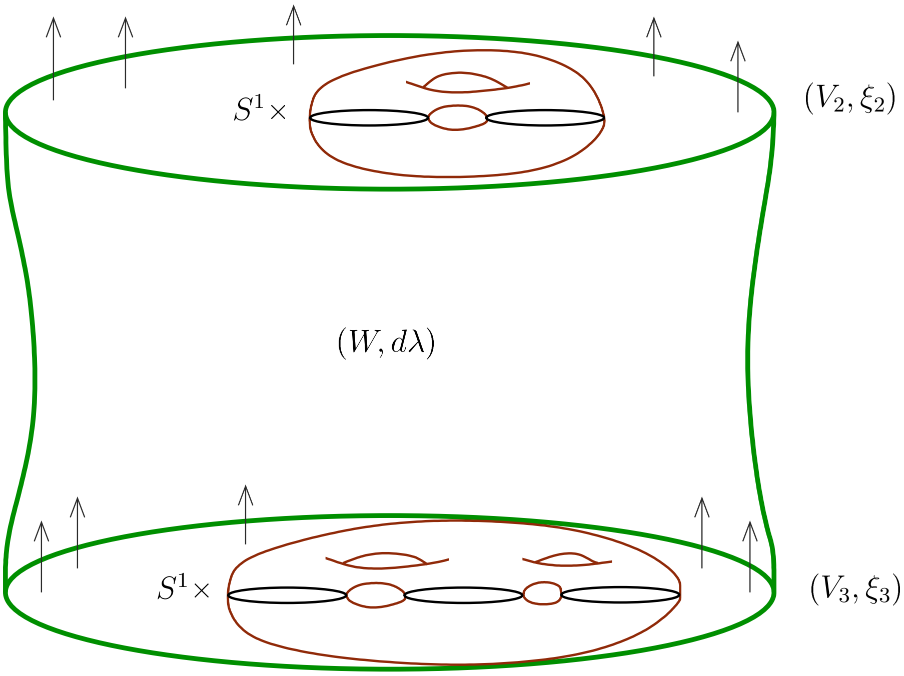

FIGURE 1.9. This exact symplectic cobordism does not exist.

CHAPTER 2

Basics on holomorphic curves

Contents

In this chapter we begin studying the analysis of J-holomorphic curves. The coverage will necessarily be a bit sparse in some places, but more detailed proofs of most things in this chapter can be found in [Wenb].

2.1. Linearized Cauchy-Riemann operators

In order to motivate the study of linear Cauchy-Riemann type operators, we begin with a formal discussion of the nonlinear Cauchy-Riemann equation and its linearization.

Fix a Riemann surface (Σ, j) and almost complex manifold (W, J). The nonlinear Cauchy-Riemann equation for maps $u : \Sigma \to W$ then takes the form

$$Tu \circ j = J \circ Tu,$$

which in any choice of local holomorphic coordinates (s, t) on suitably small neighborhoods in Σ is equivalent to

$$\partial_s u + J(u)\partial_t u = 0,$$

where we've explicitly written the dependence of $J : T_{u(z)}W \to T_{u(z)}W$ on $u(z)$ at each point $z \in \Sigma$ in order to emphasize the nonlinearity of the equation. The linearized equation at a given solution $u : \Sigma \to W$ is obtained by considering a smooth 1-parameter family of solutions $u_\rho : \Sigma \to W$ for $\rho \in (-\epsilon, \epsilon)$, with $u_0 = u$. Writing $\partial_\rho u_\rho|_{\rho=0} = \eta \in \Gamma(u^*TW)$, choosing a connection ∇ on W and taking the covariant derivative of the nonlinear equation with respect to the parameter gives

$$0 = \nabla_\rho \left[\partial_s u_\rho + J(u_\rho)\partial_t u_\rho\right]\big|_{\rho=0} = \nabla_\rho \partial_s u_\rho\big|_{\rho=0} + J(u)\nabla_\rho \partial_t u_\rho\big|_{\rho=0} + (\nabla_\eta J)\partial_t u.$$

Note that since $\partial_s u + J(u)\partial_t u = 0$, the expression on the right does not depend on the choice of connection. In particular, if we choose ∇ to be symmetric, then we can replace $\nabla_\rho \partial_s$ and $\nabla_\rho \partial_t$ with $\nabla_s \partial_\rho$ and $\nabla_t \partial_\rho$ respectively, so that the linearized equation takes the more appealing form

$$\nabla_s \eta + J(u)\nabla_t \eta + (\nabla_\eta J)\partial_t u = 0,$$

or in coordinate-free terms,

$$\nabla \eta + J(u)\nabla \eta \circ j + (\nabla_\eta J) \circ Tu \circ j = 0.$$

This is a globally well-defined linear first-order PDE for sections $\eta \in \Gamma(u^*TW)$. We will often abbreviate it in the form $\mathbf{D}_u \eta = 0$, defining the so-called **linearized Cauchy-Riemann operator at** u by

$$\begin{aligned} \mathbf{D}_u : \Gamma(u^*TW) &\to \Omega^{0,1}(\Sigma, u^*TW) \\ \eta &\mapsto \nabla \eta + J(u)\nabla \eta \circ j + (\nabla_\eta J) \circ Tu \circ j. \end{aligned} \tag{2.1}$$

Here we have used a bit of standard notation from complex geometry: $\Omega^{0,1}(\Sigma, u^*TW)$ denotes the space of u^*TW-valued $(0,1)$-forms on Σ, where "$(0,1)$" means 1-forms that are complex-*antilinear*.[1] Equivalently, elements of $\Omega^{0,1}(\Sigma, u^*TW)$ are smooth sections of $\overline{\mathrm{Hom}}_{\mathbb{C}}(T\Sigma, u^*TW) = T^{0,1}\Sigma \otimes_{\mathbb{C}} u^*TW$, where $T^{0,1}\Sigma$ denotes the $(0,1)$-part of the complexified cotangent bundle.[2]

The linearized Cauchy-Riemann operator arises in the following application. Suppose we wish to understand the structure of some space of the form

$$\{u : \Sigma \to W \mid Tu \circ j = J \circ Tu \text{ plus further conditions}\}, \tag{2.2}$$

where the "further conditions" (which we will for now leave unspecified) may impose constraints on e.g. the regularity of u, as well as its boundary and/or asymptotic behavior. The standard approach in global analysis can be summarized as follows:

Step 1: Construct a smooth Banach manifold $\mathcal{B}$ of maps $u : \Sigma \to W$ such that all the solutions we're interested in will be elements of $\mathcal{B}$. The tangent spaces $T_u\mathcal{B}$ are then Banach spaces of sections of u^*TW.

Step 2: Construct a smooth Banach space bundle $\mathcal{E} \to \mathcal{B}$ such that for each $u \in \mathcal{B}$, the fiber $\mathcal{E}_u$ is a Banach space of sections of the vector bundle

$$\overline{\mathrm{Hom}}_{\mathbb{C}}(T\Sigma, u^*TW) \to \Sigma$$

of complex-antilinear bundle maps $(T\Sigma, j) \to (u^*TW, J)$. Since our purpose is to study a first-order PDE, we need the sections in $\mathcal{E}_u$ to be "one step less regular" than the maps in $\mathcal{B}$, e.g. if $\mathcal{B}$ consists of maps of Sobolev class $W^{k,p}$, then the sections in $\mathcal{E}_u$ should be of class $W^{k-1,p}$.

[1] Complex-linear 1-forms are similarly called $(1,0)$-forms.

[2] In more straightforward terms, $T^{0,1}\Sigma \to \Sigma$ is a complex line bundle whose fiber at any given point $z \in \Sigma$ is the space of complex-antilinear maps $T_z\Sigma \to \mathbb{C}$. Similarly, fibers of $T^{1,0}\Sigma \to \Sigma$ are spaces of complex-linear maps $T_z\Sigma \to \mathbb{C}$. The direct sum of these two bundles is the complexification of $T^*\Sigma$, whose fiber at $z \in \Sigma$ consists of all *real*-linear maps $T_z\Sigma \to \mathbb{C}$.

Step 3: Show that

$$\bar{\partial}_J : \mathcal{B} \to \mathcal{E} : u \mapsto du + J(u) \circ du \circ j$$

defines a smooth section of $\mathcal{E} \to \mathcal{B}$, whose zero set is precisely the space of solutions (2.2).

Step 4: Show that under suitable assumptions (e.g. on regularity and asymptotic behavior), one can arrange such that for every $u \in \bar{\partial}_J^{-1}(0)$, the **linearization** of $\bar{\partial}_J$,[3]

$$D\bar{\partial}_J(u) : T_u\mathcal{B} \to \mathcal{E}_u$$

is a Fredholm operator and is generically surjective. (In geometric terms, this would mean that $\bar{\partial}_J$ is *transverse to the zero section.*)

Step 5: Using the implicit function theorem in Banach spaces (see [Lan93]), the surjectivity and Fredholm property of $D\bar{\partial}_J(u)$ imply that $\bar{\partial}_J^{-1}(0)$ is a smooth finite-dimensional manifold, with its tangent space at each $u \in \bar{\partial}_J^{-1}(0)$ canonically identified with $\ker D\bar{\partial}_J(u)$, hence the dimension of $\bar{\partial}_J^{-1}(0)$ near u equals the Fredholm index of $D\bar{\partial}_J(u)$.

In this context, the linearization of the section $\bar{\partial}_J$ at a point $u \in \bar{\partial}_J^{-1}(0)$ will be given by the natural extension of $\mathbf{D}_u : \Gamma(u^*TW) \to \Omega^{0,1}(\Sigma, u^*TW)$ to a suitable Banach space setting, e.g. if $\mathcal{B}$ consists of maps $\Sigma \to W$ of Sobolev class $W^{k,p}$, then $\mathbf{D}_u$ will be extended to a map from the $W^{k,p}$-sections of u^*TW to the $W^{k-1,p}$-sections of $\overline{\text{Hom}}_{\mathbb{C}}(T\Sigma, u^*TW)$.

DEFINITION 2.1.1. Fix a complex vector bundle E over a Riemann surface (Σ, j). A (real) linear **Cauchy-Riemann type operator** on E is a real-linear first-order differential operator

$$\mathbf{D} : \Gamma(E) \to \Omega^{0,1}(\Sigma, E)$$

such that for every $f \in C^\infty(\Sigma, \mathbb{R})$ and $\eta \in \Gamma(E)$,

$$\mathbf{D}(f\eta) = (\bar{\partial} f)\eta + f\mathbf{D}\eta, \tag{2.3}$$

where $\bar{\partial} f$ denotes the complex-valued $(0,1)$-form $df + i\, df \circ j$.

Observe that $\mathbf{D}$ is complex linear if and only if the Leibniz rule (2.3) also holds for all smooth complex-valued functions f, not just real-valued. It is a standard result in complex geometry that choosing a complex-linear Cauchy-Riemann type operator $\mathbf{D}$ on E is equivalent to endowing it with the structure of a *holomorphic* vector bundle, where local sections η are defined to be holomorphic if and only if $\mathbf{D}\eta = 0$. Indeed, every holomorphic bundle comes with a canonical Cauchy-Riemann operator that is expressed as $\bar{\partial}$ in holomorphic trivializations, and in the

[3]The **linearization** of a section $s : B \to E$ of a smooth vector bundle $E \to B$ at a point $x \in s^{-1}(0) \subset B$ is a linear map $Ds(x) : T_xB \to E_x$ that can be computed by choosing any connection ∇ on E and setting $Ds(x)v := \nabla_v s$. The result is independent of the choice of connection since $s(x) = 0$. Equivalently, one could choose a local chart and trivialization near x, compute the differential of the section at x in coordinates, and argue in the same way that the resulting map $T_xB \to E_x$ is independent of choices.

other direction, the equivalence follows from a local existence result for solutions to the equation $\mathbf{D}\eta = 0$, proved in §2.5 below.[4]

EXERCISE 2.1.2. If $\mathbf{D}$ is a linear Cauchy-Riemann type operator on E, show that for every smooth linear bundle map $A : E \to \overline{\mathrm{Hom}}_{\mathbb{C}}(T\Sigma, E)$, or equivalently every $A \in \Omega^{0,1}(\Sigma, \mathrm{End}_{\mathbb{R}}(E))$, $\mathbf{D} + A$ also defines a linear Cauchy-Riemann type operator on E, and every linear Cauchy-Riemann type operator on E is of this form. Using this, show that in suitable local trivializations over a subset $\mathcal{U} \subset \Sigma$ identified biholomorphically with an open set in $\mathbb{C}$, every Cauchy-Riemann type operator $\mathbf{D}$ takes the form

$$\mathbf{D} = \bar{\partial} + A : C^\infty(\mathcal{U}, \mathbb{C}^m) \to C^\infty(\mathcal{U}, \mathbb{C}^m),$$

where $\bar{\partial} = \partial_s + i\partial_t$ in complex coordinates $z = s + it$ and $A \in C^\infty(\mathcal{U}, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^m))$.

EXERCISE 2.1.3. Show that for any (not necessarily complex) connection ∇ on a complex vector bundle E over a Riemann surface Σ, $\mathbf{D}\eta := \nabla\eta + i\,\nabla\eta \circ j$ defines a linear Cauchy-Riemann type operator on E. Deduce from this that the operator $\mathbf{D}_u$ of (2.1) is a real-linear Cauchy-Riemann type operator on u^*TW.

EXERCISE 2.1.4. Suppose $\mathbf{D}$ is a linear Cauchy-Riemann type operator on a bundle E over a Riemann surface (Σ, j), and (Σ', j') is another Riemann surface with a holomorphic map $\varphi : (\Sigma', j') \to (\Sigma, j)$. Show that the pullback bundle φ^*E over (Σ', j') admits a unique linear Cauchy-Riemann type operator $\varphi^*\mathbf{D} : \Gamma(\varphi^*E) \to \Omega^{0,1}(\Sigma', \varphi^*E)$ satisfying the condition

$$(\varphi^*\mathbf{D})(\eta \circ \varphi) = \varphi^*(\mathbf{D}\eta) \qquad \text{for all } \eta \in \Gamma(E).$$

Prove moreover that if $\mathbf{D}$ is given by $\mathbf{D}\eta = \nabla\eta + i\,\nabla\eta \circ j + A\eta$ for some connection ∇ on E and $A \in \Omega^{0,1}(\Sigma, \mathrm{End}_{\mathbb{R}}(E))$, then $\varphi^*\mathbf{D}$ is given by

$$(\varphi^*\mathbf{D})\xi = \nabla\xi + i\,\nabla\xi \circ j + (\varphi^*A)\xi,$$

where ∇ in this case denotes the pullback connection on $\varphi^*E \to \Sigma'$ determined by our chosen connection on E. *Hint: Locally, you can write any section of φ^*E as a linear combination of sections of the form $\eta \circ \varphi$ for sections η of E.*

EXERCISE 2.1.5. Suppose E, F are two complex vector bundles over (Σ, j) and $\mathbf{D} : \Gamma(E \oplus F) \to \Omega^{0,1}(\Sigma, E \oplus F)$ is a linear Cauchy-Riemann type operator on $E \oplus F$, which can therefore be written in block form as

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}^E & \mathbf{D}^{EF} \\ \mathbf{D}^{FE} & \mathbf{D}^F \end{pmatrix} : \Gamma(E) \oplus \Gamma(F) \to \Omega^{0,1}(\Sigma, E) \oplus \Omega^{0,1}(\Sigma, F).$$

Show that $\mathbf{D}^E$ and $\mathbf{D}^F$ are then linear Cauchy-Riemann type operators on E and F respectively, while the off-diagonal operators $\mathbf{D}^{EF}$ and $\mathbf{D}^{FE}$ are tensorial, i.e. they can be expressed as smooth real-linear bundle maps $F \to \overline{\mathrm{Hom}}_{\mathbb{C}}(T\Sigma, E)$ and $E \to \overline{\mathrm{Hom}}_{\mathbb{C}}(T\Sigma, F)$ respectively.

[4] This statement about the existence of holomorphic vector bundle structures is true when the base is a Riemann surface, but not if it is a higher-dimensional complex manifold. In higher dimensions there are obstructions, see e.g. [Kob87].

EXERCISE 2.1.6. Suppose E is a complex vector bundle over (Σ, j), $F \subset E$ is a complex subbundle, and $\mathbf{D} : \Gamma(E) \to \Omega^{0,1}(\Sigma, E)$ is a linear Cauchy-Riemann type operator on E such that $\mathbf{D}(\Gamma(F)) \subset \Omega^{0,1}(\Sigma, F)$, so $\mathbf{D}$ restricts to a linear Cauchy-Riemann type operator on F as well. Show that the induced map

$$\Gamma(E)/\Gamma(F) \xrightarrow{\mathbf{D}} \Omega^{0,1}(\Sigma, E)/\Omega^{0,1}(\Sigma, F)$$

can then be interpreted as defining a linear Cauchy-Riemann type operator on the quotient bundle E/F, using the natural identifications

$$\Gamma(E/F) = \Gamma(E)/\Gamma(F) \qquad \text{and} \qquad \Omega^{0,1}(\Sigma, E/F) = \Omega^{0,1}(\Sigma, E)/\Omega^{0,1}(\Sigma, F).$$

2.2. Some useful Sobolev inequalities

In this section, we review a few general properties of Sobolev spaces that are essential for applications in nonlinear analysis. The results stated here are explained in more detail in Appendix A.

We consider functions with values in $\mathbb{C}$ unless otherwise specified, and defined on an open domain $\mathcal{U}$ in either $\mathbb{R}^n$ or a quotient of $\mathbb{R}^n$ on which the Lebesgue measure is well defined. Certain regularity assumptions must generally be placed on the boundary of $\overline{\mathcal{U}}$ in order for all the results stated below to hold; we will ignore this detail except to mention that the necessary assumptions are satisfied for the two classes of domains that we are most interested in, which are

$$\mathcal{U} = \mathring{\mathbb{D}} \subset \mathbb{C},$$
$$\mathcal{U} = (0, L) \times S^1 \subset \mathbb{C}/i\mathbb{Z}, \qquad 0 < L \leqslant \infty.$$

Here $\mathbb{D}$ denotes the closed unit disk, $\mathring{\mathbb{D}}$ is its interior, and the identification of $(0, L) \times S^1 = (0, L) \times (\mathbb{R}/\mathbb{Z})$ with a subset of $\mathbb{C}/i\mathbb{Z}$ arises from the obvious identification of $\mathbb{R}^2$ with $\mathbb{C}$. Certain results will be specified to hold only for *bounded* domains, which means in practice that they hold on $\mathring{\mathbb{D}}$ and $(0, L) \times S^1$ for any $L > 0$, but not on $(0, \infty) \times S^1$.

Recall that for $p \in [1, \infty)$ we define the L^p norm of a measurable function $f : \mathcal{U} \to \mathbb{R}^m$ to be

$$\|f\|_{L^p} = \left(\int_{\mathcal{U}} |f|^p \right)^{1/p}.$$

For the space L^∞ we define the norm to be the essential supremum of f over $\mathcal{U}$. Denote by

$$C_0^\infty(\mathcal{U}) \subset C^\infty(\mathcal{U})$$

the space of smooth functions with compact support in $\mathcal{U}$. We say a function f has a **weak j-th partial derivative** g if the *integration by parts* formula holds for all so-called **test functions** $\varphi \in C_0^\infty(\mathcal{U})$:

$$\int_{\mathcal{U}} g\varphi = -\int_{\mathcal{U}} f\, \partial_j \varphi.$$

Equivalently, this means that g is a partial derivative of f **in the sense of distributions** (see e.g. [**LL01**]). Higher order weak partial derivatives are defined similarly:

recall that for a multiindex $\alpha = (i_1, ...i_n)$ we denote

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{i_1} \dots \partial x_n^{i_n}},$$

where $|\alpha| := \sum_j i_j$. We then write $\partial^\alpha f = g$ if for all $\varphi \in C_0^\infty(\mathcal{U})$,

$$\int_{\mathcal{U}} g\varphi = (-1)^{|\alpha|} \int_{\mathcal{U}} f\, \partial^\alpha \varphi.$$

Now we may define $W^{k,p}(\mathcal{U})$ to be the set of functions on $\mathcal{U}$ with weak partial derivatives up to order k lying in L^p, and define the norm of such a function by

$$\|f\|_{W^{k,p}} = \sum_{|\alpha| \leqslant k} \|\partial^\alpha f\|_{L^p}.$$

This definition gives $W^{0,p}(\mathcal{U}) = L^p(\mathcal{U})$, and in general $W^{k,p}(\mathcal{U})$ can be identified with a closed subset of a product of finitely many copies of $L^p(\mathcal{U})$, one for each multiindex of order at most k. This identification shows that is a Banach space; moreover, it can be shown to be reflexive and separable for $1 < p < \infty$.

While the Sobolev spaces $W^{k,p}(\mathcal{U})$ are generally defined on *open* domains, we often consider the closure $\overline{\mathcal{U}}$ as the domain for spaces of differentiable functions $C^k(\overline{\mathcal{U}})$ and $C^\infty(\overline{\mathcal{U}})$. For instance, $C^k(\overline{\mathcal{U}})$ is the Banach space of k-times differentiable functions on $\mathcal{U}$ whose derivatives up to order k are bounded and uniformly continuous on $\mathcal{U}$; note that uniform continuity implies the existence of continuous extensions to the closure $\overline{\mathcal{U}}$. Given suitable regularity assumptions for the boundary of $\overline{\mathcal{U}}$, one can show (with some effort—cf. [**AF03**, proof of Theorem 5.19]) that $C^k(\overline{\mathcal{U}})$ is precisely the set of functions which admit k-times differentiable extensions to some open set containing $\overline{\mathcal{U}}$.

EXERCISE 2.2.1. Show that if f is a continuous function on the closed disk $\mathbb{D} \subset \mathbb{C}$ that is continuously differentiable on $\dot{\mathbb{D}} = \mathbb{D} \backslash \{0\}$ and its first derivative is Lebesgue integrable on $\dot{\mathbb{D}}$, then f also has a weak first derivative on $\mathbb{D}$, which is equal to its classical derivative almost everywhere.

The following result is an amalgamation of frequently used special cases of the Sobolev embedding theorem and the Rellich-Kondrachov compactness theorem. See Theorems A.1.6 and A.1.10 in Appendix A for the more general versions, proofs of which may be found e.g. in [**AF03**].

PROPOSITION 2.2.2 (embedding/compactness). *Assume* $1 \leqslant p < \infty$ *and* $k \in \mathbb{N}$.

(1) If $kp > n$*, then for every integer* $d \geqslant 0$*, there exists a continuous inclusion*

$$W^{k+d,p}(\mathcal{U}) \hookrightarrow C^d(\overline{\mathcal{U}}),$$

which is compact if $\mathcal{U}$ *is bounded.*

(2) If $1 \leqslant q < \infty$ *and* $m \geqslant 0$ *is another integer such that* $k \geqslant m$*,* $p \leqslant q$ *and* $k - \frac{n}{p} \geqslant m - \frac{n}{q}$*, then there exists a continuous inclusion*

$$W^{k,p}(\mathcal{U}) \hookrightarrow W^{m,q}(\mathcal{U}),$$

which is compact if $\mathcal{U}$ *is bounded and the inequality* $k - \frac{n}{p} \geqslant m - \frac{n}{q}$ *is strict.*

□

The most important case of the second inclusion is $W^{k+1,p}(\mathcal{U}) \hookrightarrow W^{k,p}(\mathcal{U})$, whose continuity is obvious, and the compactness in the case of bounded $\mathcal{U}$ can be regarded as a natural analogue of the fact (arising from the Arzelà-Ascoli theorem) that the inclusions $C^{k+1}(\overline{\mathcal{U}}) \hookrightarrow C^k(\overline{\mathcal{U}})$ are compact when $\overline{\mathcal{U}}$ is compact. A useful way to remember the hypotheses in Proposition 2.2.2 is by thinking of $W^{k,p}(\mathcal{U})$ as a space of functions that have "$k - \frac{n}{p}$ continuous derivatives".

EXERCISE 2.2.3. Show that the compactness of the inclusions in Proposition 2.2.2 fails in general for unbounded domains, e.g. for $\mathbb{R}$.

The next three results for the case $kp > n$ are proved in §A.2 as corollaries of the Sobolev embedding theorem. The first is a Sobolev analogue of the fact that the product of a C^m-function with a C^k-function for $k \geqslant m$ is also of class C^m.

PROPOSITION 2.2.4 (Banach algebra property). *Suppose $1 \leqslant p, q < \infty$, $kp > n$, $k \geqslant m$ and $k - \frac{n}{p} \geqslant m - \frac{n}{q}$. Then the product pairing $(f, g) \mapsto fg$ defines a continuous bilinear map*

$$W^{k,p}(\mathcal{U}) \times W^{m,q}(\mathcal{U}) \to W^{m,q}(\mathcal{U}).$$

In particular this applies when $m = k$ and $q = p$, hence $W^{k,p}(\mathcal{U})$ is a Banach algebra. □

The continuity statements above translate into inequalities between the norms in the respective spaces. For example, continuous inclusions $W^{k+d,p} \hookrightarrow C^d$ or $W^{k,p} \hookrightarrow W^{m,q}$ respectively imply that

$$\|f\|_{C^d} \leqslant c\|f\|_{W^{k+d,p}} \qquad \text{or} \qquad \|f\|_{W^{m,q}} \leqslant c\|f\|_{W^{k,p}}$$

for some constants $c > 0$ which may depend on d, k, p, m, q or $\mathcal{U}$, but not f. Similarly, the Banach algebra property means there is an inequality

$$\|fg\|_{W^{k,p}} \leqslant c\|f\|_{W^{k,p}}\|g\|_{W^{k,p}}$$

whenever $kp > n$, where again the constant c is independent of g and f.

We state the next result only for the case of bounded domains; it does have an extension to unbounded domains, but the statement becomes more complicated (cf. Theorem A.2.6). Given an open set $\Omega \subset \mathbb{R}^n$, we denote

$$W^{k,p}(\mathcal{U}, \Omega) := \left\{ u \in W^{k,p}(\mathcal{U}, \mathbb{R}^n) \;\middle|\; \overline{u(\mathcal{U})} \subset \Omega \right\}.$$

Note that this is an open subset if $kp > n$, due to the Sobolev embedding theorem.

PROPOSITION 2.2.5 (C^k-continuity property). *Assume $1 \leqslant p < \infty$, $kp > n$, $\mathcal{U}$ is bounded and $\Omega \subset \mathbb{R}^n$ is an open set. Then there is a well-defined and continuous map*

$$W^{k,p}(\mathcal{U}, \Omega) \xrightarrow{T} \mathscr{L}\big(C^k(\Omega, \mathbb{R}^N), W^{k,p}(\mathcal{U}, \mathbb{R}^N)\big)$$
$$T(u)f := f \circ u,$$

where the space $\mathscr{L}\big(C^k(\Omega,\mathbb{R}^N), W^{k,p}(\mathcal{U},\mathbb{R}^N)\big)$ *of bounded linear maps from* $C^k(\Omega,\mathbb{R}^N)$ *to* $W^{k,p}(\mathcal{U},\mathbb{R}^N)$ *is equipped with the operator norm. It follows in particular that the map*

$$C^k(\Omega,\mathbb{R}^N) \times W^{k,p}(\mathcal{U},\Omega) \to W^{k,p}(\mathcal{U},\mathbb{R}^N) : (f,u) \mapsto f \circ u$$

is well defined and continuous. □

REMARK 2.2.6. In the settings of Propositions 2.2.4 and 2.2.5, it is also often important to know that the classical formulas for computing derivatives of fg or $f \circ u$ via the product or chain rules remain valid for computing *weak* derivatives of functions that are not necessarily classically differentiable. This is not true in general, but does hold in these specific settings due to the fact that Sobolev spaces contain dense subspaces of smooth functions. For details, see Proposition A.2.4 and Theorem A.2.6 in Appendix A.

REMARK 2.2.7. In later chapters, it will become important to be aware that Propositions 2.2.2, 2.2.4 and 2.2.5 are the essential conditions needed for defining smooth Banach manifold structures on spaces of $W^{k,p}$-smooth maps from one manifold to another, cf. Proposition 2.7.4 and [**Eliĭ67**, **Pal68**]. This only works under the condition $kp > n$, as the smooth category is not well equipped to deal with discontinuous maps!

The following rescaling result will be needed for nonlinear regularity arguments; see Theorem A.2.9 in Appendix A for a proof.

PROPOSITION 2.2.8. *Assume* $p \in [1,\infty)$ *and* $k \in \mathbb{N}$ *satisfy* $kp > n$*, let* $\mathring{\mathbb{D}}^n$ *denote the open unit ball in* $\mathbb{R}^n$*,* $x_0 \in \mathring{\mathbb{D}}^n$ *a fixed point, and for each* $f \in W^{k,p}(\mathring{\mathbb{D}}^n)$ *and* $\epsilon > 0$ *sufficiently small define* $f_\epsilon \in W^{k,p}(\mathring{\mathbb{D}}^n)$ *by*

$$f_\epsilon(x) := f(x_0 + \epsilon x).$$

Then for any $\alpha \in (0,1)$ *satisfying* $\alpha \leqslant k - n/p$*, there exists a constant* $C > 0$ *such that for every* $f \in W^{k,p}(\mathring{\mathbb{D}}^n)$ *and all* $\epsilon > 0$ *smaller than the distance from* x_0 *to* $\partial\mathbb{D}^n$*,*

$$\|f_\epsilon - f_\epsilon(0)\|_{W^{k,p}(\mathring{\mathbb{D}}^n)} \leqslant C\epsilon^\alpha \|f - f(x_0)\|_{W^{k,p}(\mathring{\mathbb{D}}^n)}.$$

□

EXERCISE 2.2.9. Working on a 2-dimensional domain with $kp > 2$, prove directly that for any multiindex α of positive order k,

$$\|\partial^\alpha f_\epsilon\|_{L^p(\mathring{\mathbb{D}})} \leqslant \epsilon^{k-2/p} \|\partial^\alpha f\|_{L^p(\mathring{\mathbb{D}})}$$

for $f \in W^{k,p}(\mathring{\mathbb{D}})$. Find examples (e.g. in $W^{1,2}(\mathring{\mathbb{D}})$) to show that no estimate of the form

$$\|\partial^\alpha f_\epsilon\|_{L^p(\mathring{\mathbb{D}})} \leqslant C_\epsilon \|f - f(x_0)\|_{W^{k,p}(\mathring{\mathbb{D}})}$$

with $\lim_{\epsilon \to 0^+} C_\epsilon = 0$ is possible when $kp \leqslant 2$.

2.3. The fundamental elliptic estimate

We will make considerable use of the fact that the linear first-order differential operator

$$\bar{\partial} := \partial_s + i\partial_t : C^\infty(\mathbb{C}, \mathbb{C}) \to C^\infty(\mathbb{C}, \mathbb{C})$$

is *elliptic*. We will briefly touch upon the general notion of ellipticity in a bit, but in practice, the main consequence we need to be aware of is the following pair of analytical results.

THEOREM 2.3.1. *If $1 < p < \infty$, then $\bar{\partial} : W^{1,p}(\mathring{\mathbb{D}}) \to L^p(\mathring{\mathbb{D}})$ admits a bounded right inverse $T : L^p(\mathring{\mathbb{D}}) \to W^{1,p}(\mathring{\mathbb{D}})$.*

THEOREM 2.3.2. *If $1 < p < \infty$ and $k \in \mathbb{N}$, then there exists a constant $c > 0$ such that for all $f \in W_0^{k,p}(\mathring{\mathbb{D}})$,*

$$\|f\|_{W^{k,p}} \leqslant c\|\bar{\partial} f\|_{W^{k-1,p}}.$$

Here $W_0^{k,p}(\mathring{\mathbb{D}})$ denotes the $W^{k,p}$-closure of the space $C_0^\infty(\mathring{\mathbb{D}})$ of smooth functions with compact support in $\mathring{\mathbb{D}}$.

The complete proofs of the two theorems above are rather lengthy, and we shall refer to [**Wenb**, §2.6 and §2.A] for the details, but we can at least explain why they hold in the case $p = 2$. First, it is straightforward to show that the function $K \in L^1_{\mathrm{loc}}(\mathbb{C})$ defined by

$$K(z) = \frac{1}{2\pi z}$$

is a **fundamental solution** for the equation $\bar{\partial} u = f$, meaning it satisfies

$$\bar{\partial} K = \delta$$

in the sense of distributions, where δ denotes the Dirac δ-function. Hence for any $f \in C_0^\infty(\mathbb{C})$, one finds a smooth solution $u : \mathbb{C} \to \mathbb{C}$ to the equation $\bar{\partial} u = f$ as the convolution

$$u(z) = (K * f)(z) := \int_{\mathbb{C}} K(z - \zeta) f(\zeta)\, d\mu(\zeta),$$

where $d\mu(\zeta)$ denotes the Lebesgue measure with respect to the variable $\zeta \in \mathbb{C}$. It will be useful to observe that whenever f has compact support on $\mathbb{C}$, $K * f$ also has decaying behavior at infinity:

LEMMA 2.3.3. *For any $f \in C_0^\infty(\mathbb{C})$, $K * f$ satisfies $|(K * f)(z)| \leqslant \frac{C}{|z|}$ for some constant $C > 0$.*

PROOF. Choose $R > 0$ large enough so that f is supported in the disk of radius R, and suppose $|z| \geqslant 2R$. Then for all $\zeta \in \mathbb{C}$ such that $f(\zeta) \neq 0$, we have $|z - \zeta| \geqslant |z| - R \geqslant \frac{|z|}{2}$, thus

$$\begin{aligned} |(K * f)(z)| &= \frac{1}{2\pi} \left| \int_{\mathbb{C}} \frac{f(\zeta)}{z - \zeta}\, d\mu(\zeta) \right| \leqslant \frac{1}{2\pi} \int_{\mathbb{C}} \frac{|f(\zeta)|}{|z - \zeta|}\, d\mu(\zeta) \\ &\leqslant \frac{1}{\pi|z|} \int_{\mathbb{C}} |f(\zeta)|\, d\mu(\zeta) = \frac{\|f\|_{L^1}}{\pi|z|}. \end{aligned}$$

□

If $u \in C_0^\infty(\mathbb{C})$ and $\bar{\partial} u = f$, it follows from Lemma 2.3.3 that $u - K * f$ is a holomorphic function on $\mathbb{C}$ that decays at infinity, hence $u \equiv K * f$. Since $C_0^\infty(\mathring{\mathbb{D}})$ is dense in $L^p(\mathring{\mathbb{D}})$ for all $p < \infty$, Theorem 2.3.1 now follows from the claim that for all $f \in C_0^\infty(\mathring{\mathbb{D}})$, there exist estimates of the form

$$\|K * f\|_{L^p(\mathring{\mathbb{D}})} \leqslant c\|f\|_{L^p(\mathring{\mathbb{D}})}, \qquad \|\partial_j(K * f)\|_{L^p(\mathring{\mathbb{D}})} \leqslant c\|f\|_{L^p(\mathring{\mathbb{D}})}, \tag{2.4}$$

with $\partial_j = \partial_s$ or ∂_t for $j = 1, 2$ respectively, and the constant $c > 0$ independent of f.

EXERCISE 2.3.4. Use Theorem 2.3.1 and the remarks above to prove Theorem 2.3.2 for the case $k = 1$ with $f \in C_0^\infty(\mathring{\mathbb{D}})$, then extend it to $f \in W_0^{1,p}(\mathring{\mathbb{D}})$ by a density argument. Then extend it to the general case by differentiating both f and $\bar{\partial} f$.

The first estimate in (2.4) is not too hard if you remember your introductory measure theory class: it follows from a general "potential inequality" for convolution operators (see [**Wenb**, Lemma 2.6.10]), similar to Young's inequality, the key points being that K is locally of class L^1 and $\mathring{\mathbb{D}}$ has finite measure. For the second inequality, observe that $\bar{\partial}(K * f) = f$, and the rest of the first derivative of $K * f$ is determined by $\partial(K * f)$, where

$$\partial := \partial_s - i\partial_t.$$

Differentiating K in the sense of distributions provides a formula for $\partial(K * f)$ as a principal value integral, namely

$$\partial(K * f)(z) = -\frac{1}{\pi} \lim_{\epsilon \to 0^+} \int_{|\zeta - z| \geqslant \epsilon} \frac{f(\zeta)}{(z - \zeta)^2}\, d\mu(\zeta).$$

This is a so-called **singular integral operator**: it is similar to our previous convolution operator, but more difficult to handle because the kernel $\frac{1}{z^2}$ is not of class L^1_{loc} on $\mathbb{C}$. The proof of the estimate

$$\|\partial(K * f)\|_{L^p} \leqslant c\|f\|_{L^p} \quad \text{for all } f \in C_0^\infty(\mathring{\mathbb{D}}) \tag{2.5}$$

follows from a rather difficult general estimate on singular integral operators, known as the *Calderón-Zygmund inequality*, cf. [**Wenb**, §2.A] and the references therein. The good news however is that the first step in that proof is not hard: that is the case $p = 2$.

As is the case for all elliptic operators with constant coefficients, the L^2-estimate on the fundamental solution of $\bar{\partial}$ admits an easy proof using Fourier transforms. In general, a sufficiently nice function $u : \mathbb{R}^n \to \mathbb{C}$ is related to its Fourier transform $\mathscr{F} u = \hat{u} : \mathbb{R}^n \to \mathbb{C}$ by

$$u(x) = \int_{\mathbb{R}^n} \hat{u}(p) e^{2\pi i (x \cdot p)}\, d\mu(p),$$

where $x \cdot p$ denotes the standard Euclidean inner product on $\mathbb{R}^n$. It thus satisfies the identity

$$\mathscr{F}(\partial_j u)(p) = 2\pi i p_j \hat{u}(p). \tag{2.6}$$

It follows more generally that for any differential operator D of order $k \in \mathbb{N}$ with constant coefficients acting on complex-valued functions on $\mathbb{R}^n$, there is a unique polynomial $P^D : \mathbb{R}^n \to \mathbb{C}$ of degree k such that

$$\mathscr{F}(Du)(p) = P^D(p)\hat{u}(p)$$

for reasonable functions $u : \mathbb{R}^n \to \mathbb{C}$. We call D an **elliptic** operator if $P^D(p) = P^D_k(p) + O(|p|^{k-1})$ and the homogeneous kth-order part P^D_k satisfies[5]

$$P^D_k(p) \neq 0 \quad \text{for all} \quad p \neq 0.$$

Since P^D_k is homogeneous with degree k, this condition implies that P^D satisfies an estimate of the form

$$|P^D(p)| \geqslant c|p|^k \quad \text{for all } p \in \mathbb{R}^n \text{ outside of some compact subset.}$$

Now if α is any multiindex of order $|\alpha| \leqslant k$, (2.6) implies $\mathscr{F}(\partial^\alpha u)(p) = (2\pi i p)^\alpha \hat{u}(p)$ with $|(2\pi i p)^\alpha| \leqslant c|p|^{|\alpha|} \leqslant c'|P^D(p)|$ for all $|p| \gg 0$ and some constant $c' > 0$. Since $(2\pi i p)^\alpha / P^D(p)$ is now a bounded function outside of some compact subset $K \subset \mathbb{R}^n$, one therefore obtains via Plancherel's theorem a bound of the form

$$\begin{aligned} \|\partial^\alpha u\|_{L^2(\mathbb{R}^n)} &= \|\mathscr{F}(\partial^\alpha u)\|_{L^2(\mathbb{R}^n)} = \|(2\pi i p)^\alpha \hat{u}\|_{L^2(\mathbb{R}^n)} \\ &= \|(2\pi i p)^\alpha \hat{u}\|_{L^2(K)} + \|(2\pi i p)^\alpha \hat{u}\|_{L^2(\mathbb{R}^n \setminus K)} \\ &\leqslant c\|\hat{u}\|_{L^2(K)} + c\|P^D(p)\hat{u}\|_{L^2(\mathbb{R}^n \setminus K)} \leqslant c\|u\|_{L^2(\mathbb{R}^n)} + c\|Du\|_{L^2(\mathbb{R}^n)}. \end{aligned}$$

In the case of $D := \bar{\partial}$ and $\partial^\alpha := \partial$ on $\mathbb{R}^2 = \mathbb{C}$, this story becomes especially simple since

$$\mathscr{F}(\bar{\partial} u)(\zeta) = 2\pi i \zeta \hat{u}(\zeta), \qquad \mathscr{F}(\partial u)(\zeta) = 2\pi i \bar{\zeta} \hat{u}(\zeta), \tag{2.7}$$

i.e. both $\bar{\partial}$ and ∂ are first-order elliptic operators.

Proposition 2.3.5. *For all $f \in C_0^\infty(\mathbb{C})$, we have $\|\partial(K * f)\|_{L^2} = \|f\|_{L^2}$.*

Proof. We write $u = K * f$, so $\bar{\partial} u = f$, and combining (2.7) with Plancherel's theorem gives

$$\begin{aligned} \|\partial(K * f)\|_{L^2} &= \|\partial u\|_{L^2} = \|\mathscr{F}(\partial u)\|_{L^2} = \|2\pi i \bar{\zeta} \hat{u}\|_{L^2} \\ &= \left\| \frac{\bar{\zeta}}{\zeta} 2\pi i \zeta \hat{u} \right\|_{L^2} = \|2\pi i \zeta \hat{u}\|_{L^2} = \|\hat{f}\|_{L^2} = \|f\|_{L^2}. \end{aligned}$$

□

Corollary 2.3.6. *The estimate (2.5) holds in the case $p = 2$.* □

[5]In the more general setting of a differential operator sending sections of one vector bundle to sections of another, the polynomial P^D in this discussion would take values in a space of linear maps from one finite-dimensional vector space to another. One then calls D elliptic if and only if the linear transformation $P^D(p)$ is invertible for all $p \neq 0$. More details on the general notion of ellipticity can be found e.g. in [**Wenb**, §2.B].

2.4. Regularity

We will now use the estimate $\|u\|_{W^{k,p}} \leqslant c\|\bar{\partial}u\|_{W^{k-1,p}}$ from the previous section to prove three types of results about solutions to Cauchy-Riemann type equations:

(1) All solutions of reasonable Sobolev-type regularity are smooth.
(2) Every sequence of solutions satisfying uniform bounds in certain Sobolev norms has a C^∞_{loc}-convergent subsequence.
(3) All reasonable Sobolev-type topologies on spaces of solutions are (locally) equivalent to the C^∞-topology.

In the following,

$$\mathbb{D}_r \subset \mathbb{C}$$

denotes the closed disk of radius $r > 0$, and $\mathring{\mathbb{D}}_r$ denotes its interior. Note that functions of class $C^\infty(\mathbb{D}_r)$ are assumed to be smooth up to the boundary (or equivalently, on some open neighborhood of $\mathbb{D}_r$ in $\mathbb{C}$), not just on $\mathring{\mathbb{D}}_r$.

2.4.1. The linear case.

Recall from Exercise 2.1.2 that every linear Cauchy-Riemann type operator on a vector bundle of complex rank n locally takes the form $\bar{\partial} + A$, where $\bar{\partial} = \partial_s + i\partial_t$, and A is a smooth function with values in $\text{End}_{\mathbb{R}}(\mathbb{C}^n)$. Using the Sobolev embedding theorem, the following result implies by induction that weak solutions of class L^p_{loc} for $1 < p < \infty$ to linear Cauchy-Riemann type equations are always smooth. The associated local estimate will also play a major role in our proof of the Fredholm property in Chapter 4.

THEOREM 2.4.1 (Linear regularity). *Assume $1 < p < \infty$, m and k are integers with $m \geqslant k \geqslant 0$, $A : \mathbb{D} \to \text{End}_{\mathbb{R}}(\mathbb{C}^n)$ is a C^m-smooth function, $f \in W^{m,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$ and $u \in W^{k,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$ is a weak solution to the equation*

$$(\bar{\partial} + A)u = f.$$

Then:

(1) u is of class $W^{m+1,p}$ on every compact subset of $\mathring{\mathbb{D}}$.
(2) For every $r \in (0,1)$, there exists a constant $c > 0$ dependent on the Sobolev parameters k, m, p, the radius r and the zeroth-order term A, but not on u or f, such that

$$\|u\|_{W^{m+1}(\mathring{\mathbb{D}}_r)} \leqslant c\|u\|_{W^{k,p}(\mathring{\mathbb{D}})} + c\|f\|_{W^{m,p}(\mathring{\mathbb{D}})}.$$

REMARK 2.4.2. A portion of the results in this section can be generalized to allow weaker regularity hypotheses on the zeroth-order term $A : \mathbb{D} \to \text{End}_{\mathbb{R}}(\mathbb{C}^n)$. In the proof of Theorem 2.4.1 below, for instance, the importance of the assumption $m \geqslant k$ lies in the fact that there is a continuous product pairing $C^m \times W^{k,p} \to W^{k,p}$, which in the case $k \geqslant 1$ remains true (by Proposition 2.2.4) if C^m is replaced by $W^{m,q}$ for any $q \in [p, \infty)$ satisfying $mq > 2$, and for $k = 0$ it is also true if C^m is replaced by L^∞. The theorem therefore also holds for zeroth-order terms A of suitable Sobolev-type regularity, including some cases where A is not even continuous. This level of generality is occasionally useful for technical reasons, e.g. it plays a role in our proof of the similarity principle (Theorem 2.5.3) in the next section.

PROOF OF THEOREM 2.4.1 (EXCLUDING (1) FOR $k = 0$). We first prove statement (2), assuming that statement (1) is already known. It will suffice to prove the estimate for the case $m = k$, because if $m > k$, one can then repeat the same argument $m - k + 1$ times, shrinking to a slightly smaller compact subset of $\mathring{\mathbb{D}}$ each time. With this understood, let us fix an integer $k \geqslant 0$ and consider a weak solution $u \in W^{k,p}(\mathring{\mathbb{D}})$ to the equation $(\bar{\partial} + A)u = f$ with $f \in W^{k,p}(\mathring{\mathbb{D}})$ and $A \in C^k(\mathbb{D}, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n))$. For any $r \in (0, 1)$, statement (1) in the theorem implies $u \in W^{k+1,p}(\mathring{\mathbb{D}}_r)$, and our objective is to bound $\|u\|_{W^{k+1,p}(\mathring{\mathbb{D}}_R)}$ in terms of $\|u\|_{W^{k,p}(\mathring{\mathbb{D}})}$ and $\|f\|_{W^{k,p}(\mathring{\mathbb{D}})}$.

In order to apply the fundamental elliptic estimate, we need to work with functions with compact support in the interior of $\mathbb{D}$, thus we choose a smooth bump function

$$\beta \in C_0^\infty(\mathring{\mathbb{D}}, [0, 1])$$

that satisfies $\beta|_{\mathbb{D}_r} \equiv 1$. Using this choice, we now give two slightly different proofs of the required estimate. The first is based on the observation that since u is locally of class $W^{k+1,p}$ on $\mathring{\mathbb{D}}$, $\beta u \in W_0^{k+1,p}(\mathring{\mathbb{D}})$, so Theorem 2.3.2 can be applied to βu, giving

$$\begin{aligned}\|u\|_{W^{k+1,p}(\mathring{\mathbb{D}}_r)} \leqslant \|\beta u\|_{W^{k+1,p}(\mathring{\mathbb{D}})} \leqslant c\|\bar{\partial}(\beta u)\|_{W^{k,p}} &\leqslant c\|(\bar{\partial}\beta)u\|_{W^{k,p}} + c\|\beta(f - Au)\|_{W^{k,p}} \\ &\leqslant c'\|u\|_{W^{k,p}} + c'\|f\|_{W^{k,p}},\end{aligned}$$

where the use of the Leibniz rule to compute $\bar{\partial}(\beta u)$ is unproblematic since β is smooth, and we have absorbed the C^k-norms of β, $\bar{\partial}\beta$ and A into the constant $c' > 0$. Note that the latter makes use of the continuous product pairing $C^k \times W^{k,p} \to W^{k,p}$ (cf. Remark 2.4.2).

The following alternative proof of this estimate is valid only if $k \geqslant 1$ and is slightly less direct, but contains useful ideas that we will need to recycle in the proof of statement 1. By assumption, we already have a bound on $\|u\|_{W^{k,p}(\mathring{\mathbb{D}}_r)}$, so the required $W^{k+1,p}$-bound will follow if we can also find $W^{k,p}$-bounds over $\mathring{\mathbb{D}}_r$ for the weak partial derivatives $\partial_j u$, $j = 1, 2$. These functions are (according to statement 1) of class $W^{k,p}_{\mathrm{loc}}$, and since $k \geqslant 1$ and $\beta\, \partial_j u \in W_0^{k,p}(\mathring{\mathbb{D}})$, we can now apply Theorem 2.3.2 to $\beta\, \partial_j u$, giving

$$\begin{aligned}\|\partial_j u\|_{W^{k,p}(\mathring{\mathbb{D}}_r)} \leqslant \|\beta\, \partial_j u\|_{W^{k,p}(\mathring{\mathbb{D}})} &\leqslant c\left\|\bar{\partial}\,(\beta\, \partial_j u)\right\|_{W^{k-1,p}(\mathring{\mathbb{D}})} \\ &\leqslant c\|(\bar{\partial}\beta)(\partial_j u)\|_{W^{k-1,p}} + c\|\beta\, \bar{\partial}(\partial_j u)\|_{W^{k-1,p}}.\end{aligned} \tag{2.8}$$

The first term on the right hand side is bounded by $c'\|u\|_{W^{k,p}}$ for some constant $c' > 0$ that depends on the C^{k-1}-norm of $\bar{\partial}\beta$. To control the second term, we differentiate the equation $\bar{\partial}u = -Au + f$, giving

$$\bar{\partial}(\partial_j u) = -(\partial_j A)u - A\, \partial_j u + \partial_j f,$$

where the Leibniz rule has been used to compute $\partial_j(Au)$ in light of Remark A.2.5 and the continuous product pairing $C^k \times W^{k,p} \to W^{k,p}$. The $W^{k-1,p}$-norm of $\beta\, \bar{\partial}(\partial_j u)$ is now bounded by a constant times $\|u\|_{W^{k-1,p}} + \|\partial_j u\|_{W^{k-1,p}} + \|\partial_j f\|_{W^{k-1,p}} \leqslant 2\|u\|_{W^{k,p}} + \|f\|_{W^{k,p}}$, where the constant in question depends only on $\|\beta\|_{C^{k-1}}$ and $\|A\|_{C^k}$.

We now prove statement (1) in the case $k \geqslant 1$; the case $k = 0$ requires a different argument and will be dealt with as an addendum at the end of this subsection. For

$k \geqslant 1$, we can use an adaptation of the second proof of statement 2 above, where instead of proving bounds on partial derivatives $\partial_j u$, we consider the corresponding **difference quotients**

$$D_j^h u(z) := \frac{u(z + he_j) - u(z)}{h}, \qquad j = 1, 2.$$

Here $e_1 := \partial_s$, $e_2 := \partial_t$, and the domain of $D_j^h u$ can be taken to be $\mathbb{D}_r$ for any $r \in (0,1)$ if $h \in \mathbb{R}\backslash\{0\}$ is sufficiently close to 0. It suffices again to consider only the case $m = k$, so let us suppose $u, f \in W^{k,p}(\mathring{\mathbb{D}})$ and $A \in C^k(\mathbb{D})$. The difference quotients $D_j^h u$ are then also of class $W^{k,p}_{\text{loc}}$ on their domains, so for the smooth cutoff function $\beta \in C_0^\infty(\mathring{\mathbb{D}})$ with $\beta|_{\mathbb{D}_r} \equiv 1$, we can assume for all $|h| > 0$ sufficiently small that $\beta D_j^h u$ is in $W_0^{k,p}(\mathring{\mathbb{D}})$. The analogue of (2.8) in this context is then

$$\begin{aligned} \|D_j^h u\|_{W^{k,p}(\mathring{\mathbb{D}}_r)} &\leqslant \|\beta D_j^h u\|_{W^{k,p}(\mathring{\mathbb{D}})} \leqslant c \left\|\bar\partial\left(\beta\, D_j^h u\right)\right\|_{W^{k-1,p}(\mathring{\mathbb{D}})} \\ &\leqslant c\|(\bar\partial\beta)(D_j^h u)\|_{W^{k-1,p}} + c\|\beta\, \bar\partial(D_j^h u)\|_{W^{k-1,p}}. \end{aligned}$$

The first term is bounded independently of h since $\partial_j u \in W^{k-1,p}(\mathring{\mathbb{D}})$, implying a uniform $W^{k-1,p}$-bound on $D_j^h u$ as $h \to 0$; cf. Appendix A.3. To control the second term, we can apply the operator D_j^h to the equation $\bar\partial u = -Au + f$, giving

$$\bar\partial(D_j^h u) = D_j^h(\bar\partial u) = -(D_j^h A)u - A\, D_j^h u + D_j^h f.$$

Since $A \in C^k(\mathbb{D})$, $D_j^h A$ is uniformly C^{k-1}-bounded as $h \to 0$, and $\partial_j u, \partial_j f \in W^{k-1,p}(\mathring{\mathbb{D}})$ similarly implies uniform $W^{k-1,p}$-bounds on $D_j^h u$ and $D_j^h f$, thus the whole expression is uniformly $W^{k-1,p}$-bounded on some open disk containing the support of β, implying

$$\|D_j^h u\|_{W^{k,p}(\mathring{\mathbb{D}}_r)} \leqslant c$$

for some constant $c > 0$ that does not change as $h \to 0$. This implies $u \in W^{k+1,p}(\mathring{\mathbb{D}}_r)$ via a standard application of the Banach-Alaoglu theorem. Indeed, the latter implies that if there is a uniform bound on $\|D_j^h u\|_{L^p}$ as $h \to 0$, then any decaying sequence $h_\nu \to 0$ has a subsequence for which $D_j^{h_\nu} u$ is weakly L^p-convergent. The limit of this subsequence belongs to $L^p(\mathring{\mathbb{D}}_r)$, and it is straightforward to show using the definition of weak derivatives that this limit is $\partial_j u$. One finds a similar result in the presence of uniform $W^{k,p}$-bounds for any $k \in \mathbb{N}$ by applying this argument to higher-order derivatives of $\partial_j u$; for details, see Theorem A.3.1 in Appendix A.3. □

EXERCISE 2.4.3. Deduce from Theorem 2.4.1 the following corollaries for a sequence of weak solutions $u_\nu \in W^{k,p}(\mathring{\mathbb{D}})$ to $(\bar\partial + A_\nu)u_\nu = f_\nu$, assuming $f_\nu \in W^{m,p}(\mathring{\mathbb{D}})$ and $A_\nu \in C^m(\mathring{\mathbb{D}})$ for all $\nu \in \mathbb{N}$, with $m \geqslant k \geqslant 0$ and $1 < p < \infty$.

(a) If $\|u_\nu\|_{W^{k,p}(\mathring{\mathbb{D}})}$, $\|f_\nu\|_{W^{m,p}(\mathring{\mathbb{D}})}$ and $\|A_\nu\|_{C^m(\mathbb{D})}$ are uniformly bounded, then u_ν is also uniformly $W^{k+1,p}$-bounded on compact subsets of $\mathring{\mathbb{D}}$.

(b) If u_ν is $W^{k,p}$-convergent, f_ν is $W^{m,p}$-convergent and A_ν is C^m-convergent on $\mathbb{D}$, then u_ν is also $W^{m+1,p}_{\text{loc}}$-convergent on $\mathring{\mathbb{D}}$.

REMARK 2.4.4. Combining the Sobolev embedding theorem with the Arzelà-Ascoli theorem, the result of Exercise 2.4.3(a) proves that if the f_ν and A_ν are C^∞-bounded on $\mathbb{D}$, then a $W^{k,p}$-bounded sequence of solutions u_ν has a C^∞_{loc}-convergent subsequence. Part (b) implies moreover that for every $k \geqslant 0$ and $p \in (1, \infty)$, the $W^{k,p}$-topology on spaces of solutions to linear Cauchy-Riemann type equations is locally equivalent to the C^∞-topology.

EXERCISE 2.4.5. Use Theorem 2.4.1 to generalize Theorem 2.3.1 to the existence of a bounded right inverse for

$$\bar{\partial} : W^{k,p}(\mathring{\mathbb{D}}) \to W^{k-1,p}(\mathring{\mathbb{D}}).$$

for every $k \in \mathbb{N}$ and $1 < p < \infty$. *Hint: For any $R > 1$, there exists a bounded linear extension operator $E : W^{k,p}(\mathring{\mathbb{D}}) \to W^{k,p}(\mathring{\mathbb{D}}_R)$ with the property $(Ef)|_{\mathring{\mathbb{D}}} = f$ for all $f \in W^{k,p}(\mathring{\mathbb{D}})$; see Theorem A.1.4 and Corollary A.1.5.*

It remains to prove the case $k = 0$ of Theorem 2.4.1(1). As preparation for this, we start with a classical result about "weakly holomorphic" functions:

LEMMA 2.4.6. *If $u \in L^1(\mathring{\mathbb{D}})$ satisfies $\bar{\partial} u = 0$ in the sense of distributions, then u is smooth and holomorphic on the open disk $\mathring{\mathbb{D}}$.*

PROOF. Taking real and imaginary parts, it suffices to prove that the same statement holds for the Laplace equation. By mollification, any weakly harmonic function can be approximated in L^1 with smooth harmonic functions. The latter satisfy the mean value property, which behaves well under L^1-convergence, so the result follows from the mean value characterization of harmonic functions; see [**Wenb**, Lemma 2.6.26] for more details. □

LEMMA 2.4.7. *Suppose $1 < p < \infty$ and $u \in L^1(\mathring{\mathbb{D}})$ is a weak solution to $\bar{\partial} u = f$ for some $f \in L^p(\mathring{\mathbb{D}})$. Then u is of class $W^{1,p}$ on every compact subset of $\mathring{\mathbb{D}}$.*

PROOF. Let $T : L^p(\mathring{\mathbb{D}}) \to W^{1,p}(\mathring{\mathbb{D}})$ denote the bounded right inverse of $\bar{\partial} : W^{1,p}(\mathring{\mathbb{D}}) \to L^p(\mathring{\mathbb{D}})$ provided by Theorem 2.3.1. Then $u - Tf \in L^1(\mathring{\mathbb{D}})$ is a weak solution to $\bar{\partial}(u - Tf) = 0$ and is thus smooth by Lemma 2.4.6. In particular, $u - Tf$ restricts to $\mathring{\mathbb{D}}_r$ for every $r < 1$ as a function of class $W^{1,p}$, implying that u also has a restriction in $W^{1,p}(\mathring{\mathbb{D}}_r)$. □

PROOF OF THEOREM 2.4.1(1) FOR $k = 0$. Suppose $(\bar{\partial} + A)u = f$, where A is continuous on $\mathbb{D}$ and $u, f \in L^p(\mathring{\mathbb{D}})$. Then $\bar{\partial} u = -Au + f \in L^p(\mathring{\mathbb{D}})$, so Lemma 2.4.7 implies $u \in W^{1,p}_{\text{loc}}(\mathring{\mathbb{D}})$. If $m \geqslant 1$, one can now shrink the disk slightly and plug in the case $k = 1$ of the theorem to conclude $u \in W^{m+1,p}_{\text{loc}}(\mathring{\mathbb{D}})$. □

COROLLARY 2.4.8. *If $A : \mathbb{D} \to \text{End}_{\mathbb{R}}(\mathbb{C}^n)$ is of class C^m for $0 \leqslant m \leqslant \infty$, then every weak solution $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ to $(\bar{\partial} + A)u = 0$ of class L^p_{loc} for a given $p \in (1, \infty)$ is also in $W^{k,q}_{\text{loc}}(\mathring{\mathbb{D}})$ for every $k \leqslant m + 1$ and $q \in (1, \infty)$. In particular, u is of class C^m.*

PROOF. Assume for simplicity $m < \infty$, as the case $m = \infty$ will then immediately follow. Theorem 2.4.1(1) implies $u \in W^{m+1,p}(\mathring{\mathbb{D}}_r)$ for any $r < 1$. If $p > 2$, this implies via the Sobolev embedding theorem that $u \in C^m(\mathbb{D}_r)$. In particular, u is then continuous and bounded on the closed disk $\mathbb{D}_r$, so it is in $L^q(\mathring{\mathbb{D}}_r)$ for every $q \in (1, \infty)$, and feeding it into Theorem 2.4.1(1) again gives the desired result on $\mathbb{D}_r$. Since $r < 1$ was arbitrary, the result is therefore true on any compact subset of $\mathring{\mathbb{D}}$.

To finish, it will now suffice to show that if $u \in L^p(\mathring{\mathbb{D}})$ for some $p \leqslant 2$, then u is also in $L^q_{\mathrm{loc}}(\mathring{\mathbb{D}})$ for some $q > 2$. Here Theorem 2.4.1(1) again implies $u \in W^{1,p}(\mathring{\mathbb{D}}_r)$ for any $r < 1$, and according to the Sobolev embedding theorem, there is a continuous inclusion $W^{1,p} \hookrightarrow L^q$ whenever $p \leqslant q < p^*$, where $p^* \in (p, \infty]$ is determined by $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{2}$; see Theorem A.1.6. Since $p > 1$, this implies $\frac{1}{p^*} < \frac{1}{2}$ and thus $p^* > 2$, so we can choose any $q \in (2, p^*)$ and conclude $u \in L^q(\mathring{\mathbb{D}}_r)$. □

2.4.2. The nonlinear case: bootstrapping. The regularity argument in the previous subsection was inductive in nature: if the data A and f in the equation $(\bar{\partial} + A)u = f$ are smooth, then assuming $u \in W^{k,p}$ leads via elliptic estimates to the conclusion that u is actually of class $W^{k+1,p}$, so by induction, u is smooth. This technique is known in the PDE literature as *elliptic bootstrapping*. We will now prove a similar bootstrapping result for the nonlinear Cauchy-Riemann equation.

Locally, every J-holomorphic curve can be regarded as a map $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ satisfying $\partial_s u(z) + J(u(z))\partial_t u(z) = 0$ in coordinates $z = s + it \in \mathbb{D} \subset \mathbb{C}$, where J is an almost complex structure on $\mathbb{C}^n$, or equivalently, a function[6]

$$J : \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n) := \left\{ K \in \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n) \;\middle|\; K^2 = -\mathbb{1} \right\}.$$

A nonlinear analogue of the equation considered in Theorem 2.4.1 is then the *inhomogeneous* nonlinear Cauchy-Riemann equation

$$\partial_s u + J(u)\partial_t u = f \tag{2.9}$$

for functions $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$, where $J : \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$ and $f : \mathring{\mathbb{D}} \to \mathbb{C}^n$ are given.

REMARK 2.4.9. It is worth mentioning that while other nonlinear analogues of the equation $(\bar{\partial} + A)u = 0$ are possible, the equations of interest can all be reduced to (2.9) in practice. For example, in Floer homology and Gromov-Witten theory, one often considers equations that locally take the form

$$\partial_s u(z) + J(z, u(z))\partial_t u(z) = f(z, u(z)),$$

where $J : \mathring{\mathbb{D}} \times \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$ is now allowed to depend explicitly on points in the domain of $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$, and $f : \mathring{\mathbb{D}} \times \mathbb{C}^n \to \mathbb{C}^n$ is likewise a function of both z and the value $u(z)$. As was observed by Gromov already in [Gro85, 1.4.C], the solutions $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ to this equation are equivalent to honest $\bar{J}$-holomorphic curves in $\mathring{\mathbb{D}} \times \mathbb{C}^n$

[6]Here the reader should beware of a minor ambiguity in notation: while we used $\mathcal{J}(M)$ in Chapter 1 to mean the space of smooth almost complex structures on a manifold M, one can just as sensibly define $\mathcal{J}(V)$ for each real $2n$–dimensional vector space V to be the space of *linear* complex structures on V, topologized as a subset of the finite-dimensional vector space $\mathrm{End}_{\mathbb{R}}(V) \cong \mathbb{R}^{2n \times 2n}$. It is not hard to show that $\mathcal{J}(V)$ is then a smooth submanifold of $\mathrm{End}_{\mathbb{R}}(V)$; in fact, the ability to choose J-complex bases for each $J \in \mathcal{J}(V)$ gives $\mathcal{J}(V)$ a natural identification with the homogeneous space $\mathrm{Aut}_{\mathbb{R}}(V)/\mathrm{Aut}_{\mathbb{C}}(V, J) \cong \mathrm{GL}(2n, \mathbb{R})/\mathrm{GL}(n, \mathbb{C})$. In the present discussion, the notation $\mathcal{J}(\mathbb{C}^n)$ views $\mathbb{C}^n$ as a real $2n$-dimensional vector space rather than as a manifold.

of the form $\bar{u}(z) := (z, u(z))$ if one defines an almost complex structure $\bar{J}$ in block form on $\mathring{\mathbb{D}} \times \mathbb{C}^n$ by

$$\bar{J}(z, x) := \begin{pmatrix} i & 0 \\ f(z,x)i & J(z,x) \end{pmatrix}.$$

For this reason, all theorems about regularity of "honest" J-holomorphic curves imply similar results about the nonlinear inhomogeneous equations in Floer homology and Gromov-Witten theory.

The following is *not* the most general theorem provable about regularity for J-holomorphic curves, though it is the one that is most closely analogous to the linear result in the previous subsection, and is also the one that we will need most often. A partial improvement with weaker hypotheses will be discussed in §2.4.3 below.

THEOREM 2.4.10 (Nonlinear regularity, $kp > 2$ version). *Assume $1 < p < \infty$, and m and k are integers with $m \geqslant k$ and $kp > 2$.*

(1) If $J : \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$ is of class C^m and $u \in W^{k,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$ is a weak solution to the equation

$$\partial_s u + J(u)\partial_t u = f$$

for some $f \in W^{m,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$, then u is of class $W^{m+1,p}$ on every compact subset of $\mathring{\mathbb{D}}$.

(2) Consider a C^m_{loc}-convergent sequence $J_\nu \to J$ of C^m-smooth almost complex structures $\mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$, together with sequences $f_\nu \in W^{m,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$ and $u_\nu \in W^{k,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$ such that for each $\nu \in \mathbb{N}$, u_ν is a weak solution to the equation

$$\partial_s u_\nu + J_\nu(u_\nu)\partial_t u_\nu = f_\nu.$$

(a) If the norms $\|f_\nu\|_{W^{m,p}}$ and $\|u_\nu\|_{W^{k,p}}$ on $\mathring{\mathbb{D}}$ are uniformly bounded as $\nu \to \infty$, then u_ν is also uniformly $W^{m+1,p}$-bounded on every compact subset of $\mathring{\mathbb{D}}$.

(b) If f_ν is $W^{m,p}$-convergent and u_ν is $W^{k,p}$-convergent on $\mathring{\mathbb{D}}$, then u_ν is also $W^{m+1,p}$-convergent on every compact subset of $\mathring{\mathbb{D}}$.

Combining this result with the Sobolev embedding theorem and the Arzelà-Ascoli theorem yields:

COROLLARY 2.4.11. *If J is a smooth almost complex structure on $\mathbb{C}^n$, then every J-holomorphic map $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ that is of class $W^{k,p}$ for some $k \in \mathbb{N}$ and $p \in (1, \infty)$ with $kp > 2$ is smooth. Moreover, if $J_\nu \to J$ is a C^∞_{loc}-convergent sequence of almost complex structures on $\mathbb{C}^n$ and $u_\nu : \mathring{\mathbb{D}} \to \mathbb{C}^n$ is a sequence of J_ν-holomorphic maps, then for any $k \in \mathbb{N}$ and $p \in (1, \infty)$ with $kp > 2$, uniform $W^{k,p}$-bounds for u_ν imply C^∞_{loc}-convergence of a subsequence of u_ν, and similarly, $W^{k,p}$-convergence of u_ν implies C^∞_{loc}-convergence.* □

REMARK 2.4.12. We will take pains to avoid dealing with non-smooth almost complex structures in this book, but in some applications they are unavoidable for technical reasons. In such cases, one gets the most mileage out of Theorem 2.4.10 by choosing $p > 2$, as the Sobolev embedding theorem then implies that J-holomorphic curves of class $W^{1,p}$ have at least as many continuous derivatives as J does. If one instead starts with a curve u of class $W^{k,p}_{\text{loc}}$ for some $p \leqslant 2$ but $kp > 2$, then since $k \geqslant 2$, one can use the Sobolev embedding theorem to argue (cf. Corollary 2.4.8) that u is therefore also of class $W^{1,q}_{\text{loc}}$ for some $q > 2$, which leads to the same result. To summarize: if J is of class C^m, then any J-holomorphic curve of class $W^{k,p}_{\text{loc}}$ for some k, p with $kp > 2$ is also of class $W^{m+1,q}_{\text{loc}}$ for every $q \in (1, \infty)$, and in particular it is of class C^m.

Our proof of Theorem 2.4.10 will follow a similar outline to the proof of Theorem 2.4.1, which can be interpreted as the special case where $J_\nu \equiv i$ for all ν. The reason it works more generally is that if we zoom in on a sufficiently small neighborhood of one point in $\mathbb{C}^n$, then J can be viewed as a C^m-small perturbation of i. To make this precise, we shall use the following rescaling trick.

Associate to any C^m-smooth map $J : \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$ the function

$$Q := i - J \in C^m(\mathbb{C}^n, \operatorname{End}_{\mathbb{R}}(\mathbb{C}^n)).$$

In terms of Q, the equation $\partial_s u + J(u)\partial_t u = f$ then becomes

$$\bar{\partial} u - Q(u)\partial_t u = f. \tag{2.10}$$

For any given point $z_0 \in \mathring{\mathbb{D}}$, we can assume without loss of generality after an affine change of coordinates on $\mathbb{C}^n$ that $u(z_0) = 0$ and $J(0) = i$, so in particular $Q(0) = 0$. For any $\epsilon \in (0, \operatorname{dist}(z_0, \partial\mathbb{D}))$ and a fixed constant $\alpha \in (0,1)$ to be specified further below, we now associate to J, u and f the functions

$$\begin{aligned}
\widehat{J} &: \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n), & \widehat{J}(x) &:= J(\epsilon^\alpha x), \\
\widehat{Q} &: \mathbb{C}^n \to \operatorname{End}_{\mathbb{R}}(\mathbb{C}^n), & \widehat{Q}(x) &:= Q(\epsilon^\alpha x) = i - \widehat{J}(x), \\
\hat{u} &: \mathring{\mathbb{D}} \to \mathbb{C}^n, & \hat{u}(z) &:= \frac{1}{\epsilon^\alpha} u(z_0 + \epsilon z), \\
\hat{f} &: \mathring{\mathbb{D}} \to \mathbb{C}^n, & \hat{f}(z) &:= \epsilon^{1-\alpha} f(z_0 + \epsilon z).
\end{aligned} \tag{2.11}$$

Then u satisfies (2.10) if and only if $\hat{u}$ satisfies

$$\bar{\partial}\hat{u} - \widehat{Q}(\hat{u})\partial_t \hat{u} = \hat{f}. \tag{2.12}$$

The rescaled almost complex structure has the convenient feature that since $J(0)$ is the standard complex structure i, choosing $\epsilon > 0$ small makes $\widehat{J}$ arbitrarily C^m-close to i on the compact set[7] $\mathbb{D}^{2n} \subset \mathbb{C}^n$, which means $\|\widehat{Q}\|_{C^m(\mathbb{D}^{2n})}$ can be made arbitrarily small. By Proposition 2.2.8, $\|\hat{u}\|_{W^{k,p}(\mathring{\mathbb{D}})}$ will likewise stay under control for $\epsilon \to 0$ if we choose $\alpha \in (0,1)$ such that $\alpha \leqslant k - 2/p$, and in fact, choosing α to be slightly smaller then ensures that we can make $\|\hat{u}\|_{W^{k,p}}$ an arbitrarily small multiple of $\|u\|_{W^{k,p}}$ by choosing $\epsilon > 0$ small. Since $kp > 2$, this will also make $\|\hat{u}\|_{C^0}$ arbitrarily small, and we can therefore assume that $\hat{u}$ has image in $\mathbb{D}^{2n}$. By the assumption $m \geqslant k$ and the continuity of the map $C^k \times W^{k,p} \to W^{k,p}$ in Proposition 2.2.5, the function

[7] Here $\mathbb{D}^{2n}$ denotes the closed unit ball in $\mathbb{C}^n = \mathbb{R}^{2n}$.

$\mathbb{D} \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n) : z \mapsto \hat{Q}(\hat{u}(z))$ can then likewise be assumed to be arbitrarily $W^{k,p}$-small by choosing $\epsilon > 0$ small. The effect is that (2.12) can now be viewed as a $W^{k,p}$-close approximation of the linear equation $\bar{\partial}\hat{u} = \hat{f}$.

The price we pay for this rescaling is that if we are able to prove e.g. a uniform bound on the norms $\|\hat{u}_\nu\|_{W^{k+1,p}(\mathring{\mathbb{D}}_r)}$ for some sequence $u_\nu \in W^{k,p}(\mathring{\mathbb{D}})$ and $r \in (0,1)$, then the resulting $W^{k+1,p}$-bound for u_ν will be valid only on the ϵ-disk around the point z_0. But this point was chosen arbitrarily in $\mathring{\mathbb{D}}$, so the result is a uniform bound over some neighborhood of *any* interior point of $\mathbb{D}$, and since a compact subset of $\mathring{\mathbb{D}}$ can be covered by finitely many such neighborhoods, that is enough to achieve uniform bounds over compact subsets.

REMARK 2.4.13. The rescaling trick described above is one of several reasons why the condition $kp > 2$ will be needed in the proof of Theorem 2.4.10, while it was irrelevant in the linear case. Another reason is of course the Sobolev embedding theorem, which guarantees that the solutions we consider are always continuous. We will see when we study compactness in Chapter 7 that the statement in Theorem 2.4.10 about uniform bounds is generally false when $kp \leqslant 2$, and even when we extend the statement about smoothness to allow $kp \leqslant 2$ in §2.4.3, continuity will have to be imposed as an explicit extra hypothesis.

PROOF OF THEOREM 2.4.10. We will prove statement (2a) assuming that statement (1) is already known, and leave the rest as exercises.

Since $m \geqslant k$, it suffices to prove the statement for the case $k = m$, as otherwise the argument can always be repeated on slightly smaller disks at each step to increase k until it reaches m. We therefore assume that a C^k_{loc}-convergent sequence $J_\nu \to J$ of functions $\mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$ and sequences $u_\nu, f_\nu \in W^{k,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$ satisfying uniform bounds

$$\|u_\nu\|_{W^{k,p}} \leqslant M, \qquad \|f_\nu\|_{W^{k,p}} \leqslant M$$

are given such that $\partial_s u_\nu + J_\nu(u_\nu)\partial_t u_\nu = f_\nu$, and we need to establish that u_ν is also uniformly $W^{k+1,p}$-bounded over compact subsets. (Note that we can assume due to statement 1 in the theorem that each u_ν is of class $W^{k+1,p}_{\mathrm{loc}}$.) It suffices in fact to prove that every *subsequence* of u_ν has a further subsequence for which such uniform bounds hold; indeed, if the bound for the whole sequence did not exist, then we would be able to find a subsequence with norms blowing up to infinity over some compact subset, and no further subsequence of this subsequence could satisfy a uniform bound. With this understood, we can appeal to the compactness of the inclusion $W^{k,p}(\mathring{\mathbb{D}}) \hookrightarrow C^0(\mathbb{D})$ for $kp > 2$ (see Proposition 2.2.2), and replace u_ν with a subsequence (still denoted by u_ν) that is C^0-convergent on $\mathbb{D}$ to some continuous map $u : \mathbb{D} \to \mathbb{C}^n$.

For any given point $z_0 \in \mathring{\mathbb{D}}$, we can now apply a converging sequence of affine transformations to $\mathbb{C}^n$ in order to assume without loss of generality

$$u_\nu(z_0) = 0 \text{ for all } \nu, \qquad \text{and} \qquad J(0) = i.$$

We then choose

$$\alpha \in (0,1) \quad \text{with} \quad \alpha < k - \frac{2}{p}, \tag{2.13}$$

and apply the rescaling trick outlined above to replace u_ν, f_ν and J_ν with the corresponding rescalings $\hat{u}_\nu$, $\hat{f}_\nu$ and $\hat{J}_\nu$ as defined in (2.11), defining also the related functions $\hat{Q}_\nu = i - \hat{J}_\nu$. We then have the equation $\bar{\partial}\hat{u}_\nu - \hat{Q}_\nu(\hat{u}_\nu)\partial_t\hat{u}_\nu = \hat{f}_\nu$, with C^k-convergence $\hat{Q}_\nu \to \hat{Q}$ over $\mathbb{D}^{2n}$, where $\hat{Q}$ may be assumed arbitrarily C^k-small on this set by choosing $\epsilon > 0$ small. Since $\hat{u}_\nu(0) = u_\nu(z_0) = 0$ for all ν, we can choose some $\beta > \alpha$ that also satisfies the conditions in (2.13) and then apply Propostion 2.2.8 to obtain a bound

$$\|\hat{u}_\nu\|_{W^{k,p}} \leqslant C\epsilon^{\beta-\alpha}\|u_\nu\|_{W^{k,p}} \leqslant C\epsilon^{\beta-\alpha}M \tag{2.14}$$

for some constant $C > 0$ that is independent of ν and ϵ. We can therefore impose an arbitrarily small uniform $W^{k,p}$-bound (and therefore a similarly small C^0-bound) on $\hat{u}_\nu$ by choosing $\epsilon > 0$ small enough. For f_ν, it will suffice to know that the uniform bound $\|f_\nu\|_{W^{k,p}} \leqslant M$ implies a similar uniform bound

$$\|\hat{f}_\nu\|_{W^{k,p}} \leqslant M_\epsilon$$

for some constant $M_\epsilon > 0$ which may depend on ϵ, but not on ν. Our goal is now to prove that for some fixed choice of the rescaling parameter $\epsilon > 0$, $\|\partial_j\hat{u}_\nu\|_{W^{k,p}(\mathring{\mathbb{D}}_r)}$ is uniformly bounded for $j = 1, 2$ and some $r \in (0, 1)$.

The argument begins exactly the same as in the linear case: choose a smooth bump function

$$\beta \in C_0^\infty(\mathring{\mathbb{D}}, [0, 1])$$

that satisfies $\beta|_{\mathbb{D}_r} \equiv 1$. We then have $\beta\,\partial_j\hat{u}_\nu \in W_0^{k,p}(\mathring{\mathbb{D}})$, so by Theorem 2.3.2,

$$\|\partial_j\hat{u}_\nu\|_{W^{k,p}(\mathring{\mathbb{D}}_r)} \leqslant \|\beta\,\partial_j\hat{u}_\nu\|_{W^{k,p}(\mathring{\mathbb{D}})} \leqslant c\left\|\bar{\partial}\left(\beta\,\partial_j\hat{u}_\nu\right)\right\|_{W^{k-1,p}(\mathring{\mathbb{D}})}. \tag{2.15}$$

If this were still the proof of Theorem 2.4.1, we would now apply the Leibniz rule to write $\bar{\partial}(\beta\,\partial_j\hat{u}_\nu)$ as a sum of two terms, but the nonlinear case requires something slightly cleverer at this step. Let us instead derive a PDE satisfied by $\beta\,\partial_j\hat{u}_\nu$. Differentiating the equation $\bar{\partial}\hat{u}_\nu = \hat{Q}_\nu(\hat{u}_\nu)\partial_t\hat{u}_\nu + \hat{f}_\nu$ gives

$$\bar{\partial}(\partial_j\hat{u}_\nu) = \partial_j(\bar{\partial}\hat{u}_\nu) = \hat{Q}_\nu(\hat{u}_\nu)\partial_t\partial_j\hat{u}_\nu + D\hat{Q}_\nu(\hat{u}_\nu)\left(\partial_j\hat{u}_\nu, \partial_t\hat{u}_\nu\right) + \partial_j\hat{f}_\nu,$$

where $D\hat{Q}_\nu$ denotes the differential of $\hat{Q}_\nu$. In this calculation we have assumed that the product and chain rules are universally valid, but this requires some care since we are dealing with weak rather than classical derivatives: in fact, the chain rule can be used for differentiating $\hat{Q}_\nu(\hat{u}_\nu)$ according to Theorem A.2.6 since $\hat{u}_\nu$ is of class $W^{k,p}$ with $kp > 2$ and $\hat{Q}_\nu$ is of class C^k, and Proposition A.2.4 then justifies the product rule for $\hat{Q}_\nu(\hat{u}_\nu)\partial_t\hat{u}_\nu$ since $\hat{Q}_\nu(\hat{u}_\nu) \in W^{k,p}$, $\partial_t\hat{u}_\nu \in W^{k-1,p}$, and the product pairing $W^{k,p} \times W^{k-1,p} \to W^{k-1,p}$ is continuous. Returning to the formula itself, we now have

$$\begin{aligned}\bar{\partial}(\beta\,\partial_j\hat{u}_\nu) &= \beta\hat{Q}_\nu(\hat{u}_\nu)\partial_t\partial_j\hat{u}_\nu + \beta D\hat{Q}_\nu(\hat{u}_\nu)\left(\partial_j\hat{u}_\nu, \partial_t\hat{u}_\nu\right) + \beta\,\partial_j\hat{f}_\nu + (\bar{\partial}\beta)\partial_j\hat{u}_\nu \\ &= \hat{Q}_\nu(\hat{u}_\nu)\partial_t(\beta\,\partial_j\hat{u}_\nu) + D\hat{Q}_\nu(\hat{u}_\nu)\left(\beta\,\partial_j\hat{u}_\nu, \partial_t\hat{u}_\nu\right) \\ &\qquad + \beta\,\partial_j\hat{f}_\nu + (\bar{\partial}\beta)\partial_j\hat{u}_\nu - \hat{Q}_\nu(\hat{u}_\nu)(\partial_t\beta)\partial_j\hat{u}_\nu,\end{aligned}$$

so that $\beta\,\partial_j \hat{u}_\nu$ satisfies

$$\bar{\partial}(\beta\,\partial_j \hat{u}_\nu) - \widehat{Q}_\nu(\hat{u}_\nu)\partial_t(\beta\,\partial_j \hat{u}_\nu) = D\widehat{Q}_\nu(\hat{u}_\nu)(\beta\,\partial_j \hat{u}_\nu, \partial_t \hat{u}_\nu) + \left(\bar{\partial}\beta - \widehat{Q}_\nu(\hat{u}_\nu)\partial_t \beta\right)\partial_j \hat{u}_\nu + \beta\,\partial_j \hat{f}_\nu.$$

Combining this with (2.15) gives

$$\|\beta\,\partial_j \hat{u}_\nu\|_{W^{k,p}} \leqslant c\left\|\widehat{Q}_\nu(\hat{u}_\nu)\partial_t(\beta\,\partial_j \hat{u}_\nu)\right\|_{W^{k-1,p}} + c\left\|D\widehat{Q}_\nu(\hat{u}_\nu)(\beta\,\partial_j \hat{u}_\nu, \partial_t \hat{u}_\nu)\right\|_{W^{k-1,p}} + c\left\|\left(\bar{\partial}\beta - \widehat{Q}_\nu(\hat{u}_\nu)\partial_t \beta\right)\partial_j \hat{u}_\nu + \beta\,\partial_j \hat{f}_\nu\right\|_{W^{k-1,p}}. \tag{2.16}$$

It is important to note that the constant $c > 0$ in this expression comes from the elliptic estimate $\|g\|_{W^{k,p}} \leqslant c\|\bar{\partial} g\|_{W^{k-1,p}}$, so it is the same constant regardless of our choice of the scaling parameter ϵ. Let's look at each of the three terms on the right hand side separately.

Step 1: The third term.
We claim that the term on the second line of (2.16) satisfies a uniform bound. For the terms in this expression that only involve products of $\partial_j \hat{u}_\nu$ or $\partial_j \hat{f}_\nu$ with smooth functions, this follows immediately from the uniform $W^{k,p}$-bounds on $\hat{u}_\nu$ and $\hat{f}_\nu$. For the term involving $\widehat{Q}_\nu(\hat{u}_\nu)$ we observe that since $\widehat{Q}_\nu \to \widehat{Q}$ in C^k on $\mathbb{D}^{2n}$ and $\hat{u}_\nu$ can be assumed to lie in a $W^{k,p}$-small neighborhood of 0 for every ν, Proposition 2.2.5 places $\widehat{Q}_\nu(\hat{u}_\nu)$ into a $W^{k,p}$-small neighborhood of 0 for ν sufficiently large, meaning this term is uniformly $W^{k,p}$-bounded. Its product with $\partial_j \hat{u}_\nu$ is then uniformly $W^{k-1,p}$-bounded due to the continuous product pairing $W^{k,p} \times W^{k-1,p} \to W^{k-1,p}$ from Prop. 2.2.4.

Step 2: The first term.
The tricky aspect of the first term in (2.16) is that it involves kth derivatives of $\beta\,\partial_j \hat{u}_\nu$, which are actually what we were trying to bound in the first place. What saves the situation is the *smallness* of $\widehat{Q}_\nu(\hat{u}_\nu)$: indeed, we have seen above that this term can be assumed arbitrarily $W^{k,p}$-small as $\nu \to \infty$ if $\epsilon > 0$ is chosen sufficiently small. The continuous product pairing $W^{k,p} \times W^{k-1,p} \to W^{k-1,p}$ gives a bound

$$c\left\|\widehat{Q}_\nu(\hat{u}_\nu)\partial_t(\beta\,\partial_j \hat{u}_\nu)\right\|_{W^{k-1,p}} \leqslant c'\left\|\widehat{Q}_\nu(\hat{u}_\nu)\right\|_{W^{k,p}} \cdot \|\partial_t(\beta\,\partial_j \hat{u}_\nu)\|_{W^{k-1,p}} \leqslant c'\left\|\widehat{Q}_\nu(\hat{u}_\nu)\right\|_{W^{k,p}} \cdot \|\beta\,\partial_j \hat{u}_\nu\|_{W^{k,p}},$$

where $c' > 0$ is yet another constant that does not depend on ϵ. With this in mind, let us now choose $\epsilon > 0$ small enough to ensure

$$\left\|\widehat{Q}_\nu(\hat{u}_\nu)\right\|_{W^{k,p}} < \frac{1}{3c'}.$$

Step 3: The second term.
We observe first that $D\widehat{Q}_\nu \to D\widehat{Q}$ in C^{k-1}, and depending on whether $p > 2$ or $p \leqslant 2$, slightly different tricks can now be used to bound $D\widehat{Q}_\nu(\hat{u}_\nu)$. If $p > 2$, then $W^{k,p}$ has a continuous inclusion into C^{k-1} and we can therefore assume the $\hat{u}_\nu$ all lie in a fixed C^{k-1}-small neighborhood of 0, implying that $D\widehat{Q}_\nu(\hat{u}_\nu)$ is uniformly C^{k-1}-bounded. If on the other hand $p \leqslant 2$, then the condition $kp > 2$ requires $k \geqslant 2$,

and we can instead make use of a Sobolev embedding of the form $W^{k,p} \hookrightarrow W^{k-1,q}$. Indeed, choose any $q \in [p, \infty)$ such that the condition

$$0 < k - 1 - \frac{2}{q} \leqslant k - \frac{2}{p}$$

is satisfied; this is clearly possible since $k-1-\frac{2}{p} < k-\frac{2}{p}$ and $k-1-\frac{2}{\infty} = k-1 \geqslant k-\frac{2}{p}$ and $p \leqslant 2$. Proposition 2.2.2 now provides a continuous inclusion $W^{k,p} \hookrightarrow W^{k-1,q}$, and since $(k-1)q > 2$, there is also a continuous pairing $C^{k-1} \times W^{k-1,q} \to W^{k-1,q}$ from Proposition 2.2.5, implying that $D\widehat{Q}_\nu(\hat{u}_\nu)$ is uniformly $W^{k-1,q}$-bounded. In either case, the bounds can be assumed independent of the scaling parameter ϵ, and since both C^{k-1} and $W^{k-1,q}$ admit continuous product pairings with $W^{k-1,p}$, combining this with the product pairing $W^{k,p} \times W^{k-1,p} \to W^{k-1,p}$ then leads to a bound of the form

$$c\left\| D\widehat{Q}_\nu(\hat{u}_\nu)(\beta\, \partial_j \hat{u}_\nu, \partial_t \hat{u}_\nu)\right\|_{W^{k-1,p}} \leqslant c' \|\beta\, \partial_j \hat{u}_\nu\|_{W^{k,p}} \cdot \|\partial_t \hat{u}_\nu\|_{W^{k-1,p}}$$

for a constant $c' > 0$ that is independent of ν and ϵ. By (2.14), we can now choose $\epsilon > 0$ small enough so that

$$\|\partial_t \hat{u}_\nu\|_{W^{k-1,p}} \leqslant \|\hat{u}_\nu\|_{W^{k,p}} < \frac{1}{3c'}$$

for all ν.

Conclusion.
Combining the three estimates above for the terms on the right hand side of (2.16) now gives an inequality of the form

$$\|\beta\, \partial_j \hat{u}_\nu\|_{W^{k,p}} \leqslant c'' + \frac{2}{3}\|\beta\, \partial_j \hat{u}_\nu\|_{W^{k,p}},$$

where $c'' > 0$ is the bound obtained in step 1. We conclude $\|\beta\, \partial_j \hat{u}_\nu\|_{W^{k,p}} \leqslant 3c''$, and have thus found a uniform bound for $\|\hat{u}_\nu\|_{W^{k+1,p}(\mathring{\mathbb{D}}_r)}$. ☐

EXERCISE 2.4.14. Use an analogous argument via difference quotients to prove statement (1) in Theorem 2.4.10. *Hint: If you're anything like me, you might get stuck trying to estimate the difference quotient analogues of the terms in* (2.16) *that involve derivatives of $\widehat{Q}_\nu$. The difficulty is that this expression was derived using the chain rule for derivatives, and there is no similarly simple chain rule for difference quotients. The trick is to remember that difference quotients only differ from the corresponding derivatives by a remainder term. The remainder will produce extra terms in the difference quotient version of* (2.16)*, but the extra terms can be bounded.*

2.4.3. The nonlinear case: from $W^{1,p} \cap C^0$ to $W^{1,q}$. The proof of Gromov's removable singularity theorem in Chapter 7 will require a stronger variant of Theorem 2.4.10(1) for honest J-holomorphic curves (with no inhomogeneous term), in which the hypothesis $kp > 2$ is relaxed. In the absence of this condition, it is no longer automatic from the Sobolev embedding theorem that our weak solution $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ is continuous, but continuity will be needed in the proof, so we impose it as an explicit hypothesis. For this statement, we can also get away with allowing J to be continuous but not differentiable, though the conclusion in that case is correspondingly modest.

Theorem 2.4.15 (Nonlinear regularity, $kp \leqslant 2$ version). *Assume $1 < p < \infty$, $m \geqslant 0$ is an integer, and $J : \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$ is of class C^m. Then every weak solution $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ to the nonlinear Cauchy-Riemann equation $\partial_s u + J(u)\partial_t u = 0$ that is continuous and of class $W^{1,p}$ on $\mathring{\mathbb{D}}$ is also of class $W^{m+1,q}_{\mathrm{loc}}$ on $\mathring{\mathbb{D}}$ for every $q \in (1, \infty)$. In particular, u is of class C^m.*

In light of the bootstrapping result in the previous subsection, Theorem 2.4.15 will follow immediately if we can prove it in the case $m = 0$, where the statement is really that a solution of class $W^{1,p} \cap C^0$ is also of class $W^{1,q}_{\mathrm{loc}}$ for any $q > p$, in particular for some $q > 2$. The following lemma to that effect is adapted from an argument due to Sikorav, cf. [Sik94, Prop. 2.3.6(i)].

Lemma 2.4.16. *Assume $1 < p, q < \infty$ and J is a continuous almost complex structure on $\mathbb{C}^n$. If $u \in C^0(\mathbb{D}) \cap W^{1,p}(\mathring{\mathbb{D}})$ is a weak solution to the equation $\partial_s u + J(u)\partial_t u = 0$, then u is also of class $W^{1,q}$ on all compact subsets of $\mathring{\mathbb{D}}$.*

Proof. There is nothing to prove if $q \leqslant p$, so we assume throughout that $q > p$, and that $u : \mathbb{D} \to \mathbb{C}^n$ is in $W^{1,p} \cap C^0$ and is J-holomorphic. Given $z_0 \in \mathring{\mathbb{D}}$, we can assume after changing coodinates on $\mathbb{C}^n$ that $u(z_0) = 0$ and $J(0) = i$. As in the proof of Theorem 2.4.10, we then write $Q := i - J : \mathbb{C}^n \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$ and consider rescaled functions of the form

$$\begin{aligned} \widehat{J} : \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n), \qquad & \widehat{J}(x) := J(x/R), \\ \widehat{Q} : \mathbb{C}^n \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n), \qquad & \widehat{Q}(x) := Q(x/R) = i - \widehat{J}(x), \\ \hat{u} : \mathbb{D} \to \mathbb{C}^n, \qquad & \hat{u}(z) := Ru(z_0 + \epsilon z), \end{aligned} \tag{2.17}$$

where $\epsilon \in (0, 1]$ and $R \geqslant 1$ are constants, so that u is J-holomorphic if and only if $\hat{u}$ satisfies

$$\bar{\partial}\hat{u} - \widehat{Q}(\hat{u})\partial_t \hat{u} = 0. \tag{2.18}$$

Choosing $R \geqslant 1$ sufficiently large makes $\widehat{Q}$ arbitrarily C^0-small on the unit disk $\mathbb{D}^{2n} \subset \mathbb{C}^n$, and after fixing R in this way, we can (since u is continuous) choose $\epsilon \in (0, 1]$ sufficiently small to ensure $\hat{u}(\mathbb{D}) \subset \mathbb{D}^{2n}$. In this way, we are allowed to assume

$$\|\widehat{Q}(\hat{u})\|_{C^0(\mathbb{D})} < \delta$$

for some small constant $\delta > 0$, which can always be made smaller if necessary by adjusting R and ϵ. Consider the bounded linear operator

$$D_Q := \bar{\partial} - \widehat{Q}(\hat{u})\partial_t : W^{1,p}(\mathring{\mathbb{D}}, \mathbb{C}^n) \to L^p(\mathring{\mathbb{D}}, \mathbb{C}^n),$$

which has $\hat{u} \in \ker D_Q$ by (2.18), and observe that D_Q is close to $\bar{\partial} : W^{1,p}(\mathring{\mathbb{D}}) \to L^p(\mathring{\mathbb{D}})$ in the operator norm if δ is sufficiently small. Fix $r \in (0, 1)$ and a smooth compactly supported function $\beta \in C^\infty_0(\mathring{\mathbb{D}})$ with $\beta|_{\mathbb{D}_r} \equiv 1$. The Leibniz rule gives

$$D_Q(\beta\hat{u}) = \left(\bar{\partial}\beta - \widehat{Q}(\hat{u})\partial_t\beta\right)\hat{u} \in C^0(\mathbb{D}),$$

hence $D_Q(\beta\hat{u}) \in L^q(\mathring{\mathbb{D}})$. The rough outline of our argument will now be as follows. Recall from §2.3 that $\bar{\partial} : W^{1,p}(\mathring{\mathbb{D}}) \to L^p(\mathring{\mathbb{D}})$ has a bounded right inverse $T : L^p(\mathring{\mathbb{D}}) \to W^{1,p}(\mathring{\mathbb{D}})$, given by the convolution $Tf := K * f$ with a fundamental solution $K \in L^1_{\text{loc}}(\mathbb{C})$ to the $\bar{\partial}$-equation. We saw also via Lemma 2.3.3 that whenever f is smooth with compact support on $\mathbb{C}$, one has $f = T(\bar{\partial} f)$, so by density, the same is true for every $f \in W^{1,p}_0(\mathring{\mathbb{D}})$. Since $L^q(\mathring{\mathbb{D}}) \subset L^p(\mathring{\mathbb{D}})$ for $q > p$, the same convolution operator restricts to $L^q(\mathring{\mathbb{D}})$ as a bounded right inverse of $\bar{\partial} : W^{1,q}(\mathring{\mathbb{D}}) \to L^q(\mathring{\mathbb{D}})$, and also satisfies $T\bar{\partial}(\beta\hat{u}) = \beta\hat{u}$ since $\beta\hat{u} \in W^{1,p}_0(\mathring{\mathbb{D}})$. The fact that $D_Q : W^{1,p} \to L^p$ is close to $\bar{\partial}$ implies that it also has a bounded right inverse

$$T_Q : L^p(\mathring{\mathbb{D}}) \to W^{1,p}(\mathring{\mathbb{D}}),$$

which we expect should similarly restrict to L^q as a right inverse of $D_Q : W^{1,q} \to L^q$ and satisfy $\beta\hat{u} = T_Q D_Q(\beta\hat{u})$. If we can justify that expectation, then it implies $\beta\hat{u} \in W^{1,q}(\mathring{\mathbb{D}})$ and thus $\hat{u} \in W^{1,q}(\mathring{\mathbb{D}}_r)$, as we've already seen that $D_Q(\beta\hat{u})$ is in $L^q(\mathring{\mathbb{D}})$. The consequence for the original map $u \in W^{1,p}(\mathring{\mathbb{D}})$ will be that its restriction to a sufficiently small disk around the arbitrarily chosen point $z_0 \in \mathring{\mathbb{D}}$ is of class $W^{1,q}$.

To put this discussion on solid ground, let us write down $T_Q : L^p(\mathring{\mathbb{D}}) \to W^{1,p}(\mathring{\mathbb{D}})$ more explicitly. The relation $\bar{\partial} \circ T = \mathbb{1}$ gives

$$D_Q \circ T = \mathbb{1} - \widehat{Q}(\hat{u})\partial_t \circ T : L^p(\mathring{\mathbb{D}}) \to L^p(\mathring{\mathbb{D}}),$$

and this operator is clearly invertible if δ is sufficiently small; note that the necessary threshold for δ depends only on the norm of $T : L^p(\mathring{\mathbb{D}}) \to W^{1,p}(\mathring{\mathbb{D}})$, and not in any way on u, ϵ or R. In fact, we can also assume (possibly after shrinking δ further) that $\mathbb{1} - \widehat{Q}(\hat{u})\partial_t \circ T$ is an invertible operator on $L^q(\mathring{\mathbb{D}})$. A natural definition for T_Q is then

$$T_Q := T\left(\mathbb{1} - \widehat{Q}(\hat{u})\partial_t \circ T\right)^{-1} : L^p(\mathring{\mathbb{D}}) \to W^{1,p}(\mathring{\mathbb{D}}),$$

which has the desired property of restricting to $L^q(\mathring{\mathbb{D}})$ as a bounded right inverse of $D_Q : W^{1,q}(\mathring{\mathbb{D}}) \to L^q(\mathring{\mathbb{D}})$. Now using the relations $T\bar{\partial}(\beta\hat{u}) = \beta\hat{u}$ and $\bar{\partial}T = \mathbb{1}$, we compute,

$$\begin{aligned}
T_Q D_Q(\beta\hat{u}) &= T\left(\mathbb{1} - \widehat{Q}(\hat{u})\partial_t \circ T\right)^{-1} (\bar{\partial} - \widehat{Q}(\hat{u})\partial_t)(\beta\hat{u}) \\
&= T\left(\mathbb{1} - \widehat{Q}(\hat{u})\partial_t \circ T\right)^{-1} (\bar{\partial} - \widehat{Q}(\hat{u})\partial_t)T\bar{\partial}(\beta\hat{u}) \\
&= T\left(\mathbb{1} - \widehat{Q}(\hat{u})\partial_t \circ T\right)^{-1} \left(\bar{\partial}(\beta\hat{u}) - \widehat{Q}(\hat{u})\partial_t T\bar{\partial}(\beta\hat{u})\right) \\
&= T\left(\mathbb{1} - \widehat{Q}(\hat{u})\partial_t \circ T\right)^{-1} \left(\mathbb{1} - \widehat{Q}(\hat{u})\partial_t \circ T\right) \bar{\partial}(\beta\hat{u}) \\
&= T\bar{\partial}(\beta\hat{u}) = \beta\hat{u}.
\end{aligned}$$

This validates the argument outlined above: since $D_Q(\beta\hat{u})$ is in both L^p and L^q, $\beta\hat{u} = T_Q D_Q(\beta\hat{u})$ is in both $W^{1,p}$ and $W^{1,q}$, proving the first statement in the lemma. ☐

EXERCISE 2.4.17. Adapt the argument in the proof above to establish the following variant of Theorem 2.4.10(2b): for a C^m_{loc}-convergent sequence of almost complex structures $J_\nu \to J$ with $m \geqslant 0$, any C^0_{loc}-convergent sequence u_ν of J_ν-holomorphic curves that are also in $W^{1,p}_{\text{loc}}$ for some $p > 1$ actually converges in $W^{m+1,q}_{\text{loc}}$ for every $q \in (1, \infty)$. In particular, C^0_{loc}-convergence of u_ν implies C^m_{loc}-convergence.

REMARK 2.4.18. Why is there no variant of Theorem 2.4.10(2a) for $kp \leqslant 2$? Well, the first step in the proof of Theorem 2.4.10(2a) was to use the compactness of the Sobolev embedding $W^{k,p} \hookrightarrow C^0$ to replace u_ν with a C^0-convergent subsequence, without which the local rescaling trick in that proof would not have worked. If every $W^{1,p}$-bounded sequence similarly had a C^0-convergent subsequence when $p \leqslant 2$, then we could plug it into Exercise 2.4.17 and conclude that there are uniform $W^{1,q}$-bounds for some $q > 2$, so that Theorem 2.4.10(2a) would then apply. But indeed, the phenomenon of "bubbling" will demonstrate clearly in Chapter 7 that uniform $W^{1,2}$-bounds do not guarantee a C^0-convergent subsequence.

2.5. Linear local existence and the similarity principle

The following lemma can be applied in the case $A \in C^\infty(\mathbb{D}, \text{End}_{\mathbb{C}}(\mathbb{C}^n))$ to prove the aforementioned standard fact that complex-linear Cauchy-Riemann type operators induce holomorphic structures on vector bundles. The version with weakened regularity will be applied below to prove a useful "unique continuation" result about solutions to $(\bar\partial + A)f = 0$ in the real-linear case.

LEMMA 2.5.1. *Assume $2 < p < \infty$ and $A \in L^p(\mathring{\mathbb{D}}, \text{End}_{\mathbb{R}}(\mathbb{C}^n))$. Then for sufficiently small $\epsilon > 0$, the problem*

$$\begin{aligned} \bar\partial u + Au &= 0 \\ u(0) &= u_0 \end{aligned}$$

has a solution $u \in W^{1,p}(\mathring{\mathbb{D}}_\epsilon, \mathbb{C}^n)$.

REMARK 2.5.2. Note that $u : \mathring{\mathbb{D}}_\epsilon \to \mathbb{C}^n$ in the above statement is only a *weak* solution to $\bar\partial u + Au = 0$, as it is not necessarily differentiable, but by the Sobolev embedding theorem, it is at least continuous.

PROOF OF LEMMA 2.5.1. The main idea is that if we take $\epsilon > 0$ sufficiently small, then the restriction of $\bar\partial + A$ to $\mathring{\mathbb{D}}_\epsilon$ can be regarded as a small perturbation of $\bar\partial$ in the space of bounded linear operators $W^{1,p} \to L^p$. Since the latter has a bounded right inverse by Theorem 2.3.1, the same will be true for the perturbation.

Since $p > 2$, the Sobolev embedding theorem implies that functions $u \in W^{1,p}$ are also continuous and bounded by $\|u\|_{W^{1,p}}$, thus we can define a bounded linear operator

$$\Phi : W^{1,p}(\mathring{\mathbb{D}}) \to L^p(\mathring{\mathbb{D}}) \times \mathbb{C}^n : u \mapsto (\bar\partial u, u(0)).$$

Theorem 2.3.1 implies that this operator is also surjective and has a bounded right inverse, namely

$$L^p(\mathring{\mathbb{D}}) \times \mathbb{C}^n \to W^{1,p}(\mathring{\mathbb{D}}) : (f, u_0) \mapsto Tf - Tf(0) + u_0,$$

where $T : L^p(\mathring{\mathbb{D}}) \to W^{1,p}(\mathring{\mathbb{D}})$ is a right inverse of $\bar\partial$. Thus any operator sufficiently close to Φ in the norm topology also has a right inverse. Now define $\chi_\epsilon : \mathbb{D} \to \mathbb{R}$ to be the function that equals 1 on $\mathbb{D}_\epsilon$ and 0 outside of it, and let

$$\Phi_\epsilon : W^{1,p}(\mathring{\mathbb{D}}) \to L^p(\mathring{\mathbb{D}}) \times \mathbb{C}^n : u \mapsto ((\bar{\partial} + \chi_\epsilon A)u, u(0)).$$

To see that this is a bounded operator, it suffices to check that $W^{1,p} \to L^p : u \mapsto Au$ is bounded if $A \in L^p$; indeed,

$$\|Au\|_{L^p} \leqslant \|A\|_{L^p}\|u\|_{C^0} \leqslant c\|A\|_{L^p}\|u\|_{W^{1,p}},$$

again using the Sobolev embedding theorem. Now by this same trick, we find

$$\|\Phi_\epsilon u - \Phi u\| = \|\chi_\epsilon A u\|_{L^p(\mathring{\mathbb{D}})} \leqslant c\|A\|_{L^p(\mathring{\mathbb{D}}_\epsilon)}\|u\|_{W^{1,p}(\mathring{\mathbb{D}})},$$

thus $\|\Phi_\epsilon - \Phi\|$ is small if ϵ is small, and it follows that in this case Φ_ϵ is surjective. Our desired solution is therefore the restriction of any $u \in \Phi_\epsilon^{-1}(0, u_0)$ to $\mathring{\mathbb{D}}_\epsilon$. □

Here is a corollary, which says that every solution to a real-linear Cauchy-Riemann type equation looks locally like a holomorphic function in some *continuous* local trivialization.

THEOREM 2.5.3 (Similarity principle). *Suppose $A : \mathbb{D} \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$ is of class L^p for some $p > 2$ and $u \in W^{1,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$ is a weak solution to the equation $\bar{\partial} u + Au = 0$ with $u(0) = 0$. Then for sufficiently small $\epsilon > 0$, there exist maps $\Phi \in C^0(\mathbb{D}_\epsilon, \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n))$ and $f \in C^\infty(\mathring{\mathbb{D}}_\epsilon, \mathbb{C}^n)$ such that*

$$u(z) = \Phi(z) f(z), \qquad \bar{\partial} f = 0, \qquad \text{and} \qquad \Phi(0) = \mathbb{1}.$$

PROOF. The solution $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ is necessarily continuous and bounded, by the Sobolev embedding theorem. Choose a function $C : \mathbb{D} \to \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n)$ satisfying $C(z)u(z) = A(z)u(z)$ and $|C(z)| \leqslant |A(z)|$ for almost every $z \in \mathbb{D}$. Then $C \in L^p(\mathring{\mathbb{D}}, \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n))$ and u is a weak solution to $(\bar{\partial} + C)u = 0$. Note that even if A were assumed to be smooth, we do not yet know anything about the zero set of u and thus could not assume C is continuous, though we have no trouble achieving $C \in L^p(\mathring{\mathbb{D}})$ for some $p > 2$.

Since $\bar{\partial} + C$ is now complex linear, we can use Lemma 2.5.1 to find n weak solutions of class $W^{1,p}$ to $(\bar{\partial} + C)v = 0$ on $\mathring{\mathbb{D}}_\epsilon$ that define the standard complex basis of $\mathbb{C}^n$ at 0, and these solutions are continuous by the Sobolev embedding theorem. This gives rise to a map $\Phi \in W^{1,p}(\mathring{\mathbb{D}}_\epsilon, \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n))$ that satisfies $(\bar{\partial} + C)\Phi = 0$ in the sense of distributions and $\Phi(0) = \mathbb{1}$. Since Φ is continuous, we can assume without loss of generality that $\Phi(z)$ is invertible everywhere on $\mathring{\mathbb{D}}_\epsilon$. Setting $f := \Phi^{-1} u : \mathring{\mathbb{D}}_\epsilon \to \mathbb{C}^n$, the Leibniz rule then implies

$$0 = (\bar{\partial} + C)u = (\bar{\partial} + C)(\Phi f) = \left[(\bar{\partial} + C)\Phi\right] f + \Phi(\bar{\partial} f) = \Phi(\bar{\partial} f).$$

Note that the use of the Leibniz rule in this situation is justified by Proposition A.2.4 in light of the continuous product pairing $W^{1,p} \times W^{1,p} \to W^{1,p}$. It follows that $\bar{\partial} f = 0$, and f is smooth by Lemma 2.4.6. □

COROLLARY 2.5.4 (Unique continuation). *Suppose $\mathbf{D}$ is a linear Cauchy-Riemann type operator on a vector bundle E over a connected Riemann surface, and $\eta \in \Gamma(E)$ satisfies $\mathbf{D}\eta = 0$. Then either η is identically zero or its zeroes are isolated.* □

The similarity principle also has many nice applications for the nonlinear Cauchy-Riemann equation. Here is another "unique continuation" type result for the nonlinear case.

PROPOSITION 2.5.5. *Suppose J is a smooth almost complex structure on $\mathbb{C}^n$ and $u, v : \mathring{\mathbb{D}} \to \mathbb{C}^n$ are smooth J-holomorphic curves such that $u(0) = v(0) = 0$ and u and v have matching partial derivatives of all orders at 0. Then $u \equiv v$ on a neighborhood of 0.*

PROOF. Let $h = v - u : \mathring{\mathbb{D}} \to \mathbb{C}^n$. We have

$$\partial_s u + J(u(z))\partial_t u = 0 \tag{2.19}$$

and

$$\begin{aligned}
\partial_s v + J(u(z))\partial_t v &= \partial_s v + J(v(z))\partial_t v + [J(u(z)) - J(v(z))]\,\partial_t v \\
&= -\left[J(u(z) + h(z)) - J(u(z))\right]\partial_t v \\
&= -\left(\int_0^1 \frac{d}{d\tau} J(u(z) + \tau h(z))\, d\tau\right)\partial_t v \qquad (2.20)\\
&= -\left(\int_0^1 dJ(u(z) + \tau h(z)) \cdot h(z)\, d\tau\right)\partial_t v =: -A(z)h(z),
\end{aligned}$$

where the last step defines a smooth family of linear maps $A(z) \in \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$. Subtracting (2.19) from (2.20) gives the linear equation

$$\partial_s h(z) + \bar{J}(z)\partial_t h(z) + A(z)h(z) = 0,$$

where $\bar{J}(z) := J(u(z))$. This is a linear Cauchy-Riemann type equation on a trivial complex vector bundle over $\mathring{\mathbb{D}}$ with complex structure $\bar{J}(z)$ on the fiber at z. The similarity principle thus implies $h(z) = \Phi(z)f(z)$ near 0 for some holomorphic function $f(z) \in \mathbb{C}^n$ and some continuous map $\Phi(z) \in \mathrm{GL}(2n, \mathbb{R})$ representing a change of trivialization. Now if h has vanishing derivatives of all orders at 0, Taylor's formula implies

$$\lim_{z \to 0} \frac{|\Phi(z)f(z)|}{|z|^k} = 0$$

for all $k \in \mathbb{N}$, so f must also have a zero of infinite order and thus $f \equiv 0$. □

REMARK 2.5.6. For most applications of the similarity principle, the zeroth-order term $A : \mathbb{D} \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$ can be assumed smooth, but it is occasionally useful to know that weaker regularity hypotheses are also sufficient. One situation that arises very naturally in SFT, for instance, is when the equation $(\bar{\partial} + A)u = 0$ on $(\mathbb{D}, i)$ is derived from a similar equation on the half-cylinder $([0, \infty) \times S^1, i)$ via the biholomorphic transformation $[0, \infty) \times S^1 \to \mathbb{D} \backslash \{0\} : (s, t) \mapsto e^{-2\pi(s+it)}$, in which case the zeroth-order term is defined almost everywhere on $\mathbb{D}$ but may be unbounded near 0. In this context, the condition $A \in L^p$ with $p > 2$ in Theorem 2.5.3 becomes crucial, and the statement turns out to be false without it; see Exercise 4.8.6 for a hint on how to derive explicit counterexamples.

2.6. Simple curves and multiple covers

We now prove a global result about the structure of closed J-holomorphic curves. In Chapter 6 we will be able to generalize it in a straightforward way for punctured holomorphic curves with asymptotically cylindrical behavior.

THEOREM 2.6.1. *Assume (Σ, j) is a closed connected Riemann surface, (W, J) is a smooth almost complex manifold and $u : (\Sigma, j) \to (W, J)$ is a nonconstant pseudoholomorphic curve. Then there exists a factorization $u = v \circ \varphi$, where*

- *$\varphi : (\Sigma, j) \to (\Sigma', j')$ is a holomorphic map of positive degree to another closed and connected Riemann surface (Σ', j');*
- *$v : (\Sigma', j') \to (W, J)$ is a pseudoholomorphic curve which is embedded except at a finite set of self-intersections and non-immersed points.*[8]

Note that holomorphic maps $(\Sigma, j) \to (\Sigma', j')$ of degree 1 are always diffeomorphisms, so the factorization $u = v \circ \varphi$ in this case is just a reparametrization, and u is then called a **simple** curve. In all other cases, $k := \deg(\varphi) \geqslant 2$ and φ is in general a branched cover; we then call u a **k-fold branched cover** of the simple curve v.

The main idea in the proof is to construct Σ' (minus some punctures) explicitly as the image of u after removing finitely many singular points, so that we can take v to be the inclusion $\Sigma' \hookrightarrow W$. The map $\varphi : \Sigma \to \Sigma'$ is then uniquely determined. In order to carry out this program, we need some information on what the image of u can look like near each of its singularities. These come in two types, each type corresponding to one of the lemmas below, both of which should seem immediately plausible if your intuition comes from complex analysis.

LEMMA 2.6.2 (Intersections). *Suppose $u : (\Sigma, j) \to (W, J)$ and $v : (\Sigma', j') \to (W, J)$ are two nonconstant pseudoholomorphic curves with an intersection $u(z) = v(z')$. Then there exist neighborhoods $z \in \mathcal{U} \subset \Sigma$ and $z' \in \mathcal{U}' \subset \Sigma'$ such that*

$$\text{either} \quad u(\mathcal{U}) = v(\mathcal{U}') \quad \text{or} \quad u(\mathcal{U} \backslash \{z\}) \cap v(\mathcal{U}') = u(\mathcal{U}) \cap v(\mathcal{U}' \backslash \{z'\}) = \varnothing.$$

□

PROOF IN THE SPECIAL CASE $du(z) \neq 0$. While the proof of this lemma in full generality is somewhat involved, it becomes a simple application of the similarity principle (Theorem 2.5.3) if we additionally asume that either $du(z)$ or $dv(z')$ is nonzero. We can choose holomorphic local coordinates near $z \in \Sigma$ and $z' \in \Sigma'$ and smooth coordinates near $u(z) = v(z') \in W$ so that without loss of generality, $(\Sigma, j) = (\Sigma', j') = (\mathbb{D}, i)$ with $z = z' = 0$, $W = \mathbb{C}^n$ and $u(0) = v(0) = 0$. If $du(0) \neq 0$, then we can also arrange these coordinates so that

$$u(z) = (z, 0) \quad \text{and} \quad J(z, 0) = i;$$

[8] It follows from the Cauchy-Riemann equation that if $u : (\Sigma, j) \to (W, J)$ is J-holomorphic, then at each point $z \in \Sigma$, its first derivative $du(z) : T_z\Sigma \to T_{u(z)}W$ is either injective or trivial. We are referring to points with $du(z) = 0$ as **non-immersed points** of u. The term "critical points" is also commonly used for this condition, but is slightly at odds with the usual definition of that term when $\dim W \geqslant 4$ since, strictly speaking, every point is critical in the sense that $du(z)$ can never be surjective.

indeed, this is a simple matter of restricting u to a smaller disk on which it is an embedding, rescaling to replace the smaller disk with $\mathbb{D}$, then extending the resulting embedding to an embedding $\mathbb{D} \times \mathbb{D}^{2n-2}_\epsilon \hookrightarrow \mathbb{C}^n$ with its derivatives in the normal direction along $\mathbb{D} \times \{0\}$ specified to be complex linear. In these coordinates, for each $(z, w) \in \mathbb{C} \times \mathbb{C}^{n-1}$ we have

$$\begin{aligned} J(z,w) - i &= \int_0^1 \frac{d}{d\tau} J(z, \tau w)\, d\tau = \int_0^1 D_2 J(z, \tau w) w\, d\tau = \left(\int_0^1 D_2 J(z, \tau w)\, d\tau \right) w \\ &=: B(z,w) w, \end{aligned}$$

defining a smooth map $B : \mathbb{C}^n \to \mathrm{Hom}_{\mathbb{R}}(\mathbb{C}^{n-1}, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n))$.

Now writing $v(z) = (\varphi(z), f(z)) \in \mathbb{C} \times \mathbb{C}^{n-1}$, the nonlinear Cauchy-Riemann equation for v gives

$$0 = \partial_s v + J(v) \partial_t v = \partial_s v + i\, \partial_t v + [B(\varphi, f) f] \partial_t v,$$

and applying the projection $\pi : \mathbb{C} \times \mathbb{C}^{n-1} \to \mathbb{C}^{n-1}$ to this equation produces

$$0 = \bar{\partial} f + A f,$$

where $A : \mathbb{D} \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^{n-1})$ is a smooth map defined by

$$A(z) w := \pi [B(\varphi(z), f(z)) w] \partial_t v(z).$$

The similarity principle therefore implies that either f vanishes identically near 0 or its zero at the origin is isolated. □

LEMMA 2.6.3 (Branching). *Suppose $u : (\Sigma, j) \to (W, J)$ is a nonconstant pseudoholomorphic curve and $z_0 \in \Sigma$ is a non-immersed point of u. Then a neighborhood $\mathcal{U} \subset \Sigma$ of z_0 can be biholomorphically identified with the unit disk $\mathbb{D} \subset \mathbb{C}$ such that*

$$u(z) = v(z^k) \quad \text{for} \quad z \in \mathbb{D} = \mathcal{U},$$

where $k \in \mathbb{N}$, and $v : \mathbb{D} \to W$ is an injective J-holomorphic map with no non-immersed points except possibly at the origin. □

These two local results follow from a well-known formula of Micallef and White [MW95] describing the local behavior of J-holomorphic curves near non-immersed points and their intersections. The proof of that theorem is analytically quite involved, but one can also use an easier "approximate" version, which is proved in [Wen20, Appendix B.2] (see Remark 2.8.5 at the end of this chapter for further discussion of this). Since both are closely related to the phenomenon of unique continuation, you will not be surprised to learn that even beyond the "easy" case of Lemma 2.6.2 treated above, the similarity principle plays a role in the proof: the main idea is again to exploit the fact that locally J is always a small perturbation of i, hence the local behavior of J-holomorphic curves is also similar to the integrable case.

PROOF OF THEOREM 2.6.1. Let $\mathrm{Crit}(u) = \{z \in \Sigma \mid du(z) = 0\}$ denote the set of non-immersed points, and define $\Delta \subset \Sigma$ to be the set of all points $z \in \Sigma$ such that there exists $z' \in \Sigma$ and neighborhoods $z \in \mathcal{U} \subset \Sigma$ and $z' \in \mathcal{U}' \subset \Sigma$ with $u(z) = u(z')$ but $u(\mathcal{U} \backslash \{z\}) \cap u(\mathcal{U}' \backslash \{z'\}) = \varnothing$.

The lemmas quoted above imply that both of these sets are discrete. Both are therefore finite, and the set $\dot{\Sigma}' = u(\Sigma \backslash (\mathrm{Crit}(u) \cup \Delta)) \subset W$ is then a smooth submanifold of W with J-invariant tangent spaces, so it inherits a natural complex structure j' for which the inclusion $(\dot{\Sigma}', j') \hookrightarrow (W, J)$ is pseudoholomorphic. We shall now construct a new Riemann surface (Σ', j') from which $(\dot{\Sigma}', j')$ is obtained by removing a finite set of points. Let $\widehat{\Delta} = (\mathrm{Crit}(u) \cup \Delta)/\sim$, where two points in $\mathrm{Crit}(u) \cup \Delta$ are defined to be equivalent whenever they have neighborhoods in Σ with identical images under u. Then for each $[z] \in \widehat{\Delta}$, the branching lemma provides an injective J-holomorphic map $u_{[z]}$ from the unit disk $\mathbb{D}$ onto the image of a neighborhood of z under u. We define (Σ', j') by

$$\Sigma' = \dot{\Sigma}' \cup_\Phi \left(\coprod_{[z] \in \widehat{\Delta}} \mathbb{D} \right),$$

where the gluing map Φ is the disjoint union of the maps $u_{[z]} : \mathbb{D} \backslash \{0\} \to \dot{\Sigma}'$ for each $[z] \in \widehat{\Delta}$; since this map is holomorphic, the complex structure j' extends from $\dot{\Sigma}'$ to Σ'. Combining the maps $u_{[z]} : \mathbb{D} \to W$ with the inclusion $\dot{\Sigma}' \hookrightarrow W$ now defines a pseudoholomorphic map $v : (\Sigma', j') \to (W, J)$ which restricts to $\dot{\Sigma}'$ as an embedding and otherwise has at most finitely many non-immersed points and double points. Moreover, the restriction of u to $\Sigma \backslash (\mathrm{Crit}(u) \cup \Delta)$ defines a holomorphic map to $(\dot{\Sigma}', j')$ which extends by removal of singularities to a proper holomorphic map $\varphi : (\Sigma, j) \to (\Sigma', j')$ such that $u = v \circ \varphi$. Its holomorphicity implies that it has positive degree. □

2.7. Nonlinear local existence

Another consequence of the local regularity estimates for $\bar{\partial}$ is a nonlinear version of the local existence result in §2.5. One of its important consequences is the basic fact (originally a theorem of Gauss about conformal structures on surfaces) that all almost complex structures on a Riemann surface are integrable. In that context, we will sometimes also make use of the stability property written into Theorem 2.7.1 below: it implies that local holomorphic charts on sufficiently small regions can be perturbed smoothly under small perturbations of the complex structure.

For functions $f(s,t)$ on domains in $\mathbb{C}$ with complex coordinate $z = s + it$, it is often convenient to regard f formally as a function of the variables $z = s + it$ and $\bar{z} = s - it$, so that its partial derivatives are written as complex-linear combinations of

$$\frac{\partial f}{\partial z} := \frac{1}{2} (\partial_s - i\, \partial_t)\, f, \qquad \text{and} \qquad \frac{\partial f}{\partial \bar{z}} := \frac{1}{2} (\partial_s + i\, \partial_t)\, f.$$

Holomorphic functions are thus distinguished by the fact that since $\frac{\partial f}{\partial \bar{z}} \equiv 0$, their derivatives of all orders are fully determined by $\frac{\partial^k f}{\partial z^k}$ for $k \geqslant 0$. It is not hard to show that the latter also holds for J-holomorphic curves $u : (\mathbb{D}, i) \to (\mathbb{C}^n, J)$ at any point $z \in \mathbb{D}$ such that $J(u(z)) = i$; this follows by computing higher derivatives of the nonlinear Cauchy-Riemann equation at such a point.

THEOREM 2.7.1. *Assume J is a smooth almost complex structure on $\mathbb{C}^n$ with $J(0) = i$, and $a_1, \dots, a_m \in \mathbb{C}^n$ are constants for some $m \geqslant 0$. Then:*

(1) *For any $\epsilon > 0$ sufficiently small, there exists a J-holomorphic map $u : (\mathbb{D}_\epsilon, i) \to (\mathbb{C}^n, J)$ satisfying $u(0) = 0$ and $\frac{\partial^k u}{\partial z^k}(0) = a_k$ for each $k = 1, \dots, m$.*

(2) *Given a J-holomorphic map $u : (\mathbb{D}_r, i) \to (\mathbb{C}^n, J)$ on a disk of some radius $r > 0$ satisfying $u(0) = 0$, if $x_\nu \in \mathbb{C}^n$ is a sequence converging to 0, $J_\nu \to J$ is a C^∞_{loc}-convergent sequence of almost complex structures on $\mathbb{C}^n$ and $\epsilon > 0$ is sufficiently small, then there also exists for ν sufficiently large a sequence of J_ν-holomorphic maps $u_\nu : (\mathbb{D}_\epsilon, i) \to (\mathbb{C}^n, J_\nu)$ that satisfy $u_\nu(0) = x_\nu$ and are C^∞-convergent to $u|_{\mathbb{D}_\epsilon}$.*

REMARK 2.7.2. By an easy modification of the proof below, one could if desired also impose a converging sequence of constraints on finitely many derivatives of the sequence of maps u_ν in the second part of the statement.

REMARK 2.7.3. There is no uniqueness in Theorem 2.7.1, nor should one expect it: in the case $J \equiv i$, specifying $\frac{\partial^k u}{\partial z^k}(0)$ for all $k \geqslant 0$ up to some finite order still leaves an infinite-dimensional space of solutions to $\bar{\partial} u = 0$. On the other hand, specifying these derivatives for *all* $k \geqslant 0$ produces uniqueness but kills existence: there is a unique holomorphic Taylor series that has the correct derivatives, but it might have zero radius of convergence.

There are two main ingredients behind the proof of Theorem 2.7.1. One is the existence of a bounded right inverse to the operator $\bar{\partial} : W^{k,p}(\mathring{\mathbb{D}}) \to W^{k-1,p}(\mathring{\mathbb{D}})$ for every $k \in \mathbb{N}$ and $p \in (1, \infty)$, as provided by the fundamental elliptic estimates of §2.3 and Exercise 2.4.5. The other is the extension of standard differential calculus to functions defined on open subsets of Banach spaces, as presented e.g. in [**Lan93**, **Lan99**]. In particular, the surjectivity of $\bar{\partial}$ will be needed as a hypothesis for applying the implicit function theorem to a differentiable map between open subsets of infinite-dimensional Banach spaces, thereby proving that the zero-set of that map is a differentiable Banach submanifold. We will make considerably more use of that machinery in later chapters, typically in the context of smooth Banach manifolds and Banach space bundles (cf. §8.2). Since it is not entirely trivial in such settings to determine whether certain maps are differentiable, it will be useful to keep the following extension of the C^k-continuity property from Proposition 2.2.5 in mind:

PROPOSITION 2.7.4. *Under the same assumptions as in Proposition 2.2.5, if the open set $\Omega \subset \mathbb{R}^n$ is convex,*[9] *then the map*

$$\Phi : C^{k+r}(\Omega, \mathbb{R}^N) \times W^{k,p}(\mathcal{U}, \Omega) \to W^{k,p}(\mathcal{U}, \mathbb{R}^N) : (f, u) \mapsto f \circ u$$

is of class C^r for each $r \in \mathbb{N}$, and its first partial derivatives are given by

$$D_1\Phi(f, u)g = g \circ u, \qquad D_2\Phi(f, u)v = (Df \circ u)v,$$

where the second expression makes sense due to Propositions 2.2.4 and 2.2.5 and should be understood as the pointwise product of the two $W^{k,p}$-functions $Df \circ u : \mathcal{U} \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$ and $v : \mathcal{U} \to \mathbb{R}^n$.

[9]The convexity assumption on $\Omega \subset \mathbb{R}^n$ is inessential, and can be relaxed at the cost of more cumbersome notation, cf. the setup for Theorem A.2.6.

PROOF. The main step is to prove that for any fixed $f \in C^{k+1}(\Omega, \mathbb{R}^N)$, the map

$$\Psi_f : W^{k,p}(\mathcal{U}, \Omega) \to W^{k,p}(\mathcal{U}, \mathbb{R}^N) : u \mapsto f \circ u$$

is differentiable, and its derivative is given by

$$D\Psi_f(u) = \Psi_{Df}(u). \tag{2.21}$$

The latter is a continuous function $W^{k,p}(\mathcal{U}, \Omega) \to W^{k,p}(\mathcal{U}, \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N))$ by Proposition 2.2.5 since Df is of class C^k. If (2.21) is established, then by induction, Ψ_f is of class C^r whenever f is of class C^{k+r} for some $r \in \mathbb{N}$. The partial derivatives of $\Phi(f, u)$ with respect to f are easier to handle since it is a linear function of f, so for instance the stated formula for $D_1\Phi(f, u)$ is obviously correct and Proposition 2.2.5 implies that it is continuous. In this way, one can proceed inductively to show that all partial derivatives of Φ up to order r exist and are continuous; we will leave the details of this inductive argument as an exercise (cf. [**Wenb**, Lemma 2.12.7]). By a standard theorem in differential calculus (see [**Lan93**, Chapter XIII, Theorem 7.1]), it will follow that Φ is of class C^r.

The proof of (2.21) proceeds as follows. For $\eta \in W^{k,p}(\mathcal{U}, \mathbb{R}^n)$ sufficiently small, we can exploit the convexity of Ω to write

$$\begin{aligned}\Psi_f(u + \eta) &= \Psi_f(u) + [f \circ (u + \eta) - f \circ u] = \Psi_f(u) + \int_0^1 \frac{d}{dt} f \circ (u + t\eta)\, dt \\ &= \Psi_f(u) + \left(\int_0^1 Df \circ (u + t\eta)\, dt\right)\eta \\ &=: \Psi_f(u) + (Df \circ u)\eta + [\theta \circ (u + \eta, u)]\,\eta,\end{aligned} \tag{2.22}$$

where for the last step we define a function $\theta : \Omega \times \Omega \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)$ by

$$\theta(x, y) := \int_0^1 [Df((1 - t)y + tx) - Df(y)]\, dt.$$

This function is of class C^k since f is in C^{k+1}, and Proposition 2.2.5 thus implies that the map

$$\Psi_\theta : W^{k,p}(\mathcal{U}, \Omega \times \Omega) \to W^{k,p}(\mathcal{U}, \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^N)) : (u, v) \mapsto \theta \circ (u, v)$$

is continuous, implying in particular that $\theta \circ (u + \eta, u) = \Psi_\theta(u + \eta, u)$ is $W^{k,p}$-convergent to $\Psi_\theta(u, u) = \theta \circ (u, u) = 0$ as $\eta \to 0$ in $W^{k,p}$. This allows us to rewrite (2.22) as

$$\Psi_f(u + \eta) = \Psi_f(u) + \Psi_{Df}(u)\eta + \Psi_\theta(u + \eta, u)\eta$$

and interpret it as the definition of the derivative of Ψ_f at u, with $\Psi_\theta(u + \eta, u)\eta$ as the remainder term. ☐

PROOF OF THEOREM 2.7.1. Assume without loss of generality $m \geqslant 1$. We can apply a rescaling trick as in §2.4.2 to zoom in on a neighborhood of the origin in $\mathbb{C}^n$, which has the effect of identifying any smooth almost complex structure J on $\mathbb{C}^n$ satisfying $J(0) = i$ with one that is arbitrarily C^∞-close to the constant complex structure i on any given compact subset. For existence, it therefore suffices to prove the following claim: given any $a_1, \ldots, a_m \in \mathbb{C}^n$, one can find a radius $R > 0$ and a C^∞-small neighborhood $\mathcal{U}$ of i in the space of all smooth almost complex structures on the disk $\mathbb{D}^{2n}_R \subset \mathbb{C}^n$ of radius R, such that for every $J \in \mathcal{U}$ there exists a J-holomorphic map

$$u : (\mathbb{D}, i) \to (\mathring{\mathbb{D}}^{2n}_R, J)$$

satisfying $u(0) = 0$ and $\frac{\partial^k u}{\partial z^k}(0) = a_k$ for all $k = 1, \dots, m$. To start with, choose $R > 0$ large enough so that the unique holomorphic polynomial of degree m satisfying these conditions at the origin maps $\mathbb{D}$ into $\mathring{\mathbb{D}}^{2n}_R$; this polynomial is then a solution to the above problem for the case $J \equiv i$. Now pick any $k \in \mathbb{N}$ and $p \in (1, \infty)$ with $kp > 2$ and consider the sets

$$\mathcal{M} := \left\{ (J, u) \in C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)) \times W^{k+m,p}(\mathring{\mathbb{D}}, \mathring{\mathbb{D}}^{2n}_R) \;\middle|\; \partial_s u + J(u)\partial_t u = 0 \right\}$$

$$\mathcal{M}(J) := \left\{ u \in W^{k+m,p}(\mathring{\mathbb{D}}, \mathring{\mathbb{D}}^{2n}_R) \;\middle|\; (J, u) \in \mathcal{M} \right\} \qquad \text{for } J \in C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)).$$

We can present $\mathcal{M}$ as the zero set of the map

$$C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)) \times W^{k+m,p}(\mathring{\mathbb{D}}, \mathring{\mathbb{D}}^{2n}_R) \xrightarrow{F} W^{k+m-1,p}(\mathring{\mathbb{D}}, \mathbb{C}^n),$$
$$(J, u) \mapsto \partial_s u + (J \circ u)\partial_t u,$$

which we claim is of class C^1. Indeed, the map $C^{k+m+1} \times W^{k+m,p} \to W^{k+m,p}$: $(J, u) \mapsto J \circ u$ is in C^1 by Proposition 2.7.4, $u \mapsto \partial_s u$ and $u \mapsto \partial_t u$ are bounded linear maps $W^{k+m,p} \to W^{k+m-1,p}$ and thus smooth, and $(J \circ u, \partial_t u) \mapsto (J \circ u)\partial_t u$ is the continuous bilinear product pairing $W^{k+m,p} \times W^{k+m-1,p} \to W^{k+m-1,p}$, thus also smooth. Whenever $J \equiv i$ and $F(i, u) = 0$, the partial derivative of F with respect to u is

$$D_2 F(i, u) = \bar{\partial} : W^{k+m,p}(\mathring{\mathbb{D}}, \mathbb{C}^n) \to W^{k+m-1,p}(\mathring{\mathbb{D}}, \mathbb{C}^n),$$

which is surjective with a bounded right inverse by Exercise 2.4.5. It follows that $DF(i, u)$ is likewise surjective with a bounded right inverse, so that by the implicit function theorem, a neighborhood of $\{i\} \times \mathcal{M}(i)$ in $\mathcal{M}$ is a C^1-smooth Banach submanifold of $C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)) \times W^{k+m,p}(\mathring{\mathbb{D}}, \mathbb{D}^{2n}_R)$. Since functions of class $W^{k+m,p}$ in $\mathring{\mathbb{D}}$ with $kp > 2$ have well-defined derivatives up to order m at every point, it follows that the map

$$\mathcal{M} \xrightarrow{\pi} C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)) \times \mathbb{C}^{n(m+1)},$$
$$(J, u) \mapsto \left(J, u(0), \frac{\partial u}{\partial z}(0), \dots, \frac{\partial^m u}{\partial z^m}(0) \right)$$

is of class C^1, and we claim that it is a submersion near $\{i\} \times \mathcal{M}(i)$. Indeed, the tangent space $T_{(i,u)}\mathcal{M}$ is $\ker DF(i, u)$, which contains $\{0\} \oplus \ker D_2 F(i, u) = \{0\} \oplus \ker \bar{\partial}$, i.e. the set of all pairs $(0, f)$ such that $f \in W^{k+m,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$ is a holomorphic function. The map

$$\mathcal{M} \to C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)) : (J, u) \mapsto J$$

is a submersion near $\{i\} \times \mathcal{M}(i)$ if and only if for every $(i, u) \in \mathcal{M}$ and $Y \in C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n))$, $\ker DF(i, u)$ contains an element of the form (Y, f) for some $f \in W^{k+m,p}(\mathring{\mathbb{D}}, \mathbb{C}^n)$. This is true, since $DF(i,u)(Y,f) = D_1F(i,u)Y + D_2F(i,u)f$ and $D_2F(i,u) = \bar{\partial}$ is surjective. Now it suffices to observe that

$$\mathcal{M}(i) \to \mathbb{C}^{n(m+1)} : u \mapsto \left(u(0), \frac{\partial u}{\partial z}(0), \dots, \frac{\partial^m u}{\partial z^m}(0) \right)$$

is likewise a submersion, because one can find holomorphic functions on $\mathbb{C}$ having arbitrary values for their first m derivatives with respect to z at 0.

Since $R > 0$ was chosen to ensure that $\pi^{-1}(i, 0, a_1, \dots, a_m)$ is nonempty, the submersion property now enables us to find a neighborhood $\mathcal{U} \subset C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n))$ of i and a C^1-smooth map

$$C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)) \supset \mathcal{U} \to \mathcal{M} : J \mapsto (J, u_J)$$

such that $\pi(J, u_J) = (J, 0, a_1, \dots, a_m)$ for all $J \in \mathcal{U}$. For any $J \in \mathcal{U}$ that is also a smooth almost complex structure, the resulting map u_J will then be smooth by elliptic regularity, and the proof of existence is thus complete.

For the result about convergent sequences, the same rescaling trick means that it suffices to prove the result is true with $r = \epsilon = 1$ under the assumption that $u(\mathbb{D}) \subset \mathring{\mathbb{D}}^{2n}_R$ and J is arbitrarily C^∞-close to i on $\mathbb{D}^{2n}_R$. Since rescaling can also be used to make u arbitrarily close to $0 \in W^{k,p}(\mathbb{D}^n, \mathring{\mathbb{D}}^{2n}_R)$, let us assume in particular that $(J, u) \in \mathcal{M}$ lies in the neighborhood of $\{i\} \times \mathcal{M}(i)$ on which π is a submersion. The submersion property then implies the existence of a neighborhood $\mathcal{V} \subset C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)) \times \mathbb{D}^{2n}_R$ of $(J, 0)$ and a C^1-smooth map

$$C^{k+m+1}(\mathbb{D}^{2n}_R, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)) \times \mathbb{D}^{2n}_R \supset \mathcal{V} \to \mathcal{M} : (J', x) \mapsto (J', u_{(J',x)})$$

such that $u_{(J,0)} = u$ and $u_{(J',x)}(0) = x$ for each $(J', x) \in \mathcal{V}$. Given the sequences $J_\nu \to J$ and $x_\nu \to 0$, we can then set $u_\nu := u_{(J_\nu, x_\nu)}$; a priori this converges to u in the $W^{k,p}$-topology on $\mathbb{D}$, so by elliptic regularity, it is also C^∞-convergent on $\mathbb{D}_r$ for every $r < 1$. □

EXERCISE 2.7.5. The standard complex structure i on the cylinder $\mathbb{R} \times S^1$ is defined by $i\partial_s = \partial_t$ and $i\partial_t = -\partial_s$ in the obvious coordinates (s,t). The first-order differential operator $\bar{\partial} = \partial_s + i\partial_t$ is thus defined for complex-valued functions on $\mathbb{R} \times S^1$ or any subset thereof. Consider a compact subset of the form

$$Z := [a,b] \times S^1 \subset \mathbb{R} \times S^1$$

for real numbers $a < b$.

(a) Show that for each $k \in \mathbb{N}$ and $p \in (1, \infty)$, the operator $\bar{\partial} : W^{k,p}(Z) \to W^{k-1,p}(Z)$ is surjective with a bounded right inverse. *Hint: (Z,i) is biholomorphically equivalent to a subset of the unit disk.*
(b) Prove the following cylindrical analogue of the stability statement in Theorem 2.7.1: for any C^∞_{loc}-convergent sequence $j_\nu \to i$ of complex structures on $\mathbb{R} \times S^1$, there exists for large ν a sequence of holomorphic embeddings $(Z, i) \hookrightarrow (\mathbb{R} \times S^1, j_\nu)$ that is C^∞-convergent to the obvious inclusion $Z \hookrightarrow \mathbb{R} \times S^1$.

2.8. The nonlinear equation on push-offs

In §2.1 we derived the linearized Cauchy-Riemann operator $\mathbf{D}_u : \Gamma(u^*TW) \to \Omega^{0,1}(\Sigma, u^*TW)$ for a J-holomorphic curve $u : (\Sigma, j) \to (W, J)$ by linearizing the nonlinear operator $\bar{\partial}_J u := du + J(u) \circ du \circ j$ at u, where $\bar{\partial}_J$ is imagined as a section of a vector bundle $\mathcal{E} \to \mathcal{B}$ over an infinite-dimensional manifold $\mathcal{B}$ consisting of maps $\Sigma \to W$. In the big picture, $\mathbf{D}_u$ is not just a first-order *approximation* of $\bar{\partial}_J$; we will show in this section that for nearby maps of the form $u' = \exp_u(\eta) : \Sigma \to W$ with $\eta \in \Gamma(u^*TW)$ sufficiently small, the nonlinear equation $\bar{\partial}_J u' = 0$ implies a corresponding linear equation $\mathbf{D}'_u \eta = 0$ for some Cauchy-Riemann type operator $\mathbf{D}'_u$ that is a small perturbation of $\mathbf{D}_u$. This is useful for the following reason: we will devote considerable effort in the next few chapters to studying the properties of linear Cauchy-Riemann type equations and their solutions. The ability to rewrite $\bar{\partial}_J u' = 0$ as $\mathbf{D}'_u \eta = 0$ means that many of the linear results we prove imply corresponding results for the nonlinear equation.

There is a basic idea from first-year analysis in the background: if $f : \mathcal{U} \to \mathbb{R}^m$ is a smooth map on some open domain $\mathcal{U} \subset \mathbb{R}^n$ and $f(x) = 0$ for some $x \in \mathcal{U}$, then writing $\mathbf{D}_x := df(x) : \mathbb{R}^n \to \mathbb{R}^m$ for its derivative and taking $h \in \mathbb{R}^n$ sufficiently small, one has

$$f(x+h) = \int_0^1 \frac{d}{dt} f(x+th)\,dt = \int_0^1 df(x+th)h\,dt = \left(\int_0^1 df(x+th)\,dt\right) h =: \mathbf{D}'_x h,$$

where the integral in parentheses defines a linear operator $\mathbf{D}'_x : \mathbb{R}^n \to \mathbb{R}^m$ that can be assumed arbitrarily close to $\mathbf{D}_x = df(x) = \int_0^1 df(x)\,dt$ if h is sufficiently small. A slightly subtle point here is that the definition of the map $\mathbf{D}'_x : \mathbb{R}^n \to \mathbb{R}^m$ also depends on h, but in most applications this is immaterial, because we are interested in drawing conclusions about nearby solutions $x + h$ to the nonlinear equation $f(x + h) = 0$ from general theorems about solutions to linear equations of the form $\mathbf{D}'_x h = 0$, and h is such a solution. But before we can carry out this type of computation for the nonlinear section $\bar{\partial}_J : \mathcal{B} \to \mathcal{E}$ and its linearization $\mathbf{D}_u$ at $u \in \bar{\partial}_J^{-1}(0)$, we have two problems: first, f in the computation above was a function valued in a single vector space, not a section of a vector bundle, and if it had been the latter, we would at least have needed to choose a connection to identify all the fibers in order for the computation to make sense. There is a more serious problem, however, that is unique to our infinite-dimensional setting: one could use a similarly general argument (with the aid of a connection) to rewrite $\bar{\partial}_J(\exp_u \eta)$ as $\mathbf{D}'_u \eta$ for some linear operator $\mathbf{D}'_u$, but from this perspective, it would not be obvious whether $\mathbf{D}'_u$ is also a *Cauchy-Riemann* type operator. That is something we will need to know, because the linear results proved in the next few chapters are valid specifically for Cauchy-Riemann type operators, and not necessarily for arbitrary small perturbations of them in the space of *all* bounded linear operators. The secret is to apply the integration trick used above to the finite-dimensional geometric data in the Cauchy-Riemann equation, rather than applying it directly to the infinite-dimensional section $\bar{\partial}_J$.

To set up the first result, suppose $u : (\Sigma, j) \to (W, J)$ is a J-holomorphic curve and write

$$E := u^*TW,$$

so E is a complex vector bundle over Σ with complex structure $J(u(z))$ at $z \in \Sigma$. It will be convenient in the following to write elements of the total space of vector bundles such as E as pairs (z, X) where $z \in \Sigma$ and X belongs to the fiber $E_z = T_{u(z)}W$, thus the zero-section in E consists of all points of the form $(z, 0) \in E$, and at any such point there is a canonical isomorphism

$$T_{(z,0)}E = T_z\Sigma \oplus E_z, \tag{2.23}$$

the first factor being the tangent space to the zero-section, and the second the vertical subspace of $T_{(z,0)}E$. Suppose $\mathcal{O} \subset u^*TW$ is a fiberwise-convex neighborhood of the zero-section, and

$$\Psi : \mathcal{O} \to W$$

is a smooth map whose restriction to the zero-section is u and whose derivative there restricts to the second factor in (2.23) as the identity map $E_z \to T_{u(z)}W$. One obvious way to define Ψ is as $\Psi(z, X) = \exp_{u(z)} X$ for a choice of connection on W, but the actual definition will be irrelevant in what follows. We will denote the set of sections of E with image in $\mathcal{O}$ by

$$\Gamma(\mathcal{O}) := \left\{ \eta \in \Gamma(E) \;\middle|\; (z, \eta(z)) \in \mathcal{O} \text{ for all } z \in \Sigma \right\}.$$

THEOREM 2.8.1. *Given a compact Riemann surface (Σ, j), a J-holomorphic curve $u : (\Sigma, j) \to (W, J)$ and a map $\Psi : \mathcal{O} \to W$ as described above, after possibly shrinking the neighborhood $\mathcal{O} \subset u^*TW$ of the zero-section, one can associate to each J-holomorphic curve $u' : (\Sigma, j) \to (W, J)$ of the form*

$$u'(z) = \Psi(z, \eta(z)), \qquad \eta \in \Gamma(\mathcal{O}) \subset \Gamma(u^*TW)$$

a linear Cauchy-Riemann type operator $\mathbf{D}$ on u^*TW such that $\mathbf{D}\eta = 0$. Moreover, if $\eta_k \to 0$ is a C^∞-convergent sequence of sections in $\Gamma(\mathcal{O})$ such that the maps $u'_k(z) := \Psi(z, \eta_k(z))$ are J-holomorphic curves $u'_k : (\Sigma, j) \to (W, J)$ for all k, then the associated linear Cauchy-Riemann type operators $\mathbf{D}_k$ satisfying $\mathbf{D}_k\eta_k = 0$ are also C^∞-convergent, with $\mathbf{D}_k \to \mathbf{D}_u$.

It is perhaps worth emphasizing what Theorem 2.8.1 does *not* say: there is no single linear operator $\mathbf{D}$ that makes the equations $\bar{\partial}_J u' = 0$ and $\mathbf{D}\eta = 0$ equivalent for *all* maps of the form $u'(z) = (z, \eta(z))$ with $\eta \in \Gamma(\mathcal{O})$. Instead, the operator $\mathbf{D}$ in this statement is determined by the specific solution u', and other nearby J-holomorphic curves of the form $u'_1(z) = (z, \eta_1(z))$ will not need to satisfy $\mathbf{D}\eta_1 = 0$. The point of this result is rather that we can deduce properties of the specific solution u' from the properties of solutions to the linear equation $\mathbf{D}\eta = 0$.

Before launching into the proof, we state a slightly more elaborate variant that will also come in useful. The idea is to consider a more general class of curves with images near that of u, written in the form

$$u' : (\Sigma', j') \to (W, J), \qquad u'(z) = \Psi(\varphi(z), \eta(z)),$$

where (Σ', j') is another Riemann surface, $\varphi : \Sigma' \to \Sigma$ is a smooth map that accounts for deviations of u' from u in directions tangential to u, and η is a vector field along $u \circ \varphi$ that points in directions normal to u. To make this more precise, we assume u^*TW is endowed with a splitting

$$u^*TW = T_u \oplus N_u$$

of complex vector bundles, where $T_u \subset u^*TW$ is a line bundle such that

$$\operatorname{im} du(z) \subset (T_u)_z \qquad \text{for all } z \in \Sigma.$$

If u is an immersion, then the bundle $T_u \subset u^*TW$ obviously exists and is uniquely determined by this condition; we will show in Chapter 15 that this is in fact true for *every* locally nonconstant J-holomorphic curve, even if $du(z)$ vanishes at isolated points. For now, we shall just assume the splitting exists, and refer to the complementary complex subbundle $N_u \subset u^*TW$ as the **normal bundle** of u. By Exercise 2.1.5, writing the linearized Cauchy-Riemann operator $\mathbf{D}_u$ in block form

$$\mathbf{D}_u = \begin{pmatrix} \mathbf{D}_u^T & \mathbf{D}_u^{TN} \\ \mathbf{D}_u^{NT} & \mathbf{D}_u^N \end{pmatrix} : \Gamma(T_u) \oplus \Gamma(N_u) \to \Omega^{0,1}(\Sigma, T_u) \oplus \Omega^{0,1}(\Sigma, N_u)$$

with respect to this splitting gives rise to linear Cauchy-Riemann type operators $\mathbf{D}_u^T$ and $\mathbf{D}_u^N$ on T_u and N_u respectively.

DEFINITION 2.8.2. The operator $\mathbf{D}_u^N : \Gamma(N_u) \to \Omega^{0,1}(\Sigma, N_u)$ described above is called the **normal Cauchy-Riemann operator** of u.

Given the neighborhood $\mathcal{O} \subset u^*TW$ of the zero-section in the statement of Theorem 2.8.1, let us denote the resulting neighborhood in the total space of the normal bundle by

$$\mathcal{O}^N := \mathcal{O} \cap N_u \subset N_u,$$

and also define pullbacks with respect to a smooth map $\varphi : \Sigma' \to \Sigma$ by

$$\varphi^*\mathcal{O} := \{(z, X) \mid z \in \Sigma' \text{ and } (\varphi(z), X) \in \mathcal{O}\} \subset \varphi^*u^*TW,$$
$$\varphi^*\mathcal{O}^N := \{(z, X) \mid z \in \Sigma' \text{ and } (\varphi(z), X) \in \mathcal{O}^N\} \subset \varphi^*N_u.$$

These give rise to corresponding sets of sections $\Gamma(\varphi^*\mathcal{O}) \subset \Gamma(\varphi^*u^*TW)$, $\Gamma(\varphi^*\mathcal{O}^N) \subset \Gamma(\varphi^*N_u)$. We recall from Exercise 2.1.4 the notion of the pullback $\varphi^*\mathbf{D}$ of a linear Cauchy-Riemann type operator $\mathbf{D}$ via a holomorphic map φ of Riemann surfaces.

THEOREM 2.8.3. *Given compact Riemann surfaces (Σ, j) and (Σ', j'), a nonconstant J-holomorphic curve $u : (\Sigma, j) \to (W, J)$, and a map $\Psi : \mathcal{O} \to W$ and splitting $u^*TW = T_u \oplus N_u$ as described above, after possibly shrinking the neighborhood $\mathcal{O} \subset u^*TW$ of the zero-section, one can associate to any J-holomorphic curve $u' : (\Sigma', j') \to (W, J)$ of the form*

$$u'(z) = \Psi(\varphi(z), \eta(z)), \qquad \varphi \in C^\infty(\Sigma', \Sigma), \ \eta \in \Gamma(\varphi^*\mathcal{O}^N) \subset \Gamma(\varphi^*N_u)$$

a linear Cauchy-Riemann type operator $\mathbf{D}^N$ on φ^*N_u such that $\mathbf{D}^N\eta = 0$. Moreover, this association has the following continuity property: suppose we are given C^∞-convergent sequences of

- *complex structures $j'_k \to j'$ on Σ',*

- *smooth maps $\varphi_k \to \varphi$ from Σ' to Σ, and*
- *sections $\eta_k \to 0$ of $\varphi_k^* E$,*

such that $\varphi : (\Sigma', j') \to (\Sigma, j)$ is holomorphic and $u_k'(z) := \Psi(\varphi_k(z), \eta_k(z))$ defines a sequence of J-holomorphic curves $u_k' : (\Sigma', j_k') \to (W, J)$. Then the associated linear Cauchy-Riemann type operators $\mathbf{D}_k^N$ on $\varphi_k^* N_u$ over (Σ', j_k') satisfying $\mathbf{D}_k^N \eta_k = 0$ are also C^∞-convergent, with $\mathbf{D}_k^N \to \varphi^* \mathbf{D}^N$.

REMARK 2.8.4. Theorems 2.8.1 and 2.8.3 can also be applied in some situations where Σ and Σ' are not compact; notably, we will later use them in cases where these are half-cylinders of the form $([R, \infty) \times S^1, i)$. A crucial detail in that setting is that the ambient almost complex structure is translation-invariant, and therefore satisfies global C^∞-bounds along the image of u, so that any quantitative convergence estimate on a domain of the form $[N - r, N + r] \times S^1 \subset [R, \infty) \times S^1$ becomes equally valid for any N, and is therefore valid on the entire half-cylinder.

There is a common setup for the proofs of Theorems 2.8.1 and 2.8.3. We continue to abbreviate $E := u^*TW$ where $u : (\Sigma, j) \to (W, J)$ is a J-holomorphic curve, and we assume $\mathcal{O} \subset E$ is a fiberwise-convex neighborhood of the zero-section $\Sigma \subset E$, $\Psi : \mathcal{O} \to W$ is a smooth map satisfying $\Psi|_\Sigma = u$ as described in the paragraph preceding Theorem 2.8.1, (Σ', j') is a Riemann surface, $\varphi : \Sigma' \to \Sigma$ is a smooth map, η is a section of $\varphi^* E$ such that $(\varphi(z), \eta(z)) \in \mathcal{O}$ for every $z \in \Sigma'$, and $u' : \Sigma' \to W$ is the map $u'(z) = \Psi(\varphi(z), \eta(z))$. Choose a linear connection ∇ on E; this extends (2.23) to a splitting

$$T_{(z,v)} E \cong T_z\Sigma \oplus E_z = T_z\Sigma \oplus T_{u(z)} W \tag{2.24}$$

at every point $(z, v) \in E$, where the factor $T_z\Sigma$ corresponds to the horizontal subspace and E_z to the vertical subspace. We shall always use this splitting when talking about tangent vectors to the total space of E, so for instance, the derivative of a path $\gamma(t) = (x(t), v(t)) \in E$ is now given by

$$\dot\gamma(t) = (\dot x(t), \nabla_t v(t)) \in T_{x(t)}\Sigma \oplus E_{x(t)} = T_{\gamma(t)} E.$$

We can now write the derivative of $\Psi : \mathcal{O} \to W$ at a point (z, X) in block form as

$$d\Psi(z, X) = \begin{pmatrix} \alpha(z, X) & \beta(z, X) \end{pmatrix} : T_z\Sigma \oplus E_z \to T_{\Psi(z,X)} W,$$

and observe that the stated assumptions on Ψ imply

$$\alpha(z, 0) = du(z) : T_z\Sigma \to E_z, \qquad \text{and} \qquad \beta(z, 0) = \mathbb{1} : E_z \to E_z.$$

We shall assume for the rest of the argument that the neighborhood $\mathcal{O} \subset E$ is small enough so that $\beta(z, X)$ is invertible for every $(z, X) \in \mathcal{O}$. In this case, we can define another smooth function F on $\mathcal{O}$ by

$$F(z, X) := \beta(z, X)^{-1} \circ \alpha(z, X) \in \mathrm{Hom}_{\mathbb{R}}(T_z\Sigma, E_z),$$

and it satisfies $F(z, 0) = du(z)$. This function—strictly speaking, it is a section of some vector bundle—is one of several we shall encounter with the property that for each fixed $z \in \Sigma$, the function on $\mathcal{O}_z := \mathcal{O} \cap E_z$ defined by $v \mapsto F(z, v)$ takes values in a fixed vector space, in this particular case $\mathrm{Hom}_{\mathbb{R}}(T_z\Sigma, E_z)$. Whenever this happens, the convexity of $\mathcal{O}_z$ allows us to apply the fundamental theorem of calculus and write

$$\begin{aligned} F(z,X) &= F(z,0) + \int_0^1 \frac{d}{d\tau} F(z,\tau X)\, d\tau = F(z,0) + \int_0^1 D_2F(z,\tau X) X\, d\tau \\ &=: F(z,0) + F'(z,X)X, \end{aligned}$$

where we have defined a new smooth function F' on $\mathcal{O}$ by

$$F'(z,X) := \int_0^1 D_2F(z,\tau X)\, d\tau \in \operatorname{Hom}_{\mathbb{R}}\left(E_z, \operatorname{Hom}_{\mathbb{R}}(T_z\Sigma, E_z)\right),$$

so in particular, $F'(z,0) = D_2F(z,0)$ is the derivative of F in vertical directions at a point in the zero-section. In this particular example, the result is the formula

$$F(z,X) = du(z) + F'(z,X)X.$$

Here is another important example in which this trick can be applied: we can smoothly associate to each $(z,X) \in \mathcal{O}$ a complex structure on E_z defined by

$$\widehat{J}(z,X) := \beta(z,X)^{-1} \circ J(\Psi(z,X)) \circ \beta(z,X) \in \operatorname{End}_{\mathbb{R}}(E_z),$$

and since $\widehat{J}(z,0) = J(u(z))$, we then have

$$\widehat{J}(z,X) = J(u(z)) + \widehat{J}'(z,X)X,$$

where $\widehat{J}'(z,0)$ is the vertical derivative of $\widehat{J}$ at a point in the zero-section. We will use analogous notation in some other examples below.

Applying the nonlinear Cauchy-Riemann operator to the map $u'(z) = (\varphi(z), \eta(z))$ now gives

(2.25)

$$\begin{aligned} \bar\partial_J u' &= du' + J(u') \circ du' \circ j' \\ &= d\Psi(\varphi,\eta) \circ d(\varphi,\eta) + J(u') \circ d\Psi(\varphi,\eta) \circ d(\varphi,\eta) \circ j' \\ &= \begin{pmatrix} \alpha(\varphi,\eta) & \beta(\varphi,\eta) \end{pmatrix} \begin{pmatrix} d\varphi \\ \nabla\eta \end{pmatrix} + J(u') \begin{pmatrix} \alpha(\varphi,\eta) & \beta(\varphi,\eta) \end{pmatrix} \begin{pmatrix} d\varphi \circ j' \\ \nabla\eta \circ j' \end{pmatrix} \\ &= \alpha(\varphi,\eta) \circ d\varphi + \beta(\varphi,\eta) \circ \nabla\eta + J(u') \circ \alpha(\varphi,\eta) \circ d\varphi \circ j' \\ &\qquad + J(u') \circ \beta(\varphi,\eta) \circ \nabla\eta \circ j' \\ &= \beta(\varphi,\eta) \circ \Big[\nabla\eta + \widehat{J}(\varphi,\eta) \circ \nabla\eta \circ j' \\ &\qquad + F(\varphi,\eta) \circ d\varphi + \widehat{J}(\varphi,\eta) \circ F(\varphi,\eta) \circ d\varphi \circ j' \Big], \\ &= \beta(\varphi,\eta) \circ \Big[\nabla\eta + J(u\circ\varphi) \circ \nabla\eta \circ j' \\ &\qquad + F(\varphi,\eta) \circ d\varphi + \widehat{J}(\varphi,\eta) \circ F(\varphi,\eta) \circ d\varphi \circ j' + \left(\widehat{J}'(\varphi,\eta)\eta\right) \circ \nabla\eta \circ j' \Big]. \end{aligned}$$

Since $\beta(\varphi,\eta)$ is everywhere invertible, this implies that u' is J-holomorphic if and only if the expression in square brackets vanishes. That expression is a section of the fixed vector bundle $\operatorname{Hom}_{\mathbb{R}}(T\Sigma', \varphi^*E)$ for every choice of section η. If we now pick $(\Sigma', j') = (\Sigma, j)$ and $\varphi = \mathrm{Id}$ and choose a section $\eta \in \Gamma(E)$, then since $\beta(z,0) \equiv \mathbb{1}$ for all $z \in \Sigma$, plugging in the maps $u_\rho(z) := (z, \rho\eta(z))$ for $\rho \in (-\epsilon, \epsilon)$ and differentiating with respect to ρ at $\rho = 0$ gives a formula for $\mathbf{D}_u\eta$, namely

$$
\begin{aligned}
\mathbf{D}_u\eta = \partial_\rho \Big[&\nabla(\rho\eta) + J(u) \circ \nabla(\rho\eta) \circ j + F(\cdot, \rho\eta) + \widehat{J}(\cdot, \rho\eta) \circ F(\cdot, \rho\eta) \circ j \\
&+ \left(\widehat{J}'(\cdot, \rho\eta)\rho\eta\right) \circ \nabla(\rho\eta) \circ j \Big]\Big|_{\rho=0} \\
= &\nabla\eta + J(u) \circ \nabla\eta \circ j + F'(\cdot, 0)\eta + J(u) \circ \big(F'(\cdot, 0)\eta\big) \circ j + \left(\widehat{J}'(\cdot, 0)\eta\right) \circ du \circ j.
\end{aligned}
\tag{2.26}
$$

We would now like to interpret the expression in square brackets at the end of (2.25) as a linear Cauchy-Riemann type operator acting on $\eta \in \Gamma(\varphi^*E)$. We can abbreviate the whole expression as

$$\mathbf{D}_0\eta + \widetilde{C}(\cdot, \eta)$$

by defining the Cauchy-Riemann type operator $\mathbf{D}_0\eta := \nabla\eta + J(u \circ \varphi) \circ \nabla\eta \circ j'$ and the function on $\varphi^*\mathcal{O} \subset \varphi^*E$ given by

$$
\begin{aligned}
\widetilde{C}(z, X) := &F(\varphi(z), X) \circ d\varphi(z) + \widehat{J}(\varphi(z), X) \circ F(\varphi(z), X) \circ d\varphi(z) \circ j'(z) \\
&+ \left(\widehat{J}'(\varphi(z), X)X\right) \circ \nabla\eta(z) \circ j'(z) \in \operatorname{Hom}_{\mathbb{R}}(T_z\Sigma', (\varphi^*E)_z).
\end{aligned}
\tag{2.27}
$$

While the values of $\widetilde{C}(z, X)$ according to this definition are real-linear maps in general, the particular values $\widetilde{C}(z, \eta(z))$ are guaranteed to be complex antilinear if $\bar{\partial}_J u' = 0$, because in this case $\mathbf{D}_0\eta + \widetilde{C}(\cdot, \eta) = 0$, where $\mathbf{D}_0\eta$ is a section of $\overline{\operatorname{Hom}}_{\mathbb{C}}(T\Sigma', \varphi^*E)$. It follows that if we now replace $\widetilde{C}$ with its complex-antilinear part

$$C(z, X) := \frac{1}{2}\left[\widetilde{C}(z, X) + J(u(\varphi(z))) \circ \widetilde{C}(z, X) \circ j'(z)\right] \in \overline{\operatorname{Hom}}_{\mathbb{C}}(T_z\Sigma', (\varphi^*E)_z),$$

then it is still true that $\bar{\partial}_J u' = 0$ if and only if $\mathbf{D}_0\eta + C(\cdot, \eta) = 0$. If we could now prove $C(z, 0) = 0$ for all $z \in \Sigma'$, then the usual integration trick would allow us to write $C(z, X) = C'(z, X)X$ and thus define $A(z)X := C'(z, \eta(z))X$ as a linear zeroth-order term making $\mathbf{D}_0 + A$ into a linear Cauchy-Riemann type operator that annihilates η. This will not always work, but it works in two special cases that are relevant for Theorems 2.8.1 and 2.8.3.

Focusing for the moment on Theorem 2.8.1, let us assume $(\Sigma', j') = (\Sigma, j)$ and φ is the identity map, so $\widetilde{C}$ is now a function on $\mathcal{O}$, and its definition simplifies to

$$\widetilde{C}(z, X) = F(z, X) + \widehat{J}(z, X) \circ F(z, X) \circ j(z) + \left(\widehat{J}'(z, X)X\right) \circ \nabla\eta(z) \circ j(z).$$

Since $F(z, 0) = du(z)$ is complex linear and $\widehat{J}(z, 0) = J(u(z))$, we have

$$\widetilde{C}(z, 0) = du(z) + J(u(z)) \circ du(z) \circ j(z) = 0,$$

implying that $C(z, 0)$ vanishes as well, thus we can write

$$C(z, X) = C'(z, X)X \qquad \text{for} \qquad C'(z, X) := \int_0^1 D_2C(z, \tau X)\, d\tau$$

and define $A(z) := C'(z, \eta(z))$, so that whenever $\bar{\partial}_J u' = 0$, the section $\eta \in \Gamma(\mathcal{O})$ must satisfy the linear Cauchy-Riemann type equation

$$\nabla\eta + J(u) \circ \nabla\eta \circ j + A\eta = 0.$$

Suppose now that $\eta_k \in \Gamma(\mathcal{O})$ is a sequence of sections converging in C^∞ to 0 such that $\bar{\partial} u'_k = 0$ for $u'_k(z) := (z, \eta_k(z))$. Carrying out the construction above then gives a sequence of operators of the form $\mathbf{D}_k := \mathbf{D}_0 + A_k$, where $A_k(z) = C'_k(z, \eta_k(z))$, $C'_k(z, X) = \int_0^1 D_2 C_k(z, \tau X)\, d\tau$, C_k is the complex-antilinear part of $\widetilde{C}_k$, and $\widetilde{C}_k$ is given by

$$\widetilde{C}_k(z, X) = F(z, X) + \widehat{J}(z, X) \circ F(z, X) \circ j(z) + \left(\widehat{J}'(z, X)X\right) \circ \nabla\eta_k(z) \circ j(z).$$

As $\eta_k \to 0$, the latter converges to

$$\widetilde{C}_\infty(z, X) := F(z, X) + \widehat{J}(z, X) \circ F(z, X) \circ j(z),$$

so C_k converges to the complex-antilinear part C_∞ or $\widetilde{C}_\infty$ and A_k converges to $C'_\infty(\cdot, 0)$, which is just the complex-antilinear part of $\widetilde{C}'_\infty(\cdot, 0)$. For the latter, we have

$$\begin{aligned}\widetilde{C}'_\infty(z, 0)X &= F'(z, 0)X + \widehat{J}(z, 0) \circ \left(F'(z, 0)X\right) \circ j(z) + \left(\widehat{J}'(z, 0)X\right) \circ F(z, 0) \circ j(z)\\ &= F'(z, 0)X + J(u(z)) \circ \left(F'(z, 0)X\right) \circ j(z) + \left(\widehat{J}'(z, 0)X\right) \circ du(z) \circ j(z),\end{aligned}$$

which is precisely the zeroth-order term that appears in our formula (2.26) for $\mathbf{D}_u$. This proves that $\widetilde{C}'_\infty(z, 0)X$ is already complex antilinear, and thus matches $C'_\infty(z, 0)X$, and it follows that our sequence of Cauchy-Riemann type operators $\mathbf{D}_k$ converges to $\mathbf{D}_u$. The proof of Theorem 2.8.1 is thus complete.

For the situation in Theorem 2.8.3, we are given a splitting $u^*TW = T_u \oplus N_u$ and can choose the connection ∇ to respect it, in which case the operator $\mathbf{D}_0\eta = \nabla\eta + J(u \circ \varphi) \circ \nabla\eta \circ j'$ splits into a direct sum of two linear Cauchy-Riemann type operators $\mathbf{D}_0^T$ and $\mathbf{D}_0^N$ on $\varphi^* T_u$ and $\varphi^* N_u$ respectively. Taking $\eta \in \Gamma(\varphi^*\mathcal{O}^N) \subset \Gamma(\varphi^* N_u)$ and writing $\pi_N : u^*TW \to N_u$ for the fiberwise-linear projection along T_u, the vanishing of the expression in square brackets at the end of (2.25) then implies that

$$\mathbf{D}_0^N + \pi_N C(\cdot, \eta) = 0.$$

Using the relation $d\varphi \circ j' = j(\varphi) \circ (d\varphi - \bar{\partial}_j \varphi)$, (2.27) becomes

$$\begin{aligned}\widetilde{C}(z, X) = {} & F(\varphi(z), X) \circ d\varphi(z) + \widehat{J}(\varphi(z), X) \circ F(\varphi(z), X) \circ j(\varphi(z)) \circ d\varphi(z)\\ & - \widehat{J}(\varphi(z), X) \circ F(\varphi(z), X) \circ j(\varphi(z)) \circ \bar{\partial}_j \varphi(z)\\ & + \left(\widehat{J}'(\varphi(z), X)X\right) \circ \nabla\eta(z) \circ j'(z),\end{aligned} \tag{2.28}$$

which satisfies

$$\begin{aligned}\widetilde{C}(z, 0) &= du(\varphi(z)) \circ d\varphi(z) + J(u(\varphi(z))) \circ du(\varphi(z)) \circ j(\varphi(z)) \circ d\varphi(z)\\ &\qquad - J(u(\varphi(z))) \circ du(\varphi(z)) \circ j(\varphi(z)) \circ \bar{\partial}_j \varphi(z)\\ &= du(\varphi(z)) \circ \bar{\partial}_j \varphi(z).\end{aligned}$$

The expression is complex antilinear, so $C(z,0)$ is exactly the same, and as luck would have it, its image is contained in the subbundle φ^*T_u, thus composing it with the projection π_N kills it. We conclude that $\pi_N C(\cdot,\eta)$ can be written in the form $\pi_N C'(\cdot,\eta)\eta =: A^N\eta$, defining a linear Cauchy-Riemann type operator $\mathbf{D}^N := \mathbf{D}_0^N + A^N$ on φ^*N_u with $A^N(z) := \pi_N C'(z,\eta(z))$ such that $\mathbf{D}^N\eta = 0$.

Now suppose we have C^∞-convergent sequences $j'_k \to j'$, $\varphi_k \to \varphi$ and $\eta_k \to 0$, where $\varphi_k : \Sigma' \to \Sigma$ are smooth maps, $\varphi : (\Sigma', j') \to (\Sigma, j)$ is holomorphic, $\eta_k \in \Gamma(\varphi_k^*\mathcal{O}^N) \subset \Gamma(\varphi_k^* N_u)$ and $u'_k(z) := \Psi(\varphi_k(z), \eta_k(z))$ defines a J-holomorphic curve $(\Sigma', j'_k) \to (W, J)$ for every k. The fact that φ is holomorphic implies

$$\bar\partial_{j'_k,j}\varphi_k := d\varphi_k + j(\varphi_k)\circ d\varphi_k \circ j'_k \to 0$$

in C^∞. For each k, the construction above now gives a linear Cauchy-Riemann type operator $\mathbf{D}_k^N$ on $\varphi_k^* N_u$ that annihilates η_k, given by the formula

$$\mathbf{D}_k^N\eta = \nabla\eta + J(u(\varphi_k))\circ\nabla\eta\circ j'_k + A_k^N\eta,$$

where $A_k^N(z) = \pi_N C'_k(z,\eta_k)$ and C_k is the complex-antilinear part of $\widetilde{C}_k$, which according to (2.28) satisfies

$$\begin{aligned}\widetilde{C}_k(z,X) = F(\varphi_k(z),X)\circ d\varphi_k(z) + \widehat{J}(\varphi_k(z),X)\circ F(\varphi_k(z),X)\circ j(\varphi_k(z))\circ d\varphi_k(z)\\ - \widehat{J}(\varphi_k(z),X)\circ F(\varphi_k(z),X)\circ j(\varphi_k(z))\circ\bar\partial_{j'_k,j}\varphi_k(z)\\ + \left(\widehat{J}'(\varphi_k(z),X)X\right)\circ\nabla\eta_k(z)\circ j'_k(z).\end{aligned}$$

The parts involving $\bar\partial_{j'_k,j}\varphi_k$ and $\nabla\eta_k$ disappear as $k\to\infty$, leaving

$$\widetilde{C}_\infty(z,X) := F(\varphi(z),X)\circ d\varphi(z) + \widehat{J}(\varphi(z),X)\circ F(\varphi(z),X)\circ j(\varphi(z))\circ d\varphi(z),$$

and thus

$$\begin{aligned}\widetilde{C}'_\infty(z,0)X = \Big[F'(\varphi(z),0)X + J(u(\varphi(z)))\circ\left(F'(\varphi(z),0)X\right)\circ j(\varphi(z))\\ + \left(\widehat{J}'(\varphi(z),0)X\right)\circ du(\varphi(z))\circ j(\varphi(z))\Big]\circ d\varphi(z).\end{aligned}$$

If you apply the pullback operator φ^* to the zeroth-order term in our formula (2.26) for $\mathbf{D}_u$, the expression you end up with is precisely this one, which implies that $\mathbf{D}_k^N$ converges to $\varphi^*\mathbf{D}_u^N$ as $k\to\infty$, thus completing the proof of Theorem 2.8.3.

REMARK 2.8.5. Various more technical versions of Theorems 2.8.1 and 2.8.3 are possible under weaker regularity hypotheses. For example, if the map $\varphi : \Sigma' \to \Sigma$ in Theorem 2.8.3 is assumed to be of class C^m for some finite $m \geqslant 1$, then $\varphi^*N_u \to \Sigma'$ is no longer a smooth vector bundle, but is instead a bundle of class C^m. On these, one can define the notion of a connection or linear Cauchy-Riemann type operator *of class* C^{m-1}; the latter looks locally like $\bar\partial + A$ for a zeroth-order term that is a function of class C^{m-1}. (For the reason why are saying C^{m-1} here instead of C^m, see Exercise 4.1.3.) Inspecting the proof of Theorem 2.8.3, one finds that it still works if φ is only of class C^m, and the resulting linear Cauchy-Riemann type operator $\mathbf{D}^N$ on φ^*N_u is of class C^{m-1}; moreover, $\mathbf{D}^N$ can be assumed arbitrarily C^{m-1}-close to $\varphi_0^*\mathbf{D}_u^N$ if η is sufficiently C^m-small and φ is sufficiently C^m-close to a holomorphic map $\varphi_0 : (\Sigma', j') \to (\Sigma, j)$.

One application of this generalization is to prove the "approximate" version of the theorem of Micallef-White [MW95] mentioned in §2.6, which lies in the background of the dichotomy between simple and multiply covered holomorphic curves. To show for instance that two connected J-holomorphic curves u and v with non-identical images have only *isolated* intersections, the hardest step is to understand the following local picture: for some almost complex structure J on $\mathbb{C}^n$ that matches i at the origin, suppose u and v are both J-holomorphic maps $\mathbb{D} \to \mathbb{C} \times \mathbb{C}^{n-1}$ of the form

$$u(z) = (f(z), \widehat{u}(z)), \qquad v(z) = (g(z), \widehat{v}(z))$$

with $u(0) = v(0) = 0$ such that $f, g : \mathbb{D} \to \mathbb{C}$ vanish to the same order $k \in \mathbb{N}$ at 0, while $\widehat{u}$ and $\widehat{v}$ each vanish to some strictly higher order. The assumption about f and g means that after suitable reparametrizations of u and v near the origin, they can be rewritten in the form

$$u(z) = (z^k, \widehat{u}'(z)), \qquad v(z) = (z^k, \widehat{u}'(z) + \eta(z)), \tag{2.29}$$

for functions $\widehat{u}'$ and η that also vanish to order greater than k at 0. The intersections of u and v in this neighborhood of the origin are thus in bijective correspondence with the zeroes of η and its reparametrizations $\eta_j(z) := \eta(e^{2\pi i j/k} z)$ for $j \in \mathbb{Z}$. A variant of Theorem 2.8.3 then shows that each of the functions η_j is annihilated by some Cauchy-Riemann type operator, and is therefore subject to the similarity principle, so its zero set is discrete unless $\eta_j \equiv 0$. This requires weakened regularity in Theorem 2.8.3, however, because the reparametrizations leading to (2.29) can be shown to be of class C^1, but need not be smooth. The Cauchy-Riemann type operators that annihilate the η_j will therefore be only of class C^0 in general; fortunately, the hypotheses of the similarity principle (Theorem 2.5.3) only require them to be of class L^p for some $p > 2$.

For a detailed version of the argument just outlined, see [**Wen20**, Appendix B.2].

CHAPTER 3

Reeb orbits and asymptotic operators

Contents

We now begin with the analysis of the particular class of J-holomorphic curves that are important in SFT. The next three chapters will focus on the linearized problem, the goal being to prove that this linearization is Fredholm and to compute its index. Using this, along with the implicit function theorem and the Sard-Smale theorem (on genericity of smooth nonlinear Fredholm maps), we will later be able to show that moduli spaces of asymptotically cylindrical J-holomorphic curves are smooth finite-dimensional manifolds under suitable genericity assumptions.

The focus of the present chapter is on the geometric data that controls the asymptotic behavior of a punctured holomorphic curve near infinity; that is to say, the focus is on closed Reeb orbits.

3.1. Stable Hamiltonian structures and Reeb orbits

The basic notions of contact forms and Reeb vector fields were introduced in §1.3, and the most important aspects of SFT make sense primarily in that setting. However, much of the analysis underlying SFT can be formulated under more general assumptions, and this wider perspective is sometimes useful. The geometric motivation for the following definition will be elucidated in §6.1; for now, we will view it simply as a generalization of the notion of the Reeb vector field on a contact manifold.

DEFINITION 3.1.1. A **stable Hamiltonian structure** on an oriented manifold M of dimension $2n-1$ is a pair $\mathcal{H} = (\omega, \lambda)$ consisting of a closed 2-form ω with maximal rank and a 1-form λ such that

$$\lambda \wedge \omega^{n-1} > 0 \qquad \text{and} \qquad \ker \omega \subset \ker d\lambda,$$

where the kernel of a 2-form $\eta \in \Omega^2(M)$ is by definition the kernel of the bundle map $TM \to T^*M : X \mapsto \eta(X, \cdot)$. Every stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ determines a co-oriented **hyperplane field**

$$\xi := \ker \lambda$$

and a so-called **Reeb vector field** $R \in \mathfrak{X}(M)$, which is uniquely specified via the conditions

$$\omega(R, \cdot) \equiv 0 \qquad \text{and} \qquad \lambda(R) \equiv 1.$$

The maximal rank condition on ω in Definition 3.1.1 is equivalent to $\ker \omega \subset TM$ being a smooth 1-dimensional distribution, and the condition $\lambda \wedge \omega^{n-1} > 0$ then implies that λ is nonzero on that distribution. Requiring $\omega(R, \cdot) \equiv 0$ thus determines the direction of the Reeb vector field, while $\lambda(R) \equiv 1$ normalizes it; note that R and ξ are necessarily transverse to each other. Given that λ is nowhere zero, the condition $\lambda \wedge \omega^{n-1} > 0$ is also equivalent to $\omega|_\xi$ being nondegenerate, thus giving $\xi \to M$ the structure of a symplectic vector bundle. Since ω is closed, Cartan's magic formula gives

$$\mathcal{L}_R \omega = d\iota_R \omega + \iota_R d\omega \equiv 0,$$

so that ω is invariant under the flow of R. The extra condition $\ker \omega \subset \ker d\lambda$ is likewise equivalent to $d\lambda(R, \cdot) \equiv 0$ and thus implies

$$\mathcal{L}_R \lambda = d\iota_R \lambda + \iota_R d\lambda \equiv 0,$$

so that the flow of R also preserves λ, and consequently the hyperplane distribution ξ. It follows in particular that the linearized Reeb flow determines a smooth family of symplectic isomorphisms between the fibers of ξ along any orbit of R.

The most popular example of a stable Hamiltonian structure is $(d\alpha, \alpha)$ whenever α is a contact form, and the Reeb vector field in this case is exactly what was defined in §1.3. We shall generally refer to this as **the contact case**, or say that we are working in **the contact setting** when we want to consider stable Hamiltonian structures of this specific form. We will discuss some other nontrivial examples of stable Hamiltonian structures in §6.1.

Let us now clarify precisely what is meant when we refer to a "closed" Reeb orbit in this book.

DEFINITION 3.1.2. Assume M is a manifold with a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ and induced Reeb vector field R. Identifying S^1 with the quotient $\mathbb{R}/\mathbb{Z}$, we shall denote by

$$\widetilde{\mathcal{P}}(\mathcal{H}) \subset C^\infty(S^1, M)$$

the topological space consisting of all smooth maps $\gamma : S^1 \to M$ for which the derivative $\dot{\gamma}$ is a constant positive multiple of the vector field R. The constant $T > 0$ in the equation

$$\dot{\gamma}(t) = T \cdot R(\gamma(t))$$

is in this case called the **period** of $\gamma \in \widetilde{\mathcal{P}}(\mathcal{H})$ (see Remark 3.1.5 below for an explanation of this terminology). There is a natural S^1-action on $\widetilde{\mathcal{P}}(\mathcal{H})$ defined by $\phi \cdot \gamma := \gamma(\cdot + \phi)$ for $\phi \in S^1$, and we will denote the quotient space by

$$\mathcal{P}(\mathcal{H}) := \widetilde{\mathcal{P}}(\mathcal{H})/S^1.$$

An equivalence class $[\gamma] \in \mathcal{P}(\mathcal{H})$ will be called a **closed Reeb orbit**, and sometimes also an **unparametrized** closed orbit if we want to distinguish it from its **parametrizations**, by which we mean its representatives in $\widetilde{\mathcal{P}}(\mathcal{H})$. In the contact case $\mathcal{H} = (d\alpha, \alpha)$, we will also sometimes abbreviate

$$\widetilde{\mathcal{P}}(\alpha) := \widetilde{\mathcal{P}}(d\alpha, \alpha), \qquad \mathcal{P}(\alpha) := \mathcal{P}(d\alpha, \alpha).$$

Note that the period $T > 0$ of an orbit $[\gamma] \in \mathcal{P}(\mathcal{H})$ can easily be recovered from any of its parametrizations $\gamma : S^1 \to M$ since $\lambda(R) \equiv 1$ implies

$$\int_{S^1} \gamma^*\lambda = \int_{S^1} \lambda(T \cdot R(\gamma(t)))\, dt = T.$$

In our exposition, we will often blur the distinction between a parametrization $\gamma : S^1 \to M$ of a closed orbit and the equivalence class that it represents. When we define properties of an orbit in terms of a chosen parametrization, it will be obvious in most cases that the definition is independent of this choice. Here is a typical example of such a definition:

DEFINITION 3.1.3. Assume $\gamma : S^1 \to M$ is a parametrization of a closed Reeb orbit with period $T > 0$ for some stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ on M with Reeb vector field R and flow φ_R^t. The orbit is called **nondegenerate** if the spectrum of the symplectic linear transformation

$$\xi_{\gamma(0)} \xrightarrow{d\varphi_R^T(\gamma(0))} \xi_{\gamma(0)}$$

does not contain 1.

EXERCISE 3.1.4. Show that nondegenerate Reeb orbits are isolated elements of $\mathcal{P}(\mathcal{H})$.

REMARK 3.1.5. We should now point out a few important differences between SFT and Hamiltonian Floer homology as sketched in §1.2. The generators of the latter correspond to 1-periodic orbits of a time-dependent vector field X_t, thus one can regard them literally as solutions $\gamma : S^1 \to M$ to $\dot{\gamma}(t) = X_t(\gamma(t))$, with no need to introduce any scaling factors or equivalence relations. By contrast, SFT considers orbits of *all* periods, and since the Reeb vector field R is time-independent, solutions always come in S^1-families of parametrizations that are related to each other by constant shifts. An alternative formulation of Definition 3.1.2 one sometimes encounters in the literature is to let the definition of S^1 vary depending on the period of the orbit, so a parametrization of a T-periodic orbit would then be given by a map

$$x : \mathbb{R}/T\mathbb{Z} \to M \qquad \text{satisfying} \qquad \dot{x}(t) = R(x(t)).$$

One recovers a parametrization of this type from $\gamma : S^1 \to M$ in Definition 3.1.2 via the relation

$$\gamma(t) = x(Tt).$$

Notice that a closed orbit is not completely determined by its image in M, e.g. every orbit with parametrization $\gamma : S^1 \to M$ also has multiple covers γ^k for $k \in \mathbb{N}$, which are closed Reeb orbits with parametrization

$$\gamma^k(t) := \gamma(kt),$$

and γ is not equivalent to γ^k according to Definition 3.1.2 unless $k = 1$.

3.2. The linearization in Morse homology

Since Morse homology is the prototype for all Floer-type theories, we can gain useful intuition by recalling how the analysis works for the linearization of the gradient flow problem in Morse theory. The basic features of the problem were discussed already in §1.2.

Assume (M, g) is a closed n-dimensional Riemannian manifold, $f : M \to \mathbb{R}$ is a smooth function, and for two critical points $x_+, x_- \in \mathrm{Crit}(f)$, consider the moduli space of parametrized gradient flow lines

$$\mathcal{M}(x_-, x_+) := \left\{ u \in C^\infty(\mathbb{R}, M) \;\middle|\; \dot{u} + \nabla f(u) = 0, \; \lim_{s \to \pm\infty} u(s) = x_\pm \right\}.$$

The map $\mathcal{M}(x_-, x_+) \to M : u \mapsto u(0)$ gives a natural identification of $\mathcal{M}(x_-, x_+)$ with the intersection between the unstable manifold of x_- and the stable manifold of x_+ for the negative gradient flow. We say the pair (g, f) is **Morse-Smale** if f is Morse and all such intersections between stable and unstable manifolds of two critical points are transverse. In this case $\mathcal{M}(x_-, x_+)$ is a smooth manifold with

$$\dim \mathcal{M}(x_-, x_+) = \mathrm{Morse}(x_-) - \mathrm{Morse}(x_+), \tag{3.1}$$

because the unstable manifold of x_- has dimension $\mathrm{Morse}(x_-)$ and the stable manifold of x_+ has codimension $\mathrm{Morse}(x_+)$. All of this can be proved using finite-dimensional differential topology, but we will see that the dimension computation as just described cannot generalize to the study of Floer trajectories or holomorphic curves in symplectizations, because the right hand side of (3.1) in those cases becomes $\infty - \infty$. Let us therefore discuss how (3.1) can be proved using a nonlinear functional-analytic approach that does generalize. For more details on the following discussion, see [Sch93].

Following the strategy laid out in §2.1, $\mathcal{M}(x_-, x_+)$ can be identified with the zero set of a smooth section

$$\boldsymbol{\sigma} : \mathcal{B} \to \mathcal{E} : u \mapsto \dot{u} + \nabla f(u),$$

where $\mathcal{B}$ is a Banach manifold of maps $u : \mathbb{R} \to M$ satisfying $\lim_{s\to\pm\infty} u(s) = x_\pm$, and $\mathcal{E} \to \mathcal{B}$ is a smooth Banach space bundle whose fibers $\mathcal{E}_u$ contain $\Gamma(u^*TM)$. The linearization $D\boldsymbol{\sigma}(u) : T_u\mathcal{B} \to \mathcal{E}_u$ of this section at a zero $u \in \boldsymbol{\sigma}^{-1}(0)$ defines a first-order linear differential operator

$$\mathbf{D}_u : \Gamma(u^*TM) \to \Gamma(u^*TM)$$

which takes the form

$$\mathbf{D}_u\eta = \nabla_s\eta + \nabla_\eta\nabla f$$

for any choice of symmetric connection ∇ on M. Taking suitable Sobolev completions of $\Gamma(u^*TM)$, we are therefore led to consider bounded linear operators[1] of the form

$$\mathbf{D}_u = \nabla_s + \nabla\nabla f : W^{k,p}(u^*TM) \to W^{k-1,p}(u^*TM) \tag{3.2}$$

for $k \in \mathbb{N}$ and $1 < p < \infty$, and the first task is to prove that whenever x_+ and x_- satisfy the Morse condition, this is a Fredholm operator of index $\operatorname{ind}\mathbf{D}_u = \operatorname{Morse}(x_-) - \operatorname{Morse}(x_+)$.

Choose coordinates near x_+ in which g looks like the standard Euclidean inner product at x_+. This induces a trivialization of u^*TM over $[T, \infty)$ for $T > 0$ sufficiently large, and we are free to assume that the connection ∇ is the standard one determined by these coordinates on $[T, \infty)$. Using the trivialization to identify sections $\eta \in \Gamma(u^*TM)$ over $[T, \infty)$ with functions $\eta : [T, \infty) \to \mathbb{R}^n$, $\mathbf{D}_u$ now acts on η as

$$(\mathbf{D}_u\eta)(s) = \partial_s\eta(s) + A(s)\eta(s), \tag{3.3}$$

where $A(s) \in \mathbb{R}^{n\times n}$ is the matrix of the linear transformation $dX(s) : \mathbb{R}^n \to \mathbb{R}^n$, with $X(s) \in \mathbb{R}^n$ being the coordinate representation of $\nabla f(u(s)) \in T_{u(s)}M$. As $s \to \infty$, the zeroth-order term in this expression converges to a symmetric matrix

$$A_+ := \lim_{s\to\infty} A(s),$$

which is the coordinate representation of the Hessian $\nabla^2 f(x_+)$. Any choice of coordinates near x_- produces a similar formula for $\mathbf{D}_u$ over $(-\infty, -T]$, $A(s)$ converging as $s \to -\infty$ to another symmetric matrix A_- representing $\nabla^2 f(x_-)$. Both the Morse condition and the dimension $\operatorname{Morse}(x_-) - \operatorname{Morse}(x_+)$ can now be expressed entirely

[1]We are ignoring an analytical subtlety: since $u^*TM \to \mathbb{R}$ has no canonical trivialization and $\mathbb{R}$ is noncompact, it is not completely obvious what the definition of the Sobolev space $W^{k,p}(u^*TM)$ should be. We will return to this issue in a more general context in the next chapter.

in terms of these two matrices: $x_\pm$ is Morse if and only if $A_\pm$ is invertible, and the Fredholm index of $\mathbf{D}_u$ will then be

$$\text{Morse}(x_-) - \text{Morse}(x_+) = \dim E^-(A_-) - \dim E^-(A_+),$$

where for any symmetric matrix A we denote by $E^-(A)$ the direct sum of all its eigenspaces with negative eigenvalue. The main linear functional-analytic result underlying Morse homology can now be stated as follows (cf. [Sch93]):

PROPOSITION 3.2.1. *Assume $k \in \mathbb{N}$ and $1 < p < \infty$. Suppose $E \to \mathbb{R}$ is a smooth vector bundle with trivializations fixed in neighborhoods of $-\infty$ and $+\infty$, and $\mathbf{D} : W^{k,p}(E) \to W^{k-1,p}(E)$ is a first-order differential operator which asymptotically takes the form (3.3) near $\pm\infty$ with respect to the chosen trivializations, where $A(s)$ is a smooth family of n-by-n matrices with well-defined asymptotic limits $A_\pm := \lim_{s\to\pm\infty} A(s)$ which are symmetric. If A_+ and A_- are also invertible, then $\mathbf{D}$ is Fredholm and*

$$\text{ind}(\mathbf{D}) = \dim E^-(A_-) - \dim E^-(A_+). \tag{3.4}$$

□

REMARK 3.2.2. The hypothesis that $A_\pm$ is invertible in Prop. 3.2.1 cannot be lifted: indeed, suppose $\mathbf{D}$ is Fredholm but e.g. A_+ has 0 in its spectrum. Then one can easily perturb $A(s)$ and hence A_+ in two distinct ways producing two distinct values of $\dim E^-(A_+)$, pushing the zero eigenvalue either up or down. This produces two perturbed Fredholm operators that have different indices according to (3.4), but they also belong to a continuous family of Fredholm operators, and must therefore have the same index, giving a contradiction.

The formula (3.4) makes sense of course because $E^-(A_\pm)$ are both finite-dimensional vector spaces, but in Floer-type theories, we typically encounter critical points with infinite Morse index. With this in mind, it is useful to note that (3.4) can be rewritten without explicitly referencing $E^-(A_+)$ or $E^-(A_-)$. Indeed, choose a continuous path of symmetric matrices $\{B_t\}_{t\in[-1,1]}$ connecting $B(-1) := A_-$ to $B(1) := A_+$. The following fact is non-obvious, but it can be proved via simplifications of the arguments in this chapter (cf. Theorem 3.5.1). The spectrum of B_t varies continuously with t in the following sense: one can choose a family of continuous functions

$$\{\lambda_j : [-1,1] \to \mathbb{R}\}_{j\in I}$$

for the index set $I = \{1, \dots, n\}$ such that for every $t \in [-1,1]$, the set of eigenvalues of B_t counted with multiplicity is $\{\lambda_j(t)\}_{j\in I}$. The **spectral flow** from A_- to A_+ is then defined as a signed count of the number of paths of eigenvalues that cross from one side of zero to the other, namely

$$\mu^{\text{spec}}(A_-, A_+) := \#\left\{j \in I \;\middle|\; \lambda_j(-1) < 0 < \lambda_j(1)\right\} - \#\left\{j \in I \;\middle|\; \lambda_j(-1) > 0 > \lambda_j(1)\right\}.$$

The index formula (3.4) now becomes

$$\text{ind}(\mathbf{D}) = \mu^{\text{spec}}(A_-, A_+).$$

This description of the index has the advantage that it could potentially make sense and give a well-defined integer even if $A_\pm$ were symmetric operators on an infinite-dimensional Hilbert space: they might both have infinitely many positive and negative eigenvalues, but only finitely many that change sign along a path from A_- to A_+. We will make this discussion precise in the following sections.

3.3. The Hessian of the contact action functional

We will view SFT as an infinite-dimensional analogue of Morse homology in which closed nondegenerate Reeb orbits take the place of Morse critical points. The role of the Hessian is then played by a certain self-adjoint differential operator on the contact bundle along each closed orbit.

Before explaining this, let's quickly revisit the Floer homology for a time-dependent Hamiltonian $\{H_t : M \to \mathbb{R}\}_{t \in S^1}$ on a symplectic manifold (M, ω). In Chapter 1, we introduced the symplectic action functional $\mathcal{A}_H : C^\infty_{\text{contr}}(S^1, M) \to \mathbb{R}$ and wrote down the formula

$$\nabla \mathcal{A}_H(\gamma) = J_t(\gamma)\,(\dot{\gamma} - X_t(\gamma)) \in \Gamma(\gamma^* TM) =: T_\gamma C^\infty_{\text{contr}}(S^1, M)$$

for the "unregularized" gradient of $\mathcal{A}_H$ at a contractible loop $\gamma \in C^\infty_{\text{contr}}(S^1, M)$. Here X_t denotes the Hamiltonian vector field, and J_t is a time-dependent family of compatible almost complex structures, which determines the L^2-product

$$\langle \eta_1, \eta_2 \rangle_{L^2} = \int_{S^1} \omega(\eta_1(t), J_t \eta_2(t))\, dt.$$

The critical points of $\mathcal{A}_H$ are the loops γ such that $\nabla \mathcal{A}_H(\gamma) = 0$. Formally, the Hessian of $\mathcal{A}_H$ at $\gamma \in \text{Crit}(\mathcal{A}_H)$ is the "linearization of $\nabla \mathcal{A}_H$ at γ," which gives a linear operator

$$\mathbf{A}_\gamma := \nabla^2 \mathcal{A}_H(\gamma) : \Gamma(\gamma^* TM) \to \Gamma(\gamma^* TM).$$

To write it down, one can choose any connection ∇ on M, and choose for $\eta \in \Gamma(\gamma^* TM)$ a smooth family $\{\gamma_\rho : S^1 \to M\}_{\rho \in (-\epsilon, \epsilon)}$ with $\gamma_0 = \gamma$ and $\partial_\rho \gamma_\rho|_{\rho=0} = \eta$, and then compute

$$\mathbf{A}_\gamma \eta := \left. \nabla_\rho \left[\nabla \mathcal{A}_H(\gamma_\rho) \right] \right|_{\rho=0}.$$

The result is independent of the choice of connection since $\nabla \mathcal{A}_H(\gamma) = 0$.

EXERCISE 3.3.1. Show that if the connection ∇ on M is chosen to be symmetric, then $\mathbf{A}_\gamma \eta = J_t(\nabla_t \eta - \nabla_\eta X_t)$.

To adapt this discussion for SFT, fix a $(2n-1)$-dimensional manifold M with stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$, induced hyperplane field $\xi = \ker \lambda \subset TM$ and Reeb vector field R, and a complex structure $J : \xi \to \xi$ compatible with the symplectic structure $\omega|_\xi$. Let

$$\pi_\xi : TM \to \xi$$

denote the projection along R. Let us also impose the simplifying assumption that the closed 2-form ω is exact, and write

$$\omega = d\beta, \qquad \text{for some } \beta \in \Omega^1(M).$$

(For hints on what to do if ω is not exact, see Remark 3.3.7.) A choice of primitive β enables us to write down an action functional

$$\mathcal{A}_\beta : C^\infty(S^1, M) \to \mathbb{R} : \gamma \mapsto \int_{S^1} \gamma^*\beta. \tag{3.5}$$

In the contact case $\mathcal{H} = (d\alpha, \alpha)$, the contact form α gives a canonical choice of primitive for $\omega = d\alpha$, and $\mathcal{A}_\alpha = \mathcal{A}_\beta$ is then called the **contact action functional**.

One computes the first variation of $\mathcal{A}_\beta$ at $\gamma \in C^\infty(S^1, M)$ in the direction of $\eta \in \Gamma(\gamma^*TM)$ by choosing a smooth family $\{\gamma_\rho \in C^\infty(S^1, M)\}_{\rho \in (-\epsilon,\epsilon)}$ with $\gamma_0 = \gamma$ and $\partial_\rho \gamma_\rho|_{\rho=0} = \eta$: we find

$$\begin{aligned} d\mathcal{A}_\beta(\gamma)\eta &:= \frac{d}{d\rho} \int_{S^1} \gamma_\rho^*\beta \Big|_{\rho=0} = \int_{S^1} \partial_\rho \left[\beta(\dot\gamma_\rho(t))\right]\Big|_{\rho=0} dt \\ &= \int_{S^1} d\beta\Big(\partial_\rho\gamma_\rho(t)|_{\rho=0}, \dot\gamma(t)\Big)\, dt + \int_{S^1} \partial_t \Big[\beta\big(\partial_\rho \gamma_\rho(t)|_{\rho=0}\big)\Big]\, dt \\ &= \int_{S^1} \omega(\eta, \dot\gamma)\, dt = -\int_{S^1} \omega(\pi_\xi\dot\gamma, \eta)\, dt. \end{aligned}$$

The functional has a built-in degeneracy since it is parametrization-invariant; in particular, $d\mathcal{A}_\beta(\gamma)\eta = 0$ whenever η points in the direction of the Reeb vector field, a symptom of the fact that closed parametrized Reeb orbits always come in families related to each other by shifts in the parametrization. A loop $\gamma : S^1 \to M$ is critical for $\mathcal{A}_\beta$ if and only if $\dot\gamma$ is everywhere tangent to R, allowing for an infinite-dimensional family of distinct perturbations—however, there exist preferred parametrizations, namely those for which $\dot\gamma$ is a *constant positive* multiple of R, meaning

$$\dot\gamma = T \cdot R(\gamma), \qquad T := \int_{S^1} \gamma^*\lambda > 0. \tag{3.6}$$

Such a loop determines a closed Reeb orbit in the sense of Definition 3.1.2, and it corresponds to a T-periodic solution $x : \mathbb{R} \to M$ to $\dot x = R(x)$, where $\gamma(t) = x(Tt)$.

The discussion above indicates that we cannot derive a "Hessian" of $\mathcal{A}_\beta$ in the same straightforward way as in Floer homology, as the resulting operator will always have nontrivial kernel due to the degeneracy in the R direction. To avoid this, we shall consider only preferred parametrizations $\gamma : S^1 \to M$ of the form (3.6), and perturbations in directions tangent to ξ, which is transverse to every Reeb orbit. For $\eta \in \Gamma(\gamma^*\xi)$, we then have

$$d\mathcal{A}_\beta(\gamma)\eta = \int_{S^1} \omega(-J\pi_\xi\dot\gamma, J\eta)\, dt = \langle -J\pi_\xi\dot\gamma, \eta\rangle_{L^2},$$

where we define an L^2-product for sections of $\gamma^*\xi$ by

$$\langle \eta, \eta' \rangle_{L^2} := \int_{S^1} \omega(\eta, J\eta')\, dt. \tag{3.7}$$

It therefore seems sensible to write

$$\nabla \mathcal{A}_\beta(\gamma) := -J\pi_\xi\dot\gamma \in \Gamma(\gamma^*\xi),$$

and we shall define the Hessian at a critical point $\gamma \in \widetilde{\mathcal{P}}(\mathcal{H})$ as the linearization of $\nabla \mathcal{A}_\beta$ in ξ directions, that is,

$$\nabla^2 \mathcal{A}_\beta(\gamma) : \Gamma(\gamma^*\xi) \to \Gamma(\gamma^*\xi).$$

Given $\eta \in \Gamma(\gamma^*\xi)$, choose a smooth family $\{\gamma_\rho : S^1 \to M\}_{\rho \in (-\epsilon,\epsilon)}$ with $\gamma_0 = \gamma$ and $\partial_\rho \gamma_\rho|_{\rho=0} = \eta$, and fix a symmetric connection ∇ on M. Let us first use this connection to differentiate the family of sections $\pi_\xi \dot{\gamma}_\rho \in \Gamma(\gamma_\rho^*\xi)$ with respect to the parameter:

$$\begin{aligned} \nabla_\rho \left(\pi_\xi \dot{\gamma}_\rho\right)\big|_{\rho=0} &= \nabla_\rho \left[\partial_t \gamma_\rho - \lambda(\partial_t \gamma_\rho) R(\gamma_\rho)\right]\big|_{\rho=0} \\ &= \nabla_t \eta - \lambda(\dot{\gamma}) \nabla_\eta R - \partial_\rho \left[\lambda(\dot{\gamma}_\rho)\right]\big|_{\rho=0} \cdot R(\gamma). \end{aligned}$$

The latter expression is *a priori* an element of $\Gamma(\gamma^*TM)$, but since $\pi_\xi \dot{\gamma}_\rho$ belongs to the subspace $\Gamma(\gamma_\rho^*\xi) \subset \Gamma(\gamma_\rho^*TM)$ for every ρ and $\pi_\xi \dot{\gamma}$ vanishes, this derivative is independent of the choice of connection, and also takes its value in the subspace $\Gamma(\gamma^*\xi)$. Moreover, it can be simplified in light of the condition $d\lambda(R, \cdot) \equiv 0$, which implies

$$0 = T \cdot d\lambda(\eta, R(\gamma)) = d\lambda(\partial_\rho \gamma_\rho, \partial_t \gamma_\rho)\big|_{\rho=0} = \partial_\rho \left[\lambda(\dot{\gamma}_\rho)\right]\big|_{\rho=0} - \partial_t \left[\lambda(\eta)\right] = \partial_\rho \left[\lambda(\dot{\gamma}_\rho)\right]\big|_{\rho=0},$$

so that the previous computation becomes

$$\nabla_\rho \left(\pi_\xi \dot{\gamma}_\rho\right)\big|_{\rho=0} = \nabla_t \eta - T \nabla_\eta R \in \Gamma(\gamma^*\xi),$$

and thus

$$\nabla_\rho \left(-J \pi_\xi \dot{\gamma}_\rho\right)\big|_{\rho=0} = -J \left(\nabla_t \eta - T \nabla_\eta R\right) \in \Gamma(\gamma^*\xi).$$

This motivates the following definition.

DEFINITION 3.3.2. Given a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ on M with $\xi = \ker \lambda$, a parametrized orbit $\gamma \in \widetilde{\mathcal{P}}(\mathcal{H})$ of the Reeb vector field R with period $T > 0$, and an ω-compatible complex structure J on the bundle $\gamma^*\xi \to S^1$, the **asymptotic operator associated to** γ is the first-order differential operator on $\gamma^*\xi$ defined by

$$\mathbf{A}_\gamma : \Gamma(\gamma^*\xi) \to \Gamma(\gamma^*\xi) : \eta \mapsto -J(\nabla_t \eta - T \nabla_\eta R),$$

where ∇ is any choice of symmetric connection on M.

REMARK 3.3.3. If $\gamma_0, \gamma_1 : S^1 \to M$ are two parametrizations of the same orbit, related by $\gamma_1(t) = \gamma_0(t + c)$ for a constant shift $c \in S^1$, then one easily checks that their asymptotic operators $\mathbf{A}_{\gamma_0}$ and $\mathbf{A}_{\gamma_1}$ are conjugate via the isomorphism $\Gamma(\gamma_0^*\xi) \to \Gamma(\gamma_1^*\xi) : \eta \mapsto \eta(\cdot + c)$. For this reason, it makes sense in most situations to regard an asymptotic operator as something associated to an *unparametrized* closed Reeb orbit, even if concrete realizations of the operator require a choice of parametrization. The most important properties of the operator are all independent of this choice.

EXERCISE 3.3.4. For a closed Reeb orbit $\gamma : S^1 \to M$ and the pullback of the bundle $\gamma^*\xi \to S^1$ via the cover $\mathbb{R} \to S^1 = \mathbb{R}/\mathbb{Z}$, show that solutions to the linear equation $\nabla_t \eta - T \nabla_\eta R = 0$ on the pullback (with ∇ a symmetric connection on M) are given by operating on $\xi_{\gamma(0)}$ with the linearized Reeb flow. Conclude that γ is nondegenerate if and only if $\ker \mathbf{A}_\gamma$ is trivial. *Hint: Try differentiating families of solutions to the equation $\dot{x} = TR(x)$.*

REMARK 3.3.5. Another way of phrasing the result of Exercise 3.3.4 is that $\mathbf{A}_\gamma$ can be written as $-J\widehat{\nabla}_t$, where $\widehat{\nabla}_t$ is the unique *symplectic connection* on the symplectic vector bundle $\gamma^*\xi \to S^1$ for which parallel transport is given by the linearized Reeb flow.

REMARK 3.3.6 (sign conventions). You might be slightly concerned about the sign difference between the formulas for asymptotic operators in Exercise 3.3.1 and Definition 3.3.2. The former comes from Floer homology and the latter from SFT, two subfields of symplectic topology in which slightly different conventions are considered standard.[2] The discrepancy seems to originate from the fact that while our account of Floer homology has referred always to the *negative* gradient flow of $\mathcal{A}_H$, SFT is actually defined via the *positive* gradient flow of the contact action functional $\mathcal{A}_\alpha$. The words "gradient flow" in SFT must in any case be interpreted very loosely. If

$$u : [0, \infty) \times S^1 \to \mathbb{R} \times M$$

is the cylindrical end of a finite-energy J-holomorphic curve for some $J \in \mathcal{J}(\alpha)$ as we described in Chapter 1, then $u(s,t)$ does not satisfy anything so straightforward as $\partial_s - \nabla\mathcal{A}_\alpha(u(s,\cdot)) = 0$, but it does satisfy

$$\pi_\xi \partial_s u + J\pi_\xi \partial_t u = 0,$$

which can be interpreted as the projection of a positive gradient flow equation to the contact bundle. This observation is a local symptom of a more important global fact that follows from Stokes' theorem: in the contact setting, any asymptotically cylindrical J-holomorphic curve $u : \dot{\Sigma} \to \mathbb{R}\times M$ with positive and negative punctures $\Gamma^\pm$ asymptotic to orbits $\{\gamma_z\}_{z\in\Gamma^\pm}$ satisfies

$$\sum_{z\in\Gamma^+} \mathcal{A}_\alpha(\gamma) - \sum_{z\in\Gamma^-} \mathcal{A}_\alpha(\gamma) = \int_{\dot{\Sigma}} u^* d\alpha \geqslant 0.$$

This generalizes the basic fact in Floer homology that flow lines decrease action and, conversely, have their energy controlled by the action.

REMARK 3.3.7. The action functional $\mathcal{A}_\beta$ cannot be defined in the way described above if ω is not exact, but since a closed 2-form is necessarily exact in the neighborhood of any embedded circle, one can define local versions of $\mathcal{A}_\beta$ that make sense in some neighborhood of any given element of $C^\infty(S^1, M)$: the computation above shows that the gradient $\nabla\mathcal{A}_\beta$ is in any case globally defined and independent of any choices of local primitives for ω. For this reason, the asymptotic operator $\mathbf{A}_\gamma$ of an orbit γ always makes sense and can be interpreted as the Hessian of an action functional defined for loops close to γ. Various global (though typically non-canonical) definitions of an action functional in the non-exact case are also possible:

[2]The literature on embedded contact homology (ECH) is a special case: while ECH is defined within the same analytical framework as SFT, papers such as [Hut14, HT07] omit the initial minus sign in their definitions of asymptotic operators. Some of the results in §3.7 relating eigenvalues of asymptotic operators to winding numbers therefore work out differently in the ECH context.

e.g. if $[\omega]|_{\pi_2(M)} = 0$ and one chooses to focus on *contractible* orbits specifically, then one can follow the example of the symplectic action functional from §1.2 and define $\mathcal{A}(\gamma) := \int_{\mathbb{D}^2} \bar{\gamma}^*\omega$ for any choice of smooth map $\bar{\gamma} : \mathbb{D} \to M$ matching γ at the boundary.

3.4. Asymptotic operators on Hermitian bundles

We would now like to develop some of the general properties of asymptotic operators. Recall that on any symplectic vector bundle (E, ω), a compatible complex structure J determines a Hermitian inner product

$$\langle v, w \rangle = \omega(v, Jw) + i\omega(v, w),$$

and conversely, any Hermitian inner product on a complex vector bundle determines a symplectic structure via the same relation. For this reason, we shall refer to any vector bundle E with a compatible pair (J, ω) of complex and symplectic structures as a **Hermitian vector bundle**. A **unitary trivialization** of such a bundle is a trivialization that identifies fibers with $\mathbb{R}^{2n} = \mathbb{C}^n$ such that J and ω become the standard complex structure $J_0 := i$ and symplectic structure $\omega_0 := g_0(J_0 \cdot, \cdot)$ respectively; here g_0 denotes the standard Euclidean inner product.

DEFINITION 3.4.1. Fix a Hermitian vector bundle (E, J, ω) over S^1. A **smooth asymptotic operator on** (E, J, ω) is any real-linear differential operator of the form $-J\nabla_t : \Gamma(E) \to \Gamma(E)$, where ∇ is a symplectic connection on E.

Remark 3.3.5 shows that the asymptotic operator $\mathbf{A}_\gamma$ for a closed Reeb orbit γ is also a smooth asymptotic operator on $\gamma^*\xi$ in the sense of Definition 3.4.1, where the bundle $\gamma^*\xi \to S^1$ carries a Hermitian structure determined by the stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ and a choice of ω-compatible complex structure J on ξ. Since $\mathbf{A}_\gamma$ was derived as the Hessian of an action functional, it should not be surprising that it turns out to be symmetric:

EXERCISE 3.4.2. Show that any smooth asymptotic operator on a Hermitian vector bundle (E, J, ω) over S^1 is symmetric with respect to the real L^2 bundle metric

$$\langle \eta_1, \eta_2 \rangle_{L^2} := \int_{S^1} \omega(\eta_1(t), J\eta_2(t))\, dt.$$

EXERCISE 3.4.3. Show that Hermitian vector bundles (E, J, ω) over S^1 are always globally trivializable, and a choice of global unitary trivialization identifies each smooth asymptotic operator on (E, J, ω) with an operator of the form

$$\mathbf{A} : C^\infty(S^1, \mathbb{R}^{2n}) \to C^\infty(S^1, \mathbb{R}^{2n}) : \eta \mapsto -J_0\partial_t\eta - S(t)\eta$$

for some smooth loop $S : S^1 \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$, where we denote

$$\mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}) := \left\{ B \in \mathrm{End}(\mathbb{R}^{2n}) = \mathbb{R}^{2n\times 2n} \;\middle|\; B^{\mathrm{T}} = B \right\}.$$

Hint: Use the fact that the difference between two connections is a bundle map, and deduce the symmetry of $S(t)$ from Exercise 3.4.2.

REMARK 3.4.4. Most of the Hermitian vector bundles (E, J, ω) that will arise in our applications can be regarded naturally as symplectic vector bundles (E, ω) on which an auxiliary compactible complex structure J has been chosen. This auxiliary data is convenient to have in the picture for at least two reasons: one is that it is naturally present in the context of pseudoholomorphic curves, and the other is the observation in Exercise 3.4.2 that J turns a symplectic connection into a symmetric operator, for which there is a well-developed spectral theory. However, we will see that the definition of the Conley-Zehnder index for Reeb orbits depends only on symplectic data, and not on the auxiliary choice of a complex structure. We will frequently use the fact that on every symplectic vector bundle (E, ω), there is a contractible space of choices of J such that (E, J, ω) is a Hermitian vector bundle, and moreover, the homotopy classes of symplectic trivializations of (E, ω) are in bijective correspondence with the homotopy classes of unitary trivializations of (E, J, ω). This follows essentially from the fact that the natural inclusion $\mathrm{U}(n) \hookrightarrow \mathrm{Sp}(2n)$ is a homotopy equivalence (see e.g. [MS17, Prop. 2.2.4]).

For functional-analytic purposes, we shall regard asymptotic operators on Hermitian bundles (E, J, ω) as bounded real-linear operators

$$\mathbf{A} : H^1(E) \to L^2(E),$$

where H^1 is an abbreviation for the Sobolev class $W^{1,2}$. (For details on Sobolev norms for spaces of sections of vector bundles over a closed manifold, see §A.4.) Note that since the difference between any two smooth asymptotic operators is tensorial, that difference extends to a bounded linear operator on $L^2(E)$; as an operator $H^1(E) \to L^2(E)$, it is therefore the composition of a bounded operator with the compact inclusion $H^1(E) \hookrightarrow L^2(E)$, implying that it is compact. This property will play an essential role when we study the spectrum of asymptotic operators in §3.5.

For technical reasons, we will sometimes need to consider a larger class of asymtotic operators whose zeroth-order terms are not necessarily smooth, nor even continuous. One way to weaken our regularity assumptions without invalidating the discussion in the previous paragraph is the following:

DEFINITION 3.4.5. An **asymptotic operator** (of class L^∞) on a Hermitian vector bundle (E, J, ω) over S^1 is a bounded linear operator $\mathbf{A} : H^1(E) \to L^2(E)$ that is identified under any choice of global unitary trivialization with an operator of the form

$$H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n}) : \eta \mapsto -J_0 \partial_t \eta - S(t)\eta$$

for some function $S \in L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$. The space

$$\mathcal{A}(E) \subset \mathscr{L}(H^1(E), L^2(E))$$

of all asymptotic operators on E is thus an affine space over the space $L^\infty(\mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(E))$ of symmetric real-linear bundle maps $E \to E$ of class L^∞, and we assign to it the corresponding L^∞-topology. We also denote

$$\mathcal{A}^*(E) := \left\{ \mathbf{A} \in \mathcal{A}(E) \;\middle|\; \ker \mathbf{A} = \{0\} \right\},$$

and call the operators in this subset **nondegenerate**.

We will assume henceforward that all asymptotic operators we consider are of class L^∞ unless otherwise noted, though most examples that arise in geometric settings (e.g. the operator corresponding to a closed Reeb orbit) will be smooth.

EXERCISE 3.4.6. Generalize Exercise 3.4.2 to prove that asymptotic operators of class L^∞ are also L^2-symmetric.

LEMMA 3.4.7. *All asymptotic operators* $\mathbf{A} \in \mathcal{A}(E)$ *are Fredholm with index* 0.

PROOF. Choosing a global unitary trivialization, it suffices to consider an operator of the form $-J_0\partial_t - S : H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n})$ for some $S : S^1 \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ of class L^∞, and since the operator $H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n}) : \eta \mapsto S\eta$ is compact, we can regard the zeroth-order term as a compact perturbation and thus restrict attention to the operator $-J_0\,\partial_t : H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n})$. Since J_0 defines an isomorphism, it suffices actually to show that the ordinary differential operator

$$\partial_t : H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n})$$

is Fredholm with index 0. The kernel of this operator is the space of constant functions $S^1 \to \mathbb{R}^{2n}$, which has dimension $2n$. To compute the dimension of the cokernel, we observe that if $f = \partial_t F$ for some $F \in H^1(S^1, \mathbb{R}^{2n})$, then Proposition A.1.11 implies that F is absolutely continuous and has classical derivative equal to f almost everywhere, so that by periodicity and the fundamental theorem of calculus, $\int_{S^1} f(t)\,dt = 0$. Conversely, if $\int_{S^1} f(t)\,dt = 0$ with $f \in L^2(S^1, \mathbb{R}^{2n})$, then the function $F(s) := \int_0^s f(t)\,dt$ is periodic in s and (by Corollary A.1.12) defines an element of $H^1(S^1, \mathbb{R}^{2n})$ satisfying $\partial_t F = f$. Hence the image of ∂_t is exactly the set

$$\mathrm{im}(\partial_t) = \left\{ f \in L^2(S^1, \mathbb{R}^{2n}) \;\middle|\; \int_{S^1} f(t)\,dt = 0 \right\},$$

which has codimension $2n$. □

COROLLARY 3.4.8. *An asymptotic operator* $\mathbf{A} \in \mathcal{A}(E)$ *is nondegenerate if and only if it defines an isomorphism* $H^1(E) \to L^2(E)$. □

Observe that for the L^∞-topology on $\mathcal{A}(E)$ specified in Definition 3.4.5, the inclusion of $\mathcal{A}(E)$ into the space of bounded linear operators $H^1(E) \to L^2(E)$ is continuous, so the fact that invertibility is an open condition implies:

COROLLARY 3.4.9. *The subset* $\mathcal{A}^*(E) \subset \mathcal{A}(E)$ *is open.* □

Since smooth asymptotic operators on a bundle (E, ω, J) are defined in terms of symplectic connections, they also determine (and are determined by) symplectic parallel transport maps. This notion can be extended to asymptotic operators of class L^∞, but since the differential equation $(-J_0\partial_t - S(t))\Psi(t) = 0$ may in this case have discontinuous coefficients, it requires a slight generalization of the standard existence/uniqueness theorem for ODEs.

EXERCISE 3.4.10. In this exercise we consider linear ordinary differential equations with coefficients of class L^1_{loc}.

(a) Suppose $I \subset \mathbb{R}$ is a compact interval, $P \subset L^1(I, \mathrm{End}(\mathbb{R}^n))$ is a subset such that $M := \sup \left\{ \|A\|_{L^1} \;\middle|\; A \in P \right\} < 1$, and for $R > 0$, X_R denotes the complete metric space

$$X_R := \left\{ \varphi \in C^0(P \times I, \mathbb{R}^n) \;\middle|\; \|\varphi\|_{C^0} \leqslant R \right\}.$$

Show that for any $x_0 \in \mathbb{R}^n$, any $t_0 \in I$ and any $R \geqslant \frac{|x_0|}{1-M}$, the formula

$$(T\varphi)(A,t) := x_0 + \int_{t_0}^t A(s)\varphi(A,s)\,ds$$

defines a contraction map $T : X_R \to X_R$, which therefore has a unique fixed point.

(b) Deduce from the contraction in part (a) that for any open interval $\mathcal{U} \subset \mathbb{R}$ and constants $t_0 \in \mathcal{U}$, $x_0 \in \mathbb{R}^n$, there exists a continuous map

$$L^1_{\mathrm{loc}}(\mathcal{U}, \mathrm{End}(\mathbb{R}^n)) \times \mathcal{U} \to \mathbb{R}^n : (A,t) \mapsto x_A(t)$$

such that for each $A \in L^1_{\mathrm{loc}}(\mathcal{U}, \mathrm{End}(\mathbb{R}^n))$, $x_A : \mathcal{U} \to \mathbb{R}^n$ satisfies the initial value problem

$$\dot{x}(t) = A(t)x(t) \text{ for almost all } t, \qquad x(t_0) = x_0, \tag{3.8}$$

and is the unique solution to this problem that is absolutely continuous on compact subsets.

(c) Show that if $A : \mathcal{U} \to \mathrm{End}(\mathbb{R}^n)$ is assumed to be of class L^p_{loc} with $1 \leqslant p \leqslant \infty$, then the solution $x : \mathcal{U} \to \mathbb{R}^n$ to (3.8) is of class $W^{1,p}_{\mathrm{loc}}$. *Hint: For a useful characterization of $W^{1,p}_{\mathrm{loc}}(\mathbb{R})$, see Corollary A.1.12.*

PROPOSITION 3.4.11. *On any Hermitian vector bundle (E, ω, J) over $S^1 = \mathbb{R}/\mathbb{Z}$, there is a natural bijective correspondence between the following objects:*

- *Asymptotic operators $\mathbf{A}$ of class L^∞;*
- *Continuous families $\{\Psi(t)\}_{t \in \mathbb{R}}$ of Sobolev class $W^{1,\infty}_{\mathrm{loc}}$ consisting of symplectic linear maps $\Psi(t) : E_{[0]} \to E_{[t]}$ such that $\Psi(0) = \mathbb{1}$ and $\Psi(t+1) = \Psi(t)\Psi(1)$ for every $t \in \mathbb{R}$.*[3]

The correspondence between $\mathbf{A}$ and Ψ is determined by the property that for every $v_0 \in E_0$, the function $v(t) := \Psi(t)v_0 \in E_t$ satisfies the differential equation $\mathbf{A}v = 0$ almost everywhere.

PROOF. After choosing a global unitary trivialization, an asymptotic operator $\mathbf{A} = -J_0\partial_t - S(t)$ determines according to Exercise 3.4.10 a unique function $\Psi : \mathbb{R} \to \mathrm{End}(\mathbb{R}^{2n})$ that is absolutely continuous on compact subsets and satisfies the initial value problem

$$\partial_t \Psi(t) = J_0 S(t) \Psi(t), \qquad \Psi(0) = \mathbb{1},$$

where the differential equation is equivalent to $\mathbf{A}\Psi = 0$ and is assumed to hold almost everywhere. Since the function $\mathbb{R} \to \mathrm{End}(\mathbb{R}^{2n}) : t \mapsto J_0 S(t)$ is of class L^∞ and 1-periodic, Ψ is of class $W^{1,\infty}_{\mathrm{loc}}$, and periodicity implies the relation $\Psi(t+1) =$

[3]Saying that the family $\{\Psi(t)\}_{t \in \mathbb{R}}$ is of class $W^{1,\infty}_{\mathrm{loc}}$ means in this context that any choice of smooth trivialization identifies $\{\Psi(t)\}_{t \in \mathbb{R}}$ with a function $\mathbb{R} \to \mathrm{End}(\mathbb{R}^{2n})$ that is of class $W^{1,\infty}_{\mathrm{loc}}$.

$\Psi(t)\Psi(1)$ due to uniqueness of solutions. It remains to show that for all t, $\Psi(t)$ belongs to the linear symplectic group

$$\mathrm{Sp}(2n) := \left\{ B \in \mathrm{GL}(2n,\mathbb{R}) \;\middle|\; \omega_0(Bv,Bw) = \omega_0(v,w) \text{ for all } v,w \in \mathbb{R}^{2n} \right\}.$$

Writing ω_0 in terms of the standard Euclidean inner product g_0 as $\omega_0(v,w) = g_0(J_0 v, w)$, one finds that a matrix $B \in \mathrm{GL}(2n,\mathbb{R})$ belongs to $\mathrm{Sp}(2n)$ if and only if the relation $B^{\mathrm{T}} J_0 B = J_0$ holds. To prove $\Psi(t) \in \mathrm{Sp}(2n)$, one can thus use the differential equation to show that

$$\frac{d}{dt}\Psi^{\mathrm{T}} J_0 \Psi = \Psi^{\mathrm{T}}(S^{\mathrm{T}} - S)\Psi \tag{3.9}$$

holds almost everywhere; since the right hand side vanishes and $\Psi^{\mathrm{T}} J_0 \Psi$ is an absolutely continuous function of t equal to J_0 at $t = 0$, it follows that $\Psi(t)^{\mathrm{T}} J_0 \Psi(t) = J_0$ for all t.

Conversely, suppose $\Psi \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}, \mathrm{End}(\mathbb{R}^{2n}))$ satisfies $\Psi(0) = \mathbb{1}$, $\Psi(t+1) = \Psi(t)\Psi(1)$ and $\Psi(t) \in \mathrm{Sp}(2n)$ for all t. Then by Corollary A.1.12, Ψ is absolutely continuous on compact subsets and thus differentiable almost everywhere, so there is a unique $S : \mathbb{R} \to \mathrm{End}(\mathbb{R}^{2n})$ of class L^∞_{loc} determined almost everywhere by setting $S(t) := -J_0 \dot{\Psi}(t)\Psi(t)^{-1}$. The relation $\Psi(t+1) = \Psi(t)\Psi(1)$ now implies $\dot{\Psi}(t+1) = \dot{\Psi}(t)\Psi(1)$ and thus

$$S(t+1) = -J_0\dot{\Psi}(t+1)\Psi(t+1)^{-1} = -J_0\dot{\Psi}(t)\Psi(1)\Psi(1)^{-1}\Psi(t)^{-1} = S(t),$$

so S is periodic, and the equation $\partial_t \Psi = J_0 S \Psi$ is satisfied almost everywhere by construction. The condition $\Psi(t) \in \mathrm{Sp}(2n)$ then implies $S^{\mathrm{T}} - S = 0$ almost everywhere due to (3.9), hence $\mathbf{A} = -J_0\partial_t - S$ is an asymptotic operator. ☐

We shall refer to the family of symplectic linear maps $\{\Psi(t)\}_{t\in\mathbb{R}}$ induced by an asymptotic operator $\mathbf{A} \in \mathcal{A}(E)$ as the **parallel transport map** of $\mathbf{A}$.

REMARK 3.4.12. The choice to allow discontinuous asymptotic operators in this discussion has the following advantage: every family $\{\Psi(t)\}_{t\in[0,1]}$ of class $W^{1,\infty}$ consisting of symplectic linear maps $\Psi(t) : E_0 \to E_t$ has a unique extension to a family $\{\Psi(t)\}_{t\in\mathbb{R}}$ of class $W^{1,\infty}_{\mathrm{loc}}$ that satisfies the condition $\Psi(t+1) = \Psi(t)\Psi(1)$, thus every such family arises as the parallel transport of some asymptotic operator. This is true in particular for every smooth family $\{\Psi(t)\}_{t\in[0,1]}$, with no need to worry about whether the extension over $\mathbb{R}$ is differentiable at the integers.

3.5. Spectral flow

The goal of this section is to define a notion of spectral flow for asymptotic operators on Hermitian vector bundles over S^1. After fixing a global unitary trivialization, we can restrict our attention to operators $\mathbf{A}$ that act on the space of loops $\eta : S^1 \to \mathbb{R}^{2n}$ by

$$(\mathbf{A}\eta)(t) := -J_0\,\partial_t\eta(t) - S(t)\eta(t), \tag{3.10}$$

where $S : S^1 \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ is a function of class L^∞. We will sometimes refer to operators in this form as **trivialized asymptotic operators**. Regarding $\mathbf{A}$ as an unbounded linear operator on $L^2(S^1, \mathbb{R}^{2n})$ with dense domain $H^1(S^1, \mathbb{R}^{2n})$, we will see that its spectrum consists of isolated real eigenvalues with finite multiplicity. We shall prove:

THEOREM 3.5.1. *Assume* $[-1,1] \to L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})) : s \mapsto S_s$ *is a smooth path, and consider the corresponding* 1*-parameter family of unbounded linear operators*

$$\mathbf{A}_s = -J_0\partial_t - S_s(t) : L^2(S^1, \mathbb{R}^{2n}) \supset H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n}).$$

Then there exists a set of continuous functions

$$\{\lambda_j : [-1,1] \to \mathbb{R}\}_{j \in \mathbb{Z}}$$

such that for every $s \in [-1,1]$*, the spectrum of* $\mathbf{A}_s$ *consists of the numbers* $\{\lambda_j(s)\}_{j\in\mathbb{Z}}$*, each of which is an eigenvalue with finite multiplicity equal to the number of times it is repeated as* j *varies in* $\mathbb{Z}$*.*

Moreover, if additionally $\mathbf{A}_- := \mathbf{A}_{-1}$ *and* $\mathbf{A}_+ := \mathbf{A}_1$ *both have trivial kernel, then there are at most finitely many values of* $j \in \mathbb{Z}$ *for which* $\lambda_j(-1)$ *and* $\lambda_j(1)$ *have different signs, and the integer*

$$\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+) := \#\left\{ j \in \mathbb{Z} \;\middle|\; \lambda_j(-1) < 0 < \lambda_j(1) \right\} - \#\left\{ j \in \mathbb{Z} \;\middle|\; \lambda_j(-1) > 0 > \lambda_j(1) \right\}$$

depends only on $\mathbf{A}_-$ *and* $\mathbf{A}_+$*.*

REMARK 3.5.2. Differentiability of the path $[-1,1] \to L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})) : s \mapsto S_s$ means what you think it means: for every $s \in [-1,1]$, the functions $\frac{S_{s+h} - S_s}{h}$ are L^∞-convergent as $h \to 0$. In practice, we will only need to consider two general classes of smooth paths in Theorem 3.5.1: first, if $S_-, S_+ \in L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$ are given, then the linear interpolation

$$S_s := \frac{1}{2}(1-s)S_- + \frac{1}{2}(1+s)S_+$$

has a constant derivative $\frac{1}{2}(S_+ - S_-) \in L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$ with respect to s and is thus smooth. This example shows that *every* pair of asymptotic operators can be connected by a path that is smooth in the sense of Theorem 3.5.1. The second class of examples will be especially useful for defining generic perturbations of paths of asymptotic operators: it arises from smooth functions $S : [-1,1] \times [0,1] \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$, where for each $s \in [-1,1]$, $S_s := S(s, \cdot)$ need not be periodic but is equal almost everywhere to a uniquely determined element of $L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$. To see that $s \mapsto S_s$ is a smooth map $[-1,1] \to L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$, we observe first that it is continuous since $S(s,t)$ is uniformly continuous on the compact domain $[-1,1] \times [0,1]$, implying that $S_{s+h} \to S_s$ uniformly as $h \to 0$. To prove differentiability at a given point $s \in [-1,1]$, one can use the fundamental theorem of calculus to write $\frac{S_{s+h}(t) - S_s(t)}{h} = \int_0^1 \partial_s S(s + \tau h, t)\, d\tau$ and appeal again to uniform continuity to show that this converges uniformly in t to $\partial_s S(s,t)$ as $h \to 0$. Since $\partial_s S : [-1,1] \times [0,1] \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ is also a uniformly continuous function, it follows that $s \mapsto S_s$ is of class C^1, and smoothness then follows by induction.

REMARK 3.5.3. There is a natural continuous linear inclusion of $L^\infty(S^1, \mathrm{End}(\mathbb{R}^{2n}))$ as a closed subspace of the space of bounded linear operators on $L^2(S^1, \mathbb{R}^{2n})$, identifying each function $S \in L^\infty(S^1, \mathrm{End}(\mathbb{R}^{2n}))$ with the multiplication operator $\eta \mapsto S\eta$.

The smoothness of $s \mapsto S_s$ in Theorem 3.5.1 thus makes $\mathbf{A}_s$ a smooth path in the Banach space of bounded linear operators from $H^1(S^1, \mathbb{R}^{2n})$ to $L^2(S^1, \mathbb{R}^{2n})$.

We will start by giving a more abstract definition of spectral flow as an intersection number between a path of symmetric index 0 Fredholm operators and the subvariety of noninvertible operators. This relies on the general fact that spaces of operators with kernel and cokernel of fixed finite dimensions form smooth finite-codimensional submanifolds in the Banach space of all bounded linear operators. We explain this fact in §3.5.1, and then specialize to the case of symmetric index 0 operators to define the abstract version of spectral flow in §3.5.2. In §3.5.3, we show that the spectra of such operators vary continuously under small perturbations, and in §3.5.4 we specialize further to operators of the form (3.10), and explain how to interpret the abstract definition of spectral flow in terms of eigenvalues crossing the origin in $\mathbb{R}$, leading to a proof of Theorem 3.5.1.

Spectral flow can be defined more generally for certain classes of self-adjoint elliptic partial differential operators (see e.g. [**APS76**, **RS95**]), and standard proofs of its existence typically rely on perturbation results as in [**Kat95**] for the spectra of self-adjoint operators. In the following presentation, we have chosen to avoid making explicit use of self-adjointness, and instead focus on the Fredholm property; in this way, the discussion is mostly self-contained, and does not require any results from [**Kat95**].

3.5.1. Geometry in the space of Fredholm operators. Fix a field

$$\mathbb{F} := \mathbb{R} \text{ or } \mathbb{C}.$$

Given Banach spaces X and Y over $\mathbb{F}$, denote by $\mathscr{L}_{\mathbb{F}}(X, Y)$ the Banach space of bounded $\mathbb{F}$-linear maps from X to Y, with $\mathscr{L}_{\mathbb{F}}(X) := \mathscr{L}_{\mathbb{F}}(X, X)$, and let

$$\mathrm{Fred}_{\mathbb{F}}(X, Y) \subset \mathscr{L}_{\mathbb{F}}(X, Y)$$

denote the open subset consisting of Fredholm operators. Recall that an operator $\mathbf{T} \in \mathscr{L}_{\mathbb{F}}(X, Y)$ is **Fredholm** if its image is closed,[4] and its kernel and cokernel (i.e. the quotient $\operatorname{coker} \mathbf{T} := Y / \operatorname{im} \mathbf{T}$) are both finite dimensional. Its **index** is defined as

$$\mathrm{ind}_{\mathbb{F}}(\mathbf{T}) := \dim_{\mathbb{F}} \ker \mathbf{T} - \dim_{\mathbb{F}} \operatorname{coker} \mathbf{T} \in \mathbb{Z}.$$

The index defines a continuous and thus locally constant function $\mathrm{Fred}_{\mathbb{F}}(X, Y) \to \mathbb{Z}$, and for each $i \in \mathbb{Z}$, we shall denote

$$\mathrm{Fred}^i_{\mathbb{F}}(X, Y) := \left\{ \mathbf{T} \in \mathrm{Fred}_{\mathbb{F}}(X, Y) \;\middle|\; \mathrm{ind}(\mathbf{T}) = i \right\}.$$

We will often have occasion to use the following general construction. Given $\mathbf{T}_0 \in \mathrm{Fred}_{\mathbb{F}}(X, Y)$, one can choose splittings into closed linear subspaces

$$X = V \oplus K, \qquad Y = W \oplus C$$

such that $K = \ker \mathbf{T}_0$, $W = \operatorname{im} \mathbf{T}_0$, the quotient projection $\pi_C : Y \to \operatorname{coker} \mathbf{T}_0$ restricts to $C \subset Y$ as an isomorphism, and $\mathbf{T}_0|_V$ defines an isomorphism from V

[4] It is not strictly necessary to require that $\operatorname{im} \mathbf{T} \subset Y$ be closed, as this follows from the finite-dimensionality of the kernel and cokernel, cf. [**AA02**, Cor. 2.17].

to W. Using these splittings, any other $\mathbf{T} \in \mathrm{Fred}_{\mathbb{F}}(X,Y)$ can be written in block form as

$$\mathbf{T} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix},$$

with $\mathbf{T}_0$ itself written in this way as $\begin{pmatrix} \mathbf{A}_0 & 0 \\ 0 & 0 \end{pmatrix}$ for some Banach space isomorphism $\mathbf{A}_0 : V \to W$. Let $\mathcal{O} \subset \mathrm{Fred}_{\mathbb{F}}(X,Y)$ denote the open neighborhood of $\mathbf{T}_0$ for which the block $\mathbf{A}$ is invertible, and define a map

$$\Phi : \mathcal{O} \to \mathrm{Hom}_{\mathbb{F}}(\ker \mathbf{T}_0, \operatorname{coker} \mathbf{T}_0) : \mathbf{T} \mapsto \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}. \tag{3.11}$$

LEMMA 3.5.4. *The map Φ in (3.11) is smooth, and holomorphic in the case $\mathbb{F} = \mathbb{C}$, and its derivative at $\mathbf{T}_0$ defines a surjective bounded linear operator $\mathscr{L}_{\mathbb{F}}(X,Y) \to \mathrm{Hom}_{\mathbb{F}}(\ker \mathbf{T}_0, \operatorname{coker} \mathbf{T}_0)$ of the form*

$$d\Phi(\mathbf{T}_0)\mathbf{H} = \pi_C \mathbf{H}|_{\ker \mathbf{T}_0} \in \mathrm{Hom}_{\mathbb{F}}(\ker \mathbf{T}_0, \operatorname{coker} \mathbf{T}_0),$$

where π_C denotes the quotient projection $Y \to \operatorname{coker} \mathbf{T}_0$. Moreover, there exists a smooth (and holomorphic if $\mathbb{F} = \mathbb{C}$) function $\Psi : \mathcal{O} \to \mathscr{L}_{\mathbb{F}}(X)$ such that for every $\mathbf{T} \in \mathcal{O}$, $\Psi(\mathbf{T}) : X \to X$ maps $\ker \Phi(\mathbf{T}) \subset \ker \mathbf{T}_0$ isomorphically to $\ker \mathbf{T}$.

PROOF. Smoothness, holomorphicity[5] and the formula for the derivative are easily verified from the given formula for Φ; in particular, since the blocks $\mathbf{B}$ and $\mathbf{C}$ both vanish for $\mathbf{T} = \mathbf{T}_0$, we have

$$d\Phi(\mathbf{T}_0) : \mathscr{L}_{\mathbb{F}}(X,Y) \to \mathrm{Hom}_{\mathbb{F}}(K,C)$$
$$\begin{pmatrix} \mathbf{A}' & \mathbf{B}' \\ \mathbf{C}' & \mathbf{D}' \end{pmatrix} \mapsto \mathbf{D}'.$$

The map $\Psi : \mathcal{O} \to \mathscr{L}_{\mathbb{F}}(X)$ is defined in terms of the splitting $X = V \oplus K$ by

$$\Psi(\mathbf{T}) = \begin{pmatrix} \mathbb{1} & -\mathbf{A}^{-1}\mathbf{B} \\ 0 & \mathbb{1} \end{pmatrix}.$$

This is an isomorphism for each $\mathbf{T}$, with inverse given by

$$\Psi(\mathbf{T})^{-1} = \begin{pmatrix} \mathbb{1} & \mathbf{A}^{-1}\mathbf{B} \\ 0 & \mathbb{1} \end{pmatrix}.$$

Then $\mathbf{T}\Psi(\mathbf{T}) = \begin{pmatrix} \mathbf{A} & 0 \\ \mathbf{C} & \Phi(\mathbf{T}) \end{pmatrix}$, and since $\mathbf{A}$ is invertible, $\ker \mathbf{T}\Psi(\mathbf{T}) = \{0\} \oplus \ker \Phi(\mathbf{T})$. □

PROPOSITION 3.5.5. *For each $i \in \mathbb{Z}$ and each nonnegative integer $k \geqslant i$, the subset*

$$\mathrm{Fred}^{i,k}_{\mathbb{F}}(X,Y) := \left\{ \mathbf{T} \in \mathrm{Fred}^{i}_{\mathbb{F}}(X,Y) \;\middle|\; \dim_{\mathbb{F}} \ker \mathbf{T} = k \text{ and } \dim_{\mathbb{F}} \operatorname{coker} \mathbf{T} = k - i \right\}$$

[5]Holomorphicity in this infinite-dimensional setting means the same thing as usual: $\mathscr{L}_{\mathbb{C}}(X,Y)$ and $\mathrm{Hom}_{\mathbb{C}}(\ker \mathbf{T}_0, \operatorname{coker} \mathbf{T}_0)$ both have natural complex structures if $\mathbf{T}_0 \in \mathrm{Fred}_{\mathbb{C}}(X,Y)$, and we require $d\Phi(\mathbf{T})$ to commute with them for all $\mathbf{T} \in \mathcal{O}$.

admits the structure of a smooth (and complex-analytic if $\mathbb{F} = \mathbb{C}$) finite-codimensional Banach submanifold of $\mathscr{L}_{\mathbb{F}}(X, Y)$, with

$$\operatorname{codim}_{\mathbb{F}} \operatorname{Fred}_{\mathbb{F}}^{i,k}(X, Y) = k(k - i).$$

Moreover, the set

$$X^{i,k} := \left\{ (\mathbf{T}, x) \in \operatorname{Fred}_{\mathbb{F}}^{i,k}(X, Y) \times X \;\middle|\; x \in \ker \mathbf{T} \right\}$$

is a smooth (and holomorphic if $\mathbb{F} = \mathbb{C}$) subbundle of the trivial vector bundle $\operatorname{Fred}_{\mathbb{F}}^{i,k}(X, Y) \times X \to \operatorname{Fred}_{\mathbb{F}}^{i,k}(X, Y)$.

PROOF. Applying the implicit function theorem to the map Φ from Lemma 3.5.4 endows a neighborhood of $\mathbf{T}_0$ in $\Phi^{-1}(0) \subset \operatorname{Fred}_{\mathbb{F}}(X, Y)$ with the structure of a smooth Banach submanifold with

$$\operatorname{codim}_{\mathbb{F}} \Phi^{-1}(0) = \dim_{\mathbb{F}} \operatorname{Hom}_{\mathbb{F}}(\ker \mathbf{T}_0, \operatorname{coker} \mathbf{T}_0) = k(k - i).$$

If $\mathbb{F} = \mathbb{C}$, then Φ is also holomorphic and $\Phi^{-1}(0)$ is thus a complex-analytic submanifold near $\mathbf{T}_0$. Now observe that for every $\mathbf{T} \in \mathcal{O}$,

$$\dim_{\mathbb{F}} \ker \mathbf{T} = \dim_{\mathbb{F}} \ker \Phi(\mathbf{T}) \leqslant \dim_{\mathbb{F}} \ker \mathbf{T}_0 = k,$$

with equality if and only if $\Phi(\mathbf{T}) = 0$, hence, since the index is locally constant, we get $\Phi^{-1}(0) = \operatorname{Fred}_{\mathbb{F}}^{i,k}(X, Y)$ in a neighborhood of $\mathbf{T}_0$.

The vector bundle structure of $X^{i,k}$ can be understood using the smooth (and holomorphic if $\mathbb{F} = \mathbb{C}$) function $\Psi : \mathcal{O} \to \mathscr{L}_{\mathbb{F}}(X)$ from Lemma 3.5.4. This can be interpreted as a smooth (or holomorphic) bundle isomorphism on the trivial X-bundle over $\mathcal{O}$, whose restriction to $\mathcal{O} \cap \operatorname{Fred}^{i,k}(X, Y)$ sends the trivial subbundle with fiber $\ker \mathbf{T}_0 \subset X$ isomorphically to $X^{i,k}$, i.e. this restriction is the inverse of a local trivialization of $X^{i,k}$. □

For real-linear operators of index 0, one can use Prop. 3.5.5 to define the following "relative" invariant. Suppose $\{\mathbf{T}(s) \in \operatorname{Fred}_{\mathbb{R}}^{0}(X, Y)\}_{s \in [-1,1]}$ is a continuous path in the space of Fredholm operators such that $\mathbf{T}_{\pm} := \mathbf{T}(\pm 1) : X \to Y$ are both Banach space isomorphisms. We can then define

$$\mu_{\mathbb{Z}_2}^{\text{spec}}(\{\mathbf{T}(s)\}) \in \mathbb{Z}_2$$

as the parity of the number of times that a generic smooth perturbation of the path $s \mapsto \mathbf{T}(s)$ passes through operators with nontrivial kernel. This depends only on the homotopy class (with fixed end points) of the path—indeed, observe first that generic paths $\{\mathbf{T}(s) \in \operatorname{Fred}_{\mathbb{R}}^{0}(X, Y)\}_{s \in [-1,1]}$ are transverse to $\operatorname{Fred}_{\mathbb{R}}^{0,k}(X, Y)$ for every $k \in \mathbb{N}$, which implies via the codimension formula in Prop. 3.5.5 that they never intersect $\operatorname{Fred}_{\mathbb{R}}^{0,k}(X, Y)$ for $k \geqslant 2$, and their intersections with $\operatorname{Fred}_{\mathbb{R}}^{0,1}(X, Y)$ are transverse and thus isolated. Second, transversality also holds for generic homotopies

$$[0, 1] \times [-1, 1] \to \operatorname{Fred}_{\mathbb{R}}^{0}(X, Y) : (\tau, s) \mapsto \mathbf{T}_{\tau}(s)$$

with fixed end points between any pair of generic paths $\mathbf{T}_0(s)$ and $\mathbf{T}_1(s)$, so that the set of intersections with $\operatorname{Fred}_{\mathbb{R}}^{0,k}(X, Y)$ is again empty for $k \geqslant 2$ and forms a smooth 1-dimensional submanifold in $[0, 1] \times [-1, 1]$ for $k = 1$. This submanifold, moreover, is disjoint from $[0, 1] \times \{-1, 1\}$ since $\mathbf{T}_{\tau}(\pm 1) = \mathbf{T}_{\pm}$, and it is also compact since the set of $\mathbf{T} \in \mathrm{Fred}^0_{\mathbb{R}}(X, Y)$ with nontrivial kernel is a closed subset. We therefore obtain a compact 1-dimensional cobordism between the intersection sets of $\mathbf{T}_0$ and $\mathbf{T}_1$ respectively with $\mathrm{Fred}^{0,1}_{\mathbb{R}}(X, Y)$, implying that the count of intersections modulo 2 does not depend on the choice of generic path within a given homotopy class.

EXERCISE 3.5.6. Convince yourself that the standard results (as in e.g. [Hir94, §3.2] about generic transversality of intersections between smooth maps $f : M \to N$ and submanifolds $A \subset N$ continue to hold—with minimal modifications to the proofs—when N is an infinite-dimensional Banach manifold and $A \subset N$ has finite codimension.

EXERCISE 3.5.7. In the finite-dimensional case, all operators are Fredholm and there is only one homotopy class of paths of Fredholm operators $\{A(s)\}_{s \in [-1,1]}$ between two given isomorphisms $A_\pm \in \mathrm{GL}(n, \mathbb{R})$, so we can abbreviate the invariant defined above as $\mu^{\mathrm{spec}}_{\mathbb{Z}_2}(A_-, A_+) := \mu^{\mathrm{spec}}(\{A(s)\}) \in \mathbb{Z}_2$. Show that $\mu^{\mathrm{spec}}_{\mathbb{Z}_2}(A_-, A_+) = 0$ if and only if $\det A_+$ and $\det A_-$ have the same sign.

3.5.2. Symmetric operators of index zero. We now add the following assumptions to the setup from the previous subsection:

- Y is a Hilbert space $\mathcal{H}$ over $\mathbb{F}$, with inner product denoted by $\langle\ ,\ \rangle_{\mathcal{H}}$;
- X is a dense $\mathbb{F}$-linear subspace $\mathcal{D} \subset \mathcal{H}$, carrying a Banach space structure for which the inclusion $\mathcal{D} \hookrightarrow \mathcal{H}$ is a compact linear operator.

The notation $\mathcal{D} = X$ is motivated by the fact that if $\mathbf{T} \in \mathscr{L}_{\mathbb{F}}(\mathcal{D}, \mathcal{H})$, then we can also regard $\mathbf{T}$ as an **unbounded operator** on $\mathcal{H}$ with domain $\mathcal{D}$ and thus consider the spectrum of $\mathbf{T}$, see §3.5.3 below.

Since $\mathcal{H}$ is a Hilbert space, the space $\mathscr{L}_{\mathbb{F}}(\mathcal{H})$ of bounded linear operators from $\mathcal{H}$ to itself contains a distinguished closed linear subspace

$$\mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{H}) \subset \mathscr{L}_{\mathbb{F}}(\mathcal{H}),$$

consisting of self-adjoint operators. For operators that are bounded from $\mathcal{D}$ to $\mathcal{H}$ but not necessarily defined or bounded on $\mathcal{H}$, there is also the space of **symmetric** operators

$$\mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}) := \left\{ \mathbf{T} \in \mathscr{L}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}) \;\middle|\; \langle x, \mathbf{T}y \rangle_{\mathcal{H}} = \langle \mathbf{T}x, y \rangle_{\mathcal{H}} \text{ for all } x, y \in \mathcal{D} \right\}.$$

Important examples of symmetric operators are those which are self-adjoint (see Remark 3.5.11 below), though for our purposes, it will suffice to restrict attention to symmetric operators that are also Fredholm with index 0. It turns out that the space of symmetric operators in $\mathrm{Fred}^{0,1}_{\mathbb{F}}(\mathcal{D}, \mathcal{H})$ is a canonically co-oriented hypersurface in $\mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H})$, so that the invariant $\mu^{\mathrm{spec}}_{\mathbb{Z}_2}(\{\mathbf{T}(s)\})$ defined above has a natural integer-valued lift when $\mathbf{T}_\pm$ are symmetric. We will need a slightly more specialized version of this statement in order to give a general definition of spectral flow.

In the following, we let

$$\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}) := \mathrm{Fred}^0_{\mathbb{F}}(\mathcal{D}, \mathcal{H}) \cap \mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H})$$

denote the space of symmetric Fredholm operators with index 0, and for $k \in \mathbb{N}$,

$$\mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}) := \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}) \cap \mathrm{Fred}^{0,k}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}).$$

Given $\mathbf{T}_{\mathrm{ref}} \in \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H})$, consider the space

$$\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) := \left\{ \mathbf{T}_{\mathrm{ref}} + \mathbf{K} : \mathcal{D} \to \mathcal{H} \;\middle|\; \mathbf{K} \in \mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{H}) \right\}.$$

Note that the restriction of each $\mathbf{K} \in \mathscr{L}_{\mathbb{F}}(\mathcal{H})$ to $\mathcal{D}$ is a compact operator $\mathcal{D} \to \mathcal{H}$, thus every operator in $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ is a compact perturbation of $\mathbf{T}_{\mathrm{ref}}$, giving rise to a natural inclusion $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) \hookrightarrow \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D}, \mathcal{H})$. The space $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ is also affine over $\mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{H})$, and can thus be regarded naturally as a smooth Banach manifold locally modeled on $\mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{H})$; in particular, its tangent spaces are

$$T_{\mathbf{T}} \left(\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) \right) = \mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{H}).$$

A remark about the case $\mathbb{F} = \mathbb{C}$ is in order: $\mathscr{L}^{\mathrm{sym}}_{\mathbb{C}}(\mathcal{D}, \mathcal{H})$ is a *real*-linear and not a complex subspace of $\mathscr{L}_{\mathbb{C}}(\mathcal{D}, \mathcal{H})$, thus $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{C}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ is a real Banach manifold but does not carry a natural complex structure.

LEMMA 3.5.8. *For any $\mathbf{T} \in \mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H})$ that is Fredholm with index 0, $\ker \mathbf{T}$ is the orthogonal complement of $\operatorname{im} \mathbf{T}$ in $\mathcal{H}$, hence there exist splittings into closed linear subspaces*

$$\mathcal{D} = V \oplus K, \qquad \mathcal{H} = W \oplus C$$

where $K = C = \ker \mathbf{T}$, $W = \operatorname{im} \mathbf{T}$ and $V = W \cap \mathcal{D}$.

PROOF. If $x \in K := \ker \mathbf{T}$, then symmetry implies $\langle x, \mathbf{T}y \rangle_{\mathcal{H}} = \langle \mathbf{T}x, y \rangle_{\mathcal{H}} = 0$ for all $y \in \mathcal{D}$, hence $K \subset W^{\perp}$, where $W := \operatorname{im} \mathbf{T}$. But since $\operatorname{ind} \mathbf{T} = 0$, the dimension of $\ker \mathbf{T}$ equals the codimension of $\operatorname{im} \mathbf{T}$, implying that K already has the largest possible dimension for a subspace that intersects W trivially, and therefore $W \oplus K = \mathcal{H}$. Since K is also a subspace of $\mathcal{D}$ and the latter is a subspace of $\mathcal{H}$, any $x \in \mathcal{D}$ can be written uniquely as $x = v + k$ where $k \in K$ and $v \in W \cap \mathcal{D} =: V$. The continuous inclusion of $\mathcal{D}$ into $\mathcal{H}$ and the fact that W is closed in $\mathcal{H}$ imply that V is a closed subspace of $\mathcal{D}$. ☐

We now have the following modification of Prop. 3.5.5.

PROPOSITION 3.5.9. *For each integer $k \geqslant 0$, the subset*

$$\mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) := \left\{ \mathbf{T} \in \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) \;\middle|\; \dim_{\mathbb{F}} \ker \mathbf{T} = k \right\}$$

is a smooth finite-codimensional Banach submanifold of $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$, with

$$\operatorname{codim}_{\mathbb{R}} \mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) = \begin{cases} k(k+1)/2 & \text{if } \mathbb{F} = \mathbb{R}, \\ k^2 & \text{if } \mathbb{F} = \mathbb{C}, \end{cases}$$

and

$$\mathcal{D}^{\mathrm{sym},k} := \left\{ (\mathbf{T}, x) \in \mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) \times \mathcal{D} \;\middle|\; x \in \ker \mathbf{T} \right\}$$

is a smooth subbundle of the trivial vector bundle $\mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) \times \mathcal{D} \to \mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$. Moreover, the smooth submanifold $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) \subset \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ with codimension one carries a canonical co-orientation.

PROOF. Given $\mathbf{T}_0 \in \mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$, fix the splittings $\mathcal{D} = V \oplus K$ and $\mathcal{H} = W \oplus K$ as in Lemma 3.5.8. Using these in the construction of the map Φ from (3.11) produces a neighborhood $\mathcal{O} \subset \mathrm{Fred}^0_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ of $\mathbf{T}_0$ such that, by Lemma 3.5.4, $\{\mathbf{T} \in \mathcal{O} \mid \dim_{\mathbb{F}} \ker \mathbf{T} = k\} = \Phi^{-1}(0)$, where

$$\Phi : \mathcal{O} \to \mathrm{End}_{\mathbb{F}}(K) : \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \mapsto \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}.$$

Since the splittings are orthogonal, an element $\mathbf{T} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \in \mathcal{O}$ is symmetric if and only if

$$\begin{aligned} \langle x, \mathbf{A}y \rangle_{\mathcal{H}} &= \langle \mathbf{A}x, y \rangle_{\mathcal{H}} && \text{for all } x, y \in V, \\ \langle x, \mathbf{D}y \rangle_{\mathcal{H}} &= \langle \mathbf{D}x, y \rangle_{\mathcal{H}} && \text{for all } x, y \in K, \\ \langle x, \mathbf{B}y \rangle_{\mathcal{H}} &= \langle \mathbf{C}x, y \rangle_{\mathcal{H}} && \text{for all } x \in V,\ y \in K, \\ \langle x, \mathbf{C}y \rangle_{\mathcal{H}} &= \langle \mathbf{B}x, y \rangle_{\mathcal{H}} && \text{for all } x \in K,\ y \in V, \end{aligned}$$

and it follows then that $\Phi(\mathbf{T}) \in \mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K)$, where $\mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K) \subset \mathrm{End}_{\mathbb{F}}(K)$ is the real vector space of symmetric (or Hermitian when $\mathbb{F} = \mathbb{C}$) linear maps on $(K, \langle\ ,\ \rangle_{\mathcal{H}})$. We thus have $\mathcal{O} \cap \mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}}) = \Phi^{-1}(0)$ with Φ regarded as a smooth map $\mathcal{O} \cap \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}}) \to \mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K)$. The derivative at $\mathbf{T}_0$ again takes the form

$$d\Phi(\mathbf{T}_0) : \mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{H}) \to \mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K) : \begin{pmatrix} \mathbf{A}' & \mathbf{B}' \\ \mathbf{C}' & \mathbf{D}' \end{pmatrix} \mapsto \mathbf{D}',$$

where now the block matrix represents an element of $\mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{H})$ with respect to the splitting $\mathcal{H} = W \oplus K$. This operator is evidently surjective, hence by the implicit function theorem, $\Phi^{-1}(0)$ is a smooth Banach submanifold with codimension equal to $\dim_{\mathbb{R}} \mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K)$. The vector bundle structure of $\mathcal{D}^{\mathrm{sym},k}$ can be defined using the map Ψ from Lemma 3.5.4 just as in the non-symmetric case.

Finally, we observe that in the case $k = 1$, the above identifies $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ locally with the zero set of a submersion to $\mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K)$, which is a real 1-dimensional vector space since K is a 1-dimensional vector space over $\mathbb{F}$. The canonical isomorphism

$$\mathbb{R} \to \mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K) : a \mapsto a\mathbb{1}$$

thus determines a co-orientation on $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$. □

The canonical co-orientation of $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ makes it natural to define signed intersection numbers between $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ and smooth paths in the ambient space $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$. The codimensions of $\mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ for each $k \geqslant 2$ are still at least 3, hence large enough to ensure that generic paths or homotopies of paths will never intersect them. The following notion is therefore independent of choices.

DEFINITION 3.5.10. Suppose $\mathbf{T}_+, \mathbf{T}_- \in \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ are both Banach space isomorphisms $\mathcal{D} \to \mathcal{H}$. The **spectral flow**

$$\mu^{\mathrm{spec}}(\mathbf{T}_-, \mathbf{T}_+) \in \mathbb{Z}$$

from $\mathbf{T}_-$ to $\mathbf{T}_+$ is then defined as the signed count of intersections of $\mathbf{T} : [-1,1] \to \mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ with $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},1}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$, where the latter is assumed to carry the co-orientation given by Prop. 3.5.9, and $\mathbf{T} : [-1,1] \to \mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ is any smooth path that is transverse to $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},k}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ for every $k \geqslant 1$ and satisfies $\mathbf{T}(\pm 1) = \mathbf{T}_\pm$.

Note that since $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ is an affine space over $\mathscr{L}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{H})$, all paths in $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ from $\mathbf{T}_-$ to $\mathbf{T}_+$ are homotopic, so one can argue as we did for $\mu_{\mathbb{Z}_2}^{\mathrm{spec}}$ at the end of §3.5.1 that $\mu^{\mathrm{spec}}(\mathbf{T}_-, \mathbf{T}_+)$ is independent of the choice of path.

3.5.3. Perturbation of eigenvalues. Continuing in the setting of the previous subsection, we shall now regard each $\mathbf{T} \in \mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ as an unbounded operator on $\mathcal{H}$ with domain $\mathcal{D}$, see e.g. [RS80, Chapter VIII]. Notice that for each scalar $\lambda \in \mathbb{F}$, the operator $\mathbf{T} - \lambda$ also belongs to $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$. The **spectrum**

$$\sigma(\mathbf{T}) \subset \mathbb{F}$$

of $\mathbf{T}$ is defined as the set of all $\lambda \in \mathbb{F}$ for which $\mathbf{T} - \lambda : \mathcal{D} \to \mathcal{H}$ does not admit a bounded inverse. In particular, $\lambda \in \sigma(\mathbf{T})$ is an **eigenvalue** of $\mathbf{T}$ whenever $\mathbf{T} - \lambda : \mathcal{D} \to \mathcal{H}$ has nontrivial kernel, and the dimension of this kernel is called the **multiplicity** of the eigenvalue. We call λ a **simple eigenvalue** if it has multiplicity 1. By a standard argument familiar to both mathematicians and physicists, the eigenvalues of a symmetric complex-linear operator are always real.

REMARK 3.5.11. The **adjoint** of $\mathbf{T}$ is defined as an unbounded operator $\mathbf{T}^*$ with domain $\mathcal{D}^*$ satisfying

$$\langle x, \mathbf{T}y \rangle_{\mathcal{H}} = \langle \mathbf{T}^* x, y \rangle_{\mathcal{H}} \quad \text{for all } x \in \mathcal{D}^*,\ y \in \mathcal{D},$$

where $\mathcal{D}^*$ is the set of all $x \in \mathcal{H}$ such that there exists $z \in \mathcal{H}$ satisfying $\langle x, \mathbf{T}y \rangle_{\mathcal{H}} = \langle z, y \rangle_{\mathcal{H}}$ for all $y \in \mathcal{D}$. One says that $\mathbf{T}$ is **self-adjoint** if $\mathbf{T} = \mathbf{T}^*$, which means both that $\mathbf{T}$ is symmetric and $\mathcal{D} = \mathcal{D}^*$. In many applications (e.g. in Exercise 3.5.23), the latter amounts to a condition on "regularity of weak solutions". This condition implies that the inclusion $\ker \mathbf{T} \hookrightarrow (\operatorname{im} \mathbf{T})^\perp$—valid for all symmetric operators—is also surjective, so if $\mathbf{T} : \mathcal{D} \to \mathcal{H}$ is Fredholm, it is then automatic that $\operatorname{ind}(\mathbf{T}) = 0$.

PROPOSITION 3.5.12. *Assume* $\mathbf{T}_0 \in \mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$. *Then:*

(1) *Every* $\lambda \in \sigma(\mathbf{T}_0)$ *is an eigenvalue with finite multiplicity.*
(2) *The spectrum* $\sigma(\mathbf{T}_0)$ *is a discrete subset of* $\mathbb{R}$.
(3) *Suppose* $\lambda_0 \in \sigma(\mathbf{T}_0)$ *is an eigenvalue with multiplicity* $m \in \mathbb{N}$ *and* $\epsilon > 0$ *is chosen such that no other eigenvalues lie in* $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$. *Then* $\mathbf{T}_0$ *has a neighborhood* $\mathcal{O} \subset \mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ *such that for all* $\mathbf{T} \in \mathcal{O}$,

$$\sum_{\lambda \in \sigma(\mathbf{T}) \cap [\lambda_0 - \epsilon, \lambda_0 + \epsilon]} m(\lambda) = m,$$

where $m(\lambda) \in \mathbb{N}$ *denotes the multiplicity of* $\lambda \in \sigma(\mathbf{T})$.

PROOF. For every $\lambda \in \mathbb{F}$, $\mathbf{T}_0 - \lambda$ is a Fredholm operator with index 0, so it is a Banach space isomorphism $\mathcal{D} \to \mathcal{H}$ and thus has a bounded inverse if and only if its kernel is trivial. The Fredholm property also implies that the kernel is finite dimensional whenever it is nontrivial, so this proves (1).

For (2) and (3), let us assume $\mathbb{F} = \mathbb{C}$, as the case $\mathbb{F} = \mathbb{R}$ will follow by taking complexifications of real vector spaces. We claim therefore that $\sigma(\mathbf{T}_0)$ is a discrete subset of $\mathbb{C}$. To see this, suppose $\lambda_0 \in \mathbb{R}$ is an eigenvalue of $\mathbf{T}_0$ with multiplicity m, so

$$\mathbf{T}_0 - \lambda_0 \in \operatorname{Fred}_{\mathbb{C}}^{\mathrm{sym},m}(\mathcal{D}, \mathcal{H}).$$

By Lemma 3.5.8, there are splittings $\mathcal{D} = V \oplus K$ and $\mathcal{H} = W \oplus K$ with $K = \ker(\mathbf{T}_0 - \lambda_0)$, $W = \operatorname{im}(\mathbf{T}_0 - \lambda_0)$ and $V = W \cap \mathcal{D}$. Any scalar $\lambda \in \mathbb{C}$ appears in block-diagonal form $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ with respect to these splittings, and the block form for $\mathbf{T}_0$ is thus

$$\mathbf{T}_0 = \begin{pmatrix} \mathbf{A}_0 + \lambda_0 & 0 \\ 0 & \lambda_0 \end{pmatrix}$$

for some Banach space isomorphism $\mathbf{A}_0 : V \to W$. Writing nearby operators $\mathbf{T} \in \operatorname{Fred}_{\mathbb{C}}(\mathcal{D}, \mathcal{H})$ as $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$, we can imitate the construction in (3.11) to produce neighborhoods $\mathcal{O}(\mathbf{T}_0) \subset \operatorname{Fred}_{\mathbb{C}}(\mathcal{D}, \mathcal{H})$ of $\mathbf{T}_0$ and $\mathbb{D}_\epsilon(\lambda_0) \subset \mathbb{C}$ of λ_0, admitting a holomorphic map

$$\Phi : \mathcal{O}(\mathbf{T}_0) \times \mathbb{D}_\epsilon(\lambda_0) \to \operatorname{End}_{\mathbb{C}}(K) : (\mathbf{T}, \lambda) \mapsto (\mathbf{D} - \lambda) - \mathbf{C}\,(\mathbf{A} - \lambda)^{-1}\,\mathbf{B}$$

such that $\ker(\mathbf{T} - \lambda) \cong \ker \Phi(\mathbf{T}, \lambda)$. The set of eigenvalues of $\mathbf{T}_0$ near λ_0 is then the zero set of the holomorphic function

$$\mathbb{D}_\epsilon(\lambda_0) \to \mathbb{C} : \lambda \mapsto \det \Phi(\mathbf{T}_0, \lambda). \tag{3.12}$$

This function cannot be identically zero since there are no eigenvalues outside of $\mathbb{R}$, thus the zero at λ_0 is isolated, proving (2).

To prove (3), note finally that if the neighborhood $\mathcal{O}(\mathbf{T}_0) \subset \operatorname{Fred}_{\mathbb{C}}(\mathcal{D}, \mathcal{H})$ of $\mathbf{T}_0$ is sufficiently small, then for every $\mathbf{T} \in \mathcal{O}(\mathbf{T}_0)$, the holomorphic function

$$f_{\mathbf{T}} : \mathbb{D}_\epsilon(\lambda_0) \to \mathbb{C} : \lambda \mapsto \det \Phi(\mathbf{T}, \lambda)$$

has the same algebraic count of zeroes in $\mathbb{D}_\epsilon(\lambda_0)$, all of which lie in $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ if $\mathbf{T}$ is symmetric. Observe moreover that since

$$\partial_\lambda \Phi(\mathbf{T}_0, \lambda_0) = -\mathbb{1} \in \operatorname{End}_{\mathbb{C}}(K),$$

we are free to assume after possibly shrinking ϵ and $\mathcal{O}(\mathbf{T}_0)$ that $\partial_\lambda \Phi(\mathbf{T}, \lambda)$ is always a nonsingular transformation in $\operatorname{End}_{\mathbb{C}}(K)$. Since $\Phi(\mathbf{T}, \lambda)$ is in $\operatorname{End}_{\mathbb{C}}^{\mathrm{sym}}(K)$ and thus diagonalizable whenever $\mathbf{T}$ is symmetric and $\lambda \in \mathbb{R}$, it follows via Exercise 3.5.13 below that the order of any zero $f_{\mathbf{T}}(\lambda) = 0$ is precisely the multiplicity of λ as an eigenvalue of $\mathbf{T}$. □

EXERCISE 3.5.13. Suppose $\mathcal{U} \subset \mathbb{C}$ is an open subset, $A : \mathcal{U} \to \mathbb{C}^{n \times n}$ is a holomorphic map and $z_0 \in \mathcal{U}$ is a point at which $A(z_0)$ is noninvertible but diagonalizable, and $A'(z_0) \in \mathrm{GL}(n, \mathbb{C})$. Show that $\dim_{\mathbb{C}} \ker A(z_0)$ is the order of the zero of the holomorphic function $\det A : \mathcal{U} \to \mathbb{C}$ at z_0.

The next result implies that for a generic path of symmetric index 0 operators as appears in our definition of $\mu^{\text{spec}}(\mathbf{T}_-, \mathbf{T}_+)$, the spectral flow is indeed a signed count of eigenvalues crossing 0.

PROPOSITION 3.5.14. *Suppose* $\{\mathbf{T}_s \in \text{Fred}^{\text{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\text{ref}})\}_{s\in(-1,1)}$ *is a smooth path and* $\lambda_0 \in \mathbb{R}$ *is a simple eigenvalue of* $\mathbf{T}_0$. *Then:*

(1) For sufficiently small $\epsilon > 0$, *there exists a unique smooth function* $\lambda : (-\epsilon, \epsilon) \to \mathbb{R}$ *such that* $\lambda(0) = \lambda_0$ *and* $\lambda(s)$ *is a simple eigenvalue of* $\mathbf{T}_s$ *for each* $s \in (-\epsilon, \epsilon)$.

(2) The derivative $\lambda'(0)$ *is nonzero if and only if the intersection of the path* $\{\mathbf{T}_s - \lambda_0 \in \text{Fred}^{\text{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\text{ref}})\}_{s\in(-1,1)}$ *with* $\text{Fred}^{\text{sym},1}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\text{ref}})$ *at* $s = 0$ *is transverse, and the sign of* $\lambda'(0)$ *is then the sign of the intersection.*

PROOF. Using the same construction as in the proof of Proposition 3.5.12, we can find small numbers $\epsilon > 0$ and $\delta > 0$ such that

$$\big\{(s, \lambda) \in (-\epsilon, \epsilon) \times (\lambda_0 - \delta, \lambda_0 + \delta) \;\big|\; \lambda \in \sigma(\mathbf{T}_s)\big\} = \Phi^{-1}(0),$$

where

$$\Phi : (-\epsilon, \epsilon) \times (\lambda_0 - \delta, \lambda_0 + \delta) \to \text{End}^{\text{sym}}_{\mathbb{F}}(K) : (s, \lambda) \mapsto (\mathbf{D}_s - \lambda) - \mathbf{C}_s (\mathbf{A}_s - \lambda)^{-1} \mathbf{B}_s,$$

and we write $\mathbf{T}_s = \begin{pmatrix} \mathbf{A}_s & \mathbf{B}_s \\ \mathbf{C}_s & \mathbf{D}_s \end{pmatrix}$ with respect to splittings $\mathcal{D} = V \oplus K$ and $\mathcal{H} = W \oplus K$ with $K = \ker(\mathbf{T}_0 - \lambda_0)$, $W = \text{im}(\mathbf{T}_0 - \lambda_0)$ and $V = W \cap \mathcal{D}$. In saying this, we've implicitly used the assumption that λ_0 is a simple eigenvalue, as it follows that $\dim_{\mathbb{F}} \ker(\mathbf{T} - \lambda)$ cannot be larger than 1 for any $\mathbf{T}$ near $\mathbf{T}_0$ and λ near λ_0, so that $\Phi^{-1}(0)$ catches all nearby eigenvalues. Simplicity also means that $\text{End}^{\text{sym}}_{\mathbb{F}}(K)$ is real 1-dimensional, and we have

$$\partial_s \Phi(0, \lambda_0) = \partial_s \mathbf{D}_s|_{s=0}, \qquad \partial_\lambda \Phi(0, \lambda_0) = -1.$$

The implicit function theorem thus gives $\Phi^{-1}(0)$ near $(0, \lambda_0)$ the structure of a smooth 1-manifold with tangent space at $(0, \lambda_0)$ spanned by the vector

$$\partial_s + (\partial_s \mathbf{D}_s|_{s=0})\, \partial_\lambda,$$

where we are identifying $\partial_s \mathbf{D}_s|_{s=0} \in \text{End}^{\text{sym}}_{\mathbb{F}}(K)$ with a real number via the natural isomorphism $\text{End}^{\text{sym}}_{\mathbb{F}}(K) = \mathbb{R}$. Therefore $\Phi^{-1}(0)$ can be written as the graph of a uniquely determined smooth function λ, whose derivative at zero is a multiple of $\partial_s \mathbf{D}_s|_{s=0}$. This proves both statements in the proposition, since by the proof of Proposition 3.5.9, the intersection of $\{\mathbf{T}_s\}_{s\in(-1,1)}$ with $\text{Fred}^{\text{sym},1}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\text{ref}})$ is transverse if and only if $\partial_s \mathbf{D}_s|_{s=0} \neq 0$, and its sign is then the sign of $\partial_s \mathbf{D}_s|_{s=0}$. ☐

The purpose of the next lemma is to prevent eigenvalues from escaping to $\pm\infty$ under smooth families of operators in $\text{Fred}^{\text{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\text{ref}})$.

LEMMA 3.5.15. *Suppose* $\{\mathbf{K}_s \in \mathscr{L}^{\text{sym}}_{\mathbb{F}}(\mathcal{H})\}_{s\in(a,b)}$ *is a smooth path of symmetric bounded linear operators, and* $\lambda : (a, b) \to \mathbb{R}$ *is a smooth function such that for every* $s \in (a, b)$, $\lambda(s)$ *is a simple eigenvalue of* $\mathbf{T}_s := \mathbf{T}_{\text{ref}} + \mathbf{K}_s \in \text{Fred}^{\text{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\text{ref}})$. *Then*

$$|\dot{\lambda}(s)| \leqslant \|\partial_s \mathbf{K}_s\|_{\mathscr{L}(\mathcal{H})} \qquad \textit{for all } s \in (a, b).$$

PROOF. The operators $\{\mathbf{T}_s - \lambda(s)\}_{s\in(a,b)}$ form a smooth path in the manifold $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$, so Proposition 3.5.9 implies that the family of 1-dimensional eigenspaces $\ker(\mathbf{T}_s - \lambda(s)) \subset \mathcal{D}$ forms a smooth vector bundle over (a,b). We can therefore pick a smooth family of eigenvectors $x(s) \in \ker(\mathbf{T}_s - \lambda(s))$ for $s \in (a,b)$ and normalize them so that $\|x(s)\|_{\mathcal{H}} = 1$ for all s. Then $0 = \partial_s\langle x(s), x(s)\rangle_{\mathcal{H}} = \langle \dot{x}(s), x(s)\rangle_{\mathcal{H}} + \langle x(s), \dot{x}(s)\rangle_{\mathcal{H}}$ and $\lambda(s) = \langle x(s), \mathbf{T}_s x(s)\rangle_{\mathcal{H}}$, so writing $\dot{\mathbf{K}}_s := \partial_s \mathbf{K}_s = \partial_s \mathbf{T}_s$, we have

$$\begin{aligned}\dot{\lambda}(s) &= \partial_s\langle x(s), \mathbf{T}_s x(s)\rangle_{\mathcal{H}} = \langle x(s), \dot{\mathbf{K}}_s x(s)\rangle_{\mathcal{H}} + \langle \dot{x}(s), \mathbf{T}_s x(s)\rangle_{\mathcal{H}} + \langle x(s), \mathbf{T}_s \dot{x}(s)\rangle_{\mathcal{H}} \\ &= \langle x(s), \dot{\mathbf{K}}_s x(s)\rangle_{\mathcal{H}},\end{aligned}$$

as the last two terms in the first line become $\lambda(s)\left[\langle \dot{x}(s), x(s)\rangle_{\mathcal{H}} + \langle x(s), \dot{x}(s)\rangle_{\mathcal{H}}\right] = 0$ since $\mathbf{T}_s$ is symmetric and $\mathbf{T}_s x(s) = \lambda(s)x(s)$. We obtain

$$|\dot{\lambda}(s)| \leqslant \|x(s)\|_{\mathcal{H}} \|\dot{\mathbf{K}}_s\|_{\mathscr{L}(\mathcal{H})} \|x(s)\|_{\mathcal{H}} = \|\dot{\mathbf{K}}_s\|_{\mathscr{L}(\mathcal{H})}.$$

□

3.5.4. Homotopies of eigenvalues. Specializing further, we now set $\mathcal{H}$ and $\mathcal{D}$ equal to the specific real Hilbert spaces

$$\mathcal{H} := L^2(S^1, \mathbb{R}^{2n}), \qquad \mathcal{D} := H^1(S^1, \mathbb{R}^{2n})$$

and consider paths of asymptotic operators $\mathbf{A}_s : H^1(S^1,\mathbb{R}^{2n}) \to L^2(S^1,\mathbb{R}^{2n})$. Concretely, this means setting $\mathbf{T}_{\mathrm{ref}} := -J_0\,\partial_t$ and restricting to compact perturbations $\mathbf{K} \in \mathscr{L}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{H})$ of the form $\mathbf{K}\eta := -S\eta$ for $S : S^1 \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ of class L^∞. The resulting operators $\mathbf{A} = -J_0\,\partial_t - S(t)$ belong to $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ by Lemma 3.4.7, and by Remark 3.5.3, any smooth path $s \mapsto S_s$ in $L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$ gives rise to a smooth path of operators $\mathbf{A}_s = -J_0\partial_t - S_s$ in $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$.

REMARK 3.5.16. We defined the topology of $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ in §3.5.2 by regarding it as an affine space over $\mathscr{L}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{H})$, which means in practice that a family of trivialized asymptotic operators $s \mapsto \mathbf{A}_s$ is considered continuous if and only if $\mathbf{A}_s = -J_0\partial_t - S_s$ for zeroth-order terms S_s that define a continuous family of bounded linear operators on $L^2(S^1,\mathbb{R}^{2n})$. Since the natural inclusion $L^\infty(S^1, \mathrm{End}(\mathbb{R}^{2n})) \hookrightarrow \mathscr{L}(L^2(S^1,\mathbb{R}^{2n}))$ has closed image (cf. Remark 3.5.3), this is equivalent to the continuity of the map $s \mapsto S_s$ into L^∞, which means continuity in the topology of the space of asymptotic operators as specified in Definition 3.4.5.

The proof of Theorem 3.5.1 requires only one more technical ingredient, whose proof is given in Appendix C and should probably be skipped on first reading unless you have already read Chapter 9 or seen similar applications of the Sard-Smale theorem elsewhere. You might however find the result plausible in accordance with the notion that maps from 2-dimensional domains, such as a map of the form

$$(-1,1) \times \mathbb{R} \to \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}}) : (s,\lambda) \mapsto \mathbf{T}_s - \lambda$$

should *generically* not intersect submanifolds that have codimension 3 or more, such as $\mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{R}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ when $k \geqslant 2$.

LEMMA 3.5.17 (see Appendix C). *Fix a smooth path* $[-1,1] \to L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$: $s \mapsto S_s$ *and consider the* 1*-parameter family of symmetric index* 0 *Fredholm operators*

$$\mathbf{A}_s := -J_0\,\partial_t - S_s : H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n})$$

for $s \in [-1,1]$*, assuming* $\mathbf{A}_{\pm 1}$ *are isomorphisms. Then after replacing* S_s *by a family of the form* $\widetilde{S}_s(t) := S_s(t) + B(s,t)$ *for some smooth function* $B : [-1,1] \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ *that vanishes for* $s = \pm 1$ *and may be assumed arbitrarily* C^∞*-small, one can arrange that the following conditions hold:*

(1) For each $s \in (-1,1)$*, all eigenvalues of* $\mathbf{A}_s$ *are simple.*
(2) All intersections of the smooth path

$$(-1,1) \to \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) : s \mapsto \mathbf{A}_s$$

with $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{R}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ *are transverse.*

□

PROOF OF THEOREM 3.5.1. Given a family $\{\mathbf{A}_s\}_{s\in[-1,1]}$ as specified in the theorem, use Lemma 3.5.17 to obtain a C^∞-small zeroth-order perturbation making all eigenvalues simple for $s \in (-1,1)$ and all intersections with $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{R}}(\mathcal{D}, \mathcal{H})$ transverse. Proposition 3.5.14 then implies that the eigenvalues depend smoothly on s, and Lemma 3.5.15 imposes a uniform bound on their derivatives with respect to s so that each one varies only in a bounded subset of $\mathbb{R}$ for $s \in (-1,1)$. The smooth families of eigenvalues for $s \in (-1,1)$ therefore extend to continuous families for $s \in [-1,1]$ since the space of noninvertible Fredholm operators with index 0 is closed. Proposition 3.5.12 ensures moreover that these continuous families hit every eigenvalue with the correct multiplicity at $s = \pm 1$, and by Proposition 3.5.14, the formula for $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+)$ stated in the theorem is correct for the perturbed family with simple eigenvalues and transverse crossings. To obtain the same result for the original family, suppose we have a sequence of perturbations $\{\mathbf{A}^\nu_s = \mathbf{A}_s + B^\nu(s,\cdot)\}_{s\in[-1,1]}$ such that $B^\nu : [-1,1] \times S^1 \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ is C^∞-convergent to 0 as $\nu \to \infty$. Lemma 3.5.15 then provides a uniform C^1-bound for each sequence of smooth families of eigenvalues, so they have C^0-convergent subsequences as $\nu \to \infty$, giving rise to the continuous families in the statement of the theorem. □

REMARK 3.5.18. It is important to understand that the definition of spectral flow depends on the particular co-orientation of $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ that arose in the proof of Prop. 3.5.9. We saw in Prop. 3.5.14 that this is indeed the *right* co-orientation to use if we want to interpret signed intersections with $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ as signed crossing numbers of eigenvalues. In the non-symmetric setting of §3.5.1, one can show that $\mathrm{Fred}^{0,1}_{\mathbb{R}}(X, Y)$ is also co-orientable—this is obvious in the finite-dimensional case since $\mathrm{Fred}^{0,1}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^n)$ is then a regular level set of the determinant function. Moreover, $\mathrm{Fred}^{0,1}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^n)$ is connected (see Exercise 3.5.19 below), so the co-orientation is unique up to a sign. One can therefore lift the $\mathbb{Z}_2$-valued spectral flow of §3.5.1 to $\mathbb{Z}$, but as in Exercise 3.5.7, the result will be a different and much less interesting invariant than $\mu^{\mathrm{spec}}(A_-, A_+)$, as its value will always be either 0 (if $\det A_-$ and $\det A_+$ have the same sign) or ± 1 (if they don't). The reason for the discrepancy is that the canonical co-orientation of $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{R}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ must generally differ on some connected components from any possible co-orientation of the larger hypersurface $\mathrm{Fred}^{0,1}_{\mathbb{R}}(\mathcal{D}, \mathcal{H}) \subset \mathrm{Fred}^{0}_{\mathbb{R}}(\mathcal{D}, \mathcal{H})$.

EXERCISE 3.5.19. Show that the space $\mathrm{Fred}^{0,1}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2)$ of rank 1 matrices in $\mathbb{R}^{2\times 2}$ is connected, but the space $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2)$ of *symmetric* rank 1 matrices is not, and that the canonical co-orientation of $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2)$ coming from Prop. 3.5.9 differs on some components from any possible co-orientation of $\mathrm{Fred}^{0,1}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2) \subset \mathbb{R}^{2\times 2}$. *Hint: A non-symmetric* 2*-by-*2 *matrix may have rank* 1 *even if both of its eigenvalues are* 0*. For symmetric matrices, this cannot happen.*

EXERCISE 3.5.20. Find a smooth path $A : [-1, 1] \to \mathbb{R}^{2\times 2}$ of symmetric matrices such that $A_{\pm} := A(\pm 1)$ are both invertible and $\mu^{\mathrm{spec}}(A_-, A_+) = 2$, but A_+ and A_- can also be connected by a smooth path of (not necessarily symmetric) invertible matrices in $\mathbb{R}^{2\times 2}$.

DEFINITION 3.5.21. The **spectral flow** between two asymptotic operators $\mathbf{A}_{\pm}$ with trivial kernel on a Hermitian vector bundle (E, J, ω) over S^1 is defined by choosing a unitary trivialization to identify both with operators of the form $\mathbf{A}^0_{\pm} = -J_0\partial_t - S_{\pm}(t)$, and then setting $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+) := \mu^{\mathrm{spec}}(\mathbf{A}^0_-, \mathbf{A}^0_+)$, with the latter defined via Theorem 3.5.1.

You should take a moment to convince yourself that the definition of $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+)$ does not depend on the choice of unitary trivialization.

We can now clarify what is meant when we say that critical points of the action functional in SFT or Floer homology have "infinite Morse index" and "infinite Morse co-index":

PROPOSITION 3.5.22. *Every asymptotic operator has infinitely many eigenvalues of both signs.*

PROOF. For $\mathbf{A}_0 := -J_0\partial_t : H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n})$, the eigenvalues can be computed explicitly (see the proof of Theorem 3.7.1 below), so one verifies easily that there are infinitely many of both signs. It is therefore also true for $\mathbf{A}_0 + \epsilon$ for any $\epsilon \in \mathbb{R}$, and this operator has trivial kernel whenever $\epsilon \notin 2\pi\mathbb{Z}$. For any other trivialized asymptotic operator $\mathbf{A}$ with $0 \notin \sigma(\mathbf{A})$, the result then follows from Theorem 3.5.1 since $\mu^{\mathrm{spec}}(\mathbf{A}_0 + \epsilon, \mathbf{A})$ is finite, and this is precisely the signed count of eigenvalues which change sign. The condition $0 \notin \sigma(\mathbf{A})$ can then be lifted by replacing $\mathbf{A}$ with $\mathbf{A} + \epsilon$. □

EXERCISE 3.5.23. Prove:

(a) Asymptotic operators are self-adjoint (as unbounded operators on L^2 with domain H^1) in the sense of Remark 3.5.11.
(b) For any asymptotic operator $\mathbf{A}$ on a bundle E, $L^2(E)$ admits an orthonormal basis of eigenfunctions of $\mathbf{A}$. *Hint: Choose* $\lambda \in \mathbb{R}\backslash\sigma(\mathbf{A})$ *and notice that the resolvent* $(\lambda - \mathbf{A})^{-1}$ *defines a compact operator from* $L^2(E)$ *to itself.*

The following result about the distribution of eigenvalues in the spectrum $\sigma(\mathbf{A}) \subset \mathbb{R}$ will be needed when we discuss exponential decay estimates for solutions to linear Cauchy-Riemann type equations in §4.6.

PROPOSITION 3.5.24. *For any asymptotic operator* $\mathbf{A}$ *and each* $L > 0$*, there exists an upper bound on the number of eigenvalues of* $\mathbf{A}$ *that can lie in any closed interval of length* L *in* $\mathbb{R}$*.*

PROOF. We claim that if the statement holds for *some* asymptotic operator, then it holds for *every* one. Indeed, consider a family $\mathbf{A}_s = -J_0\partial_t - S_s$ defined via a smooth path $[-1,1] \to L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$, and perturb S_s for $s \in (-1,1)$ so that the transversality results of Lemma 3.5.17 hold. The functions $\lambda_j : [-1,1] \to \mathbb{R}$ in the statement of Theorem 3.5.1 are then smooth on $(-1,1)$, and by Lemma 3.5.15, they all satisfy a bound of the form $|\dot{\lambda}_j(s)| \leqslant C$ for some constant $C > 0$ that is independent of j and s. It follows that for any $c \in \mathbb{R}$, the number (counting multiplicity) of eigenvalues of $\mathbf{A}_1$ in $[c - L/2, c + L/2]$ cannot exceed the number (counting multiplicity) of eigenvalues of $\mathbf{A}_{-1}$ in the larger interval $[c - L/2 - 2C, c + L/2 + 2C]$. The result now follows from the fact that it holds for the trivial asymptotic operator $\mathbf{A}_0 := -J_0\partial_t$; as mentioned in the proof of Proposition 3.5.22 above, the eigenvalues of this operator are precisely the integer multiples of 2π, so they are evenly distributed in $\mathbb{R}$. □

3.6. The Conley-Zehnder index of a nondegenerate orbit

We are now in a position to define a suitable replacement for the Morse index in the context of SFT. In the abstract bundle setting, it takes the form of a locally constant function

$$\mu^\tau_{\mathrm{CZ}} : \mathcal{A}^*(E) \to \mathbb{Z}$$

associated to each Hermitian vector bundle (E, ω, J) over S^1 with symplectic trivialization τ, and it will have the important property that its values fully classify the connected components of the space of nondegenerate asymptotic operators. Recall from §3.4 that the space $\mathcal{A}(E)$ of asymptotic operators is an affine space over the Banach space of symmetric bundle endomorphisms of class L^∞, and the nondegenerate operators form an open subset $\mathcal{A}^*(E) \subset \mathcal{A}(E)$ characterized by the condition $\ker \mathbf{A} = \{0\}$. Since the spectrum $\sigma(\mathbf{A}) \subset \mathbb{R}$ consists entirely of eigenvalues, nondegeneracy of $\mathbf{A} \in \mathcal{A}(E)$ is equivalent to the condition

$$0 \notin \sigma(\mathbf{A}).$$

The following general class of nondegenerate asymptotic operators will be used for normalization purposes.

EXAMPLE 3.6.1. Suppose $S \in \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^2)$ is a constant 2-by-2 symmetric matrix with negative determinant. Then the trivialized asymptotic operator $\mathbf{A} = -J_0\partial_t - S : H^1(S^1, \mathbb{R}^2) \to L^2(S^1, \mathbb{R}^2)$ is nondegenerate. To see this, observe that since $\mathbb{R}^2$ is spanned by an orthogonal pair of eigenvectors, one can assume after a suitable change of basis that $J_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $S = \begin{pmatrix} a & 0 \\ 0 & -b \end{pmatrix}$ for some $a, b > 0$. The matrix appearing in the equation $\dot{\eta} = J_0 S \eta$ is then $J_0 S = \begin{pmatrix} 0 & b \\ a & 0 \end{pmatrix}$, which has the real nonzero eigenvalues $\pm\sqrt{ab}$. It follows that the two eigenvalues of $e^{J_0 S}$ lie in

$(0,1)$ and $(1,\infty)$, so there can be no 1-periodic solutions of the equation $\dot{\eta} = J_0 S\eta$, and thus no nontrivial solutions $\eta \in H^1(S^1,\mathbb{R}^2)$ to $\mathbf{A}\eta = 0$.

In higher dimensions, the same result holds for $\mathbf{A} = -J_0\partial_t - S : H^1(S^1,\mathbb{R}^{2n}) \to L^2(S^1,\mathbb{R}^{2n})$ whenever the constant matrix $S \in \operatorname{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ is unitarily equivalent to a diagonal matrix with n positive and n negative eigenvalues (counting multiplicity). Indeed, this condition is equivalent to saying that $\mathbb{R}^{2n} = \mathbb{C}^n$ can be decomposed into orthogonal complex 1-dimensional subspaces such that S restricts to an orientation-reversing isomorphism on each. The solutions to the equation $\dot{\eta} = J_0 S\eta$ are then linear combinations of solutions for the $n = 1$ case described in the previous paragraph.

Choosing the identification $\mathbb{R}^{2n} = \mathbb{C}^n$ so that $J_0 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$, the canonical example of a matrix with the properties described above is $S = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$.

EXERCISE 3.6.2. Show that the space of matrices in $\operatorname{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ unitarily equivalent to a diagonal matrix with n positive and n negative eigenvalues (counting multiplicity) is connected.

DEFINITION 3.6.3. The **Conley-Zehnder index** associates to every trivialized nondegenerate asymptotic operator $\mathbf{A} = -J_0\partial_t - S(t) : H^1(S^1,\mathbb{R}^{2n}) \to L^2(S^1,\mathbb{R}^{2n})$ an integer

$$\mu_{\mathrm{CZ}}(\mathbf{A}) \in \mathbb{Z}$$

determined uniquely by the following properties:

(1) $\mu_{\mathrm{CZ}}(\mathbf{A}) := 0$ for any operator of the form $\mathbf{A} = -J_0\partial_t - S$ where $S \in \operatorname{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ is a constant matrix unitarily equivalent to one that is diagonal with n positive and n negative eigenvalues (counting multiplicity).

(2) For any two nondegenerate operators $\mathbf{A}_\pm$,

$$\mu_{\mathrm{CZ}}(\mathbf{A}_-) - \mu_{\mathrm{CZ}}(\mathbf{A}_+) := \mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+).$$

Example 3.6.1 and Exercise 3.6.2 show that this definition does not depend on the choice of a constant matrix $S \in \operatorname{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ with n positive and n negative eigenvalues, as any two asymptotic operators constructed in this way are homotopic through a family of nondegenerate asymptotic operators, and therefore have zero spectral flow between them.

DEFINITION 3.6.4. Given a nondegenerate asymptotic operator $\mathbf{A} \in \mathcal{A}^*(E)$ on a Hermitian bundle (E, J, ω) over S^1 and a choice of symplectic trivialization τ for (E, J), the **Conley-Zehnder index** of $\mathbf{A}$ with respect to τ is the integer

$$\mu^\tau_{\mathrm{CZ}}(\mathbf{A}) \in \mathbb{Z}$$

defined by choosing any unitary trivialization homotopic to τ (cf. Remark 3.4.4) in order to write $\mathbf{A}$ as an operator $H^1(S^1,\mathbb{R}^{2n}) \to L^2(S^1,\mathbb{R}^{2n})$, and then plugging in Definition 3.6.3.

Suppose next that M is an odd-dimensional manifold with a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$, with its associated hyperplane field $\xi = \ker\lambda$ and Reeb vector field R. For a parametrized closed Reeb orbit $\gamma : S^1 \to M$ and any choice of ω-compatible complex structure J on the bundle $\gamma^*\xi \to S^1$, Exercise 3.3.4 shows that γ is nondegenerate if and only if the associated asymptotic operator $\mathbf{A}_\gamma$ has trivial kernel.

DEFINITION 3.6.5. Given a nondegenerate closed Reeb orbit $\gamma \in \widetilde{\mathcal{P}}(\mathcal{H})$ and a symplectic trivialization τ of $\gamma^*\xi \to S^1$, the **Conley-Zehnder index** of γ relative to τ is defined as

$$\mu_{\mathrm{CZ}}^\tau(\gamma) := \mu_{\mathrm{CZ}}^\tau(\mathbf{A}_\gamma),$$

where $\mathbf{A}_\gamma$ is determined as in Definition 3.3.2 via an auxiliary choice of ω-compatible complex structure J on the bundle $\gamma^*\xi \to S^1$.

It is clear from Definition 3.6.3 that any two trivialized asymptotic operators that are homotopic through a family of nondegenerate operators have the same Conley-Zehnder index, as the existence of such a homotopy implies that the spectral flow between them is zero. For this reason, $\mu_{\mathrm{CZ}}^\tau(\gamma)$ in Definition 3.6.5 depends on the trivialization τ only up to homotopy: any homotopy of trivializations gives rise to a homotopy of trivialized asymptotic operators that are all nondegenerate. For the same reason, $\mu_{\mathrm{CZ}}^\tau(\gamma)$ does not depend on the choice of complex structure J that is used in defining $\mathbf{A}_\gamma$, as the set of all such choices is contractible.

It is customary elsewhere in the literature (see e.g. [**CZ84**, **SZ92**]) to adopt a somewhat different perspective on the Conley-Zehnder index, in which it defines an integer-valued invariant of connected components of the space of **nondegenerate symplectic arcs**

$$\left\{ \Psi \in C^0([0,1], \mathrm{Sp}(2n)) \;\middle|\; \Psi(0) = \mathbb{1} \text{ and } 1 \notin \sigma(\Psi(1)) \right\},$$

where $\mathrm{Sp}(2n) \subset \mathrm{GL}(2n, \mathbb{R})$ denotes the group of linear transformations that preserve the standard symplectic form on $\mathbb{R}^{2n}$. The connection between this notion and our definitions above arises from the parallel transport map of an asymptotic operator (see Proposition 3.4.11), and is elucidated in Definition 3.6.9 below.

EXERCISE 3.6.6. For an asymptotic operator $\mathbf{A} \in \mathcal{A}(E)$ with parallel transport map $\{\Psi(t)\}_{t\in\mathbb{R}}$, show that $\mathbf{A}$ is nondegenerate if and only if $1 \notin \sigma(\Psi(1))$.

EXERCISE 3.6.7. Show that if $[0,1] \to \mathcal{A}(E) : s \mapsto \mathbf{A}_s$ is a continuous family of asymptotic operators with parallel transport maps $\{\Psi_s(t)\}_{t\in\mathbb{R}}$, then $\Psi_s(t)$ depends continuously on $(s,t) \in [0,1]\times\mathbb{R}$. *Hint: Exercise 3.4.10 contains a useful result about the continuous dependence of solutions to ODEs on parameters in the equation.*

EXERCISE 3.6.8. Show that any smooth family $\{\Psi_s(t)\}_{(s,t)\in[0,1]^2}$ of symplectic linear maps $\Psi_s(t) : E_0 \to E_t$ on a Hermitian bundle (E, ω, J) over S^1 uniquely determines a continuous family of asymptotic operators $\{\mathbf{A}_s \in \mathcal{A}(E)\}_{s\in[0,1]}$ whose parallel transport maps over the interval $[0,1]$ are $\{\Psi_s(t)\}_{t\in[0,1]}$. *Hint: See Remarks 3.4.12 and 3.5.2 for a few useful observations.*

DEFINITION 3.6.9. The **Conley-Zehnder index** $\mu_{\mathrm{CZ}}(\Psi) \in \mathbb{Z}$ of a nondegenerate symplectic arc $\Psi : [0,1] \to \mathrm{Sp}(2n)$ is defined as $\mu_{\mathrm{CZ}}(\mathbf{A})$ for any choice of trivialized asymptotic operator $\mathbf{A}$ whose parallel transport restricted to the interval $[0,1]$ is homotopic to Ψ through a family of nondegenerate symplectic arcs.

If you are wondering why Definition 3.6.9 does not simply choose $\mathbf{A}$ to be an asymptotic operator whose parallel transport is Ψ, the answer is that such an operator might not exist since we only assumed Ψ to be of class C^0 and not $W^{1,\infty}$. But by Proposition 3.4.11 and Remark 3.4.12, such an operator will always exist after perturbing $\Psi : [0,1] \to \mathrm{Sp}(2n)$ on $(0,1)$ to make it smooth. That the resulting index is independent of this choice of perturbation then follows from Exercises 3.6.6 and 3.6.8, supplemented by the fact that a continuous homotopy between two smooth paths always admits a C^0-small perturbation to a smooth homotopy.

REMARK 3.6.10. For the asymptotic operator $\mathbf{A}_\gamma$ of a Reeb orbit γ, the corresponding symplectic parallel transport map is given by the linearized Reeb flow along γ, restricted to ξ (cf. Exercise 3.3.4 and Remark 3.3.5). Thus if one prefers as in [SZ92] to speak in terms of nondegenerate symplectic arcs instead of asymptotic operators, $\mu^\tau_{\mathrm{CZ}}(\gamma)$ is equivalently the Conley-Zehnder index of the linearized Reeb flow along γ, expressed via a choice of symplectic trivialization as a nondegenerate arc in $\mathrm{Sp}(2n-2)$.

EXERCISE 3.6.11. Show that if $\mathbf{A}_1$ and $\mathbf{A}_2$ are nondegenerate asymptotic operators on Hermitian bundles E_1 and E_2 respectively, then $\mathbf{A}_1 \oplus \mathbf{A}_2$ defines a nondegenerate asymptotic operator on $E_1 \oplus E_2$, and given trivializations τ_j for $j = 1, 2$,

$$\mu^{\tau_1 \oplus \tau_2}_{\mathrm{CZ}}(\mathbf{A}_1 \oplus \mathbf{A}_2) = \mu^{\tau_1}_{\mathrm{CZ}}(\mathbf{A}_1) + \mu^{\tau_2}_{\mathrm{CZ}}(\mathbf{A}_2).$$

REMARK 3.6.12. Contact geometry in dimension one is not very interesting, but we will nonetheless occasionally need to allow $n = 1$ in the above discussion. On S^1 with its standard orientation, any 1-form that is everywhere positive is contact, and conversely, every stable Hamiltonian structure is of the form $(0, \alpha)$ with $\alpha > 0$. The induced contact structure is then a rank 0 bundle, which has a unique trivialization, and closed Reeb orbits γ are just covers of S^1. The asymptotic operators $\mathbf{A}_\gamma$ for these orbits are thus trivial operators on a 0-dimensional vector space, and in light of the direct sum formula in Exercise 3.6.11, the only reasonable convention is to set

$$\mu_{\mathrm{CZ}}(\gamma) = \mu_{\mathrm{CZ}}(\mathbf{A}_\gamma) = 0.$$

This will lead to the correct Fredholm index formula for punctured holomorphic curves in 2-dimensional symplectic cobordisms, which is just a fancy way of talking about holomorphic branched covers between punctured Riemann surfaces (cf. Proposition 15.3.1).

Here is the main result about Conley-Zehnder indices.

THEOREM 3.6.13. *On any Hermitian bundle $(E, J, \omega) \to S^1$ with symplectic trivialization τ, two nondegenerate asymptotic operators $\mathbf{A}_\pm \in \mathcal{A}^*(E)$ lie in the same connected component of $\mathcal{A}^*(E)$ if and only if $\mu^\tau_{CZ}(\mathbf{A}_+) = \mu^\tau_{CZ}(\mathbf{A}_-)$.*

Similarly, two nondegenerate symplectic arcs $\Psi_\pm : [0,1] \to \mathrm{Sp}(2n)$ are homotopic through a family of nondegenerate symplectic arcs if and only if $\mu_{CZ}(\Psi_+) = \mu_{CZ}(\Psi_-)$.

PROOF. In one direction, both statements are immediate from the definitions. For the other direction of the first statement, we trivialize the bundle and aim to show that if $\mathbf{A}_\pm = -J_0 \partial_t - S_\pm(t)$ satisfy $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+) = 0$, then they are connected

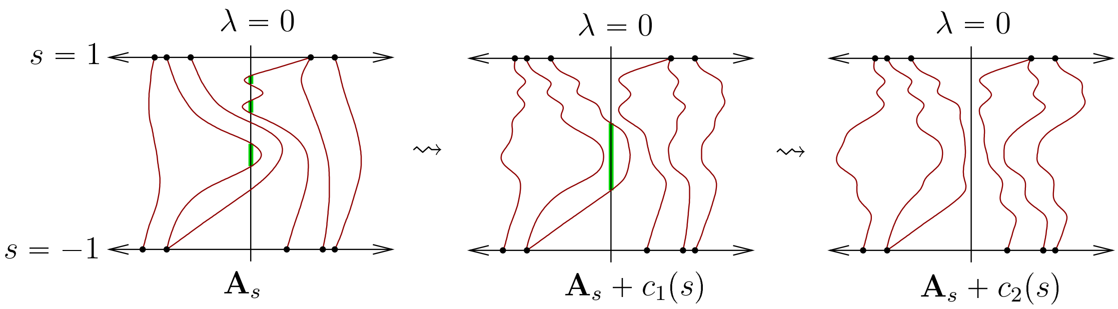

FIGURE 3.1. Modifying a path of asymptotic operators $\{\mathbf{A}_s\}_{s\in[-1,1]}$ with zero spectral flow to produce a path of nondegenerate operators.

by a path of trivialized asymptotic operators for which no eigenvalues cross 0. To see this, we can first choose any path $\{\mathbf{A}_s\}_{s\in[-1,1]}$ of asymptotic operators with $\mathbf{A}_{\pm1} = \mathbf{A}_\pm$ such that (after perturbing it via Lemma 3.5.17) all eigenvalues of $\mathbf{A}_s$ for $-1 < s < 1$ are simple and their crossings with 0 are transverse with respect to the parameter s. Any consecutive pair of crossings with opposite signs can then be eliminated (see Figure 3.1) by changing $\{\mathbf{A}_s\}_{s\in[-1,1]}$ to $\{\mathbf{A}_s + c(s)\}_{s\in[-1,1]}$ for a suitable choice of smooth function $c : [-1,1] \to \mathbb{R}$. Since the spectral flow is zero, one can repeat this modification until one obtains a path with no crossings.

The statement about symplectic arcs follows from the statement about asymptotic operators via Exercise 3.6.7. □

3.7. Winding numbers of eigenfunctions

To compute Conley-Zehnder indices, Exercise 3.6.11 shows that it suffices if we know how to compute them for operators on Hermitian line bundles. The next two theorems provide a useful tool for this.

THEOREM 3.7.1. *Let* $\mathbf{A} = -J_0\partial_t - S(t) : H^1(S^1,\mathbb{R}^2) \to L^2(S^1,\mathbb{R}^2)$, *where* $S \in L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$. *For each* $\lambda \in \sigma(\mathbf{A})$, *denote the corresponding eigenspace by* $E_\lambda \subset H^1(S^1,\mathbb{R}^2)$.

(1) *Every nontrivial eigenfunction* $e_\lambda \in E_\lambda$ *is a continuous nowhere-zero loop in* $\mathbb{R}^2$ *and thus has a well-defined winding number* $\mathrm{wind}(e_\lambda) \in \mathbb{Z}$.
(2) *Any two nontrivial eigenfunctions in the same eigenspace* E_λ *have the same winding number.*
(3) *If* $\lambda, \mu \in \sigma(\mathbf{A})$ *satisfy* $\lambda < \mu$, *then any two nontrivial eigenfunctions* $e_\lambda \in E_\lambda$ *and* $e_\mu \in E_\mu$ *satisfy* $\mathrm{wind}(e_\lambda) \leqslant \mathrm{wind}(e_\mu)$.
(4) *For every* $k \in \mathbb{Z}$, $\mathbf{A}$ *has exactly two eigenvalues (counting multiplicity) for which the corresponding eigenfunctions have winding number equal to* k.

PROOF. We follow the proof given in [**HWZ95**].

Statement (1) follows from the generalized existence/uniqueness result of Exercise 3.4.10 for solutions to the (possibly discontinuous) linear ODE $\partial_t\eta = J_0(S+\lambda)\eta$. In particular, any solution that equals zero at a point must be identically zero, since the trivial function is also a solution.

To prove (2), let η_0 and η_1 be nontrivial eigenfunctions for the same eigenvalue λ. If their winding numbers are different, then there exists $t_0 \in S^1$ at which $\eta_1(t_0)$ is a nonzero real multiple of $\eta_0(t_0)$, so after rescaling, we can assume $\eta_0(t_0) = \eta_1(t_0)$. But η_0 and η_1 are both solutions to the same linear ODE, so this implies $\eta_0(t) = \eta_1(t)$ for all t and thus contradicts the assumption on the winding numbers.

We prove the rest first for the case $S = 0$ and the operator $\mathbf{A}_0 = -J_0\partial_t$. Let us identify $\mathbb{R}^2$ with $\mathbb{C}$ so that J_0 becomes i, and the equation $\mathbf{A}_0\eta = \lambda\eta$ becomes

$$-i\,\partial_t\eta = \lambda\eta, \qquad \text{hence} \qquad \eta(t) = \eta(0)e^{i\lambda t}.$$

This is a well-defined function $S^1 \to \mathbb{C}$ if and only if $\lambda \in 2\pi\mathbb{Z}$, thus $\sigma(\mathbf{A}_0) = 2\pi\mathbb{Z}$, and the winding number of an eigenfunction with eigenvalue $2\pi k$ is k. Statements (2) and (3) for the operator $\mathbf{A}_0$ are now obvious, and (4) follows from the observation that for each $\lambda = 2\pi k$, the eigenspace E_λ has complex dimension one and thus real dimension two, so in this case each eigenvalue is to be counted with multiplicity two.

For the case of an arbitrary trivialized asymptotic operator $\mathbf{A}$, observe first that each eigenspace $E_\lambda \subset H^1(S^1, \mathbb{R}^2)$ is at most 2-dimensional, as the uniqueness of solutions in Exercise 3.4.10 gives a linear injection

$$E_\lambda \hookrightarrow \mathbb{R}^2 : \eta \mapsto \eta(0).$$

Now choose a smooth path of asymptotic operators $\{\mathbf{A}_s\}_{s\in[0,1]}$ from $\mathbf{A}_0 = -J_0\partial_t$ to $\mathbf{A}_1 = \mathbf{A}$, and perturb it as in Lemma 3.5.17 so that all eigenvalues of $\mathbf{A}_s$ for $s \in (0,1)$ are simple. The same argument as in the proof of Theorem 3.5.1 (combining Propositions 3.5.12 and 3.5.14 and Lemma 3.5.15) produces a discrete set of continuous functions $\{\lambda_j : [0,1] \to \mathbb{R}\}_{j\in\mathbb{Z}}$ whose values at each $s \in [0,1]$ are the eigenvalues of $\mathbf{A}_s$ (counted with multiplicity), and since eigenvalues for $s \in (0,1)$ are simple, on the open interval these functions are all smooth and no two of them ever coincide (see Figure 3.2). It follows that the λ_j can be ordered to ensure $\lambda_j(s) \leqslant \lambda_k(s)$ for $j < k$ and all $s \in [0,1]$, with strict inequality $\lambda_j(s) < \lambda_k(s)$ when $s \in (0,1)$. Proposition 3.5.9 now implies that the 1-dimensional eigenspaces corresponding to the eigenvalues $\lambda_j(s)$ for $s \in (0,1)$ also vary smoothly in $H^1(S^1, \mathbb{R}^2)$ with s, so in light of the continuous inclusion $H^1(S^1) \hookrightarrow C^0(S^1)$ from the Sobolev embedding theorem, one can span these eigenspaces with continuous families of nontrivial eigenfunctions $\eta_j(s) \in H^1(S^1, \mathbb{R}^2)$. If we normalize them so that $\|\eta_j(s)\|_{L^2} = 1$ for all j and s, then they also extend continuously to $s = 0$ and $s = 1$ as nontrivial eigenfunctions of $\mathbf{A}_0$ and $\mathbf{A}_1$ respectively, and continuity implies that $\operatorname{wind}(\eta_j(s)) \in \mathbb{Z}$ is independent of s. Finally, we observe that pairs of the functions $\lambda_j : [0,1] \to \mathbb{R}$ may coincide at $s = 1$ since eigenvalues of $\mathbf{A}_1$ need not be simple (this occurs once in Figure 3.2), but the bound $\dim E_\lambda \leqslant 2$ implies that no more than two of these functions can ever have the same value. At $s = 0$, our computation of the spectrum of $\mathbf{A}_0$ shows that exactly two of them attain each value in $2\pi\mathbb{Z}$. It follows that for *every* $s \in [0,1]$, including $s = 1$, the function $\mathbb{Z} \to \mathbb{Z} : j \mapsto \operatorname{wind}(\eta_j(s))$ is monotone increasing and attains every value exactly twice. □

The theorem implies the existence of a well-defined and nondecreasing function

$$\sigma(\mathbf{A}) \to \mathbb{Z} : \lambda \mapsto \operatorname{wind}(\lambda),$$

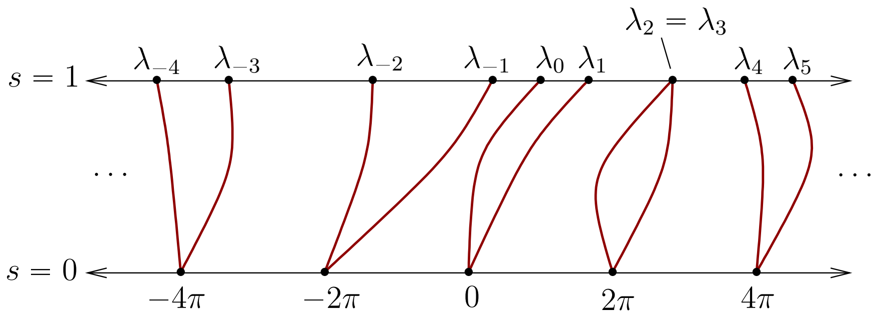

FIGURE 3.2. Generic deformation of eigenvalues from $\mathbf{A}_0 = -J_0\partial_t$ to an arbitrary asymptotic operator $\mathbf{A}$ in the proof of Theorem 3.7.1.

where $\text{wind}(\lambda)$ is defined as $\text{wind}(e_\lambda)$ for any nontrivial $e_\lambda \in E_\lambda$, and this function attains every value exactly twice (counting multiplicity of eigenvalues). Since eigenvalues of $\mathbf{A}$ are isolated, we can therefore associate to any asymptotic operator $\mathbf{A}$ on the trivial Hermitian line bundle the integers

$$\begin{aligned} \alpha_+(\mathbf{A}) &= \min_{\lambda \in \sigma(\mathbf{A}) \cap (0,\infty)} \text{wind}(\lambda) \in \mathbb{Z}, \\ \alpha_-(\mathbf{A}) &= \max_{\lambda \in \sigma(\mathbf{A}) \cap (-\infty,0)} \text{wind}(\lambda) \in \mathbb{Z}, \\ p(\mathbf{A}) &= \alpha_+(\mathbf{A}) - \alpha_-(\mathbf{A}) \geqslant 0. \end{aligned} \tag{3.13}$$

We refer to $\alpha_\pm(\mathbf{A})$ as the (positive and negative) **extremal winding numbers** of $\mathbf{A}$. If $\mathbf{A}$ is nondegenerate, then Theorem 3.7.1 implies that $p(\mathbf{A})$ is either 0 or 1, and it is in this case called the **parity** of $\mathbf{A}$; the following result justifies this terminology.

THEOREM 3.7.2. *If $\mathbf{A}$ is a nondegenerate asymptotic operator on the trivial Hermitian line bundle $S^1 \times \mathbb{R}^2 \to S^1$, then*

$$\mu_{CZ}(\mathbf{A}) = 2\alpha_-(\mathbf{A}) + p(\mathbf{A}) = 2\alpha_+(\mathbf{A}) - p(\mathbf{A}).$$

PROOF. The operator $\mathbf{A}_0 = -J_0\partial_t - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ satisfies $\mu_{\text{CZ}}(\mathbf{A}_0) = 0$ by definition, and it has two constant eigenfunctions with eigenvalues of opposite signs, hence

$$\alpha_-(\mathbf{A}_0) = \alpha_+(\mathbf{A}_0) = 0,$$

consistent with the stated formula. The general case then follows by choosing a generic (in the sense of Lemma 3.5.17) path from $\mathbf{A}_0$ to $\mathbf{A}$ and observing that all three expressions in the stated formula change in the same way whenever a simple eigenvalue crosses zero. □

For any Hermitian line bundle (E, J, ω) over S^1 with an asymptotic operator $\mathbf{A}$, we can similarly choose a symplectic trivialization τ to define $\alpha_\pm^\tau(\mathbf{A}) \in \mathbb{Z}$ and $p(\mathbf{A}) = \alpha_+^\tau(\mathbf{A}) - \alpha_-^\tau(\mathbf{A}) \geqslant 0$; note that the dependence on τ cancels out in the last formula, so that $p(\mathbf{A})$ is independent of choices. We then can associate to any closed

Reeb orbit $\gamma \in \mathcal{P}(\mathcal{H})$ with a trivialization τ of $\gamma^*\xi$ the integers $\alpha_\pm^\tau(\gamma)$ and $p(\gamma)$, such that if γ is nondegenerate, then $p(\gamma) \in \{0,1\}$ and

$$\mu_{\mathrm{CZ}}^\tau(\gamma) = 2\alpha_-^\tau(\gamma) + p(\gamma) = 2\alpha_+^\tau(\gamma) - p(\gamma).$$

EXERCISE 3.7.3. Given a Hermitian vector bundle $(E, J, \omega) \to S^1$ with two unitary trivializations $\tau_j : E \to S^1 \times \mathbb{R}^{2n}$ for $j = 1, 2$, denote by

$$\deg(\tau_1 \circ \tau_2^{-1}) \in \mathbb{Z}$$

the winding number of $\det g : S^1 \to \mathrm{U}(1) \subset \mathbb{C}\backslash\{0\}$, where $g : S^1 \to \mathrm{U}(n)$ is the transition map appearing in the formula $\tau_1 \circ \tau_2^{-1}(t, v) = (t, g(t)v)$. Show that for any nondegenerate asymptotic operator $\mathbf{A}$ on (E, J, ω),

$$\mu_{\mathrm{CZ}}^{\tau_2}(\mathbf{A}) = \mu_{\mathrm{CZ}}^{\tau_1}(\mathbf{A}) + 2\deg(\tau_2 \circ \tau_1^{-1}).$$

Exercise 3.7.3 provides the useful formula

$$\mu_{\mathrm{CZ}}^{\tau_2}(\gamma) = \mu_{\mathrm{CZ}}^{\tau_1}(\gamma) + 2\deg(\tau_2 \circ \tau_1^{-1})$$

for any two symplectic trivializations τ_1, τ_2 of ξ along a nondegenerate Reeb orbit γ, where $\deg(\tau_2 \circ \tau_1^{-1})$ can be defined in this case after homotopies of τ_1 and τ_2 to unitary trivializations. In particular, this shows that the **parity**

$$\mu_{\mathrm{CZ}}^{\mathbb{Z}_2}(\gamma) := [\mu_{\mathrm{CZ}}^\tau(\gamma)] \in \mathbb{Z}_2$$

of the orbit does not depend on a choice of trivialization. We sometimes refer to **even orbits** and **odd orbits** accordingly.

To any closed Reeb orbit of period $T > 0$ parametrized by a loop $\gamma : S^1 \to M$ with $\dot\gamma = T \cdot R(\gamma)$, one can associate a Reeb orbit of period kT for each $k \in \mathbb{N}$, parametrized by

$$\gamma^k : S^1 \to M : t \mapsto \gamma(kt).$$

We say γ^k is the **k-fold cover** of γ, and define the **covering multiplicity**

$$\mathrm{cov}(\gamma) \in \mathbb{N}$$

of an orbit γ to be the largest number k such that γ is the k-fold cover of another orbit. We say γ is **simple** or **simply covered** if $\mathrm{cov}(\gamma) = 1$, and otherwise call it **multiply covered**. Notice that sections $\eta \in \Gamma(\gamma^*\xi)$ also have k-fold covers $\eta^k \in \Gamma((\gamma^k)^*\xi)$, defined by $\eta^k(t) = \eta(kt)$.

If $\mathbf{A}_\gamma$ has parallel transport map $\{\Psi_\gamma(t)\}_{t\in\mathbb{R}}$, then the parallel transport map of $\mathbf{A}_{\gamma^k}$ for each $k \in \mathbb{N}$ is given by

$$\Psi_{\gamma^k}(t) = \Psi_\gamma(kt).$$

If $\mathbf{A}_\gamma$ is given in some choice of unitary trivialization of $\gamma^*\xi$ by $-J_0\partial_t - S(t)$, then using the pullback of the same trivialization on $(\gamma^k)^*\xi$, one now deduces via Proposition 3.4.11 that $\mathbf{A}_{\gamma^k}$ is given by

$$\mathbf{A}_{\gamma^k} = -J_0\partial_t - kS(kt).$$

This implies:

PROPOSITION 3.7.4. *Given a Reeb orbit γ and $k \in \mathbb{N}$, the k-fold cover of each eigenfunction e_λ of $\mathbf{A}_\gamma$ with $\mathbf{A}_\gamma e_\lambda = \lambda e_\lambda$ is an eigenfunction of $\mathbf{A}_{\gamma^k}$ satisfying $\mathbf{A}_{\gamma^k} e_\lambda^k = k\lambda e_\lambda^k$.* □

EXERCISE 3.7.5. Assume $\dim M = 3$.

(a) If γ is a closed Reeb orbit in M and τ is the pullback under $S^1 \to S^1 : t \mapsto kt$ of a trivialization of $\gamma^*\xi \to S^1$, deduce from Theorem 3.7.1 that a nontrivial eigenfunction e_λ of $\mathbf{A}_{\gamma^k}$ is a k-fold cover if and only if $\mathrm{wind}^\tau(e_\lambda)$ is divisible by k.

(b) Under the same assumptions, show that for any nontrivial eigenfunction e_λ of $\mathbf{A}_{\gamma^k}$,

$$\mathrm{cov}(e_\lambda) := \max\{m \in \mathbb{N} \mid e_\lambda \text{ is an } m\text{-fold cover}\} = \gcd(k, \mathrm{wind}^\tau(e_\lambda)).$$

(c) Show that if γ is a nondegenerate Reeb orbit with even Conley-Zehnder index, then so are all of its multiple covers.

3.8. Elliptic and hyperbolic orbits

In this section we develop a few more techniques for the computation of Conley-Zehnder indices, focusing mainly (but not exclusively) on the low-dimensional case. We will need to use the following notation for the "floor" and "ceiling" of a real number $\theta \in \mathbb{R}$,

$$\mathbb{Z} \ni \lfloor \theta \rfloor \leqslant \theta \leqslant \lceil \theta \rceil \in \mathbb{Z} \qquad \text{for} \qquad \theta \in \mathbb{R},$$

where by definition $\lceil \theta \rceil = \lfloor \theta \rfloor + 1$ whenever $\theta \notin \mathbb{Z}$, and $\lceil \theta \rceil = \lfloor \theta \rfloor$ for $\theta \in \mathbb{Z}$.

DEFINITION 3.8.1. Assume M is a 3-manifold with a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ and the associated data $\xi = \ker \lambda$ and R, $\gamma : S^1 \to M$ parametrizes a nondegenerate Reeb orbit of period $T \equiv \lambda(\dot{\gamma}) > 0$, and $\varphi_*^T : \xi_{\gamma(0)} \to \xi_{\gamma(0)}$ denotes the restriction of the linearized time-T Reeb flow to $\xi_{\gamma(0)}$. Let $\lambda_1, \lambda_2 \in \mathbb{C}$ denote the two eigenvalues of φ_*^T, which satisfy $\lambda_1\lambda_2 = 1$ since φ_*^T is symplectic, and $\lambda_1 \neq 1 \neq \lambda_2$ since γ is nondegenerate. Then γ is called

(1) **positive hyperbolic** if $\lambda_1, \lambda_2 > 0$;
(2) **negative hyperbolic** if $\lambda_1, \lambda_2 < 0$;
(3) **elliptic** if $\lambda_1, \lambda_2 \notin \mathbb{R}$.

Similarly, a nondegenerate symplectic arc $\Psi : [0, 1] \to \mathrm{Sp}(2)$ or a nondegenerate asymptotic operator $\mathbf{A}$ on a Hermitian line bundle $(E, \omega, J) \to S^1$ with parallel transport map $\{\Psi(t)\}_{t \in \mathbb{R}}$ can be called **positive/negative hyperbolic** or **elliptic**[6] according to the properties of the eigenvalues $\lambda_1, \lambda_2 \in \sigma(\Psi(1))$.

Observe that every nondegenerate orbit must satisfy exactly one of the three conditions in Definition 3.8.1, each of which encodes qualitative aspects of the dynamics in the neighborhood of that orbit. This is a large subject, which we will not get into here except to make some observations about the invariance of these properties under deformations. In the elliptic case, the two eigenvalues λ_1, λ_2 necessarily form a conjugate pair on the unit circle $\mathrm{U}(1) \subset \mathbb{C}$, and in both other cases, they must both lie on the same side of 0 in $\mathbb{R}\backslash\{0\}$. Since $\mathrm{U}(1) \cup (-\infty, 0)$ and $(0, \infty)\backslash\{1\}$ are each open and closed subsets of $(\mathbb{R}\backslash\{0, 1\}) \cup \mathrm{U}(1) \subset \mathbb{C}$ (see Figure 3.3), it follows that

[6]Caution: the use of the word "elliptic" in this context is unrelated to its meaning in the theory of partial differential operators (which will be relevant from Chapter 4 onwards). Every asymptotic operator is elliptic in the latter sense, but not in the dynamical sense under consideration here.

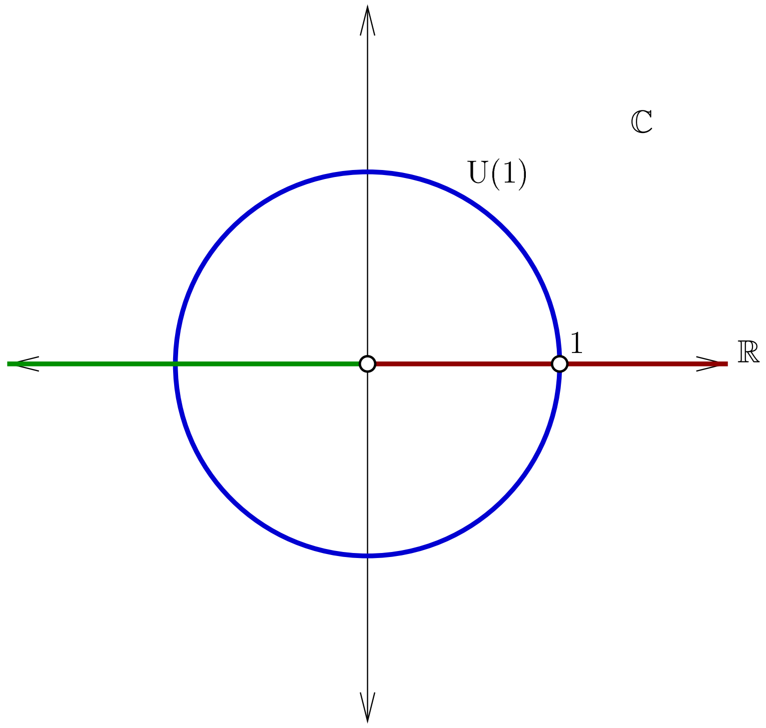

FIGURE 3.3. Every symplectic linear map in dimension two has spectrum contained in $(\mathbb{R}\backslash\{0\}) \cup \mathrm{U}(1) \subset \mathbb{C}$, but eigenvalues cannot move between $(0, \infty)$ and $(-\infty, 0) \cup \mathrm{U}(1)$ without crossing 1.

under any smooth deformation of stable Hamiltonian structures $\{\mathcal{H}_s\}_{s\in[0,1]}$ with an accompanying smooth family $\{\gamma_s \in \mathcal{P}(\mathcal{H}_s)\}_{s\in[0,1]}$ of nondegenerate Reeb orbits, the orbits cannot deform from positive hyperbolic to either of the other two categories, i.e. γ_0 is positive hyperbolic if and only if γ_1 is. Indeed, since eigenvalues of the linearized flow deform continuously as functions of s, they could not pass from the positive real line to the circle or negative real line without crossing 1, which would mean degeneracy. We will see in Theorem 3.9.1 below that this invariance property is related to the odd/even parity of the Conley-Zehnder index. It's worth looking first at a couple of concrete examples.

EXAMPLE 3.8.2. On the trivial Hermitian line bundle over S^1, consider an asymptotic operator of the form

$$\mathbf{A} = -J_0\partial_t - \epsilon$$

for $\epsilon \in \mathbb{R}$. The spectrum and eigenfunctions of this operator were computed for $\epsilon = 0$ in the proof of Theorem 3.7.1; for general $\epsilon \in \mathbb{R}$, the eigenfunctions are the same, but the spectrum is shifted to $2\pi\mathbb{Z} - \epsilon$, implying that $\mathbf{A}$ is degenerate if and only if $\epsilon \in 2\pi\mathbb{Z}$. If $\epsilon \notin 2\pi\mathbb{Z}$, then inspecting the winding of the eigenfunctions and applying Theorem 3.7.2 gives

$$\mu_{\mathrm{CZ}}(\mathbf{A}) = 2\lfloor \epsilon/2\pi \rfloor + 1.$$

The parallel transport map $\Psi : \mathbb{R} \to \mathrm{Sp}(2)$ for this operator is given by

$$\Psi(t) = e^{\epsilon t J_0} = \begin{pmatrix} \cos(\epsilon t) & -\sin(\epsilon t) \\ \sin(\epsilon t) & \cos(\epsilon t) \end{pmatrix},$$

so $\sigma(\Psi(1)) = \{e^{i\epsilon}, e^{-i\epsilon}\}$, and $\mathbf{A}$ is therefore elliptic whenever $\epsilon \notin \pi\mathbb{Z}$, and negative hyperbolic for $\epsilon \in \pi\mathbb{Z}\backslash 2\pi\mathbb{Z}$.

EXERCISE 3.8.3. Show that the asymptotic operators of Example 3.8.2 arise in the following concrete example of a closed Reeb orbit: $\gamma : S^1 \to S^1 \times \mathbb{R}^2 : t \mapsto (t, 0)$ with positive contact form $\alpha = f(\rho)\, d\theta + g(\rho)\, d\phi$ written in positively-oriented coordinates $(\theta, (\rho, \phi)) \in S^1 \times \mathbb{R}^2$, where (ρ, ϕ) are the standard polar coordinates on $\mathbb{R}^2$ and $f, g : [0, \infty) \to \mathbb{R}$ are suitably chosen functions. Assuming $f(0) > 0$ and $g(0) = 0$, find an explicit formula for the offset $\epsilon \in \mathbb{R}$ in terms of the ratio $f''(0)/g''(0)$.

EXAMPLE 3.8.4. The asymptotic operator

$$\mathbf{A} = -J_0 \partial_t - \begin{pmatrix} \epsilon & 0 \\ 0 & -\epsilon \end{pmatrix}$$

for $\epsilon > 0$ has parallel transport map $\Psi(t) = \begin{pmatrix} \cosh(\epsilon t) & \sinh(\epsilon t) \\ \sinh(\epsilon t) & \cosh(\epsilon t) \end{pmatrix}$ and thus $\sigma(\Psi(1)) = \{e^{\epsilon t}, e^{-\epsilon t}\}$, so it is positive hyperbolic. It also satisfies $\mu_{\mathrm{CZ}}(\mathbf{A}) = 0$ by the definition of the Conley-Zehnder index.

Observe that by changing global trivializations as in Exercise 3.7.3, one can produce from this example a positive hyperbolic asymptotic operator with arbitrary even Conley-Zehnder index; indeed, changing trivializations alters the path $\Psi : [0, 1] \to \mathrm{Sp}(2)$, but does not change $\Psi(1)$.

3.9. Some computational tools

We can now establish a useful topological criterion for computing the Conley-Zehnder index of a nondegenerate symplectic arc $\Psi : [0, 1] \to \mathrm{Sp}(2)$. Given $v \in \mathbb{R}^2\backslash\{0\}$, use the canonical identification $\mathbb{R}^2 = \mathbb{C}$ to write $\Psi(t)v = r_v(t)e^{i\theta_v(t)}$ for some continuous functions $r_v(t) > 0$ and $\theta_v(t) \in \mathbb{R}$. The **winding interval** of Ψ is defined as the set

$$\Delta(\Psi) := \left\{ \frac{\theta_v(1) - \theta_v(0)}{2\pi} \;\middle|\; v \in \mathbb{R}^2\backslash\{0\} \right\} \subset \mathbb{R}.$$

Notice that $\theta_v(1) - \theta_v(0)$ depends only on the normalized vector $v/|v| \subset S^1 \subset \mathbb{R}^2$, and this dependence is continuous, thus $\Delta(\Psi) \subset \mathbb{R}$ is indeed a connected and compact set, i.e. a closed bounded interval.

THEOREM 3.9.1. *Given a nondegenerate symplectic arc $\Psi : [0, 1] \to \mathrm{Sp}(2)$, exactly one of the following holds:*

- *Ψ is elliptic or negative hyperbolic and there exists an integer $k \in \mathbb{Z}$ with*

$$\mu_{CZ}(\Psi) = 2k + 1 \qquad \text{and} \qquad \Delta(\Psi) \subset (k, k + 1).$$

- Ψ *is positive hyperbolic and there exists an integer* $k \in \mathbb{Z}$ *with*

$$\mu_{CZ}(\Psi) = 2k \qquad \text{and} \qquad \Delta(\Psi) \cap \mathbb{Z} = \{k\}.$$

PROOF. Observe first that by explicit calculation, the stated formula relating $\mu_{\mathrm{CZ}}(\Psi)$ and the winding interval $\Delta(\Psi)$ is correct for each of the models in Examples 3.8.2 and 3.8.4, which cover all possible values of the Conley-Zehnder index.

Next, if $\Psi : [0,1] \to \mathrm{Sp}(2)$ is an arbitrary continuous map with $\Psi(0) = \mathbb{1}$, then the condition $\Delta(\Psi) \cap \mathbb{Z} \neq \varnothing$ is equivalent to the existence of a vector $v \in \mathbb{R}^2 \backslash \{0\}$ for which $\Psi(1)v$ is a positive multiple of v, meaning $\Psi(1)$ has a positive eigenvalue. This is true if and only if Ψ is either degenerate or positive hyperbolic.

Now suppose Ψ is an arbitrary nondegenerate symplectic arc. If $\mu_{\mathrm{CZ}}(\Psi)$ is odd, then Theorem 3.6.13 provides a homotopy $\{\Psi_s : [0,1] \to \mathrm{Sp}(2)\}_{s \in [0,1]}$ through nondegenerate symplectic arcs with $\Psi_1 = \Psi$ so that Ψ_0 is the parallel transport of one of the elliptic models in Example 3.8.2. Since $\sigma(\Psi_0(1)) \subset \mathrm{U}(1)\backslash\{1\}$ and $\sigma(\Psi_s(1))$ cannot contain 1 for any $s \in [0,1]$, it follows that $\sigma(\Psi(1))$ is also contained in either the unit circle or the negative real line, so Ψ is elliptic or negative hyperbolic. If instead $\mu_{\mathrm{CZ}}(\Psi)$ is even, then a similar argument using the positive hyperbolic models of Example 3.8.4 implies that Ψ is positive hyperbolic.

Returning to the case $\mu_{\mathrm{CZ}}(\Psi) \notin 2\mathbb{Z}$, we now know that the nondegenerate symplectic arcs Ψ_s in the homotopy of the previous paragraph are never positive hyperbolic, thus $\Delta(\Psi_s) \cap \mathbb{Z} = \varnothing$ for every s. Since the winding intervals $\Delta(\Psi_s)$ depend continuously on s, it follows that $\Delta(\Psi)$ is contained within the same open unit interval $(k, k+1)$ as $\Delta(\Psi_0)$, so the stated formula for $\mu_{\mathrm{CZ}}(\Psi)$ now follows from the fact that it holds for the models in Example 3.8.2.

Finally, if $\mu_{\mathrm{CZ}}(\Psi) \in 2\mathbb{Z}$, then all Ψ_s in the homotopy are positive hyperbolic, implying that $\Psi_s(1)$ for each s has two simple eigenvalues $\lambda_s^- \in (0,1)$ and $\lambda_s^+ \in (1,\infty)$, whose corresponding eigenvectors $v_s^\pm$ span $\mathbb{R}^2$. Since the eigenvalues are simple, all of this data varies continuously with s, and one therefore obtains two homotopies of paths $\{v_s^\pm(t) := \Psi_s(t)v_s^\pm \in \mathbb{R}^2\backslash\{0\}\}_{s\in[0,1]}$, such that $v_s^+(t)$ and $v_s^-(t)$ are linearly independent for all s and t, and the normalized paths $t \mapsto v_s^\pm(t)/|v_s^\pm(t)|$ are loops. This implies that their total winding is the same for all $s \in [0,1]$ and for both signs, thus $\Delta(\Psi)$ contains only one integer, and it is the same integer that $\Delta(\Psi_0)$ contains. Once again, the stated formula for $\mu_{\mathrm{CZ}}(\Psi)$ now follows from the fact that it holds for the model in Example 3.8.4. $\square$

Using the direct sum property in Exercise 3.6.11, one derives from Theorem 3.9.1 the following alternative characterization of the Conley-Zehnder index in higher dimensions (cf. [**FH93**, Proposition 5] or [**Sch95**, Theorem 3.3.7]):

COROLLARY 3.9.2. *Using the canonical identification* $\mathbb{R}^2 = \mathbb{C}$*, consider the paths in* $\mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$ *defined by*

$$\alpha(t) := \begin{pmatrix} e^{\pi i t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\pi i t} \end{pmatrix}, \qquad \beta(t) := \begin{pmatrix} \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} & 0 & \cdots & 0 \\ 0 & e^{\pi i t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\pi i t} \end{pmatrix},$$

and the loop

$$\sigma(t) := \begin{pmatrix} e^{2\pi i t} & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Defining the standard symplectic form on $\mathbb{C}^n$ *by* $\omega_0 = \operatorname{Re}\langle i\cdot, \cdot\rangle$ *and identifying* $\mathbb{R}^{2n}$ *with* $\mathbb{C}^n$ *makes all of these into paths in* $\operatorname{Sp}(2n)$*. Then every nondegenerate symplectic arc* $\Psi : [0,1] \to \operatorname{Sp}(2n)$ *is homotopic through nondegenerate symplectic arcs to exactly one of the arcs* $\Phi_k(t) := \sigma(t)^k\alpha(t)$ *or* $\Psi_k(t) := \sigma(t)^k\beta(t)$ *for some* $k \in \mathbb{Z}$*, which satisfy*

$$\mu_{CZ}(\Phi_k) = 2k + n, \qquad \mu_{CZ}(\Psi_k) = 2k + n - 1.$$

□

3.10. Multiple covers and the monodromy angle

In many applications, it is important to understand how the Conley-Zehnder index scales when an orbit γ is replaced by its multiple covers γ^k for $k \in \mathbb{N}$. A first guess would be $\mu_{\mathrm{CZ}}(\gamma^k) = k\mu_{\mathrm{CZ}}(\gamma)$, which turns out to be true in the hyperbolic cases, but the behavior of elliptic orbits is more complicated. Let us frame the discussion in terms of asymptotic operators, and associate to each $k \in \mathbb{N}$ and each operator $\mathbf{A} \in \mathcal{A}(E)$ on a Hermitian vector bundle $(E, \omega, J) \to S^1$ its **k-fold cover**

$$\mathbf{A}^k \in \mathcal{A}(\pi_k^* E), \qquad \text{where} \qquad \pi_k : S^1 \to S^1 : t \mapsto kt,$$

defined via the condition that if $\mathbf{A}$ has parallel transport map $\{\Psi(t)\}_{t\in\mathbb{R}}$, then the parallel transport map of $\mathbf{A}^k$ is $\{\Psi_k(t)\}_{t\in\mathbb{R}}$ with $\Psi_k(t) := \Psi(kt)$. In particular, if $\mathbf{A} = \mathbf{A}_\gamma$ for a Reeb orbit γ, then $\mathbf{A}^k = \mathbf{A}_{\gamma^k}$. If we choose a unitary trivialization of E to write $\mathbf{A} = -J_0\partial_t - S(t) : H^1(S^1, \mathbb{R}^2) \to L^2(S^1, \mathbb{R}^2)$, then using the pullback of this trivialization on $\pi_k^* E$ identifies $\mathbf{A}^k$ with

$$\mathbf{A}^k = -J_0\partial_t - kS(kt) : H^1(S^1, \mathbb{R}^2) \to L^2(S^1, \mathbb{R}^2).$$

LEMMA 3.10.1. *For every* $k \in \mathbb{N}$ *and every trivialized asymptotic operator* $\mathbf{A} : H^1(S^1, \mathbb{R}^2) \to L^2(S^1, \mathbb{R}^2)$*,* $\alpha_-(\mathbf{A}^k) \geqslant k\alpha_-(\mathbf{A})$ *and* $\alpha_+(\mathbf{A}^k) \leqslant k\alpha_+(\mathbf{A})$*.*

PROOF. By definition, $\alpha_-(\mathbf{A})$ is the winding number of some eigenfunction e_λ of $\mathbf{A}$ with eigenvalue $\lambda < 0$. By Proposition 3.7.4, the k-fold cover e_λ^k is likewise an eigenfunction of $\mathbf{A}^k$ with eigenvalue $k\lambda < 0$, so its winding $\operatorname{wind}(e_\lambda^k) = k \operatorname{wind}(e_\lambda) = k\alpha_-(\mathbf{A})$ provides a lower bound for $\alpha_-(\mathbf{A}^k)$. A similar argument shows that $k\alpha_+(\mathbf{A})$ is an upper bound for $\alpha_+(\mathbf{A}^k)$. □

LEMMA 3.10.2. *Suppose* (E, ω, J) *is a Hermitian line bundle over* S^1 *and* $\mathbf{A} \in \mathcal{A}(E)$ *is nondegenerate.*

(1) If $\mathbf{A}$ *is positive hyperbolic, then* $\mathbf{A}^k$ *is also positive hyperbolic for every* $k \in \mathbb{N}$*.*

(2) If $\mathbf{A}$ *is negative hyperbolic, then its covers* $\mathbf{A}^k$ *for* $k \in \mathbb{N}$ *odd are also negative hyperbolic, but its double cover* $\mathbf{A}^2$ *is either positive hyperbolic or degenerate.*

(3) If $\mathbf{A}$ is elliptic, then either $\mathbf{A}^k$ is also elliptic for every $k \in \mathbb{N}$ or it is elliptic for all k outside of a subgroup $m\mathbb{Z} \subset \mathbb{Z}$ for some integer $m \geqslant 2$, and one of the following is true:

(i) m is odd and $\mathbf{A}^{km}$ is degenerate for all $k \in \mathbb{N}$;

(ii) $\mathbf{A}^{km}$ is negative hyperbolic for all k odd and degenerate for all k even.

PROOF. All three statements follow easily from properties of the spectrum of the parallel transport map $\Psi(1) : E_0 \to E_0$ and the fact that $\Psi(k) = \Psi(1)^k$ for every $k \in \mathbb{N}$. In the elliptic case in particular, if $\sigma(\Psi(1)) = \{e^{2\pi i\theta}, e^{-2\pi i\theta}\}$, then $\sigma(\Psi(k)) = \{e^{2\pi ki\theta}, e^{-2\pi ik\theta}\}$ contains no real numbers for any $k \in \mathbb{N}$ if θ is irrational, and otherwise there is degeneracy or negative hyperbolicity only for $k \in m\mathbb{Z}$ where $m \in \mathbb{N}$ is the smallest natural number such that $m\theta \in \frac{1}{2}\mathbb{Z}$. If $m\theta \in \mathbb{Z}$, then m is necessarily odd and we have degeneracy for all $k \in m\mathbb{Z}$. The remaining possibility is that $m\theta$ is a half-integer but not an integer, in which case $\sigma(\Psi(km))$ is $\{-1\}$ for all odd k and $\{1\}$ for all even k. □

THEOREM 3.10.3. *Let E denote the trivial Hermitian line bundle $S^1 \times \mathbb{R}^2 \to S^1$. There exists a unique function*

$$\theta : \mathcal{A}(E) \to \mathbb{R},$$

*called the **monodromy angle**, such that*

$$\alpha_-(\mathbf{A}) \leqslant \theta(\mathbf{A}) \leqslant \alpha_+(\mathbf{A}) \qquad \text{and} \qquad \theta(\mathbf{A}^k) = k\theta(\mathbf{A})$$

for all $\mathbf{A} \in \mathcal{A}(E)$ and $k \in \mathbb{N}$. Moreover, θ has the following properties:

(1) θ is continuous with respect to the L^∞-topology on $\mathcal{A}(E)$ (see Definition 3.4.5);

(2) $\mathbf{A} \in \mathcal{A}(E)$ is elliptic if and only if $\theta(\mathbf{A}) \notin \frac{1}{2}\mathbb{Z}$;

(3) $\mathbf{A} \in \mathcal{A}(E)$ is negative hyperbolic if and only if $\theta(\mathbf{A}) \in \frac{1}{2}\mathbb{Z}\backslash\mathbb{Z}$;

(4) $\mathbf{A} \in \mathcal{A}(E)$ is either degenerate or positive hyperbolic if and only if $\theta(\mathbf{A}) \in \mathbb{Z}$.

PROOF. We proceed in seven steps.

Step 1: Existence and uniqueness.

We claim that for each trivialized asymptotic operator $\mathbf{A}$, there is a unique $\theta \in \mathbb{R}$ such that

$$\alpha_-(\mathbf{A}^k) \leqslant k\theta \leqslant \alpha_+(\mathbf{A}^k)$$

for every $k \in \mathbb{N}$. Indeed, this condition means $\theta \in \bigcap_{k\in\mathbb{N}}[\alpha_-(\mathbf{A}^k)/k, \alpha_+(\mathbf{A}^k)/k]$. Choose any strictly increasing sequence $k_j \in \mathbb{N}$ such that k_{j+1} is divisible by k_j for all j; then writing $k_{j+1}/k_j =: m \in \mathbb{N}$ for a given j, Lemma 3.10.1 implies

(3.14)
$$\frac{\alpha_-(\mathbf{A}^{k_j})}{k_j} = \frac{m\alpha_-(\mathbf{A}^{k_j})}{k_{j+1}} \leqslant \frac{\alpha_-(\mathbf{A}^{k_{j+1}})}{k_{j+1}} \leqslant \frac{\alpha_+(\mathbf{A}^{k_{j+1}})}{k_{j+1}} \leqslant \frac{m\alpha_+(\mathbf{A}^{k_j})}{k_{j+1}} = \frac{\alpha_-(\mathbf{A}^{k_j})}{k_j},$$

so $\left[\frac{\alpha_-(\mathbf{A}^{k_{j+1}})}{k_{j+1}}, \frac{\alpha_+(\mathbf{A}^{k_{j+1}})}{k_{j+1}}\right] \subset \left[\frac{\alpha_-(\mathbf{A}^{k_j})}{k_j}, \frac{\alpha_+(\mathbf{A}^{k_j})}{k_j}\right]$ for all j, meaning that the intervals $\left[\frac{\alpha_-(\mathbf{A}^{k_j})}{k_j}, \frac{\alpha_+(\mathbf{A}^{k_j})}{k_j}\right]$ form a nested sequence. Since every asymptotic operator $\mathbf{A}$ has either trivial kernel or a unique winding number associated to nontrivial eigenfunctions with eigenvalue 0, $\alpha_+(\mathbf{A}) - \alpha_-(\mathbf{A})$ can never be greater than 2, implying that the lengths of the intervals in our nested sequence tend to 0 as $j \to \infty$. It follows that there is a unique real number

$$\theta \in \bigcap_{j=1}^{\infty} \left[\frac{\alpha_-(\mathbf{A}^{k_j})}{k_j}, \frac{\alpha_+(\mathbf{A}^{k_j})}{k_j} \right].$$

Now if there exists a $k \in \mathbb{N}$ such that $\theta \notin [\alpha_-(\mathbf{A}^k)/k, \alpha_+(\mathbf{A}^k)/k]$, then the intervals $[\alpha_-(\mathbf{A}^k)/k, \alpha_+(\mathbf{A}^k)/k]$ and $[\alpha_-(\mathbf{A}^{k_j})/k_j, \alpha_+(\mathbf{A}^{k_j})/k_j]$ must also be disjoint for all j sufficiently large. But the latter is impossible since by (3.14), $\left[\frac{\alpha_-(\mathbf{A}^N)}{N}, \frac{\alpha_+(\mathbf{A}^N)}{N}\right]$ must be contained in both of these intervals whenever $N \in \mathbb{N}$ is divisible by both k and k_j, so this proves the claim. Defining $\theta(\mathbf{A}) := \theta$, the resulting function $\theta : \mathcal{A}(E) \to \mathbb{R}$ now manifestly has both of the properties $\alpha_-(\mathbf{A}) \leqslant \theta \leqslant \alpha_+(\mathbf{A})$ and $\theta(\mathbf{A}^k) = k\theta(\mathbf{A})$, and it is the only function that does so.

Step 2: Continuity.
To see that θ is continuous, fix $\mathbf{A}_0 \in \mathcal{A}(E)$ and, for a given $\epsilon > 0$, choose $k \in \mathbb{N}$ sufficiently large so that $\frac{\alpha_+(\mathbf{A}_0^k) - \alpha_-(\mathbf{A}_0^k)}{k} < \epsilon$. Write $\alpha_\pm(\mathbf{A}_0^k) = \text{wind}(e_{\lambda_0^\pm})$, where $e_{\lambda_0^\pm}$ are specific eigenfunctions of $\mathbf{A}_0^k$ with eigenvalues $\lambda_0^+ > 0$ and $\lambda_0^- < 0$. Then for any $\mathbf{A} \in \mathcal{A}(E)$ sufficiently close to $\mathbf{A}_0$, we can also assume $\mathbf{A}^k$ is close to $\mathbf{A}_0^k$, so Proposition 3.5.12 implies that $\mathbf{A}^k$ also has eigenvalues $\lambda^+ > 0$ and $\lambda^- < 0$ close to λ_0^+ and λ_0^- respectively, whose corresponding eigenfunctions $e_{\lambda^\pm}$ are close to $e_{\lambda_0^\pm}$ in the H^1-topology and therefore also in C^0, implying they have the same winding numbers. This proves

$$\alpha_-(\mathbf{A}_0^k) \leqslant \alpha_-(\mathbf{A}^k) \leqslant \alpha_+(\mathbf{A}^k) \leqslant \alpha_+(\mathbf{A}_0^k),$$

so the condition $\theta(\mathbf{A}) \in [\alpha_-(\mathbf{A}^k)/k, \alpha_+(\mathbf{A}^k)/k]$ implies that $\theta(\mathbf{A})$ and $\theta(\mathbf{A}_0)$ both belong to $[\alpha_-(\mathbf{A}_0^k)/k, \alpha_+(\mathbf{A}_0^k)/k]$, and thus $|\theta(\mathbf{A}) - \theta(\mathbf{A}_0)| < \epsilon$.

Step 3: Positive hyperbolic implies $\theta \in \mathbb{Z}$.
By Theorems 3.7.2 and 3.9.1, $\mathbf{A}$ is positive hyperbolic if and only if it is nondegenerate with $p(\mathbf{A}) := \alpha_+(\mathbf{A}) - \alpha_-(\mathbf{A}) = 0$, so $\theta(\mathbf{A}) \in [\alpha_-(\mathbf{A}), \alpha_+(\mathbf{A})]$ must be an integer.

Step 4: Degenerate implies $\theta \in \mathbb{Z}$.
We claim that if $\mathbf{A} \in \mathcal{A}(E)$ is degenerate, then it lies in the closure of the set of positive hyperbolic operators in $\mathcal{A}(E)$, in which case steps 2 and 3 imply that $\theta(\mathbf{A})$ is an integer. Thinking in terms of parallel transport maps, the claim follows easily from the fact that any 2-by-2 symplectic matrix with spectrum $\{1\}$ is equivalent after a change of basis to one of the form $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ for some $a \in \mathbb{R}$, which can be perturbed within $\text{Sp}(2)$ to $\begin{pmatrix} e^\epsilon & a \\ 0 & e^{-\epsilon} \end{pmatrix}$ for $\epsilon > 0$ small, and the latter can then be realized using Proposition 3.4.11 as the end point of the parallel transport of a nearby positive hyperbolic asymptotic operator.

Step 5: Odd index implies $\theta \notin \mathbb{Z}$.
Suppose $\mathbf{A} \in \mathcal{A}(E)$ is nondegenerate with $\mu_{\text{CZ}}(\mathbf{A})$ odd, so Theorem 3.7.2 implies that $[\alpha_-(\mathbf{A}), \alpha_+(\mathbf{A})]$ is a unit interval, and we claim that $\theta(\mathbf{A})$ lies in the *interior* of this interval. Suppose to the contrary that $\theta(\mathbf{A}) = \alpha_-(\mathbf{A})$. One can use a change

of trivialization to shift the winding numbers of $\alpha_\pm(\mathbf{A})$ by any desired integer, in which case $\theta(\mathbf{A})$ gets adjusted by the same shift, and $\mu_{\mathrm{CZ}}(\mathbf{A})$ is shifted by twice the same integer (cf. Exercise 3.7.3), so we can use this trick to assume without loss of generality that $[\alpha_-(\mathbf{A}), \alpha_+(\mathbf{A})] = [0,1]$, $\mu_{\mathrm{CZ}}(\mathbf{A}) = 1$ and $\theta(\mathbf{A}) = 0$. Then for every $k \in \mathbb{N}$, Lemma 3.10.1 implies

$$0 = k\theta(\mathbf{A}) = k\alpha_-(\mathbf{A}) \leqslant \alpha_-(\mathbf{A}^k) \leqslant \theta(\mathbf{A}^k) = k\theta(\mathbf{A}) = 0,$$

thus $\alpha_-(\mathbf{A}^k) = 0$ as well. Now let $\Psi : \mathbb{R} \to \mathrm{Sp}(2)$ denote the parallel transport map of $\mathbf{A}$, which by Theorem 3.9.1 satisfies either $\sigma(\Psi(1)) \subset (-\infty, 0)$ or $\sigma(\Psi(1)) \subset \mathrm{U}(1)\backslash\{1,-1\}$. Since $\Psi(k) = \Psi(1)^k$ for every $k \in \mathbb{N}$, in either case there exist arbitrarily large values of k for which $\sigma(\Psi(k))$ is also contained in either $(-\infty, 0)$ or $\mathrm{U}(1)\backslash\{1,-1\}$, which means there are arbitrarily large nondegenerate covers $\mathbf{A}^k$ for which $\mu_{\mathrm{CZ}}(\mathbf{A}^k)$ is also odd, implying in this situation that $\mu_{\mathrm{CZ}}(\mathbf{A}^k) = 2\alpha_-(\mathbf{A}^k) + 1 = 1$. But if Ψ_k denotes the parallel transport of $\mathbf{A}^k$, Theorem 3.9.1 then implies that the winding interval $\Delta(\Psi_k)$ is a compact subinterval of $(0,1)$ for arbitrarily large values of $k \in \mathbb{N}$, which is impossible since $\Delta(\Psi) = [a,b]$ for $0 < a \leqslant b < 1$ implies $\Delta(\Psi_k) \subset [ka, kb]$ for all k, and the latter can no longer be contained in $(0,1)$ when $k > 1/a$.

If we instead assume $\theta(\mathbf{A}) = \alpha_+(\mathbf{A})$, then after a different change of trivialization we can assume without loss of generality that $[\alpha_-(\mathbf{A}), \alpha_+(\mathbf{A})] = [-1, 0]$ and $\theta(\mathbf{A}) = 0$, so in this case $\mu_{\mathrm{CZ}}(\mathbf{A}^k) = -1$ for arbitrarily large values of k, and one obtains a similar contradiction by looking at the winding intervals $\Delta(\Psi_k) \subset (-1, 0)$.

Step 6: Negative hyperbolic is equivalent to $\theta \in \frac{1}{2}\mathbb{Z}\backslash\mathbb{Z}$.
If $\mathbf{A} \in \mathcal{A}(E)$ is negative hyperbolic, then $\mathbf{A}^2$ is either degenerate or positive hyperbolic by Lemma 3.10.2. By the results of steps 3 and 4, it follows that $\theta(\mathbf{A}) \in \frac{1}{2}\mathbb{Z}$. But since $\mu_{\mathrm{CZ}}(\mathbf{A})$ is odd by Theorem 3.9.1, step 5 implies $\theta(\mathbf{A}) \notin \mathbb{Z}$.

Step 7: Elliptic implies $\theta \notin \frac{1}{2}\mathbb{Z}$.
If $\mathbf{A} \in \mathcal{A}(E)$ is elliptic, then Lemma 3.10.2 implies that $\mathbf{A}^2$ is either elliptic or negative hyperbolic, so step 5 and Theorem 3.9.1 imply $\theta(\mathbf{A}^2) = 2\theta(\mathbf{A}) \notin \mathbb{Z}$ and thus $\theta(\mathbf{A}) \notin \frac{1}{2}\mathbb{Z}$. ☐

Since $[\alpha_-(\mathbf{A}), \alpha_+(\mathbf{A})]$ is always either a single point or a unit interval when $\mathbf{A}$ is nondegenerate, Theorem 3.10.3 gives rise to the formulas

$$\alpha_-(\mathbf{A}) = \lfloor\theta(\mathbf{A})\rfloor, \qquad \alpha_+(\mathbf{A}) = \lceil\theta(\mathbf{A})\rceil, \qquad \text{if} \quad \ker \mathbf{A} = \{0\}. \tag{3.15}$$

Recall that a contact form α is called nondegenerate if all of its closed Reeb orbits are nondegenerate, and this condition holds for generic contact forms (see Remark 1.3.13). In this situation, Lemma 3.10.2 implies that all covers of an elliptic orbit are also elliptic, so one deduces from Theorem 3.10.3 that the corresponding monodromy angle must be irrational. Combining these observations with the relation between $\mu_{\mathrm{CZ}}(\mathbf{A})$ and $\alpha_\pm(\mathbf{A})$ in Theorem 3.7.1, one now obtains the following result for multiply covered Reeb orbits:

COROLLARY 3.10.4. *Suppose γ is a nondegenerate Reeb orbit in a 3-manifold M with stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ such that the multiple covers γ^k are also nondegenerate for every $k \in \mathbb{N}$. Choose a symplectic trivialization τ of $\xi = \ker\lambda$*

along γ, and use the same notation to denote the trivializations along γ^k defined by pulling back τ along the covering map $S^1 \to S^1 : t \mapsto kt$.

- *If γ is (positive or negative) hyperbolic, then*
$$\mu^\tau_{CZ}(\gamma^k) = k\mu^\tau_{CZ}(\gamma)$$
for every $k \in \mathbb{N}$.
- *If γ is elliptic, then there exists an irrational number $\theta \in \mathbb{R}\backslash\mathbb{Q}$ such that*
$$\mu^\tau_{CZ}(\gamma^k) = 2\lfloor k\theta\rfloor + 1 = 2\lceil k\theta\rceil - 1$$
for every $k \in \mathbb{N}$.

□

REMARK 3.10.5 (sign conventions). Our definition of the Conley-Zehnder index for nondegenerate symplectic arcs agrees with definitions given in most other sources (such as [FH93,Sch95,Sal99]), but one should be aware of occasional discrepancies. The index μ_τ in [SZ92] differs from our μ_{CZ} by a sign: the reason (as helpfully pointed out by [Sch95, p. 84]) is that Salamon and Zehnder define the standard complex structure on $\mathbb{R}^{2n}$ as $\begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}$ instead of $\begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$, thus reversing its symplectic structure and, in particular, changing the orientation of $\mathbb{R}^2$, so that all winding numbers reverse sign. From the perspective of Floer homology, for which μ_τ was developed, the result is sensible: as mentioned in Remark 3.3.6, the asymptotic operator in Floer homology has a different sign than in SFT, so reversing the sign of the Conley-Zehnder index is the right thing to do if you want to regard it as a relative Morse index for the action functional. It is inconvenient however in other respects, e.g. when trying to compute μ_{CZ} in terms of winding numbers, thus later papers on Floer homology have often used definitions of $\mu_{\text{CZ}}(\Psi)$ that are equivalent to ours, but introduced modified indices for orbits in order to absorb the sign difference, e.g. [Sal99] defines $\mu_H(\gamma) := n - \mu_{\text{CZ}}(\Psi)$ for the linearized flow Ψ along an orbit γ. For the reasons why the latter is a natural convention in that context, see Theorem 11.3.1 and Remark 11.3.2.

3.11. CZ = Morse for geodesics

For any Riemannian manifold M, the unit cotangent bundle ST^*M carries a canonical contact form whose Reeb orbits are lifts of geodesics on M. A basic result used often in applications is that for a natural choice of trivialization along the lift of any closed co-orientable geodesic on M, the Conley-Zehnder index of the lift equals the *Morse index* of the geodesic. In this section, we will prove a slight generalization of this statement that is valid for arbitrary geodesics without any orientability hypothesis. The proof is based roughly on [Web02], which proves a similar result in the context of perturbed Hamiltonian systems.

3.11.1. The unit cotangent bundle. In the following, we assume $(M, \langle\ ,\ \rangle)$ is a Riemannian manifold with
$$\dim M = n \in \mathbb{N}.$$

For each $q \in M$, the Riemannian metric induces musical isomorphisms

$$\flat : T_q M \to T_q^* M : X \mapsto X_\flat := \langle X, \cdot \rangle, \qquad \sharp := \flat^{-1} : T_q^* M \to T_q M : \eta \mapsto \eta^\sharp,$$

and we will also use $\langle \ , \ \rangle$ to denote the induced bundle metric on T^*M, for which $\flat$ and $\sharp$ are isometries. When viewing the cotangent bundle T^*M as a manifold, we shall denote its elements as pairs

$$(q, p) \in T^*M,$$

where $q \in M$ and $p \in T_q^*M$. The **unit cotangent bundle** is then defined by

$$ST^*M := \left\{ (q,p) \in T^*M \;\middle|\; |p| = 1 \right\},$$

where $|p| := \sqrt{\langle p, p \rangle}$. In order to describe the tangent space to T^*M at a point (q,p), we use the Levi-Cività connection ∇ on M to split $T_{(q,p)}(T^*M)$ into horizontal and vertical subspaces, which are naturally identified with T_qM and T_q^*M respectively, producing a natural isomorphism

$$T_{(q,p)}(T^*M) = T_q M \oplus T_q^* M.$$

With this understood, we will always write elements of $T_{(q,p)}(T^*M)$ as pairs (X, η) where $X \in T_qM$ is the horizontal part and $\eta \in T_q^*M$ the vertical part, so e.g. the derivative of a smooth path $\gamma(t) = (q(t), p(t)) \in T^*M$ is now expressed in terms of the covariant derivative induced on T^*M by the Levi-Cività connection, namely as

$$\dot{\gamma}(t) = (\dot{q}(t), \nabla_t p(t)) \in T_{q(t)}M \oplus T_{q(t)}^*M = T_{(q(t),p(t))}(T^*M).$$

For $(q,p) \in ST^*M$, we have

$$T_{(q,p)}(ST^*M) = \left\{ (X, \eta) \in T_qM \oplus T_q^*M \;\middle|\; \langle p, \eta \rangle = 0 \right\} \subset T_{(q,p)}(T^*M).$$

The tautological Liouville form λ_{std} on T^*M is given by

$$(\lambda_{\mathrm{std}})_{(q,p)}(X, \eta) := p(X),$$

and is thus often abbreviated in the literature as "$p\,dq$"; indeed, if one makes any choice of local coordinates $q^1, \dots, q^n$ on a region $\mathcal{U} \subset M$ and denotes by $(q^1, \dots, q^n, p_1, \dots, p_n)$ the induced chart on $T^*\mathcal{U} \subset T^*M$ for which a point $q \in \mathcal{U}$ with coordinates $(q^1, \dots, q^n)$ gives rise to elements $(q, \sum_i p_i\, dq^i) \in T^*\mathcal{U}$ with coordinates $(q^1, \dots, q^n, p_1, \dots, p_n)$, then λ_{std} in these coordinates takes the form $\sum_i p_i\, dq^i$. Note that while we have made use of the splitting $T_{(q,p)}(T^*M) = T_qM \oplus T_q^*M$ determined by the Levi-Cività connection in order to write down the formula for λ_{std} above, its definition does not actually depend on the connection or the metric. It will be convenient however to make further use of this splitting in order to write down the symplectic form $d\lambda_{\mathrm{std}}$. We start by defining a Riemannian metric $\langle \ , \ \rangle$ on T^*M for which the horizontal and vertical parts of $T_{(q,p)}(T^*M)$ are orthogonal to each other and each carry the inner products that we previously chose on T_qM and T_q^*M respectively. We also use the splitting to define an almost complex structure J_{std} on T^*M in block matrix form by

$$J_{\mathrm{std}} := \begin{pmatrix} 0 & \sharp \\ -\flat & 0 \end{pmatrix} : T_{(q,p)}(T^*M) \to T_{(q,p)}(T^*M),$$

and observe that J_{std} preserves the metric defined above, and maps the horizontal and vertical subbundles of $T(T^*M)$ isometrically to each other. Both of these definitions depend on the Riemannian metric of M, but the pairing $\langle J_{\mathrm{std}}\cdot, \cdot\rangle$ turns out to be independent of this data:

EXERCISE 3.11.1. Prove

$$d\lambda_{\mathrm{std}} = \langle J_{\mathrm{std}}\cdot, \cdot\rangle \qquad \text{on } T^*M.$$

In particular, $d\lambda_{\mathrm{std}}$ is a symplectic form, J_{std} is a $d\lambda_{\mathrm{std}}$-compatible almost complex structure, and for each $(q,p) \in T^*M$, the horizontal and vertical subspaces of $T_{(q,p)}(T^*M)$ are Lagrangian. *Hint: Given $\xi, \eta \in T_{(q,p)}(T^*M)$, make an intelligent choice of smooth map $\gamma(s,t) = (q(s,t), p(s,t)) \in T^*M$ for $(s,t) \in (-\epsilon,\epsilon)^2$ such that $\gamma(0,0) = (q,p)$, $\partial_s\gamma(0,0) = \xi$ and $\partial_t\gamma(0,0) = \eta$, and use the formula $d\lambda_{\mathrm{std}}(\xi,\eta) = \partial_s\left[\lambda_{\mathrm{std}}(\partial_t\gamma)\right] - \partial_t\left[\lambda_{\mathrm{std}}(\partial_s\gamma)\right]\big|_{s=t=0}$. It suffices to consider separately the cases where each of ξ,η is horizontal or vertical.*

EXERCISE 3.11.2. Show that the Liouville vector field V on T^*M charaterized by the relation $d\lambda_{\mathrm{std}}(V,\cdot) = \lambda_{\mathrm{std}}$ is given by $V(q,p) = (0,p)$.

Exercise 3.11.2 shows that the unit cotangent bundle ST^*M is a contact-type hypersurface in $(T^*M, d\lambda_{\mathrm{std}})$, thus the restriction

$$\alpha_{\mathrm{std}} := \lambda_{\mathrm{std}}|_{T(ST^*M)} \in \Omega^1(ST^*M)$$

is a contact form. The induced contact structure $\xi_{\mathrm{std}} = \ker\alpha_{\mathrm{std}} \subset T(ST^*M)$ is given by

$$(\xi_{\mathrm{std}})_{(q,p)} = \left\{(X,\eta) \in T_qM \oplus T_q^*M \;\middle|\; p(X) = \langle p,\eta\rangle = 0\right\} \subset T_{(q,p)}(ST^*M). \tag{3.16}$$

It is easy to check that this subbundle is preserved by J_{std}, hence J_{std} also defines a compatible complex structure on the symplectic vector bundle $(\xi_{\mathrm{std}}, d\alpha_{\mathrm{std}})$. Moreover, the intersections of ξ_{std} with the horizontal and vertical subbundles of $T(T^*M)$ define a natural splitting

$$\xi_{\mathrm{std}} = \xi^h_{\mathrm{std}} \oplus \xi^v_{\mathrm{std}}, \tag{3.17}$$

where $(\xi^h_{\mathrm{std}})_{(q,p)}$ is the horizontal lift of $\ker p \subset T_qM$ and $(\xi^v_{\mathrm{std}})_{(q,p)}$ is the orthogonal complement of p in the vertical subspace of $T_{(q,p)}(T^*M)$ under its canonical identification with T_q^*M. Both summands in this splitting are Lagrangian subbundles of $(\xi_{\mathrm{std}}, d\alpha_{\mathrm{std}})$, and we have

$$J_{\mathrm{std}}(\xi^h_{\mathrm{std}}) = \xi^v_{\mathrm{std}}.$$

EXERCISE 3.11.3. Show that as complex vector bundles with complex structure J_{std}, both $T(T^*M) \to T^*M$ and $\xi_{\mathrm{std}} \to ST^*M$ are complex-isomorphic to their respective dual bundles. This implies in particular that both have vanishing first Chern class. *Hint: Use the horizontal/vertical splitting.*

Using the formula for $d\lambda_{\mathrm{std}}$ in Exercise 3.11.1, one computes that the Reeb vector field $R_{\mathrm{std}} := R_{\alpha_{\mathrm{std}}}$ of α_{std} is

$$R_{\mathrm{std}}(q,p) = (p^\sharp, 0).$$

A smooth path $\gamma(t) = (q(t), p(t)) \in ST^*M$ therefore satisfies $\dot{\gamma} = R_{\mathrm{std}}(\gamma)$ if and only if $\dot{q}(t) = p(t)^\sharp$ and $\nabla_t p(t) = 0$, which implies $\nabla_t \dot{q}(t) = 0$, hence the path $t \mapsto q(t)$ is a geodesic in $(M, \langle\ ,\ \rangle)$. We've proved:

PROPOSITION 3.11.4. *Every geodesic* $q : (a, b) \to M$ *on* $(M, \langle\ ,\ \rangle)$ *with unit speed* $|\dot{q}| \equiv 1$ *has a natural lift* $\gamma : (a, b) \to ST^*M$ *to a Reeb orbit in* $(ST^*M, \alpha_{\mathrm{std}})$*, given by* $\gamma(t) := (q(t), \dot{q}(t)_\flat)$*. Conversely, every Reeb orbit in* (ST^*M, α) *is in this sense the lift of a unique geodesic on* $(M, \langle\ ,\ \rangle)$ *with unit speed.* □

Since we are interested mainly in *closed* Reeb orbits, we now consider closed geodesics on $(M, \langle\ ,\ \rangle)$. The first thing that must be pointed out is that since geodesics $q : \mathbb{R} \to M$ satisfy a second-order equation, the condition $q(T) = q(0)$ for some $T > 0$ does not suffice to ensure that q is T-periodic, though enhancing it with $\dot{q}(T) = \dot{q}(0)$ would suffice. The more convenient approach for our purposes is to parametrize geodesics on $S^1 = \mathbb{R}/\mathbb{Z}$, though this will necessitate dropping the preference for unit speed, as a closed geodesic may in principle have any length. This corresponds to the fact that in order to consider Reeb orbits of all possible periods, we must allow parametrizations $\gamma : S^1 \to M$ in which $\dot{\gamma}$ is proportional but not necessarily equal to $R_{\mathrm{std}}(\gamma)$ (cf. Definition 3.1.2). We thus associate to any nonconstant periodic solution of the geodesic equation

$$q : S^1 \to M, \qquad \nabla_t \dot{q} \equiv 0$$

with constant speed $T := |\dot{q}(t)| > 0$ the closed Reeb orbit with parametrization

$$\gamma : S^1 \to ST^*M, \qquad \gamma(t) := (q(t), \dot{q}(t)_\flat / T),$$

which satisfies $\dot{\gamma} = T \cdot R_{\mathrm{std}}(\gamma)$. Conversely, projecting any closed Reeb orbit $\gamma : S^1 \to ST^*M$ to M gives rise to a periodic geodesic $q : S^1 \to M$ whose constant speed is the period of γ.

REMARK 3.11.5. Like closed Reeb orbits, nonconstant periodic geodesics always come in S^1-families of parametrizations related to each other by constant shifts. But for geodesics, an additional reparametrization is possible that makes no sense for Reeb orbits: one can replace $q : S^1 \to M$ with the geodesic $q^- : S^1 \to M$ defined by $q^-(t) := q(-t)$. Obviously q and q^- have the same image in M, but their lifts as Reeb orbits in ST^*M are typically *disjoint* from each other, e.g. even at time $t = 0$ where $q(0) = q^-(0)$, the vertical components of their lifts occupy antipodal points in the fiber of ST^*M over $q(0)$.

3.11.2. The energy functional and its Hessian. One of the great historical successes of Morse theory (see [Mil63]) was to prove existence results for geodesics by exploiting the fact that geodesics, unlike Reeb orbits (cf. Prop. 3.5.22), can be regarded as critical points of a suitable functional with *finite* Morse index. The functional in question is called **energy**, and in the setting of periodic orbits, one defines it by

$$\mathcal{E} : C^\infty(S^1, M) \to \mathbb{R} : q \mapsto \frac{1}{2} \int_{S^1} |\dot{q}(t)|^2 \, dt.$$

Its first variation at $q \in C^\infty(S^1, M)$ is a linear operator $d\mathcal{E}(q) : \Gamma(q^*TM) \to \mathbb{R}$ taking the form

$$d\mathcal{E}(q)X = -\int_{S^1} \langle \nabla_t \dot{q}(t), X(t) \rangle \, dt =: \langle \nabla\mathcal{E}(q), X \rangle_{L^2},$$

where we define the L^2-gradient of $\mathcal{E}$ at q by

$$\nabla\mathcal{E}(q) := -\nabla_t \dot{q} \in \Gamma(q^*TM)$$

and observe that this gradient vanishes if and only if q is a periodic geodesic. To write down the Hessian of $\mathcal{E}$ at a critical point q, we consider a smooth family of loops $\{q_\rho : S^1 \to M\}_{\rho \in (-\epsilon,\epsilon)}$ with $q_0 = q$ and $X := \partial_\rho q_\rho|_{\rho=0} \in \Gamma(q^*TM)$, and compute

$$\nabla_\rho \nabla\mathcal{E}(q_\rho)\big|_{\rho=0} = -\nabla_\rho \nabla_t \dot{q}\big|_{\rho=0} = -\nabla_t \nabla_\rho \dot{q}\big|_{\rho=0} - R(X, \dot{q})\dot{q} = -\nabla_t^2 X - R(X, \dot{q})\dot{q},$$

where R denotes the Riemann curvature tensor of $(M, \langle\ ,\ \rangle)$, and we've used the symmetry of the connection to replace $\nabla_\rho \dot{q}_\rho|_{\rho=0}$ with $\nabla_t \partial_\rho q_\rho|_{\rho=0} = \nabla_t X$. This determines the second-order differential operator

$$\nabla^2\mathcal{E}(q) : \Gamma(q^*TM) \to \Gamma(q^*TM) : X \mapsto -\nabla_t^2 X - R(X, \dot{q})\dot{q},$$

which we will call the **Hessian** of $\mathcal{E}$ at the critical point $q \in C^\infty(S^1, M)$.

EXERCISE 3.11.6. Check that $\nabla^2\mathcal{E}(q)$ is symmetric with respect to the L^2-pairing $\langle X, Y \rangle_{L^2} := \int_{S^1} \langle X(t), Y(t) \rangle \, dt$ on $\Gamma(q^*TM)$.

We are not interested in constant geodesics for this discussion, so let us restrict attention to critical points q of $\mathcal{E}$ such that $\dot{q}$ is nowhere zero. In this case, $\nabla^2\mathcal{E}(q)$ always has the nontrivial section $\dot{q} \in \Gamma(q^*TM)$ in its kernel; this results from the fact that the energy functional is invariant under the S^1-action defined on $C^\infty(S^1, M)$ by shifting parametrizations. As we will see below, however, it can and frequently does happen that this one section spans the entire kernel. Let us henceforward denote by

$$N_q \subset q^*TM$$

the normal bundle of any nonconstant periodic geodesic $q : S^1 \to M$, defined literally as the orthogonal complement of the line bundle $\mathbb{R}\langle \dot{q} \rangle \subset q^*TM$ spanned by $\dot{q}$.

LEMMA 3.11.7. *Given a nonconstant periodic geodesic $q : S^1 \to M$, the Hessian $\nabla^2\mathcal{E}(q)$ respects the splitting $q^*TM = \mathbb{R}\langle \dot{q} \rangle \oplus N_q$, and its kernel on the first factor is the 1-dimensional space spanned by $\dot{q} \in \Gamma(q^*TM)$.*

PROOF. An arbitrary section of the 1-dimensional subbundle $\mathbb{R}\langle \dot{q} \rangle \subset q^*TM$ can be written as $f\dot{q}$ for some smooth function $f : S^1 \to \mathbb{R}$, and since $\nabla_t \dot{q} \equiv 0$ and the Riemann tensor is antisymmetric, we then have

$$\nabla^2\mathcal{E}(q) f\dot{q} = -\nabla_t \nabla_t (f\dot{q}) - R(f\dot{q}, \dot{q})\dot{q} = -\nabla_t(\dot{f}\dot{q}) - fR(\dot{q}, \dot{q})\dot{q} = -\ddot{f}\dot{q} \in \mathbb{R}\langle \dot{q} \rangle.$$

This proves that $\nabla^2\mathcal{E}(q)$ preserves the subspace $\Gamma(\mathbb{R}\langle \dot{q} \rangle) \subset \Gamma(q^*TM)$, and moreover, $f\dot{q}$ belongs to $\ker \nabla^2\mathcal{E}(q)$ if and only if $f : S^1 \to \mathbb{R}$ has vanishing second derivative, which is ruled out by periodicity unless f is constant. Now, observe that a section $X \in \Gamma(q^*TM)$ takes values in the normal subbundle N_q if and only if it is L^2-orthogonal to $f\dot{q}$ for every smooth function $f : S^1 \to \mathbb{R}$. The fact that $\nabla^2\mathcal{E}(q)$ is L^2-symmetric then implies that it also preserves $\Gamma(N_q) \subset \Gamma(q^*TM)$. □

The lemma implies that the interesting information in $\nabla^2\mathcal{E}(q)$ is carried by its well-defined restriction

$$\nabla^2_N\mathcal{E}(q) : \Gamma(N_q) \to \Gamma(N_q),$$

another symmetric operator which we will call the **normal Hessian** of q.

EXERCISE 3.11.8. Show that the extension of $\nabla^2_N\mathcal{E}(q)$ to a bounded linear operator $H^2(N_q) \to L^2(N_q)$ is Fredholm with index 0. *Hint: If N_q is orientable, then it is trivial and one can thus reduce the problem as in Lemma 3.4.7 to calculating the kernel and cokernel of $\partial_t^2 : H^2(S^1, \mathbb{R}^{n-1}) \to L^2(S^1, \mathbb{R}^{n-1})$, which is doable via either integration or Fourier series. If N_q is not orientable, then it suffices to supplement this result with the following claim: the operator*

$$\partial_t^2 : \left\{ f \in H^2(S^1, \mathbb{R}) \;\middle|\; f(t+1/2) = -f(t) \right\} \to \left\{ f \in L^2(S^1, \mathbb{R}) \;\middle|\; f(t+1/2) = -f(t) \right\}$$

is invertible.

The results of §3.5.3 now imply that as an unbounded operator on $L^2(N_q)$ with dense and compactly embedded domain $H^2(N_q) \subset L^2(N_q)$, the spectrum of $\nabla^2_N\mathcal{E}(q)$ consists of isolated real eigenvalues, each with finite multiplicity. These conclusions could also alternatively be deduced from the observation that $\nabla^2_N\mathcal{E}(q)$ with domain $H^2(N_q) \subset L^2(N_q)$ is a self-adjoint operator on $L^2(N_q)$ with compact resolvent. But there is now a crucial difference in comparison with asymptotic operators:

LEMMA 3.11.9. *The spectrum $\sigma(\nabla^2_N\mathcal{E}(q)) \subset \mathbb{R}$ is bounded from below, so in particular, $\nabla^2_N\mathcal{E}(q)$ has at most finitely many negative eigenvalues, counting multiplicitiy.*

PROOF. It suffices to show that $\nabla^2_N\mathcal{E}(q) + \lambda$ is a positive operator for every $\lambda > 0$ sufficiently large. Indeed, for $X \in \Gamma(N_q)$, integration by parts gives

$$\begin{aligned}
\left\langle X, (\nabla^2_N\mathcal{E}(q) + \lambda)X \right\rangle_{L^2} &= -\langle X, \nabla_t^2 X\rangle_{L^2} - \langle X, R(X, \dot{q})\dot{q}\rangle_{L^2} + \lambda\|X\|^2_{L^2} \\
&= \|\nabla_t X\|^2_{L^2} + \lambda\|X\|^2_{L^2} - \langle X, R(X, \dot{q})\dot{q}\rangle_{L^2} \\
&\geqslant \|\nabla_t X\|^2_{L^2} + (\lambda - c)\|X\|^2_{L^2}
\end{aligned}$$

for some constant $c > 0$, since $X \mapsto R(X, \dot{q})\dot{q}$ is an L^2-bounded operator. □

DEFINITION 3.11.10. The **Morse index**

$$\mathrm{Morse}(q) \geqslant 0$$

of a nonconstant periodic geodesic $q : S^1 \to M$ on $(M, \langle\ ,\ \rangle)$ is defined as the number of negative eigenvalues (counting multiplicity) of the normal Hessian $\nabla^2_N\mathcal{E}(q)$. We call q a **nondegenerate** geodesic if $\nabla^2_N\mathcal{E}(q)$ does not have 0 as an eigenvalue.

EXAMPLE 3.11.11. If $(M, \langle\ ,\ \rangle)$ has negative sectional curvature everywhere, then all of its nonconstant periodic geodesics are nondegenerate and have Morse index 0. Indeed, for any geodesic $q : S^1 \to M$ and $X \in \Gamma(N_q)$, the same calculation as above via integration by parts gives

$$\langle X, \nabla^2_N\mathcal{E}(q)X\rangle_{L^2} = \|\nabla_t X\|^2_{L^2} - \langle X, R(X, \dot{q})\dot{q}\rangle_{L^2}.$$

Since X and $\dot{q}$ are everywhere orthogonal, the second term is nonnegative due to the sectional curvature assumption, and vanishes if and only if $X \equiv 0$. It follows that $\nabla_N^2 \mathcal{E}(q)$ has no nonpositive eigenvalues.

3.11.3. Conley-Zehnder indices of lifted geodesics. We are now ready to state the main result of this section. Given a nonconstant periodic geodesic $q : S^1 \to M$, denote its lift to the unit cotangent bundle by $\gamma = (q, p) : S^1 \to ST^*M$, where we recall that $p(t) \in T^*_{q(t)}M$ is a unit vector obtained by acting on $\dot{q}(t)$ with the musical isomorphism $\flat : T_{q(t)}M \to T^*_{q(t)}M$ and then normalizing it. By (3.17), the Hermitian bundle $(\gamma^*\xi_{\text{std}}, J_{\text{std}}, d\alpha_{\text{std}})$ splits into a direct sum of two Lagrangian subbundles $\gamma^*\xi^h_{\text{std}}$ and $\gamma^*\xi^v_{\text{std}} = J_{\text{std}}(\gamma^*\xi^h_{\text{std}})$, the first of which is simply the horizontal lift of N_q. Given a symplectic trivialization τ of $(\gamma^*\xi_{\text{std}}, d\alpha_{\text{std}})$, we denote by

$$\mu^\tau(\gamma^*\xi_{\text{std}}, \gamma^*\xi^h_{\text{std}})$$

the **Maslov index** of the resulting loop of Lagrangian subspaces in $\mathbb{R}^{2n-2}$ (see e.g. [MS17, §2.3]).

THEOREM 3.11.12. *Assume $(M, \langle\ ,\ \rangle)$ is a Riemannian manifold, and $q : S^1 \to M$ is a nonconstant periodic geodesic with lift $\gamma = (q, p) : S^1 \to ST^*M$. Then q is a nondegenerate geodesic if and only if γ is a nondegenerate Reeb orbit, and in this case,*

$$\mu^\tau_{CZ}(\gamma) = \text{Morse}(q) + \mu^\tau(\gamma^*\xi_{\text{std}}, \gamma^*\xi^h_{\text{std}})$$

for any choice of symplectic trivialization τ of $\gamma^*\xi_{\text{std}}$.

EXERCISE 3.11.13. Convince yourself that both sides of the formula in Theorem 3.11.12 transform the same way if the choice of symplectic trivialization τ is changed.

The most interesting special case of Theorem 3.11.12 is when the geodesic $q : S^1 \to M$ is **co-orientable**, meaning that its normal bundle $N_q \to S^1$ is orientable. This is guaranteed to hold if M itself is orientable, and it also holds for the double cover of any (not necessarily co-orientable) periodic geodesic. In this case N_q is a trivial bundle, so there is a preferred class of unitary trivializations of $\gamma^*\xi_{\text{std}}$ defined by horizontally lifting any global real frame of N_q to a frame on $\gamma^*\xi^h_{\text{std}}$ and then regarding it as a complex frame on $\gamma^*\xi_{\text{std}}$. For any trivialization τ of this form, the loop of Lagrangian subspaces defining $\mu^\tau(\gamma^*\xi_{\text{std}}, \gamma^*\xi^h_{\text{std}})$ is constant, so the Maslov index in the above formula vanishes, giving:

COROLLARY 3.11.14. *In the setting of Theorem 3.11.12, assume additionally that the geodesic $q : S^1 \to M$ is co-orientable and τ is obtained in the natural way from a trivialization of the normal bundle $N_q \to S^1$. Then*

$$\mu^\tau_{CZ}(\gamma) = \text{Morse}(q).$$

□

The proof of Theorem 3.11.12 requires deriving a new formula for the asymptotic operator of γ so that an explicit relationship can be seen between its kernel and that of $\nabla_N^2 \mathcal{E}(q)$. The formula in Definition 3.3.2 is non-ideal for this purpose, as it requires a choice of symmetric connection on ST^*M that is not canonical in the present setting, and it also ignores certain useful information that is specific to the unit cotangent bundle. In our present context, the contact action functional $\mathcal{A}_{\mathrm{std}} := \mathcal{A}_{\alpha_{\mathrm{std}}} : C^\infty(S^1, ST^*M) \to \mathbb{R}$ at an arbitrary loop $\gamma = (q,p) : S^1 \to ST^*M$ takes the form

$$\mathcal{A}_{\mathrm{std}}(\gamma) = \int_{S^1} \gamma^*\alpha_{\mathrm{std}} = \int_{S^1} \alpha_{\mathrm{std}}(\dot\gamma(t))\,dt = \int_{S^1} p(t)\big(\dot q(t)\big)\,dt,$$

and its gradient $\nabla\mathcal{A}_{\mathrm{std}}(\gamma) = -J_{\mathrm{std}}\pi_{\xi_{\mathrm{std}}}\dot\gamma \in \Gamma(\gamma^*\xi_{\mathrm{std}})$ can be written as

$$\begin{aligned}\nabla\mathcal{A}_{\mathrm{std}}(\gamma) &= -J_{\mathrm{std}}\left[\dot\gamma - \alpha_{\mathrm{std}}(\dot\gamma)R_{\mathrm{std}}(\gamma)\right] \\ &= \begin{pmatrix} 0 & -\sharp \\ \flat & 0\end{pmatrix}\left[\begin{pmatrix}\dot q \\ \nabla_t p\end{pmatrix} - p(\dot q)\begin{pmatrix} p^\sharp \\ 0\end{pmatrix}\right] = \begin{pmatrix} 0 & -\sharp \\ \flat & 0\end{pmatrix}\begin{pmatrix}\dot q - p(\dot q)p^\sharp \\ \nabla_t p\end{pmatrix} \\ &= \begin{pmatrix} -\nabla_t p^\sharp \\ \dot q_\flat - p(\dot q)p\end{pmatrix}.\end{aligned}$$

Next, assume γ is a critical point of $\mathcal{A}_{\mathrm{std}}$ parametrized so that $T := p(\dot q) > 0$ is constant, so $q : S^1 \to M$ is a geodesic of constant speed T and $p = \frac{1}{T}\dot q_\flat$. To compute the Hessian $\mathbf{A}_\gamma := \nabla^2\mathcal{A}_{\mathrm{std}}(\gamma) : \Gamma(\gamma^*\xi_{\mathrm{std}}) \to \Gamma(\gamma^*\xi_{\mathrm{std}})$, we consider a smooth 1-parameter family of loops $\gamma_\rho = (q_\rho, p_\rho) : S^1 \to ST^*M$ with $\gamma_0 = \gamma$ and $\partial_\rho\gamma_\rho|_{\rho=0} = \big(\,\partial_\rho q_\rho|_{\rho=0}\,,\,\nabla_\rho p_\rho|_{\rho=0}\big) =: (X,\eta) \in \Gamma(\gamma^*\xi_{\mathrm{std}})$, and compute

$$\mathbf{A}_\gamma\begin{pmatrix} X \\ \eta\end{pmatrix} = \nabla_\rho\begin{pmatrix} -\nabla_t p_\rho^\sharp \\ (\dot q_\rho)_\flat - p_\rho(\dot q_\rho)p_\rho\end{pmatrix}\Bigg|_{\rho=0} = \begin{pmatrix} -\nabla_\rho\nabla_t p_\rho^\sharp \\ (\nabla_\rho\dot q_\rho)_\flat - \nabla_\rho\left[p_\rho(\dot q_\rho)p_\rho\right]\end{pmatrix}\Bigg|_{\rho=0}.$$

The top entry can be rewritten using the Riemann tensor of $(M, \langle\ ,\ \rangle)$ as

$$-\nabla_\rho\nabla_t p_\rho^\sharp\big|_{\rho=0} = -\nabla_t\eta^\sharp - R(X,\dot q)p^\sharp.$$

For the bottom entry, we observe that $p(X) \equiv \langle p,\eta\rangle = \frac{1}{T}\eta(\dot q) \equiv 0$ due to the assumption that (X,η) takes values in the contact subbundle, and since $\nabla_t p \equiv 0$, we also have $0 \equiv \partial_t\left[p(X)\right] = p(\nabla_t X)$. Using the symmetry of the connection, we then find

$$(\nabla_\rho\dot q_\rho)_\flat - \nabla_\rho\left[p_\rho(\dot q_\rho)p_\rho\right]\big|_{\rho=0} = \nabla_t X_\flat - \underbrace{\left[\eta(\dot q) - p(\nabla_t X)\right]}_{=0} p - T\eta = \nabla_t X_\flat - T\eta.$$

The fact that $\nabla_t\dot q$ and $\nabla_t p$ vanish also implies incidentally that the covariant derivative operators on q^*TM and q^*T^*M induced by the Levi-Cività connection preserve the subbundles $N_q \subset q^*TM$ and $(N_q)_\flat \subset q^*T^*M$, which are the orthogonal complements of $\mathbb{R}\langle\dot q\rangle$ and $\mathbb{R}\langle p\rangle$ respectively. They also are canonically identified with the summands $\gamma^*\xi^h_{\mathrm{std}}$ and $\gamma^*\xi^v_{\mathrm{std}}$ in the horizontal/vertical splitting of $\gamma^*\xi_{\mathrm{std}}$, so the latter now acquires via this splitting a natural metric connection. Putting all of this together gives

$$\mathbf{A}_\gamma\begin{pmatrix} X \\ \eta\end{pmatrix} = \begin{pmatrix} -\nabla_t\eta^\sharp - R(X,\dot q)p^\sharp \\ \nabla_t X_\flat - T\eta\end{pmatrix},$$

which can also be written in terms of the connection on $\gamma^*\xi_{\mathrm{std}}$ described above as

$$\mathbf{A}_\gamma = -J_{\mathrm{std}}\nabla_t + \begin{pmatrix} -\frac{1}{T}R(\cdot,\dot q)\dot q & 0 \\ 0 & -T \end{pmatrix}. \tag{3.18}$$

Exercise 3.11.15. Use the classical symmetries of the Riemann tensor to show that $X \mapsto R(X,\dot q)\dot q$ defines a symmetric bundle map $N_q \to N_q$.

We can now fit the formulas above for $\nabla_N^2\mathcal{E}(q)$ and $\mathbf{A}_\gamma$ into the following more abstract context. Suppose (E,J,ω) is a Hermitian vector bundle over S^1, $\ell \subset E$ is a Lagrangian subbundle and we define a complementary Lagrangian subbundle by $\ell^\perp := J(\ell)$, hence

$$E = \ell \oplus \ell^\perp.$$

In terms of the bundle metric $\langle\ ,\ \rangle := \omega(\cdot, J\cdot)$ on E, the subbundles ℓ and $\ell^\perp$ are orthogonal, and J defines bundle isometries between them in both directions. For any choice of metric connection ∇ on ℓ, we can therefore use J to define a corresponding metric connection on $\ell^\perp$, and then define a connection on E via the splitting $\ell \oplus \ell^\perp$. For this connection on E, the data J, ω and $\langle\ ,\ \rangle$ are all parallel, and the subbundles ℓ and $\ell^\perp$ are preserved. We can then associate to ∇ and any symmetric linear bundle map $S \in \Gamma(\mathrm{End}^{\mathrm{sym}}(\ell))$ an L^2-symmetric second-order differential operator

$$\mathbf{H} := -\nabla_t^2 + S : \Gamma(\ell) \to \Gamma(\ell), \tag{3.19}$$

and additionally for any constant $T > 0$, an L^2-symmetric first-order differential operator

$$\mathbf{A} := -J\nabla_t + \begin{pmatrix} \frac{1}{T}S & 0 \\ 0 & -T \end{pmatrix} : \Gamma(E) \to \Gamma(E). \tag{3.20}$$

Observe that since ∇ respects the symplectic stucture on E, $-J\nabla_t$ is an asymptotic operator on (E,J,ω), and therefore so is any perturbation of $-J\nabla_t$ by a symmetric zeroth-oder term, in particular $\mathbf{A}$. The concrete example to keep in mind is of course when $q : S^1 \to M$ is a geodesic with speed T and $\gamma : S^1 \to ST^*M$ is its lift, so one can set $E := \gamma^*\xi_{\mathrm{std}}$, $\ell := \gamma^*\xi^h_{\mathrm{std}}$, $J := J_{\mathrm{std}}$ and $S := -R(\cdot,\dot q)\dot q$, with ∇ defined on ξ^h_{std} by horizontally lifting the restriction of the Levi-Civitá connection to N_q. By the same arguments that we've already used in this special case, $\mathbf{H}$ extends to a Fredholm operator $H^2(\ell) \to L^2(\ell)$ with index 0 and, as an unbounded operator on $L^2(\ell)$ with dense domain $H^2(\ell)$, its spectrum also consists of isolated real eigenvalues with finite multiplicity.

Lemma 3.11.16. *For any metric connection ∇ on ℓ and any $S \in \Gamma(\mathrm{End}^{\mathrm{sym}}(\ell))$ and $T > 0$, assume the operators $\mathbf{H}$ and $\mathbf{A}$ are defined as in (3.19) and (3.20) respectively. Then there is a natural isomorphism*

$$\ker \mathbf{A} \xrightarrow{\cong} \ker \mathbf{H} : \begin{pmatrix} \varphi \\ \psi \end{pmatrix} \mapsto \varphi.$$

Proof. Let us denote the isomorphisms $\ell \xrightarrow{J} \ell^\perp$ and $\ell^\perp \xrightarrow{J} \ell$ by $\sharp : \ell \to \ell^\perp$ and $-\flat : \ell^\perp \to \ell$ respectively, so that $\flat = (\sharp)^{-1}$ and J appears in block form with respect to the splitting $E = \ell \oplus \ell^\perp$ as $\begin{pmatrix} 0 & \sharp \\ -\flat & 0 \end{pmatrix}$. Then $\mathbf{A}\begin{pmatrix} \varphi \\ \psi \end{pmatrix} = 0$ means

$$-\nabla_t \psi^\sharp + \frac{1}{T} S\varphi = 0 \qquad \text{and} \qquad \nabla_t \varphi_\flat - T\psi = 0,$$

hence φ also satisfies $-\nabla_t^2 \varphi + S\varphi = \mathbf{H}\varphi = 0$. Conversely, if $\varphi \in \Gamma(\ell)$ is any solution to this equation, then setting $\psi := \frac{1}{T}\nabla_t \varphi_\flat \in \Gamma(\ell^\perp)$ gives $(\varphi, \psi) \in \ker \mathbf{A}$. □

LEMMA 3.11.17. *For any compact subset $K \subset \Gamma(\mathrm{End}^{\mathrm{sym}}(\ell))$, there exists a constant $c \geqslant 0$ such that for every $S \in K$ and every metric connection ∇ on ℓ, the operator $\mathbf{H}$ defined in (3.19) has no eigenvalues in $(-\infty, -c)$.*

PROOF. By the same calculation as in the proof of Lemma 3.11.9, it suffices to pick $c := \max_{S \in K} \|S\|_{L^\infty}$. □

It follows from Lemma 3.11.17 and the discreteness of the spectrum $\sigma(\mathbf{H})$ that $\mathbf{H}$ always has at most finitely many negative eigenvalues, counted with multiplicity. Let us define $\mathbf{H}$ to be **nondegenerate** if and only if $0 \notin \sigma(\mathbf{H})$, and call the count (with multiplicity) of negative eigenvalues in $\sigma(\mathbf{H})$ the **Morse index**

$$\mathrm{Morse}(\mathbf{H}) \geqslant 0$$

of the operator $\mathbf{H}$. Lemma 3.11.16 implies that for any constant $T > 0$, $\mathbf{H}$ is nondegenerate if and only if the associated asymptotic operator $\mathbf{A}$ defined in (3.20) is nondegenerate. The next statement therefore implies Theorem 3.11.12:

PROPOSITION 3.11.18. *For any metric connection ∇ on ℓ and any section $S \in \Gamma(\mathrm{End}^{\mathrm{sym}}(\ell))$ such that $\mathbf{H} := -\nabla_t^2 + S$ is nondegenerate, and any constant $T > 0$, the asymptotic operator $\mathbf{A}$ defined in (3.20) satisfies*

$$\mu_{CZ}^\tau(\mathbf{A}) = \mathrm{Morse}(\mathbf{H}) + \mu^\tau(E, \ell)$$

for every choice of symplectic trivialization τ of (E, ω), where $\mu^\tau(E, \ell)$ denotes the Maslov index of the loop of Lagrangian subspaces defined by writing $\ell \subset E$ in the trivialization τ.

There are two main steps in the proof: the first is to show that the formula is correct whenever S is replaced by $S + c$ for a sufficiently large constant $c \gg 0$, and we will then use spectral flow to derive the general case from this. For $\mathbf{H} + c$ with $c \gg 0$, Lemma 3.11.17 implies that the Morse index vanishes, so we are only left with the Maslov index and thus need to prove:

LEMMA 3.11.19. *For any metric connection ∇ on ℓ and any $S \in \Gamma(\mathrm{End}^{\mathrm{sym}}(\ell))$, there exists a constant $c_0 \in \mathbb{R}$ such that for every $c \geqslant c_0$ and $T > 0$, the asymptotic operator on (E, J, ω) defined by*

$$\mathbf{A}_c := -J\nabla_t + \begin{pmatrix} \frac{1}{T}(S + c) & 0 \\ 0 & -T \end{pmatrix}$$

satisfies

$$\mu_{CZ}^\tau(\mathbf{A}_c) = \mu^\tau(E, \ell)$$

for an arbitrary choice of symplectic trivialization τ.

PROOF. Let us replace τ with a homotopic unitary trivialization and use this to identify E with the trivial bundle $S^1 \times \mathbb{R}^{2n}$ with standard complex structure $J = J_0$ and standard Euclidean inner product. By the classification of homotopy classes of loops of Lagrangian subspaces (see [MS17, §2.3]), we can also assume the subbundle $\ell \subset E$ is given by

$$\ell_t = e^{\pi imt}\mathbb{R} \oplus \mathbb{R}^{n-1} \subset \mathbb{C}^n = \mathbb{R}^{2n}, \tag{3.21}$$

where $m := \mu^\tau(E, \ell) \in \mathbb{Z}$. Let ∇^0 denote the unique metric connection on ℓ that respects the splitting in (3.21) and is the trivial connection on the second summand. Given an arbitrary metric connection ∇ on ℓ, $\nabla^s := s\nabla + (1-s)\nabla^0$ then defines a smooth 1-parameter family of metric connections interpolating between ∇^0 and $\nabla^1 = \nabla$. Given $S \in \Gamma(\mathrm{End}^{\mathrm{sym}}(\ell))$ and constants $c, T > 0$, we can then consider the family of asymptotic operators

$$\mathbf{A}_c^s := -J_0\nabla_t^s + \begin{pmatrix} \frac{1}{T}(sS + c) & 0 \\ 0 & -T \end{pmatrix}, \qquad s \in [0,1].$$

Lemmas 3.11.16 and 3.11.17 imply that these are nondegenerate for every $s \in [0,1]$ if c is chosen to be sufficiently large, thus for the purposes of computing $\mu_{\mathrm{CZ}}^\tau(\mathbf{A}_c)$, we are free to replace $\mathbf{A}_c = \mathbf{A}_c^1$ with

$$\mathbf{A}_c^0 = -J_0\nabla_t^0 + \begin{pmatrix} \frac{c}{T} & 0 \\ 0 & -T \end{pmatrix},$$

whose kernel is isomorphic to that of $-(\nabla^0)^2 + c$. By the usual integration by parts calculation, the latter is a positive-definite operator for every $c > 0$, so we are now free to rescale both $c > 0$ and $T > 0$ at will in order to replace $\mathbf{A}_c^0$ with the simpler operator

$$-J_0\nabla_t^0 + \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}. \tag{3.22}$$

This operator respects the canonical splitting of $E = S^1 \times \mathbb{R}^{2n}$ into a direct sum of n trivial complex line bundles, and in light of the way our connection ∇^0 and the splitting $E = \ell \oplus \ell^\perp$ were defined, its restriction to each summand other than the first one matches the "standard" asymptotic operator that is used in normalizing the Conley-Zehnder index (cf. Definition 3.6.3), and thus has index 0. The Conley-Zehnder index of $\mathbf{A}_c$ therefore matches that of

$$\widehat{\mathbf{A}} := -J_0\widehat{\nabla}_t + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which is defined on the trivial complex line bundle $S^1 \times \mathbb{C}$, but with the block decomposition defined with respect to the splitting $S^1 \times \mathbb{C} = \widehat{\ell} \oplus i\widehat{\ell}$ with $\widehat{\ell}_t := e^{\pi imt}\mathbb{R} \subset \mathbb{C}$, and $\widehat{\nabla}$ denoting the unique connection such that $t \mapsto e^{\pi imt}$ and $t \mapsto ie^{\pi imt}$ locally define parallel sections. Local solutions to the equation $\widehat{\mathbf{A}}\eta = 0$ can easily be found in the form $\eta = fe^{\pi imt} + gie^{\pi imt}$ for real-valued functions f and g; the resulting parallel transport map $\Psi : \mathbb{R} \to \mathrm{Sp}(2)$ can then be written as

$$\Psi(t) = \begin{pmatrix} \cos(\pi mt) & -\sin(\pi mt) \\ \sin(\pi mt) & \cos(\pi mt) \end{pmatrix} \begin{pmatrix} \cosh(t) & -\sinh(t) \\ -\sinh(t) & \cosh(t) \end{pmatrix}.$$

One now uses Theorem 3.9.1 to compute $\mu_{CZ}(\Psi) = m$. □

Proposition 3.11.18 (and therefore also Theorem 3.11.12) will now follow if we can show that for any choice of the data ∇, S and T, the operators $\mathbf{H}$ from (3.19) and $\mathbf{A}$ from (3.20) are related to the adjusted asymptotic operator $\mathbf{A}_c$ from Lemma 3.11.19 by

$$\mu^{\tau}_{CZ}(\mathbf{A}) - \mu^{\tau}_{CZ}(\mathbf{A}_c) = \mathrm{Morse}(\mathbf{H})$$

whenever $c > 0$ is sufficiently large. In other words, we need to show that the spectral flow from $\mathbf{A}$ to $\mathbf{A}_c$ satisfies

$$\mu^{\mathrm{spec}}(\mathbf{A}, \mathbf{A}_c) = \mathrm{Morse}(\mathbf{H}). \tag{3.23}$$

Since the parameter $T > 0$ can always be changed without causing degeneracy, let us set $T := 1$ without loss of generality and thus consider the family of asymptotic operators

$$\mathbf{A}_s := -J\nabla_t + \begin{pmatrix} S + s & 0 \\ 0 & -1 \end{pmatrix}, \qquad s \in [0, c]$$

as a deformation from $\mathbf{A}$ to $\mathbf{A}_c$. The intuitive reason for (3.23) should be clear: as $\mathbf{H}$ is deformed to $\mathbf{H} + c$ via the family $\{\mathbf{H} + s\}_{s\in[0,c]}$, exactly $\mathrm{ind}(\mathbf{H})$ eigenvalues (counting multiplicity) pass from the negative to the positive real axis, so one expects the same for $\{\mathbf{A}_s\}_{s\in[0,c]}$ in light of the isomorphisms $\ker(\mathbf{H}+s) \cong \ker \mathbf{A}_s$ arising from Lemma 3.11.16. Making this intuition precise involves a couple of tricky issues: one is that the family $\{\mathbf{A}_s\}_{s\in[0,c]}$ cannot really be assumed to be generic, so it may involve non-simple crossings of eigenvalues, and another complication is that there is no obvious correspondence in general between eigenfunctions of $\mathbf{H}+s$ and $\mathbf{A}_s$ when the eigenvalues are nonzero. What we do know so far is that the spectrum of $\mathbf{H}$ is discrete, thus eigenvalues of the family $\{\mathbf{A}_s\}_{s\in[0,c]}$ change signs only finitely many times. We also know from Proposition 3.5.12 that for any parameter $s_0 \in [0, c]$ at which such a crossing occurs, the total number of eigenvalues (counting multiplicity) of any sufficiently small perturbation of $\mathbf{A}_{s_0}$ in a sufficiently small interval around 0 will be exactly $\dim\ker \mathbf{A}_{s_0}$. The relation (3.23) will thus hold if and only if $\dim\ker \mathbf{A}_{s_0}$ is always the actual contribution of the crossing at $s = s_0$ to the spectral flow of the family $\{\mathbf{A}_s\}_{s\in[0,c]}$, which is equivalent to the condition that for every s close to s_0, all eigenvalues of $\mathbf{A}_s$ in a small interval around 0 are negative for $s < s_0$ and positive for $s > s_0$. Theorem 3.11.12 thus follows from:

LEMMA 3.11.20. *Suppose $s_0 > 0$ and $I \subset \mathbb{R}$ is a compact interval containing 0 in its interior such that $\sigma(\mathbf{A}_{s_0}) \cap I = \{0\}$. Then for any $\epsilon > 0$ sufficiently small, all eigenvalues of $\mathbf{A}_s$ in I are negative for $s_0 - \epsilon < s < s_0$ and positive for $s_0 < s < s_0 + \epsilon$.*

To prove this, we begin by transforming it into a statement about compact self-adjoint operators. Assume s_0 and $I \subset \mathbb{R}$ are as stated in the lemma, then pick any $s_1 \in \mathbb{R}\backslash\sigma(\mathbf{A}_{s_0})$ and any $\epsilon > 0$ sufficiently small so that $\sigma(\mathbf{A}_s)$ also does not contain s_1 for any $s \in (s_0 - \epsilon, s_0 + \epsilon)$. Then $s_1 - \mathbf{A}_s$ is an invertible operator $H^1(E) \to L^2(E)$ for every s in this interval, and composing its inverse with the compact inclusion $H^1(E) \hookrightarrow L^2(E)$ gives rise to a family of compact self-adjoint operators

$$\mathbf{K}_s : L^2(E) \to L^2(E) : \eta \mapsto (s_1 - \mathbf{A}_s)^{-1}\eta, \qquad s \in (s_0 - \epsilon, s_0 + \epsilon).$$

A section $\eta \in \Gamma(E)$ is then an eigenfunction of $\mathbf{A}_s$ with eigenvalue $\lambda \in \mathbb{R}$ if and only if it is an eigenfunction of $\mathbf{K}_s$ with eigenvalue $\mu := \frac{1}{s_1 - \lambda}$, and we note that μ is a strictly increasing and smooth function of λ. Moreover, the family of bounded linear operators $\mathbf{A}_s \in \mathscr{L}(H^1(E), L^2(E))$ depends linearly on the parameter s, and the dependence of $\mathbf{K}_s \in \mathscr{L}(L^2(E))$ on s is therefore also smooth, with its derivative at $s = s_0$ given by

$$\dot{\mathbf{K}}_{s_0} := \frac{d}{ds}\mathbf{K}_s\bigg|_{s=s_0} = \mathbf{K}_{s_0}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\mathbf{K}_{s_0},$$

where the matrix in the middle can be interpreted simply as the orthogonal projection on $L^2(E)$ defined by projecting sections of $E = \ell \oplus \ell^\perp$ to the first summand. Wrting $J = \begin{pmatrix} 0 & \sharp \\ -\flat & 0 \end{pmatrix}$ as in the proof of Lemma 3.11.16, any nontrivial section $\begin{pmatrix} \varphi \\ \psi \end{pmatrix} \in \ker \mathbf{A}_{s_0}$ is a nontrivial solution to the system of ODEs

$$-\nabla_t \psi^\sharp + (S + s_0)\varphi = 0 \qquad \nabla_t \varphi_\flat - \psi = 0,$$

thus $\varphi \in \Gamma(\ell)$ is necessarily a nontrivial section, and moreover, $\begin{pmatrix} \varphi \\ \psi \end{pmatrix}$ is also an eigenfunction of $\mathbf{K}_{s_0}$ with eigenvalue $1/s_1$, implying

$$\begin{aligned}\left\langle \begin{pmatrix} \varphi \\ \psi \end{pmatrix}, \dot{\mathbf{K}}_{s_0}\begin{pmatrix} \varphi \\ \psi \end{pmatrix} \right\rangle_{L^2} &= \left\langle \mathbf{K}_{s_0}\begin{pmatrix} \varphi \\ \psi \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\mathbf{K}_{s_0}\begin{pmatrix} \varphi \\ \psi \end{pmatrix} \right\rangle_{L^2} \\ &= \frac{1}{s_1^2}\left\langle \begin{pmatrix} \varphi \\ \psi \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \varphi \\ \psi \end{pmatrix} \right\rangle_{L^2} = \frac{1}{s_1^2}\|\varphi\|_{L^2}^2 > 0.\end{aligned}$$

The following functional-analytic result therefore implies Lemma 3.11.20:

PROPOSITION 3.11.21. *Assume $\mathcal{H}$ is a Hilbert space, $J \subset \mathbb{R}$ is an open interval, $\{\mathbf{K}_s \in \mathscr{L}(\mathcal{H})\}_{s \in J}$ is a smooth family of compact self-adjoint operators, $\lambda_0 \in \sigma(\mathbf{K}_{s_0})$ is a nonzero eigenvalue for some $s_0 \in J$ and $I \subset \mathbb{R}$ is a compact interval containing λ_0 in its interior such that $\sigma(\mathbf{K}_{s_0}) \cap I = \{\lambda_0\}$. Suppose additionally that the derivative $\dot{\mathbf{K}}_{s_0} := \partial_s \mathbf{K}_s|_{s=s_0} \in \mathscr{L}(\mathcal{H})$ satisfies*

$$\langle x, \dot{\mathbf{K}}_{s_0} x \rangle > 0 \quad \text{for all} \quad x \in \ker(\mathbf{K}_{s_0} - \lambda_0)\backslash\{0\}.$$

Then for any $\epsilon > 0$ sufficiently small and every $s \in (s_0 - \epsilon, s_0 + \epsilon)$, the eigenvalues of $\mathbf{K}_s$ in I are all greater than λ_0 for $s > s_0$ and less than λ_0 for $s < s_0$.

PROOF. We claim first that if $\epsilon > 0$ is sufficiently small, then there exists a smooth family of orthogonal projections $\{\pi_s \in \mathscr{L}(\mathcal{H})\}_{s \in (s_0 - \epsilon, s_0 + \epsilon)}$ such that for every $s \in (s_0 - \epsilon, s_0 + \epsilon)$,

$$\operatorname{im} \pi_s = \bigoplus_{\lambda \in \sigma(\mathbf{K}_s) \cap I} \ker(\lambda - \mathbf{K}_s).$$

This follows from the holomorphic functional calculus for bounded linear operators (see e.g. [**BS18**, §5.2.4]). An explicit formula for π_s is given by the operator-valued Cauchy integral

$$\pi_s = \frac{1}{2\pi i}\int_C (z - \mathbf{K}_s)^{-1}\, dz,$$

where $C \subset \mathbb{C}$ can be chosen to be any oriented loop around λ_0 that does not touch or encircle any other eigenvalue of $\mathbf{K}_{s_0}$, and $\epsilon > 0$ is assumed small enough so that $\sigma(\mathbf{K}_s)$ remains disjoint from C when $|s - s_0| < \epsilon$.

Since π_s is self-adjoint and has finite-dimensional image, the complementary projection $\mathbb{1} - \pi_s$ is Fredholm, implying that the codimension of its image (which is the dimension of $\operatorname{im} \pi_s$) is an upper semi-continuous function of s. It follows that if $\epsilon > 0$ is sufficiently small, we can assume $\dim \operatorname{im} \pi_s \leqslant \dim \operatorname{im} \pi_{s_0}$ for all $s \in (s_0 - \epsilon, s_0 + \epsilon)$. On the other hand, π_{s_0} is clearly injective on the finite-dimensional space $\operatorname{im} \pi_{s_0}$, and therefore so is π_s for all s close enough to s_0, thus we can always assume in fact that π_s defines an *isomorphism*

$$\ker(\lambda_0 - \mathbf{K}_{s_0}) \xrightarrow{\pi_s} \bigoplus_{\lambda \in \sigma(\mathbf{K}_s) \cap I} \ker(\lambda - \mathbf{K}_s).$$

With this understood, the proposition now follows by observing that for every $x \in \ker(\lambda_0 - \mathbf{K}_{s_0})$ with $\|x\| = 1$, the smooth path $x_s := \frac{\pi_s x}{\|\pi_s x\|} \in \bigoplus_{\lambda \in \sigma(\mathbf{K}_s) \cap I} \ker(\lambda - \mathbf{K}_s) \subset \mathcal{H}$ satisfies $\langle x_{s_0}, \mathbf{K}_{s_0} x_{s_0} \rangle = \lambda_0$, $\langle \partial_s x_s, x_s \rangle = \frac{1}{2} \partial_s \langle x_s, x_s \rangle = 0$, and

$$\begin{aligned} \left.\frac{d}{ds} \langle x_s, \mathbf{K}_s x_s \rangle \right|_{s = s_0} &= \langle x, \dot{\mathbf{K}}_{s_0} x \rangle + 2 \langle \partial_s x_s |_{s = s_0}, \mathbf{K}_{s_0} x \rangle \\ &= \langle x, \dot{\mathbf{K}}_{s_0} x \rangle + 2\lambda_0 \langle \partial_s x_s |_{s = s_0}, x \rangle = \langle x, \dot{\mathbf{K}}_{s_0} x \rangle > 0. \end{aligned}$$

□

With Lemma 3.11.19 and (3.23) now established, the proof of Theorem 3.11.12 is complete.

3.12. Morse-Bott families

3.12.1. Clean intersection conditions. Nondegeneracy is a generic condition for contact forms, but the explicit examples one can construct are typically not generic, as they tend to have natural symmetries. In this context, it is useful to allow a relaxation of the nondegeneracy condition.

The correct notion is inspired in part by the idea of "clean" intersections. Recall that for a smooth map $f : M \to Q$ and a submanifold $N \subset Q$, f is said to intersect N **transversely** at $p \in M$ (written "$f \pitchfork N$") if $f(p) \in N$ and $\operatorname{im} df(p) + T_{f(p)} N = T_{f(p)} Q$. This condition implies via a standard application of the implicit function theorem that a neighborhood Σ of p in $f^{-1}(N)$ is a smooth submanifold of M with $T_p \Sigma = df(p)^{-1}\big(T_{f(p)} N\big)$. More generally, one says that f has a **clean intersection** with N at p if $f(p) \in N$ and there exists a smooth submanifold $\Sigma \subset M$ containing p such that

$$f(\Sigma) \subset N \qquad \text{and} \qquad T_p \Sigma = df(p)^{-1}\big(T_{f(p)} N\big). \tag{3.24}$$

Notice that for any submanifold $\Sigma \subset M$ through p satisfying $f(\Sigma) \subset N$, $T_p \Sigma$ is obviously contained in $df(p)^{-1}\big(T_{f(p)} N\big)$, so the nontrivial detail in (3.24) is that the subspace $df(p)^{-1}\big(T_{f(p)} N\big)$ is not any larger than it absolutely must be, given that the submanifold $\Sigma \subset f^{-1}(N) \subset Q$ exists. One often sees this condition stated in the context where M is another submanifold of Q and $f : M \to Q$ is the inclusion

$M \hookrightarrow Q$: the two submanifolds $M, N \subset Q$ are then said to intersect cleanly at $p \in M \cap N$ if there is a submanifold $\Sigma \subset Q$ containing p such that

$$\Sigma \subset M \cap N \qquad \text{and} \qquad T_p\Sigma = T_pM \cap T_pN. \tag{3.25}$$

Once again, the existence of the submanifold $\Sigma \subset M \cap N$ implies $\dim(T_pM \cap T_pN) \geqslant \dim \Sigma$, and the clean intersection condition then asks for $\dim(T_pM \cap T_pN)$ to be as small as possible, given the circumstances.

It turns out that sets of clean intersections are always smooth submanifolds, though in contrast to the transverse case, their dimensions cannot be easily predicted:

PROPOSITION 3.12.1. *If $f : M \to Q$ intersects $N \subset Q$ cleanly at $p \in M$, then the set $f^{-1}(N) \subset M$ is a smooth submanifold of M in some neighborhood of p, with $T_p\big(f^{-1}(N)\big) = df(p)^{-1}\big(T_{f(p)}N\big)$.*

PROOF. Choose a smooth function $g : \mathcal{U} \to \mathbb{R}^k$ on a neighborhood $\mathcal{U} \subset Q$ of $f(p)$ that has 0 as a regular value with $g^{-1}(0) = N \cap \mathcal{U}$, and let $f_{\mathcal{U}}$ denote the restriction of $f : M \to Q$ to the domain $f^{-1}(\mathcal{U}) \subset M$, which is an open neighborhood of p. Then $f^{-1}(N \cap \mathcal{U})$ is the zero set of the function $h := g \circ f_{\mathcal{U}} : f^{-1}(\mathcal{U}) \to \mathbb{R}^k$, and $\ker dh(p) = df(p)^{-1}\big(T_{f(p)}N\big)$. If we could assume $f \pitchfork N$, then $dh(p) : T_pM \to \mathbb{R}^k$ would now be surjective, but instead of assuming that, let us simply choose a linear projection $\pi : \mathbb{R}^k \to I := \operatorname{im} dh(p) \subset \mathbb{R}^k$ and consider the smooth map $\pi \circ h : f^{-1}(\mathcal{U}) \to I$. By construction, $d(\pi \circ h)(p) = \pi \circ dh(p)$ is a surjective map $T_pM \to I$ that has the same kernel as $dh(p)$, so the implicit function theorem implies that some neighborhood $\widetilde{\Sigma}$ of p in $(\pi \circ h)^{-1}(0) \subset M$ is a smooth submanifold of M with $T_p\widetilde{\Sigma} = \ker dh(p) = df(p)^{-1}\big(T_{f(p)}N\big)$. But the condition (3.24) provides another submanifold $\Sigma \subset M$ that has this same tangent space at p and is manifestly also contained in $\widetilde{\Sigma}$, since $\Sigma \subset f^{-1}(N)$ implies $h|_\Sigma \equiv 0$ and therefore also $(\pi \circ h)|_\Sigma \equiv 0$. It follows that neighborhoods of p in Σ and $\widetilde{\Sigma}$ are identical, and since $\widetilde{\Sigma}$ contains $f^{-1}(N) \cap f^{-1}(\mathcal{U})$, every point of $f^{-1}(N)$ sufficiently close to p therefore belongs to Σ. ☐

In the case of cleanly intersecting submanifolds $M, N \subset Q$, Proposition 3.12.1 implies that $M \cap N$ is also a submanifold and satisfies

$$T_p(M \cap N) = T_pM \cap T_pN \qquad \text{for all } p \in M \cap N.$$

Note however that nothing in this discussion gives any precise prediction for the dimension of $M \cap N$: it may in fact have different dimensions on different connected components, and these dimensions may be larger than in the case $M \pitchfork N$. Indeed, in the absence of the condition $T_pM + T_pN = T_pQ$, the intersection of T_pM and T_pN is allowed to have larger dimension than it would in the transverse case.

EXERCISE 3.12.2. Show that the clean intersection condition for a map $f : M \to Q$ and submanifold $N \subset Q$ is open, i.e. if it is satisfied at $p \in f^{-1}(N)$, then it is also satisfied at all points of $f^{-1}(N)$ in some neighborhood of p.

In Morse theory, nondegeneracy of a critical point $p \in M$ of a function $f : M \to \mathbb{R}$ is a transversality condition: it means that the section $df \in \Gamma(T^*M)$ is transverse to the zero-section of T^*M at p, and Morse functions are thus characterized by the property that the intersection of df with the zero-section is transverse everywhere. If one relaxes this to require only a clean intersection, the function $f : M \to \mathbb{R}$ is called **Morse-Bott** instead of Morse: the critical set $\operatorname{Crit}(f) := \{p \in M \mid df(p) = 0\}$ will then be a smooth submanifold of M, possibly with different dimensions on different connected components, whose tangent space at every point $p \in \operatorname{Crit}(f)$ satisfies

$$T_p \operatorname{Crit}(f) = \ker \nabla(df)(p),$$

where $\nabla(df)(p) : T_pM \to T_p^*M$ denotes the linearization of the section $df \in \Gamma(T^*M)$ at $p \in (df)^{-1}(0)$. If one choose a Riemannian metric, one can identify $\nabla(df)(p)$ with the usual Hessian $\nabla^2 f(p) : T_pM \to T_pM$ and thus write the Morse-Bott condition as

$$T_p \operatorname{Crit}(f) = \ker \nabla^2 f(p).$$

In order to state the analogue of this condition for a closed Reeb orbit $\gamma : S^1 \to M$ with respect to a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$, we suppose that γ belongs to a smoothly parametrized family of orbits

$$\{\gamma_\rho \in \widetilde{\mathcal{P}}(\mathcal{H})\}_{\rho \in \mathbb{D}^k}$$

with $\gamma_0 = \gamma$, where $\mathbb{D}^k \subset \mathbb{R}^k$ is the unit disk of some dimension $k \geqslant 0$. We have two immediate observations about such families:

PROPOSITION 3.12.3. *In any connected smooth family of closed orbits for a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$, all orbits have the same period.*

PROOF. This follows from the condition $d\lambda(R, \cdot) \equiv 0$ on the stable Hamiltonian structure, which implies for any $\rho \in \mathbb{D}^k$ that

$$\int_{S^1} \gamma_\rho^* \lambda - \int_{S^1} \gamma^* \lambda = 0,$$

since by Stokes' theorem, the latter is the integral of $d\lambda$ over an annulus tangent to R. □

REMARK 3.12.4. The analogue of Proposition 3.12.3 in Morse theory is the fact that for any smooth function $f : M \to \mathbb{R}$ containing a connected smooth submanifold $\Sigma \subset M$ on which df vanishes, f is constant. (Indeed, Proposition 3.12.3 could also have been proven by comparing the locally-defined action functional $\mathcal{A}_\beta$ from §3.3, whose critical points are closed Reeb orbits, with the functional $\Phi(\gamma) := \int_{S^1} \gamma^* \lambda$, whose values at critical points of $\mathcal{A}_\beta$ are the periods of those orbits—the crucial observation is then that since $d\lambda(R, \cdot) \equiv 0$, Φ is automatically critical at any critical point of $\mathcal{A}_\beta$.) Note however that this result depends crucially on the *smoothness* of the family of critical points. It is not true for smooth functions $f : M \to \mathbb{R}$ in general that for every critical point $p \in \operatorname{Crit}(f)$, f must have the same value at all other critical points near p. We will see in §11.3.2 that for any exact symplectic manifold Σ with a smooth function $f : \Sigma \to \mathbb{R}$, there is a natural way to define a contact form on $S^1 \times \Sigma$ that admits a bijective correspondence between critical points $p \in \Sigma$ of f and closed Reeb orbits of the form $S^1 \times \{p\} \subset S^1 \times \Sigma$; one can use this correspondence to construct examples of degenerate Reeb orbits that have other closed orbits of different periods in arbitrarily small neighborhoods.

Our second observation about the family of orbits $\{\gamma_\rho\}_{\rho\in\mathbb{D}^k}$ is that for every $v \in \mathbb{R}^k$, the section $\eta_v \in \Gamma(\gamma^*TM)$ defined by

$$\eta_v(t) := \partial_s \gamma_{sv}(t)\big|_{s=0} \tag{3.26}$$

is determined by the linearized Reeb flow, i.e. if γ has period $T > 0$ and φ_R^t denotes the flow of R, then $\eta_v(t) = d\varphi_R^{Tt}(\gamma(0))\eta_v(0)$. After shifting the parametrization of each orbit γ_ρ by a constant that depends smoothly on ρ, we can arrange for $\eta_v(0)$ to lie in $\xi_{\gamma(0)}$ for every $v \in \mathbb{R}^k$, in which case η_v will be a section of $\gamma^*\xi$ since the linearized Reeb flow preserves ξ. Exercise 3.3.4 implies in this situation that η_v is annihilated by the asymptotic operator $\mathbf{A}_\gamma$, hence (3.26) defines a linear map

$$\mathbb{R}^k \to \ker \mathbf{A}_\gamma : v \mapsto \eta_v.$$

This scenario is only interesting if the nearby orbits $\gamma_\rho : S^1 \to M$ for $\rho \neq 0$ are all different from γ, so a natural extra assumption to impose is that the linear map $\mathbb{R}^k \to \ker \mathbf{A}_\gamma$ should be injective, implying $\dim \ker \mathbf{A}_\gamma \geqslant k$. The Morse-Bott condition can then be summarized as the requirement that $\dim \ker \mathbf{A}_k$ should be as small as possible, given that a k-dimensional family of orbits near γ exists:

DEFINITION 3.12.5. Suppose M is equipped with a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$, and write $\xi = \ker \lambda$ and R for the associated hyperplane field and Reeb vector field. A closed Reeb orbit with parametrization $\gamma : S^1 \to M$ is called **Morse-Bott** if for some integer $k \geqslant 0$, there exists a smooth family of parametrized Reeb orbits $\{\gamma_\rho : S^1 \to M\}_{\rho\in\mathbb{D}^k}$ with $\gamma_0 = \gamma$ such that (3.26) defines a vector space isomorphism

$$\mathbb{R}^k \xrightarrow{\cong} \ker \mathbf{A}_\gamma : v \mapsto \eta_v$$

The stable Hamiltonian structure $\mathcal{H}$ is likewise called Morse-Bott (or in the contact case $\mathcal{H} = (d\alpha, \alpha)$, one calls α a **Morse-Bott contact form**) if all of its closed Reeb orbits are Morse-Bott.

The reader should take a moment to verify that the condition in this definition depends only on the underlying *unparametrized* Reeb orbit in $\mathcal{P}(\mathcal{H})$, and not on the choice of parametrization.

REMARK 3.12.6. The case $k = 0$ is allowed in Definition 3.12.5: the condition then means merely that $\ker \mathbf{A}_\gamma$ is trivial. Nondegenerate orbits are therefore also considered Morse-Bott.

Definition 3.12.5 can be found in various alternative forms in the literature, typically involving *Morse-Bott submanifolds* of M. We'll come to this in Proposition 3.12.9 below, but first, let's establish that the set of all closed Morse-Bott Reeb orbits really is smooth:

LEMMA 3.12.7. *For a Morse-Bott orbit γ belonging to a family $\{\gamma_\rho \in \widetilde{\mathcal{P}}(\mathcal{H})\}_{\rho\in\mathbb{D}^k}$ with $\gamma_0 = \gamma$ as described in Definition 3.12.5, the image of the map $\mathbb{D}^k \to \mathcal{P}(\mathcal{H}) : \rho \mapsto [\gamma_\rho]$ contains a neighborhood of $[\gamma]$ in the space $\mathcal{P}(\mathcal{H})$ of unparametrized closed orbits.*

PROOF. Write $x_0 := \gamma(0)$, let φ_R^t denote the time-t flow of R, and for a small neighborhood $\mathcal{U} \subset M$ of x_0 and a small interval $(T_0 - \epsilon, T_0 + \epsilon)$ around the period $T_0 > 0$ of γ, consider the smooth map

$$\Phi : (T_0 - \epsilon, T_0 + \epsilon) \times \mathcal{U} \to M \times M : (T, x) \mapsto (x, \varphi_R^T(x)).$$

Writing $\Delta \subset M \times M$ for the diagonal, a point $(T, x) \in (T_0 - \epsilon, T_0 + \epsilon) \times \mathcal{U}$ satisfies $\Phi(T, x) \in \Delta$ if and only if x is the starting point of a parametrized T-periodic orbit of R. In particular, the $(k+1)$-dimensional submanifold $\Sigma \subset (T_0 - \epsilon, T_0 + \epsilon) \times \mathcal{U}$ consisting of points $(T_0, \gamma_\rho(t))$ for $\rho \in \mathbb{D}^k$ and $t \in S^1$ near 0 satisfies $\Phi(\Sigma) \subset \Delta$, and we claim that all intersections of Φ with Δ near (T_0, x_0) belong to Σ. This will follow from Proposition 3.12.1 after showing that (T_0, x_0) is a clean intersection of Φ with Δ. To show the latter, we compute

$$d\Phi(T_0, x_0) : \mathbb{R} \oplus T_{x_0}M \to T_{(x_0,x_0)}(M \times M) : (s, X) \mapsto (X, d\varphi_R^T(x_0)X + sR(x_0)),$$

so $d\Phi(T_0, x_0)(s, X)$ is tangent to Δ if and only if $d\varphi_R^T(x_0)X + sR(x_0) = X$. Using the splitting $T_{x_0}M = \mathbb{R}R(x_0) \oplus \xi_{x_0}$ to decompose $X = tR(x_0) + X_\xi$ for $t \in \mathbb{R}$ and $X_\xi \in \xi_{x_0}$, this relation becomes the two equations

$$tR(x_0) + sR(x_0) = tR(x_0),$$
$$d\varphi_R^T(x_0)X_\xi = X_\xi,$$

implying that $s = 0$ while $t \in \mathbb{R}$ is arbitrary, and X_ξ is fixed by $d\varphi_R^T(x_0)$. Vectors X_ξ with the latter property are precisely the initial values of solutions $\eta \in \Gamma(\gamma^*\xi)$ to $\mathbf{A}_\gamma\eta = 0$, so by assumption, they form a k-dimensional subspace, and we conclude that $d\Phi(T_0, x_0)^{-1}\big(T_{(x_0,x_0)}\Delta\big)$ has dimension $k+1$. This matches the dimension of our submanifold Σ and thus proves the clean intersection condition. □

LEMMA 3.12.8. *Suppose γ is a closed Reeb orbit that is a d-fold cover of an embedded Reeb orbit, and also that γ belongs to a family $\{\gamma_\rho : S^1 \to M\}_{\rho \in \mathbb{D}^k}$ with $\gamma_0 = \gamma$ satisfying the conditions in Definition 3.12.5. Then for some neighborhood $\mathcal{D}^k \subset \mathbb{D}^k$ of 0, there exists a free smooth $\mathbb{Z}_d$-action on $\mathcal{D}^k \times S^1$ such that the map $\mathcal{D}^k \times S^1 \to M : (\rho, t) \mapsto \gamma_\rho(t)$ descends to a smooth embedding*

$$(\mathcal{D}^k \times S^1)/\mathbb{Z}_d \hookrightarrow M.$$

PROOF. Denote $F(\rho, t) := \gamma_\rho(t)$. By assumption, $\gamma(t + j/d) = \gamma(t)$ for each $j \in \mathbb{Z}_d$, and we claim that there exists a unique germ of a diffeomorphism $f_j : \mathbb{D}^k \to \mathbb{D}^k$ defined near 0 satisfying $f_j(0) = 0$ such that γ_ρ and $\gamma_{f_j(\rho)}(\cdot + j/d)$ are the same orbit up to a constant shift, which can be assumed arbitrarily small for ρ close to 0. The existence of a continuous function f_j with this property follows from Lemma 3.12.7, since $\gamma_\rho(\cdot - j/d)$ converges to $\gamma(\cdot - j/d) = \gamma$ as $\rho \to 0$, and it follows in particular that we can write

$$\gamma_{f_j(\rho)}(t + j/d + \tau_j(\rho)) = \gamma_\rho(t)$$

for a uniquely determined germ of a continuous function $\tau_j : \mathbb{D}^k \to \mathbb{R}$ defined near $0 \in \mathbb{D}^k$ such that $\tau_j(0) = 0$. One can use the implicit function theorem to show that f_j and τ_j are in fact both smooth, and they now give rise to a germ of a diffeomorphism $\psi_j : \mathbb{D}^k \times S^1 \to \mathbb{D}^k \times S^1$ defined near $\{0\} \times S^1$ by

$$\psi_j(\rho, t) := (f_j(\rho), t + j/d + \tau_j(\rho)),$$

which is the unique such map satisfying

$$F \circ \psi_j = F \qquad \text{and} \qquad \psi_j(0,t) = (0, t + j/d) \text{ for all } t \in S^1.$$

The uniqueness of f_j and ψ_j implies that they satisfy $f_{j+\ell} = f_j \circ f_\ell$ and $\psi_{j+\ell} = \psi_j \circ \psi_\ell$ for all $j, \ell \in \mathbb{Z}_d$, so they define smooth $\mathbb{Z}_d$-actions, and after replacing $\mathbb{D}^k$ with a sufficiently small $\mathbb{Z}_d$-invariant neighborhood $\mathcal{D}^k \subset \mathbb{D}^k$ of 0, we can now take $\mathcal{D}^k \times S^1$ as both domain and target for each ψ_j. Since the nontrivial sections of $\ker \mathbf{A}_\gamma$ are nowhere zero, the map F is an immersion, and the fact that γ is a d-fold cover of an embedded orbit implies that F becomes an embedding after letting it descend to the quotient $(\mathcal{D}^k \times S^1)/\mathbb{Z}_d$. □

Lemma 3.12.8 implies that the images of the orbits in a Morse-Bott family parametrized by $\mathbb{D}^k$ form a $(k+1)$-dimensional submanifold $S \subset M$. We call this a **Morse-Bott submanifold** of closed Reeb orbits: its tangent space at a point $\gamma(t) \in S$ is spanned by the Reeb vector field plus the set of all values at t of nontrivial sections $\eta \in \ker \mathbf{A}_\gamma \subset \Gamma(\gamma^*\xi)$. Since the latter are determined from their initial values via the linearized Reeb flow, they are precisely the set of vectors in $\xi_{\gamma(t)}$ that are fixed by $d\varphi_R^T(\gamma(t))$, and we therefore obtain the following equivalent formulation of the Morse-Bott condition from Definition 3.12.5:

PROPOSITION 3.12.9. *A submanifold $S \subset M$ invariant under the time-T flow φ_R^T of the Reeb vector field is a Morse-Bott submanifold of T-periodic orbits if and only if for all points $x \in S$,*

$$\ker(d\varphi_R^T(x) - \mathbb{1}) = T_x S.$$

Moreover, a T-periodic orbit is Morse-Bott if and only if its image belongs to a φ_R^T-invariant submanifold satisfying this condition. □

Here is a summary of the picture assembled from the results above:

THEOREM 3.12.10. *For any stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ on a manifold M, the Morse-Bott condition defines open subsets*

$$\widetilde{\mathcal{P}}^{\mathrm{MB}}(\mathcal{H}) \subset \widetilde{\mathcal{P}}(\mathcal{H}), \qquad \mathcal{P}^{\mathrm{MB}}(\mathcal{H}) = \widetilde{\mathcal{P}}^{\mathrm{MB}}(\mathcal{H})/S^1 \subset \widetilde{\mathcal{P}}(\mathcal{H})/S^1 = \mathcal{P}(\mathcal{H})$$

in the spaces of parametrized/unparametrized closed Reeb orbits respectively. Each connected component $\widetilde{\mathcal{P}}_0^{\mathrm{MB}}(\mathcal{H}) \subset \widetilde{\mathcal{P}}^{\mathrm{MB}}(\mathcal{H})$ consists of orbits that all have the same period, and it has a natural smooth manifold structure such that the map

$$\widetilde{\mathcal{P}}_0^{\mathrm{MB}}(\mathcal{H}) \to M : \gamma \mapsto \gamma(0)$$

is a smooth local diffeomorphism onto a smooth submanifold of M, i.e. a Morse-Bott submanifold. □

DEFINITION 3.12.11. We refer to the connected components of the open subsets $\widetilde{\mathcal{P}}^{\mathrm{MB}}(\mathcal{H}) \subset \widetilde{\mathcal{P}}(\mathcal{H})$ or $\mathcal{P}^{\mathrm{MB}}(\mathcal{H}) \subset \mathcal{P}(\mathcal{H})$ as (connected) **Morse-Bott families** of closed parametrized or unparametrized orbits respectively.

Note that by the Arzelà-Ascoli theorem, if M is compact, then every subset of $\widetilde{\mathcal{P}}(\mathcal{H})$ consisting of orbits with uniformly bounded periods is likewise compact, so Theorem 3.12.10 implies:

COROLLARY 3.12.12. *If M is compact and $\mathcal{H}$ is Morse-Bott, then the set of all periods of orbits in $\mathcal{P}(\mathcal{H})$ forms a discrete subset $\mathcal{T}(\mathcal{H})$ of $[0,\infty)$, and for each $T \in \mathcal{T}(\mathcal{H})$, the fixed point set of the time-T flow $\varphi_R^T : M \to M$ is a disjoint union of compact smooth submanifolds, each foliated by the images of closed Reeb orbits.*

PROOF. The only detail not covered by the previous results above is that the set $\mathcal{T}(\mathcal{H})$ cannot contain any sequence converging to 0. This is a standard fact about nowhere-zero vector fields on compact manifolds: one deduces it from the observation that every point $x \in M$ has a "flow-box" neighborhood $\mathcal{U} \subset M$ consisting of a union of segments of embedded orbits of some positive length $\epsilon_x > 0$, so that every closed orbit passing through $\mathcal{U}$ must have period at least ϵ_x. Taking a finite subcover of the resulting open cover of M then gives a positive lower bound for the periods of all closed orbits. □

EXAMPLE 3.12.13 (pre-quantization bundles). Suppose (X,ω) is a symplectic manifold such that $\omega/2\pi$ represents an integral cohomology class $[\omega/2\pi] \in H^2(X;\mathbb{Z})$; equivalently, this means that the integrals $\int_\Sigma f^*\omega$ for arbitrary closed oriented surfaces Σ and smooth maps $f : \Sigma \to X$ are always multiples of 2π. For each $k \in \mathbb{Z}$, one can then construct a Hermitian line bundle $E_k \to X$ with $c_1(E_k) = -k[\omega/2\pi]$, and its orthonormal frame bundle is a principal U(1)-bundle $\pi : M_k \to X$. Any choice of principal connection 1-form $A \in \Omega^1(M_k, \mathfrak{u}(1))$ then gives rise to a curvature 2-form $F_A = dA \in \Omega^2(M_k, \mathfrak{u}(1))$ that is (since U(1) is abelian) the pullback $\pi^*\Omega_A$ of a globally-defined curvature 2-form $\Omega_A \in \Omega^2(X, \mathfrak{u}(1))$, and according to Chern-Weil theory, $c_1(E_k) = \left[-\frac{1}{2\pi i}\Omega_A\right]$. This means Ω_A will always be cohomologous to $ik\omega$, in which case it is possible to choose a specific connection 1-form A such that Ω_A and $ik\omega$ are equal; writing $A = i\lambda$ to define a real-valued 1-form $\lambda \in \Omega^1(M_k)$, we then have

$$d\lambda = k\pi^*\omega. \tag{3.27}$$

The pair $\mathcal{H} := (\pi^*\omega, \lambda)$ is now a stable Hamiltonian structure on M_k whose induced hyperplane field $\xi = \ker\lambda$ is the horizontal subbundle defined by the connection, and the Reeb vector field is the fundamental vector field for the canonical generator of $\mathfrak{u}(1) = i\mathbb{R}$, so every fiber of $\pi : M_k \to X$ is a closed Reeb orbit of period 2π. Since the flow is periodic, the Morse-Bott condition is trivial to verify, and the entirety of M_k is a Morse-Bott submanifold.

When $k \geqslant 1$ in this example, (3.27) implies that λ is a contact form, and the case $k = 1$ specifically is known as the **Boothby-Wang construction** (cf. [Gei08, §7.2]). The contact manifold $(M_1, \xi = \ker\lambda)$ is also sometimes called the **prequantization** of the symplectic manifold (X,ω).

Thinking in terms of Morse-Bott submanifolds can sometimes obscure one or two subtleties that lurk behind the Morse-Bott condition. One is that in a connected Morse-Bott family $\mathcal{P}_0^{\mathrm{MB}}(\mathcal{H}) \subset \mathcal{P}^{\mathrm{MB}}(\mathcal{H})$, not all orbits need have the same covering multiplicity; equivalently, while all the orbits in $\mathcal{P}_0^{\mathrm{MB}}(\mathcal{H})$ have the same period, they need not have the same *minimal* period, meaning the period of the underlying simple orbit that they cover (cf. Exercise 3.12.14 below). As a consequence, the set $\mathcal{P}^{\mathrm{MB}}(\mathcal{H}) = \widetilde{\mathcal{P}}^{\mathrm{MB}}(\mathcal{H})/S^1$ of *unparametrized* Morse-Bott orbits is not always a smooth manifold, but can instead be an orbifold, because the order of the stabilizer subgroups of the natural S^1-action on $\widetilde{\mathcal{P}}^{\text{MB}}(\mathcal{H})$ can vary from point to point.

EXERCISE 3.12.14. Let (r, ϕ) denote polar coordinates on $\mathbb{R}^2$, with ϕ normalized to take values in $\mathbb{R}/\mathbb{Z}$ instead of $\mathbb{R}/2\pi\mathbb{Z}$, and let θ denote a similar coordinate on $S^1 = \mathbb{R}/\mathbb{Z}$. Consider a contact form on $S^1 \times \mathbb{R}^2$ of the form

$$\alpha = f(r)\, d\theta + g(r)\, d\phi$$

for smooth functions $f, g : [0, \infty) \to \mathbb{R}$, whose behavior near $r = 0$ is chosen to ensure that α has a smooth extension over the coordinate singularity at $S^1 \times \{0\}$.[7] The latter requires $g(0) = 0$ among other things, thus $f(0)$ must be nonzero.

(a) Show that α is a contact form on $S^1 \times (\mathbb{R}^2 \backslash \{0\})$ if and only if the function $fg' - f'g$ is nowhere zero, and write down a formula for its Reeb vector field under this assumption.

(b) Assuming the contact condition in part (a), show that for every $r_0 > 0$ such that $f'(r_0)/g'(r_0) \in \mathbb{Q} \cup \{\infty\}$, the 2-torus $\{r = r_0\}$ is foliated by closed Reeb orbits, and this torus is a Morse-Bott submanifold if and only if $\frac{d}{dr} \frac{f'(r)}{g'(r)}\Big|_{r=r_0} \neq 0$.

(c) Show that the contact condition is satisfied at $S^1 \times \{0\}$ if and only if $g''(0) \neq 0$, and in this case, $S^1 \times \{0\}$ is always a closed orbit.

(d) Assuming the condition in part (c), show that the d-fold cover of the embedded Reeb orbit at $S^1 \times \{0\}$ is nondegenerate if and only if $d\frac{f''(0)}{g''(0)} \notin \mathbb{Z}$.

(e) The result of part (d) implies that if α takes the form

$$\alpha = (1 + ar^2)\, d\theta + br^2\, d\phi$$

for constants $a, b \in \mathbb{R}$ that do not both vanish, then all covers of the orbit at $S^1 \times \{0\}$ are nondegenerate if and only if a/b is an irrational number. Assuming a/b is *not* irrational, show that for some $d \in \mathbb{N}$ depending on the values of a and b, the covers of this orbit up to multiplicity $d - 1$ are all nondegenerate, but the d-fold cover belongs to a 2-dimensional Morse-Bott family in which all other members of the family are embedded orbits.

People sometimes casually assume that all closed 2-dimensional Morse-Bott submanifolds are tori as in Exercise 3.12.14, since the existence of a nowhere-vanishing vector field dictates that they must have Euler characteristic zero. However, there is one other closed surface with vanishing Euler characteristic:

EXERCISE 3.12.15. Construct an explicit example of a contact 3-manifold containing a Morse-Bott Klein bottle.[8] *Hint: Start with a contact form on $\mathbb{T}^3$ defined by similar formulas as in Exercise 3.12.14, then divide it by a cleverly chosen $\mathbb{Z}_2$-action.*

[7] The precise condition required for this is that both of the functions $(r, \phi) \mapsto f(r)$ and $(r, \phi) \mapsto g(r)/r^2$ have smooth extensions over the coordinate singularity in $\mathbb{R}^2$.

[8] It is sometimes tempting to believe that oriented manifolds cannot contain closed non-orientable hypersurfaces—that is in fact true in $\mathbb{R}^n$, but it is not true in general. Consider for example $\mathbb{RP}^2 \subset \mathbb{RP}^3$.

3.12.2. The perturbed Conley-Zehnder indices. Degenerate Reeb orbits do not have well-defined Conley-Zehnder indices, but various generalizations of the usual definition can be used to extract meaningful information in the Morse-Bott case. The simplest and most useful for our purposes is as follows: for any (possibly degenerate) asymptotic operator $\mathbf{A}$ and trivialization τ, we can define two *perturbed* Conley-Zehnder indices

$$\mu_{\mathrm{CZ}}^{\tau,+}(\mathbf{A}) := \mu_{\mathrm{CZ}}^{\tau}(\mathbf{A} + \epsilon), \qquad \mu_{\mathrm{CZ}}^{\tau,-}(\mathbf{A}) := \mu_{\mathrm{CZ}}^{\tau}(\mathbf{A} - \epsilon), \tag{3.28}$$

where the scalar perturbation $\epsilon > 0$ is assumed arbitrarily small, and similarly for an arbitrary closed Reeb orbit γ,

$$\mu_{\mathrm{CZ}}^{\tau,\pm}(\gamma) := \mu_{\mathrm{CZ}}^{\tau,\pm}(\mathbf{A}_\gamma).$$

The discreteness of the spectrum ensures that $\mu_{\mathrm{CZ}}^{\tau,+}(\gamma)$ and $\mu_{\mathrm{CZ}}^{\tau,-}(\gamma)$ are both independent of the choice of small $\epsilon > 0$, and if γ is nondegenerate, then both are equal to $\mu_{\mathrm{CZ}}^{\tau}(\gamma)$. More generally, the relationship between the Conley-Zehnder index and spectral flow implies that the two perturbed indices differ from each other by

$$\mu_{\mathrm{CZ}}^{\tau,-}(\gamma) - \mu_{\mathrm{CZ}}^{\tau,+}(\gamma) = \dim\ker \mathbf{A}_\gamma;$$

if γ belongs to a Morse-Bott family $\mathcal{P}_0^{\mathrm{MB}}(\mathcal{H}) \subset \mathcal{P}(\mathcal{H})$, this difference is precisely the dimension of $\mathcal{P}_0^{\mathrm{MB}}(\mathcal{H})$. If one allows γ to move continuously through a Morse-Bott family of orbits, then the fact that $\dim\ker \mathbf{A}_\gamma$ stays constant implies that no nonzero eigenvalues of $\mathbf{A}_\gamma$ can change sign, and there is thus no spectral flow for either of the perturbed operators $\mathbf{A}_\gamma \pm \epsilon$. This implies that the perturbed Conley-Zehnder indices of a Morse-Bott orbit depend only on the Morse-Bott family to which it belongs:

PROPOSITION 3.12.16. *For any continuous path $\{\gamma_\rho\}_{\rho\in[0,1]}$ in a family of Morse-Bott orbits and any accompanying continuous family of trivializations $\{\tau_\rho\}_{\rho\in[0,1]}$ of $\gamma_\rho^*\xi$, $\mu_{CZ}^{\tau_0,\pm}(\gamma_0) = \mu_{CZ}^{\tau_1,\pm}(\gamma_1)$.* □

We can quickly deduce from this two useful facts about the topology of the bundle $\xi \to M$ along any Morse-Bott family:

COROLLARY 3.12.17. *For any connected Morse-Bott family $\mathcal{P}_0^{\mathrm{MB}}(\mathcal{H}) \subset \mathcal{P}(\mathcal{H})$ of closed orbits, any choice of trivialization τ_0 for ξ along one of the orbits $\gamma_0 \in \mathcal{P}_0^{\mathrm{MB}}(\mathcal{H})$ uniquely determines (up to homotopy) a trivialization τ along every other orbit $\gamma \in \mathcal{P}_0^{\mathrm{MB}}(\mathcal{H})$; the trivialization τ is characterized by the condition that it can be connected to τ_0 via a continuous family of trivializations along any continuous family of orbits in $\mathcal{P}_0^{\mathrm{MB}}(\mathcal{H})$ connecting γ_0 to γ.*

PROOF. It should be clear that given the trivialization τ_0 along γ_0, any continuous path in $\mathcal{P}_0^{\mathrm{MB}}(\mathcal{H})$ from γ_0 to γ uniquely determines a homotopy class of trivializations τ along γ. What needs to be proved is that τ (up to homotopy) does not depend on the choice of this path. To see this, observe that by Proposition 3.12.16, $\mu_{\mathrm{CZ}}^{\tau,\pm}(\gamma) = \mu_{\mathrm{CZ}}^{\tau_0,\pm}(\gamma_0)$, so the perturbed Conley-Zehnder indices are independent of any choice of path, and by Exercise 3.7.3, these determine τ up to homotopy. □

COROLLARY 3.12.18. *For any smoothly parametrized loop of parametrized Morse-Bott Reeb orbits $\{\gamma_\rho \in \widetilde{\mathcal{P}}(\mathcal{H})\}_{\rho\in S^1}$, the pullback of the vector bundle $\xi \to M$ via the map $\mathbb{T}^2 \to M : (\rho, t) \mapsto \gamma_\rho(t)$ is a trivial bundle.* □

3.12.3. The Robbin-Salamon index. There is another variant of the Conley-Zehnder index that is sometimes used for degenerate orbits, and while we will not need it in this book, its definition is useful to be aware of. Recall from §3.6 that the Conley-Zehnder index of a trivialized asymptotic operator $\mathbf{A} = -J_0\partial_t - S(t)$ can also be defined as something that assigns an integer to every continuous *nondegenerate symplectic arc*

$$\Psi : [0, 1] \to \mathrm{Sp}(2n), \qquad \Psi(0) = \mathbb{1}, \quad 1 \notin \sigma(\Psi(1)),$$

where the word "nondegenerate" in this case refers to the condition $1 \notin \sigma(\Psi(1))$. The relationship between Ψ and $\mathbf{A}$ comes from solving the differential equation $(-J_0\partial_t - S(t))\Psi(t) = 0$; equivalently, if $\mathbf{A}$ is the asymptotic operator of a nondegenerate orbit γ expressed in a trivialization τ, then Ψ represents the linearized Reeb flow along γ, and $\mu^\tau_{\mathrm{CZ}}(\gamma) = \mu_{\mathrm{CZ}}(\Psi)$. Robbin and Salamon introduced in [RS93] a generalization of μ_{CZ} that assigns a half-integer to arbitrary continuous symplectic arcs $\Psi : [a, b] \to \mathrm{Sp}(2n)$,

$$\mu_{\mathrm{RS}}(\Psi) \in \frac{1}{2}\mathbb{Z}.$$

This index is invariant under homotopies of Ψ with fixed end points, matches $\mu_{\mathrm{CZ}}(\Psi)$ when $\Psi(a) = \mathbb{1}$ and $1 \notin \sigma(\Psi(b))$, and additionally has the appealing property that it respects *concatenation* of arcs, meaning that for any $c \in (a, b)$,

$$\mu_{\mathrm{RS}}(\Psi) = \mu_{\mathrm{RS}}\big(\Psi|_{[a,c]}\big) + \mu_{\mathrm{RS}}\big(\Psi|_{[c,b]}\big).$$

Since this is defined in particular for arbitrary symplectic arcs $\Psi : [0, 1] \to \mathrm{Sp}(2n)$ with $\Psi(0) = \mathbb{1}$, one can then define the **Robbin-Salamon index** of an arbitrary closed orbit γ with respect to a trivialization τ as

$$\mu^\tau_{\mathrm{RS}}(\gamma) := \mu_{\mathrm{RS}}(\Psi) \in \frac{1}{2}\mathbb{Z}$$

by taking Ψ to be the linearized Reeb flow along γ expressed in the trivialization.

Using the perturbed Conley-Zehnder indices we defined in the previous subsection, a quick and dirty definition for the Robbin-Salamon index of an orbit γ can be stated as

$$\mu^\tau_{\mathrm{RS}}(\gamma) := \frac{\mu^{\tau,+}_{\mathrm{CZ}}(\gamma) + \mu^{\tau,-}_{\mathrm{CZ}}(\gamma)}{2}. \tag{3.29}$$

Since this looks very different from the original definition given in [RS93], we will now discuss the main ideas behind the original definition, and then show in Proposition 3.12.23 how they eventually lead to (3.29). The discussion will be framed at first in terms of spectral flow; at the end we will also sketch how the presentation of [RS93] via symplectic arcs can be derived from this (see Proposition 3.12.24).

Suppose we would like to generalize the integer-valued spectral flow $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+)$ from Theorem 3.5.1 so that $\mathbf{A}_-$ and $\mathbf{A}_+$ are allowed to have nontrivial kernel. Our presentation in §3.5.2 defined $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+)$ initially as a homotopy-invariant intersection number between a path from $\mathbf{A}_-$ to $\mathbf{A}_+$ in the space of symmetric Fredholm operators and the co-oriented codimension 1 submanifold consisting of operators with 1-dimensional kernel. This worked well because $\mathbf{A}_-$ and $\mathbf{A}_+$ themselves did not belong to that submanifold—however, defining intersection numbers is generally a much more delicate business if intersections are allowed to occur at the boundary. In certain situations, one can still define a homotopy-invariant count in which transverse boundary intersections contribute $\pm 1/2$ instead of ± 1, and the total intersection number therefore belongs to $\frac{1}{2}\mathbb{Z}$ instead of $\mathbb{Z}$. The prerequisiste for being able to do this is that behavior at the boundary is well enough understood to predict how boundary intersections can change under small perturbations.

In our situation, we can think in the following terms. For a smooth path $\{\mathbf{A}_s\}_{s\in[-1,1]}$ of asymptotic operators from $\mathbf{A}_- := \mathbf{A}_{-1}$ to $\mathbf{A}_+ := \mathbf{A}_1$, any sensible definition of spectral flow should obviously satisfy

$$\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+) = \mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_s) + \mu^{\mathrm{spec}}(\mathbf{A}_s, \mathbf{A}_+) \qquad \text{for } s \in (-1,1). \tag{3.30}$$

Suppose first that the $\mathbf{A}_-$ and $\mathbf{A}_+$ are nondegenerate. A point $s \in (-1,1)$ where $\ker \mathbf{A}_s \neq \{0\}$ will be called a **crossing**; for an arbitrary path, there may in general be infinitely many crossings, and the dimension of $\ker \mathbf{A}_s$ at a crossing cannot be predicted. Proposition 3.5.14 and Lemma 3.5.17 show however that *generic* smooth paths are much nicer: for these, there are only finitely many crossings, and they are all so-called **simple crossings**, meaning that the kernel for each is only 1-dimensional and each has a smooth family of simple eigenvalues $\lambda(s) \in \sigma(\mathbf{A}_s) \subset \mathbb{R}$ that passes transversely through 0 at the parameter value where the crossing occurs. Equivalently, a simple crossing is a transverse intersection of the path $s \mapsto \mathbf{A}_s$ with the manifold of symmetric Fredholm operators with 1-dimensional kernel. The sign of the number

$$\Lambda(\mathbf{A}_s, \dot{\mathbf{A}}_s) := \dot{\lambda}(s) \in \mathbb{R}\backslash\{0\}$$

determines whether a crossing contributes $+1$ or -1 to the spectral flow; the notation is meant to emphasize that it is determined by $\mathbf{A}_s$ and its derivative $\dot{\mathbf{A}}_s := \partial_s \mathbf{A}_s$. If we also formally define

$$\operatorname{sign} \Lambda(\mathbf{A}_s, \dot{\mathbf{A}}_s) := 0 \qquad \text{whenever } \ker \mathbf{A}_s = \{0\},$$

we obtain the formula

$$\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+) = \sum_{-1<s<1} \operatorname{sign} \Lambda(\mathbf{A}_s, \dot{\mathbf{A}}_s),$$

in which the sum is finite because there are only finitely-many values of s for which $\mathbf{A}_s$ is degenerate. Now, if $s \in (-1,1)$ is one of the parameter values at which a crossing occurs, then one reasonable way to to ensure the relation (3.30) is by letting this crossing contribute $\frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_s, \dot{\mathbf{A}}_s) = \pm\frac{1}{2}$ to each of $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_s)$ and $\mu^{\mathrm{spec}}(\mathbf{A}_s, \mathbf{A}_+)$. Pursuing this idea leads to a more general definition of $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+)$ as

$$\begin{aligned} \mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+) := \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_{-1}, \dot{\mathbf{A}}_{-1}) + \sum_{-1<s<1} \operatorname{sign}\Lambda(\mathbf{A}_s, \dot{\mathbf{A}}_s) \\ + \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_1, \dot{\mathbf{A}}_1) \in \frac{1}{2}\mathbb{Z}, \end{aligned} \tag{3.31}$$

which is still a finite sum and matches the previous formula if $\mathbf{A}_-$ and $\mathbf{A}_+$ are both nondegenerate, but also makes sense if either of the end points has 1-dimensional kernel, so long as the path $\mathbf{A}_s$ is chosen generically so as to make all crossings simple.

Before we can accept (3.31) as a reasonable definition, we need to check that it does not depend on the choice of (generic) path from $\mathbf{A}_-$ to $\mathbf{A}_+$. What makes this possible is the following easily verified property of $\operatorname{sign}\Lambda(\mathbf{A}_s, \dot{\mathbf{A}}_s)$:

LEMMA 3.12.19. *Suppose $\{\mathbf{A}_s\}_{s\in(-\delta,\delta)}$ and $\{\mathbf{A}'_s\}_{s\in(-\delta,\delta)}$ are two smooth paths of asymptotic operators that each have a simple crossing at $\mathbf{A}_0 = \mathbf{A}'_0 =: \mathbf{A}$. Then there exists an $\epsilon > 0$ such that $\mathbf{A}_s$ and $\mathbf{A}'_s$ are all nondegenerate for every $s \in [-\epsilon, \epsilon]\backslash\{0\}$, and for any s in this set,*

$$\mu^{\text{spec}}(\mathbf{A}_s, \mathbf{A}'_s) = \begin{cases} \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}, \dot{\mathbf{A}}'_0) - \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}, \dot{\mathbf{A}}_0) & \text{if } s > 0, \\ \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}, \dot{\mathbf{A}}_0) - \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}, \dot{\mathbf{A}}'_0) & \text{if } s < 0. \end{cases}$$

PROOF. For both paths there is a smooth family of simple eigenvalues $\lambda(s)$ or $\lambda'(s)$ that passes transversely through 0 at $s = 0$, and we can assume for $\epsilon > 0$ sufficiently small that all other eigenvalues stay outside some fixed neighborhood of 0. The spectral flow is therefore determined by the signs of $\lambda(s)$ and $\lambda'(s)$ for $s \neq 0$ close to 0, which are in turn determined by the signs of the derivatives $\dot{\lambda}(0)$ and $\dot{\lambda}'(0)$. □

LEMMA 3.12.20. *For asymptotic operators $\mathbf{A}_-$ and $\mathbf{A}_+$ with kernels of dimension at most 1, the definition of $\mu^{\text{spec}}(\mathbf{A}_-, \mathbf{A}_+)$ in (3.31) is independent of the choice of smooth path from $\mathbf{A}_-$ to $\mathbf{A}_+$ with only simple crossings.*

PROOF. The result is already known in the case where $\mathbf{A}_-$ and $\mathbf{A}_+$ are nondegenerate. For a general path γ of asymptotic operators with only simple crossings, let us denote the quantity on the right hand side of (3.31) by $\mu^{\text{spec}}(\gamma) \in \frac{1}{2}\mathbb{Z}$. Given two paths α, β from $\mathbf{A}_-$ and $\mathbf{A}_+$ parametrized on the interval $[-1, 1]$ with simple crossings, pick a small number $\epsilon > 0$ and break up each into three segments by defining

$$\alpha_- := \alpha|_{[-1,-1+\epsilon]}, \quad \alpha_0 := \alpha|_{[-1+\epsilon,1-\epsilon]}, \quad \alpha_+ := \alpha|_{[1-\epsilon,1]},$$
$$\beta_- := \beta|_{[-1,-1+\epsilon]}, \quad \beta_0 := \beta|_{[-1+\epsilon,1-\epsilon]}, \quad \beta_+ := \beta|_{[1-\epsilon,1]}.$$

Choose additionally two generic smooth paths

$$\gamma_- \text{ from } \alpha(-1+\epsilon) \text{ to } \beta(-1+\epsilon), \qquad \gamma_+ \text{ from } \alpha(1-\epsilon) \text{ to } \beta(1-\epsilon).$$

For $\epsilon > 0$ sufficiently small, we are free to assume the end points of γ_- and γ_+ are all nondegenerate operators, and Lemma 3.12.19 gives

$$\mu^{\text{spec}}(\gamma_-) = \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_-, \dot{\beta}(-1)) - \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_-, \dot{\alpha}(-1)),$$
$$\mu^{\text{spec}}(\gamma_+) = \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_+, \dot{\alpha}(1)) - \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_+, \dot{\beta}(1)).$$

Using this together with (3.30) and the fact that the result is already known in the nondegenerate case, we find

$$\begin{aligned}\mu^{\text{spec}}(\alpha) &= \mu^{\text{spec}}(\alpha_-) + \mu^{\text{spec}}(\alpha_0) + \mu^{\text{spec}}(\alpha_+)\\ &= \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_-, \dot{\alpha}(-1)) + \mu^{\text{spec}}(\alpha_0) + \frac{1}{2}\Lambda(\mathbf{A}_+, \dot{\alpha}(1))\\ &= \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_-, \dot{\alpha}(-1)) + \mu^{\text{spec}}(\gamma_-) + \mu^{\text{spec}}(\beta_0) - \mu^{\text{spec}}(\gamma_+) + \frac{1}{2}\Lambda(\mathbf{A}_+, \dot{\alpha}(1))\\ &= \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_-, \dot{\beta}(-1)) + \mu^{\text{spec}}(\beta_0) + \frac{1}{2}\operatorname{sign}\Lambda(\mathbf{A}_+, \dot{\beta}(1))\\ &= \mu^{\text{spec}}(\beta_-) + \mu^{\text{spec}}(\beta_0) + \mu^{\text{spec}}(\beta_+)\\ &= \mu^{\text{spec}}(\beta).\end{aligned}$$

□

It remains to extend the definition of $\mu^{\text{spec}}(\mathbf{A}_-, \mathbf{A}_+)$ for cases where $\mathbf{A}_+$ or $\mathbf{A}_-$ has kernel of dimension greater than 1. This presents a slight conundrum, because according to Lemma 3.5.17, generic paths of asymptotic operators should in this case not pass through $\mathbf{A}_\pm$ at all—this means that it will no longer be possible frame the discussion in terms of generic paths. The trick will instead be to consider a condition that is open and dense within the space of paths that are *constrained* to pass through a specific operator, whose kernel may have arbitrary dimension.

The analysis needed for this was set up in §3.5.1–§3.5.3. Suppose $E \to S^1$ is a Hermitian vector bundle, and recall that $\mathcal{A}(E)$ denotes the space of asymptotic operators of class L^∞ on E, which is an affine space over $L^\infty(\operatorname{End}_{\mathbb{R}}^{\text{sym}}(E))$. If $\mathbf{A}_0 \in \mathcal{A}(E)$ has nontrivial kernel $K := \ker \mathbf{A}_0 \subset H^1(E)$, then arguing as in the proof of Proposition 3.5.12, we find an open neighborhood $\mathcal{O}(\mathbf{A}_0) \subset \mathcal{A}(E)$ of $\mathbf{A}_0$, an interval $(-\delta, \delta) \subset \mathbb{R}$ and a smooth map

$$\Phi : \mathcal{O}(\mathbf{A}_0) \times (-\delta, \delta) \to \operatorname{End}_{\mathbb{R}}^{\text{sym}}(K)$$

such that

$$\ker \Phi(\mathbf{A}, \lambda) \cong \ker(\mathbf{A} - \lambda) \qquad \text{for all } (\mathbf{A}, \lambda) \in \mathcal{O}(\mathbf{A}_0) \times (-\delta, \delta).$$

Concretely, if we write $\mathbf{A} \in \mathcal{O}(\mathbf{A}_0)$ in block form as

$$\mathbf{A} = \begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{pmatrix} : (K^\perp \cap H^1(E)) \oplus K \to K^\perp \oplus K,$$

then Φ is given by

$$\Phi(\mathbf{A}, \lambda) = (\mathbf{d} - \lambda) - \mathbf{c}(\mathbf{a} - \lambda)^{-1}\mathbf{b} \in \operatorname{End}^{\text{sym}}(K). \tag{3.32}$$

The partial differential of Φ at $\mathbf{A}_0$ in the direction of $\dot{\mathbf{A}} \in T_{\mathbf{A}_0}\mathcal{A}(E) = L^\infty(\operatorname{End}_{\mathbb{R}}^{\text{sym}}(E))$ thus defines a symmetric linear operator on $\ker \mathbf{A}_0$,

$$\Gamma(\mathbf{A}_0, \dot{\mathbf{A}}) := D_1\Phi(\mathbf{A}_0)\dot{\mathbf{A}} \in \operatorname{End}^{\text{sym}}(\ker \mathbf{A}_0),$$

given by

$$\Gamma(\mathbf{A}_0, \dot{\mathbf{A}}) := \Pi \circ \dot{\mathbf{A}}\big|_{\ker \mathbf{A}_0},$$

where $\Pi : L^2(E) \to \ker \mathbf{A}_0$ denotes the orthogonal projection. This object is called the **crossing form**: it associates to each asymptotic operator $\mathbf{A}$ and each tangent vector $\dot{\mathbf{A}} \in T_{\mathbf{A}}\mathcal{A}(E) = L^\infty(\mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(E))$ to the space of asymptotic operators a symmetric linear map $\Gamma(\mathbf{A}, \dot{\mathbf{A}})$ on $\ker \mathbf{A}$, or equivalently, the real-valued quadratic form $\langle \cdot, \Gamma(\mathbf{A}, \dot{\mathbf{A}}) \cdot \rangle_{L^2}$ on $\ker \mathbf{A}$. The **signature**

$$\operatorname{sign} \Gamma(\mathbf{A}, \dot{\mathbf{A}}) \in \mathbb{Z}$$

of this quadratic form is by definition the number of positive eigenvalues of $\Gamma(\mathbf{A}, \dot{\mathbf{A}})$ minus the number of negative eigenvalues (counted with multiplicity). For a smooth path $\{\mathbf{A}_s\}_{s \in [-1,1]}$ with a crossing at $s = s_0$, writing $\dot{\mathbf{A}}_{s_0} := \partial_s \mathbf{A}_s\big|_{s=s_0}$, the crossing is called **regular** if the crossing form $\Gamma(\mathbf{A}_{s_0}, \dot{\mathbf{A}}_{s_0}) : \ker \mathbf{A}_{s_0} \to \ker \mathbf{A}_{s_0}$ is invertible.

We can now generalize Lemma 3.12.19 as follows:

LEMMA 3.12.21. *Suppose $\{\mathbf{A}_s\}_{s \in (-\delta,\delta)}$ and $\{\mathbf{A}'_s\}_{s \in (-\delta,\delta)}$ are two smooth paths of asymptotic operators that each have a regular crossing at $\mathbf{A}_0 = \mathbf{A}'_0 =: \mathbf{A}$. Then there exists an $\epsilon > 0$ such that $\mathbf{A}_s$ and $\mathbf{A}'_s$ are all nondegenerate for every $s \in [-\epsilon, \epsilon] \backslash \{0\}$, and for any s in this set,*

$$\mu^{\mathrm{spec}}(\mathbf{A}_s, \mathbf{A}'_s) = \begin{cases} \frac{1}{2} \operatorname{sign} \Gamma(\mathbf{A}, \dot{\mathbf{A}}'_0) - \frac{1}{2} \operatorname{sign} \Gamma(\mathbf{A}, \dot{\mathbf{A}}_0) & \text{if } s > 0, \\ \frac{1}{2} \operatorname{sign} \Gamma(\mathbf{A}, \dot{\mathbf{A}}_0) - \frac{1}{2} \operatorname{sign} \Gamma(\mathbf{A}, \dot{\mathbf{A}}'_0) & \text{if } s < 0. \end{cases}$$

PROOF. Write $K := \ker \mathbf{A}$ and $V := K^\perp \cap H^1(E)$, where $K^\perp \subset L^2(E)$ is the L^2-orthogonal complement, and suppose the crossing form $\Gamma(\mathbf{A}, \dot{\mathbf{A}}_0)$ has k_+ positive and k_- negative eigenvalues (counting multiplicity), with $k := \dim K = k_+ + k_-$. There is a function $\Phi : \mathcal{O}(\mathbf{A}) \times (-\delta, \delta) \to \mathrm{End}^{\mathrm{sym}}(K)$ defined for some neighborhood $\mathcal{O}(\mathbf{A}) \subset \mathcal{A}(E)$ of $\mathbf{A}$ and some $\delta > 0$ such that $\Phi(\mathbf{A}_s, \lambda) : K \to K$ has kernel isomorphic to that of $\mathbf{A}_s - \lambda$ for each s and λ close to 0. Following (3.32), we can write

$$\Phi(\mathbf{A}_s, \lambda) = \mathbf{d}_s - \mathbf{c}_s(\mathbf{a}_s - \lambda)^{-1}\mathbf{b}_s - \lambda =: \Psi(s, \lambda) - \lambda,$$

where $\mathbf{a}_s$, $\mathbf{b}_s$, $\mathbf{c}_s$ and $\mathbf{d}_s$ are smooth families of operators representing the components of $\mathbf{A}_s$ in its block decomposition as an operator $V \oplus K \to K^\perp \oplus K$. We see from this formula that the functions Φ and Ψ both depend smoothly on s and real-analytically on λ; they also admit complex-analytic extensions that are defined for complex λ near 0 and take values in the complexification of $\mathrm{End}^{\mathrm{sym}}(K)$. We can therefore also view

$$f_s(\lambda) := \det \Phi(\mathbf{A}_s, \lambda) \in \mathbb{C}$$

as a smoothly s-dependent family of holomorphic functions defined for $\lambda \in \mathbb{C}$ in some neighborhood of 0, and by Exercise 3.5.13, the zeroes of f_s are precisely the eigenvalues of $\mathbf{A}_s$, with multiplicity matching the order of the zero. Now, since $\Psi(0, \lambda) = 0$ for every λ, the fundamental theorem of calculus gives

$$\Psi(s, \lambda) = \int_0^1 \frac{d}{d\tau} \Psi(\tau s, \lambda)\, d\tau = s \int_0^1 \partial_s \Psi(\tau s, \lambda)\, d\tau =: s\widehat{\Psi}(s, \lambda),$$

thus defining a new function $\widehat{\Psi}(s,\lambda)$ that is likewise smooth in s and holomorphic in λ, such that

$$\widehat{\Psi}(0,0) = \partial_s \Psi(0,0) = \partial_s \mathbf{d}_s\big|_{s=0} = D_1\Phi(\mathbf{A},0)\dot{\mathbf{A}}_0 = \Gamma(\mathbf{A},\dot{\mathbf{A}}_0).$$

Writing

$$f_s(s\lambda) = \det\left(s\widehat{\Psi}(s,s\lambda) - s\lambda\right) = s^k \det\left(\widehat{\Psi}(s,s\lambda) - \lambda\right) =: s^k g_s(\lambda)$$

defines a new smoothly s-dependent family of holomorphic functions g_s such that $g_0(\lambda) = \det\left(\Gamma(\mathbf{A},\dot{\mathbf{A}}_0) - \lambda\right)$, so our assumptions on $\Gamma(\mathbf{A},\dot{\mathbf{A}}_0)$ imply via Exercise 3.5.13 that g_0 has exactly k_+ positive and k_- negative zeroes, counting multiplicity. It follows that for any $L > 0$ larger than the magnitudes of all the eigenvalues of $\Gamma(\mathbf{A},\dot{\mathbf{A}}_0)$, the interval $[-L,L]$ also contains exactly k_+ positive and k_- negative zeroes (counting multiplicity) of g_s for each $s \neq 0$ close enough to 0. Turning this into a statement about the zeroes of f_s and applying Exercise 3.5.13 once more, we deduce that for some $\epsilon > 0$ small, the interval $[-\epsilon,\epsilon]$ contains k_+ positive and k_- negative eigenvalues of $\mathbf{A}_s$ for each $s > 0$ sufficiently small, or the same statement with the signs reversed if $s < 0$.

Applying the same arguments to $\mathbf{A}'_s$, we can now fix $\epsilon > 0$ small and choose for any $s \neq 0$ close enough to 0 a path from $\mathbf{A}_s$ to $\mathbf{A}'_s$ that stays in a small enough neighborhood of $\mathbf{A}$ to guarantee that no eigenvalues cross $\pm\epsilon$. The spectral flow is thus determined entirely by the comparative distribution of positive and negative eigenvalues of $\mathbf{A}_s$ and $\mathbf{A}'_s$ close to 0, which is determined by the crossing form as described above. □

With Lemma 3.12.21 in hand, the same calculation as in the proof of Lemma 3.12.20 (with Λ replaced by Γ) makes the following definition independent of choices:

DEFINITION 3.12.22. Given arbitrary asymptotic operators $\mathbf{A}_\pm \in \mathcal{A}(E)$ on a Hermitian vector bundle $E \to S^1$, the (half-integer-valued) **spectral flow** $\mu^{\text{spec}}(\mathbf{A}_-,\mathbf{A}_+) \in \frac{1}{2}\mathbb{Z}$ from $\mathbf{A}_-$ to $\mathbf{A}_+$ is defined by

$$\mu^{\text{spec}}(\mathbf{A}_-,\mathbf{A}_+) := \frac{1}{2}\operatorname{sign}\Gamma(\mathbf{A}_{-1},\dot{\mathbf{A}}_{-1}) + \sum_{-1<s<1}\operatorname{sign}\Gamma(\mathbf{A}_s,\dot{\mathbf{A}}_s) + \frac{1}{2}\operatorname{sign}\Gamma(\mathbf{A}_1,\dot{\mathbf{A}}_1),$$

where $\{\mathbf{A}_s \in \mathcal{A}(E)\}_{s\in[-1,1]}$ is any choice of smooth path from $\mathbf{A}_- = \mathbf{A}_{-1}$ to $\mathbf{A}_+ = \mathbf{A}_1$ with only regular crossings, and $\dot{\mathbf{A}}_s := \partial_s\mathbf{A}_s \in T_{\mathbf{A}_s}\mathcal{A}(E) = L^\infty(\operatorname{End}_{\mathbb{R}}^{\text{sym}}(E))$.

This generalized notion of spectral flow leads to an obvious half-integer-valued generalization of the Conley-Zehnder index for arbitrary asymptotic operators $\mathbf{A} \in \mathcal{A}(E)$ with respect to a trivialization τ of E: the **Robbin-Salamon index** $\mu^\tau_{\text{RS}}(\mathbf{A})$ is uniquely determined by the conditions

$$\mu^\tau_{\text{RS}}(\mathbf{A}_-) - \mu^\tau_{\text{RS}}(\mathbf{A}_+) = \mu^{\text{spec}}(\mathbf{A}_-,\mathbf{A}_+) \qquad \text{for all } \mathbf{A}_\pm \in \mathcal{A}(E) \tag{3.33}$$

and

$$\mu^\tau_{\text{RS}}(\mathbf{A}) = \mu^\tau_{\text{CZ}}(\mathbf{A}) \qquad \text{for all } \mathbf{A} \text{ nondegenerate.} \tag{3.34}$$

If E is the trivial bundle $S^1 \times \mathbb{R}^{2n}$, one can of course omit the trivialization from the notation and simply write $\mu_{\mathrm{RS}}(\mathbf{A})$. For a closed Reeb orbit γ, even a degenerate one, one can now define

$$\mu_{\mathrm{RS}}^{\tau}(\gamma) := \mu_{\mathrm{RS}}^{\tau}(\mathbf{A}_{\gamma}).$$

PROPOSITION 3.12.23. *The definition of $\mu_{RS}^{\tau}(\gamma) \in \frac{1}{2}\mathbb{Z}$ in terms of spectral flow is equivalent to* (3.29).

PROOF. An equivalent statement purely in terms of asymptotic operators would be that for any $\mathbf{A} \in \mathcal{A}(E)$ and trivialization τ of E,

$$\mu_{\mathrm{RS}}^{\tau}(\mathbf{A}) = \frac{\mu_{\mathrm{CZ}}^{\tau}(\mathbf{A} - \epsilon) + \mu_{\mathrm{CZ}}^{\tau}(\mathbf{A} + \epsilon)}{2}$$

for $\epsilon > 0$ sufficiently small. Using (3.33) and (3.34) to characterize μ_{RS}, this formula follows by computing

$$\mu^{\mathrm{spec}}(\mathbf{A} - \epsilon, \mathbf{A}) = \frac{1}{2} \dim \ker \mathbf{A} = \mu^{\mathrm{spec}}(\mathbf{A}, \mathbf{A} + \epsilon),$$

for which one can use the path $s \mapsto \mathbf{A} + s$, as it has a single regular crossing at $s = 0$ whose crossing form is positive definite. □

As mentioned at the beginning of this subsection, the original definition of the Robbin-Salamon index in [**RS93**] frames it in terms of symplectic arcs $\Psi : [a, b] \to \mathrm{Sp}(2n)$ rather than spectral flow. A general definition of $\mu_{\mathrm{RS}}(\Psi)$ can be expressed as follows.[9] For each $k \in \mathbb{N}$, the Lie group $\mathrm{Sp}(2n)$ has a smooth submanifold

$$\mathrm{Sp}_k(2n) := \left\{ A \in \mathrm{Sp}(2n) \;\middle|\; \dim \ker(A - \mathbb{1}) = k \right\},$$

i.e. the set of all symplectic linear transformations for which 1 is an eigenvalue of geometric multiplicity k.[10] The union of these submanifolds for all $k > 0$ is an algebraic variety called the **Maslov cycle**. The stratum $\mathrm{Sp}_1(2n) \subset \mathrm{Sp}(2n)$ has codimension 1, while the others all have strictly larger codimension, so generic arcs in $\mathrm{Sp}(2n)$ can be made to miss $\mathrm{Sp}_k(2n)$ for $k \geqslant 2$ and intersect $\mathrm{Sp}_1(2n)$ transversely, and these intersections can be counted with signs (including factors of 1/2 for intersections at an end point). For arcs with end points lying in $\mathrm{Sp}_k(2n)$ for $k \geqslant 2$, one needs a more constrained notion of generic intersections, which are also called *regular crossings* and can be characterized via the nonsingularity of a quadratic form, the *crossing form* $\Gamma(\Psi, \dot{\Psi})$. Analogously to Lemma 3.12.21, one can use the signature of this crossing form to characterize, for any two symplectic arcs Ψ, Ψ' in $\mathrm{Sp}(2n)$ with regular crossings at the same point $\Psi(0) = \Psi'(0) \in \mathrm{Sp}_k(2n)$, the signed count

[9]The definition by Robbin and Salamon derives their generalized Conley-Zehnder index from a Maslov index for paths in the space of pairs of Lagrangian subspaces of $\mathbb{R}^{2n}$. The relation between this and the Conley-Zehnder index for symplectic arcs is elucidated in [**RS93**, §5].

[10]Caution: For symplectic matrices, the geometric and algebraic multiplicity of an eigenvalue need not match. The strata $\mathrm{Sp}_k(2n)$ are defined specifically in terms of geometric multiplicity, not algebraic.

of intersections with $\mathrm{Sp}_1(2n)$ for a generic symplectic arc from $\Psi(\pm\epsilon)$ to $\Psi'(\pm\epsilon)$ for $\epsilon > 0$ small (cf. the bottom of page 832 in [RS93]). As a consequence, the index

$$\begin{aligned} \mu_{\mathrm{RS}}(\Psi) := \frac{1}{2}\operatorname{sign}\Gamma(\Psi(a),\dot{\Psi}(a)) + \sum_{a<t<b}\operatorname{sign}\Gamma(\Psi(t),\dot{\Psi}(t)) \\ + \frac{1}{2}\operatorname{sign}\Gamma(\Psi(b),\dot{\Psi}(b)) \in \frac{1}{2}\mathbb{Z} \end{aligned} \tag{3.35}$$

for a smooth symplectic arc $\Psi : [a,b] \to \mathrm{Sp}(2n)$ with only regular crossings depends only on its homotopy class with fixed end points. Observe that for the Conley-Zehnder index, we always consider symplectic arcs Ψ starting at $\mathbb{1}$, which lies in the smallest stratum $\mathrm{Sp}_{2n}(2n)$ of the Maslov cycle, thus the standard theory of transverse intersections is not sufficient for understanding the Conley-Zehnder index from this perspective—it is essential to also understand regular crossings and the crossing form.

PROPOSITION 3.12.24. *Under the natural correspondence between trivialized asymptotic operators* $\mathbf{A}$ *and symplectic arcs* $\Psi : [0,1] \to \mathrm{Sp}(2n)$ *with* $\Psi(0) = \mathbb{1}$*, the definition of* $\mu_{RS}(\Psi)$ *in* (3.35) *matches our definition of* $\mu_{RS}(\mathbf{A})$ *in terms of spectral flow.*

PROOF SKETCH. We claim first that the result holds for $n = 1$ and the symplectic arc $\Psi_0 : [0,1] \to \mathrm{Sp}(2)$ given by

$$\Psi_0(t) := \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}.$$

Indeed, the corresponding asymptotic operator $\mathbf{A}_0$ is nondegenerate, so $\mu_{\mathrm{RS}}(\mathbf{A}_0) = \mu_{\mathrm{CZ}}(\mathbf{A}_0)$, and Theorem 3.9.1 gives $\mu_{\mathrm{CZ}}(\mathbf{A}_0) = 0$. In order to compute (3.35), we observe that Ψ_0 has no crossings other than the required one at $t = 0$, so the claim will follow if we can show that the crossing form there has signature 0. Without needing to compute the crossing form, we can deduce this from homotopy invariance: the obvious extension $\widehat{\Psi}_0$ of Ψ_0 to an arc $[-1,1] \to \mathrm{Sp}(2)$ also has only the one crossing at $t = 0$, and we claim that it is homotopic with fixed end points to an arc $\widehat{\Psi}_0' : [-1,1] \to \mathrm{Sp}(2)$ that has no crossings at all, implying

$$\operatorname{sign}\Gamma(\mathbb{1},\dot{\Psi}_0(0)) = \mu_{\mathrm{RS}}(\widehat{\Psi}_0) = \mu_{\mathrm{RS}}(\widehat{\Psi}_0') = 0.$$

To see the homotopy in question, let $R(\theta) : \mathbb{R}^2 \to \mathbb{R}^2$ denote the rotation by angle θ and observe that for each $\tau \in [-1,1]$, we have

$$R(\pi/2)^{-1}\widehat{\Psi}_0(\tau)R(\pi/2) = -J_0\widehat{\Psi}_0(\tau)J_0 = \widehat{\Psi}_0(-\tau).$$

For each $\tau \in (0,1)$, we can therefore construct a symplectic arc from $\widehat{\Psi}(-1)$ to $\widehat{\Psi}(1)$ that follows initially a segment of $\widehat{\Psi}$ from $t = -1$ to $t = -\tau$, then conjugates $\widehat{\Psi}(-\tau)$ with rotations for angles ranging from 0 to $\pi/2$, and finally follows the opposite segment of $\widehat{\Psi}$ from $t = \tau$ to $t = 1$. The matrices along this arc never have 1 as an eigenvalue, so there are no crossings, and letting $\tau \to 0$ shows that the arc is also homotopic to $\widehat{\Psi}$ with fixed end points, as claimed.

The direct sum property gives an easy extension of the first claim removing the condition $n = 1$; it implies namely that for any $n \in \mathbb{N}$, the arc

$$\Psi_0 : [0,1] \to \mathrm{Sp}(2n) : t \mapsto \begin{pmatrix} e^t & 0 & \cdots & 0 & 0 \\ 0 & e^{-t} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & e^t & 0 \\ 0 & 0 & \cdots & 0 & e^{-t} \end{pmatrix}$$

and corresponding asymptotic operator $\mathbf{A}_0$ satisfy $\mu_{\mathrm{RS}}(\Psi_0) = 0 = \mu_{\mathrm{CZ}}(\mathbf{A}_0) = \mu_{\mathrm{RS}}(\mathbf{A}_0)$.

Now fix $\mathbf{A}_0$ and Ψ_0 as above, and suppose $\mathbf{A}$ is an arbitrary trivialized asymptotic operator, with $\Psi : [0,1] \to \mathrm{Sp}(2n)$ as the corresponding symplectic arc starting at $\Psi(0) = \mathbb{1}$. Choose a smooth path of asymptotic operators $\{\mathbf{A}_s\}_{s\in[0,1]}$ from $\mathbf{A}_0$ to $\mathbf{A}_1 := \mathbf{A}$, and perturb it if necessary so that all crossings are regular. This path gives rise to a homotopy of symplectic arcs

$$\{\Psi_s : [0,1] \to \mathrm{Sp}(2n)\}_{s\in[0,1]}$$

from Ψ_0 to $\Psi_1 := \Psi$, all satisfying $\Psi_s(0) = \mathbb{1}$, and fixing $t = 1$ then gives rise to another symplectic arc

$$\alpha : [0,1] \to \mathrm{Sp}(2n) : s \mapsto \Psi_s(1).$$

The homotopy invariance of (3.35) with fixed end points now gives

$$\mu_{\mathrm{RS}}(\Psi) = \mu_{\mathrm{RS}}(\Psi_0) + \mu_{\mathrm{RS}}(\alpha) = \mu_{\mathrm{RS}}(\alpha),$$

so $\mu_{\mathrm{RS}}(\Psi)$ is a signed count of crossings for the symplectic arc $\alpha(s)$, with the final crossing (if any) at $s = 1$ weighted by $1/2$. But the 1-eigenspace of $\alpha(s)$ is naturally isomorphic to the kernel of $\mathbf{A}_s$, so crossings of α are equivalent to crossings of the path of asymptotic operators $s \mapsto \mathbf{A}_s$, and after comparing the definitions of the crossing forms for asymptotic operators and symplectic arcs, one finds

$$\mu_{\mathrm{RS}}(\alpha) = -\mu^{\mathrm{spec}}(\mathbf{A}_0, \mathbf{A}_1) = -\mu_{\mathrm{RS}}(\mathbf{A}_0) + \mu_{\mathrm{RS}}(\mathbf{A}) = \mu_{\mathrm{RS}}(\mathbf{A}).$$

□

REMARK 3.12.25. The relationship between spectral flow and other definitions of the Conlex-Zehnder index is treated at length in [RS95], but unlike what we have done in this section, that paper declines to give any definition for $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+)$ outside the case where both $\mathbf{A}_+$ and $\mathbf{A}_-$ are nondegenerate. Remark 5.31 in [RS95] states a reason for this that will be relevant in the next chapter: the most important application of spectral flow for asymptotic operators is to compute the Fredholm indices of certain linear Cauchy-Riemann type operators on bundles over punctured domains, but those operators fail to be Fredholm when the asymptotic operators are degenerate. In my personal opinion, this fact also makes the Robbin-Salamon index itself somewhat less useful, and it is the main reason why we will instead use the perturbed indices $\mu_{\mathrm{CZ}}^{\tau,\pm}(\gamma)$ from §3.12.2 in the rest of this book.

CHAPTER 4

Fredholm theory with cylindrical ends

Contents

In this chapter we will study the class of linear Cauchy-Riemann type operators that arise by linearizing the nonlinear equation for moduli spaces in SFT. We saw in the previous chapter that linearizing certain PDEs over noncompact domains naturally leads one to consider a class of symmetric *asymptotic operators* (e.g. the Hessian of a Morse function at its critical points), which have trivial kernel if and only if a nondegeneracy (i.e. Morse) condition is satisfied. Our main goal in this chapter is to show that the linear Cauchy-Riemann type operators in SFT are Fredholm if their asymptotic operators are nondegenerate.

4.1. Cauchy-Riemann operators with punctures

The setup throughout this chapter will be as follows.

Assume (Σ, j) is a closed connected Riemann surface of genus $g \geqslant 0$, $\Gamma \subset \Sigma$ is a finite set partitioned into two subsets

$$\Gamma = \Gamma^+ \cup \Gamma^-,$$

and $\dot{\Sigma} := \Sigma \backslash \Gamma$ denotes the resulting punctured Riemann surface. We shall fix a choice of **holomorphic cylindrical coordinate** near each puncture $z \in \Gamma^\pm$, meaning the following. Given $R \geqslant 0$, let $(Z^R_\pm, i)$ denote the half-cylinders

$$Z^R_+ := [R, \infty) \times S^1, \qquad Z^R_- := (-\infty, -R] \times S^1, \qquad Z_\pm := Z^0_\pm,$$

with complex structure $i\partial_s = \partial_t$, $i\partial_t = -\partial_s$ in coordinates $(s,t) \in \mathbb{R} \times S^1$. The standard half-cylinders $Z_\pm$ are each biholomorphically equivalent to the punctured

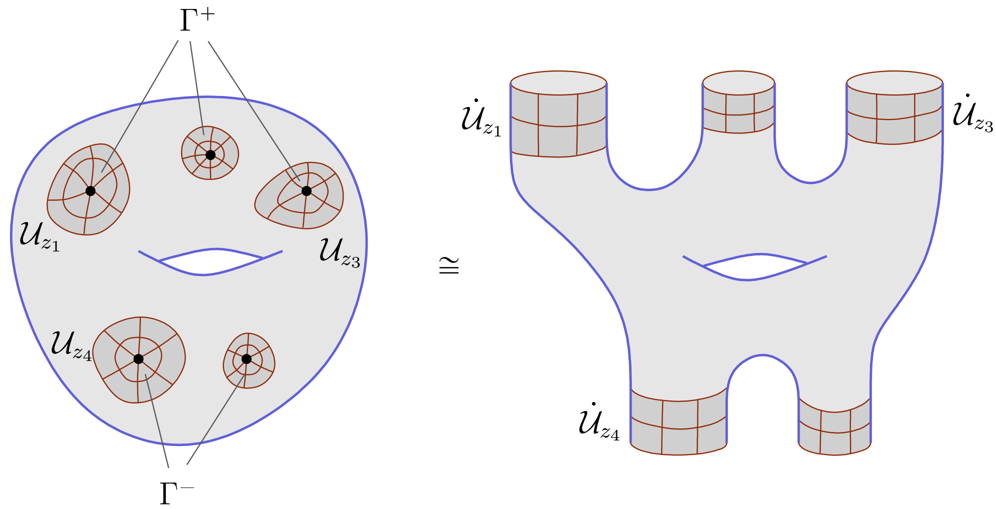

FIGURE 4.1. A Riemann surface with genus 1 and five punctures, depicted at the right as three positive and two negative cylindrical ends.

disk $\dot{\mathbb{D}} := \mathbb{D}\backslash\{0\}$ via the maps

$$\psi_\pm : Z_\pm \to \dot{\mathbb{D}} : (s,t) \mapsto e^{\mp 2\pi(s+it)}.$$

For $z \in \Gamma^\pm$, we choose a closed neighborhood $\mathcal{U}_z \subset \Sigma$ of z with a biholomorphic map

$$\varphi_z : (\dot{\mathcal{U}}_z, j) \to (Z_\pm, i),$$

where $\dot{\mathcal{U}}_z := \mathcal{U}_z\backslash\{z\}$, such that $\psi_\pm \circ \varphi_z : \dot{\mathcal{U}}_z \to \dot{\mathbb{D}}$ extends holomorphically to $\mathcal{U}_z \to \mathbb{D}$ with $z \mapsto 0$. One can always find such coordinates by choosing holomorphic coordinates near z. We can thus view the punctured neighborhoods $\dot{\mathcal{U}}_z \subset \dot{\Sigma}$ as **cylindrical ends** $Z_\pm$; see Figure 4.1.

Suppose (E, J) is a complex vector bundle of rank n over $(\dot{\Sigma}, j)$. An **asymptotically Hermitian structure** on (E, J) is a choice of Hermitian vector bundles (E_z, J_z, ω_z) of rank n over S^1 associated to each puncture $z \in \Gamma^\pm$, together with choices of complex bundle isomorphisms

$$E|_{\dot{\mathcal{U}}_z} \to \mathrm{pr}_2^* E_z$$

covering $\varphi_z : \dot{\mathcal{U}}_z \to Z_\pm$, where $\mathrm{pr}_2 : Z_\pm \to S^1$ denotes the natural projection to the S^1 factor. This isomorphism induces from any unitary trivialization τ of (E_z, J_z, ω_z) a trivialization

$$\tau : E|_{\dot{\mathcal{U}}_z} \to Z_\pm \times \mathbb{R}^{2n} \tag{4.1}$$

identifying J with $J_0 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$ over the cylindrical end. We will call this trivialization of E over $\dot{\mathcal{U}}_z$ an **asymptotic trivialization** near z. The bundle (E_z, J_z, ω_z) will be referred to as the **asymptotic bundle** associated to (E, J) near z.

Fixing asymptotic trivializations near every puncture, we can now define Sobolev spaces of sections of E by

$$W^{k,p}(E) := \left\{ \eta \in W^{k,p}_{\mathrm{loc}}(E) \;\middle|\; \eta_z \in W^{k,p}(\mathring{Z}_\pm, \mathbb{R}^{2n}) \text{ for every } z \in \Gamma^\pm \right\},$$

where $\eta_z : Z_\pm \to \mathbb{R}^{2n}$ denotes the expression of $\eta|_{\dot{\mathcal{U}}_z}$ in terms of the asymptotic trivialization, and we use the standard area form $ds \wedge dt$ on $Z_\pm$ in defining the norm on $W^{k,p}(\mathring{Z}_\pm, \mathbb{R}^{2n})$. Since S^1 is compact, the definition of this space does not depend on the choice of asymptotic trivialization, and moreover, one can pick a finite collection of charts and local trivializations covering $\dot{\Sigma}$ away from the punctures, supplemented by an asymptotic trivialization near each puncture, to define a norm on $W^{k,p}(E)$ that is (up to equivalence) independent of choices and makes $W^{k,p}(E)$ a Banach space. (For details on the construction of Sobolev norms for spaces of sections of vector bundles, see Appendices A.4 and A.5.) One must still be a bit careful with the noncompact ends, however:

EXERCISE 4.1.1. Convince yourself that different choices of asymptotically Hermitian structure on $E \to \dot{\Sigma}$ can give rise to *inequivalent* definitions of the space $W^{k,p}(E)$.

Any linear Cauchy-Riemann type operator on E has as its target the complex vector bundle

$$F := \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, E),$$

so sections of F are the same thing as E-valued $(0,1)$-forms. An asymptotic trivialization τ as in (4.1) then also induces a trivialization

$$F|_{\dot{\mathcal{U}}_z} \to Z_\pm \times \mathbb{R}^{2n} : \lambda \mapsto \tau(\lambda(\partial_s)),$$

where ∂_s is the coordinate vector field on $\dot{\mathcal{U}}_z$ arising from its identification with $Z_\pm$. This trivialization yields a corresponding definition for the Sobolev spaces $W^{k,p}(F)$, which depend on the asymptotically Hermitian structure of E but not on the choices of asymptotic trivializations. Having made these choices, a Cauchy-Riemann type operator $\mathbf{D} : \Gamma(E) \to \Gamma(F)$ always appears over $\dot{\mathcal{U}}_z$ as a linear map on $C^\infty(Z_\pm, \mathbb{R}^{2n})$ of the form

$$\mathbf{D}\eta(s,t) = \bar{\partial}\eta(s,t) + S(s,t)\eta(s,t), \tag{4.2}$$

where $\bar{\partial} := \partial_s + J_0\partial_t$ and $S \in C^\infty(Z_\pm, \mathrm{End}(\mathbb{R}^{2n}))$.

Since it is occasionally useful for technical reasons, we will in this chapter permit the bundle $E \to \dot{\Sigma}$ to be of class C^{m+1} for $m < \infty$, meaning it can be covered by local trivializations such that all transition maps are of class C^{m+1}, but possibly not smooth.[1] On such a bundle, the spaces $C^k(E)$ and $W^{k,p}(E)$ are well defined for each $k \leqslant m+1$ due to the continuous product pairings $C^{m+1} \times C^k \to C^k$ and $C^{m+1} \times W^{k,p} \to W^{k,p}$.

[1] This situation arises if one considers J-holomorphic curves $u : (\Sigma, j) \to (M, J)$ with respect to an almost complex structure J that is of class C^{m+1} but not smooth. According to Theorem 2.4.10 and the Sobolev embedding theorem, u is then a C^{m+1}-smooth map, so the pullback bundle $u^*TM \to \dot{\Sigma}$ is of class C^{m+1}, and since a derivative of J appears in the formula for the linearized operator $\mathbf{D}_u$, the latter is a Cauchy-Riemann type operator of class C^m.

DEFINITION 4.1.2. Suppose $E \to \dot{\Sigma}$ is of class C^{m+1} for some $m \in \{0, 1, 2, \dots, \infty\}$. A **linear Cauchy-Riemann type operator of class** C^m on E is then a first-order differential operator $\mathbf{D} : C^{m+1}(E) \to C^m(F)$ that takes the form $\mathbf{D} = \bar{\partial} + S$ in local trivializations with zeroth-order terms S of class C^m.

EXERCISE 4.1.3. Check that if the zeroth-order term of a Cauchy-Riemann type operator is of class C^m in a given trivialization, then this remains true after transforming it by a transition map of class C^{m+1}, though it does not remain true in general if the transition map is only of class C^m.

DEFINITION 4.1.4. Suppose $E \to \dot{\Sigma}$ is an asymptotically Hermitian vector bundle of class C^{m+1} for some $m \in \{0, 1, 2, \dots, \infty\}$, $\mathbf{A}_z$ is an asymptotic operator on (E_z, J_z, ω_z) and $\mathbf{D}$ is a linear Cauchy-Riemann type operator of class C^m on E. We say that $\mathbf{D}$ is C^m**-asymptotic to** $\mathbf{A}_z$ at z if $\mathbf{D}$ appears in the form (4.2) with respect to an asymptotic trivialization near z, with

$$\|S - S_\infty\|_{C^k(Z^R_\pm)} \to 0 \quad \text{as} \quad R \to \infty$$

for all $k \leqslant m$, where $S_\infty(s, t) := S_\infty(t)$ is a C^m-smooth loop of symmetric matrices such that $\mathbf{A}_z$ appears in the corresponding unitary trivialization of (E_z, J_z, ω_z) as $-J_0 \partial_t - S_\infty$.

Recall that an asymptotic operator is called **nondegenerate** if 0 is not in its spectrum, which means it defines an isomorphism $H^1(S^1) \to L^2(S^1)$. We will sometimes omit the prefix "C^m-" in the term "C^m-asymptotic"; when this happens, the case $m = \infty$ is meant. The objective of this chapter is to prove the following:

THEOREM 4.1.5. *Suppose $m \in \mathbb{N} \cup \{\infty\}$, (E, J) is an asymptotically Hermitian vector bundle of class C^{m+1} over $(\dot{\Sigma}, j)$, $\mathbf{A}_z$ is a nondegenerate asymptotic operator on the associated asymptotic bundle (E_z, J_z, ω_z) for each $z \in \Gamma$, and $\mathbf{D}$ is a linear Cauchy-Riemann type operator of class C^m that is C^m-asymptotic to $\mathbf{A}_z$ at each puncture z. Then for every $k \in \{1, \dots, m + 1\}$ and $p \in (1, \infty)$,*

$$\mathbf{D} : W^{k,p}(E) \to W^{k-1,p}(F)$$

is Fredholm. Moreover, $\operatorname{ind} \mathbf{D}$ and $\ker \mathbf{D}$ are each independent of k and p, the latter being a space of C^m-smooth sections whose derivatives up to order m decay exponentially fast to 0 on the cylindrical ends.

REMARK 4.1.6. We assume $m \geqslant 1$ in Theorem 4.1.5 for safety's sake, but most steps in the proof will also work for $m = 0$, the only exception being the exponential decay estimate carried out in §4.6. Even without this, our proof that $\mathbf{D}$ is Fredholm remains valid for $m = 0$ if $p \geqslant 2$ (see Remark 4.7.4). In any case, the applications in this book will only require the case $m = \infty$.

REMARK 4.1.7. A further possible generalization would be to consider linear Cauchy-Riemann type operators of Sobolev class $W^{m,q}$, as is done in [**MS12**, Appendix C.1]. As observed in Remark 2.4.2, most of the regularity theory for Cauchy-Riemann type operators can be extended to this setting, thus producing a generalization of Theorem 4.1.5. We will not pursue this here since it would make several details more complicated, especially in the asymptotic analysis, and we do not have any applications in mind for it that cannot also be handled by other means.

The index of $\mathbf{D}$ is determined by a generalization of the Riemann-Roch formula involving the Conley-Zehnder indices $\mu^\tau_{\mathrm{CZ}}(\mathbf{A}_z)$ that were introduced in the previous chapter. We will postpone serious discussion of the index formula until Chapter 5, but here is the statement:

THEOREM 4.1.8. *In the setting of Theorem 4.1.5,*

$$\operatorname{ind} \mathbf{D} = n\chi(\dot{\Sigma}) + 2c_1^\tau(E) + \sum_{z\in\Gamma^+} \mu^\tau_{CZ}(\mathbf{A}_z) - \sum_{z\in\Gamma^-} \mu^\tau_{CZ}(\mathbf{A}_z),$$

where τ is an arbitrary choice of asymptotic trivializations, $c_1^\tau(E) \in \mathbb{Z}$ is the relative first Chern number *of E with respect to τ, and the sum is independent of this choice.*

REMARK 4.1.9. The index formula reveals that the nondegeneracy condition on the asymptotic operators in Theorem 4.1.5 cannot be relaxed. Indeed, if $\mathbf{D}$ were Fredholm but had a degenerate asymptotic operator $\mathbf{A}_z$ at some puncture $z \in \Gamma$, then $\mathbf{D}$ could be perturbed to make $\mathbf{A}_z$ nondegenerate with at least two distinct possible values of its Conley-Zehnder index. This would produce two arbitrarily small perturbations of $\mathbf{D}$ that have different Fredholm indices according to Theorem 4.1.8, in which case $\mathbf{D}$ itself cannot be Fredholm. This is a marked contrast with the theory of linearized Cauchy-Riemann operators on *closed* Riemann surfaces: in the closed case, all Cauchy-Riemann type operators on the same bundle E are Fredholm and have the same index, because the difference between any two of them is a zeroth-order operator, which is compact due to the compactness of the inclusions $W^{k,p}(E) \hookrightarrow W^{k-1,p}(E)$. The difference in the punctured case is that since $\dot{\Sigma}$ is not compact, neither is the inclusion $W^{k,p}(E) \hookrightarrow W^{k-1,p}(E)$, hence zeroth-order terms can affect both the Fredholm property and the index.

Standing assumptions.
For the entirety of this chapter, $\dot{\Sigma} = \Sigma\backslash(\Gamma^+ \cup \Gamma^-)$ is a punctured Riemann surface as described above with fixed choices of holomorphic cylindrical coordinates near each puncture, $E \to \dot{\Sigma}$ is an asymptotically Hermitian vector bundle of complex rank $n \in \mathbb{N}$ and of class C^{m+1} for some $m \in \{0, 1, 2, \ldots, \infty\}$, and $\mathbf{D}$ is a linear Cauchy-Riemann type operator of class C^m on E which is C^m-asymptotic at each puncture $z \in \Gamma$ to an asymptotic operator $\mathbf{A}_z$. We will not always need to assume that the $\mathbf{A}_z$ are nondegenerate or that $m > 0$, so these conditions will be specified whenever they are relevant. The Sobolev parameters k and p will always lie in the ranges $1 \leqslant k \leqslant m + 1$ and $1 < p < \infty$ unless otherwise indicated.

For subdomains $\Sigma_0 \subset \dot{\Sigma}$, we will sometimes denote the $W^{k,p}$-norm on sections of E restricted to Σ_0 by

$$\|\eta\|_{W^{k,p}(\Sigma_0)} := \|\eta\|_{W^{k,p}(E|_{\Sigma_0})},$$

and we will use the same notation for sections of other bundles such as $F = \overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, E)$ over this domain when there is no danger of confusion. The space

$$W_0^{k,p}(\Sigma_0) \subset W^{k,p}(E)$$

is defined in this case as the $W^{k,p}$-closure of the space of smooth sections of E with compact support in Σ_0.

4.2. A lemma on semi-Fredholm operators

The standard approach for proving that elliptic operators are Fredholm begins by proving that they are **semi-Fredholm**, meaning $\dim \ker \mathbf{D} < \infty$ and $\operatorname{im} \mathbf{D}$ is closed. We saw in §2.4 that all Cauchy-Riemann type operators satisfy a local estimate of the form $\|\eta\|_{W^{k,p}} \leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}} + c\|\eta\|_{W^{k-1,p}}$, and we will see later in this chapter that a global version of this estimate also holds if the asymptotic operators at all punctures are nondegenerate. Recalling that the inclusion $W^{k,p} \hookrightarrow W^{k-1,p}$ is compact for functions on a bounded domain, such estimates can be used to establish the hypotheses of the following abstract functional-analytic result.

LEMMA 4.2.1. *Suppose X, Y and Z are Banach spaces, $\mathbf{T} \in \mathscr{L}(X, Y)$, $\mathbf{K} \in \mathscr{L}(X, Z)$ is a compact operator, and there is a constant $c > 0$ such that for all $x \in X$,*

$$\|x\|_X \leqslant c\|\mathbf{T}x\|_Y + c\|\mathbf{K}x\|_Z. \tag{4.3}$$

Then $\ker \mathbf{T}$ *is finite dimensional and* $\operatorname{im} \mathbf{T}$ *is closed.*

PROOF. A vector space is finite dimensional if and only if the closed unit ball in that space is a compact set, so we begin by proving the latter holds for $\ker \mathbf{T}$. Suppose $x_k \in \ker \mathbf{T}$ is a bounded sequence. Then since $\mathbf{K}$ is a compact operator, $\mathbf{K}x_k$ has a convergent subsequence in Z, which is therefore Cauchy. But (4.3) then implies that the corresponding subsequence of x_k in X is also Cauchy, and thus converges.

Since we now know $\ker \mathbf{T}$ is finite dimensional, we also know there is a closed complement $V \subset X$ with $\ker \mathbf{T} \oplus V = X$. Then the restriction $\mathbf{T}|_V$ has the same image as $\mathbf{T}$, thus if $y \in \overline{\operatorname{im} \mathbf{T}}$, there is a sequence $x_k \in V$ such that $\mathbf{T}x_k \to y$. We claim that x_k is bounded. If not, then $\mathbf{T}(x_k/\|x_k\|_X) \to 0$ and $\mathbf{K}(x_k/\|x_k\|_X)$ has a convergent subsequence, so (4.3) implies that a subsequence of $x_k/\|x_k\|_X$ also converges to some $x_\infty \in V$ with $\|x_\infty\| = 1$ and $\mathbf{T}x_\infty = 0$, a contradiction since $\mathbf{T}|_V : V \to Y$ is injective. But now since x_k is bounded, $\mathbf{K}x_k$ also has a convergent subsequence and $\mathbf{T}x_k$ converges by assumption, thus (4.3) yields also a convergent subsequence of x_k, whose limit x satisfies $\mathbf{T}x = y$. This completes the proof that $\operatorname{im} \mathbf{T}$ is closed. □

4.3. Some global regularity estimates

The following lemma is an immediate consequence of the local elliptic regularity result of Theorem 2.4.1, after covering a compact subset with finitely many local holomorphic coordinate charts and trivializations.

LEMMA 4.3.1. *Suppose $\mathbf{D}$ is of class C^m with $0 \leqslant m \leqslant \infty$, $1 \leqslant k \leqslant m+1$, $1 < p < \infty$, and $\Sigma_0 \subset \Sigma_1 \subset \dot{\Sigma}$ are open subsets with compact closure such that $\bar{\Sigma}_0 \subset \Sigma_1$. Then there exists a constant $c > 0$ such that for every $\eta \in W^{k,p}(E)$,*

$$\|\eta\|_{W^{k,p}(\Sigma_0)} \leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(\Sigma_1)} + c\|\eta\|_{W^{k-1,p}(\Sigma_1)}.$$

□

If $\Gamma = \varnothing$, then Lemma 4.3.1 suffices already for proving that $\mathbf{D}$ is semi-Fredholm, as one can then set $\Sigma_0 = \Sigma_1 := \Sigma$, observe that the inclusion $W^{k,p}(\Sigma) \hookrightarrow W^{k-1,p}(\Sigma)$ is a compact operator, and plug the estimate into Lemma 4.2.1. The estimate is insufficient however if there are punctures, because one has to chop off the cylindrical ends of $\dot{\Sigma}$ in order to obtain a domain with compact closure. Our next task is therefore to prove a truly *global* estimate that pays attention to neighborhoods of the punctures. Recall that in §2.4.1, we proved that weak solutions of class $\eta \in L^p_{\mathrm{loc}}$ for a given $p \in (1,\infty)$ to a linear Cauchy-Riemann type equation $\mathbf{D}\eta = \xi$ with $\xi \in W^{m,p}_{\mathrm{loc}}$ are always of class $W^{m+1,p}_{\mathrm{loc}}$. This local statement does not imply $\eta \in W^{m+1,p}$ in general, since it says nothing about any decay conditions at infinity that would be needed to produce finite L^p-norms. That is the purpose of the next result:

LEMMA 4.3.2. *Suppose $\mathbf{D}$ is of class C^m with $0 \leqslant m \leqslant \infty$, $1 < p < \infty$ and $1 \leqslant k \leqslant m+1$. If $\eta \in L^p(E)$ is a weak solution to $\mathbf{D}\eta = \xi$ with $\xi \in W^{k-1,p}(F)$, then $\eta \in W^{k,p}(E)$.*

PROOF. By induction, it suffices to show that if $\eta \in W^{k-1,p}$ and $\mathbf{D}\eta = \xi \in W^{k-1,p}$ then $\eta \in W^{k,p}$. Theorem 2.4.1 implies that this is true locally, so the task is to bound the $W^{k,p}$-norm of η on the cylindrical ends. Pick an asymptotic trivialization and write $\mathbf{D}$ on one of the ends $Z_\pm \cong \dot{\mathcal{U}}_z$ as $\bar{\partial} + S(s,t)$. Let us assume for concreteness that the puncture is a positive one, and now consider the norms of η on the bounded sets

$$Z_N := (N, N+1) \times S^1 \quad \text{and} \quad Z'_N := (N-1, N+2) \times S^1.$$

Since Z_N has closure in Z'_N, Lemma 4.3.1 gives

$$\begin{aligned}
\|\eta\|_{W^{k,p}(Z_N)} &\leqslant c\|\bar{\partial}\eta\|_{W^{k-1,p}(Z'_N)} + c\|\eta\|_{W^{k-1,p}(Z'_N)} \\
&= c\|\xi - S\eta\|_{W^{k-1,p}(Z'_N)} + c\|\eta\|_{W^{k-1,p}(Z'_N)} \\
&\leqslant c\|\xi\|_{W^{k-1,p}(Z'_N)} + c'\|\eta\|_{W^{k-1,p}(Z'_N)},
\end{aligned}$$

where in the last line we've incorporated $\|S\|_{C^{k-1}(Z'_N)}$ into the constant $c' > 0$. An important detail here is that the constants in these estimates can be assumed independent of N: indeed, the C^{k-1}-norm of S on $[N-1, N+2] \times S^1$ is bounded uniformly in N since $S(s,t)$ is asymptotically C^{k-1}-convergent to some $S_\infty(t)$, and the constant that arises by applying Lemma 4.3.1 with $\mathbf{D} := \bar{\partial}$ does not care if the domain is shifted by a translation. We can therefore take the sum of this estimate for all $N \in \mathbb{N}$ and relabel the constants, producing

$$\|\eta\|_{W^{k,p}(\mathring{Z}^1_+)} \leqslant c\|\xi\|_{W^{k-1,p}(\mathring{Z}_+)} + c\|\eta\|_{W^{k-1,p}(\mathring{Z}_+)}. \tag{4.4}$$

□

COROLLARY 4.3.3. *If $\mathbf{D}$ is of class C^m with $0 \leqslant m \leqslant \infty$ and $1 < p < \infty$, every weak solution $\eta \in L^p(E)$ of $\mathbf{D}\eta = 0$ is in $\bigcap_{k \leqslant m+1} \bigcap_{p \leqslant q < \infty} W^{k,q}(E)$; in particular, η is of class C^m, and its derivatives up to order m decay to zero at infinity.*

PROOF. This is essentially a global version of Corollary 2.4.8, and is proved via a very similar argument. For simplicity we assume $m < \infty$, as the case $m = \infty$ will then follow. If $p > 2$, then the Sobolev embedding theorem (Theorem A.1.6 and its adaptation for bundles sketched in §A.5) gives continuous inclusions $W^{m+1,p}(E) \hookrightarrow C^m(E)$ and $W^{m+1,p}(E) \hookrightarrow W^{m,q}(E)$ for all $q \in [p,\infty]$. The latter can be fed back into Lemma 4.3.2 to conclude $\eta \in W^{m+1,q}(E)$ for every $q \in [p,\infty)$, and the derivatives up to order m decay at infinity since the constant $c > 0$ in the Sobolev inequality

$$\|\eta\|_{C^m(Z^R_\pm)} \leqslant c\|\eta\|_{W^{m+1,p}(Z^R_\pm)}$$

does not depend on R, while the finiteness of $\|\eta\|_{W^{m+1,p}(Z_\pm)}$ implies that the right hand side converges to 0 as $R \to \infty$.

If $p \leqslant 2$, then since $\eta \in W^{1,p}(E)$, the Sobolev embedding theorem gives $\eta \in L^q(E)$ for every $q \in [p, p^*)$ where $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{2}$, and Lemma 4.3.2 then gives $\eta \in W^{m+1,q}(E)$ for all q in this range. Since $p > 1$, $\frac{1}{p^*} < \frac{1}{2}$, thus some of the q in this interval satisfy $q > 2$, and one can then repeat the argument of the previous paragraph to establish the result for all $q \geqslant p$, as well as the C^m-decay. □

REMARK 4.3.4. Corollary 4.3.3 is valid without any nondegeneracy assumption on asymptotic operators, but it is also not as strong a result as one would like. It will imply that the kernel of $\mathbf{D} : W^{k,p}(E) \to W^{k-1,p}(F)$ is independent of k, but we do not yet have enough knowledge of the asymptotic decay of sections $\eta \in \ker \mathbf{D}$ to determine whether they are also in $L^q(E)$ for $1 < q < p$, and for this reason, it is not yet clear whether $\ker \mathbf{D}$ depends on p. (This problem did not arise in our earlier local results, e.g. in Corollary 2.4.8, because we were working on domains with finite measure in local coordinates.) The latter will be deduced in §4.6 from an exponential decay estimate that makes explicit use of the nondegeneracy assumption.

One can now supplement Lemma 4.3.1 with (4.4) to produce a global estimate of the form

$$\|\eta\|_{W^{k,p}(E)} \leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(E)} + c\|\eta\|_{W^{k-1,p}(E)}$$

for all $\eta \in W^{k,p}(E)$, but this is also not quite what we need. The trouble is that since $\dot{\Sigma}$ is generally noncompact, the inclusion $W^{k,p}(E) \hookrightarrow W^{k-1,p}(E)$ is not a compact operator. To prove the semi-Fredholm property, we will need to replace the $W^{k-1,p}$-norm of η in this estimate with the norm of its restriction to a compact subset of $\dot{\Sigma}$, and this will be where the nondegeneracy assumption becomes essential.

4.4. Translation-invariant operators on the cylinder

In this section, we establish a special case of Theorem 4.1.5 that serves as the asymptotic analogue of the fundamental elliptic estimates from Chapter 2. Beyond those local estimates, this is the main analytical ingredient that makes all Floer-type theories in symplectic geometry work.

THEOREM 4.4.1. *Suppose $\mathbf{A} = -J_0\partial_t - S(t)$ is a nondegenerate asymptotic operator on the trivial Hermitian vector bundle $S^1 \times \mathbb{R}^{2n} \to S^1$, with $S : S^1 \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ of class C^m, $0 \leqslant m \leqslant \infty$. Then the operator*

$$\partial_s - \mathbf{A} = \partial_s + J_0\partial_t + S(t) : W^{k,p}(\mathbb{R} \times S^1, \mathbb{R}^{2n}) \to W^{k-1,p}(\mathbb{R} \times S^1, \mathbb{R}^{2n})$$

is an isomorphism if $1 \leqslant k \leqslant m+1$ and $1 < p < \infty$.

REMARK 4.4.2. The same reasoning as in Remark 4.1.9 concludes via the index formula of the next chapter that the converse of Theorem 4.4.1 also holds: if $\mathbf{A}$ is degenerate, then $\partial_s - \mathbf{A} : W^{k,p}(\mathbb{R} \times S^1) \to W^{k-1,p}(\mathbb{R} \times S^1)$ is not an isomorphism, in fact it is not even Fredholm.

Thanks to Lemma 4.3.2, it suffices to prove the case $k = 1$ of Theorem 4.4.1, as the rest will then follow via regularity. A detailed general proof for $k = 1$ can be found in [**Sal99**, Lemma 2.4]. We give below a different proof for the case $k = 1$ and $p = 2$, using a trick suggested by Sam Lisi. The case $p \neq 2$ can be deduced from this in conjunction with the basic local L^p-estimate from Chapter 2 (namely Theorem 2.3.2).

The trick behind the proof below is to take the Fourier transform of both sides of the equation $(\partial_s - \mathbf{A})u = f$ with respect to *the $\mathbb{R}$-coordinate only*. Concretely, let $\mathscr{S}(\mathbb{R} \times S^1)$ denote the space of smooth functions $u : \mathbb{R} \times S^1 \to \mathbb{C}^N$ for some $N \in \mathbb{N}$ whose derivatives of all orders have rapid decay at infinity, meaning the function $(s,t) \mapsto |s|^k \partial^\alpha u(s,t)$ is bounded on $\mathbb{R} \times S^1$ for all $k \in \mathbb{N}$ and all multiindices α. A minor variation on the usual argument for the Fourier transform then shows that the complex-linear transformations $u \mapsto \mathscr{F}u = \hat{u}$ and $v \mapsto \mathscr{F}^* v = \check{v}$ defined by

$$\hat{u}(\sigma,t) := \int_{-\infty}^{\infty} u(s,t) e^{-2\pi i s\sigma}\, ds, \qquad \check{v}(s,t) := \int_{-\infty}^{\infty} v(\sigma,t) e^{2\pi i s\sigma}\, d\sigma$$

are bijections $\mathscr{S}(\mathbb{R} \times S^1) \to \mathscr{S}(\mathbb{R} \times S^1)$ and are inverse to each other.

PROPOSITION 4.4.3. *Let $\langle\ ,\ \rangle_{L^2}$ denote the standard complex L^2-product for functions $\mathbb{R} \times S^1 \to \mathbb{C}^N : (s,t) \mapsto u(s,t)$, defined in terms of the standard Hermitian inner product on $\mathbb{C}^N$ and the measure $ds\,dt$. The operator $\mathscr{F}$ then has the following properties:*

(1) $\langle \hat{u}, \hat{v} \rangle_{L^2} = \langle u, v \rangle_{L^2}$ for all $u, v \in \mathscr{S}(\mathbb{R} \times S^1)$;
(2) $\widehat{\partial_s u}(\sigma,t) = 2\pi i \sigma \hat{u}(\sigma,t)$ for all $u \in \mathscr{S}(\mathbb{R} \times S^1)$;
(3) $\widehat{\partial_t u}(\sigma,t) = \partial_t \hat{u}(\sigma,t)$ for all $u \in \mathscr{S}(\mathbb{R} \times S^1)$;
(4) For any continuous function $\Phi : S^1 \to \operatorname{End}_{\mathbb{C}}(\mathbb{C}^N)$ and every $u \in \mathscr{S}(\mathbb{R} \times S^1)$, $\widehat{\Phi u} = \Phi \hat{u}$, where we denote $(\Phi u)(s,t) := \Phi(t) u(s,t)$.

□

Since $\mathscr{S}(\mathbb{R} \times S^1)$ contains $C_0^\infty(\mathbb{R} \times S^1)$ and is thus dense in $L^2(\mathbb{R} \times S^1)$, the first property in Proposition 4.4.3 implies in particular that $\mathscr{F}$ and $\mathscr{F}^*$ extend uniquely to isometries on $L^2(\mathbb{R} \times S^1)$. Adding the second and third properties gives a useful new characterization of the Sobolev space $H^1(\mathbb{R} \times S^1) := W^{1,2}(\mathbb{R} \times S^1)$:

EXERCISE 4.4.4. Show that a function $u \in L^2(\mathbb{R} \times S^1)$ is in $H^1(\mathbb{R} \times S^1)$ if and only if its Fourier transform $\hat{u}$ with respect to the $\mathbb{R}$-factor has both of the following properties:

- The function $(\sigma,t) \mapsto |\sigma| \hat{u}(\sigma,t)$ is also in $L^2(\mathbb{R} \times S^1)$;
- The function $\hat{u}(\sigma,t)$ has a weak partial derivative $\partial_t \hat{u}$ in $L^2(\mathbb{R} \times S^1)$.

Show moreover that the usual $W^{1,2}$-norm is then equivalent to

$$\|u\|_{H^1} := \|\hat{u}\|_{L^2} + \big\| |\sigma| \cdot \hat{u} \big\|_{L^2} + \|\partial_t \hat{u}\|_{L^2},$$

and that the second and third properties in Proposition 4.4.3 also hold (in the sense of weak derivatives) for all $u \in H^1(\mathbb{R} \times S^1)$. *Hint: $C_0^\infty(\mathbb{R} \times S^1)$ is also dense in $H^1(\mathbb{R} \times S^1)$; see Theorem A.5.1.*

PROOF OF THEOREM 4.4.1 FOR $k = 1$ AND $p = 2$. Since $\mathbf{A} = -J_0\partial_t - S(t)$ is not generally a complex-linear operator, we start by complifying it, i.e. we consider the natural extension of $\partial_s + J_0\partial_t + S : H^1(\mathbb{R} \times S^1, \mathbb{R}^{2n}) \to L^2(\mathbb{R} \times S^1, \mathbb{R}^{2n})$ to a complex-linear operator

$$\partial_s - \mathbf{A} = \partial_s + J_0\partial_t + S : H^1(\mathbb{R} \times S^1, \mathbb{C}^{2n}) \to L^2(\mathbb{R} \times S^1, \mathbb{C}^{2n}).$$

Observe that $(\partial_s - \mathbf{A})\bar{u} = \overline{(\partial_s - \mathbf{A})u}$ for all $u \in H^1(\mathbb{R} \times S^1, \mathbb{C}^{2n})$, thus it will suffice to prove that this complexification is an isomorphism, as this will imply the same result for the underlying real-linear operator. With this in mind, all functions for the remainder of this proof will be assumed to take values in $\mathbb{C}^{2n}$.

Since $\mathbf{A} = -J_0\partial_t - S(t)$ only involves a derivative with respect to t and a (complexified) zeroth-order term, it commutes with the Fourier transform operator $\mathscr{F}$, so that applying $\mathscr{F}$ to both sides of $(\partial_s - \mathbf{A})u = f$ and applying Proposition 4.4.3 and Exercise 4.4.4 transforms it into the equation

$$(2\pi i\sigma + J_0\partial_t + S)\hat{u} = \hat{f} \quad \text{almost everywhere.} \tag{4.5}$$

We need to show that for every $\hat{f} \in L^2(\mathbb{R} \times S^1)$, this equation has an almost everywhere unique solution $\hat{u} : \mathbb{R} \times S^1 \to \mathbb{C}^{2n}$ such that the norms $\|\hat{u}\|_{L^2}$, $\big\||\sigma| \cdot \hat{u}\big\|_{L^2}$ and $\|\partial_t\hat{u}\|_{L^2}$ are all finite and satisfy bounds in terms of $\|\hat{f}\|_{L^2}$.

It will be convenient to abbreviate

$$\hat{u}_\sigma := \hat{u}(\sigma, \cdot) : S^1 \to \mathbb{C}^{2n}$$

for functions $\hat{u} : \mathbb{R} \times S^1 \to \mathbb{C}^{2n}$ and $\sigma \in \mathbb{R}$. The equation (4.5) then becomes

$$(2\pi i\sigma - \mathbf{A})\hat{u}_\sigma = (2\pi i\sigma + J_0\partial_t + S)\hat{u}_\sigma = \hat{f}_\sigma \tag{4.6}$$

for each individual $\sigma \in \mathbb{R}$. Note that for $\hat{f} \in L^2(\mathbb{R} \times S^1)$, Fubini's theorem implies $\hat{f}_\sigma \in L^2(S^1)$ for almost every $\sigma \in \mathbb{R}$. For these particular values of σ, (4.6) does have a unique solution $\hat{u}_\sigma \in H^1(S^1)$: indeed, $\mathbf{A}$ is nondegenerate by assumption, thus it has no imaginary eigenvalues, implying that the operator $(2\pi i\sigma - \mathbf{A}) : H^1(S^1) \to L^2(S^1)$ has a bounded inverse for every $\sigma \in \mathbb{R}$, which we shall denote by

$$R_\sigma = (2\pi i\sigma - \mathbf{A})^{-1} : L^2(S^1) \to H^1(S^1).$$

It follows that there exists an almost everywhere unique function $\hat{u} : \mathbb{R} \times S^1 \to \mathbb{C}^{2n}$ such that for almost every $\sigma \in \mathbb{R}$, $\hat{u}_\sigma = R_\sigma\hat{f}_\sigma \in H^1(S^1)$ satisfies (4.6). It is not immediately obvious whether this implies that $\hat{u}$ also satisfies (4.5), but before addessing this, let us check that $\hat{u}$ satisfies all the required bounds.

As preparation, observe first that since $\mathbf{A}$ is symmetric, for every $\lambda \in \mathbb{R}$ and $\eta \in H^1(S^1)$ we have

$$\begin{aligned}\|(i\lambda - \mathbf{A})\eta\|_{L^2}^2 = \langle (i\lambda - \mathbf{A})\eta, (i\lambda - \mathbf{A})\eta\rangle_{L^2} = \lambda^2\|\eta\|_{L^2}^2 + \|\mathbf{A}\eta\|_{L^2}^2 \\ - i\lambda\left(\langle \eta, \mathbf{A}\eta\rangle_{L^2} - \langle \mathbf{A}\eta, \eta\rangle_{L^2}\right) = \lambda^2\|\eta\|_{L^2}^2 + \|\mathbf{A}\eta\|_{L^2}^2,\end{aligned}$$

giving rise to two estimates,

$$\|(i\lambda - \mathbf{A})\eta\|_{L^2} \geqslant |\lambda| \cdot \|\eta\|_{L^2} \quad \text{and} \quad \|(i\lambda - \mathbf{A})\eta\|_{L^2} \geqslant \|\mathbf{A}\eta\|_{L^2},$$

valid for all $\eta \in H^1(S^1)$. The first of these is equivalent to

$$\|R_\sigma \eta\|_{L^2} \leqslant \frac{1}{2\pi|\sigma|}\|\eta\|_{L^2} \quad \text{for all} \quad \eta \in L^2(S^1), \tag{4.7}$$

and combining the second estimate with the inequality $\|\mathbf{A}\eta\|_{L^2} \geqslant c\|\eta\|_{H^1}$ arising from the fact that $\mathbf{A}$ is invertible, we obtain $\|(2\pi i\sigma - \mathbf{A})\eta\|_{L^2} \geqslant c\|\eta\|_{H^1}$, and thus (after renaming the constant),

$$\|R_\sigma \eta\|_{H^1} \leqslant c\|\eta\|_{L^2} \quad \text{for all} \quad \eta \in L^2(S^1), \tag{4.8}$$

where the constant $c > 0$ is independent of $\sigma \in \mathbb{R}$.

Feeding (4.8) into Fubini's theorem now yields

$$\begin{aligned}\int_{-\infty}^{\infty} \|\hat{u}_\sigma\|^2_{L^2(S^1)}\, d\sigma + \int_{-\infty}^{\infty} \|\partial_t \hat{u}_\sigma\|^2_{L^2(S^1)}\, d\sigma &= \int_{-\infty}^{\infty} \|\hat{u}_\sigma\|^2_{H^1(S^1)}\, d\sigma \\ &= \int_{-\infty}^{\infty} \|R_\sigma \hat{f}_\sigma\|^2_{H^1(S^1)}\, d\sigma \leqslant c^2 \int_{-\infty}^{\infty} \|\hat{f}_\sigma\|^2_{L^2(S^1)}\, d\sigma = c^2\|\hat{f}\|^2_{L^2(\mathbb{R}\times S^1)},\end{aligned}$$

where the first integral on the left hand side is simply $\|\hat{u}\|^2_{L^2(\mathbb{R}\times S^1)}$. The second integral on the left hand side tells us moreover that the function $(\sigma, t) \mapsto \partial_t \hat{u}_\sigma(t)$ on $\mathbb{R} \times S^1$ (defined for almost every σ) has L^2-norm bounded by $c\|\hat{f}\|_{L^2}$, thus it is locally integrable on $\mathbb{R} \times S^1$. It is now another straightforward exercise in Fubini's theorem to show that this function is in fact the weak partial derivative $\partial_t \hat{u}$, so that (4.5) then follows from the fact that (4.6) is satisfied for almost all σ. Finally, (4.7) implies

$$\begin{aligned}\big\||\sigma| \cdot \hat{u}\big\|^2_{L^2(\mathbb{R}\times S^1)} &= \int_{-\infty}^{\infty} |\sigma|^2 \cdot \|\hat{u}_\sigma\|^2_{L^2(S^1)}\, d\sigma = \int_{-\infty}^{\infty} |\sigma|^2 \cdot \|R_\sigma \hat{f}_\sigma\|^2_{L^2(S^1)}\, d\sigma \\ &\leqslant \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} \|\hat{f}_\sigma\|^2_{L^2(S^1)}\, d\sigma = \frac{1}{(2\pi)^2}\|\hat{f}\|^2_{L^2(\mathbb{R}\times S^1)},\end{aligned}$$

which completes the proof that $f \mapsto u$ is a bounded linear map $L^2(\mathbb{R} \times S^1) \to H^1(\mathbb{R} \times S^1)$. ☐

4.5. Proof of the semi-Fredholm property

The following consequence of Theorem 4.4.1 is more obviously an asymptotic variant of the fundamental elliptic estimate from Chapter 2. Its key feature for our purposes is that, in contrast e.g. to Lemma 4.3.1, it does not mention the $W^{k-1,p}$-norm of η. Recall that $W^{k,p}_0(\mathring{Z}^R_\pm)$ denotes the $W^{k,p}$-closure of $C^\infty_0(\mathring{Z}^R_\pm)$, so such functions remain in $W^{k,p}$ if they are extended as zero to larger domains containing $\mathring{Z}^R_\pm$. Note that functions of class $W^{k,p}_0$ on $\mathring{Z}^R_\pm$ need not vanish near infinity—in fact, C^∞_0 is dense in $W^{k,p}(\mathbb{R} \times S^1)$, see Theorem A.5.1.

LEMMA 4.5.1. *Assume* $\mathbf{D}$ *is of class* C^m *with* $0 \leqslant m \leqslant \infty$, $1 \leqslant k \leqslant m+1$, $1 < p < \infty$, *and* $z \in \Gamma^\pm$ *is a puncture such that the asymptotic operator* $\mathbf{A}_z$ *is nondegenerate. Then in holomorphic cylindrical coordinates on* $Z_\pm^R \subset \dot{\mathcal{U}}_z$ *for every* $R \geqslant 0$ *sufficiently large, there exists a constant* $c > 0$ *such that*

$$\|\eta\|_{W^{k,p}(\mathring{Z}_\pm^R)} \leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(\mathring{Z}_\pm^R)} \qquad \text{for all} \qquad \eta \in W_0^{k,p}(\mathring{Z}_\pm^R).$$

PROOF. Write $\mathbf{D} = \partial_s + J_0\partial_t + S(s,t)$ and $\mathbf{D}_0 = \partial_s + J_0\partial_t + S_\infty(t)$ in an asymptotic trivialization on $\dot{\mathcal{U}}_z = Z_\pm$, where the nondegenerate asymptotic operator is $\mathbf{A} = -J_0\partial_t - S_\infty(t)$ and we assume

$$\|S - S_\infty\|_{C^{k-1}(Z_\pm^R)} \to 0 \quad \text{as} \quad R \to \infty.$$

For $\eta \in W_0^{k,p}(\mathring{Z}_\pm^R)$, there is a canonical extension $\eta \in W^{k,p}(\mathbb{R} \times S^1)$ that equals zero outside $Z_\pm^R$, so Theorem 4.4.1 implies an estimate

$$\|\eta\|_{W^{k,p}(\mathring{Z}_\pm^R)} = \|\eta\|_{W^{k,p}(\mathbb{R}\times S^1)} \leqslant c\|\mathbf{D}_0\eta\|_{W^{k-1,p}(\mathbb{R}\times S^1)} = c\|\mathbf{D}_0\eta\|_{W^{k-1,p}(\mathring{Z}_\pm^R)}$$

for some constant $c > 0$. Rewriting this in terms of $\mathbf{D}$ gives

$$\begin{aligned}\|\eta\|_{W^{k,p}(\mathring{Z}_\pm^R)} &\leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(\mathring{Z}_\pm^R)} + c\|(S_\infty - S)\eta\|_{W^{k-1,p}(\mathring{Z}_\pm^R)} \\ &\leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(\mathring{Z}_\pm^R)} + c'\|S_\infty - S\|_{C^{k-1}(Z_\pm^R)} \cdot \|\eta\|_{W^{k,p}(\mathring{Z}_\pm^R)},\end{aligned}$$

where we've used the continuity of the product pairing $C^{k-1} \times W^{k-1,p} \to W^{k-1,p}$ and the inclusion $W^{k,p} \hookrightarrow W^{k-1,p}$. Importantly, the constant $c' > 0$ in this estimate does not depend on R, thus we are free to choose $R > 0$ large enough so that $\|S_\infty - S\|_{C^{k-1}(Z_\pm^R)} \leqslant \frac{1}{2c'}$, in which case $\|\eta\|_{W^{k,p}(\mathring{Z}_\pm^R)}$ can be pulled over to the left hand side, giving

$$\frac{1}{2}\|\eta\|_{W^{k,p}(\mathring{Z}_\pm^R)} \leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(\mathring{Z}_\pm^R)}.$$

□

We can now prove a global estimate suitable for feeding into Lemma 4.2.1. Let

$$\Sigma^R \subset \dot{\Sigma}$$

denote the truncated open subset obtained by deleting the ends $Z_\pm^R \subset \dot{\mathcal{U}}_z$ from $\dot{\Sigma}$ for all $z \in \Gamma$. For any given subset $\Sigma_1 \subset \Sigma$, we also define corresponding punctured and truncated subsets respectively by

$$\dot{\Sigma}_1 := \Sigma_1 \cap \dot{\Sigma}, \qquad \Sigma_1^R := \Sigma_1 \cap \Sigma^R,$$

so Σ_1^R has compact closure in $\dot{\Sigma}$ for each $R \geqslant 0$ (see Figure 4.2). On first reading, you may prefer to assume $\Sigma_0 = \Sigma_1 := \Sigma$ in the following lemma, as this is the case we will use for proving the semi-Fredholm property. We are stating it somewhat more generally for the sake of other applications.

LEMMA 4.5.2. *Assume* $\mathbf{D}$ *is of class* C^m, $1 \leqslant k \leqslant m+1$, $1 < p < \infty$, $\Sigma_0 \subset \Sigma_1 \subset \Sigma$ *are open subsets such that*

$$\bar{\Sigma}_0 \subset \Sigma_1, \qquad (\bar{\Sigma}_0 \backslash \Sigma_0) \cap \Gamma = \varnothing, \qquad \text{and} \qquad (\bar{\Sigma}_1 \backslash \Sigma_1) \cap \Gamma = \varnothing,$$

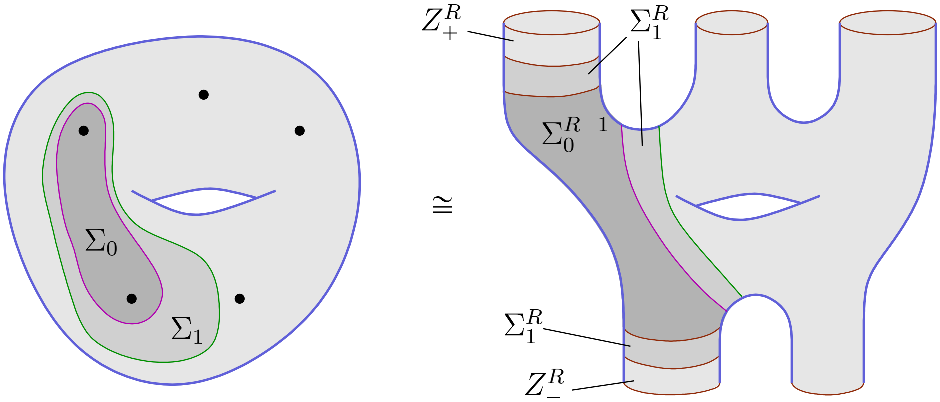

FIGURE 4.2. A punctured Riemann surface with subsets $\Sigma_0 \subset \bar{\Sigma}_0 \subset \Sigma_1 \subset \Sigma$ and their truncations $\Sigma_0^{R-1} \subset \overline{\Sigma_0^{R-1}} \subset \Sigma_1^R \subset \dot{\Sigma}$ as in Lemma 4.5.2.

and the asymptotic operators $\mathbf{A}_z$ *are nondegenerate for all* $z \in \Gamma \cap \Sigma_0$*. Then for any* $R > 0$ *sufficiently large, there exists a constant* $c > 0$ *such that*

$$\|\eta\|_{W^{k,p}(\dot{\Sigma}_0)} \leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(\dot{\Sigma}_1)} + c\|\eta\|_{W^{k-1,p}(\Sigma_1^R)}$$

for all $\eta \in W^{k,p}(\dot{\Sigma}_1)$*.*

PROOF. Fix $R > 1$ large enough so that the end $Z_\pm^{R-1} \subset \dot{\mathcal{U}}_z$ is disjoint from both $\bar{\Sigma}_0 \backslash \Sigma_0$ and $\bar{\Sigma}_1 \backslash \Sigma_1$ for every $z \in \Gamma^+ \cup \Gamma^-$, and so that Lemma 4.5.1 is valid on $Z_\pm^{R-1}$ whenever $z \in \Gamma \cup \Sigma_0$. The closure of Σ_0^{R-1} is then contained in Σ_1^R (see Figure 4.2), so we can choose another open set $\mathcal{V}$ with

$$\overline{\Sigma_0^{R-1}} \subset \mathcal{V} \subset \overline{\mathcal{V}} \subset \Sigma_1^R$$

and a smooth cutoff function $\beta \in C_0^\infty(\mathcal{V})$ such that $\beta \equiv 1$ on a neighborhood of $\overline{\Sigma_0^{R-1}}$. Letting

$$\dot{\mathcal{U}}_\Gamma^{R-1} \subset \dot{\Sigma}$$

denote the union of all the ends $\mathring{Z}_\pm^{R-1} \subset \dot{\mathcal{U}}_z$ for $z \in \Gamma \cap \Sigma_0$, we can now write any $\eta \in W^{k,p}(\dot{\Sigma}_1)$ as $\beta\eta + (1-\beta)\eta$, where $\beta\eta$ vanishes outside of $\mathcal{V}$ while $(1-\beta)\eta$ vanishes outside of $\dot{\mathcal{U}}_\Gamma^{R-1}$ and belongs to $W_0^{k,p}(\dot{\mathcal{U}}_\Gamma^{R-1})$. Lemma 4.3.1 then gives

$$\begin{aligned}\|\beta\eta\|_{W^{k,p}(\dot{\Sigma}_0)} = \|\beta\eta\|_{W^{k,p}(\mathcal{V})} &\leqslant c\|\mathbf{D}(\beta\eta)\|_{W^{k-1,p}(\Sigma_1^R)} + c\|\beta\eta\|_{W^{k-1,p}(\Sigma_1^R)} \\ &\leqslant c'\|\mathbf{D}\eta\|_{W^{k-1,p}(\Sigma_1^R)} + c'\|\eta\|_{W^{k-1,p}(\Sigma_1^R)},\end{aligned}$$

where the C^k-norm of β has been absorbed into the constant $c' > 0$. Similarly, Lemma 4.5.1 gives

$$\begin{aligned}\|(1-\beta)\eta\|_{W^{k,p}(\dot{\Sigma}_0)} = \|(1-\beta)\eta\|_{W^{k,p}(\dot{\mathcal{U}}_\Gamma^{R-1})} &\leqslant c\|\mathbf{D}[(1-\beta)\eta]\|_{W^{k-1,p}(\dot{\mathcal{U}}_\Gamma^{R-1})} \\ &\leqslant c'\|\mathbf{D}\eta\|_{W^{k-1,p}(\dot{\mathcal{U}}_\Gamma^{R-1})} + c'\|\eta\|_{W^{k-1,p}(\Sigma_1^R)},\end{aligned}$$

where the constant $c' > 0$ now contains information about the C^{k-1}-norms of $1-\beta$ and $\bar{\partial}\beta$ over $\dot{\mathcal{U}}_\Gamma^{R-1}$, with the important detail that the latter is only nonzero in the annuli $(R-1,R)\times S^1 \subset \dot{\mathcal{U}}_{R-1}$ and thus vanishes outside of Σ_1^R. Putting these estimates together and relabeling the constants, we obtain

$$\begin{aligned}\|\eta\|_{W^{k,p}(\dot\Sigma_0)} &= \|\beta\eta + (1-\beta)\eta\|_{W^{k,p}(\dot\Sigma_0)} \leqslant \|\beta\eta\|_{W^{k,p}(\dot\Sigma_0)} + \|(1-\beta)\|_{W^{k,p}(\dot\Sigma_0)}\\ &\leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(\dot\Sigma_1)} + c\|\eta\|_{W^{k-1,p}(\Sigma_1^R)}.\end{aligned}$$

□

COROLLARY 4.5.3. *Under the assumptions of Theorem 4.1.5, the operator* $\mathbf{D} : W^{k,p}(E) \to W^{k-1,p}(F)$ *has finite-dimensional kernel and closed image.*

PROOF. Choosing $\Sigma_0 = \Sigma_1 := \Sigma$ in Lemma 4.5.2 gives an estimate

$$\|\eta\|_{W^{k,p}(\dot\Sigma)} \leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(\dot\Sigma)} + c\|\eta\|_{W^{k-1,p}(\Sigma^R)}$$

for every $R \gg 1$ sufficiently large. The closure of the truncated surface Σ^R is a compact manifold with smooth boundary, thus the inclusion $W^{k,p}(\Sigma^R) \hookrightarrow W^{k-1,p}(\Sigma^R)$ compact, and so therefore is the map

$$W^{k,p}(\dot\Sigma) \to W^{k-1,p}(\Sigma^R) : \eta \mapsto \eta|_{\Sigma^R}.$$

We have thus established the hypotheses of Lemma 4.2.1. □

4.6. Exponential decay

We would now like to show that the kernel of $\mathbf{D} : W^{k,p}(E) \to W^{k-1,p}(F)$ is the *same* finite-dimensional vector space for every choice of the Sobolev parameters $k \in \{1,\dots,m+1\}$ and $p \in (1,\infty)$. We know already from Corollary 2.4.8 that this is true locally: if η is annihilated by $\mathbf{D}$ and belongs to $W^{k,p}(E)$ for any given $k \in \{1,\dots,m+1\}$ and $p \in (1,\infty)$, then $\eta \in W^{m+1,q}_{\text{loc}}$ for every $q \in (1,\infty)$. We also know from Corollary 4.3.3 that $\eta \in W^{m+1,q}(E)$ for every $q \in [p,\infty)$, but there is some uncertainty as to whether η must also decay fast enough at infinity to belong to $W^{m+1,q}(E)$ for $1 < q < p$. We shall prove in this section that if $m \geqslant 1$, then this is true at any end for which the asymptotic operator is nondegenerate. This will follow from the fact that nondegeneracy forces bounded solutions to decay exponentially fast.

4.6.1. Morse homology. To see why exponential behavior at infinity should be expected, let's consider first the analogy with Morse homology that was discussed in §3.2. The linearized operator for the gradient flow equation acts on sections of γ^*TM for a gradient flow line $\gamma : \mathbb{R} \to M$, and after choosing a global trivialization of γ^*TM, it takes the form[2]

$$\mathbf{D} : C^\infty(\mathbb{R},\mathbb{R}^n) \to C^\infty(\mathbb{R},\mathbb{R}^n) : \eta \mapsto \partial_s\eta - A(s)\eta$$

for some function $A : \mathbb{R} \to \mathbb{R}^{n\times n}$ that has a symmetric limit $A_+ := \lim_{s\to+\infty} A(s)$ corresponding to the Hessian at the critical point $x_+ := \lim_{s\to+\infty}\gamma(s)$. Let us choose a new trivialization in which A_+ is diagonal, and consider only $s \gg 1$ for which $A(s)$

[2]The linearized gradient flow operator was written as $\mathbf{D}\eta = \partial_s\eta + A(s)\eta$ when it first appeared in (3.3), but we have inserted a sign here in order to make the discussion more precisely analogous with Cauchy-Riemann type equations and asymptotic operators.

is an arbitrarily good approximation of A_+. In this regime, the linearized equation $\mathbf{D}\eta = 0$ becomes

$$\partial_s \eta \approx \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \eta.$$

In the "translation-invariant" case where $A(s)$ is a constant for large s, this equation is satisfied exactly, and one can then make some immediate pronouncements about the qualitative behavior of solutions as s becomes large: they are linear combinations of functions of the form $e^{\lambda s} v$ for eigenvectors $v \in \mathbb{R}^n$ of A_+ with corresponding eigenvalues $\lambda \in \mathbb{R}$. Some of these grow exponentially, some decay, and in general (if x_+ is a degenerate critical point) there may also be some that are constant, with $\lambda = 0 \in \sigma(A_+)$. If however we impose the condition that x_+ is nondegenerate and consider only the solutions that are bounded as $s \to \infty$, then every solution decays exponentially, i.e. it satisfies a bound of the form

$$|\eta(s)| \leqslant C e^{\lambda s}$$

for some constant $C > 0$, where the decay rate $\lambda < 0$ can be chosen to be the largest negative eigenvalue of A_+. Moreover, the normalized solution $\eta(s)/|\eta(s)|$ converges as $s \to \infty$ to an eigenvector of A_+.

In order to extend this discussion beyond the translation-invariant case, we can examine the functions

$$v(s) := \frac{\eta(s)}{|\eta(s)|} \in \mathbb{R}^n, \qquad \mu(s) := \langle v(s), A(s)v(s) \rangle \in \mathbb{R},$$

where $\langle \ , \ \rangle$ denotes the Euclidean inner product on $\mathbb{R}^n$. Our intuition from the previous paragraph suggests that as $s \to \infty$, $v(s)$ should converge to an eigenvector of A_+ while $\mu(s)$ converges to the corresponding eigenvalue. If we can prove the latter, then the desired exponential decay estimate will follow: indeed, the positive function $|\eta(s)|^2$ satisfies

$$\frac{d}{ds}|\eta(s)|^2 = 2\langle \eta(s), \dot{\eta}(s) \rangle = 2\langle \eta(s), A(s)\eta(s) \rangle = 2\mu(s)|\eta(s)|^2,$$

and solving this differential equation gives $|\eta(s)|^2 = e^{2\int_0^s \mu(\tau)\, d\tau} |\eta(0)|^2$, or equivalently,

$$|\eta(s)| = e^{\int_0^s \mu(\tau)\, d\tau} |\eta(0)|. \tag{4.9}$$

If $\lim_{\tau \to \infty} \mu(\tau)$ is an eigenvalue $\lambda \in \sigma(A_+)$, then λ will necessarily be negative unless the solution η is unbounded or x_+ is degenerate, and any $-\delta \in (\lambda, 0)$ can then be assumed to be greater than $\mu(\tau)$ for τ sufficiently large, implying that the integral in the exponent is at most $-\delta s$ plus a constant. This implies an estimate of the form $|\eta(s)| \leqslant C e^{-\delta s}$.

The crucial task is thus to prove that $\mu(s)$ converges to an eigenvalue as $s \to \infty$. Let us assume only what is necessary: the critical point x_+ need not be nondegenerate, but the spectrum $\sigma(A_+) \subset \mathbb{R}$ will in any case be a discrete set, and rather than focusing only on *bounded* solutions, suppose that $\eta : \mathbb{R} \to \mathbb{R}^n$ is any solution to $(\partial_s - A)\eta = 0$ satisfying an estimate of the form

$$|\eta(s)| \leqslant Ce^{\lambda_0 s} \tag{4.10}$$

for some $\lambda_0 \in \mathbb{R}$. Note that λ_0 in this discussion need not be negative; if we take $\lambda_0 = 0$, we are simply imposing the condition that η is bounded as $s \to \infty$, but we can also allow unbounded solutions with a controlled exponential growth rate by taking $\lambda_0 > 0$. Since the matrices $A(s)$ satisfy a uniform bound as $s \to \infty$, it is not hard to show that every solution satisfies this condition for some λ_0 sufficiently large, though the analog of this fact for Cauchy-Riemann type equations will be much less obvious—regardless, solutions that do not satisfy this condition for any $\lambda_0 \in \mathbb{R}$ cannot belong to any Banach space that we'd ever want to consider, so we shall simply ignore them.

LEMMA 4.6.1. *If η satisfies (4.10), then there exists a sequence $s_k \to \infty$ such that* $\liminf_{k\to\infty} \mu(s_k) \leqslant \lambda_0$.

PROOF. If not, then we have $\mu(s) \geqslant \lambda_0 + \epsilon$ for all $s \geqslant R$ if $R, \epsilon > 0$ are sufficiently large and small respectively. In this case $\int_0^s \mu(\tau)\, d\tau$ will inevitably become larger than $\lambda_0 s$ plus any given constant for s large, and (4.10) thus contradicts (4.9). □

Assuming (4.10), one useful way to proceed is now by estimating the distance of $\mu(s)$ from the spectrum $\sigma(A_+)$, which defines a continuous nonnegative function

$$d(s) := \operatorname{dist}(\mu(s), \sigma(A_+)) := \inf_{\lambda\in\sigma(A_+)} |\mu(s) - \lambda|.$$

It turns out that this function arises naturally when one examines the behavior of the derivative $\dot{\mu}(s)$. It will be useful to keep in mind that since $v(s) = \eta(s)/|\eta(s)|$ has constant norm, the derivative of $\langle v, v\rangle$ always vanishes, implying

$$\langle v, \dot{v}\rangle \equiv 0. \tag{4.11}$$

Moreover, since $\dot{\eta} = A\eta$ and $\partial_s|\eta|^2 = 2\mu|\eta|^2$, we have

$$A\eta = \partial_s\left(\sqrt{|\eta|^2}v\right) = \frac{2\mu|\eta|^2}{2|\eta|}v + |\eta|\dot{v} = |\eta|\left(\mu v + \dot{v}\right),$$

implying that $v(s)$ satisfies the differential equation

$$\dot{v}(s) = [A(s) - \mu(s)]\, v(s), \tag{4.12}$$

and also therefore

$$\langle v, (A - \mu)v\rangle \equiv 0 \tag{4.13}$$

in light of (4.11). In the translation-invariant case $A \equiv A_+$, appealing to the symmetry of A_+, we now find

$$\dot{\mu} = 2\langle \dot{v}, A_+ v\rangle = 2\langle (A_+ - \mu)v, A_+ v\rangle = 2\langle (A_+ - \mu)v, (A_+ - \mu)v\rangle = 2|(A_+ - \mu)v|^2,$$

where we have used (4.13) to permit the insertion of an extra μ on the right of the inner product. The function $|(A_+ - \mu)v|^2$ can only vanish when $\mu(s)$ is an eigenvalue of A_+, and it can otherwise be bounded from below in terms of the distance $d(s)$ from the spectrum of A_+. Indeed, we recall that for any self-adjoint linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with $c \in \mathbb{R}\backslash\sigma(T)$,

$$\|(T-c)^{-1}\| = \max_{\lambda \in \sigma(T)} \left| \frac{1}{\lambda - c} \right| = \frac{1}{\operatorname{dist}(c, \sigma(T))}, \tag{4.14}$$

as one easily sees by writing down $(T-c)^{-1}$ in an orthonormal basis for which T is diagonal. This yields the estimate $|(A_+ - \mu(s))v(s)|^2 \geqslant d(s)^2|v(s)|^2 = d(s)^2$, and thus

$$\dot{\mu}(s) \geqslant 2d(s)^2,$$

which will be enough information to conclude that the distance between $\mu(s)$ and the spectrum $\sigma(A_+)$ tends to 0 as $s \to \infty$.

Before making this conclusion more precise, let's remove the extra assumption that A is translation-invariant. Write

$$A(s) = A_+ + \Delta(s),$$

where $|\Delta(s)| \to 0$ as $s \to \infty$, and for good measure let's also assume $|\dot{\Delta}(s)| \to 0$ as $s \to \infty$. Since we can now expect the family of operators $A_+ - \mu(s)$ to play a special role, let us abbreviate it by

$$B(s) := A_+ - \mu(s),$$

and note that by (4.14), it satisfies $|B(s)v(s)| \geqslant d(s)$. The differential equation (4.12) then becomes $\dot{v} = (B + \Delta)v$, which we can use in the following to replace any appearance of $\dot{v}(s)$ with something that depends only on $v(s)$ and $B(s)$. We shall also abbreviate by $\epsilon(s)$ any term whose limit as $s \to \infty$ is 0, so for instance $\langle v, \Delta v\rangle$ and $\langle v, \dot{\Delta} v\rangle$ are both $\epsilon(s)$. To estimate $\dot{\mu}$, we can now write $\mu = \langle v, A_+ v\rangle + \langle v, \Delta v\rangle$ and take advantage of the symmetry of A_+ and the relation $\langle \dot{v}, v\rangle \equiv 0$ to write

$$\begin{aligned}
\dot{\mu} &= 2\langle \dot{v}, A_+ v\rangle + \langle \dot{v}, \Delta v\rangle + \langle v, \Delta \dot{v}\rangle + \langle v, \dot{\Delta} v\rangle \\
&= 2\langle \dot{v}, (A_+ - \mu)v\rangle + \langle \dot{v}, \Delta v\rangle + \langle v, \Delta \dot{v}\rangle + \epsilon(s) \\
&= 2\langle (B+\Delta)v, (B+\Delta)v\rangle + \langle (B+\Delta)v, \Delta v\rangle + \langle v, \Delta(B+\Delta)v\rangle + \epsilon(s) \\
&\geqslant 2|Bv|^2 - \epsilon(s)|Bv| - \epsilon(s) \\
&= |Bv| \cdot \big(2|Bv| - \epsilon(s)\big) - \epsilon(s),
\end{aligned}$$

where in the second to last line, we've appealed to the decay condition on Δ to combine into $\epsilon(s)|Bv|$ all terms that contain a single factor of Bv, while the additional $\epsilon(s)$ contains all terms in which Bv does not appear. Let us combine this result with the aforementioned consequence of (4.14) and summarize:

$$\dot{\mu}(s) \geqslant |B(s)v(s)| \cdot \big(2|B(s)v(s)| - \epsilon(s)\big) - \epsilon(s) \quad \text{where} \quad |B(s)v(s)| \geqslant d(s). \tag{4.15}$$

The discreteness of the spectrum $\sigma(A_+) \subset \mathbb{R}$ makes it easy to use (4.15) for ruling out any oscillatory behavior in the function $\mu(s)$. For instance:

LEMMA 4.6.2. *If there exists a sequence $s_k \to \infty$ for which a limit $\lambda := \lim_{k\to\infty} \mu(s_k)$ exists, then λ is an eigenvalue of A_+ and $\lim_{s\to\infty} \mu(s) = \lambda$.*

PROOF. By assumption, $\mu(s_k)$ is a Cauchy sequence, thus one can use the mean value theorem to find another sequence $\tau_k \to \infty$ with $\mu(\tau_k) \to \lambda$ and $\dot{\mu}(\tau_k) \to 0$. By (4.15), this is impossible unless $d(\tau_k) \to 0$, implying $\lambda \in \sigma(A_+)$. Assuming this, suppose $t_k \to 0$ is another sequence for which $\mu(t_k)$ does not converge to λ, implying that μ oscillates between a sequence of values converging to λ and another sequence of values that stay some fixed distance away from it. Since $\sigma(A_+)$ is discrete, we can extract from this a third sequence $\sigma_k \to \infty$ for which $\mu(\sigma_k)$ takes a fixed value $\lambda' \in \mathbb{R} \backslash \sigma(A_+)$ but $\dot{\mu}(\sigma_k) \leqslant 0$, and this contradicts (4.15) when k becomes large enough. □

LEMMA 4.6.3. $\liminf_{s \to \infty} \mu(s) \leqslant \lambda_0$.

PROOF. Lemma 4.6.1 implies that if this does not hold, then there is oscillatory behavior, i.e. there exist sequences $s_k, t_k \to \infty$ with $s_1 < t_1 < s_2 < t_2 < \ldots$ such that $\liminf_{k \to \infty} \mu(s_k) \leqslant \lambda_0$ but $\mu(t_k) \geqslant \lambda_0 + \epsilon$ for some $\epsilon > 0$ and all k. The same argument as in Lemma 4.6.2 then leads to a third sequence that violates (4.15) since $\sigma(A_+)$ is discrete. □

These two lemmas imply that unless $\lim_{s \to \infty} \mu(s)$ exists and is an eigenvalue, we must have $\lim_{s \to -\infty} \mu(s) = -\infty$.

LEMMA 4.6.4. *The function $\mu(s)$ does not diverge to $-\infty$ as $s \to \infty$.*

PROOF. Since $\sigma(A_+)$ is bounded, $\mu(s) \to -\infty$ would imply $d(s) \to \infty$, so that (4.15) would force $\dot{\mu}(s)$ to become positive for all s large. That is absurd. □

We summarize:

COROLLARY 4.6.5. *If η satisfies the bound (4.10), then the function $\mu(s)$ appearing in (4.9) converges as $s \to \infty$ to an eigenvalue $\lambda \leqslant \lambda_0$ of A_+.* □

4.6.2. Holomorphic curves. You should now find the following result plausible. The proof will require differentiating the zeroth-order term of the operator $\mathbf{D}$ near a puncture, so we have to require $\mathbf{D}$ to be of class at least C^1; this is the only step in our proof of the Fredholm property at which the case $m = 0$ must be excluded. Corollary 2.4.8 implies that weak solutions to $\mathbf{D}\eta = 0$ are then also classical solutions of class C^1, so we shall assume this in the statement.

THEOREM 4.6.6. *Assume $\mathbf{D}$ is of class C^1 and is C^1-asymptotic to the asymptotic operator $\mathbf{A}_z$ at a puncture $z \in \Gamma^{\pm}$, and suppose η is a continuously differentiable section of $E|_{\dot{\mathcal{U}}_z}$ satisfying $\mathbf{D}\eta = 0$ which, in holomorphic cylindrical coordinates and an asymptotic trivialization on $Z_{\pm} \cong \dot{\mathcal{U}}_z$, satisfies the estimate*

$$\|\eta(s, \cdot)\|_{L^2(S^1)} \leqslant Ce^{\lambda_0 s}$$

for some $C > 0$ and $\lambda_0 \in \mathbb{R}$. Then there exists an eigenvalue $\lambda \in \sigma(\mathbf{A}_z)$ with $\pm\lambda \leqslant \pm\lambda_0$ and a real-valued C^1-function $\mu(s)$ with

$$\lim_{s \to \pm\infty} \mu(s) = \lambda$$

such that the formula

$$\|\eta(s, \cdot)\|_{L^2(S^1)} = e^{\int_0^s \mu(\tau)\, d\tau} \|\eta(0, \cdot)\|_{L^2(S^1)}$$

holds for all s.

The proof below is adapted from an argument in [HWZ96], and it follows essentially the same outline as in the previous discussion of Morse homology. We will treat $s \mapsto \eta(s, \cdot)$ as a path in the Hilbert space $L^2(S^1)$ and $\mathbf{A}_z$ as an unbounded self-adjoint operator on that space, with dense domain $H^1(S^1)$ (cf. Exercise 3.5.23).

The fact that $L^2(S^1)$ is infinite dimensional and $\mathbf{A}_z$ is unbounded introduces two new technical complications: first, we need a generalization of (4.14) for relating the norm of the resolvent $(\mathbf{A}_z - \lambda)^{-1}$ to the distance of λ from $\sigma(\mathbf{A}_z)$. Such a generalization follows easily from the spectral theorem for unbounded self-adjoint operators (see [RS80, Theorem VIII.4]), according to which $\mathbf{A}_z$ can be identified unitarily with an operator of the form $\mathbf{T} : L^2(X, \mu) \supset \mathcal{D}(\mathbf{T}) \to L^2(X, \mu) : f \mapsto gf$ for some finite measure space (X, μ) and some function $g : X \to \mathbb{R}$. The domain $\mathcal{D}(\mathbf{T})$ is then the space of all $f \in L^2(X, \mu)$ for which gf also belongs to $L^2(X, \mu)$, and the spectrum of $\mathbf{T}$ is the essential range of g, i.e. the set of all $\lambda \in \mathbb{R}$ such that $g^{-1}((\lambda - \epsilon, \lambda + \epsilon))$ has positive measure for every $\epsilon > 0$. If $\lambda \notin \sigma(\mathbf{A}_z)$, then this presentation identifies $(\mathbf{A}_z - \lambda)^{-1}$ with multiplication on $L^2(X, \mu)$ by the function $\frac{1}{g-\lambda}$, whose L^∞-norm is $1/\operatorname{dist}(\lambda, \sigma(\mathbf{A}_z))$, thus giving

$$\|(\mathbf{A}_z - \lambda)^{-1}\| = \frac{1}{\operatorname{dist}(\lambda, \sigma(\mathbf{A}_z))}. \tag{4.16}$$

The second complication is that while $\sigma(\mathbf{A}_z)$ is still discrete, it is unbounded both above and below, whereas the boundedness of $\sigma(A_+)$ was used in Lemma 4.6.4 in order to prevent $\mu(s)$ from diverging to $-\infty$. However, the proof of that lemma would still have worked if we only knew that a sequence $\lambda_k \to -\infty$ could be found such that $\operatorname{dist}(\lambda_k, \sigma(\mathbf{A}_z))$ stays uniformly bounded away from zero. For asymptotic operators, Proposition 3.5.24 guarantees this.

PROOF OF THEOREM 4.6.6. To simplify the notation, let us assume the puncture $z \in \Gamma$ is positive, as the proof in the negative case would be completely analogous. We will also assume η is not identically zero, since there is otherwise nothing to prove. After fixing an asymptotic trivialization, we write $\mathbf{D} = \bar{\partial} + S(s,t)$ and $\mathbf{A} := \mathbf{A}_z = -J_0\partial_t - S_\infty(t)$, with $\|S - S_\infty\|_{C^1(Z_+^R)} \to 0$ as $R \to \infty$. For each $s \geqslant 0$, abbreviate

$$\eta_s := \eta(s, \cdot) \in C^1(S^1), \qquad \dot{\eta}_s := \partial_s \eta(s, \cdot) \in C^0(S^1)$$

and define a first-order differential operator on functions $S^1 \to \mathbb{R}^{2n}$ by

$$\mathbf{A}_s := -J_0\partial_t - S(s, \cdot),$$

so the equation $\mathbf{D}\eta = 0$ can now be written as

$$\dot{\eta}_s = \mathbf{A}_s \eta_s.$$

Writing $\|\cdot\|$ and $\langle\ ,\ \rangle$ for the L^2-norm and inner product respectively on S^1, let

$$v_s := \frac{\eta_s}{\|\eta_s\|} \in C^1(S^1), \qquad \mu(s) := \langle v_s, \mathbf{A}_s v_s \rangle \in \mathbb{R}.$$

Note that by the similarity principle, zeroes of η are isolated, thus $\|\eta_s\|$ is always positive and v_s is therefore well defined. Since η is of class C^1, we can differentiate the function $s \mapsto \langle \eta_s, \eta_s \rangle$ and find $\partial_s \|\eta_s\|^2 = 2\mu(s)\|\eta_s\|^2$, implying

$$\|\eta_s\| = e^{\int_0^s \mu(\tau)\, d\tau} \|\eta_0\|.$$

A quick computation based on $\dot{\eta}_s = \mathbf{A}_s \eta_s$ and $\partial_s \|\eta_s\|^2 = 2\mu \|\eta_s\|^2$ also gives the equation

$$\dot{v}_s := \partial_s v_s = (\mathbf{A}_s - \mu(s))\, v_s, \tag{4.17}$$

and since $\|v_s\| = 1$ is independent of s, we have

$$\langle v_s, \dot{v}_s \rangle = \langle v_s, (\mathbf{A}_s - \mu(s))\, v_s \rangle = 0. \tag{4.18}$$

Using the assumption $\|\eta_s\| \leqslant Ce^{\lambda_0 s}$, the proof of Lemma 4.6.1 can be repeated verbatim to show that $\liminf_{k\to\infty} \mu(s_k) \leqslant \lambda_0$ for some sequence $s_k \to \infty$.

There is a small technical complication in computing $\dot{\mu}(s)$: since $v(s,t) := v_s(t)$ is only a C^1-function and $\mathbf{A}_s$ is a differential operator, differentiating $\langle v_s, \mathbf{A}_s v_s \rangle$ with respect to s under the integral sign produces second derivatives of v that might not exist. However, this problem goes away if we define $\Delta_s := S(s,\cdot) - S_\infty$ and write

$$\mathbf{A}_s = \mathbf{A} + \Delta_s,$$

as the symmetry of $\mathbf{A}$ makes it possible to compute both partial derivatives of the function $(x,y) \mapsto \langle v_x, \mathbf{A} v_y \rangle$ by moving $\mathbf{A}$ to whichever side of the inner product is not being differentiated. In particular, these two partial derivatives are both continuous, so the function is of class C^1, and so therefore is $\mu(s)$, with

$$\dot{\mu}(s) = 2\langle \dot{v}_s, \mathbf{A}_s v \rangle + \langle \dot{v}_s, \Delta_s v_s \rangle + \langle v_s, \Delta_s \dot{v}_s \rangle + \langle v_s, \dot{\Delta}_s v_s \rangle,$$

where we abbreviate $\dot{\Delta}_s := \partial_s \Delta_s$. The family of self-adjoint operators

$$\mathbf{B}_s := \mathbf{A} - \mu(s)$$

on $L^2(S^1)$ with domain $H^1(S^1)$ satisfies

$$\|\mathbf{B}_s v_s\| \geqslant d(s) := \operatorname{dist}(\mu(s), \sigma(\mathbf{A})) \tag{4.19}$$

by (4.16), and since $\mathbf{A}_s - \mu(s) = \mathbf{B}_s + \Delta_s$, we can abbreviate $\epsilon(s)$ for any term that converges to 0 as $s \to \infty$ and exploit (4.17) and (4.18) to rewrite $\dot{\mu}(s)$ as

$$\begin{aligned}
\dot{\mu}(s) &= 2\langle \dot{v}_s, (\mathbf{A} - \mu(s)) v_s \rangle + \langle \dot{v}_s, \Delta_s v_s \rangle + \langle v_s, \Delta_s \dot{v}_s \rangle + \epsilon(s) \\
&= 2\langle (\mathbf{B}_s + \Delta_s) v_s, (\mathbf{B}_s + \Delta_s) v_s \rangle + \langle (\mathbf{B}_s + \Delta_s) v_s, \Delta_s v_s \rangle \\
&\quad + \langle v_s, \Delta_s (\mathbf{B}_s + \Delta_s) v_s \rangle + \epsilon(s) \\
&\geqslant 2\|\mathbf{B}_s v_s\|^2 - \epsilon(s)\|\mathbf{B}_s v_s\| - \epsilon(s) = \|\mathbf{B}_s v_s\| \cdot (2\|\mathbf{B}_s v_s\| - \epsilon(s)) - \epsilon(s).
\end{aligned} \tag{4.20}$$

With this relation and (4.19) in place, the proofs of Lemmas 4.6.2 and 4.6.3 can be repeated with trivial modifications to show that unless $\lim_{s\to\infty} \mu(s)$ exists and is an eigenvalue of $\mathbf{A}$, $\mu(s)$ must diverge to $-\infty$ as $s \to \infty$. But if the latter happens, then after using Proposition 3.5.24 to pick a sequence $\lambda_k \to -\infty$ with $\operatorname{dist}(\lambda_k, \sigma(\mathbf{A})) \geqslant \epsilon$ for some fixed $\epsilon > 0$, we can find a sequence $s_k \to \infty$ such that $\mu(s_k) = \lambda_k$ and $\dot{\mu}(s_k) \leqslant 0$ for all k, which similarly contradicts (4.19) and (4.20). □

If the asymptotic operator $\mathbf{A}_z$ is nondegenerate, then we can study bounded solutions to $\mathbf{D}\eta = 0$ by setting $\lambda_0 := 0$ and observe that the eigenvalue λ appearing in Theorem 4.6.6 will then be negative (for a positive puncture) or positive (for a negative puncture). This implies:

COROLLARY 4.6.7. *In the setting of Theorem 4.6.6, if $\mathbf{A}_z$ is nondegenerate and $\delta > 0$ is strictly less than any positive eigenvalue of $\mp\mathbf{A}_z$, then every globally bounded solution η to $\mathbf{D}\eta = 0$ satisfies an asymptotic decay estimate of the form*

$$\|\eta(s,\cdot)\|_{L^2} \leqslant Ce^{\mp\delta s}$$

in holomorphic cylindrical coordinates near $z \in \Gamma^\pm$, for some constant $C > 0$. □

Here is the most important consequence for the purposes of Fredholm theory.

COROLLARY 4.6.8. *In the setting of Theorem 4.1.5,*

$$\ker \mathbf{D} \subset \bigcap_{\ell \leqslant m+1} \bigcap_{1<q<\infty} W^{\ell,q}(E),$$

hence $\ker \mathbf{D}$ *is the same finite-dimensional vector space for all choices of k and p.*

PROOF. We can assume $m < \infty$ without loss of generality. Every $\eta \in W^{k,p}(E)$ annihilated by $\mathbf{D}$ is then locally of class $W^{m+1,q}$ for every $q \in (1,\infty)$ by Corollary 2.4.8, and Corollary 4.3.3 implies that it is also in $W^{m+1,q}(E)$ for all $q \in [p,\infty)$, so it is also in $C^m(E)$ and thus bounded. It will therefore suffice to prove that the restriction of η to each cylindrical end $Z_\pm \subset \dot{\mathcal{U}}_z$ is in $L^q(\mathring{Z}_\pm)$ for $q > 1$ arbitrarily close to 1; this information can then be fed into the hypotheses of Corollary 4.3.3 to conclude that $\eta \in W^{m+1,q}(\mathring{Z}_\pm)$ for every $q \in (1,\infty)$.

Let us consider for concreteness a positive end Z_+, and fix $q \in (1,2]$. Since S^1 has finite measure and $q \leqslant 2$, there is a constant $c > 0$ such that $\|f\|_{L^q(S^1)} \leqslant c\|f\|_{L^2(S^1)}$ for all measurable functions f on S^1. Choosing $\delta > 0$ as in Corollary 4.6.7, we then have

$$\begin{aligned}\|\eta\|^q_{L^q(\mathring{Z}_+)} &= \int_0^\infty \|\eta(s,\cdot)\|^q_{L^q(S^1)}\,ds \leqslant c^q \int_0^\infty \|\eta(s,\cdot)\|^q_{L^2(S^1)}\,ds \\ &\leqslant c^q C^q \int_0^\infty e^{-q\delta s}\,ds < \infty.\end{aligned}$$

□

REMARK 4.6.9. If one does not care too much about the sharpness of the exponential decay rate, then one gets a slightly quicker proof of Corollary 4.6.7 by differentiating the function $\alpha(s) := \frac{1}{2}\|\eta_s\|^2$ twice. Indeed, if $\mathbf{A}_z$ is nondegenerate, then for any $\delta > 0$ small enough so that $[-\delta,\delta] \cap \sigma(\mathbf{A}_z) = \emptyset$, one has $\|\mathbf{A}_s\eta_s\| \geqslant \delta\|\eta_s\|$ for all $s \gg 0$, which can be used to derive a differential inequality of the form

$$\ddot{\alpha}(s) \geqslant 4\delta^2\alpha(s) \qquad \text{for all } s \geqslant R, \tag{4.21}$$

valid for some large constant $R > 0$ that depends on δ and the operator $\mathbf{D}$, but not on the solution η. This inequality says that α must be "at least as convex" as any actual solution β to the corresponding differential equation $\ddot{\beta}(s) = 4\delta^2\beta(s)$, and choosing β to be the unique solution that matches α at $s = R$ and is bounded as $s \to \infty$, one deduces $\alpha \leqslant \beta$ for all $s \geqslant R$, implying $\alpha(s) \leqslant \alpha(R)e^{-2\delta(s-R)}$ and thus

$$\|\eta_s\| \leqslant e^{-\delta(s-R)}\|\eta_R\| \qquad \text{for all } s \geqslant R.$$

The details of this argument can be found in [Sch95, Lemma 3.1.23]. On the other hand, we will occasionally need to use the fact that the estimate of the decay rate in our version is generally sharper, e.g. at a positive puncture one can take any $\delta > 0$ for which the interval $[-\delta, 0]$ contains no eigenvalues, even if $[-\delta, \delta]$ does contain positive eigenvalues. The latter would kill the alternative argument, because using the second derivative has the effect of erasing the distinction between positive and negative eigenvalues.

4.6.3. Exponential weights. We can now say more precisely what is meant by the statement in Theorem 4.1.5 that elements of $\ker \mathbf{D}$ have exponentially decaying derivatives up to order m. This is best explained in the language of *exponentially weighted* Sobolev spaces, which will also become important in Chapter 8 when we study the corresponding nonlinear problem.

For $k \geqslant 0$, $1 \leqslant p \leqslant \infty$ and $\delta \in \mathbb{R}$, define

$$W^{k,p,\delta}(\mathring{Z}^R_\pm, \mathbb{R}^{2n}) := \left\{ f : \mathring{Z}^R_\pm \to \mathbb{R}^{2n} \;\middle|\; f = e^{\mp\delta s} g \text{ for some } g \in W^{k,p}(\mathring{Z}^R_\pm, \mathbb{R}^{2n}) \right\},$$

with the case $k = 0$ abbreviated by $L^{p,\delta} := W^{0,p,\delta}$. This is a Banach space with respect to the norm

$$\|f\|_{W^{k,p,\delta}(\mathring{Z}^R_\pm)} := \|e^{\pm\delta s} f\|_{W^{k,p}(\mathring{Z}^R_\pm)},$$

and in fact there is an obvious isometry $W^{k,p}(\mathring{Z}_\pm) \to W^{k,p,\delta}(\mathring{Z}_\pm) : f \mapsto e^{\mp\delta s} f$. We typically consider $W^{k,p,\delta}(\mathring{Z}_\pm)$ for $\delta > 0$, which forces functions in this space to decay exponentially at infinity. Concretely, if $p > 2$, then the inclusion $W^{m+1,p} \hookrightarrow C^m$ implies that functions $f \in W^{m+1,p,\delta}(\mathring{Z}^R_\pm)$ take the form $e^{\mp\delta s} g$ where g is of class C^m with a global C^m-bound. It follows that every derivative $\partial^\alpha f$ of order $|\alpha| \leqslant m$ is the product of $e^{\mp\delta s}$ with a globally bounded function, producing an estimate of the form

$$|\partial^\alpha f(s,t)| \leqslant C e^{\mp\delta s} \qquad \text{for all } |\alpha| \leqslant m.$$

The statement about decaying derivatives in Theorem 4.1.5 is therefore a consequence of the following:

PROPOSITION 4.6.10. *In the setting of Theorem 4.1.5, the restriction of any $\eta \in \ker \mathbf{D}$ to the end $Z_\pm \cong \dot{\mathcal{U}}_z$ near $z \in \Gamma^\pm$ belongs to $W^{\ell,q,\delta}(\mathring{Z}_\pm)$ for every $\ell \leqslant m+1$ and $1 < q < \infty$ and any $\delta > 0$ smaller than every positive eigenvalue of $\mp\mathbf{A}_z$.*

PROOF. We consider only the case $z \in \Gamma^+$, as the argument for negative punctures is analogous. The assumption is then that $-\delta$ lies in the open interval between 0 and the largest negative eigenvalue $\lambda \in \sigma(\mathbf{A}_z)$, and we can also choose a slightly larger number δ_1 with $\lambda < -\delta_1 < -\delta < 0$. Working in an asymptotic trivialization near z and writing $\mathbf{D} = \bar{\partial} + S(s,t)$, consider the function

$$\widehat{\eta}(s,t) := e^{\delta s}\eta(s,t),$$

which satisfies $\widehat{\mathbf{D}}\widehat{\eta} = 0$ for the Cauchy-Riemann type operator $\widehat{\mathbf{D}} := \mathbf{D} - \delta$, which is C^m-asymptotic to $\widehat{\mathbf{A}}_z := \mathbf{A}_z + \delta$ at z. The latter is nondegenerate since $-\delta \notin \sigma(\mathbf{A}_z)$, and since our previous regularity results imply that η is bounded, $\widehat{\eta}$ satisfies an estimate of the form $\|\widehat{\eta}(s,\cdot)\|_{L^2(S^1)} \leqslant Ce^{\delta s}$ for some $C > 0$. Since $\lambda + \delta$ is the largest eigenvalue of $\widehat{\mathbf{A}}_z$ less than or equal to δ and $\lambda + \delta < -(\delta_1 - \delta) < 0$, Corollary 4.6.7 now provides a decay estimate of the form

$$\|\widehat{\eta}(s,\cdot)\|_{L^2(S^1)} \leqslant C' e^{-(\delta_1 - \delta)s}$$

for some constant $C' > 0$. This implies via Fubini's theorem as in the proof of Corollary 4.6.8 that $\widehat{\eta}$ is of class L^2 on $Z_+ \cong \dot{\mathcal{U}}_z$, so we can now plug in the previous regularity estimates to conclude that $\widehat{\eta}$ is of class $W^{\ell,q}$ for every $\ell \leqslant m+1$ and $q \in (1,\infty)$, which is equivalent to η having finite $W^{\ell,q,\delta}$-norm. □

Working with exponential weights is quite easy in practice if one remembers the trick of replacing η with $\widehat{\eta}(s,t) = e^{\delta s}\eta(s,t)$ as used in the proof above. Another application is the result of the following exercise, which tells us that there is considerable freedom in the choice of topology for the space of solutions to the equation $\mathbf{D}\eta = 0$. We will see this phenomenon again in §6.7 for the nonlinear case, where it will imply that the geometrically "natural" topology on a moduli space of punctured J-holomorphic curves is equivalent to the more technical weighted Sobolev topologies that are needed for analyzing the moduli space's local structure.

EXERCISE 4.6.11. Suppose $\mathbf{D}$ and $\mathbf{D}_\nu = \mathbf{D} + S_\nu$ for $\nu \in \mathbb{N}$ are Cauchy-Riemann type operators of class C^m with $1 \leqslant m \leqslant \infty$, all of them C^m-asymptotic to nondegenerate asymptotic operators and satisfying $\lim_{\nu\to\infty} \|S_\nu\|_{C^m} = 0$. Show that for any $\delta > 0$ sufficiently small, if η_ν is a C^0_{loc}-convergent sequence of bounded solutions to $\mathbf{D}_\nu\eta_\nu = 0$, then it also converges in $W^{k,q}(E)$ and in the $W^{k,q,\delta}$-norm on the cylindrical ends for any $k \leqslant m+1$ and $q \in (1,\infty)$. *Hint: Replace $\eta_\nu(s,t)$ with $e^{\delta s}\eta_\nu(s,t)$ on each cylindrical end, then apply the global regularity estimate of Lemma 4.5.2. You might find it helpful to first convince yourself that the constants in that estimate can be assumed independent of ν.*

4.7. Formal adjoints and proof of the Fredholm property

In order to show that coker $\mathbf{D}$ is also finite dimensional, we will apply the above arguments to the formal adjoint of $\mathbf{D}$, an operator whose kernel is naturally isomorphic to the cokernel of $\mathbf{D}$. Let us choose Hermitian bundle metrics $\langle\ ,\ \rangle_E$ on E and $\langle\ ,\ \rangle_F$ on F, and fix an area form $d\,\mathrm{vol}$ on $\dot{\Sigma}$ that takes the form $d\,\mathrm{vol} = ds \wedge dt$ on the cylindrical ends. The **formal adjoint** of $\mathbf{D}$ is then defined as the unique first-order linear differential operator

$$\mathbf{D}^* : C^{m+1}(F) \to C^m(E)$$

that satisfies the relation

$$\langle \lambda, \mathbf{D}\eta \rangle_{L^2(F)} = \langle \mathbf{D}^*\lambda, \eta \rangle_{L^2(E)} \quad \text{for all} \quad \eta \in C^{m+1}_0(E),\ \lambda \in C^{m+1}_0(F),$$

where C_0^k indicates the space of C^k-smooth sections with compact support, and we use the real-valued L^2-pairings

$$\langle \eta, \xi \rangle_{L^2(E)} := \operatorname{Re} \int_{\dot{\Sigma}} \langle \eta, \xi \rangle_E \, d\,\mathrm{vol}, \qquad \text{for} \qquad \eta, \xi \in \Gamma(E),$$

$$\langle \alpha, \lambda \rangle_{L^2(F)} := \operatorname{Re} \int_{\dot{\Sigma}} \langle \alpha, \lambda \rangle_F \, d\,\mathrm{vol}, \qquad \text{for} \qquad \alpha, \lambda \in \Gamma(F).$$

The word "formal" refers to the fact that we are not viewing $\mathbf{D}^*$ as the adjoint of an unbounded operator on a Hilbert space (cf. [RS80]); that would be a stronger condition.

EXERCISE 4.7.1. Show that $\mathbf{D}^*$ is well defined and, for suitable choices of complex local trivializations of E and F and holomorphic coordinates on open subsets $\mathcal{U} \subset \dot{\Sigma}$, can be written locally as

$$\mathbf{D}^* = -\partial + A : C^{m+1}(\mathcal{U}, \mathbb{R}^{2n}) \to C^m(\mathcal{U}, \mathbb{R}^{2n})$$

for some $A \in C^m(\mathcal{U}, \operatorname{End}(\mathbb{R}^{2n}))$, where $\partial := \partial_s - J_0 \partial_t$.

The formula in the above exercise reveals that $\mathbf{D}^*$ is also an elliptic operator,[3] and thus has the same local properties as $\mathbf{D}$; indeed, $-\partial + A$ can be transformed into $\bar{\partial} + B$ for some zeroth-order term B if we conjugate it by a suitable complex-antilinear change of trivialization. In particular, our local estimates for $\mathbf{D}$ and their consequences, notably Lemma 4.3.1, are all equally valid for $\mathbf{D}^*$.

To obtain suitable asymptotic estimates for $\mathbf{D}^*$, let us fix asymptotic trivializations τ of E, use the corresponding trivializations of F over the ends as described in §4.1, and choose the bundle metrics such that both appear standard in these trivializations over the ends. We will say that the bundle metrics are **compatible with the asymptotically Hermitian structure** of E whenever they are chosen in this way outside of a compact subset of $\dot{\Sigma}$. We can then express $\mathbf{D}$ as $\bar{\partial} + S(s,t)$ on $\dot{\mathcal{U}}_z = Z_{\pm}$, and integrate by parts to obtain

$$\mathbf{D}^* = -\partial + S(s,t)^{\mathrm{T}}.$$

To identify this expression with a Cauchy-Riemann type operator, let

$$C := \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$$

denote the $\mathbb{R}$-linear transformation on $\mathbb{R}^{2n} = \mathbb{C}^n$ representing complex conjugation. Then since C anticommutes with J_0, we have

$$\begin{aligned}(C^{-1}\mathbf{D}^* C)\eta &= -C\partial_s(C\eta) + CJ_0\partial_t(C\eta) + CS(s,t)^{\mathrm{T}}C\eta \\ &= -\partial_s\eta - J_0\partial_t\eta + CS(s,t)^{\mathrm{T}}C\eta = -(\bar{\partial}\eta - CS(s,t)^{\mathrm{T}}C\eta) \\ &=: -(\bar{\partial} + \bar{S}(s,t))\eta,\end{aligned}$$

[3] Technically, this property of the formal adjoint is part of the definition of ellipticity: we call a differential operator elliptic whenever (1) it has the properties necessary for proving fundamental estimates using Fourier transforms as we did with $\bar{\partial}$ in §2.3, and (2) its formal adjoint also has these properties. The former requires the principal symbol of the operator to be everywhere injective, and the latter requires it to be surjective.

where we've defined $\bar{S}(s,t) := -CS(s,t)^{\mathrm{T}}C$. Now if the asymptotic operator $\mathbf{A}_z$ at $z \in \Gamma^{\pm}$ is written in the chosen trivialization as $\mathbf{A} := -J_0\partial_t - S_\infty(t)$, the asymptotic convergence of $S(s,t)$ implies that similarly

$$\|\bar{S} - \bar{S}_\infty\|_{C^k(Z^R_\pm)} \to 0 \quad \text{as} \quad R \to \infty$$

for all $k \leqslant m$, where

$$\bar{S}_\infty(t) := -CS_\infty(t)C.$$

This defines a trivialized asymptotic operator $\bar{\mathbf{A}} = -J_0\partial_t - \bar{S}_\infty(t)$ to which $-\mathbf{D}^*$ is (after a suitable change of trivialization) asymptotic at the puncture z; in particular, our proof of the global regularity result, Lemma 4.3.2, now also works for $\mathbf{D}^*$. Finally, notice that $\mathbf{A}$ and $-\bar{\mathbf{A}}$ are conjugate: indeed,

$$(C^{-1}\bar{\mathbf{A}}C)\eta = -CJ_0\partial_t(C\eta) + CCS_\infty(t)C(C\eta) = J_0\partial_t\eta + S_\infty(t)\eta = -\mathbf{A}\eta.$$

This implies that $\mathbf{A}$ is nondegenerate if and only if $\bar{\mathbf{A}}$ is; applying this assumption for all of the $\mathbf{A}_z$, the proofs of Lemma 4.5.1 and Lemma 4.5.2 now also go through for $\mathbf{D}^*$.

We've proved:

PROPOSITION 4.7.2. *Suppose $\mathbf{D}^*$ is defined with respect to Hermitian bundle metrics on E and $F = \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, E)$ that are compatible with the asymptotically Hermitian structure of E. If additionally all the asymptotic operators $\mathbf{A}_z$ are nondegenerate, then*

$$\mathbf{D}^* : W^{k,p}(F) \to W^{k-1,p}(E)$$

is semi-Fredholm. Moreover, if $\mathbf{D}$ is of class C^m with $1 \leqslant m \leqslant \infty$, then $\ker \mathbf{D}^*$ is contained in $W^{\ell,q}(F)$ for every $\ell \leqslant m+1$ and $q \in (1,\infty)$, and is thus independent of the choice of k and p. □

Since $\ker \mathbf{D}^*$ is now known to be finite dimensional, the next result completes the proof of the Fredholm property for $\mathbf{D}$ by showing that its image has finite codimension. It should be emphasized that both the statement and the proof of this result depend on the fact that $\ker \mathbf{D}^*$ is the same space for all choices of Sobolev parameters, so e.g. it is automatically a subspace of $W^{k-1,p}(F)$.

LEMMA 4.7.3. *If $\mathbf{D}$ is of class C^m with $1 \leqslant m \leqslant \infty$, all its asymptotic operators are nondegenerate, and $\mathbf{D}^*$ is defined under the same assumptions as in Prop. 4.7.2, then for $\mathbf{D} : W^{k,p}(E) \to W^{k-1,p}(E)$ with $1 \leqslant k \leqslant m+1$,*

$$W^{k-1,p}(F) = \operatorname{im} \mathbf{D} + \ker \mathbf{D}^*.$$

PROOF. Consider first the case $k = 1$. Since $\mathbf{D} : W^{1,p}(E) \to L^p(F)$ is semi-Fredholm, its image is closed, hence $\operatorname{im} \mathbf{D} + \ker \mathbf{D}^*$ is a closed subspace of $L^p(F)$. Recall that our standing assumptions include $1 < p < \infty$, so for a closed subspace $V \subset L^p(F)$ in general, one has $V \neq L^p(F)$ if and only if there exists a nontrivial bounded linear functional on $L^p(F)$ that annihilates V; this is conventionally considered a standard application of the Hahn-Banach theorem, though it can also be deduced from the uniform convexity of $L^p(F)$ without appealing to Hahn-Banach (see e.g. [**Wen26**, Exercise 14.9]). It follows via the Riesz representation theorem that if $\operatorname{im} \mathbf{D} + \ker \mathbf{D}^* \neq L^p(F)$, then there is a nontrivial element $\alpha \in (L^p(F))^* \cong L^q(F)$ for $\frac{1}{p} + \frac{1}{q} = 1$ such that

$$\langle \mathbf{D}\eta + \lambda, \alpha \rangle_{L^2(F)} = 0 \quad \text{for all} \quad \eta \in W^{1,p}(E), \ \lambda \in \ker \mathbf{D}^*. \tag{4.22}$$

Choosing $\lambda = 0$, this implies in particular

$$\langle \mathbf{D}\eta, \alpha \rangle_{L^2(F)} = 0 \quad \text{for all} \quad \eta \in W^{1,p}(E).$$

Since one can plug in arbitrary smooth compactly supported sections in trivialized neighborhoods for η, this means that α is a weak solution of class L^q to the formal adjoint equation $\mathbf{D}^*\alpha = 0$, so $\alpha \in \ker \mathbf{D}^*$. This contradicts (4.22) if we plug in $\eta = 0$ and $\lambda = \alpha$, thus completing the proof for $k = 1$.

For $k \geqslant 2$, suppose $\alpha \in W^{k-1,p}(F) \subset L^p(F)$ is given: then the case $k = 1$ provides elements $\eta \in W^{1,p}(E)$ and $\lambda \in \ker \mathbf{D}^*$ such that $\mathbf{D}\eta + \lambda = \alpha$. Since $\lambda \in W^{m+1,q}(F)$ for all $q \in (1, \infty)$, we have $\mathbf{D}\eta = \alpha - \lambda \in W^{k-1,p}(F)$ and thus by Lemma 4.3.2, $\eta \in W^{k,p}(E)$, completing the proof for all $k \leqslant m + 1$. $\square$

REMARK 4.7.4. If $\mathbf{D}$ is only of class C^0 but not C^1, then we do not have the exponential decay results from the previous section, but Lemma 4.7.3 still holds for $p \geqslant 2$ if $\ker \mathbf{D}^*$ is understood to be the kernel of the specific operator $\mathbf{D}^* : W^{1,q}(F) \to L^q(E)$ for $\frac{1}{p} + \frac{1}{q} = 1$. Indeed, Lemma 4.3.2 implies since $p \geqslant q$ that $\ker \mathbf{D}^*$ is then also a subspace of $W^{k-1,p}(F)$.

The proof of the Fredholm property for $\mathbf{D}$ is now complete, but in order to see that its index does not depend on k or p, we still need to see that this is true for $\dim \operatorname{coker} \mathbf{D}$. This follows from the corresponding fact about $\ker \mathbf{D}^*$, via a slight strengthening of Lemma 4.7.3:

PROPOSITION 4.7.5. *Under the same assumptions as in Lemma 4.7.3 for the operators* $\mathbf{D} : W^{k,p}(E) \to W^{k-1,p}(F)$ *and* $\mathbf{D}^* : W^{k,p}(F) \to W^{k-1,p}(E)$, *we have* $W^{k-1,p}(F) = \operatorname{im} \mathbf{D} \oplus \ker \mathbf{D}^*$ *and* $W^{k-1,p}(E) = \operatorname{im} \mathbf{D}^* \oplus \ker \mathbf{D}$. *In particular, the projections defined by these splittings give isomorphisms*

$$\operatorname{coker} \mathbf{D} \cong \ker \mathbf{D}^* \quad \text{and} \quad \operatorname{coker} \mathbf{D}^* \cong \ker \mathbf{D},$$

thus $\mathbf{D}^* : W^{k,p}(F) \to W^{k-1,p}(E)$ *is a Fredholm operator with*

$$\operatorname{ind} \mathbf{D}^* = -\operatorname{ind} \mathbf{D}.$$

PROOF. By Lemma 4.7.3, the first splitting follows if we can show that $\operatorname{im} \mathbf{D} \cap \ker \mathbf{D}^* = \{0\}$. Recall first (see §A.5) that the smooth functions with compact support form a dense subspace of $W^{k,p}(\dot{\Sigma})$ for every $k \geqslant 0$ and $p \in [1, \infty)$, so the definition of the formal adjoint implies via density and Hölder's inequality that if $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$,

$$\langle \lambda, \mathbf{D}\eta \rangle_{L^2(F)} = \langle \mathbf{D}^*\lambda, \eta \rangle_{L^2(E)} \quad \text{for all} \quad \eta \in W^{1,p}(E), \ \lambda \in W^{1,q}(F). \tag{4.23}$$

Now suppose $\lambda \in \operatorname{im} \mathbf{D} \cap \ker \mathbf{D}^*$ and write $\lambda = \mathbf{D}\eta$, assuming $\eta \in W^{k,p}(E)$. Our regularity and asymptotic results imply that since $\mathbf{D}^*\lambda = 0$, $\lambda \in W^{1,q}(F)$, where q can be chosen to satisfy $\frac{1}{p} + \frac{1}{q} = 1$. We can therefore apply (4.23) and obtain

$$\langle \lambda, \lambda \rangle_{L^2(F)} = \langle \lambda, \mathbf{D}\eta \rangle_{L^2(F)} = \langle \mathbf{D}^*\lambda, \eta \rangle_{L^2(E)} = 0,$$

hence $\lambda = 0$.

The proof that $W^{k-1,p}(E) = \operatorname{im} \mathbf{D}^* \oplus \ker \mathbf{D}$ is analogous. □

This result hints at the fact that $\mathbf{D}^*$ is—under some natural extra assumptions—globally equivalent to another Cauchy-Riemann type operator. To see this, let us impose a further constraint on the relation between the Hermitian bundle metrics $\langle\ ,\ \rangle_E$ and $\langle\ ,\ \rangle_F$. Note that since the area form $d\,\mathrm{vol}$ is necessarily j-invariant, it induces a Hermitian bundle metric on $T\dot{\Sigma}$, namely

$$\langle X, Y\rangle_\Sigma := d\,\mathrm{vol}(X, jY) + i\, d\,\mathrm{vol}(X, Y),$$

which matches the standard bundle metric in the trivializations over the ends defined via the cylindrical coordinates. This induces real-linear isomorphisms from $T\dot{\Sigma}$ to the complex-linear and -antilinear parts of the complexified cotangent bundle,

$$T\dot{\Sigma} \to \Lambda^{1,0}T^*\dot{\Sigma} : X \mapsto X^{1,0} := \langle X, \cdot\rangle_\Sigma,$$
$$T\dot{\Sigma} \to \Lambda^{0,1}T^*\dot{\Sigma} : X \mapsto X^{0,1} := \langle \cdot, X\rangle_\Sigma,$$

where the first isomorphism is complex antilinear and the second is complex linear. We use these to define Hermitian bundle metrics on $\Lambda^{1,0}T^*\dot{\Sigma}$ and $\Lambda^{0,1}T^*\dot{\Sigma}$ in terms of the metric on $T\dot{\Sigma}$; note that this is a straightforward definition for $\Lambda^{0,1}T^*\dot{\Sigma}$, but since the isomorphism to $\Lambda^{1,0}T^*\dot{\Sigma}$ is complex *antilinear*, we really mean

$$\langle X^{1,0}, Y^{1,0}\rangle_\Sigma := \langle Y, X\rangle_\Sigma \quad \text{for} \quad X, Y \in T\dot{\Sigma}.$$

Now observe that as a vector bundle with complex structure $\lambda \mapsto J \circ \lambda$, $F = \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, E)$ is naturally isomorphic to the complex tensor product

$$F \cong \Lambda^{0,1}T^*\Sigma \otimes E.$$

We can therefore make a natural choice for $\langle\ ,\ \rangle_F$ as the tensor product metric determined by $\langle\ ,\ \rangle_\Sigma$ and $\langle\ ,\ \rangle_E$. It is easy to check that this choice is compatible with the asymptotically Hermitian structure of E.

Next, we notice that the area form $d\,\mathrm{vol}$ also induces a natural complex bundle isomorphism

$$E \to \mathrm{Hom}_{\mathbb{C}}(T\dot{\Sigma}, F).$$

Indeed, the right hand side is canonically isomorphic to the complex tensor product

$$\mathrm{Hom}_{\mathbb{C}}(T\dot{\Sigma}, F) \cong \Lambda^{1,0}T^*\dot{\Sigma} \otimes F \cong \Lambda^{1,0}T^*\dot{\Sigma} \otimes \Lambda^{0,1}T^*\dot{\Sigma} \otimes E,$$

and $\Lambda^{1,0}T^*\dot{\Sigma} \otimes \Lambda^{0,1}T^*\dot{\Sigma}$ is isomorphic to the trivial complex line bundle $\epsilon^1 := \dot{\Sigma} \times \mathbb{C} \to \dot{\Sigma}$ via

$$\Lambda^{1,0}T^*\dot{\Sigma} \otimes \Lambda^{0,1}T^*\dot{\Sigma} \to \epsilon^1 : X^{1,0} \otimes Y^{0,1} \mapsto X^{1,0}(Y) = \langle X, Y\rangle_\Sigma.$$

EXERCISE 4.7.6. Assuming $\langle\ ,\ \rangle_F$ is chosen as the tensor product metric described above, show that under the natural identification of E with $\mathrm{Hom}_{\mathbb{C}}(T\dot{\Sigma}, F)$,

$$-\mathbf{D}^* : \Gamma(F) \to \Omega^{1,0}(\dot{\Sigma}, F)$$

satisfies the Leibniz rule

$$-\mathbf{D}^*(f\lambda) = (\partial f)\lambda + f(-\mathbf{D}^*\lambda)$$

for all $f \in C^1(\dot{\Sigma}, \mathbb{R})$, where $\partial f \in \Omega^{1,0}(\dot{\Sigma})$ denotes the complex-valued $(1,0)$-form $df - i\, df \circ j$.

We might summarize this exercise by saying that $-\mathbf{D}^*$ is an "anti-Cauchy-Riemann type" operator on F. But such an object is easily transformed into an honest Cauchy-Riemann type operator: let $\bar{F}$ denote the **conjugate bundle** to F, which we define as the same real vector bundle F but with the sign of its complex structure reversed, so $\lambda \mapsto -J \circ \lambda$. Now there is a canonical isomorphism

$$\operatorname{Hom}_{\mathbb{C}}(T\dot{\Sigma}, F) \cong \overline{\operatorname{Hom}_{\mathbb{C}}}(T\dot{\Sigma}, \bar{F}),$$

and the same operator defines a real-linear map

$$-\mathbf{D}^* : \Gamma(\bar{F}) \to \Omega^{0,1}(\dot{\Sigma}, \bar{F})$$

which satisfies our usual Leibniz rule for Cauchy-Riemann type operators.

Its asymptotic behavior also fits into the scheme we've been describing: we have already seen this by computing $\mathbf{D}^*$ on the ends with respect to asymptotic trivializations. To express this in trivialization-invariant language, observe that each of the Hermitian bundles (E_z, J_z, ω_z) over S^1 for $z \in \Gamma$ has a conjugate bundle $\bar{E}_z$ with complex structure $-J_z$ and symplectic structure $-\omega_z$; its natural Hermitian inner product is then the complex conjugate of the one on E_z. The asymptotic operator $\mathbf{A}_z$ on E_z can be expressed as $-J_z \hat{\nabla}_t$, where $\hat{\nabla}_t$ is a symplectic connection on (E_z, ω_z). Then $\hat{\nabla}_t$ is also a symplectic connection on $(\bar{E}_z, -\omega_z)$, so we naturally obtain an asymptotic operator on $\bar{E}_z$ in the form

$$\bar{\mathbf{A}}_z := -\mathbf{A}_z : \Gamma(\bar{E}_z) \to \Gamma(\bar{E}_z), \tag{4.24}$$

where the sign reversal arises from the reversal of the complex structure. One can check that if we choose a unitary trivialization of E_z and the conjugate trivialization of $\bar{E}_z$, this relationship between $\mathbf{A}_z$ and $\bar{\mathbf{A}}_z$ produces precisely the relationship between $\mathbf{A} = -J_0\partial_t - S_\infty(t)$ and $\bar{\mathbf{A}} = -J_0\partial_t - \bar{S}_\infty(t)$ that we saw previously, with $\bar{S}_\infty(t) = -CS_\infty(t)C$. Let us summarize all this with a theorem.

THEOREM 4.7.7. *Assume $\langle\ ,\ \rangle_F$ is chosen to be the tensor product metric on $F = \Lambda^{0,1}T^*\Sigma \otimes E$ induced by $\langle\ ,\ \rangle_E$ and the area form $d\,\mathrm{vol}$. Then under the isomorphism induced by $d\,\mathrm{vol}$ from E to $\operatorname{Hom}_{\mathbb{C}}(T\dot{\Sigma}, F)$ and the natural identification of the latter with its conjugate $\overline{\operatorname{Hom}_{\mathbb{C}}}(T\dot{\Sigma}, \bar{F})$, the operator $-\mathbf{D}^* : \Gamma(F) \to \Gamma(E)$ defines a linear Cauchy-Riemann type operator on the conjugate bundle $\bar{F}$,*

$$-\mathbf{D}^* : \Gamma(\bar{F}) \to \Omega^{0,1}(\dot{\Sigma}, \bar{F}),$$

and it is asymptotic at each puncture $z \in \Gamma$ to the conjugate asymptotic operator (4.24). □

4.8. The asymptotic formula

The theorem below is an addendum to the exponential decay discussion in §4.6, and in particular, it provides a vast improvement on Theorem 4.6.6. Such an improvement is not needed for the Fredholm theory in the present chapter, but we will have occasion to use it for the corresponding nonlinear result about exponential convergence in §6.7, and it will become crucially important when we discuss low-dimensional transversality results and intersection theory in Chapters 15 and 16. One can think of it as an asymptotic analogue of the similarity principle: one of its immediate consequences (Corollary 4.8.3 below) is that nontrivial solutions to linear Cauchy-Riemann type equations cannot have zeroes accumulating near infinity.

Like Theorem 4.6.6, the result can be stated for Cauchy-Riemann type operators of any class C^m with $m \geqslant 1$, but for simplicity, we will focus on the case $m = \infty$ in this section. It will be useful to introduce a strenghthening of our usual condition on the asymptotic behavior of such operators.

DEFINITION 4.8.1. Under the same assumptions as in Definition 4.1.4, we say that the operator $\mathbf{D}$ is **$C^{m,\delta}$-asymptotic to $\mathbf{A}_z$** at $z \in \Gamma^\pm$ for some constant $\delta \geqslant 0$ if $\mathbf{D}$ and $\mathbf{A}_z$ appear in an asymptotic trivialization near z as $\mathbf{D} = \bar{\partial} + S(s,t)$ and $\mathbf{A}_z = -J_0\partial_t - S_\infty(t)$ respectively with

$$\|e^{\pm\delta s}\left(S - S_\infty\right)\|_{C^k(Z^R_\pm)} \to 0 \quad \text{as} \quad R \to \infty$$

for all $k \leqslant m$.

For $\delta > 0$, this condition means that the convergence of $\mathbf{D}$ to its translation-invariant asymptotic form on the cylindrical end near z is exponentially fast. We will see in §6.7 that this assumption is always satisfied with $\delta > 0$ sufficiently small for the Cauchy-Riemann type operators that we care most about, namely those that arise by linearizing the nonlinear Cauchy-Riemann equation along J-holomorphic curves asymptotic to nondegenerate or Morse-Bott Reeb orbits. The following is an amalgamation of results due originally to Hofer-Wysocki-Zehnder [**HWZ96**], Mora [**Mor03**] and Siefring [**Sie08**].

THEOREM 4.8.2. *Assume $\mathbf{D}$ is of class C^∞ and is C^∞-asymptotic to the asymptotic operator $\mathbf{A}_z$ at the puncture $z \in \Gamma^\pm$, and $\eta \in \Gamma(E)$ is a nontrivial solution to the equation $\mathbf{D}\eta = 0$ satisfying an asymptotic estimate of the form*

$$\|\eta(s,\cdot)\|_{L^2(S^1)} \leqslant Ce^{\lambda_0 s} \qquad \text{on } Z_\pm \cong \dot{\mathcal{U}}_z$$

with respect to an asymptotic trivialization near z, for suitable constants $C > 0$ and $\lambda_0 \in \mathbb{R}$. Then there exists an eigenvalue $\lambda \in \sigma(\mathbf{A}_z)$ satisfying

$$\lambda \leqslant \lambda_0 \text{ if } z \in \Gamma^+, \qquad \text{or} \qquad \lambda \geqslant \lambda_0 \text{ if } z \in \Gamma^-,$$

such that on $\dot{\mathcal{U}}_z \cong Z_\pm$, η is given by the formula

$$\eta(s,t) = e^{\int_0^s \mu(\tau)\,d\tau} v(s,t) \tag{4.25}$$

for smooth functions $\mu(s)$ and $v(s,t)$ satisfying

$$\mu(s) \to \lambda, \qquad \text{and} \qquad \frac{d^k}{ds^k}\mu(s) \to 0 \qquad \text{as } s \to \pm\infty$$

for every $k \in \mathbb{N}$, and

$$\|v(s,\cdot)\|_{L^2(S^1)} \equiv \text{const} \neq 0 \qquad \text{and} \qquad v(s,\cdot) \to E_\lambda \text{ in } C^\infty(S^1) \text{ as } s \to \pm\infty,$$

where $E_\lambda \subset C^\infty(S^1)$ denotes the λ-eigenspace of $\mathbf{A}_z$. Moreover, if $\mathbf{D}$ is additionally $C^{\infty,\delta}$-asymptotic to $\mathbf{A}_z$ at z for some $\delta > 0$, then $v(s,\cdot)$ converges in $C^\infty(S^1)$ as $s \to \pm\infty$ to a single nontrivial eigenfunction in $e_\lambda \in E_\lambda$, and we can write

$$\eta(s,t) = e^{\lambda s}\left[e_\lambda(t) + r(s,t)\right], \tag{4.26}$$

for some remainder function $r(s,t)$ that decays along with its derivatives of all orders to 0 uniformly in t as $s \to \pm\infty$.

COROLLARY 4.8.3. *In the setting of Theorem 4.8.2, the solution η cannot have a sequence of zeroes on $\dot{\mathcal{U}}_z \cong Z_\pm$ accumulating at infinity.*

PROOF. By the uniqueness of solutions to linear ODEs, the nontrivial functions in the eigenspace E_λ to which the function $v(s,\cdot)$ in (4.25) converges uniformly are all nowhere zero, and since $\dim E_\lambda < \infty$, the eigenfunctions $e_\lambda \in E_\lambda$ satisfying $\|e_\lambda\|_{L^2} = \text{const} = \|v(s,\cdot)\|_{L^2(S^1)}$ form a compact set and thus satisfy a uniform bound away from zero. This implies that $v(s,\cdot)$ must also be nowhere zero for $s \gg 0$. □

As observed in the proof of Corollary 4.8.3, the eigenspace E_λ is finite dimensional, so its intersection with any sphere of constant radius in $L^2(S^1)$ is compact, and the convergence $v(s,\cdot) \to E_\lambda$ of the loops arising in (4.25) thus means that every sequence $s_k \to \infty$ has a subsequence for which $v(s_k,\cdot)$ converges to some specific nontrivial eigenfunction $e_\lambda \in E_\lambda$. In the generic situation, $\dim E_\lambda = 1$ and the intersection of E_λ with the unit sphere contains only two eigenfunctions—it then follows that $\lim_{s\to\infty} v(s,\cdot) = e_\lambda$. Example 4.8.5 below shows however that this limit does not always exist if $\dim E_\lambda > 1$, i.e. one might find multiple sequences $s_k, t_k \to \infty$ such that $v(s_k,\cdot)$ and $v(t_k,\cdot)$ converge to distinct eigenfunctions. Example 4.8.5 also shows that if $\mathbf{A}_z$ is degenerate and $\eta \to 0$ at infinity, then this convergence is not always guaranteed to be exponentially fast. However, the stronger formula (4.26) does guarantee such a result, and we will make use of this fact in §6.7:

COROLLARY 4.8.4. *Under the assumptions of Theorem 4.8.2, including the hypothesis that $\mathbf{D}$ is $C^{\infty,\delta}$-asymptotic to $\mathbf{A}_z$ for some $\delta > 0$, any solution η to the equation $\mathbf{D}\eta = 0$ that converges to 0 near $z \in \Gamma^\pm$ also satisfies an exponential decay condition $|\eta(s,t)| \leqslant Ce^{-\epsilon s}$ for suitable constants $C > 0$ and $\pm\epsilon > 0$.*

PROOF. There is nothing to prove if $\eta \equiv 0$, so assume otherwise. The hypothesis then implies that η is bounded near z, so (4.26) holds for some $\lambda \in \sigma(\mathbf{A}_z)$ with $\pm\lambda \leqslant 0$. If $\lambda = 0$, the formula then contradicts the assumption that η decays to 0 near z, since the eigenfunction e_λ must be nowhere zero. We conclude $\pm\lambda < 0$, which gives exponential decay. □

EXAMPLE 4.8.5. Assume $f : [0,\infty) \to \mathbb{C}\backslash\{0\}$ is a smooth function, and define a complex-linear Cauchy-Riemann type operator on the trivial complex line bundle over the half-cylinder $Z_+ = [0,\infty) \times S^1$ by

$$\mathbf{D} = \bar\partial + S(s,t), \qquad \text{for} \qquad S(s,t) := -\frac{f'(s)}{f(s)} \in \mathbb{C} = \operatorname{End}_{\mathbb{C}}(\mathbb{C}).$$

The function

$$\eta(s,t) := f(s)$$

then satisfies $\mathbf{D}\eta = 0$, and if f is chosen so that all derivatives of f'/f converge to 0 as $s \to \infty$, then $\mathbf{D}$ is C^∞-asymptotic to the trivial asymptotic operator $\mathbf{A}_0 := -J_0\partial_t$, which is degenerate. One can now observe various peculiar phenomena via different choices of f:

(1) If $f(s) = 1/s$ for all $s \geqslant 1$, then $S(s,t) = 1/s$ and η is thus a nontrivial solution to $\mathbf{D}\eta = 0$ that decays to 0 as $s \to \infty$, but not exponentially fast. This shows that in general, (4.25) in the case $\lim_{s\to\pm\infty}\mu(s) = 0$ does not imply any exponential decay estimate, so Corollary 4.8.4 fails without the extra hypothesis requiring exponential convergence of the operator to its asymptotic form.

(2) If $f(s) = e^{i\ln(s)}$ for $s \geqslant 1$, then $S(s,t) = -i/s$, and η is now a nontrivial solution for which $\|\eta(s,\cdot)\|_{L^2(S^1)} = 1$ for all s and $\eta(s,\cdot)$ does not converge as $s \to \infty$ to any specific eigenfunction, but instead wanders slowly through the unit sphere in the 2-dimensional eigenspace $E_0 = \ker \mathbf{A}_0$, which consists of constant functions.

The next exercise reveals a connection between the exponential convergence hypothesis behind (4.26) and the necessity of the condition $p > 2$ in the similarity principle (Theorem 2.5.3).

EXERCISE 4.8.6. Consider the coordinate transformation $z = e^{-2\pi(s+it)}$ which identifies points $(s,t) \in Z_+$ in the half-cylinder with points $z \in \dot{\mathbb{D}}$ in the punctured disk. Associate to any function $\eta : Z_+ \to \mathbb{R}^{2n}$ the function $\widehat{\eta} : \dot{\mathbb{D}} \to \mathbb{R}^{2n}$ defined by

$$\widehat{\eta}(z) := \eta(s,t),$$

and to each $\widehat{S} : \dot{\mathbb{D}} \to \mathrm{End}(\mathbb{R}^{2n})$ the function $S : Z_+ \to \mathrm{End}(\mathbb{R}^{2n})$ defined by

$$S(s,t) := 2\pi\bar{z}\widehat{S}(z).$$

Assume in the following that S and η are smooth; under suitable asymptotic hypotheses, $\widehat{S}$ and $\widehat{\eta}$ can then be regarded as locally integrable functions on $\mathbb{D}$ defined almost everywhere.

(a) Show that the equations $(\bar{\partial} + S)\eta = 0$ on Z_+ and $(\bar{\partial} + \widehat{S})\widehat{\eta} = 0$ on $\dot{\mathbb{D}}$ are equivalent.

(b) Deduce the following special case of the asymptotic formula (4.26) from the similarity principle: If S satisfies an exponential decay estimate $|S(s,t)| \leqslant Ce^{-\delta s}$ for some constants $C, \delta > 0$, then every bounded solution $\eta : Z_+ \to \mathbb{R}^{2n}$ to $(\bar{\partial} + S)\eta = 0$ satisfies

$$\eta(s,t) = e^{\lambda s}\left[e_\lambda(t) + r(s,t)\right]$$

for some nontrivial eigenfunction $e_\lambda \in E_\lambda$ of the trivial asymptotic operator $\mathbf{A}_0 := -J_0\partial_t$ with nonpositive eigenvalue $\lambda \in \sigma(\mathbf{A}_0) = 2\pi\mathbb{Z}$, and a remainder function $r(s,t)$ that decays to 0 uniformly in t as $s \to \infty$.

(c) Extract from Example 4.8.5 an example of a function $\widehat{S} : \mathbb{D} \to \mathrm{End}(\mathbb{R}^{2n})$ that is smooth on $\dot{\mathbb{D}}$ and of class L^2 on $\mathbb{D}$ but not in L^p for any $p > 2$, such that the equation $(\bar{\partial} + \widehat{S})\widehat{\eta} = 0$ admits multiple weak solutions $\widehat{\eta} : \mathbb{D} \to \mathbb{R}^{2n}$ that each separately have the following properties:

(a) $\widehat{\eta} \in W^{1,p}(\mathbb{D})$ for every $p \in (1, \infty)$, and its continuous extension over $\mathbb{D}$ has an isolated zero at $z = 0$, but this zero has order 0 and $\widehat{\eta}$ does not satisfy an estimate of the form $|\widehat{\eta}(z)| \leqslant C|z|^{\alpha}$ for any constants $C, \alpha > 0$.

(b) $\widehat{\eta}$ belongs to $W^{1,2}(\mathbb{D})$ but not to $W^{1,p}(\mathbb{D})$ for any $p > 2$, and it is bounded on $\dot{\mathbb{D}}$ but has no continuous extension to $z = 0$.

REMARK 4.8.7. The necessity of an exponential convergence hypothesis on $\mathbf{D}$ in order to obtain (4.26) has sometimes gone unappreciated in the literature, e.g. this author has written the occasional paper [**Wen10**] in which the formula was carelessly stated without any mention of such a hypothesis. However, this oversight is typically harmless for two reasons: one is that, as already mentioned above, the hypothesis is actually satisfied for most of the operators that are of interest in practice. But even without that, most of the standard applications of (4.26), including the one in [**Wen10**], can also be derived from the weaker formula (4.25); a nice example of this is Corollary 4.8.3 above.

We now explain the main elements of the proof of Theorem 4.8.2, referring to [**HWZ96**, **Sie08**] for some of the details. We shall reuse the notation $\mathbf{A} := \mathbf{A}_z$, $\mathbf{A}_s := -J_0\partial_t - S(s, \cdot) = \mathbf{A} + \Delta_s$, $\eta_s := \eta(s, \cdot)$, $\dot{\eta}_s := \partial_s\eta(s, \cdot)$, $v_s := \eta_s/\|\eta_s\|$ and so forth from the proof of Theorem 4.6.6, with $\| \ \|$ and $\langle \ , \ \rangle$ always denoting the L^2-norm and inner product respectively on S^1 whenever there are no other L^p-norms in the picture to cause confusion. Assume for simplicity that z is a positive puncture, so since $\mathbf{D}$ is of class C^∞, the solution $\eta|_{\dot{\mathcal{U}}_z}$ is a smooth function on $Z_+ = [0, \infty) \times S^1$. The result of Theorem 4.6.6 tells us

$$\|\eta_s\| = e^{\int_0^s \mu(\tau)\, d\tau}\|\eta_0\|$$

where

$$\mu(s) = \langle v_s, \mathbf{A}_s v_s \rangle \to \lambda \in \sigma(\mathbf{A}) \qquad \text{as } s \to \infty$$

and $\lambda \leqslant \lambda_0$. We shall also write

$$v(s,t) := v_s(t) = \frac{\eta(s,t)}{\|\eta_s\|},$$

thus defining v as another smooth function on Z_+. We then have

$$\eta(s,t) = e^{\int_0^s \mu(\tau)\, d\tau} w(s,t), \qquad \text{where} \qquad w(s,t) := \|\eta_0\| \cdot v(s,t),$$

so (4.25) will follow with w in the role of v if we can show that the derivatives of $\mu(s)$ decay and $v(s, \cdot) \to E_\lambda$ as $s \to \infty$.

The equation (4.17) satisfied by v_s translates into a linear Cauchy-Riemann type equation for v, namely

$$(\bar{\partial} + S(s,t) + \mu(s))v(s,t) = 0. \tag{4.27}$$

The function $\mu(s)$ is smooth since η and v are, so the Cauchy-Riemann type operator $\mathbf{D} + \mu$ in this equation is of class C^∞, but we have to be a bit careful for two reasons: its asymptotic operator is $\mathbf{A} - \lambda$, which is degenerate, and it may be only C^0-asymptotic to this operator since we have no knowledge as yet about the behavior of the derivatives of μ as $s \to \infty$.

LEMMA 4.8.8. *The functions* μ *and* v *on* $[0,\infty)$ *and* Z_+ *respectively have globally bounded derivatives of all orders.*

PROOF. For $r > 0$ and $N \geqslant r$, define the domains

$$I_N^r := (N-r, N+r), \qquad Z_N^r := I_N^r \times S^1.$$

By the Sobolev embedding theorem, it will suffice to prove that for some $r > 0$ and $p \in (1,\infty)$ and arbitrarily large N, there exist bounds $\|\mu\|_{W^{k,p}(I_N^r)} \leqslant c$ and $\|v\|_{W^{k,p}(I_N^r)} \leqslant c$ for arbitrarily large $k \in \mathbb{N}$, with constants $c > 0$ that may depend on k, p, r but not on N. (The word "uniform" should be understood in the following to mean "independent of N".) In light of (4.27), the basic regularity estimate from Lemma 4.3.1 gives

$$\begin{aligned} \|v\|_{W^{k,p}(Z_N^r)} &\leqslant c\|\bar{\partial}v\|_{W^{k-1,p}(Z_N^R)} + c\|v\|_{W^{k-1,p}(Z_N^R)} \\ &= c\|(S+\mu)v\|_{W^{k-1,p}(Z_N^R)} + c\|v\|_{W^{k-1,p}(Z_N^R)} \end{aligned} \tag{4.28}$$

for any $k \in \mathbb{N}$, $p \in (1,\infty)$ and $R > r$, with constants $c > 0$ that do not depend on N. Since $\|v_s\| = 1$ for every s, Fubini's theorem implies an L^2-bound for v over Z_N^R that is independent of N, and since there is clearly also a uniform C^0-bound on μ, this estimate gives a uniform $W^{1,2}$-bound for v on each Z_N^r. By the Sobolev embedding theorem (the $kp = n$ case of Theorem A.1.6), we therefore have uniform L^p-bounds on v over each Z_N^r for every $p \in [2,\infty)$, and applying (4.28) again then gives uniform $W^{1,p}$-bounds for v over each Z_N^r after slightly shrinking r.

Before this can be improved further, we need some more knowledge of the derivatives of μ. According to the top line of (4.20),

$$\dot{\mu}(s) = 2\|\dot{v}_s\|^2 + \langle \dot{v}_s, \Delta_s v_s\rangle + \langle v_s, \Delta_s \dot{v}_s\rangle + \epsilon(s), \tag{4.29}$$

where $|\epsilon(s)|$ can be assumed arbitrarily small for $s \gg 0$. One can use the Cauchy-Schwarz inequality and the boundedness of Δ_s and $\|v_s\|$ to extract from this a bound of the form

$$|\dot{\mu}(s)| \leqslant c_1\|\dot{v}_s\|^2 + c_2$$

for suitable constants $c_1, c_2 > 0$, and for any $p \in (1,\infty)$, this implies

$$|\dot{\mu}(s)|^p \leqslant c_1\|\dot{v}_s\|_{L^2}^{2p} + c_2$$

after modifying the constants. Now if $\frac{1}{p} + \frac{1}{q} = 1$, we can use Hölder's inequality for functions on S^1 to write

$$\|\dot{v}_s\|_{L^2}^{2p} = \left(\int_{S^1} |\partial_s v(s,t)|^2\, dt\right)^p \leqslant \|1\|_{L^q(S^1)}^p \cdot \int_{S^1} |\partial_s v(s,t)|^{2p}\, dt,$$

which turns the previous estimate into

$$|\dot{\mu}(s)|^p \leqslant c_1 \int_{S^1} |\partial_s v(s,t)|^{2p}\, dt + c_2,$$

so integrating this with respect to s over the interval I_N^r then produces a uniform bound for $\|\mu\|_{W^{1,p}}$ on that interval in terms of the uniform bounds on $\|v\|_{W^{1,2p}}$ that we already obtained over Z_N^r. Feeding this information into (4.28) for $k = 2$ and $p > 2$, the $W^{1,p}$-norm of $(S+\mu)v$ on the right hand side can now be bounded uniformly due to the Banach algebra property of $W^{1,p}$, and we conclude that after shrinking $r > 0$ slightly further, v satisfies a uniform $W^{2,p}$-bound over Z^r_N for all N.

This back-and-forth process of improving the Sobolev bounds on v and μ can be continued inductively for all higher derivatives; for details, see [**HWZ96**, Lemma 3.3] or [**Sie08**, Lemma A.5]. □

COROLLARY 4.8.9. *For all $k \in \mathbb{N}$, $\partial_s^k \mu(s) \to 0$ as $s \to \infty$.*

PROOF. Consider an arbitrary sequence $s_j \to \infty$ and define the sequence of functions $\mu_j(s) := \mu(s_j + s)$. Lemma 4.8.8 implies that μ_j satisfies uniform C^∞-bounds, so by Arzelà-Ascoli, it has a C^∞_{loc}-convergent subsequence, whose limit can only be the constant function $\mu_\infty(s) := \lambda$, implying $\partial_s^k \mu(s_j) \to \partial_s^k \mu_\infty(0) = 0$. □

COROLLARY 4.8.10. *For every multiindex α, $\partial_s \partial^\alpha v(s, \cdot)$ converges uniformly to 0 as $s \to \infty$.*

PROOF. Every term in (4.29) except for $\|\dot{v}_s\|^2$ is already known to decay to 0 as $s \to \infty$: for $\dot{\mu}(s)$ this comes from Corollary 4.8.9, and for the other terms on the right hand side it follows from the fact that v has bounded derivatives while $\Delta_s \to 0$. This proves

$$\|\partial_s v(s, \cdot)\| \to 0 \quad \text{as } s \to \infty,$$

and it follows via the boundedness of all derivatives of v and the Arzelà-Ascoli theorem that every sequence $s_j \to \infty$ has a subsequence for which the functions $\partial_s v_j(s,t) := \partial_s v(s_j + s, t)$ converge in C^∞_{loc} to 0. This prevents the existence of a multiindex α and sequence $s_j \to \infty$ for which $\|\partial_s \partial^\alpha(s_j, \cdot)\|_{C^0(S^1)}$ is bounded away from 0. □

LEMMA 4.8.11. *Every sequence $s_k \to \infty$ has a subsequence such that $v(s_k, \cdot)$ converges in $C^\infty(S^1)$ to a nontrivial eigenfunction e_λ of $\mathbf{A}$ with eigenvalue λ.*

PROOF. Using Lemma 4.8.8, the Arzelà-Ascoli theorem provides a subsequence such that $v_k(s,t) := v(s_k + s, t)$ converges in C^∞_{loc} on Z_+ to a smooth function v_∞ which will satisfy the limit of the corresponding translations of the equation $(\bar{\partial} + S(s,t) + \mu(s))v = 0$, namely

$$(\bar{\partial} + S_\infty(t) + \lambda)v_\infty = (\partial_s - \mathbf{A} + \lambda)v_\infty = 0.$$

By Corollary 4.8.10, v_∞ must also satisfy $\partial_s v_\infty \equiv 0$, hence it is a function of the form $v_\infty(s,t) = e_\lambda(t)$ with $(-\mathbf{A} + \lambda)e_\lambda = 0$. Finally, v_∞ must be nontrivial because for each s, $\|v_\infty(s, \cdot)\| = \lim_{k\to\infty} \|v(s_k + s, \cdot)\| = 1$. □

The proof of (4.25) with the associated conditions on μ and v is now complete.

We will give only a brief sketch of the proof of (4.26), referring to [**Sie08**, Appendix A] for the details. The improvement in the scaling factor from $e^{\int_0^s \mu(\tau)\, d\tau}$ to $e^{\lambda s}$ is based on showing that if Δ_s and its derivatives decay exponentially as $s \to \infty$, then so does $\lambda - \mu(s)$, i.e. there is an estimate of the form

$$|\lambda - \mu(s)| \leqslant Ce^{-\delta s}, \tag{4.30}$$

for constants $C, \delta > 0$ that need not be related to the constants in the hypothesis of the theorem. (In the following we reserve the freedom to change the value of such constants whenever convenient, and we will sometimes do so without mentioning it.) To see why this helps, notice that it implies a similar decay estimate for the integral $\int_s^\infty |\lambda - \mu(\tau)|\, d\tau$ as a function of s, and thus

$$\left|\lambda s - \int_0^s \mu(\tau)\, d\tau\right| \leqslant \int_0^s |\lambda - \mu(\tau)|\, d\tau$$
$$= \int_0^\infty |\lambda - \mu(\tau)|\, d\tau - \int_s^\infty |\lambda - \mu(\tau)|\, d\tau \leqslant I + Ce^{-\delta s},$$

where $I := \int_0^\infty |\lambda - \mu(\tau)|\, d\tau$ is another constant independent of s. This implies

$$\eta(s,t) = e^{\int_0^s \mu(\tau)\, d\tau}\|\eta_0\| v(s,t) = e^{\lambda s} e^{R(s)} \|\eta_0\| v(s,t)$$

for a function $R(s) \in \mathbb{R}$ that satisfies $|R(s)| \leqslant I + Ce^{-\delta s}$ and thus $\lim_{s\to\infty} R(s) = I$. We can now define a new remainder function $R_1(s) \in \mathbb{R}$ with $\lim_{s\to\infty} R_1(s) = 0$ such that $e^{R(s)} = e^I + R_1(s)$. Rescaling v by defining $\widetilde{v} := e^I \|\eta_0\| v$, this implies

$$\eta(s,t) = e^{\lambda s}\left[\widetilde{v}(s,t) + r(s,t)\right]$$

for some remainder $r(s,t)$ that decays to 0 as $s \to \infty$. It then remains only to prove that $v(s,\cdot)$ converges as $s \to \infty$ to a single eigenfunction $e_\lambda \in E_\lambda$, which implies the same statement about $\widetilde{v}$. Writing

$$P : L^2(S^1) \to L^2(S^1)$$

for the orthogonal projection onto the eigenspace E_λ, it suffices to prove that Pv_s has a well-defined L^2-limit as $s \to \infty$, since we already know $v(s,\cdot) \to E_\lambda$ and thus $(\mathbb{1} - P)v_s \to 0$. For this purpose, we can regard $s \mapsto v_s$ as a continuously differentiable path in $L^2(S^1)$ and, for any $a, b \geqslant 0$, write

$$Pv_b - Pv_a = \int_a^b \partial_s(Pv_s)\, ds = \int_a^b P\dot{v}_s\, ds \in L^2(S^1).$$

The crucial detail is now that the exponential decay of Δ_s implies

$$\|P\dot{v}_s\| \leqslant Ce^{-\delta s}, \tag{4.31}$$

so that if $b > a \geqslant R \gg 0$, then

$$\|Pv_b - Pv_a\| \leqslant \int_a^b \|P\dot{v}_s\|\, ds \leqslant C\int_a^b e^{-\delta s}\, ds \leqslant C\int_R^\infty e^{-\delta s}\, ds,$$

which can also be made arbitrarily small by taking R large. This implies that for every $s_k \to \infty$, Pv_{s_k} is an L^2-Cauchy sequence, and therefore converges.

For the proofs of (4.30) and (4.31), see Lemmas A.8 through A.10 in [**Sie08**].

CHAPTER 5

The index formula

Contents

5.1. Riemann-Roch with punctures

As in the previous chapter, let $\mathbf{D}$ denote a linear Cauchy-Riemann type operator of class C^m $(1 \leqslant m \leqslant \infty)$ on an asymptotically Hermitian vector bundle E of complex rank $n \in \mathbb{N}$ over a punctured Riemann surface $(\dot{\Sigma} = \Sigma \backslash (\Gamma^+ \cup \Gamma^-), j)$, and assume that $\mathbf{D}$ is asymptotic at each puncture $z \in \Gamma$ to a nondegenerate asymptotic operator $\mathbf{A}_z$ on the asymptotic bundle (E_z, J_z, ω_z) over S^1. Writing

$$F := \overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, E)$$

for the bundle of complex-antilinear homomorphisms $T\dot{\Sigma} \to E$, the main result of the previous chapter was that

$$\mathbf{D} : W^{k,p}(E) \to W^{k-1,p}(F)$$

is Fredholm for any $k \in \{1, \ldots, m+1\}$ and $p \in (1, \infty)$, and its kernel and index do not depend on k or p. The main goal of this chapter is to compute $\operatorname{ind}(\mathbf{D}) \in \mathbb{Z}$.

The index will depend on the Conley-Zehnder indices $\mu_{\mathrm{CZ}}^{\tau}(\mathbf{A}_z) \in \mathbb{Z}$ introduced in Chapter 3, but since these depend on arbitrary choices of unitary trivializations τ, we need a way of selecting preferred trivializations. The most natural condition is to require that every (E_z, J_z, ω_z) be endowed with a unitary trivialization such that the corresponding asymptotic trivializations of (E, J) extend to a global trivialization;[1] if there is only one puncture z, for instance, then this condition determines $\mu_{\mathrm{CZ}}^{\tau}(\mathbf{A}_z)$ uniquely. This convention has been used to state the formula for $\operatorname{ind}(\mathbf{D})$ in several of the standard references, e.g. in [HWZ99]. We would prefer however to state a

[1] Note that (E, J) is always globally trivializable unless $\Gamma = \varnothing$, as a punctured surface can be retracted to its 1-skeleton.

formula which is also valid when $\Gamma = \emptyset$ and $E \to \Sigma$ is nontrivial. One way to do this is by allowing completely arbitrary asymptotic trivializations, but introducing a topological invariant to measure their failure to extend globally over E.

DEFINITION 5.1.1. Fix a compact oriented surface S with boundary. The **relative first Chern number** associates to every complex vector bundle (E, J) over S and trivialization τ of $E|_{\partial S}$ an integer

$$c_1^\tau(E) \in \mathbb{Z}$$

satisfying the following properties:

(1) If $(E, J) \to S$ is a line bundle, then $c_1^\tau(E)$ is the signed count of zeroes for a generic section $\eta \in \Gamma(E)$ that appears as a nonzero constant at ∂S with respect to τ.
(2) For any two bundles (E_1, J_1) and (E_2, J_2) with trivializations τ_1 and τ_2 respectively over ∂S,

$$c_1^{\tau_1 \oplus \tau_2}(E_1 \oplus E_2) = c_1^{\tau_1}(E_1) + c_1^{\tau_2}(E_2).$$

Note that in the first point above, counting zeroes "with signs" actually means adding up their *orders* in the sense of complex analysis, so e.g. the function $z \mapsto z^k$ has a zero of order k at the origin if $k \geqslant 1$, while $z \mapsto \bar{z}^k$ has a zero of order $-k$.[2] It follows from standard arguments in differential topology (see [Mil97]) that this count of zeroes is invariant under homotopies of sections that are nowhere zero at ∂S, thus $c_1^\tau(E)$ for a line bundle does not depend on the choice of section, though it does depend (up to homotopy) on the choice of boundary trivialization τ. It is also not hard to show via genericity arguments that a higher rank complex vector bundle over a compact surface can always be split into a direct sum of line bundles, and while this splitting is not uniquely determined, changing the topology of any summand forces correspondinging changes in other summands such that the sum of their relative first Chern numbers remains unchanged. It follows that the conditions stated above uniquely determine $c_1^\tau(E)$ for all complex vector bundles over compact oriented surfaces. The definition clearly matches the usual first Chern number $c_1(E) \in \mathbb{Z}$ when $\partial S = \emptyset$, in which case there is no need for any choice of trivializations. It also extends in a natural way to the category of asymptotically Hermitian vector bundles with asymptotic trivializations:

DEFINITION 5.1.2. For an asymptotically Hermitian vector bundle $E \to \dot{\Sigma}$ over a punctured Riemann surface $\dot{\Sigma} = \Sigma \backslash (\Gamma^+ \cup \Gamma^-)$, with trivializations on the cylindrical ends determined by a choice τ of unitary trivialization for the asymptotic bundle (E_z, J_z, ω_z) associated to each puncture $z \in \Gamma^\pm$, we define

$$c_1^\tau(E) := c_1^\tau(E|_{\bar{\Sigma}^R}) \qquad \text{for any } R \geqslant 0,$$

[2]The precise definition can be phrased in terms of winding numbers: for a function $f : \mathbb{C} \to \mathbb{C}$ with an isolated zero $f(z_0) = 0$, the zero has order $k \in \mathbb{Z}$ if the loop $\theta \mapsto f(z_0 + \epsilon e^{i\theta}) \in \mathbb{C}\backslash\{0\}$ has winding number k for all $\epsilon > 0$ sufficiently small. Note that this changes by a sign if the function is composed with an orientation-reversing homeomorphism of its domain, thus $c_1^\tau(E)$ depends on the orientation of S.

where $\bar{\Sigma}^R \subset \dot{\Sigma}$ is the compact oriented surface with boundary obtained by deleting all of the positive and negative cylindrical ends $(R, \infty) \times S^1$ and $(-\infty, -R) \times S^1$ respectively.

EXERCISE 5.1.3. Given two distinct choices of boundary trivializations τ_1 and τ_2 for a complex vector bundle E of rank n over a compact oriented surface, show that

$$c_1^{\tau_2}(E) = c_1^{\tau_1}(E) - \deg(\tau_2 \circ \tau_1^{-1}),$$

where $\deg(\tau_2 \circ \tau_1^{-1}) \in \mathbb{Z}$ denotes the sum over all boundary components of the winding numbers of the determinants of the transition maps $S^1 \to \mathrm{U}(m)$.

REMARK 5.1.4. To apply Exercise 5.1.3 in the setting of an asymptotically Hermitian bundle over a punctured Riemann surface $\dot{\Sigma}$, the winding number at each negative end must be computed by traversing $\{-R\} \times S^1$ in the *wrong direction*, or equivalently, by inserting a minus sign in front of the usual winding number along $\{-R\} \times S^1$. This is consistent with the orientation induced on $\{-R\} \times S^1$ as a boundary component of the compact subdomain $\Sigma^R \subset \dot{\Sigma}$.

EXERCISE 5.1.5. Combining Exercise 5.1.3 and Remark 5.1.4 above with Exercise 3.7.3, show that for our asymptotically Hermitian vector bundle E with Cauchy-Riemann type operator $\mathbf{D}$ and asymptotic operators $\mathbf{A}_z$, the number

$$2c_1^\tau(E) + \sum_{z \in \Gamma^+} \mu_{\mathrm{CZ}}^\tau(\mathbf{A}_z) - \sum_{z \in \Gamma^-} \mu_{\mathrm{CZ}}^\tau(\mathbf{A}_z)$$

is independent of the choice of asymptotic trivializations τ.

The above exercise shows that the right hand side of the following index formula is independent of all choices.

THEOREM 5.1.6. *The Fredholm index of* $\mathbf{D}$ *is given by*

$$\operatorname{ind} \mathbf{D} = n\chi(\dot{\Sigma}) + 2c_1^\tau(E) + \sum_{z \in \Gamma^+} \mu_{CZ}^\tau(\mathbf{A}_z) - \sum_{z \in \Gamma^-} \mu_{CZ}^\tau(\mathbf{A}_z),$$

where $n = \operatorname{rank}_{\mathbb{C}} E$ *and* τ *is an arbitrary choice of asymptotic trivializations.*

REMARK 5.1.7. The case $n = 0$ is allowed in the above formula: Then $c_1^\tau(E) = 0$ and all the Conley-Zehnder indices vanish by convention (cf. Remark 3.6.12), while on the left hand side, $\mathbf{D}$ is the unique linear operator between two 0-dimensional vector spaces—which is Fredholm with index 0. This case will be relevant to the dimension of the moduli space of holomorphic branched covers of a punctured Riemann surface; see Prop. 15.3.1.

NOTATION. Throughout this chapter, we shall denote the integer on the right hand side in Theorem 5.1.6 by

$$I(\mathbf{D}) := n\chi(\dot{\Sigma}) + 2c_1^\tau(E) + \sum_{z \in \Gamma^+} \mu_{\mathrm{CZ}}^\tau(\mathbf{A}_z) - \sum_{z \in \Gamma^-} \mu_{\mathrm{CZ}}^\tau(\mathbf{A}_z) \in \mathbb{Z}.$$

Our goal is thus to prove that $\operatorname{ind}(\mathbf{D}) = I(\mathbf{D})$.

When $\Gamma = \emptyset$, Theorem 5.1.6 is equivalent to the classical Riemann-Roch formula, which is more often stated for *holomorphic* vector bundles over a closed Riemann surface (Σ, j) with genus g as

$$\operatorname{ind}_{\mathbb{C}}(\mathbf{D}_0) = n(1-g) + c_1(E), \tag{5.1}$$

where $\mathbf{D}_0$ is the canonical complex-linear Cauchy-Riemann type operator determined by the holomorphic structure. An arbitrary real-linear Cauchy-Riemann type operator is then of the form $\mathbf{D} = \mathbf{D}_0 + B$, where the zeroth-order term $B \in \Gamma(\operatorname{Hom}_{\mathbb{R}}(E, F))$ defines a compact perturbation since the inclusion $W^{k,p}(\Sigma) \hookrightarrow W^{k-1,p}(\Sigma)$ is compact. It follows that $\mathbf{D}$ has the same *real* Fredholm index as $\mathbf{D}_0$, namely twice the complex index shown on the right hand side of (5.1), which matches what we see in Theorem 5.1.6.

REMARK 5.1.8. Now seems a good moment to clarify explicitly that all dimensions (and therefore also Fredholm indices) in this book are *real* dimensions, not complex dimensions, unless otherwise specified.

Reduction to the complex-linear case does not work in general if there are punctures. It remains true that arbitrary Cauchy-Riemann type operators can be written as $\mathbf{D} = \mathbf{D}_0 + B$ where $\mathbf{D}_0$ is complex linear, but the perturbation introduced by the zeroth-order term B is not compact since $W^{k,p}(\dot{\Sigma}) \hookrightarrow W^{k-1,p}(\dot{\Sigma})$ is not compact when $\Gamma \neq \emptyset$. Another indication that this idea cannot work is the fact that while the formula in Theorem 5.1.6 always gives an *even* integer when $\Gamma = \emptyset$, it can be odd when there are punctures, in which case $\mathbf{D}$ clearly cannot have the same index as any complex-linear operator. Our proof will therefore have to deal with more than just the complex category.

The punctured version of Theorem 5.1.6 was first proved by Schwarz in his thesis [**Sch95**], its main purpose at the time being to help define algebraic operations (notably the *pair-of-pants product*) in Hamiltonian Floer homology. Schwarz's proof used a "linear gluing" construction that gives a relation between indices of operators on bundles over surfaces obtained by gluing together constituent surfaces along matching cylindrical ends. Since any surface with ends can be "capped off" to form a closed surface, one obtains the general index formula if one already knows how to compute it for closed surfaces and for planes (i.e. caps). For the latter, it is simple enough to write down model Cauchy-Riemann operators on planes and compute their kernels and cokernels explicitly, so in this way, the general case is reduced to the classical Riemann-Roch formula. An analogous linear gluing argument for compact surfaces with boundary is used in [**MS12**, Appendix C] to reduce the general Riemann-Roch formula to an explicit computation for Cauchy-Riemann operators on the disk with a totally real boundary condition.

In this chapter, we will follow a different path, and use an argument that was first sketched by Taubes for the closed case in [**Tau96a**, §7], with an additional argument for the punctured case that was suggested by Chris Gerig [**Ger18**]. The argument is (in my opinion) analytically somewhat easier than the more standard approaches, and in addition to proving the formula we need for punctured surfaces, it produces a new proof in the closed case without assuming the classical Riemann-Roch formula. It also provides a gentle preview of two analytical phenomena that will later assume prominent roles in our discussion of SFT: *bubbling* and *gluing*.

To see the idea behind Taubes's argument, we can start by noticing an apparent numerical coincidence in the closed case. Assume (E, J) is a complex line bundle over a closed Riemann surface (Σ, j), and $\mathbf{D} : \Gamma(E) \to \Gamma(F) = \Omega^{0,1}(\Sigma, E)$ is a Cauchy-Riemann type operator. We know that $\operatorname{ind}(\mathbf{D}) = \operatorname{ind}(\mathbf{D} + B)$ for any zeroth-order term $B \in \Gamma(\operatorname{Hom}_{\mathbb{R}}(E, F))$. But E and F are both complex vector bundles, so B can always be split uniquely into its complex-linear and complex-antilinear parts, i.e. there is a natural splitting of $\operatorname{Hom}_{\mathbb{R}}(E, F)$ into a direct sum of complex line bundles[3]

$$\operatorname{Hom}_{\mathbb{R}}(E, F) = \operatorname{Hom}_{\mathbb{C}}(E, F) \oplus \overline{\operatorname{Hom}}_{\mathbb{C}}(E, F).$$

Out of curiosity, let's compute the first Chern number of the second factor; this will be the signed count of zeroes of a generic complex-*antilinear* zeroth-order perturbation. To start with, note that

$$\overline{\operatorname{Hom}}_{\mathbb{C}}(E, F) \cong \overline{\operatorname{Hom}}_{\mathbb{C}}(E, \mathbb{C}) \otimes F,$$

and then observe that $\overline{\operatorname{Hom}}_{\mathbb{C}}(E, \mathbb{C})$ and E are isomorphic: indeed, any Hermitian bundle metric $\langle\ ,\ \rangle_E$ on E gives rise to a bundle isomorphism[4]

$$E \to \overline{\operatorname{Hom}}_{\mathbb{C}}(E, \mathbb{C}) : \eta \mapsto \langle \cdot, \eta \rangle_E.$$

We thus have $\overline{\operatorname{Hom}}_{\mathbb{C}}(E, F) \cong E \otimes F$, so $c_1(\overline{\operatorname{Hom}}_{\mathbb{C}}(E, F)) = c_1(E) + c_1(F)$. We can compute $c_1(F)$ by the same trick, since

$$F = \overline{\operatorname{Hom}}_{\mathbb{C}}(T\Sigma, E) \cong \overline{\operatorname{Hom}}_{\mathbb{C}}(T\Sigma, \mathbb{C}) \otimes E \cong T\Sigma \otimes E,$$

so $c_1(F) = c_1(T\Sigma) + c_1(E) = \chi(\Sigma) + c_1(E)$ by the Poincaré-Hopf theorem, and thus

$$c_1(\overline{\operatorname{Hom}}_{\mathbb{C}}(E, F)) = \chi(\Sigma) + 2c_1(E).$$

Since we're looking at a line bundle over a surface without punctures, this number is the same as $I(\mathbf{D})$. This coincidence is too improbable to ignore, and indeed, it turns out not to be coincidental. Here is an informal statement of a result that we will later prove a more precise version of in order to deduce Theorem 5.1.6.

"THEOREM". *Given a Cauchy-Riemann type operator* $\mathbf{D} : H^1(E) \to L^2(F)$ *on a line bundle* (E, J) *over a closed Riemann surface* (Σ, j), *choose a complex-antilinear zeroth-order perturbation* $B \in \Gamma(\overline{\operatorname{Hom}}_{\mathbb{C}}(E, F))$ *whose zeroes are all non-degenerate. Then for sufficiently large* $r > 0$, $\ker(\mathbf{D} + rB)$ *is approximately spanned by* 1*-dimensional spaces of sections with support localized near the positive zeroes of* B. *In particular,* $\dim \ker(\mathbf{D} + rB)$ *equals the number of positive zeroes of* B.

To deduce $\operatorname{ind}(\mathbf{D}) = I(\mathbf{D})$ from this, we need to apply the same trick to the formal adjoint $\mathbf{D}^*$. As we will review in §5.2, $-\mathbf{D}^*$ can be regarded under certain natural assumptions as a Cauchy-Riemann type operator on the bundle $\bar{F}$ conjugate

[3] Here the complex structure on $\operatorname{Hom}_{\mathbb{R}}(E, F)$ and its subbundles is defined in terms of the complex structure of F, i.e. it sends $B \in \operatorname{Hom}_{\mathbb{R}}(E, F)$ to $J \circ B \in \operatorname{Hom}_{\mathbb{R}}(E, F)$.

[4] We are assuming as usual that Hermitian inner products are complex antilinear in the first argument and linear in the second.

to F, and the formal adjoint of $\mathbf{D} + rB$ then gives rise to a Cauchy-Riemann type operator of the form

$$-\mathbf{D}^* + rB' : \Gamma(\bar{F}) \to \Gamma(\bar{E}) = \Omega^{0,1}(\Sigma, \bar{F}),$$

where $B' : \bar{F} \to \bar{E}$ is also complex antilinear and has the same zeroes as B, but with opposite signs. Applying the above "theorem" to $-\mathbf{D}^*$ thus identifies $\ker(\mathbf{D} + rB)^*$ for sufficiently large $r > 0$ with a space whose dimension equals the number of *negative* zeroes of B. This gives

$$\begin{aligned}\operatorname{ind}(\mathbf{D}) &= \operatorname{ind}(\mathbf{D} + rB) = \dim\ker(\mathbf{D} + rB) - \dim\ker(\mathbf{D} + rB)^* \\ &= c_1(\overline{\operatorname{Hom}}_{\mathbb{C}}(E, F)) = I(\mathbf{D}).\end{aligned}$$

It's worth mentioning that the "large perturbation" argument we've just sketched is only one simple example of an idea with a long and illustrious history. Another simple example is the observation by Witten [**Wit82**] that after choosing a Morse function on a Riemannian manifold, certain large deformations of the de Rham complex lead to an approximation of the Morse complex, with generators of the de Rham complex having support concentrated near the critical points of the Morse function—this yields a somewhat novel proof of de Rham's theorem. A much deeper example is Taubes's isomorphism [**Tau96b**] between the Seiberg-Witten invariants of symplectic 4-manifolds and certain holomorphic curve invariants: here also, the idea is to consider a large compact perturbation of the Seiberg-Witten equations and show that, in the limit where the perturbation becomes infinitely large, solutions of the Seiberg-Witten equations localize near J-holomorphic curves. For a more recent exploration of this idea in the context of Dirac operators, see [**Mar17**].

Before proceeding with the details, let us fix three simplifying assumptions that can be imposed without loss of generality:

ASSUMPTION 5.1.9. *E and $\mathbf{D}$ are of class C^∞.*

This can always be achieved by an arbitrarily small perturbation, and small perturbations do not change the Fredholm index.

ASSUMPTION 5.1.10. *(E, J) has complex rank* 1.

Indeed, an asymptotically Hermitian bundle E of complex rank $n \in \mathbb{N}$ always admits a decomposition into asymptotically Hermitian line bundles $E = E_1 \oplus \ldots \oplus E_n$, producing a corresponding splitting of the target bundle $F = F_1 \oplus \ldots \oplus F_n$. The operator $\mathbf{D}$ need not respect these splittings, but it is always *homotopic through Fredholm operators* to one that does: we saw in Theorem 3.6.13 that the asymptotic operators $\mathbf{A}_z$ are homotopic through nondegenerate asymptotic operators to any other operators $\mathbf{A}'_z$ that have the same Conley-Zehnder indices, so one can choose $\mathbf{A}'_z$ to respect the splitting. Any homotopy of Cauchy-Riemann operators following such a homotopy of nondegenerate asymptotic operators then produces a continuous family of Fredholm operators by the main result of Chapter 4, implying that their indices do not change. The general index formula then follows from the line bundle case since any two Cauchy-Riemann type Fredholm operators $\mathbf{D}_1$ and $\mathbf{D}_2$ on bundles over the same Riemann surface satisfy

$$\operatorname{ind}(\mathbf{D}_1 \oplus \mathbf{D}_2) = \operatorname{ind}(\mathbf{D}_1) + \operatorname{ind}(\mathbf{D}_2) \quad \text{and} \quad I(\mathbf{D}_1 \oplus \mathbf{D}_2) = I(\mathbf{D}_1) + I(\mathbf{D}_2).$$

ASSUMPTION 5.1.11. *$k = 1$ and $p = 2$.*

This means we will concretely be considering the operator

$$\mathbf{D} : H^1(E) \to L^2(F),$$

where H^1 as usual is an abbreviation for $W^{1,2}$. This assumption is clearly harmless since we know that $\operatorname{ind} \mathbf{D}$ does not depend on the choice of k and p.

5.2. Some remarks on the formal adjoint

For the beginning of this section, we can drop the assumption that (E, J) is a line bundle, and assume $\operatorname{rank}_{\mathbb{C}} E = n \in \mathbb{N}$, though later we will again set $n = 1$.

Recall from §4.7 that if we fix global Hermitian bundle metrics $\langle\ ,\ \rangle_E$ and $\langle\ ,\ \rangle_F$ on (E, J) and (F, J) respectively, and an area form $d\,\mathrm{vol}$ on $\dot{\Sigma}$ that matches $ds \wedge dt$ on the cylindrical ends, then $\mathbf{D}$ has a *formal adjoint*

$$\mathbf{D}^* : \Gamma(F) \to \Gamma(E)$$

satisfying

$$\langle \lambda, \mathbf{D}\eta \rangle_{L^2(F)} = \langle \mathbf{D}^*\lambda, \eta \rangle_{L^2(E)} \quad \text{for all} \quad \eta \in H^1(E),\ \lambda \in H^1(F).$$

Here, the real-valued L^2-pairings are defined by

$$\langle \eta, \xi \rangle_{L^2(E)} := \operatorname{Re} \int_{\dot{\Sigma}} \langle \eta, \xi \rangle_E \, d\,\mathrm{vol} \quad \text{for} \quad \eta, \xi \in \Gamma(E),$$

and similarly for sections of F. The essential features of the formal adjoint are that $\ker \mathbf{D}^* \cong \operatorname{coker} \mathbf{D}$ and $\operatorname{coker} \mathbf{D}^* \cong \ker \mathbf{D}$, hence $\operatorname{ind}(\mathbf{D}^*) = -\operatorname{ind}(\mathbf{D})$. Recall moreover that $d\,\mathrm{vol}$ induces a natural Hermitian bundle metric on $\dot{\Sigma}$ by

$$\langle \cdot, \cdot \rangle_\Sigma = d\,\mathrm{vol}(\cdot, j\cdot) + i\, d\,\mathrm{vol}(\cdot, \cdot),$$

which determines a bundle isomorphism

$$T\dot{\Sigma} \to \Lambda^{0,1}T^*\dot{\Sigma} : X \mapsto X^{0,1} := \langle \cdot, X \rangle_\Sigma,$$

as well as a complex-*antilinear* isomorphism

$$T\dot{\Sigma} \to \Lambda^{1,0}T^*\dot{\Sigma} : X \mapsto X^{1,0} := \langle X, \cdot \rangle_\Sigma.$$

If $\langle\ ,\ \rangle_F$ is then chosen to be the tensor product metric determined via the natural isomorphism

$$F = \overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, E) \cong \Lambda^{0,1}T^*\dot{\Sigma} \otimes E \cong T\dot{\Sigma} \otimes E,$$

then E admits a natural isomorphism to $\Lambda^{1,0}T^*\dot{\Sigma} \otimes F$ such that

$$-\mathbf{D}^* : \Gamma(F) \to \Gamma(E) \cong \Omega^{1,0}(\dot{\Sigma}, F)$$

becomes an *anti-Cauchy-Riemann* type operator, i.e. it satisfies the Leibniz rule

$$-\mathbf{D}^*(f\lambda) = (\partial f)\lambda + f(-\mathbf{D}^*\lambda)$$

for all $f \in C^\infty(\dot{\Sigma}, \mathbb{R})$, with $\partial f := df - i\, df \circ j \in \Omega^{1,0}(\dot{\Sigma})$. Equivalently, $-\mathbf{D}^*$ defines a Cauchy-Riemann type operator on the **conjugate** bundle $\bar{F} \to \dot{\Sigma}$, defined as the real bundle $F \to \dot{\Sigma}$ but with the sign of its complex structure reversed; we shall distinguish this Cauchy-Riemann operator from $-\mathbf{D}^*$ by writing it as

$$-\bar{\mathbf{D}}^* : \Gamma(\bar{F}) \to \Omega^{0,1}(\dot{\Sigma}, \bar{F}),$$

though it is technically the same operator. The identity map defines a natural complex-antilinear isomorphism between any complex vector bundle and its conjugate bundle; we shall denote this isomorphism generally by

$$E \to \bar{E} : v \mapsto \bar{v},$$

so in particular it satisfies $\overline{cv} = \bar{c}\bar{v}$ for all scalars $c \in \mathbb{C}$, and similarly

$$\bar{\mathbf{D}}^* \bar{\lambda} = \overline{\mathbf{D}^* \lambda}$$

for $\lambda \in \Gamma(F)$. The asymptotic operators for $-\bar{\mathbf{D}}^*$ are

$$\bar{\mathbf{A}}_z = -\mathbf{A}_z : \Gamma(\bar{E}_z) \to \Gamma(\bar{E}_z).$$

LEMMA 5.2.1. *If τ is a choice of asymptotic trivialization on E and $\bar{\tau}$ denotes the* conjugate *asymptotic trivialization,*[5] *then*

$$c_1^{\bar{\tau}}(\bar{E}) = -c_1^{\tau}(E), \qquad \textit{and} \qquad \mu_{CZ}^{\bar{\tau}}(\bar{\mathbf{A}}_z) = -\mu_{CZ}^{\tau}(\mathbf{A}_z) \textit{ for all } z \in \Gamma.$$

PROOF. Assuming E is a line bundle, suppose η is a generic section of E that matches a nonzero constant with respect to τ on the cylindrical ends, so $c_1^{\tau}(E)$ is the signed count of zeroes of η. Then $\bar{\eta} \in \Gamma(\bar{E})$ is similarly a nonzero constant on the ends with respect to $\bar{\tau}$, but the signs of its zeroes are opposite those of η because they are defined as winding numbers with respect to *conjugate* local trivializations. This proves $c_1^{\bar{\tau}}(\bar{E}) = -c_1^{\tau}(E)$.

The Conley-Zehnder indices can be computed from the formula

$$\mu_{\mathrm{CZ}}^{\tau}(\mathbf{A}_z) = \alpha_+^{\tau}(\mathbf{A}_z) + \alpha_-^{\tau}(\mathbf{A}_z),$$

see Theorem 3.7.2. Here $\alpha_-^{\tau}(\mathbf{A}_z)$ is the largest possible winding number relative to τ of an eigenfunction for $\mathbf{A}_z$ with negative eigenvalue, and $\alpha_+^{\tau}(\mathbf{A}_z)$ is the smallest possible winding number with positive eigenvalue. The eigenfunctions of $\bar{\mathbf{A}}_z = -\mathbf{A}_z$ are the same, but the signs of their eigenvalues are reversed, and the signs of their winding numbers are also reversed because they must be measured relative to the conjugate trivialization, thus

$$\alpha_{\pm}^{\bar{\tau}}(\bar{\mathbf{A}}_z) = -\alpha_{\mp}^{\tau}(\mathbf{A}_z),$$

implying

$$\mu_{\mathrm{CZ}}^{\bar{\tau}}(\bar{\mathbf{A}}_z) = \alpha_+^{\bar{\tau}}(\bar{\mathbf{A}}_z) + \alpha_-^{\bar{\tau}}(\bar{\mathbf{A}}_z) = -\alpha_-^{\tau}(\mathbf{A}_z) - \alpha_+^{\tau}(\mathbf{A}_z) = -\mu_{\mathrm{CZ}}^{\tau}(\mathbf{A}_z).$$

The above calculations are all valid for line bundles, but the general case follows by taking direct sums. □

We are now able to show that Theorem 5.1.6 is consistent with what we already know about the formal adjoint.

PROPOSITION 5.2.2. *$I(-\bar{\mathbf{D}}^*) = -I(\mathbf{D})$.*

[5] If $\tau : E|_{\mathcal{U}} \to \mathcal{U} \times \mathbb{C}^n$ is a local trivialization of E with $\tau(v) = (z, w)$, the conjugate trivialization $\bar{\tau} : \bar{E}|_{\mathcal{U}} \to \mathcal{U} \times \mathbb{C}^n$ is defined by $\bar{\tau}(\bar{v}) = (z, \bar{w})$.

PROOF. Under the isomorphism $F \cong \Lambda^{0,1}T^*\dot{\Sigma} \otimes E \cong T\dot{\Sigma} \otimes E$, an asymptotic trivialization τ on E induces an asymptotic trivialization $\partial_s \otimes \tau$ on F, where ∂_s denotes the asymptotic trivialization of $T\dot{\Sigma}$ defined via an outward pointing vector field on the cylindrical ends. Counting zeroes of vector fields then proves $c_1^{\partial_s}(T\dot{\Sigma}) = \chi(\dot{\Sigma})$, so

$$c_1^{\partial_s \otimes \tau}(F) = c_1^{\partial_s \otimes \tau}(T\dot{\Sigma} \otimes E) = n c_1^{\partial_s}(T\dot{\Sigma}) + c_1^{\tau}(E) = n\chi(\dot{\Sigma}) + c_1^{\tau}(E).$$

Applying Lemma 5.2.1 to the conjugate bundle then gives

$$c_1^{\overline{\partial_s \otimes \tau}}(\bar{F}) = -n\chi(\dot{\Sigma}) - c_1^{\tau}(E).$$

The unitary trivializations of the asymptotic bundles $\bar{E}_z$ corresponding to $\overline{\partial_s \otimes \tau}$ are simply $\bar{\tau}$, thus using Lemma 5.2.1 again for the Conley-Zehnder terms,

$$\begin{aligned} I(-\bar{\mathbf{D}}^*) &= n\chi(\dot{\Sigma}) + 2c_1^{\overline{\partial_s \otimes \tau}}(\bar{F}) + \sum_{z \in \Gamma^+} \mu_{\mathrm{CZ}}^{\bar{\tau}}(\bar{\mathbf{A}}_z) - \sum_{z \in \Gamma^-} \mu_{\mathrm{CZ}}^{\bar{\tau}}(\bar{\mathbf{A}}_z) \\ &= -n\chi(\dot{\Sigma}) - 2c_1^{\tau}(E) - \sum_{z \in \Gamma^+} \mu_{\mathrm{CZ}}^{\tau}(\mathbf{A}_z) + \sum_{z \in \Gamma^-} \mu_{\mathrm{CZ}}^{\tau}(\mathbf{A}_z) \\ &= -I(\mathbf{D}). \end{aligned}$$

□

We next consider the effect of an antilinear zeroth-order perturbation on the formal adjoint. By "antilinear zeroth-order perturbation," we generally mean a smooth section

$$B \in \Gamma(\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F)).$$

It is perhaps easier to understand B in terms of the conjugate bundle $\bar{E}$: indeed, there exists a unique

$$\beta \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F))$$

such that

$$B\eta = \beta\bar{\eta},$$

and this correspondence defines a bundle isomorphism $\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F) \cong \mathrm{Hom}_{\mathbb{C}}(\bar{E}, F)$.

EXERCISE 5.2.3. Assume X and Y are complex vector bundles over the same base.

(a) Show that $\bar{X} \otimes \bar{Y}$ is canonically isomorphic to the conjugate bundle of $X \otimes Y$.
(b) Show that $\mathrm{Hom}_{\mathbb{C}}(\bar{X}, \bar{Y})$ is canonically isomorphic to the conjugate bundle of $\mathrm{Hom}_{\mathbb{C}}(X, Y)$, and $\overline{\mathrm{Hom}}_{\mathbb{C}}(\bar{X}, \bar{Y})$ is canonically isomorphic to the conjugate bundle of $\overline{\mathrm{Hom}}_{\mathbb{C}}(X, Y)$.
(c) Show that $\Lambda^{0,1}X^* := \overline{\mathrm{Hom}}_{\mathbb{C}}(X, \mathbb{C})$ is canonically isomorphic to the conjugate bundle of $\Lambda^{1,0}X^* := \mathrm{Hom}_{\mathbb{C}}(X, \mathbb{C})$.

Define the Cauchy-Riemann type operator

$$\mathbf{D}_B := \mathbf{D} + B : \Gamma(E) \to \Gamma(F) = \Omega^{0,1}(\dot{\Sigma}, E),$$

so $\mathbf{D}_B\eta = \mathbf{D}\eta + \beta\bar{\eta}$. To write down $\mathbf{D}_B^*$, observe that since $\beta : \bar{E} \to F$ is a complex-linear bundle map between Hermitian bundles, it has a complex-linear adjoint

$$\beta^\dagger : F \to \bar{E} \quad \text{such that} \quad \langle \beta^\dagger\lambda, \bar{\eta} \rangle_{\bar{E}} = \langle \lambda, \beta\bar{\eta} \rangle_F \text{ for } \lambda \in F,\ \bar{\eta} \in \bar{E}.$$

Here the bundle metric on $\bar{E}$ is defined by $\langle \bar{\eta}, \bar{\xi} \rangle_{\bar{E}} := \langle \xi, \eta \rangle_E$. We then have

$$\begin{aligned}\operatorname{Re}\langle \lambda, B\eta \rangle_F = \operatorname{Re}\langle \lambda, \beta\bar{\eta} \rangle_F = \operatorname{Re}\langle \beta^\dagger \lambda, \bar{\eta} \rangle_{\bar{E}} = \operatorname{Re}\langle \eta, \overline{\beta^\dagger \lambda} \rangle_E = \operatorname{Re}\langle \overline{\beta^\dagger \lambda}, \eta \rangle_E \\ = \operatorname{Re}\langle \bar{\beta}^\dagger \bar{\lambda}, \eta \rangle_E,\end{aligned}$$

where $\bar{\beta}^\dagger \in \Gamma(\operatorname{Hom}_{\mathbb{C}}(\bar{F}, E))$ denotes the image of $\beta^\dagger \in \Gamma(\operatorname{Hom}_{\mathbb{C}}(F, \bar{E}))$ under the complex-antilinear identity map from $\operatorname{Hom}_{\mathbb{C}}(F, \bar{E})$ to its conjugate bundle (see Exercise 5.2.3). The formal adjoint of $\mathbf{D}_B$ is thus

$$\mathbf{D}_B^* = \mathbf{D}^* + B^* : \Gamma(F) \to \Gamma(E),$$

where $B^* : F \to E$ is defined by

$$B^* \lambda := \bar{\beta}^\dagger \bar{\lambda}.$$

To write down the resulting Cauchy-Riemann type operator on $\bar{F}$, we replace $B^* : F \to E$ with $\bar{B}^* : \bar{F} \to \bar{E}$, defined by

$$\bar{B}^* \bar{\lambda} := \overline{B^* \lambda} = \beta^\dagger \lambda,$$

giving a Cauchy-Riemann operator

$$-\bar{\mathbf{D}}_B^* = -\bar{\mathbf{D}}^* + (-\bar{B}^*) : \Gamma(\bar{F}) \to \Gamma(\bar{E}) = \Omega^{0,1}(\dot{\Sigma}, \bar{F}).$$

The point of writing down this formula is to make the following observations:

LEMMA 5.2.4. *The zeroth-order perturbation $-\bar{B}^* : \bar{F} \to \bar{E}$ appearing in $-\bar{\mathbf{D}}_B^*$ has the following properties:*

(1) $-\bar{B}^* : \bar{F} \to \bar{E}$ is complex antilinear;
(2) There is a natural complex bundle isomorphism $\overline{\operatorname{Hom}_{\mathbb{C}}}(\bar{F}, \bar{E}) \cong \operatorname{Hom}_{\mathbb{C}}(F, \bar{E})$ that identifies $-\bar{B}^*$ with $-\beta^\dagger$;
(3) If $n = 1$ and $B \in \Gamma(\overline{\operatorname{Hom}_{\mathbb{C}}}(E, F))$ has only nondegenerate zeroes, then $-\bar{B}^* \in \Gamma(\overline{\operatorname{Hom}_{\mathbb{C}}}(\bar{F}, \bar{E}))$ has the same zeroes but with opposite signs.

PROOF. The first two statements follow immediately from the fact that $-\bar{B}^*$ is the composition of the canonical conjugation map $\bar{F} \to F$ with the complex-linear bundle map $-\beta^\dagger : F \to \bar{E}$. For the third, it suffices to compare what $\beta \in \Gamma(\operatorname{Hom}_{\mathbb{C}}(\bar{E}, F))$ and $-\beta^\dagger : \Gamma(\operatorname{Hom}_{\mathbb{C}}(F, \bar{E}))$ look like in local trivializations near a zero: one is minus the complex conjugate of the other, hence their zeroes count with opposite signs. □

5.3. The index zero case on a torus

As a warmup for the general case, we now fill in the details of Taubes's proof of Theorem 5.1.6 in the case

$$\dot{\Sigma} = \mathbb{T}^2 := \mathbb{C}/(\mathbb{Z} \oplus i\mathbb{Z})$$

and $E = \mathbb{T}^2 \times \mathbb{C}$, i.e. a trivial line bundle. In this case $I(\mathbf{D}) = \chi(\mathbb{T}^2) + 2c_1(E) = 0$, so our aim is to prove $\operatorname{ind}(\mathbf{D}) = 0$. What we will show in fact is that $\mathbf{D}$ is homotopic through a continuous family of Fredholm operators to one that is an isomorphism. Since E and F are now both trivial, it will suffice to consider the operator

$$\mathbf{D} := \bar{\partial} = \partial_s + i\partial_t : H^1(\mathbb{T}^2, \mathbb{C}) \to L^2(\mathbb{T}^2, \mathbb{C}),$$

whose formal adjoint is $\mathbf{D}^* := -\partial = -\partial_s + i\partial_t$. An antilinear zeroth-order perturbation is then equivalent to a choice of function $\beta : \mathbb{T}^2 \to \mathbb{C}$, giving rise to a family of operators

$$\mathbf{D}_r\eta := \bar{\partial}\eta + r\beta\bar{\eta}$$

for $r \in \mathbb{R}$, where $\bar{\eta} : \mathbb{T}^2 \to \mathbb{C}$ now denotes the straightforward complex conjugate of η. Let us assume that $\beta : \mathbb{T}^2 \to \mathbb{C}$ is nowhere zero; note that this would not be possible in more general situations, but is possible here because $\operatorname{Hom}_{\mathbb{C}}(\bar{E}, F)$ is a trivial bundle.

LEMMA 5.3.1. *$\mathbf{D}_r$ is injective for all $r > 0$ sufficiently large.*

PROOF. Elliptic regularity implies that any $\eta \in \ker \mathbf{D}_r$ is smooth, so we shall restrict our attention to smooth functions $\eta : \mathbb{T}^2 \to \mathbb{C}$. We start by comparing the two second-order differential operators

$$\mathbf{D}^*\mathbf{D} \text{ and } \mathbf{D}_r^*\mathbf{D}_r : C^\infty(\mathbb{T}^2, \mathbb{C}) \to C^\infty(\mathbb{T}^2, \mathbb{C}).$$

Both are nonnegative L^2-symmetric operators, and in fact, the first is simply the Laplacian

$$\mathbf{D}^*\mathbf{D} = -\partial\bar{\partial} = (-\partial_s + i\partial_t)(\partial_s + i\partial_t) = -\partial_s^2 - \partial_t^2 = -\Delta.$$

The formal adjoint of $\mathbf{D}_r$ takes the form

$$\mathbf{D}_r^*\eta = \mathbf{D}^*\eta + rB^*\eta = \mathbf{D}^*\eta + r\beta\bar{\eta},$$

thus for any $\eta \in C^\infty(\mathbb{T}^2, \mathbb{C})$,

$$\begin{aligned}
\mathbf{D}_r^*\mathbf{D}_r\eta &= (\mathbf{D}^* + rB^*)(\mathbf{D} + rB)\eta \\
&= \mathbf{D}^*\mathbf{D}\eta + r\left(\beta\overline{\bar{\partial}\eta} - \partial(\beta\bar{\eta})\right) + r^2B^*B\eta \\
&= \mathbf{D}^*\mathbf{D}\eta + r\left(\beta\partial\bar{\eta} - (\partial\beta)\bar{\eta} - \beta\partial\bar{\eta}\right) + r^2B^*B\eta \\
&= \mathbf{D}^*\mathbf{D}\eta + r^2B^*B\eta - r(\partial\beta)\bar{\eta}.
\end{aligned} \tag{5.2}$$

This is a *Weitzenböck formula*: its main message is that the Laplacian $\mathbf{D}^*\mathbf{D}$ and the related operator $\mathbf{D}_r^*\mathbf{D}_r$ differ from each other only by a zeroth-order term that will be positive definite if r is sufficiently large. Indeed, since β is nowhere zero, we have $|B\eta| \geqslant c|\eta|$ for some constant $c > 0$, thus

$$\begin{aligned}
\|\mathbf{D}_r\eta\|_{L^2}^2 &= \langle\eta, \mathbf{D}_r^*\mathbf{D}_r\eta\rangle_{L^2} = \langle\eta, \mathbf{D}^*\mathbf{D}\eta\rangle_{L^2} + r^2\langle\eta, B^*B\eta\rangle_{L^2} - r\langle\eta, (\partial\beta)\bar{\eta}\rangle_{L^2} \\
&= \|\mathbf{D}\eta\|_{L^2}^2 + r^2\|B\eta\|_{L^2}^2 - r\langle\eta, (\partial\beta)\bar{\eta}\rangle_{L^2} \\
&\geqslant \left(r^2c^2 - r\|\partial\beta\|_{C^0}\right)\|\eta\|_{L^2}^2.
\end{aligned}$$

We conclude that as soon as $r > 0$ is large enough to make the quantity in parentheses positive, $\mathbf{D}_r\eta$ cannot vanish unless $\|\eta\|_{L^2} = 0$. □

PROOF OF THEOREM 5.1.6 FOR $E = \mathbb{T}^2 \times \mathbb{C}$. The lemma above shows that one can add a large antilinear perturbation to $\mathbf{D} = \bar{\partial}$ making the deformed operator $\mathbf{D}_r$ injective. By Lemma 5.2.4, the same argument applies to the formal adjoint $\mathbf{D}^*$, implying that for sufficiently large $r > 0$, $\mathbf{D}_r^*$ is injective and thus $\mathbf{D}_r$ is also surjective, and therefore an isomorphism. This proves $\operatorname{ind}(\mathbf{D}) = \operatorname{ind}(\mathbf{D}_r) = 0$. □

Let's consider which particular details of the setup made the proof above possible. First, the zeroth-order perturbation is complex antilinear. We used this, if only implicitly, in deriving the Weitzenböck formula (5.2): the key step is in the third line, where the two terms involving $\partial\bar{\eta}$ cancel each other out and leave nothing but zeroth-order terms remaining. This would not have happened if e.g. $B : E \to F$ had been complex linear—we would then have seen terms depending on the first derivative of η in $\mathbf{D}_r^*\mathbf{D}_r\eta - \mathbf{D}^*\mathbf{D}\eta$, and this would have killed the whole argument. The fact that this cancelation happens when the perturbation is antilinear probably looks like magic at this point, but there is a principle behind it; we will discuss it further in §5.4 below, see Remark 5.4.4.

The second crucial fact we used was that $\beta : \mathbb{T}^2 \to \mathbb{C}$ is nowhere zero, in order to obtain the lower bound on $\|B\eta\|_{L^2}$ in terms of $\|\eta\|_{L^2}$. This cannot always be achieved—it is possible in this special case only because E and F are both trivial bundles and thus so is $\operatorname{Hom}_{\mathbb{C}}(\bar{E}, F)$. On more general bundles, the best we could hope for would be to pick $\beta \in \Gamma(\operatorname{Hom}_{\mathbb{C}}(\bar{E}, F))$ with finitely many zeroes, all nondegenerate. In this case the above argument fails, but it still tells us something. Suppose $\Sigma_\epsilon \subset \mathbb{T}^2$ is a region disjoint from the isolated zeroes of β. Then there exists a constant $c_\epsilon > 0$, dependent on the region Σ_ϵ, such that

$$\|\beta\bar{\eta}\|^2_{L^2(\mathbb{T}^2)} \geqslant \|\beta\bar{\eta}\|^2_{L^2(\Sigma_\epsilon)} \geqslant c_\epsilon\|\eta\|^2_{L^2(\Sigma_\epsilon)},$$

so instead of the estimate at the end of the proof above implying $\mathbf{D}_r$ is injective, we obtain one of the form

$$\|\mathbf{D}_r\eta\|^2_{L^2(\mathbb{T}^2)} \geqslant c_\epsilon r^2\|\eta\|_{L^2(\Sigma_\epsilon)} - cr\|\eta\|^2_{L^2(\mathbb{T}^2)}.$$

To see what this means, imagine we have sequences $r_\nu \to \infty$ and $\eta_\nu \in \ker \mathbf{D}_{r_\nu}$, normalized so that $\|\eta_\nu\|_{L^2} = 1$ for all ν. The estimate above then implies

$$\|\eta_\nu\|^2_{L^2(\Sigma_\epsilon)} \leqslant \frac{c}{c_\epsilon r_\nu} \to 0 \quad \text{as} \quad \nu \to \infty,$$

so while all sections η_ν have the same amount of "energy" (as measured via their L^2-norms), the energy is escaping from Σ_ϵ as r_ν increases. This is true for *any* domain Σ_ϵ disjoint from the zeroes, so we conclude that in the limit as $r \to \infty$, sections in $\ker \mathbf{D}_r$ have their energy concentrated in infinitesimally small neighborhoods of the zeroes of β. We will see in the following how to extract useful information from this concentration of energy.

5.4. A Weitzenböck formula for Cauchy-Riemann operators

The Weitzenböck formula (5.2) can be generalized to a useful relation between any two Cauchy-Riemann type operators that differ by an *antilinear* zeroth-order term. To see this, we start with a short digression on holomorphic and antiholomorphic vector bundles.

A smooth function $f : \mathbb{C} \supset \mathcal{U} \to \mathbb{C}$ is called **antiholomorphic** if it satisfies $(\partial_s - i\partial_t)f = 0$, which means its differential anticommutes with the complex structure on $\mathbb{C}$. The class of antiholomorphic functions is not closed under composition, but it is closed under products, hence one can define an **antiholomorphic structure** on a complex vector bundle to be a system of local trivializations for which all transition maps are antiholomorphic. Given the standard correspondence between holomorphic structures and Cauchy-Riemann type operators (see §2.5), it is easy to establish a similar correspondence between antiholomorphic structures and (complex-linear) **anti-Cauchy-Riemann type** operators, i.e. those which satisfy

$$\mathbf{D}(f\eta) = (\partial f)\eta + f\mathbf{D}\eta$$

for all $f \in C^\infty(\dot{\Sigma}, \mathbb{C})$, where $\partial f := df - i\, df \circ j \in \Omega^{1,0}(\dot{\Sigma})$. We've seen one important example of such an operator already: if $\mathbf{D} : \Gamma(E) \to \Gamma(F)$ is complex linear, then $-\mathbf{D}^*$ can be interpreted as a complex-linear anti-Cauchy-Riemann operator on F, and thus endows F with an antiholomorphic structure. Another example occurs naturally on conjugate bundles: if E has a holomorphic structure, then $\bar{E}$ inherits from this an antiholomorphic structure. This is immediate from the fact that $f : \mathbb{C} \supset \mathcal{U} \to \mathbb{C}$ is holomorphic if and only if $\bar{f} : \mathcal{U} \to \mathbb{C}$ is antiholomorphic. If $\mathbf{D} : \Gamma(E) \to \Gamma(F) = \Omega^{0,1}(\dot{\Sigma}, E)$ is the corresponding complex-linear Cauchy-Riemann type operator on E, we shall denote the resulting anti-Cauchy-Riemann operator by

$$\bar{\mathbf{D}} : \Gamma(\bar{E}) \to \Gamma(\bar{F}) = \Omega^{1,0}(\dot{\Sigma}, \bar{E}),$$

where by definition $\bar{\mathbf{D}}\bar{\eta} = \overline{\mathbf{D}\eta}$.

Exercise 5.4.1. Show that if X and Y are antiholomorphic vector bundles over the same base, then $X \otimes Y$ and $\mathrm{Hom}_{\mathbb{C}}(X, Y)$ both naturally inherit antiholomorphic bundle structures such that the obvious Leibniz rules are satisfied. *Remark: The proof of this is exactly the same as for holomorphic bundles, one only needs to change some signs.*

The next result is the main tool needed for our proof of the index formula.

Proposition 5.4.2. *Assume $E \to \dot{\Sigma}$ is an asymptotically Hermitian line bundle, $\mathbf{D} : \Gamma(E) \to \Gamma(F) = \Omega^{0,1}(\Sigma, E)$ is a linear Cauchy-Riemann type operator C^0-asymptotic to asymptotic operators $\{\mathbf{A}_z\}_{z \in \Gamma}$ at the punctures, and $B : E \to F$ is a complex-antilinear bundle map. We consider the family of Cauchy-Riemann type operators*

$$\mathbf{D}_r := \mathbf{D} + rB : \Gamma(E) \to \Gamma(F) \quad \textit{for} \quad r \in \mathbb{R},$$

and denote by $\mathbf{D}_r^* = \mathbf{D}^* + rB^* : \Gamma(F) \to \Gamma(E)$ their formal adjoints with respect to fixed choices of area forms and bundle metrics compatible with the asymptotically Hermitian structure of E. Then there exists a real-linear bundle map $B_1 : E \to E$ such that for all $r \in \mathbb{R}$ and $\eta \in \Gamma(E)$,

$$\mathbf{D}_r^*\mathbf{D}_r\eta = \mathbf{D}^*\mathbf{D}\eta + r^2 B^* B\eta + rB_1\eta.$$

Moreover, if B is C^1-bounded as a section of $\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F)$, then B_1 is C^0-bounded as a section of $\mathrm{End}_{\mathbb{R}}(E)$.

Proof. We consider first the case where $\mathbf{D}$ is complex linear. The operators $\bar{\mathbf{D}}$ and $-\mathbf{D}^*$ are then complex-linear anti-Cauchy-Riemann operators on $\bar{E}$ and F respectively, so as a corollary of the linear local existence result in §2.5, they determine antiholomorphic vector bundle structures on $\bar{E}$ and F. By Exercise 5.4.1, these induce an antiholomorphic vector bundle structure on $\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F)$, giving rise to a complex-linear anti-Cauchy-Riemann operator ∂_H on $\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F)$ that satisfies the Leibniz rule

$$-\mathbf{D}^*(\Phi\bar{\eta}) = (\partial_H \Phi)\bar{\eta} + \Phi(\bar{\mathbf{D}}\bar{\eta}) \quad \text{for all} \quad \bar{\eta} \in \Gamma(\bar{E}),\ \Phi \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F)).$$

Writing $B\eta = \beta\bar{\eta}$ and $B^*\lambda = \bar{\beta}^\dagger\bar{\lambda}$ for $\beta \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F))$ and its complex adjoint $\beta^\dagger \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(F, \bar{E}))$, we have

$$\begin{aligned}
\mathbf{D}_r^*\mathbf{D}_r\eta &= (\mathbf{D}^* + rB^*)(\mathbf{D} + rB)\eta \\
&= \mathbf{D}^*\mathbf{D}\eta + r\bar{\beta}^\dagger\overline{\mathbf{D}\eta} - r(-\mathbf{D}^*)(\beta\bar{\eta}) + r^2 B^*B\eta \\
&= \mathbf{D}^*\mathbf{D}\eta + r\bar{\beta}^\dagger\bar{\mathbf{D}}\bar{\eta} - r(\partial_H\beta)\bar{\eta} - r\beta\bar{\mathbf{D}}\bar{\eta} + r^2 B^*B\eta \\
&= \mathbf{D}^*\mathbf{D}\eta + r^2 B^*B\eta - r(\partial_H\beta)\bar{\eta} + r\left(\bar{\beta}^\dagger - \beta\right)\bar{\mathbf{D}}\bar{\eta}.
\end{aligned}$$

Here β and $\bar{\beta}^\dagger$ are both viewed as complex-linear bundle maps $\bar{F} \to E$, the latter in the obvious way, and the former acting as $\mathbb{1} \otimes \beta$ on $\bar{F} = \Lambda^{1,0}T^*\dot{\Sigma} \otimes \bar{E}$ with target $\Lambda^{1,0}T^*\dot{\Sigma} \otimes F = \Lambda^{1,0}T^*\dot{\Sigma} \otimes \Lambda^{0,1}T^*\dot{\Sigma} \otimes E = E$. Choosing unitary local trivializations, β and $\bar{\beta}^\dagger$ are represented by the same complex-valued function: indeed, the latter is the transpose of the former as n-by-n complex matrices, but since $n = 1$, this means they are identical, and the last term in the formula above therefore vanishes, leaving

$$\mathbf{D}_r^*\mathbf{D}_r\eta = \mathbf{D}^*\mathbf{D}\eta + r^2 B^*B\eta - r(\partial_H\beta)\bar{\eta}. \tag{5.3}$$

If $\mathbf{D}$ is not complex linear, then we define its complex-linear part $\mathbf{D}_{\mathbb{C}} : \Gamma(E) \to \Gamma(F)$ by

$$\mathbf{D}_{\mathbb{C}}\eta := \frac{1}{2}\left(\mathbf{D}\eta - J\mathbf{D}(J\eta)\right)$$

and observe that this also satisfies the Leibniz rule $\mathbf{D}_{\mathbb{C}}(f\eta) = (\bar{\partial}f)\eta + f\mathbf{D}\eta$ for all $f \in C^\infty(\dot{\Sigma})$, so it is a complex-linear Cauchy-Riemann type operator and $\mathbf{D} = \mathbf{D}_{\mathbb{C}} + A$ for some complex-antilinear bundle map $A : E \to F$. Writing $A\eta := \alpha\bar{\eta}$ for $\alpha \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F))$, we can then apply (5.3) to both $\mathbf{D} = \mathbf{D}_{\mathbb{C}} + A$ and $\mathbf{D}_r = \mathbf{D}_{\mathbb{C}} + (A + rB)$, giving

$$\mathbf{D}^*\mathbf{D} - \mathbf{D}_{\mathbb{C}}^*\mathbf{D}_{\mathbb{C}} = A^*A\eta - (\partial_H\alpha)\bar{\eta}$$

and

$$\mathbf{D}_r^*\mathbf{D}_r - \mathbf{D}_{\mathbb{C}}^*\mathbf{D}_{\mathbb{C}} = (A + rB)^*(A + rB)\eta - (\partial_H\alpha)\bar{\eta} - r(\partial_H\beta)\bar{\eta}.$$

Subtracting the first relation from the second gives

$$\mathbf{D}_r^*\mathbf{D}_r - \mathbf{D}^*\mathbf{D} = r^2 B^*B + r\left[(A^*B + B^*A)\eta - (\partial_H\beta)\bar{\eta}\right],$$

so we can define $B_1\eta$ as the expression in brackets at the right.

Concerning bounds on $\|B_1\|_{C^0}$: choose an asymptotic trivialization on the cylindrical end $Z_\pm \cong \dot{\mathcal{U}}_z$ near one of the punctures z, identifying $\mathbf{D}$ on this region with $\bar{\partial} + S : C^\infty(Z_\pm, \mathbb{C}) \to C^\infty(Z_\pm, \mathbb{C})$ for a smooth function $S : Z_\pm \to \mathrm{End}(\mathbb{C})$ which satisfies $\lim_{s\to\pm\infty} S(s,t) = S_\infty(t)$ for a loop $S_\infty : S^1 \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{C})$ determined by the asymptotic operator $\mathbf{A}_z$. The conjugate operator $\bar{\mathbf{D}}$ is then given by $\partial + \bar{S}$ over $Z_\pm$, and since the bundle metrics were assumed compatible with the asymptotically Hermitian structure, we can assume they are standard in our chosen trivialization, so that $\mathbf{D}^*$ becomes identified with $-\partial + S^{\mathrm{T}}$. The antilinear bundle map $B : E \to F$ is identified likewise with a function $B : Z_\pm \to \overline{\mathrm{End}}_{\mathbb{C}}(\mathbb{C})$ of the form $B(s,t)v = \beta(s,t)\bar{v}$ for a function $\beta : Z_\pm \to \mathbb{C}$. The complex-linear part of $\mathbf{D}$ over $Z_\pm$ is given by $\mathbf{D}_{\mathbb{C}} = \bar{\partial} + S_{\mathbb{C}}$, where

$$S_{\mathbb{C}} := \frac{1}{2}\left(S - iSi\right), \qquad \text{hence} \qquad A = \frac{1}{2}(S + iSi).$$

The latter clearly satisfies a global bound on $Z_\pm$ in light of the asymptotic convergence of S to S_∞, thus a C^0-bound on B implies a C^0-bound on $A^*B + B^*A$.

A coordinate formula for $\partial_H\beta$ can be derived from the corresponding formulas for $\mathbf{D}^*$ and $\bar{\mathbf{D}}$ via the Leibniz rule $-\mathbf{D}^*(\beta\bar{\eta}) = (\partial_H\beta)\bar{\eta} + \beta\bar{\mathbf{D}}\bar{\eta}$: indeed,

$$\begin{aligned} -\mathbf{D}^*(\beta\bar{\eta}) &= -(-\partial + S^{\mathrm{T}})(\beta\bar{\eta}) = (\partial - S^{\mathrm{T}})(\beta\bar{\eta}) = (\partial\beta)\bar{\eta} + \beta(\partial\bar{\eta}) - S^{\mathrm{T}}\beta\bar{\eta} \\ &= (\partial_H\beta)\bar{\eta} + \beta\bar{\mathbf{D}}\bar{\eta} = (\partial_H\beta)\bar{\eta} + \beta(\partial + \bar{S})\bar{\eta} = (\partial_H\beta)\bar{\eta} + \beta(\partial\bar{\eta}) + \beta\bar{S}\bar{\eta} \end{aligned}$$

implying

$$\partial_H\beta = \partial\beta + \beta\bar{S} - S^{\mathrm{T}}\beta.$$

This expression is C^0-bounded in terms of the C^1-norm of B. □

REMARK 5.4.3. The above proof used the assumption $n = 1$ in order to conclude $\bar{\beta}^\dagger - \beta \equiv 0$. For higher rank bundles, this imposes a nontrivial condition that must be satisfied in order for the Weitzenböck formula to hold, cf. [GW17].

REMARK 5.4.4. We can now pick out a geometric reason for the miraculous cancelation in the Weitzenböck formula: the perturbation B is described by a complex bundle map $\bar{E} \to F$, where $\bar{E}$ and F both have natural antiholomorphic bundle structures defined via the complex-linear parts of $\bar{\mathbf{D}}$ and $-\mathbf{D}^*$ respectively. A complex-linear perturbation $B : E \to F$ would not work because E is holomorphic rather than antiholomorphic: while $\bar{\mathbf{D}}$ can be fit into the same Leibniz rule with $-\mathbf{D}^*$, the same is not true of $\mathbf{D}$.

5.5. Large antilinear perturbations and energy concentration

We continue in the setting of Proposition 5.4.2 and consider

$$\mathbf{D}_r := \mathbf{D} + rB : \Gamma(E) \to \Gamma(F)$$

for $r \in \mathbb{R}$, where $B\eta = \beta\bar{\eta}$ for a fixed section $\beta \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F))$. After a compact perturbation of $\mathbf{D}$, we can without loss of generality also impose the following assumptions on $\mathbf{D}$, β and the area form $d\,\mathrm{vol}$:

(i) All zeroes of β are nondegenerate.
(ii) Both $|\beta|$ and $1/|\beta|$ are bounded outside of a compact subset of $\dot{\Sigma}$.
(iii) Near each point $\zeta \in \dot{\Sigma}$ with $\beta(\zeta) = 0$, there exists a neighborhood $\mathcal{D}(\zeta) \subset \dot{\Sigma}$ of ζ, a holomorphic coordinate chart identifying $(\mathcal{D}(\zeta), j, \zeta)$ with the unit disk $(\mathbb{D}, i, 0)$, and a local trivialization of E over $\mathcal{D}(\zeta)$ that identifies $\mathbf{D}$ with $\bar{\partial} = \partial_s + i\partial_t : C^\infty(\mathbb{D}, \mathbb{C}) \to C^\infty(\mathbb{D}, \mathbb{C})$ and β with one of the functions

$$\beta(z) = z \qquad \text{or} \qquad \beta(z) = \bar{z},$$

the former if ζ is a positive zero and the latter if it is negative.

(iv) In the holomorphic coordinate on $\mathcal{D}(\zeta)$ described above, $d\,\mathrm{vol}$ is the standard Lebesgue measure.

As in the torus case discussed in §5.3, we will see that the Weitzenböck formula implies a concentration of energy near the zeroes of β for sections $\eta \in \ker \mathbf{D}_r$ as $r \to \infty$. To understand what really happens in this limit, we will use a rescaling trick. Denote the zero set of β by

$$Z(\beta) = Z^+(\beta) \cup Z^-(\beta) \subset \dot{\Sigma},$$

partitioned into the positive and negative zeroes. For any $\eta \in \Gamma(E)$, $\zeta \in Z^\pm(\beta)$ and $r > 0$, we then define a rescaled function

$$\eta^{(\zeta,r)} : \mathbb{D}_{\sqrt{r}} \to \mathbb{C} : z \mapsto \frac{1}{\sqrt{r}}\eta(z/\sqrt{r}),$$

where the right hand side denotes the local representation of η on $\mathcal{D}(\zeta)$ in the chosen coordinate and trivialization. Notice that the equation $\mathbf{D}_r\eta = 0$ appears in this local representation as either

$$\bar{\partial}\eta + rz\bar{\eta} = 0 \quad \text{or} \quad \bar{\partial}\eta + r\bar{z}\bar{\eta} = 0 \qquad \text{on } \mathcal{D}(\zeta), \tag{5.4}$$

depending on the sign of ζ, and the function $f := \eta^{(\zeta,r)}$ then satisfies

$$\bar{\partial} f + z\bar{f} = 0 \quad \text{or} \quad \bar{\partial} f + \bar{z}\bar{f} = 0 \qquad \text{on } \mathbb{D}_{\sqrt{r}}.$$

We will take a closer look at these two PDEs in §5.6 below. But first, observe that by change of variables,

$$\|\eta^{(\zeta,r)}\|_{L^2(\mathbb{D}_{\sqrt{r}})} = \|\eta\|_{L^2(\mathcal{D}(\zeta))}.$$

LEMMA 5.5.1. *Assume $r_\nu \to \infty$, and $\eta_\nu \in \ker \mathbf{D}_{r_\nu}$ is a sequence satisfying a uniform L^2-bound. Then after passing to a subsequence, the rescaled functions $\eta_\nu^\zeta := \eta_\nu^{(\zeta,r_\nu)} : \mathbb{D}_{\sqrt{r_\nu}} \to \mathbb{C}$ for each $\zeta \in Z^\pm(\beta)$ converge in $C^\infty_{\mathrm{loc}}(\mathbb{C})$ to smooth functions $\eta_\infty^\zeta \in L^2(\mathbb{C})$ satisfying*

$$\bar{\partial}\eta_\infty^\zeta + z\overline{\eta_\infty^\zeta} = 0 \quad \text{if } \zeta \in Z^+(\beta),$$
$$\bar{\partial}\eta_\infty^\zeta + \bar{z}\overline{\eta_\infty^\zeta} = 0 \quad \text{if } \zeta \in Z^-(\beta).$$

Moreover, if $\xi_\nu \in \ker \mathbf{D}_{r_\nu}$ is another sequence with these same properties and convergence $\xi_\nu^\zeta \to \xi_\infty^\zeta$, then

$$\lim_{\nu\to\infty} \langle \eta_\nu, \xi_\nu \rangle_{L^2(E)} = \sum_{\zeta \in Z(\beta)} \langle \eta_\infty^\zeta, \xi_\infty^\zeta \rangle_{L^2(\mathbb{C})}.$$

PROOF. The uniform L^2-bound implies uniform bounds on $\|\eta_\nu^\zeta\|_{L^2(\mathbb{D}_R)}$ for every $R > 0$, where ν here is assumed sufficiently large so that $R < \sqrt{r_\nu}$. Since η_ν^ζ satisfies a Cauchy-Riemann type equation on $\mathbb{D}_R$, the usual elliptic estimates (see Chapter 2) then imply uniform H^k-bounds for every $k \in \mathbb{N}$ on every compact subset in the interior of $\mathbb{D}_R$, hence η_ν^ζ has a C^∞_{loc}-convergent subsequence on $\mathbb{C}$, and the limit η_∞^ζ clearly satisfies the stated PDE. The uniform L^2-bound also implies a uniform bound on $\|\eta_\nu^\zeta\|_{L^2(\mathbb{D}_{\sqrt{r_\nu}})}$ and thus an R-independent uniform bound on $\|\eta_\nu^\zeta\|_{L^2(\mathbb{D}_R)}$ as $\nu \to \infty$, implying that η_∞^ζ is in $L^2(\mathbb{C})$.

The limit of $\langle\eta_\nu,\xi_\nu\rangle_{L^2(E)}$ is now proved using the Weitzenböck formula. Let

$$\dot{\Sigma}_\epsilon := \dot{\Sigma} \setminus \bigcup_{\zeta\in Z(\beta)} \mathcal{D}(\zeta),$$

so there exists a constant $c > 0$ such that β satisfies $|\beta(z)\bar{v}| \geqslant c|v|$ for all $v \in E_z$, $z \in \dot{\Sigma}_\epsilon$. (Note that this depends on the assumption of $1/|\beta|$ being bounded outside of a compact subset.) Now by Proposition 5.4.2,

$$\begin{aligned}
0 = \|\mathbf{D}_{r_\nu}\eta_\nu\|^2_{L^2(\dot{\Sigma})} &= \langle\eta_\nu, \mathbf{D}^*_{r_\nu}\mathbf{D}_{r_\nu}\eta_\nu\rangle_{L^2(\dot{\Sigma})} \\
&= \langle\eta_\nu, \mathbf{D}^*\mathbf{D}\eta_\nu\rangle_{L^2(\dot{\Sigma})} + r_\nu^2\langle\eta_\nu, B^*B\eta_\nu\rangle_{L^2(\dot{\Sigma})} + r_\nu\langle\eta_\nu, B_1\eta_\nu\rangle_{L^2(\dot{\Sigma})} \\
&\geqslant \|\mathbf{D}\eta_\nu\|^2_{L^2(\dot{\Sigma})} + r_\nu^2 c^2\|\eta_\nu\|^2_{L^2(\dot{\Sigma}_\epsilon)} - r_\nu c'\|\eta_\nu\|^2_{L^2(\dot{\Sigma})} \\
&\geqslant r_\nu^2 c^2\|\eta_\nu\|^2_{L^2(\dot{\Sigma}_\epsilon)} - r_\nu c'\|\eta_\nu\|^2_{L^2(\dot{\Sigma})}
\end{aligned}$$

for some constant $c' > 0$ independent of ν. This implies

$$\|\eta_\nu\|^2_{L^2(\dot{\Sigma}_\epsilon)} \leqslant \frac{c'}{c^2 r_\nu}\|\eta_\nu\|^2_{L^2(\dot{\Sigma})} \to 0 \quad \text{as} \quad \nu\to\infty$$

since $\|\eta_\nu\|_{L^2(\dot{\Sigma})}$ is uniformly bounded. The same estimate applies to ξ_ν, so that $\langle\eta_\nu,\xi_\nu\rangle_{L^2(\dot{\Sigma}_\epsilon)} \to 0$ and thus by change of variables,

$$\begin{aligned}
\lim_{\nu\to\infty}\langle\eta_\nu,\xi_\nu\rangle_{L^2(\dot{\Sigma})} &= \lim_{\nu\to\infty}\sum_{\zeta\in Z(\beta)}\langle\eta_\nu,\xi_\nu\rangle_{L^2(\mathcal{D}(\zeta))} = \lim_{\nu\to\infty}\sum_{\zeta\in Z(\beta)}\langle\eta_\nu^\zeta,\xi_\nu^\zeta\rangle_{L^2(\mathbb{D}_{\sqrt{r_\nu}})} \\
&= \sum_{\zeta\in Z(\beta)}\langle\eta_\infty^\zeta,\xi_\infty^\zeta\rangle_{L^2(\mathbb{C})}.
\end{aligned}$$

□

5.6. Two Cauchy-Riemann type problems on the plane

The rescaling trick in the previous section produced smooth solutions $f : \mathbb{C} \to \mathbb{C}$ of class $L^2(\mathbb{C})$ to the two equations

$$\bar{\partial} f + z\bar{f} = 0, \qquad \bar{\partial} f + \bar{z}\bar{f} = 0.$$

It turns out that we can say precisely what all such solutions are. Write $\mathbf{D}_+ f := \bar{\partial} f + z\bar{f}$ and $\mathbf{D}_- f := \bar{\partial} f + \bar{z}\bar{f}$. Both operators differ from the complex-linear operator $\bar{\partial}$ by antilinear perturbations, so they satisfy Weitzenböck formulas relating $\mathbf{D}^*_\pm\mathbf{D}_\pm$ to the Laplacian $-\Delta = \bar{\partial}^*\bar{\partial} = -\partial_s^2 - \partial_t^2$. Indeed, applying (5.3) in these special cases gives

$$\mathbf{D}^*_+\mathbf{D}_+ f = -\Delta f + |z|^2 f - 2\bar{f} \qquad \text{and} \qquad \mathbf{D}^*_-\mathbf{D}_- f = -\Delta f + |z|^2 f.$$

To make use of this, recall that a smooth function $u : \mathcal{U} \to \mathbb{R}$ on an open subset $\mathcal{U} \subset \mathbb{C}$ is called **subharmonic** if it satisfies

$$-\Delta u \leqslant 0.$$

Subharmonic functions satisfy a **mean value property**:

$$-\Delta u \leqslant 0 \text{ on } \mathcal{U} \qquad \Rightarrow \qquad u(z_0) \leqslant \frac{1}{\pi r^2}\int_{\mathbb{D}_r(z_0)} u(z)\,d\mu(z) \quad \text{for all} \quad \mathbb{D}_r(z_0) \subset \mathcal{U},$$

where $\mathbb{D}_r(z_0) \subset \mathbb{C}$ denotes the disk of radius $r > 0$ about a point $z_0 \in \mathcal{U}$, and $d\mu(z)$ is the Lebesgue measure on $\mathbb{C}$; see e.g. [Eva98, p. 85].

EXERCISE 5.6.1. Show that for any smooth complex-valued function f on an open subset of $\mathbb{C}$,

$$\Delta|f|^2 = 2\operatorname{Re}\langle f, \Delta f\rangle + 2|\nabla f|^2,$$

where $\langle\ ,\ \rangle$ denotes the standard Hermitian inner product on $\mathbb{C}$ and $|\nabla f|^2 := |\partial_s f|^2 + |\partial_t f|^2$.

PROPOSITION 5.6.2. *The equation $\bar{\partial} f + \bar{z}\bar{f} = 0$ does not admit any nontrivial smooth solutions $f \in L^2(\mathbb{C}, \mathbb{C})$.*

PROOF. If $f : \mathbb{C} \to \mathbb{C}$ is smooth with $\mathbf{D}_- f = 0$, the Weitzenböck formula for $\mathbf{D}_-$ implies $\Delta f = |z|^2 f$. Then by Exercise 5.6.1,

$$\Delta|f|^2 = 2\operatorname{Re}\langle f, |z|^2 f\rangle + 2|\nabla f|^2 = 2|z|^2|f|^2 + 2|\nabla f|^2,$$

implying that $|f|^2 : \mathbb{C} \to \mathbb{R}$ is subharmonic. Now if $f(z_0) \neq 0$ for some $z_0 \in \mathbb{C}$, the mean value property implies

$$\int_{\mathbb{D}_r(z_0)} |f(z)|^2 \, d\mu(z) \geqslant \pi r^2 |f(z_0)|^2 \to \infty \quad \text{as} \quad r \to \infty,$$

so $f \notin L^2(\mathbb{C})$. □

PROPOSITION 5.6.3. *Every smooth solution $f \in L^2(\mathbb{C}, \mathbb{C})$ to the equation $\bar{\partial} f + z\bar{f} = 0$ is a constant real multiple of $f_0(z) := e^{-\frac{1}{2}|z|^2}$.*

PROOF. We claim first that every smooth solution in $L^2(\mathbb{C}, \mathbb{C})$ of $\mathbf{D}_+ f = 0$ is purely real valued. The Weitzenböck formula for this case gives $\Delta f = |z|^2 f - 2\bar{f}$, and taking the difference between this equation and its complex conjugate then implies that $u := \operatorname{Im} f : \mathbb{C} \to \mathbb{R}$ satisfies

$$\Delta u = (|z|^2 + 2)u.$$

Now by Exercise 5.6.1,

$$\Delta(u^2) = 2|\nabla u|^2 + 2(|z|^2 + 2)u^2 \geqslant 0,$$

so $u^2 : \mathbb{C} \to \mathbb{R}$ is subharmonic, and the mean value property implies as in the proof of Prop. 5.6.2 that $u \notin L^2(\mathbb{C})$ and hence $f \notin L^2(\mathbb{C})$ unless $u \equiv 0$. This proves the claim.

It is easy to check however that f_0 is a solution and is in $L^2(\mathbb{C})$. Since it is also nowhere zero, every other solution f must then take the form $f(z) = v(z) f_0(z)$ for some *real-valued* function $v : \mathbb{C} \to \mathbb{R}$. Since $\mathbf{D}_+$ is a Cauchy-Riemann type operator, the Leibniz rule then implies $\bar{\partial} v \equiv 0$. But the only globally holomorphic functions with trivial imaginary parts are constant. □

5.7. A linear gluing argument

Now we're getting somewhere.

LEMMA 5.7.1. *Suppose the assumptions of §5.5 hold and* $\beta \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F))$ *has* $I_+ \geqslant 0$ *positive and* $I_- \geqslant 0$ *negative zeroes. Then for all* $r > 0$ *sufficiently large,*

$$\dim \ker \mathbf{D}_r \leqslant I_+ \qquad \text{and} \qquad \dim \operatorname{coker} \mathbf{D}_r \leqslant I_-.$$

In particular, for sufficiently large r, $\mathbf{D}_r$ *is injective if all zeroes of* β *are negative and surjective if all zeroes are positive.*

PROOF. Arguing by contradiction, suppose there exists a sequence $r_\nu \to \infty$ such that $\dim \ker \mathbf{D}_{r_\nu} > I_+$, and pick $(I_+ + 1)$ sequences of sections $\eta^1_\nu, \dots, \eta^{I_+ + 1}_\nu \in \ker \mathbf{D}_{r_\nu}$ which form L^2-orthonormal sets for each ν. By Lemma 5.5.1, we can then extract a subsequence such that rescaling near the zeroes of β produces C^∞_{loc}-convergent sequences whose limits form an $(I_+ + 1)$-dimensional orthonormal set in

$$\bigoplus_{\zeta \in Z(\beta)} L^2(\mathbb{C}, \mathbb{C}),$$

where the component functions $f \in L^2(\mathbb{C}, \mathbb{C})$ for $\zeta \in Z^+(\beta)$ satisfy $\bar{\partial} f + z\bar{f} = 0$, while those for $\zeta \in Z^-(\beta)$ satisfy $\bar{\partial} f + \bar{z}\bar{f} = 0$. Proposition 5.6.2 now implies that the component functions for $\zeta \in Z^-(\beta)$ are all trivial, and by Proposition 5.6.3, the components for $\zeta \in Z^+(\beta)$ belong to 1-dimensional subspaces $\ker \mathbf{D}_+ \subset L^2(\mathbb{C})$ generated by the function $e^{-\frac{1}{2}|z|^2}$. We conclude that the limiting orthonormal set lives in a precisely I_+-dimensional subspace

$$\bigoplus_{\zeta \in Z^+(\beta)} \ker \mathbf{D}_+ \subset \bigoplus_{\zeta \in Z(\beta)} L^2(\mathbb{C}, \mathbb{C}),$$

and this is a contradiction since there are $I_+ + 1$ elements in the set.

Applying the same argument to the formal adjoint implies similarly $\dim \ker \mathbf{D}^*_r \leqslant I_-$ for r sufficiently large. ☐

We would next like to turn the two inequalities in the above lemma into equalities, which means showing that the I_+-dimensional subspace of $\bigoplus_{\zeta \in Z^+(\beta)} L^2(\mathbb{C}, \mathbb{C})$ generated by solutions of $\bar{\partial} f + z\bar{f} = 0$ is isomorphic to $\ker \mathbf{D}_r$ for r sufficiently large. This requires a simple example of a *linear gluing* argument, the point of which is to reverse the "convergence after rescaling" process that we saw in Lemma 5.5.1. The first step is a **pregluing** construction which turns elements of $\bigoplus_{\zeta \in Z^+(\beta)} \ker \mathbf{D}_+$ into *approximate* solutions to $\mathbf{D}_r \eta = 0$ for large r. To this end, fix a smooth bump function

$$\rho \in C^\infty_0(\mathring{\mathbb{D}}, [0, 1]), \qquad \rho|_{\mathbb{D}_{1/2}} \equiv 1$$

and define for each $\zeta \in Z^+(\beta)$ and $r > 0$ a linear map

$$\Phi^\zeta_r : \ker \mathbf{D}_+ \to \Gamma(E)$$

such that $\Phi^\zeta_r(f)$ is a section with support in $\mathcal{D}(\zeta)$ whose expression in our fixed coordinate and trivialization on that neighborhood is the function

$$f^\zeta_r(z) = \rho(z)\sqrt{r} f(\sqrt{r} z).$$

Adding up the Φ_r^ζ for all $\zeta \in Z^+(\beta)$ then produces a linear map

$$\Phi_r : \bigoplus_{\zeta \in Z^+(\beta)} \ker \mathbf{D}_+ \to \Gamma(E)$$

whose image consists of sections supported near $Z^+(\beta)$, each a linear combination of cut-off Gaussians with energy concentrated in smaller neighborhoods of $Z^+(\beta)$ for larger r. These sections are manifestly not in $\ker \mathbf{D}_r$ since they vanish on open subsets and thus violate unique continuation, but they are close, in a quantitative sense:

LEMMA 5.7.2. *For each $r > 0$, there exists a constant $c_r > 0$ such that*

$$\|\mathbf{D}_r \Phi_r(f)\|_{L^2} \leqslant c_r \|f\|_{L^2} \quad \text{for all} \quad f \in \bigoplus_{\zeta \in Z^+(\beta)} \ker \mathbf{D}_+,$$

and $c_r \to 0$ as $r \to \infty$. Moreover, for every pair $f, g \in \bigoplus_{\zeta \in Z^+(\beta)} \ker \mathbf{D}_+$,

$$\langle \Phi_r(f), \Phi_r(g) \rangle_{L^2} \to \langle f, g \rangle_{L^2}$$

as $r \to \infty$.

PROOF. First, observe that any $f \in \bigoplus_{\zeta \in Z^+(\beta)} \ker \mathbf{D}_+$ is described by a collection of functions $\{f_\zeta \in L^2(\mathbb{C})\}_{\zeta \in Z^+(\beta)}$ which take the form

$$f_\zeta(z) = K_\zeta e^{-\frac{1}{2}|z|^2},$$

for some constants $K_\zeta \in \mathbb{R}$. Since each f_ζ is in $\ker \mathbf{D}_+$, we plug in the local formula (5.4) for $\mathbf{D}_r$ and find

$$\begin{aligned} \mathbf{D}_r \left(\Phi_r(f)|_{\mathcal{D}(\zeta)} \right)(z) &= \bar\partial \rho(z) \cdot \sqrt{r} f_\zeta(\sqrt{r} z) + \rho(z) \cdot r \bar\partial f_\zeta(\sqrt{r} z) \\ &\quad + r z \rho(z) \sqrt{r} f_\zeta(\sqrt{r} z) \\ &= \bar\partial \rho(z) \cdot \sqrt{r} f_\zeta(\sqrt{r} z) + \rho(z) r \cdot \mathbf{D}_+ f_\zeta(\sqrt{r} z) \\ &= \bar\partial \rho(z) \cdot \sqrt{r} K_\zeta e^{-\frac{1}{2} r |z|^2}. \end{aligned} \tag{5.5}$$

Now since $\bar\partial \rho = 0$ in $\mathbb{D}_{1/2}$, we obtain

$$\begin{aligned} \|\mathbf{D}_r \Phi_r(f)\|_{L^2}^2 &= \sum_{\zeta \in Z^+(\beta)} \int_{\mathcal{D}(\zeta)} |\mathbf{D}_r \Phi_r(f)(z)|^2 \, d\mu(z) \\ &= \sum_{\zeta \in Z^+(\beta)} \int_{\mathbb{D} \setminus \mathbb{D}_{1/2}} |\bar\partial \rho(z)|^2 r K_\zeta^2 e^{-r|z|^2} \, d\mu(z) \\ &\leqslant I r e^{-r/4} \sum_{\zeta \in Z^+(\beta)} K_\zeta^2, \end{aligned}$$

where we abbreviate $I := \int_{\mathbb{D} \setminus \mathbb{D}_{1/2}} \left| \bar\partial \rho(z) \right|^2 \, d\mu(z)$. The norm of f is given by

$$\|f\|_{L^2}^2 = \sum_{\zeta \in Z^+(\beta)} \int_{\mathbb{C}} K_\zeta^2 e^{-|z|^2} \, d\mu(z) = \left(\int_{\mathbb{C}} e^{-|z|^2} \, d\mu(z) \right) \sum_{\zeta \in Z^+(\beta)} K_\zeta^2.$$

We conclude that there is a bound of the form

$$\|\mathbf{D}_r \Phi_r(f)\|_{L^2} \leqslant C \sqrt{r} e^{-r/2} \|f\|_{L^2},$$

which proves the first statement since $\sqrt{r}e^{-r/2} \to 0$ as $r \to \infty$.

The second statement follows by a change of variable, since

$$\begin{aligned}\langle \Phi_r(f), \Phi_r(g) \rangle_{L^2} &= \sum_{\zeta \in Z^+(\beta)} \langle \Phi_r(f)|_{\mathcal{D}(\zeta)}, \Phi_r(g)|_{\mathcal{D}(\zeta)} \rangle_{L^2(\mathcal{D}(\zeta))} \\ &= \sum_{\zeta \in Z^+(\beta)} \int_{\mathbb{D}} \rho^2(z) r f_\zeta(\sqrt{r}z) g_\zeta(\sqrt{r}z)\, d\mu(z) \\ &= \sum_{\zeta \in Z^+(\beta)} \int_{\mathbb{D}_{\sqrt{r}}} \rho^2\left(\frac{z}{\sqrt{r}}\right) f_\zeta(z) g_\zeta(z)\, d\mu(z).\end{aligned}$$

The functions f_ζ and g_ζ are both real multiples of $e^{-\frac{1}{2}|z|^2}$, so this last integral for each $\zeta \in Z^+(\beta)$ is bounded between $\int_{\mathbb{D}_{\sqrt{r}/2}} f_\zeta(z) g_\zeta(z)\, d\mu(z)$ and $\int_{\mathbb{D}_{\sqrt{r}}} f_\zeta(z) g_\zeta(z)\, d\mu(z)$, both of which converge to $\int_{\mathbb{C}} f_\zeta(z) g_\zeta(z)\, d\mu(z)$ as $r \to \infty$, thus

$$\lim_{r\to\infty} \langle \Phi_r(f), \Phi_r(g) \rangle_{L^2} = \langle f, g \rangle_{L^2}.$$

□

To turn approximate solutions into actual solutions, let

$$\Pi_r : L^2(E) \to \ker \mathbf{D}_r$$

denote the orthogonal projection. We will prove:

PROPOSITION 5.7.3. *If all zeroes of β are positive, then the linear map*

$$\Pi_r \circ \Phi_r : \bigoplus_{\zeta \in Z^+(\beta)} \ker \mathbf{D}_+ \to \ker \mathbf{D}_r$$

is injective for all $r > 0$ sufficiently large.

This statement says in effect that whenever $r > 0$ is large enough and $\eta := \Phi_r(f) \in \Gamma(E)$ is in the image of the pregluing map, with f normalized by $\|f\|_{L^2} = 1$, we can find a "correction" $\xi \in (\ker \mathbf{D}_r)^\perp$ such that

$$\eta + \xi \neq 0 \quad \text{but} \quad \mathbf{D}_r(\eta + \xi) = 0.$$

An element $\xi \in (\ker \mathbf{D}_r)^\perp$ with the second property certainly exists, and in fact it's unique: indeed, the assumption $Z^-(\beta) = \emptyset$ implies via Lemma 5.7.1 that $\mathbf{D}_r$ is surjective and thus restricts to an isomorphism from $(\ker \mathbf{D})^\perp \cap H^1(E)$ to $L^2(F)$, with a bounded right inverse

$$\mathbf{Q}_r : L^2(F) \to H^1(E) \cap (\ker \mathbf{D})^\perp,$$

hence $\xi := -\mathbf{Q}_r(\mathbf{D}_r\eta)$. We know moreover from Lemma 5.7.2 that $\|\eta\|_{L^2}$ is close to $\|f\|_{L^2} = 1$, so to prove $\eta + \xi \neq 0$, it would suffice to show $\|\xi\|_{L^2}$ is small, which sounds likely since we also know $\|\mathbf{D}_r\eta\|_{L^2}$ is small and $\mathbf{Q}_r$ is a bounded operator. To make this reasoning precise, we just need to have some control over $\|\mathbf{Q}_r\|$ as $r \to \infty$, or equivalently, a quantitative measure of the injectivity of $\mathbf{D}_r|_{(\ker \mathbf{D}_r)^\perp \cap H^1(E)}$. This requires one last appeal to the Weitzenböck formula.

LEMMA 5.7.4. *Assume all zeroes of β are positive. Then there exist constants $c > 0$ and r_0 such that for all $r > r_0$,*

$$\|\eta\|_{L^2} \leqslant c\|\mathbf{D}_r\eta\|_{L^2} \quad \text{for all} \quad \eta \in H^1(E) \cap (\ker \mathbf{D}_r)^\perp.$$

PROOF. Let us instead prove that if zeroes of β are all *negative*, then the same bound holds for all $\eta \in H^1(E)$. The stated result follows from this by considering the formal adjoint and using Exercise 5.7.5 below. Note that by density, it suffices to prove the estimate holds for all $\eta \in C_0^\infty(E)$.

Assume therefore that $Z^+(\beta) = \varnothing$ and, arguing by contradiction, suppose there exist sequences $r_\nu \to \infty$ and $\eta_\nu \in C_0^\infty(E)$ with $\|\eta_\nu\|_{L^2} = 1$ and

$$\|\mathbf{D}_{r_\nu}\eta_\nu\|_{L^2} \to 0.$$

The usual rescaling trick and application of the Weitzenböck formula then produces for each $\zeta \in Z^-(\beta)$ a sequence of functions $\eta_\nu^\zeta := \eta_\nu^{(\zeta,r_\nu)} : \mathbb{D}_{\sqrt{r_\nu}} \to \mathbb{C}$ which satisfy

$$\sum_{\zeta \in Z^-(\beta)} \|\eta_\nu^\zeta\|^2_{L^2(\mathbb{D}_{\sqrt{r_\nu}})} \to 1 \quad \text{and} \quad \|\mathbf{D}_-\eta_\nu^\zeta\|_{L^2(\mathbb{D}_{\sqrt{r_\nu}})} \to 0$$

as $\nu \to \infty$. Indeed, defining $\dot{\Sigma}_\epsilon$ as in the proof of Lemma 5.5.1, a similar application of the Weitzenböck formula yields

$$\|\mathbf{D}_{r_\nu}\eta_\nu\|^2_{L^2(\dot{\Sigma})} \geqslant r_\nu^2 c^2 \|\eta_\nu\|^2_{L^2(\dot{\Sigma}_\epsilon)} - r_\nu c' \|\eta_\nu\|^2_{L^2(\dot{\Sigma})} = r_\nu^2 c^2 \|\eta_\nu\|^2_{L^2(\dot{\Sigma}_\epsilon)} - r_\nu c',$$

for some $c' > 0$. Thus we obtain

$$\|\eta_\nu\|^2_{L^2(\dot{\Sigma}_\epsilon)} \leqslant \frac{\|\mathbf{D}_{r_\nu}\eta_\nu\|^2_{L^2(\dot{\Sigma})}}{c^2 r_\nu^2} + \frac{c'}{r_\nu c^2} \to 0 \quad \text{as} \quad \nu \to \infty,$$

so there is again concentration of energy near the zeroes of the antilinear perturbation: in particular,

$$\begin{aligned} 1 &= \lim_{\nu\to\infty} \|\eta_\nu\|^2_{L^2(\dot{\Sigma})} \\ &= \lim_{\nu\to\infty} \|\eta_\nu\|^2_{L^2(\dot{\Sigma}_\epsilon)} + \lim_{\nu\to\infty} \sum_{\zeta\in Z^-(\beta)} \|\eta_\nu\|^2_{L^2(\mathcal{D}(\zeta))} \\ &= \lim_{\nu\to\infty} \sum_{\zeta\in Z^-(\beta)} \|\eta_\nu^\zeta\|^2_{L^2(\mathbb{D}_{\sqrt{r_\nu}})}. \end{aligned}$$

Moreover, we have

$$\mathbf{D}_-\eta_\nu^\zeta(z) = \frac{1}{r_\nu}\bar{\partial}\eta_\nu\left(\frac{z}{\sqrt{r_\nu}}\right) + \frac{\bar{z}}{\sqrt{r_\nu}}\bar{\eta}_\nu\left(\frac{z}{\sqrt{r_\nu}}\right) = \frac{1}{r_\nu}\mathbf{D}_{r_\nu}\eta_\nu\left(\frac{z}{\sqrt{r_\nu}}\right).$$

Taking the square of the norms on each side, we may integrate and use change of variables to obtain

$$\|\mathbf{D}_-\eta_\nu^\zeta\|_{L^2(\mathbb{D}_{\sqrt{r_\nu}})} = \frac{1}{\sqrt{r_\nu}}\|\mathbf{D}_{r_\nu}\eta_\nu\|_{L^2(\mathcal{D}(\zeta))} \to 0 \quad \text{as} \quad \nu \to \infty.$$

The elliptic estimates from Chapter 2 now provide uniform H^k-bounds for each η_ν^ζ on compact subsets of $\mathbb{C}$ for every $k \in \mathbb{N}$, so that a subsequence converges in $C^\infty_{\text{loc}}(\mathbb{C})$ to a smooth map $\eta_\infty^\zeta \in L^2(\mathbb{C}, \mathbb{C})$ satisfying $\mathbf{D}_-\eta_\infty^\zeta = 0$. But $\sum_{\zeta\in Z^-(\beta)} \|\eta_\infty^\zeta\|^2_{L^2(\mathbb{C})} = 1$, so at least one of these solutions is nontrivial and thus contradicts Proposition 5.6.2. ☐

EXERCISE 5.7.5. Show that for any Fredholm Cauchy-Riemann type operator $\mathbf{D}$ on E, the following two estimates are equivalent, with the same constant $c > 0$ in both:

(i) $\|\eta\|_{L^2(E)} \leqslant c\|\mathbf{D}\eta\|_{L^2(F)}$ for all $\eta \in H^1(E) \cap (\ker \mathbf{D})^\perp$;
(ii) $\|\lambda\|_{L^2(F)} \leqslant c\|\mathbf{D}^*\lambda\|_{L^2(E)}$ for all $\lambda \in H^1(F) \cap (\ker \mathbf{D}^*)^\perp$.

Hint: Elliptic regularity implies that for $\mathbf{D}$ *and* $\mathbf{D}^*$ *as bounded linear operators* $H^1 \to L^2$, $(\ker \mathbf{D})^\perp = \operatorname{im} \mathbf{D}^*$ *and* $(\ker \mathbf{D}^*)^\perp = \operatorname{im} \mathbf{D}$.

PROOF OF PROPOSITION 5.7.3. If the statement is not true, then there exist sequences $r_\nu \to \infty$ and

$$f_\nu \in \bigoplus_{\zeta \in Z^+(\beta)} \ker \mathbf{D}_+$$

such that $\|f_\nu\|_{L^2} = 1$ and $\eta_\nu := \Phi_{r_\nu}(f_\nu) \in (\ker \mathbf{D}_{r_\nu})^\perp$ for all ν. Lemmas 5.7.2 and 5.7.4 then provide estimates of the form

- $\|\eta_\nu\|_{L^2} \to 1$,
- $\|\mathbf{D}_{r_\nu}\eta_\nu\|_{L^2} \to 0$, and
- $\|\eta_\nu\|_{L^2} \leqslant c\|\mathbf{D}_{r_\nu}\eta_\nu\|_{L^2}$

as $\nu \to \infty$, with $c > 0$ independent of ν. These imply:

$$1 = \lim_{\nu\to\infty} \|\eta_\nu\|_{L^2} \leqslant \lim_{\nu\to\infty} c\|\mathbf{D}_{r_\nu}\eta_\nu\|_{L^2} = 0.$$

□

We've proved:

PROPOSITION 5.7.6. *Suppose the assumptions of §5.5 hold and that the section* $\beta \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F))$ *has* $I_+ \geqslant 0$ *positive and* $I_- \geqslant 0$ *negative zeroes. If* $I_- = 0$*, then* $\mathbf{D}_r$ *is surjective with* $\dim\ker \mathbf{D}_r = I_+$ *for all* $r > 0$ *sufficiently large. If* $I_+ = 0$*, then* $\mathbf{D}_r$ *is injective with* $\dim\operatorname{coker} \mathbf{D}_r = I_-$ *for all* $r > 0$ *sufficiently large. In either case,*

$$\operatorname{ind}(\mathbf{D}_r) = I_+ - I_-$$

for all $r > 0$ *sufficiently large.* □

5.8. Antilinear deformations of asymptotic operators

Proposition 5.7.6 suffices to prove the index formula in the closed case, but there is an additional snag if $\Gamma \neq \varnothing$: since $H^1(\dot{\Sigma}) \hookrightarrow L^2(\dot{\Sigma})$ is not a compact inclusion, we have no guarantee that $\mathbf{D}$ and $\mathbf{D}_r := \mathbf{D} + rB$ will have the same index, and generally they will not. A solution to this problem has been pointed out by Chris Gerig [**Ger18**], using a special class of asymptotic operators that also originate in the work of Taubes (see [**Tau10**, Lemma 2.3]).

In general, the only obvious way to guarantee $\operatorname{ind}(\mathbf{D}) = \operatorname{ind}(\mathbf{D}_r)$ for large $r > 0$ is if we can arrange for *every* operator in the family $\{\mathbf{D}_r\}_{r \geqslant 0}$ to be Fredholm, which is not automatic since the zeroth-order perturbation $B : E \to F$ is required to be bounded away from zero near ∞ and must therefore change the asymptotic operators at the punctures. We are therefore led to ask:

QUESTION. *For what nondegenerate asymptotic operators* $\mathbf{A} : H^1(E) \to L^2(E)$ *on a Hermitian line bundle* $(E, J, \omega) \to S^1$ *can one find complex-antilinear bundle maps* $B : E \to E$ *such that*

$$\mathbf{A}_r := \mathbf{A} - rB : H^1(E) \to L^2(E)$$

is an isomorphism for every $r \geqslant 0$*?*

It turns out that it will suffice to find, for each unitary trivialization τ and every $k \in \mathbb{Z}$, a particular pair $(\mathbf{A}_k, B_k)$ such that $\mathbf{A}_k - rB_k$ is nondegenerate for all $r \geqslant 0$ and $\mu^\tau_{\mathrm{CZ}}(\mathbf{A}_k) = k$. To see why, let us proceed under the assumption that such pairs can be found, and use them to compute the index:

LEMMA 5.8.1. *Given* $\mathbf{D}$ *as in Theorem 5.1.6, fix asymptotic trivializations* τ *and suppose that for each puncture* $z \in \Gamma$ *there exists a smooth asymptotic operator* $\mathbf{A}'_z$ *on* (E_z, J_z, ω_z) *with* $\mu^\tau_{\mathrm{CZ}}(\mathbf{A}'_z) = \mu^\tau_{\mathrm{CZ}}(\mathbf{A}_z)$*, such that if* $\mathbf{A}'_z$ *is written with respect to* τ *as* $-J_0\partial_t - S_z(t)$*, then the deformed asymptotic operator*

$$H^1(S^1, \mathbb{R}^2) \to L^2(S^1, \mathbb{R}^2) : \eta \mapsto -J_0\partial_t\eta - S_z(t)\eta - r\beta_z(t)\bar{\eta} \tag{5.6}$$

is nondegenerate for some smooth loop $\beta_z : S^1 \to \mathbb{C}\backslash\{0\}$ *and every* $r \geqslant 0$*. Then*

$$\operatorname{ind}(\mathbf{D}) = \chi(\dot{\Sigma}) + 2c_1^\tau(E) + \sum_{z\in\Gamma^+} \operatorname{wind}(\beta_z) - \sum_{z\in\Gamma^-} \operatorname{wind}(\beta_z).$$

PROOF. Since $\mu^\tau_{\mathrm{CZ}}(\mathbf{A}_z) = \mu^\tau_{\mathrm{CZ}}(\mathbf{A}'_z)$, we can deform $\mathbf{A}_z$ to $\mathbf{A}'_z$ continuously through a family of nondegenerate asymptotic operators. It follows that we can deform $\mathbf{D}$ through a continuous family of Fredholm Cauchy-Riemann type operators to a new operator $\mathbf{D}'$ whose asymptotic operators are $\mathbf{A}'_z$ for $z \in \Gamma$, and $\operatorname{ind}(\mathbf{D}') = \operatorname{ind}(\mathbf{D})$. After a further deformation that preserves the Fredholm property, we are free to assume in fact that $\mathbf{D}'$ is written with respect to the trivialization τ on the cylindrical end near $z \in \Gamma^\pm$ as the translation-invariant operator

$$\partial_s + J_0\partial_t + S_z(t).$$

Now choose $\beta \in \Gamma(\operatorname{Hom}_{\mathbb{C}}(\bar{E}, F))$ with nondegenerate zeroes such that the deformed operators $\mathbf{D}_r\eta := \mathbf{D}'\eta + r\beta\bar{\eta}$ appear in trivialized form on the cylindrical end near $z \in \Gamma^\pm$ as

$$\mathbf{D}_r\eta = \partial_s\eta + J_0\partial_t\eta + S_z(t)\eta + r\beta_z(t)\bar{\eta}.$$

This means $\mathbf{D}_r$ is asymptotic at z to (5.6), which is nondegenerate for every $r \geqslant 0$, implying $\mathbf{D}_r$ is Fredholm for every $r \geqslant 0$ and thus

$$\operatorname{ind}(\mathbf{D}) = \operatorname{ind}(\mathbf{D}_r).$$

The trivializations τ induce trivializations over the cylindrical ends for $\bar{E}$ and $F = \Lambda^{0,1}T^*\dot{\Sigma}\otimes E$, and the expression for β in the resulting asymptotic trivialization of $\operatorname{Hom}_{\mathbb{C}}(\bar{E}, F)$ near $z \in \Gamma$ is $\beta_z(t)$. It follows that the signed count of zeroes of β is

$$\begin{aligned} i(\mathbf{D}) &:= c_1^\tau(\operatorname{Hom}_{\mathbb{C}}(\bar{E}, F)) + \sum_{z\in\Gamma^+} \operatorname{wind}(\beta_z) - \sum_{z\in\Gamma^-} \operatorname{wind}(\beta_z) \\ &= \chi(\dot{\Sigma}) + 2c_1^\tau(E) + \sum_{z\in\Gamma^+} \operatorname{wind}(\beta_z) - \sum_{z\in\Gamma^-} \operatorname{wind}(\beta_z), \end{aligned}$$

where the computation $c_1^\tau(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F)) = \chi(\dot{\Sigma}) + 2c_1^\tau(E)$ follows from the natural isomorphism

$$\begin{aligned}\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F) &\cong \bar{E}^* \otimes F \cong E \otimes F \cong E \otimes \Lambda^{0,1}T^*\dot{\Sigma} \otimes E \cong \Lambda^{0,1}T^*\dot{\Sigma} \otimes E \otimes E \\ &\cong T\dot{\Sigma} \otimes E \otimes E.\end{aligned}$$

We are free to assume that all zeroes of β are either positive or negative, depending on the sign of $i(\mathbf{D})$. Proposition 5.7.6 then implies $\mathrm{ind}(\mathbf{D}_r) = i(\mathbf{D})$ for large r. ☐

Notice that instead of nondegenerate families $\mathbf{A} - rB$ parametrized by $r \in [0, \infty)$, it is just as well to find such families which are nondegenerate and have the right Conley-Zehnder index for all $r > 0$, as the $r \geqslant 1$ portion of this family can be rewritten as $(\mathbf{A} - B) - rB$ for $r \geqslant 0$. The following lemma thus completes the proof of Theorem 5.1.6.

LEMMA 5.8.2. *For every $k \in \mathbb{Z}$, the trivial Hermitian line bundle over S^1 admits a smooth asymptotic operator $\mathbf{A}_k$ and a smooth loop $\beta_k : S^1 \to \mathbb{C}\backslash\{0\}$ such that the deformed asymptotic operators*

$$\mathbf{A}_{k,r}\eta := \mathbf{A}_k\eta - r\beta_k\bar{\eta}$$

are nondegenerate for every $r > 0$ and satisfy

$$\mu_{CZ}(\mathbf{A}_{k,r}) = \mathrm{wind}(\beta_k) = k.$$

PROOF. We claim that the choices

$$\mathbf{A}_k\eta := -J_0\partial_t\eta - \pi k\eta \quad \text{and} \quad \beta_k(t) := e^{2\pi i k t}$$

do the trick. We prove this in three steps.

Step 1: $k = 0$. The above formula gives $\mathbf{A}_{0,r} = -J_0\partial_t\eta - r\bar{\eta}$, in which the $r = 1$ case is precisely the operator that we used in Chapter 3 to normalize the Conley-Zehnder index, hence $\mu_{\mathrm{CZ}}(\mathbf{A}_{0,1}) = 0$ by definition. More generally, all of these operators can be expressed in the form $\mathbf{A} := -J_0\partial_t - S$ where $S \in \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^2)$ is a constant nonsingular 2-by-2 symmetric matrix that anticommutes with J_0. We claim that *all* asymptotic operators of this form are nondegenerate. Indeed, the conditions $S^{\mathrm{T}} = S$ and $SJ_0 = -J_0S$ for $J_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ imply that S takes the form $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ with $\det S = -a^2 - b^2 \neq 0$, and moreover S is of this form if and only if J_0S also is. In particular, J_0S is traceless, symmetric, and nonsingular. Solutions of $\mathbf{A}\eta = 0$ then satisfy $\dot{\eta} = J_0S\eta$, which has no periodic solutions since J_0S has one positive and one negative eigenvalue, hence $\ker \mathbf{A} = \{0\}$.

Step 2: Even k. There is a cheap trick to deduce the case $k = 2m$ for any $m \in \mathbb{N}$ from the $k = 0$ case. Recall that by Exercise 3.7.3 in Chapter 3, conjugating $\mathbf{A}_{0,r}$ by a change of trivialization changes its Conley-Zehnder index by twice the degree of that change. In particular, the operator

$$\widetilde{\mathbf{A}}_{0,r}\eta := e^{2\pi i m t}\mathbf{A}_{0,r}(e^{-2\pi i m t}\eta)$$

is also a nondegenerate asymptotic operator, but with $\mu_{\mathrm{CZ}}(\widetilde{\mathbf{A}}_{0,r}) = \mu_{\mathrm{CZ}}(\mathbf{A}_{0,r}) + 2m = k$. Explicitly, we compute

$$\widetilde{\mathbf{A}}_{0,r}\eta = -J_0\partial_t\eta - \pi k\eta - rke^{2\pi ikt}\bar{\eta},$$

so $\mathbf{A}_{k,r} = \widetilde{\mathbf{A}}_{0,r/k}$ is also nondegenerate for every $r > 0$.

Step 3: Odd k. Another cheap trick relates each $\mathbf{A}_{k,r}$ to $\mathbf{A}_{2k,r}$ after an adjustment in r. Given an arbitrary asymptotic operator $\mathbf{A} = -J_0\partial_t - S(t)$ and $m \in \mathbb{N}$, define

$$\mathbf{A}^m := -J_0\partial_t - mS(mt).$$

Geometrically, if $\mathbf{A}$ is a trivialized representation for the asymptotic operator of a Reeb orbit $\gamma : S^1 \to M$, then $\mathbf{A}^m$ is the operator for the m-fold covered orbit $\gamma^m : S^1 \to M : t \mapsto \gamma(mt)$. It is easy to check in particular that if we define $\eta^m(t) := \eta(mt)$ for any given loop $\eta : S^1 \to \mathbb{R}^2$, then

$$\mathbf{A}^m\eta^m = m(\mathbf{A}\eta)^m,$$

so this gives an embedding of $\ker \mathbf{A}$ into $\ker \mathbf{A}^m$, implying that whenever $\mathbf{A}^m$ is nondegenerate for some $m \in \mathbb{N}$, so is $\mathbf{A}$. To make use of this, observe that

$$\mathbf{A}^2_{k,r}\eta = -J_0\partial_t\eta - \pi 2k\eta - 2re^{4\pi ikt}\bar{\eta} = \mathbf{A}_{2k,2r}\eta,$$

so $\mathbf{A}^2_{k,r}$ is nondegenerate for all $r > 0$ by Step 2, and therefore so is $\mathbf{A}_{k,r}$. □

The proof of Theorem 5.1.6 is now complete.

EXERCISE 5.8.3. Derive a Weitzenböck formula for asymptotic operators and use it to show that for any smooth asymptotic operator $\mathbf{A}$ on the trivial Hermitian line bundle and any smooth $\beta : S^1 \to \mathbb{C}\backslash\{0\}$, the deformed operators $\mathbf{A}_r\eta := \mathbf{A}\eta - r\beta\bar{\eta}$ are all nondegenerate for $r > 0$ sufficiently large. Deduce from this that $\mu_{\mathrm{CZ}}(\mathbf{A}_r) = \mathrm{wind}(\beta)$ for large $r > 0$.

REMARK 5.8.4. The proof of the index formula explained in this chapter was recently extended by Dylan Cant [Can22] to the setting of "relative" SFT, i.e. Cauchy-Riemann type operators on bundles over punctured Riemann surfaces with boundary, subject to totally real boundary conditions.

CHAPTER 6

Symplectic cobordisms and moduli spaces

Contents

In this chapter, we introduce the moduli spaces of holomorphic curves that are used to define SFT.

Recall that in Chapter 1, we motivated the notion of a contact manifold by considering hypersurfaces M in a symplectic manifold (W, ω) that satisfy a *convexity* (also known as "contact-type") condition. The point of that condition was that it presents M as one member of a smooth 1-parameter family of hypersurfaces that all have the same Hamiltonian dynamics. That 1-parameter family furnishes the basic model of what we call the *symplectization* of M with its induced contact structure. Stable Hamiltonian structures—which already made a somewhat unmotivated appearance in Chapter 3—were originally introduced in [**HZ94**] as a generalization of the contact-type condition on hypersurfaces, and they were later recognized to furnish the most natural geometric setting for punctured holomorphic curves. This setting has the advantage of allowing us to view seemingly distinct theories such as

Hamiltonian Floer homology as special cases of SFT—and even if we are only interested in contact manifolds, the generalization sometimes makes computations easier than they might be in a purely contact setting. We therefore begin this chapter by discussing the geometric motivation behind stable Hamiltonian structures, and how they naturally arise on hypersurfaces in symplectic manifolds. Once the geometric setting is understood, we shall proceed to define the moduli spaces of punctured holomorphic curves for SFT, and establish a few of their basic properties, in particular the dichotomy between *simple* curves and *multiple covers*, and an asymptotic regularity result that forces exponential convergence near the punctures.

6.1. Stable Hamiltonian structures

6.1.1. Hamiltonian structures and dynamics. For any smooth hypersurface M in a $2n$-dimensional symplectic manifold (W, ω), the restriction $\omega_M := \omega|_{TM} \in \Omega^2(M)$ is a closed 2-form of maximal rank on M. Its 1-dimensional kernel is the characteristic line field $\ker \omega_M \subset TM$, whose integral curves are the orbits on M of any Hamiltonian vector field generated by a function $H : W \to \mathbb{R}$ that has M as a regular level set. The following definition is a way of formulating this notion without needing to mention the ambient manifold W.

DEFINITION 6.1.1. A **Hamiltonian structure** on a smooth $(2n-1)$-manifold M is a closed 2-form $\omega \in \Omega^2(M)$ with maximal rank. The 1-dimensional distribution

$$\ell_\omega := \ker \omega \subset TM$$

is then called the **characteristic line field** of ω.

Notice that ω descends to a nondegenerate 2-form on the quotient bundle TM/ℓ_ω, making the latter into a symplectic vector bundle over M. Since symplectic linear maps preserve orientation, it follows that TM/ℓ_ω is canonically oriented, so if M is orientable, then ℓ_ω is necessarily also orientable. We will typically consider situations in which M is given with an orientation, so that ℓ_ω inherits an orientation.[1] A nowhere zero section $X \in \Gamma(\ell_\omega)$ that is oriented positively can then be called a **Hamiltonian vector field** on (M, ω).

The set of all possible Hamiltonian vector fields on (M, ω) forms an open and convex subset of the infinite-dimensional vector space $\Gamma(\ell_\omega)$. In order to select a favored element in this space and discuss Hamiltonian flows on M, one needs to choose some auxiliary data.

DEFINITION 6.1.2. Given an oriented manifold M with a Hamiltonian structure ω, a **framing** of ω is a choice of 1-form $\lambda \in \Omega^1(M)$ such that λ is positive on the oriented line field ℓ_ω. The pair (ω, λ) will be referred to in this case as a **framed Hamiltonian structure** on M.[2]

[1] Our convention for orienting quotient spaces is that if V is an oriented vector space and $W \subset V$ is an oriented subspace, then for any positive basis $(w_1, \ldots, w_k, v_1, \ldots, v_m)$ of V such that $(w_1, \ldots, w_k)$ is a positive basis of W, the quotient projection sends $(v_1, \ldots, v_m)$ to a positive basis of V/W.

[2] This terminology is widespread but not entirely standardized, e.g. [Eli07] uses the word "framing" to mean what we would call a "stable framing" (see Definition 6.1.15) together with an extra choice of ω-compatible complex structure J on $\xi = \ker \lambda$.

EXERCISE 6.1.3. Fix an oriented $(2n-1)$-manifold M with Hamiltonian structure ω.

(a) Show that the space of all framings of ω is convex, and use a partition of unity to show that framings always exist.
(b) Show that $\lambda \in \Omega^1(M)$ is a framing of ω if and only if $\lambda \wedge \omega^{n-1} > 0$.

A framing λ associates to a Hamiltonian structure ω two useful pieces of auxiliary data: one is the so-called **Reeb vector field** R, which is the particular Hamiltonian vector field determined by the conditions

$$\omega(R, \cdot) \equiv 0 \quad \text{and} \quad \lambda(R) \equiv 1.$$

Secondly, λ determines a complementary vector bundle for ℓ_ω, namely

$$\xi := \ker \lambda \subset TM.$$

This is a co-oriented hyperplane distribution transverse to ℓ_ω, thus $\omega|_\xi$ is nondegenerate and gives $\xi \to M$ the structure of a symplectic vector bundle.

EXAMPLE 6.1.4. If $\alpha \in \Omega^1(M)$ is a contact form on M, then $(d\alpha, \alpha)$ is a framed Hamiltonian structure whose associated vector field R and hyperplane distribution ξ are the usual Reeb vector field from contact geometry (see Definition 1.3.6) and the contact structure defined via α.

As in the contact-geometric setting, the Reeb vector field of an arbitrary framed Hamiltonian structure (ω, λ) satisfies

$$\mathcal{L}_R \omega = d\iota_R \omega + \iota_R d\omega = 0,$$

thus its flow $\varphi^t : M \to M$ preserves ω. Unlike the contact setting, φ^t need not satisfy any particular properties in relation to λ, so it need not preserve ξ. However, for any integral curve $\gamma \subset M$ of ℓ_ω, the linearized flow of R along γ preserves R and thus descends to the quotient bundle TM/ℓ_ω, on which it preserves the symplectic structure since $\mathcal{L}_R \omega = 0$. Defining

$$\pi_\xi : TM \to \xi$$

as the fiberwise linear projection along ℓ_ω, π_ξ descends to a natural bundle isomorphism $TM/\ell_\omega \xrightarrow{\cong} \xi$, so the observations above prove:

PROPOSITION 6.1.5. *Suppose (ω, λ) is a framed Hamiltonian structure on M with associated Reeb vector field R and flow φ^t, and $\gamma : (a, b) \to M$ is a solution to the equation $\dot{\gamma} = R(\gamma)$. Then for any $t_0, t_1 \in (a, b)$, the linear map*

$$\pi_\xi \circ d\varphi^{t_1 - t_0}(\gamma(t_0)) : \xi_{\gamma(t_0)} \to \xi_{\gamma(t_1)}$$

is a symplectic isomorphism. In particular, there exists a unique symplectic connection ∇^ω on the bundle ξ along each integral curve of ℓ_ω such that parallel transport along the path γ is given by the composition of the projection π_ξ with the linearized Reeb flow. □

EXERCISE 6.1.6. Show that if ∇ is any symmetric connection on M, then the symplectic connection ∇^ω on $\gamma^*\xi$ in Proposition 6.1.5 is given by the formula

$$\nabla_t^\omega \eta = \pi_\xi \left(\nabla_t \eta - \nabla_\eta R \right).$$

Hint: It suffices to show that the right hand side defines a connection on $\gamma^*\xi$ whose parallel sections are the same as those of ∇^ω.

LEMMA 6.1.7. *For any solution $\gamma : (a, b) \to M$ of $\dot{\gamma} = R(\gamma)$, any $\eta \in \Gamma(\gamma^*\xi)$ and any symmetric connection ∇ on M,*

$$\lambda(\nabla_t \eta - \nabla_\eta R) = -d\lambda(R(\gamma), \eta).$$

In particular, it follows that the projection π_ξ can be omitted from the formula in Exercise 6.1.6 if $d\lambda(R, \cdot) \equiv 0$.

PROOF. Consider a smooth 1-parameter family $\{\gamma_\rho : (a, b) \to M\}_{\rho \in (-\epsilon, \epsilon)}$ with $\gamma_0 = \gamma$ and $\partial_\rho \gamma_\rho|_{\rho=0} = \eta$. Repeating the calculation that preceded Definition 3.3.2 in the present more general context, one finds

$$\nabla_\rho \left(\pi_\xi \dot{\gamma}_\rho \right) = \nabla_t \eta - \nabla_\eta R - d\lambda(\eta, R(\gamma)) \cdot R(\gamma),$$

and the fact that $\pi_\xi \dot{\gamma}_\rho$ is in $\Gamma(\gamma_\rho^*\xi)$ for every ρ while $\pi_\xi \dot{\gamma} = 0$ implies that the right hand side is a section of $\gamma^*\xi$. Evaluating λ on this expression then gives the stated formula. □

DEFINITION 6.1.8. A periodic orbit $\gamma : \mathbb{R} \to M$ with period $T > 0$ of the Reeb vector field R for a framed Hamiltonian structure (ω, λ) is called **nondegenerate** if the symplectic linear map $\pi_\xi \circ d\varphi^T(\gamma(0)) : \xi_{\gamma(0)} \to \xi_{\gamma(0)}$ does not have 1 as an eigenvalue. Equivalently, this means that the bundle $\gamma^*\xi \to \mathbb{R}$ does not admit any T-periodic sections that are parallel with respect to the symplectic connection ∇^ω described in Proposition 6.1.5.

In the case $(\omega, \lambda) = (d\alpha, \alpha)$ for α a contact form, this notion of nondegeneracy is equivalent to the notion we defined for Reeb vector fields of contact forms in §1.3, and it implies that a T-periodic orbit γ is always *isolated*, in the sense that there cannot exist a sequence of T_j-periodic orbits $\gamma_j : \mathbb{R} \to M$ disjoint from γ for which $T_j \to T$ and $\gamma_j \to \gamma$ in C^∞ (or any other reasonable topology).

As in Chapter 3, nondegeneracy can also be rephrased in terms of asymptotic operators. If $\gamma : S^1 \to M$ satisfies

$$\dot{\gamma} = T \cdot R(\gamma)$$

for some $T > 0$ and $J : \xi \to \xi$ is a choice of complex structure compatible with ω, then $(\gamma^*\xi, J, \omega|_\xi)$ is a Hermitian vector bundle over S^1, and we define the asymptotic operator associated to γ by

$$\mathbf{A}_\gamma := -J\nabla_t^\omega : \Gamma(\gamma^*\xi) \to \Gamma(\gamma^*\xi).$$

Here ∇^ω is the symplectic connection defined on ξ along integral curves of ℓ_ω via Proposition 6.1.5. Exercise 3.4.2 implies that $\mathbf{A}_\gamma$ is a symmetric operator with respect to the natural real L^2-product on $\Gamma(\gamma^*\xi)$ determined by the bundle metric $\omega(\cdot, J\cdot)$, and this symmetry can be explained by interpreting $\mathbf{A}_\gamma$ as the Hessian of a (locally defined) action functional as in §3.3. The definition of nondegeneracy for the orbit γ can now be reformulated as the condition that the asymptotic operator $\mathbf{A}_\gamma$ is nondegenerate in the sense of Chapter 3, i.e. its kernel is trivial. In this case, we define the **Conley-Zehnder index** of γ with respect to any choice of symplectic trivialization τ for $\gamma^*\xi$ as

$$\mu^\tau_{\mathrm{CZ}}(\gamma) := \mu^\tau_{\mathrm{CZ}}(\mathbf{A}_\gamma).$$

An explicit formula for $\mathbf{A}_\gamma$ comes from Exercise 6.1.6: for any symmetric connection ∇ on M, we have

$$\mathbf{A}_\gamma \eta = -J\pi_\xi \left(\nabla_t \eta - T\nabla_\eta R\right).$$

Note that by Lemma 6.1.7, the projection π_ξ cannot always be omitted from this formula, though it can in the contact case.

6.1.2. Collar neighborhoods and cobordisms. If (W, ω) is a symplectic manifold, any hypersurface $M \subset W$ naturally inherits the Hamiltonian structure $\omega_M := \omega|_{TM}$, and Exercise 6.1.3 implies that if M is oriented (which we shall always assume), then it can be endowed with a framing as auxiliary data. We would now like to examine how the symplectic structure in a neighborhood of M is determined by the Hamiltonian structure on M.

PROPOSITION 6.1.9. *Suppose M is a smooth oriented hypersurface in a symplectic manifold (W, ω), and associate to any given vector field $V \in \Gamma(TW|_M)$ along M the 1-form*

$$\lambda := \omega(V, \cdot)|_{TM} \in \Omega^1(M).$$

Then V is positively transverse[3] *to M if and only if λ is a framing of the Hamiltonian structure $\omega_M := \omega|_{TM} \in \Omega^2(M)$. Moreover, if this holds and M is compact and contained in the interior of W, then a neighborhood $\mathcal{N}(M) \subset W$ of M admits a symplectomorphism*

$$(\mathcal{N}(M), \omega) \cong ((-\epsilon, \epsilon) \times M, \omega_M + d(r\lambda))$$

identifying $M \subset \mathcal{N}(M)$ with $\{0\} \times M$ and V with ∂_r, where r denotes the coordinate on the first factor of $(-\epsilon, \epsilon) \times M$.

PROOF. Pick a Hamiltonian vector field $X \in \Gamma(\ell_\omega)$ on M. If V is tangent to M at some point $x \in M$, then clearly $\lambda(X(x)) = \omega(V(x), X(x)) = -\omega(X(x), V(x)) = 0$ since $X(x) \in \ker \omega_M$. If on the other hand V is transverse to M at x, then $\lambda(X(x)) = -\omega(X(x), V(x))$ cannot vanish, as this would imply $\omega(X(x), \cdot) = 0$, violating the assumption that ω is nondegenerate. To check the sign, choose a basis $(Y_1, \ldots, Y_{2n-2})$ of $\xi := \ker \lambda$ at x that is positively oriented with respect to the volume form $\omega^{n-1}|_\xi$, and observe that the orientation of ℓ_ω is defined to make $(X(x), Y_1, \ldots, Y_{2n-2})$ a positively oriented basis of $T_x M$. The orientation of the basis $(V(x), X(x), Y_1, \ldots, Y_{2n-2})$ of $T_x W$ is therefore positive or negative depending on whether $V(x)$ is positively or negatively transverse to M. In either case, $\omega^n(V(x), X(x), Y_1, \ldots, Y_{2n-2})$ is the product of a positive combinatorial factor with $\omega(V(x), X(x))$ and $\omega^{n-1}(Y_1, \ldots, Y_{2n-2})$ since $\omega(V(x), Y_j) = \lambda(Y_j) = 0$ and

[3] In this context, we say that V is **positively transverse** to M if for every point $x \in M$ and positively oriented basis $(Y_1, \ldots, Y_{2n-1})$ of $T_x M$, the basis $(V(x), Y_1, \ldots, Y_{2n-1})$ of $T_x W$ is also positively oriented.

$\omega(X(x), Y_j) = 0$ for all $j = 1, \dots, 2n-2$. Since $\omega^{n-1}(Y_1, \dots, Y_{2n-2})$ is positive by the definition of the orientation on ξ, the sign of $\lambda(X(x)) = \omega(V(x), X(x))$ is therefore positive if and only if the basis $(V(x), X(x), Y_1, \dots, Y_{2n-2})$ is positively oriented.

Now assume $\lambda = \omega(V, \cdot)|_{TM}$ is a framing and let R denote the associated Reeb vector field. To find the desired tubular neighborhood of M in W, we shall use the Moser deformation trick. We first extend V arbitrarily to a smooth vector field on a neighborhood of M and use its flow φ_V^t to define an embedding

$$(-\epsilon, \epsilon) \times M : (r, x) \mapsto \varphi_V^r(x)$$

for $\epsilon > 0$ sufficiently small. This identifies a neighborhood of M with $(-\epsilon, \epsilon) \times M$ such that M becomes $\{0\} \times M$ and V becomes ∂_r. Under this identification, ω matches the 2-form $\omega_0 := \omega_M + d(r\lambda)$ along $M = \{0\} \times M$; indeed, the latter is $\omega_M + dr \wedge \lambda$ along this hypersurface, so it satisfies

$$\omega_0|_{TM} = \omega_M = \omega|_{TM}, \qquad \text{and} \qquad \omega_0(\partial_r, \cdot)|_{TM} = \lambda = \omega(V, \cdot)|_{TM} = \omega(\partial_r, \cdot)|_{TM}.$$

This proves that ω_0 is also a symplectic form on some neighborhood of M, and so is $\omega_t := t\omega + (1-t)\omega_0$ for every $t \in [0, 1]$, which also matches ω and ω_0 along M.

We claim that there is a 1-form β on $(-\epsilon, \epsilon) \times M$ satisfying

$$\omega = \omega_0 + d\beta \qquad \text{and} \qquad \beta|_M = 0.$$

Indeed, a formula for β can be written using a chain homotopy induced by the obvious deformation retraction of $(-\epsilon, \epsilon) \times M$ to M. In general, if $h : [0, 1] \times N \to Q$ is a smooth homotopy between two maps $f_i = h(i, \cdot) : N \to Q$ for $i = 0, 1$, one can define a chain homotopy $P : \Omega^*(Q) \to \Omega^{*-1}(N)$ by

$$(P\alpha)_x(Y_1, \dots, Y_{m-1}) := \int_0^1 (h^*\alpha)_{(t,x)}(\partial_t, Y_1, \dots, Y_{m-1})\, dt, \qquad \text{for } \alpha \in \Omega^m(Q).$$

The chain homotopy relation $f_1^*\alpha - f_0^*\alpha = d(P\alpha) + P(d\alpha)$ can be checked by integrating both sides over an arbitrary compact oriented m-dimensional submanifold $\Sigma \subset N$ with boundary: the left hand side then becomes the integral of $h^*\alpha$ over $-(\{0\} \times \Sigma) \coprod (\{1\} \times \Sigma)$, while on the right hand side, $\int_\Sigma d(P\alpha) = \int_{\partial\Sigma} P\alpha = \int_{[0,1]\times\partial\Sigma} h^*\alpha$ sees the rest of the integral of $h^*\alpha$ over $\partial([0, 1] \times \Sigma)$, and the relation then follows from Stokes' theorem since $\int_\Sigma P(d\alpha) = \int_{[0,1]\times\Sigma} h^* d\alpha$. Now apply this formula in the present setting with $h(t, r, x) := (tr, x)$ defining a homotopy between the identity map on $(-\epsilon, \epsilon) \times M$ and the projection to $M = \{0\} \times M$: this produces from any closed 2-form α on $(-\epsilon, \epsilon) \times M$ vanishing along M a 1-form $P\alpha$ that also vanishes along M and satisfies $\alpha = d(P\alpha)$ due to the chain homotopy relation. The claim is thus proven by setting $\beta := P(\omega - \omega_0)$.

We can now write

$$\omega_t = \omega_0 + t\, d\beta$$

for each $t \in [0, 1]$. If there exists a smooth isotopy ψ^t on some neighborhood of M satisfying $(\psi^t)^*\omega_t = \omega_0$ for every $t \in [0, 1]$, then it is generated by a time-dependent vector field Y_t which must satisfy

$$0 = \frac{d}{dt}(\psi^t)^*\omega_t = (\psi^t)^* \left(\mathcal{L}_{Y_t}\omega_t + \partial_t\omega_t\right),$$

and thus

$$0 = \mathcal{L}_{Y_t}\omega_t + \partial_t(t\, d\beta) = d\iota_{Y_t}\omega_t + d\beta.$$

This relation then holds if we pick Y_t to be the unique vector field satisfying $\omega_t(Y_t, \cdot) = -\beta$, which is clearly possible on some neighborhood of M due to the nondegeneracy of ω_t. Moreover, Y_t then vanishes along M, so its flow up to time $t = 1$ is well defined on a possibly smaller neighborhood of M, and we obtain a diffeomorphism of such a neighborhood that fixes M and identifies ω with ω_0. □

REMARK 6.1.10. The statement about the tubular neighborhood in Proposition 6.1.9 has obvious analogues if M is a boundary component of W instead of lying in the interior. Here one obtains a collar of the form $(-\epsilon, 0] \times M$ if the given orientation of M matches the boundary orientation of ∂W, which is true if and only if the transverse vector field V points *outward*. If instead V points inward, these two orientations are opposite and the collar is of the form $[0, \epsilon) \times M$.

EXAMPLE 6.1.11. In the case $(\omega, \lambda) = (d\alpha, \alpha)$ for a contact form α, the symplectic form on the tubular neighborhood in Proposition 6.1.9 can be rewritten as $d(e^t\alpha)$ by defining the coordinate $t := \ln(r + 1)$. The proposition is easier to prove in this case: one can construct the neighborhood simply by flowing along V, with no need for the Moser deformation trick (cf. Exercise 1.3.2).

DEFINITION 6.1.12. Given two closed $(2n - 1)$-dimensional oriented manifolds $M_\pm$ with Hamiltonian structures $\omega_\pm$, a **symplectic cobordism from** (M_-, ω_-) **to** (M_+, ω_+) is a compact symplectic $2n$-manifold W whose boundary admits an orientation-preserving diffeomorphism to $-M_- \coprod M_+$ identifying $\omega|_{T(\partial W)}$ with ω_- on M_- and ω_+ on M_+. Here the minus sign in front of M_- denotes an orientation reversal, i.e. the given orientation of M_- is opposite the boundary orientation of ∂W.

If the Hamiltonian structures $\omega_\pm$ are additionally endowed with framings $\lambda_\pm$, then we can also refer to (W, ω) as a **symplectic cobordism from** $(M_-, \mathcal{H}_-)$ **to** $(M_+, \mathcal{H}_+)$, where we abbreviate the framed Hamiltonian structures $\mathcal{H}_\pm := (\omega_\pm, \lambda_\pm)$ on $M_\pm$.

We will sometimes refer to the boundary components M_+ and M_- of a symplectic cobordism (W, ω) as its **positive** and **negative boundary** respectively. In the case where $\mathcal{H}_\pm = (d\alpha_\pm, \alpha_\pm)$ for contact forms $\alpha_\pm$ on $M_\pm$, (W, ω) is what we have previously called a symplectic cobordism from $(M_-, \xi_- := \ker \alpha_-)$ to $(M_+, \xi_+ := \ker \alpha_+)$, and the positive/negative boundaries were previously called the convex/concave boundaries (see §1.4). Note however that convexity and concavity impose nontrivial conditions on (W, ω) near its boundary, e.g. that $\omega|_{\partial W}$ must be exact, whereas *any* compact symplectic manifold with boundary can be viewed as a symplectic cobordism between two manifolds with Hamiltonian structures (either of which may be empty). Moreover, if $\dim W \geqslant 4$, then no component of ∂W can be both convex and concave; see [**Wen18**, Proposition 8.10] for a simple proof of this based on Stokes' theorem. For cobordisms between Hamiltonian structures, however, the labeling of each boundary component as positive or negative is a choice that can be freely reversed—the only caveat is that if we are considering *framed* Hamiltonian structures, then each orientation reversal requires replacing the corresponding framing λ with $-\lambda$.

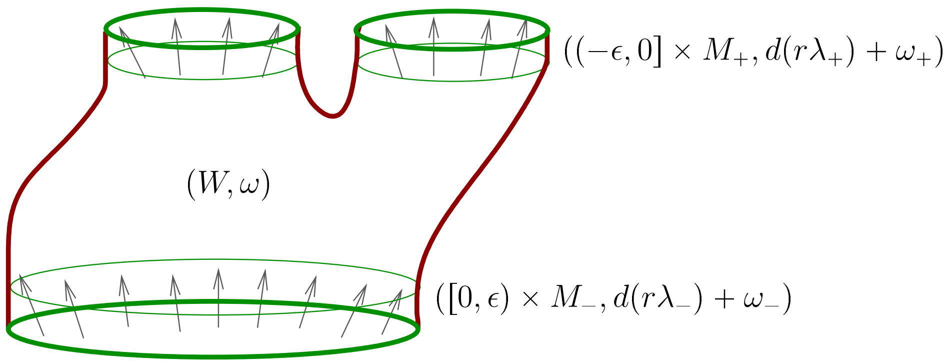

FIGURE 6.1. A symplectic cobordism with positive and negative boundary components $\partial W = -M_- \coprod M_+$ inheriting Hamiltonian structures $\omega_\pm$, shown with their symplectic collar neighborhoods determined by choices of framings $\lambda_\pm$.

From the perspective of SFT, the main difference between the positive and negative boundaries of a cobordism (W, ω) is the form of the collar neighborhoods $\mathcal{N}(M_\pm) \subset W$ that they inherit from Proposition 6.1.9 and Remark 6.1.10, namely

$$(6.1)\qquad \begin{aligned} (\mathcal{N}(M_+), \omega) &\cong ((-\epsilon, 0] \times M_+, \omega_+ + d(r\lambda_+)) \,, \\ (\mathcal{N}(M_-), \omega) &\cong ([0, \epsilon) \times M_-, \omega_- + d(r\lambda_-)) \,. \end{aligned}$$

REMARK 6.1.13. While it may happen that the framings $\lambda_\pm$ of $(M_\pm, \omega_\pm)$ in the above picture are contact forms, one cannot generally expect the induced contact structures to be determined uniquely up to isotopy unless there is also a convexity or concavity condition. For a concrete example, consider the torus $\mathbb{T}^3$ with the sequence of contact forms

$$\alpha_k := \cos(2\pi k\rho)\, d\theta + \sin(2\pi k\rho)\, d\phi$$

for $k \in \mathbb{N}$, written in coordinates $(\rho, \phi, \theta) \in S^1 \times S^1 \times S^1$. We will show in Chapter 11 that the contact structures $\xi_k := \ker \alpha_k$ are not contactomorphic for different values of k. But all of them can be deformed through families of contact structures given by

$$\xi_k^s := \ker \left[(1 - s)\alpha_k + s\, d\rho\right], \qquad s \in [0, 1),$$

so that by Gray's stability theorem, they are all isotopic to arbitrarily small perturbations of the same integrable distribution $\xi^1 := \ker d\rho$. Now pick an area form σ on the closed disk $\mathbb{D}^2$ and consider the symplectic manifold $(W, \omega) := (\mathbb{D}^2 \times \mathbb{T}^2, \sigma \oplus (d\phi \wedge d\theta))$. Identifying $\partial\mathbb{D}^2$ with S^1 in the canonical way, the boundary of W becomes $\mathbb{T}^3$ with Hamiltonian structure $\omega|_{T(\partial W)} = d\phi \wedge d\theta$, and $d\rho$ can be chosen as a framing. It follows that for any $s < 1$ close enough to 1 and any $k \in \mathbb{N}$, the contact form $(1 - s)\alpha_k + s\, d\rho$ is also a framing of this same Hamiltonian structure, even though the isomorphism class of the induced contact structure depends on k.[4]

[4]Apart from being an example of a symplectic cobordism with non-convex framed Hamiltonian boundary, the construction in Remark 6.1.13 amounts to the observation, originating in [Gir94],

6.1.3. Stability. We now introduce an extra condition on framed Hamiltonian structures that will be crucial for the analysis of punctured holomorphic curves.

Definition 6.1.14. A hypersurface M in the interior of a symplectic manifold (W, ω) is called **stable** if a neighborhood of M admits a **stabilizing vector field** V, meaning that V is transverse to M and the 1-parameter family of hypersurfaces

$$M_t := \varphi_V^t(M), \qquad -\epsilon < t < \epsilon$$

generated by the flow φ_V^t of V has the property that each of the diffeomorphisms $M \to M_t$ defined by flowing along V preserves characteristic line fields. The definition has obvious analogues for cases where M is a boundary component of W with V pointing in or outwards.

Definition 6.1.15. A framing λ of a Hamiltonian structure ω on M is called **stable** if

$$d\lambda(R, \cdot) \equiv 0$$

for the associated Reeb vector field R, or equivalently, $\ker \omega \subset \ker d\lambda$. The pair (ω, λ) is in this case called a **stable Hamiltonian structure** (or "SHS" for short).

Stable hypersurfaces first appeared in [**HZ94**] as a class of regular energy surfaces in Hamiltonian systems for which one could reasonably expect the existence of periodic orbits. Indeed, we saw in §1.3 that Liouville vector fields transverse to a hypersurface are stabilizing vector fields, thus contact-type hypersurfaces are also stable. Relatedly, $(d\alpha, \alpha)$ is a stable Hamiltonian structure whenever α is a contact form; we will take a look at some less familiar examples in §6.3. The first appearance of stable Hamiltonian structures as such (though initially without this terminology) was in [**BEH$^+$03**], where they furnished the natural setting for the compactness results of symplectic field theory. They have been studied more systematically in [**CV15**].

Proposition 6.1.16. *A hypersurface M in a symplectic manifold (W, ω) is stable if and only if the Hamiltonian structure $\omega_M := \omega|_{TM}$ on M admits a stable framing.*

Proof. Suppose V is a stabilizing vector field for M with flow φ_V^t, and $\lambda := \omega(V, \cdot)|_{TM}$ is the induced framing of ω_M, with associated Reeb vector field R. Then R generates the kernel of $(\varphi_V^t)^*\omega|_{TM}$ for all t close to 0, implying

$$0 = \mathcal{L}_V \omega(R, \cdot)\big|_{TM} = d\iota_V \omega(R, \cdot)\big|_{TM} = d\lambda(R, \cdot)|_{TM},$$

so λ is a stable framing.

Conversely, if λ is any stable framing of ω_M with Reeb vector field R, then Proposition 6.1.9 identifies a neighborhood $(\mathcal{N}(M), \omega)$ of M with $((-\epsilon, \epsilon) \times M, \omega_M + d(r\lambda))$, and on $M_t := \{t\} \times M$ for every $t \in (-\epsilon, \epsilon)$ we have

$$\omega(R, \cdot)|_{TM_t} = (\omega_M + t\, d\lambda)(R, \cdot) = 0.$$

This shows that R generates the characteristic line field of M_t for every t, thus ∂_r is a stabilizing vector field. ∎

that all of the contact structures ξ_k on $\mathbb{T}^3$ are *weakly* symplectically fillable, and in fact the same symplectic manifold can be regarded as a weak filling of all of them.

We can immediately observe two convenient properties of stable Hamiltonian structures that do not hold without the stability condition: first, the Reeb flow preserves λ, since

$$\mathcal{L}_R\lambda = d\iota_R\lambda + \iota_R d\lambda = d(1) + 0 = 0.$$

The linearized Reeb flow therefore preserves ξ, so there is no longer a need to compose it with the projection $\pi_\xi : TM \to \xi$ when defining the natural symplectic connection ∇^ω along orbits and the notion of nondegeneracy. Similarly, Lemma 6.1.7 now removes the need for including π_ξ in the formula of Exercise 6.1.6 for ∇^ω, and this leads to a simplified formula for the asymptotic operator at a T-periodic orbit γ:

$$\mathbf{A}_\gamma\eta = -J\left(\nabla_t\eta - T\nabla_\eta R\right).$$

DEFINITION 6.1.17. A **symplectic cobordism with stable boundary** is a symplectic cobordism from $(M_-, \mathcal{H}_-)$ to $(M_+, \mathcal{H}_+)$ in the sense of Definition 6.1.12, where $M_\pm$ are closed oriented manifolds endowed with stable Hamiltonian structures $\mathcal{H}_\pm = (\omega_\pm, \lambda_\pm)$.

6.2. Almost complex manifolds with cylindrical ends

6.2.1. Symplectizations and energy. In §1.3, we called the noncompact cylindrical symplectic manifold $(\mathbb{R}\times M, d(e^r\alpha))$ the *symplectization* of the contact manifold $(M, \xi = \ker\alpha)$, and observed (see Exercise 1.3.9) that up to symplectomorphism, it only depends on ξ and not on α. We also defined a natural class of compatible almost complex structures $\mathcal{J}(\alpha)$ on $\mathbb{R}\times M$. If M is endowed with a framed Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ instead of a contact form α, then there is no single symplectic structure on $\mathbb{R}\times M$ that can be called canonical, but there is a natural *class* of symplectic structures arising from the model collar neighborhoods we wrote down in Propostion 6.1.9. Indeed, fix $\epsilon > 0$ small and define

$$\mathcal{T} := \left\{\varphi \in C^\infty(\mathbb{R}, (-\epsilon, \epsilon)) \;\middle|\; \varphi' > 0\right\}, \tag{6.2}$$

which has an obvious identification with the set of all "level-preserving" embeddings $\mathbb{R}\times M \hookrightarrow (-\epsilon, \epsilon)\times M$. If $\epsilon > 0$ is small enough for $\omega + d(r\lambda)$ to be symplectic on $(-\epsilon, \epsilon)\times M$, then pulling it back via the embedding defined via any choice of $\varphi \in \mathcal{T}$ gives rise to a symplectic form

$$\omega_\varphi := \omega + d\left(\varphi(r)\lambda\right) \tag{6.3}$$

on $\mathbb{R}\times M$.

There is a much more obvious generalization of the space $\mathcal{J}(\alpha)$ to the framed Hamiltonian setting.

DEFINITION 6.2.1. Given a framed Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ with associated Reeb vector field R and hyperplane distribution ξ, denote by

$$\mathcal{J}(\mathcal{H}) \subset \mathcal{J}(\mathbb{R}\times M)$$

the space of smooth almost complex structures J on $\mathbb{R}\times M$ with the following properties:

- J is invariant under the $\mathbb{R}$-action on $\mathbb{R}\times M$ by translation of the first factor;

- $J\partial_r = R$ and $JR = -\partial_r$, where r denotes the natural coordinate on the first factor;
- $J(\xi) = \xi$ and $J|_\xi$ is compatible[5] with the symplectic vector bundle structure $\omega|_\xi$.

Notice that if $\mathcal{H} = (d\alpha, \alpha)$ for a contact form α, then $\mathcal{J}(\mathcal{H})$ matches the space $\mathcal{J}(\alpha)$ defined in Chapter 1. One of the crucial reasons to consider only *stable* Hamiltonian structures will be the following easy observation:

PROPOSITION 6.2.2. *Given a framed Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ and an almost complex structure J on $\mathbb{R} \times M$, let us say that J is* ***tamed by*** $\mathcal{H}$ *if the number $\epsilon > 0$ in (6.2) can be chosen such that the symplectic form ω_φ of (6.3) tames J for every $\varphi \in \mathcal{T}$. The following conditions are then equivalent:*

(1) Every $J \in \mathcal{J}(\mathcal{H})$ is tamed by $\mathcal{H}$.
(2) There exists a $J \in \mathcal{J}(\mathcal{H})$ that is tamed by $\mathcal{H}$.
(3) The framing λ is stable.

PROOF. Consider the splitting $T(\mathbb{R} \times M) = \varepsilon \oplus \xi$, where $\xi = \ker \lambda$ and ε is the subbundle spanned by ∂_r and the Reeb vector field R. For any $J \in \mathcal{J}(\mathcal{H})$, these two subbundles are both complex, and ε comes with a canonical trivialization identifying $J|_\varepsilon$ with i. If λ is stable and $\varphi \in \mathcal{T}$, then writing $\omega_\varphi = \omega + \varphi(r)\, d\lambda + \varphi'(r)\, dr \wedge \lambda$, we notice that ε and ξ are also ω_φ-symplectic orthogonal complements. Tameness then follows from the fact that $J|_\varepsilon = i$ is tamed by $\omega_\varphi|_\varepsilon = dr \wedge \lambda|_\varepsilon$ and $J|_\xi$ is tamed by $\omega_\varphi|_\xi = (\omega + \varphi(r)\, d\lambda)|_\xi$, where the latter necessarily holds for any $\epsilon > 0$ sufficiently small since $\omega|_\xi$ tames $J|_\xi$ and tameness is an open condition.

Conversely, suppose $J \in \mathcal{J}(\mathcal{H})$ and λ is not stable, so there exists a point $x \in M$ where $d\lambda(R, v) > 0$ for some $v \in \xi_x$. At $(0, x) \in \mathbb{R} \times M$, we can pick a constant $c > 0$ and write

$$\begin{aligned}\omega_\varphi(R + cJv, J(R + cJv)) &= \omega_\varphi(\partial_r, R) + c^2 \omega_\varphi(v, Jv) - c\omega_\varphi(R, v) \\ &= \varphi'(0) + c^2 \left(\omega + \varphi(0)\, d\lambda\right)(v, Jv) - c\varphi(0)\, d\lambda(R, v).\end{aligned}$$

Choosing $\varphi \in \mathcal{T}$ so that $\varphi(0) = \epsilon/2$, the sum of the second and third terms becomes negative for any $c > 0$ sufficiently small, and since $\varphi \in \mathcal{T}$ can also be chosen to make $\varphi'(0)$ as small as we like, there exists a choice for which the total is negative, meaning ω_φ does not tame J. ☐

Given a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ and $J \in \mathcal{J}(\mathcal{H})$, we define the **energy** of a J-holomorphic curve $u : (\Sigma, j) \to (\mathbb{R} \times M, J)$ by

$$E(u) := \sup_{\varphi \in \mathcal{T}} \int_\Sigma u^* \omega_\varphi,$$

[5]A question frequently asked by beginners in this field is: Would it not suffice to assume $J|_\xi$ is only *tamed* by $\omega|_\xi$ and not necessarily compatible? The short answer is that the standard analytical treatment of punctured holomorphic curves depends on this compatibility assumption in essential ways, mainly because without it, asymptotic operators would not be symmetric (cf. Exercise 3.4.2). If one wishes to relax this assumption, then several fundamental results need to be reproved, e.g. the Fredholm property for Cauchy-Riemann type operators, and their proofs are not obvious. See §6.9 for further discussion.

where the parameter $\epsilon > 0$ in the definition of $\mathcal{T}$ is assumed small enough so that ω_φ tames J for every $\varphi \in \mathcal{T}$. Tameness then implies $E(u) \geqslant 0$, with equality if and only if u is constant. In the contact case, this notion of energy is not identical to the "Hofer energy" that we defined in Chapter 1, nor to Hofer's original definition from [**Hof93**], but all three are equivalent for our purposes, in the sense that uniform bounds on any of them imply uniform bounds on the others.

EXAMPLE 6.2.3. If $x : \mathbb{R} \to M$ is a periodic orbit of R with period $T > 0$, then we can parametrize it as the loop $\gamma : S^1 \to M : t \mapsto x(tT)$ satisfying $\dot\gamma = T \cdot R(\gamma)$ and associate to this loop the map

$$u_\gamma : \mathbb{R} \times S^1 \to \mathbb{R} \times M : (s,t) \mapsto (Ts, \gamma(t)).$$

Then u_γ is J-holomorphic for any $J \in \mathcal{J}(\mathcal{H})$, and is called the **trivial cylinder** (or sometimes also the **orbit cylinder**) over γ. Its energy can be computed via Stokes's theorem: since $\int_{\mathbb{R}\times S^1} u_\gamma^*\omega = 0$ and $\int_{S^1} \gamma^*\lambda = T$, we have

$$E(u_\gamma) = \sup_{\varphi \in \mathcal{T}} \int_{\mathbb{R}\times S^1} u^* d(\varphi(r)\lambda) = 2\epsilon T.$$

EXERCISE 6.2.4. Given *any* orbit $x : \mathbb{R} \to M$ of R, show that the map

$$u : \mathbb{C} \to \mathbb{R} \times M : s + it \mapsto (s, x(t))$$

is J-holomorphic for every $J \in \mathcal{J}(\mathcal{H})$, but its energy is infinite. *Remark: Here it does not matter whether the orbit is periodic. If it is, then the parametrization $x : \mathbb{R} \to M$ covers it infinitely many times.*

REMARK 6.2.5. For an instructive concrete example of Exercise 6.2.4, take $M = S^1$ with its trivial Hamiltonian structure ($\omega := 0 \in \Omega^2(S^1)$ has maximal rank) and the framing $\lambda := dt \in \Omega^1(S^1)$ with respect to the obvious coordinate $t \in S^1 = \mathbb{R}/\mathbb{Z}$. Then $x(t) := t$ is a Reeb orbit, $\mathcal{J}(\omega,\lambda)$ contains only the standard complex structure of $\mathbb{R} \times S^1$, and u becomes the holomorphic map $\mathbb{C} \to \mathbb{R} \times S^1 : s + it \mapsto (s,t)$, which, under the biholomorphic identification $\psi : \mathbb{R} \times S^1 \to \mathbb{C}\backslash\{0\} : (s,t) \mapsto e^{2\pi(s+it)}$, becomes the complex-valued function $\psi \circ u(z) = e^{2\pi z}$ on $\mathbb{C}$. This function has an essential singularity at ∞. More generally, one can show that a holomorphic map $u : \mathbb{D}\backslash\{0\} \to \mathbb{R} \times S^1$ has infinite energy if and only if the singularity of $\psi \circ u : \mathbb{D}\backslash\{0\} \to \mathbb{C}$ at 0 is essential (cf. Exercise 7.1.4).

The trivial cylinders in Example 6.2.3 have several desirable properties, e.g. the map $u_\gamma : \mathbb{R} \times S^1 \to \mathbb{R} \times M$ is proper, and its composition with the projection $\mathbb{R} \times M \to M$ converges asymptotically to a loop near each of the punctures in $\mathbb{R} \times S^1 \cong S^2\backslash\{0,\infty\}$. We will see in Chapter 7 that under generic assumptions about the dynamics of the Reeb vector field, *all* punctured holomorphic curves with finite energy have these two properties. By contrast, the plane $u : \mathbb{C} \to \mathbb{R} \times M$ in Exercise 6.2.4 is not a proper map, and its projection to M may have dense image (if the orbit is not periodic) on a neighborhood of the puncture in $\mathbb{C} \cong S^2\backslash\{\infty\}$. We shall generally exclude curves with infinite energy from consideration.

6.2.2. The linearization along a trivial cylinder. The following demonstrates why asymptotic operators are relevant in SFT: Let us compute the linearized Cauchy-Riemann operator

$$\mathbf{D}_{u_\gamma} : \Gamma(u_\gamma^* T(\mathbb{R} \times M)) \to \Omega^{0,1}(\mathbb{R} \times S^1, u_\gamma^* T(\mathbb{R} \times M))$$

for the trivial cylinder in Example 6.2.3. We derived a general formula for $\mathbf{D}_u$ in §2.1, but in the present situation, we will get more useful information by computing $\mathbf{D}_{u_\gamma}$ directly. To do this, consider the natural splitting of complex subbundles

$$T(\mathbb{R} \times M) = \varepsilon \oplus \xi,$$

where ε denotes the line bundle spanned by ∂_r and R, which comes with a global trivialization identifying $J|_\varepsilon$ with the standard complex structure i. Under the resulting splittings $u_\gamma^* T(\mathbb{R} \times M) = u_\gamma^* \varepsilon \oplus u_\gamma^* \xi$ and $\overline{\mathrm{Hom}}_{\mathbb{C}}(T(\mathbb{R} \times S^1), u_\gamma^* T(\mathbb{R} \times M)) = \overline{\mathrm{Hom}}_{\mathbb{C}}(T(\mathbb{R} \times S^1), u_\gamma^* \varepsilon) \oplus \overline{\mathrm{Hom}}_{\mathbb{C}}(T(\mathbb{R} \times S^1), u_\gamma^* \xi)$, we can write $\mathbf{D}_{u_\gamma}$ in block form

$$\mathbf{D}_{u_\gamma} = \begin{pmatrix} \mathbf{D}_{u_\gamma}^\varepsilon & \mathbf{D}_{u_\gamma}^{\varepsilon\xi} \\ \mathbf{D}_{u_\gamma}^{\xi\varepsilon} & \mathbf{D}_{u_\gamma}^{\xi} \end{pmatrix}.$$

Exercise 6.2.6. Suppose $\mathbf{D} : \Gamma(E) \to \Omega^{0,1}(\dot{\Sigma}, E)$ is a linear Cauchy-Riemann type operator on a vector bundle E with a complex-linear splitting $E = E_1 \oplus E_2$, and

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{pmatrix}$$

is the resulting block decomposition of $\mathbf{D}$. Use the Leibniz rule satisfied by $\mathbf{D}$ to show that $\mathbf{D}_{11}$ and $\mathbf{D}_{22}$ are also Cauchy-Riemann type operators on E_1 and E_2 respectively, while the off-diagonal terms are tensorial, i.e. they commute with multiplication by smooth real-valued functions and thus define bundle maps $\mathbf{D}_{12} : E_2 \to \Lambda^{0,1} T^* \dot{\Sigma} \otimes E_1$ and $\mathbf{D}_{21} : E_1 \to \Lambda^{0,1} T^* \dot{\Sigma} \otimes E_2$.

Now observe that if $u = (u_{\mathbb{R}}, u_M) : \mathbb{R} \times S^1 \to \mathbb{R} \times M$ is another cylinder near u_γ, the nonlinear operator $(\bar{\partial}_J u)\partial_s = \partial_s u + J\,\partial_t u \in \Gamma(u^* T(\mathbb{R} \times M)) = \Gamma(u^*\varepsilon \oplus u^*\xi)$ takes the form

$$(\bar{\partial}_J u)\partial_s = \begin{pmatrix} \partial_s u_{\mathbb{R}} - \lambda(\partial_t u_M) + i\,(\partial_t u_{\mathbb{R}} + \lambda(\partial_s u_M)) \\ \pi_\xi\, \partial_s u_M + J \pi_\xi\, \partial_t u_M \end{pmatrix},$$

where we are using the canonical trivialization of $u^*\varepsilon$ via ∂_r and R to express the top block as a complex-valued function. As observed already in Chapter 3, the bottom block of this expression can be interpreted in terms of the gradient flow of an action functional, in this case the locally defined functional $\mathcal{A}_\beta : C^\infty(S^1) \to \mathbb{R}$ from (3.5) in §3.3, with $\nabla \mathcal{A}_\omega(\gamma) = -J\pi_\xi\, \partial_t \gamma$. Linearizing in the direction of a section $\eta^\xi \in \Gamma(u_\gamma^* \xi)$ and taking the ξ component thus yields an expression involving the Hessian of $\mathcal{A}_\omega$ at the critical point γ, namely

$$(\mathbf{D}_{u_\gamma}^\xi \eta^\xi)\partial_s = (\partial_s - \mathbf{A}_\gamma)\eta^\xi.$$

To compute the blocks $\mathbf{D}_{u_\gamma}^\varepsilon$ and $\mathbf{D}_{u_\gamma}^{\xi\varepsilon}$, notice that $\mathbf{D}_{u_\gamma}\eta^\varepsilon = 0$ whenever η^ε is a constant linear combination of ∂_r and R, as η^ε is then the derivative of a smooth family of J-holomorphic reparametrizations of u_γ. This is enough to prove $\mathbf{D}_{u_\gamma}^{\xi\varepsilon} = 0$, since the latter is tensorial by Exercise 6.2.6, and expressing arbitrary sections of $u_\gamma^*\varepsilon$ as $f\partial_r + gR$, we can apply the Leibniz rule for $\mathbf{D}^\varepsilon_{u_\gamma}$ and conclude

$$(\mathbf{D}^\varepsilon_{u_\gamma}\eta^\varepsilon)\partial_s = (\partial_s + i\,\partial_t)\eta^\varepsilon$$

in the canonical trivialization. The remaining off-diagonal term can be computed as follows: assume $u^\rho = (u^\rho_{\mathbb{R}}, u^\rho_M) : \mathbb{R}\times S^1 \to \mathbb{R}\times M$ is a smooth 1-parameter family of maps for $\rho \in \mathbb{R}$ near 0 such that $u_0 = u_\gamma$ and $\eta^\xi = \partial_\rho u^\rho|_{\rho=0} \in \Gamma(u_\gamma^*\xi)$, which implies

$$\partial_\rho u^\rho_{\mathbb{R}}\big|_{\rho=0} = \lambda\left(\partial_\rho u^\rho_M\big|_{\rho=0}\right) = 0.$$

Differentiating the real and imaginary parts in the top block of $(\bar\partial_J u^\rho)\partial_s$ with respect to the parameter at $\rho = 0$ then gives

$$\partial_\rho\left(\partial_s u^\rho_{\mathbb{R}} - \lambda(\partial_t u^\rho_M)\right)\big|_{\rho=0} = -\partial_\rho\left[\lambda(\partial_t u^\rho)\right]\big|_{\rho=0} = -d\lambda(\eta, \partial_t u_\gamma) = T\cdot d\lambda(R(\gamma), \eta),$$

and

$$\partial_\rho\left(\partial_t u^\rho_{\mathbb{R}} + \lambda(\partial_s u^\rho)\right)\big|_{\rho=0} = \partial_\rho\left[\lambda(\partial_s u^\rho)\right]\big|_{\rho=0} = d\lambda(\eta, \partial_s u_\gamma) = 0.$$

This proves:

PROPOSITION 6.2.7. *For any framed Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ and $J \in \mathcal{J}(\mathcal{H})$, the J-holomorphic trivial cylinder $u_\gamma : \mathbb{R}\times S^1 \to \mathbb{R}\times M$ for a T-periodic orbit $\gamma : S^1 \to M$ has linearized Cauchy-Riemann operator $\mathbf{D}_{u_\gamma} : \Gamma(u_\gamma^*\varepsilon \oplus u_\gamma^*\xi) \to \Omega^{0,1}(\mathbb{R}\times S^1, u_\gamma^*\varepsilon \oplus u_\gamma^*\xi)$ given by*

$$\left(\mathbf{D}_{u_\gamma}\eta\right)\partial_s = \partial_s\eta + \begin{pmatrix} i\partial_t & T\cdot d\lambda(R(\gamma),\cdot) \\ 0 & -\mathbf{A}_\gamma \end{pmatrix}\eta.$$

In particular, if (ω, λ) is a stable Hamiltonian structure, then the off-diagonal term vanishes and $\mathbf{D}_{u_\gamma}$ becomes equivalent to an operator from $\Gamma(u_\gamma^*\varepsilon \oplus u_\gamma^*\xi)$ to itself taking the form $\partial_s - (-i\partial_t \oplus \mathbf{A}_\gamma)$, where $-i\partial_t \oplus \mathbf{A}_\gamma$ defines an asymptotic operator on the direct sum of the trivial Hermitian line bundle over S^1 with $\gamma^*\xi$. □

Proposition 6.2.7 places the linearization $\mathbf{D}_{u_\gamma}$ into the analytical context of the Fredholm theory from Chapters 4 and 5, though it does so if and only if the framing λ of ω is stable. This is the second reason why we shall almost always assume our Hamiltonian structures are stable from now on.

6.2.3. Completed cobordisms. Assume (W, ω) is a symplectic cobordism from $(M_-, \mathcal{H}_-)$ to $(M_+, \mathcal{H}_+)$, where $\mathcal{H}_\pm = (\omega_\pm, \lambda_\pm)$ are framed Hamiltonian structures. For most purposes, (W, ω) is not a suitable setting for J-holomorphic curves, as it lacks any mechanism to control the behavior of curves that touch the boundary. We will therefore remove the boundary by attaching *cylindrical ends*, and then impose a finite energy condition to control the behavior of curves near infinity. As a smooth manifold, the **completion** of W is defined by

$$\widehat{W} := \left((-\infty, 0]\times M_-\right) \cup_{M_-} W \cup_{M_+} \left([0,\infty)\times M_+\right),$$

where the smooth structure on a neighborhood of $M_\pm = \{0\}\times M_\pm \subset W$ is defined with reference to the collar neighborhoods of ∂W in (6.1). Modifying (6.2) by

$$\mathcal{T}_0 := \left\{\varphi \in C^\infty(\mathbb{R}, (-\epsilon, \epsilon)) \;\middle|\; \varphi' > 0 \text{ and } \varphi(r) = r \text{ for } r \text{ near } 0\right\} \tag{6.4}$$

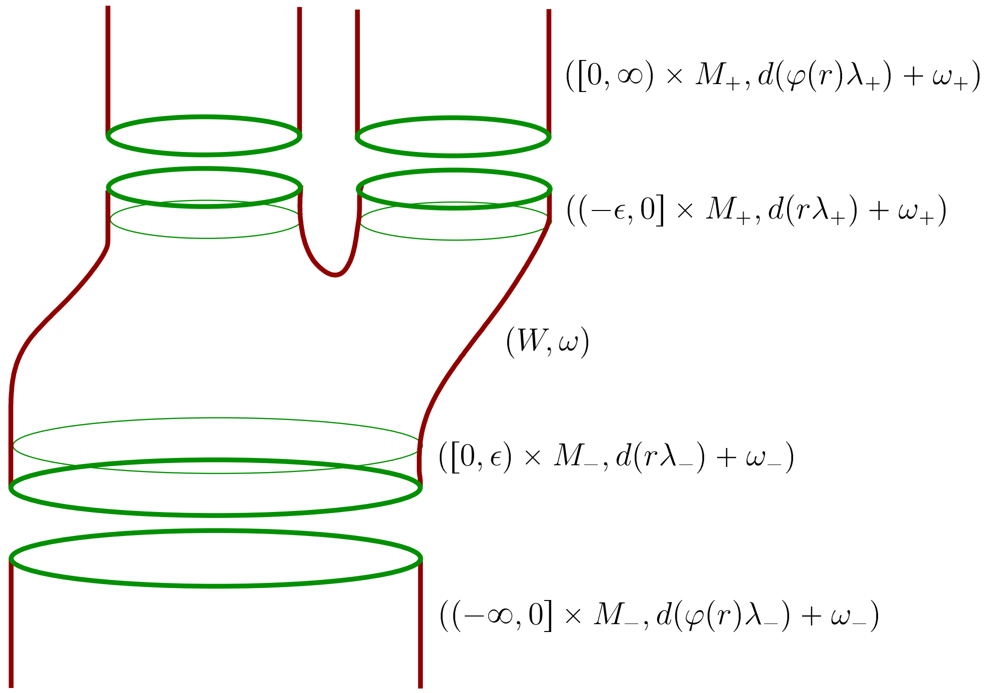

FIGURE 6.2. The completion $(\widehat{W},\omega_\varphi)$ of a symplectic cobordism between two manifolds with framed Hamiltonian structures.

for a fixed $\epsilon > 0$ sufficiently small, we can then use any $\varphi \in \mathcal{T}_0$ to define a symplectic form on $\widehat{W}$ by

$$\omega_\varphi := \begin{cases} d\left(\varphi(r)\lambda_+\right) + \omega_+ & \text{on } [0,\infty)\times M_+, \\ \omega & \text{on } W, \\ d\left(\varphi(r)\lambda_-\right) + \omega_- & \text{on } (-\infty,0]\times M_-, \end{cases}$$

see Figure 6.2. For each $r_0 \geqslant 0$, we define the compact submanifold

$$W^{r_0} := \left([-r_0,0]\times M_-\right)\cup_{M_-} W \cup_{M_+} \left([0,r_0]\times M_+\right),$$

and can view (W^{r_0},ω_φ) as a symplectic cobordism from $(M_-^{r_0},\mathcal{H}_-^{r_0})$ to $(M_+^{r_0},\mathcal{H}_+^{r_0})$ where $M_\pm^{r_0} := \{\pm r_0\}\times M_\pm \subset \widehat{W}$ and the framed Hamiltonian structures $\mathcal{H}_\pm^{r_0} = (\omega_\pm^{r_0},\lambda_\pm^{r_0})$ are given by

$$\omega_\pm^{r_0} := \omega_\varphi\big|_{TM_\pm^{r_0}} = \omega_\pm + \varphi(\pm r_0)\, d\lambda_\pm, \quad \text{and} \quad \lambda_\pm^{r_0} := \omega_\varphi(\partial_r,\cdot)\big|_{TM_\pm^{r_0}} = \varphi'(\pm r_0)\ \lambda_\pm.$$

Notice that if the $\mathcal{H}_\pm$ are stable, then (W^{r_0},ω_φ) also becomes a symplectic cobordism with stable boundary for arbitrary choices $\varphi \in \mathcal{T}_0$.

Since $\widehat{W}$ is noncompact, almost complex structures J on $\widehat{W}$ will need to satisfy conditions near infinity in order for moduli spaces of J-holomorphic curves to be well behaved, but we would like to preserve the freedom of choosing arbitrary compatible or tame almost complex structures in compact subsets.

CONVENTION 6.2.8. For the next definition and, in fact, for the rest of the book, we fix a choice of function $\psi \in \mathcal{T}_0$ and number $r_0 \geqslant 0$. The space of almost complex structures that we work with will depend on these choices, but we will suppress this in the notation in order to avoid unnecessary clutter. In most situations, it is natural to set $r_0 := 0$, in which case the choice of $\psi \in \mathcal{T}_0$ is irrelevant, but occasionally it is useful to have a bit more flexibility: allowing $r_0 > 0$ gives us the freedom to use almost complex structures that need not be translation-invariant on the entirety of the cylindrical ends attached to $\widehat{W}$, but instead only on the portion outside of the enlarged coboridism W^{r_0}.

DEFINITION 6.2.9. Given fixed choices of $r_0 \geqslant 0$ and $\psi \in \mathcal{T}_0$ as described above, let

$$\mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-) \subset \mathcal{J}(\widehat{W})$$

denote the space of smooth almost complex structures J on $\widehat{W}$ such that:

- J on $[r_0, \infty) \times M_+$ matches an element of $\mathcal{J}(\mathcal{H}_+)$;[6]
- J on $(-\infty, -r_0] \times M_-$ matches an element of $\mathcal{J}(\mathcal{H}_-)$;
- J on W^{r_0} is tamed by ω_ψ.

Let

$$\mathcal{J}(\omega, \mathcal{H}_+, \mathcal{H}_-) \subset \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$$

denote the subset for which J is additionally compatible with ω_ψ on W^{r_0}.

Using the fixed choices ψ and r_0 described in Convention 6.2.8, we can set

$$\mathcal{T} := \left\{ \varphi \in \mathcal{T}_0 \;\middle|\; \varphi \equiv \psi \text{ on } [-r_0, r_0] \right\}, \tag{6.5}$$

and conclude from Proposition 6.2.2 that if the framed Hamiltonian structures $\mathcal{H}_\pm$ are both stable, then any given $J \in \mathcal{J}(\omega, \mathcal{H}_+, \mathcal{H}_-)$ is tamed by ω_φ for every $\varphi \in \mathcal{T}$ whenever the number $\epsilon > 0$ in (6.4) is chosen sufficiently small. In this case, it is sensible to define the **energy** of a J-holomorphic curve $u : (\Sigma, j) \to (\widehat{W}, J)$ by

$$E(u) := \sup_{\varphi \in \mathcal{T}} \int_\Sigma u^* \omega_\varphi.$$

REMARK 6.2.10. For any closed manifold M with stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ and a choice of strictly increasing function $\varphi : [0, 1] \to (-\epsilon, \epsilon)$ for $\epsilon > 0$ sufficiently small, one can consider the cobordism

$$([0, 1] \times M, \omega + d(\varphi(r)\lambda)).$$

This has stable boundary, and one would like to regard it as the "trivial cobordism from $(M, \mathcal{H})$ to itself" and identify its completion with the symplectization of $(M, \mathcal{H})$, though strictly speaking, this is wrong: the stable Hamiltonian structures $\mathcal{H}_\pm$ that it induces on $M_- := \{0\} \times M$ and $M_+ := \{1\} \times M$ are in general different from $\mathcal{H}$, and one cannot technically regard $\mathcal{J}(\mathcal{H})$ as contained in any space of the form $\mathcal{J}(\omega, \mathcal{H}_+, \mathcal{H}_-)$ without inventing questionable new notions such as the

[6] While it might seem natural to instead require $J|_{[r_0,\infty)\times M_+} \in \mathcal{J}(\mathcal{H}_+^{r_0})$, the resulting space of almost complex structures would be equivalent to replacing (W, ω) by the larger cobordism (W^{r_0}, ω_ψ) and then repeating this definition with r_0 set to 0. As stated, the definition allows a bit more freedom in applications, which will be useful in Chapter 9 when we need to make perturbations of J on compact subsets to achieve transversality. A similar remark applies to the conditions at the negative end.

"infinitesimal trivial cobordism" $[0, 0] \times M$. It is nonetheless true for fairly trivial reasons that most results about $\mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ or $\mathcal{J}(\omega, \mathcal{H}_+, \mathcal{H}_-)$ apply equally well to $\mathcal{J}(\mathcal{H})$, and we shall use this fact in the following without always mentioning it.

6.3. Examples of stable Hamiltonian structures

6.3.1. The contact case. The following example has been mentioned a few times already and is the one we will work with most often in this book. If α is a contact form on M, then $\mathcal{H} := (d\alpha, \alpha)$ is a stable Hamiltonian structure whose Reeb vector field is the usual contact-geometric notion of a Reeb vector field $R = R_\alpha$. The space $\mathcal{J}(\mathcal{H})$ in this case matches what was called $\mathcal{J}(\alpha)$ in Chapter 1. For two contact manifolds $(M_\pm, \xi_\pm = \ker \alpha_\pm)$, a symplectic cobordism (W, ω) from (M_-, ξ_-) to (M_+, ξ_+) as defined in §1.4 can also be regarded as a symplectic cobordism with stable boundary from $(M_-, \mathcal{H}_-)$ to $(M_-, \mathcal{H}_+)$, where we choose a Liouville vector field V near ∂W to write $\alpha_\pm := \omega(V, \cdot)|_{TM_\pm}$ and $\mathcal{H}_\pm := (d\alpha_\pm, \alpha_\pm)$. Conversely, any symplectic cobordism from $(M_-, \mathcal{H}_-)$ to $(M_+, \mathcal{H}_+)$ with $\mathcal{H}_\pm = (d\alpha_\pm, \alpha_\pm)$ given by contact forms is also a symplectic cobordism in the contact sense from $(M_-, \xi_- = \ker \alpha_-)$ to $(M_+, \xi_+ = \ker \alpha_+)$. One can see this from the collar neighborhoods (6.1), in which ω takes the form $d\alpha_\pm + d(r\alpha_\pm) = d\left((r + 1)\alpha_\pm\right)$, hence it has primitives in these collars whose restrictions to the boundary are contact forms for $\xi_\pm$.

6.3.2. The Floer case. The next example allows one to treat Hamiltonian Floer homology, for most purposes, as a special case of SFT.

Suppose (W, Ω) is a closed symplectic manifold and $H : S^1 \times W \to \mathbb{R}$ is a smooth function, and denote $H_t := H(t, \cdot) : W \to \mathbb{R}$. The time-dependent Hamiltonian vector field X_t defined by $dH_t = -\Omega(X_t, \cdot)$ can then be viewed as defining a *symplectic connection* on the trivial symplectic fiber bundle

$$M := S^1 \times W \xrightarrow{t} S^1,$$

i.e. the flow of $R(t, x) := \partial_t + X_t(x)$ defines symplectic parallel transport maps between fibers. The horizontal subbundle for this connection is the "symplectic orthogonal complement" of the vertical subbundle with respect to the closed 2-form

$$\omega := \Omega + dt \wedge dH.$$

In other words, ω restricts to the fibers of $M \to S^1$ as Ω, and the subbundle $\{Y \in TM \mid \omega(Y, \cdot)|_{T(\{\text{const}\} \times W)} = 0\}$ is generated by R, so ω is the **connection 2-form** defining the connection, cf. [MS17]. Setting $\lambda := dt$ then makes $\mathcal{H} := (\omega, \lambda)$ a stable Hamiltonian structure with Reeb vector field R, and its closed orbits in homotopy classes that project to S^1 with degree one are in 1-to-1 correspondence with the 1-periodic Hamiltonian orbits on W. Notice that this is very different from the contact case: instead of being a contact structure, $\xi = \ker dt$ is an integrable distribution whose integral submanifolds are the fibers of $M \to S^1$.

EXERCISE 6.3.1. Show that the notions of nondegeneracy for closed Reeb orbits on M and for 1-periodic Hamiltonian orbits on W (see §1.2) coincide.

EXERCISE 6.3.2. Work out the relationship between the locally defined action functional $\mathcal{A}_\beta$ from (3.5) in this example and the symplectic action functional for

Hamiltonian systems that we discussed in §1.2. (Try not to worry too much about signs.)

A choice of $J \in \mathcal{J}(\mathcal{H})$ is equivalent to a choice of smooth S^1-parametrized family of compatible almost complex structures $\{J_t\}_{t \in S^1}$ on (W, ω), and J-holomorphic curves $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ can then be written as

$$u = (f, v) : \dot{\Sigma} \to \left(\mathbb{R} \times S^1\right) \times W,$$

where $f : (\dot{\Sigma}, j) \to (\mathbb{R} \times S^1, i)$ is holomorphic. In particular, if $(\dot{\Sigma}, j) = (\mathbb{R} \times S^1, i)$ and f is taken to have an extension to $S^2 \to S^2$ of degree one, then u can be reparametrized so that f is the identity map, hence $u = (\mathrm{Id}, v) : \mathbb{R} \times S^1 \to (\mathbb{R} \times S^1) \times W$ is a section of the trivial fiber bundle $(\mathbb{R} \times S^1) \times W \to \mathbb{R} \times S^1$, and one can check that the equation satisfied by $v : \mathbb{R} \times S^1 \to W$ is precisely the Floer equation

$$\partial_s v + J_t(v)(\partial_t v - X_t(v)) = 0.$$

This setup admits various easy generalizations that produce other interesting variants of Floer homology. One can, for instance, replace the trivial fibration $M = S^1 \times W \to S^1$ with the mapping torus of a given symplectomorphism $\phi : (W, \omega) \to (W, \omega)$, producing a theory in which closed Reeb orbits are equivalent to fixed points of (some Hamiltonian perturbation of) ϕ. This theory is known as *symplectic Floer homology*, see e.g. [DS94, Sei02]. One can also consider closed Reeb orbits whose projections to S^1 have degree greater than 1: this produces a theory based on the *periodic* (but not necessarily fixed) points of the symplectomorphism ϕ. A particular variant of this, specialized to the case $\dim W = 2$, is known as *periodic Floer homology*; see [HS05]. In a slightly different direction, Heegaard Floer homology, a topological invariant of 3-manifolds inspired by Floer's Lagrangian intersection theory, can be reformulated as a theory that counts punctured holomorphic curves with Legendrian boundary in the symplectization of $\Sigma \times [0, 1]$ with a very simple stable Hamiltonian structure, where Σ is a Heegaard surface for the given 3-manifold; see [Lip06]. As a general rule, it is possible (though not always helpful) to reformulate almost any Floer-type theory based on a perturbed holomorphic curve equation within the geometric setup for SFT.

For another interesting example of stable Hamiltonian structures separate from the contact and Floer cases, see [BEH⁺03, Example 2.2 and Remark 5.9].

6.4. Moduli spaces of asymptotically cylindrical curves

Fix a closed manifold M with stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ and $J \in \mathcal{J}(\mathcal{H})$, along with a Riemann surface $(\dot{\Sigma} = \Sigma \backslash \Gamma, j)$ with positive and/or negative punctures $\Gamma = \Gamma^+ \cup \Gamma^-$ and choices of holomorphic cylindrical coordinates $(s, t) \in Z_\pm \cong \dot{\mathcal{U}}_z$ near each puncture $z \in \Gamma^\pm$. Here we are again using the notation

$$Z_+ = [0, \infty) \times S^1, \qquad Z_- = (-\infty, 0] \times S^1,$$

with the choice of Z_+ or Z_- depending on the sign of the puncture (cf. §4.1).

DEFINITION 6.4.1. A smooth map $u : \dot{\Sigma} \to \mathbb{R} \times M$ is called **asymptotically cylindrical** if for each $z \in \Gamma^\pm$, there exists a closed Reeb orbit $\gamma_z : S^1 \to M$ with associated trivial cylinder $u_{\gamma_z} : \mathbb{R} \times S^1 \to \mathbb{R} \times M$, and constants $s_0 \in \mathbb{R}$ and $t_0 \in S^1$ such that

$$u(s - s_0, t - t_0) = \exp_{u_{\gamma_z}(s,t)} h_z(s,t) \quad \text{for } (s,t) \in Z_\pm \cong \dot{\mathcal{U}}_z \text{ with } |s| \gg 0, \tag{6.6}$$

where $h_z(s,t)$ is a vector field along u_{γ_z} satisfying

$$h_z(\cdot + s, \cdot) \to 0 \quad \text{in } C^\infty(Z_\pm) \quad \text{as } s \to \pm\infty.$$

Here we assume that the exponential map and all norms involved in describing the C^∞-convergence of $h_z(\cdot + s, \cdot)$ are invariant under the $\mathbb{R}$-translation action on $\mathbb{R} \times M$. We call γ_z the **asymptotic orbit** of u at the puncture z, and call the vector field h_z along u_{γ_z} appearing in (6.6) the **asymptotic representative** of u at z.

Note that if fixed parametrizations $S^1 \to M$ have been chosen for each closed Reeb orbit, then the decay condition in Definition 6.4.1 implies that both h_z and the constants s_0 and t_0 are uniquely determined by u and the choice of holomorphic cylindrical coordinate system near z. The following exercise shows that the asymptotically cylindrical condition itself is also independent of the choices of holomorphic cylindrical coordinates.

EXERCISE 6.4.2. Consider S^1 with the trivial stable Hamiltonian structure $\mathcal{H}$ (see Remark 6.2.5) and the standard complex structure $i \in \mathcal{J}(\mathcal{H})$ on its symplectization $\mathbb{R} \times S^1$. The biholomorphic map $\mathbb{R} \times S^1 \to \mathbb{C}^* = S^2 \backslash \{0, \infty\} : (s,t) \mapsto e^{2\pi(s+it)}$ can be used to identify the latter with a twice-punctured Riemann sphere.

(a) Show that a holomorphic map $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times S^1, i)$ is asymptotically cylindrical if and only if it extends over the punctures to a holomorphic map $(\Sigma, j) \to (S^2, i)$. Find a relationship between its asymptotic orbits and the presence of critical points of the extension at Γ.
(b) Deduce that for any two choices of holomorphic cylindrical coordinates near a puncture of $\dot{\Sigma}$, the resulting coordinate transformation satisfies the conditions of an asymptotically cylindrical map.
(c) Conclude that the notion of an asymptotically cylindrical map in Definition 6.4.1 does not depend on the choices of holomorphic cylindrical coordinates.

These notions extend in a straightforward way to the setting of a completed symplectic cobordism $\widehat{W}$ with fixed choices of $\psi \in \mathcal{T}_0$, $r_0 \geqslant 0$ and $J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ as in Definition 6.2.9. We shall denote by $\xi_\pm$ and $R_\pm$ the hyperplane distribution and Reeb vector field respectively determined by stable Hamiltonian structures $\mathcal{H}_\pm = (\omega_\pm, \lambda_\pm)$ on the boundary components $M_\pm \subset \partial W$. An asymptotically cylindrical map $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ is then a proper map that sends neighborhoods of positive/negative punctures to the positive/negative cylindrical ends of $\widehat{W}$, where they asymptotically approach trivial cylinders over closed orbits of $R_\pm$ in $\{\pm\infty\} \times M_\pm$; see Figure 6.3.

It is easy to check that asymptotically cylindrical J-holomorphic curves always have finite energy. The converse turns out to be true as well if all Reeb orbits are nondegenerate or Morse-Bott; we will prove the nondegenerate case of this statement in Chapter 7.

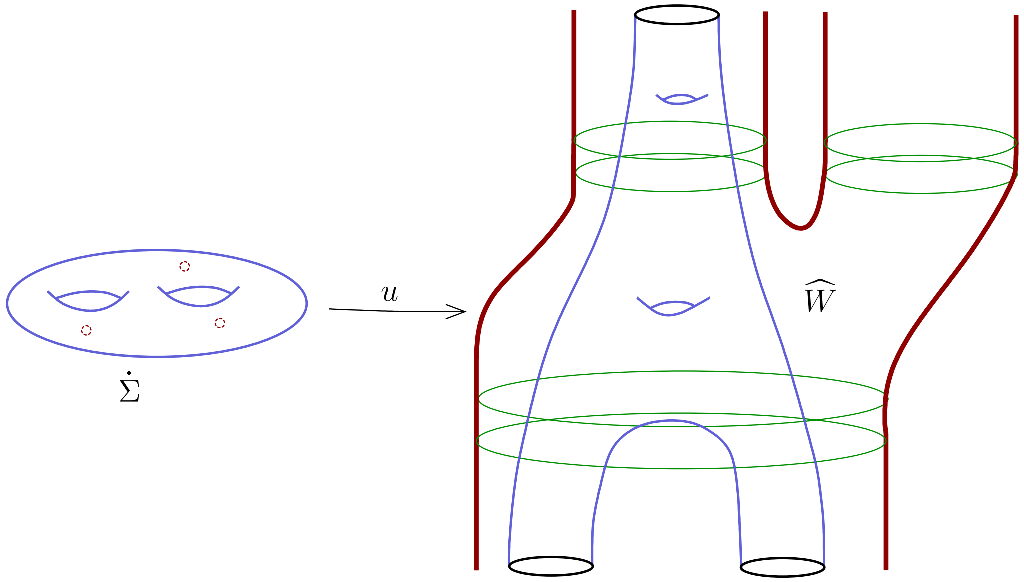

FIGURE 6.3. An asymptotically cylindrical holomorphic curve in $(\widehat{W}, J)$ with genus 2, one positive puncture and two negative punctures.

Every asymptotically cylindrical curve $u : \dot{\Sigma} \to \widehat{W}$ has a well-defined **relative homology class**, meaning the following. Denote the asymptotic orbits of u at its punctures $z \in \Gamma^\pm$ by γ_z, and let

$$\bar{\boldsymbol{\gamma}}^\pm \subset M_\pm$$

denote the closed 1-dimensional submanifold defined as the union over $z \in \Gamma^\pm$ of the images of the orbits γ_z. Let $\bar{\Sigma}$ denote the surface with boundary obtained from $\dot{\Sigma}$ by appending $\{\pm\infty\} \times S^1$ to each of its cylindrical ends, and let $\overline{W}$ likewise denote the compactification of $\widehat{W}$ obtained by attaching $\{\pm\infty\} \times M_\pm$ to its cylindrical ends. Both of these are compact oriented topological manifolds with boundary whose interiors are $\dot{\Sigma}$ and $\widehat{W}$ respectively, and $\partial\overline{W}$ has a natural identification with $\partial W = -M_- \coprod M_+$. Then $u : \dot{\Sigma} \to \widehat{W}$ has a unique continuous extension

$$\bar{u} : (\bar{\Sigma}, \partial\bar{\Sigma}) \to (\overline{W}, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$$

and thus represents a relative homology class

$$[u] := u_*[\bar{\Sigma}] \in H_2(\overline{W}, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-) = H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-),$$

where $[\bar{\Sigma}] \in H_2(\bar{\Sigma}, \partial\bar{\Sigma})$ denotes the relative fundamental class of $\bar{\Sigma}$, and we can use the obvious deformation retraction of $\overline{W}$ to W in order to consider homology classes in W instead of $\overline{W}$. If we consider curves in a symplectization $\mathbb{R} \times M$ instead of the completed cobordism $\widehat{W}$, then $\overline{W}$ becomes $[-\infty, \infty] \times M$ and it is convenient to retract this to $\{0\} \times M \cong M$, thus writing

$$[u] \in H_2([-\infty, \infty] \times M, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-) = H_2(M, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-).$$

We now proceed to define moduli spaces. Fix integers $g, m, k_+, k_- \geqslant 0$ along with ordered sets of Reeb orbits

$$\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \dots, \gamma_{k_\pm}^\pm),$$

where each $\gamma_i^\pm$ is a closed orbit of $R_\pm$ in $M_\pm$. Denote the union of the images of the $\gamma_i^\pm$ by $\bar{\boldsymbol{\gamma}}^\pm \subset M_\pm$, and choose a relative homology class

$$A \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$$

whose image under the boundary map $H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-) \xrightarrow{\partial} H_1(\bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ defined via the long exact sequence of the pair $(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ is

$$\partial A = \sum_{i=1}^{k_+} [\gamma_i^+] - \sum_{i=1}^{k_-} [\gamma_i^-] \in H_1(\bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-). \tag{6.7}$$

The **moduli space** of **unparametrized** J-holomorphic curves of **genus** g with m **marked points**, **homologous to** A and **asymptotic to** $(\boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ is then defined as a set of equivalence classes of tuples

$$\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) = \{(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)\} \,/\sim,$$

where:

(1) (Σ, j) is a closed connected Riemann surface of genus g;
(2) $\Gamma^+ = (z_1^+, \dots, z_{k_+}^+)$, $\Gamma^- = (z_1^-, \dots, z_{k_-}^-)$ and $\Theta = (\zeta_1, \dots, \zeta_m)$ are disjoint ordered sets of distinct points in Σ;
(3) $u : (\dot{\Sigma} := \Sigma \backslash (\Gamma^+ \cup \Gamma^-), j) \to (\widehat{W}, J)$ is an asymptotically cylindrical J-holomorphic map with $[u] = A$, asymptotic at $z_i^\pm \in \Gamma^\pm$ to $\gamma_i^\pm$ for $i = 1, \dots, k_\pm$;
(4) Equivalence

$$(\Sigma_0, j_0, \Gamma_0^+, \Gamma_0^-, \Theta_0, u_0) \sim (\Sigma_1, j_1, \Gamma_1^+, \Gamma_1^-, \Theta_1, u_1)$$

means the existence of a biholomorphic map $\psi : (\Sigma_0, j_0) \to (\Sigma_1, j_1)$, taking $\Gamma_0^\pm$ to $\Gamma_1^\pm$ and Θ_0 to Θ_1 with the ordering preserved, such that

$$u_1 \circ \psi = u_0.$$

When there is no need to specify the relative homology class A, we can abbreviate

$$\mathcal{M}_{g,m}(J, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) := \bigcup_A \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-),$$

where A ranges over the set of all classes in $H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ satisfying the condition (6.7). We shall often abuse notation by abbreviating elements $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)]$ in this moduli space by

$$u \in \mathcal{M}_{g,m}(J, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-).$$

The **automorphism group**

$$\mathrm{Aut}(u) = \mathrm{Aut}(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)$$

of u is defined as the group of biholomorphic maps $\psi : (\Sigma, j) \to (\Sigma, j)$ which act as the identity on $\Gamma^+ \cup \Gamma^- \cup \Theta$ and satisfy $u = u \circ \psi$. Clearly, the isomorphism class of this group depends only on the equivalence class $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \in$

$\mathcal{M}_{g,m}(J, \gamma^+, \gamma^-)$, and we will see in §6.8 below that it is always finite unless $u : \dot{\Sigma} \to \widehat{W}$ is constant. The significance of the marked points is that they determine an **evaluation map**

$$\mathrm{ev} : \mathcal{M}_{g,m}(J, \gamma^+, \gamma^-) \to \widehat{W}^{\times m} : [(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \mapsto (u(\zeta_1), \dots, u(\zeta_m))$$

where $\Theta = (\zeta_1, \dots, \zeta_m)$. For most of our applications, we will be free to assume $m = 0$, as marked points are not needed for defining the most basic versions of SFT. The evaluation map does play a prominent role however in more algebraically elaborate versions of the theory, and especially in the Gromov-Witten invariants (the "closed case" of SFT).

REMARK 6.4.3. The definition of $\mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ given above permits elements $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \in \mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ for which $u : \dot{\Sigma} \to \widehat{W}$ is a constant map if $\Gamma^+ = \Gamma^- = \varnothing$ and $A = 0 \in H_2(\widehat{W})$, but in this case, it is conventional to impose an extra **stability** condition, namely that constant maps are allowed only if

$$\chi(\Sigma \backslash \Theta) < 0.$$

Several details in our study of $\mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ and its compactification will only make sense under this extra assumption, which is harmless since, in practice, we our usually only interested in nonconstant curves. One consequence is that if u is constant, then the group $\mathrm{Aut}(\Sigma, j, \Theta)$ of biholomorphic maps on (Σ, j) fixing Θ is finite (cf. Prop. 8.3.1 in Chapter 8), so in conjunction with Theorem 6.8.1 below, this implies that the automorphism group $\mathrm{Aut}(u)$ for an element $u \in \mathcal{M}_{g,m}(J, \gamma^+, \gamma^-)$ is *always* finite.

6.5. The topology of the moduli space

The elliptic regularity results from Chapters 2 and 4 give us a wide range of freedom in defining the topology of $\mathcal{M}_{g,m}(J, \gamma^+, \gamma^-)$, as they imply that most reasonable choices we could conceivably make on this front will turn out to be equivalent. The notion of convergence described below is geared toward what is easiest to prove in practice, and thus uses mainly the C^∞_{loc}-topology, though some care must then be taken to ensure that nothing wild happens near infinity. For that purpose, we shall make use of the natural translation action

$$\tau_c : \mathbb{R} \times M_\pm \to \mathbb{R} \times M_\pm : (r, x) \mapsto (r + c, x) \qquad \text{for } c \in \mathbb{R},$$

which is holomorphic for the unique translation-invariant almost complex structure on $\mathbb{R} \times M_\pm$ determined by restricting $J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ to the positive or negative cylindrical end. Note that on sufficiently small neighborhoods of each puncture, an asymptotically cylindrical J-holomorphic map $u : \dot{\Sigma} \to \widehat{W}$ can always be regarded as a pseudoholomorphic map from a subset of the standard cylinder $\mathbb{R} \times S^1$ to $\mathbb{R} \times M_\pm$.

REMARK 6.5.1. If $j_\nu \to j$ is a C^∞-convergent sequence of complex structures on a closed Riemann surface Σ, then for any point $z \in \Sigma$, Theorem 2.7.1 provides a C^∞-convergent sequence of holomorphic embeddings $(\mathbb{D}, i) \hookrightarrow (\Sigma, j_\nu)$ that send $0 \in \mathbb{D}$ to the point $z \in \Sigma$. In this situation, we therefore lose no generality in assuming after a convergent sequence of reparametrizations that $j_\nu \equiv j$ on some fixed neighborhood of each puncture in $\dot{\Sigma}$, permitting the choice of a single cylindrical coordinate system $(s,t) \in Z_\pm$ near each puncture that is holomorphic for every j_ν.

DEFINITION 6.5.2. Fix a closed manifold M with stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$, and suppose $\gamma : S^1 \to M$ is a closed Reeb orbit with period $T > 0$. A sequence of smooth maps $u_\nu : Z_\pm \to \mathbb{R} \times M$ will be said to **converge asymptotically to** γ if for every sequence $R_\nu \to \pm\infty$, there exist sequences $\theta_\nu \in S^1$ and $r_\nu \in \mathbb{R}$ such that the sequence of maps

$$[-R_\nu, R_\nu] \times S^1 \to \mathbb{R} \times M : (s,t) \mapsto \tau_{r_\nu} \circ u_\nu(s + R_\nu, t + \theta_\nu)$$

converges in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to the trivial cylinder $u_\gamma(s,t) = (Ts, \gamma(t))$. For a completed symplectic cobordism $\widehat{W}$, a similar definition makes sense for any sequence of maps $u_\nu : Z_\pm \to \widehat{W}$ that all have images in the positive or negative cylindrical end.

The freedom to choose *arbitrary* sequences $R_\nu \to \pm\infty$ in this definition ensures that no important information is lost due to the fact that we are considering convergence only on compact subsets. To express this more quantitatively, recall from Chapter 4 the notation

$$Z^R_+ = [R, \infty) \times S^1, \qquad Z^R_- = (-\infty, -R] \times S^1.$$

EXERCISE 6.5.3. Assume that $u_\nu : Z_\pm \to \mathbb{R} \times M$ converges in $C^\infty_{\text{loc}}(Z_\pm)$ to an asymptotically cylindrical map $u : Z_\pm \to \mathbb{R} \times M$ asymptotic to the orbit γ with trivial cylinder $u_\gamma(s,t) := (Ts, \gamma(t))$, and that u_ν also satisfies the asymptotic convergence condition described in Definition 6.5.2. Show that for any choice of translation-invariant metric on $\mathbb{R} \times M$, u_ν then satisfies a formula of the form

$$u_\nu(s - s_0, t - t_0) = \exp_{u_\gamma(\varphi_\nu(s,t))} \eta_\nu(s,t) \qquad \text{for } \pm s \gg 0,$$

where $s_0 \in \mathbb{R}$ and $t_0 \in S^1$ are constants, $\varphi_\nu : \mathbb{R} \times S^1 \to \mathbb{R} \times S^1$ is a sequence smooth maps, and $\eta_\nu \in \Gamma(\varphi^* u_\gamma^* \xi)$ is a sequence of smooth sections satisfying

$$\sup_\nu \|\varphi_\nu - \text{Id}\|_{C^m(Z^R_\pm)} \to 0 \quad \text{and} \quad \sup_\nu \|\eta_\nu\|_{C^m(Z^R_\pm)} \to 0 \qquad \text{as } R \to \infty$$

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

DEFINITION 6.5.4. Suppose $J_\nu \to J$ is a C^∞-convergent sequence of almost complex structures belonging to $\mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ on the completed symplectic cobordism $\widehat{W}$. We will say that a sequence $[(\Sigma_\nu, j_\nu, \Gamma^+_\nu, \Gamma^-_\nu, \Theta_\nu, u_\nu)] \in \mathcal{M}_{g,m}(J_\nu, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ **converges** to $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \in \mathcal{M}_{g,m}(J, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ as $\nu \to \infty$ if for sufficiently large ν, the equivalence classes in the sequence admit representatives of the form $(\Sigma, j'_\nu, \Gamma^+, \Gamma^-, \Theta, u'_\nu)$ such that

(1) $j'_\nu \to j$ in $C^\infty_{\text{loc}}(\dot{\Sigma})$;
(2) $u'_\nu \to u$ in $C^\infty_{\text{loc}}(\dot{\Sigma}, \widehat{W})$;
(3) $j'_\nu \equiv j$ on a neighborhood of the punctures (cf. Remark 6.5.1);
(4) Each puncture $z \in \Gamma^\pm$ has a holomorphic cylindrical coordinate neighborhood $(\mathcal{U} \backslash \{z\}, j) \cong (Z_\pm, i)$, on which the j'_ν all match, such that the resulting reparametrizations $u'_\nu : Z_\pm \to \widehat{W}$ all have images in the positive/negative

cylindrical end and converge asymptotically in the sense of Definition 6.5.2 to the corresponding asymptotic orbit of u.

By Exercise 6.5.3 above, this notion of convergence also implies *uniform* convergence on the compactified surface $\bar{\Sigma}$,

$$\bar{u}'_\nu \to \bar{u} \qquad \text{in } C^0(\bar{\Sigma}, \overline{W}),$$

so that convergence in $\mathcal{M}_{g,m}(J, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ preserves relative homology classes.

REMARK 6.5.5. There is an obvious generalization of Definition 6.5.4 to situations where not only the almost complex structures but also the symplectic form ω and stable Hamiltonian structures $\mathcal{H}_\pm$ are allowed to vary in C^∞-convergent families.

Taking a fixed almost complex structure $J_\nu := J$ in Definition 6.5.4 produces a notion of convergent sequences in the moduli space $\mathcal{M}_{g,m}(J, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$, and one can show that there is a unique metrizable topology for which this is the definition of convergence. We will not prove this, since we do not really need to know it in such generality—in practice, we will eventually focus on cases in which $\mathcal{M}_{g,m}(J, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ can also be given the structure of a smooth manifold or orbifold, and we will then see directly that the resulting notion of convergence is equivalent to what is defined above.

Since relative homology classes are preserved under the notion of convergence defined above, the moduli spaces $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ for individual classes $A \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ are open and closed subsets of $\mathcal{M}_{g,m}(J, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$. Each may also have multiple connected components and contain curves in different homotopy classes of asymptotically cylindrical maps $\dot{\Sigma} \to \widehat{W}$, but there are at least two good reasons to keep track of the relative homology class A in particular. One is that A, together with the sets of orbits $\boldsymbol{\gamma}^\pm$ and the parameters g and m, determines the dimension of $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ as a smooth orbifold under suitable transversality assumptions; we will discuss this in Chapter 8. The other reason is the following energy bound, which is an easy exercise in Stokes' theorem, and will play a crucial role in the compactness theory of the next chapter:

PROPOSITION 6.5.6. *For each choice of the data $g, m, A, \boldsymbol{\gamma}^\pm$, the energy $E(u)$ defined in §6.2.3 satisfies a uniform bound for all $u \in \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$.* □

EXAMPLE 6.5.7. If W is taken to be a one-point space, then asymptotically cylindrical maps $u : \dot{\Sigma} \to W$ are constant and thus cannot have any punctures, so elements of $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ can be regarded as tuples (Σ, j, Θ) consisting of a closed genus g Riemann surface (Σ, j) and an ordered set of m points $\Theta \subset \Sigma$. Such a tuple is also called a **marked Riemann surface**, and the symbol

$$\mathcal{M}_{g,m}$$

is used to abbreviate the moduli space in this case. Two marked Riemann surfaces $(\Sigma_0, j_0, \Theta_0)$ and $(\Sigma_1, j_1, \Theta_1)$ are **equivalent** (and thus represent the same element of $\mathcal{M}_{g,m}$) if there exists a biholomorphic map $\psi : (\Sigma_0, j_0) \to (\Sigma_1, j_1)$ sending Θ_0 to Θ_1 with the ordering preserved, and the automorphism group

$$\mathrm{Aut}(\Sigma, j, \Theta)$$

consists of all biholomorphic maps $(\Sigma, j) \to (\Sigma, j)$ that fix Θ. The topology of $\mathcal{M}_{g,m}$ is determined by the condition that $[(\Sigma_\nu, j_\nu, \Theta_\nu)] \to [(\Sigma, j, \Theta)]$ if and only if each $(\Sigma_\nu, j_\nu, \Theta_\nu)$ is equivalent for sufficiently large ν to a marked Riemann surface of the form (Σ, j'_ν, Θ) such that j'_ν is C^∞-convergent to j. Finally, we call a marked Riemann surface (Σ, j, Θ) **stable** if $\chi(\Sigma \backslash \Theta) < 0$; note that the condition in Remark 6.4.3 about maps $u : \dot{\Sigma} \to \widehat{W}$ being nonconstant can never be achieved in this case.

6.6. Orbits in families

For situations in which Reeb orbits are not isolated—for instance when they exist in smooth Morse-Bott families—some small modifications to the definitions of the previous sections are appropriate. Instead of ordered sets of fixed Reeb orbits $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \ldots, \gamma_{k_\pm}^\pm)$, we can consider ordered sets

$$\mathcal{P}^\pm = (\mathcal{P}_1^\pm, \ldots, \mathcal{P}_{k_\pm}^\pm)$$

in which each $\mathcal{P}_i^\pm$ denotes a connected component in the space $\mathcal{P}(\mathcal{H}_\pm)$ of unparametrized closed Reeb orbits on $(M_\pm, \mathcal{H}_\pm)$, and then define

$$\mathcal{M}_{g,m}(J, \mathcal{P}^+, \mathcal{P}^-) := \bigcup_{\boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-} \mathcal{M}_{g,m}(J, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-),$$

where the union is over the set of all pairs of tuples $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \ldots, \gamma_{k_\pm}^\pm)$ such that $\gamma_i^\pm \in \mathcal{P}_i^\pm$ for every $i = 1, \ldots, k_\pm$. The previously defined notion of convergence allows for convergent sequences $u_\nu \to u$ in $\mathcal{M}_{g,m}(J, \mathcal{P}^+, \mathcal{P}^-)$ in which the asymptotic orbits vary with ν but converge as $\nu \to \infty$.

Some thought is required before $\mathcal{M}_{g,m}(J, \mathcal{P}^+, \mathcal{P}^-)$ can be split up into open and closed subsets $\mathcal{M}_{g,m}(J, A, \mathcal{P}^+, \mathcal{P}^-)$ that keep track of relative homology classes. An obvious first guess would be that one should let $\bar{\mathcal{P}}^\pm \subset M_\pm$ denote the union of the images of all the orbits in the families $\mathcal{P}_1^\pm, \ldots, \mathcal{P}_{k_\pm}^\pm$, and then let

$$[u] \in H_2(W, \bar{\mathcal{P}}^+ \cup \bar{\mathcal{P}}^-)$$

denote the image of the usual relative homology class $[u] \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ under the map $H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-) \to H_2(W, \bar{\mathcal{P}}^+ \cup \bar{\mathcal{P}}^-)$ induced by inclusions. The sets $\bar{\mathcal{P}}_i^\pm \subset M_\pm$ are nice objects if the orbits are Morse-Bott—they are then smooth submanifolds—and the class $[u] \in H_2(W, \bar{\mathcal{P}}^+ \cup \bar{\mathcal{P}}^-)$ defined in this way is clearly constant on each connected component of $\mathcal{M}_{g,m}(J, \mathcal{P}^+, \mathcal{P}^-)$. But there is a problem: The analogue of Proposition 6.5.6 in this context will not hold, as $[u] \in H_2(W, \bar{\mathcal{P}}^+ \cup \bar{\mathcal{P}}^-)$ does not constrain the topology of the map u tightly enough to achieve a uniform energy bound. To see this, consider the $k = 0$ case of the pre-quantization bundle in Example 3.12.13: this gives a manifold of the form $M = X \times S^1$ with a stable Hamiltonian structure (ω, dt), where ω is a symplectic form on X (pulled back under the obvious projection $M \to X$) and t denotes the coordinate on S^1. Every fiber of the trivial S^1-fibration $M \to X$ is now a closed Reeb orbit, so the whole of M is a Morse-Bott submanifold and $H_2(M, \bar{\mathcal{P}}^+ \cup \bar{\mathcal{P}}^-)$ will therefore always be trivial. But since $[\omega]$ is nontrivial on $H_2(X)$, there might very well exist sequences of J-holomorphic curves with fixed asymptotic orbits but unbounded energy, e.g. one sees this already by considering closed curves with image in a single fiber. The problem here is that too much information was lost when we chose to quotient out all 2-chains in Morse-Bott submanifolds.

For a more useful approach, we observe that in any manifold M with stable Hamiltonian structure $\mathcal{H}$, every subset $\mathcal{P}_0 \subset \mathcal{P}(\mathcal{H})$ determines a distinguished subgroup

$$G_{\mathcal{P}_0} \subset H_2(M)$$

that is generated by all maps $\mathbb{T}^2 \to M$ of the form $(\rho, t) \mapsto \gamma_\rho(t)$ for families of parametrized orbits $\{\gamma_\rho : S^1 \to M\}_{\rho \in S^1}$ representing loops in $\mathcal{P}_0$. In fancier terms, we can let $\widetilde{\mathcal{P}}_0 \subset \widetilde{\mathcal{P}}(\mathcal{H})$ denote the set of all parametrizations of orbits in $\mathcal{P}_0$ and define $G_{\mathcal{P}_0}$ as the image of the composition

$$H_1(\widetilde{\mathcal{P}}) = H_1(\widetilde{\mathcal{P}}) \otimes H_1(S^1) \xrightarrow{\times} H_2(\widetilde{\mathcal{P}} \times S^1) \to H_2(M),$$

where the last homomorphism is induced by the evaluation map $\widetilde{\mathcal{P}} \times S^1 \to M : (\gamma, t) \mapsto \gamma(t)$. For the chosen connected families of orbits $\mathcal{P}_1^\pm, \ldots, \mathcal{P}_{k_\pm}^\pm \subset \mathcal{P}(\mathcal{H}_\pm)$ in $M_\pm$, let us abbreviate $\mathcal{P}^\pm := \mathcal{P}_1^\pm \cup \ldots \cup \mathcal{P}_{k_\pm}^\pm$ and define

$$G_{\mathcal{P}} \subset H_2(W)$$

as the image of the map

$$G_{\mathcal{P}^+} \oplus G_{\mathcal{P}^-} \hookrightarrow H_2(M_+) \oplus H_2(M_-) = H_2(M_+ \coprod M_-) \to H_2(W)$$

induced by the inclusion $M_+ \coprod M_- = \partial W \hookrightarrow W$.

Next, choose a "base point" in each family $\mathcal{P}_i^\pm$, meaning a specific orbit

$$\gamma_i^\pm \in \mathcal{P}_i^\pm,$$

and write $\bar{\boldsymbol{\gamma}}^\pm \subset M_\pm$ as before for the 1-dimensional submanifold formed by the union of the images of these orbits. The inclusion $(W, \varnothing) \hookrightarrow (W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ induces an isomorphism of $H_2(W)$ with the set of classes $A \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ that are annihilated by the connecting homomorphism in the long exact sequence of the pair $(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$, so we can also view $G_{\mathcal{P}}$ as a subgroup of $H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$. Each $u \in \mathcal{M}_{g,m}(J, \mathcal{P}^+, \mathcal{P}^-)$ then determines a class

$$[u] \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)/G_{\mathcal{P}},$$

which is defined by gluing to each boundary component of the map $\bar{u} : \bar{\Sigma} \to \overline{W}$ an annulus in $\partial\overline{W} = M_+ \coprod M_-$ determined by a path in the respective connected family $\mathcal{P}_i^\pm$ from the asymptotic orbit of u to the chosen base point $\gamma_i^\pm \in \mathcal{P}_i^\pm$. The resulting class in $H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ will generally depend on the choices of paths in the families $\mathcal{P}_i^\pm$, but the ambiguity lies in the subgroup $G_{\mathcal{P}}$ and thus disappears after projecting to the quotient. For each $A \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)/G_{\mathcal{P}}$, we define

$$\mathcal{M}_{g,m}(J, A, \mathcal{P}^+, \mathcal{P}^-) := \left\{ u \in \mathcal{M}_{g,m}(J, \mathcal{P}^+, \mathcal{P}^-) \;\middle|\; [u] = A \right\},$$

which is again an open and closed subset of $\mathcal{M}_{g,m}(J, \mathcal{P}^+, \mathcal{P}^-)$.

The following lemma permits a generalization of the uniform energy bound from Proposition 6.5.6 to the Morse-Bott setting, and the detail about c_1 will be similarly important when we write down dimension formulas in Chapter 8.

LEMMA 6.6.1. *Suppose $\mathcal{P} \subset \mathcal{P}(\mathcal{H})$ is a union of Morse-Bott families of Reeb orbits in a manifold M with stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$, and choose any ω-compatible complex structure on the vector bundle $\xi = \ker \lambda \subset TM$. Then the cohomology classes $[\omega] \in H^2_{\mathrm{dR}}(M)$ and $c_1(\xi) \in H^2(M)$ both vanish when evaluated on the subgroup $G_{\mathcal{P}} \subset H_2(M)$.*

PROOF. For $[\omega]$, it suffices to observe that $\int_{\mathbb{T}^2} f^*\omega = 0$ for any smooth map $f : \mathbb{T}^2 \to M$ such that $f(\rho, \cdot) : S^1 \to M$ parametrizes an orbit in $\mathcal{P}$ for every $\rho \in S^1$, and this holds because the Reeb vector field takes values in $\ker \omega$. The statement about $c_1(\xi)$ follows from Corollary 3.12.18, as the bundle $f^*\xi \to \mathbb{T}^2$ for any such map f is trivial. □

6.7. Asymptotic regularity

For the analytic setup in later chapters, we will need to use exponentially weighted Sobolev spaces, thus we need to check that all asymptotically cylindrical holomorphic curves actually belong to such spaces. At the local level, this is already clear: since we are using smooth almost complex structures, the results of §2.4 imply that all J-holomorphic curves are smooth, and in particular, they are of class $W^{k,p}_{\mathrm{loc}}$ for every $k \in \mathbb{N}$ and $p \in (1, \infty)$. Similarly, convergence of a sequence of J-holomorphic curves in C^∞_{loc} is equivalent to convergence in $W^{k,p}_{\mathrm{loc}}$ for every k and p. It remains only to check that suitable decay conditions are satisfied on each of the cylindrical ends.

We recall the following notation from §4.6: for Sobolev parameters k, p and a real number $\delta \in \mathbb{R}$, the **exponentially weighted Sobolev space** of functions of class $W^{k,p,\delta}$ on the half-cylinder $\mathring{Z}_+ = (0, \infty) \times S^1$ or $\mathring{Z}_- = (-\infty, 0) \times S^1$ is

$$W^{k,p,\delta}(\mathring{Z}_\pm) := \left\{ e^{\mp \delta s} f \;\middle|\; f \in W^{k,p}(\mathring{Z}_\pm) \right\}.$$

This is a Banach space with respect to the norm

$$\|f\|_{W^{k,p,\delta}} := \|e^{\pm \delta s} f\|_{W^{k,p}},$$

and if $\delta > 0$, then its elements satisfy a forced exponential decay condition as $s \to \pm\infty$.

Recall also from Definition 6.4.1 the notion of the *asymptotic representative* of a holomorphic curve at a puncture.

PROPOSITION 6.7.1. *Assume $\mathcal{H} = (\omega, \lambda)$ is a stable Hamiltonian structure on a manifold M, $J \in \mathcal{J}(\mathcal{H})$, $\gamma : S^1 \to M$ is a Morse-Bott Reeb orbit and $\delta \in (0, 2\pi)$ is small enough so that the interval $(0, \delta]$ contains no eigenvalues of $\mp \mathbf{A}_\gamma$ for the asymptotic operator $\mathbf{A}_\gamma$ of γ. Fix an $\mathbb{R}$-invariant Riemannian metric on $\mathbb{R} \times M$ for the purpose of defining Sobolev norms on vector fields along asymptotically cylindrical maps $Z_\pm \to \mathbb{R} \times M$.*

Suppose $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is J-holomorphic and asymptotically cylindrical with a puncture $z \in \Gamma^\pm$ that is asymptotic to γ. Then its asymptotic repesentative at z with respect to any choice of holomorphic cylindrical coordinates belongs to $W^{k,p,\delta}(Z_\pm)$ for every $k \in \mathbb{N}$ and $p \in (1, \infty)$.

Further, suppose $J_\nu \in \mathcal{J}(\mathcal{H})$ is a C^∞-convergent sequence with $J_\nu \to J$, and $u_\nu : (\dot{\Sigma}, j_\nu) \to (\mathbb{R} \times M, J_\nu)$ is a sequence of asymptotically cylindrical J_ν-holomorphic curves that converge to $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ in the sense described in Definition 6.5.4, i.e. u_ν and j_ν are C^∞_{loc}-convergent away from the punctures and u_ν is converges asymptotically at each puncture to the respective asymptotic orbit of u. Then the asymptotic representatives of u_ν at z converge in $W^{k,p,\delta}(Z_\pm)$ for every $k \in \mathbb{N}$ and $p \in (1, \infty)$ to the asymptotic repesentative of u at z.

REMARK 6.7.2. The obvious analogue of Proposition 6.7.1 for curves in completed cobordisms also holds, with no meaningful change to the proof.

REMARK 6.7.3. The second statement in Proposition 6.7.1 obscures a slightly subtle detail about the convergence of u_ν to u. According to (6.6), the asymptotic representatives h_ν of u_ν at z are defined to satisfy the relation

$$u_\nu(s - s_\nu, t - t_\nu) = \exp_{u_{\gamma_\nu}(s,t)} h_\nu(s, t)$$

for all $(s, t) \in Z_\pm \cong \dot{\mathcal{U}}_z$ close enough to infinity, where $\gamma_\nu : S^1 \to M$ is a sequence of chosen parametrizations (converging to γ) of the asymptotic orbits of u_ν, and the shift parameters $s_\nu \in \mathbb{R}$ and $t_\nu \in S^1$ are uniquely determined by the map u_ν, the parametrizations $\gamma_\nu : S^1 \to M$, and the condition that h_ν vanish at infinity. These shift parameters s_ν, t_ν will in general vary with ν, though of course they converge to the corresponding constants for u as $\nu \to \infty$. This fact implies that even in the nondegenerate case, where $\gamma_\nu = \gamma$ for all ν, the maps u_ν cannot be regarded simply as $W^{k,p,\delta}$-small perturbations of u as $\nu \to \infty$, and we will have to be careful about this point when we define a suitable Banach manifold setting for these maps in Chapter 8.

Morally, the reason Proposition 6.7.1 holds is that asymptotic representatives, as sections of the bundle $u_\gamma^* T(\mathbb{R} \times M)$ near the ends, satisfy a linear Cauchy-Riemann type equation and are thus subject to the exponential decay results in §4.6. We will use Theorem 2.8.1 to show this, but there is a complication: The linear Cauchy-Riemann type operator obtained in this way is degenerate at infinity, even if γ is a nondegenerate orbit—degeneracy is guaranteed specifically in directions tangent to the trivial cylinder over γ, thus leaving open the possibility that the decay of a solution near infinity might fail to be exponentially fast (cf. Example 4.8.5). If γ is nondegenerate, then we can fix this by showing first that the component of h normal to the trivial cylinder has exponential decay, so that we can then apply the stronger version of the asymptotic formula in Theorem 4.8.2. The case of a degenerate Morse-Bott orbit requires more involved arguments which we will not discuss here, but instead refer to the original paper [**HWZ96**].

PROOF OF PROPOSITION 6.7.1 WHEN γ IS NONDEGENERATE. We consider for simplicity just the case of a positive puncture. The choice of holomorphic cylindrical coordinate neighborhood $Z_+ \cong \dot{\mathcal{U}}_z \subset \dot{\Sigma}$ can be adjusted by constant shifts in the coordinates so that without loss of generality,

$$u(s, t) = \exp_{u_\gamma(s,t)} h(s, t)$$

for $s \gg 0$, where the assumption on $h \in \Gamma(u_\gamma^* T(\mathbb{R} \times M))$ is that its derivatives of all orders converge to 0 uniformly in t as $s \to \infty$. Recall that the stable Hamiltonian structure determines a natural splitting

$$T(\mathbb{R} \times M) = \varepsilon \oplus \xi$$

of complex vector bundles, where ε denotes the canonically trivial complex line bundle spanned by the vector fields ∂_r and R. The trivial cylinder $u_\gamma : \mathbb{R} \times S^1 \to \mathbb{R} \times M$ is everywhere tangent to ε, thus $u_\gamma^* \xi$ can be regarded as the normal bundle of u_γ, and for $R > 0$ sufficiently large, there are uniquely determined maps

$$\varphi : Z_+^R \to \mathbb{R} \times S^1, \qquad \eta \in \Gamma(\varphi^* u_\gamma^* \xi)$$

such that for $s \geqslant R$,

$$u(s,t) = \exp_{u_\gamma(\varphi(s,t))} \eta(s,t),$$

while $\varphi|_{Z_+^r}$ can be assumed arbitrarily C^∞-close to the inclusion $Z_+^r \hookrightarrow \mathbb{R} \times S^1$ for r sufficiently large, and $\eta|_{Z_+^r}$ converges with all its derivatives uniformly to 0 as $r \to \infty$. In this situation, we can apply Theorem 2.8.3 (with Remark 2.8.4) to find a linear Cauchy-Riemann type operator $\mathbf{D}^N$ on $\varphi^* u_\gamma^* \xi$ that annihilates η, and moreover, the restriction of $\mathbf{D}^N$ to Z_+^r for sufficiently large $r > 0$ can be assumed arbitrarily close to the normal Cauchy-Riemann operator of the trivial cylinder u_γ. By Proposition 6.2.7, the latter takes the form $\partial_s - \mathbf{A}_\gamma$, thus proving that $\mathbf{D}^N$ is C^∞-asymptotic to the nondegenerate asymptotic operator $\mathbf{A}_\gamma$. This is enough to deduce from Theorem 4.8.2 that η and its derivatives of all orders are bounded by functions of the form $Ce^{-\delta s}$ for any $\delta > 0$ smaller than the smallest positive eigenvalue of $-\mathbf{A}_\gamma$.

We now turn attention back to the asymptotic representative $h \in \Gamma(u_\gamma^* T(\mathbb{R} \times M))$, which by Theorem 2.8.1 (and Remark 2.8.4), likewise satisfies $\mathbf{D}h = 0$ for some linear Cauchy-Riemann type operator $\mathbf{D}$ that is close to $\mathbf{D}_{u_\gamma}$ on Z_+^r for $r \gg 0$, and is therefore C^∞-asymptotic to $-i\partial_t \oplus \mathbf{A}_\gamma$. The exponential decay of η implies moreover that the convergence of $\mathbf{D}|_{Z_+^r}$ toward $\mathbf{D}_{u_\gamma}$ as $r \to \infty$ is exponentially fast, thus $\mathbf{D}$ is in fact $C^{\infty,\delta}$-asymptotic to $-i\partial_t \oplus \mathbf{A}_\gamma$, so that the stronger asymptotic formula (4.26) in Theorem 4.8.2 applies. Since h decays to 0 at infinity, the eigenvalue of $-i\partial_t \oplus \mathbf{A}_\gamma$ appearing in this formula cannot be 0, but must instead be strictly negative, and we conclude that all derivatives of h are bounded by functions of the form $Ce^{-\delta s}$, where one can take any $\delta > 0$ that is smaller than the smallest strictly positive eigenvalue of $i\partial_t \oplus (-\mathbf{A}_\gamma)$.

The case of a converging sequence is handled similarly: the point is that the resulting sequence of asymptotic representatives h_ν then satisfies a converging sequence of linear Cauchy-Riemann type equations $\mathbf{D}_\nu h_\nu = 0$ that are C^∞-asymptotic to $-i\partial_t \oplus \mathbf{A}_\gamma$. Since the h_ν are already known to decay exponentially, we can then multiply them by an exponential weighting function $e^{\epsilon s}$ for sufficiently small $\epsilon > 0$ in order to replace $-i\partial_t \oplus \mathbf{A}_\gamma$ with the nondegenerate asymptotic operator $(-i\partial_t \oplus \mathbf{A}_\gamma) + \epsilon$, and then apply Exercise 4.6.11. □

6.8. Simple curves and multiple covers revisited

In §2.6, we proved that closed J-holomorphic curves are all either embedded in the complement of a finite set or are multiple covers of curves with this property. The same thing holds in the punctured case:

THEOREM 6.8.1. *Assume* $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ *is a nonconstant asymptotically cylindrical* J*-holomorphic curve whose asymptotic orbits are all nondegenerate or Morse-Bott, where* $\dot{\Sigma} = \Sigma \backslash \Gamma$ *for some closed Riemann surface* (Σ, j) *and finite subset* $\Gamma \subset \Sigma$*. Then there exists a factorization* $u = v \circ \varphi$*, where*

- $\varphi : (\Sigma, j) \to (\Sigma', j')$ *is a holomorphic map of positive degree to another closed and connected Riemann surface* (Σ', j')*;*
- $v : (\dot{\Sigma}', j') \to (\widehat{W}, J)$ *is an asymptotically cylindrical* J*-holomorphic curve which is embedded except at a finite set of non-immersed points and self-intersections, where* $\dot{\Sigma}' := \Sigma' \backslash \Gamma'$ *with* $\Gamma' := \varphi(\Gamma)$ *and* $\Gamma = \varphi^{-1}(\Gamma')$*.*

As in the closed case, we call u a **simple** curve if the holomorphic map $\varphi : (\Sigma, j) \to (\Sigma', j')$ is a diffeomorphism, and u is otherwise a **k-fold multiple cover** of v with $k := \deg(\varphi) \geqslant 2$.

The proof of this theorem is an almost verbatim repeat of the proof of Theorem 2.6.1 in Chapter 2, but with one new ingredient added. Recall that in the closed case, our proof required two lemmas which described the local picture of a J-holomorphic curve $u : \dot{\Sigma} \to \widehat{W}$ near either a double point $u(z_0) = u(z_1)$ for $z_0 \neq z_1$ or a non-immersed point $du(z_0) = 0$. Both statements were completely local and thus equally valid for non-closed curves, but we now need similar statements to describe what kinds of singularities can appear in the neighborhood of a puncture. The following lemma is due to Siefring [Sie08], and follows from a "relative asymptotic formula" that describes the exponential decay of asymptotic representatives somewhat more precisely than Proposition 6.7.1 (cf. Lemma 16.2.1).

LEMMA 6.8.2 (Asymptotics). *Assume* $u : (\dot{\Sigma} = \Sigma \backslash \Gamma, j) \to (\widehat{W}, J)$ *is asymptotically cylindrical and is asymptotic at* $z_0 \in \Gamma$ *to a nondegenerate or Morse-Bott Reeb orbit. Then a punctured neighborhood* $\dot{\mathcal{U}}_{z_0} \subset \dot{\Sigma}$ *of* z_0 *can be identified biholomorphically with the punctured disk* $\dot{\mathbb{D}} = \mathbb{D} \backslash \{0\}$ *such that*

$$u(z) = v(z^k) \quad \textit{for} \quad z \in \dot{\mathbb{D}} \cong \dot{\mathcal{U}}_{z_0},$$

where, $k \in \mathbb{N}$ *and* $v : (\dot{\mathbb{D}}, i) \to (\widehat{W}, J)$ *is an embedded and asymptotically cylindrical* J*-holomorphic curve. Moreover, if* $u' : (\dot{\Sigma}' = \Sigma' \backslash \Gamma', j') \to (\widehat{W}, J)$ *is another asymptotically cylindrical curve with a puncture* $z_0' \in \Gamma'$*, then the images of* u *near* z_0 *and* u' *near* z_0' *are either identical or disjoint.* □

EXERCISE 6.8.3. With Lemma 6.8.2 in hand, adapt the proof of Theorem 2.6.1 in Chapter 2 to prove Theorem 6.8.1. If you get stuck, see [Nel15, §3.2].

PROPOSITION 6.8.4. *If* $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \in \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ *is represented by a simple curve, then* $\operatorname{Aut}(u)$ *is trivial. If it is represented by a* k*-fold cover of a simple curve, then* $|\operatorname{Aut}(u)| \leqslant k$*. In particular,* $\operatorname{Aut}(u)$ *is always finite.*[7]

[7] See Remark 6.4.3 for the case where u is constant.

PROOF. If u is simple, then it is a diffeomorphism onto its image in a small neighbourhood of some point, and any map φ satisfying $u = u \circ \varphi$ would be the identity on such a neighbourhood. By unique continuation, we conclude that $\mathrm{Aut}(u)$ is trivial. In general, if $u = v \circ \varphi$ for some simple

$$v : \Sigma' \to W$$

and

$$\varphi : \Sigma \to \Sigma'$$

a k-fold branched cover, we have

$$\mathrm{Aut}(u) = \{f : \Sigma \to \Sigma \mid v \circ \varphi \circ f = v \circ \varphi\}.$$

By a similar argument as in the previous case, knowing that v is simple implies we only need to look at solutions to

$$\varphi \circ f = \varphi.$$

Remove the set of branch points B from Σ' together with the set $\varphi^{-1}(B)$ from Σ, so that φ becomes an honest covering map. Any $\varphi \in \mathrm{Aut}(u)$ then defines a deck transformation of the cover, and for a cover of degree k, there are at most k such transformations. □

6.9. Possible generalizations

In this section, I would like to add a few remarks on the set of assumptions involved in our geometric setup, and which of them could possibly be relaxed. A certain amount of what I have to say on this subject is speculative, and should perhaps be taken with a grain of salt; in any case, the reader who is only interested in the standard setup for SFT may feel free to skip it.

6.9.1. Asymptotically cylindrical ends. When (W, ω) is a symplectic cobordism with stable boundary $(M_\pm, \mathcal{H}_\pm)$ and $J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ belongs to our distinguished class of almost complex structures, the completion $(\widehat{W}, J)$ is what is known as an **almost complex manifold with cylindrical ends**. In particular, it has the feature that J is translation-invariant on both ends outside of some compact subset. For certain applications, it is natural to consider a weaker variant of this condition, in which J is not translation-invariant and thus does not belong to $\mathcal{J}(\mathcal{H}_\pm)$ on any neighborhood of infinity, but has *asymptotic approach* to something that is translation-invariant. The precise condition suggested in [**BEH**[+]**03**] was as follows: if $\tau_c : \mathbb{R} \times M_\pm \to \mathbb{R} \times M_\pm$ denotes the translation map $(r, x) \mapsto (r + c, x)$ for $c \in \mathbb{R}$, then there exist $J_\pm \in \mathcal{J}(\mathcal{H}_\pm)$ such that

$$\tau_c^* J\big|_{[0,\infty)\times M_+} \to J_+ \text{ as } c \to \infty \quad \text{and} \quad \tau_c^* J\big|_{(-\infty,0]\times M_-} \to J_- \text{ as } c \to -\infty, \tag{6.8}$$

with uniform convergence of all derivatives. If $(\widehat{W}, J)$ satisfies this condition, it is known as an **almost complex manifold with asymptotically cylindrical ends**. It remains unclear whether any reasonable theory of J-holomorphic curves exists at this level of generality, though Bao [**Bao15**] has shown that the compactness results from [**BEH**[+]**03**] do extend under a stricter hypothesis that the convergence in (6.8) is exponentially fast. It seems very likely that the rest of the results in this book will also hold under Bao's hypothesis, but proving this would require some extra analytical effort that we would prefer to avoid, and it is in any case unnecessary for the development of symplectic and contact invariants. One concrete application of the compactness results from [**Bao15**] is to show that certain configurations of nodal J-holomorphic curves in an almost complex 4-manifold have the same geometric structure as the neighborhood of a singular point in a Lefschetz fibration; see [**Wen18**, Appendix A].

6.9.2. Tame but not compatible. In the analysis of closed J-holomorphic curves on a symplectic manifold (W, ω), it almost never matters whether J is assumed to be *compatible* with ω or only *tamed* by it. One encounters occasional situations in which a lemma is easier to prove under one of those assumptions than the other, e.g. tameness has the obvious advantage of being an open condition, while certain formulas take appealingly simpler forms in the compatible case. But almost everything that is important in the theory works either way.

For an odd-dimensional manifold M with a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$, we have defined the special class of translation-invariant almost complex structures $J \in \mathcal{J}(\mathcal{H})$ on $\mathbb{R} \times M$ with the property that $J|_\xi$ is compatible with $\omega|_\xi$, and there is a temptation to believe that replacing "compatible" with "tame" in this definition would be harmless. That is false. This is to say, while it seems possible that the analytical foundations of SFT might still work when $J|_\xi$ is only tamed by but not compatible with $\omega|_\xi$, this is by no means obvious; some nontrivial work would need to be done to prove it, and that work has not been done. The difficulty concerns the asymptotic operators

$$\mathbf{A}_\gamma = -J\nabla_t^\omega : \Gamma(\gamma^*\xi) \to \Gamma(\gamma^*\xi)$$

associated to closed Reeb orbits γ. We have seen in Proposition 6.2.7 that $\mathbf{A}_\gamma$ appears in the linearized Cauchy-Riemann operator for the trivial cylinder over γ, and for that reason, it will also appear in asymptotic expressions of linearized Cauchy-Riemann operators for arbitrary asymptotically cylindrical curves. When we study the local structure of the moduli space in the next few chapters, we will need those linearized Cauchy-Riemann operators to be Fredholm, and our proof of this in Chapter 4 made essential use of the fact that $\mathbf{A}_\gamma$ is L^2-symmetric. We have also invoked the symmetry of $\mathbf{A}_\gamma$ whenever we discussed exponential convergence of solutions at infinity, as in §4.6, §4.8 and §6.7, and the existing proofs of Lemma 6.8.2, which we used for establishing the dichotomy between simple and multiply covered curves, also require it.

The symmetry of $\mathbf{A}_\gamma$ was proved in Exercise 3.4.2, but this required $\omega|_\xi$ to be J-invariant, i.e. compatibility, not just tameness. Without compatibility, $\mathbf{A}_\gamma$ need not be symmetric, and its eigenvalues need not be real.

This is not necessarily a catastrophe, as the tameness of J does still give $\mathbf{A}_\gamma = -J\nabla_t^\omega$ some useful properties short of symmetry. This situation has an analogue in the finite-dimensional setting of Morse homology. The role of asymptotic operators in that setting is played by the Hessian $\nabla^2 f(x) : T_xM \to T_xM$ of a Morse function $f : M \to \mathbb{R}$ at a critical point $x \in M$, which appears in linearizations of the gradient-flow equation because $\nabla^2 f(x)$ is the linearization of the gradient vector field ∇f at a point in its zero-set. However, Morse homology can also be defined under a relaxed assumption, where instead of counting flow lines of the actual gradient of f with respect to a Riemannian metric, one counts flow lines of some other **gradient-like** vector field X on M, meaning

$$df(X) > 0 \text{ wherever } df \neq 0.$$

One can see by looking at f in local Morse coordinates that under this condition, X must vanish at the critical points of f, and for technical reasons, one usually needs to impose a more precise condition on the behavior of X near those points, e.g. that for some choice of Riemannian metric on M there exists a constant $\delta > 0$ such that

$$df(X) \geqslant \delta \left(|X|^2 + |df|^2\right).$$

If one now linearizes the flow equation for X, the term that appears near $\pm\infty$ for each flow line is no longer the Hessian of f at critical points x, but rather the linearization $DX(x) : T_xM \to T_xM$ of the vector field at points in its zero-set. Such a linearization need not be symmetric, and for smooth vector fields in general, there are few constraints on what the linear map $DX(x) : T_xM \to T_xM$ may look like, beyond saying that for generic vector fields, it will be invertible. For gradient-like vector fields, however, there are constraints, e.g. nonzero eigenvalues of $DX(x) : T_xM \to T_xM$ must always have nontrivial real part (see [CE12, Lemma 9.9]).

The relevance of gradient-like vector fields to our discussion is that if $\mathcal{H} = (\omega, \lambda)$ is a stable Hamiltonian structure on M and $J : \xi \to \xi$ is ω-tame, then the "vector field" $V(\gamma) := -J\pi_\xi\dot{\gamma}$ on $C^\infty(S^1, M)$ is gradient-like with respect to the action functional $\mathcal{A}_\beta$ of (3.5) in §3.3, because

$$d\mathcal{A}_\beta(\omega)V(\gamma) = -\int_{S^1} \omega(\pi_\xi\dot{\gamma}, -J\pi_\xi\dot{\gamma})\, dt = \int_{S^1} \omega(\pi_\xi\dot{\gamma}, J\pi_\xi\dot{\gamma})\, dt \geqslant 0,$$

with strict inequality unless γ parametrizes a Reeb orbit. The asymptotic operator $\mathbf{A}_\gamma = -J\nabla_t^\omega : L^2(\gamma^*\xi) \supset H^1(\gamma^*\xi) \to L^2(\gamma^*\xi)$ is defined as the linearization of V at a Reeb orbit γ, so one can use these observations to prove as in the finite-dimensional case that no eigenvalue of $\mathbf{A}_\gamma$ can be purely imaginary unless it is 0. This added information is enough to generalize our proof of Theorem 4.4.1 on the invertibility of translation-invariant operators $\partial_s - \mathbf{A}_\gamma$ over the cylinder, which was the main technical step in the proof of the Fredholm property in Chapter 4. There remain other things to check, especially in the realm of asymptotic decay conditions, and one should not attempt to use the machinery of SFT in this greater generality without first writing down those details. But if I had to bet the life of one of my Ph.D. students,[8] I would bet that it works.

6.9.3. Framed but not stable. Every compact symplectic manifold (W, ω) with boundary can be viewed as a symplectic cobordism between odd-dimensional manifolds $(M_\pm, \mathcal{H}_\pm)$ endowed with framed Hamiltonian structures $\mathcal{H}_\pm = (\omega_\pm, \lambda_\pm)$. The collar neighborhoods (6.1) then give rise to a reasonable notion of a symplectic completion $(\widehat{W}, \omega_\varphi)$, admitting tame almost complex structures that belong to $\mathcal{J}(\mathcal{H}_\pm)$ on the cylindrical ends. In general, the framings $\lambda_\pm$ of $\mathcal{H}_\pm$ do not need to be

[8] Needless to say, I learned this expression from my Ph.D. advisor.

stable in order for this construction to make sense, and stability imposes an extra constraint, i.e. not every Hamiltonian structure admits a stable framing. However, we saw two reasons in this chapter why the theory of J-holomorphic curves may run into trouble if stability of $\lambda_\pm$ is not also assumed. The first reason concerns the definition of energy: the symplectic structure ω_φ on $\widehat{W}$ depends in general on the arbitrary choice of a function φ in the space $\mathcal{T}_0$ defined in (6.4), and for a non-stable Hamiltonian structure, ω_φ does not tame J for every choice of φ. We will see in Chapter 7 that the ability to choose $\varphi \in \mathcal{T}_0$ arbitrarily is essential, and as a consequence, the standard compactness theory for punctured J-holomorphic curves does not hold in the setting of cobordisms with non-stable boundary. This is not to say that no interesting theory exists: the work of Fish and Hofer [**FH23**] on so-called *feral* holomorphic curves shows that there is much wilder behavior in this setting than anything discussed in the present book, but it is not without interesting applications.

In any case, compactness is not the only feature of the SFT setup, and one can imagine applications for which this aspect of the theory is unimportant, or is trivial for other geometric reasons. Thus a valid question remains: Can other aspects of the fundamentals of SFT, such as the Fredholm and transversality theory, still be defined with respect to Hamiltonian structures that are not stable?

On this question, I am slightly more optimistic, but the answer as in §6.9.2 is that if it can be done, then some nontrivial amount of work would be required in proving it. The danger here is visible in Proposition 6.2.7: If $d\lambda(R, \cdot)$ does not vanish everywhere, then the linearized Cauchy-Riemann operator for a trivial cylinder does not take the form $\partial_s - \mathbf{A}$ for an asymptotic operator $\mathbf{A}$, and as a result, the linearized operators for asymptotically cylindrical curves in general will not fit into the scheme of the Fredholm theory we established in Chapter 4. On the other hand, it is quite easy to see that the particular consequence of Theorem 4.4.1 we will need in Chapter 8 holds anyway: the linearization along the trivial cylinder takes the form

$$\partial_s - \begin{pmatrix} -i\partial_t & -B \\ 0 & \mathbf{A}_\gamma \end{pmatrix} = \begin{pmatrix} \partial_s - (-i\partial_t) & B \\ 0 & \partial_s - \mathbf{A}_\gamma \end{pmatrix} \tag{6.9}$$

with respect to the splitting $u_\gamma^* T(\mathbb{R} \times M) = u_\gamma^* \varepsilon \oplus u_\gamma^* \xi$, for some bundle map $B : u_\gamma^* \xi \to u_\gamma^* \varepsilon$. Such upper-triangular operators are invertible whenever both of their diagonal terms are. Here, the upper left block presents us with a minor headache since $-i\partial_t$ is a degenerate asymptotic operator, but this can be rectified by working in exponentially weighted Sobolev spaces, which has the effect of adding a small constant to this operator to make it nondegenerate. The result is that the analysis behind our proof of the semi-Fredholm property in Chapter 4 actually does work in this more general context. Moreover, the non-symmetric operator appearing in the first matrix in (6.9) has the same spectral properties as an asymptotic operator: each of its eigenvalues is also an eigenvalue of either $-i\partial_t$ or $\mathbf{A}_\gamma$. There are again still some things to check, but it seems likely that all of our results thus far, and some portion of the local analysis of the moduli space in Chapters 8 and 9, could admit reasonable generalizations to the non-stable setting. We will not attempt to carry out such a generalization in this book, since we have no interesting applications for it in mind—the most important Hamiltonian structures are the examples from contact geometry and Floer homology discussed in §6.3, and these are of course stable.

6.9.4. SFT without symplectic structures? Let's not carried away: Whatever subset of the results in this book remains intact after removing symplectic structures entirely from the picture, one should clearly no longer refer to it as "symplectic" field theory. Nonetheless, a large portion of the theory of moduli spaces of *closed* J-holomorphic curves is valid in arbitrary almost complex manifolds with no taming symplectic form—the usual regularity results all hold, the moduli spaces are well defined, the dichotomy between simple and multiply covered curves still makes sense, and so does the main result of Chapters 8 and 9, namely that after a generic perturbation of J, the moduli space becomes a smooth manifold whose dimension is determined by the index formula in Chapter 5. What definitely does not work is Gromov's compactness theorem: One can define a purely analytical notion of energy for a J-holomorphic curve (essentially as the L^2-norm of its derivative, see [**MS12**]), but in the absence of any taming condition, there is no reason for this energy to be bounded. As we will see in Chapter 7, without uniform energy bounds, the moduli space cannot be expected to have a natural compactification. Generalizing to an arbitrary almost complex manifold with cylindrical ends will definitely not improve this situation, so let us accept from the start that without tameness, there will be no compactness theory.

It nonetheless seems reasonable to ask whether the Fredholm and transversality theory of SFT might still hold. In fact, if $(\widehat{W}, J)$ is an almost complex manifold with cylindrical ends $[0, \infty) \times M_+$ and/or $(-\infty, 0] \times M_-$ on which J belongs to $\mathcal{J}(\mathcal{H}_\pm)$ for stable Hamiltonian structures $\mathcal{H}_\pm$ on $M_\pm$, then the Fredholm and transversality theory will be absolutely fine: there is no need to have any symplectic structure on the original compact cobordism W. A more interesting question is whether the Hamiltonian structures on the cylindrical ends can also be dispensed with, i.e. we could assume that J is translation-invariant on the cylindrical ends and maps ∂_r to some vector fields $R_\pm$ on $M_\pm$, but place no further assumptions on these vector fields or on the maximal J-invariant subbundles $\xi_\pm \subset T(\{r\} \times M_\pm)$.

One now runs into a starker version of the problem already discussed in §6.9.2: the asymptotic operators that appear as asymptotic data for linearized Cauchy-Riemann operators take the form

$$\mathbf{A}_\gamma = -J\nabla_t : \Gamma(\gamma^*\xi) \to \Gamma(\gamma^*\xi),$$

where ∇ is a connection on $\gamma^*\xi$ determined by the linearized flow of R, but ξ does not carry any symplectic structure for this connection to preserve, and as a consequence, there is now virtually no constraint on the spectral properties of $\mathbf{A}_\gamma$. In particular, $\mathbf{A}_\gamma$ can have purely imaginary eigenvalues without being degenerate, in which case, the proof of Theorem 4.4.1 on translation-invariant operators $\partial_s - \mathbf{A}_\gamma$ cannot be rescued, and the Fredholm property will fail. This does not necessarily mean that the situation is hopeless, but anything further I could say on this topic would be pure speculation.

CHAPTER 7

Asymptotics and compactness

Contents

Moduli spaces of pseudoholomorphic curves are generally not compact, but they have natural *compactifications*, obtained by allowing certain types of curves with singular behavior. For closed holomorphic curves, this fact is known as *Gromov's compactness theorem*, and our main goal in this chapter is to state its generalization to punctured curves, which is usually called the *SFT compactness theorem*. The theorem was first proved in [**BEH**$^+$**03**] (see also [**CM05**] for an alternative approach), and we do not have space here to present a complete proof, but we can still describe the main geometric and analytical ideas behind it.

The overarching theme of this chapter is the notion of *bubbling*, of which we will see several examples. Bubbling arises in a natural way from elliptic regularity: Recall that in Chapter 2, we proved that whenever $kp > 2$, any uniformly $W^{k,p}$-bounded sequence u_ν of J-holomorphic curves for a smooth almost complex structure J is also uniformly C^m_{loc}-bounded for every $m \in \mathbb{N}$ (cf. Theorem 2.4.10). The Arzelà-Ascoli theorem implies that such sequences have C^∞_{loc}-convergent subsequences, and this is true in particular whenever u_ν is uniformly C^1-bounded, as a C^1-bound implies a $W^{1,p}$-bound with $p > 2$. Let us take note of this fact for future use:

Proposition 7.0.1. *If (W, J_ν) is a sequence of almost complex manifolds with $J_\nu \to J$ in C^∞, then any uniformly C^1-bounded sequence of J_ν-holomorphic maps $u_\nu : \mathbb{D} \to W$ has a subsequence convergent in C^∞_{loc} on $\mathring{\mathbb{D}}$.* □

If one wants to prove compactness for a moduli space of J-holomorphic curves, it therefore suffices in general to establish a C^1-bound. We will work mainly in settings where a weaker condition than this holds, namely that the curves u_ν have bounded *energy* $E(u_\nu) \geqslant 0$, defined typically as the integral of a taming symplectic form over u_ν, or (in the noncompact settings that we consider) the supremum of such integrals for a distinguished class of taming symplectic forms. Observe that if $u : (\mathbb{D}, i) \to (W, J)$ is J-holomorphic and J is tamed by a symplectic form ω, then $g(X, Y) := \frac{1}{2}[\omega(X, JY) + \omega(Y, JX)]$ defines a Riemannian metric on W such that in holomorphic coordinates $s + it \in \mathbb{D}$, the equation $\partial_s u + J\,\partial_t u = 0$ implies

$$
\begin{aligned}
u^*\omega(\partial_s, \partial_t) &= \omega(\partial_s u, \partial_t u) = \frac{1}{2}\left[\omega(\partial_s u, J\partial_s u) + \omega(\partial_t u, J\partial_t u)\right] \\
&= \frac{1}{2}\left(|\partial_s u|_g^2 + |\partial_t u|_g^2\right),
\end{aligned}
\tag{7.1}
$$

This shows that a uniform bound on $E(u_\nu) = \int_{\mathbb{D}} u_\nu^*\omega$ for a sequence of local J-holomorphic curves u_ν implies a uniform local $W^{1,2}$-bound. That is just short of the $W^{1,p}$-bound for $p > 2$ that is required for producing results like Proposition 7.0.1, but it will turn out to suffice for exerting tight control over the range of interesting things that can happen when C^1-bounds fail. In such cases, the sequence u_ν will not be compact, but we will see that it becomes compact after removing finitely many points from its domain, and near those points, one can take a sequence of reparametrizations to find additional nontrivial holomorphic curves in the limit, the so-called "bubbles". This is one of the ways that the "nodal" curves in Gromov's compactness theorem can arise, and we will see the same phenomenon at work in several other contexts as well.

7.1. Removal of singularities

As an important tool for use in the rest of this chapter, we begin with the following result from [**Gro85**]:

THEOREM 7.1.1 (Gromov's removable singularity theorem). *Assume (W, ω) is a symplectic manifold with a tame almost complex structure J, and $u : \mathbb{D}\backslash\{0\} \to W$ is a J-holomorphic curve that has its image contained in a compact subset of W and satisfies*

$$
\int_{\mathbb{D}\backslash\{0\}} u^*\omega < \infty.
$$

Then u admits a smooth extension to $\mathbb{D}$.

The most interesting part of the proof establishes that u has a *continuous* extension. Once that is achieved, elliptic regularity will imply that the extension is actually smooth: indeed, for a continuous map $u : \mathbb{D} \to W$ that is smooth and J-holomorphic on $\mathbb{D}\backslash\{0\}$ and satisfies $\int_{\mathbb{D}\backslash\{0\}} u^*\omega < \infty$, (7.1) implies that the first derivative of u on $\mathbb{D}\backslash\{0\}$ is in $L^2(\mathbb{D}\backslash\{0\})$, so that Exercise 2.2.1 implies $u \in W^{1,2}(\mathring{\mathbb{D}})$. The smoothness of u then follows from Theorem 2.4.15.

The proof of continuity will use as a black box the following additional result from [**Gro85**], which is closely related to a standard result about minimal surfaces.

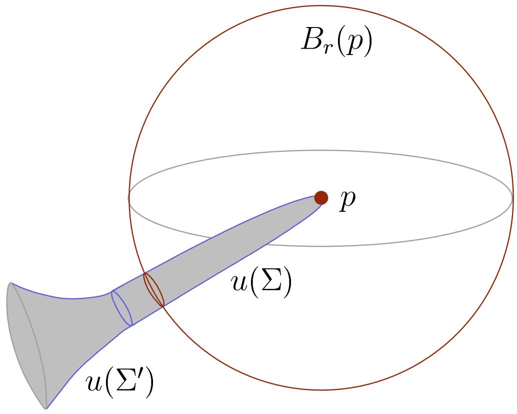

FIGURE 7.1. The intersection of a J-holomorphic curve u with an open ball $B_r(p)$ defines a proper map $\Sigma \to B_r(p)$. The monotonicity lemma prevents this map from having arbitrarily small area, assuming it passes through p.

THEOREM (Gromov's monotonicity lemma [Gro85]). *Suppose* (W, ω) *is a compact symplectic manifold (possibly with boundary),* J *is an* ω*-tame almost complex structure, and* $B_r(p) \subset W$ *denotes the open ball of radius* $r > 0$ *about* $p \in W$ *with respect to the Riemannian metric* $g(X,Y) := \frac{1}{2}\omega(X, JY) + \frac{1}{2}\omega(Y, JX)$*. Then there exist constants* $c, R > 0$ *such that for all* $r \in (0, R)$ *and* $p \in W$ *with* $B_r(p) \subset W$*, every proper non-constant* J*-holomorphic curve* $u : (\Sigma, j) \to (B_r(p), J)$ *passing through* p *satisfies*

$$\int_\Sigma u^*\omega \geqslant cr^2.$$

In the statement above, (Σ, j) is assumed to be an arbitrary (generally noncompact) Riemann surface *without boundary.* In applications, one typically has a larger (e.g. closed or punctured) domain Σ' in the picture, and Σ is defined to be the connected component of $u^{-1}(B_r(p)) \subset \Sigma'$ containing some point $z \in u^{-1}(p)$. The main message of the theorem is that u must use up at least a certain amount of energy for every ball whose center it passes through, so e.g. the portion of the curve passing through $B_r(p)$ cannot become arbitrarily "thin" as in Figure 7.1.

Returning to the removable singularity theorem, we shall use the biholomorphic map

$$Z_+ := [0, \infty) \times S^1 \to \mathbb{D}\backslash\{0\} : (s,t) \mapsto e^{-2\pi(s+it)}$$

to transform J-holomorphic maps $\mathbb{D}\backslash\{0\} \to W$ into maps $Z_+ \to W$, and the goal will be to show that whenever such a map u has precompact image and satisfies $\int_{Z_+} u^*\omega < \infty$, there exists a point $p \in W$ such that

$$u(s, \cdot) \to p \quad \text{in} \quad C^\infty(S^1, W) \text{ as } s \to \infty. \tag{7.2}$$

Fix the obvious flat metric on Z_+ and any Riemannian metric on W in order to define norms such as $|du(s,t)|$ for $(s,t) \in Z_+$. As usual, $\mathbb{D}_\epsilon \subset \mathbb{C}$ will denote the closed ϵ-disk about the origin for $\epsilon > 0$, and for a point z in either $\mathbb{C}$ or $\mathbb{R} \times S^1$, we write

$$\mathbb{D}_\epsilon(z) \subset \mathbb{C} \text{ or } \mathbb{R} \times S^1$$

for the closed disk of radius $\epsilon > 0$ about z, where in the case of the cylinder, ϵ will always be assumed small enough so that $\mathbb{D}_\epsilon(z) \subset \mathbb{C}$ projects to an embedded disk in $\mathbb{C}/i\mathbb{Z} \cong \mathbb{R} \times S^1$.

LEMMA 7.1.2. *Assume $u : (Z_+, i) \to (W, J)$ is a J-holomorphic map with precompact image satisfying $\int_{Z_+} u^*\omega < \infty$, where ω is a symplectic form on W taming the almost complex structure J. Then there exists a constant $C > 0$ such that $|du(s,t)| \leqslant C$ for all $(s,t) \in Z_+$.*

PROOF, PART 1. Arguing by contradiction, suppose there exists a sequence $z_\nu = (s_\nu, t_\nu) \in Z_+$ with $|du(z_\nu)| =: R_\nu \to \infty$. Choose a sequence of positive numbers $\epsilon_\nu > 0$ that converge to zero but not too fast, so that $\epsilon_\nu R_\nu \to \infty$. We then consider the sequence of reparametrized maps

$$v_\nu : \mathbb{D}_{\epsilon_\nu R_\nu} \to W : z \mapsto u(z_\nu + z/R_\nu).$$

These are also J-holomorphic since $z \mapsto z_\nu + z/R_\nu$ is holomorphic, and the values of v_ν depend only on the values of u over the ϵ_ν-disk about z_ν. Notice that since $s_\nu \to \infty$ and $\epsilon_\nu \to 0$, we are free to assume that all of these ϵ_ν-disks are disjoint; moreover, tameness of J implies $u^*\omega \geqslant 0$ and $v_\nu^*\omega \geqslant 0$, thus

$$\sum_\nu \int_{\mathbb{D}_{\epsilon_\nu R_\nu}} v_\nu^*\omega = \sum_\nu \int_{\mathbb{D}_{\epsilon_\nu}(z_\nu)} u^*\omega \leqslant \int_{Z_+} u^*\omega < \infty,$$

implying

$$\int_{\mathbb{D}_{\epsilon_\nu R_\nu}} v_\nu^*\omega \to 0 \quad \text{as} \quad k \to \infty. \tag{7.3}$$

We would now like to say something about a limit of the maps v_ν as $\nu \to \infty$, but this will require a brief pause in the proof, as we don't yet have quite enough information to do so. We know that the v_ν are uniformly C^0-bounded, since $u(Z_+)$ is contained in a compact subset. It would be ideal if we also had a uniform C^1-bound, as then elliptic regularity (Prop. 7.0.1) would give a C^∞_{loc} convergent subsequence on the union of all the domains $\mathbb{D}_{\epsilon_\nu R_\nu}$, i.e. on the entire plane. We have

$$dv_\nu(z) = \frac{1}{R_\nu} du(z_\nu + z/R_\nu),$$

hence $|dv_\nu(0)| = 1$, but we will need to know more about $|du|$ on the rest of $\mathbb{D}_{\epsilon_\nu}(z_\nu)$ in order to deduce a C^1-bound for v_ν on all of $\mathbb{D}_{\epsilon_\nu R_\nu}$. We'll come back to this in a moment. PROOF TO BE CONTINUED. . .

Here is the auxiliary lemma that is needed to complete the proof above. Its message for our present purposes is that the sequences z_ν and ϵ_ν can be improved by a small adjustment, so that the conditions $\epsilon_\nu \to 0$ and $\epsilon_\nu R_\nu \to \infty$ are both preserved, but we also obtain a bound on $|du_\nu|$ over $\mathbb{D}_{\epsilon_\nu}(z_\nu)$ in terms of R_ν.

LEMMA 7.1.3 (Hofer). *Suppose (X, d) is a complete metric space, $g : X \to [0, \infty)$ is continuous, $x_0 \in X$ and $\epsilon_0 > 0$. Then there exist $x \in X$ and $\epsilon > 0$ such that,*

(a) $\epsilon \leqslant \epsilon_0$,
(b) $g(x)\epsilon \geqslant g(x_0)\epsilon_0$,
(c) $d(x, x_0) \leqslant 2\epsilon_0$, *and*
(d) $g(y) \leqslant 2g(x)$ *for all* $y \in \overline{B_\epsilon(x)}$.

PROOF. If there is no $x_1 \in \overline{B_{\epsilon_0}(x_0)}$ such that $g(x_1) > 2g(x_0)$, then we can set $x = x_0$ and $\epsilon = \epsilon_0$ and are done. If such a point x_1 does exist, then we set $\epsilon_1 := \epsilon_0/2$ and repeat the process above for the pair (x_1, ϵ_1): that is, if there is no $x_2 \in \overline{B_{\epsilon_1}(x_1)}$ with $g(x_2) > 2g(x_1)$, we set $(x, \epsilon) = (x_1, \epsilon_1)$ and are finished, and otherwise define $\epsilon_2 = \epsilon_1/2$ and repeat for (x_2, ϵ_2). This process must eventually terminate, as otherwise we obtain a Cauchy sequence x_n with $g(x_n) \to \infty$, which is impossible if X is complete. □

PROOF OF LEMMA 7.1.2, PART 2. Applying Lemma 7.1.3 to $X = Z_+$ with $g(z) = |du(z)|$, we can replace the original sequences ϵ_ν and z_ν with new sequences for which all the previously stated properties still hold, but additionally,

$$|du(z)| \leqslant 2|du(z_\nu)| \quad \text{for all} \quad z \in \mathbb{D}_{\epsilon_\nu}(z_\nu).$$

Our sequence of reparametrizations v_ν then satisfies

$$|dv_\nu(z)| \leqslant 2 \quad \text{for all} \quad z \in \mathbb{D}_{\epsilon_\nu R_\nu},$$

so by elliptic regularity, v_ν has a subsequence convergent in $C^\infty_{\text{loc}}(\mathbb{C})$ to a J-holomorphic map

$$v_\infty : \mathbb{C} \to W$$

which is not constant since $|dv_\infty(0)| = \lim_{k\to\infty} |dv_\nu(0)| = 1$. Informally, we say that the blow-up of the derivatives at z_ν has caused a plane to "bubble off". However, (7.3) implies that for every $R > 0$, one can write $\epsilon_\nu R_\nu \geqslant R$ for ν sufficiently large and thus

$$\int_{\mathbb{D}_R} v_\infty^*\omega = \lim_{k\to\infty} \int_{\mathbb{D}_R} v_\nu^*\omega \leqslant \lim_{k\to\infty} \int_{\mathbb{D}_{\epsilon_\nu R_\nu}} v_\nu^*\omega = 0,$$

implying $\int_{\mathbb{C}} v_\infty^*\omega = 0$. It follows that v_∞ must be constant, so we have a contradiction. □

Continuing in the setting of Lemma 7.1.2, the next task is to obtain a uniform limit of $u(s, \cdot)$ as $s \to \infty$. Pick any sequence of nonnegative numbers $s_\nu \to \infty$, and consider the sequence of J-holomorphic half-cylinders

$$u_\nu : [-s_\nu, \infty) \times S^1 \to W : (s, t) \mapsto u(s + s_\nu, t).$$

By Lemma 7.1.2, these maps are uniformly C^1-bounded, so elliptic regularity gives a subsequence converging in C^∞_{loc} on $\mathbb{R} \times S^1$ to a J-holomorphic cylinder

$$u_\infty : \mathbb{R} \times S^1 \to W.$$

Observe that for any $c > 0$, we can write $-s_\nu/2 \leqslant -c$ for sufficiently large ν and thus compute

$$\int_{[-c,c]\times S^1} u_\infty^*\omega = \lim_{k\to\infty} \int_{[-c,c,]\times S^1} u_\nu^*\omega \leqslant \lim_{k\to\infty} \int_{[-s_\nu/2,\infty)\times S^1} u_\nu^*\omega$$
$$= \lim_{k\to\infty} \int_{[s_\nu/2,\infty)\times S^1} u^*\omega = 0$$

since $\int_{Z_+} u^*\omega < \infty$. This implies $\int_{\mathbb{R}\times S^1} u_\infty^*\omega = 0$, so u_∞ is a constant map to some point $p \in W$, hence after replacing s_ν with a subsequence,

$$u(s_\nu, \cdot) = u_\nu(0, \cdot) \to p \quad \text{in} \quad C^\infty(S^1, W) \text{ as } \nu \to \infty.$$

To finish the proof of (7.2), we need to show that one cannot find two sequences $s_\nu \to \infty$ and $s_\nu' \to \infty$ such that $u(s_\nu, \cdot) \to p$ and $u(s_\nu', \cdot) \to p'$ for distinct points $p \neq p' \in W$. This is an easy consequence of the monotonicity lemma: indeed, if two such sequences exist, then we can find a sequence $s_\nu'' \to \infty$ for which the loops $u(s_\nu'', \cdot)$ alternate between arbitrarily small neighborhoods of p and p'. Since u is continuous, it must then pass through $\partial B_{2r}(p)$ infinitely many times for $r > 0$ sufficiently small, and in fact there exists an infinite sequence of pairwise disjoint neighborhoods $\mathcal{U}_\nu \subset Z_+$ such that each

$$u|_{\mathcal{U}_\nu} : \mathcal{U}_\nu \to B_r(q_\nu)$$

is a proper map passing through some point $q_\nu \in \partial B_{2r}(p)$. The monotonicity lemma then implies

$$\int_{Z_+} u^*\omega \geqslant \sum_\nu \int_{\mathcal{U}_\nu} u^*\omega \geqslant \sum_\nu cr^2 = \infty,$$

a contradiction. This completes the proof that there is a continuous extension in the setting of the removable singularity theorem; the rest is elliptic regularity.

EXERCISE 7.1.4. Given an area form ω on $S^2 = \mathbb{C} \cup \{\infty\}$ and a finite subset $\Gamma \subset S^2$, show that a holomorphic function $f : S^2\backslash\Gamma \to \mathbb{C}$ has an essential singularity at one of its punctures if and only if $\int_{\mathbb{C}} f^*\omega = \infty$.

7.2. Finite energy and asymptotics

As further preparation for the compactness discussion, we now prove the converse of the observation that asymptotically cylindrical curves always have finite energy. We work in the setting described in §6.2.3: (W, ω) is a symplectic cobordism with stable boundary $\partial W = -M_- \coprod M_+$ carrying stable Hamiltonian structures $\mathcal{H}_\pm = (\omega_\pm, \lambda_\pm)$ with induced hyperplane distributions $\xi_\pm = \ker \lambda_\pm$ and Reeb vector fields $R_\pm$. The completion $(\widehat{W}, \omega_h)$ carries the symplectic structure

$$\omega_h := \begin{cases} d\left(h(r)\lambda_+\right) + \omega_+ & \text{on } [0,\infty) \times M_+ \\ \omega & \text{on } W, \\ d\left(h(r)\lambda_-\right) + \omega_- & \text{on } (-\infty, 0] \times M_-, \end{cases}$$

for some C^0-small smooth function $h(r)$ with $h' > 0$ that is the identity near $r = 0$, and for a fixed constant $r_0 \geqslant 0$, we define a compact subset

$$W^{r_0} := ([-r_0, 0] \times M_-) \cup_{M_-} W \cup_{M_+} ([0, r_0] \times M_+) \subset \widehat{W},$$

outside of which our ω_h-tame almost complex structures $J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ are required to be translation-invariant and compatible with $\mathcal{H}_\pm$. The **energy** of a J-holomorphic curve $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ is defined by

$$E(u) := \sup_{f \in \mathcal{T}} \int_{\dot{\Sigma}} u^* \omega_f,$$

where

$$\mathcal{T} := \left\{ f \in C^\infty(\mathbb{R}, (-\epsilon, \epsilon)) \;\middle|\; f' > 0 \text{ and } f \equiv h \text{ near } [-r_0, r_0] \right\}.$$

The constant $\epsilon > 0$ should always be assumed sufficiently small so that if $J_\pm \in \mathcal{J}(\mathcal{H}_\pm)$ and $X \in \xi_\pm$,

$$(\omega_\pm + \kappa\, d\lambda_\pm)(X, J_\pm X) > 0 \quad \text{whenever} \quad X \neq 0 \text{ and } \kappa \in (-2\epsilon, 2\epsilon). \tag{7.4}$$

This condition implies that every $J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ is tamed by ω_f for every $f \in \mathcal{T}$; cf. Proposition 6.2.2. It follows that all J-holomorphic curves satisfy $E(u) \geqslant 0$, with equality if and only if u is constant.

THEOREM 7.2.1. *Assume all closed Reeb orbits in $(M_+, \mathcal{H}_+)$ and $(M_-, \mathcal{H}_-)$ are nondegenerate, $J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$, (Σ, j) is a closed Riemann surface with $\dot{\Sigma} = \Sigma \backslash \Gamma$ for some finite subset $\Gamma \subset \Sigma$, and $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ is a J-holomorphic curve such that none of the singularities in Γ are removable and $E(u) < \infty$. Then u is asymptotically cylindrical.*

REMARK 7.2.2. The theorem also holds in the setting of a symplectization $(\mathbb{R} \times M, J)$ with $J \in \mathcal{J}(\mathcal{H})$ for a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ on M. The only real difference in this case is the slightly simpler definition of energy,

$$E(u) = \sup_{f \in \mathcal{T}} \int_{\dot{\Sigma}} u^* \omega_f,$$

where $\omega_f := d\big(f(r)\lambda\big) + \omega$ and

$$\mathcal{T} = \left\{ f \in C^\infty(\mathbb{R}, (-\epsilon, \epsilon)) \;\middle|\; f' > 0 \right\}.$$

This change necessitates a few trivial modifications to the proof of Theorem 7.2.1 given below.

Like removal of singularities, Theorem 7.2.1 is really a local result, so let us formulate a more precise and more general statement in these terms. Let

$$\dot{\mathbb{D}} := \mathbb{D} \backslash \{0\} \subset \mathbb{C}$$

and define the two biholomorphic maps

$$\begin{aligned} \varphi_+ &: Z_+ := [0, \infty) \times S^1 \to \dot{\mathbb{D}} : (s, t) \mapsto e^{-2\pi(s+it)} \\ \varphi_- &: Z_- := (-\infty, 0] \times S^1 \to \dot{\mathbb{D}} : (s, t) \mapsto e^{2\pi(s+it)}. \end{aligned} \tag{7.5}$$

THEOREM 7.2.3. *Suppose $J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ and $u : (\dot{\mathbb{D}}, i) \to (\widehat{W}, J)$ is a J-holomorphic map with $E(u) < \infty$. Then either the singularity at $0 \in \mathbb{D}$ is removable or u is a proper map. In the latter case the puncture is either positive or negative, meaning that u maps neighborhoods of 0 to neighborhoods of $\{\pm\infty\} \times M_\pm$, and the puncture has a well-defined* ***charge****, defined as*

$$Q = \lim_{\epsilon \to 0^+} \int_{\partial \mathbb{D}_\epsilon} u^* \lambda_\pm,$$

which satisfies $\pm Q > 0$. Moreover, the map

$$(u_\mathbb{R}(s,t), u_M(s,t)) := u \circ \varphi_\pm(s,t) \in \mathbb{R} \times M_\pm \quad \text{for } (s,t) \in Z_\pm \text{ near infinity}$$

satisfies

$$u_\mathbb{R}(s, \cdot) - Ts \to c \quad \text{in} \quad C^\infty(S^1) \text{ as } s \to \pm\infty$$

for $T := |Q|$ and a constant $c \in \mathbb{R}$, while for every sequence $s_\nu \to \pm\infty$, one can restrict to a subsequence such that

$$u_M(s_\nu, \cdot) \to \gamma(T \cdot) \quad \text{in} \quad C^\infty(S^1, M_\pm) \text{ as } \nu \to \infty$$

for some T-periodic Reeb orbit $\gamma : \mathbb{R}/T\mathbb{Z} \to M_\pm$. If γ is nondegenerate or Morse-Bott, then in fact

$$u_M(s, \cdot) \to \gamma(T \cdot) \quad \text{in} \quad C^\infty(S^1, M_\pm) \text{ as } s \to \pm\infty$$

We will not prove this result in its full strength, as in particular the last step (when γ is nondegenerate or Morse-Bott) requires some asymptotic elliptic regularity results that we do not have space to explain here. Note however that most of the statement above does not require any nondegeneracy assumption at all. The price for this level of generality is that if $s_\nu, s'_\nu \to \pm\infty$ are two distinct sequences, then we have no guarantee in general that the two Reeb orbits obtained as limits of subsequences of $u_M(s_\nu, \cdot)$ and $u_M(s'_\nu, \cdot)$ will be the same; an explicit example where they differ can be found in [Sie17]. If one of these orbits is assumed to be isolated, however—which is guaranteed if the orbit is nondegenerate—then we will be able to show that both are the same up to parametrization, hence *geometrically*, $u_M(s,t)$ lies in arbitrarily small neighborhoods of the orbit γ as $s \to \pm\infty$. This turns out to be also true in the more general Morse-Bott setting, though it is then much harder to prove, since γ need not be isolated. Once $u_M(s, \cdot)$ is localized near γ, one can use the nondegeneracy condition as in §6.7 to prove that $u_M(s, \cdot)$ converges exponentially fast to γ as $s \to \infty$. For details on this step, we refer to the original sources: [HWZ96, HWZ01] for the nondegenerate case, and [HWZ96, Bou02] when the Reeb vector field is Morse-Bott. Those papers deal exclusively with the contact case, but the setting of general stable Hamiltonian structures is also dealt with in [Sie08].

Ignoring the final step for now, the proof of Theorem 7.2.3 will reuse most of the techniques that we already saw in our proof of removal of singularities in §7.1. The main idea is to use a combination of the monotonicity lemma and bubbling analysis to show that unless u has a removable singularity, it is a proper map, and for any sequence $s_\nu \to \pm\infty$, the holomorphic maps defined by

$$u_\nu(s,t) = u \circ \varphi_\pm(s + s_\nu, t)$$

on a sequence of increasingly large half-cylinders must have a subsequence converging in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to either a constant map or a trivial cylinder. The first case will turn out to mean (as in Theorem 7.1.1) that the puncture is removable, and the second implies asymptotic convergence to a closed Reeb orbit.

One major difference between the proof of Theorem 7.2.3 and removal of singularities is that since $\widehat{W}$ is noncompact, sequences of curves in $\widehat{W}$ with uniformly bounded first derivatives need not be locally C^0-bounded. This issue will arise both in the bubbling argument to prove $|du_\nu(s,t)| \leqslant C$ and in the analysis of the sequence u_ν itself. In such cases, one can use the $\mathbb{R}$-translation action

$$\tau_c : \mathbb{R} \times M_\pm \to \mathbb{R} \times M_\pm : (r, x) \mapsto (r + c, x) \quad \text{for} \quad c \in \mathbb{R} \tag{7.6}$$

on suitable subsets of the cylindrical ends to replace unbounded sequences with uniformly C^1-bounded sequences of curves mapping into $\mathbb{R} \times M_+$ or $\mathbb{R} \times M_-$. These $\mathbb{R}$-translations are the reason why our definition of energy needs to be something slightly more complicated than just the symplectic area $\int_{\dot\Sigma} u^*\omega$ for a single choice of symplectic form. To understand bubbling in the presence of arbitrarily large $\mathbb{R}$-translations, we will need the following lemma.

LEMMA 7.2.4. *Suppose $J \in \mathcal{J}(\mathcal{H})$ for some stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ on a manifold M, and $u : (\dot\Sigma, j) \to (\mathbb{R} \times M, J)$ is a J-holomorphic curve satisfying*

$$E(u) < \infty \quad \text{and} \quad \int_{\dot\Sigma} u^*\omega = 0.$$

If $\dot\Sigma = \mathbb{C}$, then u is constant. If $\dot\Sigma = \mathbb{R} \times S^1$, then u either is constant or is biholomorphically equivalent to a trivial cylinder over a closed Reeb orbit.

PROOF. Denote $\xi = \ker \lambda$ and let

$$\pi_\xi : T(\mathbb{R} \times M) \to \xi$$

denote the projection along the subbundle spanned by ∂_r (the unit vector field in the $\mathbb{R}$-direction) and the Reeb vector field R. Then since ω annihilates both ∂_r and R, for any local holomorphic coordinates (s,t) on a subset of $\dot\Sigma$, the compatibility of $J|_\xi$ with $\omega|_\xi$ implies

$$u^*\omega(\partial_s, \partial_t) = \omega(\partial_s u, \partial_t u) = \omega(\partial_s u, J\partial_s u) = \omega(\pi_\xi \partial_s u, J\pi_\xi \partial_s u) \geqslant 0,$$

hence $\int_{\dot\Sigma} u^*\omega \geqslant 0$ for every J-holomorphic curve, and equality means that u is everywhere tangent to the subbundle spanned by ∂_r and R. This implies that $\operatorname{im} u$ is contained in the image of some J-holomorphic plane of the form

$$u_\gamma : \mathbb{C} \to \mathbb{R} \times M : s + it \mapsto (s, \gamma(t)),$$

where $\gamma : \mathbb{R} \to M$ is a (not necessarily periodic) orbit of R. If γ is not periodic, then u_γ is embedded, hence there exists a unique (and necessarily holomorphic) map $\Phi : (\dot\Sigma, j) \to (\mathbb{C}, i)$ such that $u = u_\gamma \circ \Phi$. If on the other hand γ is periodic with minimal period $T > 0$, then u_γ descends to an embedding of the cylinder

$$\hat{u}_\gamma : \mathbb{C}/iT\mathbb{Z} \to \mathbb{R} \times M,$$

and we can view u_γ as a covering map to this embedded cylinder. Now there exists a unique holomorphic map $\Phi : \dot{\Sigma} \to \mathbb{C}/iT\mathbb{Z}$ such that $u = \hat{u}_\gamma \circ \Phi$. If $\dot{\Sigma} = \mathbb{C}$, then $\pi_1(\mathbb{C}) = 0$ implies that Φ can be lifted to a (necessarily holomorphic) map $\widetilde{\Phi} : \mathbb{C} \to \mathbb{C}$ with $u_\gamma \circ \widetilde{\Phi} = u$. Relabelling symbols, we conclude that in general, if $\dot{\Sigma} = \mathbb{C}$, then $u = u_\gamma \circ \Phi$ for a holomorphic map $\Phi : \mathbb{C} \to \mathbb{C}$.

Let us consider all cases in which the factorzation $u = u_\gamma \circ \Phi$ exists, where $\Phi : (\dot{\Sigma}, j) \to (\mathbb{C}, i)$ is holomorphic and $\dot{\Sigma} = \Sigma \backslash \Gamma$ for a closed Riemann surface (Σ, j). We will now use the removable singularity theorem for $\Phi : \dot{\Sigma} \to S^2 \backslash \{0\}$ to show that unless Φ is constant, $\int_{\dot{\Sigma}} u^* \omega_f = \infty$ for suitable choices of $f \in \mathcal{T}$. This integral can be rewritten as

$$\int_{\dot{\Sigma}} u^* \omega_f = \int_{\dot{\Sigma}} \Phi^* u_\gamma^* \omega_f = \int_{\dot{\Sigma}} \Phi^* d\left(f(s)\, dt\right) = \int_{\dot{\Sigma}} \Phi^* \left(f'(s)\, ds \wedge dt\right) \tag{7.7}$$

since $\omega_f = d\big(f(r)\,\lambda\big) + \omega$ and $u_\gamma(s,t) = (s, \gamma(t))$. Since $f' > 0$, $f'(s)\, ds \wedge dt$ is an area form on $\mathbb{C}$ with infinite area.

We claim now that for suitable choices of $f \in \mathcal{T}$, one can find an area form Ω on $S^2 = \mathbb{C} \cup \{\infty\}$ such that $\Omega \leqslant f'(s)\, ds \wedge dt$. To see this, let us change coordinates so that ∞ becomes 0: the diffeomorphism $\Psi : \mathbb{C}^* \to \mathbb{C}^* : z \mapsto 1/z$ is holomorphic and thus satisfies $\frac{\partial \Psi}{\partial \bar{z}} = \frac{\partial \bar{\Psi}}{\partial z} = 0$, so we have

$$\begin{aligned} \Psi^* \left(f'(s)\, ds \wedge dt\right) &= -\frac{1}{2i} \Psi^* \left(f'(s)\, dz \wedge d\bar{z}\right) = -\frac{1}{2i} f'(\operatorname{Re} \Psi)\, d\Psi \wedge d\bar{\Psi} \\ &= -\frac{1}{2i} f'(\operatorname{Re} \Psi)\, \frac{\partial \Psi}{\partial z}\, dz \wedge \frac{\partial \bar{\Psi}}{\partial \bar{z}}\, d\bar{z} \\ &= -\frac{1}{2i} f'(s/|z|^2) \left(-\frac{1}{z^2}\right) \left(-\frac{1}{\bar{z}^2}\right) dz \wedge d\bar{z} \\ &= \frac{f'(s/|z|^2)}{|z|^4}\, ds \wedge dt \qquad \text{for} \qquad z = s + it \in \mathbb{C} \backslash \{0\}. \end{aligned} \tag{7.8}$$

We need to show that this 2-form can be bounded away from 0 as $z \to 0$. Let us choose $f \in \mathcal{T}$ such that

$$f(\pm r) = \pm \left(\epsilon - \frac{\epsilon}{2r}\right) \qquad \text{for} \qquad r \geqslant 1 \tag{7.9}$$

and extend f arbitrarily to $[-1, 1]$ such that $f' > 0$. We can then find a constant $c > 0$ such that f' satisfies

$$f'(r) \geqslant \min \left\{ c, \frac{\epsilon}{2r^2} \right\} \qquad \text{for all} \qquad r \in \mathbb{R}.$$

Plugging this into (7.8) gives

$$\Psi^* \left(f'(s)\, ds \wedge dt\right) \geqslant \min \left\{ \frac{c}{|z|^4}, \frac{\epsilon}{2s^2} \right\} ds \wedge dt,$$

which clearly blows up as $|z| \to 0$, proving the claim.

With this established, we observe that for any number $C > 0$, the fact that $f'(s)\, ds \wedge dt$ has infinite area implies we can choose an area form Ω on S^2 with

$$\Omega \leqslant f'(s)\, ds \wedge dt \text{ on } S^2 \backslash \{\infty\} \qquad \text{and} \qquad \int_{S^2} \Omega > C.$$

We now have two possibilities:

(1) If $\int_{\dot{\Sigma}} \Phi^*\Omega < \infty$, then Theorem 7.1.1 implies that the singularities of $\Phi : \dot{\Sigma} \to \mathbb{C} \subset S^2$ at Γ are all removable, i.e. Φ extends to a holomorphic map $(\Sigma, j) \to (S^2, i)$, which has a well-defined mapping degree $k \geqslant 0$. Then

$$\int_{\dot{\Sigma}} u^*\omega_f = \int_{\dot{\Sigma}} \Phi^* \left(f'(s)\, ds \wedge dt \right) \geqslant \int_{\dot{\Sigma}} \Phi^*\Omega = \int_{\Sigma} \Phi^*\Omega = k \int_{S^2} \Omega > kC.$$

Since $C > 0$ can be chosen arbitrarily large, this implies $\int_{\dot{\Sigma}} u^*\omega_f = \infty$ unless $k = 0$, meaning Φ is constant.

(2) If $\int_{\dot{\Sigma}} \Phi^*\Omega = \infty$ (meaning there is an essential singularity, cf. Exercise 7.1.4), then since $\Phi^* \left(f'(s)\, ds \wedge dt \right) \geqslant \Phi^*\Omega$, (7.7) implies $\int_{\mathbb{C}} u^*\omega_f = \infty$.

Since u is constant whenever Φ is, this completes the proof for $\dot{\Sigma} = \mathbb{C}$.

If $\dot{\Sigma} = \mathbb{R} \times S^1$, then it remains to deal with the case where the factorization $u = u_\gamma \circ \Phi$ does not exist because γ is periodic. If the minimal period is $T > 0$, then let us in this case redefine u_γ as an embedded J-holomorphic trivial cylinder

$$u_\gamma : \mathbb{R} \times S^1 \to \mathbb{R} \times M : (s, t) \mapsto (Ts, \gamma(Tt)).$$

Since the new u_γ is embedded, we can now write $u = u_\gamma \circ \Phi$ for a unique holomorphic map $\Phi : \mathbb{R} \times S^1 \to \mathbb{R} \times S^1$. Identifying $\mathbb{R} \times S^1$ biholomorphically with $S^2 \backslash \{0, \infty\}$, we claim that Φ extends to a holomorphic map $S^2 \to S^2$. Indeed, by the removable singularity theorem, this is true if and only if $\int_{\mathbb{R} \times S^1} \Phi^*\Omega < \infty$ for some area form Ω on S^2. Notice that $u_\gamma^*\omega_f = T^2 \cdot f'(Ts)\, ds \wedge dt$, defines an area form on $\mathbb{R} \times S^1$ with finite area for any $f \in \mathcal{T}$ since $\int_{-\infty}^{\infty} f'(s)\, ds < \infty$; this is equivalent to the observation that trivial cylinders always have finite energy. Using the biholomorphic map $(s, t) \mapsto e^{2\pi(s+it)}$ to identify $\mathbb{R} \times S^1$ with $\mathbb{C}^* = S^2 \backslash \{0, \infty\}$ and using coordinates $z = x + iy$ on the latter, another computation along the lines of (7.8) gives

$$u_\gamma^*\omega_f = \frac{T^2}{4\pi^2} \frac{f'\left(\frac{T}{2\pi} \log |z|\right)}{|z|^2}\, dx \wedge dy \qquad \text{for} \qquad z = x + iy \in \mathbb{C}^*.$$

Now suppose $f \in \mathcal{T}$ is chosen as in (7.9). Then one can check that the positive function in front of $dx \wedge dy$ in the above formula goes to $+\infty$ as $|z| \to 0$; this means that one can find an area form Ω on $\mathbb{C}$ with $\Omega \leqslant u_\gamma^*\omega_f$ on $\mathbb{C}^*$. The singularity at $+\infty \in S^2$ can be handled in a similar way, thus we can find an area form Ω on S^2 such that $\Omega \leqslant u_\gamma^*\omega_f$ on $\mathbb{R} \times S^1$. Now since $E(u) < \infty$, we have

$$\int_{\mathbb{R} \times S^1} \Phi^*\Omega \leqslant \int_{\mathbb{R} \times S^1} \Phi^* u_\gamma^*\omega_f = \int_{\mathbb{R} \times S^1} u^*\omega_f < \infty,$$

so by Theorem 7.1.1, Φ has a holomorphic extension $S^2 \to S^2$, which is then a map of degree $k \geqslant 0$ with $\Phi^{-1}(\{0, \infty\}) \subset \{0, \infty\}$. If $k = 0$ then Φ is constant, and so is u. Otherwise, Φ is surjective and thus hits both 0 and ∞, but it can only do this at either 0 or ∞, thus it either fixes both or interchanges them. After composing with a biholomorphic map of S^2 preserving $\mathbb{R} \times S^1$, we may assume without loss of generality that $\Phi(0) = 0$ and $\Phi(\infty) = \infty$. This makes Φ a polynomial with only one zero, hence as a map on $\mathbb{C} \cup \{\infty\}$, $\Phi(z) = cz^k$ for some $c \in \mathbb{C}^*$. Up to biholomorphic equivalence, $\Phi(z)$ is then z^k, which appears in cylindrical coordinates as the map $(s,t) \mapsto (ks, kt)$, so u is now the trivial cylinder

$$u(s,t) = u_\gamma(ks,kt) = (kTs, \gamma(kTt))$$

over the k-fold cover of γ. □

REMARK 7.2.5. It is useful in some applications to observe that Lemma 7.2.4 does not require M to be compact. In contrast, the compactness arguments in this chapter almost always depend on the assumption that W and $M_\pm$ are compact—without this, one would need to add some explicit assumption to guarantee local C^0-bounds on sequences of holomorphic curves, e.g. the assumption in Theorem 7.1.1 that $u(\mathbb{D}\backslash\{0\})$ is contained in a compact subset.

Before continuing, it is worth noting that neither of the two definitions of energy we've been using—one for curves in $\widehat{W}$ and the other for symplectizations—is unique, i.e. each can be tweaked in various ways such that the results of this section still hold. Indeed, the original definitions appearing in [**Hof93**, **BEH+03**] are slightly different, but equivalent to these. The next lemma illustrates one further example of this freedom, which will be useful in some of the arguments below.

LEMMA 7.2.6. *Given a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ on M, a sufficiently small constant $\epsilon > 0$ as in (7.4), and $J \in \mathcal{J}(\mathcal{H})$, consider the alternative notion of energy for J-holomorphic curves $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ defined by*

$$E_0(u) = \sup_{f \in \mathcal{T}_0} \int_{\dot{\Sigma}} u^*\omega_f$$

where $\omega_f = d\,(f(r)\,\lambda) + \omega$ and

$$\mathcal{T}_0 = \left\{ f \in C^\infty(\mathbb{R}, (a,b)) \;\middle|\; f' > 0 \right\}$$

for some constants $-\epsilon \leqslant a < b \leqslant \epsilon$. Then if $E(u)$ denotes the energy as written in Remark 7.2.2, there exists a constant $c > 0$, depending on the data a, b, ϵ and $\mathcal{H}$ but not on u, such that

$$cE(u) \leqslant E_0(u) \leqslant E(u).$$

PROOF. The second of the two inequalities is immediate since $\mathcal{T}_0 \subset \mathcal{T}$. For the first inequality, note that since $\epsilon > 0$ is small, we can assume there exists a constant $c > 1$ such that for every $X \in T(\mathbb{R} \times M)$ and every $\kappa \in [-\epsilon, \epsilon]$,

$$\frac{1}{c}(\omega + \kappa\, d\lambda)(X, JX) \leqslant \omega(X, JX) \leqslant c(\omega + \kappa\, d\lambda)(X, JX). \tag{7.10}$$

This uses (7.4) and the fact that $d\lambda$ annihilates $\ker \omega$. Now suppose $f \in \mathcal{T}$, choose a constant $\delta \in (0, b-a]$ and define $\tilde{f} \in \mathcal{T}_0$ by

$$\tilde{f}(r) = \frac{\delta}{2\epsilon} f(r) + \frac{a+b}{2}.$$

Then $\tilde{f}'(r) = \frac{\delta}{2\epsilon} f'(r)$, and given a J-holomorphic curve $u : \dot{\Sigma} \to \mathbb{R} \times M$, we can write $\omega_f = \omega + f(r)\, d\lambda + f'(r)\, dr \wedge \lambda$ and use (7.10) to estimate

$$\begin{aligned} \int_{\dot{\Sigma}} u^* \omega_f &= \int_{\dot{\Sigma}} u^* \left(\omega + f(r)\, d\lambda\right) + \int_{\dot{\Sigma}} u^* \left(f'(r)\, dr \wedge \lambda\right) \\ &\leqslant c \int_{\dot{\Sigma}} u^* \omega + \frac{2\epsilon}{\delta} \int_{\dot{\Sigma}} u^* \left(\tilde{f}'(r)\, dr \wedge \lambda\right) \\ &\leqslant c^2 \int_{\dot{\Sigma}} u^* \left(\omega + \tilde{f}(r)\, d\lambda\right) + \frac{2\epsilon}{\delta} \int_{\dot{\Sigma}} u^* \left(\tilde{f}'(r)\, dr \wedge \lambda\right). \end{aligned}$$

If $c^2 \geqslant \frac{2\epsilon}{b-a}$, then we can choose $\delta := 2\epsilon/c^2 \leqslant b - a$ and rewrite the last expression as

$$\begin{aligned} c^2 \int_{\dot{\Sigma}} u^* &\left(\omega + \tilde{f}(r)\, d\lambda\right) + \frac{2\epsilon}{\delta} \int_{\dot{\Sigma}} u^* \left(\tilde{f}'(r)\, dr \wedge \lambda\right) \\ &= c^2 \int_{\dot{\Sigma}} u^* \left(\omega + \tilde{f}(r)\, d\lambda + \tilde{f}'(r)\, dr \wedge \lambda\right) = c^2 \int_{\dot{\Sigma}} u^* \omega_{\tilde{f}} \leqslant c^2 E_0(u). \end{aligned}$$

On the other hand if $c^2 < \frac{2\epsilon}{b-a}$, we can set $\delta := b - a$ and write

$$\begin{aligned} c^2 \int_{\dot{\Sigma}} u^* \left(\omega + \tilde{f}(r)\, d\lambda\right) &+ \frac{2\epsilon}{\delta} \int_{\dot{\Sigma}} u^* \left(\tilde{f}'(r)\, dr \wedge \lambda\right) \\ &\leqslant \frac{2\epsilon}{b-a} \int_{\dot{\Sigma}} u^* \left(\omega + \tilde{f}(r)\, d\lambda + \tilde{f}'(r)\, dr \wedge \lambda\right) \\ &= \frac{2\epsilon}{b-a} \int_{\dot{\Sigma}} u^* \omega_{\tilde{f}} \leqslant \frac{2\epsilon}{b-a} E_0(u). \end{aligned}$$

□

With this preparation out of the way, we now begin in earnest with the proof of Theorem 7.2.3. Assume $u : \dot{\mathbb{D}} \to \widehat{W}$ is a J-holomorphic punctured disk satisfying $E(u) < \infty$. Using the maps $\varphi_\pm : Z_\pm \to \dot{\mathbb{D}}$ defined in (7.5), we shall write

$$u_\pm := u \circ \varphi_\pm : Z_\pm \to \widehat{W}$$

and observe that these reparametrizations have no impact on the energy, i.e.

$$E(u_\pm) = \sup_{f \in \mathcal{T}} \int_{Z_\pm} (u \circ \varphi_\pm)^* \omega_f = \sup_{f \in \mathcal{T}} \int_{\dot{\mathbb{D}}} u^* \omega_f = E(u).$$

Fix a Riemannian metric on $\widehat{W}$ that is translation-invariant on the cylindrical ends, and fix the standard metric on the half-cylinders $Z_\pm$. We will use these metrics implicitly whenever referring to quantities such as $|du_\pm(z)|$.

LEMMA 7.2.7. *There exists a constant $C > 0$ such that $|du_+(s,t)| \leqslant C$ for all $(s,t) \in Z_+$.*

PROOF. We use a bubbling argument as in the proof of Lemma 7.1.2. Suppose the contrary, so there exists a sequence $z_\nu = (s_\nu, t_\nu) \in Z_+$ with $R_\nu := |du_+(z_\nu)| \to \infty$. Choose a sequence $\epsilon_\nu > 0$ with $\epsilon_\nu \to 0$ but $\epsilon_\nu R_\nu \to \infty$, and using Lemma 7.1.3, assume without loss of generality that

$$|du_+(z)| \leqslant 2R_\nu \quad \text{for all} \quad z \in \mathbb{D}_{\epsilon_\nu}(z_\nu).$$

Define a rescaled sequence of J-holomorphic disks by

$$v_\nu : \mathbb{D}_{\epsilon_\nu R_\nu} \to \widehat{W} : z \mapsto u \circ \varphi_+(z_\nu + z/R_\nu).$$

These satisfy $|dv_\nu| \leqslant 2$ on their domains, but they are not necessarily C^1-bounded, since their images may escape to infinity. We distinguish three possibilities, at least one of which must hold:

Case 1: $v_\nu(0)$ has a bounded subsequence.

Then the corresponding subsequence of $v_\nu : \mathbb{D}_{\epsilon_\nu R_\nu} \to \widehat{W}$ is uniformly C^1-bounded on every compact subset and thus (by Proposition 7.0.1) has a further subsequence convergent in $C^\infty_{\text{loc}}(\mathbb{C})$ to a J-holomorphic plane

$$v_\infty : \mathbb{C} \to \widehat{W}$$

with $|dv_\infty(0)| = \lim_{k\to\infty} |dv_\nu(0)| = 1$. But by the same argument we used in the proof of Lemma 7.1.2, the fact that $\int_{Z_+} u_+^* \omega_f < \infty$ for any choice of $f \in \mathcal{T}$ implies

$$\int_{\mathbb{C}} v_\infty^* \omega_f = 0,$$

hence v_∞ is constant, and this is a contradiction.

Case 2: $v_\nu(0)$ has a subsequence diverging to $\{+\infty\} \times M_+$.

Restricting to this subsequence, suppose

$$v_\nu(0) \in \{r_\nu\} \times M_+,$$

so $r_\nu \to \infty$, and assume without loss of generality that $r_\nu > r_0$ for all ν. Let $\widetilde{R}_\nu \in (0, \epsilon_\nu R_\nu]$ for each ν denote the largest radius such that $v_\nu(\mathbb{D}_{\widetilde{R}_\nu}) \subset (r_0, \infty) \times M_+$. Then $\widetilde{R}_\nu \to \infty$ since $|dv_\nu|$ is bounded. Now using the $\mathbb{R}$-translation maps τ_c defined in (7.6), define

$$\tilde{v}_\nu := \tau_{-r_\nu} \circ v_\nu|_{\mathbb{D}_{\widetilde{R}_\nu}} : \mathbb{D}_{\widetilde{R}_\nu} \to \mathbb{R} \times M_+.$$

Since we're using a translation-invariant metric on $[r_0, \infty) \times M_+$, $\tilde{v}_\nu$ is now a uniformly C^1_{loc}-bounded sequence of maps into $\mathbb{R} \times M_+$. Proposition 7.0.1 thus provides a subsequence convergent in $C^\infty_{\text{loc}}(\mathbb{C})$ to a plane

$$v_\infty : \mathbb{C} \to \mathbb{R} \times M_+,$$

which is J_+-holomorphic, where $J_+ \in \mathcal{J}(\mathcal{H}_+)$ denotes the restriction of J to $[r_0, \infty) \times M_+$, extended over $\mathbb{R} \times M_+$ by $\mathbb{R}$-invariance. We claim,

$$E(v_\infty) < \infty \quad \text{and} \quad \int_{\mathbb{C}} v_\infty^* \omega_+ = 0, \tag{7.11}$$

where $E(v_\infty)$ is now defined as in Remark 7.2.2. By Lemma 7.2.6, the first part of the claim will follow if we can fix a constant $a \in (-\epsilon, \epsilon)$ and establish a uniform bound

$$\int_{\mathbb{C}} v_\infty^* \Omega_f^+ \leqslant C,$$

with $\Omega_f^+ := \omega_+ + d\big(f(r)\,\lambda_+\big)$, for all smooth and strictly increasing functions $f : \mathbb{R} \to (a, \epsilon)$. For convenience in the following, we shall assume $a > h(r_0)$. Now if f

is such a function, then for any $R > 0$,

$$\int_{\mathbb{D}_R} v_\infty^* \Omega_f^+ = \lim_{k\to\infty} \int_{\mathbb{D}_R} v_\nu^* \tau_{-r_\nu}^* \Omega_f^+ = \lim_{k\to\infty} \int_{\mathbb{D}_R} v_\nu^* \Omega_{f_\nu}^+,$$

where $f_\nu(r) := f(r - r_\nu)$. Notice that the dependence of the last integral on the function f_ν is limited to the interval $(r_0, \infty) \subset \mathbb{R}$ in its domain, since $v_\nu(\mathbb{D}_R) \subset (r_0, \infty) \times M_+$. Then since $f > a > h(r_0)$ by assumption, there exists for each ν a function $h_\nu \in \mathcal{T}$ that matches f_ν outside some neighborhood of $(-\infty, r_0]$ and thus satisfies

$$\int_{\mathbb{D}_R} v_\nu^* \Omega_{f_\nu}^+ = \int_{\mathbb{D}_R} v_\nu^* \omega_{h_\nu} \leqslant \int_{\mathbb{D}_{\epsilon_\nu R_\nu}} v_\nu^* \omega_{h_\nu} = \int_{\mathbb{D}_{\epsilon_\nu}(z_\nu)} u_+^* \omega_{h_\nu} \leqslant \int_{Z_+} u_+^* \omega_{h_\nu} \leqslant E(u).$$

This is true for every $R > 0$ and thus proves the first part of (7.11). To establish the second part, fix $R > 0$ again and pick any $f \in \mathcal{T}$. Observe that since we can assume (after perhaps passing to a subsequence) the disks $\mathbb{D}_{\epsilon_\nu}(z_\nu)$ are all disjoint,

$$\begin{aligned} 0 &= \lim_{k\to\infty} \int_{\mathbb{D}_{\epsilon_\nu}(z_\nu)} u_+^* \omega_f = \lim_{k\to\infty} \int_{\mathbb{D}_{\epsilon_\nu R_\nu}} v_\nu^* \omega_f = \lim_{k\to\infty} \int_{\mathbb{D}_{\epsilon_\nu R_\nu}} \tilde{v}_\nu^* \tau_{r_\nu}^* \omega_f \\ &\geqslant \lim_{k\to\infty} \int_{\mathbb{D}_R} \tilde{v}_\nu^* \tau_{r_\nu}^* \omega_f = \lim_{k\to\infty} \int_{\mathbb{D}_R} \tilde{v}_\nu^* \Omega_{f_\nu}^+, \end{aligned}$$

where now $f_\nu(r) := f(r + r_\nu)$. Writing $\Omega_{f_\nu}^+ = \omega_+ + d\big(f_\nu(r)\, \lambda_+\big) = \omega_+ + f_\nu(r)\, d\lambda_+ + f_\nu'(r)\, dr \wedge \lambda_+$, we can choose f such that $f'(r) = f'(r + r_\nu) \to 0$ as $\nu \to \infty$, so the third term contributes nothing to the integral. For the second term, let $f_+ := \lim_{k\to\infty} f_\nu(r) = \lim_{r\to\infty} f(r)$, so the calculation above becomes

$$0 \geqslant \int_{\mathbb{D}_R} v_\infty^* \left(\omega_+ + f_+\, d\lambda_+\right).$$

Now observe that since $f_+ \in [-\epsilon, \epsilon]$, condition (7.4) implies that the 2-form $\omega_+ + f_+\, d\lambda_+$ is nondegenerate on ξ_+, and it also annihilates ∂_r and R_+, so the vanishing of this integral implies that v_∞ is everywhere tangent to ∂_r and R_+ over $\mathbb{D}_R$. But $R > 0$ was arbitrary, so this is true on the whole plane, which is equivalent to $\int_{\mathbb{C}} v_\infty^* \omega_+ = 0$. With the claim established, we apply Lemma 7.2.4 and conclude that v_∞ is constant, contradicting the fact that $|dv_\infty(0)| = 1$.

Case 3: $v_\nu(0)$ has a subsequence diverging to $\{-\infty\} \times M_-$.
This is simply the mirror image of case 2: writing the restriction of J to $(-\infty, -r_0] \times M_-$ as J_-, one can follow the same bubbling argument but translate up and instead of down, giving rise to a limiting nonconstant J_--holomorphic plane $v_\infty : \mathbb{C} \to \mathbb{R} \times M_-$ that has finite energy but $\int_{\mathbb{C}} v_\infty^* \omega_- = 0$, in contradiction to Lemma 7.2.4. ☐

Consider now a sequence $s_\nu \to \infty$ and construct the J-holomorphic half-cylinders

$$u_\nu : [-s_\nu, \infty) \times S^1 \to \widehat{W} : (s,t) \mapsto u_+(s + s_\nu, t).$$

The derivatives $|du_\nu|$ are uniformly bounded due to Lemma 7.2.7, though again, u_ν might fail to be uniformly bounded in C^0. We distinguish three cases.

Case 1: $u_\nu(0,0)$ has a bounded subsequence.
Then the corresponding subsequence of u_ν is uniformly C^1-bounded on compact

subsets and thus has a further subsequence converging in $C^\infty_{\mathrm{loc}}(\mathbb{R} \times S^1)$ to a J-holomorphic cylinder

$$u_\infty : \mathbb{R} \times S^1 \to \widehat{W}.$$

For any $f \in \mathcal{T}$ and any $c > 0$, we have

$$\begin{aligned} \int_{[-c,c]\times S^1} u_\infty^* \omega_f &= \lim_{k\to\infty} \int_{[-c,c]\times S^1} u_\nu^* \omega_f \leqslant \lim_{k\to\infty} \int_{[-s_\nu/2,\infty)\times S^1} u_\nu^* \omega_f \\ &= \lim_{k\to\infty} \int_{[s_\nu/2,\infty)\times S^1} u_+^* \omega_f = 0 \end{aligned} \tag{7.12}$$

since $\int_{Z_+} u_+^* \omega_f < \infty$. It follows that $\int_{\mathbb{R}\times S^1} u_\infty^* \omega_f = 0$, so u_∞ is a constant map to some point $p \in \widehat{W}$, implying that after passing to a subsequence of s_ν,

$$u_+(s_\nu, \cdot) \to p \quad \text{in } C^\infty(S^1, \widehat{W}) \quad \text{as } \nu \to \infty.$$

Case 2: $u_\nu(0,0)$ has a subsequence diverging to $\{+\infty\} \times M_+$. Passing to the corresponding subsequence of u_ν, suppose

$$u_\nu(0,0) \in \{r_\nu\} \times M_+,$$

so $r_\nu \to \infty$. Since the derivatives $|du_\nu|$ are uniformly bounded, we can then find a sequence of intervals $[-R_\nu^-, R_\nu^+] \subset [-s_\nu, \infty)$ such that

$$u_\nu([-R_\nu^-, R_\nu^+] \times S^1) \subset [r_0, \infty) \times M_+ \quad \text{and} \quad R_\nu^\pm \to \infty.$$

Now the translated sequence

$$\tau_{-r_\nu} \circ u_\nu|_{[-R_\nu^-, R_\nu^+]\times S^1} : [-R_\nu^-, R_\nu^+] \times S^1 \to \mathbb{R} \times M_+$$

is uniformly C^1-bounded on compact subsets and thus has a subsequence coverging in C^∞_{loc} to a J_+-holomorphic cylinder

$$u_\infty : \mathbb{R} \times S^1 \to \mathbb{R} \times M_+,$$

where J_+ again denotes the restriction of J to $[r_0, \infty) \times M_+$, extended over $\mathbb{R} \times M_+$ by $\mathbb{R}$-translation. We claim that this cylinder satisfies

$$E(u_\infty) < \infty \quad \text{and} \quad \int_{\mathbb{R}\times S^1} u_\infty^* \omega_+ = 0.$$

The proof of this should be an easy exercise if you understood the proofs of (7.11) and (7.12) above, so I will leave it as such. Lemma 7.2.4 now implies that u_∞ is either constant or is a reparametrization of a trivial cylinder

$$u_\gamma : \mathbb{R} \times S^1 \to \mathbb{R} \times M_+ : (s,t) \mapsto (Ts, \gamma(Tt))$$

for some Reeb orbit $\gamma : \mathbb{R}/T\mathbb{Z} \to M_+$ with period $T > 0$. More precisely, all the biholomorphic reparametrizations of $\mathbb{R} \times S^1$ are of the form $(s,t) \mapsto (\pm s + a, \pm t + b)$, thus after shifting the parametrization of γ, we can write u_∞ without loss of generality in the form

$$u_\infty(s,t) = (\pm Ts + a, \gamma(\pm Tt)) \tag{7.13}$$

for some constant $a \in \mathbb{R}$ and a choice of signs to be determined below (see Lemma 7.2.11).

Case 3: $u_\nu(0,0)$ has a subsequence diverging to $\{-\infty\} \times M_-$.
Writing $J_- := J|_{(-\infty,-r_0]\times M_-} \in \mathcal{J}(\mathcal{H}_-)$ and imitating the argument for case 2, we suppose $u_\nu(0,0) \in \{-r_\nu\} \times M_-$ with $r_\nu \to \infty$ and obtain a subsequence for which $\tau_{r_\nu} \circ u_\nu$ converges in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to a J_--holomorphic cylinder $u_\infty : \mathbb{R} \times S^1 \to \mathbb{R} \times M_-$, where u_∞ is either a constant or takes the form (7.13) for some orbit Reeb $\gamma : \mathbb{R}/T\mathbb{Z} \to M_-$ of period $T > 0$.

Here is one easy consequence of the discussion so far. Use the Riemannian metric on $\widehat{W}$ to define a metric $\text{dist}_{C^0}(\cdot,\cdot)$ on the space of continuous loops $S^1 \to \widehat{W}$.

LEMMA 7.2.8. *Given $\delta > 0$, there exists $s_0 \geqslant 0$ such that for every $s \geqslant s_0$, the loop $u_+(s,\cdot) : S^1 \to \widehat{W}$ satisfies*

$$\text{dist}_{C^0}(u_+(s,\cdot), \ell_s) < \delta,$$

where for each s, $\ell_s : S^1 \to \widehat{W}$ either is constant or is a loop of the form $\ell_s(t) = (r, \gamma(\pm Tt))$ in $[r_0,\infty) \times M_+$ or $(-\infty, r_0] \times M_-$ for some constant $r \in \mathbb{R}$ and Reeb orbit $\gamma : \mathbb{R}/T\mathbb{Z} \to M_\pm$ of period $T > 0$, which may depend on s.

PROOF. If not, then there exists a sequence $s_\nu \to \infty$ such that each of the loops $u_+(s_\nu,\cdot)$ lies at C^0-distance at least δ away from any loop of the above form. However, the preceding discussion then gives a subsequence for which $u(s_\nu,\cdot)$ becomes arbitrarily C^∞-close to such a loop, so this is a contradiction. □

LEMMA 7.2.9. *If $u : \dot{\mathbb{D}} \to \widehat{W}$ is not bounded, then it is proper.*

PROOF. We use the monotonicity lemma. Suppose there exists a sequence $(s_\nu, t_\nu) \in Z_+$ such that $u_+(s_\nu, t_\nu)$ diverges to $\{+\infty\} \times M_+$. This implies $s_\nu \to \infty$, and we claim then that for every $R \geqslant r_0$, there exists $s_0 \geqslant 0$ such that

$$u_+((s_0,\infty) \times S^1) \subset (R,\infty) \times M_+.$$

If not, then we find $R \geqslant r_0$ and a sequence $(s'_\nu, t'_\nu) \in Z_+$ with $s'_\nu \to \infty$ such that $u_+(s'_\nu, t'_\nu) \notin (R,\infty) \times M_+$ for every ν. By continuity, we are free to suppose $u_+(s'_\nu, t'_\nu) \in \{R\} \times M_+$ for all ν, since Lemma 7.2.8 implies $u_+(\{s_\nu\} \times S^1) \subset (2R,\infty) \times M_+$ for ν sufficiently large. Using Lemma 7.2.8 again, we also have

$$u_+(\{s'_\nu\} \times S^1) \subset (R-1, R+1) \times M_+$$

for all ν large. Assuming $2R > R+2$ without loss of generality, we can therefore find infinitely many pairwise disjoint annuli of the form $[s'_\nu, s_j] \times S^1 \subset Z_+$ containing open sets that u maps properly to small balls centered at points in $\{R+2\} \times M_+$. Choosing any $f \in \mathcal{T}$, the monotonicity lemma implies that each of these contributes at least some fixed amount to $\int_{Z_+} u_+^* \omega_f$, contradicting the assumption that $E(u) < \infty$.[1]

A similar argument works if $u_+(s_\nu, t_\nu)$ diverges to $\{-\infty\} \times M_-$, proving that for every $R \geqslant r_0$, there exists $s_0 \geqslant 0$ with

$$u_+((s_0,\infty) \times S^1) \subset (-\infty, -R) \times M_-.$$

□

[1] The fact that $\widehat{W}$ is noncompact is not a problem for this application of the monotonicity lemma, as we are only using it in the compact subset $W^{2R} \subset \widehat{W}$.

If u is bounded, then the singularity at 0 is removable by Theorem 7.1.1. If not, then Lemma 7.2.9 implies that it maps neighborhoods of the puncture to neighborhoods of either $\{+\infty\} \times M_+$ or $\{-\infty\} \times M_-$, and we shall refer to the puncture as *positive* or *negative* accordingly.

LEMMA 7.2.10. *If the puncture is positive/negative, then the limit*

$$Q := \lim_{s\to\infty} \int_{S^1} u_+(s,\cdot)^* \lambda_\pm \in \mathbb{R}$$

exists.

PROOF. If the puncture is positive, fix $s_0 \geqslant 0$ such that $u_+([s_0,\infty) \times S^1) \subset [r_0,\infty) \times M_+$. Then by Stokes' theorem, it suffices to show that the integral $\int_{[s_0,\infty)\times S^1} u_+^* d\lambda_+$ exists, which is true if

$$\int_{[s_0,\infty)\times S^1} \left| u_+^* d\lambda_+ \right| < \infty. \tag{7.14}$$

We claim first that $\int_{[s_0,\infty)\times S^1} u_+^*\omega_+ < \infty$. Indeed, for any $s > s_0$ and $f \in \mathcal{T}$, we have

$$E(u) \geqslant \int_{[s_0,s]\times S^1} u_+^*\omega_f = \int_{[s_0,s]\times S^1} u_+^*\omega_+ + \int_{[s_0,s]\times S^1} u_+^* d\left(f(r)\,\lambda_+\right).$$

Applying Stokes' theorem, the second term becomes the sum of some number not dependent on s and the integral

$$\int_{S^1} u_+(s,\cdot)^* \left(f(r)\,\lambda_+\right) = \int_{S^1} \left[f \circ u_+(s,\cdot)\right] u_+(s,\cdot)^*\lambda_+,$$

which is bounded as $s \to \infty$ since f and $|du_+|$ are both bounded. This proves that $\int_{[s_0,s]\times S^1} u_+^*\omega_+$ is also bounded as $s \to \infty$, and since $u_+^*\omega_+ \geqslant 0$, the claim follows. Now observe that since $d\lambda_+$ annihilates the kernel of ω_+ and the latter tames J on ξ_+, there exists a constant $c > 0$ such that $|u_+^* d\lambda_+| \leqslant c|u_+^*\omega_+|$, implying (7.14).

An analogous argument works if the puncture is negative. □

The number $Q \in \mathbb{R}$ defined in the above lemma matches what we referred to in the statement of Theorem 7.2.3 as the **charge** of the puncture.

LEMMA 7.2.11. *If the puncture is nonremovable and $Q \neq 0$, then the puncture is positive/negative if and only if $Q > 0$ or $Q < 0$ respectively. In either case, given any sequence $s_\nu \to \infty$ with $u_+(s_\nu, 0) \in \{\pm r_\nu\} \times M_\pm$, one can find a sequence $R_\nu \in [0, s_\nu]$ with $R_\nu \to \infty$ such that u_+ maps $[s_\nu - R_\nu, \infty) \times S^1$ into the positive/negative cylindrical end for every ν, and the sequence of half-cylinders*

$$u_\nu : [-R_\nu, \infty) \times S^1 \to \mathbb{R} \times M_+ \quad \text{or} \quad u_\nu : (-\infty, R_\nu] \times S^1 \to \mathbb{R} \times M_-$$

defined by $u_\nu(s,t) = \tau_{\mp r_\nu} \circ u_\pm(s \pm s_\nu, t)$ has a subsequence convergent in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to a $J_\pm$-holomorphic cylinder of the form

$$u_\infty : \mathbb{R} \times S^1 \to \mathbb{R} \times M_\pm : (s,t) \mapsto (Ts + a, \gamma(Tt))$$

for some constant $a \in \mathbb{R}$ and Reeb orbit $\gamma : \mathbb{R}/T\mathbb{Z} \to M_\pm$ with period $T := \pm Q$.

PROOF. Assume the puncture is either positive or negative and $Q \neq 0$. In the discussion preceding Lemma 7.2.8, we showed that the sequence $u'(s,t) := \tau_{\mp r_\nu} \circ u_+(s+s_\nu, t)$ defined on $[-R_\nu, \infty) \times S^1$ has a subsequence convergent in C^∞_{loc} to a $J_\pm$-holomorphic cylinder $u'_\infty : \mathbb{R} \times S^1 \to \mathbb{R} \times M_\pm$ which is either constant or of the form

$$u'_\infty(s,t) = (\sigma T s + a, \gamma(\sigma T t)) \tag{7.15}$$

for some $a \in \mathbb{R}$, $\sigma = \pm 1$ and a Reeb orbit $\gamma : \mathbb{R}/T\mathbb{Z} \to M_\pm$ of period $T > 0$. We then have

$$0 \neq Q = \lim_{s\to\infty} \int_{S^1} u_+(s,\cdot)^*\lambda_\pm = \lim_{k\to\infty} \int_{S^1} u'_\nu(0,\cdot)^*\lambda_\pm = \int_{S^1} u'_\infty(0,\cdot)^*\lambda_\pm,$$

so u'_∞ cannot be constant, and from (7.15) we deduce $Q = \sigma T$, hence $u'_\infty(s,t) = (Qs + a, \gamma(Qt))$. Writing $u_+(s,t) = (u_\mathbb{R}(s,t), u_M(s,t)) \in \mathbb{R} \times M_\pm$ for s sufficiently large, it follows that every sequence $s_\nu \to \infty$ admits a subsequence for which

$$\partial_s u_\mathbb{R}(s_\nu, \cdot) \to Q \quad \text{in} \quad C^\infty(S^1, \mathbb{R}),$$

and consequently $\partial_s u_\mathbb{R}(s,\cdot) \to Q$ in $C^\infty(S^1,\mathbb{R})$ as $s \to \infty$. This proves that the sign of Q matches the sign of the puncture whenever $Q \neq 0$. The stated formula for u_∞ now follows by adjusting all the appropriate signs in the case $Q < 0$. □

LEMMA 7.2.12. *If the puncture is nonremovable, then* $Q \neq 0$.

PROOF. Assume on the contrary that u is a proper map, say with a positive puncture, but $Q = 0$. In this case, the argument of the previous lemma shows that the limiting map $u_\infty : \mathbb{R} \times S^1 \to \mathbb{R} \times M_+$ will always be *constant*, thus for every sequence $s_\nu \to \infty$, there exists a point $p \in M_+$ such that $u_+(s_\nu, 0) \in \{r_\nu\} \times M_+$ with $r_\nu \to \infty$ and

$$\tau_{-r_\nu} \circ u_+(s_\nu, \cdot) \to (0,p) \in \mathbb{R} \times M_+ \quad \text{in} \quad C^\infty(S^1, \mathbb{R} \times M_+) \text{ as } \nu \to \infty.$$

In particular, this implies that all derivatives of u_+ decay to 0 as $s \to \infty$. Intuitively, this should suggest to you that portions of u_+ near infinity will have improbably small symplectic area, perhaps violating the monotonicity lemma—this will turn out to be true, but we have to be a bit clever with our argument, since u_+ is unbounded. We will make this argument precise by translating pieces of u_+ downward so that we only compute its symplectic area in $[0,2] \times M_+$. Fix a function $f : \mathbb{R} \to (-\epsilon, \epsilon)$ with $f' > 0$ and set $\Omega^+_f = \omega_+ + d(f(r)\,\lambda_+)$.

Given a small number $\delta > 0$, we can find $s_0 \geqslant 0$ such that $|du_+(s,t)| < \delta$ for all $s \geqslant s_0$ and each of the loops $u_+(s,\cdot)$ for $s \geqslant s_0$ is δ-close to a constant in $C^1(S^1)$. Assume $u_+(s_0, 0) \in \{R\} \times M_+$ and choose $s_1 > s_0$ such that $u_+(s_1, 0) \in \{R+2\} \times M_+$, which is possible since $u_+(s,t) \to \{+\infty\} \times M_+$ as $s \to \infty$. Now consider the J_+-holomorphic annulus

$$v_\delta := \tau_{-R} \circ u_+|_{[s_0,s_1]\times S^1} : [s_0, s_1] \times S^1 \to \mathbb{R} \times M_+.$$

We claim that $\int_{[s_0,s_1]\times S^1} v_\delta^*\Omega_f^+$ can be made arbitrarily small by choosing δ suitably small. Indeed, we can use Stokes' theorem to write this integral as

$$\begin{aligned}\int_{[s_0,s_1]\times S^1} v_\delta^*\Omega_f^+ &= \int_{[s_0,s_1]\times S^1} v_\delta^*\omega_+ + \int_{[s_0,s_1]\times S^1} v_\delta^* d\left(f(r)\,\lambda_+\right) \\ &= \int_{[s_0,s_1]\times S^1} v_\delta^*\omega_+ + \int_{S^1} \left[v_\delta(s_1,\cdot)^*\left(f(r)\,\lambda_+\right) - v_\delta(s_0,\cdot)^*\left(f(r)\,\lambda_+\right)\right].\end{aligned}$$

The second term is small because $f(r)$ is bounded and $|v_\delta(s,\cdot)^*\lambda_+|$ is small in proportion to $|dv_\delta(s,t)| = |du_+(s,t)|$ for $s \geqslant s_0$. For the first term, observe that since both of the loops $v_\delta(s_i,\cdot)$ for $i = 0,1$ are nearly constant, they are contractible and can be filled in with disks $\bar{v}_i : \mathbb{D} \to \mathbb{R} \times M_+$ for which $\left|\int_{\mathbb{D}} \bar{v}_i^*\omega_+\right|$ may be assumed arbitrarily small. Moreover, since all of the loops $v_\delta(s,\cdot)$ are similarly contractible, the union of these two disks with the annulus v_δ defines a closed cycle in M_+ that is trivial in $H_2(M_+)$, hence the integral of the closed 2-form ω_+ over this cycle vanishes, implying

$$\int_{[s_0,s_1]\times S^1} v_\delta^*\omega_+ = \int_{\mathbb{D}} \bar{v}_1^*\omega_+ - \int_{\mathbb{D}} \bar{v}_0^*\omega_+,$$

which is therefore arbitrarily small, and this proves the claim.

To finish, notice that since v_δ maps its boundary components to small neighborhoods of $\{0\} \times M_+$ and $\{2\} \times M_+$, one can fix a suitable choice of radius $r_1 > 0$ such that v_δ must pass through a point in $p \in \{1\} \times M_+$ for which the boundary of v_δ is outside the ball $B_{r_1}(p)$. The monotonicity lemma then bounds the symplectic area of v_δ from below by a constant times r_1^2, but since we can also make this area arbitrarily small by choosing δ smaller, this is a contradiction.

As usual, the case of a negative puncture can be handled similarly. □

We've now proved every statement in Theorem 7.2.3 up to the final detail about the case where the asymptotic orbit is nondegenerate or Morse-Bott. The complete proof of this part requires delicate analytical results from [**HWZ96**, **HWZ01**, **HWZ96**, **Bou02**], but we can explain the first step for the nondegenerate case. In the following, we say that a closed Reeb orbit $\gamma : \mathbb{R}/T\mathbb{Z} \to M_\pm$ is **isolated** if, after rescaling the domain to write it as an element of $C^\infty(S^1, M_\pm)$, there exists a neighborhood $\gamma \in \mathcal{U} \subset C^\infty(S^1, M_\pm)$ such that all closed Reeb orbits in $\mathcal{U}$ are reparametrizations of γ.

LEMMA 7.2.13. *Suppose the puncture is nonremovable, write*

$$u_+(s,t) = (u_{\mathbb{R}}(s,t), u_M(s,t)) \in \mathbb{R} \times M_\pm$$

for $s \geqslant 0$ sufficiently large, and suppose $s_\nu \to \infty$ is a sequence and $\gamma : \mathbb{R}/T\mathbb{Z} \to M_\pm$ is a Reeb orbit such that

$$u_M(s_\nu,\cdot) \to \gamma(T\cdot) \quad \text{in} \quad C^\infty(S^1, M_\pm).$$

If γ is isolated, then for every neighborhood $\mathcal{U} \subset C^\infty(S^1, M_\pm)$ of the set of parametrizations $\{\gamma(\cdot + \theta) \mid \theta \in S^1\}$, we have $u_M(s,\cdot) \in \mathcal{U}$ for all sufficiently large s.

Proof. Note first that if γ is isolated, then its image admits a neighborhood $\operatorname{im}\gamma \subset \mathcal{V} \subset M_\pm$ such that no point in $\mathcal{V}\setminus \operatorname{im}\gamma$ is contained in another Reeb orbit of period T. Indeed, we could otherwise find a sequence of T-periodic Reeb orbits passing through a sequence of points in $\mathcal{V}\setminus \operatorname{im}\gamma$ that converge to a point in $\operatorname{im}\gamma$. Since their derivatives are determined by the Reeb vector field and are therefore bounded, the Arzelà-Ascoli theorem then gives a subsequence of these orbits converging to a reparametrization of γ, contradicting the assumption that γ is isolated.

Arguing by contradiction, suppose now that there exists a sequence $s'_\nu \to \infty$ with $u_M(s_\nu, \cdot) \notin \mathcal{U}$ for all ν. We can nonetheless restrict to a subsequence for which $u_M(s'_\nu, \cdot)$ converges to some Reeb orbit $\tilde{\gamma} : \mathbb{R}/T\mathbb{Z} \to M_\pm$. Then $\tilde{\gamma}$ is disjoint from γ, and by continuity, one can find a sequence $s''_\nu \to \infty$ for which each $u_M(s''_\nu, 0)$ lies in the region $\mathcal{V}$ some fixed distance away from $\operatorname{im}\gamma$. There must then be a subsequence for which $u_M(s''_\nu, \cdot)$ converges to another T-periodic orbit, but this is impossible since no such orbits exist in $\mathcal{V}\setminus \operatorname{im}\gamma$. □

7.3. Degenerations of holomorphic curves

To motivate the SFT compactness theorem, we shall now discuss three examples of phenomena that can prevent a sequence of holomorphic curves from having a compact subsequence. The theorem will then tell us that these three things are, in essence, the only things that can go wrong.

Throughout this section and the next, assume $J_\nu \to J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ is a C^∞-convergent sequence of tame almost complex structures on the completed cobordism $\widehat{W}$. More generally, one can also allow the data ω, h and $\mathcal{H}_\pm$ to vary in C^∞-convergent sequences, but let's not clutter the notation too much. We shall denote the restrictions of J_ν to the cylindrical ends by

$$J^+_\nu := J_\nu|_{[r_0,\infty)\times M_+} \in \mathcal{J}(\mathcal{H}_+), \qquad J^-_\nu := J_\nu|_{(-\infty,-r_0]\times M_-} \in \mathcal{J}(\mathcal{H}_-),$$

and $J^\pm$ similarly for the limiting almost complex structure J. Suppose

$$u_\nu := [(\Sigma_\nu, j_\nu, \Gamma^+_\nu, \Gamma^-_\nu, \Theta_\nu, u_\nu)] \in \mathcal{M}_{g,m}(J_\nu, A_\nu, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$$

is a sequence of J_ν-holomorphic curves in $\widehat{W}$ with fixed genus $g \geqslant 0$ and $m \geqslant 0$ marked points, varying relative homology classes $A_\nu \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ and fixed collections of asymptotic orbits $\boldsymbol{\gamma}^\pm = (\gamma^\pm_1, \ldots, \gamma^\pm_{m_\pm})$. Observe that the energies $E(u_\nu)$ depend only on the orbits $\boldsymbol{\gamma}^\pm$ and relative homology classes A_ν, so in particular, $E(u_\nu)$ is uniformly bounded whenever the relative homology class is also fixed. The fundamental question of this section is:

Question. *If $E(u_\nu)$ is uniformly bounded and no subsequence of u_ν converges to an element of $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ for any $A \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$, what can happen?*

7.3.1. Bubbling. Suppose the marked Riemann surfaces $(\Sigma_\nu, j_\nu, \Gamma^+_\nu \cup \Gamma^-_\nu \cup \Theta_\nu)$ form a convergent sequence, meaning we can assume after biholomorphic reparametrization that the surfaces $\Sigma_\nu = \Sigma$ are all identical with identical sets of punctures $\Gamma^\pm_\nu = \Gamma^\pm$ and marked points $\Theta_\nu = \Theta$, while their complex structures are C^∞-convergent

$$j_\nu \to j \in \mathcal{J}(\Sigma).$$

Suppose additionally that there exists a point $\zeta_0 \in \dot{\Sigma}$ such that $u_\nu(\zeta_0) \in \widehat{W}$ is contained in a compact subset for all ν, and that for some choice of Riemannian metrics on $\dot{\Sigma} = \Sigma \backslash \Gamma$ and $\widehat{W}$ that are translation-invariant on the cylindrical ends of both, the maps $u_\nu : \dot{\Sigma} \to \widehat{W}$ are locally C^1-bounded outside some finite subset

$$\Gamma' = \{\zeta_1, \ldots, \zeta_N\} \subset \dot{\Sigma},$$

i.e. for every compact set $K \subset \dot{\Sigma} \backslash \Gamma'$, there exists a constant $C_K > 0$ independent of ν such that

$$|du_\nu| \leqslant C_K \quad \text{on } K.$$

Then Proposition 7.0.1 gives a subsequence that converges in $C^\infty_{\text{loc}}(\dot{\Sigma}\backslash\Gamma')$ to a J-holomorphic curve

$$u_\infty : \dot{\Sigma}\backslash\Gamma' \to \widehat{W}$$

with $E(u_\infty) \leqslant \limsup E(u_\nu) < \infty$, thus all the punctures $\Gamma^+ \cup \Gamma^- \cup \Gamma'$ of u_∞ are either removable or positively or negatively asymptotic to Reeb orbits. We cannot be sure that the asymptotic behavior of u_∞ at $\Gamma^\pm$ is the same as for u_ν, but let's assume this for now. (We will dicsuss in §7.3.2 some things that can happen if this does not hold.) Then to complete the picture, we need to understand not only what u_∞ is doing at the additional punctures Γ', but also what is happening to u_ν near these points as its first derivative blows up. For this, we can apply the familiar rescaling trick: Choose for each ζ_i a sequence $z^i_\nu \to \zeta_i$ such that $|du_\nu(z^i_\nu)| =: R_\nu \to \infty$, along with a sequence $\epsilon_\nu \to 0$ with $\epsilon_\nu R_\nu \to \infty$, and using Lemma 7.1.3, assume without loss of generality that $|du_\nu(z)| \leqslant 2R_\nu$ for all z in the ϵ_ν-ball about z^i_ν. For convenience, we can choose a holomorphic coordinate system identifying a neighborhood of ζ_i with $\mathbb{D} \subset \mathbb{C}$ and placing ζ_i at the origin, so $z^i_\nu \to 0$ in these coordinates, and assume without loss of generality that they identify our chosen metric near ζ_i with the Euclidean metric. Now setting

$$v^i_\nu(z) = u_\nu(z^i_\nu + z/R_\nu) \quad \text{for} \quad z \in \mathbb{D}_{\epsilon_\nu R_\nu}$$

gives a sequence of J_ν-holomorphic maps $v^i_\nu : \mathbb{D}_{\epsilon_\nu R_\nu} \to \widehat{W}$ whose energies and first derivatives are both uniformly bounded. As in the arguments of §2, we now have three possibilities:

- If $u_\nu(z^i_\nu)$ has a bounded subsequence, then the corresponding subsequence of v^i_ν converges in $C^\infty_{\text{loc}}(\mathbb{C})$ to a J-holomorphic plane $v^i_\infty : \mathbb{C} \to \widehat{W}$ with finite energy.
- If $u_\nu(z^i_\nu)$ has a subsequence diverging to $\{\pm\infty\} \times M_\pm$, then translating v^i_ν by the $\mathbb{R}$-action produces a limiting finite-energy plane v^i_∞ in the positive/negative symplectization $\mathbb{R} \times M_\pm$.

Viewing $\mathbb{C}$ as the punctured sphere $S^2\backslash\{\infty\}$, the singularity of v^i_∞ at ∞ may be removable, in which case v^i_∞ extends to a J-holomorphic sphere and we say that u_ν has "bubbled off a sphere" at ζ_i. Alternatively, v^i_∞ may be positively or negatively asymptotic to a Reeb orbit at ∞.

Figure 7.2 shows two scenarios that could occur for a sequence in which $|du_\nu|$ blows up at three points $\Gamma' = \{\zeta_1, \zeta_2, \zeta_3\}$. Both scenarios show u_∞ with ζ_1 and ζ_2 as removable singularities and ζ_3 as a negative puncture, but the behavior of the various v_∞^i reveals a wide spectrum of possibilities. In the lower-left picture, the points $u_\nu(z_\nu^1)$ are bounded and bubble off a sphere $v_\infty^1 : S^2 \to \widehat{W}$. The picture shows that v_∞^1 passes through $u_\infty(\zeta_1)$ at some point; this does not follow from our argument so far, but in this situation, one can use a more careful analysis of u_ν near ζ_1 to show that it must be true, i.e. "bubbles connect". At ζ_3, we have $u_\nu(z_\nu^3) \to \{-\infty\} \times M_-$ and v_∞^3 is a plane in $\mathbb{R} \times M_-$ with a positive puncture asymptotic to the same orbit as ζ_3; the coincidence of these orbits is another detail that does not follow from the analysis above, but turns out to be true in the general picture. The situation at ζ_2 allows two different interpretations: v_∞^2 could be the plane with negative end in $\mathbb{R} \times M_+$, meaning $u_\nu(z_\nu^2) \to \{+\infty\} \times M_+$, and the picture then shows an additional plane in $\widehat{W}$ with a positive end approaching the same asymptotic orbit as v_∞^2 as well as a point passing through $u_\infty(\zeta_2)$. One would need to choose a different rescaled sequence near ζ_2 to find this extra plane, but as we will see, the SFT compactness theorem dictates that some such object must be there. Alternatively, $u_\nu(z_\nu^2)$ could also be bounded at ζ_2, in which case v_∞^2 must be the plane in $\widehat{W}$ with positive end, and the extra plane above this is something that one could find via a different choice of rescaled sequence. In general, the range of actual possibilities can involve arbitrarily many additional curves that could be discovered via different choices of rescaled sequences: e.g. there could be entire "bubble trees" as shown in the lower-right picture, where each v_∞^i is only one of several curves that arise as limits of different parametrizations of u_ν near ζ_i. One good place to read about the analysis of bubble trees is [**HWZ03**, §4].

EXERCISE 7.3.1. The following simple application of standard bubbling arguments will be needed in the next chapter for showing that diffeomorphism groups act *properly* on the space of stable J-holomorphic curves. Suppose Σ is a closed, connected and oriented surface, $\Theta \subset \Sigma$ is a finite subset such that

$$\chi(\Sigma \backslash \Theta) < 0,$$

$j_\nu \to j$ and $j'_\nu \to j'$ are C^∞-convergent sequences of complex structures on Σ, and

$$(\Sigma, j'_\nu) \xrightarrow{\varphi_\nu} (\Sigma, j_\nu)$$

is a sequence of biholomorphic diffeomorphisms that fix the points in Θ. The following shows that in this situation, φ_ν must have a C^∞-convergent subsequence:

(a) Prove that if Σ has positive genus, then there can be no bubbling, hence φ_ν has a C^∞-convergent subsequence whose limit is a biholomorphic map $\varphi : (\Sigma, j') \to (\Sigma, j)$. *Hint: The universal cover of Σ is contractible, so $\pi_2(\Sigma) = 0$.*

(b) Prove that if $\Sigma = S^2$ and bubbling occurs, then it occurs at no more than one point, i.e. there exists a point $\zeta \in S^2$ such that φ_ν is uniformly C^1-bounded on all compact subsets of $S^2 \backslash \{\zeta\}$. Show moreover that if $|d\varphi_\nu|$ really does blow up along some sequence approaching ζ, then a subsequence of φ_ν converges in $C^\infty_{\text{loc}}(S^2 \backslash \{\zeta\})$ to a constant, and derive a contradiction from this using the assumption that φ_ν fixes each point in Θ. *Hint: Choose an area form Ω on S^2 and look at $\int_K \varphi_\nu^* \Omega$ for compact subsets $K \subset S^2 \backslash \{\zeta\}$ as $\nu \to \infty$.*

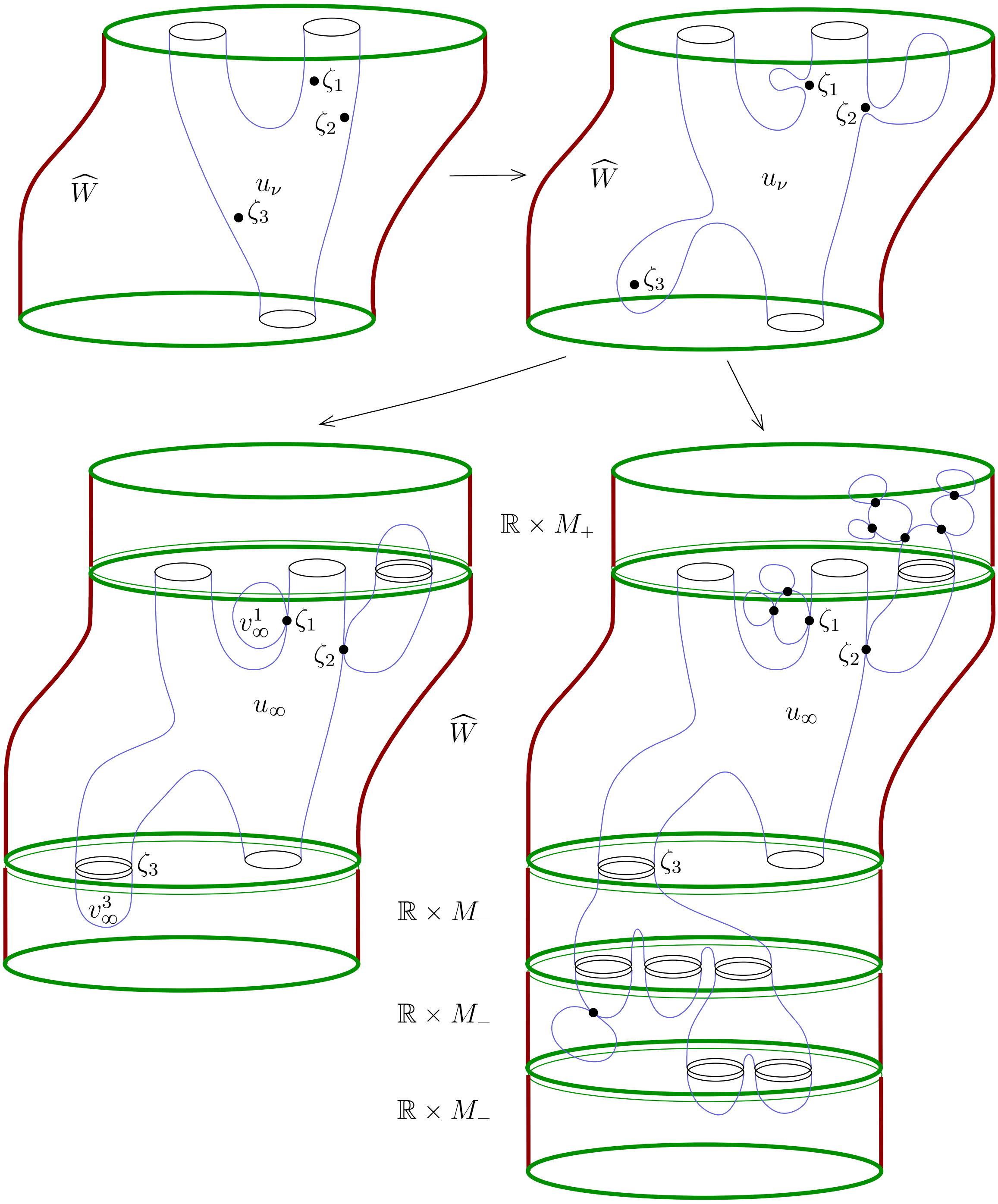

FIGURE 7.2. Two possible pictures of spheres and/or planes that can bubble off when the first derivative blows up near three points.

REMARK 7.3.2. Exercise 7.3.1 shows that whenever $\chi(\Sigma \backslash \Theta) < 0$, the group of biholomorphic automorphisms $(\Sigma, j) \to (\Sigma, j)$ fixing Θ is compact. By the Lefschetz fixed point theorem for smooth self-maps on a closed oriented manifold, the algebraic count of fixed points for a smooth map $\varphi : \Sigma \to \Sigma$ homotopic to the identity is $\chi(\Sigma)$.

Since φ is holomorphic, all such fixed points must contribute positively to this count, so the assumption $\chi(\Sigma \backslash \Theta) = \chi(\Sigma) - \#\Theta < 0$ proves that no such maps can exist, i.e. the group of biholomorphic automorphisms fixing Θ is also discrete, and therefore finite.

7.3.2. Breaking. Figure 7.2 already shows some phenomena that could be interpreted as "breaking" in the Floer-theoretic sense, but breaking can also happen when no derivatives are blowing up, simply due to the fact that our domains are noncompact. Figures 7.3 and 7.4 show three such scenarios, where we assume again that $j_\nu \to j$ and $\dot{\Sigma} = \Sigma \backslash \Gamma$ and $\widehat{W}$ carry Riemannian metrics that are translation-invariant on the cylindrical ends such that

$$|du_\nu| \leqslant C \quad \text{everywhere on } \dot{\Sigma}$$

for some constant $C > 0$ independent of ν. This is a stronger condition than we had in §7.3.1, and if there exists a point $\zeta_0 \in \dot{\Sigma}$ such that $u_\nu(\zeta_0) \in \widehat{W}$ is bounded, it implies that u_ν has a subsequence converging in $C^\infty_{\text{loc}}(\dot{\Sigma})$ to a J-holomorphic map

$$u_\infty : \dot{\Sigma} \to \widehat{W}$$

with $E(u_\infty) \leqslant \limsup E(u_\nu) < \infty$. Convergence in C^∞_{loc} is, however, not very strong: there may in general be no relation between the asymptotic behavior of u_∞ and u_ν at corresponding punctures, e.g. the top scenario in Figure 7.3 shows a case in which a negative puncture of u_ν becomes a removable singularity of u_∞. Whenever this happens, there must be more to the story: in this example, one can choose holomorphic cylindrical coordinates $(s,t) \in (-\infty, 0] \times S^1 \subset \dot{\Sigma}$ near the negative puncture of u_ν and find a sequence $s_\nu \to \infty$ such that the sequence of half-cylinders

$$(-\infty, s_\nu] \times S^1 \to \widehat{W} : (s,t) \mapsto u_\nu(s - s_\nu, t)$$

is uniformly C^1-bounded and thus converges in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to a finite-energy J-holomorphic cylinder $v_- : \mathbb{R} \times S^1 \to \widehat{W}$. In the picture, v_- turns out to have a removable singularity at $+\infty$ mapping to the same point as the removable singularity of u_∞, and its negative puncture approaches the same orbit as the negative puncture of u_ν.

More complicated things can happen in general. The bottom scenario in this same figure shows a case where all three singularities of u_∞ are removable, thus it extends to a closed curve, while at one of the positive cylindrical ends $[0, \infty) \times S^1 \subset \dot{\Sigma}$ of u_ν, we can find a sequence $s_\nu \to \infty$ such that the half-cylinders

$$[-s_\nu, \infty) \times S^1 \to \widehat{W} : (s,t) \mapsto u_\nu(s + s_\nu, t)$$

are uniformly C^1-bounded and converge in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to a J-holomorphic cylinder $v^1_+ : \mathbb{R} \times S^1 \to \widehat{W}$ with one removable singularity and one positive puncture. At the other positive end, we can perform the same trick in two distinct ways for two sequences $s_\nu \to \infty$, one diverging faster than the other: the result is a pair of J-holomorphic cylinders $v^2_+, v^3_+ : \mathbb{R} \times S^1 \to \widehat{W}$, the former with both singularities removable (thus forming a holomorphic sphere in the picture), and the latter with one removable singularity and one positive puncture.

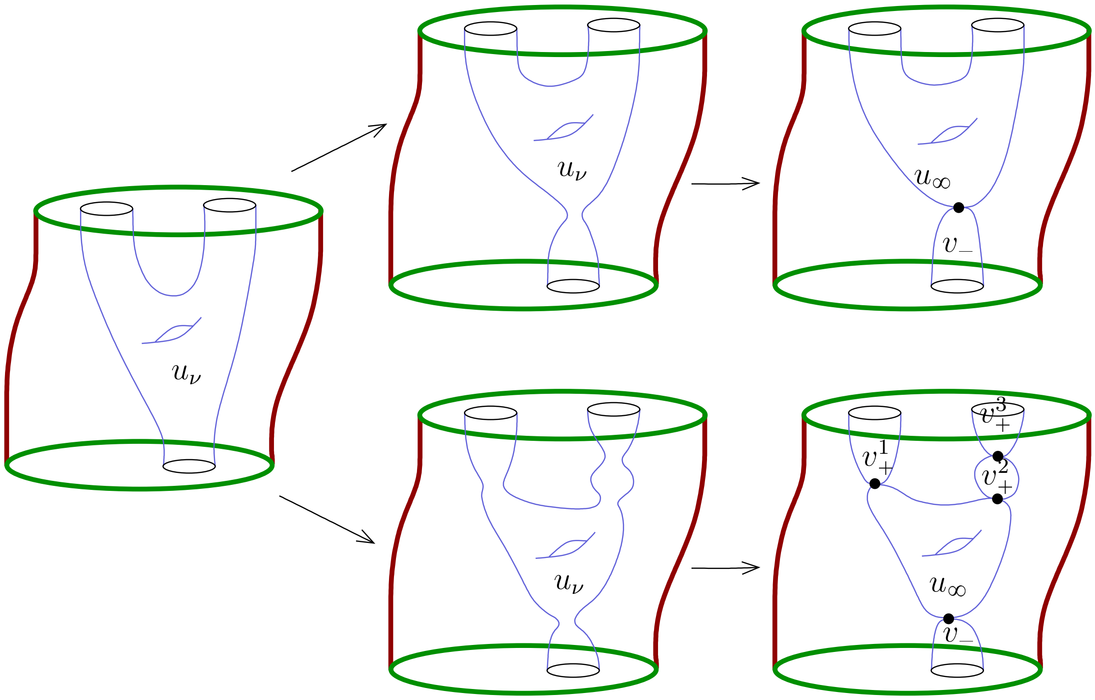

FIGURE 7.3. Even with fixed conformal structures on the domains and without bubbling, a sequence of punctured holomorphic curves in $\widehat{W}$ can break to produce multiple curves in $\widehat{W}$ with extra removable punctures. The picture shows two such scenarios.

It can get weirder. Remember that $\widehat{W}$ is also noncompact!

In each of the above scenarios, we tacitly assumed that all of the various sequences obtained by reparametrizing portions of u_ν were locally C^0-bounded, thus all of the limits were curves in $\widehat{W}$. But it may also happen that some of these sequences are C^0_{loc}-bounded while others locally diverge toward $\{\pm\infty\} \times M_\pm$; in fact, two such sequences that both diverge toward, say, $\{+\infty\} \times M_+$, might even locally diverge infinitely far from *each other*, meaning one of them approaches $\{+\infty\} \times M_+$ quantitatively faster than the other. This phenomenon leads to the notion of limiting curves with multiple *levels*.

In Figure 7.4, we see a scenario in which u_ν satisfies the same conditions as above, except that instead of $u_\nu(\zeta_0)$ being bounded, it diverges to $\{+\infty\} \times M_+$. It follows that after applying suitable $\mathbb{R}$-translations, a subsequence converges in $C^\infty_{\text{loc}}(\dot{\Sigma})$ to a J^+-holomorphic curve

$$u_\infty : \dot{\Sigma} \to \mathbb{R} \times M_+$$

with finite energy. In the example, all three of its punctures are nonremovable, but two of them approach orbits that have nothing to do with the asymptotic orbits of u_ν. Now observe that since u_ν has a negative cylindrical end $(-\infty, 0] \times S^1 \subset \dot{\Sigma}$, one can necessarily find a sequence $s_\nu \to \infty$ such that $u_\nu(-s_\nu, 0)$ is bounded, and the sequence of half-cylinders

$$(-\infty, s_\nu] \times S^1 \to \widehat{W} : (s,t) \mapsto u_\nu(s - s_\nu, t)$$

is then uniformly C^1-bounded and thus has a subsequence convergent in $C^\infty_{\text{loc}}(\mathbb{R}\times S^1)$ to a finite-energy J-holomorphic cylinder $v_0 : \mathbb{R} \times S^1 \to \widehat{W}$. In the picture, v_0 has both a positive and a negative puncture, but its negative end again approaches a different Reeb orbit from the negative ends of u_ν, so one can deduce that there must be still more happening near $-\infty$: there exists another sequence $s'_\nu \to \infty$ with $s'_\nu - s_\nu \to \infty$ such that suitable $\mathbb{R}$-translations of the half-cylinders

$$(-\infty, s_\nu] \times S^1 \to (-\infty, -r_0] \times M_- : (s,t) \mapsto u_\nu(s - s'_\nu, t)$$

define uniformly C^1-bounded maps into $\mathbb{R}\times M_-$, giving a subsequence that converges in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to a finite-energy J^--holomorphic cylinder

$$v_- : \mathbb{R} \times S^1 \to \mathbb{R} \times M_-.$$

Finally, the fact that u_∞ has a positive asymptotic orbit different from those of u_ν indicates that something more must also be happening near $+\infty$: in the example, one of the positive ends $[0,\infty) \times S^1 \subset \dot{\Sigma}$ admits a sequence $s_\nu \to \infty$ such that $u_\nu(s_\nu, 0) \in \{r_\nu\} \times M_+$ for some $r_\nu \to \infty$, and suitable $\mathbb{R}$-translations of

$$[-s_\nu, \infty) \times S^1 \to [r_0, \infty) \times M_+ : (s,t) \mapsto u_\nu(s + s_\nu, t)$$

become a uniformly C^1-bounded sequence of half-cylinders in $\mathbb{R} \times M_+$, with a subsequence converging in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to a finite-energy J^+-holomorphic cylinder

$$v^2_+ : \mathbb{R} \times S^1 \to \mathbb{R} \times M_+$$

that connects the errant asymptotic orbit of u_∞ to the corresponding orbit of u_ν. One can now perform the same trick at the other positive end of $\dot{\Sigma}$, as there necessarily also exists a sequence $s'_\nu \to \infty$ in this end such that $u_\nu(s'_\nu, 0) \in \{r_\nu\} \times M_+$ for the same sequence $r_\nu \to \infty$ as in the above discussion. The resulting limit curve $v^1_+ : \mathbb{R} \times S^1 \to \mathbb{R} \times M_+$ however is not guaranteed to be interesting: in the picture, it turns out to be a trivial cylinder.

The type of degeneration shown in Figure 7.4 happens whenever the sequence u_ν does interesting things in multiple regions of its domain that are sent increasingly far away from each other in the image. The usual picture of $\widehat{W}$ that collapses the cylindrical ends to a finite size therefore becomes increasingly inadequate for visualizing u_ν as $\nu \to \infty$: the middle picture in Figure 7.4 deals with this by expanding the scale of the cylindrical ends so that the convergence to upper and lower levels becomes visible.

7.3.3. The Deligne-Mumford space of Riemann surfaces. We next need to relax the assumption that the marked Riemann surfaces $(\Sigma_\nu, j_\nu, \Gamma^+_\nu \cup \Gamma^-_\nu \cup \Theta_\nu)$ converge. Recall from Example 6.5.7 that for integers $g \geqslant 0$ and $\ell \geqslant 0$, the moduli space of marked Riemann surfaces is the space of equivalence classes

$$\mathcal{M}_{g,\ell} = \{(\Sigma, j, \Theta)\} \big/ \sim,$$

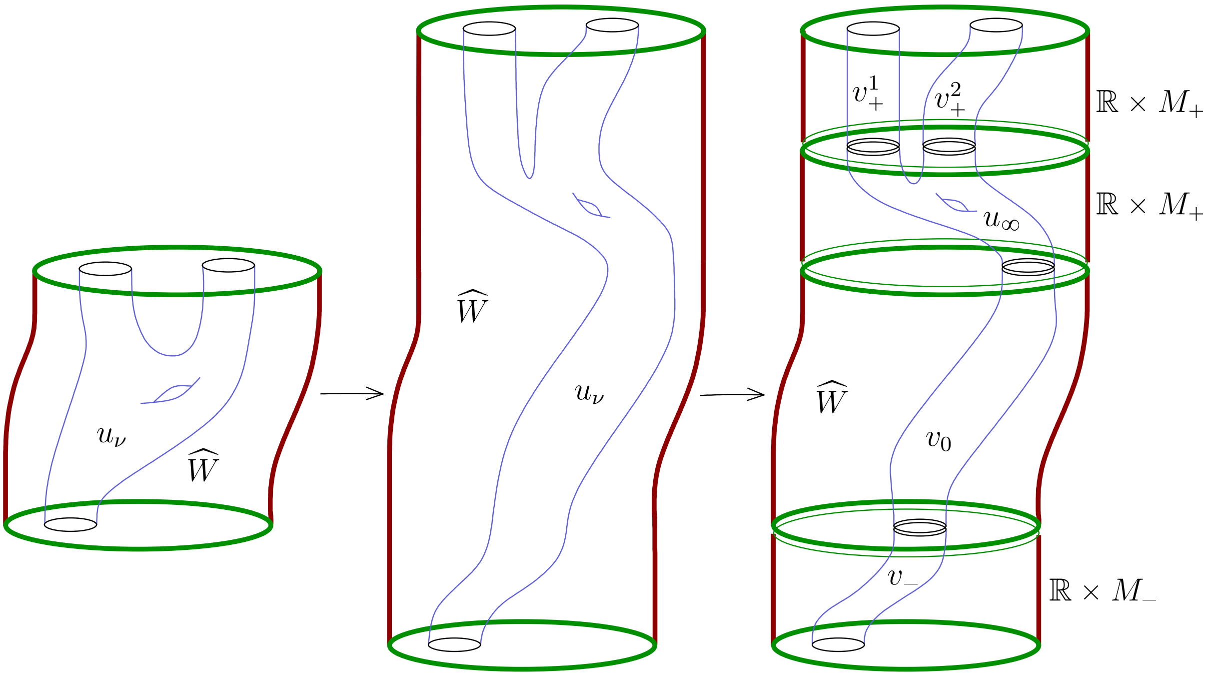

FIGURE 7.4. Different portions of a breaking sequence of curves may also become infinitely far apart in the limit, so that some live in $\widehat{W}$ while others live in the symplectization of M_+ or M_-.

where (Σ, j) is a closed connected Riemann surface of genus g, $\Theta \subset \Sigma$ is an ordered set of ℓ distinct points, and $(\Sigma, j, \Theta) \sim (\Sigma', j', \Theta')$ whenever there exists a biholomorphic map $\varphi : (\Sigma, j) \to (\Sigma', j')$ taking Θ to Θ' with the ordering preserved. This space is fairly easy to understand in the finitely many cases with $2g+\ell < 3$, e.g. since the uniformization theorem below implies that every closed genus 0 Riemann surface is biholomorphically equivalent to the standard Riemann sphere $(S^2, i) := \mathbb{C} \cup \{\infty\}$, $\mathcal{M}_{0,\ell}$ is a one-point space for each $\ell \leqslant 3$. We say that (Σ, j, Θ) is **stable** whenever $\chi(\Sigma \backslash \Theta) < 0$, which means $2g + \ell \geqslant 3$. In the stable case, Exercise 7.3.1 and Remark 7.3.2 show that every marked Riemann surface has a finite automorphism group, and results of Chapter 8 will show that $\mathcal{M}_{g,\ell}$ is a smooth orbifold of dimension $6g - 6 + 2\ell$. It is generally not compact, but it admits a natural compactification

$$\overline{\mathcal{M}}_{g,\ell} \supset \mathcal{M}_{g,\ell},$$

known as the **Deligne-Mumford compactification**. The main goal of this section is to state the Deligne-Mumford compactness theorem in a form that is especially useful for holomorphic curve theory. Standard proofs of the theorem are typically based on either algebraic or hyperbolic geometry; we will use the latter perspective to motivate the main definitions and give the main idea of the proof. For more details on Deligne-Mumford compactness from the hyperbolic perspective, see [**Hum97**, **SS92**].

We recall first the following standard result.

Theorem 7.3.3 (Uniformization theorem). *Every simply connected Riemann surface is biholomorphically equivalent to either the Riemann sphere* $S^2 = \mathbb{C} \cup \{\infty\}$*, the complex plane* $\mathbb{C}$ *or the upper half plane* $\mathbb{H} = \{\operatorname{Im} z > 0\} \subset \mathbb{C}$.

The uniformization theorem implies that every Riemann surface can be presented as a quotient of either (S^2, i), $(\mathbb{C}, i)$ or $(\mathbb{H}, i)$ by some freely acting discrete group of biholomorphic transformations. The only punctured surface $\dot{\Sigma} = \Sigma \backslash \Theta$ that has S^2 as its universal cover is S^2 itself. It is almost as easy to see which surfaces are covered by $\mathbb{C}$, as the only biholomorphic transformations on $(\mathbb{C}, i)$ with no fixed points are the translations, so every freely acting discrete subgroup of $\operatorname{Aut}(\mathbb{C}, i)$ is either trivial, a cyclic group of translations or a lattice. The resulting quotients are, respectively, $(\mathbb{C}, i)$, $(\mathbb{R} \times S^1, i) \cong (\mathbb{C}\backslash\{0\}, i)$ and the unpunctured tori (T^2, j). All *stable* marked Riemann surfaces are thus quotients of $(\mathbb{H}, i)$.

Proposition 7.3.4. *There exists on* $(\mathbb{H}, i)$ *a complete Riemannian metric* g_P *of constant curvature* -1 *that defines the same conformal structure as* i *and has the property that all conformal transformations on* $(\mathbb{H}, i)$ *are also isometries of* $(\mathbb{H}, g_P)$.

Proof. We define g_P at $z = x + iy \in \mathbb{H}$ by

$$g_P = \frac{1}{y^2} g_E,$$

where g_E is the Euclidean metric. The conformal transformations on $(\mathbb{H}, i)$ are given by fractional linear transformations

$$\begin{aligned} \operatorname{Aut}(\mathbb{H}, i) &= \left\{ \varphi(z) = \frac{az+b}{cz+d} \;\middle|\; a,b,c,d \in \mathbb{R}, \quad ad - bc = 1 \right\} \Big/ \{\pm 1\} \\ &= \operatorname{SL}(2, \mathbb{R})/\{\pm 1\} =: \operatorname{PSL}(2, \mathbb{R}), \end{aligned}$$

and one can check that each of these defines an isometry with respect to g_P. One can also compute that g_P has curvature -1, and the geodesics of g_P are precisely the lines and semicircles that meet $\mathbb{R}$ orthogonally, parametrized so that they exist for all forward and backward time, thus g_P is complete. More details on all of this can be found in the book by Hummel [**Hum97**]. □

By lifting to universal covers, this implies the following.

Corollary 7.3.5. *For every marked Riemann surface* (Σ, j, Θ) *with* $\chi(\Sigma\backslash\Theta) < 0$*, the punctured Riemann surface* $(\Sigma\backslash\Theta, j)$ *admits a unique complete Riemannian metric* g_j *of constant curvature* -1 *that defines the same conformal structure as* j*. Moreover, all biholomorphic transformations on* $(\Sigma\backslash\Theta, j)$ *are also isometries of* $(\Sigma\backslash\Theta, g_j)$. □

The metric g_j in this corollary is often called the **Poincaré metric**; we also call it "hyperbolic" because of its negative curvature. Its existence is in fact *equivalent* to the stable case of the uniformization theorem: indeed, the constant negative curvatuve of $(\Sigma\backslash\Theta, g_j)$ implies on the one hand that it is locally isometric to $(\mathbb{H}, g_P)$, and also that any two points in its universal cover can be connected by a unique geodesic with respect to the lift of g_j. This is enough information to construct a global isometry between $(\mathbb{H}, g_P)$ and the universal cover of $(\Sigma\backslash\Theta, g_j)$ by starting from one point and following geodesics. For an analytical proof of Corollary 7.3.5 in the case $\Theta = \emptyset$ without assuming Theorem 7.3.3, see [Tro92].

Every nontrivial class in $\pi_1(\Sigma\backslash\Theta)$ contains a unique geodesic for g_j. Now suppose $C \subset \Sigma\backslash\Theta$ is a union of disjoint embedded geodesics such that each connected component of $\Sigma\backslash(\Theta \cup C)$ has the homotopy type of a disk with two holes. The components are then called **pairs of pants**, and the result is called a **pair-of-pants decomposition** of $(\Sigma\backslash\Theta, j)$. Two examples for the case $g = 1$ and $\ell = 3$ are shown in Figure 7.5. The collection of geodesics $C \subset \Sigma\backslash\Theta$ can generally be chosen in many ways, giving rise to many distinct pair-of-pants decompositions of a single surface $(\Sigma\backslash\Theta, j)$, but all of them have the same number of pieces: indeed, since a pair of pants always has Euler characteristic -1, the number of pairs of pants that must be assembled to build $\Sigma\backslash\Theta$ is

$$-\chi(\Sigma\backslash\Theta) \in \mathbb{N}.$$

Note that each individual pair of pants in such a decomposition may or may not be a compact surface, as its three "boundary" components come generally in two types: the actual (*nondegenerate*) boundary components are closed geodesics, but there may also be **degenerate** components, which are actually punctures in $\Sigma\backslash\Theta$ (i.e. marked points of Σ) rather than boundary components. It is a useful convention to regard such degenerate boundary components as "closed geodesics of length 0".

With this convention understood, one can show that the lengths of the geodesic boundary components of a pair of pants uniquely determine its hyperbolic metric up to isometry (see [Hum97, Prop. IV.2.7]), and therefore also its conformal structure. Up to biholomorphic equivalence, the punctured Riemann surface $(\Sigma\backslash\Theta, j)$ can thus be characterized via finitely many parameters, namely the lengths $\ell(\gamma) > 0$ of each of the closed geodesics $\gamma \subset C$ that form the nondegenerate boundary components in a pair-of-pants decomposition, plus some information dictating how to glue pairs of pants together along these geodesics. The latter information can be expressed as a "twist" parameter $\theta(\gamma) \in S^1$ for each matching pair of boundary components to be glued together, i.e. one imagines modifying a given pair-of-pants decomposition by cutting it apart along the geodesic γ and gluing it back together via a gluing map that is rotated by an angle of $2\pi\theta$ along the closed geodesic. The tuple of pairs $(\ell(\gamma), \theta(\gamma)) \in (0, \infty) \times S^1$ associated to each of the geodesics $\gamma \subset C \subset \Sigma\backslash\Theta$ now gives a local parametrization of the moduli space $\mathcal{M}_{g,\ell}$ near any given element $[(\Sigma, j, \Theta)] \in \mathcal{M}_{g,\ell}$, known as the *Fenchel-Nielsen coordinates*. Note that since there are always exactly $-\chi(\Sigma\backslash\Theta) = 2g - 2 + \ell$ pairs of pants in a decomposition, the total number of geodesics involved is $[3(2g - 2 + \ell) - \ell]/2 = 3g - 3 + \ell$, thus one can read off the formula

$$\dim \mathcal{M}_{g,\ell} = 6g - 6 + 2\ell \quad \text{for} \quad 2g + \ell \geqslant 3$$

from this geometric picture. We will discuss an alternative analytical way to deduce this dimension formula in §8.3.

One can also see the noncompactness of $\mathcal{M}_{g,\ell}$ in this picture quite concretely: the twist parameters belong to a compact space, but each length parameter can potentially shrink to 0 or blow up to ∞ as j (and hence g_j) is deformed. It turns out

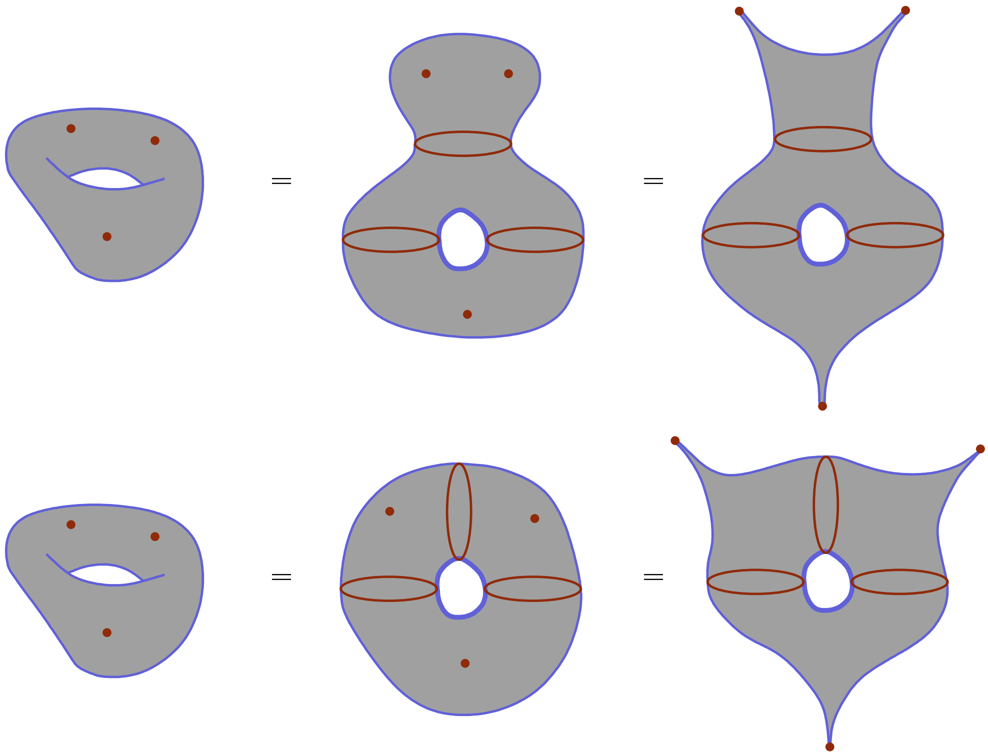

FIGURE 7.5. Two distinct pair-of-pants decompositions for the same genus 1 Riemann surface with three marked points. The decompositions are shown from two perspectives: the pictures at the right are meant to give a more accurate impression of the Poincaré metric, which becomes singular and forms a cusp at each marked point.

that the latter possibility is an illusion, but one may need to switch to a different pair-of-pants decomposition to see why:

THEOREM (Bers; see [**Hum97**, Theorem 3.7] or [**SS92**, Theorem 3.19.2]). *For every pair of integers $g \geqslant 0$ and $\ell \geqslant 0$ with $2g + \ell \geqslant 3$, there exists a constant $C = C(g, \ell) > 0$ such that every $[(\Sigma, j, \Theta)] \in \mathcal{M}_{g,\ell}$ admits a pair-of-pants decomposition in which all geodesics bounding the pairs of pants have length at most C.*

This theorem implies that from a hyperbolic perspective, the only meaningful way for stable marked Riemann surfaces to degenerate is when some of the bounding geodesics in a pair-of-pants decomposition shrink to length zero. Figure 7.6 shows several examples of degenerate Riemann surfaces that can arise in this way for $g = 1$ and $\ell = 3$, giving elements of the space that we will now define as $\overline{\mathcal{M}}_{1,3}$.

DEFINITION 7.3.6. A **marked nodal Riemann surface** with $\ell \geqslant 0$ marked points and $N \geqslant 0$ **nodes** is a tuple (S, j, Θ, Δ) consisting of:

- A closed but not necessarily connected Riemann surface (S, j);
- An ordered set of ℓ points $\Theta \subset S$;

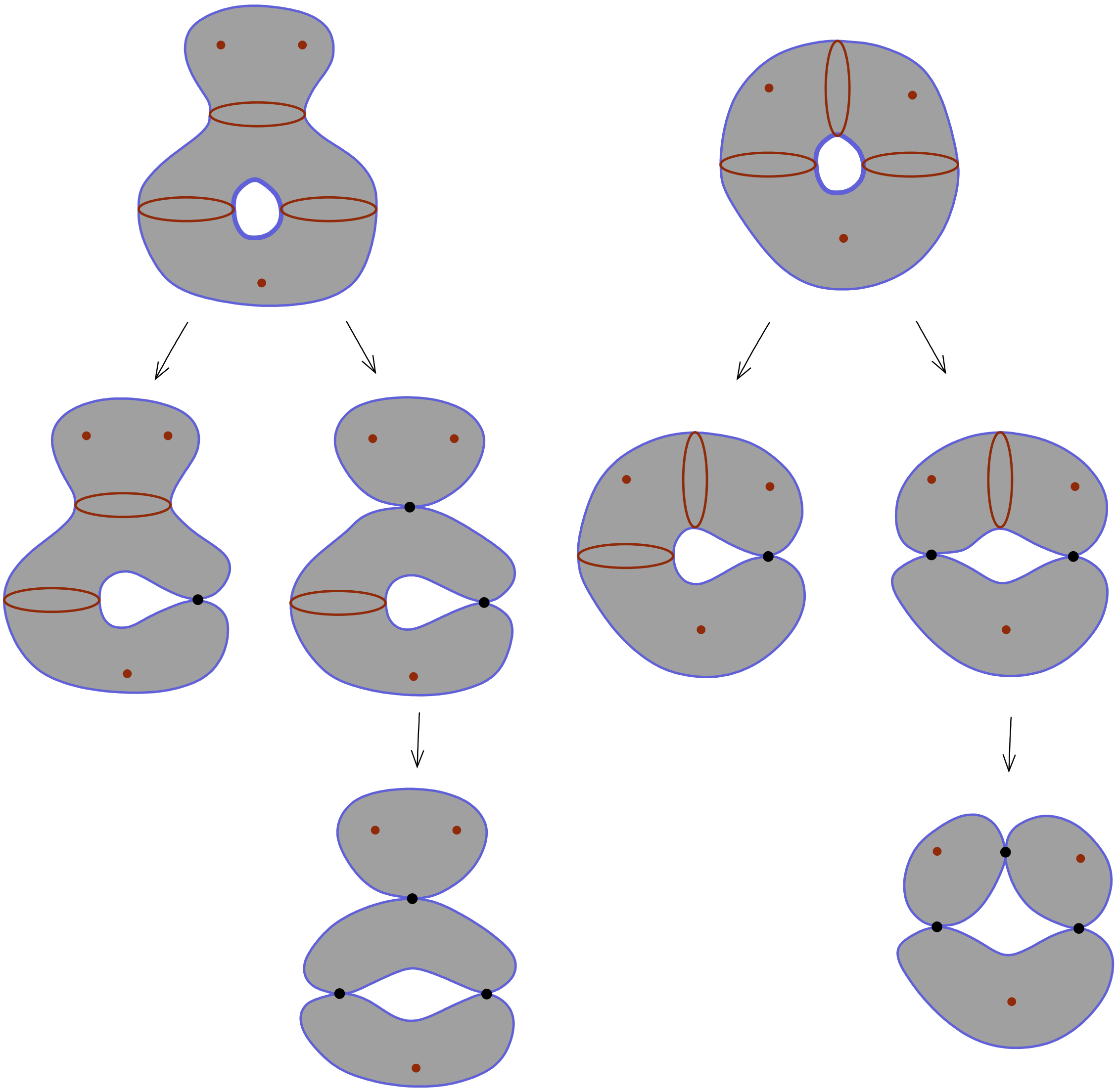

FIGURE 7.6. Starting from each of the pair-of-pants decompositions for the $g = 1$ and $\ell = 3$ case from Figure 7.5, shrinking geodesic lengths to zero produces various examples of stable nodal Riemann surfaces belonging to $\overline{\mathcal{M}}_{1,3}$.

- An unordered set of $2N$ points $\Delta \subset S \backslash \Theta$ equipped with an involution $\sigma : \Delta \to \Delta$ having no fixed points. Each pair $\{z, \sigma(z)\}$ for $z \in \Delta$ is referred to as a **node**.

Let $\widehat{S}$ denote the closed surface obtained by performing connected sums on S at each node $\{z^+, z^-\} \subset \Delta$. We then say that (S, j, Θ, Δ) is **connected** if and only if $\widehat{S}$ is connected, and the genus of $\widehat{S}$ is called the **arithmetic genus** of (S, j, Θ, Δ). We say that (S, j, Θ, Δ) is **stable** if every connected component of $S \backslash (\Theta \cup \Delta)$ has negative Euler characteristic. Finally, two nodal Riemann surfaces (S, j, Θ, Δ) and $(S', j', \Theta', \Delta')$ are considered **equivalent** if there exists a biholomorphic map φ :

$(S, j) \to (S', j')$ taking Θ to Θ' with the ordering preserved and taking Δ to Δ' such that nodes are mapped to nodes.

The nodes $\{z^+, z^-\} \subset \Delta$ are typically represented in pictures as self-intersections of S, cf. Figure 7.6. We can think of the *stable* nodal surfaces as precisely those which admit pair-of-pants decompositions. All nodal Riemann surfaces we consider will be assumed connected in the sense defined above unless otherwise indicated; note that S itself can nonetheless be disconnected, as is the case in four out of the six nodal surfaces shown in Figure 7.6.

We now introduce some further terminology and notation that will be useful in the next section as well. Whenever $\dot{\Sigma} := \Sigma \backslash \Gamma$ is obtained by puncturing a Riemann surface (Σ, j) at finitely many points $\Gamma \subset \Sigma$, we shall define the **circle compactification**

$$\bar{\Sigma} := \dot{\Sigma} \cup \bigcup_{z \in \Gamma} \delta_z,$$

where for each $z \in \Gamma$, the circle δ_z is defined as a "half-projectivization" of the tangent space at z:

$$\delta_z := (T_z\Sigma \backslash \{0\}) \Big/ \mathbb{R}_+^*,$$

with the positive real numbers $\mathbb{R}_+^*$ acting by scalar multiplication. To understand the topology of $\bar{\Sigma}$, one can equivalently define it by choosing holomorphic cylindrical coordinates $[0, \infty) \times S^1 \subset \dot{\Sigma}$ near each z, and replacing the open half-cylinder with $[0, \infty] \times S^1$, where δ_z is now the **circle at infinity** $\{\infty\} \times S^1$. There is no natural choice of global smooth structure on $\bar{\Sigma}$, but it is homeomorphic to an oriented surface with boundary and carries both smooth and conformal structures on its interior, due to the obvious identification

$$\dot{\Sigma} = \bar{\Sigma} \backslash \bigcup_{z \in \Gamma} \delta_z \subset \bar{\Sigma}.$$

The conformal structure of Σ at each $z \in \Gamma$ does induce on each of the circles δ_z an **orthogonal structure**, meaning a preferred class of homeomorphisms to S^1 that are all related to each other by rotations. One can therefore speak of **orthogonal maps** $\delta_z \to \delta_{z'}$ for $z, z' \in \Gamma$, which are always homeomorphisms and can either preserve or reverse orientation.

Now if (S, j, Θ, Δ) is a nodal Riemann surface, we let $\dot{S} = S \backslash \Delta$ and form the circle compactification $\overline{S}$, which has the topology of a compact oriented surface with boundary. Given a node $\{z^+, z^-\} \subset \Delta$, a **decoration** for $\{z^+, z^-\}$ is a choice of orientation-reversing orthogonal map

$$\Phi : \delta_{z^+} \to \delta_{z^-}.$$

We say that (S, j, Θ, Δ) is a **decorated** nodal surface if it is equipped with a choice of decoration Φ for every node, or **partially decorated** if Φ is defined for some subset of the nodes. A partial decoration Φ gives rise to another compact oriented surface

$$\widehat{S}_\Phi := \overline{S} / \sim,$$

where the equivalence relation identifies δ_{z^+} with δ_{z^-} via Φ for each decorated node $\{z^+, z^-\} \subset \Delta$. Note that if every node is decorated, then $\widehat{S}_\Phi$ has the topology of a closed connected and oriented surface whose genus defines the arithmetic genus of (S, j, Θ, Δ) according to Definition 7.3.6. We shall denote the collection of special circles in $\widehat{S}_\Phi$ where boundary components $\delta_{z^+}, \delta_{z^-} \subset \partial\overline{S}$ have been identified by

$$C_\Phi \subset \widehat{S}_\Phi.$$

Since $\widehat{S}_\Phi \backslash (\partial \widehat{S}_\Phi \cup C_\Phi)$ has a natural identification with $\dot{S}$, it inherits smooth and conformal structures which degenerate along C_Φ and $\partial\widehat{S}_\Phi$. We will say that two partially decorated marked nodal Riemann surfaces $(S, j, \Theta, \Delta, \Phi)$ and $(S', j', \Theta', \Delta', \Phi')$ are **equivalent** if (S, j, Θ, Δ) and $(S', j', \Theta', \Delta')$ are equivalent via a biholomorphic map $\varphi : (S, j) \to (S', j')$ that extends continuously from $\dot{S} \to \dot{S}'$ to a homeomorphism $\widehat{S}_\Phi \to \widehat{S}'_{\Phi'}$.

Now if $2g + \ell \geqslant 3$, the **Deligne-Mumford compactification** of the moduli space of Riemann surfaces (also known more simply as the *Deligne-Mumford moduli space*) is defined as the set $\overline{\mathcal{M}}_{g,\ell}$ of equivalence classes of stable nodal Riemann surfaces with ℓ marked points and arithmetic genus g. There is a natural inclusion

$$\mathcal{M}_{g,\ell} \subset \overline{\mathcal{M}}_{g,\ell}$$

by regarding each marked Riemann surface (Σ, j, Θ) as a nodal Riemann surface $(\Sigma, j, \Theta, \Delta)$ with $\Delta = \varnothing$.

The most important property of $\overline{\mathcal{M}}_{g,\ell}$ is that it admits the structure of a compact metrizable topological space for which the inclusion $\mathcal{M}_{g,\ell} \hookrightarrow \overline{\mathcal{M}}_{g,\ell}$ is continuous onto an open and dense subset. Intuitively, this follows from the discussion of pair-of-pants decompositions above, in particular Bers's theorem. While one can describe the topology of $\overline{\mathcal{M}}_{g,\ell}$ very naturally in terms of hyperbolic metrics and Fenchel-Nielsen coordinates, this is not the most practical characterization for our purposes, so we will now give an alternative description that does not require any hyperbolic geometry.

To get the right picture, imagine what the conformal structure of a "nearly degenerate" hyperbolic surface (Σ, g) looks like in a collar neighborhood $\mathcal{N} \subset \Sigma$ of some closed geodesic $\gamma \subset \Sigma$ with very small length bounding two pairs of pants. The conformal structure j determined by g does not see the fact that the length of the geodesic is small, as one can always rescale g to give the geodesic length 1 without changing j. But this rescaling increases lengths in both dimensions, thus it also makes the collar $\mathcal{N}$ look thick, i.e. we can now view it as a long "neck"

$$(\mathcal{N}, j) \cong ([-R, R] \times S^1, i) \subset (\Sigma, j), \qquad R \gg 0,$$

holomorphically embedded in (Σ, j) such that $\{0\} \times S^1$ parametrizes the geodesic γ. The length of this neck is in inverse proportion to the length of γ, so the degeneration of (Σ, j) to a nodal Riemann surface can be realized in purely conformal terms by letting the neck length go to ∞.

Let us turn this intuitive picture into a practical definition, a diagram of which is shown in Figure 7.7. Suppose (S, j, Θ, Δ) is a marked nodal Riemann surface and $\{z^+, z^-\} \subset \Delta$ is one of its nodes; note that the choice of which nodal point to label

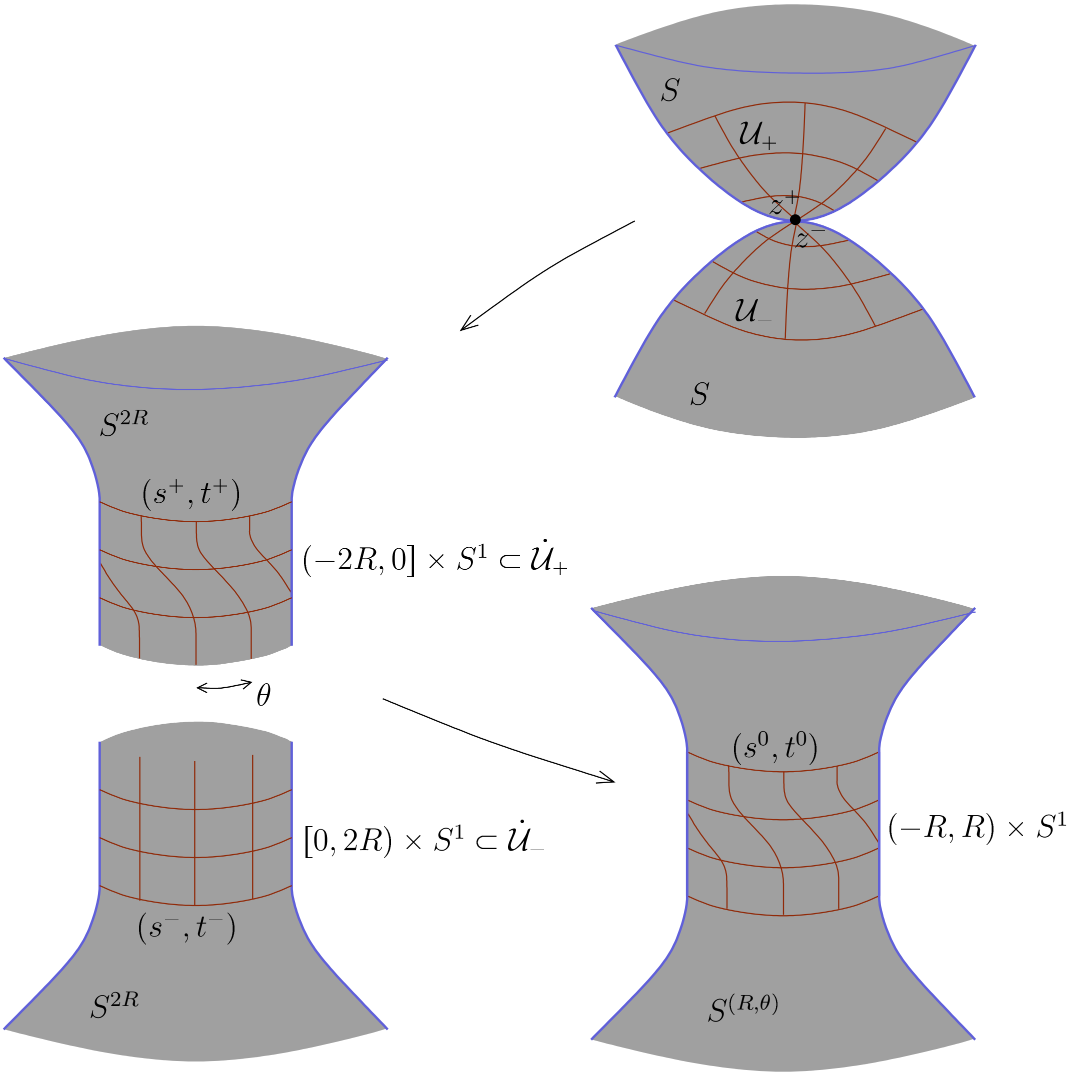

FIGURE 7.7. For each node $\{z^+, z^+\}$, one chooses holomorphic cylindrical coordinates on neighborhoods $\mathcal{U}_\pm \subset S$ or $z^\pm$ and then truncates the cylindrical ends to define S^{2R} with two collars of finite length carrying coordinates $(s^\pm, t^\pm)$. Gluing these collars together produces the surface $S^{(R,\theta)}$, which contains a neck $(-R, R) \times S^1$ with coordinates (s^0, t^0). The picture shows the coordinate grid for (s^+, t^+) skewed in order to indicate the twist parameter $\theta \in S^1$ in the attaching map.

as z^+ and which as z^- in this situation is arbitrary, but we shall make a choice and stick with it. We also need to choose two neighborhoods $\mathcal{U}_\pm \subset S \backslash \Theta$ of $z^\pm$ with holomorphic cylindrical coordinates

$$(s^\pm, t^\pm) \in Z_\mp \cong \dot{\mathcal{U}}_\pm := \mathcal{U}_\pm \backslash \{z^\pm\} \subset S$$

identifying the corresponding punctured neighborhoods biholomorphically with the standard half-cylinders $Z_- = (-\infty, 0] \times S^1$ and $Z_+ = [0, \infty) \times S^1$ respectively. Note the sign reversal, so we are regarding z^+ as a negative puncture and z^- as a positive puncture of $S\backslash\Delta$. We shall assume both neighborhoods are small enough so that they do not intersect Θ, $\Delta\backslash\{z^+, z^-\}$, or each other. Recall from Chapter 4 the notation

$$Z_+^R := [R, \infty) \times S^1, \qquad Z_-^R := (-\infty, -R] \times S^1$$

for $R \geqslant 0$, which gives rise to the truncated surface

$$S^R \subset S\backslash\{z^+, z^-\}$$

obtained by deleting $Z_-^R \subset \dot{\mathcal{U}}_+$ and $Z_+^R \subset \dot{\mathcal{U}}_-$ from $S\backslash\{z^+, z^-\}$. We now associate to each $(R, \theta) \in (0, \infty) \times S^1$ a new marked nodal surface

$$\mathbf{S}^{(R,\theta)} = (S^{(R,\theta)}, \Theta, \Delta^R)$$

defined as follows:

- $S^{(R,\theta)} := S^{2R}/\sim$, where the equivalence relation identifies the collars $(-2R, 0) \times S^1 \subset \dot{\mathcal{U}}_+$ and $(0, 2R) \times S^1 \subset \dot{\mathcal{U}}_-$ via the diffeomorphism

$$\begin{aligned} \dot{\mathcal{U}}_+ \supset (-2R, 0) \times S^1 &\to (0, 2R) \times S^1 \subset \dot{\mathcal{U}}_- \\ (s, t) &\mapsto (s + 2R, t + \theta). \end{aligned} \tag{7.16}$$

- $\Theta \subset S$ is the same finite ordered subset as before, but regarded as a subset of $S^{(R,\theta)}$ since the neighborhoods $\mathcal{U}_\pm$ were assumed to be disjoint from Θ.
- $\Delta^R := \Delta\backslash\{z^+, z^-\}$ is endowed with the restriction of the involution $\sigma : \Delta \to \Delta$ and thus regarded as a set of nodes in $S^{(R,\theta)}$, which makes sense since the neighborhoods $\mathcal{U}_\pm$ were also assumed to be disjoint from $\Delta\backslash\{z^+, z^-\}$.

On the open annulus in $S^{(R,\theta)}$ that is obtained by gluing together the two collars, the coordinate systems (s^+, t^+) and (s^-, t^-) are both defined and are related to each other by

$$s^- = s^+ + 2R, \qquad t^- = t^+ + \theta.$$

We can define a third coordinate system $(s^0, t^0) \in (-R, R) \times S^1$ on this region by

$$s^0 := s^+ + R = s^- - R, \qquad t^0 := t^+ = t^- - \theta,$$

thus identifying it with $(-R, R) \times S^1$. We will call this region the **neck** in $S^{(R,\theta)}$ corresponding to the node $\{z^+, z^-\} \in \Delta$.

One can interpret the attaching maps (7.16) in the limit $R \to \infty$ as associating to each $\theta \in S^1$ a decoration $\Phi^\theta : \delta_{z^+} \to \delta_{z^-}$ of the node $\{z^+, z^-\} \subset \Delta$, which is expressed in the chosen cylindrical coordinates as

$$\Phi^\theta(-\infty, t) := (\infty, t + \theta).$$

It is thus natural to extend the family $\{\mathbf{S}^{(R,\theta)}\}_{(R,\theta)\in(0,\infty)\times S^1}$ to include $R = \infty$ by defining $\mathbf{S}^{(\infty,\theta)}$ as the *partially decorated* marked nodal surface

$$\mathbf{S}^{(\infty,\theta)} := (S, \Theta, \Delta, \Phi^\theta).$$

More generally, one can choose disjoint neighborhoods of all nodal points and carry out this construction with pairs of parameters $(R, \theta) \in (0, \infty] \times S^1$ associated independently to each node in Δ. Choosing an ordering of the $N \in \mathbb{N}$ nodes, this produces a family of decorated marked nodal surfaces

$$\{(S^\tau, \Theta, \Delta^\tau, \Phi^\tau)\}_{\tau\in P} \qquad \text{for} \qquad \tau = (R_1, \theta_1, \ldots, R_N, \theta_N) \in P := \left((0, \infty] \times S^1\right)^N,$$

such that S^τ contains a neck of length $2R_i$ corresponding to the ith node whenever $R_i < \infty$. Here Δ^τ contains the ith node if and only if $R_i = \infty$, and Φ^τ denotes the collection of decorations Φ^{θ_i} for all such nodes; in particular, Δ^τ is empty whenever all of the real parameters $R_1, \ldots, R_N$ are finite, and it includes all nodes (equipped additionally with decorations) whenever $R_1 = \ldots = R_N = \infty$. Let us denote the set of parameter values having the latter property by

$$P_\infty := \{R_1 = \ldots = R_N = \infty\} \subset P.$$

We will say that a sequence $\tau_\nu \in P$ **converges to** P_∞ and write "$\tau_\nu \to P_\infty$" if every open neighborhood of P_∞ in P contains τ_ν for all ν sufficiently large.

REMARK 7.3.7. When we discuss the convergence of complex structures on these surfaces, it will be useful to observe that every compact subset $K \subset S \setminus \Delta$ is naturally a subset of S^τ for all τ sufficiently close to P_∞. Moreover, if K contains the truncation S^0 defined by deleting all of the chosen cylindrical coordinate neighborhoods of nodal points from S, then every point in $S^\tau \setminus K$ belongs to a neck, thus $S^\tau \setminus K$ has a natural identification with a disjoint union of subsets of $\mathbb{R} \times S^1$.

DEFINITION 7.3.8. Given a marked nodal Riemann surface (S, j, Θ, Δ), we will refer to any family of decorated marked nodal surfaces $\{(S^\tau, \Theta, \Delta^\tau, \Phi^\tau\}_{\tau \in P}$ constructed via the procedure above as a **pregluing** of (S, j, Θ, Δ). It is uniquely determined by the choices of signs for the two points in each node and the holomorphic cylindrical coordinates chosen near each such point.

DEFINITION 7.3.9. Suppose $\{(S^\tau, \Theta, \Delta^\tau, \Phi^\tau)\}_{\tau \in P}$ is a pregluing of the marked nodal Riemann surface (S, j, Θ, Δ), and j_ν is a sequence of complex structures on S^{τ_ν} for some sequence $\tau_\nu = (R_1^\nu, \theta_1^\nu, \ldots, R_N^\nu, \theta_N^\nu) \in P$ with $\tau_\nu \to P_\infty$. We will say that the sequence of marked nodal Riemann surfaces $(S^{\tau_\nu}, j_\nu, \Theta, \Delta^{\tau_\nu})$ **converges** to (S, j, Θ, Δ) if there exists a compact subset $K \subset S \setminus \Delta$ containing the truncation S^0 such that the following conditions (which make sense due to Remark 7.3.7) are satisfied:

(1) $j_\nu \to j$ in the C^∞-topology on K;
(2) For all ν large, j_ν matches the standard complex structure of $\mathbb{R} \times S^1$ outside of K.

REMARK 7.3.10. The condition $\tau_\nu \to P_\infty$ in Definition 7.3.9 means $\lim_{k \to \infty} R_i^\nu = \infty$ for every $i = 1, \ldots, N$, whereas there is no convergence condition on the parameters θ_i^ν. Imposing such a condition for a subset of the nodes in Δ gives rise to a natural notion of convergence to a *partially decorated* marked nodal Riemann surface $(S, j, \Theta, \Delta, \Phi)$, where the decoration Φ determines the required limits for the θ-parameters at the relevant nodes. This extension of the definition will be useful when we discuss convergence of holomorphic buildings in §7.5. Note that by the compactness of S^1, every sequence that is convergent to (S, j, Θ, Δ) in the sense of Definition 7.3.9 also has a subsequence that converges to $(S, j, \Theta, \Delta, \Phi)$ for some decoration Φ.

EXERCISE 7.3.11. Show that if the convergence condition in Definition 7.3.9 holds, then there also exists another (necessarily non-stable) marked nodal Riemann surface $(S', j', \Theta', \Delta')$, obtained from (S, j, Θ, Δ) by the addition of an extra spherical component with either one or two nodal points and no marked points, such that the elements in the sequence $(S^{\tau_\nu}, j_\nu, \Theta, \Delta^{\tau_\nu})$ are biholomorphically equivalent to the elements of a sequence that converges to $(S', j', \Theta', \Delta')$ in the sense of Definition 7.3.9. *Hint: A neck does not look any different from a truncated cylinder inserted between two necks. Similarly, a disk around a point can be made to look like a truncated plane attached to a neck.*

Exercise 7.3.11 is the reason why $\overline{\mathcal{M}}_{g,\ell}$ is defined to consist only of *stable* nodal surfaces; without stability, the topology determined by the next definition would not be Hausdorff. We will now define a notion of convergence on $\overline{\mathcal{M}}_{g,\ell}$ that is equivalent to the notion described in hyperbolic-geometric terms in [**BEH⁺03**, §4], and thus determines a metrizable topology on $\overline{\mathcal{M}}_{g,\ell}$.

DEFINITION 7.3.12. For $2g + \ell \geqslant 3$, a sequence $[(S_\nu, j_\nu, \Theta_\nu, \Delta_\nu)] \in \overline{\mathcal{M}}_{g,\ell}$ is **convergent** to $[(S, j, \Theta, \Delta)] \in \overline{\mathcal{M}}_{g,\ell}$ if for some pregluing $\{(S^\tau, \Theta, \Delta^\tau, \Phi^\tau)\}_{\tau \in P}$ of (S, j, Θ, Δ), there exists a sequence $\tau_\nu \in P$ and a sequence of complex structures j'_ν on S^{τ_ν} such that the marked nodal Riemann surfaces $(S_\nu, j_\nu, \Theta_\nu, \Delta_\nu)$ and $(S^{\tau_\nu}, j'_\nu, \Theta, \Delta^{\tau_\nu})$ are equivalent for all ν sufficiently large and $(S^{\tau_\nu}, j'_\nu, \Theta, \Delta^{\tau_\nu})$ converges to (S, j, Θ, Δ) in the sense of Definition 7.3.9.

One can show that the topology defined on $\overline{\mathcal{M}}_{g,\ell}$ in this way is independent of the choices underlying the pregluing construction. (A direct proof of this will also emerge from gluing arguments in §10.2.) The following important result can then be regarded as a corollary of Bers's theorem on pair-of-pants decompositions for hyperbolic Riemann surfaces.

THEOREM 7.3.13 (the Deligne-Mumford compactness theorem). *For every $g \geqslant 0$ and $\ell \geqslant 0$ with $2g + \ell \geqslant 3$, the space $\overline{\mathcal{M}}_{g,\ell}$ is compact. In particular, every sequence $[(\Sigma_\nu, j_\nu, \Theta_\nu)] \in \mathcal{M}_{g,\ell}$ has a subsequence that is biholomorphically equivalent to a sequence converging in the sense of Definition 7.3.9 to some stable nodal Riemann surface with arithmetic genus g and ℓ marked points.* □

Here is a slightly simpler description of convergence in $\overline{\mathcal{M}}_{g,\ell}$ that suffices for most applications; one can derive it easily from Definition 7.3.9 by compressing the necks.

PROPOSITION 7.3.14. *Suppose $2g + \ell \geqslant 3$, and $[(\Sigma_\nu, j_\nu, \Theta_\nu)] \in \mathcal{M}_{g,\ell}$ is a sequence converging to $[(S, j, \Theta, \Delta)] \in \overline{\mathcal{M}}_{g,\ell}$. Then the stable nodal marked Riemann surface (S, j, Θ, Δ) admits a decoration Φ such that for sufficiently large ν, there are homeomorphisms*

$$\varphi_\nu : \hat{S}_\Phi \to \Sigma_\nu,$$

smooth outside of C_Φ, which map Θ to Θ_ν preserving the ordering and satisfy

$$\varphi_\nu^* j_\nu \to j \quad \text{in} \quad C^\infty_{\text{loc}}(\hat{S}_\Phi \backslash C_\Delta).$$

□

EXERCISE 7.3.15. The space $\mathcal{M}_{0,4}$ has a natural identification with $S^2 \backslash \{0, 1, \infty\}$, defined by choosing the unique identification of any Riemann sphere carrying four marked points $(S^2, j, (z_1, z_2, z_3, z_4))$ with $\mathbb{C} \cup \{\infty\}$ such that z_1, z_2, z_3 are identified with $0, 1, \infty$ respectively, while z_4 is sent to some point in $S^2 \backslash \{0, 1, \infty\}$. Show that this extends continuously to an identification of $\overline{\mathcal{M}}_{0,4}$ with S^2. What do the three nodal curves in $\overline{\mathcal{M}}_{0,4} \backslash \mathcal{M}_{0,4}$ look like in terms of pair-of-pants decompositions?

7.4. The compactified moduli space

We now introduce the natural compactification of $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$.

7.4.1. Nodal curves. A punctured J-holomorphic **nodal curve** in $(\widehat{W}, J)$ with $m \geqslant 0$ marked points consists of the data $(S, j, \Gamma^+, \Gamma^-, \Theta, \Delta, u)$, where

- $(S, j, \Gamma \cup \Theta, \Delta)$ is a marked nodal Riemann surface such that $\Gamma^+, \Gamma^+, \Theta \subset S$ are disjoint finite sets, $\Gamma := \Gamma^+ \cup \Gamma^-$ and $\#\Theta = m$;
- $u : (\dot{S}, j) \to (\widehat{W}, J)$ for $\dot{S} := S \backslash \Gamma$ is an asymptotically cylindrical J-holomorphic map with positive punctures Γ^+ and negative punctures Γ^- such that for each node $\{z^+, z^-\} \subset \Delta$, $u(z^+) = u(z^-)$.

Equivalence of two nodal curves

$$(S_0, j_0, \Gamma_0^+, \Gamma_0^-, \Theta_0, \Delta_0, u_0) \sim (S_1, j_1, \Gamma_1^+, \Gamma_1^-, \Theta_1, \Delta_1, u_1)$$

is defined as the existence of an equivalence of marked nodal Riemann surfaces

$$\varphi : (S_0, j_0, \Gamma_0^+ \cup \Gamma_0^- \cup \Theta_0, \Delta_0) \to (S_1, j_1, \Gamma_1^+ \cup \Gamma_1^- \cup \Theta_1, \Delta_1)$$

such that $u_0 = u_1 \circ \varphi$. We say that $(S, j, \Gamma^+, \Gamma^-, \Theta, \Delta, u)$ is **connected** if the nodal Riemann surface $(S, j, \Gamma \cup \Theta, \Delta)$ is connected, and its **arithmetic genus** is then defined to be the arithmetic genus of the latter. We say that $(S, j, \Gamma^+, \Gamma^-, \Theta, \Delta, u)$ is **stable** if every connected component of $S \backslash (\Gamma \cup \Theta \cup \Delta)$ on which u is constant has negative Euler characteristic. Note that the underlying marked nodal Riemann surface $(S, j, \Gamma \cup \Theta, \Delta)$ need not be stable in general. We will see in Chapter 10 (see Proposition 10.1.1) that a nodal curve is stable if and only if its automorphism group (i.e. its group of self-equivalences) is finite.

REMARK 7.4.1. If $(\widehat{W}, J)$ is a one-point space, the nodal curves in $(\widehat{W}, J)$ are always constant and cannot have any punctures (since $\widehat{W}$ has no cylindrical ends), thus there is a natural one-to-one correspondence between nodal curves in this setting and marked nodal Riemann surfaces. Note that the two notions of stability obtained in this way for a marked nodal Riemann surface are equivalent.

Nodal curves are sometimes also referred to as *holomorphic buildings of height 1.* These are the objects that form the *Gromov compactification* of $\mathcal{M}_{g,m}(J, A)$ when W is a closed symplectic manifold. One can now roughly imagine how the compactness theorem in that setting is proved: Given a converging sequence of almost complex structures $J_\nu \to J$ and a sequence $[(\Sigma_\nu, j_\nu, \Theta_\nu, u_\nu)] \in \mathcal{M}_{g,m}(J_\nu, A_\nu)$ with uniformly bounded energy, we can first add some auxiliary marked points if necessary to assume that $2g + m \geqslant 3$. Now a subsequence of the domains $[(\Sigma_\nu, j_\nu, \Theta_\nu)] \in \mathcal{M}_{g,m}$ converges to an element of the Deligne-Mumford space $[(S, j, \Theta, \Delta)] \in \overline{\mathcal{M}}_{g,m}$. Concretely, this implies that for large ν, our sequence in $\mathcal{M}_{g,m}(J_\nu, A_\nu)$ admits representatives

$(\Sigma, j'_\nu, \Theta, u'_\nu)$, with Σ a fixed surface with fixed marked points $\Theta \subset \Sigma$, and (S, j, Θ, Δ) admits decorations Φ so that one can identify $\widehat{S}_\Phi$ with Σ and find

$$j'_\nu \to j \quad \text{in} \quad C^\infty_{\text{loc}}(\Sigma \backslash C)$$

for some collection of disjoint circles $C \subset \Sigma$. The connected components of $(\Sigma \backslash C, j)$ are then biholomorphically equivalent to the connected components of $(S \backslash \Delta, j)$, and if the newly reparametrized maps $u'_\nu : \Sigma \to W$ are uniformly C^1_{loc}-bounded on $\Sigma \backslash C$, then a subsequence converges in $C^\infty_{\text{loc}}(\Sigma \backslash C)$ to a limiting finite-energy J-holomorphic map $u_\infty : (S \backslash \Delta, j) \to (W, J)$, whose singularities at Δ are removable. In particularly nice cases, this may be the end of the story, and our subsequence of $[(\Sigma_\nu, j_\nu, \Theta_\nu, u_\nu)] \in \mathcal{M}_{g,m}(J_\nu, A_\nu)$ converges to the nodal curve $[(S, j, \Theta, \Delta, u_\infty)]$; in particular, the domain $[(S, j, \Theta, \Delta)]$ in this case is stable and is thus an element of $\overline{\mathcal{M}}_{g,m}$. But more complicated things can also happen, e.g. u'_ν might not be C^1-bounded, in which case there is bubbling. The bubbles that arise in this setting will be planes with removable punctures, i.e. spheres, so they produce extra domain components with nonnegative Euler characteristic, but since they are never constant, the limiting nodal curve is still considered stable. Similarly, since $\Sigma \backslash C$ is not compact, there can also be breaking as in Figure 7.3, producing more non-stable domain components—but again, the limiting map on these components will never be constant.

7.4.2. Holomorphic buildings. Only a small subset of the phenomena observed in §7.3 can be described via nodal curves: we've seen that in general, we also have to allow "broken" curves with multiple "levels". This notion can be formalized as follows.

DEFINITION 7.4.2. Assume $(\widehat{W}, J)$ is an almost complex manifold with cylindrical ends obtained by completing a symplectic cobordism (W, ω) with stable boundary components $(M_+, \mathcal{H}_+)$ and $(M_-, \mathcal{H}_-)$ and choosing $J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$. Given integers $g, m, N_+, N_- \geqslant 0$, a **holomorphic building of height** $N_-|1|N_+$ in $(\widehat{W}, J)$ with arithmetic genus g and m marked points is a tuple

$$\mathbf{u} = (S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\text{nd}}, \Delta^{\text{br}}, L, \Phi, u),$$

with the various data defined as follows:

- The **domain** $(S, j, \Gamma \cup \Theta, \Delta^{\text{nd}} \cup \Delta^{\text{br}})$ is a marked nodal Riemann surface of arithmetic genus g, where the finite sets $\Gamma^+, \Gamma^-, \Theta, \Delta^{\text{nd}}, \Delta^{\text{br}} \subset S$ are all disjoint, $\Gamma = \Gamma^+ \cup \Gamma^-$, $\#\Theta = m$, and the fixed-point-free involution on $\Delta^{\text{nd}} \cup \Delta^{\text{br}}$ is assumed to preserve the subsets Δ^{nd} and Δ^{br}. Matched pairs in these subsets are called the **nodes** and **breaking pairs** respectively of $\mathbf{u}$. The **marked points** of $\mathbf{u}$ are the points in Θ, while Γ^+ and Γ^- are its positive and negative **punctures** respectively.
- The **level structure** is a locally constant function

$$L : S \to \{-N_-, \ldots, -1, 0, 1, \ldots, N_+\}$$

that attains every value in $\{-N_-, \ldots, N_+\}$ except possibly 0, and satisfies:
(1) $L(z^+) = L(z^-)$ for each node $\{z^+, z^-\} \subset \Delta^{\text{nd}}$;

(2) Each breaking pair $\{z^+, z^-\} \subset \Delta^{\mathrm{br}}$ can be labelled such that $L(z^+) - L(z^-) = 1$;
(3) $L(\Gamma^+) = \{N_+\}$ and $L(\Gamma^-) = \{-N_-\}$.

- The **decoration** is a choice of orientation-reversing orthogonal map
$$\delta_{z^+} \xrightarrow{\Phi} \delta_{z^-}$$
for each breaking pair $\{z^+, z^-\} \subset \Delta^{\mathrm{br}}$.
- The **map** is an asymptotically cylindrical pseudoholomorphic curve
$$u : (\dot{S} := S \backslash (\Gamma \cup \Delta^{\mathrm{br}}), j) \to \coprod_{N \in \{-N_-, \dots, N_+\}} (\widehat{W}_N, J^N),$$
where
$$(\widehat{W}_N, J^N) := \begin{cases} (\mathbb{R} \times M_+, J^+) & \text{for } N \in \{1, \dots, N_+\}, \\ (\widehat{W}, J) & \text{for } N = 0, \\ (\mathbb{R} \times M_-, J^-) & \text{for } N \in \{-N_-, \dots, -1\}, \end{cases}$$
and u sends $\dot{S} \cap L^{-1}(N)$ into $\widehat{W}_N$ for each N, with positive punctures at Γ^+ and negative punctures at Γ^-. Moreover,
$$u(z^+) = u(z^-) \quad \text{for every node } \{z^+, z^-\} \subset \Delta^{\mathrm{nd}},$$
and for each breaking pair $\{z^+, z^-\} \subset \Delta^{\mathrm{br}}$ labelled with $L(z^+) - L(z^-) = 1$, u has a positive puncture at z^- and a negative puncture at z^+ asymptotic to the same orbit, called the **breaking orbit** at $\{z^+, z^-\}$, such that if $u_+ : \delta_{z^+} \to M_\pm$ and $u_- : \delta_{z^-} \to M_\pm$ denote the induced asymptotic parametrizations of the orbit, then
$$u_+ = u_- \circ \Phi : \delta_{z^+} \to M_\pm.$$

The following additional notation and terminology for the building $\mathbf{u}$ in Definition 7.4.2 will be useful to keep in mind. For each $N \in \{-N_-, \dots, 0, \dots, N_+\}$, denote
$$\dot{S}_N := \big(S \backslash (\Gamma \cup \Delta^{\mathrm{br}})\big) \cap L^{-1}(N),$$
and denote the restriction of u to this subset by
$$u^N : \dot{S}_N \to \begin{cases} \mathbb{R} \times M_+ & \text{if } N > 0, \\ \widehat{W} & \text{if } N = 0, \\ \mathbb{R} \times M_- & \text{if } N < 0. \end{cases}$$
Including $\Theta \cap L^{-1}(N)$ and $\Delta^{\mathrm{nd}} \cap L^{-1}(N)$ in the data defines u^N as a (generally disconnected) nodal curve with marked points, whose positive punctures are in bijective correspondence with the negative punctures of u^{N+1} if $N < N_+$. We call u_N the **Nth level** of $\mathbf{u}$, and call it an **upper** or **lower** level if $N > 0$ or $N < 0$ respectively, and the **main level** if $N = 0$.[2] By convention, every holomorphic building in $\widehat{W}$ has exactly one main level, which lives in $\widehat{W}$ itself, and arbitrary nonnegative numbers of upper and lower levels, which live in the symplectizations $\mathbb{R} \times M_\pm$. One

[2]The upper levels collectively are also sometimes called the **upper layer** of the building; similarly the **lower layer** and **main layer**, where the latter consists only of the main level.

slightly subtle detail is that it is possible for the main level to be *empty*, meaning 0 is not in the image of the level function L. The requirement that L should attain every other value from $-L_-$ to L_+ is a convention to ensure that upper and lower levels are not empty, so e.g. if a building has an empty main level and no lower levels, then the lowest nonempty upper level is always labelled 1 instead of something arbitrary.

The positive punctures of the topmost level of $\mathbf{u}$ are Γ^+, and the negative punctures of the bottommost level are Γ^-, so these give rise to lists of positive/negative asymptotic orbits $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \ldots, \gamma_{k_\pm}^\pm)$ in $M_\pm$. There is also a relative homology class

$$[\mathbf{u}] \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-).$$

To define this, notice that the its definition in §6.4 for smooth curves $u : \dot{\Sigma} \to \widehat{W}$ can be reformulated in the following way: there is a retraction $\pi : \widehat{W} \to W$ that collapses each cylindrical end to $M_\pm \subset \partial W$, and since u is asymptotically cylindrical, the map $\pi \circ u : \dot{\Sigma} \to W$ extends to a continuous map on the circle compactification,

$$\bar{u} : \bar{\Sigma} \to W,$$

whose relative homology class is $[u]$. The conditions on nodes and breaking orbits allow us to perform a similar trick for the building $\mathbf{u}$, using the map

$$\pi : \coprod_{N \in \{-N_-, \ldots, N_+\}} \widehat{W}_N \to W$$

which acts as the identity on W but collapses cylindrical ends of $\widehat{W}$ to ∂W and similarly collapses each copy of $\mathbb{R} \times M_\pm$ to $M_\pm \subset \partial W$. Extending the decorations Φ arbitrarily to decorations of the nodes Δ^{nd}, one can then take the circle compactification of $\dot{S} := S \backslash (\Gamma \cup \Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}})$ and glue matching boundary components together along Φ to form a compact surface with boundary $\overline{S}_\Phi$ such that $\pi \circ u : \dot{S} \to W$ extends to a continuous map

$$\bar{u} : \overline{S}_\Phi \to W.$$

Its relative homology class defines $[\mathbf{u}] \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$.

DEFINITION 7.4.3. We say that the building $\mathbf{u}$ is **stable** if two properties hold:

(1) Every connected component of $S \backslash (\Gamma \cup \Theta \cup \Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}})$ on which the map u is constant has negative Euler characteristic;
(2) There is no $N \in \{-N_-, \ldots, -1\} \cup \{1, \ldots, N_+\}$ for which the Nth level consists entirely of a disjoint union of trivial cylinders without any marked points or nodes.

An **equivalence** between two holomorphic buildings

$$\mathbf{u}_i = (S_i, j_i, \Gamma_i^+, \Gamma_i^-, \Theta_i, \Delta_i^{\mathrm{nd}}, \Delta_i^{\mathrm{br}}, L_i, \Phi_i, u_i), \qquad i = 0, 1$$

is defined as an equivalence of partially decorated marked nodal Riemann surfaces

$$(S_0, j_0, \Gamma_0 \cup \Theta_0, \Delta_0^{\mathrm{nd}} \cup \Delta_0^{\mathrm{br}}, \Phi_0) \xrightarrow{\varphi} (S_1, j_1, \Gamma_1 \cup \Theta_1, \Delta_1^{\mathrm{nd}} \cup \Delta_1^{\mathrm{br}}, \Phi_1)$$

such that $\varphi(\Gamma_0^\pm) = \Gamma_1^\pm$, $\varphi(\Theta_0) = \Theta_1$, $\varphi(\Delta_0^{\mathrm{nd}}) = \Delta_1^{\mathrm{nd}}$, $\varphi(\Delta_0^{\mathrm{br}}) = \Delta_1^{\mathrm{br}}$, $L_1 \circ \varphi = L_0$, and

$$u_1^0 \circ \varphi = u_0^0,$$

while

$$u_1^N \circ \varphi = u_0^N \text{ up to } \mathbb{R}\text{-translation} \quad \text{for each } N \neq 0.$$

We will show in Chapter 10 (see Proposition 10.1.1) that the stability of a building $\mathbf{u}$ is equivalent to the finiteness of its automorphism group, i.e. its group of self-equivalences.

Given lists of orbits $\boldsymbol{\gamma}^{\pm}$ and a relative homology class A, the set of equivalence classes of stable holomorphic buildings in $(\widehat{W}, J)$ with arithmetic genus g and m marked points, positively/negatively asymptotic to $\boldsymbol{\gamma}^{\pm}$ and homologous to A will be denoted by

$$\overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-).$$

Observe that for any $A \neq 0$, there is a natural inclusion $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) \subset \overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ defined by regarding J-holomorphic curves in $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ as buildings with no upper or lower levels and no nodes. Such buildings are always stable if the sets of asymptotic orbits are nonempty or $A \neq 0$, because they are not constant; if there are no orbits and $A = 0$, then one must add the condition $2g + m \geqslant 3$ to ensure stability (cf. Remark 6.4.3).

REMARK 7.4.4. Generalizing Example 6.5.7 and Remark 7.4.1, $\overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ reduces in the case $W = \{\text{pt}\}$ to the Deligne-Mumford moduli space $\overline{\mathcal{M}}_{g,m}$

7.5. The SFT topology

In §7.3.3, we described the topology of the Deligne-Mumford space $\overline{\mathcal{M}}_{g,\ell}$ in terms of families of Riemann surfaces called *pregluings*, which contain long necks $([-R, R] \times S^1, i)$ with $R \to \infty$. We shall now give a similar definition for convergence of holomorphic buildings, assuming as usual that we have C^∞-convergence $J_\nu \to J$ of a sequence of almost complex structures in $\mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$. We will focus on the special case where a sequence of smooth curves converges to a building, since everything else can be derived from this. You may want to have a look at Figure 7.8 for intuition before proceeding; the purpose of the somewhat lengthy definitions below is, essentially, to give a precise characterization of what is happening in that picture and others like it.

We first give local descriptions of what it means for a sequence of smooth curves to degenerate along a node or a breaking orbit. The next two definitions serve a purpose similar to that of Definition 6.5.2: by allowing arbitrary sequences of shifts along the cylinder, they ensure that no important information is missed when we consider convergence only on compact subsets.

DEFINITION 7.5.1. Assume $u_\nu : [-R_\nu, R_\nu] \times S^1 \to \widehat{W}$ is a sequence of smooth maps with $R_\nu \to \infty$. We will say that this sequence **degenerates to a node** at the point $p \in \widehat{W}$ if for every sequence $c_\nu \in [-R_\nu, R_\nu]$ such that

$$a_\nu := -R_\nu - c_\nu \to -\infty \qquad \text{and} \qquad b_\nu := R_\nu - c_\nu \to +\infty,$$

the sequence of maps

$$[a_\nu, b_\nu] \times S^1 \to \widehat{W} : (s, t) \mapsto u_\nu(s + c_\nu, t)$$

converges in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to the constant map with image p.

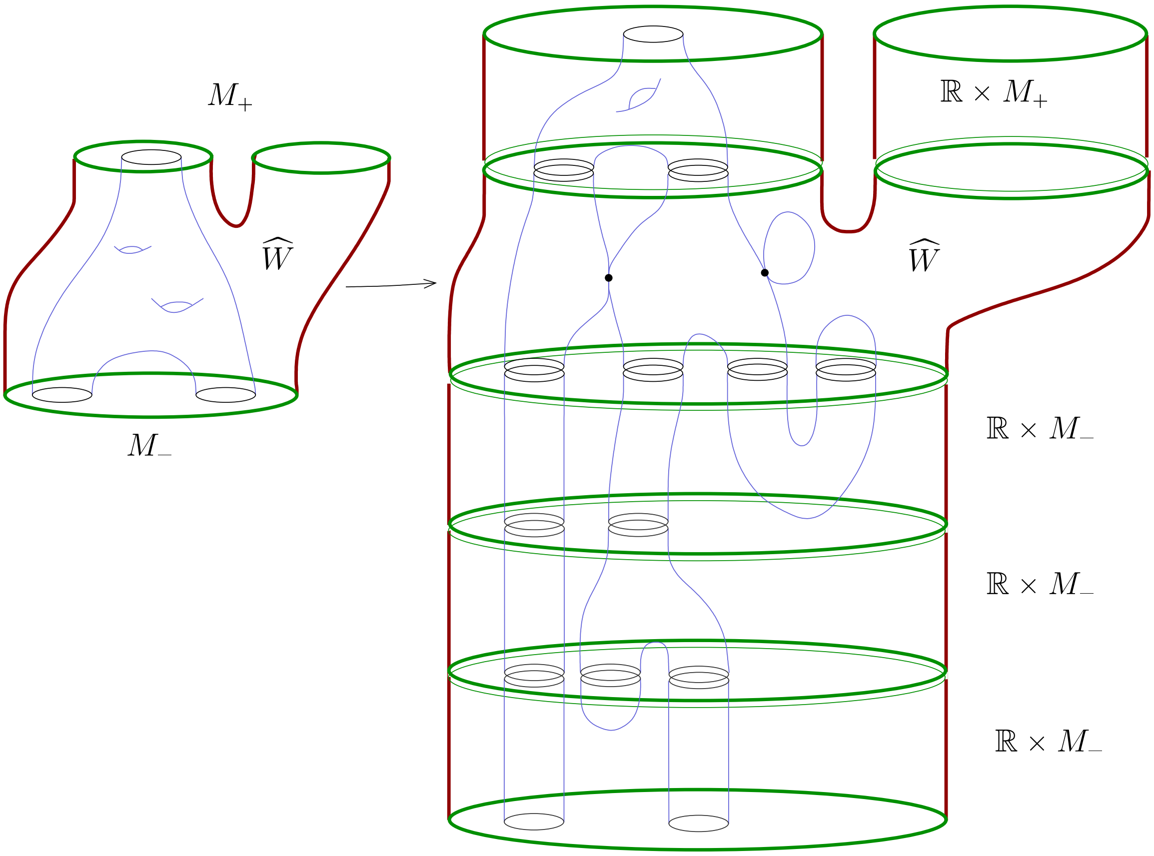

FIGURE 7.8. Convergence to a building with arithmetic genus 2, one upper level and three lower levels.

DEFINITION 7.5.2. Assume $u_\nu : [-R_\nu, R_\nu] \times S^1 \to \widehat{W}$ is a sequence of smooth maps with $R_\nu \to \infty$. We will say that this sequence **breaks along** a Reeb orbit $\gamma : S^1 \to M_\pm$ if the image of each u_ν for ν sufficiently large is contained in the (positive or negative resp.) cylindrical end and, for every sequence $c_\nu \in [-R_\nu, R_\nu]$ with

$$a_\nu := -R_\nu - c_\nu \to -\infty \qquad \text{and} \qquad b_\nu := R_\nu - c_\nu \to +\infty,$$

there exist sequences $\theta_\nu \in S^1$ and $r_\nu \in \mathbb{R}$ such that the sequence of maps

$$[a_\nu, b_\nu] \times S^1 \to \mathbb{R} \times M_\pm : (s,t) \mapsto \tau_{r_\nu} \circ u_\nu(s + c_\nu, t + \theta_\nu)$$

converges in $C^\infty_{\mathrm{loc}}(\mathbb{R} \times S^1)$ to the trivial cylinder over γ.

DEFINITION 7.5.3. Assume $\mathbf{u} = (S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\mathrm{nd}}, \Delta^{\mathrm{br}}, L, \Phi, u)$ is a holomorphic building of height $N_-|1|N_+$ in $(\widehat{W}, J)$, $\{(S^\tau, \Gamma \cup \Theta, \Delta^{\mathrm{nd}}_\tau \cup \Delta^{\mathrm{br}}_\tau, \Phi^\tau)\}_{\tau \in P}$ is a pregluing of the marked nodal Riemann surface $(S, j, \Gamma \cup \Theta, \Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}})$, and we are additionally given the following sequences:

- Parameters $\tau_\nu = (R_1^\nu, \theta_1^\nu, \ldots, R_N^\nu, \theta_N^\nu) \in P$ with $R_i^\nu < \infty$ for every $i = 1, \ldots, N$ and $\tau_\nu \to P_\infty$;

- Complex structures j_ν on the smooth surfaces S^{τ_ν} that match j near $\Gamma^+ \cup \Gamma^-$ and match the standard complex structure i on each of the necks $[-R_i^\nu, R_i^\nu] \times S^1 \subset S^{\tau_\nu}$;
- Asymptotically cylindrical holomorphic maps $u_\nu : (S^{\tau_\nu} \backslash \Gamma, j_\nu) \to (\widehat{W}, J_\nu)$.

For each $N \in \{-N_-, \ldots, N_+\}$, let $\ddot{S}_N := (S \backslash (\Gamma \cup \Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}})) \cap L^{-1}(N)$ and observe that every compact subset of $\ddot{S}_N$ can be regarded as a subset of S^{τ_ν} for ν sufficiently large (cf. Remark 7.3.7). We will say that the sequence of smooth J_ν-holomorphic curves $(S^{\tau_\nu}, j_\nu, \Gamma^+, \Gamma^-, \Theta, u_\nu)$ **converges** to the J-holomorphic building $\mathbf{u}$ if the following conditions hold:

(1) *Domains*: The marked Riemann surfaces $(S^{\tau_\nu}, j_\nu, \Gamma \cup \Theta)$ converge in the sense of Definition 7.3.9 and Remark 7.3.10 to the partially decorated marked nodal Riemann surface $(S, j, \Gamma \cup \Theta, \Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}}, \Phi)$.

(2) *Main level*:
$$u_\nu|_{\ddot{S}_0} \to u|_{\ddot{S}_0} \quad \text{in} \quad C^\infty_{\mathrm{loc}}(\ddot{S}_0, \widehat{W}),$$
and for each main-level node $\{z^+, z^-\} \subset \Delta^{\mathrm{nd}} \cap L^{-1}(0)$, the restriction of u_ν to the corresponding neck $[-R_i^\nu, R_i^\nu] \times S^1 \subset S^{\tau_\nu}$ degenerates (in the sense of Definition 7.5.1) to a node at the point $u(z^+) = u(z^-) \in \widehat{W}$.

(3) *Upper and lower levels*: For each $\pm N > 0$, $u_\nu(\ddot{S}_N)$ is contained in the positive/negative cylindrical end for all ν sufficiently large, and there exists a sequence $r_\nu^N \to \pm\infty$ such that the resulting $\mathbb{R}$-translations converge:
$$\tau_{-r_\nu^N} \circ u_\nu|_{\ddot{S}_N} \to u|_{\ddot{S}_N} \quad \text{in} \quad C^\infty_{\mathrm{loc}}(\ddot{S}_N, \mathbb{R} \times M_\pm).$$
Moreover, for each node $\{z^+, z^-\} \subset \Delta^{\mathrm{nd}} \cap L^{-1}(N)$ at level N, the restriction of $\tau_{-r_\nu^N} \circ u_\nu$ to the corresponding neck $[-R_i^\nu, R_i^\nu] \times S^1 \subset S^{\tau_\nu}$ degenerates (in the sense of Definition 7.5.1) to a node at the point $u(z^+) = u(z^-) \in \mathbb{R} \times M_\pm$.

(4) *Separation of levels*: The sequences $r_\nu^N \to \pm\infty$ satisfy
$$\lim_{\nu \to \infty} \left(r_\nu^{N+1} - r_\nu^N \right) = +\infty \quad \text{for all } N < N_+.$$

(5) *Punctures*: u_ν converges asymptotically (in the sense of Definition 6.5.2) at each puncture $z \in \Gamma^\pm$ to the corresponding asymptotic orbit of u.

(6) *Breaking orbits*: For each breaking pair $\{z^-, z^+\} \subset \Delta^{\mathrm{br}}$, the restrictions of u_ν to the corresponding neck $[-R_i^\nu, R_i^\nu] \times S^1 \subset S^{\tau_\nu}$ break (in the sense of Definition 7.5.2) along the corresponding breaking orbit of u.

Generalizing Exercise 7.3.11, we have:

EXERCISE 7.5.4. Show that if $\mathbf{u} = (S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\mathrm{nd}}, \Delta^{\mathrm{br}}, L, \Phi, u)$ is a stable holomorphic building and we are given a sequence for which the convergence condition in Definition 7.5.3 holds, then it also holds with $\mathbf{u}$ replaced by a different building that is not stable.

DEFINITION 7.5.5. A sequence
$$[(\Sigma_\nu, j_\nu, \Gamma_\nu^+, \Gamma_\nu^-, \Theta_\nu, u_\nu)] \in \mathcal{M}_{g,m}(J_\nu, A_\nu, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$$

converges to

$$[(S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\mathrm{nd}}, \Delta^{\mathrm{br}}, L, \Phi, u)] \in \overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$$

if for ν sufficiently large, the equivalence classes in the sequence admit representatives that converge to $(S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\mathrm{nd}}, \Delta^{\mathrm{br}}, L, \Phi, u)$ in the sense described in Definition 7.5.3.

We refer to [**BEH**$^+$**03**] or [**Abb14**] for the proof that the notion of convergence described above determines a metrizable topology on the compactified moduli space $\overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$. We will refer to this topology as **the SFT topology**; in the literature, it is sometimes also called the *Gromov-Hofer topology.* Just as in the uncompactified moduli space, one can show that convergence in the SFT topology preserves relative homology classes. Exercise 7.5.4 demonstrates why it is essential to include the word "stable" in the definition of the moduli space $\overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$, as without this condition, the moduli space could not be Hausdorff.

The main result of [**BEH**$^+$**03**] can now be expressed as follows:

THEOREM 7.5.6. *Fix integers $g \geqslant 0$ and $m \geqslant 0$, assume all Reeb orbits in $(M, \mathcal{H}_+)$ and $(M, \mathcal{H}_-)$ are nondegenerate and that $J_\nu \to J$ in $\mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$. Then for any sequence*

$$[(\Sigma_\nu, j_\nu, \Gamma_\nu^+, \Gamma_\nu^-, \Theta_\nu, u_\nu)] \in \mathcal{M}_{g,m}(J_\nu, A_\nu, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$$

of stable J_ν-holomorphic curves in $\widehat{W}$ with uniformly bounded energy $E(u_\nu)$, a subsequence converges to some stable holomorphic building

$$[\mathbf{u}_\infty] = [(S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\mathrm{nd}}, \Delta^{\mathrm{br}}, L, \Phi, u_\infty)] \in \overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-).$$

REMARK 7.5.7. The theorem is also true under the more general hypothesis that the Reeb vector fields are Morse-Bott. In this case, one can also allow the asymptotic Reeb orbits of the sequence to vary, as long as the sum of their periods is uniformly bounded—such a bound plays the role of an energy bound and guarantees a convergent subsequence of orbits via the Arzelà-Ascoli theorem.

REMARK 7.5.8. Stability of the limit in Theorem 7.5.6 is guaranteed for the same reasons as in our discussion of Gromov compactness in §7.4.1: stable domains degenerate to stable nodal domains as geodesics in pair-of-pants decompositions shrink to zero length, while bubbling and breaking produce additional domain components that are not stable, but on which the maps are never trivial.

7.6. The translation-invariant setting

A few minor modifications to the above discussion are necessary to compactify the moduli space of curves in a symplectization $(\mathbb{R} \times M, J)$ for $J \in \mathcal{J}(\mathcal{H})$. It is possible to view this as a special case of a completed symplectic cobordism, but this perspective produces a certain amount of extraneous data that is not meaningful. The key observation is that in the presence of an $\mathbb{R}$-action, one should really compactify $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)/\mathbb{R}$ instead of $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$. The compactification $\overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ then consists of holomorphic buildings as defined in §7.4.2, but since all levels live in the same symplectization $\mathbb{R} \times M$, there is no longer a distinguished *main level* or any meaningful notion of *upper* vs. *lower* levels; the level structure is simply a function $L : S \to \{1, \dots, N\}$ for some $N \in \mathbb{N}$, called the **height** of the building, and equivalence of buildings must permit $\mathbb{R}$-translations within each level. For these reasons, the SFT compactness theorem in symplectizations has a few qualitative differences, but is still very much analogous to what is presented in §7.5.

7.7. Stretching the neck

The main results of [**BEH+03**] include one more compactness theorem that arises often in applications: it involves holomorphic curves in a degenerating sequence of symplectic manifolds that are being "stretched" along a hypersurface. Continuing in the setting of §7.2, where $(\widehat{W}, \omega_h)$ is the completion of a symplectic cobordism (W, ω) with stable boundary, assume that an additional closed stable hypersurface

$$V \subset W$$

in the interior of W is given. Note that we are not assuming V is connected, so it may in fact be a union of many disjoint closed connected hypersurfaces. Figure 7.9 shows an example in which V has two connected components, labelled $V_1, V_2 \subset W$.

Stability means that V carries a stable Hamiltonian structure $\mathcal{H} = (\omega_V := \omega|_{TV}, \lambda)$ such that, by Proposition 6.1.9, some neighborhood $\mathcal{N}(V) \subset W$ of V admits a symplectic identification with the collar

$$(\mathcal{N}^{(\epsilon)}, \omega^{(\epsilon)}) := \big((-\epsilon, \epsilon) \times V, \omega_V + d(r\lambda)\big)$$

for $\epsilon > 0$ sufficiently small, where r denotes the coordinate on $(-\epsilon, \epsilon)$ and $V \subset \mathcal{N}(V)$ is identified with $\{0\} \times V \subset \mathcal{N}^{(\epsilon)}$. Now for $T \geqslant \epsilon$, we consider a family of smooth manifolds defined by

$$W^T := (W \backslash V) \cup \mathcal{N}^T,$$

where the enlarged collar $\mathcal{N}^T := (-T, T) \times V$ is glued in via identifications $\mathcal{N}(V) \backslash V \ni (r, x) \sim (r + T - \epsilon, x) \in \mathcal{N}^T$ for $r \in (0, \epsilon)$ and $\mathcal{N}(V) \backslash V \ni (r, x) \sim (r - T + \epsilon, x)$ for $r \in (-\epsilon, 0)$. In other words, W^T is obtained from W via a topologically trivial surgery that replaces the small collar neighborhood $\mathcal{N}(V) \subset W$ with a longer "neck" of the form $(-T, T) \times V$; see Figure 7.9. Any choice of diffeomorphism $f_T : (-T, T) \to (-\epsilon, \epsilon)$ that satisfies $f_T' \equiv 1$ outside a compact subset then determines a diffeomorphism

$$\Psi_T : W^T \to W$$

that sends $\mathcal{N}^T$ to $\mathcal{N}^{(\epsilon)} \cong \mathcal{N}(V)$ via $(r, x) \mapsto (f_T(r), x)$ and acts everywhere else as the identity map, i.e. the canonical identification of $W^T \backslash \mathcal{N}^T$ with $W \backslash \mathcal{N}(V)$. We define a symplectic form ω^T on W^T so that Ψ_T becomes a symplectomorphism, so

$$\omega^T := \Psi_T^* \omega = \omega_V + d\big(f_T(r)\lambda\big) \quad \text{on} \quad \mathcal{N}^T \subset W^T.$$

Since (W^T, ω^T) is identical to (W, ω) near the boundary, it is also a symplectic cobordism with stable boundary and has a similarly constructed completion $(\widehat{W}^T, \omega_h^T)$. Note that the function f_T is an arbitrary choice, so the symplectic form ω^T is not canonical on the neck $\mathcal{N}^T$, but it will be useful to notice that its cohomology class

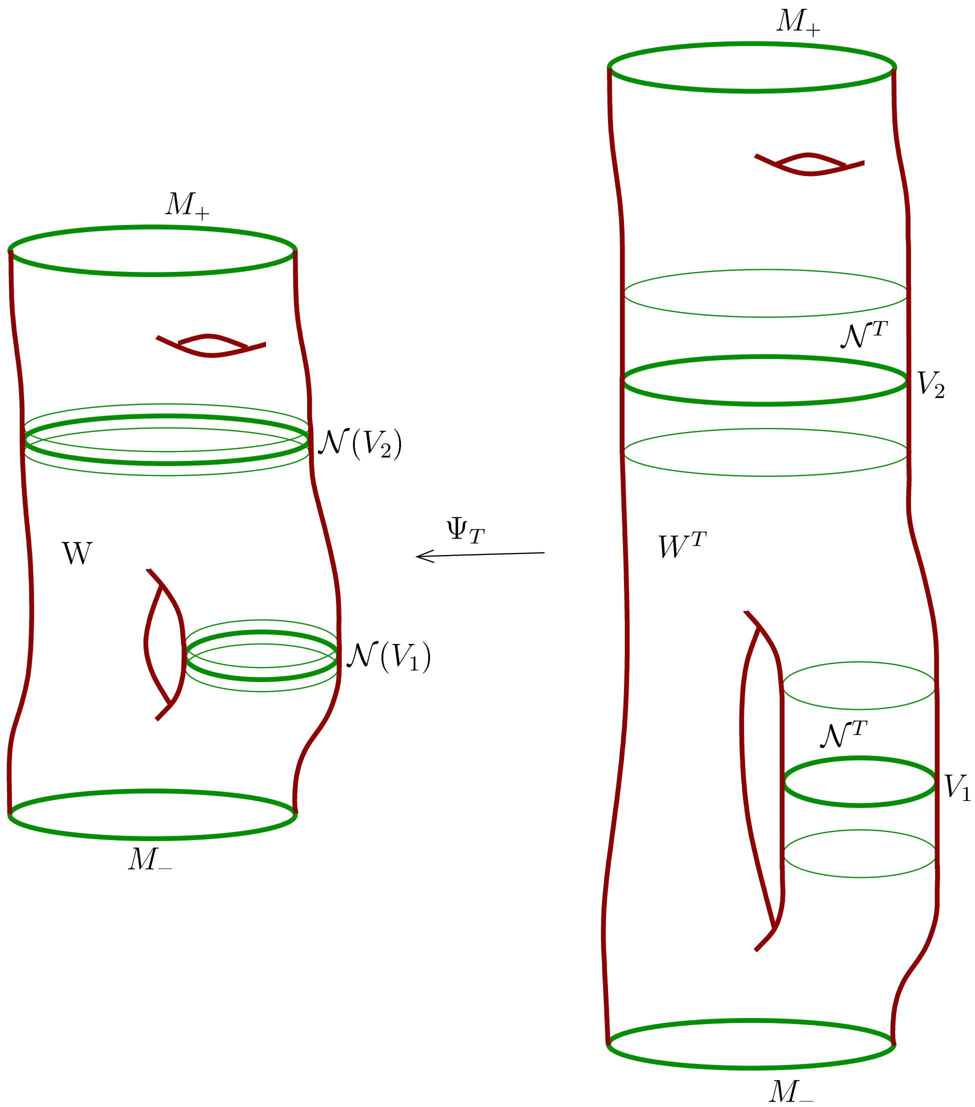

FIGURE 7.9. Stretching a symplectic cobordism W along an interior stable hypersurface $V = V_1 \coprod V_2$ produces a family of symplectic cobordisms $\{W^T\}_{T>\epsilon}$ containing long "necks" $(-T,T) \times V \cong \mathcal{N}^T \subset W^T$.

is independent of choices: in fact, any two symplectic forms ω_1^T and ω_2^T constructed in this way on W^T are related to each other by

$$\omega_1^T - \omega_2^T = d\Lambda \quad \text{for some } \Lambda \in \Omega^1(W^T) \text{ with compact support in } \mathcal{N}^T. \tag{7.17}$$

Any choice of almost complex structure $J_V \in \mathcal{J}(\mathcal{H})$ on the symplectization $\mathbb{R} \times V$ defines an ω-tame translation-invariant almost complex structure on $(-\epsilon, \epsilon) \times V \cong \mathcal{N}(V) \subset W$, which can then be extended to a global almost complex structure $J \in \mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ on $\widehat{W}$. Recall from Proposition 6.2.2 that after shrinking $\epsilon > 0$ if necessary, we can assume without loss of generality that J_V is also tamed by ω^T on $(-T,T) \times V$ for every possible choice of the function f_T. With this assumption in place, we can then define a family of tame almost complex structures $J^T \in \mathcal{J}_\tau(\omega^T, \mathcal{H}_+, \mathcal{H}_-)$ on the family of completed symplectic cobordisms $\widehat{W}^T$ by

$$J^T := \begin{cases} J_V & \text{on } \mathcal{N}^T \subset \widehat{W}^T, \\ J & \text{on } \widehat{W}^T \backslash \mathcal{N}^T = \widehat{W} \backslash \mathcal{N}(V), \end{cases} \qquad \epsilon \leqslant T < \infty.$$

This family degenerates as $T \to \infty$, i.e. pushing each J^T forward through the symplectomorphism $\Psi_T : (W^T, \omega^T) \to (W, \omega)$ produces a family in $\mathcal{J}_\tau(W, \omega)$ that has no well-defined limit along $V \subset W$ as $T \to \infty$. One should instead visualize the limit for $T \to \infty$ as a new almost complex manifold with two extra cylindrical ends. Consider the compact manifold W_V obtained by spltting W apart along V, i.e. W_V is a compact manifold whose boundary contains an open and closed subset canonically identified with the disjoint union of two copies of V, such that gluing these two copies to each other via their canonical identification reproduces W. One can now understand W_V as a symplectic cobordism from $M_- \coprod V$ to $M_+ \coprod V$, and its completion $\widehat{W}^\infty$ inherits an almost complex structure J^∞ that is defined to match J_V on both of the cylindrical ends corresponding to V and J everywhere else. In the language of [BEH+03],

$$(\widehat{W}^\infty, J^\infty)$$

is called a **split almost complex manifold** (with cylindrical ends). Note that we have made no assumptions in this discussion as to whether $V \subset W$ is connected or separates W, thus $\widehat{W}^\infty$ may have multiple connected components; in particular, if V is connected and separates W, then the negative cylindrical end $(-\infty, 0] \times V$ and positive cylindrical end $[0, \infty) \times V$ corresponding to the two copies of V in ∂W_V lie in distinct connected components of $\widehat{W}^\infty$.

More generally, suppose $J_\nu \to J$ is a C^∞-convergent sequence of almost complex structures in $\mathcal{J}_\tau(\omega, \mathcal{H}_+, \mathcal{H}_-)$ that restrict to the tubular neighborhood $\mathcal{N}(V) \subset W$ as translation-invariant structures belong to $\mathcal{J}(\mathcal{H})$. For any positive sequence $T_\nu \to \infty$, the sequence of "stretched" almost complex structures $J_\nu^{T_\nu}$ on $\widehat{W}^{T_\nu}$ "converges" to J^∞ on $\widehat{W}^\infty$, and we can then consider sequences of asymptotically cylindrical holomorphic curves

$$u_\nu : (\dot{\Sigma}_\nu, j_\nu) \to (\widehat{W}^{T_\nu}, J_\nu^{T_\nu}).$$

The cohomological relation (7.17) implies:

LEMMA 7.7.1. *The energy $E(u)$ of an asymptotically cylindrical holomorphic curve $u : (\dot{\Sigma}, j) \to (\widehat{W}^T, J^T)$ is independent of the choice of the function f_T used in defining ω^T.* □

In this setting, a new variation on the breaking phenomenon discussed in §7.3.2 emerges. Figure 7.10 shows a scenario involving a sequence of closed holomorphic tori $u_\nu : (\mathbb{T}^2, j_\nu) \to (W^{T_\nu}, J^{T_\nu})$ with uniformly bounded energy in a closed symplectic manifold (W, ω) that is being stretched along a separating stable hypersurface $V \subset W$. Here $\widehat{W} = W$ and $\widehat{W}^T = W^T$ since W has empty boundary, but cylindrical ends and asymptotically cylindrical punctured curves appear in the limit due to the fact that $T_\nu \to \infty$. Concretely, the cobordism W_V obtained by splitting W along

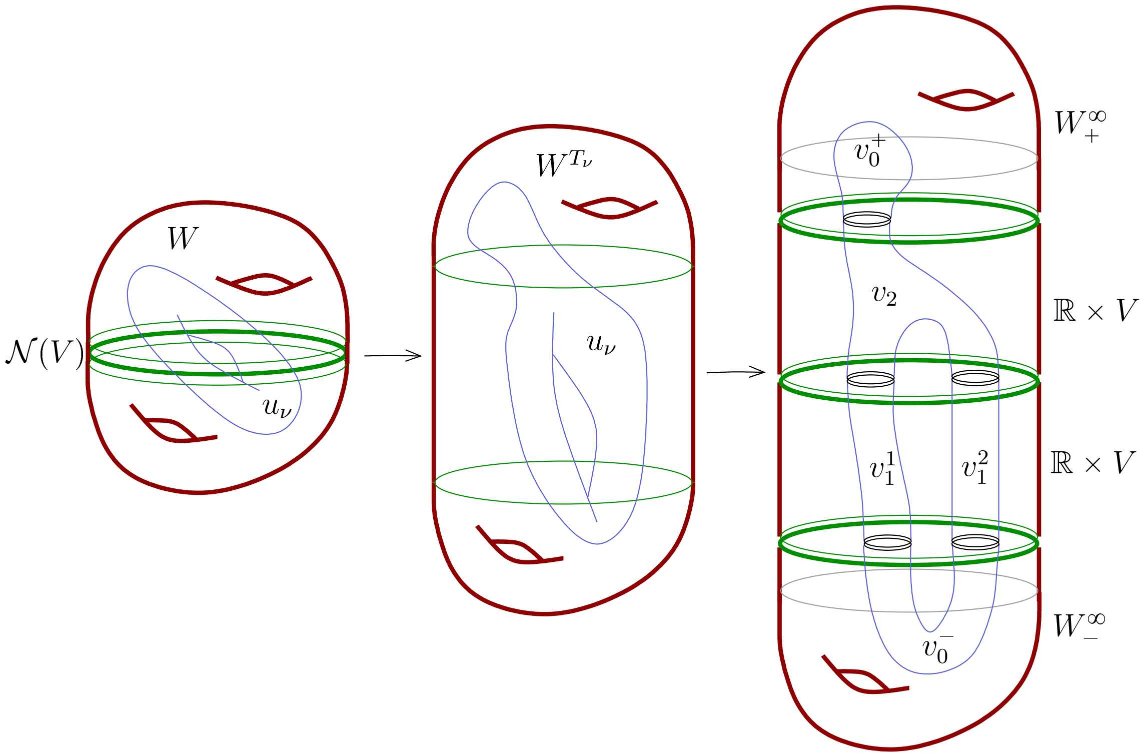

FIGURE 7.10. A sequence of holomorphic tori degenerating as a closed symplectic manifold is stretched along a separating stable hypersurface.

V has two connected components: one is a cobordism W_- from $\emptyset$ to V, the other is a cobordism W_+ from V to $\emptyset$, and the completions of these two cobordisms are labelled W_-^∞ and W_+^∞ respectively in the figure. The completions $W_\pm^\infty$ inherit natural tame almost complex structures $J_\pm^\infty$ that match J_V on the cylindrical ends and J everywhere else. In this thought-experiment, the domain $\mathbb{T}^2$ decomposes into regions

$$\mathbb{A}^- \cup \mathbb{A}_1^1 \cup \mathbb{A}_1^2 \cup \mathbb{P} \cup \mathbb{D} \subset \mathbb{T}^2$$

with disjoint interiors, on which the following behavior occurs. The region $(\mathbb{A}^-, j_\nu)$ is biholomorphically equivalent to $([-a_\nu, a_\nu] \times S^1, i)$ for some sequence $a_\nu \to \infty$, such that

$$u_\nu(\mathbb{A}^-) \subset W_- \cup ([-T_\nu, 0] \times V) \subset W^{T_\nu}$$

for every ν and, after identifying $[-T_\nu, 0] \times V$ with $[0, T_\nu] \times V \subset W_-^\infty$ via a shift in the r-coordinate, we see a sequence of J_-^∞-holomorphic maps

$$([-a_\nu, a_\nu] \times S^1, i) \to (W_-^\infty, J_-^\infty)$$

that converge in $C^\infty_{\mathrm{loc}}(\mathbb{R} \times S^1)$ to a holomorphic cylinder

$$v_0^- : (\mathbb{R} \times S^1, i) \to (W_-^\infty, J_-^\infty).$$

The regions $(\mathbb{A}_1^1, j_\nu)$ and $(\mathbb{A}_1^2, j_\nu)$ are also biholomorphically equivalent to $([-b_\nu, b_\nu] \times S^1, i)$ and $([-c_\nu, c_\nu] \times S^1, i)$ respectively for suitable sequences $b_\nu, c_\nu \to \infty$, but the restrictions of u_ν to these two domains have image in the neck $\mathcal{N}^{T_\nu} = (-T_\nu, T_\nu) \times V \subset W^{T_\nu}$ and converge in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ to a pair of J_V-holomorphic cylinders

$$v_1^1, v_1^2 : (\mathbb{R} \times S^1, i) \to (\mathbb{R} \times V, J_V).$$

Since the length of the neck is unbounded in the limit, this need not be the only interesting thing happening in the neck. The region $(\mathbb{P}, j_\nu)$ has the topology of a compact pair-of-pants, and corresponds to the curve v_2 in the picture: it is biholomorphically identified with the complement of shrinking sequences of disjoint disks around the three points $0, 1, \infty$ in S^2, on which u_ν also has image in $\mathcal{N}^{T_\nu} \subset W^{T_\nu}$, but higher up than that of $\mathbb{A}^1_{k,1}$ and $\mathbb{A}^2_{k,1}$, so that after a diverging sequence of $\mathbb{R}$-translations, these restrictions converge in C^∞_{loc} to a another curve

$$v_2 : (S^2 \backslash \{0, 1, \infty\}, i) \to (\mathbb{R} \times V, J_V).$$

Finally, the region $(\mathbb{D}, j_\nu)$ is biholomorphically identified with an expanding sequence of disks in $\mathbb{C}$, such that

$$u_\nu(\mathbb{D}) \subset W_+ \cup ([0, T_\nu] \times V) \subset W^{T_\nu}$$

and, after identifying $[0, T_\nu] \times V$ with $[-T_\nu, 0] \times V \subset W_+^\infty$ by shifting the r-coordinate, this sequence converges in $C^\infty_{\text{loc}}(\mathbb{C})$ to a J_+^∞-holomorphic plane

$$v_0^+ : (\mathbb{C}, i) \to (W_+^\infty, J_+^\infty).$$

The freedom in the choice of the functions $f_{T_\nu} : (-T_\nu, T_\nu) \to (-\epsilon, \epsilon)$ arising from Lemma 7.7.1 can be used to show that all of these limiting curves have finite energy, and are thus asymptotically cylindrical.

While the limit in Figure 7.10 appears to have four levels, W_-^∞ and W_+^∞ should really be viewed as two connected components of the *same* level, since both arise by attaching cylindrical ends to the cobordism W_V obtained by splitting W open along V. If V, for instance, had been a *nonseparating* hypersurface, then $W \backslash V$ would have only one component, so instead of W_+ and W_-, one would have a single connected symplectic cobordism from one copy of V to a second copy. For this reason, we label the components in $W_-^\infty \coprod W_+^\infty$ together as the **main level** of the building in Figure 7.10. The levels in copies of the symplectization of $(V, \mathcal{H})$ are called **insert levels**, or collectively, the **insert layer** (as distinguished from the **main layer**, which consists only of the main level). This terminology calls attention to a slightly misleading feature of Figure 7.10: Contrary to appearances, it is not generally sensible to give a linear ordering to the levels in a building that arises from stretching.

This fact is perhaps more obvious from Figure 7.11. In this scenario, W is a cobordism with both positive and negative stable boundary components M_+ and M_- respectively, while the interior stable hypersurface $V = V_1 \coprod V_2$ has two connected components, one that separates W and another that does not. (The components of the collar neighborhood $\mathcal{N}(V)$ have been labelled accordingly in the picture.) The result is that splitting W along V produces a disjoint symplectic cobordism in which one component W_1 is a cobordism from $M_- \coprod V_1$ to $V_1 \coprod V_2$, and the other component W_2 is a cobordism from V_2 to M_+. (The completions of these cobordisms are labelled $\widehat{W}_1^\infty$ and $\widehat{W}_2^\infty$ in the picture.) The holomorphic building we see in this scenario includes three upper levels (in the symplectization of M_+) and one lower level (in the symplectization of M_-), which arise due to the same phenomena that were described in §7.3, but variations on the phenomena behind Figure 7.10 also produce two insert levels living in the symplectization of V, each of which has two connected components $\mathbb{R} \times V_1$ and $\mathbb{R} \times V_2$, thus they appear in the picture as two separate pairs of levels. The main level looks like two levels in this picture due to the fact that $W \backslash V$ again has two components. While a linear ordering of the levels in Figure 7.11 is clearly impossible, one can sensibly apply a *cyclic* ordering to the union of the main level with the insert levels, i.e. so that the main layer is level $0 \in \mathbb{Z}_3$ and the insert layer consists of levels $1, 2 \in \mathbb{Z}_3$. For the upper and lower levels, it remains appropriate to assign an integer label to each, thus one can reasonably define a locally constant level function on the domain of the entire building with values in the group $\mathbb{Z} \oplus \mathbb{Z}_3$. This convention is implemented in the following definition.

DEFINITION 7.7.2. Assume $(\widehat{W}^\infty, J^\infty)$ is a split almost complex manifold with cylindrical ends, obtained as described above by splitting a symplectic cobordism (W, ω) with stable boundary components $(M_+, \mathcal{H}_+)$ and $(M_-, \mathcal{H}_-)$ along an interior stable hypersurface $(V, \mathcal{H})$. Given integers $g, m, N_+, N_-, N_0 \geqslant 0$, a **holomorphic building of height** $N_- | \overset{N_0}{\underset{1}{\vee}} | N_+$ **in** $(\widehat{W}^\infty, J^\infty)$ is a tuple[3]

$$\mathbf{u} = (S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\text{nd}}, \Delta^{\text{br}}, L, \Phi, u)$$

as described in Definition 7.4.2 with the following modifications:

- The **level structure** is a locally constant function
$$L : S \to \mathbb{Z} \oplus \mathbb{Z}_{N_0+1}$$
whose values are all of the form $(N, 0)$ for $-N_- \leqslant N \leqslant N_+$ or $(0, N)$ for $N \in \mathbb{Z}_{N_0+1}$, where all of these values are attained except possibly for $(0, 0)$, and moreover:
  (1) Each breaking pair $\{z^+, z^-\} \subset \Delta^{\text{br}}$ can be labelled such that $L(z^+) - L(z^-)$ is either $(1, 0)$ or $(0, 1)$;
  (2) $L(\Gamma^+) = \{(N_+, 0)\}$ and $L(\Gamma^-) = \{(-N_-, 0)\}$.
- The asymptotically cylindrical holomorphic **map** u sends components of $(\dot{S} := S \backslash (\Gamma \cup \Delta^{\text{br}}), j)$ with $L = (\pm N, 0)$ for $N > 0$ to $(\mathbb{R} \times M_\pm, J_\pm)$, components with $L = (0, 0)$ to $(\widehat{W}^\infty, J^\infty)$, and components with $L = (0, N)$ for $N \in \mathbb{Z}_{N_0+1} \backslash \{0\}$ to $(\mathbb{R} \times V, J_V)$. The same conditions as in Definition 7.4.2 apply for breaking pairs $\{z^+, z^-\} \in \Delta^{\text{br}}$ with $L(z^+) - L(z^-) = (1, 0)$ or $L(z^+) - L(z^-) = (0, 1)$ with the exception that in the latter case, the common asymptotic orbit of z^+ and z^- lies in V instead of M_+ or M_-.

The components of $\mathbf{u}$ with $L = (0, 0)$ form the **main layer** of the building, while components with $L = (N, 0)$ form the **upper layer** for $N > 0$ and **lower layer** for $N < 0$, and components with $L = (0, N)$ for $N \in \mathbb{Z}_{N_0+1} \backslash \{0\}$ form the **insert layer**.

[3]The case $\partial W = \varnothing$ is most commonly encountered in applications, so that there are no upper or lower levels, but only a main level and insert levels. Following [BEH+03], a building with N_0 insert levels in this setting is called a **building of height** $\overset{N_0}{\underset{1}{\vee}}$.

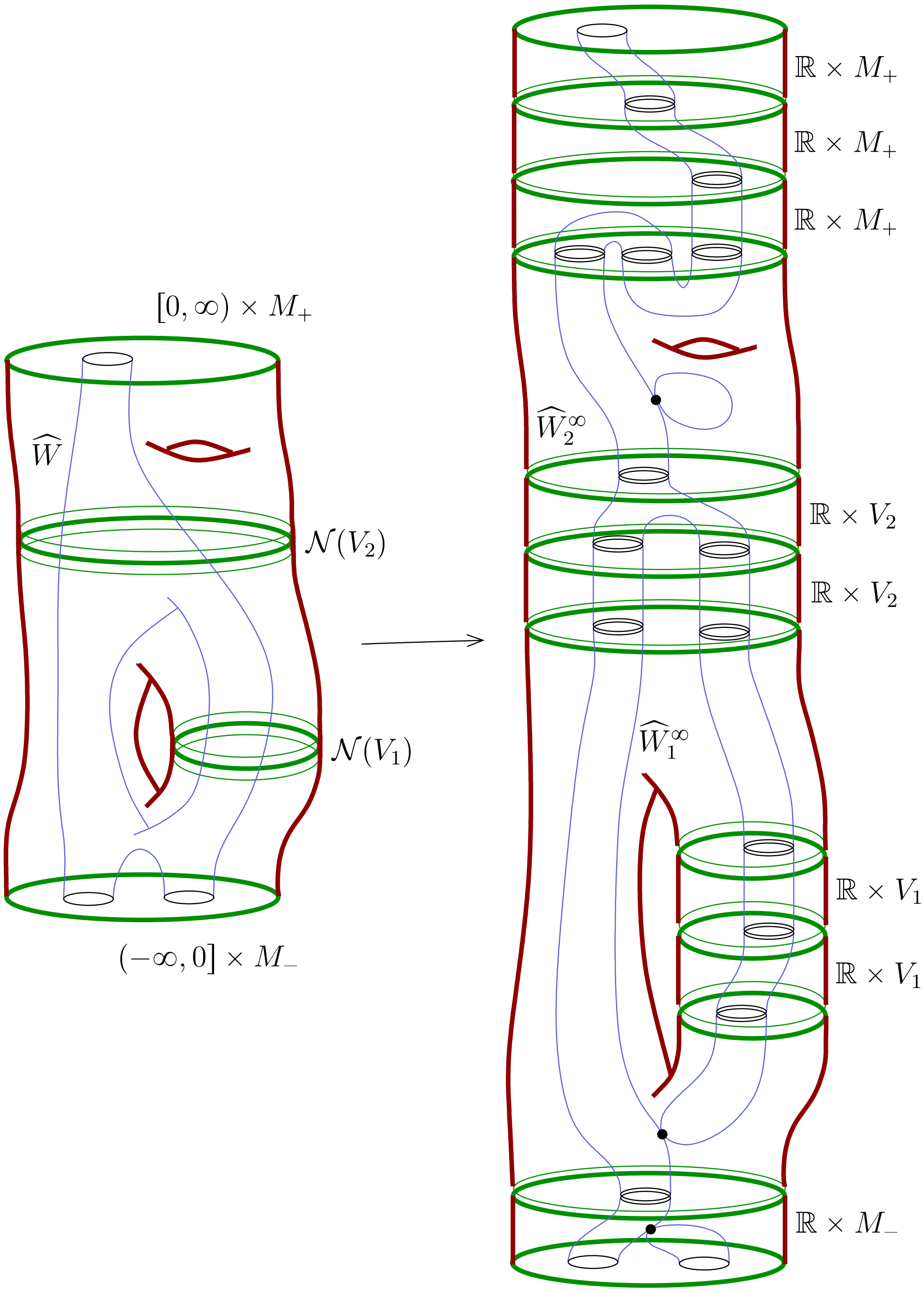

FIGURE 7.11. Degeneration to a holomorphic building in the same split symplectic cobordism as in Figure 7.9, where W is split along a disconnected stable hypersurface $V = V_1 \coprod V_2$. In this scenario, the limiting building includes three upper levels, one lower level and two insert levels.

DEFINITION 7.7.3. The building $\mathbf{u} = (S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\mathrm{nd}}, \Delta^{\mathrm{br}}, L, \Phi, u)$ in Definition 7.7.2 is called **stable** if:

(1) Every connected component of $S \backslash (\Gamma \cup \Theta \cup \Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}})$ on which the map u is constant has negative Euler characteristic;
(2) There is no upper, lower or insert level consisting only of a disjoint union of trivial cylinders without any marked points or nodes.

The notion of **equivalence** between two buildings in this setting is a straightforward generalization of the definition for buildings of height $N_-|1|N_+$, as described in §7.4. Equivalences must of course preserve the level structure, so in particular, they identify insert levels with insert levels up to $\mathbb{R}$-translation.

To define **convergence** of a sequence of smooth curves $[(\Sigma_\nu, j_\nu, \Gamma_\nu^+, \Gamma_\nu^-, \Theta_\nu, u_\nu)] \in \mathcal{M}_{g,m}(J_\nu^{T_\nu}, A_\nu, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ to a building $[(S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\mathrm{nd}}, \Delta^{\mathrm{br}}, L, \Phi, u)]$, one can similarly generalize Definition 7.5.3 and represent the equivalence classes of the sequence as $J_\nu^{T_\nu}$-holomorphic maps defined on surfaces S^{τ_ν} constructed from a pregluing of $(S, j, \Gamma \cup \Theta, \Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}})$. The novelty in this situation is that both the domains S^{τ_ν} and the targets $\widehat{W}^{T_\nu}$ contain necks with lengths diverging to ∞, but the important phenomena can still be characterized in terms of C^∞_{loc}-convergence up to $\mathbb{R}$-translation. The main result is then:

THEOREM 7.7.4. *Suppose all closed Reeb orbits in $(M_\pm, \mathcal{H}_\pm)$ and $(V, \mathcal{H})$ are nondegenerate or Morse-Bott, and $J_\nu \to J$ is a C^∞-convergent sequence of almost complex structures in $\mathcal{J}_\tau(\widehat{W}, \mathcal{H}_+, \mathcal{H}_-)$ that restrict to the tubular neighborhood $\mathcal{N}(V) \subset \widehat{W}$ as translation-invariant almost complex structures belonging to $\mathcal{J}(\mathcal{H})$. Then every sequence of stable $J_\nu^{T_\nu}$-holomorphic curves in $\widehat{W}^{T_\nu}$ with uniformly bounded genus and energy and fixed asymptotic orbits has a subsequence convergent to a stable holomorphic building of height $N_-|\overset{N_0}{\underset{1}{\vee}}|N_+$ for some $N_+, N_0, N_- \geqslant 0$.*

EXAMPLE 7.7.5. The most popular setting for applications of Theorem 7.7.4 arises from Lagrangian submanifolds $L \subset W$. By the Weinstein neighborhood theorem, L always has a neighborhood W_- symplectomorphic to a neighborhood of the zero-section in T^*L, so $M := \partial W_-$ is a contact-type hypersurface contactomorphic to the unit cotangent bundle of L. Stretching the neck then yields T^*L as the completion of W_-, and $W \backslash L$ as the completion of $W_+ := W \backslash \mathring{W}_-$. This construction has often been used in order to study Lagrangian submanifolds via SFT-type methods, see e.g. [EGH00, Theorem 1.7.5] and [Eva10, CM18].

CHAPTER 8

Smoothness of the moduli space

Contents

8.1. The main result on regular curves

8.2. Functional-analytic setup

8.3. Moduli of complex structures

8.3.1. Teichmüller space and automorphism groups.

PROPOSITION 8.3.1. *to be written*

8.3.2. Teichmüller slices.

8.3.3. Adding marked points.

8.4. Fredholm regularity and the implicit function theorem

8.5. Evaluation and forgetful maps

CHAPTER 9

Transversality

Contents

9.1. A paradigm for genericity arguments

LEMMA 9.1.1. *to be written*

9.2. Generic transversality in cobordisms

9.2.1. A theorem for somewhere injective curves.

9.2.2. The universal moduli space.

9.2.3. Applying the Sard-Smale theorem.

9.2.4. From C_ϵ to C^∞.

9.3. Generic transversality in symplectizations

9.3.1. Main results in the $\mathbb{R}$-invariant setting.

9.3.2. Injective points of the projected curve.

9.3.3. Smoothness of the universal moduli space.

9.4. Transversality of constraints

9.4.1. The evaluation and forgetful maps.

9.4.2. Constraints on derivatives.

CHAPTER 10

Gluing

Contents

10.1. Moduli spaces of holomorphic buildings

10.1.1. Notation and conventions.

10.1.2. Automorphism groups.

PROPOSITION 10.1.1. *write this*

PROOF. write this □

10.1.3. The deformation operator.

10.2. Deligne-Mumford space is an orbifold

CHAPTER 11

Cylindrical contact homology and the tight 3-tori

Contents

11.1. Contact structures on $\mathbb{T}^3$ and Giroux torsion

11.2. Definition of cylindrical contact homology

11.2.1. Preliminary remarks.

11.2.2. A compactness result for cylinders.

11.2.3. The chain complex.

11.2.4. The homology.

11.2.5. Chain maps.

11.2.6. Chain homotopies.

11.2.7. Proof of invariance.

11.3. Computing $HC_*(\mathbb{T}^3, \xi_k)$

11.3.1. The Morse-Bott setup.

11.3.2. A digression on the Floer equation.

THEOREM 11.3.1. *to be written*

REMARK 11.3.2 (sign conventions). to be written

11.3.3. Admissible data for $(\mathbb{T}^3, \xi_k)$.

CHAPTER 12

Coherent orientations

Contents

12.1. Gluing maps and coherence

12.2. Permutations of punctures and bad orbits

12.3. Orienting moduli spaces in general

12.4. The determinant line bundle

12.5. Determinant bundles of moduli spaces

12.6. An algorithm for coherent orientations

12.7. Permutations and bad orbits revisited

CHAPTER 13

The generating function of SFT

Contents

13.1. Some important caveats on transversality

13.2. Auxiliary data, grading and supercommutativity

13.3. The definition of H and commutators

13.4. Interlude: Orbifolds and branched manifolds

13.4.1. How to count zeroes in an orbifold.

13.4.2. Multivalued perturbations.

13.4.3. The inhomogeneous Cauchy-Riemann equation.

13.5. Cylindrical contact homology revisited

13.6. Combinatorics of gluing

13.7. Some remarks on torsion, coefficients, and conventions

13.7.1. What if $H_1(M)$ has torsion?

13.7.2. Combinatorial conventions.

13.7.3. Coefficients: $\mathbb{Q}$, $\mathbb{Z}$ or $\mathbb{Z}_2$?

CHAPTER 14

Contact invariants

Contents

14.1. The Eliashberg-Givental-Hofer package

14.1.1. Full SFT as a Weyl superalgebra.

14.1.2. The semiclassical limit: rational SFT.

14.1.3. The contact homology algebra.

14.1.4. Algebraic overtwistedness.

14.2. SFT generating functions for cobordisms

14.2.1. Weak, strong and stable cobordisms.

14.2.2. Counting disconnected index 0 curves.

14.3. Full SFT as a BV_∞-algebra

14.3.1. Cobordism maps and invariance.

14.3.2. Algebraic torsion.

CHAPTER 15

Transversality and counting singularities in dimension four

Contents

15.1. Automatic transversality

15.1.1. The statement.

THEOREM 15.1.1. *to be written*

15.1.2. The normal Cauchy-Riemann operator.

15.1.3. Counting zeroes on line bundles with punctures.

15.1.4. Proof of Theorem 15.1.1 and generalizations.

15.1.5. An alternative proof for $g = \#\Gamma_0 = 0$.

15.2. Curves in symplectizations of 3-manifolds

15.3. Implict function theorems for local foliations

15.3.1. Immersed curves with no automorphisms.

15.3.2. Embedded curves asymptotic to distinct simple orbits.

15.3.3. Immersed multiply covered curves.

PROPOSITION 15.3.1. *to be written*

15.4. Consequences for coherent orientations

CHAPTER 16

Intersection theory for punctured holomorphic curves

Contents

16.1. Prologue

16.2. Homotopy-invariant intersection numbers

LEMMA 16.2.1. *to be written*

16.3. The adjunction formula

16.4. Local foliations: the general case

CHAPTER 17

Torsion computations and applications

Contents

17.1. Some J-holomorphic foliations

17.1.1. Gradient flow lines as holomorphic cylinders.

17.1.2. Domains with $\mathbb{T}^2$-symmetry.

17.1.3. Open books with holomorphic pages.

17.2. Contact homology of overtwisted contact manifolds

17.3. Examples with higher-order algebraic torsion

17.4. Rigorous obstructions to fillings and cobordisms

APPENDIX A

Sobolev spaces

Contents

In this appendix, we review some of the standard properties of Sobolev spaces, in particular using them to prove Propositions 2.2.4, 2.2.5 and 2.2.8 from §2.2, and elucidating the construction of Sobolev spaces of sections on vector bundles. A good reference for the necessary background material is [**AF03**].

A.1. Approximation, extension and embedding theorems

Unless otherwise noted, all functions in the following are assumed to be defined on a nonempty open subset

$$\mathcal{U} \subset \mathbb{R}^n$$

with its standard Lebesgue measure, and taking values in a finite-dimensional normed vector space that will usually not need to be specified, though occasionally we will assume it is $\mathbb{R}$ or $\mathbb{C}$ so that one can define products of functions. The domain $\mathcal{U}$ will also sometimes have additional conditions specified such as boundedness or regularity at the boundary, though we will try not to add too many more restrictions than are really needed. The most useful assumption to impose on $\mathcal{U}$ is known as the **strong local Lipschitz condition**: if $\mathcal{U}$ is bounded, then it means simply that near every boundary point of $\mathcal{U}$, one can find smooth local coordinates in which $\mathcal{U}$ looks like the region bounded by the graph of a Lipschitz-continuous function, and in this case we call $\mathcal{U}$ a **bounded Lipschitz domain**. If $\mathcal{U}$ is unbounded, then one needs to impose extra conditions guaranteeing e.g. uniformity of Lipschitz constants, and the precise definition becomes a bit lengthy (see [**AF03**, §4.9]). For our purposes, all we really need to know about the strong local Lipschitz condition is that that it is satisfied both by bounded Lipschitz domains and by relatively tame unbounded domains such as $(0, 1) \times (0, \infty) \subset \mathbb{R}^2$ which have smooth boundary with finitely many corners. We will repeatedly need to use the generalized version of **Hölder's inequality**, which states that for any finite collection of measurable functions $f_1, \dots, f_m$,

$$\left\| \prod_{i=1}^m |f_i| \right\|_{L^p} \leqslant \prod_{i=1}^m \|f_i\|_{L^{p_i}} \quad \text{for } 1 \leqslant p \leqslant p_1, \dots, p_m \leqslant \infty \text{ with } \frac{1}{p} = \sum_{i=1}^m \frac{1}{p_i}. \tag{A.1}$$

This is an easy corollary of the standard version,

$$\big\| |f| \cdot |g| \big\|_{L^1} \leqslant \|f\|_{L^p} \cdot \|g\|_{L^q} \quad \text{whenever } 1 \leqslant p, q \leqslant \infty \text{ and } 1 = \frac{1}{p} + \frac{1}{q}.$$

For an integer $k \geqslant 0$ and real number $p \in [1, \infty]$, we define $W^{k,p}(\mathcal{U})$ as in §2.2 to be the Banach space of all $f \in L^p(\mathcal{U})$ which have weak partial derivatives $\partial^\alpha f \in L^p(\mathcal{U})$ for all $|\alpha| \leqslant k$. For $p = 2$, these spaces are also often denoted by

$$H^k(\mathcal{U}) := W^{k,2}(\mathcal{U}),$$

and they admit Hilbert space structures with inner product

$$\langle f, g \rangle_{H^k} = \sum_{|\alpha| \leqslant k} \langle \partial^\alpha f, \partial^\alpha g \rangle_{L^2}.$$

We denote by

$$W_0^{k,p}(\mathcal{U}) \subset W^{k,p}(\mathcal{U}), \qquad H_0^k(\mathcal{U}) \subset H^k(\mathcal{U})$$

the closed subspaces defined as the closures of $C_0^\infty(\mathcal{U})$ with respect to the relevant norms. Since $C_0^\infty(\mathcal{U})$ is dense in $L^p(\mathcal{U})$ for $1 \leqslant p < \infty$ (see e.g. [**LL01**, §2.19]), there is no difference between $W^{0,p}(\mathcal{U})$ and $W_0^{0,p}(\mathcal{U})$ for $p < \infty$, but in general $W_0^{k,p}(\mathcal{U}) \neq W^{k,p}(\mathcal{U})$ for $k \geqslant 1$, with a few notable exceptions such as the case $\mathcal{U} = \mathbb{R}^n$ (cf. Corollary A.1.2 below). Let

$$W_{\text{loc}}^{k,p}(\mathcal{U}) := \big\{ \text{functions } f \text{ on } \mathcal{U} \;\big|\; f \in W^{k,p}(\mathcal{V}) \text{ for all open subsets } \mathcal{V} \subset \mathcal{U} \text{ with compact closure } \overline{\mathcal{V}} \subset \mathcal{U} \big\},$$

and say that a sequence $f_j \in W_{\text{loc}}^{k,p}(\mathcal{U})$ converges in $W_{\text{loc}}^{k,p}$ to $f \in W_{\text{loc}}^{k,p}(\mathcal{U})$ if the restrictions to all precompact open subsets $\mathcal{V} \subset \overline{\mathcal{V}} \subset \mathcal{U}$ converge in $W^{k,p}(\mathcal{V})$. Recall that for $k \in \{0, 1, 2, \dots, \infty\}$, $C^k(\mathcal{U})$ denotes the space of functions on $\mathcal{U}$ with continuous derivatives up to order k, while

$$C^k(\overline{\mathcal{U}}) \subset C^k(\mathcal{U})$$

is the space of $f \in C^k(\mathcal{U})$ such that for all $|\alpha| \leqslant k$, $\partial^\alpha f$ is bounded and uniformly continuous.

THEOREM A.1.1 ([**AF03**, §3.17, 3.22]). *For any open subset $\mathcal{U} \subset \mathbb{R}^n$, and any $k \geqslant 0$, $1 \leqslant p < \infty$, the subspace*

$$C^\infty(\mathcal{U}) \cap W^{k,p}(\mathcal{U}) \subset W^{k,p}(\mathcal{U})$$

is dense. Moreover, if $\mathcal{U} \subset \mathbb{R}^n$ satisfies the strong local Lipschitz condition, then the space

$$\left\{ f \in C^\infty(\mathcal{U}) \;\middle|\; f = \tilde{f}|_{\mathcal{U}} \text{ for some } \tilde{f} \in C_0^\infty(\mathbb{R}^n) \right\}$$

is also dense in $W^{k,p}(\mathcal{U})$, so in particular,

$$C^\infty(\overline{\mathcal{U}}) \cap W^{k,p}(\mathcal{U}) \subset W^{k,p}(\mathcal{U})$$

is dense. □

COROLLARY A.1.2. *The space $C_0^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$ for every $k \geqslant 0$ and $p \in [1, \infty)$.* □

Here is another useful characterization of $W_0^{k,p}(\mathcal{U})$:

THEOREM A.1.3 ([**AF03**, §5.29]). *Assume $\mathcal{U} \subset \mathbb{R}^n$ is an open subset satisfying the strong local Lipschitz condition. Then a function $f \in W^{k,p}(\mathcal{U})$ belongs to $W_0^{k,p}(\mathcal{U})$ if and only if the function $\widetilde{f}$ on $\mathbb{R}^n$ defined to match f on $\mathcal{U}$ and 0 everywhere else belongs to $W^{k,p}(\mathbb{R}^n)$.* □

While it is obvious from the definitions that functions in $W_0^{k,p}(\mathcal{U})$ always admit extensions of class $W^{k,p}$ over $\mathbb{R}^n$, this is much less obvious for functions in $W^{k,p}(\mathcal{U})$ in general, and it is not true without sufficient assumptions about the regularity of $\partial\mathcal{U}$. For our purposes it suffices to consider the following case.

THEOREM A.1.4 ([**AF03**, §5.22]). *Assume $\mathcal{U} \subset \mathbb{R}^n$ is a bounded open subset such that $\partial\overline{\mathcal{U}}$ is a submanifold of class C^m for some $m \in \{1, 2, 3, \ldots, \infty\}$. Then there exists a linear operator E that maps functions defined almost everywhere on $\mathcal{U}$ to functions defined almost everywhere on $\mathbb{R}^n$ and has the following properties:*

- *For every function f on $\mathcal{U}$, $Ef|_{\mathcal{U}} \equiv f$ almost everywhere;*
- *For every nonnegative integer $k \leqslant m$ and every $p \in [1, \infty)$, E defines a bounded linear operator $W^{k,p}(\mathcal{U}) \to W^{k,p}(\mathbb{R}^n)$.*

□

COROLLARY A.1.5. *Suppose $\mathcal{U}, \mathcal{U}' \subset \mathbb{R}^n$ are open subsets such that $\mathcal{U}$ has compact closure contained in $\mathcal{U}'$. If $\mathcal{U}$ satisfies the hypothesis of Theorem A.1.4, then the resulting extension operator E can be chosen such that it maps each $W^{k,p}(\mathcal{U})$ for $k \leqslant m$ and $1 \leqslant p < \infty$ into $W_0^{k,p}(\mathcal{U}')$.*

PROOF. Choose a smooth function $\rho : \mathcal{U}' \to [0, 1]$ that has compact support and equals 1 on $\overline{\mathcal{U}}$, then replace the operator E given by Theorem A.1.4 with the operator $f \mapsto \rho \cdot Ef$. □

To state the Sobolev embedding theorem in its proper generality, recall that for $0 < \alpha \leqslant 1$, the **Hölder seminorm** of a function f on $\mathcal{U}$ is defined by

$$|f|_{C^\alpha} := |f|_{C^\alpha(\mathcal{U})} := \sup_{x \neq y \in \mathcal{U}} \frac{|f(x) - f(y)|}{|x - y|^\alpha},$$

and $C^{k,\alpha}(\mathcal{U})$ is then defined as the Banach space of functions $f \in C^k(\overline{\mathcal{U}})$ for which the norm

$$\|f\|_{C^{k,\alpha}} := \|f\|_{C^k} + \max_{|\beta|=k} |\partial^\beta f|_{C^\alpha}$$

is finite. In reading the following statement, it is important to remember that elements of $W^{k,p}(\mathcal{U})$ are technically not functions, but rather *equivalence classes* of functions defined almost everywhere. Thus when we say e.g. that there is an inclusion $W^{k,p}(\mathcal{U}) \hookrightarrow C^{m,\alpha}(\mathcal{U})$, the literal meaning is that for every function f representing an element of $W^{k,p}(\mathcal{U})$, one can change the values of f in a unique way on some set of measure zero in $\mathcal{U}$ so that after this change, $f \in C^{m,\alpha}(\mathcal{U})$. Continuity of the inclusion means that there is a bound of the form

$$\|f\|_{C^{m,\alpha}} \leqslant c\|f\|_{W^{k,p}}$$

for all $f \in W^{k,p}(\mathcal{U})$, where $c > 0$ is a constant which may in general depend on m, α, k, p and $\mathcal{U}$, but not on f.

THEOREM A.1.6 ([AF03, §4.12]). *Assume $\mathcal{U} \subset \mathbb{R}^n$ is an open subset satisfying the strong local Lipschitz condition, $k \geqslant 1$ is an integer and $1 \leqslant p < \infty$.*

(1) If $0 < k - n/p \leqslant 1$, then there exist continuous inclusions

$$W^{k,p}(\mathcal{U}) \hookrightarrow C^{0,\alpha}(\mathcal{U}) \qquad \text{for each } \alpha \in (0,1) \text{ with } \alpha \leqslant k - n/p,$$
$$W^{k,p}(\mathcal{U}) \hookrightarrow L^{q}(\mathcal{U}) \qquad \text{for each } q \in [p,\infty].$$

(2) If $kp < n$ and $p^* > p$ is defined by the condition

$$\frac{1}{p^*} = \frac{1}{p} - \frac{k}{n},$$

then there exist continuous inclusions

$$W^{k,p}(\mathcal{U}) \hookrightarrow L^{q}(\mathcal{U}), \qquad \text{for each } q \in [p,p^*].$$

(3) If $kp = n$, then there exist continuous inclusions

$$W^{k,p}(\mathcal{U}) \hookrightarrow L^{q}(\mathcal{U}), \qquad \text{for each } q \in [p,\infty).$$

Moreover, the spaces $W^{k,p}_0(\mathcal{U})$ admit similar inclusions under no assumption on the open subset $\mathcal{U} \subset \mathbb{R}^n$. □

Under the same assumption on the domain $\mathcal{U}$, one can apply Theorem A.1.6 to successive derivatives of functions in $W^{k,p}(\mathcal{U})$ and thus obtain the following inclusions for any integer $d \geqslant 0$:

$$W^{k+d,p}(\mathcal{U}) \hookrightarrow C^{d,\alpha}(\mathcal{U}) \quad \text{if } 0 < k - n/p \leqslant 1,\ 0 < \alpha < 1 \text{ and } \alpha \leqslant k - n/p, \tag{A.2}$$

$$W^{k+d,p}(\mathcal{U}) \hookrightarrow W^{d,q}(\mathcal{U}) \quad \text{if } kp > n \text{ and } p \leqslant q \leqslant \infty, \tag{A.3}$$

$$W^{k+d,p}(\mathcal{U}) \hookrightarrow W^{d,q}(\mathcal{U}) \quad \text{if } kp < n \text{ and } p \leqslant q \leqslant p^*, \text{ with } \frac{1}{p^*} = \frac{1}{p} - \frac{k}{n}, \tag{A.4}$$

$$W^{k+d,p}(\mathcal{U}) \hookrightarrow W^{d,q}(\mathcal{U}) \quad \text{if } kp = n \text{ and } p \leqslant q < \infty. \tag{A.5}$$

REMARK A.1.7. The embedding theorem suggests that one should intuitively think of $W^{k,p}(\mathcal{U})$ as consisting of functions with "$k - n/p$ continuous derivatives," where the number $k - n/p$ may in general be a non-integer and/or negative. This provides a useful mnemonic for results about embeddings of one Sobolev space into another, such as the following.

COROLLARY A.1.8. *Assume $\mathcal{U} \subset \mathbb{R}^n$ is an open subset satisfying the strong local Lipschitz condition, $1 \leqslant p, q < \infty$, and $k, m \geqslant 0$ are integers satisfying*

$$k \geqslant m, \qquad p \leqslant q, \qquad \text{and} \qquad k - \frac{n}{p} \geqslant m - \frac{n}{q}.$$

Then there exists a continuous inclusion $W^{k,p}(\mathcal{U}) \hookrightarrow W^{m,q}(\mathcal{U})$. □

EXERCISE A.1.9. Derive Corollary A.1.8 from Theorem A.1.6 by checking that under the stated conditions, there is a continuous inclusion $W^{k-m,p}(\mathcal{U}) \hookrightarrow L^q(\mathcal{U})$. Show also that the hypothesis $p \leqslant q$ is unnecessary if $\mathcal{U} \subset \mathbb{R}^n$ has finite measure.

By the Arzelà-Ascoli theorem, the natural inclusion

$$C^{k,\alpha'}(\mathcal{U}) \hookrightarrow C^{k,\alpha}(\mathcal{U})$$

for $\alpha < \alpha'$ is a compact operator whenever $\mathcal{U} \subset \mathbb{R}^n$ is bounded. It follows that if $\mathcal{U} \subset \mathbb{R}^n$ in (A.2) is bounded and α is *strictly* less than the extremal value $k - n/p$, then the inclusion (A.2) is also compact. A similar statement holds for the inclusion (A.4) when $p \leqslant q < p^*$, and this is known as the **Rellich-Kondrachov compactness theorem**. We summarize these as follows:

THEOREM A.1.10 ([**AF03**, §6.3]). *Assume $\mathcal{U} \subset \mathbb{R}^n$ is a bounded Lipschitz domain, $k \geqslant 1$ and $d \geqslant 0$ are integers and $1 \leqslant p < \infty$.*

(1) If $kp > n$ and $k - n/p < 1$, then the inclusions

$$W^{k+d,p}(\mathcal{U}) \hookrightarrow C^{d,\alpha}(\mathcal{U}) \qquad \text{for } \alpha \in (0, k-n/p),$$
$$W^{k+d,p}(\mathcal{U}) \hookrightarrow W^{d,q}(\mathcal{U}) \qquad \text{for } q \in [p,\infty)$$

are compact.

(2) If $kp \leqslant n$ and $p^* \in (p,\infty]$ is defined by the condition $1/p^* = 1/p - k/n$, then the inclusions

$$W^{k+d,p}(\mathcal{U}) \hookrightarrow W^{d,q}(\mathcal{U}) \qquad \text{for } q \in [p,p^*)$$

are compact.

In particular, the continuous inclusion $W^{k,p}(\mathcal{U}) \hookrightarrow W^{m,q}(\mathcal{U})$ in Corollary A.1.8 is compact whenever the inequality $k - n/p \geqslant m - n/q$ is strict. □

On connected 1-dimensional domains $\mathcal{U} \subset \mathbb{R}$, the spaces $W^{1,p}(\mathcal{U})$ admit an alternative characterization in terms of classical derivatives defined almost everywhere:

PROPOSITION A.1.11. *For $-\infty < a < b < \infty$, every absolutely continuous function on $[a,b]$ belongs to $W^{1,1}((a,b))$ and has a weak derivative that is equal to its classical derivative almost everywhere. Conversely, every function in $W^{1,1}((a,b))$ is equal almost everywhere to an absolutely continuous function defined on $[a,b]$.*

PROOF. Let us denote the classical derivative of a function f by f'_c and the weak derivative by f'_w whenever there is danger of confusion. If f is absolutely continuous on $[a,b]$, then for every test function $\varphi \in C_0^\infty((a,b))$, $f\varphi$ defines an absolutely continuous function on $[a,b]$ that vanishes at the end points, so the fundamental theorem of calculus implies $\int_{[a,b]} (f\varphi)'_c = \int_{[a,b]} f'_c \varphi + \int_{[a,b]} f\varphi' = 0$, proving that the almost everywhere defined function $f'_c \in L^1([a,b])$ is also the weak derivative f'_w, and thus $f \in W^{1,1}((a,b))$.

Conversely, suppose $f \in W^{1,1}((a,b))$, so it has a weak derivative $f'_w \in L^1((a,b))$. We can then define an absolutely continuous function g on $[a,b]$ by $g(x) := \int_a^x f'_w$, which is differentiable almost everywhere and satisfies $g'_c = f'_w$. By the argument of the previous paragraph, g'_c is also a weak derivative g'_w, thus $g - f$ is a function on

(a,b) with vanishing weak derivative, implying via [LL01, Theorem 6.11] that $g-f$ is equal almost everywhere to a constant. ☐

Corollary A.1.12. *For $-\infty < a < b < \infty$ and $1 \leqslant p \leqslant \infty$, $W^{1,p}((a,b))$ has a canonical identification with the space of absolutely continuous functions on $[a,b]$ whose classical derivatives belong to $L^p([a,b])$.* ☐

A.2. Products, compositions, and rescaling

We now restate and prove Propositions 2.2.4, 2.2.5 and 2.2.8 from §2.2. These are all corollaries of the Sobolev embedding theorem, so in particular they hold for the same class of domains $\mathcal{U} \subset \mathbb{R}^n$, and the restrictions on $\mathcal{U}$ can be dropped at the cost of replacing each space $W^{k,p}$ by $W^{k,p}_0$.

We begin by generalizing Prop. 2.2.4, hence we consider Sobolev spaces of functions valued in $\mathbb{R}$ or $\mathbb{C}$ so that pointwise products of functions are well defined almost everywhere. We say that there is a **continuous product map**,

$$W^{k_1,p_1}(\mathcal{U}) \times \ldots \times W^{k_m,p_m}(\mathcal{U}) \to W^{k,p}(\mathcal{U}),$$

or a continuous product **pairing** in the case $m = 2$, if for every set of functions $f_i \in W^{k_i,p_i}(\mathcal{U})$ with $i = 1, \ldots, m$, the pointwise product function $f_1 \cdot \ldots \cdot f_m$ is in $W^{k,p}(\mathcal{U})$ and there is an estimate of the form

$$\|f_1 \cdot \ldots \cdot f_m\|_{W^{k,p}} \leqslant c\|f_1\|_{W^{k_1,p_1}} \cdot \ldots \cdot \|f_m\|_{W^{k_m,p_m}}$$

for some constant $c > 0$ not depending on $f_1, \ldots, f_m$. The case $m = 2$, $k_1 = k_2 = k$ and $p_1 = p_2 = p$ is especially interesting, as the space $W^{k,p}(\mathcal{U})$ is then a **Banach algebra**. More generally, one can ask under what circumstances multiplication by functions of class $W^{k,p}$ defines a bounded linear operator on functions of class $W^{m,q}$. A hint about this comes from the world of classically differentiable functions: multiplication by C^k-smooth functions defines a continuous map $C^m \to C^m$ if and only if $k \geqslant m$. The corresponding answer in Sobolev spaces turns out to be that functions of class $W^{k,p}$ need to have strictly more than zero derivatives in the sense of Remark A.1.7, and at least as many derivatives as functions of class $W^{m,q}$.

Theorem A.2.1. *Assume $\mathcal{U} \subset \mathbb{R}^n$ is an open subset satisfying the strong local Lipschitz condition, $1 \leqslant p, q < \infty$, and $k, m \geqslant 0$ are integers satisfying*

$$k \geqslant m, \qquad kp > n, \qquad \text{and} \qquad k - \frac{n}{p} \geqslant m - \frac{n}{q}.$$

Then there exists a continuous product pairing

$$W^{k,p}(\mathcal{U}, \mathbb{C}) \times W^{m,q}(\mathcal{U}, \mathbb{C}) \to W^{m,q}(\mathcal{U}, \mathbb{C}) : (f,g) \mapsto fg.$$

The following preparatory lemma will be useful both for proving the product estimate and for further results below. It is an easy consequence of Theorem A.1.6 and Hölder's inequality.

Lemma A.2.2. *Assume $\mathcal{U} \subset \mathbb{R}^n$ is an open subset satisfying the strong local Lipschitz condition, $m \geqslant 2$ is an integer, and we are given positive numbers*

$p_1, \ldots, p_m \geqslant 1$ and integers $k_1, \ldots, k_m \geqslant 0$. Let $I := \{i \in \{1, \ldots, m\} \mid k_i p_i \leqslant n\}$. Then for any $q \geqslant 1$ satisfying

$$\sum_{i \in I} \left(\frac{1}{p_i} - \frac{k_i}{n} \right) < \frac{1}{q} \leqslant \sum_{i=1}^{m} \frac{1}{p_i},$$

there is a continuous product map

$$W^{k_1,p_1}(\mathcal{U}) \times \ldots \times W^{k_m,p_m}(\mathcal{U}) \to L^q(\mathcal{U}).$$

PROOF. By the generalized Hölder inequality (A.1), it suffices to show that for any $q \geqslant 1$ in the stated range, one can find numbers $q_1, \ldots, q_m \in [q, \infty]$ satisfying $1/q = 1/q_1 + \ldots + 1/q_m$ for which Theorem A.1.6 provides continuous inclusions

$$W^{k_i,p_i}(\mathcal{U}) \hookrightarrow L^{q_i}(\mathcal{U})$$

for each $i = 1, \ldots, m$. Whenever $k_i p_i > n$, this inclusion is valid with q_i chosen freely from the interval $[p_i, \infty]$, so $1/q_i$ can then take any value subject to the constraint

$$0 \leqslant \frac{1}{q_i} \leqslant \frac{1}{p_i}.$$

If on the other hand $k_i p_i \leqslant n$, then we can arrange $1/q_i$ to take any value in the range

$$\frac{1}{p_i} - \frac{k_i}{n} < \frac{1}{q_i} \leqslant \frac{1}{p_i}.$$

Adding these up, the range of values for $\sum_i \frac{1}{q_i}$ that we can achieve in this way covers the stated interval. □

PROOF OF THEOREM A.2.1. By density of smooth functions, it suffices to prove that an estimate of the form

$$\|fg\|_{W^{m,q}} \leqslant c\|f\|_{W^{k,p}}\|g\|_{W^{m,q}}$$

holds for all $f \in C^\infty(\mathcal{U}) \cap W^{k,p}(\mathcal{U})$ and $g \in C^\infty(\mathcal{U}) \cap W^{m,q}(\mathcal{U})$. Equivalently, we need to show that for all f and g of this type and every multiindex α of degree $|\alpha| \leqslant m$, there is a constant $c > 0$ independent of f and g such that

$$\|\partial^\alpha(fg)\|_{L^q} \leqslant c\|f\|_{W^{k,p}}\|g\|_{W^{m,q}}.$$

Since f and g are smooth, we are free to use the product rule in computing $\partial^\alpha(fg)$, which will then be a linear combination of terms of the form $\partial^\beta f \cdot \partial^\gamma g$ where $|\alpha| = |\beta| + |\gamma|$, hence we have reduced the problem to proving a bound

$$\|\partial^\beta f \cdot \partial^\gamma g\|_{L^q} \leqslant c\|f\|_{W^{k,p}}\|g\|_{W^{m,q}}$$

for every pair of multiindices β, γ with $|\beta| + |\gamma| \leqslant m$. Since $\partial^\beta f \in W^{k-|\beta|,p}(\mathcal{U})$ and $\partial^\gamma f \in W^{m-|\gamma|,q}(\mathcal{U})$, the result follows if we can assume that for every pair of integers $a, b \geqslant 0$ satisfying $a + b \leqslant m$, there exists a continuous product pairing

$$W^{k-a,p}(\mathcal{U}) \times W^{m-b,q}(\mathcal{U}) \to L^q(\mathcal{U}). \tag{A.6}$$

If $(k - a)p > n$, then $W^{k-a,p} \hookrightarrow L^\infty$ and (A.6) is immediate since $W^{m-b,q} \hookrightarrow L^q(\mathcal{U})$. For the remaining cases, we shall apply Lemma A.2.2, noting that the condition $1/q \leqslant 1/p + 1/q$ is trivially satisfied.

If $(m-b)q > n$ but $(k-a)p \leqslant n$, then the hypotheses of the lemma are satisfied if and only if

$$\frac{1}{p} - \frac{k-a}{n} < \frac{1}{q}.$$

Since $\frac{1}{p} - \frac{k}{n} \leqslant \frac{1}{q} - \frac{m}{n}$ by assumption, we have

$$\frac{1}{p} - \frac{k-a}{n} = \frac{1}{p} - \frac{k}{n} + \frac{a}{n} \leqslant \frac{1}{q} - \frac{m}{n} + \frac{a}{n} \leqslant \frac{1}{q}$$

since $a \leqslant m$, and equality holds only if $a = m$, $b = 0$ and $k - n/p = m - n/q$, which implies $mq > n$. In this case $W^{m-b,q} = W^{m,q} \hookrightarrow L^\infty$, and the pairing (A.6) follows because $W^{k-a,p} = W^{k-m,p}$ embeds continuously into L^q: the latter follows from Theorem A.1.6 since $\frac{1}{p} - \frac{k-m}{n} = \frac{1}{q}$.

Finally, when $(k-a)p \leqslant n$ and $(m-b)q \leqslant n$, the hypotheses of the lemma are satisfied since

$$\left(\frac{1}{p} - \frac{k-a}{n}\right) + \left(\frac{1}{q} - \frac{m-b}{n}\right) \leqslant \frac{1}{p} - \frac{k}{n} + \frac{1}{q} - \frac{m}{n} + \frac{m}{n} = \left(\frac{1}{p} - \frac{k}{n}\right) + \frac{1}{q} < \frac{1}{q},$$

where we've used the assumption $kp > n$ and the fact that $a + b \leqslant m$. ☐

REMARK A.2.3. A much simpler argument shows similarly that for any open domain $\mathcal{U} \subset \mathbb{R}^n$, any integer $k \geqslant 1$ and any $p \in [1, \infty)$, there is a continuous product pairing

$$C^k(\overline{\mathcal{U}}, \mathbb{C}) \times W^{k,p}(\mathcal{U}, \mathbb{C}) \times W^{k,p}(\mathcal{U}, \mathbb{C}).$$

As in Theorem A.2.1, this follows from the density of $C^\infty \cap W^{k,p} \subset W^{k,p}$ after showing that all $f \in C^k(\overline{\mathcal{U}})$ and $g \in C^\infty(\mathcal{U}) \cap W^{k,p}(\mathcal{U})$ satisfy an estimate of the form $\|fg\|_{W^{k,p}} \leqslant c\|f\|_{C^k}\|g\|_{W^{k,p}}$. The latter follows easily from the definition of the $W^{k,p}$-norm.

In general it is not straightforward to say when the usual product rule $\partial_i(fg) = \partial_i f \cdot g + f \cdot \partial_i g$ does or does not hold in the sense of weak derivatives. If g and $\partial_i g$ are locally integrable and f is smooth, then there is no trouble: the formula can be derived in this case directly from the definition of weak derivatives, using the observation that for any test function $\varphi \in C_0^\infty(\mathcal{U})$, φf is also in $C_0^\infty(\mathcal{U})$ and satisfies the product rule. If on the other hand f and g are not continuous but have well-defined weak derivatives and a locally integrable product, then there is no guarantee in general that any of $\partial_i(fg)$, $\partial_i f \cdot g$ or $f \cdot \partial_i g$ should be well-defined locally integrable functions. Theorem A.2.1 provides a means of resolving this question whenever f and g belong to suitable Sobolev spaces.

PROPOSITION A.2.4. *Suppose k, m, p, q and $\mathcal{U} \subset \mathbb{R}^n$ satisfy the same conditions as in Theorem A.2.1, and $m \geqslant 1$. Then for every $f \in W^{k,p}(\mathcal{U}, \mathbb{C})$ and $g \in W^{m,q}(\mathcal{U}, \mathbb{C})$, the weak partial derivatives of $fg \in W^{m,q}(\mathcal{U}, \mathbb{C})$ are given almost everywhere by the usual Leibniz rule $\partial_i(fg) = \partial_i f \cdot g + f \cdot \partial_i g$.*

PROOF. Choose sequences of smooth functions f_j, g_j with $f_j \to f$ in $W^{k,p}$ and $g_j \to g$ in $W^{m,q}$. Then since $k \geqslant m \geqslant 1$, there is also L^p-convergence $\partial_i f_j \to \partial_i f$ and L^q-convergence $\partial_i g_j \to \partial_i g$, so after restricting to a subsequence, we may assume that all four of the sequences f_j, $\partial_i f_j$, g_j and $\partial_i g_j$ converge pointwise almost everywhere. The continuity of the product pairing $W^{k,p} \times W^{m,q} \to W^{m,q}$ now implies $W^{m,q}$-convergence $f_j g_j \to fg$ and thus L^q-convergence

$$\partial_i(f_j g_j) = \partial_i f_j \cdot g_j + f_j \cdot \partial_i g_j \to \partial_i(fg).$$

The result follows since $\partial_i f_j \cdot g_j + f_j \cdot \partial_i g_j$ also converges pointwise almost everywhere to $\partial_i f \cdot g + f \cdot \partial_i g$. □

REMARK A.2.5. A slight simplification of the same argument as in Proposition A.2.4 shows that the product rule also holds (without any assumption on the open domain $\mathcal{U} \subset \mathbb{R}^n$) for $f \in C^m(\overline{\mathcal{U}}, \mathbb{C})$ and $g \in W^{m,p}(\mathcal{U}, \mathbb{C})$ for any $p \in [1, \infty)$ if $m \geqslant 1$. The key facts here are the continuity of the product pairing $C^m \times W^{m,p} \to W^{m,p}$ and the density of C^1 in $W^{m,p}$, so that f and g can be approximated by pairs for which the classical product rule holds. Both results can also be extended in a similar manner to prove the expected formula for $\partial^\alpha(fg)$ for any multiindex α of order $|\alpha| \leqslant m$.

The next result generalizes Proposition 2.2.5 and concerns the following question: if $f : \mathcal{U} \to \mathbb{R}^m$ is a function of class $W^{k,p}$ whose graph lies in some open subset $\mathcal{V} \subset \mathcal{U} \times \mathbb{R}^m$, and $\Psi : \mathcal{V} \to \mathbb{R}^N$ is another function, under what conditions can we conclude that the function

$$\mathcal{U} \to \mathbb{R}^N : x \mapsto \Psi(x, f(x))$$

is in $W^{k,p}(\mathcal{U}, \mathbb{R}^N)$? We will abbreviate this function in the following by $\Psi \circ (\mathrm{Id} \times f)$, and we would also like to know whether it depends continuously (in the $W^{k,p}$-topology) on f and Ψ. The following theorem is stated rather generally, but on first reading you may prefer to assume $\mathcal{U} \subset \mathbb{R}^n$ is bounded, in which case some of the hypotheses become vacuous. We will say that an open subset $\mathcal{V} \subset \mathcal{U} \times \mathbb{R}^m$ is a **star-shaped neighborhood of** $f : \mathcal{U} \to \mathbb{R}^m$ if it contains the graph of f and

$$(x, v) \in \mathcal{V} \quad \Rightarrow \quad (x, tv + (1-t)f(x)) \in \mathcal{V} \text{ for all } t \in [0, 1].$$

THEOREM A.2.6. *Assume $\mathcal{U} \subset \mathbb{R}^n$ is an open subset satisfying the strong local Lipschitz condition, $p \in [1, \infty)$ and $k \in \mathbb{N}$ satisfy $kp > n$, and $\mathcal{V} \subset \mathcal{U} \times \mathbb{R}^m$ is a star-shaped neighborhood of some function $f_0 \in W^{k,p}(\mathcal{U}, \mathbb{R}^m)$. Assume also $\mathcal{O}^{k,p}(\mathcal{U}; \mathcal{V}) \subset W^{k,p}(\mathcal{U}, \mathbb{R}^m)$ is an open neighborhood of f_0 such that*

$$(x, f(x)) \in \mathcal{V} \quad \text{for all } x \in \mathcal{U} \text{ and } f \in \mathcal{O}^{k,p}(\mathcal{U}; \mathcal{V}),$$

and $C^k_0(\overline{\mathcal{V}}, \mathbb{R}^N) \subset C^k(\overline{\mathcal{V}}, \mathbb{R}^N)$ is a closed linear subspace such that all $\Psi \in C^k_0(\overline{\mathcal{V}}, \mathbb{R}^n)$ have the following properties:[1]

(1) There exists a bounded subset $\mathcal{K} \subset \mathcal{U}$ such that $\Psi(x, v)$ is independent of x for all $x \in \mathcal{U} \backslash \mathcal{K}$;

(2) $\Psi \circ (\mathrm{Id} \times f_0) \in L^p(\mathcal{U}, \mathbb{R}^N)$.

Then there is a well-defined and continuous map

$$\mathcal{O}^{k,p}(\mathcal{U}; \mathcal{V}) \xrightarrow{T} \mathscr{L}\big(C^k_0(\overline{\mathcal{V}}, \mathbb{R}^N), W^{k,p}(\mathcal{U}, \mathbb{R}^N)\big),$$
$$T(f)\Psi := \Psi \circ (\mathrm{Id} \times f),$$

[1]Both of the conditions on $\Psi \in C^k_0(\overline{\mathcal{V}}, \mathbb{R}^n)$ are vacuous if $\mathcal{U} \subset \mathbb{R}^n$ is bounded.

so in particular the map

$$C_0^k(\overline{\mathcal{V}}, \mathbb{R}^N) \times \mathcal{O}^{k,p}(\mathcal{U};\mathcal{V}) \to W^{k,p}(\mathcal{U}, \mathbb{R}^N) : (\Psi, f) \mapsto \Psi \circ (\mathrm{Id} \times f),$$

is well defined and continuous. Moreover, for each $\Psi \in C_0^k(\overline{\mathcal{V}}, \mathbb{R}^N)$ *and* $f \in W^{k,p}(\mathcal{U}, \mathbb{R}^N)$, *the weak partial derivatives of* $\Psi \circ (\mathrm{Id} \times f)$ *are given almost everywhere by the classical formula*

$$\partial_j \left[\Psi \circ (\mathrm{Id} \times f)\right](x) = \partial_j \Psi(x, f(x)) + D_2\Psi(x, f(x))\partial_j f(x),$$

where $\partial_j \Psi$ *denotes the partial derivative of* $\Psi(x,v)$ *with respect to the* j*th coordinate in* $x \in \mathbb{R}^n$, *and* $D_2\Psi$ *is its differential with respect to* $v \in \mathbb{R}^m$.

PROOF. We will show first that if $f \in \mathcal{O}^{k,p}(\mathcal{U};\mathcal{V})$ is smooth, then $\Psi \circ (\mathrm{Id} \times f)$ belongs to $W^{k,p}(\mathcal{U}, \mathbb{R}^N)$ for every $\Psi \in C_0^k(\overline{\mathcal{V}}, \mathbb{R}^N)$. Since $\mathcal{V}$ is a star-shaped neighborhood of f_0, we have

$$\begin{aligned} |\Psi(x, f(x)) - \Psi(x, f_0(x))| &= \left| \int_0^1 \frac{d}{dt} \Psi\big(x, tf(x) + (1-t)f_0(x)\big)\, dt \right| \\ &\leqslant \left(\int_0^1 \left|D_2\Psi\big(x, tf(x) + (1-t)f_0(x)\big)\right| dt \right) \cdot |f(x) - f_0(x)| \\ &\leqslant \|\Psi\|_{C^1(\mathcal{V})} \cdot |f(x) - f_0(x)| \end{aligned}$$

for all $x \in \mathcal{U}$, implying

$$\|\Psi \circ (\mathrm{Id} \times f) - \Psi \circ (\mathrm{Id} \times f_0)\|_{L^p} \leqslant \|\Psi\|_{C^1(\mathcal{V})} \cdot \|f - f_0\|_{L^p}, \tag{A.7}$$

hence $\Psi \circ (\mathrm{Id} \times f) \in L^p(\mathcal{U}, \mathbb{R}^N)$.

For $\ell = 1, \ldots, k$, we can regard the ℓth derivative of Ψ with respect to variables in $\mathbb{R}^m$ as a bounded and uniformly continuous map from $\mathcal{V}$ into the vector space of symmetric ℓ-multilinear maps from $\mathbb{R}^m$ to $\mathbb{R}^N$, denoting this by

$$D_2^\ell \Psi : \mathcal{V} \to \mathrm{Hom}((\mathbb{R}^m)^{\otimes \ell}, \mathbb{R}^N).$$

Denote the partial derivatives with respect to variables in $\mathcal{U} \subset \mathbb{R}^n$ by

$$D_1^\beta \Psi : \mathcal{V} \to \mathbb{R}^N,$$

where β is a multiindex in n variables. Now for any multiindex α with $|\alpha| \leqslant k$, the derivative $\partial^\alpha(\Psi \circ (\mathrm{Id} \times f))$ is a linear combination of product functions of the form

$$(D_1^\gamma D_2^\ell \Psi \circ (\mathrm{Id} \times f))(\partial^{\beta_1} f, \ldots, \partial^{\beta_\ell} f) : \mathcal{U} \to \mathbb{R}^N, \tag{A.8}$$

where $\ell + |\gamma| \in \{1, \ldots, |\alpha|\}$ and $|\beta_1| + \ldots + |\beta_\ell| = |\alpha| - |\gamma|$. If $\ell = 0$ but $|\gamma| > 0$, then this expression is clearly in $L^p(\mathcal{U}, \mathbb{R}^N)$ since it is continuous and $D_1^\gamma \Psi(x,v) = 0$ for $x \in \mathcal{U} \backslash \mathcal{K}$, where $\mathcal{K}$ is bounded. For $\ell \geqslant 1$, it satisfies

$$\left\|(D_1^\gamma D_2^\ell \Psi \circ (\mathrm{Id} \times f))(\partial^{\beta_1} f, \ldots, \partial^{\beta_\ell} f)\right\|_{L^p(\mathcal{U})} \leqslant \|D_1^\gamma D_2^\ell \Psi\|_{C^0(\mathcal{V})} \cdot \left\| \prod_{j=1}^{\ell} |\partial^{\beta_j} f| \right\|_{L^p(\mathcal{U})} \tag{A.9}$$

if the product on the right hand side has finite L^p-norm. The latter is trivially true if $\ell = 1$. To deal with the $\ell \geqslant 2$ case, note that $\partial^{\beta_j} f \in W^{k-|\beta_j|,p}(\mathcal{U})$ for each $j = 1, \ldots, \ell$, so the necessary bound will follow from the existence of a continuous product map

$$W^{k-m_1,p}(\mathcal{U}) \times \ldots \times W^{k-m_\ell,p}(\mathcal{U}) \to L^p(\mathcal{U})$$

for $m_j := |\beta_j|$, and we claim that such a product map does exist whenever $kp > n$ and $m_1, \dots, m_\ell \geqslant 0$ are integers satisfying $m_1 + \dots + m_\ell \leqslant k$. To see this, note first that since $W^{k-m_j,p} \hookrightarrow L^\infty$ whenever $(k - m_j)p > n$, it suffices to prove the claim under the assumption that $(k - m_j)p \leqslant n$ for every $j = 1, \dots, \ell$. In this case, Lemma A.2.2 provides the desired product map if the condition

$$\sum_{j=1}^{\ell} \left(\frac{1}{p} - \frac{k - m_j}{n} \right) < \frac{1}{p} \leqslant \sum_{j=1}^{\ell} \frac{1}{p}$$

is satisfied. And it is: using $kp > n$, $\ell \geqslant 2$ and $m_1 + \dots + m_\ell \leqslant k$, we find

$$\begin{aligned} \sum_{j=1}^{\ell} \left(\frac{1}{p} - \frac{k - m_j}{n} \right) &= \ell \left(\frac{1}{p} - \frac{k}{n} \right) + \frac{m_1 + \dots + m_\ell}{n} \\ &\leqslant \frac{1}{p} + (\ell - 1) \left(\frac{1}{p} - \frac{k}{n} \right) < \frac{1}{p}. \end{aligned}$$

This proves that $\Psi \circ (\mathrm{Id} \times f) \in W^{k,p}(\mathcal{U}, \mathbb{R}^N)$.

An important detail in both of the estimates (A.7) and (A.9) is that on the right hand side, the term depending on Ψ is bounded by something linearly proportional to $\|\Psi\|_{C^k(\mathcal{V})}$, and the same is true of other estimates mentioned below that can be derived in a similar manner. We will not comment on this point any further, but it is the reason why rather than just proving that the map $(\Psi, f) \mapsto \Psi \circ (\mathrm{Id} \times f)$ is continuous, we will obtain the stronger result that the map sending f to the linear operator $\Psi \mapsto \Psi \circ (\mathrm{Id} \times f)$ is continuous with respect to the operator norm.

Next, suppose $f \in \mathcal{O}^{k,p}(\mathcal{U}; \mathcal{V})$ is not necessarily smooth but $f_i \in \mathcal{O}^{k,p}(\mathcal{U}; \mathcal{V})$ is a sequence of smooth functions converging to f in $W^{k,p}$, while $\Psi_i \in C_0^k(\overline{\mathcal{V}}, \mathbb{R}^N)$ converges to $\Psi \in C_0^k(\overline{\mathcal{V}}, \mathbb{R}^N)$ in C^k. Then the same argument we used to estimate $\|\Psi \circ (\mathrm{Id} \times f) - \Psi \circ (\mathrm{Id} \times f_0)\|_{L^p}$ shows that $\Psi_i \circ (\mathrm{Id} \times f_i) \to \Psi \circ (\mathrm{Id} \times f)$ in L^p, and since f_i is also C^0-convergent, the compactly supported functions $D_1^\gamma \Psi_i \circ (\mathrm{Id} \times f_i)$ converge to $D_1^\gamma \Psi \circ (\mathrm{Id} \times f)$ in L^p for each multiindex with $1 \leqslant |\gamma| \leqslant k$. For $\ell \geqslant 1$ and $|\gamma| + \ell \leqslant k$, $D_1^\gamma D_2^\ell \Psi_i \circ (\mathrm{Id} \times f_i)$ converges to $D_1^\gamma D_2^\ell \Psi \circ (\mathrm{Id} \times f)$ in $C^0(\overline{\mathcal{U}}, \mathbb{R}^N)$, and each of the derivatives $\partial^{\beta_j} f_i$ appearing in (A.8) also converges in $L^p(\mathcal{U})$. In light of the continuous product maps discussed above, it follows that each derivative $\partial^\alpha(\Psi_i \circ (\mathrm{Id} \times f_i))$ for $|\alpha| \leqslant k$ is L^p-convergent, and its limit is necessarily (by Exercise A.2.7 below) the corresponding weak derivative $\partial^\alpha(\Psi \circ (\mathrm{Id} \times f))$, hence $\Psi \circ (\mathrm{Id} \times f) \in W^{k,p}(\mathcal{U}, \mathbb{R}^N)$ and $\Psi_i \circ (\mathrm{Id} \times f_i) \xrightarrow{W^{k,p}} \Psi \circ (\mathrm{Id} \times f)$. Since all sequences in this discussion can also be replaced with subsequences that are pointwise almost everywhere convergent, this also proves that the classical formula for $\partial^\alpha(\Psi_i \circ (\mathrm{Id} \times f_i))$ for each $|\alpha| \leqslant k$ remains valid for computing the corresponding weak derivative $\partial^\alpha(\Psi \circ (\mathrm{Id} \times f))$. With this understood, one can now repeat the arguments of this paragraph for an arbitrary $W^{k,p}$-convergent sequence $f_i \to f$ without assuming the f_i are smooth, thus proving the continuity of the map $(\Psi, f) \mapsto \Psi \circ (\mathrm{Id} \times f)$. □

EXERCISE A.2.7. Show that if f_i is a sequence of smooth functions on an open set $\mathcal{U} \subset \mathbb{R}^n$ with $f_i \xrightarrow{L^p} f$ and $\partial^\alpha f_i \xrightarrow{L^p} g$ for some multiindex α and functions $f, g \in L^p(\mathcal{U})$, then $\partial^\alpha f = g$ in the sense of distributions.

The following result on coordinate transformations of the domain can be proved in an analogous way to Theorem A.2.6, though it is considerably easier since there is no need to worry about Sobolev product maps (and thus no need to assume $kp > n$ or impose regularity conditions on the domain).

THEOREM A.2.8 ([AF03, §3.41]). *Assume $k \in \mathbb{N}$, $1 \leqslant p \leqslant \infty$, and $\mathcal{U}, \mathcal{U}' \subset \mathbb{R}^n$ are open subsets with a C^k-smooth diffeomorphism $\varphi : \mathcal{U} \to \mathcal{U}'$ such that all derivatives of φ and φ^{-1} up to order k are bounded and uniformly continuous. Then there is a well-defined Banach space isomorphism*

$$W^{k,p}(\mathcal{U}') \to W^{k,p}(\mathcal{U}) : f \mapsto f \circ \varphi.$$

□

Next, we restate and prove Proposition 2.2.8. Denote by $\mathring{\mathbb{D}}^n$ and $\mathring{\mathbb{D}}^n_\epsilon(x_0)$ the open balls of radius 1 and ϵ about the origin and a point x_0 respectively in $\mathbb{R}^n$.

THEOREM A.2.9. *Assume $p \in [1, \infty)$ and $k \in \mathbb{N}$ satisfy $kp > n$, and for a given point $x_0 \in \mathring{\mathbb{D}}^n$ with $\epsilon_0 := \operatorname{dist}(x_0, \partial\mathbb{D}^n)$, associate to each $f \in W^{k,p}(\mathring{\mathbb{D}}^n)$ and $\epsilon \in (0, \epsilon_0)$ the function $f_\epsilon \in W^{k,p}(\mathring{\mathbb{D}}^n)$ defined by*

$$f_\epsilon(x) := f(x_0 + \epsilon x).$$

Then for each $\alpha \in (0,1)$ satisfying $\alpha \leqslant k - \frac{n}{p}$, there exists a constant $C > 0$ such that the estimate

$$\|f_\epsilon - f_\epsilon(0)\|_{W^{k,p}} \leqslant C\epsilon^\alpha \|f - f(x_0)\|_{W^{k,p}}$$

holds for all $f \in W^{k,p}(\mathring{\mathbb{D}}^n)$ and $\epsilon \in (0, \epsilon_0)$.

PROOF. To estimate $\|f_\epsilon - f_\epsilon(0)\|_{L^p}$, we use the fact that $f - f(x_0) \in W^{k,p}$ is Hölder continuous, i.e. Theorem A.1.6 embeds $W^{k,p}$ continuously into $C^{0,\alpha}$ for any $\alpha \in (0,1)$ with $\alpha \leqslant k - n/p$, thus f satisfies

$$|f(x) - f(x_0)| \leqslant c\|f - f(x_0)\|_{W^{k,p}(\mathring{\mathbb{D}}^n)} \cdot |x - x_0|^\alpha \qquad \text{for all } x \in \mathring{\mathbb{D}}^n_{\epsilon_0}(x_0)$$

for some constant $c > 0$. We therefore have

$$\begin{aligned} \|f_\epsilon - f_\epsilon(0)\|^p_{L^p} &= \int_{\mathbb{D}^n} |f(x_0 + \epsilon x) - f(x_0)|^p \leqslant c^p \|f - f(x_0)\|^p_{W^{k,p}} \int_{\mathbb{D}^n} |\epsilon x|^{\alpha p} \\ &= c^p \|f - f(x_0)\|^p_{W^{k,p}} \cdot \epsilon^{\alpha p} \int_{\mathbb{D}^n} |x|^{\alpha p} =: C^p \epsilon^{\alpha p} \|f - f(x_0)\|^p_{W^{k,p}} \end{aligned}$$

for a suitable constant $C > 0$, implying $\|f_\epsilon - f_\epsilon(0)\|_{L^p} \leqslant C\epsilon^\alpha \|f - f(x_0)\|_{W^{k,p}}$.

Next, consider a multiindex β of order $|\beta| = m \in \{1, \ldots, k\}$. The functions $\partial^\beta(f - f(x_0)) = \partial^\beta f$ and $\partial^\beta(f_\epsilon - f_\epsilon(0)) = \partial^\beta f_\epsilon$ for each $\epsilon \in (0, \epsilon_0)$ are then in $W^{k-m,p}(\mathring{\mathbb{D}}^n)$, and we need to establish bounds on $\|\partial^\beta f_\epsilon\|_{L^p}$ in terms of the $W^{k,p}$-norm of $f - f(x_0)$. If $m < k$, then Theorem A.1.6 gives a continuous inclusion

$$W^{k-m,p}(\mathring{\mathbb{D}}^n) \hookrightarrow L^q(\mathring{\mathbb{D}}^n) \tag{A.10}$$

for any $q \in [p, \infty)$ satisfying $1/q \geqslant 1/p - (k - m)/n$. The same is also trivially true in the case $m = k$, since q and p must then be equal. Notice that if $(k - m)p \geqslant n$, then q is allowed to be arbitrarily large. We will therefore assume in general that (A.10) holds with $q \in [p, \infty)$ satisfying

$$\frac{1}{q} + \frac{1}{r} = \frac{1}{p},$$

where $r = \frac{n}{k-m} \in (0, \infty]$ if $(k-m)p < n$ and otherwise $r = p + \delta$ for some $\delta > 0$ which may be chosen arbitrarily small. Given this, we apply change of variables and Hölder's inequality to find

$$\begin{aligned}
\|\partial^\beta f_\epsilon\|^p_{L^p(\mathring{\mathbb{D}}^n)} &= \epsilon^{mp} \int_{\mathbb{D}^n} |\partial^\beta f(x_0 + \epsilon x)|^p = \epsilon^{mp-n} \int_{\mathbb{D}^n_\epsilon(x_0)} |\partial^\beta f(x)|^p \\
&\leqslant \epsilon^{mp-n} \|\partial^\beta f\|^p_{L^q(\mathring{\mathbb{D}}^n_\epsilon)} \|1\|^p_{L^r(\mathring{\mathbb{D}}^n_\epsilon)} \\
&\leqslant \epsilon^{mp-n} \left[\mathrm{Vol}(\mathbb{D}^n_\epsilon(x_0))\right]^{p/r} \|\partial^\beta f\|^p_{L^q(\mathring{\mathbb{D}}^n)} \\
&\leqslant c\epsilon^{mp-n} \left[\mathrm{Vol}(\mathbb{D}^n_\epsilon(x_0))\right]^{p/r} \|\partial^\beta f\|^p_{W^{k-m,p}(\mathring{\mathbb{D}}^n)} \\
&\leqslant c\epsilon^{mp-n} \left[\mathrm{Vol}(\mathbb{D}^n_\epsilon(x_0))\right]^{p/r} \|f - f(x_0)\|^p_{W^{k,p}(\mathring{\mathbb{D}}^n)}
\end{aligned}$$

for some constant $c > 0$. Writing $\mathrm{Vol}(\mathbb{D}^n_\epsilon(x_0)) = C\epsilon^n$ for a suitable constant $C > 0$, the exponent on ϵ in this expression becomes $mp - n + \frac{np}{r}$. If $(k-m)p < 0$, this is exactly $kp - n = (k - n/p)p$, and otherwise, taking $r - p > 0$ to be arbitrarily small makes it less than but arbitrarily close to mp. Since $\alpha \leqslant k - n/p$ and $\alpha < 1 \leqslant m$, we are now free to replace this exponent with αp and rewrite the established estimate as $\|\partial^\beta f_\epsilon\|_{L^p} \leqslant C\epsilon^\alpha \|f - f(x_0)\|_{W^{k,p}}$. □

A.3. Difference quotients

If f is a function on $\mathbb{R}^n$, then for every $i = 1, \ldots, n$ and $h \in \mathbb{R}\backslash\{0\}$, the **difference quotient**

$$D^h_i f(x_1, \ldots, x_n) := \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}$$

defines a function $D^h_i f$ on $\mathbb{R}^n$. The **total difference quotient** of f is then the n-tuple of functions

$$D^h f := (D^h_1 f, \ldots, D^h_n f),$$

so for example if $f : \mathbb{R}^n \to \mathbb{R}^m$, then $D^h f : \mathbb{R}^n \to \mathbb{R}^{mn}$. The transformation $f \mapsto D^h_i f$ is obviously linear for any fixed number h, and it satisfies a Leibniz rule

$$D^h_i(fg) = D^h_i f \cdot g + f \cdot D^h_i g$$

whenever pointwise products of f and g can be defined (e.g. if both are real or complex valued). It also commutes with differentiation

$$D^h_i(\partial_j f) = \partial_j(D^h_i f)$$

on any function f for which $\partial_j f$ can be defined (weakly or strongly). Clearly if $f \in W^{k,p}(\mathbb{R}^n)$, then $D^h f \in W^{k,p}(\mathbb{R}^n)$ for every $h \in \mathbb{R}\backslash 0$, and if f is supported in an open subset $\mathcal{U} \subset \mathbb{R}^n$, then $D^h f$ is supported in an arbitrarily small neighborhood of $\overline{\mathcal{U}}$ for sufficiently small $|h|$. Moreover, if f is a function defined only on $\mathcal{U} \subset \mathbb{R}^n$, then on any open subset $\mathcal{V} \subset \mathcal{U}$ with compact closure in $\mathcal{U}$, $D^h f$ can be defined on $\mathcal{V}$ for any $h \in \mathbb{R}\backslash\{0\}$ satisfying

$$|h| < \operatorname{dist}(\mathcal{V}, \mathbb{R}^n\backslash\mathcal{U}) := \inf\left\{|x-y| \;\middle|\; x \in \mathcal{V} \text{ and } y \in \mathbb{R}^n\backslash\mathcal{U}\right\}.$$

The following result about difference quotients is useful for proving local regularity of solutions to PDEs, as in §2.4.

THEOREM A.3.1. *Assume $\mathcal{V} \subset \mathcal{U} \subset \mathbb{R}^n$ are open subsets with $\mathcal{V}$ having compact closure contained in $\mathcal{U}$, $1 \leqslant p < \infty$, and $k \in \mathbb{N}$.*

(1) If $f \in W^{k,p}(\mathcal{U})$, then $D^h f$ converges to ∇f in $W^{k-1,p}$ on $\mathcal{V}$ as $h \to 0$, and

$$\|D^h f\|_{W^{k-1,p}(\mathcal{V})} \leqslant \|\nabla f\|_{W^{k-1,p}(\mathcal{U})}$$

for all $h \neq 0$ with $|h| < \operatorname{dist}(\mathcal{V}, \mathbb{R}^n\backslash\mathcal{U})$.

(2) Suppose $p > 1$, $f \in W^{k-1,p}(\mathcal{U})$ and the difference quotients $D^h f$ satisfy a uniform bound

$$\|D^h f\|_{W^{k-1,p}(\mathcal{V})} \leqslant C$$

for all $h \neq 0$ with $|h|$ sufficiently small. Then $f|_{\mathcal{V}} \in W^{k,p}(\mathcal{V})$ and its first derivative satisfies $\|\nabla f\|_{W^{k-1,p}(\mathcal{V})} \leqslant m_{k,p} C$, where $m_{k,p} \in \mathbb{N}$ is a constant depending only on the definition of the $W^{k-1,p}$-norm.

The next few results are intended as preparation for the proof of Theorem A.3.1.

LEMMA A.3.2. *For any open subset $\mathcal{U} \subset \mathbb{R}^n$ and continuously differentiable function f on $\mathcal{U}$, the difference quotients $D_i^h f$ converge to $\partial_i f$ uniformly on compact subsets as $h \to 0$.*

PROOF. Fix a compact subset $\mathcal{K} \subset \mathcal{U}$. Then for every $x \in \mathcal{K}$ and $h \in \mathbb{R}\backslash\{0\}$ sufficiently small, the mean value theorem gives

$$D_i^h f(x) = \partial_i f(x')$$

where

$$x' := (x_1, \ldots, x_{i-1}, x_i + th, x_{i+1}, \ldots, x_n) \in \mathcal{U}$$

for some $t \in [0,1]$, so in particular, $|x' - x| \leqslant |h|$. We then have $|\partial_i f(x) - D_i^h f(x)| = |\partial_i f(x) - \partial_i f(x')|$, and the result follows since both x and x' may be assumed to lie in a compact subset of $\mathcal{U}$, on which $\partial_i f$ is uniformly continuous. □

PROPOSITION A.3.3. *Suppose $1 \leqslant p < \infty$, $\mathcal{U} \subset \mathbb{R}^n$ is an open subset and $f \in W^{1,p}(\mathcal{U})$. Then for any open subset $\mathcal{V} \subset \mathcal{U}$ with compact closure in $\mathcal{U}$, $\|D^h f\|_{L^p(\mathcal{V})} \leqslant \|\nabla f\|_{L^p(\mathcal{U})}$ for every $h \neq 0$ with $|h| < \operatorname{dist}(\mathcal{V}, \mathbb{R}^n\backslash\mathcal{U})$, and $D^h f \to \nabla f$ in L^p on $\mathcal{V}$ as $h \to 0$.*

PROOF. We show first that for any $f \in W^{1,p}(\mathcal{U})$,

$$\|D_i^h f\|_{L^p(\mathcal{V})} \leqslant \|\partial_i f\|_{L^p(\mathcal{U})}, \qquad i = 1, \ldots, n \tag{A.11}$$

for every $\mathcal{V} \subset \mathcal{U}$ with compact closure in $\mathcal{U}$ and every $h \neq 0$ with $|h| < \operatorname{dist}(\mathcal{V}, \mathbb{R}^n\backslash\mathcal{U})$. Indeed, if $f \in W^{1,p}(\mathcal{U}) \cap C^\infty(\mathcal{U})$, then denoting the standard basis of $\mathbb{R}^n$ by $(e_1, \ldots, e_n)$, we have

$$\left|D_i^h f(x)\right| = \left|\frac{f(x+he_i)-f(x)}{h}\right| = \left|\frac{1}{h}\int_0^1 \frac{d}{dt} f(x+the_i)\,dt\right|$$
$$= \left|\int_0^1 \partial_i f(x+the_i)\,dt\right| \leqslant \int_0^1 |\partial_i f(x+the_i)|\,dt.$$

Then since any measurable function $g : [0,1] \to \mathbb{R}$ satisfies

$$\left(\int_0^1 |g(t)|\,dt\right)^p \leqslant \int_0^1 |g(t)|^p\,dt$$

by Jensen's inequality, this gives

$$\|D_i^h f\|_{L^p(\mathcal{V})}^p = \int_{\mathcal{V}} \left|D_i^h f(x)\right|^p d\mu(x) \leqslant \int_{\mathcal{V}} \left(\int_0^1 |\partial_i f(x+the_i)|\,dt\right)^p d\mu(x)$$
$$\leqslant \int_{\mathcal{V}} \int_0^1 |\partial_i f(x+the_i)|^p\,dt\,d\mu(x) = \int_0^1 \int_{\mathcal{V}} |\partial_i f(x+the_i)|^p\,d\mu(x)\,dt$$
$$\leqslant \int_0^1 \|\partial_i f\|_{L^p(\mathcal{U})}^p\,dt = \|\partial_i f\|_{L^p(\mathcal{U})}^p.$$

This estimate extends to every $f \in W^{1,p}(\mathcal{U})$ by density of smooth functions.

Next, suppose $f \in W^{1,p}(\mathcal{U})$ and $\epsilon > 0$ is given. Choose a smooth approximation $f_\epsilon \in W^{1,p}(\mathcal{U}) \cap C^\infty(\mathcal{U})$ with $\|f - f_\epsilon\|_{W^{1,p}(\mathcal{U})} < \epsilon/3$. By Lemma A.3.2, $D_i^h f_\epsilon \to \partial_i f_\epsilon$ in C^0_{loc} on $\mathcal{U}$ as $h \to 0$, and since $\mathcal{V}$ has finite measure, this implies we can find $\delta > 0$ such that $|h| < \delta$ implies $\|D_i^h f_\epsilon - \partial_i f_\epsilon\|_{L^p(\mathcal{V})} < \epsilon/3$. Now by (A.11),

$$\|D_i^h f_\epsilon - D_i^h f\|_{L^p(\mathcal{V})} \leqslant \|\partial_i f_\epsilon - \partial_i f\|_{L^p(\mathcal{U})} \leqslant \|f_\epsilon - f\|_{W^{1,p}(\mathcal{U})} < \epsilon/3,$$

so combining these estimates gives $\|D_i^h f - \partial_i f\|_{L^p(\mathcal{V})} < \epsilon$ whenever $|h| < \delta$. □

The proof of the next proposition will require the following standard result from real analysis, known as the **Banach-Alaoglu theorem**. It follows easily from the separability of L^p-spaces for $p < \infty$ together with the duality of L^p and L^q for $1/p + 1/q = 1$; see for instance [**LL01**, §2.18].

THEOREM A.3.4 (Banach-Alaoglu). *For any measurable subset $\mathcal{U} \subset \mathbb{R}^n$, if $1 < p < \infty$, then every bounded sequence $f_j \in L^p(\mathcal{U})$ has a weakly convergent subsequence, i.e. after passing to a subsequence, one can find a function $f_\infty \in L^p(\mathcal{U})$ such that for every $\varphi \in L^q(\mathcal{U})$ with $1/p + 1/q = 1$, $\int_{\mathcal{U}} f_j\varphi \to \int_{\mathcal{U}} f_\infty\varphi$.* □

REMARK A.3.5. One popular way of summarizing the Banach-Alaoglu theorem is the statement that "closed balls in L^p are weakly compact"; indeed, if $f_j \in L^p(\mathcal{U})$ satisfies the bound $\|f_j\|_{L^p} \leqslant C$, then the weak limit f_∞ provided by Theorem A.3.4 also satisfies $\|f_\infty\|_{L^p} \leqslant C$. The latter follows from the general fact that for any sequence $f_j \in L^p(\mathcal{U})$ converging weakly to some $f_\infty \in L^p(\mathcal{U})$,

$$\|f_\infty\|_{L^p(\mathcal{U})} \leqslant \liminf \|f_j\|_{L^p(\mathcal{U})}.$$

The proof of this is not hard; see e.g. [**LL01**, §2.11].

PROPOSITION A.3.6. *Suppose $\mathcal{V} \subset \mathcal{U} \subset \mathbb{R}^n$ are open subsets such that $\mathcal{V}$ has compact closure contained in $\mathcal{U}$, $1 < p < \infty$, f is a measurable function on $\mathcal{U}$ with $\|f\|_{L^p(\mathcal{V})} < \infty$, and there exist constants $C > 0$ and $\delta > 0$ such that*

$$\|D_i^h f\|_{L^p(\mathcal{V})} \leqslant C \qquad \text{whenever } 0 < |h| < \delta.$$

Then $f|_{\mathcal{V}}$ has a weak partial derivative $\partial_i f \in L^p(\mathcal{V})$ satisfying $\|\partial_i f\|_{L^p(\mathcal{V})} \leqslant C$.

PROOF. For any sequence $h_j \to 0$ of sufficiently small nonzero real numbers, the sequence $D_i^{h_j} f$ satisfies $\|D_i^{h_j} f\|_{L^p(\mathcal{V})} \leqslant C$, thus the Banach-Alaoglu theorem implies that after passing to a subsequence, one finds a function $g \in L^p(\mathcal{V})$ with $\|g\|_{L^p(\mathcal{V})} \leqslant C$ such that

$$\int_{\mathcal{V}} (D_i^{h_j} f)\varphi \to \int_{\mathcal{V}} g\varphi$$

for all $\varphi \in L^q(\mathcal{V})$, where $1/p + 1/q = 1$. In particular, this is true for all test functions $\varphi \in C_0^\infty(\mathcal{V})$, and in this case there is an "integration by parts" relation

$$\begin{aligned}\int_{\mathcal{V}} (D_i^{h_j} f)\varphi &= \int_{\mathcal{V}} \frac{f(x + h_j e_i) - f(x)}{h_j} \varphi(x)\, d\mu(x) \\ &= -\int_{\mathcal{V}} f(x) \frac{\varphi(x - h_j e_i) - \varphi(x)}{-h_j}\, d\mu(x) = -\int_{\mathcal{V}} f D_i^{-h_j} \varphi.\end{aligned}$$

By Lemma A.3.2, $D_i^{-h_j}\varphi \to \partial_i \varphi$ uniformly on $\mathcal{V}$ and thus also in $L^q(\mathcal{V})$, so taking the limit of the integrals, we've shown

$$\int_{\mathcal{V}} g\varphi = -\int_{\mathcal{V}} f\partial_i\varphi \qquad \text{for all } \varphi \in C_0^\infty(\mathcal{V}),$$

or in other words, $\partial_i f = g \in L^p(\mathcal{V})$. ☐

PROOF OF THEOREM A.3.1. The two statements in the theorem follow by applying Propositions A.3.3 and A.3.6 respectively to $\partial^\alpha f$ for every multiindex α with $|\alpha| \leqslant k-1$, using the fact that $D^h(\partial^\alpha f) = \partial^\alpha(D^h f)$. For the bound on $\|\nabla f\|_{W^{k-1,p}(\mathcal{V})}$, we observe that by assumption,

$$\|D^h f\|_{W^{k-1,p}(\mathcal{V})} = \sum_{|\alpha| \leqslant k-1} \|\partial^\alpha(D^h f)\|_{L^p(\mathcal{V})} = \sum_{|\alpha| \leqslant k-1} \|D^h(\partial^\alpha f)\|_{L^p(\mathcal{V})} \leqslant C,$$

thus each individual term in this sum satisfies $\|D^h(\partial^\alpha f)\|_{L^p(\mathcal{V})} \leqslant C$, implying $\|\nabla(\partial^\alpha f)\|_{L^p(\mathcal{V})} \leqslant C$ and thus

$$\begin{aligned}\|\nabla f\|_{W^{k-1,p}(\mathcal{V})} &= \sum_{|\alpha| \leqslant k-1} \|\partial^\alpha(\nabla f)\|_{L^p(\mathcal{V})} = \sum_{|\alpha| \leqslant k-1} \|\nabla(\partial^\alpha f)\|_{L^p(\mathcal{V})} \\ &\leqslant \sum_{|\alpha| \leqslant k-1} C =: m_{k,p} C.\end{aligned}$$

☐

A.4. Spaces of sections of vector bundles

In this section, fix a field

$$\mathbb{F} := \mathbb{R} \text{ or } \mathbb{C},$$

assume M is a smooth n-dimensional manifold, possibly with boundary, and $\pi : E \to M$ is a smooth vector bundle of rank m over $\mathbb{F}$. This comes with a "bundle atlas" $\mathcal{A}(\pi)$, a set whose elements $\alpha \in \mathcal{A}(\pi)$ each consist of the following data:

(1) An open subset $\mathcal{U}_\alpha \subset M$;
(2) A smooth local coordinate chart $\varphi_\alpha : \mathcal{U}_\alpha \xrightarrow{\cong} \Omega_\alpha$, where Ω_α is an open subset of $\mathbb{R}^n_+ := \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_n \geqslant 0\}$;
(3) A smooth local trivialization $\Phi_\alpha : E|_{\mathcal{U}_\alpha} \xrightarrow{\cong} \mathcal{U}_\alpha \times \mathbb{F}^m$.

Smoothness of φ_α and Φ_α means as usual that for every pair $\alpha, \beta \in \mathcal{A}(\pi)$, the coordinate transformations

$$\varphi_{\beta\alpha} := \varphi_\beta^{-1} \circ \varphi_\alpha : \Omega_{\alpha\beta} \xrightarrow{\cong} \Omega_{\beta\alpha}, \qquad \Omega_{\alpha\beta} := \varphi_\alpha(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$$

and transition maps

$$g_{\beta\alpha} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to \mathrm{GL}(m, \mathbb{F}) \quad \text{such that} \quad \Phi_\beta \circ \Phi_\alpha^{-1}(x, v) = (x, g_{\beta\alpha}(x)v) \quad \text{for } x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta,\ v \in \mathbb{F}^m$$

are smooth, and we shall assume the bundle atlas is maximal in the sense that any triple $(\mathcal{U}, \varphi, \Phi)$ that is smoothly compatible with every $\alpha \in \mathcal{A}(\pi)$ also belongs to $\mathcal{A}(\pi)$.

Any $\alpha \in \mathcal{A}(\pi)$ now associates to sections $\eta : M \to E$ their local coordinate representatives

$$\eta^\alpha := \mathrm{pr}_2 \circ \Phi_\alpha \circ \eta \circ \varphi_\alpha^{-1} : \Omega_\alpha \to \mathbb{F}^m,$$

where $\mathrm{pr}_2 : \mathcal{U}_\alpha \times \mathbb{F}^m \to \mathbb{F}^m$ is the projection, and the representatives with respect to two distinct $\alpha, \beta \in \mathcal{A}(\pi)$ are related by

$$\eta^\beta = (g_{\beta\alpha} \circ \varphi_\beta^{-1})(\eta^\alpha \circ \varphi_{\alpha\beta}) \quad \text{on } \Omega_{\beta\alpha} \subset \Omega_\beta.$$

For $p \in [1, \infty]$ and each integer $k \geqslant 0$, we then define the topological vector space of sections of class $W^{k,p}_{\mathrm{loc}}$ by

$$W^{k,p}_{\mathrm{loc}}(E) := \big\{\eta : M \to E \bigm| \text{sections such that } \eta^\alpha \in W^{k,p}_{\mathrm{loc}}(\mathring{\Omega}_\alpha, \mathbb{F}^m) \text{ for all } \alpha \in \mathcal{A}(\pi)\big\},$$

where convergence $\eta_j \to \eta$ in $W^{k,p}_{\mathrm{loc}}(E)$ means that $\eta_j^\alpha \to \eta^\alpha$ in $W^{k,p}_{\mathrm{loc}}(\mathring{\Omega}_\alpha, \mathbb{F}^m)$ for all $\alpha \in \mathcal{A}(\pi)$. Note that Ω_α is not necessarily an open subset of $\mathbb{R}^n$ since it may contain points in $\partial\mathbb{R}^n_+ = \mathbb{R}^{n-1} \times \{0\}$, but its interior $\mathring{\Omega}_\alpha$ is open in $\mathbb{R}^n$, and $W^{k,p}_{\mathrm{loc}}(\mathring{\Omega}_\alpha)$ is thus defined as in §A.1. Strictly speaking, elements of $\eta \in W^{k,p}_{\mathrm{loc}}(E)$ are not sections but *equivalence classes* of sections defined almost everywhere—the latter notion is defined with respect to any measure arising from a smooth volume element on M, and it does not depend on this choice.

It turns out that $W^{k,p}_{\mathrm{loc}}(E)$ can be given the structure of a Banach space if M is compact. This follows from the fact that M can then be covered by a finite subset of the atlas $\mathcal{A}(\pi)$, but we must be a little bit careful: not all charts in $\mathcal{A}(\pi)$ are equally suitable for defining $W^{k,p}$-norms on sections, because e.g. even a nice smooth section $\eta \in \Gamma(E)$ may have $\|\eta^\alpha\|_{W^{k,p}(\mathring{\Omega}_\alpha)} = \infty$ if $\Omega_\alpha \subset \mathbb{R}^n_+$ is unbounded. One way to deal with this is as follows: we will say that $\alpha \in \mathcal{A}(\pi)$ is a **precompact chart** if there exists $\alpha' \in \mathcal{A}(\pi)$ and a compact subset $\mathcal{K} \subset M$ such that

$$\mathcal{U}_\alpha \subset \mathcal{K} \subset \mathcal{U}_{\alpha'}.$$

When this is the case, $\Omega_\alpha \subset \mathbb{R}^n_+$ is necessarily bounded, and the transition maps between two precompact charts necessarily have bounded derivatives of all orders, as they are restrictions to precompact subsets of maps that are smooth on larger domains. If M is compact, then one can always find a finite subset $I \subset \mathcal{A}(\pi)$ consisting of precompact charts such that $M = \bigcup_{\alpha \in I} \mathcal{U}_\alpha$.

DEFINITION A.4.1. Suppose $E \to M$ is a smooth vector bundle over a compact manifold M, and $I \subset \mathcal{A}(\pi)$ is a finite set of precompact charts such that $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ is an open cover of M. We then define $W^{k,p}(E)$ as the vector space of all sections $\eta : M \to E$ for which the norm

$$\|\eta\|_{W^{k,p}} := \|\eta\|_{W^{k,p}(E)} := \sum_{\alpha \in I} \|\eta^\alpha\|_{W^{k,p}(\mathring{\Omega}_\alpha)}$$

is finite.

The norm in the above definition depends on auxiliary choices, but it is easy to see that the resulting definition of the space $W^{k,p}(E)$ and its topology do not. In fact:

PROPOSITION A.4.2. *If M is compact, then $W^{k,p}(E) = W^{k,p}_{\mathrm{loc}}(E)$, and a sequence η_j converges to η in $W^{k,p}_{\mathrm{loc}}(E)$ if and only if the norm given in Definition A.4.1 satisfies $\|\eta_j - \eta\|_{W^{k,p}(E)} \to 0$.*

The proposition is an immediate consequence of the following.

LEMMA A.4.3. *Suppose M is a smooth manifold, $\pi : E \to M$ is a smooth vector bundle, $\{\beta\} \cup J \subset \mathcal{A}(\pi)$ is a finite collection of charts such that $M = \bigcup_{\alpha \in J} \mathcal{U}_\alpha$ and all coordinate transformations and transition maps relating any two charts in the collection $\{\beta\} \cup J$ have bounded derivatives of all orders (e.g. it suffices to assume all are precompact). Then there exists a constant $c > 0$ such that*

$$\|\eta^\beta\|_{W^{k,p}(\mathring{\Omega}_\beta)} \leqslant c \sum_{\alpha \in J} \|\eta^\alpha\|_{W^{k,p}(\mathring{\Omega}_\alpha)}$$

for all sections $\eta : M \to E$ with $\eta^\alpha \in W^{k,p}(\mathring{\Omega}_\alpha)$ for every $\alpha \in J$.

PROOF. Choose a partition of unity $\{\rho_\alpha : M \to [0,1]\}_{\alpha \in J}$ subordinate to the finite open cover $\{\mathcal{U}_\alpha\}_{\alpha \in J}$. Now $\eta = \sum_{\alpha \in J} \rho_\alpha \eta$, and each $\rho_\alpha \eta$ is supported in $\mathcal{U}_\alpha$, so $(\rho_\alpha \eta)^\beta$ has support in $\Omega_{\beta\alpha} = \varphi_\beta(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$. Thus using Theorem A.2.8 with the fact that $g_{\beta\alpha}$, φ_β^{-1}, $\varphi_{\alpha\beta}$ and $\varphi_{\beta\alpha} = \varphi_{\alpha\beta}^{-1}$ are all smooth functions with bounded derivatives of all orders on the domains in question, we find

$$\begin{aligned}\|\eta^\beta\|_{W^{k,p}(\mathring{\Omega}_\beta)} &= \Big\|\sum_{\alpha\in J}(\rho_\alpha\eta)^\beta\Big\|_{W^{k,p}(\mathring{\Omega}_\beta)} \leqslant \sum_{\alpha\in J}\|(\rho_\alpha\eta)^\beta\|_{W^{k,p}(\mathring{\Omega}_{\beta\alpha})}\\ &= \sum_{\alpha\in J}\|(\rho_\alpha\circ\varphi_\beta^{-1})(g_{\beta\alpha}\circ\varphi_\beta^{-1})(\eta^\alpha\circ\varphi_{\alpha\beta})\|_{W^{k,p}(\mathring{\Omega}_{\beta\alpha})}\\ &\leqslant c\sum_{\alpha\in J}\|\eta^\alpha\|_{W^{k,p}(\mathring{\Omega}_{\alpha\beta})} \leqslant c\sum_{\alpha\in J}\|\eta^\alpha\|_{W^{k,p}(\mathring{\Omega}_\alpha)}.\end{aligned}$$

□

COROLLARY A.4.4. *If M is compact, then the norm on $W^{k,p}(E)$ given by Definition A.4.1 is independent of all auxiliary choices up to equivalence of norms.* □

THEOREM A.4.5. *For any smooth vector bundle $\pi : E \to M$ over a compact manifold M, $W^{k,p}(E)$ is a Banach space.*

PROOF. If $\eta_j \in W^{k,p}(E)$ is a Cauchy sequence, then for some chosen finite collection $I \subset \mathcal{A}(\pi)$ of precompact charts covering M, the sequences η_j^α for $\alpha \in I$ are Cauchy in $W^{k,p}(\mathring{\Omega}_\alpha)$ and thus have limits $\xi^{(\alpha)} \in W^{k,p}(\mathring{\Omega}_\alpha, \mathbb{F}^m)$. Choosing a partition of unity $\{\rho_\alpha : M \to [0,1]\}_{\alpha\in I}$ subordinate to $\{\mathcal{U}_\alpha\}_{\alpha\in I}$, we can now associate to each $\alpha \in I$ a section $\eta_{\infty,\alpha} \in W^{k,p}(E)$ characterized uniquely by the condition that it vanishes outside of $\mathcal{U}_\alpha$ and is represented in the trivialization on $\mathcal{U}_\alpha$ by

$$\eta^\alpha_{\infty,\alpha} = (\rho_\alpha \circ \varphi_\alpha^{-1})\xi^{(\alpha)}.$$

We claim that $\rho_\alpha\eta_j \to \eta_{\infty,\alpha}$ in $W^{k,p}(E)$ for each $\alpha \in I$. Indeed, we have

$$(\rho_\alpha\eta_j)^\alpha = (\rho_\alpha\circ\varphi_\alpha^{-1})\eta_j^\alpha \to (\rho_\alpha\circ\varphi_\alpha^{-1})\xi^{(\alpha)} = \eta^\alpha_{\infty,\alpha} \quad \text{in } W^{k,p}(\mathring{\Omega}_\alpha)$$

since $\eta_j^\alpha \to \xi^{(\alpha)}$. For all other $\beta \in I$ not equal to α, $(\rho_\alpha\eta_j)^\beta - \eta^\beta_{\infty,\alpha} \in W^{k,p}(\mathring{\Omega}_\beta, \mathbb{F}^m)$ has support in $\Omega_{\beta\alpha} = \varphi_\beta(\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$, thus

$$\|(\rho_\alpha\eta_j)^\beta - \eta^\beta_{\infty,\alpha}\|_{W^{k,p}(\mathring{\Omega}_\beta)} = \|(\rho_\alpha\eta_j)^\beta - \eta^\beta_{\infty,\alpha}\|_{W^{k,p}(\mathring{\Omega}_{\beta\alpha})} \leqslant c\|(\rho_\alpha\eta_j)^\alpha - \eta^\alpha_{\infty,\alpha}\|_{W^{k,p}(\mathring{\Omega}_\alpha)},$$

where the inequality comes from Lemma A.4.3 after replacing M with $\mathcal{U}_\alpha$, and $\mathcal{U}_\beta$ with $\mathcal{U}_\beta \cap \mathcal{U}_\alpha$ (note that the lemma does not require M to be compact). With the claim established, we have

$$\eta_j = \sum_{\alpha\in I}\rho_\alpha\eta_j \to \sum_{\alpha\in I}\eta_{\infty,\alpha} \quad \text{in } W^{k,p}(E).$$

□

REMARK A.4.6. One can use exactly the same approach to show that when M is compact, the space $C^k(E)$ of C^k-smooth sections $\eta : M \to E$ has a canonical (up to equivalence of norms) Banach space structure for each finite integer $k \geqslant 0$ such that convergence in the C^k-norm is equivalent to uniform convergence of all derivatives up to order k.

EXERCISE A.4.7. For $\mathcal{U} \subset \mathbb{R}^n$ an open subset, the space $W^{k,p}_{\mathrm{loc}}(\mathcal{U})$ was defined in §A.1, but one can give it an alternative definition in the present context by viewing functions on $\mathcal{U}$ as sections of a trivial vector bundle over $\mathcal{U}$, with the latter viewed as a noncompact smooth n-manifold. Show that these two definitions of $W^{k,p}_{\mathrm{loc}}(\mathcal{U})$ are equivalent.

EXERCISE A.4.8. Suppose $\mathcal{U} \subset \mathbb{R}^n$ is a bounded open subset with smooth boundary, so its closure $\overline{\mathcal{U}} \subset \mathbb{R}^n$ is a smooth compact submanifold with boundary, and let $E \to \overline{\mathcal{U}}$ be a trivial vector bundle. Show that there is a canonical Banach space isomorphism between $W^{k,p}(\mathcal{U})$ as defined in §A.1 and $W^{k,p}(E)$ as defined in the present section. *Hint: Recall that sections in $W^{k,p}(E)$ are only required to be defined almost everywhere, so in particular if the domain M is a manifold with boundary, they need not be well defined on ∂M.*

In light of Exercise A.4.8, the natural generalization of $W^{k,p}_0(\mathcal{U})$ in the present setting is

$$W^{k,p}_0(E) := \overline{C^\infty_0(E|_{M \setminus \partial M})},$$

i.e. it is the closure in the $W^{k,p}$-norm of the space of smooth sections that vanish near the boundary. Density of smooth sections will imply that this is the same as $W^{k,p}(E)$ if M is closed, but in general $W^{k,p}_0(E)$ is a closed subspace of $W^{k,p}(E)$.

The partition of unity argument in Theorem A.4.5 contains all the essential ideas needed to generalize results about Sobolev spaces on domains in $\mathbb{R}^n$ to compact manifolds. We now state the essential results, leaving the proofs as exercises.

THEOREM A.4.9. *Assume M is a smooth compact n-manifold, possibly with boundary, $\pi : E \to M$ is a smooth vector bundle of finite rank, $k \geqslant 0$ is an integer and $1 \leqslant p < \infty$. Then the Banach space $W^{k,p}(E)$ has the following properties.*

(1) The space $\Gamma(E)$ of smooth sections is dense in $W^{k,p}(E)$.

(2) If $kp > n$, then for each integer $d \geqslant 0$, there exists a continuous and compact inclusion

$$W^{k+d,p}(E) \hookrightarrow C^d(E).$$

(3) The natural inclusion

$$W^{k+1,p}(E) \hookrightarrow W^{k,p}(E)$$

is compact.

(4) Suppose $F, G \to M$ are smooth vector bundles such that there exists a smooth bundle map

$$E \otimes F \to G : \eta \otimes \xi \mapsto \eta \cdot \xi.$$

Then if $kp > n$ and $0 \leqslant m \leqslant k$, there exists a continuous product pairing

$$W^{k,p}(E) \times W^{m,p}(F) \to W^{m,p}(G) : (\eta, \xi) \mapsto \eta \cdot \xi.$$

In particular, products of $W^{k,p}$ sections give $W^{k,p}$ sections whenever $kp > n$.

(5) Suppose $F \to M$ is another smooth vector bundle, $\mathcal{V} \subset E$ is an open subset that intersects every fiber of E, and we consider the spaces

$$W^{k,p}(\mathcal{V}) := \{\eta \in W^{k,p}(E) \mid \eta(M) \subset \mathcal{V}\}$$

and

$$C^k_M(\mathcal{V}, F) := \left\{\Phi : \mathcal{V} \to F \mid \textit{fiber-preserving maps of class } C^k\right\},$$

where the latter is assigned the topology of C^k*-convergence on compact subsets. If* $kp > n$*, then* $W^{k,p}(\mathcal{V})$ *is an open subset of* $W^{k,p}(E)$*, and the map*

$$C^k_M(\mathcal{V}, F) \times W^{k,p}(\mathcal{V}) \to W^{k,p}(F) : (\Phi, \eta) \mapsto \Phi \circ \eta$$

is well defined and continuous.

(6) If N *is another smooth compact manifold and* $\varphi : N \to M$ *is a smooth diffeomorphism, then there is a Banach space isomorphism*

$$W^{k,p}(E) \to W^{k,p}(\varphi^* E) : \eta \mapsto \eta \circ \varphi.$$

□

Remark A.4.10. It is sometimes useful to extend the definitions and results of this section to vector bundles that are not smooth, e.g. vector bundles of class C^k or $W^{k,p}$, for which all transition maps are required to be of class C^k or $W^{k,p}$ respectively. The latter makes sense in general only if $kp > n$, so that transition maps are at least continuous. Given a bundle of this type, one can enhance the arguments of this section with the aid of Theorem A.2.1 to show that $W^{m,p}(E)$ is a well-defined Banach space for every $m \leqslant k$, though it would not be well defined if $m > k$. Such spaces arise frequently in global analysis, e.g. if f is a non-smooth element in the Banach manifold $\mathcal{B}$ of $W^{k,p}$-smooth maps of M into another manifold N, then $f^*TN \to M$ is in general a vector bundle of class $W^{k,p}$, and $T_f\mathcal{B} = W^{k,p}(f^*TN)$.

A.5. Some remarks on domains with cylindrical ends

For bundles $\pi : E \to M$ with M noncompact, $W^{k,p}(E)$ is not generally well defined without making additional choices. When $M = \dot{\Sigma} = \Sigma \backslash \Gamma$ is a punctured Riemann surface and $\pi : E \to \dot{\Sigma}$ is equipped with an asymptotically Hermitian structure $\{(E_z, J_z, \omega_z)\}_{z\in\Gamma}$ as defined in Chapter 4, one nice way to define $W^{k,p}(E)$ was introduced in §4.1: one takes it to be the space of sections in $W^{k,p}_{\text{loc}}(E)$ whose $W^{k,p}$-norms on each cylindrical end are finite with respect to a choice of asymptotic trivialization. This definition requires the convenient fact that complex vector bundles over S^1 are always trivial, though one can also do without this by using the ideas in the previous section. Indeed, any collection of local trivializations on the asymptotic bundle $E_z \to S^1$ covering S^1 gives rise via the asymptotically Hermitian structure to a collection of trivializations on E covering the corresponding cylindrical end $\dot{\mathcal{U}}_z$. The key fact is then that S^1 is compact, hence one can always choose such a covering to be finite: combining this with a finite covering of $\dot{\Sigma}$ in the complement of its cylindrical ends by precompact charts, we obtain a covering of $\dot{\Sigma}$ by a finite collection of bundle charts that are not all precompact, but nonetheless have the property that all transition maps have bounded derivatives of all orders. This is enough to define a $W^{k,p}$-norm for sections of $E \to \dot{\Sigma}$ as in Definition A.4.1 and to prove that it does not depend on the choices of charts or local trivializations, though it does depend on the asymptotically Hermitian structure.

With this definition understood, one can easily generalize the Sobolev embedding theorem and other important statements in Theorem A.4.9 to the setting of an asymptotically Hermitian bundle over a punctured Riemann surface. We shall leave the details of this generalization as an exercise, but take the opportunity to point out a few important differences from the compact case.

First, since $\dot{\Sigma}$ is not compact, neither are the inclusions

$$W^{k+d,p}(E) \hookrightarrow C^d(E), \qquad W^{k+1,p}(E) \hookrightarrow W^{k,p}(E).$$

The proof of compactness fails due to the fact that cylindrical ends require local trivializations over unbounded domains of the form $(0,\infty) \times (0,1) \subset \mathbb{R}^2$, for which Theorem A.1.10 does not hold. And indeed, considering unbounded shifts on the infinite cylinder $\dot{\Sigma} = \mathbb{R} \times S^1$, it is easy to find a sequence of $W^{k,p}$-bounded functions with $kp > 2$ that do not have a C^0-convergent subsequence. That is the bad news.

The good news is that if $\eta \in W^{k+d,p}(E)$ for $kp > 2$, then one can say considerably more about η than just that it is C^d-smooth. Indeed, restricting to one of the cylindrical ends $[0,\infty) \times S^1 \subset \dot{\Sigma}$, notice that the finiteness of the $W^{k+d,p}$-norm over $\dot{\Sigma}$ implies

$$\|\eta\|_{W^{k+d,p}((R,\infty)\times S^1)} \to 0 \quad \text{as} \quad R \to \infty.$$

Since these domains are all naturally diffeomorphic for different values of R, the C^d-norm of η over $(R,\infty) \times S^1$ is bounded by the $W^{k+d,p}$-norm via a constant that does not depend on R, so this implies an asymptotic decay condition

$$\|\eta\|_{C^d([R,\infty)\times S^1)} \to 0 \quad \text{as} \quad R \to \infty$$

for every $\eta \in W^{k+d,p}(E)$.

Here is another useful piece of good news: since $\dot{\Sigma}$ does not have boundary, $W^{k,p}(E) = W_0^{k,p}(E)$.

THEOREM A.5.1. *Given an asymptotically Hermitian bundle E over a punctured Riemann surface $\dot{\Sigma}$, the space $C_0^\infty(E)$ of smooth sections with compact support is dense in $W^{k,p}(E)$ for all $k \geqslant 0$ and $1 \leqslant p < \infty$.*

PROOF. We can assume as in Definition A.4.1 that the $W^{k,p}$-norm for sections η of E is given by

$$\|\eta\|_{W^{k,p}} = \sum_{\alpha \in I} \|\eta^\alpha\|_{W^{k,p}(\Omega_\alpha)},$$

where $I \subset \mathcal{A}(\pi)$ is a finite collection of bundle charts

$$\alpha = \left(\varphi_\alpha : \mathcal{U}_\alpha \xrightarrow{\cong} \Omega_\alpha\,,\ \Phi_\alpha : E|_{\mathcal{U}_\alpha} \xrightarrow{\cong} \mathcal{U}_\alpha \times \mathbb{C}^n\right)$$

such that each of the open sets $\Omega_\alpha \subset \mathbb{C}$ is either bounded or (for charts over the cylindrical ends) of the form

$$\Omega_\alpha = (0,\infty) \times \omega_\alpha \subset \mathbb{R}^2 = \mathbb{C}$$

for some bounded open subset $\omega_\alpha \subset \mathbb{R}$. Now given $\eta \in W^{k,p}(E)$, Theorem A.1.1 provides for each $\alpha \in I$ a sequence $\eta_j^\alpha \in W^{k,p}(\Omega_\alpha)$ of smooth functions with bounded support such that $\eta_j^\alpha \to \eta^\alpha$ in $W^{k,p}(\Omega_\alpha)$. Choose a partition of unity $\{\rho_\alpha : \dot{\Sigma} \to [0,1]\}_{\alpha \in I}$ subordinate to the open cover $\{\mathcal{U}_\alpha\}_{\alpha \in I}$ and let

$$\eta_j := \sum_{\alpha \in I} \rho_\alpha(\eta_j^\alpha \circ \varphi_\alpha) \in W^{k,p}(E).$$

These sections are smooth and have compact support since the η_j^α have bounded support in Ω_α, and they converge in $W^{k,p}$ to η. $\square$

APPENDIX B

The Floer C_ϵ space

The C_ϵ-topology for functions was introduced by Floer [**Flo88b**] to provide a Banach manifold of perturbed geometric structures without departing from the smooth category: it is a way to circumvent the annoying fact that spaces of smooth functions which arise naturally in geometric settings are not Banach spaces. The construction of C_ϵ spaces generally depends on several arbitrary choices and is thus far from canonical, but this detail is unimportant since the C_ϵ space itself is never the main object of interest. What is important is merely the properties that it has, namely that it not only embeds continuously into C^∞ and contains an abundance of nontrivial functions, but also is a separable Banach space and can therefore be used in the Sard-Smale theorem for genericity arguments. We shall prove these facts in this appendix.

Fix a smooth finite-rank vector bundle $\pi : E \to M$ over a finite-dimensional compact manifold M, possibly with boundary. For each integer $k \geqslant 0$, we denote by $C^k(E)$ the Banach space of C^k-smooth sections of E; note that the norm on $C^k(E)$ depends on various auxiliary choices but is well defined up to equivalence of norms since M is compact. Now if $\boldsymbol{\epsilon} = (\epsilon_k)_{k=0}^\infty$ is a sequence of positive numbers with $\epsilon_k \to 0$, set

$$C_{\boldsymbol{\epsilon}}(E) = \left\{ \eta \in \Gamma(E) \;\middle|\; \|\eta\|_{C_{\boldsymbol{\epsilon}}} < \infty \right\},$$

where the $C_{\boldsymbol{\epsilon}}$-norm is defined by

$$\|\eta\|_{C_{\boldsymbol{\epsilon}}} = \sum_{k=0}^{\infty} \epsilon_k \|\eta\|_{C^k}. \tag{B.1}$$

The norm for $C_{\boldsymbol{\epsilon}}(E)$ is somewhat more delicate than for $C^k(E)$, e.g. its equivalence class is not obviously independent of auxiliary choices. This remark is meant as a sanity check, but it should not cause extra concern since, in practice, the space $C_{\boldsymbol{\epsilon}}(E)$ is typically regarded as an auxiliary choice in itself. In many applications, one fixes an open subset $\mathcal{U} \subset M$ and considers the closed subspace

$$C_{\boldsymbol{\epsilon}}(E;\mathcal{U}) = \left\{ \eta \in C_{\boldsymbol{\epsilon}}(E) \;\middle|\; \eta|_{M \setminus \mathcal{U}} \equiv 0 \right\}.$$

Remark B.0.1. The requirement for M to be compact can be relaxed as long as $\mathcal{U} \subset M$ has compact closure: e.g. in one situation of frequent interest in this book, we take M to be the noncompact completion of a symplectic cobordism. In this case $C_{\boldsymbol{\epsilon}}(E;\mathcal{U})$ can be defined as a closed subspace of $C_{\boldsymbol{\epsilon}}(E|_{M_0})$ where $M_0 \subset M$ is any compact manifold with boundary that contains the closure of $\mathcal{U}$. For this reason, we lose no generality in continuing under the assumption that M is compact.

In order to prove things about $C_\epsilon(E)$, we will need to specify a more precise definition of the C^k-norms. To this end, define a sequence of vector bundles $E^{(k)} \to M$ for integers $k \geqslant 0$ inductively by

$$E^{(0)} := E, \qquad E^{(k+1)} := \mathrm{Hom}(TM, E^{(k)}).$$

Choose connections and bundle metrics on both TM and E; these induce connections and bundle metrics on each of the $E^{(k)}$, so that for any section $\xi \in \Gamma(E^{(k)})$, the covariant derivative $\nabla\xi$ is now a section of $E^{(k+1)}$. In particular for $\eta \in \Gamma(E)$, we can define the "kth covariant derivative" of η as a section

$$\nabla^k \eta \in \Gamma(E^{(k)}).$$

Using the bundle metrics to define C^0-norms for sections of $E^{(k)}$, we can then define

$$\|\eta\|_{C^k(E)} = \sum_{m=0}^{k} \|\nabla^m \eta\|_{C^0(E^{(m)})},$$

where by convention $\nabla^0 \eta := \eta$. We will assume throughout the following that the C^k-norms appearing in (B.1) are defined in this way.

THEOREM B.0.2. *$C_\epsilon(E)$ is a Banach space.*

PROOF. We need to show that C_ϵ-Cauchy sequences converge in the C_ϵ-norm. It is clear from the definitions that if $\eta_j \in C_\epsilon(E)$ is Cauchy, then η_j is also C^k-Cauchy for every $k \geqslant 0$, hence its derivatives $\nabla^k \eta_j$ for every k are C^0-convergent to continuous sections ξ^k of $E^{(k)}$. This convergence implies that $\xi^{k+1} = \nabla \xi^k$ in the sense of distributions, hence by the equivalence of classical and distributional derivatives (see e.g. [**LL01**, §6.10]), $\eta_\infty := \xi^0$ is smooth with $\nabla^k \eta_\infty = \xi^k$, so that $\nabla^k \eta_j \to \nabla^k \eta_\infty$ in $C^0(E^{(k)})$ for all k.

We claim $\eta_\infty \in C_\epsilon(E)$. Choose $N > 0$ such that $\|\eta_i - \eta_j\|_{C_\epsilon} < 1$ for all $i, j \geqslant N$. Then for every $m \in \mathbb{N}$ and every $i \geqslant N$,

$$\begin{aligned}\sum_{k=0}^{m} \epsilon_k \|\eta_i\|_{C^k} &\leqslant \sum_{k=0}^{m} \epsilon_k \|\eta_i - \eta_N\|_{C^k} + \sum_{k=0}^{m} \epsilon_k \|\eta_N\|_{C^k} \\ &\leqslant \|\eta_i - \eta_N\|_{C_\epsilon} + \|\eta_N\|_{C_\epsilon} < 1 + \|\eta_N\|_{C_\epsilon}.\end{aligned}$$

Fixing m and letting $i \to \infty$, we then have

$$\sum_{k=0}^{m} \epsilon_k \|\eta_\infty\|_{C^k} \leqslant 1 + \|\eta_N\|_{C_\epsilon}$$

for all m, so we can now let $m \to \infty$ and conclude $\|\eta_\infty\|_{C_\epsilon} \leqslant 1 + \|\eta_N\|_{C_\epsilon} < \infty$.

The argument that $\|\eta_j - \eta_\infty\|_{C_\epsilon} \to 0$ as $j \to \infty$ is similar: pick $\epsilon > 0$ and N such that $\|\eta_i - \eta_j\|_{C_\epsilon} < \epsilon$ for all $i, j \geqslant N$. Then for a fixed $m \in \mathbb{N}$, we can let $i \to \infty$ in the expression $\sum_{k=0}^{m} \epsilon_k \|\eta_i - \eta_j\|_{C^k} < \epsilon$, giving

$$\sum_{k=0}^{m} \epsilon_k \|\eta_\infty - \eta_j\|_{C^k} \leqslant \epsilon.$$

This is true for every m, so we can take $m \to \infty$ and conclude $\|\eta_\infty - \eta_j\|_{C_\epsilon} \leqslant \epsilon$ for all $j \geqslant N$. □

To show that $C_\epsilon(E)$ is also separable, we will follow a hint[1] from [HS95] and embed it isometrically into another Banach space that can be more easily shown to be separable. For each integer $k \geqslant 0$, define the vector bundle

$$F^{(k)} = E^{(0)} \oplus \ldots \oplus E^{(k)},$$

and let X_ϵ denote the vector space of all sequences

$$\xi := (\xi^0, \xi^1, \xi^2, \ldots) \in \prod_{k=0}^{\infty} C^0(F^{(k)})$$

such that

$$\|\xi\|_{X_\epsilon} := \sum_{k=0}^{\infty} \epsilon_k \|\xi^k\|_{C^0} < \infty.$$

EXERCISE B.0.3. Adapt the proof of Theorem B.0.2 to show that X_ϵ is also a Banach space.

LEMMA B.0.4. *X_ϵ is separable.*

PROOF. Since $C^0(F^{(k)})$ is separable for each $k \geqslant 0$, we can fix countable dense subsets $P^k \subset C^0(F^{(k)})$. The set

$$P := \left\{ (\xi^0, \ldots, \xi^N, 0, 0, \ldots) \in X_\epsilon \;\middle|\; N \geqslant 0 \text{ and } \xi^k \in P^k \text{ for all } k = 0, \ldots, N \right\}$$

is then countable and dense in X_ϵ. □

THEOREM B.0.5. *$C_\epsilon(E)$ is separable.*

PROOF. Consider the injective linear map

$$C_\epsilon(E) \hookrightarrow X_\epsilon : \eta \mapsto \left(\eta, (\eta, \nabla\eta), (\eta, \nabla\eta, \nabla^2\eta), \ldots \right).$$

This is an isometric embedding and thus presents $C_\epsilon(E)$ as a closed linear subspace of X_ϵ, hence the theorem follows from Lemma B.0.4 and the fact that subspaces of separable metric spaces are always separable. □

Note that given any open subset $\mathcal{U} \subset M$, Theorems B.0.2 and B.0.5 also hold for $C_\epsilon(E; \mathcal{U})$, as a closed subspace of $C_\epsilon(E)$. So far in this discussion, however, there has been no guarantee that $C_\epsilon(E)$ or $C_\epsilon(E; \mathcal{U})$ contains anything other than the zero-section, though it is clear that in theory, one should always be able to enlarge the space by choosing new sequences ϵ_k that converge to zero faster. The following result says that $C_\epsilon(E; \mathcal{U})$ can always be made large enough to be useful in applications.

THEOREM B.0.6. *Given an open subset $\mathcal{U} \subset M$, the sequence ϵ_k can be chosen to have the following properties:*

(1) $C_\epsilon(E; \mathcal{U})$ is dense in the space of continuous sections vanishing outside $\mathcal{U}$.

(2) Given any point $p \in \mathcal{U}$, a neighborhood $\mathcal{N}_p \subset \mathcal{U}$ of p, a number $\delta > 0$ and a continuous section η_0 of E, there exists a section $\eta \in \Gamma(E)$ and a smooth compactly supported function $\beta : \mathcal{N}_p \to [0, 1]$ such that

$$\beta\eta \in C_\epsilon(E; \mathcal{U}), \qquad \beta(p)\eta(p) = \eta_0(p), \qquad \text{and} \qquad \|\eta - \eta_0\|_{C^0} < \delta.$$

[1]Thanks to Sam Lisi for explaining to me what the hint in [HS95] was referring to.

PROOF. Note first that it suffices to find two separate sequences ϵ_k and ϵ'_k that have the first and second property respectively, as the sequence of minima $\min(\epsilon_k, \epsilon'_k)$ will then have both properties.

The following construction for the first property is based on a suggestion by Barney Bramham. Observe first that the space $C^0(E;\mathcal{U})$ of continuous sections vanishing outside $\mathcal{U}$ is a closed subspace of $C^0(E)$ and is thus separable, so we can choose a countable C^0-dense subset $P \subset C^0(E;\mathcal{U})$. Moreover, the space of *smooth* sections vanishing outside $\mathcal{U}$ is dense in $C^0(E;\mathcal{U})$, hence we can assume without loss of generality that the sections in P are smooth. Now write $P = \{\eta_1, \eta_2, \eta_3, \dots\}$ and define $\epsilon_k > 0$ for every integer $k \geqslant 0$ to have the property

$$\epsilon_k < \frac{1}{2^k} \min\left\{ \frac{1}{\|\eta_1\|_{C^k}}, \dots, \frac{1}{\|\eta_k\|_{C^k}} \right\}.$$

Then every η_j is in $C_{\boldsymbol{\epsilon}}(E;\mathcal{U})$, as

$$\|\eta_j\|_{C_{\boldsymbol{\epsilon}}} < \sum_{k=0}^{j-1} \epsilon_k \|\eta_j\|_{C^k} + \sum_{k=j}^{\infty} \frac{1}{2^k} < \infty.$$

The second property is essentially local, so it can be deduced from Lemma B.0.7 below. $\square$

LEMMA B.0.7. *Suppose $\beta : \mathring{\mathbb{D}}^n \to [0,1]$ is a smooth function with compact support on the open unit ball $\mathring{\mathbb{D}}^n \subset \mathbb{R}^n$ and $\beta(0) = 1$. One can choose a sequence of positive numbers $\epsilon_k \to 0$ such that for every $\eta_0 \in \mathbb{R}^m$ and $r > 0$, the function $\eta : \mathbb{R}^n \to \mathbb{R}^m$ defined by*

$$\eta(p) := \beta(p/r)\eta_0$$

satisfies $\sum_{k=0}^{\infty} \epsilon_k \|\eta\|_{C^k} < \infty$.

PROOF. Define $\epsilon_k > 0$ so that for $k \geqslant 1$,

$$\epsilon_k = \frac{1}{k^k \|\beta\|_{C^k}}.$$

Then

$$\sum_{k=1}^{\infty} \epsilon_k \|\eta\|_{C^k} \leqslant \sum_{k=1}^{\infty} \frac{1}{k^k \|\beta\|_{C^k}} \frac{\|\beta\|_{C^k}}{r^k} = \sum_{k=1}^{\infty} \left(\frac{1/r}{k} \right)^k < \infty.$$

$\square$

APPENDIX C

Genericity in the space of asymptotic operators

The purpose of this appendix is to prove Lemma 3.5.17, which was needed for our definition of spectral flow in §3.5. The proof combines some ideas from that section with the technique used in Chapter 9 to prove generic transversality of moduli spaces via the Sard-Smale theorem. Some knowledge of that technique should thus be considered a prerequisite for this appendix; if you have never seen it before and were directed here after reading the statement of Lemma 3.5.17, you might want to skip this for now and come back after you've read as far as Chapter 9.

Recalling the notation from Chapter 3, we fix the real Hilbert spaces

$$\mathcal{H} = L^2(S^1, \mathbb{R}^{2n}), \qquad \mathcal{D} = H^1(S^1, \mathbb{R}^{2n}),$$

the symmetric index 0 Fredholm operator

$$\mathbf{T}_{\mathrm{ref}} = -J_0\,\partial_t : \mathcal{D} \to \mathcal{H}$$

and, given a bounded family of symmetric matrices $S \in L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$, refer to any operator of the form

$$\mathbf{A} = -J_0\,\partial_t - S : \mathcal{D} \to \mathcal{H}$$

as an **asymptotic operator**. Such operators belong to the space of symmetric compact perturbations of $\mathbf{T}_{\mathrm{ref}}$,

$$\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) = \left\{ \mathbf{T}_{\mathrm{ref}} + \mathbf{K} : \mathcal{D} \to \mathcal{H} \;\middle|\; \mathbf{K} \in \mathscr{L}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{H}) \right\},$$

which we regard as a smooth Banach manifold via its obvious identification with the space $\mathscr{L}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{H})$ of symmetric bounded linear operators on $\mathcal{H}$. For $k \in \mathbb{N}$, we denote by

$$\mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{R}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) \subset \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$$

the finite-codimensional submanifold determined by the condition $\dim_{\mathbb{R}} \ker \mathbf{A} = \dim_{\mathbb{R}} \operatorname{coker} \mathbf{A} = k$.

Here is the statement of Lemma 3.5.17 again.

Lemma. *Fix a smooth path $[-1,1] \to L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})) : s \mapsto S_s$ and consider the 1-parameter family of symmetric index 0 Fredholm operators*

$$\mathbf{A}_s := -J_0\,\partial_t - S_s : H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n})$$

for $s \in [-1,1]$, assuming $\mathbf{A}_{\pm 1}$ are isomorphisms. Then after replacing S_s by a family of the form $\widetilde{S}_s(t) := S_s(t) + B(s,t)$ for some smooth function $B : [-1,1] \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ that vanishes for $s = \pm 1$ and may be assumed arbitrarily C^∞-small, one can arrange that the following conditions hold:

(1) *For each* $s \in (-1,1)$, *all eigenvalues of* $\mathbf{A}_s$ *are simple.*
(2) *All intersections of the smooth path*

$$(-1,1) \to \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}}) : s \mapsto \mathbf{A}_s$$

with $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{R}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ *are transverse.*

We shall now prove this by constructing a Floer-type space of $C_{\boldsymbol{\epsilon}}$-smooth (see Appendix B) perturbed families of asymptotic operators, and using the Sard-Smale theorem to find a countable collection of comeager subsets whose intersection contains perturbations achieving the desired conditions.

Choose a sequence of positive numbers $\boldsymbol{\epsilon} = (\epsilon_k)_{k=0}^{\infty}$ with $\epsilon_k \to 0$ to define a separable Banach space

$$\mathcal{A}_{\boldsymbol{\epsilon}} := \left\{ B \in C^\infty([-1,1]\times S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})) \;\middle|\; \|B\|_{C_{\boldsymbol{\epsilon}}} < \infty \text{ and } B(\pm 1,\cdot) \equiv 0 \right\},$$

and assume via Theorem B.0.6 that $\mathcal{A}_{\boldsymbol{\epsilon}}$ is dense in the Banach space of continuous functions $[-1,1]\times S^1 \to \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ vanishing at $\{\pm 1\}\times S^1$. We then consider perturbed 1-parameter families of asymptotic operators of the form

$$\mathbf{A}^B_s := \mathbf{A}_s + B(s,\cdot) : \mathcal{D} \to \mathcal{H}$$

for $B \in \mathcal{A}_{\boldsymbol{\epsilon}}$, $s \in [-1,1]$. Remarks 3.5.2 and 3.5.3 imply that the perturbed family defines a smooth path in $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ as long as the original path $s \mapsto S_s$ is smooth in $L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}(\mathbb{R}^{2n}))$. For each $k \in \mathbb{N}$ and $B \in \mathcal{A}_{\boldsymbol{\epsilon}}$, define the set

$$\mathcal{V}^k(B) = \left\{ (s,\lambda) \in (-1,1)\times\mathbb{R} \;\middle|\; \dim_{\mathbb{R}} \ker\left(\mathbf{A}^B_s - \lambda\right) = k \right\}.$$

To show that eigenvalues are generically simple, we need to show that for a comeager set of choices of $B \in \mathcal{A}_{\boldsymbol{\epsilon}}$, $\mathcal{V}^k(B)$ is empty for all $k \geqslant 2$. Given $(s_0,\lambda_0) \in \mathcal{V}^k(B)$, recall from §3.5 that there exist decompositions

$$\mathcal{D} = V \oplus K, \qquad \mathcal{H} = W \oplus K$$

where $K = \ker\left(\mathbf{A}^B_{s_0} - \lambda_0\right)$, $W = \mathrm{im}\left(\mathbf{A}^B_{s_0} - \lambda_0\right)$ is the L^2-orthogonal complement of K, and $V = W \cap \mathcal{D}$, so that any symmetric bounded linear operator $\mathbf{T}$ in a sufficiently small neighborhood $\mathcal{O} \subset \mathscr{L}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D},\mathcal{H})$ of $\mathbf{A}^B_{s_0} - \lambda_0$ can be written in block form

$$\mathbf{T} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$$

with $\mathbf{A} : V \to W$ invertible. This gives rise to a smooth map

$$\Phi : \mathcal{O} \to \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(K) : \mathbf{T} \mapsto \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$$

whose zero set is precisely the set of nearby symmetric operators with k-dimensional kernel. A neighborhood of (s_0,λ_0) in $\mathcal{V}^k(B)$ can thus be identified with the zero set of the map

$$\Psi_B(s,\lambda) := \Phi(\mathbf{A}^B_s - \lambda) \in \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(K),$$

defined for $(s,\lambda) \in (-1,1)\times\mathbb{R}$ sufficiently close to (s_0,λ_0). Notice that the derivative $d\Psi_B(s,\lambda) : \mathbb{R}\oplus\mathbb{R} \to \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(K)$ is Fredholm since its domain and target are both finite dimensional, and it can only ever be surjective when $k = \dim_{\mathbb{R}} K = 1$.

The following space will now play the role of a "universal moduli space" as in Chapter 9: let

$$\mathcal{V}^k = \left\{ (s, \lambda, B) \in (-1, 1) \times \mathbb{R} \times \mathcal{A}_\epsilon \;\middle|\; (s, \lambda) \in \mathcal{V}^k(B) \right\}.$$

The proof that this is a smooth Banach manifold depends on the following algebraic lemma.

LEMMA C.0.1. *Fix an asymptotic operator* $\mathbf{A} = -J_0\,\partial_t - S$ *and a linear transformation*

$$\Upsilon : \ker \mathbf{A} \to \ker \mathbf{A}$$

that is symmetric with respect to the L^2*-product. Then there exists a continuous loop* $B : S^1 \to \operatorname{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ *such that*

$$\langle \eta, B\xi \rangle_{L^2} = \langle \eta, \Upsilon\xi \rangle_{L^2}$$

for all $\eta, \xi \in \ker \mathbf{A}$.

PROOF. Note first that every nontrivial loop $\eta \in \ker \mathbf{A} \subset H^1(S^1, \mathbb{R}^{2n})$ is continuous and nowhere zero due to the generalized existence/uniqueness result for solutions to linear ODEs in Exercise 3.4.10. It follows that if we fix a basis $(\eta_1, \ldots, \eta_k)$ for $\ker \mathbf{A}$, then the vectors $\eta_1(t), \ldots, \eta_k(t) \in \mathbb{R}^{2n}$ are also linearly independent for all $t \in S^1$ and thus span a continuous S^1-family of k-dimensional subspaces $V_t \subset \mathbb{R}^{2n}$, each equipped with a distinguished basis. There is therefore a unique continuous S^1-family of linear transformations $\widehat{B}(t) : V_t \to V_t$ such that for every $\eta \in \ker \mathbf{A}$, $\widehat{B}(t)\eta(t) = (\Upsilon\eta)(t)$ for all t. Extend $\widehat{B}(t)$ arbitrarily to a continuous family of linear maps on $\mathbb{R}^{2n}$.

The matrices $\widehat{B}(t) \in \operatorname{End}(\mathbb{R}^{2n})$ need not be symmetric, but they do satisfy

$$\langle \eta, \widehat{B}\xi \rangle_{L^2} = \langle \eta, \Upsilon\xi \rangle_{L^2} \qquad \text{for all } \eta, \xi \in \ker \mathbf{A}.$$

Since Υ is symmetric, this implies moreover that for all $\eta, \xi \in \ker \mathbf{A}$,

$$\langle \eta, \Upsilon\xi \rangle_{L^2} = \langle \xi, \Upsilon\eta \rangle_{L^2} = \langle \xi, \widehat{B}\eta \rangle_{L^2} = \langle \eta, \widehat{B}^{\mathrm{T}}\xi \rangle_{L^2}.$$

The loop $B := \frac{1}{2}(\widehat{B} + \widehat{B}^{\mathrm{T}})$ thus has the desired properties. □

Now using the previously described construction in the space of symmetric Fredholm operators, a neighborhood of any point (s_0, λ_0, B_0) in $\mathcal{V}^k$ can be identified with the zero set of a smooth map of the form

$$\Psi(s, \lambda, B) := \Psi_B(s, \lambda) \in \operatorname{End}^{\mathrm{sym}}_{\mathbb{R}}(K),$$

defined for all (s, λ, B) sufficiently close to (s_0, λ_0, B_0) in $(-1, 1) \times \mathbb{R} \times \mathcal{A}_\epsilon$, where $K = \ker\left(\mathbf{A}^{B_0}_{s_0} - \lambda_0\right)$. The partial derivative of Ψ with respect to the third variable at (s_0, λ_0, B_0) is then a linear map

$$\mathbf{L} := D_3\Psi(s_0, \lambda_0, B_0) : \mathcal{A}_\epsilon \to \operatorname{End}^{\mathrm{sym}}_{\mathbb{R}}(K)$$

of the form

$$\mathbf{L}B : K \to K : \eta \mapsto \pi_K(B(s_0, \cdot)\eta), \tag{C.1}$$

where $\pi_K : W \oplus K \to K$ is the orthogonal projection. We claim that $\mathbf{L}$ is surjective. Indeed, for any $\Upsilon \in \operatorname{End}^{\mathrm{sym}}_{\mathbb{R}}(K)$, Lemma C.0.1 provides a continuous loop $C_0 : S^1 \to \operatorname{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ such that

$$\pi_K(C_0\eta) = \Upsilon\eta \quad \text{for all } \eta \in K,$$

and this can be extended to a continuous function $C : [-1,1] \times S^1 \to \operatorname{End}^{\mathrm{sym}}(\mathbb{R}^{2n})$ satisfying $C(s_0, \cdot) \equiv C_0$ and $C(\pm 1, \cdot) \equiv 0$ since $s_0 \neq \pm 1$. The function C might fail to be of class C_ϵ, but since it can be approximated arbitrarily well in the C^0-norm by functions in $\mathcal{A}_\epsilon$, we conclude that the image of $\mathbf{L}$ is dense in $\operatorname{End}^{\mathrm{sym}}_{\mathbb{R}}(K)$. Since the latter is finite dimensional, the claim follows.

The implicit function theorem now gives $\mathcal{V}^k$ the structure of a smooth Banach submanifold of $(-1,1) \times \mathbb{R} \times \mathcal{A}_\epsilon$, and it is separable since the latter is also separable. Consider the projection

$$\pi : \mathcal{V}^k \to \mathcal{A}_\epsilon : (s, \lambda, B) \mapsto B, \tag{C.2}$$

which is a smooth map of separable Banach manifolds whose fibers $\pi^{-1}(B)$ are the spaces $\mathcal{V}^k(B)$. Using Lemma 9.1.1, the fact that each map Ψ_B is Fredholm implies that π is also a Fredholm map, so the Sard-Smale theorem implies that the regular values of π form a comeager subset

$$\mathcal{A}_\epsilon^{\mathrm{reg},k} \subset \mathcal{A}_\epsilon.$$

The intersection

$$\mathcal{A}_\epsilon^{\mathrm{reg}} := \bigcap_{k \in \mathbb{N}} \mathcal{A}_\epsilon^{\mathrm{reg},k}$$

is then another comeager subset of $\mathcal{A}_\epsilon$, with the property that for each $B \in \mathcal{A}_\epsilon^{\mathrm{reg}}$ and every $k \in \mathbb{N}$ and $(s, \lambda) \in \mathcal{V}^k(B)$, $d\Psi_B(s, \lambda)$ is (by Lemma 9.1.1) surjective. As was observed previously, this is impossible for dimensional reasons if $k \geqslant 2$, implying that $\mathcal{V}^k(B)$ is then empty.

To find perturbations that also achieve the transversality condition, we use a similar argument: define for each $B \in \mathcal{A}_\epsilon$ the subset

$$\mathcal{V}^0(B) = \left\{ s \in (-1,1) \;\middle|\; \dim_{\mathbb{R}} \ker \mathbf{A}_s^B = 1 \right\},$$

along with the corresponding universal set

$$\mathcal{V}^0 = \left\{ (s, B) \in (-1,1) \times \mathcal{A}_\epsilon \;\middle|\; s \in \mathcal{V}^0(B) \right\}.$$

A neighborhood of any (s_0, B_0) in $\mathcal{V}^0$ is then the zero set of a smooth map of the form

$$\Psi(s, B) = \Phi(\mathbf{A}_s^B) \in \operatorname{End}^{\mathrm{sym}}_{\mathbb{R}}(\ker \mathbf{A}_{s_0}^{B_0}),$$

defined for all $(s, B) \in (-1,1) \times \mathcal{A}_\epsilon$ close enough to (s_0, B_0). For a fixed $B \in \mathcal{A}_\epsilon$ near B_0 and $s_1 \in \mathcal{V}^0(B)$ near s_0, a neighborhood of s_1 in $\mathcal{V}^0(B)$ is then the zero set of $\Psi_B(s) := \Psi(s, B)$, and the intersection of the path $s \mapsto \mathbf{A}_s^B \in \operatorname{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ with $\operatorname{Fred}^{\mathrm{sym},1}_{\mathbb{R}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ at $s = s_1$ is transverse if and only if

$$d\Psi_B(s_1) : \mathbb{R} \to \operatorname{End}^{\mathrm{sym}}_{\mathbb{R}}(\ker \mathbf{A}_{s_0}^{B_0})$$

is surjective. At (s_0, B_0), the partial derivative of Ψ with respect to B is again the same operator

$$\mathbf{L} = D_2\Psi(s_0, B_0) : \mathcal{A}_{\boldsymbol{\epsilon}} \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\ker \mathbf{A}_{s_0}^{B_0})$$

as in (C.1), which we've already seen is surjective due to Lemma C.0.1. Thus one can apply the Sard-Smale theorem to the projection

$$\mathcal{V}^0 \to \mathcal{A}_{\boldsymbol{\epsilon}} : (s, B) \mapsto B,$$

obtaining a comeager subset $\mathcal{A}_{\boldsymbol{\epsilon}}^{\mathrm{reg},0} \subset \mathcal{A}_{\boldsymbol{\epsilon}}$ such that all paths $\mathbf{A}_s + B(s, \cdot)$ for $B \in \mathcal{A}_{\boldsymbol{\epsilon}}^{\mathrm{reg},0}$ satisfy the required transversality condition. The comeager subset $\mathcal{A}_{\boldsymbol{\epsilon}}^{\mathrm{reg},0} \cap \mathcal{A}_{\boldsymbol{\epsilon}}^{\mathrm{reg}} \subset \mathcal{A}_{\boldsymbol{\epsilon}}$ thus consists of perturbed families of operators for which all desired conditions are satisfied, and it contains a sequence converging in the C^∞-topology to 0. This concludes the proof of Lemma 3.5.17.

Bibliography

[Abb14] C. Abbas, *An introduction to compactness results in symplectic field theory*, Springer, Heidelberg, 2014. MR3157146

[AS19] A. Abbondandolo and F. Schlenk, *Floer homologies, with applications*, Jahresber. Dtsch. Math.-Ver. **121** (2019), no. 3, 155–238.

[AA02] Y. A. Abramovich and C. D. Aliprantis, *An invitation to operator theory*, Graduate Studies in Mathematics, vol. 50, American Mathematical Society, Providence, RI, 2002.

[AF03] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, 2nd ed., Pure and Applied Mathematics (Amsterdam), vol. 140, Elsevier/Academic Press, Amsterdam, 2003.

[ABW10] P. Albers, B. Bramham, and C. Wendl, *On nonseparating contact hypersurfaces in symplectic 4-manifolds*, Algebr. Geom. Topol. **10** (2010), no. 2, 697–737.

[APS76] M. F. Atiyah, V. K. Patodi, and I. M. Singer, *Spectral asymmetry and Riemannian geometry. III*, Math. Proc. Cambridge Philos. Soc. **79** (1976), no. 1, 71–99.

[AD14] M. Audin and M. Damian, *Morse theory and Floer homology*, Universitext, Springer, London; EDP Sciences, Les Ulis, 2014. Translated from the 2010 French original by Reinie Erné.

[Avd21] R. Avdek, *Liouville hypersurfaces and connect sum cobordisms*, J. Symplectic Geom. **19** (2021), no. 4, 865–957.

[Bao15] E. Bao, *On J-holomorphic curves in almost complex manifolds with asymptotically cylindrical ends*, Pacific J. Math. **278** (2015), no. 2, 291–324.

[Bou02] F. Bourgeois, *A Morse-Bott approach to contact homology*, Ph.D. Thesis, Stanford University, 2002.

[BEH+03] F. Bourgeois, Y. Eliashberg, H. Hofer, K. Wysocki, and E. Zehnder, *Compactness results in symplectic field theory*, Geom. Topol. **7** (2003), 799–888.

[BS18] T. Bühler and D. A. Salamon, *Functional analysis*, Graduate Studies in Mathematics, vol. 191, American Mathematical Society, Providence, RI, 2018.

[Can22] D. Cant, *A dimension formula for relative symplectic field theory*, Ph.D. Thesis, Stanford University, 2022.

[CE12] K. Cieliebak and Y. Eliashberg, *From Stein to Weinstein and back: symplectic geometry of affine complex manifolds*, American Mathematical Society Colloquium Publications, vol. 59, American Mathematical Society, Providence, RI, 2012.

[CM05] K. Cieliebak and K. Mohnke, *Compactness for punctured holomorphic curves*, J. Symplectic Geom. **3** (2005), no. 4, 589–654. Conference on Symplectic Topology.

[CM18] ———, *Punctured holomorphic curves and Lagrangian embeddings*, Invent. Math. **212** (2018), no. 1, 213–295.

[CV15] K. Cieliebak and E. Volkov, *First steps in stable Hamiltonian topology*, J. Eur. Math. Soc. (JEMS) **17** (2015), no. 2, 321–404.

[CZ84] C. Conley and E. Zehnder, *Morse-type index theory for flows and periodic solutions for Hamiltonian equations*, Comm. Pure Appl. Math. **37** (1984), no. 2, 207–253.

[CE20] J. Conway and J. B. Etnyre, *Contact surgery and symplectic caps*, Bull. Lond. Math. Soc. **52** (2020), no. 2, 379–394.

[DK90] S. K. Donaldson and P. B. Kronheimer, *The geometry of four-manifolds*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1990. Oxford Science Publications.

[DS94] S. Dostoglou and D. A. Salamon, *Self-dual instantons and holomorphic curves*, Ann. of Math. (2) **139** (1994), no. 3, 581–640.

[Eli89] Y. Eliashberg, *Classification of overtwisted contact structures on* 3*-manifolds*, Invent. Math. **98** (1989), no. 3, 623–637.

[Eli07] Y. Eliashberg, *Symplectic field theory and its applications*, International Congress of Mathematicians. Vol. I, Eur. Math. Soc., Zürich, 2007, pp. 217–246.

[EGH00] Y. Eliashberg, A. Givental, and H. Hofer, *Introduction to symplectic field theory*, Geom. Funct. Anal., Special Volume (2000), 560–673.

[Elĭ67] H. I. Elĭasson, *Geometry of manifolds of maps*, J. Differential Geometry **1** (1967), 169–194.

[EH02] J. B. Etnyre and K. Honda, *On symplectic cobordisms*, Math. Ann. **323** (2002), no. 1, 31–39.

[Eva98] L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998.

[Eva10] J. D. Evans, *Lagrangian spheres in del Pezzo surfaces*, J. Topol. **3** (2010), no. 1, 181–227.

[FH23] J. W. Fish and H. Hofer, *Feral curves and minimal sets*, Ann. of Math. (2) **197** (2023), no. 2, 533–738.

[Flo88a] A. Floer, *Morse theory for Lagrangian intersections*, J. Differential Geom. **28** (1988), no. 3, 513–547.

[Flo88b] ———, *The unregularized gradient flow of the symplectic action*, Comm. Pure Appl. Math. **41** (1988), no. 6, 775–813.

[Flo88c] ———, *An instanton-invariant for* 3*-manifolds*, Comm. Math. Phys. **118** (1988), no. 2, 215–240.

[FH93] A. Floer and H. Hofer, *Coherent orientations for periodic orbit problems in symplectic geometry*, Math. Z. **212** (1993), no. 1, 13–38.

[FOOO09] K. Fukaya, Y.-G. Oh, H. Ohta, and K. Ono, *Lagrangian intersection Floer theory: anomaly and obstruction. Part I*, AMS/IP Studies in Advanced Mathematics, vol. 46, American Mathematical Society, Providence, RI; International Press, Somerville, MA, 2009.

[Gei08] H. Geiges, *An introduction to contact topology*, Cambridge Studies in Advanced Mathematics, vol. 109, Cambridge University Press, Cambridge, 2008.

[Ger18] C. Gerig, *Seiberg-Witten and Gromov invariants for self-dual harmonic 2-forms*, Ph.D. Thesis, University of California, Berkeley, 2018.

[GW17] C. Gerig and C. Wendl, *Generic transversality for unbranched covers of closed pseudo-holomorphic curves*, Comm. Pure Appl. Math. **70** (2017), no. 3, 409–443.

[Gir94] E. Giroux, *Une structure de contact, même tendue, est plus ou moins tordue*, Ann. Sci. École Norm. Sup. (4) **27** (1994), no. 6, 697–705 (French, with English summary).

[Gro85] M. Gromov, *Pseudoholomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), no. 2, 307–347.

[Hir94] M. W. Hirsch, *Differential topology*, Springer-Verlag, New York, 1994.

[Hof93] H. Hofer, *Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three*, Invent. Math. **114** (1993), no. 3, 515–563.

[Hof06] H. Hofer, *A general Fredholm theory and applications*, Current developments in mathematics, 2004, Int. Press, Somerville, MA, 2006, pp. 1–71.

[HS95] H. Hofer and D. A. Salamon, *Floer homology and Novikov rings*, The Floer memorial volume, Progr. Math., vol. 133, Birkhäuser, Basel, 1995, pp. 483–524.

[HWZ95] H. Hofer, K. Wysocki, and E. Zehnder, *Properties of pseudo-holomorphic curves in symplectisations. II. Embedding controls and algebraic invariants*, Geom. Funct. Anal. **5** (1995), no. 2, 270–328.

[HWZ96] ———, *Properties of pseudoholomorphic curves in symplectisations. I. Asymptotics*, Ann. Inst. H. Poincaré Anal. Non Linéaire **13** (1996), no. 3, 337–379.

[HWZ01] ______, *The asymptotic behavior of a finite energy plane* (2001). FIM preprint, available at http://e-collection.library.ethz.ch/view/eth:25385.

[HWZ96] ______, *Properties of pseudoholomorphic curves in symplectisations. IV. Asymptotics with degeneracies*, Contact and symplectic geometry (Cambridge, 1994), 1996, pp. 78–117.

[HWZ99] ______, *Properties of pseudoholomorphic curves in symplectizations. III. Fredholm theory*, Topics in nonlinear analysis, 1999, pp. 381–475.

[HWZ03] ______, *Finite energy foliations of tight three-spheres and Hamiltonian dynamics*, Ann. of Math. (2) **157** (2003), no. 1, 125–255.

[HZ94] H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser Verlag, Basel, 1994.

[Hum97] C. Hummel, *Gromov's compactness theorem for pseudo-holomorphic curves*, Progress in Mathematics, vol. 151, Birkhäuser Verlag, Basel, 1997.

[Hut14] M. Hutchings, *Lecture Notes on Embedded Contact Homology*, Contact and Symplectic Topology, Bolyai Society Mathematical Studies, vol. 26, Springer, 2014, pp. 389–484.

[HS05] M. Hutchings and M. Sullivan, *The periodic Floer homology of a Dehn twist*, Algebr. Geom. Topol. **5** (2005), 301–354.

[HT07] M. Hutchings and C. H. Taubes, *Gluing pseudoholomorphic curves along branched covered cylinders. I*, J. Symplectic Geom. **5** (2007), no. 1, 43–137.

[Kat95] T. Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[Kob87] S. Kobayashi, *Differential geometry of complex vector bundles*, Publications of the Mathematical Society of Japan, vol. 15, Princeton University Press, Princeton, NJ, 1987. Kanô Memorial Lectures, 5.

[Lan93] S. Lang, *Real and functional analysis*, 3rd ed., Springer-Verlag, New York, 1993.

[Lan99] ______, *Fundamentals of differential geometry*, Springer-Verlag, New York, 1999.

[LW11] J. Latschev and C. Wendl, *Algebraic torsion in contact manifolds*, Geom. Funct. Anal. **21** (2011), no. 5, 1144–1195. With an appendix by Michael Hutchings.

[Laz20] O. Lazarev, *Maximal contact and symplectic structures*, J. Topol. **13** (2020), no. 3, 1058–1083.

[LL01] E. H. Lieb and M. Loss, *Analysis*, 2nd ed., Graduate Studies in Mathematics, vol. 14, American Mathematical Society, Providence, RI, 2001.

[Lip06] R. Lipshitz, *A cylindrical reformulation of Heegaard Floer homology*, Geom. Topol. **10** (2006), 955–1096. [Paging previously given as 955–1097].

[Lut77] R. Lutz, *Structures de contact sur les fibrés principaux en cercles de dimension trois*, Ann. Inst. Fourier (Grenoble) **27** (1977), no. 3, ix, 1–15 (French, with English summary).

[Mar17] M. Maridakis, *Spinor pairs and the concentration principle for Dirac operators*, Trans. Amer. Math. Soc. **369** (2017), no. 3, 2231–2254.

[MS12] D. McDuff and D. Salamon, *J-holomorphic curves and symplectic topology*, 2nd ed., American Mathematical Society Colloquium Publications, vol. 52, American Mathematical Society, Providence, RI, 2012.

[MS17] ______, *Introduction to symplectic topology*, 3rd ed., Oxford University Press, 2017.

[MW95] M. J. Micallef and B. White, *The structure of branch points in minimal surfaces and in pseudoholomorphic curves*, Ann. of Math. (2) **141** (1995), no. 1, 35–85.

[Mil63] J. Milnor, *Morse theory*, Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51, Princeton University Press, Princeton, N.J., 1963.

[Mil97] J. W. Milnor, *Topology from the differentiable viewpoint*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997. Based on notes by David W. Weaver; Revised reprint of the 1965 original.

[Mor03] E. Mora, *Pseudoholomorphic cylinders in symplectisations*, Ph.D. Thesis, New York University, 2003.

[Nel15] J. Nelson, *Automatic transversality in contact homology I: regularity*, Abh. Math. Semin. Univ. Hambg. **85** (2015), no. 2, 125–179.

[Pal68] R. S. Palais, *Foundations of global non-linear analysis*, W. A. Benjamin, Inc., New York-Amsterdam, 1968.

[Rab78] P. H. Rabinowitz, *Periodic solutions of Hamiltonian systems*, Comm. Pure Appl. Math. **31** (1978), no. 2, 157–184.

[RS80] M. Reed and B. Simon, *Methods of modern mathematical physics. I*, 2nd ed., Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1980. Functional analysis.

[RS93] J. Robbin and D. Salamon, *The Maslov index for paths*, Topology **32** (1993), no. 4, 827–844.

[RS95] ______, *The spectral flow and the Maslov index*, Bull. London Math. Soc. **27** (1995), no. 1, 1–33.

[Rob70] R. C. Robinson, *A global approximation theorem for Hamiltonian systems*, Global Analysis (Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968), Amer. Math. Soc., Providence, R.I., 1970, pp. 233–243.

[RT95] Y. Ruan and G. Tian, *A mathematical theory of quantum cohomology*, J. Differential Geom. **42** (1995), no. 2, 259–367.

[RT97] ______, *Higher genus symplectic invariants and sigma models coupled with gravity*, Invent. Math. **130** (1997), 455–516.

[Sal99] D. Salamon, *Lectures on Floer homology*, Symplectic geometry and topology (Park City, UT, 1997), IAS/Park City Math. Ser., vol. 7, Amer. Math. Soc., Providence, RI, 1999, pp. 143–229.

[SZ92] D. Salamon and E. Zehnder, *Morse theory for periodic solutions of Hamiltonian systems and the Maslov index*, Comm. Pure Appl. Math. **45** (1992), no. 10, 1303–1360.

[Sch93] M. Schwarz, *Morse homology*, Progress in Mathematics, vol. 111, Birkhäuser Verlag, Basel, 1993.

[Sch95] ______, *Cohomology operations from S^1-cobordisms in Floer homology*, Ph.D. Thesis, ETH Zürich, 1995.

[Sei02] P. Seidel, *Symplectic Floer homology and the mapping class group*, Pacific J. Math. **206** (2002), no. 1, 219–229.

[Sei08] ______, *Fukaya categories and Picard-Lefschetz theory*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2008.

[SS92] M. Seppälä and T. Sorvali, *Geometry of Riemann surfaces and Teichmüller spaces*, North-Holland Mathematics Studies, vol. 169, North-Holland Publishing Co., Amsterdam, 1992.

[Sie08] R. Siefring, *Relative asymptotic behavior of pseudoholomorphic half-cylinders*, Comm. Pure Appl. Math. **61** (2008), no. 12, 1631–1684.

[Sie17] ______, *Finite-energy pseudoholomorphic planes with multiple asymptotic limits*, Math. Ann. **368** (2017), no. 1-2, 367–390.

[Sik94] J.-C. Sikorav, *Some properties of holomorphic curves in almost complex manifolds*, Holomorphic curves in symplectic geometry, Progr. Math., vol. 117, Birkhäuser, Basel, 1994, pp. 165–189.

[Tau96a] C. H. Taubes, *Counting pseudo-holomorphic submanifolds in dimension* 4, J. Differential Geom. **44** (1996), no. 4, 818–893.

[Tau96b] C. H. Taubes, SW $\Rightarrow$ Gr*: from the Seiberg-Witten equations to pseudo-holomorphic curves*, J. Amer. Math. Soc. **9** (1996), no. 3, 845–918.

[Tau10] C. H. Taubes, *Embedded contact homology and Seiberg-Witten Floer cohomology I*, Geom. Topol. **14** (2010), no. 5, 2497–2581.

[Tro92] A. J. Tromba, *Teichmüller theory in Riemannian geometry*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1992.

[Web02] J. Weber, *Perturbed closed geodesics are periodic orbits: index and transversality*, Math. Z. **241** (2002), no. 1, 45–82.

[Wei78] A. Weinstein, *Periodic orbits for convex Hamiltonian systems*, Ann. of Math. (2) **108** (1978), no. 3, 507–518.

[Wen10] C. Wendl, *Automatic transversality and orbifolds of punctured holomorphic curves in dimension four*, Comment. Math. Helv. **85** (2010), no. 2, 347–407.

[Wen13] ———, *Non-exact symplectic cobordisms between contact* 3*-manifolds*, J. Differential Geom. **95** (2013), no. 1, 121–182.

[Wen18] ———, *Holomorphic curves in low dimensions: from symplectic ruled surfaces to planar contact manifolds*, Lecture Notes in Mathematics, vol. 2216, Springer-Verlag, 2018.

[Wen20] ———, *Lectures on Contact 3-Manifolds, Holomorphic Curves and Intersection Theory*, Cambridge Tracts in Mathematics, vol. 220, Cambridge University Press, Cambridge, 2020.

[Wena] ———, *Signs (or how to annoy a symplectic topologist)*. Blog post, available at https://symplecticfieldtheorist.wordpress.com/2015/08/23/.

[Wenb] ———, *Lectures on holomorphic curves in symplectic and contact geometry*. Version 3.3, http://www.math.hu-berlin.de/~wendl/pub/jhol_bookv33.pdf.

[Wen26] ———, *Functional Analysis* (2026). Notes from the course at HU Berlin, available at https://www.mathematik.hu-berlin.de/~wendl/Winter2025/FunkAna/lecturenotes.pdf.

[Wit82] E. Witten, *Supersymmetry and Morse theory*, J. Differential Geom. **17** (1982), no. 4, 661–692 (1983).