Lectures on Symplectic Field Theory

Chris Wendl

Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
E-mail address: `wendl@math.hu-berlin.de`

Contents

LECTURE 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 lecture, 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. 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 1.2. 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 [Wend, §2.2], among other places.

PROPOSITION 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 $[\omega] \in H^2_{\mathrm{dR}}(M)$ 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 1.4. *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 1.4, 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 1.5 (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 [**Wenc**], 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 1.6. Regarding $C^\infty_{\mathrm{contr}}(S^1, M)$ as a Fréchet manifold with tangent spaces $T_\gamma C^\infty_{\mathrm{contr}}(S^1, M) = \Gamma(\gamma^* TM)$, show that the first variation of the action functional $\mathcal{A}_H$ is

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

for $\eta \in \Gamma(\gamma^* TM)$. In particular, 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

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

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 $\mathrm{ind}(x) - \mathrm{ind}(y) = 1$. The real magic however is contained in the following statement about the case $\mathrm{ind}(x) - \mathrm{ind}(y) = 2$:

PROPOSITION 1.7. *For generic choices of f and g and any two critical points $x, y \in \mathrm{Crit}(f)$ with $\mathrm{ind}(x) - \mathrm{ind}(y) = 2$, $\mathcal{M}(x, y)/\mathbb{R}$ is homeomorphic to a finite collection of circles and open intervals whose end points are canonically identified with the finite set*

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

We say that $\mathcal{M}(x, y)$ has a natural **compatification** $\overline{\mathcal{M}}(x, y)$, 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

$$\partial^2 = 0$$

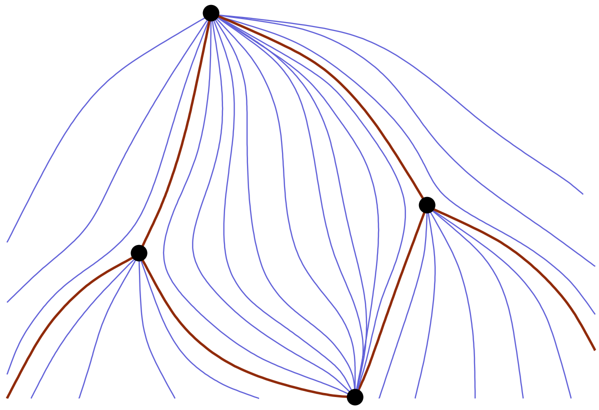

FIGURE 1.1. One-parameter families of gradient flow lines on a Riemannian manifold degenerate to broken flow lines.

as a consequence of the fact that compact 1-manifolds always have zero boundary points when counted with appropriate signs.[2] 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 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.

[2]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.

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 Lecture 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 $C^\infty_{\mathrm{contr}}(S^1, M)$ 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 $C^\infty_{\mathrm{contr}}(S^1, M)$ formally as a manifold with tangent spaces

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

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 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 1.6 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 $u : \mathbb{R} \to C^\infty_{\mathrm{contr}}(S^1, M)$ 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

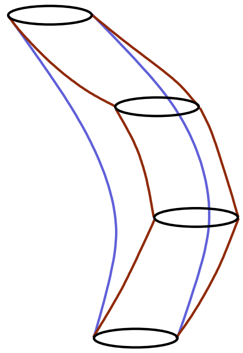

FIGURE 1.2. A family of smooth Floer trajectories can degenerate into a broken Floer trajectory.

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 pseudo-holomorphic 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 Lecture 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 1.8. 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 1.9. 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 1.10. 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 1.10 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. 1.24 below) in Lecture 9.

PROPOSITION 1.11. *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 nonegenerate, 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 1.12. 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 [**AS**] for a survey.

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

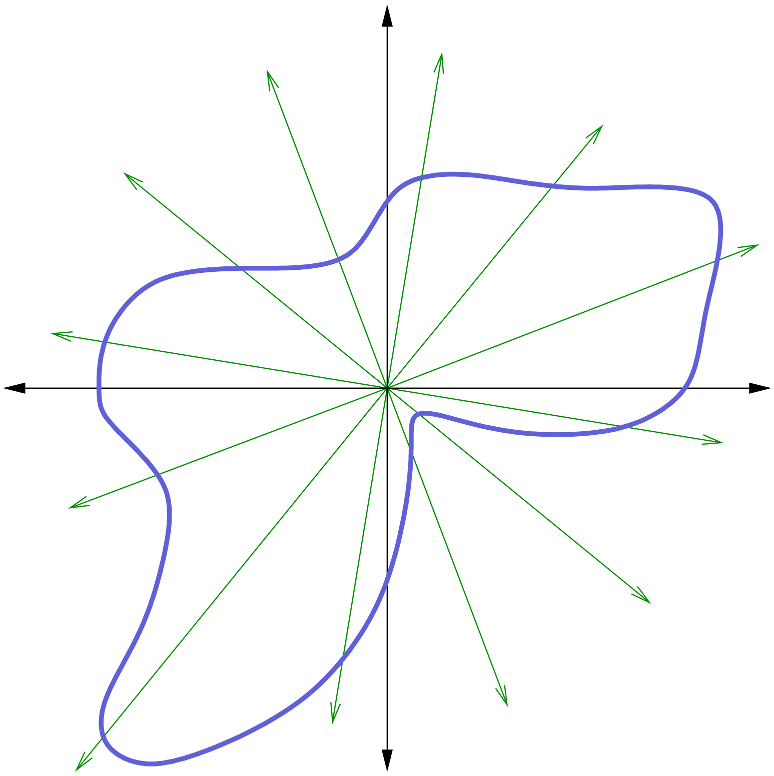

FIGURE 1.3. A star-shaped hypersurface in Euclidean space

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 dual 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**. 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 1.13. 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 1.14. 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

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

hence $\Phi^*\omega = d(e^r\alpha)$.

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 1.15 (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**.[3] We call the pair (M, ξ) a **contact manifold**, and refer to a diffeomorphism $\varphi : M \to M'$ as a

[3]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 these notes; for our purposes, the co-orientation is *always* considered to be part of the data of a contact structure.

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 1.16 (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 [**Wend**, 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 1.14 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 1.17. 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 1.18. Given a contact form α on M, the **Reeb vector field** is the unique vector field R_α that satisfies

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

EXERCISE 1.19. 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 1.20 (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).

EXERCISE 1.21. Recall that for any smooth manifold M, the cotangent bundle T^*M carries a tautological 1-form $\lambda \in \Omega^1(T^*M)$ that locally takes the form $\lambda = \sum_{j=1}^n p_j\, dq_j$ in any choice of local coordinates $(q_1, \dots, q_n)$ on a neighbood $\mathcal{U} \subset M$, with $(p_1, \dots, p_n)$ denoting the induced coordinates on the cotangent fibers over $\mathcal{U}$. This is a Liouville form, with $d\lambda$ 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 $(T^*M, d\lambda)$, and the Liouville vector field on T^*M dual to λ 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 along $S_\xi M$ with $e^r\alpha$.

REMARK 1.22. Exercise 1.21 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. 1.11 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**[4] 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

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

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

[4]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.

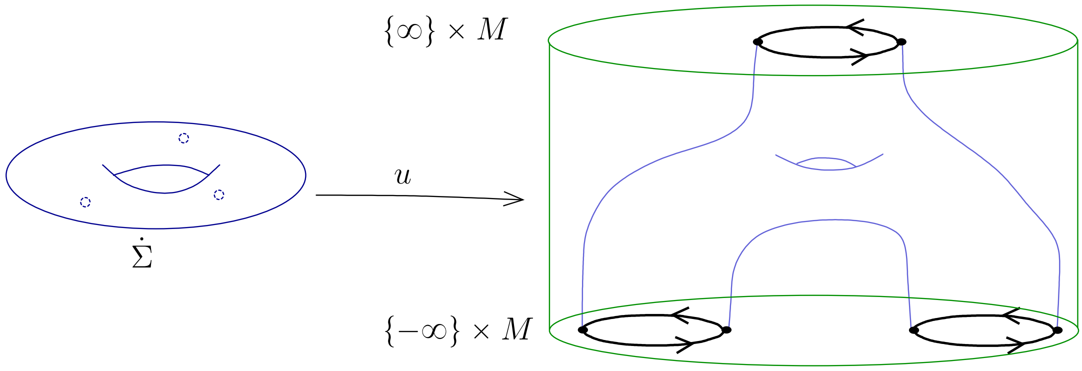

Figure 1.4. An asymptotically cylindrical holomorphic curve in a symplectization, with genus 1, one positive puncture and two negative punctures.

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 1.23. 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. 1.11 will be proved in Lecture 9. 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 1.19),

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

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 1.25 below). The following result is the contact analogue of Proposition 1.11.

PROPOSITION 1.24. *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 1.24, this implies the existence of a contractible periodic Reeb orbit and thus proves the Weinstein conjecture in these settings.

REMARK 1.25. 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 Lecture 8 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 2020. 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[5] 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 in this case that the Liouville vector field chosen near ∂W must point *outward* at

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

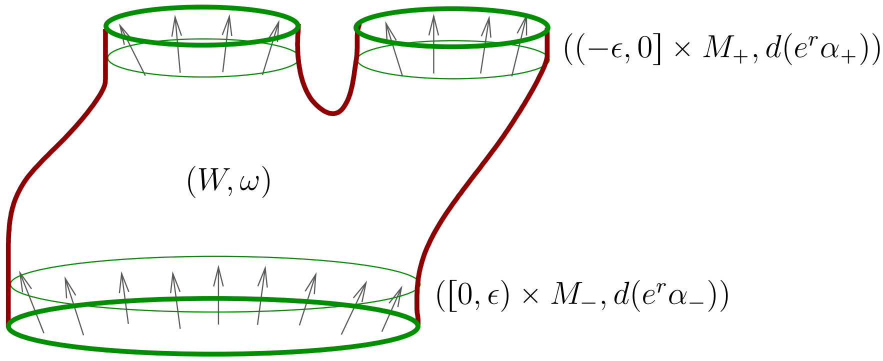

FIGURE 1.5. A symplectic cobordism with concave boundary (M_-, ξ_-) and convex boundary (M_+, ξ_+), with symplectic collar neighborhoods defined by flowing along Liouville vector fields near the boundary.

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 1.14, 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 1.26. 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

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

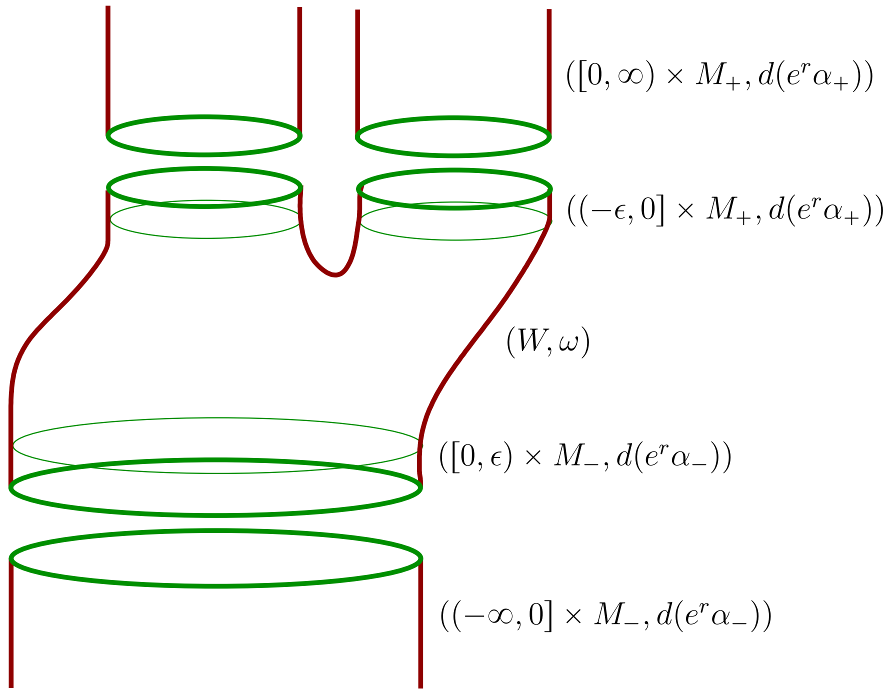

FIGURE 1.6. The completion of a symplectic cobordism

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 $(\widehat{W},\widehat{\omega})$ also inherits a global Liouville form.

EXERCISE 1.27. 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**, **CE**, **Laz**]).

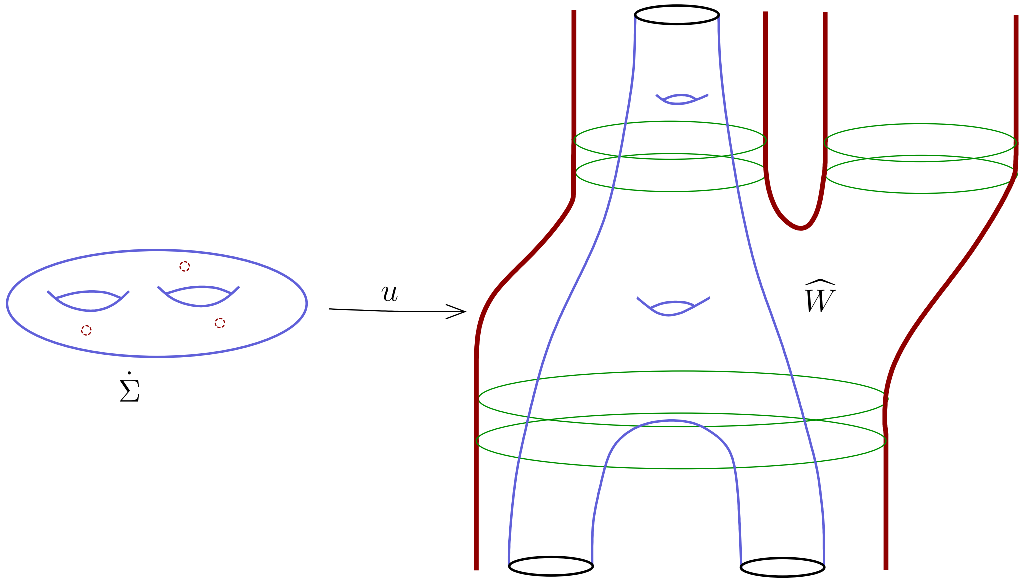

Figure 1.7. An asymptotically cylindrical holomorphic curve in a completed symplectic cobordism, with genus 2, one positive puncture and two negative punctures.

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,[6] 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. 1.3, 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

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

[6] 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 lectures we will see some reasons to worry that it might not (see §6.7.2).

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

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

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

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

It will be a straightforward matter to generalize Proposition 1.24 and show that finite energy implies asymptotically cylindrical behavior in completed cobordisms.

EXERCISE 1.28. 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 soon. 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[7] 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.[8]

[7] Actually I should be making a distinction here between "good" and "bad" Reeb orbits, but let's discuss that later; see Lecture 11.

[8] Similar combinatorial factors are hidden behind the symbol "#" in our definitions of ∂_{CH} and $\mathbf{H}$, and will be discussed in earnest in Lecture 12.

"THEOREM" 1.29. *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 1.28 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" 1.29 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

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

with $\mathcal{M}(\gamma; \gamma_1, \ldots, \gamma_m)$ 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 $\gamma_1, \ldots, \gamma_m$, 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" 1.30. *$\partial_{\mathrm{CH}}^2 = 0$, 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

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

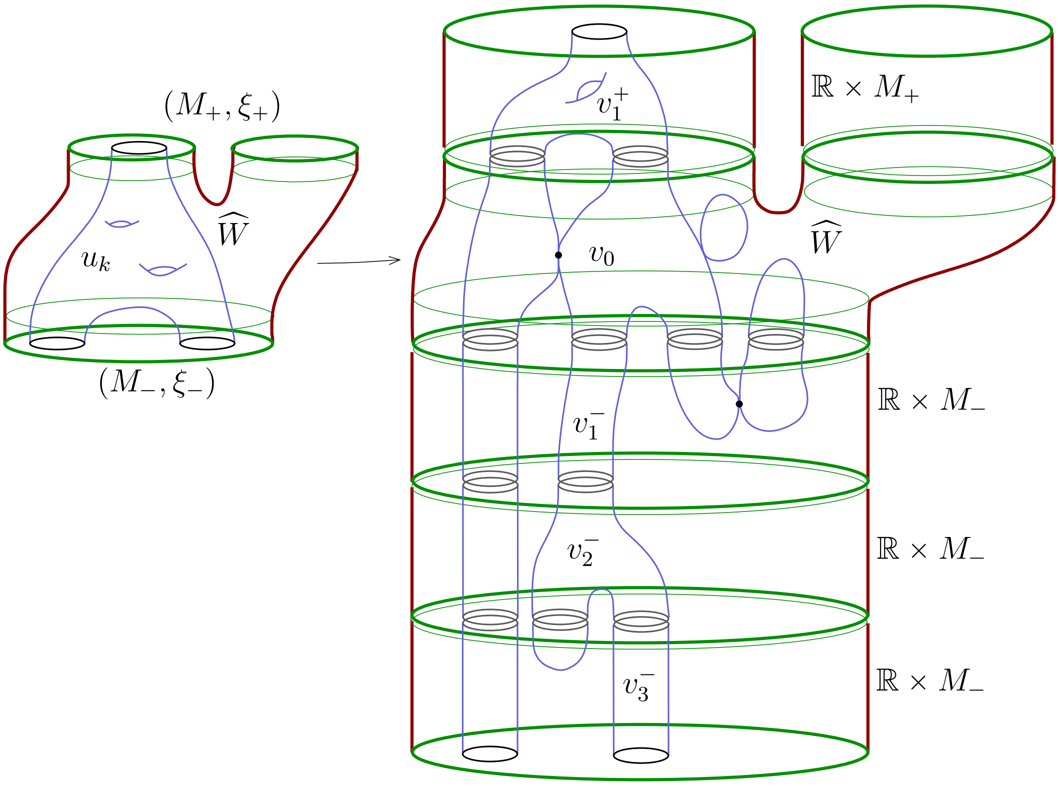

FIGURE 1.8. Degeneration of a sequence u_k of finite energy punctured holomorphic curves with genus 2, one positive puncture and two negative punctures in a symplectic cobordism. The limiting holomorphic building $(v_1^+, v_0, v_1^-, v_2^-, v_3^-)$ in this example has one upper level living in the symplectization $\mathbb{R} \times M_+$, 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.

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" 1.31. $[\mathbf{H}, \mathbf{H}] = 0$, *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 Lecture 12.

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 1.32. *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 Lecture 10 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).

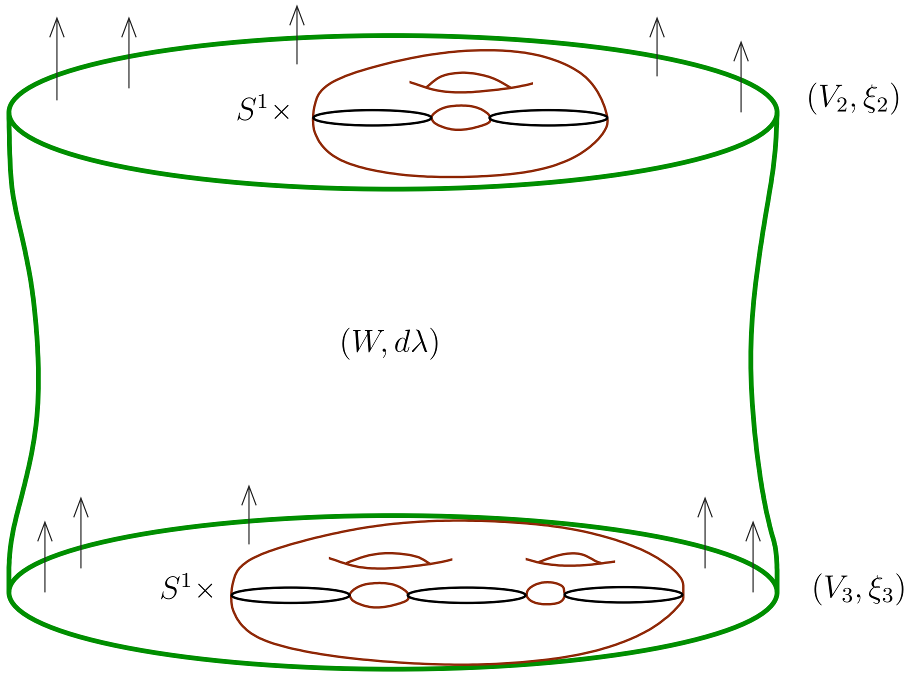

FIGURE 1.9. This exact symplectic cobordism does not exist.

THEOREM 1.33. *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 [**Wen13b**, **Avd**] 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 Lecture 13 and compute it for these examples in Lecture 16.

LECTURE 2

Basics on holomorphic curves

Contents

In this lecture 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 everything in this lecture can be found in [**Wend**].

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 $\eta \in \Gamma(u^*TM)$. 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: $\Omega^{0,1}(\Sigma, u^*TM)$ 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 $\Omega^{0,1}(\Sigma, u^*TM)$ 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 $W^{k-1,p}$.

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).

---

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

DEFINITION 2.1. 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 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 2.2. If $\mathbf{D}$ is a linear Cauchy-Riemann type operator on E, prove that every other such operator is of the form $\mathbf{D} + A$ where $A : E \to \overline{\mathrm{Hom}}_{\mathbb{C}}(T\Sigma, E)$ is

[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.

[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**].

a smooth linear bundle map. 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 2.3. Verify that the linearized operator $\mathbf{D}_u$ of (2.1) is a real-linear Cauchy-Riemann type operator.

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.

Throughout this section 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) that $C^k(\overline{\mathcal{U}})$ is precisely the set of functions which admit k-times differentiable extensions to some open set containing $\overline{\mathcal{U}}$.

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.6 and A.10 in Appendix A for the more general versions, proofs of which may be found e.g. in [**AF03**].

PROPOSITION 2.4 (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.

$\square$

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 2.4 is by thinking of $W^{k,p}(\mathcal{U})$ as a space of functions that have "$k - \frac{n}{p}$ continuous derivatives".

EXERCISE 2.5. Show that the compactness of the inclusions in Proposition 2.4 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 2.6 (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.18). 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 2.7 (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 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 2.8. In the settings of Propositions 2.6 and 2.7, 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.16 and Theorem A.18 in Appendix A.

REMARK 2.9. Though we will not yet use it in this lecture, Propositions 2.4, 2.6 and 2.7 are the essential conditions needed in order to define smooth Banach manifold structures on spaces of $W^{k,p}$-smooth maps from one manifold to another, cf. [**Elĭ67**, **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.21 in Appendix A for a proof.

PROPOSITION 2.10. *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 2.11. 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**. There is no need to discuss here precisely what ellipticity means in full generality (see [**Wend**, §2.B] if you're curious about this); in practice, the main consequence is the following pair of analytical results.

THEOREM 2.12. *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 2.13. *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 [**Wend**, §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 is not hard to show from this formula that whenever $f \in C_0^\infty$, $K * f$ has decaying behavior at infinity (see [**Wend**, Lemma 2.6.13]). Thus if $u \in C_0^\infty$ and $\bar{\partial} u = f$, it follows 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 2.12 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 2.14. Use Theorem 2.12 and the remarks above to prove Theorem 2.13 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 [**Wend**, 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. [**Wend**, §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 $\hat{u} : \mathbb{R}^n \to \mathbb{C}$ 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

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

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

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

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 $\widehat{\partial^\alpha u}(p) = (2\pi i p)^\alpha \hat{u}(p)$ 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

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

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

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

[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 $p \neq 0$

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

PROPOSITION 2.15. *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

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

□

COROLLARY 2.16. *The estimate (2.5) holds in the case $p = 2$.* □

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 2.2 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 $\mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$. 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 Lecture 4.

THEOREM 2.17 (Linear regularity). *Assume $1 < p < \infty$, m and k are integers with $m \geqslant k \geqslant 0$, $A : \mathbb{D} \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$ 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*

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

PROOF (EXCLUDNG THE CASE $k = 0$ OF (1)). We begin with a proof of 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 2.13 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 $c' > 0$.

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 2.13 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.17 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.22 in Appendix A.3. □

EXERCISE 2.18. Deduce from Theorem 2.17 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 $W^{m+1,p}_{\mathrm{loc}}$-convergent on $\mathring{\mathbb{D}}$.

REMARK 2.19. Combining the Sobolev embedding theorem with the Arzelà-Ascoli theorem, the result of Exercise 2.18(a) proves that if the f_ν and A_ν are C^∞-bounded on $\mathbb{D}$, then the solutions u_ν have 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 2.20. Use Theorem 2.17 to generalize Theorem 2.12 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.4 and Corollary A.5.*

It remains to prove the case $k = 0$ of Theorem 2.17(1). As preparation for this, we start with a classical result about "weakly holomorphic" functions:

LEMMA 2.21. *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 [**Wend**, Lemma 2.6.26] for more details. □

LEMMA 2.22. *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 2.12. 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 2.21. 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 2.17(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 2.22 implies $u \in W^{1,p}_{\mathrm{loc}}(\mathring{\mathbb{D}})$. If $m \geqslant 1$, one can now shrink the disk slightly and plug in the case $k = 1$ of the theorem to conclude $u \in W^{m+1,p}_{\mathrm{loc}}(\mathring{\mathbb{D}})$. □

COROLLARY 2.23. *If $A : \mathbb{D} \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$ 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 $W^{k,q}_{\mathrm{loc}}(\mathring{\mathbb{D}})$ 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 2.17(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 2.17(1) again gives the desired result on $\mathring{\mathbb{D}}_r$. 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 $L^q_{\text{loc}}(\mathring{\mathbb{D}})$ for some $q > 2$. Here Theorem 2.17(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.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 $u \in L^q(\mathring{\mathbb{D}}_r)$. $\square$

2.4.2. The nonlinear case. 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]

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

Since it is useful for certain applications and does not make the proofs any harder, we will in this section consider solutions $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ to a more general *inhomogeneous* nonlinear Cauchy-Riemann equation

$$\partial_s u(z) + J(z, u(z))\partial_t u(z) = f(z), \qquad \text{or for short:} \qquad \partial_s u + J(z, u)\partial_t u = f,$$

where $f : \mathring{\mathbb{D}} \to \mathbb{C}^n$ is a given function and $J : \mathring{\mathbb{D}} \times \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$ is now allowed to depend on points both in the target $\mathbb{C}^n$ and in the domain $\mathring{\mathbb{D}}$. The nonlinear analogue of Theorem 2.17 is then the following.

THEOREM 2.24 (Nonlinear regularity). *Assume $1 < p < \infty$, and m and k are integers with $m \geqslant k$ and $kp > 2$.*

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

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

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 $\mathring{\mathbb{D}}$.

[6]Here the reader should beware of a minor ambiguity in notation: while we used $\mathcal{J}(M)$ in Lecture 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 $\operatorname{End}_{\mathbb{R}}(V) \cong \mathbb{R}^{2n \times 2n}$. It is not hard to show that $\mathcal{J}(V)$ is then a smooth submanifold of $\operatorname{End}_{\mathbb{R}}(V)$; 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 $\operatorname{Aut}_{\mathbb{R}}(V)/\operatorname{Aut}_{\mathbb{C}}(V, J) \cong \mathrm{GL}(2n, \mathbb{R})/\mathrm{GL}(n, \mathbb{C})$. 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.

(2) *Consider a C^m_{loc}-convergent sequence $J_\nu \to J$ of C^m-smooth maps $\mathring{\mathbb{D}} \times \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$, 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*

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

  (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 2.25. *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.* $\square$

EXERCISE 2.26. Use Theorem 2.24 to show that on a symplectic manifold (M, ω) endowed with smooth families of almost complex structures $\{J_t \in \mathcal{J}(M, \omega)\}_{t \in S^1}$ and Hamiltonians $\{H_t \in C^\infty(M, \mathbb{R})\}_{t \in S^1}$, weak solutions to the Floer equation (1.4) that are locally of class $W^{k,p}$ with $kp > 2$ are also smooth.

REMARK 2.27. 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 2.24 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 2.23) 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 2.24 will follow a similar outline to the proof of Theorem 2.17, 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 $J : \mathring{\mathbb{D}} \times \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$ the function

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

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

$$\bar{\partial} u - Q(z,u)\partial_t u = f. \tag{2.9}$$

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 $J(z_0, 0) = i$, so in particular $Q(z_0, 0) = 0$. 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

$$\begin{aligned}
\widehat{J} : \mathbb{D} \times \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n), \qquad & \widehat{J}(z,x) := J(z_0 + \epsilon z, \epsilon^\alpha x), \\
\widehat{Q} : \mathbb{D} \times \mathbb{C}^n \to \operatorname{End}_{\mathbb{R}}(\mathbb{C}^n), \qquad & \widehat{Q}(z,x) := Q(z_0 + \epsilon z, \epsilon^\alpha x) = i - \widehat{J}(z,x), \\
\hat{u} : \mathring{\mathbb{D}} \to \mathbb{C}^n, \qquad & \hat{u}(z) := \frac{1}{\epsilon^\alpha} u(z_0 + \epsilon z), \\
\hat{f} : \mathring{\mathbb{D}} \to \mathbb{C}^n, \qquad & \hat{f}(z) := \epsilon^{1-\alpha} f(z_0 + \epsilon z).
\end{aligned} \tag{2.10}$$

Then u satisfies (2.9) if and only if $\hat{u}$ satisfies

$$\bar{\partial}\hat{u} - \widehat{Q}(z,\hat{u})\partial_t \hat{u} = \hat{f}. \tag{2.11}$$

The rescaled almost complex structure has the convenient feature that since $J(z_0, 0)$ is the standard complex structure i, choosing $\epsilon > 0$ small makes $\widehat{J}$ arbitrarily C^m-close to i on the compact set[7] $\mathbb{D} \times \mathbb{D}^{2n} \subset \mathbb{C} \times \mathbb{C}^n$, which means $\|\widehat{Q}\|_{C^m(\mathbb{D}\times\mathbb{D}^{2n})}$ can be made arbitrarily small. By Proposition 2.10, $\|\hat{u}\|_{W^{k,p}(\mathring{\mathbb{D}})}$ 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 the map $z \mapsto (z, \hat{u}(z))$ for $z \in \mathbb{D}$ has image in $\mathbb{D} \times \mathbb{D}^{2n}$. By the assumption $m \geqslant k$ and the continuity of the map $C^k \times W^{k,p} \to W^{k,p}$ in Proposition 2.7, the function $\mathbb{D} \to \operatorname{End}_{\mathbb{R}}(\mathbb{C}^n) : z \mapsto \widehat{Q}(z, \hat{u}(z))$ can then likewise be assumed to be arbitrarily $W^{k,p}$-small by choosing $\epsilon > 0$ small. The effect is that (2.11) 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 then a uniform bound over some neighborhood of *every* 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 2.28. The rescaling trick described above is one of several reasons why the condition $kp > 2$ is needed in Theorem 2.24, while it was irrelevant in the linear case. We will see when we study compactness in Lecture 9 that the result is false in general without this assumption.

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

PROOF OF THEOREM 2.24. 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 $\mathring{\mathbb{D}} \times \mathbb{C}^n \to \mathcal{J}(\mathbb{C}^n)$ 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 $\partial_s u_\nu + J_\nu(z, u_\nu)\partial_t u_\nu = f_\nu$, 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 2.4), 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

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

We then choose

$$\alpha \in (0,1) \qquad \text{with} \qquad \alpha < k - \frac{2}{p}, \tag{2.12}$$

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 $\widehat{J}_\nu$ as defined in (2.10), defining also the related functions $\widehat{Q}_\nu = i - \widehat{J}_\nu$. We then have the equation $\bar{\partial}\hat{u}_\nu - \widehat{Q}_\nu(z, \hat{u}_\nu)\partial_t \hat{u}_\nu = \hat{f}_\nu$, with C^k-convergence $\widehat{Q}_\nu \to \widehat{Q}$ over $\mathbb{D} \times \mathbb{D}^{2n}$, where $\widehat{Q}$ 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 (2.12) and then apply Propostion 2.10 to obtain a bound

$$\|\hat{u}_\nu\|_{W^{k,p}} \leqslant C\epsilon^{\beta-\alpha}\|u_\nu\|_{W^{k,p}} \leqslant C\epsilon^{\beta-\alpha}M \tag{2.13}$$

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 2.13,

$$\|\partial_j \hat{u}_\nu\|_{W^{k,p}(\mathring{\mathbb{D}}_r)} \leqslant \|\beta\,\partial_j \hat{u}_\nu\|_{W^{k,p}(\mathring{\mathbb{D}})} \leqslant c \left\|\bar\partial\left(\beta\,\partial_j \hat{u}_\nu\right)\right\|_{W^{k-1,p}(\mathring{\mathbb{D}})}. \tag{2.14}$$

If this were still the proof of Theorem 2.17, 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 $\bar\partial \hat{u}_\nu = \widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t \hat{u}_\nu + \hat{f}_\nu$ gives

$$\begin{aligned}
\bar\partial(\partial_j \hat{u}_\nu) &= \partial_j(\bar\partial \hat{u}_\nu) \\
&= \widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t\partial_j \hat{u}_\nu + \partial_j \widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t \hat{u}_\nu + D_2\widehat{Q}_\nu(z,\hat{u}_\nu)\left(\partial_j \hat{u}_\nu, \partial_t \hat{u}_\nu\right) + \partial_j \hat{f}_\nu,
\end{aligned}$$

where $\partial_j \widehat{Q}_\nu$ means a partial derivative of $\widehat{Q}_\nu(s+it,x)$ with respect to s or t, and $D_2\widehat{Q}_\nu$ is its partial differential with respect to $x \in \mathbb{C}^n$. 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 $\widehat{Q}_\nu(z,\hat{u}_\nu)$ according to Theorem A.18 since $\hat{u}_\nu$ is of class $W^{k,p}$ with $kp > 2$ and $\widehat{Q}_\nu$ is of class C^k, and Proposition A.16 then justifies the product rule for $\widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t \hat{u}_\nu$ since $\widehat{Q}_\nu(z,\hat{u}_\nu) \in W^{k,p}$, $\partial_t \hat{u}_\nu \in W^{k-1,p}$, 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

$$\begin{aligned}
\bar\partial(\beta\,\partial_j \hat{u}_\nu) &= \beta\widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t\partial_j \hat{u}_\nu + \beta\partial_j\widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t \hat{u}_\nu + \beta D_2\widehat{Q}_\nu(z,\hat{u}_\nu)\left(\partial_j \hat{u}_\nu, \partial_t \hat{u}_\nu\right) \\
&\qquad + \beta\,\partial_j \hat{f}_\nu + (\bar\partial\beta)\partial_j \hat{u}_\nu \\
&= \widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t(\beta\,\partial_j \hat{u}_\nu) + D_2\widehat{Q}_\nu(z,\hat{u}_\nu)\left(\beta\,\partial_j \hat{u}_\nu, \partial_t \hat{u}_\nu\right) + \beta\partial_j\widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t \hat{u}_\nu \\
&\qquad + \beta\,\partial_j \hat{f}_\nu + (\bar\partial\beta)\partial_j \hat{u}_\nu - \widehat{Q}_\nu(z,\hat{u}_\nu)(\partial_t\beta)\partial_j \hat{u}_\nu,
\end{aligned}$$

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

$$\begin{aligned}
\bar\partial(\beta\,\partial_j \hat{u}_\nu) - \widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t(\beta\,\partial_j \hat{u}_\nu) &= D_2\widehat{Q}_\nu(z,\hat{u}_\nu)(\beta\,\partial_j \hat{u}_\nu, \partial_t \hat{u}_\nu) \\
&\quad + \left(\bar\partial\beta - \widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t\beta\right)\partial_j \hat{u}_\nu + \beta\left(\partial_j\widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t \hat{u}_\nu + \partial_j \hat{f}_\nu\right).
\end{aligned}$$

Combining this with (2.14) gives

$$\begin{aligned}
\|\beta\,\partial_j \hat{u}_\nu\|_{W^{k,p}} &\leqslant c\left\|\widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t(\beta\,\partial_j \hat{u}_\nu)\right\|_{W^{k-1,p}} + c\left\|D_2\widehat{Q}_\nu(z,\hat{u}_\nu)(\beta\,\partial_j \hat{u}_\nu, \partial_t \hat{u}_\nu)\right\|_{W^{k-1,p}} \\
&\quad + c\left\|\left(\bar\partial\beta - \widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t\beta\right)\partial_j \hat{u}_\nu + \beta\left(\partial_j\widehat{Q}_\nu(z,\hat{u}_\nu)\partial_t \hat{u}_\nu + \partial_j \hat{f}_\nu\right)\right\|_{W^{k-1,p}}.
\end{aligned} \tag{2.15}$$

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 (2.15) 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 $\widehat{Q}_\nu(z, \hat{u}_\nu)$ we observe that since $\widehat{Q}_\nu \to \widehat{Q}$ in C^k on $\mathbb{D} \times \mathbb{D}^{2n}$ and $\hat{u}_\nu$ can be assumed to lie in a $W^{k,p}$-small neighborhood of 0 for every ν, Proposition 2.7 places $\widehat{Q}_\nu(z, \hat{u}_\nu)$ into a $W^{k,p}$-small neighborhood of the function $z \mapsto \widehat{Q}(z, 0)$ 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. 2.6.

It remains to find a uniform $W^{k-1,p}$-bound for $\partial_j \widehat{Q}_\nu(z, \hat{u}_\nu) \partial_t \hat{u}_\nu$. For this, slightly different arguments are in order depending on whether $p > 2$ or $p \leqslant 2$. If $p > 2$, then $W^{k,p}$ has a continuous inclusion into C^{k-1}, thus $\hat{u}_\nu$ for every ν lies in a C^{k-1}-small neighborhood of 0 while $\partial_j \widehat{Q}_\nu$ converges in C^{k-1} to $\partial_j \widehat{Q}$, implying that $\partial_j \widehat{Q}_\nu(z, \hat{u}_\nu)$ is uniformly C^{k-1}-bounded. A $W^{k-1,p}$-bound on $\partial_j \widehat{Q}_\nu(z, \hat{u}_\nu) \partial_t \hat{u}_\nu$ then comes from the (obviously) continuous product pairing $C^{k-1} \times W^{k-1,p} \to W^{k-1,p}$. If on the other hand $p \leqslant 2$, then we necessarily have $k \geqslant 2$ since $kp > 2$, 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 $k - 1 - \frac{2}{\infty} = k - 1 \geqslant k - \frac{2}{p}$ for $p \leqslant 2$. Proposition 2.4 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 2.7, implying that $\partial_j \widehat{Q}_\nu(z, \hat{u}_\nu)$ is uniformly $W^{k-1,q}$-bounded. Since $(k-1)q > 2$ and $k - 1 - \frac{2}{q} \geqslant k - 1 - \frac{2}{p}$, Proposition 2.6 now gives a continuous product pairing $W^{k-1,q} \times W^{k-1,p} \to W^{k-1,p}$, which provides the desired $W^{k-1,p}$-bound on $\partial_j \widehat{Q}_\nu(z, \hat{u}_\nu) \partial_t \hat{u}_\nu$.

Step 2: The first term.
The tricky aspect of the first term in (2.15) 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 $\widehat{Q}_\nu(z, \hat{u}_\nu)$: 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

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

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

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

Step 3: The second term.
We observe first that since $D_2\widehat{Q}_\nu \to D_2\widehat{Q}$ in C^{k-1}, the same trick that was used to bound $\partial_j\widehat{Q}_\nu(z,\hat{u}_\nu)$ in step 1 furnishes $D_2\widehat{Q}_\nu(z,\hat{u}_\nu)$ with a uniform C^{k-1}-bound if $p > 2$, and a uniform $W^{k-1,q}$-bound for some $q \geqslant p$ with $(k-1)q > 2$ if $p \leqslant 2$, where in both cases the bounds can be assumed independent of the scaling parameter ϵ. 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

$$c\left\| D_2\widehat{Q}_\nu(z,\hat{u}_\nu)(\beta\,\partial_j\hat{u}_\nu, \partial_t\hat{u}_\nu)\right\|_{W^{k-1,p}} \leqslant c' \|\beta\,\partial_j\hat{u}_\nu\|_{W^{k,p}} \cdot \|\partial_t\hat{u}_\nu\|_{W^{k-1,p}}$$

for a constant $c' > 0$ that is independent of ν and ϵ. By (2.13), 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 (2.15) 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 2.29. Use an analogous argument via difference quotients to prove statement (1) in Theorem 2.24. *Hint: If you're anything like me, you might get stuck trying to estimate the difference quotient analogues of the terms in* (2.15) *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* (2.15)*, but the extra terms can be bounded.*

2.5. Linear local existence and applications

The following lemma can be applied in the case $A \in C^\infty(\mathbb{D}, \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n))$ 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 2.30. *Assume* $2 < p < \infty$ *and* $A \in L^p(\mathring{\mathbb{D}}, \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n))$. *Then for sufficiently small* $\epsilon > 0$, *the problem*

$$\bar{\partial}u + Au = 0$$
$$u(0) = u_0$$

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

Remark 2.31. 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 2.30. 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 2.12, 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 2.12 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 2.32 (Similarity principle). *Suppose $A : \mathbb{D} \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$ is smooth and $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ satisfies 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. After shrinking the domain if necessary, we may assume without loss of generality that the smooth solution $u : \mathring{\mathbb{D}} \to \mathbb{C}^n$ is bounded. Choose a map $C : \mathbb{D} \to \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n)$ satisfying $C(z)u(z) = A(z)u(z)$ and $|C(z)| \leqslant |A(z)|$ for almost every $z \in \mathbb{D}$. Then $C \in L^\infty(\mathring{\mathbb{D}}, \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n))$ and u is a weak solution to $(\bar\partial + C)u = 0$. Note that since we do not know anything about the zero set of u, we cannot assume C is continuous, but we have no trouble assuming $C \in L^p(\mathring{\mathbb{D}})$ for every $p > 2$.

Since $\bar{\partial}+C$ is now complex linear, we can use Lemma 2.30 to find a complex basis of $W^{1,p}$-smooth weak solutions to $(\bar{\partial}+C)v = 0$ on $\mathring{\mathbb{D}}_\epsilon$ that define the standard 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.16 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 2.21. $\square$

COROLLARY 2.33 (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.* $\square$

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 2.34. *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

$$\partial_s u + J(u(z))\partial_t u = 0 \tag{2.16}$$

and

$$\begin{aligned}
\partial_s v + J(u(z))\partial_t v &= \partial_s v + J(v(z))\partial_t v + \left[J(u(z)) - J(v(z))\right] \partial_t v \\
&= -\left[J(u(z) + h(z)) - J(u(z))\right] \partial_t v \\
&= -\left(\int_0^1 \frac{d}{d\tau} J(u(z) + \tau h(z))\, d\tau\right) \partial_t v \qquad (2.17)\\
&= -\left(\int_0^1 dJ(u(z) + \tau h(z)) \cdot h(z)\, d\tau\right) \partial_t v =: -A(z)h(z),
\end{aligned}$$

where the last step defines a smooth family of linear maps $A(z) \in \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$. Subtracting (2.16) from (2.17) 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 $f \equiv 0$. $\square$

2.6. Simple curves and multiple covers

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

THEOREM 2.35. *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 2.36 (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*

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

$\square$

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 2.32) if we additionally asume that either $du(z)$ or $dv(z')$ is

[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.

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

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

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 2.37 (Branching). *Suppose $u : (\Sigma, j) \to (W, J)$ is a nonconstant pseudo-holomorphic 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*

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

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]. 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 2.36 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 2.35. 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 $u(\mathcal{U} \backslash \{z\}) \cap u(\mathcal{U}' \backslash \{z'\}) = \emptyset$.

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. $\square$

LECTURE 3

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 lectures 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.

3.1. 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

$$\dim \mathcal{M}(x_-, x_+) = \operatorname{ind}(x_-) - \operatorname{ind}(x_+), \tag{3.1}$$

because the unstable manifold of x_- has dimension $\operatorname{ind}(x_-)$ and the stable manifold of x_+ has codimension $\operatorname{ind}(x_+)$. 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 $\operatorname{ind} \mathbf{D}_u = \operatorname{ind}(x_-) - \operatorname{ind}(x_+)$.

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 $s \to \infty$,

[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 lecture.

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 $\operatorname{ind}(x_-) - \operatorname{ind}(x_+)$ can now be expressed entirely 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

$$\operatorname{ind}(x_-) - \operatorname{ind}(x_+) = \dim E^-(A_-) - \dim E^-(A_+),$$

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 3.1. *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*

$$\operatorname{ind}(\mathbf{D}) = \dim E^-(A_-) - \dim E^-(A_+). \tag{3.4}$$

□

REMARK 3.2. The hypothesis that $A_\pm$ is invertible in Prop. 3.1 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 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 (cf. Theorem 3.19)

$$\mu^{\text{spec}}(A_-, A_+) := \#\left\{j \in I \;\middle|\; \lambda_j(-1) < 0 < \lambda_j(1)\right\} - \#\left\{j \in I \;\middle|\; \lambda_j(-1) > 0 > \lambda_j(1)\right\}.$$

The index formula (3.4) now becomes

$$\operatorname{ind}(\mathbf{D}) = \mu^{\text{spec}}(A_-, A_+).$$

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 next section.

3.2. 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 Lecture 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

$$\nabla \mathcal{A}_H(\gamma) = J_t(\gamma)\left(\dot{\gamma} - X_t(\gamma)\right) \in \Gamma(\gamma^* TM) =: T_\gamma C^\infty_{\text{contr}}(S^1, M)$$

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 $\gamma \in \operatorname{Crit}(\mathcal{A}_H)$ 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

$$\mathbf{A}_\gamma \eta := \nabla_\rho \left[\nabla \mathcal{A}_H(\gamma_\rho)\right]\big|_{\rho=0}.$$

The result is independent of the choice of connection since $\nabla \mathcal{A}_H(\gamma) = 0$.

EXERCISE 3.3. 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 contact manifold (M, ξ) with contact form α, induced Reeb vector field R_α, and a complex structure $J : \xi \to \xi$ compatible with the symplectic structure $d\alpha|_\xi$. Let

$$\pi_\xi : TM \to \xi$$

denote the projection along R_α. The **contact action functional** is defined by

$$\mathcal{A}_\alpha : C^\infty(S^1, M) \to \mathbb{R} : \gamma \mapsto \int_{S^1} \gamma^*\alpha.$$

The first variation of this functional for $\gamma \in C^\infty(S^1, M)$ and $\eta \in \Gamma(\gamma^* TM)$ is

$$d\mathcal{A}_\alpha(\gamma)\eta = \int_{S^1} d\alpha(\eta, \dot\gamma)\, dt = -\int_{S^1} d\alpha(\pi_\xi \dot\gamma, \eta)\, dt.$$

The functional has a built-in degeneracy since it is parametrization-invariant; in particular, $d\mathcal{A}_\alpha(\gamma)\eta = 0$ whenever η points in the direction of the Reeb vector field, a symptom of the fact that closed Reeb orbits always come in families related to each other by reparametrization. A loop $\gamma : S^1 \to M$ is critical for $\mathcal{A}_\alpha$ if and only if $\dot\gamma$ is everywhere tangent to R_α, allowing for an infinite-dimensional family of distinct perturbations—however, there exist preferred parametrizations, namely those for which $\dot\gamma$ is a *constant* multiple of R_α, meaning

$$\dot\gamma = T \cdot R_\alpha(\gamma), \qquad T := \mathcal{A}_\alpha(\gamma). \tag{3.5}$$

Such a loop corresponds to a T-periodic solution $x : \mathbb{R} \to M$ to $\dot x = R_\alpha(x)$, where $\gamma(t) = x(Tt)$.

The discussion above indicates that we cannot derive a "Hessian" of $\mathcal{A}_\alpha$ 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_α direction. To avoid this, we shall consider only preferred parametrizations $\gamma : S^1 \to M$ of the form (3.5), and perturbations in directions tangent to ξ, which is transverse to every Reeb orbit. For $\eta \in \Gamma(\gamma^*\xi)$, we then have

$$d\mathcal{A}_\alpha(\gamma)\eta = \int_{S^1} d\alpha(-J\pi_\xi\dot\gamma, J\eta)\, dt = \langle -J\pi_\xi\dot\gamma, \eta\rangle_{L^2},$$

where we define an L^2-product for sections of $\gamma^*\xi$ by

$$\langle \eta, \eta' \rangle_{L^2} := \int_{S^1} d\alpha(\eta, J\eta')\, dt. \tag{3.6}$$

It therefore seems sensible to write

$$\nabla\mathcal{A}_\alpha(\gamma) := -J\pi_\xi\dot\gamma \in \Gamma(\gamma^*\xi),$$

and we shall define the Hessian at a critical point γ as the linearization of $\nabla\mathcal{A}_\alpha$ in ξ directions, that is,

$$\nabla^2\mathcal{A}_\alpha(\gamma) : \Gamma(\gamma^*\xi) \to \Gamma(\gamma^*\xi).$$

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:

$$\begin{aligned}\nabla_\rho\left(\pi_\xi\dot\gamma_\rho\right)\big|_{\rho=0} &= \nabla_\rho\left[\partial_t\gamma_\rho - \alpha(\partial_t\gamma_\rho)R_\alpha(\gamma_\rho)\right]\big|_{\rho=0} \\ &= \nabla_t\eta - \alpha(\dot\gamma)\nabla_\eta R_\alpha - \partial_\rho\left[\alpha(\partial_t\gamma_\rho)\right]\big|_{\rho=0}\cdot R_\alpha(\gamma).\end{aligned}$$

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 relation

$$0 = T\cdot d\alpha(\eta, R_\alpha(\gamma)) = d\alpha(\partial_\rho\gamma_\rho, \partial_t\gamma_\rho)\big|_{\rho=0} = \partial_\rho\left[\alpha(\dot\gamma_\rho)\right]\big|_{\rho=0} - \partial_t\left[\alpha(\eta)\right] = \partial_\rho\left[\alpha(\dot\gamma_\rho)\right]\big|_{\rho=0},$$

implying

$$\nabla_\rho\left(\pi_\xi\dot\gamma_\rho\right)\big|_{\rho=0} = \nabla_t\eta - T\nabla_\eta R_\alpha \in \Gamma(\gamma^*\xi),$$

and thus

$$\nabla_\rho\left(-J\pi_\xi\dot\gamma_\rho\right)\big|_{\rho=0} = -J\left(\nabla_t\eta - T\nabla_\eta R_\alpha\right) \in \Gamma(\gamma^*\xi).$$

This calculation motivates the following definition.

DEFINITION 3.4. Given a loop $\gamma : S^1 \to M$ parametrizing a closed Reeb orbit in $(M, \xi = \ker\alpha)$ with period $T \equiv \alpha(\dot\gamma)$, the **asymptotic operator associated to** γ is the first-order differential operator on $\gamma^*\xi$ defined by

$$\mathbf{A}_\gamma : \Gamma(\gamma^*\xi) \to \Gamma(\gamma^*\xi) : \eta \mapsto -J(\nabla_t\eta - T\nabla_\eta R_\alpha)$$

EXERCISE 3.5. Show that $\mathbf{A}_\gamma$ is symmetric with respect to the L^2 inner product (3.6) on $\Gamma(\gamma^*\xi)$. Moreover, γ is nondegenerate (see §1.3) if and only if $\ker\mathbf{A}_\gamma$ is trivial. *Hint for nondegeneracy: Consider the pullback of $\gamma^*\xi$ via the cover $\mathbb{R} \to S^1 = \mathbb{R}/\mathbb{Z}$, and show that solutions to $\nabla_t\eta - T\nabla_\eta R_\alpha = 0$ on the pullback are given by operating on $\xi_{\gamma(0)}$ with the linearized Reeb flow. To see this, try differentiating families of solutions to the equation $\dot x = TR_\alpha(x)$.*

REMARK 3.6. The Reeb vector field R_α of a contact form α satisfies

$$\begin{aligned}\mathcal{L}_{R_\alpha}\alpha &= d\iota_{R_\alpha}\alpha + \iota_{R_\alpha}d\alpha = d(1) + d\alpha(R_\alpha,\cdot) \equiv 0, \text{ and} \\ \mathcal{L}_{R_\alpha}d\alpha &= d\iota_{R_\alpha}d\alpha + \iota_{R_\alpha}d^2\alpha = d\big(d\alpha(R_\alpha,\cdot)\big) \equiv 0,\end{aligned}$$

thus its flow preserves $\xi = \ker\alpha$ along with its symplectic vector bundle structure $d\alpha|_\xi$. Another way of phrasing the hint in the the above exercise is then as follows: $\mathbf{A}_\gamma$ can also be written as $-J\widehat\nabla_t$, where $\widehat\nabla_t$ is the unique *symplectic connection* on $(\gamma^*\xi, d\alpha)$ for which parallel transport is given by the linearized Reeb flow.

REMARK 3.7 (sign conventions). You might be slightly concerned about the sign difference between the formulas for asymptotic operators in Exercise 3.3 and Definition 3.4. 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 $\mathcal{A}_\alpha$. 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 Lecture 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: 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.

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 3.8. 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 3.6 shows that the asymptotic operator $\mathbf{A}_\gamma$ for a closed Reeb orbit γ is also a smooth asymptotic operator on $(\gamma^*\xi, J, d\alpha)$ in the sense of Definition 3.8.

[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 §3.5 relating eigenvalues of asymptotic operators to winding numbers therefore work out differently in the ECH context.

EXERCISE 3.9. 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 3.10. 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 $S : S^1 \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$, where we denote

$$\mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n}) := \left\{ B \in \mathrm{End}(\mathbb{R}^{2n}) = \mathbb{R}^{2n \times 2n} \;\middle|\; B^{\mathrm{T}} = B \right\}.$$

Hint: Use the fact that the difference between two connections is a bundle map, and deduce the symmetry of $S(t)$ from Exercise 3.9.

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 §3.3.

For technical reasons, we will sometimes need to consider a larger class of asymptotic operators whose zeroth-order terms are not necessarily smooth, nor even continuous. The weakest regularity condition we can impose without invalidating the discussion in the previous paragraph is the following:

DEFINITION 3.11. 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 $S \in L^\infty(S^1, \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n}))$. 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 3.12. Generalize Exercise 3.9 to prove that asymptotic operators of class L^∞ are also L^2-symmetric.

LEMMA 3.13. *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 $S : S^1 \to \operatorname{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n})$ 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.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.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

$$\operatorname{im}(\partial_t) = \left\{ f \in L^2(S^1, \mathbb{R}^{2n}) \;\middle|\; \int_{S^1} f(t)\,dt = 0 \right\},$$

which has codimension $2n$. □

COROLLARY 3.14. *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 3.11, 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 3.15. *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 3.16. 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 $x_0 \in \mathbb{R}^n$ 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 $T : X_R \to X_R$ and 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

$$\dot{x}(t) = A(t)x(t) \text{ for almost all } t, \qquad x(t_0) = x_0, \tag{3.7}$$

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 (3.7) is of class $W^{1,p}_{\mathrm{loc}}$. *Hint: For a useful characterization of $W^{1,p}_{\mathrm{loc}}(\mathbb{R})$, see Corollary A.12.*

PROPOSITION 3.17. *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 3.16 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

$$\frac{d}{dt} \Psi^{\mathrm{T}} J_0 \Psi = \Psi^{\mathrm{T}} (S^{\mathrm{T}} - S) \Psi \tag{3.8}$$

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.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 $L^{\infty}_{\mathrm{loc}}$ 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 (3.8), hence $\mathbf{A} = -J_0 \partial_t - S$ is an asymptotic operator. $\square$

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 3.18. 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.

3.3. 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

$$(\mathbf{A}\eta)(t) := -J_0 \, \partial_t \eta(t) - S(t)\eta(t), \tag{3.9}$$

where $S : S^1 \to \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n})$ 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 3.19. *Assume* $[-1,1] \to L^\infty(S^1, \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})) : s \mapsto S_s$ *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 the number* $\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+) \in \mathbb{Z}$ *defined by*

$$\#\left\{j \in \mathbb{Z} \;\middle|\; \lambda_j(-1) < 0 < \lambda_j(1)\right\} - \#\left\{j \in \mathbb{Z} \;\middle|\; \lambda_j(-1) > 0 > \lambda_j(1)\right\}$$

depends only on $\mathbf{A}_-$ *and* $\mathbf{A}_+$*.*

REMARK 3.20. Differentiability of the path $[-1,1] \to L^\infty(S^1, \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})) : s \mapsto S_s$ 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 3.19: first, if $S_-, S_+ \in L^\infty(S^1, \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n}))$ are given, then the linear interpolation

$$S_s := \frac{1}{2}(1-s)S_- + \frac{1}{2}(1+s)S_+$$

has a constant derivative $\frac{1}{2}(S_+ - S_-) \in L^\infty(S^1, \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n}))$ 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 3.19. The second class of examples will be especially useful for defining generic perturbations of paths of asymptotic operators: it arises from smooth functions $S : [-1,1] \times [0,1] \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$, 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 $L^\infty(S^1, \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n}))$. To see that $s \mapsto S_s$ is a smooth map $[-1,1] \to L^\infty(S^1, \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n}))$, 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 $\partial_s S : [-1,1] \times [0,1] \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ 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 3.21. 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 3.19 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 §3.3.1, and then specialize to the case of symmetric index 0 operators to define the abstract version of spectral flow in §3.3.2. In §3.3.3, we show that the spectra of such operators vary continuously under small perturbations, and in §3.3.4 we specialize further to operators of the form (3.9) 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 3.19.

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, in particular, does not require any results from [**Kat95**].

3.3.1. 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

$$\operatorname{Fred}_{\mathbb{F}}(X,Y) \subset \mathscr{L}_{\mathbb{F}}(X,Y)$$

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

$$\operatorname{ind}_{\mathbb{F}}(\mathbf{T}) := \dim_{\mathbb{F}} \ker \mathbf{T} - \dim_{\mathbb{F}} \operatorname{coker} \mathbf{T} \in \mathbb{Z}.$$

The index defines a continuous and thus locally constant function $\operatorname{Fred}_{\mathbb{F}}(X,Y) \to \mathbb{Z}$, and for each $i \in \mathbb{Z}$, we shall denote

$$\operatorname{Fred}^i_{\mathbb{F}}(X,Y) := \left\{ \mathbf{T} \in \operatorname{Fred}_{\mathbb{F}}(X,Y) \;\middle|\; \operatorname{ind}(\mathbf{T}) = i \right\}.$$

We will often have occasion to use the following general construction. Given $\mathbf{T}_0 \in \operatorname{Fred}_{\mathbb{F}}(X,Y)$, 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

$$\Phi : \mathcal{O} \to \mathrm{Hom}_{\mathbb{F}}(\ker \mathbf{T}_0, \operatorname{coker} \mathbf{T}_0) : \mathbf{T} \mapsto \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}. \tag{3.10}$$

LEMMA 3.22. *The map Φ in (3.10) 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

$$\begin{aligned} d\Phi(\mathbf{T}_0) : \mathscr{L}_{\mathbb{F}}(X,Y) &\to \mathrm{Hom}_{\mathbb{F}}(K,C) \\ \begin{pmatrix} \mathbf{A}' & \mathbf{B}' \\ \mathbf{C}' & \mathbf{D}' \end{pmatrix} &\mapsto \mathbf{D}'. \end{aligned}$$

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, $\ker \mathbf{T}\Psi(\mathbf{T}) = \{0\} \oplus \ker \Phi(\mathbf{T})$. $\square$

PROPOSITION 3.23. *For each $i \in \mathbb{Z}$ and each nonnegative integer $k \geqslant i$, the subset*

$$\mathrm{Fred}_{\mathbb{F}}^{i,k}(X,Y) := \left\{ \mathbf{T} \in \mathrm{Fred}_{\mathbb{F}}^{i}(X,Y) \;\middle|\; \dim_{\mathbb{F}} \ker \mathbf{T} = k \text{ and } \dim_{\mathbb{F}} \operatorname{coker} \mathbf{T} = k - i \right\}$$

[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*

$$\operatorname{codim}_{\mathbb{F}} \operatorname{Fred}^{i,k}_{\mathbb{F}}(X, Y) = k(k - i).$$

Moreover, the set

$$X^{i,k} := \left\{ (\mathbf{T}, x) \in \operatorname{Fred}^{i,k}_{\mathbb{F}}(X, Y) \times X \;\middle|\; x \in \ker \mathbf{T} \right\}$$

is a smooth (and holomorphic if $\mathbb{F} = \mathbb{C}$*) subbundle of the trivial vector bundle* $\operatorname{Fred}^{i,k}_{\mathbb{F}}(X, Y) \times X \to \operatorname{Fred}^{i,k}_{\mathbb{F}}(X, Y)$*.*

PROOF. Applying the implicit function theorem to the map Φ from Lemma 3.22 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 $\Phi^{-1}(0) = \operatorname{Fred}^{i,k}_{\mathbb{F}}(X, Y)$ 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 3.22. 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 $V^{i,k}$, i.e. this restriction is the inverse of a local trivialization of $V^{i,k}$. □

For real-linear operators of index 0, one can use Prop. 3.23 to define the following "relative" invariant. Suppose $\{\mathbf{T}(s) \in \operatorname{Fred}^0_{\mathbb{R}}(X, Y)\}_{s \in [-1,1]}$ 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

$$\mu^{\text{spec}}_{\mathbb{Z}_2}(\{\mathbf{T}(s)\}) \in \mathbb{Z}_2$$

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 $\{\mathbf{T}(s) \in \operatorname{Fred}^0_{\mathbb{R}}(X, Y)\}_{s \in [-1,1]}$ are transverse to $\operatorname{Fred}^{0,k}_{\mathbb{R}}(X, Y)$ for every $k \in \mathbb{N}$, which implies via the codimension formula in Prop. 3.23 that they never intersect $\operatorname{Fred}^{0,k}_{\mathbb{R}}(X, Y)$ for $k \geqslant 2$, and their intersections with $\operatorname{Fred}^{0,1}_{\mathbb{R}}(X, Y)$ are transverse and thus isolated. Second, transversality also holds for generic homotopies

$$[0, 1] \times [-1, 1] \to \operatorname{Fred}^0_{\mathbb{R}}(X, Y) : (\tau, s) \mapsto \mathbf{T}_{\tau}(s)$$

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 $\operatorname{Fred}^{0,k}_{\mathbb{R}}(X, Y)$ 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 3.24. 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 3.25. 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.

3.3.2. 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 §3.3.3 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 3.29 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 3.26. *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. 3.23.

PROPOSITION 3.27. *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 $\mathbf{T}_0 \in \mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},k}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$, fix the splittings $\mathcal{D} = V \oplus K$ and $\mathcal{H} = W \oplus K$ as in Lemma 3.26. Using these in the construction of the map Φ from (3.10) produces a neighborhood $\mathcal{O} \subset \mathrm{Fred}_{\mathbb{F}}^{0}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ of $\mathbf{T}_0$ such that, by Lemma 3.22, $\{\mathbf{T} \in \mathcal{O} \mid \dim_{\mathbb{F}} \ker \mathbf{T} = k\} = \Phi^{-1}(0)$, 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 $\Phi(\mathbf{T}) \in \mathrm{End}_{\mathbb{F}}^{\mathrm{sym}}(K)$, where $\mathrm{End}_{\mathbb{F}}^{\mathrm{sym}}(K) \subset \mathrm{End}_{\mathbb{F}}(K)$ 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 $\mathcal{O} \cap \mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},k}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}}) = \Phi^{-1}(0)$ with Φ regarded as a smooth map $\mathcal{O} \cap \mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}}) \to \mathrm{End}_{\mathbb{F}}^{\mathrm{sym}}(K)$. The derivative at $\mathbf{T}_0$ again takes the form

$$d\Phi(\mathbf{T}_0) : \mathscr{L}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{H}) \to \mathrm{End}_{\mathbb{F}}^{\mathrm{sym}}(K) : \begin{pmatrix} \mathbf{A}' & \mathbf{B}' \\ \mathbf{C}' & \mathbf{D}' \end{pmatrix} \mapsto \mathbf{D}',$$

where now the block matrix represents an element of $\mathscr{L}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{H})$ 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 $\dim_{\mathbb{R}} \mathrm{End}_{\mathbb{F}}^{\mathrm{sym}}(K)$. The vector bundle structure of $\mathcal{D}^{\mathrm{sym},k}$ can be defined using the map Ψ from Lemma 3.22 just as in the non-symmetric case.

Finally, we observe that in the case $k = 1$, the above identifies $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},1}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ locally with the zero set of a submersion to $\mathrm{End}_{\mathbb{F}}^{\mathrm{sym}}(K)$, which is a real 1-dimensional vector space since K is a 1-dimensional vector space over $\mathbb{F}$. The canonical isomorphism

$$\mathbb{R} \to \mathrm{End}_{\mathbb{F}}^{\mathrm{sym}}(K) : a \mapsto a\mathbb{1}$$

thus determines a co-orientation on $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},1}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$. $\square$

The canonical co-orientation of $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},1}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ makes it natural to define signed intersection numbers between $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},1}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ and smooth paths in the ambient space $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$. The codimensions of $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},k}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ 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 3.28. Suppose $\mathbf{T}_+, \mathbf{T}_- \in \mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ 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 $\mathbf{T} : [-1,1] \to \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ with $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$, where the latter is assumed to carry the co-orientation given by Prop. 3.27, and $\mathbf{T} : [-1,1] \to \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ is any smooth path that is transverse to $\mathrm{Fred}^{\mathrm{sym},k}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ for every $k \geqslant 1$ and satisfies $\mathbf{T}(\pm 1) = \mathbf{T}_{\pm}$.

Note that since $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ is an affine space over $\mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{H})$, all paths in $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ from $\mathbf{T}_-$ to $\mathbf{T}_+$ are homotopic, so one can argue as we did for $\mu^{\mathrm{spec}}_{\mathbb{Z}_2}$ at the end of §3.3.1 that $\mu^{\mathrm{spec}}(\mathbf{T}_-,\mathbf{T}_+)$ is independent of the choice of path.

3.3.3. Perturbation of eigenvalues. Continuing in the setting of the previous subsection, we shall now regard each $\mathbf{T} \in \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ 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 $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$. 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 3.29. 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 3.41), the latter amounts to a condition on "regularity of weak solutions". This condition implies that the inclusion $\ker \mathbf{T} \hookrightarrow (\mathrm{im}\, \mathbf{T})^{\perp}$—valid for all symmetric operators—is also surjective, so if $\mathbf{T} : \mathcal{D} \to \mathcal{H}$ is Fredholm, it is then automatic that $\mathrm{ind}(\mathbf{T}) = 0$.

PROPOSITION 3.30. *Assume $\mathbf{T}_0 \in \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$. 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 $\mathcal{O} \subset \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ 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 3.26, 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 (3.10) 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

$$\Phi : \mathcal{O}(\mathbf{T}_0) \times \mathbb{D}_\epsilon(\lambda_0) \to \operatorname{End}_{\mathbb{C}}(K) : (\mathbf{T}, \lambda) \mapsto (\mathbf{D} - \lambda) - \mathbf{C}\left(\mathbf{A} - \lambda\right)^{-1} \mathbf{B}$$

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

$$\mathbb{D}_\epsilon(\lambda_0) \to \mathbb{C} : \lambda \mapsto \det \Phi(\mathbf{T}_0, \lambda). \tag{3.11}$$

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 3.31 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 3.31. 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 $\mu^{\mathrm{spec}}(\mathbf{T}_-, \mathbf{T}_+)$, the spectral flow is indeed a signed count of eigenvalues crossing 0.

PROPOSITION 3.32. *Suppose $\{\mathbf{T}_s \in \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})\}_{s\in(-1,1)}$ 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 $\{\mathbf{T}_s - \lambda_0 \in \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})\}_{s\in(-1,1)}$ with $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ 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 3.30, we can find small numbers $\epsilon > 0$ and $\delta > 0$ such that

$$\left\{(s, \lambda) \in (-\epsilon, \epsilon) \times (\lambda_0 - \delta, \lambda_0 + \delta) \;\middle|\; \lambda \in \sigma(\mathbf{T}_s)\right\} = \Phi^{-1}(0),$$

where

$$\Phi : (-\epsilon, \epsilon) \times (\lambda_0 - \delta, \lambda_0 + \delta) \to \mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K) : (s, \lambda) \mapsto (\mathbf{D}_s - \lambda) - \mathbf{C}_s\,(\mathbf{A}_s - \lambda)^{-1}\,\mathbf{B}_s,$$

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 $K = \ker(\mathbf{T}_0 - \lambda_0)$, $W = \mathrm{im}(\mathbf{T}_0 - \lambda_0)$ 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 $\mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K)$ 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 $\partial_s \mathbf{D}_s|_{s=0} \in \mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K)$ with a real number via the natural isomorphism $\mathrm{End}^{\mathrm{sym}}_{\mathbb{F}}(K) = \mathbb{R}$. 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 3.27, the intersection of $\{\mathbf{T}_s\}_{s\in(-1,1)}$ with $\mathrm{Fred}^{\mathrm{sym},1}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ 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 $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$.

LEMMA 3.33. *Suppose $\{\mathbf{K}_s \in \mathscr{L}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{H})\}_{s\in(a,b)}$ 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 $\mathbf{T}_s := \mathbf{T}_{\mathrm{ref}} + \mathbf{K}_s \in \mathrm{Fred}^{\mathrm{sym}}_{\mathbb{F}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$. Then*

$$|\dot{\lambda}(s)| \leqslant \|\partial_s \mathbf{K}_s\|_{\mathscr{L}(\mathcal{H})} \qquad \text{for all } s \in (a, b).$$

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 3.27 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

$$|\dot{\lambda}(s)| \leqslant \|x(s)\|_{\mathcal{H}} \|\dot{\mathbf{K}}_s\|_{\mathscr{L}(\mathcal{H})} \|x(s)\|_{\mathcal{H}} = \|\dot{\mathbf{K}}_s\|_{\mathscr{L}(\mathcal{H})}.$$

□

3.3.4. 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 $S : S^1 \to \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n})$ 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 3.13, and by Remark 3.21, any smooth path $s \mapsto S_s$ in $L^\infty(\mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n}))$ 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 3.34. We defined the topology of $\mathrm{Fred}^{\mathrm{sym}}_{\mathbb{R}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ in §3.3.2 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 3.21), 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 3.11.

The proof of Theorem 3.19 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 Lecture 8 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 3.35 (see Appendix C). *Fix a smooth path* $[-1,1] \to L^\infty(\mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})) : s \mapsto S_s$ *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* $B : [-1,1] \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ *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

$$(-1,1) \to \mathrm{Fred}_{\mathbb{R}}^{\mathrm{sym}}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}}) : s \mapsto \mathbf{A}_s$$

with $\mathrm{Fred}_{\mathbb{R}}^{\mathrm{sym},1}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ *are transverse.*

□

PROOF OF THEOREM 3.19. Given a family $\{\mathbf{A}_s\}_{s\in[-1,1]}$ as specified in the theorem, use Lemma 3.35 to obtain a C^∞-small zeroth-order perturbation making all eigenvalues simple for $s \in (-1,1)$ and all intersections with $\mathrm{Fred}_{\mathbb{R}}^{\mathrm{sym},1}(\mathcal{D},\mathcal{H})$ transverse. Proposition 3.32 then implies that the eigenvalues depend smoothly on s, and Lemma 3.33 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 3.30 ensures moreover that these continuous families hit every eigenvalue with the correct multiplicity at $s = \pm 1$, and by Proposition 3.32, 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 $\{\mathbf{A}_s^\nu = \mathbf{A}_s + B^\nu(s,\cdot)\}_{s\in[-1,1]}$ such that $B^\nu : [-1,1] \times S^1 \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ is C^∞-convergent to 0 as $\nu \to \infty$. Lemma 3.33 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 3.36. It is important to understand that the definition of spectral flow depends on the particular co-orientation of $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},1}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ that arose in the proof of Prop. 3.27. We saw in Prop. 3.32 that this is indeed the *right* co-orientation to use if we want to interpret signed intersections with $\mathrm{Fred}_{\mathbb{F}}^{\mathrm{sym},1}(\mathcal{D},\mathcal{H},\mathbf{T}_{\mathrm{ref}})$ as signed crossing numbers of eigenvalues. In the non-symmetric setting of §3.3.1, one can show that $\mathrm{Fred}_{\mathbb{R}}^{0,1}(X,Y)$ is also co-orientable—this is obvious in the finite-dimensional case since $\mathrm{Fred}_{\mathbb{R}}^{0,1}(\mathbb{R}^n,\mathbb{R}^n)$ is then a regular level set of the determinant function. Moreover, $\mathrm{Fred}_{\mathbb{R}}^{0,1}(\mathbb{R}^n,\mathbb{R}^n)$ is connected (see Exercise 3.37 below), so the co-orientation is unique up to a sign. One can therefore lift the $\mathbb{Z}_2$-valued spectral flow of §3.3.1 to $\mathbb{Z}$, but as in Exercise 3.25, 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 3.37. 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. 3.27 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 3.38. 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 3.39. 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 3.19.

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 3.40. *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 3.54 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 3.19 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 3.41. Prove:

(a) Asymptotic operators are self-adjoint (as unbounded operators on L^2 with domain H^1) in the sense of Remark 3.29.
(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.*

3.4. The Conley-Zehnder index

We are now in a position to define a suitable replacement for the Morse index in the context of SFT. It will take 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 will have the important property that its values fully classify the connected components of the space of nondegenerate asymptotic operators. Recall from §3.2 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 3.42. Suppose $S \in \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^2)$ 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 $S \in \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n})$ 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 3.43. Show that the space of matrices on $\mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n})$ unitarily equivalent to a diagonal matrix with n positive and n negative eigenvalues (counting multiplicity) is connected.

DEFINITION 3.44. 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 $S \in \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ 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 3.42 and Exercise 3.43 show that this definition does not depend on the choice of a constant matrix $S \in \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ 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 3.45. 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

$$\mu_{\mathrm{CZ}}^{\tau}(\mathbf{A}) \in \mathbb{Z}$$

defined by choosing any unitary trivialization homotopic to τ 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 3.44.

If γ is a nondegenerate Reeb orbit γ in a $(2n-1)$-dimensional contact manifold $(M, \xi = \ker\alpha)$, then for any symplectic trivialization τ of $\gamma^*\xi \to S^1$, the **Conley-Zehnder index** of γ relative to τ is defined as

$$\mu_{\mathrm{CZ}}^{\tau}(\gamma) := \mu_{\mathrm{CZ}}^{\tau}(\mathbf{A}_\gamma).$$

It is clear from the definition 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, the definition above for $\mu_{\mathrm{CZ}}^{\tau}(\mathbf{A})$ 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.

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 3.17), and is elucidated in Definition 3.49 below.

EXERCISE 3.46. 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 3.47. 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 3.16 contains a useful result about the continuous dependence of solutions to ODEs on parameters in the equation.*

EXERCISE 3.48. 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 3.18 and 3.20 for a few useful observations.*

DEFINITION 3.49. The **Conley-Zehnder index** $\mu_{\mathrm{CZ}}(\Psi) \in \mathbb{Z}$ of a nondegenerate symlectic 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 3.49 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 3.17 and Remark 3.18, 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 3.46 and 3.48, supplemented by the fact that a continuous homotopy between two smooth paths always admits a C^0-small perturbation to a smooth homotopy.

REMARK 3.50. 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 3.5 and Remark 3.6). Thus if one prefers as in [SZ92] to speak in terms of nondegenerate symplectic arcs instead of asymptotic operators, $\mu_{\mathrm{CZ}}^\tau(\gamma)$ 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 3.51. 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 $j = 1,2$,

$$\mu_{\mathrm{CZ}}^{\tau_1\oplus\tau_2}(\mathbf{A}_1 \oplus \mathbf{A}_2) = \mu_{\mathrm{CZ}}^{\tau_1}(\mathbf{A}_1) + \mu_{\mathrm{CZ}}^{\tau_2}(\mathbf{A}_2).$$

REMARK 3.52. 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, so the induced contact structure is a rank 0 bundle, it 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 3.51, the only reasonable convention is to set

$$\mu_{\mathrm{CZ}}(\gamma) = \mu_{\mathrm{CZ}}(\mathbf{A}_\gamma) = 0.$$

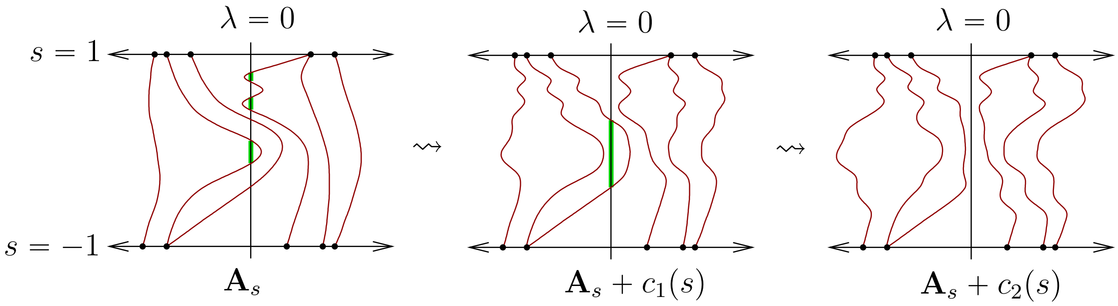

FIGURE 3.1. Modifying a path of asymptotic operators $\{\mathbf{A}_s\}_{s\in[-1,1]}$ with zero spectral flow to produce a path of nondegenerate operators.

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 14.36).

Here is the main result about Conley-Zehnder indices.

THEOREM 3.53. *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 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 $\mathbf{A}_{\pm 1} = \mathbf{A}_\pm$ such that (after perturbing it via Lemma 3.35) 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 3.47. $\square$

3.5. Winding numbers of eigenfunctions

To compute Conley-Zehnder indices, Exercise 3.51 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 3.54. *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 $S \in L^\infty(S^1, \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n}))$. 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 [**HWZ95a**].

Statement (1) follows from the generalized existence/uniqueness result of Exercise 3.16 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 3.16 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 3.35 so that all eigenvalues of $\mathbf{A}_s$ for $s \in (0,1)$ are simple. The same argument as in the proof of Theorem 3.19 (combining Propositions 3.30 and 3.32 and Lemma 3.33) 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 3.27 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

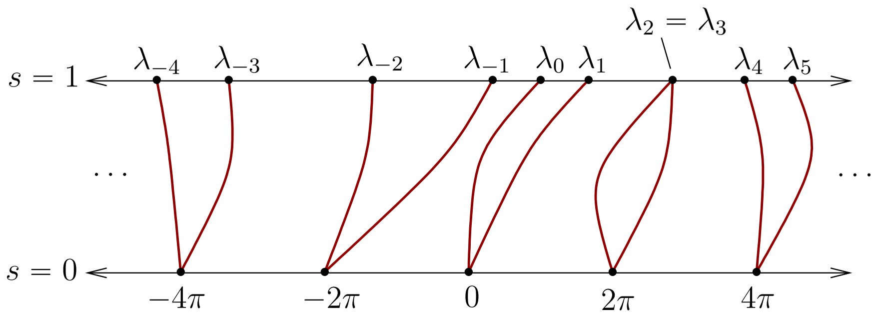

FIGURE 3.2. Generic deformation of eigenvalues from $\mathbf{A}_0 = -J_0\partial_t$ to an arbitrary asymptotic operator $\mathbf{A}$ in the proof of Theorem 3.54.

$\eta_j(s) \in H^1(S^1, \mathbb{R}^2)$. 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 $\text{wind}(\eta_j(s)) \in \mathbb{Z}$ 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 $\mathbb{Z} \to \mathbb{Z} : j \mapsto \text{wind}(\eta_j(s))$ is monotone increasing and attains every value exactly twice. □

The theorem implies the existence of a well-defined and nondecreasing function

$$\sigma(\mathbf{A}) \to \mathbb{Z} : \lambda \mapsto \text{wind}(\lambda),$$

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

$$\begin{aligned} \alpha_+(\mathbf{A}) &= \min_{\lambda \in \sigma(\mathbf{A}) \cap (0,\infty)} \text{wind}(\lambda) \in \mathbb{Z}, \\ \alpha_-(\mathbf{A}) &= \max_{\lambda \in \sigma(\mathbf{A}) \cap (-\infty,0)} \text{wind}(\lambda) \in \mathbb{Z}, \\ p(\mathbf{A}) &= \alpha_+(\mathbf{A}) - \alpha_-(\mathbf{A}) \geqslant 0. \end{aligned} \tag{3.12}$$

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 3.54 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 3.55. *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 $\mu_{\mathrm{CZ}}(\mathbf{A}_0) = 0$ 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 3.35) 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 $\alpha^\tau_\pm(\mathbf{A}) \in \mathbb{Z}$ 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 Reeb orbit γ in a contact 3-manifold $(M, \xi = \ker\alpha)$ with a trivialization τ of $\gamma^*\xi$ the integers $\alpha^\tau_\pm(\gamma)$ and $p(\gamma)$, such that if γ is nondegenerate, then $p(\gamma) \in \{0, 1\}$ and

$$\mu^\tau_{\mathrm{CZ}}(\gamma) = 2\alpha^\tau_-(\gamma) + p(\gamma) = 2\alpha^\tau_+(\gamma) - p(\gamma).$$

EXERCISE 3.56. 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),

$$\mu^{\tau_2}_{\mathrm{CZ}}(\mathbf{A}) = \mu^{\tau_1}_{\mathrm{CZ}}(\mathbf{A}) + 2\deg(\tau_2 \circ \tau_1^{-1}).$$

Exercise 3.56 provides the useful formula

$$\mu^{\tau_2}_{\mathrm{CZ}}(\gamma) = \mu^{\tau_1}_{\mathrm{CZ}}(\gamma) + 2\deg(\tau_2 \circ \tau_1^{-1})$$

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**

$$\mu^{\mathbb{Z}_2}_{\mathrm{CZ}}(\gamma) := [\mu^\tau_{\mathrm{CZ}}(\gamma)] \in \mathbb{Z}_2$$

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 $\dot\gamma = T \cdot R_\alpha(\gamma)$, 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 it is **multiply covered** if $k \geqslant 2$. We say γ is **simply covered** if it is not the k-fold cover of another Reeb orbit for any $k \geqslant 2$. 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 3.17 that $\mathbf{A}_{\gamma^k}$ is given by

$$\mathbf{A}_{\gamma^k} = -J_0\partial_t - kS(kt).$$

This implies:

PROPOSITION 3.57. *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 $\mathbf{A}_{\gamma^k} e_\lambda^k = k\lambda e_\lambda^k$.* $\square$

EXERCISE 3.58. 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 3.54 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.

3.6. 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 3.59. Assume $(M, \xi = \ker\alpha)$ is a 3-dimensional contact manifold, $\gamma : S^1 \to M$ parametrizes a nondegenerate Reeb orbit of period $T \equiv \alpha(\dot\gamma) > 0$, 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 $\lambda_1 \ne 1 \ne \lambda_2$ 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 3.59, 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 $\mathrm{U}(1) \cap (-\infty, 0)$ 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 under any smooth deformation of contact forms $\{\alpha_s\}_{s\in[0,1]}$ for which γ is a nondegenerate Reeb orbit for every α_s, it cannot deform from positive hyperbolic to either of the other two categories, i.e. if it is positive hyperbolic at $s = 0$, then it remains so at $s = 1$. 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 3.63 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 3.60. 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 3.54; 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 3.55 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 3.61. Show that the asymptotic operators of Example 3.60 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 $f''(0)/g''(0)$.

[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 Lecture 4 onwards). Every asymptotic operator is elliptic in the latter sense, but not in the dynamical sense under consideration here.

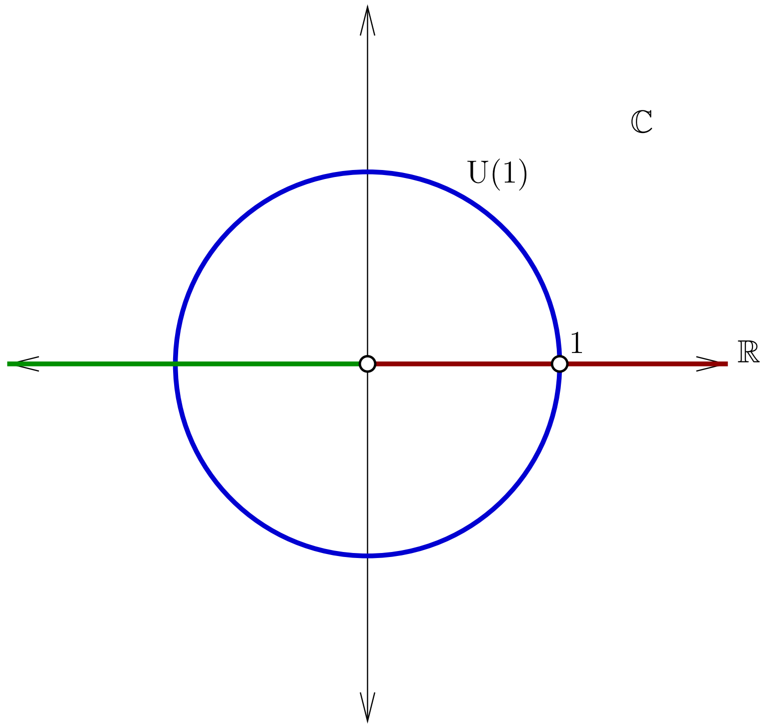

FIGURE 3.3. Every symplectic linear map in dimension two has spectrum contained in $(\mathbb{R}\backslash\{0\}) \cup \mathrm{U}(1) \subset \mathbb{C}$, but eigenvalues cannot move between $(0, \infty)$ and $(-\infty, 0) \cup \mathrm{U}(1)$ without crossing 1.

EXAMPLE 3.62. 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 3.56, 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 $\Psi(1)$.

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 3.63. *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 3.60 and 3.62, 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 $\Delta(\Psi) \cap \mathbb{Z} \neq \emptyset$ 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 3.53 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 3.60. 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 3.62 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 $\Delta(\Psi_s) \cap \mathbb{Z} = \emptyset$ 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 3.60.

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 3.62. □

Using the direct sum property in Exercise 3.51, one derives from Theorem 3.63 the following alternative characterization of the Conley-Zehnder index in higher dimensions (cf. [**FH93**, Proposition 5] or [**Sch95**, Theorem 3.3.7]):

COROLLARY 3.64. *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* $\omega_0 = \mathrm{Re}\langle i\cdot, \cdot\rangle$ *and identifying* $\mathbb{R}^{2n}$ *with* $\mathbb{C}^n$ *makes all of these into paths in* $\mathrm{Sp}(2n)$. *Then every nondegenerate symplectic arc* $\Psi : [0,1] \to \mathrm{Sp}(2n)$ *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*

$$\mu_{CZ}(\Phi_k) = 2k + n, \qquad \mu_{CZ}(\Psi_k) = 2k + n - 1.$$

□

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 3.65. *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 3.57, the k-fold cover e_λ^k is likewise an eigenfunction of $\mathbf{A}^k$ with eigenvalue $k\lambda < 0$, so its winding $\mathrm{wind}(e_\lambda^k) = k\,\mathrm{wind}(e_\lambda) = k\alpha_-(\mathbf{A})$ 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 3.66. *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 3.67. *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 3.11);
(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 3.65 implies

(3.13)
$$\frac{\alpha_-(\mathbf{A}^{k_j})}{k_j} = \frac{m\alpha_-(\mathbf{A}^{k_j})}{k_{j+1}} \leqslant \frac{\alpha_-(\mathbf{A}^{k_{j+1}})}{k_{j+1}} \leqslant \frac{\alpha_+(\mathbf{A}^{k_{j+1}})}{k_{j+1}} \leqslant \frac{m\alpha_+(\mathbf{A}^{k_j})}{k_{j+1}} = \frac{\alpha_-(\mathbf{A}^{k_j})}{k_j},$$

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 (3.13), $\left[\frac{\alpha_-(\mathbf{A}^N)}{N}, \frac{\alpha_+(\mathbf{A}^N)}{N}\right]$ 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 $\alpha_\pm(\mathbf{A}_0^k) = \operatorname{wind}(e_{\lambda_0^\pm})$, 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 3.30 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 3.55 and 3.63, $\mathbf{A}$ 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 $\operatorname{Sp}(2)$ to $\begin{pmatrix} e^\epsilon & a \\ 0 & e^{-\epsilon} \end{pmatrix}$ for $\epsilon > 0$ small, and the latter can then be realized using Proposition 3.17 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 $\mu_{CZ}(\mathbf{A})$ odd, so Theorem 3.55 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 $\mu_{CZ}(\mathbf{A})$ is shifted by twice the same integer (cf. Exercise 3.56), so we can use this trick to assume without loss of generality that $[\alpha_-(\mathbf{A}), \alpha_+(\mathbf{A}] = [0, 1]$, $\mu_{CZ}(\mathbf{A}) = 1$ and $\theta(\mathbf{A}) = 0$. Then for every $k \in \mathbb{N}$, Lemma 3.65 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 3.63 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 $\mu_{CZ}(\mathbf{A}^k)$ is also odd, implying in this situation that $\mu_{CZ}(\mathbf{A}^k) = 2\alpha_-(\mathbf{A}^k) + 1 = 1$. But if Ψ_k denotes the parallel transport of $\mathbf{A}^k$, Theorem 3.63 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 $\mu_{CZ}(\mathbf{A}^k) = -1$ 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 3.66. By the results of steps 3 and 4, it follows that $\theta(\mathbf{A}) \in \frac{1}{2}\mathbb{Z}$. But since $\mu_{CZ}(\mathbf{A})$ is odd by Theorem 3.63, 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 3.66 implies that $\mathbf{A}^2$ is either elliptic or negative hyperbolic, so step 5 and Theorem 3.63 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 3.67 gives rise to the formulas

$$\alpha_-(\mathbf{A}) = \lfloor \theta(\mathbf{A}) \rfloor, \qquad \alpha_+(\mathbf{A}) = \lceil \theta(\mathbf{A}) \rceil, \qquad \text{if} \quad \ker \mathbf{A} = \{0\}. \tag{3.14}$$

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 1.25). In this situation, Lemma 3.66 implies that all covers of an elliptic orbit are also elliptic, so one deduces from Theorem 3.67 that the corresponding monodromy angle must be irrational. Combining these observations with the relation between $\mu_{CZ}(\mathbf{A})$ and $\alpha_\pm(\mathbf{A})$ in Theorem 3.54, one now obtains the following result for multiply covered Reeb orbits:

COROLLARY 3.68. *Suppose γ is a nondegenerate Reeb orbit in a contact 3-manifold $(M, \xi = \ker\alpha)$ such that the multiple covers γ^k are also nondegenerate for every $k \in \mathbb{N}$. Choose a symplectic trivialization τ of ξ 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*
$$\mu_{CZ}^\tau(\gamma^k) = k\mu_{CZ}^\tau(\gamma)$$
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*
$$\mu_{CZ}^\tau(\gamma^k) = 2\lfloor k\theta \rfloor + 1 = 2\lceil k\theta \rceil - 1$$
for every $k \in \mathbb{N}$.

□

REMARK 3.69 (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 3.7, 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 $\mu_{CZ}(\Psi)$ that are equivalent to ours, but introduced modified indices for orbits in order to absorb the sign difference, e.g. [**Sal99**] defines $\mu_H(\gamma) := n - \mu_{CZ}(\Psi)$ for the linearized flow Ψ along an orbit γ. For the reasons why the latter is a natural convention in that context, see Theorem 10.30 and Remark 10.31.

LECTURE 4

Fredholm theory with cylindrical ends

Contents

In this lecture 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 lecture 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 goal in this lecture 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 lecture 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 $(Z_\pm^R, i)$ denote the half-cylinders

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

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 disk $\dot{\mathbb{D}} := \mathbb{D} \backslash \{0\}$ via the maps

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

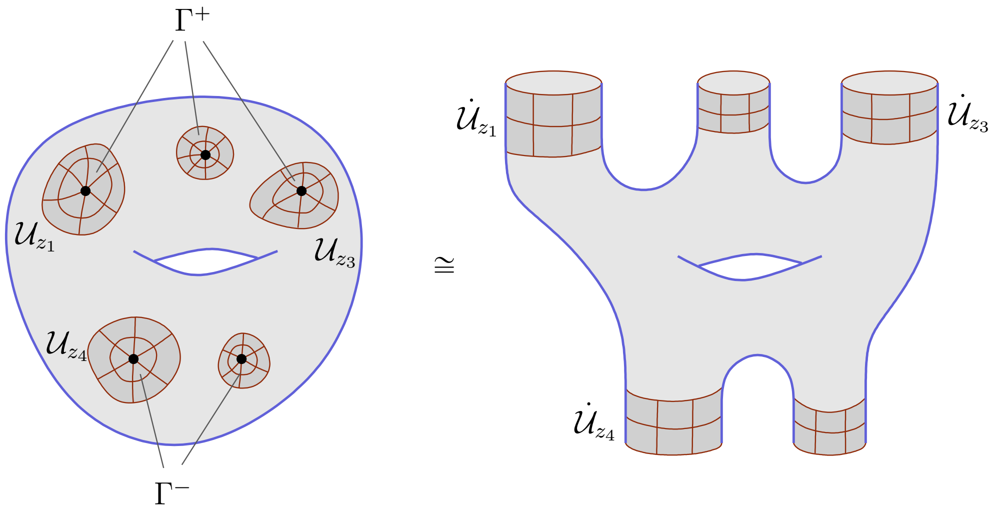

FIGURE 4.1. A Riemann surface with genus 1 and five punctures, depicted at the right as three positive and two negative cylindrical ends.

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 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. 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 lecture 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 $C^{m+1} \times W^{k,p} \to W^{k,p}$.

DEFINITION 4.2. 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.

[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 2.24 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.

EXERCISE 4.3. 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 4.4. Suppose $E \to \dot{\Sigma}$ is an asymptotically Hermitian vector bundle of class C^{m+1} for some $m \in \{0, 1, 2, \ldots, \infty\}$, $\mathbf{A}_z$ 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

$$\|S - S_\infty\|_{C^k(Z_\pm^R)} \to 0 \quad \text{as} \quad R \to \infty$$

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 lecture is to prove the following:

THEOREM 4.5. *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 $k \in \{1, \ldots, m+1\}$ 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 4.6. We assume $m \geqslant 1$ in Theorem 4.5 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 4.31). In any case, the applications in this book will only require the case $m = \infty$.

The index of $\mathbf{D}$ is determined by a generalization of the Riemann-Roch formula involving the Conley-Zehnder indices $\mu_{\mathrm{CZ}}^\tau(\mathbf{A}_z)$ that were introduced in the previous lecture. We will postpone serious discussion of the index formula until Lecture 5, but here is the statement:

THEOREM 4.7. *In the setting of Theorem 4.5,*

$$\operatorname{ind} \mathbf{D} = n\chi(\dot{\Sigma}) + 2c_1^\tau(E) + \sum_{z \in \Gamma^+} \mu_{CZ}^\tau(\mathbf{A}_z) - \sum_{z \in \Gamma^-} \mu_{CZ}^\tau(\mathbf{A}_z),$$

where 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 4.8. The index formula reveals that the nondegeneracy condition on the asymptotic operators in Theorem 4.5 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 4.7, 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 lecture, $\dot{\Sigma} = \Sigma \backslash (\Gamma^+ \cup \Gamma^-)$ 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 $F = \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, E)$ 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 lecture 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 4.9. *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 2.17, after covering a compact subset with finitely many local holomorphic coordinate charts and trivializations.

LEMMA 4.10. *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 4.10 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 4.9. 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 4.11. *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 2.17 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 4.10 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 4.10 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 4.12. *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 2.23 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.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 4.11 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 4.11 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 4.13. Corollary 4.12 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 2.23, 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 4.10 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 4.5 that serves as the asymptotic analogue of the fundamental elliptic estimates from Lecture 2. Beyond those local estimates, this is the main analytical ingredient that makes all Floer-type theories in symplectic geometry work.

THEOREM 4.14. *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 $S : S^1 \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ 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 4.15. The same reasoning as in Remark 4.8 concludes via the index formula of the next lecture that the converse of Theorem 4.14 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 4.11, it suffices to prove the case $k = 1$ of Theorem 4.14, 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 Lecture 2 (namely Theorem 2.13).

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 4.16. *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 $\Phi : S^1 \to \mathrm{End}_{\mathbb{C}}(\mathbb{C}^N)$ 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 4.16 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 4.17. 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 4.16 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.38.*

PROOF OF THEOREM 4.14 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 4.16 and Exercise 4.17 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 4.14 is more obviously an asymptotic variant of the fundamental elliptic estimate from Lecture 2. Its key feature for our purposes is that, in contrast e.g. to Lemma 4.10, 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.38.

Lemma 4.18. *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* $Z^R_\pm \subset \dot{\mathcal{U}}_z$ *for every* $R \geqslant 0$ *sufficiently large, there exists a constant* $c > 0$ *such that*

$$\|\eta\|_{W^{k,p}(\mathring{Z}^R_\pm)} \leqslant c\|\mathbf{D}\eta\|_{W^{k-1,p}(\mathring{Z}^R_\pm)} \qquad \textit{for all} \qquad \eta \in W^{k,p}_0(\mathring{Z}^R_\pm).$$

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

$$\|S - S_\infty\|_{C^{k-1}(Z^R_\pm)} \to 0 \quad \text{as} \quad R \to \infty.$$

For $\eta \in W^{k,p}_0(\mathring{Z}^R_\pm)$, there is a canonical extension $\eta \in W^{k,p}(\mathbb{R} \times S^1)$ that equals zero outside $Z^R_\pm$, so Theorem 4.14 implies an estimate

$$\|\eta\|_{W^{k,p}(\mathring{Z}^R_\pm)} = \|\eta\|_{W^{k,p}(\mathbb{R}\times S^1)} \leqslant c\|\mathbf{D}_0\eta\|_{W^{k-1,p}(\mathbb{R}\times S^1)} = c\|\mathbf{D}_0\eta\|_{W^{k-1,p}(\mathring{Z}^R_\pm)}$$

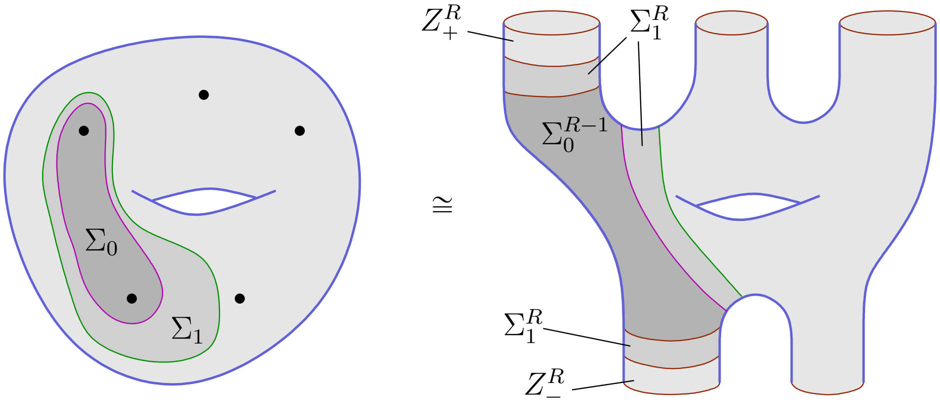

FIGURE 4.2. A punctured Riemann surface with subsets $\Sigma_0 \subset \bar{\Sigma}_0 \subset \Sigma_1 \subset \Sigma$ and their truncations $\Sigma_0^{R-1} \subset \overline{\Sigma_0^{R-1}} \subset \Sigma_1^R \subset \dot{\Sigma}$ as in Lemma 4.19.

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 4.9. 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 4.19. *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*

$$\bar{\Sigma}_0 \subset \Sigma_1, \qquad (\bar{\Sigma}_0 \backslash \Sigma_0) \cap \Gamma = \varnothing, \qquad \textit{and} \qquad (\bar{\Sigma}_1 \backslash \Sigma_1) \cap \Gamma = \varnothing,$$

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 4.18 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 4.10 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 4.18 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 4.20. *Under the assumptions of Theorem 4.5, 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 4.19 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 4.9. ☐

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 2.23 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 $\eta \in W^{m+1,q}_{\mathrm{loc}}$ for every $q \in (1, \infty)$. We also know from Corollary 4.12 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. It will follow from the fact that nondegeneracy forces bounded solutions to decay exponentially fast.

To see what nondegeneracy has to do with exponential decay conditions, let's consider for a moment the analogy with Morse homology that was discussed in §3.1. 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

$$\mathbf{D} : C^\infty(\mathbb{R}, \mathbb{R}^n) \to C^\infty(\mathbb{R}, \mathbb{R}^n) : \eta \mapsto \partial_s \eta + A(s)\eta$$

for some function $A : \mathbb{R} \to \mathbb{R}^{n \times n}$ that has a symmetric and invertible limit $A_+ := \lim_{s \to \infty} A(s)$. Let us choose a new trivialization in which A_+ is diagonal, and consider only $s \gg 1$ for which $A(s)$ is an arbitrarily good approximation of A_+. In this regime, the linearized equation becomes

$$\partial_s \eta \approx -\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \eta.$$

If this were the precise equation, then we could make some immediate pronouncements about the qualitative behavior of solutions as s becomes large: they are linear combinations of exponential functions, some growing and some decaying. Note that since $0 \notin \sigma(A_+)$, there is no middle ground between growth and decay: no nontrivial solutions can have a finite but nonzero limit as $s \to \infty$. We are not interested in the solutions that grow exponentially at infinity, as these do not have geometric meaning or belong to any reasonable Banach space of solutions we'd like to consider. Those that do belong to such spaces have exponential decay

$$|\eta(s)| \leqslant Ce^{-\lambda s}$$

for some constant $C > 0$, where for the decay rate $\lambda > 0$ one can choose any number less than the smallest positive eigenvalue of A_+.

It is not so straightforward to make this heuristic argument precise, because as long as $A(s)$ is not exactly but only approximately equal to A_+, it will not respect the splitting of $\mathbb{R}^n$ into positive and negative eigenspaces of A_+, i.e. there will be cross terms. One can therefore expect a decaying solution $\eta(s)$ to have a nontrivial but decaying component spanned by eigenvectors with negative eigenvalue, for which the tendency toward exponential growth is balanced by the cross terms, and it is not easy to say how fast this component decays.

One solution to this problem is to differentiate the equation one more time and produce a second-order differential inequality, which has the effect of erasing the distinction between positive and negative eigenvalues. Concretely, let us consider the function

$$\alpha(s) := \frac{1}{2}|\eta(s)|^2 = \frac{1}{2}\langle \eta(s), \eta(s)\rangle,$$

where $\langle\ ,\ \rangle$ is the Euclidean inner product on $\mathbb{R}^n$. Since $\dot{\eta} = -A\eta$, its first derivative is given by

$$\dot{\alpha} = -\langle \eta, A\eta\rangle.$$

Recalling that A is not symmetric but converges as $s \to \infty$ to something symmetric, we can then write its second derivative as

$$\ddot{\alpha} = \langle A\eta, A\eta\rangle - \langle \eta, \dot{A}\eta\rangle + \langle \eta, A(A\eta)\rangle = 2|A\eta|^2 + \langle (A^{\mathrm{T}} - A)\eta, A\eta\rangle - \langle \eta, \dot{A}\eta\rangle.$$

Now consider what happens in the region $s \geqslant R$ for some $R \gg 1$. For any $\lambda > 0$ such that $\sigma(A_+) \cap [-\lambda, \lambda] = \varnothing$, we have $|A_+ v| \geqslant \lambda|v|$ for all $v \in \mathbb{R}^n$, and since $A(s) \to A_+$ as $s \to \infty$, we can assume (after a slight adjustment to λ) if $R > 0$ is chosen large enough that

$$|A(s)v| \geqslant \lambda|v| \quad \text{for all} \quad v \in \mathbb{R}^n,\ s \geqslant R$$

also holds, implying $|A(s)\eta(s)|^2 \geqslant \lambda^2|\eta(s)|^2 = 2\lambda^2\alpha(s)$. If we also assume that A is "C^1-asymptotic" to A_+, meaning $\|A - A_+\|_{C^1([R,\infty))} \to 0$ as $R \to \infty$, then $\|\dot{A}(s)\|$ can be assumed arbitrarily small for all $s \geqslant R$, and similarly, the symmetry of A_+ means that $\|A^{\mathrm{T}}(s) - A(s)\|$ can be assumed arbitrarily small. The result is that for arbitrarily small values $\epsilon > 0$, one can choose $R \gg 1$ large enough so that α satisfies the differential inequality

$$\ddot{\alpha} \geqslant (4\lambda^2 - \epsilon)\alpha.$$

If we now replace the original interval $[-\lambda, \lambda]$ with a slightly larger one that still does not contain any eigenvalues of A_+, we can repeat the argument and replace $4\lambda^2 - \epsilon$ in this expression with $4\lambda^2$, establishing the relation

$$\ddot{\alpha}(s) \geqslant 4\lambda^2\alpha(s) \quad \text{for all} \quad s \geqslant R.$$

This inequality says that the function $\alpha(s)$ should be "at least as convex" as an actual solution to the differential equation

$$\ddot{\beta}(s) = 4\lambda^2\beta(s). \tag{4.9}$$

Solutions to the latter are exponential functions, either growing or decaying, but in principle we can ignore the growing solutions since we want to assume η has reasonable behavior at $+\infty$. Let us therefore compare the function α with the unique decaying solution to (4.9) that has the same value as α at an initial point, namely

$$\beta(s) := \alpha(R)e^{-2\lambda(s-R)}.$$

The function $f(s) := \alpha(s) - \beta(s)$ then satisfies

$$\ddot{f}(s) \geqslant 4\lambda^2 f(s) \text{ for all } s \geqslant R, \qquad \text{and} \qquad f(R) = 0. \tag{4.10}$$

EXERCISE 4.21. Show that every C^2-function $f : [R, \infty) \to \mathbb{R}$ satisfying both conditions in (4.10) also satisfies either $\lim_{s\to\infty} f(s) = \infty$ or $f \leqslant 0$. *Hint: Remember the mean value theorem?*

If $\lim_{s\to\infty} f(s) = \infty$ in this situation, then $|\eta(s)|^2$ must also blow up as $s \to \infty$, so that our solution η cannot be of class L^p or L^∞ or any other class of functions we are likely to want to consider. According to Exercise 4.21, the only alternative is then that $\alpha(s) \leqslant \beta(s)$ holds for all $s \geqslant R$, meaning

$$|\eta(s)| \leqslant |\eta(R)| e^{-\lambda(s-R)}.$$

It's worth noting that the constant R in this result does not depend on η; it is determined by our choice of $\lambda > 0$ and the rate at which $A(s)$ approaches A_+ in the original differential equation. In other words, this argument does not just deliver an exponential decay result for a single solution—it provides a uniform bound for all solutions.

You should now find the following result plausible, and the proof will seem familiar, though several details are technically trickier than in the Morse homology setting. Note that since the proof requires differentiating the zeroth-order term of the operator near a puncture, we have to require $m \geqslant 1$ instead of the usual assumption $0 \leqslant m \leqslant \infty$. This is the only step in our proof of the Fredholm property at which the case $m = 0$ must be excluded.

LEMMA 4.22. *Assume* $\mathbf{D}$ *is of class* C^m *with* $1 \leqslant m \leqslant \infty$, $z \in \Gamma^\pm$ *is a puncture for which the asymptotic operator* $\mathbf{A}_z$ *is nondegenerate, and* $\lambda > 0$ *is a constant such that*

$$\sigma(\mathbf{A}_z) \cap [-\lambda, \lambda] = \varnothing.$$

Then in holomorphic cylindrical coordinates on $Z_\pm \cong \dot{\mathcal{U}}_z$, *there exists a constant* $R > 0$ *such that every weak solution* $\eta \in L^\infty(E|_{\dot{\mathcal{U}}_z})$ *to* $\mathbf{D}\eta = 0$ *satisfies*

$$\|\eta(\pm s, \cdot)\|_{L^2(S^1)} \leqslant \|\eta(\pm R, \cdot)\|_{L^2(S^1)} \cdot e^{-\lambda(s-R)} \qquad \textit{for all} \qquad s \geqslant R.$$

PROOF. To simplify the notation, let us assume the puncture $z \in \Gamma$ is positive, as the proof in the negative case will be completely analogous. 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 $\|S - S_\infty\|_{C^m(Z_+^R)} \to 0$ as $R \to \infty$. Let us also write the L^2-product for functions $\mathring{Z}_+ \to \mathbb{R}^{2n}$ as

$$\langle u, v \rangle_{L^2(\mathring{Z}_+)} := \int_{[0,\infty)\times S^1} \langle u(s,t), v(s,t) \rangle \, ds\, dt,$$

where $\langle\ ,\ \rangle$ with no subscript denotes Euclidean inner product on $\mathbb{R}^{2n}$. For a given constant $\lambda > 0$ with $\sigma(\mathbf{A}) \cap [-\lambda, \lambda] = \varnothing$, choose two slightly larger constants

$$\lambda_2 > \lambda_1 > \lambda \qquad \text{satisfying} \qquad \sigma(\mathbf{A}) \cap [-\lambda_2, \lambda_2] = \varnothing.$$

Then $\mathbf{A}$ satisfies the estimate

$$\|\mathbf{A}\eta\|_{L^2(S^1)} \geqslant \lambda_2 \|\eta\|_{L^2(S^1)} \qquad \text{for all} \qquad \eta \in H^1(S^1, \mathbb{R}^{2n}).$$

Indeed, according to Exercise 3.41(b), $L^2(S^1, \mathbb{R}^{2n})$ is spanned by an orthonormal basis of eigenfunctions $\{e_j \in H^1(S^1)\}_{j\in\mathbb{Z}}$ satisfying $\mathbf{A}e_j = \mu_j e_j$ for $\mu_j \in \sigma(\mathbf{A})$, so for every $\eta \in H^1(S^1) \subset L^2(S^1)$, we have $\eta = \sum_{j\in\mathbb{Z}} c_j e_j$ for $c_j := \langle e_j, \eta\rangle_{L^2}$ and thus

$$\langle \mathbf{A}\eta, \mathbf{A}\eta\rangle_{L^2} = \sum_{j\in\mathbb{Z}} c_j^2 \mu_j^2 \geqslant \lambda_2^2 \sum_{j\in\mathbb{Z}} c_j^2 = \lambda_2^2 \|\eta\|_{L^2}^2.$$

This estimate for $\mathbf{A}$ says that the composition of $\mathbf{A}^{-1} : L^2(S^1) \to H^1(S^1)$ with the inclusion $i : H^1(S^1) \hookrightarrow L^2(S^1)$ has $\|i \circ \mathbf{A}^{-1}\|_{\mathscr{L}(L^2)} \leqslant \frac{1}{\lambda_2}$ for the operator norm on $\mathscr{L}(L^2(S^1))$. For $s \geqslant 0$, let

$$\mathbf{A}_s := -J_0\partial_t - S(s,\cdot) : H^1(S^1, \mathbb{R}^{2n}) \to L^2(S^1, \mathbb{R}^{2n}),$$

and observe that since $S(s,\cdot) \to S_\infty$ uniformly as $s \to \infty$,

$$\begin{aligned} \|(\mathbf{A} - \mathbf{A}_s)\eta\|_{L^2} &= \|(S_\infty - S(s,\cdot))\eta\|_{L^2} \leqslant \|S_\infty - S(s,\cdot)\|_{C^0} \cdot \|\eta\|_{L^2} \\ &\leqslant \|S_\infty - S(s,\cdot)\|_{C^0} \cdot \|\eta\|_{H^1} \end{aligned}$$

implies $\lim_{s\to\infty} \mathbf{A}_s = \mathbf{A}$, with convergence in the operator norm on $\mathscr{L}(H^1(S^1), L^2(S^1))$. It follows that for $s > 0$ sufficiently large, $\mathbf{A}_s$ is also invertible and $\lim_{s\to\infty} \mathbf{A}_s^{-1} = \mathbf{A}^{-1}$ in $\mathscr{L}(L^2(S^1), H^1(S^1))$, in which case the norms $\|i \circ \mathbf{A}_s^{-1}\|_{\mathscr{L}(L^2)}$ also converge to $\|i \circ \mathbf{A}^{-1}\|_{\mathscr{L}(L^2)}$. This argument proves that we can also assume

$$\|\mathbf{A}_s\eta\|_{L^2(S^1)} \geqslant \lambda_1 \|\eta\|_{L^2(S^1)} \quad \text{for all} \quad \eta \in H^1(S^1, \mathbb{R}^{2n}) \text{ and } s \geqslant R \tag{4.11}$$

if $R > 0$ is sufficiently large.

We will impose two further conditions requiring R to be large. The first is motivated by the fact that $S(s,t)$ is not generally symmetric but $S_\infty(t)$ is. In light of the asymptotic convergence of $S(s,t)$, we can for any $\epsilon > 0$ assume after making $R > 0$ large enough that

$$\|S - S^{\mathrm{T}}\|_{C^0(Z_+^R)} < \epsilon. \tag{4.12}$$

Since $m \geqslant 1$, we can also assume $\|S - S_\infty\|_{C^1(Z_+^R)}$ is arbitrarily small, implying in particular that if R is large enough, then

$$\|\partial_s S\|_{C^0(Z_+^R)} < \epsilon. \tag{4.13}$$

In the following we shall exploit the freedom to make R larger in order to make ϵ smaller as needed.

Now suppose $u \in L^\infty(\mathring{Z}_+, \mathbb{R}^{2n})$ satisfies $(\bar\partial + S)u = 0$. Then u is also locally of class L^p for every $p \in (1,\infty)$, so by Corollary 2.23, u is in $W^{m+1,p}_{\mathrm{loc}}(\mathring{Z}_+)$ and is C^m-smooth, implying in particular that u is locally of class $W^{2,p}$ and continuously differentiable. Abbreviate $u_s := u(s,\cdot)$ and consider the function $\alpha : [0,\infty) \to \mathbb{R}$ defined by

$$\alpha(s) := \frac{1}{2}\|u_s\|_{L^2(S^1)}^2.$$

Writing the equation $\bar\partial u + Su = 0$ as $\partial_s u(s,\cdot) = \mathbf{A}_s u_s$, differentiating under the integral sign gives

$$\dot\alpha(s) = \langle u_s, \mathbf{A}_s u_s\rangle_{L^2(S^1)} = \int_{S^1} \langle u(s,t), -J_0\partial_t u(s,t) - S(s,t)u(s,t)\rangle \, dt. \tag{4.14}$$

We have to be a bit more careful in computing $\ddot{\alpha}(s)$, as the term $-J_0\partial_t u$ on the right hand side of (4.14) might not be differentiable, though since $u \in W^{2,p}_{\text{loc}}(\mathring{Z}_+)$, there is a well-defined weak derivative $\partial_s\partial_t u \in L^p_{\text{loc}}(\mathring{Z}_+)$. Since 2-dimensional domains admit continuous product pairings $W^{2,p} \times W^{1,p} \to W^{1,p}$ and $C^1 \times W^{1,p} \to W^{1,p}$, Proposition A.16 (and Remark A.16) permit the computation of the weak derivative of the integrand via the Leibniz rule,

$$\frac{\partial}{\partial s}\langle u, -J_0\partial_t u - Su\rangle = \langle \partial_s u, -J_0\partial_t u - Su\rangle + \langle u, -J_0\partial_s\partial_t u - (\partial_s S)u - S(\partial_s u)\rangle,$$

and the resulting function is in $L^p_{\text{loc}}(\mathring{Z}_+)$, hence locally integrable. It is an easy exercise in Fubini's theorem (see Exericse 4.23(a) below) to check that in this situation, *weak* differentiation under the integral sign is also allowed, so we obtain a weak second derivative of α,

$$\begin{aligned}\ddot{\alpha}(s) = \int_{S^1} \Big(&\langle \partial_s u(s,t), -J_0\partial_t u(s,t) - S(s,t)u(s,t)\rangle \\ &+ \langle u(s,t), -J_0\partial_s\partial_t u(s,t) - \partial_s S(s,t)u(s,t) - S(s,t)\partial_s u(s,t)\rangle\Big)\,dt.\end{aligned}$$

We would like to apply integration by parts to remove the second derivative of u from this expression, but again we must be careful about regularity since $\partial_s\partial_t u$ may be only a weak derivative. Observe first that since $u \in W^{2,p}_{\text{loc}}(\mathring{Z}_+)$, $\partial_s u \in W^{1,p}_{\text{loc}}(\mathring{Z}_+)$ and the product pairing $W^{2,p} \times W^{1,p} \to W^{1,p}$ over 2-dimensional domains is continuous, the function $\partial_t\langle u, -J_0\partial_s u\rangle$ is in $L^p_{\text{loc}}(\mathring{Z}_+)$ and equal to $\langle \partial_t u, -J_0\partial_s u\rangle + \langle u, -J_0\partial_s\partial_t u\rangle$ by Proposition A.16. Fubini's theorem then implies that for almost every $s \geqslant 0$, both $\langle u(s,\cdot), -J_0\partial_s u(s,\cdot)\rangle$ and $\partial_t\langle u(s,\cdot), -J_0\partial_s u(s,\cdot)\rangle$ are in $L^p(S^1)$, and according to Exercise 4.23(b), the latter is the weak derivative of the former as functions on S^1. It follows via Proposition A.11 that for these values of s, $\langle u(s,\cdot), -J_0\partial_s u(s,\cdot)\rangle : S^1 \to \mathbb{R}$ is equal almost everywhere to an absolutely continuous function whose derivative almost everywhere is equal to its weak derivative, and the integral of its weak derivative over S^1 therefore vanishes. This justifies the use of integration by parts to write

$$\begin{aligned}\int_{S^1}\langle u(s,t), -J_0\partial_s\partial_t u(s,t)\rangle\,dt &= \int_{S^1}\langle \partial_t u(s,t), J_0\partial_s u(s,t)\rangle\,dt \\ &= \int_{S^1}\langle -J_0\partial_t u(s,t), \partial_s u(s,t)\rangle\,dt = \langle (\mathbf{A}_s + S(s,\cdot))u_s, \mathbf{A}_s u_s\rangle_{L^2(S^1)}\end{aligned}$$

for almost every s. For these values of s, our formula for $\ddot{\alpha}(s)$ now becomes

$$\begin{aligned}\ddot{\alpha}(s) &= \langle \mathbf{A}_s u_s, \mathbf{A}_s u_s\rangle_{L^2} + \langle (\mathbf{A}_s + S(s,\cdot))u_s, \mathbf{A}_s u_s\rangle_{L^2} - \langle u_s, \partial_s S(s,\cdot)u_s + S(s,\cdot)\mathbf{A}_s u_s\rangle_{L^2} \\ &= 2\|\mathbf{A}_s u_s\|^2_{L^2} + \left\langle [S(s,\cdot) - S(s,\cdot)^{\mathrm{T}}]u_s, \mathbf{A}_s u_s\right\rangle_{L^2} - \langle u_s, \partial_s S(s,\cdot)u_s\rangle_{L^2},\end{aligned}$$

where the L^2-products here are all for functions on S^1. Conveniently, the right hand side of this expression is a continuous function of s, implying that $\dot{\alpha}$ is a C^1-function and this continuous expression for $\ddot{\alpha}$ is its classical derivative.

Using the conditions (4.11), (4.12) and (4.13), we can now conclude that $\alpha : (0, \infty) \to \mathbb{R}$ is a C^2-function which for $s \geqslant R$ satisfies

$$\begin{aligned}\ddot{\alpha}(s) &\geqslant 2\|\mathbf{A}_s u_s\|_{L^2}^2 - \epsilon\|u_s\|_{L^2} \cdot \|\mathbf{A}_s u_s\|_{L^2} - \epsilon\|u_s\|_{L^2}^2 \\ &= \|\mathbf{A}_s u_s\|_{L^2} \cdot (2\|\mathbf{A}_s u_s\|_{L^2} - \epsilon\|u_s\|_{L^2}) - \epsilon\|u_s\|_{L^2}^2 \\ &\geqslant \lambda_1\|u_s\|_{L^2} \cdot (2\lambda_1\|u_s\|_{L^2} - \epsilon\|u_s\|_{L^2}) - \epsilon\|u_s\|_{L^2}^2 \\ &= \left(2\lambda_1^2 - \epsilon\lambda_1 - \epsilon\right) \cdot \|u_s\|_{L^2}^2 = 4\left(\lambda_1^2 - \frac{\epsilon\lambda_1}{2} - \frac{\epsilon}{2}\right)\alpha(s).\end{aligned}$$

Let us now increase R in order to shrink ϵ so that without loss of generality,

$$\lambda_1^2 - \frac{\epsilon\lambda_1}{2} - \frac{\epsilon}{2} \geqslant \lambda^2,$$

noting that this choice depends only on the function $S(s,t)$ and not on the solution u. The last inequality then becomes

$$\ddot{\alpha}(s) \geqslant 4\lambda^2\alpha(s) \quad \text{for all} \quad s \geqslant R.$$

Recall now that $u : \mathring{Z}_+ \to \mathbb{R}^{2n}$ was assumed to be globally of class L^∞, so the function $\alpha(s)$ is bounded as $s \to \infty$. Plugging $f(s) := \alpha(s) - \alpha(R)e^{-2\lambda(s-R)}$ into Exercise 4.21 then gives

$$\frac{1}{2}\|u_s\|_{L^2}^2 = \alpha(s) \leqslant \alpha(R)e^{-2\lambda(s-R)} = \frac{1}{2}\|u_R\|_{L^2}^2 e^{-2\lambda(s-R)},$$

or equivalently,

$$\|u_s\|_{L^2} \leqslant \|u_R\|_{L^2} e^{-\lambda(s-R)} \quad \text{for all} \quad s \geqslant R.$$

□

EXERCISE 4.23. Consider a locally integrable function $f : \mathring{Z}_+ \to \mathbb{R}$ on the half-cylinder $\mathring{Z}_+ = (0, \infty) \times S^1$. Show:

(a) If f has a locally integrable weak partial derivative $\partial_s f : \mathring{Z}_+ \to \mathbb{R}$ and we define $F : (0, \infty) \to \mathbb{R}$ almost everywhere by $F(s) := \int_{S^1} f(s,t)\,dt$, then F has weak derivative $F'(s) := \int_{S^1} \partial_s f(s,t)\,dt$.

(b) If f has a locally integrable weak partial derivative $\partial_t f : \mathring{Z}_+ \to \mathbb{R}$, then for almost every $s > 0$, the function $f(s, \cdot) : S^1 \to \mathbb{R}$ has locally integrable weak derivative $\partial_t f(s, \cdot) : S^1 \to \mathbb{R}$. *Hint: Consider test functions on Z_+ of the form $\beta(s)\varphi(t)$ for $\beta \in C_0^\infty((0, \infty))$ and $\varphi \in C^\infty(S^1)$.*

EXERCISE 4.24. Show that the constant $R > 0$ in Lemma 4.22 can be chosen so that the result remains true with the same value of R after adjusting the functions S and S_∞ by a sufficiently C^1-small perturbation.

COROLLARY 4.25. *In the setting of Theorem 4.5,*

$$\ker \mathbf{D} \subset \bigcap_{\ell \leqslant m+1} \bigcap_{1 < q < \infty} W^{\ell,q}(E),$$

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 2.23, and Corollary 4.12 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 therefore suffices to prove that the restriction of η to each cylindrical end $Z_\pm^R \subset \dot{\mathcal{U}}_z$ is in $L^q(\mathring{Z}_\pm^R)$ for $q > 1$ arbitrarily close to 1 and $R > 0$ sufficiently large. Let us consider for concreteness a positive end Z_+^R 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 $\lambda > 0$ such that the corresponding asymptotic operator has no eigenvalues in $[-\lambda,\lambda]$, Lemma 4.22 implies

$$\begin{aligned}\|\eta\|^q_{L^q(\mathring{Z}_+^R)} &= \int_R^\infty \|\eta(s,\cdot)\|^q_{L^q(S^1)}\,ds \leqslant c^q \int_R^\infty \|\eta(s,\cdot)\|^q_{L^2(S^1)}\,ds \\ &\leqslant c^q \|\eta(R,\cdot)\|^q_{L^2(S^1)} \int_R^\infty e^{-q\lambda(s-R)}\,ds < \infty.\end{aligned}$$

$\square$

We can now say more precisely what is meant by the statement in Theorem 4.5 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 later when we study the corresponding nonlinear problem. For $k \geqslant 0$, $1 \leqslant p \leqslant \infty$ and $\lambda \in \mathbb{R}$, define

$$W^{k,p,\lambda}(\mathring{Z}_\pm^R,\mathbb{R}^{2n}) := \left\{ f : \mathring{Z}_\pm^R \to \mathbb{R}^{2n} \;\middle|\; f = e^{\mp\lambda s} g \text{ for some } g \in W^{k,p}(\mathring{Z}_\pm^R,\mathbb{R}^{2n}) \right\},$$

with the case $k = 0$ abbreviated by $L^{p,\lambda} := W^{0,p,\lambda}$. This is a Banach space with respect to the norm

$$\|f\|_{W^{k,p,\lambda}(\mathring{Z}_\pm^R)} := \|e^{\pm\lambda s} f\|_{W^{k,p}(\mathring{Z}_\pm^R)},$$

and in fact there is an obvious isometry $W^{k,p}(\mathring{Z}_\pm) \to W^{k,p,\lambda}(\mathring{Z}_\pm) : f \mapsto e^{\mp\lambda s} f$. We typically consider $W^{k,p,\lambda}(\mathring{Z}_\pm)$ for $\lambda > 0$, 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 $f \in W^{m+1,p,\lambda}(\mathring{Z}_\pm^R)$ take the form $e^{\mp\lambda s} g$ 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 $e^{\mp\lambda s}$ with a globally bounded function, producing an estimate of the form

$$|\partial^\alpha f(s,t)| \leqslant C e^{\mp\lambda s} \quad \text{for all } |\alpha| \leqslant m.$$

The statement about decaying derivatives in Theorem 4.5 is therefore a consequence of the following:

PROPOSITION 4.26. *Under the same assumptions as in Lemma 4.22, for every $k \leqslant m+1$, $q \in (1,\infty)$ and every $R > 0$ sufficiently large, weak solutions $\eta \in L^\infty(E|_{\dot{\mathcal{U}}_z})$ to $\mathbf{D}\eta = 0$ near a positive puncture $z \in \Gamma^+$ satisfy an estimate of the form*

$$\|\eta\|_{W^{k,q,\lambda}(\mathring{Z}_+^1)} \leqslant c\|\eta\|_{L^\infty([0,R]\times S^1)},$$

where the constant $c > 0$ depends on k, q, R and λ but not on η. A similar result holds for negative punctures.

PROOF. We again assume for concreteness that the puncture is positive, and choose an asymptotic trivialization to express sections η of $E|_{\dot{\mathcal{U}}_z}$ as functions $u : Z_+ \to \mathbb{R}^{2n}$. On any domain $\mathcal{U} \subset Z_+$ with compact closure in the interior of Z_+, repeated application of Lemma 4.10 gives bounds on $\|u\|_{W^{k,q}(\mathcal{U})}$ for $k \leqslant m+1$ in terms of the L^q-norm of u on any strictly larger domain with compact closure, and the latter is bounded in turn by the L^∞-norm on the same domain. It therefore suffices to find a bound on $\|u\|_{W^{k,q,\lambda}(\mathring{Z}_+^r)}$, where $r > 1$ can be chosen to be as large as is needed.

For the case $k = 0$ and $q \leqslant 2$, we derive a bound on $\|u\|_{L^{q,\lambda}(\mathring{Z}_+^r)}$ using Lemma 4.22 as follows. Choose $\lambda_1 > \lambda$ so that the condition $\sigma(\mathbf{A}) \cap [-\lambda_1, \lambda_1] = \varnothing$ still holds, and then choose $r > 1$ large enough for the exponential bound in Lemma 4.22 to hold on Z_+^r with decay rate λ_1. Using the continuous inclusion $L^2(S^1) \hookrightarrow L^q(S^1)$, we then have

$$\begin{aligned}
\|u\|^q_{L^{q,\lambda}(\mathring{Z}_+^r)} &= \int_{Z_+^r} e^{q\lambda s}|u(s,t)|^q \, ds\, dt = \int_r^\infty e^{q\lambda s}\|u(s,\cdot)\|^q_{L^q(S^1)}\, ds \\
&\leqslant c\int_r^\infty e^{q\lambda s}\|u(s,\cdot)\|^q_{L^2(S^1)}\, ds \leqslant c\|u(r,\cdot)\|^q_{L^2(S^1)} \cdot \int_r^\infty e^{q\lambda s}e^{-q\lambda_1(s-r)}\, ds \\
&\leqslant ce^{q\lambda_1 r}\|u\|^q_{L^\infty([0,r]\times S^1)} \cdot \int_r^\infty e^{-q(\lambda_1-\lambda)s}\, ds \\
&= \frac{ce^{q\lambda_1 r}}{q(\lambda_1-\lambda)}e^{-q(\lambda_1-\lambda)r} \cdot \|u\|^q_{L^\infty([0,r]\times S^1)}.
\end{aligned}$$

To improve this to a $W^{k,q,\lambda}$-bound for $k \geqslant 1$, we observe that the function $e^{\lambda s}u$, which is now known to be in $L^q(\mathring{Z}_+)$, also satisfies a Cauchy-Riemann type equation, namely

$$\mathbf{D}_\lambda(e^{\lambda s}u) := (\bar{\partial} + S_\lambda)(e^{\lambda s}u) = 0, \qquad \text{where} \qquad S_\lambda(s,t) := S(s,t) - \lambda,$$

which is C^m-asymptotic to the shifted asymptotic operator $\mathbf{A}_\lambda := \mathbf{A} + \lambda$. Since no eigenvalues of $\mathbf{A}$ lie in $[-\lambda, \lambda]$, the operator $\mathbf{A}_\lambda$ is also nondegenerate. Repeated application of Lemma 4.19 therefore provides a $W^{k,q}$-bound on $e^{\lambda s}u$ for every $k \leqslant m+1$ in terms of its L^q-bound on a sufficiently large truncation, which is also bounded by the truncated L^∞-norm of u. Finally, one can apply the Sobolev embedding theorem as in Corollary 4.12 and repeat the application of Lemma 4.19 as needed to produce a $W^{k,p}$-bound on $e^{\lambda s}u$ for every $p \in [q, \infty)$ and $k \leqslant m+1$. □

The bound in Proposition 4.26 tells us that we have considerable freedom in our choice of topology for the space of solutions to the equation $\mathbf{D}\eta = 0$: any sequence of (not necessarily uniformly) bounded solutions that converge uniformly on compact subsets must also converge in the much stronger topology of the $W^{k,p,\lambda}$-norm for every $k \leqslant m+1$, $p \in (1, \infty)$ and $\lambda > 0$ smaller than the absolute value of every eigenvalue of every asymptotic operator. We will see this phenomenon again in 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 carrying out the analysis.

EXERCISE 4.27. 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 asymptotic operators, such that the asymptotic operators for $\mathbf{D}$ are all nondegenerate and $\lim_{\nu\to\infty} \|S_\nu\|_{C^m} = 0$. Show that if $1 < p < \infty$ and η_ν is a uniformly L^p-bounded sequence of weak solutions to $\mathbf{D}_\nu \eta_\nu = 0$ that is uniformly convergent on compact subsets, then for every $k \leqslant m+1$, $q \in (1,\infty)$ and $\lambda > 0$ sufficiently small, η_ν converges in $W^{k,q}(E)$ and in the $W^{k,q,\lambda}$-norm on the cylindrical ends to a solution η of $\mathbf{D}\eta = 0$. *Hint: Exercise 4.24 should help you prove a uniform L^p-bound for $e^{\pm\lambda s}\eta_\nu$ on the cylindrical ends $\dot{\mathcal{U}}_z \cong Z_\pm$. Once you have that, you can eliminate the exponential weights from the picture by adding constants to $\mathbf{D}$ and $\mathbf{D}_\nu$.*

4.7. Formal adjoints and proof of the Fredholm property

In order to show that $\operatorname{coker} \mathbf{D}$ 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

$$\langle \lambda, \mathbf{D}\eta \rangle_{L^2(F)} = \langle \mathbf{D}^*\lambda, \eta \rangle_{L^2(E)} \qquad \text{for all} \qquad \eta \in C_0^{m+1}(E),\ \lambda \in C_0^{m+1}(F),$$

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 4.28. 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[2] 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 4.10, 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

$$\begin{aligned}(C^{-1}\mathbf{D}^*C)\eta &= -C\partial_s(C\eta) + CJ_0\partial_t(C\eta) + CS(s,t)^{\mathrm{T}}C\eta \\ &= -\partial_s\eta - J_0\partial_t\eta + CS(s,t)^{\mathrm{T}}C\eta = -(\bar{\partial}\eta - CS(s,t)^{\mathrm{T}}C\eta) \\ &=: -(\bar{\partial} + \bar{S}(s,t))\eta,\end{aligned}$$

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 4.11, 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 4.18 and Lemma 4.19 now also go through for $\mathbf{D}^*$.

We've proved:

[2] 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 this property. The former requires the principal symbol of the operator to be everywhere injective, and the latter requires it to be surjective.

PROPOSITION 4.29. *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 4.30. *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. 4.29, then for* $\mathbf{D} : W^{k,p}(E) \to W^{k-1,p}(E)$ *with* $1 \leqslant k \leqslant m+1$,

$$W^{k-1,p}(F) = \mathrm{im}\,\mathbf{D} + \ker \mathbf{D}^*.$$

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 $\mathrm{im}\,\mathbf{D} + \ker \mathbf{D}^*$ is a closed subspace of $L^p(F)$. Then if $\mathrm{im}\,\mathbf{D} + \ker \mathbf{D}^* \neq L^p(F)$, the Hahn-Banach theorem[3] provides a nontrivial element $\alpha \in (L^p(F))^* \cong L^q(F)$ for $\frac{1}{p} + \frac{1}{q} = 1$ such that

$$\langle \mathbf{D}\eta + \lambda, \alpha \rangle_{L^2(F)} = 0 \quad \text{for all} \quad \eta \in W^{1,p}(E),\ \lambda \in \ker \mathbf{D}^*. \tag{4.15}$$

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 (4.15) 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 4.11, $\eta \in W^{k,p}(E)$, completing the proof for all $k \leqslant m+1$. □

REMARK 4.31. 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 4.30 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 4.11 implies since $p \geqslant q$ that $\ker \mathbf{D}^*$ is then also a subspace of $W^{k-1,p}(F)$.

[3]In the case $p = 2$, one can forego the Hahn-Banach theorem and simply take an L^2-orthogonal complement.

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 4.30:

PROPOSITION 4.32. *Under the same assumptions as in Lemma 4.30 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 4.30, 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 $\frac{1}{p} + \frac{1}{q} = 1$,

$$\langle \lambda, \mathbf{D}\eta \rangle_{L^2(F)} = \langle \mathbf{D}^*\lambda, \eta \rangle_{L^2(E)} \quad \text{for all} \quad \eta \in W^{1,p}(E),\ \lambda \in W^{1,q}(F). \tag{4.16}$$

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 (4.16) 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 structure 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

$$F = \Lambda^{0,1}T^*\Sigma \otimes E.$$

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

$$\mathrm{Hom}_{\mathbb{C}}(T\dot{\Sigma}, F) = \Lambda^{1,0}T^*\dot{\Sigma} \otimes F = \Lambda^{1,0}T^*\dot{\Sigma} \otimes \Lambda^{0,1}T^*\dot{\Sigma} \otimes E,$$

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 4.33. 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

$$\mathrm{Hom}_{\mathbb{C}}(T\dot{\Sigma}, F) = \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, \bar{F}),$$

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 $-J_z\widehat{\nabla}_t$, where $\widehat{\nabla}_t$ is a symplectic connection on (E_z, ω_z). Then $\widehat{\nabla}_t$ 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

$$\bar{\mathbf{A}}_z := -\mathbf{A}_z : \Gamma(\bar{E}_z) \to \Gamma(\bar{E}_z), \tag{4.17}$$

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 4.34. *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 $\mathrm{Hom}_{\mathbb{C}}(T\dot{\Sigma}, F)$ and the natural identification of the latter with its conjugate $\overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, \bar{F})$, 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 (4.17). $\square$

LECTURE 5

The index formula

Contents

5.1. Riemann-Roch with punctures

As in the previous lecture, 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

$$F := \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, E)$$

for the bundle of complex-antilinear homomorphisms $T\dot{\Sigma} \to E$, the main result of the previous lecture 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 lecture is to compute $\mathrm{ind}(\mathbf{D}) \in \mathbb{Z}$.

The index will depend on the Conley-Zehnder indices $\mu^\tau_{\mathrm{CZ}}(\mathbf{A}_z) \in \mathbb{Z}$ introduced in Lecture 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 $\mu^\tau_{\mathrm{CZ}}(\mathbf{A}_z)$ uniquely. This convention has been used to state the formula for $\mathrm{ind}(\mathbf{D})$ 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. 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 $\partial S = \emptyset$, and moreover, it extends in an obvious way to the category of asymptotically Hermitian vector bundles with asymptotic trivializations.

EXERCISE 5.2. Given two distinct choices of asymptotic trivializations τ_1 and τ_2 for an asymptotically Hermitian bundle E of rank n, 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 punctures of the winding numbers of the determinants of the transition maps $S^1 \to \mathrm{U}(m)$.[3]

[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.

[3] Caution: to compute this winding number at a negative puncture using cylindrical coordinates $(s, t) \in (-\infty, 0] \times S^1$, one must traverse $\{-s\} \times S^1$ for $s \gg 1$ in the *wrong direction*, as this is

EXERCISE 5.3. Combining Exercise 5.2 above with Exercise 3.56, 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 5.4. *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 5.5. The case $n = 0$ is also allowed in the above formula: then $c_1^\tau(E) = 0$ and all the Conley-Zehnder indices vanish by convention (cf. Remark 3.52), 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. 14.36.

NOTATION. Throughout this lecture, we shall denote the integer on the right hand side in Theorem 5.4 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 $\Gamma = \varnothing$, Theorem 5.4 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

$$\operatorname{ind}_{\mathbb{C}}(\mathbf{D}_0) = n(1-g) + c_1(E). \tag{5.1}$$

This formula assumes that the Cauchy-Riemann type operator $\mathbf{D}_0$ is complex linear, but an arbitrary real-linear Cauchy-Riemann 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 5.4.

REMARK 5.6. 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 stated.

---

consistent with the orientation induced on $\{-s\} \times S^1$ as a boundary component of a large compact subdomain of $\dot{\Sigma}$.

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 5.4 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 5.4 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 lecture, 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. 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 $\mathrm{ind}(\mathbf{D}) = \mathrm{ind}(\mathbf{D} + B)$ for any zeroth-order term $B \in \Gamma(\mathrm{Hom}_{\mathbb{R}}(E, F))$. 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 $\mathrm{Hom}_{\mathbb{R}}(E, F)$ into a direct sum of complex line bundles[4]

$$\mathrm{Hom}_{\mathbb{R}}(E, F) = \mathrm{Hom}_{\mathbb{C}}(E, F) \oplus \overline{\mathrm{Hom}}_{\mathbb{C}}(E, F).$$

[4]Here the complex structure on $\mathrm{Hom}_{\mathbb{R}}(E, F)$ and its subbundles is defined in terms of the complex structure of F, i.e. it sends $B \in \mathrm{Hom}_{\mathbb{R}}(E, F)$ to $J \circ B \in \mathrm{Hom}_{\mathbb{R}}(E, F)$.

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

$$\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F) = \overline{\mathrm{Hom}}_{\mathbb{C}}(E, \mathbb{C}) \otimes F,$$

and then observe that $\overline{\mathrm{Hom}}_{\mathbb{C}}(E, \mathbb{C})$ and E are isomorphic: indeed, any Hermitian bundle metric $\langle\ ,\ \rangle_E$ on E gives rise to a bundle isomorphism[5]

$$E \to \overline{\mathrm{Hom}}_{\mathbb{C}}(E, \mathbb{C}) : \eta \mapsto \langle \cdot, \eta \rangle_E.$$

We thus have $\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F) \cong E \otimes F$, so $c_1(\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F)) = c_1(E) + c_1(F)$. We can compute $c_1(F)$ by the same trick since

$$F = \overline{\mathrm{Hom}}_{\mathbb{C}}(T\Sigma, E) = \overline{\mathrm{Hom}}_{\mathbb{C}}(T\Sigma, \mathbb{C}) \otimes E \cong T\Sigma \otimes E,$$

so $c_1(F) = c_1(T\Sigma) + c_1(E) = \chi(\Sigma) + c_1(E)$ by the Poincaré-Hopf theorem, and thus

$$c_1(\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F)) = \chi(\Sigma) + 2c_1(E).$$

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 5.4.

"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 $B \in \Gamma(\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F))$ 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 $\mathrm{ind}(\mathbf{D}) = I(\mathbf{D})$ 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 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

$$\begin{aligned} \mathrm{ind}(\mathbf{D}) = \mathrm{ind}(\mathbf{D} + rB) &= \dim \ker(\mathbf{D} + rB) - \dim \ker(\mathbf{D} + rB)^* \\ &= c_1(\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F)) = I(\mathbf{D}). \end{aligned}$$

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

[5] We are assuming as usual that Hermitian inner products are complex antilinear in the first argument and linear in the second.

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 5.7. *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 5.8. *(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 3.53 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 Lecture 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 $\mathbf{D}_2$ 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 5.9. *$k = 1$ 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 the end of Lecture 4 that if we fix global Hermitian structures $\langle\ ,\ \rangle_E$ 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

$$\langle \eta, \xi \rangle_{L^2(E)} := \mathrm{Re} \int_{\dot{\Sigma}} \langle \eta, \xi \rangle_E \, d\,\mathrm{vol} \quad \text{for} \quad \eta, \xi \in \Gamma(E),$$

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

$$F = \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, E) = \Lambda^{0,1}T^*\dot{\Sigma} \otimes E = T\dot{\Sigma} \otimes E,$$

then E admits a natural isomorphism to $\Lambda^{1,0}T^*\dot{\Sigma} \otimes F$ such that

$$-\mathbf{D}^* : \Gamma(F) \to \Gamma(E) = \Omega^{1,0}(\dot{\Sigma}, F)$$

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 5.10. *If τ is a choice of asymptotic trivialization on E and $\bar{\tau}$ denotes the* conjugate *asymptotic trivialization*[6], *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 3.55. 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. $\square$

We are now able to show that Theorem 5.4 is consistent with what we already know about the formal adjoint.

PROPOSITION 5.11. $I(-\bar{\mathbf{D}}^*) = -I(\mathbf{D})$.

PROOF. Under the isomorphism $F = \Lambda^{0,1}T^*\dot{\Sigma} \otimes E = T\dot{\Sigma} \otimes E$, 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 5.10 to the conjugate bundle then gives

$$c_1^{\overline{\partial_s \otimes \tau}}(\bar{F}) = -n\chi(\dot{\Sigma}) - c_1^{\tau}(E).$$

[6] 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 $\bar{\tau}(\bar{v}) = (z, \bar{w})$.

The unitary trivializations of the asymptotic bundles $\bar{E}_z$ corresponding to $\overline{\partial_s \otimes \tau}$ are simply $\bar{\tau}$, thus using Lemma 5.10 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 $\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F) = \mathrm{Hom}_{\mathbb{C}}(\bar{E}, F)$.

EXERCISE 5.12. 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

$$\begin{aligned} \mathrm{Re}\langle \lambda, B\eta \rangle_F = \mathrm{Re}\langle \lambda, \beta\bar{\eta} \rangle_F = \mathrm{Re}\langle \beta^\dagger \lambda, \bar{\eta} \rangle_{\bar{E}} = \mathrm{Re}\langle \eta, \overline{\beta^\dagger \lambda} \rangle_E = \mathrm{Re}\langle \overline{\beta^\dagger \lambda}, \eta \rangle_E \\ = \mathrm{Re}\langle \bar{\beta}^\dagger \bar{\lambda}, \eta \rangle_E, \end{aligned}$$

where $\bar{\beta}^\dagger \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{F}, E))$ denotes the image of $\beta^\dagger \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(F, \bar{E}))$ under the complex-antilinear identity map from $\mathrm{Hom}_{\mathbb{C}}(F, \bar{E})$ to its conjugate bundle (see Exercise 5.12). 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 5.13. *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 $\overline{\mathrm{Hom}}_{\mathbb{C}}(\bar{F}, \bar{E}) = \mathrm{Hom}_{\mathbb{C}}(F, \bar{E})$ that identifies $-\bar{B}^*$ with $-\beta^\dagger$;
(3) If $n = 1$ and $B \in \Gamma(\overline{\mathrm{Hom}}_{\mathbb{C}}(E, F))$ has only nondegenerate zeroes, then $-\bar{B}^* \in \Gamma(\overline{\mathrm{Hom}}_{\mathbb{C}}(\bar{F}, \bar{E}))$ 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 $\beta \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F))$ and $-\beta^\dagger : \Gamma(\mathrm{Hom}_{\mathbb{C}}(F, \bar{E}))$ 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 5.4 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 $\mathrm{ind}(\mathbf{D}) = 0$. 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 $\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F)$ is a trivial bundle.

LEMMA 5.14. *$\mathbf{D}_r$ is injective for all $r > 0$ sufficiently large.*

PROOF. Elliptic regularity implies 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 $\eta \in C^\infty(\mathbb{T}^2, \mathbb{C})$,

$$\begin{aligned} \mathbf{D}_r^*\mathbf{D}_r\eta &= (\mathbf{D}^* + rB^*)(\mathbf{D} + rB)\eta \\ &= \mathbf{D}^*\mathbf{D}\eta + r\left(\beta\overline{\bar{\partial}\eta} - \partial(\beta\bar{\eta})\right) + r^2 B^*B\eta \\ &= \mathbf{D}^*\mathbf{D}\eta + r\left(\beta\partial\bar{\eta} - (\partial\beta)\bar{\eta} - \beta\partial\bar{\eta}\right) + r^2B^*B\eta \\ &= \mathbf{D}^*\mathbf{D}\eta + r^2B^*B\eta - r(\partial\beta)\bar{\eta}. \end{aligned} \tag{5.2}$$

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 5.4 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 5.13, 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 5.18.

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 $-\mathbf{D}^*$ is 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 5.15. 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 5.16. *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 $\eta \in \Gamma(E)$,

$$\mathbf{D}_r^* \mathbf{D}_r \eta = \mathbf{D}^* \mathbf{D} \eta + r^2 B^* B \eta + r B_1 \eta.$$

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 5.15, 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

$$
\begin{aligned}
\mathbf{D}_r^*\mathbf{D}_r\eta &= (\mathbf{D}^* + rB^*)(\mathbf{D} + rB)\eta \\
&= \mathbf{D}^*\mathbf{D}\eta + r\bar{\beta}^\dagger\overline{\mathbf{D}\eta} - r(-\mathbf{D}^*)(\beta\bar{\eta}) + r^2B^*B\eta \\
&= \mathbf{D}^*\mathbf{D}\eta + r\bar{\beta}^\dagger\bar{\mathbf{D}}\bar{\eta} - r(\partial_H\beta)\bar{\eta} - r\beta\bar{\mathbf{D}}\bar{\eta} + r^2B^*B\eta \\
&= \mathbf{D}^*\mathbf{D}\eta + r^2B^*B\eta - r(\partial_H\beta)\bar{\eta} + r\left(\bar{\beta}^\dagger - \beta\right)\bar{\mathbf{D}}\bar{\eta}.
\end{aligned}
$$

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

$$
\mathbf{D}_r^*\mathbf{D}_r\eta = \mathbf{D}^*\mathbf{D}\eta + r^2B^*B\eta - r(\partial_H\beta)\bar{\eta}. \tag{5.3}
$$

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

$$
\mathbf{D}_r^*\mathbf{D}_r - \mathbf{D}^*\mathbf{D} = r^2B^*B + r\left[(A^*B + B^*A)\eta - (\partial_H\beta)\bar{\eta}\right],
$$

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

$$S_\mathbb{C} := \frac{1}{2}(S - iSi), \qquad \text{hence} \qquad A = \frac{1}{2}(S + iSi).$$

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 5.17. 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. [**GW**].

REMARK 5.18. 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 5.16 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 5.19. *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 $C^\infty_{\text{loc}}(\mathbb{C})$ 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 Lecture 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

$$\dot{\Sigma}_\epsilon := \dot{\Sigma} \backslash \bigcup_{\zeta \in Z(\beta)} \mathcal{D}(\zeta),$$

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 5.16,

$$\begin{aligned} 0 &= \|\mathbf{D}_{r_\nu}\eta_\nu\|^2_{L^2(\dot{\Sigma})} = \langle \eta_\nu, \mathbf{D}^*_{r_\nu}\mathbf{D}_{r_\nu}\eta_\nu \rangle_{L^2(\dot{\Sigma})} \\ &= \langle \eta_\nu, \mathbf{D}^*\mathbf{D}\eta_\nu \rangle_{L^2(\dot{\Sigma})} + r_\nu^2 \langle \eta_\nu, B^* B \eta_\nu \rangle_{L^2(\dot{\Sigma})} + r_\nu \langle \eta_\nu, B_1 \eta_\nu \rangle_{L^2(\dot{\Sigma})} \\ &\geqslant \|\mathbf{D}\eta_\nu\|^2_{L^2(\dot{\Sigma})} + r_\nu^2 c^2 \|\eta_\nu\|^2_{L^2(\dot{\Sigma}_\epsilon)} - r_\nu c' \|\eta_\nu\|^2_{L^2(\dot{\Sigma})} \\ &\geqslant r_\nu^2 c^2 \|\eta_\nu\|^2_{L^2(\dot{\Sigma}_\epsilon)} - r_\nu c' \|\eta_\nu\|^2_{L^2(\dot{\Sigma})} \end{aligned}$$

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 5.20. Show that for any smooth complex-valued function f on an open subset of $\mathbb{C}$,

$$\Delta|f|^2 = 2\,\mathrm{Re}\langle f, \Delta f\rangle + 2|\nabla f|^2,$$

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 5.21. *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 $\mathbf{D}_- f = 0$, then the Weitzenböck formula for $\mathbf{D}_-$ implies $\Delta f = |z|^2 f$. Then by Exercise 5.20,

$$\Delta|f|^2 = 2\,\mathrm{Re}\langle f, |z|^2 f\rangle + 2|\nabla f|^2 = 2|z|^2|f|^2 + 2|\nabla f|^2,$$

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

$$\int_{\mathbb{D}_r(z_0)} |f(z)|^2\, d\mu(z) \geqslant \pi r^2 |f(z_0)|^2 \to \infty \qquad \text{as} \qquad r \to \infty,$$

so $f \notin L^2(\mathbb{C})$. □

PROPOSITION 5.22. *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 $u := \mathrm{Im}\, f : \mathbb{C} \to \mathbb{R}$ satisfies

$$\Delta u = (|z|^2 + 2)u.$$

Now by Exercise 5.20,

$$\Delta(u^2) = 2|\nabla u|^2 + 2(|z|^2 + 2)u^2 \geqslant 0,$$

so $u^2 : \mathbb{C} \to \mathbb{R}$ is subharmonic, and the mean value property implies as in the proof of Prop. 5.21 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 5.23. *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 $\eta_\nu^1, \dots, \eta_\nu^{I_+ + 1} \in \ker \mathbf{D}_{r_\nu}$ which form L^2-orthonormal sets for each ν. By Lemma 5.19, 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 $\zeta \in Z^+(\zeta)$ satisfy $\bar{\partial} f + z\bar{f} = 0$, while those for $\zeta \in Z^-(\zeta)$ satisfy $\bar{\partial} f + \bar{z}\bar{f} = 0$. Proposition 5.21 now implies that the component functions for $\zeta \in Z^-(\zeta)$ are all trivial, and by Proposition 5.22, the components for $\zeta \in Z^+(\zeta)$ 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 5.19. 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

$$\rho \in C_0^\infty(\mathring{\mathbb{D}}, [0, 1]), \qquad \rho|_{\mathbb{D}_{1/2}} \equiv 1$$

and define for each $\zeta \in Z^+(\beta)$ and $r > 0$ a linear map

$$\Phi_r^\zeta : \ker \mathbf{D}_+ \to \Gamma(E)$$

such that $\Phi_r^\zeta(f)$ is a section with support in $\mathcal{D}(\zeta)$ whose expression in our fixed coordinate and trivialization on that neighborhood is the function

$$f_r^\zeta(z) = \rho(z)\sqrt{r} f(\sqrt{r} z).$$

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 5.24. *For each $r > 0$, there exists a constant $c_r > 0$ such that*

$$\|\mathbf{D}_r \Phi_r(f)\|_{L^2} \leqslant c_r \|f\|_{L^2} \quad \textit{for all} \quad f \in \bigoplus_{\zeta \in Z^+(\beta)} \ker \mathbf{D}_+,$$

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 5.25. *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 $Z^-(\beta) = \varnothing$ implies via Lemma 5.23 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 5.24 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 5.26. *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 5.27 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 5.19, 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 Lecture 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 5.21. □

EXERCISE 5.27. 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 5.25. 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 5.24 and 5.26 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 5.28. *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 5.28 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, 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 5.29. *Given* $\mathbf{D}$ *as in Theorem 5.4, 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* $\mu^\tau_{CZ}(\mathbf{A}'_z) = \mu^\tau_{CZ}(\mathbf{A}_z)$*, 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*

$$\mathrm{ind}(\mathbf{D}) = \chi(\dot{\Sigma}) + 2c_1^\tau(E) + \sum_{z\in\Gamma^+} \mathrm{wind}(\beta_z) - \sum_{z\in\Gamma^-} \mathrm{wind}(\beta_z).$$

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 $\mathrm{ind}(\mathbf{D}') = \mathrm{ind}(\mathbf{D})$. 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 $\beta \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F))$ 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

$$\mathrm{ind}(\mathbf{D}) = \mathrm{ind}(\mathbf{D}_r).$$

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 $\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F)$ near $z \in \Gamma$ is $\beta_z(t)$. It follows that the signed count of zeroes of β is

$$\begin{aligned} i(\mathbf{D}) &:= c_1^\tau(\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F)) + \sum_{z\in\Gamma^+} \mathrm{wind}(\beta_z) - \sum_{z\in\Gamma^-} \mathrm{wind}(\beta_z) \\ &= \chi(\dot{\Sigma}) + 2c_1^\tau(E) + \sum_{z\in\Gamma^+} \mathrm{wind}(\beta_z) - \sum_{z\in\Gamma^-} \mathrm{wind}(\beta_z), \end{aligned}$$

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

$$\begin{aligned}\mathrm{Hom}_{\mathbb{C}}(\bar{E}, F) &= \bar{E}^* \otimes F = E \otimes F = E \otimes \Lambda^{0,1}T^*\dot{\Sigma} \otimes E = \Lambda^{0,1}T^*\dot{\Sigma} \otimes E \otimes E \\ &= T\dot{\Sigma} \otimes E \otimes E.\end{aligned}$$

We are free to assume that all zeroes of β are either positive or negative, depending on the sign of $i(\mathbf{D})$. Proposition 5.28 then implies $\mathrm{ind}(\mathbf{D}_r) = i(\mathbf{D})$ for large r. $\square$

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 5.4.

LEMMA 5.30. *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 Lecture 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 $S \in \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^2)$ 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 3.56 in Lecture 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 $\mathbf{A}_{k,r}$. $\square$

The proof of Theorem 5.4 is now complete.

EXERCISE 5.31. 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 $r > 0$.

LECTURE 6

Symplectic cobordisms and moduli spaces

Contents

In this lecture, we introduce the moduli spaces of holomorphic curves that are used to define SFT.

Recall that in Lecture 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. A useful generalization of this notion was introduced in [HZ94], and was later recognized to be the most natural geometric setting for punctured holomorphic curves. It 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 lecture by introducing *stable Hamiltonian structures.* 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. 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 6.2. 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]

EXERCISE 6.3. 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.

[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 $(w_1, \dots, w_k, v_1, \dots, v_m)$ of V such that $(w_1, \dots, w_k)$ is a positive basis of W, the quotient projection sends $(v_1, \dots, v_m)$ 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 6.15) together with an extra choice of ω-compatible complex structure J on $\xi = \ker \lambda$.

(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 6.4. 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 1.18) 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 6.5. *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 6.6. Show that if ∇ is any symmetric connection on M, then the symplectic connection ∇^ω on $\gamma^*\xi$ in Proposition 6.5 is given by the formula

$$\nabla^\omega_t \eta = \pi_\xi \left(\nabla_t \eta - \nabla_\eta R \right).$$

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 6.7. *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 6.6 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 3.4 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. $\square$

DEFINITION 6.8. 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 6.5.

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 the contact case, 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 6.5. Exercise 3.9 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 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 Lecture 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

$$\mu_{\mathrm{CZ}}^\tau(\gamma) := \mu_{\mathrm{CZ}}^\tau(\mathbf{A}_\gamma).$$

An explicit formula for $\mathbf{A}_\gamma$ comes from Exercise 6.6: 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 6.7, the projection π_ξ cannot always be omitted from this formula, though it can in the contact case.

In Lecture 3, the symmetry of the asymptotic operator in the contact setting was explained by interpreting it as the Hessian of the contact action functional $\mathcal{A}_\alpha(\gamma) := \int_{S^1} \gamma^*\alpha$. A similar interpretation is possible in this more general setting, though the action functional may be only locally defined. Indeed, while $\omega \in \Omega^2(M)$ need not be globally exact, it is necessarily exact on a neighborhood of the image of any given loop $\gamma_0 : S^1 \to M$, so one can pick a primitive β of ω on this neighborhood and, for a sufficiently small neighborhood $\mathcal{U}(\gamma_0) \subset C^\infty(S^1, M)$ of γ_0, consider the action functional

$$\mathcal{A}_\omega : \mathcal{U}(\gamma_0) \to \mathbb{R} : \gamma \mapsto \int_{S^1} \gamma^*\beta. \tag{6.1}$$

Its first variation at $\gamma \in \mathcal{U}(\gamma_0)$ in the direction $\eta \in \Gamma(\gamma^*\xi)$ is then

$$d\mathcal{A}_\omega(\gamma)\eta = -\int_{S^1} \omega(\dot{\gamma}, \eta)\, dt = \langle -J\pi_\xi\dot{\gamma}, \eta \rangle_{L^2},$$

where $\langle\ ,\ \rangle_{L^2}$ denotes the real L^2-product on $\gamma^*\xi$ defined by integrating $\omega(\cdot, J\cdot)$. This leads us to interpret $-J\pi_\xi\dot{\gamma}$ as a "gradient" $\nabla\mathcal{A}_\omega(\gamma)$, and if $\dot{\gamma} = T \cdot R(\gamma)$, then differentiating this gradient in the direction of $\eta \in \Gamma(\gamma^*\xi)$ gives $\mathbf{A}_\gamma\eta$.

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 6.3 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 6.9. *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 $(-\epsilon, \epsilon) \times M$.

[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_xM, the basis $(V(x), Y_1, \ldots, Y_{2n-1})$ of T_xW is also positively oriented.

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 $\omega(X(x), Y_j) = 0$ for all $j = 1, \ldots, 2n-2$. Since $\omega^{n-1}(Y_1, \ldots, Y_{2n-2})$ 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 $(V(x), X(x), Y_1, \ldots, Y_{2n-2})$ 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 $\omega_t := t\omega + (1-t)\omega$ for every $t \in [0, 1]$, which also matches ω along M. The latter implies that ω_t represents the same cohomology class in $H^2_{\mathrm{dR}}((-\epsilon, \epsilon) \times M) = H^2_{\mathrm{dR}}(M)$ for every $t \in [0, 1]$, thus we can find a smooth family of 1-forms β_t on $(-\epsilon, \epsilon) \times M$ satisfying

$$\omega_t = \omega_0 + d\beta_t \qquad \text{and} \qquad \beta_t|_M = 0 \text{ for all } t \in [0, 1].$$

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

$$0 = \mathcal{L}_{Y_t} \omega_t + \partial_t (d\beta_t) = d\iota_{Y_t} \omega_t + d\dot{\beta}_t$$

for $\dot{\beta}_t := \partial_t \beta_t$. This relation is then satisfied if we pick Y_t to be the unique vector field satisfying $\omega_t(Y_t, \cdot) = -\dot{\beta}_t$, 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. $\square$

REMARK 6.10. The statement about the tubular neighborhood in Proposition 6.9 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 6.11. In the case $(\omega, \lambda) = (d\alpha, \alpha)$ for a contact form α, the symplectic form on the tubular neighborhood in Proposition 6.9 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 1.14).

DEFINITION 6.12. 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 $-\lambda$.

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 6.9 and Remark 6.10, namely

$$
\begin{aligned}
(\mathcal{N}(M_+), \omega) &\cong ((-\epsilon, 0] \times M_+, \omega_+ + d(r\lambda_+)), \\
(\mathcal{N}(M_-), \omega) &\cong ([0, \epsilon) \times M_-, \omega_- + d(r\lambda_-)).
\end{aligned}
\tag{6.2}
$$

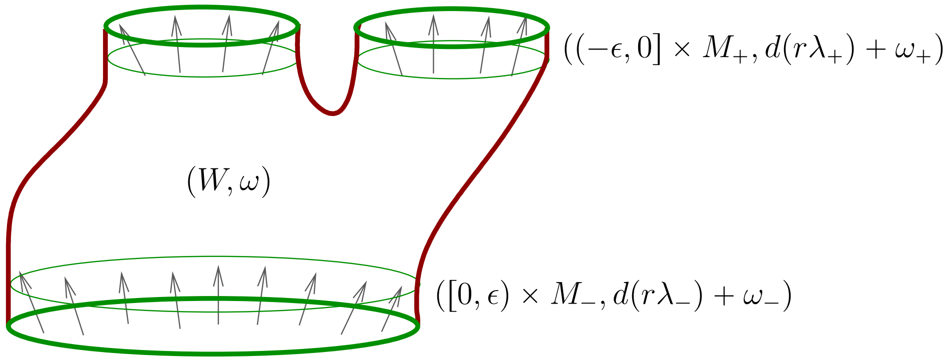

FIGURE 6.1. A symplectic cobordism with positive and negative boundary components $\partial W = -M_- \coprod M_+$ inheriting Hamiltonian structures $\omega_\pm$, shown with their symplectic collar neighborhoods determined by choices of framings $\lambda_\pm$.

REMARK 6.13. 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 Lecture 10 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]

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 6.14. 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,

[4]Apart from being an example of a symplectic cobordism with non-convex framed Hamiltonian boundary, the construction in Remark 6.13 amounts to the observation, originating in [Gir94], 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.

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 6.15. 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 6.16. *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 6.9 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. ☐

We can immediately observe two convenient features 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 6.7 now removes the need for including π_ξ in the formula of Exercise 6.6 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 6.17. A **symplectic cobordism with stable boundary** is a symplectic cobordism from $(M_-, \mathcal{H}_-)$ to $(M_+, \mathcal{H}_+)$ in the sense of Definition 6.12, 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. 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 1.21) 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 6.9. Indeed, fix $\epsilon > 0$ small and define

$$\mathcal{T} := \left\{\varphi \in C^\infty(\mathbb{R}, (-\epsilon, \epsilon)) \;\middle|\; \varphi' > 0\right\}, \tag{6.3}$$

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

$$\omega_\varphi := \omega + d\left(\varphi(r)\lambda\right) \tag{6.4}$$

on $\mathbb{R} \times M$.

There is a much more obvious generalization of the space $\mathcal{J}(\alpha)$ to the framed Hamiltonian setting.

DEFINITION 6.18. 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 $\omega|_\xi$.

[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 3.5).

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 Lecture 1. One of the crucial reasons to consider only *stable* Hamiltonian structures will be the following easy observation:

PROPOSITION 6.19. *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 (6.3) can be chosen such that the symplectic form ω_φ of (6.4) 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

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

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 Lecture 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.

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 §6.7 for further discussion.

EXAMPLE 6.20. 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 6.21. 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 6.22. For an instructive concrete example of Exercise 6.21, 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 9.5).

The trivial cylinders in Example 6.20 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 Lecture 9 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 6.21 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.

In order to see 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 6.20. 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 6.23. 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 Lecture 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 $\mathcal{A}_\omega : C^\infty(S^1) \to \mathbb{R}$ from §6.1.1, 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 6.23, and expressing arbitrary sections of $u_\gamma^* \varepsilon$ as $f \partial_r + gR$, we can apply the Leibniz rule for $\mathbf{D}_{u_\gamma}^\varepsilon$ and conclude

$$(\mathbf{D}_{u_\gamma}^\varepsilon \eta^\varepsilon)\partial_s = (\partial_s + i\,\partial_t)\eta^\varepsilon$$

in the canonical trivialization. The remaining off-diagonal term can be computed as follows: assume $u^\rho = (u_\mathbb{R}^\rho, u_M^\rho) : \mathbb{R} \times S^1 \to \mathbb{R} \times M$ 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

$$\partial_\rho u_\mathbb{R}^\rho\big|_{\rho=0} = \lambda\left(\partial_\rho u_M^\rho\big|_{\rho=0}\right) = 0.$$

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 6.24. *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 6.24 places the linearization $\mathbf{D}_{u_\gamma}$ into the analytical context of the Fredholm theory from Lectures 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.

6.2.2. 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 (6.2). Modifying (6.3) by

$$\mathcal{T}_0 := \left\{\varphi \in C^\infty(\mathbb{R}, (-\epsilon, \epsilon)) \;\middle|\; \varphi' > 0 \text{ and } \varphi(r) = r \text{ for } r \text{ near } 0\right\} \tag{6.5}$$

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

$$W^{r_0} := \left([-r_0, 0] \times M_-\right) \cup_{M_-} W \cup_{M_+} \left([0, r_0] \times M_+\right),$$

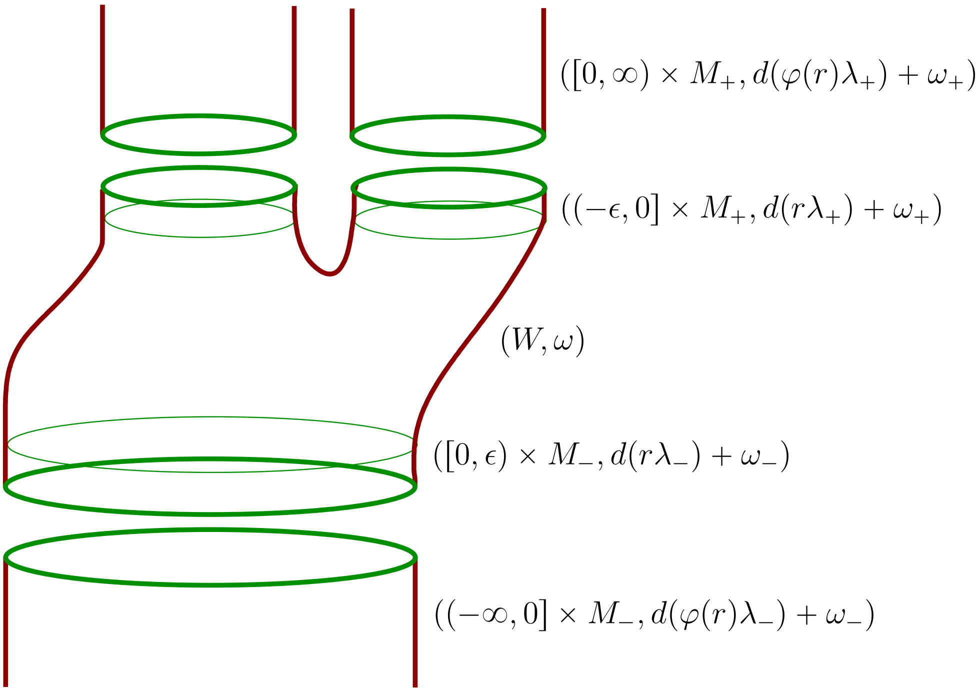

Figure 6.2. The completion $(\widehat{W},\omega_\varphi)$ of a symplectic cobordism between two manifolds with framed Hamiltonian structures.

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

$$\omega_\pm^{r_0} := \omega_\varphi\big|_{TM_\pm^{r_0}} = \omega_\pm + \varphi(\pm r_0)\, d\lambda_\pm, \qquad \text{and} \qquad \lambda_\pm^{r_0} := \omega_\varphi(\partial_r,\cdot)\big|_{TM_\pm^{r_0}} = \varphi'(\pm r_0)\ \lambda_\pm.$$

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.

Definition 6.25. Given $\psi\in\mathcal{T}_0$ and $r_0 \geqslant 0$, let

$$\mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-) \subset \mathcal{J}(\widehat{W})$$

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]

[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 Lecture 8 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.

- 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

$$\mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-) \subset \mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$$

denote the subset for which J is additionally compatible with ω_ψ on W^{r_0}.

Setting

$$\mathcal{T}(\psi, r_0) := \left\{\varphi \in \mathcal{T}_0 \;\middle|\; \varphi \equiv \psi \text{ on } [-r_0, r_0]\right\}, \tag{6.6}$$

Proposition 6.19 implies that if the framed Hamiltonian structures $\mathcal{H}_\pm$ are both stable, then any given $J \in \mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ is tamed by ω_φ for every $\varphi \in \mathcal{T}(\psi, r_0)$ whenever the number $\epsilon > 0$ in (6.5) 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

$$E(u) := \sup_{\varphi \in \mathcal{T}(\psi, r_0)} \int_\Sigma u^* \omega_\varphi.$$

REMARK 6.26. 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 $\mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ without inventing questionable new notions such as the "infinitesimal trivial cobordism" $[0,0] \times M$. It is nonetheless true for fairly trivial reasons that most results about $\mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ or $\mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ 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 Lecture 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 (6.2), 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 6.27. 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 6.28. Work out the relationship between the locally defined action functional $\mathcal{A}_\omega$ from §6.1.1 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 ϕ :

$(W, \omega) \to (W, \omega)$, 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 6.29. 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

$$u(s - s_0, t - t_0) = \exp_{u_{\gamma_z}(s,t)} h_z(s, t) \quad \text{for } (s, t) \in Z_\pm \cong \dot{\mathcal{U}}_z \text{ with } |s| \gg 0, \tag{6.7}$$

where $h_z(s, t)$ is a vector field along u_{γ_z} satisfying

$$h_z(\cdot + s, \cdot) \to 0 \quad \text{in } C^\infty(Z_\pm) \qquad \text{as } s \to \pm\infty.$$

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 (6.7) the **asymptotic representative** of u at z.

Note that the decay condition in Definition 6.29 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.

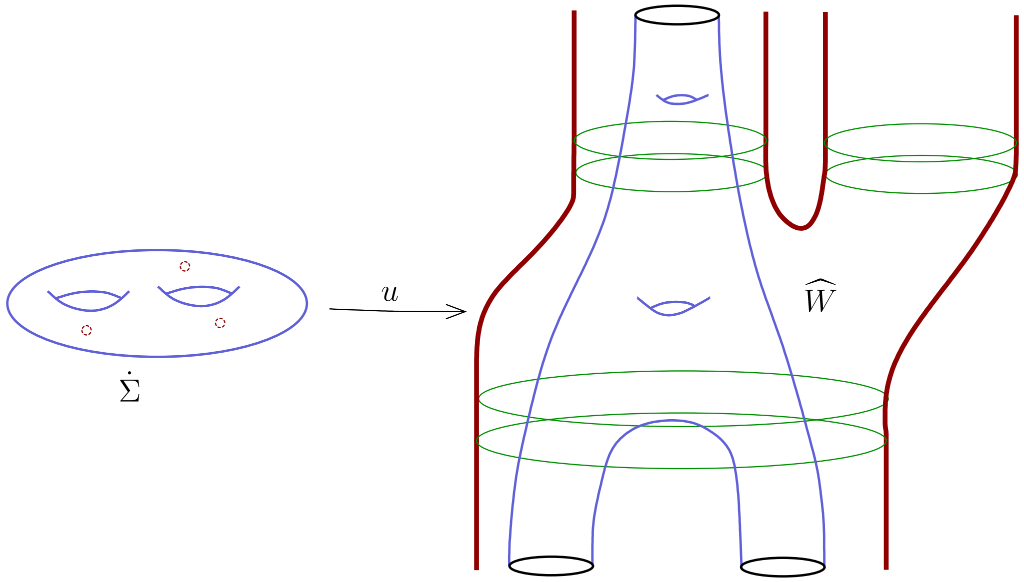

Figure 6.3. An asymptotically cylindrical holomorphic curve in $(\widehat{W}, J)$ with genus 2, one positive puncture and two negative punctures.

Exercise 6.30. Consider S^1 with the trivial stable Hamiltonian structure $\mathcal{H}$ (see Remark 6.22) 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 6.29 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 $J \in \mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$. 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. We will prove in Lecture 9 that the converse is also true whenever all Reeb orbits are nondegenerate.

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

$$\partial A = \sum_{i=1}^{k_+} [\gamma_i^+] - \sum_{i=1}^{k_-} [\gamma_i^-] \in H_1(\bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-).$$

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

$$u_1 \circ \psi = u_0.$$

We shall often abuse notation by abbreviating elements $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)]$ in this moduli space by

$$u \in \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-).$$

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 $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \in \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$, and we will see in §6.6 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**

$$\mathrm{ev} : \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) \to \widehat{W}^{\times m} : [(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \mapsto (u(\zeta_1), \dots, u(\zeta_m))$$

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 6.31. The definition of $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ given above permits elements $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \in \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ 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 $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ 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, so in conjunction with Theorem 6.34 below, this implies that the automorphism group $\mathrm{Aut}(u)$ for an element $u \in \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ is *always* finite.

The elliptic regularity results from Lectures 2 and 4 give us a wide range of freedom in defining the topology of $\mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$, as they imply that most reasonable choices we could conceivably make on this front will turn out to be equivalent. Let us define the notion of convergence in $\mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ as follows:

$$[(\Sigma_\nu, j_\nu, \Gamma_\nu^+, \Gamma_\nu^-, \Theta_\nu, u_\nu)] \to [(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \quad \text{as } \nu \to \infty$$

means that 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^∞;
(2) $u'_\nu \to u$ in $C^\infty_{\text{loc}}(\dot{\Sigma}, \widehat{W})$;
(3) $\bar{u}'_\nu \to \bar{u}$ in $C^0(\bar{\Sigma}, W)$.

It will also turn out that the third condition is unnecessary if one adds an assumption to make sure that all closed Reeb orbits are isolated, though the proof of this is nontrivial and requires the bubbling analysis that we will discuss in Lecture 9. One can show that there is a unique metrizable topology on $\mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ for which this is the notion 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 $\mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ 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.

6.5. Asymptotic regularity

For the analytic setup in the next lecture, 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 $W^{k,p}_{\text{loc}}$ 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 $W^{k,p}_{\text{loc}}$ for every k and p. It remains only to check that suitable decay conditions are satisfied on each of the cylindrical ends. This requires some of the same ideas that were used in §4.6 to prove exponential decay for solutions of linear Cauchy-Riemann type equations.

The following result originates (in a somewhat less general context) in [**HWZ96**], and the complete proof would be too lengthy to present here, but we will give a sketch. 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 6.29 the notion of the *asymptotic representative* of a holomorphic curve at a puncture.

PROPOSITION 6.32. *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 nondegenerate Reeb orbit and $\delta > 0$ is small enough so that the asymptotic operator $\mathbf{A}_\gamma$ has no eigenvalues in $[-\delta, \delta]$.*

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 with convergence

$$j_\nu \to j \text{ in } C^\infty(\Sigma), \quad u_\nu \to u \text{ in } C^\infty_{\mathrm{loc}}(\dot{\Sigma}), \quad \text{and} \quad \bar{u}_\nu \to \bar{u} \text{ in } C^0(\bar{\Sigma}),$$

where j_ν also matches j on some fixed neighborhood of the puncture z for every ν. 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 6.33. The obvious analogue of Proposition 6.32 for curves in completed cobordisms also holds, with no meaingful change to the proof.

SKETCH OF THE PROOF OF PROPOSITION 6.32. After a slight reparametrization of u near infinity, one can assume the asymptotic representative $h \in \Gamma(u_\gamma^* T(\mathbb{R} \times M)|_{Z_\pm})$ for u at $z \in \Gamma^\pm$ takes values in the subbundle $u_\gamma^* \xi$, which can be regarded as the normal bundle of the trivial cylinder. The following general principle then applies: whenever $v : (S, j) \to (W, J)$ and $w : (S', j') \to (W, J)$ are two immersed J-holomorphic curves such that w is obtained by exponentiating a section η of the normal bundle of v, the section η satisfies a linear Cauchy-Riemann type equation defined on that normal bundle. The proof of this is similar to the argument in Proposition 2.34; for a complete account, see e.g. [**Wen20**, Proposition B.28]. In the present context, it means that our asymptotic representative h is annihilated by some linear Cauchy-Riemann type operator $\mathbf{D}$ that is defined on the bundle $u_\gamma^* \xi$ over one cylindrical end, and the asymptotic behavior of u implies that that operator is C^∞-asymptotic to the nondegenerate asymptotic operator $\mathbf{A}_\gamma$. Proposition 4.26 then implies $h \in W^{k,p,\delta}(Z_\pm)$ for all $k \in \mathbb{N}$ and $p \in (1, \infty)$.

The case of a converging sequence is handled similarly; cf. Exercises 4.24 and 4.27. $\square$

6.6. 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 6.34. *Assume $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ is a nonconstant asymptotically cylindrical J-holomorphic curve whose asymptotic orbits are all nondegenerate,* *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 2.35 in Lecture 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 6.32 (cf. Lemma 15.4).

LEMMA 6.35 (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 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*

$$u(z) = v(z^k) \quad \text{for} \quad z \in \dot{\mathbb{D}} = \dot{\mathcal{U}}_{z_0},$$

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 6.36. With Lemma 6.35 in hand, adapt the proof of Theorem 2.35 in Lecture 2 to prove Theorem 6.34. If you get stuck, see [Nel15, §3.2].

PROPOSITION 6.37. *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* $\mathrm{Aut}(u)$ *is trivial. If it is represented by a* k*-fold cover of a simple curve, then* $|\mathrm{Aut}(u)| \leqslant k$. *In particular,* $\mathrm{Aut}(u)$ *is always finite.*[7]

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

$$v : \Sigma' \to W$$

[7] cf. Remark 6.31 for the case where u is constant.

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. $\square$

6.7. 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.

6.7.1. Asymptotically cylindrical ends. When (W, ω) is a symplectic cobordism with stable boundary $(M_\pm, \mathcal{H}_\pm)$ and $J \in \mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ 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 $\tau_c : \mathbb{R} \times S^1 \to \mathbb{R} \times S^1$ 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].

6.7.2. 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 6.24 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 two lectures, we will need those linearized Cauchy-Riemann operators to be Fredholm, and our proof of this in Lecture 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 and §6.5, and the existing proofs of Lemma 6.35, 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 3.5, 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 $\nabla^2 f(x) : T_x M \to T_x M$ of a Morse function $f : M \to M$ 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 $DX(x) : T_x M \to T_x M$ 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 $DX(x) : T_x M \to T_x M$ 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 $DX(x)$: $T_x M \to T_x M$ 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 $\mathcal{A}_\omega$ of §6.1.1, because

$$d\mathcal{A}_\omega(\omega)V(\gamma) = -\int_{S^1} \omega(\dot{\gamma}, -J\pi_\xi\dot{\gamma})\, dt = \int_{S^1} \omega(\pi_\xi\dot{\gamma}, J\pi_\xi\dot{\gamma})\, dt \geqslant 0,$$

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 4.14 on the invertibility of translation-invariant operators $\partial_s - \mathbf{A}_\gamma$ over the cylinder, which was the main technical step in our proof of the Fredholm property in Lecture 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.

6.7.3. 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 (6.2) 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 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 lecture why the theory of J-holomorphic curves may run

[8]Needless to say, I learned this expression from my Ph.D. advisor.

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 (6.5), and for a non-stable Hamiltonian structure, ω_φ does not tame J for every choice of φ. We will see in Lecture 9 that the ability to choose $\varphi \in \mathcal{T}_0$ arbitrarily is essential, and as a consequence, there really is no reasonable compactness theory for J-holomorphic curves on cobordisms with non-stable boundary.

But 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 §6.7.2 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 6.24: 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 Lecture 4. On the other hand, it is quite easy to see that the particular consequence of Theorem 4.14 we will need in the next lecture 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 we will see in the next lecture how to rectify this 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 Lecture 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 in the next two lectures 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.

6.7.4. 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 the next lecture, namely that after a generic perturbation of J, the moduli space becomes a smooth manifold whose dimension is determined by the index formula in Lecture 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 without any taming condition there is no reason for this energy to be bounded. As we will see in Lecture 9, 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 $\xi_\pm \subset T(\{\mathrm{r}\} \times M_\pm)$.

One now runs into a starker version of the problem already discussed in §6.7.2: 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 4.14 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.

LECTURE 7

Smoothness of the moduli space

Contents

In this lecture, we study the local structure of the moduli space

$$\mathcal{M}(J) := \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$$

introduced in §6.4. We assume as before that (W, ω) is a $2n$-dimensional symplectic cobordism with stable boundary $\partial W = -M_- \coprod M_+$ inheriting stable Hamiltonian structures $\mathcal{H}_\pm = (\omega_\pm, \lambda_\pm)$ with induced Reeb vector fields $R_\pm$ and hyperplane distributions $\xi_\pm = \ker \lambda_\pm$, while $g, m, k_+, k_- \geqslant 0$ are integers, $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \ldots, \gamma_{k_\pm}^\pm)$ are ordered sets of closed $R_\pm$-orbits in $M_\pm$, and $A \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ is a relative homology class with $\partial A = \sum_i [\gamma_i^+] - \sum_i [\gamma_i^-] \in H_1(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$. The noncompact completion of (W, ω) is denoted by $(\widehat{W}, \omega_\psi)$ for some fixed function $\psi : \mathbb{R} \to (-\epsilon, \epsilon)$ that scales the symplectic form on the cylindrical ends, and $r_0 \geqslant 0$ is a fixed constant which determines the size of the ends $[r_0, \infty) \times M_+$ and $(-\infty, -r_0] \times M_-$ on which we require our almost complex structures

$$J \in \mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$$

to be $\mathbb{R}$-invariant. The complement of these ends has closure

$$W^{r_0} := ([-r_0, 0] \times M_-) \cup_{M_-} W \cup_{M_+} ([0, r_0] \times M_+) .$$

We will often make use of the fact that since J matches translation-invariant almost complex structures in $\mathcal{J}(\mathcal{H}_\pm)$ outside of W^{r_0}, there are natural complex vector bundle splittings

$$T(\mathbb{R} \times M_\pm) = \varepsilon_\pm \oplus \xi_\pm,$$

where $\varepsilon_\pm$ denotes the canonically trivial line bundle spanned by ∂_r and the Reeb vector field $R_\pm$.

We will sometimes also consider the slightly simpler situation of J-holomorphic curves in $(\mathbb{R} \times M, J)$, where M is a closed manifold with a fixed stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ and $\mathbb{R}$-invariant almost complex structure $J \in \mathcal{J}(\mathcal{H})$. In this case, we shall write

$$\widehat{W} := \mathbb{R} \times M, \qquad M_\pm := M, \qquad \text{and} \qquad \mathcal{H}_\pm = (\omega_\pm, \lambda_\pm) := \mathcal{H} = (\omega, \lambda),$$

and regard A as an element of $H_2(M, \bar{\gamma}^+ \cup \bar{\gamma}^-)$.

7.1. The main result on regular curves

The major theorem we need to prove in this lecture gives $\mathcal{M}(J)$ a smooth structure in the presence of suitable transversality conditions. It will be followed in the next lecture by two important results examining when these transversality conditions can be achieved. Before stating the first result, there is a word that may need some clarification: orbifolds, introduced originally by Satake [Sat56] under a different name, are Hausdorff topological spaces that are locally homeomorphic to open subsets of vector spaces divided by finite group actions. More precisely, if we say that a space M admits the structure of an **n-dimensional orbifold** with **local isotropy group** G at a given point $x \in M$, this implies that there is a homeomorphism

$$\mathcal{U} \xrightarrow{\varphi} \mathcal{O}/G,$$

where $\mathcal{U} \subset M$ is a neighborhood of x, G is a finite group with a linear action[1] on $\mathbb{R}^n$, $\mathcal{O} \subset \mathbb{R}^n$ is a G-invariant open neighborhood of 0 and $\varphi(x) = 0$. It is important to understand that the isotropy group can vary from point to point, but it is always required to be finite, and one can easily show that if a given point $x \in M$ has isotropy group G, then x has a neighborhood in which all points have isotropy groups isomorphic to subgroups of G. This implies that the set of points with trivial isotropy group is open, and this subset is then a manifold. More should be said about the precise meaning of *smoothness* on an orbifold and what a smooth map between orbifolds is. Definitions for these notions may be found e.g. in [ALR07, Dav, FO99]), and they vary slightly among different sources; an elegant presentation in the language of *groupoids* is given in [McD06, §2]. For the sake of applications treated in this book, it will not be necessary to know the precise definitions, as we will only need to consider orbifolds with trivial isotropy, i.e. manifolds. For the following theorem, it will suffice to understand that an orbifold is what arises if you divide a manifold by a smooth and proper Lie group action that is not necessarily free but has at most a finite stabilizer subgroup at every point.

THEOREM 7.1. *Suppose either $J \in \mathcal{J}(\mathcal{H})$ or $J \in \mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$, and that the orbits $\gamma_i^\pm$ are all nondegenerate. Then the moduli space $\mathcal{M}(J)$ contains an open subset*

$$\mathcal{M}^{\mathrm{reg}}(J) \subset \mathcal{M}(J)$$

[1] Some sources require the G-action on $\mathbb{R}^n$ to be linear, and some only require it to be a smooth action that fixes the origin. The two notions are equivalent: if the G-action is smooth but not linear, one can use averaging to construct a G-invariant Riemannian metric on $\mathcal{O}$ and then observe that the exponential map identifies a neighborhood of 0 in $T_0\mathbb{R}^n$ equivariantly with $\mathcal{O}$, where the G-action on $T_0\mathbb{R}^n$ is defined by linearizing the original G-action at the origin.

consisting of so-called Fredholm regular *curves, which naturally admits the structure of a smooth orbifold of dimension*

$$\dim \mathcal{M}^{\mathrm{reg}}(J) = (n-3)(2-2g-k_+-k_-) + 2c_1^\tau(A) + \sum_{i=1}^{k_+} \mu_{CZ}^\tau(\gamma_i^+) - \sum_{i=1}^{k_-} \mu_{CZ}^\tau(\gamma_i^-) + 2m,$$

where $\dim W = 2n$*,* τ *is a choice of unitary trivialization for* $(\xi_\pm, J, \omega_\pm)$ *along each of the asymptotic orbits* $\gamma_i^\pm$*, and* $c_1^\tau(A)$ *denotes the relative first Chern number of the complex vector bundle* $(u^*T\widehat{W}, J) \to \dot{\Sigma}$ *with respect to the asymptotic trivialization determined by* τ *and the splitting* $T(\mathbb{R} \times M_\pm) = \varepsilon_\pm \oplus \xi_\pm$*. The local isotropy group of* $\mathcal{M}^{\mathrm{reg}}(J)$ *at* u *is* $\mathrm{Aut}(u)$*, hence the moduli space is a manifold near any regular element with trivial automorphism group.*

EXERCISE 7.2. Verify that the number in the dimension formula above is independent of the choice of trivializations τ, and that $c_1^\tau(u^*T\widehat{W})$ depends only on the relative homology class A.

The integer in the above dimension formula is often called the **virtual dimension** of $\mathcal{M}(J)$ and denoted by

$$\text{vir-dim}\, \mathcal{M}(J) := (n-3)(2-2g-k_+-k_-) + 2c_1^\tau(A) + \sum_{i=1}^{k_+} \mu_{\mathrm{CZ}}^\tau(\gamma_i^+) - \sum_{i=1}^{k_-} \mu_{\mathrm{CZ}}^\tau(\gamma_i^-) + 2m.$$

Ignoring the marked points, the virtual dimension of a space $\mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ containing a curve $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ with punctures $z \in \Gamma^\pm$ and nondegenerate asymptotic orbits $\{\gamma_z\}_{z\in\Gamma^\pm}$ is sometimes also called the **index** of u,

$$\mathrm{ind}(u) := (n-3)\chi(\dot{\Sigma}) + 2c_1^\tau(u^*T\widehat{W}) + \sum_{z\in\Gamma^+} \mu_{\mathrm{CZ}}^\tau(\gamma_z) - \sum_{z\in\Gamma^-} \mu_{\mathrm{CZ}}^\tau(\gamma_z) \in \mathbb{Z},$$

and we will see that it is in fact the Fredholm index of an operator closely related to the linearized Cauchy-Riemann operator $\mathbf{D}_u$ at u. The word "virtual" refers to the fact that in general, the regularity condition may fail, and thus $\mathcal{M}(J)$ might not be smooth, or if it is, it might actually be of a different dimension (see Example 8.8 below), but in an ideal world where transversality is always satisfied, its dimension would be vir-dim $\mathcal{M}(J)$. This notion makes sense in finite-dimensional contexts as well: if $f : \mathbb{R}^n \to \mathbb{R}^m$ is a smooth map, then we would say that $f^{-1}(0)$ has virtual dimension $n-m$, even though $f^{-1}(0)$ might in general be all sorts of strange things other than a smooth $(n-m)$-dimensional manifold. In particular, $n-m$ could be negative, in which case $f^{-1}(0)$ would be empty if transversality were satisfied, but in general this need not be the case. It is true however that f can always be *perturbed* to a map whose zero set is an $(n-m)$-dimensional manifold (or empty if $n-m < 0$). The same is true in principle of the nonlinear Cauchy-Riemann equation, and in the next lecture, we will prove that the transversality condition behind the definition of $\mathcal{M}^{\mathrm{reg}}(J) \subset \mathcal{M}(J)$ can be achieved at least for the *somewhere injective* curves in $\mathcal{M}(J)$ after allowing C^∞-small perturbations of J. In general, it is a formidably difficult problem to find perturbations that are sufficiently generic to achieve transversality while also respecting all symmetries inherent in the setup as well as the extra structure provided by the *compatification* of $\mathcal{M}(J)$, which is usually crucial for meangingful applications. Such issues require more sophisticated methods than we will discuss here, but a good place to read about them is [**FFGW16**].

Sections 7.2 and 7.3 below serve as preparation for the proof of Theorem 7.1, which is carried out in §7.4. Section 7.5 then proves a few more specialized results about the space $\mathcal{M}^{\mathrm{reg}}(J)$ that are useful in applications.

7.2. Functional-analytic setup

The first step in proving Theorem 7.1 is to define a suitable Banach manifold of maps for candidate solutions of the nonlinear Cauchy-Riemann equation to live in. Before doing this, a brief digression on the definition of the word "manifold" is in order.

REMARK 7.3. In finite dimensions, the standard convention (which we follow in this book) is to require all manifolds to be not only locally Euclidean but also Hausdorff and second countable. Those last two conditions are important, for instance because without both of them, the standard classification of compact 1-manifolds as disjoint unions of circles and intervals becomes false, and many fundamental results in differential topology (including the proof that $\partial^2 = 0$ in all Floer-type theories) depend on that classification. On the other hand, conventions vary on the precise set of topological adjectives that should be associated to the word "manifold" in infinite dimensions. In the classic book by Lang [**Lan99**], smooth Banach manifolds are not assumed to be Hausdorff, second countable or paracompact except when each of those conditions is specifically needed. So for instance, they are not needed in the implicit function theorem, but this observation comes with caveats: if you use the implicit function theorem to cut a finite-dimensional submanifold out of a non-Hausdorff Banach manifold, then the submanifold might not be Hausdorff either. In practice, all of the important manifolds we encounter in this book, both finite and infinite dimensional, will be metrizable and separable, from which it follows that they are both Hausdorff and second countable. Some care is warranted because for an infinite-dimensional Banach space, separability is typically not obvious, and without the second countability axiom, fundamental results like the Sard-Smale theorem (see §8.1) become false.

Fix $k \in \mathbb{N}$ and $p \in (1, \infty)$ with $kp > 2$, a small number $\delta \geqslant 0$, and a Riemannian metric on $\widehat{W}$ that is translation-invariant in the cylindrical ends. Fix also a closed connected surface Σ of genus g, and disjoint finite ordered sets of distinct points

$$\Gamma^{\pm} = (z_1^{\pm}, \dots, z_{k_{\pm}}^{\pm}), \qquad \Theta = (\zeta_1, \dots, \zeta_m)$$

in Σ, together with disjoint neighborhoods

$$\mathcal{U}_j^{\pm} \subset \Sigma$$

of each $z_j^\pm \in \Gamma^\pm$ with complex structures j_Γ and biholomorphic identifications of $(\mathcal{U}_j^\pm, j_\Gamma, z_j)$ with $(\mathbb{D}, i, 0)$ for each $j = 1, \ldots, k_\pm$. This determines holomorphic cylindrical coordinates identifying each of the punctured neighborhoods

$$\dot{\mathcal{U}}_j^\pm \subset \dot{\Sigma} := \Sigma \backslash (\Gamma^+ \cup \Gamma^-)$$

biholomorphically with the half-cylinder $Z_\pm$.

For reasons that will become clear when we study the linearized Cauchy-Riemann operator in the punctured setting, we will need to consider exponentially weighted Sobolev spaces. Suppose $E \to \dot{\Sigma}$ is an asymptotically Hermitian vector bundle: then the Banach space

$$W^{k,p,\delta}(E) \subset W^{k,p}_{\mathrm{loc}}(E)$$

is defined to consist of sections $\eta \in W^{k,p}_{\mathrm{loc}}(E)$ whose representatives $f : Z_\pm \to \mathbb{C}^m$ in cylindrical coordinates $(s,t) \in Z_\pm$ and asymptotic trivializations at the ends satisfy

$$\|e^{\pm\delta s} f\|_{W^{k,p}(Z_\pm)} < \infty. \tag{7.1}$$

The norm of a section $\eta \in W^{k,p,\delta}(E)$ is defined by adding the $W^{k,p}$-norm of η over a large compact subdomain in $\dot{\Sigma}$ to the weighted norms (7.1) for each cylindrical end. If $\delta = 0$, this just produces the usual $W^{k,p}(E)$, but for $\delta > 0$, sections in $W^{k,p,\delta}(E)$ are guaranteed to have exponential decay at infinity.

REMARK 7.4. It is occasionally useful to observe that the definition of $W^{k,p,\delta}(E)$ also makes sense when $\delta < 0$. In this case, sections in $W^{k,p,\delta}(E)$ are of class $W^{k,p}_{\mathrm{loc}}$ but need not be globally in $W^{k,p}(E)$, as they are also allowed to have exponential *growth* at infinity.

We now want to define a Banach manifold of maps $u : \dot{\Sigma} \to \widehat{W}$ that will contain all the asymptotically cylindrical J-holomorphic curves with our particular choice of asymptotic orbits. Recall from §6.4 that the asymptotically cylindrical condition means

$$u(s - s_0, t - t_0) = \exp_{(T_j^\pm s, \gamma_j^\pm(t))} h_j(s,t) \qquad \text{for sufficiently large } |s| \tag{7.2}$$

in holomorphic cylindrical coordinates $(s,t) \in Z_\pm$ near each puncture $z_j^\pm \in \Gamma^\pm$, where $T_j^\pm > 0$ is the period of the orbit $\gamma_j^\pm : S^1 \to M_\pm$, $h_j(s,t)$ is a vector field along the trivial cylinder that decays along with all its derivatives as $s \to \pm\infty$, and $s_0 \in \mathbb{R}$ and $t_0 \in S^1$ are constants. Equivalently, every asymptotically cylindrical map can be assumed to satisfy

$$u(s,t) = \exp_{(T_j^\pm s + a, \gamma_j^\pm(t+b))} h_j(s,t), \qquad \lim_{s \to \pm\infty} h(s,t) = 0$$

for some constants $a \in \mathbb{R}$ and $b \in S^1$. Here the exponential map is defined with respect to a connection that is $\mathbb{R}$-invariant on the cylindrical ends; since W, M_+ and M_- are compact, none of the resulting definitions depend on this choice so long as the $\mathbb{R}$-invariance condition is satisfied. Let us fix a neighborhood

$$\mathcal{D} \subset T\widehat{W}$$

of the zero-section that is also $\mathbb{R}$-invariant on the cylindrical ends and small enough to have the property that for the bundle projection $\pi : T\widehat{W} \to \widehat{W}$, the map

$$(\pi, \exp)|_{\mathcal{D}} : \mathcal{D} \hookrightarrow W \times W$$

is a well-defined diffeomorphism onto a neighborhood of the diagonal. We then define the space

$$\mathcal{B}^{k,p,\delta} := W^{k,p,\delta}(\dot{\Sigma}, \widehat{W}\,; \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) \subset C^0(\dot{\Sigma}, \widehat{W})$$

to consist of all continuous maps $u : \dot{\Sigma} \to \widehat{W}$ of the form

$$u(z) = \exp_{f(z)} h(z),$$

where:

- $f : \dot{\Sigma} \to \widehat{W}$ is smooth and, in our fixed cylindrical coordinates $(s,t) \in Z_\pm$ on neighborhoods of the punctures $z_j^\pm \in \Gamma^\pm$, takes the form
$$f(s,t) = (T_j^\pm s + a, \gamma_j^\pm(t + b)) \quad \text{for } |s| \text{ sufficiently large},$$
where $a \in \mathbb{R}$ and $b \in S^1$ are arbitrary constants and $T_j^\pm > 0$ is the period of the Reeb orbit $\gamma_j^\pm : S^1 \to M_\pm$;
- $h \in W^{k,p,\delta}(f^*T\widehat{W})$ and $h(z) \in \mathcal{D}$ for all $z \in \dot{\Sigma}$.

Though it is not immediate since $\dot{\Sigma}$ is noncompact, one can adapt the ideas in [**Elĭ67**] or [**Pal68**] to give $\mathcal{B}^{k,p,\delta}$ the structure of a smooth, separable and metrizable Banach manifold. The key point is the condition $kp > 2$, which guarantees the continuous inclusion $W^{k,p,\delta}(f^*T\widehat{W}) \hookrightarrow C^0(f^*T\widehat{W})$ as well as Banach algebra and C^k-continuity properties, cf. Propositions 2.4, 2.7 and 2.8 in Lecture 2. These properties are needed in order to show that the transition maps between pairs of charts of the form $\exp_f h \mapsto h$ are smooth.

The tangent space to $\mathcal{B}^{k,p,\delta}$ at $u \in \mathcal{B}^{k,p,\delta}$ can be written as

$$T_u\mathcal{B}^{k,p,\delta} = W^{k,p,\delta}(u^*T\widehat{W}) \oplus V_\Gamma,$$

where $V_\Gamma \subset \Gamma(u^*T\widehat{W})$ is a non-canonical choice of a $2(k_+ + k_-)$-dimensional vector space of smooth sections asymptotic at the punctures to constant linear combinations of the vector fields spanning the canonical trivialization of the first factor in $T(\mathbb{R} \times M_\pm) = \varepsilon_\pm \oplus \xi_\pm$, i.e. they point in the ∂_r- and $R_\pm$-directions. The space V_Γ appears due to the fact that two distinct elements of $\mathcal{B}^{k,p,\delta}$ are generally asymptotic to collections of trivial cylinders that differ from each other by $k_+ + k_-$ pairs of constant shifts $(a,b) \in \mathbb{R} \times S^1$.

Fix $J \in \mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ and a smooth complex structure j on Σ that matches j_Γ in the neighborhoods $\mathcal{U}_j^\pm$ of the punctures. The nonlinear Cauchy-Riemann operator is then defined as a smooth section

$$\bar{\partial}_{j,J} : \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta} : u \mapsto du + J(u) \circ du \circ j$$

of a Banach space bundle

$$\mathcal{E}^{k-1,p,\delta} \to \mathcal{B}^{k,p,\delta}$$

with fibers

$$\mathcal{E}_u^{k-1,p,\delta} = W^{k-1,p,\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u^*T\widehat{W})).$$

The zero set of $\bar{\partial}_{j,J}$ is the set of all maps $u \in \mathcal{B}^{k,p,\delta}$ that are pseudoholomorphic from $(\dot{\Sigma}, j)$ to $(\widehat{W}, J)$. Note that the smoothness of $\bar{\partial}_{j,J}$ depends mainly on the fact that J is smooth. Indeed, in local coordinates $\bar{\partial}_{j,J}$ looks like $u \mapsto \partial_s u + (J \circ u)\partial_t u$, in which the most obviously nonlinear ingredient is $u \mapsto J \circ u$. If J were only of class C^k, then the C^k-continuity property would imply that the map $u \mapsto J \circ u$ sends maps of class $W^{k,p}$ continuously to maps of class $W^{k,p}$, and one can use an inductive argument to show that this map then becomes r-times differentiable if J is of class C^{k+r}; see [**Wend**, Lemma 2.12.5]. Moreover, $\bar{\partial}_{j,J}(u)$ satisfies the same exponential weighting condition as u at the cylindrical ends due to the fact that J is $\mathbb{R}$-invariant near infinity.

For $u \in \bar{\partial}_{j,J}^{-1}(0)$, the linearization $D\bar{\partial}_{j,J}(u) : T_u\mathcal{B}^{k,p,\delta} \to \mathcal{E}_u^{k-1,p,\delta}$ defines a bounded linear operator

$$\mathbf{D}_u : W^{k,p,\delta}(u^*T\widehat{W}) \oplus V_\Gamma \to W^{k-1,p,\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u^*T\widehat{W})).$$

We derived a formula for this operator in Lecture 2 and showed that it is of Cauchy-Riemann type. Since V_Γ is finite dimensional, $\mathbf{D}_u$ will be Fredholm if and only if its restriction to the first factor is Fredholm; denote this restriction by

$$\mathbf{D}_\delta : W^{k,p,\delta}(u^*T\widehat{W}) \to W^{k-1,p,\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u^*T\widehat{W})),$$

where we've chosen the notation to emphasize the dependence of this operator on the choice of exponential weight $\delta \geqslant 0$ in the definition of our Banach space. We will see presently why it's important to pay attention to this detail. We first take note of the following consequence of the computation in Proposition 6.24 for $\mathbf{D}_u$ in the case where u is a trivial cylinder:

PROPOSITION 7.5. *The Cauchy-Riemann type operator $\mathbf{D}_u$ on $u^*T\widehat{W}$ is asymptotic at its punctures $z_j^\pm \in \Gamma^\pm$ for $j = 1, \dots, k_\pm$ to the asymptotic operators $(-i\partial_t) \oplus \mathbf{A}_{\gamma_j^\pm}$ on $(\gamma_j^\pm)^*(\varepsilon_\pm \oplus \xi_\pm)$.* □

Perhaps you can now see a problem: even if the orbits $\gamma_j^\pm$ are all nondegenerate, the asymptotic operators $(-i\partial_t) \oplus \mathbf{A}_{\gamma_j^\pm}$ are degenerate, as they have nontrivial kernel consisting of constant sections in the first (trivial) factor of $(\gamma_j^\pm)^*(\varepsilon_\pm \oplus \xi_\pm)$. This implies in particular that

$$\mathbf{D}_0 : W^{k,p}(u^*T\widehat{W}) \to W^{k-1,p}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u^*T\widehat{W}))$$

is *not* Fredholm, except of course in the special case where there are no punctures.

The situation is saved by the exponential weight:

LEMMA 7.6. *For every $\delta > 0$ sufficiently small, the operator $\mathbf{D}_\delta$ is Fredholm and has index*

$$\operatorname{ind}(\mathbf{D}_\delta) = n\chi(\Sigma) - (n+1)\#\Gamma + 2c_1^\tau(u^*T\widehat{W}) + \sum_{j=1}^{k_+} \mu_{CZ}^\tau(\gamma_j^+) - \sum_{j=1}^{k_-} \mu_{CZ}^\tau(\gamma_j^-).$$

Moreover, every element of $\mathcal{M}(J)$ can be represented by a map $u \in \mathcal{B}^{k,p,\delta}$.

PROOF. The second claim follows from Proposition 6.32 in the previous lecture.

To see that $\mathbf{D}_\delta : W^{k,p,\delta} \to W^{k-1,p,\delta}$ is Fredholm and to compute its index, we can identify it with a Cauchy-Riemann type operator from $W^{k,p}$ to $W^{k-1,p}$. Indeed, pick any smooth function $f : \dot{\Sigma} \to \mathbb{R}$ with $f(s,t) = \mp\delta s$ on the cylindrical ends near $\Gamma^\pm$, define Banach space isomorphisms

$$\Phi_\delta : W^{k,p} \to W^{k,p,\delta} : \eta \mapsto e^f \eta,$$
$$\Psi_\delta : W^{k-1,p} \to W^{k-1,p,\delta} : \theta \mapsto e^f \theta,$$

and consider the bounded linear map

$$\mathbf{D}'_\delta := \Psi_\delta^{-1} \mathbf{D}_\delta \Phi_\delta : W^{k,p}(u^*T\widehat{W}) \to W^{k-1,p}(\overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u^*T\widehat{W})).$$

Using the Leibniz rule for $\mathbf{D}_\delta$, it is straightforward to show that $\mathbf{D}'_\delta$ is also a linear Cauchy-Riemann type operator. Moreover, suppose $\mathbf{D}_\delta$ takes the form $\bar{\partial} + S(s,t)$ in coordinates and trivialization on the cylindrical end near $z_j^\pm$, where $S(s,t) \to S_\infty(t)$ as $s \to \pm\infty$ and $\mathbf{A}_{\gamma_j^\pm} = -i\partial_t - S_\infty(t)$. Then $\mathbf{D}'_\delta$ on this same end takes the form

$$\mathbf{D}'_\delta \eta = e^{\pm\delta s}(\bar{\partial} + S(s,t))(e^{\mp\delta s}\eta) = \bar{\partial}\eta + (S(s,t) \mp \delta)\eta$$

and is therefore asymptotic to the perturbed asymptotic operator

$$\widetilde{\mathbf{A}}_j^\pm := \Big((-i\partial_t) \oplus \mathbf{A}_{\gamma_j^\pm} \Big) \pm \delta.$$

The latter is the direct sum of two asymptotic operators $-i\partial_t \pm \delta$ on the trivial line bundle and $\mathbf{A}_{\gamma_j^\pm} \pm \delta$ on $(\gamma_j^\pm)^*\xi_\pm$ respectively. Since $\gamma_j^\pm$ is nondegenerate by assumption and the spectrum of $\mathbf{A}_{\gamma_j^\pm}$ is discrete, we can assume $\ker(\mathbf{A}_{\gamma_j^\pm} \pm \delta)$ remains trivial if $\delta > 0$ is sufficiently small, and the Conley-Zehnder index of this perturbed operator will be the same as without the perturbation. On the other hand, the spectrum of $-i\partial_t$ consists of the integer multiples of 2π, thus $-i\partial_t \pm \delta$ also becomes nondegenerate for any $\delta > 0$ small. Its Conley-Zehnder index can be deduced from the winding numbers of its eigenfunctions using Theorem 3.55 in Lecture 3: $-i\partial_t$ has a 2-dimensional nullspace consisting of sections with winding number 0, and this becomes an eigenspace for the smallest positive eigenvalue if the puncture is positive or the largest negative eigenvalue if the puncture is negative. Theorem 3.55 thus gives

$$\mu_{\mathrm{CZ}}(-i\partial_t \pm \delta) = \mp 1,$$

and therefore,

$$\mu^\tau_{\mathrm{CZ}}(\widetilde{\mathbf{A}}_j^\pm) = \mp 1 + \mu^\tau_{\mathrm{CZ}}(\gamma_j^\pm).$$

Plugging this into the general index formula from Lecture 5 then gives the stated result. □

Putting back the missing $2(\#\Gamma)$ dimensions in the domain of $\mathbf{D}_u$, we have:

COROLLARY 7.7. *For all $\delta > 0$ sufficiently small, the linearized Cauchy-Riemann operator $\mathbf{D}_u : T_u\mathcal{B}^{k,p,\delta} \to \mathcal{E}_u^{k-1,p,\delta}$ is Fredholm with index*

$$\mathrm{ind}(\mathbf{D}_u) = n\chi(\Sigma) - (n-1)\#\Gamma + 2c_1^\tau(u^*T\widehat{W}) + \sum_{j=1}^{k_+} \mu^\tau_{CZ}(\gamma_j^+) - \sum_{j=1}^{k_-} \mu^\tau_{CZ}(\gamma_j^-).$$

7.3. Moduli of complex structures

Since the moduli space $\mathcal{M}(J)$ is not defined with reference to any fixed complex structure on the domains $\dot{\Sigma}$, we must build this freedom into the setup.

7.3.1. Teichmüller space and automorphism groups. For any integers $g, \ell \geqslant 0$, the **moduli space of Riemann surfaces** of genus g with ℓ marked points is a space of equivalence classes

$$\mathcal{M}_{g,\ell} = \{(\Sigma, j, \Theta)\} \big/ \sim$$

where (Σ, j) is a closed, oriented and connected surface with genus g, $\Theta \subset \Sigma$ is an ordered set of ℓ points and equivalence is defined via biholomorphic maps that preserve the marked points Θ along with their ordering. We shall refer to triples (Σ, j, Θ) that represent equivalence classes in this space as **pointed Riemann surfaces**. The space $\mathcal{M}_{g,\ell}$ has been studied extensively in algebraic geometry, though it can also be understood using the same global analytic methods that we have been applying for $\mathcal{M}(J)$. It is known in particular that $\mathcal{M}_{g,\ell}$ is always a smooth orbifold, and for any $[(\Sigma, j, \Theta)] \in \mathcal{M}_{g,\ell}$, it satisfies

$$\dim \operatorname{Aut}(\Sigma, j, \Theta) - \dim \mathcal{M}_{g,\ell} = 3\chi(\Sigma) - 2\ell, \tag{7.3}$$

where $\operatorname{Aut}(\Sigma, j, \Theta)$ is the group of biholomorphic transformations of (Σ, j) that fix the points in Θ, which is a smooth finite-dimensional Lie group. We will sketch proofs of these facts below. Observe that there are only four special cases in which $\chi(\Sigma \backslash \Theta)$ is nonnegative, and for these, we will see that it is not difficult to give explicit descriptions of both $\operatorname{Aut}(\Sigma, j, \Theta)$ and $\mathcal{M}_{g,\ell}$. All other cases satisfy the following condition:

DEFINITION 7.8. A pointed Riemann surface (Σ, j, Θ) is **stable** if $\chi(\Sigma \backslash \Theta) < 0$.

In the stable cases, Proposition 7.9 below shows that $\operatorname{Aut}(\Sigma, j, \Theta)$ is always a finite group, thus (7.3) turns into the well-known dimension formula

$$\dim \mathcal{M}_{g,\ell} = -3\chi(\Sigma) + 2\ell = 6g - 6 + 2\ell. \tag{7.4}$$

This is also the dimension of the **Teichmüller space**

$$\mathcal{T}(\Sigma, \Theta) := \mathcal{J}(\Sigma) \big/ \operatorname{Diff}_0(\Sigma, \Theta),$$

where $\mathcal{J}(\Sigma)$ denotes the space of all smooth complex structures on Σ compatible with its orientation, and

$$\operatorname{Diff}_0(\Sigma, \Theta) \subset \operatorname{Diff}(\Sigma, \Theta)$$

is the identity component of the group $\operatorname{Diff}(\Sigma, \Theta)$ of orientation-preserving diffeomorphisms that fix Θ, acting on $\mathcal{J}(\Sigma)$ by $\varphi \cdot j := \varphi^* j$.[2] It is a classical result that $\mathcal{T}(\Sigma, \Theta)$ is always a smooth manifold of the same dimension as $\mathcal{M}_{g,\ell}$, and indeed, the latter can be presented (see Exercise 7.11 below) as the quotient of the former by the discrete action of the **mapping class group**

$$M(\Sigma, \Theta) := \operatorname{Diff}(\Sigma, \Theta) \big/ \operatorname{Diff}_0(\Sigma, \Theta).$$

[2]Strictly speaking, acting by pullbacks defines a *right* group action $\mathcal{J}(\Sigma) \times \operatorname{Diff}(\Sigma, \theta) \to \mathcal{J}(\Sigma) : (j, \varphi) \mapsto \varphi^* j$, and if I wanted to turn it into a left group action, I should define the action of φ on j to be $(\varphi^{-1})^* j$. I hope you will forgive me for consistently ignoring this technicality.

We will see below that $\mathcal{T}(\Sigma, \Theta)$ admits easy explicit descriptions in the four non-stable cases, whereas in the stable cases, the fact that $\mathcal{T}(\Sigma, \Theta)$ is a manifold and $\mathcal{M}_{g,\ell}$ is an orbifold can be derived from the following:

PROPOSITION 7.9. *If* $\chi(\Sigma \backslash \Theta) < 0$*, then the action of* $\mathrm{Diff}_0(\Sigma, \Theta)$ *on* $\mathcal{J}(\Sigma)$ *is free and proper. Moreover, the action of* $\mathrm{Diff}(\Sigma, \Theta)$ *on* $\mathcal{J}(\Sigma)$ *is also proper, and its stabilizer subgroup* $\mathrm{Aut}(\Sigma, j, \Theta)$ *for each* $j \in \mathcal{J}(\Sigma)$ *is finite.*

SKETCH OF THE PROOF. If $\varphi \in \mathrm{Diff}_0(\Sigma, \Theta)$ satisfies $\varphi^* j = j$ for some $j \in \mathcal{J}(\Sigma)$, then $\varphi : (\Sigma, j) \to (\Sigma, j)$ is holomorphic, thus if $\varphi \neq \mathrm{Id}$, its fixed points are isolated and count positively for the purposes of the Lefschetz fixed point theorem, which says

$$\# \mathrm{Fix}(\varphi) = \chi(\Sigma)$$

since φ is homotopic to the identity. But φ has at least $\ell = \#\Theta$ fixed points, so this implies $\chi(\Sigma) - \ell = \chi(\Sigma \backslash \Theta) \geqslant 0$ and thus contradicts the stability assumption. This proves that the $\mathrm{Diff}_0(\Sigma, \Theta)$-action on $\mathcal{J}(\Sigma)$ is free, and also that the group $\mathrm{Aut}(\Sigma, j, \Theta)$ is discrete for each $j \in \mathcal{J}(\Sigma)$, as $\mathrm{Aut}(\Sigma, j, \Theta) \cap \mathrm{Diff}_0(\Sigma, \Theta) = \{\mathrm{Id}\}$.

The proof of properness requires some ideas from the compactness theory of holomorphic curves, which we have not discussed yet but will do so in Lecture 9, so for now we will give only an outline of the argument. (The missing details are filled in by Exercise 9.21.) Properness of the action of $\mathrm{Diff}_0(\Sigma, \Theta)$ on $\mathcal{J}(\Sigma)$ means that the map

$$\mathrm{Diff}_0(\Sigma, \Theta) \times \mathcal{J}(\Sigma) \to \mathcal{J}(\Sigma) \times \mathcal{J}(\Sigma) : (\varphi, j) \mapsto (\varphi^* j, j)$$

is proper, thus we need to show that if $\varphi_\nu \in \mathrm{Diff}_0(\Sigma, \Theta)$ and $j_\nu \in \mathcal{J}(\Sigma)$ are sequences with C^∞-convergence $j_\nu \to j$ and $j'_\nu := \varphi_\nu^* j_\nu \to j'$, then φ_ν also has a C^∞-convergent subsequence. The maps $\varphi_\nu : (\Sigma, j'_\nu) \to (\Sigma, j_\nu)$ in this situation are holomorphic curves of degree 1 with respect to converging sequences of domain and target complex structures, and they satisfy the uniform "energy" bound

$$E(\varphi_\nu) := \int_\Sigma \varphi_\nu^* d\,\mathrm{vol} = \int_{\varphi_\nu(\Sigma)} d\,\mathrm{vol} = \int_\Sigma d\,\mathrm{vol} = \mathrm{Vol}(\Sigma)$$

for any choice of area form $d\,\mathrm{vol}$ on Σ, which can be interpreted as a symplectic form taming the complex structures j_ν and j'_ν. By elliptic regularity (Corollary 2.25), φ_ν will have a C^∞-convergent subsequence if it satisfies a uniform C^1-bound, is this implies uniform $W^{1,p}$-bounds for $p > 2$. If no such bound holds, then one can reparametrize φ_ν around a sequence of points where its first derivative is blowing up and find a nonconstant holomorphic sphere $(S^2, i) \to (\Sigma, j)$ "bubbling off" in the limit. This is impossible if Σ has positive genus, since the map $S^2 \to \Sigma$ would have to have degree at least 1, but $\pi_2(\Sigma) = 0$. If Σ is also a sphere, then there is a different contradiction when $\#\Theta \geqslant 3$, because the fact that $\deg(\varphi_\nu) = 1$ implies that the bubble must absorb all available energy—it follows that there can be at most one bubble, which is found by reparametrizing the sequence around some specific limit point $\zeta_\infty \in S^2$, and on $S^2 \backslash \{\zeta_\infty\}$ the sequence φ_ν converges in C^∞_{loc} to a constant map. The latter is impossible since each φ_ν fixes at least two distinct marked points in $S^2 \backslash \{\zeta_\infty\}$.

We observe finally that the compactness argument in the previous paragraph works equally well if $\mathrm{Diff}_0(\Sigma, \Theta)$ is replaced by $\mathrm{Diff}(\Sigma, \Theta)$, and one can then apply it to any sequence in $\mathrm{Aut}(\Sigma, j, \Theta)$, proving that this group is compact. Since it is also discrete, $\mathrm{Aut}(\Sigma, j, \Theta)$ is therefore finite. $\square$

REMARK 7.10. The case $g = 1$ with no marked points is not stable, but the compactness part of the proof above still works in this case, proving that $\mathrm{Aut}(\mathbb{T}^2, j)$ is compact for every $j \in \mathcal{J}(\mathbb{T}^2)$.

EXERCISE 7.11. Show that the action of $\mathrm{Diff}(\Sigma, \Theta)$ on $\mathcal{J}(\Sigma)$ descends to a proper action of the mapping class group $M(\Sigma, \Theta)$ on $\mathcal{T}(\Sigma, \Theta)$ whose stabilizer subgroup at each $[j] \in \mathcal{T}(\Sigma, \Theta)$ is

$$\mathrm{Aut}(\Sigma, j, \Theta)/\,\mathrm{Aut}_0(\Sigma, j, \Theta), \qquad \text{where} \qquad \mathrm{Aut}_0(\Sigma, j, \Theta) := \mathrm{Aut}(\Sigma, j, \Theta) \cap \mathrm{Diff}_0(\Sigma, \Theta).$$

Show also that this stabilizer is always finite, and $\mathcal{T}(\Sigma, \Theta)/M(\Sigma, \Theta)$ is naturally homeomorphic to $\mathcal{M}_{g,\ell}$. *Hint: There is nothing to prove if $g = 0$ and $\ell \leqslant 2$, as standard results on mapping class groups imply $M(\Sigma, \Theta)$ is trivial in these cases.*

Another way to see why $\mathcal{T}(\Sigma, \Theta)$ is a smooth manifold is by interpreting (7.3) as a formula for a Fredholm index. Indeed, consider first the case $\ell = 0$. The right hand side is then $\chi(\Sigma) + 2c_1(T\Sigma)$, which is, according to Riemann-Roch, the index of the operator

$$\mathbf{D}_{\mathrm{Id}} : W^{k,p}(T\Sigma) \to W^{k-1,p}(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma)),$$

defined as the linearization at the identity map of the nonlinear Cauchy-Riemann operator for holomorphic maps $(\Sigma, j) \to (\Sigma, j)$. This is also known as the **canonical Cauchy-Riemann operator** for (Σ, j), i.e. it is the one that looks like the standard $\bar{\partial}$ in any choice of holomorphic local trivialization of $T\Sigma$, so its kernel is the space of holomorphic vector fields. Viewing it as a linearization, we see that there is a natural inclusion $T_{\mathrm{Id}}\,\mathrm{Aut}(\Sigma, j) \hookrightarrow \ker \mathbf{D}_{\mathrm{Id}}$, and it is not hard to show that this inclusion is in fact an isomorphism. For this it suffices to show that $\mathrm{Aut}(\Sigma, j)$ and $\ker \mathbf{D}_{\mathrm{Id}}$ have the same dimension, and in fact both can be computed explicitly in all cases. For $g = 0$ and $g = 1$, $\mathrm{Aut}(\Sigma, j)$ has real dimension 6 or 2 respectively, as follows from explicit descriptions of these groups which we shall review below. In the genus 0 case, we have $c_1(T\Sigma) = 2$ and $\mathrm{ind}\,\mathbf{D}_{\mathrm{Id}} = 6$, so $\dim \ker \mathbf{D}_{\mathrm{Id}} \geqslant 6$, but the reverse inequality also holds because picking any three distinct points $\zeta_1, \zeta_2, \zeta_3 \in \Sigma$ gives a linear map

$$\begin{aligned} \ker \mathbf{D}_{\mathrm{Id}} &\to T_{\zeta_1}\Sigma \oplus T_{\zeta_2}\Sigma \oplus T_{\zeta_3}\Sigma \\ X &\mapsto (X(\zeta_1), X(\zeta_2), X(\zeta_3)) \end{aligned}$$

that must be injective since all zeroes of the holomorphic vector field $X \in \ker \mathbf{D}_{\mathrm{Id}}$ count positively.[3] This proves $\dim \ker \mathbf{D}_{\mathrm{Id}} = 6 = \mathrm{ind}\,\mathbf{D}_{\mathrm{Id}}$, hence $\mathbf{D}_{\mathrm{Id}}$ is surjective and $T_{\mathrm{Id}}\,\mathrm{Aut}(S^2, j) = \ker \mathbf{D}_{\mathrm{Id}}$. In the genus one case, $\mathbf{D}_{\mathrm{Id}}$ has index 0 and is not surjective, but a similar argument using $c_1(T\mathbb{T}^2) = 0$ shows that $\dim \ker \mathbf{D}_{\mathrm{Id}} \leqslant 2$, which is therefore an equality since $\dim \ker \mathbf{D}_{\mathrm{Id}} \geqslant \dim \mathrm{Aut}(\mathbb{T}^2, j) = 2$. For genus at least 2, the first Chern number becomes negative and the positivity of zeroes of

[3]In keeping with Remark 5.6, it seems important to clarify here that although $\mathbf{D}_{\mathrm{Id}}$ is naturally a complex-linear operator, all dimensions and Fredholm indices in this discussion are real—which is why they are all even.

holomorphic vector fields thus implies that $\mathbf{D}_{\mathrm{Id}}$ is injective, hence $\mathrm{Aut}(\Sigma, j)$ is also 0-dimensional, meaning it is discrete—this gives an alternative proof of the discreteness of automorphism groups in the stable case, without need of the Lefschetz fixed point theorem (cf. Proposition 7.9).

The cokernel of $\mathbf{D}_{\mathrm{Id}}$ likewise has a natural identification with $T_{[j]}\mathcal{T}(\Sigma)$, resulting from the observation that the image of $\mathbf{D}_{\mathrm{Id}}$ is essentially the tangent space to the orbit of j under the natural $\mathrm{Diff}_0(\Sigma)$-action on $\mathcal{J}(\Sigma)$. To see this, suppose $\varphi_\rho \in \mathrm{Diff}_0(\Sigma)$ is a smooth 1-parameter family of diffeomorphisms with $\varphi_0 = \mathrm{Id}$, and write $X := \partial_\rho \varphi_\rho|_{\rho=0} \in \Gamma(T\Sigma)$. Choose an area form $d\,\mathrm{vol}$ on Σ and let ∇ denote the Levi-Cività connection with respect to the Riemannian metric $g := d\,\mathrm{vol}(\cdot, j\cdot)$; observe that since ∇ respects the metric, it also respects the conformal structure of (Σ, j) and therefore satisfies $\nabla j \equiv 0$. It follows then from (2.1) that $\mathbf{D}_{\mathrm{Id}} : \Gamma(T\Sigma) \to \Omega^{0,1}(\Sigma, T\Sigma)$ is given by

$$\mathbf{D}_{\mathrm{Id}} X = \nabla X + j \circ \nabla X \circ j,$$

and a quick computation using the symmetry of ∇ then yields the formula

$$\begin{aligned}\partial_\rho \left(\varphi_\rho^* j\right)\big|_{\rho=0} &= \partial_\rho \left(T\varphi_\rho^{-1} \circ j \circ T\varphi_\rho\right)\big|_{\rho=0} = -\nabla X \circ j + j \circ \nabla X \\ &= j(\nabla X + j \circ \nabla X \circ j) = j\mathbf{D}_{\mathrm{Id}} X = \mathbf{D}_{\mathrm{Id}}(jX) \in \Gamma(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma)).\end{aligned}$$

We should note that this way of identifying $\mathrm{im}\,\mathbf{D}_{\mathrm{Id}}$ with a $\mathrm{Diff}_0(\Sigma)$-orbit ignores a few technical details that deserve some care: strictly speaking, if we are viewing $\mathbf{D}_{\mathrm{Id}}$ as an operator $W^{k,p}(T\Sigma) \to W^{k-1,p}(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma))$, then $\mathrm{Diff}_0(\Sigma)$ should be replaced by a neighborhood of Id in the Banach manifold of $W^{k,p}$-smooth maps $\Sigma \to \Sigma$, with k and p chosen large enough to be able to say that these maps are at least C^1-diffeomorphisms. The resulting orbit of j then lives not in $\mathcal{J}(\Sigma)$ but in the space $\mathcal{J}^{k-1,p}(\Sigma)$ of almost complex structures of class $W^{k-1,p}$. (For a more careful treatment of Teichmüller theory from this perspective, see [Tro92].) Just as the choice of k and p does not affect the index or kernel of $\mathbf{D}_{\mathrm{Id}}$, these analytical details have no meaningful impact on the conclusion of the discussion, which is a characterization of $\mathrm{coker}\,\mathbf{D}_{\mathrm{Id}}$.

This whole discussion remains valid if marked points are included: the main difference is then that the Cauchy-Riemann operator on $T\Sigma$ should be restricted to a space of vector fields that vanish at Θ, defining a 2ℓ-codimensional subspace

$$W^{k,p}_\Theta(T\Sigma) := \left\{ X \in W^{k,p}(T\Sigma) \;\middle|\; X|_\Theta = 0 \right\},$$

which is the tangent space at Id to the Banach manifold of $W^{k,p}$-smooth maps $\Sigma \to \Sigma$ that fix the marked points. Restricting the domain in this way decreases the Fredholm index of $\mathbf{D}_{\mathrm{Id}}$ by 2ℓ, producing the right hand side of (7.3). The same arguments as sketched above then give natural isomorphisms

$$\ker \mathbf{D}_{\mathrm{Id}} = T_{\mathrm{Id}} \mathrm{Aut}(\Sigma, j, \Theta) =: \mathfrak{aut}(\Sigma, j, \Theta), \qquad \text{and} \qquad \mathrm{coker}\,\mathbf{D}_{\mathrm{Id}} = T_{[j]}\mathcal{T}(\Sigma, \Theta).$$

7.3.2. Teichmüller slices. The following notion will give a practical means of building variations in j into the analytic description of $\mathcal{M}(J)$.

DEFINITION 7.12. Given a pointed Riemann surface (Σ, j_0, Θ), a **Teichmüller slice through** (j_0, Θ) (or simply a "Teichmüller slice through j_0" if $\Theta = \varnothing$) is a smooth family of complex structures $\{j_\tau \in \mathcal{J}(\Sigma)\}_{\tau\in P}$ parametrized by some smooth finite-dimensional manifold P such that:

(1) $j_{\tau_0} = j_0$ for some $\tau_0 \in P$;
(2) The map $P \to \mathcal{T}(\Sigma,\Theta) : \tau \mapsto [j_\tau]$ is a diffeomorphism onto an open neighborhood of $[j_0]$;
(3) For every $\tau \in P$, the linear map

$$T_\tau P \to \Gamma(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma, j)) : \gamma'(0) \mapsto \partial_s j_{\gamma(s)}\big|_{s=0} \tag{7.5}$$

defined by choosing smooth paths $\gamma(s) \in P$ through $\gamma(0) = \tau$ sends $T_\tau P$ injectively onto to a complement of the image of the canonical Cauchy-Riemann operator $\mathbf{D}_{\mathrm{Id}} : W^{k,p}_\Theta(T\Sigma) \to W^{k-1,p}(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma, j))$ of (Σ, j).

We shall typically identify a Teichmüller slice with its image

$$\mathcal{T} := \bigcup_{\tau\in P}\{j_\tau\} \subset \mathcal{J}(\Sigma),$$

which we regard as a smoothly embedded finite-dimensional submanifold of $\mathcal{J}(\Sigma)$ that contains j_0 and has tangent spaces

$$T_j\mathcal{T} \subset \Gamma(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma, j))$$

given by the image of the map (7.5); symbolically, the third condition then says

$$\mathbf{D}_{\mathrm{Id}}(W^{k,p}_\Theta(T\Sigma)) \oplus T_j\mathcal{T} = W^{k-1,p}(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma, j)) \quad \text{for each } j \in \mathcal{T}.$$

EXERCISE 7.13. Verify that the image of the map (7.5) is automatically a section of $\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma, j)$, not just $\mathrm{End}_{\mathbb{R}}(T\Sigma)$.

EXERCISE 7.14. Show that the definition of a Teichmüller slice does not depend on the choices of Sobolev parameters $k \in \mathbb{N}$ and $p \in (1,\infty)$. *Hint: Elliptic regularity.*

The important consequence of the identification $T_{[j_0]}\mathcal{T}(\Sigma,\Theta) = \operatorname{coker}\mathbf{D}_{\mathrm{Id}}$ is that in practice, Teichmüller slices are relatively easy to construct:

PROPOSITION 7.15. *Suppose P is a finite-dimensional manifold and $\{j_\tau \in \mathcal{J}(\Sigma)\}_{\tau\in P}$ is a smooth family of complex structures such that $j_{\tau_0} = j_0$ for some $\tau_0 \in \Theta$ and the third condition in Definition 7.12 is satisfied for $\tau = \tau_0$. Then $\{j_\tau\}_{\tau\in P}$ becomes a Teichmüller slice after replacing P with any sufficiently small open neighborhood of τ_0 in P.* □

It is easy to see that Teichmüller slices always exist, e.g. first choose $T_{j_0}\mathcal{T} \subset \Gamma(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma))$ to be any complement of $\operatorname{im}\mathbf{D}_{\mathrm{Id}}$, let $\mathcal{O} \subset T_{j_0}\mathcal{T}$ be a suitably small neighborhood of 0 in this vector space, and define a family $\{j_y \in \mathcal{J}(\Sigma)\}_{y\in\mathcal{O}}$ by

$$j_y = \left(\mathbb{1} + \frac{1}{2}j_0 y\right) j_0 \left(\mathbb{1} + \frac{1}{2}j_0 y\right)^{-1}, \tag{7.6}$$

where the inverse on the right is defined as long as $y \in \mathcal{O}$ is sufficiently small.

We will often need Teichmüller slices that satisfy a few additional properties, and for this purpose, it is useful to consider the stable and non-stable cases separately. Non-stable pointed Riemann surfaces include the spheres with at most two marked points and the torus with no marked points. For spheres (S^2, j, Θ), the uniformization theorem (see Theorem 9.22) implies that (S^2, j) is always biholomorphically equivalent to the standard Riemann sphere (S^2, i), i.e. the extended complex plane $\mathbb{C} \cup \{\infty\}$. Moreover, if there are no more than three marked points, then we can choose biholomorphic maps to assign each of them wherever we like: the standard convention is to make Θ a subset of the points ∞, 0 and 1 in that order. It follows that $\mathcal{M}_{0,\ell}$ is a one-point space for $\ell \leqslant 2$, and since the mapping class group is trivial in these cases, so is $\mathcal{T}(S^2, \Theta)$ for $\#\Theta \leqslant 2$ (cf. Exercise 7.11). Precise descriptions of the automorphism groups in these cases are well known: $\mathrm{Aut}(S^2, i)$ is the 6-dimensional group of fractional linear transformations

$$\mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\} : z \mapsto \frac{az+b}{cz+d}, \qquad a, b, c, d \in \mathbb{C} \text{ with } ad - bc = 1,$$

often denoted by $\mathrm{PSL}(2, \mathbb{C}) := \mathrm{SL}(2, \mathbb{C})/\{\pm\mathbb{1}\}$, while $\mathrm{Aut}(S^2, i, (\infty)) = \mathrm{Aut}(\mathbb{C}, i)$ is the 4-dimensional group of affine transformations on $\mathbb{C}$, and $\mathrm{Aut}(S^2, i, (0, \infty)) = \mathrm{GL}(1, \mathbb{C}) = \mathbb{C}^*$. Equation (7.3) now translates into the observation that when $\#\Theta \leqslant 2$, $\mathcal{T}(S^2, \Theta)$ is discrete and $\mathrm{Aut}(S^2, i, \Theta) = \mathrm{Aut}(S^2 \backslash \Theta, i)$ is a Lie group of dimension $6 - 2(\#\Theta)$.

REMARK 7.16. The main points in the previous paragraph can also be deduced from holomorphic curve theory without any knowledge of the uniformization theorem or the mapping class group. In particular, the fact that $\mathcal{T}(S^2, \Theta)$ is discrete for $\#\Theta \leqslant 3$ follows already from the observation that $\mathbf{D}_{\mathrm{Id}} : W^{k,p}_{\Theta}(TS^2) \to W^{k-1,p}(\overline{\mathrm{End}}_{\mathbb{C}}(TS^2))$ in these cases is surjective. We proved this previously for the case $\Theta = \varnothing$, and the other cases follow similarly: our explicit descriptions of $\mathrm{Aut}(S^2, i, \Theta)$ give easy lower bounds for $\dim \ker \mathbf{D}_{\mathrm{Id}}$, and these match the upper bounds one gets by counting zeroes of holomorphic vector fields. Feeding this knowledge into the implicit function theorem and applying a bit of compactness theory as in Proposition 7.9, one can even turn this into a proof (without need of uniformization) that $\mathcal{T}(S^2, \Theta)$ is a one-point space. Notice that all complex structures compatible with a given orientation on a surface are also compatible with a symplectic form, thus they form a contractible space. Given $j \in \mathcal{J}(S^2)$, it follows that there exists a smooth path $\{j_s \in \mathcal{J}(S^2)\}_{s \in [0,1]}$ with $j_0 = i$ and $j_1 = j$, and the idea is then to show that the parametric moduli space

$$\Big\{(s, \varphi) \;\Big|\; s \in [0, 1] \text{ and } \varphi : (S^2, i) \to (S^2, j_s) \text{ holomorphic} \\ \text{with } \deg(\varphi) = 1 \text{ and } \varphi(\zeta) = \zeta \text{ for } \zeta = 0, 1, \infty\Big\}$$

is a smooth, compact and connected 1-manifold with boundary for which the projection $(s, \varphi) \mapsto s$ is a diffeomorphism to $[0, 1]$. This gives rise to a smooth 1-parameter family of diffeomorphisms $\varphi_s \in \mathrm{Diff}_0(S^2, \Theta)$ such that $\varphi_0 = \mathrm{Id}$ and $\varphi_s^* j_s = i$ for every s, proving $[j] = [i] \in \mathcal{T}(S^2, \Theta)$ if $\Theta \subset \{0, 1, \infty\}$.

REMARK 7.17. No nontrivial transformation in $\mathrm{PSL}(2, \mathbb{C})$ fixes more than two distinct points on S^2, thus $\mathrm{Aut}(S^2, i, \Theta)$ is always a trivial group when $\#\Theta \geqslant 3$, and Exercise 7.11 then implies that $\mathcal{M}_{0,\ell}$ is a manifold. One can also see this more directly, as every pointed Riemann sphere (S^2, j, Θ) with $\ell \geqslant 3$ marked points is equivalent to a unique one of the form (S^2, i, Θ') where $\Theta' = (0, 1, \infty, \zeta_4, \ldots, \zeta_\ell)$. This gives a natural homeomorphism between $\mathcal{M}_{0,\ell}$ and the set of ordered $(\ell-3)$-tuples of distinct points in $S^2 \backslash \{0, 1, \infty\}$.

The torus with no marked points is more interesting, as it is the only case for which both the Teichmüller space and the automorphism group have positive dimensions (equal, according to (7.3)). Applying uniformization again, the universal cover of any genus one Riemann surface $(\mathbb{T}^2, j)$ is the standard complex plane $(\mathbb{C}, i)$, which means that $(\mathbb{T}^2, j)$ is always biholomorphic to the quotient of $(\mathbb{C}, i)$ by a lattice $\Lambda \subset \mathbb{C}$. We can assume without loss of generality that Λ is generated by $1 \in \mathbb{C}$ and some point λ in the upper half plane $\mathbb{H} \subset \mathbb{C}$. Acting on i with the real-linear transformation that sends $1 \mapsto 1$ and $\lambda \mapsto i$ then identifies $(\mathbb{C}/\Lambda, i)$ biholomorphically with $(\mathbb{T}^2, j_\lambda)$, where $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ and j_λ is a translation-invariant complex structure compatible with the orientation of $\mathbb{T}^2$. In fact, every such translation-invariant complex structure on $\mathbb{T}^2$ can be obtained in this way, thus we have now parametrized the set of all translation-invariant complex structures on $\mathbb{T}^2$ by the upper half plane.

LEMMA 7.18. *For $\lambda, \lambda' \in \mathbb{H}$, the translation-invariant complex structures j_λ and $j_{\lambda'}$ on $\mathbb{T}^2$ represent the same element in the Teichmüller space $\mathcal{T}(\mathbb{T}^2)$ if and only if $\lambda = \lambda'$.*

PROOF. If $[j_\lambda] = [j_{\lambda'}] \in \mathcal{T}(\mathbb{T}^2)$, then $j_\lambda = \varphi^* j_{\lambda'}$ for some $\varphi \in \mathrm{Diff}_0(\mathbb{T}^2)$, which can be lifted to a diffeomorphism of $\mathbb{C}$ that fixes the lattice $\mathbb{Z} + i\mathbb{Z}$. Transforming via the real-linear map that sends $1 \mapsto 1$ and $\lambda \mapsto i$, this gives rise to a map $\psi : \mathbb{C} \to \mathbb{C}$ that is biholomorphic with respect to the standard complex structure and satisfies $\psi(0) = 0$, $\psi(1) = 1$ and $\psi(\lambda) = \lambda'$. But the only biholomorphic map on $\mathbb{C}$ that fixes two points is the identity. □

Since every $(\mathbb{T}^2, j)$ is biholomorphically equivalent to at least one of the $(\mathbb{T}^2, j_\lambda)$ by uniformization, the lemma implies that the smooth family

$$\mathcal{T} := \left\{ j_\lambda \in \mathcal{J}(\mathbb{T}^2) \right\}_{\lambda \in \mathbb{H}} \tag{7.7}$$

globally parametrizes the Teichmüller space $\mathcal{T}(\mathbb{T}^2)$. It should now seem unsurprising that $\mathcal{T}$ is also a Teichmüller slice in the sense of Definition 7.12 through each j_λ. One can prove this by writing down the canonical Cauchy-Riemann operator for $(\mathbb{C}/\Lambda, i)$ as the standard $\bar{\partial}$ on a trivial complex line bundle, so for a natural choice of L^2-pairings, its formal adjoint is $-\partial$ on a similarly trivial line bundle. The kernel of the latter is the space of constant functions $\mathbb{C}/\Lambda \to \mathbb{C}$, which has a natural identification with the tangent space $T_{j_\lambda}\mathcal{T}$ for each $\lambda \in \mathbb{H}$.

One can form a similarly explicit picture of the groups $\mathrm{Aut}(\mathbb{T}^2, j_\lambda)$ for each $\lambda \in \mathbb{H}$. First, let $\zeta = [(0,0)] \in \mathbb{R}^2/\mathbb{Z}^2 = \mathbb{T}^2$, and consider the subgroup

$$\mathrm{Aut}(\mathbb{T}^2, j_\lambda, \zeta) = \left\{ \psi \in \mathrm{Aut}(\mathbb{T}^2, j_\lambda) \;\middle|\; \psi(\zeta) = \zeta \right\}.$$

We know this subgroup is finite since $(\mathbb{T}^2, j_\lambda, (\zeta))$ is stable, but we can be more precise. Identifying $(\mathbb{T}^2, j_\lambda)$ with $(\mathbb{C}/\Lambda, i)$, any $\psi \in \mathrm{Aut}(\mathbb{T}^2, j_\lambda)$ that fixes ζ lifts via the projection $\mathbb{C} \to \mathbb{C}/\Lambda$ to a biholomorphic map $\widetilde{\psi} : (\mathbb{C}, i) \to (\mathbb{C}, i)$ that fixes the origin. All such maps are of the form $\widetilde{\psi}(z) = cz$ for some $c \in \mathbb{C}^*$, implying that ψ is the projection to $\mathbb{T}^2 = \mathbb{C}/(\mathbb{Z}+i\mathbb{Z})$ of a real-linear map on $\mathbb{C}$ which preserves $\mathbb{Z}+i\mathbb{Z}$. In other words, $\psi \in \mathrm{SL}(2,\mathbb{Z})$, and this discussion implies

$$\mathrm{Aut}(\mathbb{T}^2, j_\lambda, \zeta) = \left\{ \psi \in \mathrm{SL}(2,\mathbb{Z}) \;\middle|\; \psi^* j_\lambda = j_\lambda \right\}.$$

It is not hard to show explicitly that this group is finite, though its order will depend in general on λ. For finiteness, note that it is discrete, and after a suitable change of basis transforming j_λ to the standard complex structure i, can be identified with a subgroup of $\mathrm{SL}(2,\mathbb{R}) \cap \mathrm{GL}(1,\mathbb{C}) = \mathrm{U}(1)$, which is compact. Now observe that since every j_λ is translation-invariant, $\mathrm{Aut}(\mathbb{T}^2, j_\lambda)$ also contains a complentary subgroup which is a copy of $\mathbb{T}^2$ acting on $\mathbb{T}^2$ by translations. Clearly any $\psi \in \mathrm{Aut}(\mathbb{T}^2, j_\lambda)$ can be composed with a translation to produce an element of $\mathrm{Aut}(\mathbb{T}^2, j_\lambda, \zeta)$, so this presents the full automorphism group as a semidirect product

$$\mathrm{Aut}(\mathbb{T}^2, j_\lambda) = \mathbb{T}^2 \rtimes \mathrm{Aut}(\mathbb{T}^2, j_\lambda, \zeta).$$

The main observation we would like to take away from this decription is the following: for every $\lambda \in \mathbb{H}$, the action of $\mathrm{Aut}(\mathbb{T}^2, j_\lambda)$ on $\mathcal{J}(\Sigma)$ preserves the specific Teichmüller slice defined in (7.7). That is a property we will want to have more generally:

PROPOSITION 7.19. *Given any pointed Riemann surface (Σ, j_0, Θ), there exists a Teichmüller slice $\mathcal{T}$ through (j_0, Θ) with the following properties:*

(1) $\mathcal{T}$ is preserved under the group action

$$\begin{aligned} \mathrm{Aut}(\Sigma, j_0, \Theta) \times \mathcal{J}(\Sigma) &\to \mathcal{J}(\Sigma) \\ (\psi, j) &\mapsto \psi^* j; \end{aligned}$$

(2) For every $j \in \mathcal{T}$, $\mathrm{Aut}(\Sigma, j, \Theta) \cap \mathrm{Diff}_0(\Sigma, \Theta) = \mathrm{Aut}(\Sigma, j_0, \Theta) \cap \mathrm{Diff}_0(\Sigma, \Theta)$;

(3) Every $j \in \mathcal{T}$ equals j_0 near Θ.

PROOF. There is nothing to prove if $\Sigma = S^2$ and $\#\Theta \leqslant 2$, since $\mathcal{T}(\Sigma, \Theta)$ is trivial in these cases. For the case $\Sigma = \mathbb{T}^2$ and $\Theta = \varnothing$, the third condition we want to satisfy is vacuous, and the first two are satisfied by the explicit Teichmüller slice defined in (7.7); in particular, $\mathrm{Aut}(\mathbb{T}^2, j_\lambda) \cap \mathrm{Diff}_0(\mathbb{T}^2)$ is the same group of translations for all $\lambda \in \mathbb{H}$. It therefore remains only to deal with cases where (Σ, j_0, Θ) is stable, so that

$$G := \mathrm{Aut}(\Sigma, j_0, \Theta)$$

is a finite group and $G \cap \mathrm{Diff}_0(\Sigma, \Theta)$ is trivial, making the second condition vacuous. To produce a slice $\mathcal{T}$ that is G-invariant, it suffices to find a G-invariant complement $T_{j_0}\mathcal{T}$ for $\operatorname{im} \mathbf{D}_{\mathrm{Id}}$ and then follow the prescription of (7.6). Note that $\operatorname{im} \mathbf{D}_{\mathrm{Id}}$ is itself G-invariant, so for this purpose, it is enough to choose a G-invariant L^2-pairing on the sections of $\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma)$, and define $T_{j_0}\mathcal{T}$ as the L^2-orthogonal complement of $\operatorname{im} \mathbf{D}_{\mathrm{Id}}$. Elliptic regularity for weak solutions implies that the elements of $T_{j_0}\mathcal{T}$ are then smooth. A suitable L^2-pairing can be defined naturally in terms of a G-invariant Riemannian metric on Σ, which is easy to construct by an averaging procedure since G is finite.

Having constructed $T_{j_0}\mathcal{T}$ to be G-invariant, we can now modify it as follows. Fix a small G-invariant neighborhood $\mathcal{U} \subset \Sigma$ of Θ and a smooth G-invariant cutoff function $\beta : \Sigma \to [0,1]$ such that $1 - \beta$ is supported in $\mathcal{U}$ and $\beta = 0$ near Θ. Taking $\mathcal{U}$ sufficiently small, we can assume that the transformation

$$y \mapsto \beta y$$

maps any bounded subset of the finite-dimensional space $T_{j_0}\mathcal{T}$ to something arbitrarily close in the L^p-norm, so that without loss of generality, the space $\{\beta y \mid y \in T_{j_0}\mathcal{T}\}$ is also a G-invariant complenent to the image of $\mathbf{D}_{\mathrm{Id}} : W^{1,p}_{\Theta}(T\Sigma) \to L^p(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma))$. Since βy always vanishes near Θ, applying (7.6) to this new subspace now produces a G-invariant Teichmüller slice consisting of complex structures that match j_0 near Θ. $\square$

REMARK 7.20. If (Σ, j_0, Θ) is stable, then the cutoff trick in the proof above extends easily to give an $\mathrm{Aut}(\Sigma, j_0, \Theta)$-invariant Teichmüller slice $\mathcal{T}$ such that each $j \in \mathcal{T}$ can be assumed to match j_0 near any finite set fixed in advance (not just Θ). If we are willing to allow Teichmüller slices that are not $\mathrm{Aut}(\Sigma, j_0, \Theta)$-invariant, then this remains true in the non-stable cases as well, and is sometimes useful, e.g. this freedom is exploited in the proof of automatic transversality for non-immersed J-holomorphic curves in 4-manifolds (see [**Wen10a**, Lemma 3.15]).

LEMMA 7.21. *For any Teichmüller slice $\mathcal{T}$ through (j_0, Θ) that is preserved by the action of $G := \mathrm{Aut}(\Sigma, j_0, \Theta)$, the subgroup $G_0 := \mathrm{Aut}(\Sigma, j_0, \Theta) \cap \mathrm{Diff}_0(\Sigma, \Theta)$ acts trivially on $\mathcal{T}$, hence the G-action descends to an action of $H := G/G_0$ on $\mathcal{T}$, and the natural map*

$$\mathcal{T}/H \to \mathcal{M}_{g,\ell}$$

induced by restricting the quotient projection $\mathcal{J}(\Sigma) \to \mathcal{J}(\Sigma)/\mathrm{Diff}(\Sigma, \Theta)$ to $\mathcal{T}$ is a local homeomorphism near $[j_0]$.

PROOF. The triviality of the action by G_0 on $\mathcal{T}$ is immediate from the fact that no two points in $\mathcal{T}$ represent the same element of Teichmüller space. Exercise 7.11 identifies $\mathcal{M}_{g,\ell}$ with $\mathcal{T}(\Sigma, \Theta)/M(\Sigma, \Theta)$, where the mapping class group $M(\Sigma, \Theta)$ acts on $\mathcal{T}(\Sigma, \Theta)$ with finite stabilizer group H at $[j_0] \in \mathcal{T}(\Sigma, \Theta)$. Since the mapping class group is discrete and H is finite, a neighborhood of $[(\Sigma, j_0, \Theta)]$ in $\mathcal{M}_{g,\ell}$ then has a natural identification with $\mathcal{U}/H$ for some H-invariant neighborhood $\mathcal{U} \subset \mathcal{T}(\Sigma, \Theta)$ of $[j_0]$. This neighborhood is identified with a neighborhood $\mathcal{U}' \subset \mathcal{T}$ of j_0 via the map $\mathcal{T} \to \mathcal{T}(\Sigma, \Theta) : j \mapsto [j]$, but the latter is also H-equivariant, so it descends to a homeomorphism $\mathcal{U}'/H \to \mathcal{U}/H$. $\square$

7.3.3. Adding marked points. The lemma below is not necessary for the proof of Theorem 7.1, but will be needed in §7.5 for results involving forgetful maps.

According to the dimension formula (7.4), each marked point added to a stable Riemann surface increases the dimension of Teichmüller space by 2. It will occasionally be useful to have a more precise description of how Teichmüller slices can be modified under the addition of marked points. Notice that for any nested pair $\Theta \subset \Theta' \subset \Sigma$ of finite ordered subsets, $\mathrm{Diff}_0(\Sigma, \Theta')$ is a subgroup of $\mathrm{Diff}_0(\Sigma, \Theta)$, giving rise to a canonical projection map between Teichmüller spaces

$$\mathcal{T}(\Sigma, \Theta') = \mathcal{J}(\Sigma)/\mathrm{Diff}_0(\Sigma, \Theta') \to \mathcal{J}(\Sigma)/\mathrm{Diff}_0(\Sigma, \Theta) = \mathcal{T}(\Sigma, \Theta).$$

The next lemma implies that this map is always a smooth submersion. For simplicity of notation, we shall consider only the case where Θ consists of the first ℓ elements in the ordered set Θ' for some $\ell \geqslant 0$.

LEMMA 7.22. *Suppose (Σ, j_0, Θ) is a pointed Riemann surface with $\ell \geqslant 0$ marked points $\Theta = (w_1, \dots, w_\ell)$, and given another ordered set of distinct points $(\zeta_1, \dots, \zeta_k)$ in $\Sigma \backslash \Theta$, let $\Theta' = (w_1, \dots, w_\ell, \zeta_1, \dots, \zeta_k)$.*

(1) If (Σ, j_0, Θ) is not stable and $k = 1$, then every Teichmüller slice through (j_0, Θ) is also a Teichmüller slice through (j_0, Θ') and vice versa.

(2) If (Σ, j_0, Θ) is stable, then for any neighborhood $\mathcal{V} \subset \Sigma \backslash \Theta$ of $\{\zeta_1, \dots, \zeta_k\}$, there exist Teichmüller slices $\mathcal{T}$ through (j_0, Θ) and $\mathcal{T}'$ through (j_0, Θ') that each satisfy the conditions of Proposition 7.19, where $\mathcal{T}'$ is of the form

$$\mathcal{T}' = \{\varphi_\tau^* j \in \mathcal{J}(\Sigma)\}_{(j,\tau) \in \mathcal{T} \times (\mathcal{U}_1 \times \dots \times \mathcal{U}_k)}$$

for some smooth family of diffeomorphisms $\{\varphi_\tau : \Sigma \to \Sigma\}_{\tau \in \mathcal{U}_1 \times \dots \times \mathcal{U}_k}$ parametrized by the product of a set of disjoint neighborhoods

$$\zeta_1 \in \mathcal{U}_1 \subset \mathcal{V} \subset \Sigma \backslash \Theta, \ \dots \ \zeta_k \in \mathcal{U}_k \subset \mathcal{V} \subset \Sigma \backslash \Theta,$$

such that each φ_τ is supported in $\mathcal{V}$, $\varphi_\tau = \mathrm{Id}$ for $\tau = (\zeta_1, \dots, \zeta_k)$, and

$$\varphi_\tau(\zeta_i) = \zeta_i' \text{ for each } i = 1, \dots, k \quad \text{if} \quad \tau = (\zeta_1', \dots, \zeta_k').$$

Notice that in the construction for the stable case, $\mathcal{T}'$ contains $\mathcal{T}$ as the subfamily $\{\varphi_\tau^* j \mid j \in \mathcal{T}, \ \tau = (\zeta_1, \dots, \zeta_k)\}$, and under the natural identifications of $\mathcal{T}$ and $\mathcal{T}'$ with subsets of $\mathcal{T}(\Sigma, \Theta)$ and $\mathcal{T}(\Sigma, \Theta')$ respectively, the canonical map $\mathcal{T}(\Sigma, \Theta') \to \mathcal{T}(\Sigma, \Theta)$ becomes identified with the natural projection map $\mathcal{T} \times \mathcal{U}_1 \times \dots \mathcal{U}_k \to \mathcal{T}$. The main point of the construction is that since

$$(\Sigma, \varphi_\tau^* j, (w_1, \dots, w_\ell, \zeta_1, \dots, \zeta_k)) \sim (\Sigma, j, (w_1, \dots, w_\ell, \varphi_\tau(\zeta_1), \dots, \varphi_\tau(\zeta_k))),$$

the extra $2k$ dimensions in $\mathcal{T}'$ compared with $\mathcal{T}$ now have an obvious geometric interpretation as the freedom to move the extra k marked points.

PROOF OF LEMMA 7.22. We handle the non-stable cases first: if $\Sigma \cong S^2$, $\ell \leqslant 2$ and $k = 1$, then there is nothing to prove because $\mathcal{T}(\Sigma, \Theta)$ and $\mathcal{T}(\Sigma, \Theta')$ are both trivial. The remaining non-stable case is $\Sigma \cong \mathbb{T}^2$ with $\ell = 0$ and $k = 1$, so $\mathcal{T}(\Sigma, \Theta) = \mathcal{T}(\mathbb{T}^2)$ and $\mathcal{T}(\Sigma, \Theta') = \mathcal{T}(\mathbb{T}^2, \zeta)$ are both 2-dimensional manifolds, and we can assume j_0 is one of the translation-invariant complex structures j_λ introduced above Lemma 7.18. The key observation here is that the image of the canonical Cauchy-Riemann operator

$$\mathbf{D}_{\mathrm{Id}} : W^{k,p}(T\mathbb{T}^2) \to W^{k-1,p}(\overline{\mathrm{End}}_{\mathbb{C}}(T\mathbb{T}^2))$$

does not change if we restrict its domain to $W^{k,p}_{\Theta'}(T\mathbb{T}^2)$, as $\ker \mathbf{D}_{\mathrm{Id}} = T_{\mathrm{Id}} \operatorname{Aut}(\mathbb{T}^2, j_\lambda)$ is precisely the space of constant vector fields, defining a closed complement of the subspace $W^{k,p}_{\Theta'}(T\mathbb{T}^2)$ of vector fields that vanish at ζ:

$$W^{k,p}(T\mathbb{T}^2) = W^{k,p}_{\Theta'}(T\mathbb{T}^2) \oplus \ker \mathbf{D}_{\mathrm{Id}}.$$

As a consequence, subspaces complementary to $\operatorname{im} \mathbf{D}_{\mathrm{Id}}$ are the same with repect to both domains, so that the defining property of a Teichmüller slice is the same for $(\mathbb{T}^2, j_\lambda)$ as it is for $(\mathbb{T}^2, j_\lambda, \zeta)$.

Now assume (Σ, j_0, Θ) is stable, so the automorphism groups

$$G := \operatorname{Aut}(\Sigma, j_0, \Theta) \quad \text{and} \quad G' := \operatorname{Aut}(\Sigma, j_0, \Theta')$$

are finite, with $G' \subset G$. By Prop. 7.19 and Remark 7.20, we can find a G-invariant Teichmüller slice $\mathcal{T}$ through (j_0, Θ) such that every $j \in \mathcal{T}$ matches j_0 on a neighborhood of Θ'. In particular, there are fixed disjoint neighborhoods $\mathcal{U}_i \subset \Sigma$ of the points ζ_i for $i = 1, \ldots, k$ on which all $j \in \mathcal{T}$ match j_0. Let us fix such a set of neighborhoods, which we can assume to be contained in any given neighborhood $\mathcal{V} \subset \Sigma \backslash \Theta$ of $\{\zeta_1, \ldots, \zeta_k\}$, and also fix holomorphic coordinates identifying each biholomorphically with $(\mathbb{D}, i)$ so that ζ_i is the origin. Since G' acts j_0-holomorphically and fixes each ζ_i, we can also assume that the G'-action on Σ preserves these coordinate neighborhoods and acts on each by rotations with respect to the coordinates. Now choose a smooth cutoff function $\beta : [0, 1] \to [0, 1]$ with $\beta(s) = 1$ near $s = 0$ and $\beta(s) = 0$ near $s = 1$, and for each $i = 1, \ldots, k$, define

$$V_i \subset \Gamma(T\Sigma)$$

to be the vector space of all smooth vector fields that are supported in $\mathcal{U}_i$ and take the form

$$X(z) = \beta(|z|) X_0$$

in the chosen coordinates on this neighborhood, where $X_0 \in \mathbb{C}$ is an arbitrary constant. The direct sum

$$V' = V_1 \oplus \ldots \oplus V_k$$

is then a $2k$-dimensional subspace of $\Gamma(T\Sigma)$ satisfying

$$W^{k,p}_{\Theta}(T\Sigma) = W^{k,p}_{\Theta'}(T\Sigma) \oplus V', \tag{7.8}$$

and since elements of G' act on the coordinate neighborhoods by rotations, the induced action of G' on $W^{k,p}_{\Theta}(T\Sigma)$ defined by pullbacks preserves this splitting. Now choose a G'-invariant neighborhood $\mathcal{O}' \subset V'$ of the origin and define a smooth family of diffeomorphisms

$$\{\varphi_X \in \operatorname{Diff}(\Sigma, \Theta)\}_{X \in \mathcal{O}'},$$

where for each $X \in \mathcal{O}'$, φ_X is the time-1 flow of X, which is supported in $\mathcal{U}_1 \cup \ldots \cup \mathcal{U}_k$ and therefore fixes Θ. Since each $X \in V'$ is holomorphic near the points ζ_i, we can assume after possibly shrinking $\mathcal{O}'$ that every φ_X is j_0-holomorphic on some smaller neighborhood of these points for $X \in \mathcal{O}'$. We can then define $\mathcal{T}'$ as the smoothly parametrized family

$$\mathcal{T}' = \{\varphi_X^* j \in \mathcal{J}(\Sigma)\}_{(j,X) \in \mathcal{T} \times \mathcal{O}'}.$$

Since every $j \in \mathcal{T}$ matches j_0 near Θ' and every φ_X has support in $\mathcal{U}_1 \cup \ldots \cup \mathcal{U}_k$ and is holomorphic on a smaller neighborhood of $\{\zeta_1, \ldots, \zeta_k\}$, every $j \in \mathcal{T}'$ has the desired property that it matches j_0 near Θ'. Moreover, $\mathcal{T}'$ is G'-invariant since the same is true of both $\mathcal{T}$ and the collection of diffeomorphisms $\{\varphi_X\}_{X \in \mathcal{O}'}$. To see that $\mathcal{T}'$ really is a Teichmüller slice through (j_0, Θ'), which correponds to the parameter

$(j_0, 0) \in \mathcal{T} \times \mathcal{O}'$, we observe that differentiating $\varphi_X^* j$ with respect to variations in $j \in \mathcal{T}$ at $(j, X) = (j_0, 0)$ produces $T_{j_0}\mathcal{T}$, while fixing $j = j_0$ and differentiating with respect to variations in $X \in \mathcal{O}'$ produces $\mathbf{D}_{\mathrm{Id}}(V')$, hence

$$T_{j_0}\mathcal{T}' = T_{j_0}\mathcal{T} \oplus \mathbf{D}_{\mathrm{Id}}(V').$$

This is indeed a complementary subspace to the image of $\mathbf{D}_{\mathrm{Id}} : W^{k,p}_{\Theta'}(T\Sigma) \to W^{k-1,p}(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma))$ in light of (7.8), since the image of $\mathbf{D}_{\mathrm{Id}}$ on the larger domain $W^{k,p}_{\Theta}(T\Sigma)$ is complementary to $T_{j_0}\mathcal{T}$.

Finally, we observe that by construction, the map

$$\mathcal{O}' \to \mathcal{U}_1 \times \ldots \times \mathcal{U}_k : X \mapsto (\varphi_X(\zeta_1), \ldots, \varphi_X(\zeta_k))$$

sends some neighborhood of 0 diffeomorphically to a neighborhood of $(\zeta_1, \ldots, \zeta_k)$, so after shrinking neighborhoods as necessary, we can reparametrize the family $\{\varphi_X\}_{X \in \mathcal{O}'}$ according to the images of the points $\zeta_1, \ldots, \zeta_k$ as in the statement of the lemma. $\square$

7.4. Fredholm regularity and the implicit function theorem

We are now in a position to define the necessary regularity condition and prove that a neighborhood of any given regular element $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)]$ in $\mathcal{M}(J)$ is an orbifold of the stated dimension. After reparametrizing, we can assume without loss of generality that Σ, $\Gamma^\pm$ and Θ are precisely the data that were fixed in §7.2, and $j_0 \in \mathcal{J}(\Sigma)$ matches j_Γ on our fixed coordinate neighborhoods of $\Gamma^\pm$. Combining the local regularity results of Lecture 2 with the exponential convergence result in §6.5, we also have

$$u_0 \in \mathcal{B}^{k,p,\delta}$$

for every $k \in \mathbb{N}$ and $p \in (1, \infty)$ with $kp > 2$ and $\delta > 0$ sufficiently small, so u_0 belongs to the zero set of the smooth section $\bar{\partial}_{j_0,J} : \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$. Let us fix the parameters k, p, δ, and also assume that δ is small enough for our computation of the index of $\mathbf{D}_u : T_u\mathcal{B}^{k,p,\delta} \to \mathcal{E}_u^{k-1,p,\delta}$ in Corollary 7.7 to be valid.

In order to build variations in j_0 into this setup, we choose a Teichmüller slice

$$\mathcal{T} \subset \mathcal{J}(\Sigma)$$

through $(j_0, \Gamma \cup \Theta)$ as provided by Prop. 7.19, so in particular, $\mathcal{T}$ is invariant under the action of the group

$$G := \mathrm{Aut}(\Sigma, j_0, \Gamma \cup \Theta)$$

of biholomorphic maps fixing $\Gamma \cup \Theta$, and every $j \in \mathcal{T}$ matches j_Γ near Γ. Equation (7.3) now becomes

$$\dim G - \dim \mathcal{T} = 3\chi(\Sigma) - 2(k_+ + k_- + m). \tag{7.9}$$

There is a natural extension of the nonlinear operator $\bar{\partial}_{j_0,J}$ in §7.2 to a smooth section

$$\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta} : (j, u) \mapsto du + J(u) \circ du \circ j$$

of a Banach space bundle $\mathcal{E}^{k-1,p,\delta} \to \mathcal{T} \times \mathcal{B}^{k,p,\delta}$ with fibers

$$\mathcal{E}^{k-1,p,\delta}_{(j,u)} = W^{k-1,p,\delta}\big(\overline{\mathrm{Hom}}_{\mathbb{C}}((T\dot{\Sigma}, j), (u^*T\widehat{W}, J))\big).$$

Since the action of G on $\mathcal{J}(\Sigma)$ preserves $\mathcal{T}$, each $\varphi \in G$ defines a diffeomorphism

$$\mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{T} \times \mathcal{B}^{k,p,\delta} : (j,u) \mapsto \varphi \cdot (j,u) := (\varphi^* j, u \circ \varphi)$$

which is covered by a bundle map $\mathcal{E}^{k-1,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$ given by

$$\mathcal{E}^{k-1,p,\delta}_{(j,u)} \to \mathcal{E}^{k-1,p,\delta}_{(\varphi^* j, u\circ\varphi)} : \lambda \mapsto \varphi \cdot \lambda := \lambda \circ d\varphi.$$

The section $\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$ is equivariant with respect to this G-action. Its zero set is the set of all pairs $(j,u) \in \mathcal{T} \times \mathcal{B}^{k,p,\delta}$ for which $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ is pseudoholomorphic, and it contains (j_0, u_0) by construction. It also is preserved by the action of G, thus each $\varphi \in G$ acts on $\bar{\partial}_J^{-1}(0)$ by

$$\bar{\partial}_J^{-1}(0) \to \bar{\partial}_J^{-1}(0) : (j,u) \mapsto (\varphi^* j, u \circ \varphi). \tag{7.10}$$

The stabilizer of this action at (j_0, u_0) is $\mathrm{Aut}(u_0)$, a finite group whenever u_0 is not constant (see Proposition 6.37). Observe that any two elements in the same G-orbit of $\bar{\partial}_J^{-1}(0)$ define equivalent elements of the moduli space $\mathcal{M}(J)$, as they are related to each other by a biholomorphic reparametrization that fixes the punctures and marked points.

PROPOSITION 7.23. *The map*

$$\bar{\partial}_J^{-1}(0)/G \to \mathcal{M}(J) : [(j,u)] \mapsto [(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)]$$

defines a homeomorphism from an open neighborhood of $[(j_0, u_0)]$ *to an open neighborhood of* $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)]$.

The tricky aspect of this statement is that arbitrary elements of $\mathcal{M}(J)$ near u_0 may be given with representatives $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ for which j is close to j_0 but does not belong to the chosen Teichmüller slice $\mathcal{T}$. We will therefore have to replace such curves by reparametrizations, making use of the following lemma:

LEMMA 7.24. *Suppose* $\varphi_\nu \in \mathrm{Diff}(\Sigma, \Gamma)$ *is a sequence with convergence* $\varphi_\nu \to \mathrm{Id}$ *in* $C^\infty(\Sigma)$, *and let* $\bar{\Sigma}$ *denote the compactification of* $\dot{\Sigma}$ *to a compact topological surface with boundary as described in §6.4. Then the restricted maps* $\dot{\Sigma} \xrightarrow{\varphi_\nu} \dot{\Sigma}$ *converge in* $C^\infty_{\mathrm{loc}}(\dot{\Sigma})$ *and have continuous extensions* $\bar{\Sigma} \xrightarrow{\bar{\varphi}_\nu} \bar{\Sigma}$ *that converge in* $C^0(\bar{\Sigma})$.

PROOF. The C^∞_{loc}-convergence on $\dot{\Sigma}$ is clear. For C^0-convergence on $\bar{\Sigma}$, it suffices to prove uniform convergence on each of the cylindrical ends. Let us choose local coordinates near one of the punctures, identifying a neighborhood with the unit disk $\mathbb{D} \subset \mathbb{C}$ such that the puncture becomes $0 \in \mathbb{D}$. Since $\varphi_\nu \in \mathrm{Diff}(\Sigma, \Gamma)$ and $\varphi_\nu \to \mathrm{Id}$, we can then assume $\varphi_\nu(\mathbb{D}_{1/2}) \subset \mathbb{D}$ for ν large and $\varphi_\nu(0) = 0$, hence

$$\varphi_\nu(z) = A_\nu(z) z \quad \text{for} \quad z \in \mathbb{D}_{1/2},$$

where $A_\nu : \mathbb{D}_{1/2} \to \mathrm{End}_{\mathbb{R}}(\mathbb{C})$ is a C^∞-convergent sequence with $A_\nu(0) = d\varphi_\nu(0)$ and $A_\nu \to \mathbb{1}$. Let us now rewrite this in the cylindrical coordinates

$$s + it \in \mathbb{C}/i\mathbb{Z} \cong \mathbb{R} \times S^1$$

related to z by $z = e^{-2\pi(s+it)}$: for $s > 0$ sufficiently large, φ_ν then takes the form

$$\begin{aligned}\varphi_\nu(s+it) = -\frac{1}{2\pi}\log\left(A_\nu(z)z\right) = -\frac{1}{2\pi}\log\left(e^{-2\pi s}A_\nu(z)\frac{z}{|z|}\right) \\ = s - \frac{1}{2\pi}\log\left(A_\nu(z)e^{-2\pi it}\right) \to s+it,\end{aligned}$$

where since $A_\nu e^{-2\pi it}$ stays a positive distance away from 0 for large ν, the convergence is uniform in s and t. □

PROOF OF PROPOSITION 7.23. There are four issues to address: that the map in question is both injective and surjective on sufficiently small neighborhoods, and also that the map and its inverse are both continuous.

For injectivity, let us abbreviate

$$\Theta' := \Gamma \cup \Theta,$$

and let $H \subset M(\Sigma, \Theta')$ denote the stabilizer of $[j_0] \in \mathcal{T}(\Sigma, \Theta')$ under the action of $M(\Sigma, \Theta')$ on $\mathcal{T}(\Sigma, \Theta')$. According to Exercise 7.11, the inclusion $G \hookrightarrow \mathrm{Diff}(\Sigma, \Theta')$ descends to an isomorphism $G/G_0 \to H$, where

$$G_0 := G \cap \mathrm{Diff}_0(\Sigma, \Theta').$$

Suppose (j, u), $(j', u') \in \bar{\partial}_J^{-1}(0)$ represent the same element of $\mathcal{M}(J)$, meaning there exists a diffeomorphism $\varphi \in \mathrm{Diff}(\Sigma, \Theta')$ with $j' = \varphi^* j$ and $u' = u \circ \varphi$. Assume both are sufficiently close to (j_0, u_0) so that $[j], [j'] \in \mathcal{T}(\Sigma, \theta)$ both lie in some H-invariant neighborhood $\mathcal{U} \subset \mathcal{T}(\Sigma, \Theta')$ of $[j_0]$ that is small enough to be disjoint from $[\psi] \cdot \mathcal{U}$ for every $[\psi] \in M(\Sigma, \Theta') \backslash H$. The condition $\varphi^* j = j'$ then implies $[\varphi] \in H$, hence $[\varphi] = [\psi] \in M(\Sigma, \Theta')$ for some $\psi \in G$. The latter then satisfies

$$[\psi^* j] = [\psi] \cdot [j] = [\varphi] \cdot [j] = [\varphi^* j] = [j'] \in \mathcal{T}(\Sigma, \Theta'),$$

but since $\mathcal{T}$ is G-invariant and the map $\mathcal{T} \to \mathcal{T}(\Sigma, \Theta') : j \mapsto [j]$ is injective, $\psi^* j \in \mathcal{T}$ then implies $\psi^* j = j'$. The map $\varphi \circ \psi^{-1}$ therefore belongs to $\mathrm{Aut}(\Sigma, j, \Theta') \cap \mathrm{Diff}_0(\Sigma, \Theta')$, which according to Proposition 7.19 can be assumed to be a subgroup of G. This proves $\varphi \in G$, hence (j, u) and (j', u') are equivalent in $\bar{\partial}_J^{-1}(0)/G$.

We next prove surjectivity. According to our definition of the topology of $\mathcal{M}(J)$ in §6.4, any sequence in $\mathcal{M}(J)$ converging to $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)]$ can be written as $[(\Sigma, j_\nu, \Gamma^+, \Gamma^-, \Theta, u_\nu)] \in \mathcal{M}(J)$, where

$$j_\nu \to j_0 \text{ in } C^\infty(\Sigma), \qquad u_\nu \to u_0 \text{ in } C^\infty_{\mathrm{loc}}(\dot{\Sigma}), \qquad \text{and} \qquad \bar{u}_\nu \to \bar{u}_0 \text{ in } C^0(\bar{\Sigma}).$$

Claim: These sequences can always be modified, without changing the equivalence classes in $\mathcal{M}(J)$, so as to assume $j_\nu \in \mathcal{T}$ for all ν sufficiently large.

Consider first the stable case, i.e. $\chi(\Sigma \backslash \Theta') < 0$. The convergence $j_\nu \to j_0$ implies $[j_\nu] \to [j_0]$ in $\mathcal{T}(\Sigma, \Theta')$, so by the local homeomorphism of $\mathcal{T}(\Sigma, \Theta')$ with $\mathcal{T}$ near j_0, there exists a sequence $\varphi_\nu \in \mathrm{Diff}_0(\Sigma, \Theta')$ for ν sufficiently large satisfying

$$j'_\nu := \varphi_\nu^* j_\nu \in \mathcal{T} \qquad \text{and} \qquad j'_\nu \to j_0.$$

Since the action of $\mathrm{Diff}_0(\Sigma, \Theta')$ on $\mathcal{J}(\Sigma)$ is proper, φ_ν then has a C^∞-convergent subsequence whose limit is in G_0, and the freeness of the action then implies that this limit is the identity map. Since that is the only possible limit for convergent subsequences of φ_ν, this proves $\varphi_\nu \to \mathrm{Id}$, and we are then free to replace j_ν and u_ν by $j'_\nu \in \mathcal{T}$ and $u'_\nu := u_\nu \circ \varphi_\nu$ respectively. Lemma 7.24 implies that u'_ν still converges to u_0 in both $C^\infty_{\mathrm{loc}}(\dot{\Sigma})$ and $C^0(\bar{\Sigma})$, so this proves the claim.

In the non-stable cases $\chi(\Sigma\backslash\Theta') \geqslant 0$, the argument above requires the following small modification. Expand Θ' to a slightly larger finite set $\Theta'' \subset \Sigma$ so that $\chi(\Sigma\backslash\Theta'') = -1$; this requires adding one extra marked point in the case $\Sigma = \mathbb{T}^2$ with $\Gamma = \Theta = \varnothing$, and up to three extra points if $\Sigma = S^2$. In either case, the group of homotopically trivial automorphisms G_0 is exactly large enough to find $\psi_\nu \in G_0$ sending the extra marked points to any desired images, thus we can choose ψ_ν to make $\varphi_\nu \circ \psi_\nu$ fix all points in Θ''. By the second condition in Proposition 7.19, the ψ_ν also belong to $\mathrm{Aut}(\Sigma, j'_\nu, \Theta')$, thus $\varphi_\nu \circ \psi_\nu$ is now a sequence of biholomorphic maps $(\Sigma, j'_\nu) \to (\Sigma, j_\nu)$ that fix Θ'', and the argument of the previous paragraph implies $\varphi_\nu \circ \psi_\nu \to \mathrm{Id}$. Replacing j_ν and u_ν with $j''_\nu := (\varphi_\nu \circ \psi_\nu)^* j_\nu$ and $u''_\nu := u_\nu \circ (\varphi_\nu \circ \psi_\nu)$ then proves the claim.

With the claim in place, relabel the modified sequences as j_ν and u_ν, and notice that the maps $u_\nu : (\dot{\Sigma}, j_\nu) \to (\widehat{W}, J)$ now satisfy the hypotheses of Proposition 6.32, implying that u_ν converges to u_0 in the topology of $\mathcal{B}^{k,p,\delta}$, hence $(j_\nu, u_\nu) \in \bar{\partial}_J^{-1}(0)$ and $(j_\nu, u_\nu) \to (j_0, u_0)$. This proves both surjectivity and the continuity of the inverse map.

The continuity of the map $\bar{\partial}_J^{-1}(0)/G \to \mathcal{M}(J)$ is an easy consequence of elliptic regularity: specifically, Corollary 2.25 implies that any $W^{k,p}_{\mathrm{loc}}$-convergent sequence of J-holomorphic curves is also C^∞_{loc}-convergent, and C^0-convergence on the ends follows from $W^{k,p,\delta}$-convergence since the latter implies $W^{k,p}$-convergence and $kp > 2$. $\square$

DEFINITION 7.25. We say that $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)]$ is **Fredholm regular** if there exists a choice of Teichmüller slice $\mathcal{T}$ through (j_0, Γ, Θ) such that the linearization

$$D\bar{\partial}_J(j_0, u_0) : T_{j_0}\mathcal{T} \oplus T_{u_0}\mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0)},$$
$$(y, \eta) \mapsto J(u_0) \circ du_0 \circ y + \mathbf{D}_{u_0}\eta$$

is surjective.

REMARK 7.26. It is not hard to show that the surjectivity condition in this definition does not actually depend on the choice of Teichmüller slice. The key point is that $T_{j_0}\mathcal{T}$ is complementary to the image of the canonical Cauchy-Riemann operator $\mathbf{D}_{\mathrm{Id}} : W^{k,p}_{\Gamma\cup\Theta}(T\Sigma) \to W^{k-1,p}(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma))$. Composing the latter with $du : T\dot{\Sigma} \to u^*T\widehat{W}$ then produces a subspace of $\mathcal{E}^{k-1,p,\delta}_{(j_0,u_0)}$ that is automatically contained in $\operatorname{im} \mathbf{D}_{u_0}$, because the smooth sections in this subspace can all be obtained by differentiating $\bar{\partial}_J(j_0, u)$ for a family of reparametrizations $u = u_0 \circ \varphi$ with $\varphi \in \mathrm{Diff}(\Sigma, \Gamma \cup \Theta)$. As a consequence, the operator in Definition 7.25 has the same image as its obvious extension

$$W^{k-1,p,\delta}(\overline{\mathrm{End}}_{\mathbb{C}}(T\dot{\Sigma})) \oplus T_{u_0}\mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0)},$$
$$(y, \eta) \mapsto J(u_0) \circ du_0 \circ y + \mathbf{D}_{u_0}\eta,$$

which does not involve any Teichmüller slice. (For more detailed versions of this argument, see [**Wend**, Lemma 4.3.2] or [**Wen10a**, Lemma 3.11]).

REMARK 7.27. The G-equivariance of $\bar{\partial}_J$ implies that if $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)] \in \mathcal{M}(J)$ is Fredholm regular, then so is every curve $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)] \in \mathcal{M}(J)$ for which $(j, u) \in \bar{\partial}_J^{-1}(0)$ belongs to the same G-orbit as (j_0, u_0).

The operator in Definition 7.25 is Fredholm, and we can plug in the formula (7.9) for $\dim T_{j_0}\mathcal{T} = \dim \mathcal{T}$ and Corollary 7.7 for $\operatorname{ind} \mathbf{D}_u$ to compute

$$\begin{aligned}
\operatorname{ind} D\bar{\partial}_J(j_0, u_0) &= \dim T_{j_0}\mathcal{T} + \operatorname{ind} \mathbf{D}_{u_0} \\
&= \dim G - 3\chi(\Sigma) + 2(\#\Gamma) + 2m + n\chi(\Sigma) - (n-1)\#\Gamma \\
&\qquad + 2c_1^\tau(u^* T\widehat{W}) + \sum_{j=1}^{k_+} \mu_{\mathrm{CZ}}^\tau(\gamma_j^+) - \sum_{j=1}^{k_-} \mu_{\mathrm{CZ}}^\tau(\gamma_j^-) \\
&= \dim G + (n-3)\chi(\dot{\Sigma}) + 2c_1^\tau(A) + \sum_{j=1}^{k_+} \mu_{\mathrm{CZ}}^\tau(\gamma_j^+) - \sum_{j=1}^{k_-} \mu_{\mathrm{CZ}}^\tau(\gamma_j^-) + 2m \\
&= \dim G + \operatorname{vir-dim} \mathcal{M}(J).
\end{aligned}$$

It follows that if $D\bar{\partial}_J(j_0, u_0)$ is surjective, then its kernel has finite dimension equal to $\dim G + \operatorname{vir-dim} \mathcal{M}(J)$ and thus admits a closed complement, so $D\bar{\partial}_J(j_0, u_0)$ then admits a bounded right inverse. We can therefore apply the infinite-dimensional version of the implicit function theorem (see e.g. [Lan99, §I.5]) to conclude:

LEMMA 7.28. *If $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)] \in \mathcal{M}(J)$ is Fredholm regular, then some G-invariant neighborhood $\mathcal{V} \subset \bar{\partial}_J^{-1}(0)$ of (j_0, u_0) is a smooth submanifold of $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$ with dimension $\dim G + \operatorname{vir-dim} \mathcal{M}(J)$.* □

It should now be clear how we intend to get from here to the statement that $\mathcal{M}^{\mathrm{reg}}(J)$ is an orbifold of dimension equal to $\operatorname{vir-dim} \mathcal{M}(J)$. Proposition 7.23 identifies a neighborhood of $u_0 \in \mathcal{M}^{\mathrm{reg}}(J)$ with a neighborhood of $[(j_0, u_0)]$ in $\bar{\partial}_J^{-1}(0)/G$, where a G-invariant neighborhood of (j_0, u_0) in $\bar{\partial}_J^{-1}(0)$ can be assumed to be a smooth manifold of the correct dimension, and the action of G on $\bar{\partial}_J^{-1}(0)$ is proper with finite stabilizer subgroups given by the automorphism groups of each curve. There is just one caveat: to conclude that $\bar{\partial}_J^{-1}(0)/G$ is a smooth orbifold, we need to know that the G-action on $\bar{\partial}_J^{-1}(0)$ is smooth, and this is less obvious than it looks. The naive way one would try to prove it is by showing that the G-action on the infinite-dimensional Banach manifold $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$ is smooth—if it is, then its restriction to the smooth submanifold $\bar{\partial}_J^{-1}(0)$ is also smooth. There is just one problem: if G is not discrete, then the map

$$G \times (\mathcal{T} \times \mathcal{B}^{k,p,\delta}) \xrightarrow{\Phi} \mathcal{T} \times \mathcal{B}^{k,p,\delta} : (\varphi, (j, u)) \mapsto (\varphi^* j, u \circ \varphi)$$

is definitely *not* smooth—in fact it is not even differentiable, at least not with respect to variations in G. One can see this by trying to compute its partial derivative at a point $(\mathrm{Id}, (j, u))$ with respect to variations in G: given a vector field $X \in \Gamma(T\Sigma)$ in the Lie algebra of G, choosing a smooth family $\varphi_\rho \in G$ with $\varphi_0 = \mathrm{Id}$ and $\partial_\rho \varphi_\rho|_{\rho=0} = X$ gives

$$d\Phi(\mathrm{Id}, (j, u))(X, (0, 0)) = \left(\partial_\rho(\varphi_\rho^* j)|_{\rho=0}, \partial_\rho(u \circ \varphi_\rho)|_{\rho=0}\right) = (\mathbf{D}_{\mathrm{Id}}(jX), du(X)).$$

The trouble here is that for arbitrary elements $(j, u) \in \mathcal{T} \times \mathcal{B}^{k,p,\delta}$, the expression on the right hand side does not belong to $T_{(j,u)}(\mathcal{T} \times \mathcal{B}^{k,p,\delta})$; if u is of class $W^{k,p}_{\mathrm{loc}}$ but not $W^{k+1,p}_{\mathrm{loc}}$, then the section $du(X)$ of $u^*T\widehat{W}$ will not be of class $W^{k,p}_{\mathrm{loc}}$ and thus does not belong to $T_u\mathcal{B}^{k,p,\delta}$. This annoying issue arises whenever one attempts to differentiate non-discrete reparametrization actions on infinite-dimensional Banach manifolds of maps, and the desire to view such maps as "smooth" was part of the motivation for the development of *sc-calculus* by Hofer-Wysocki-Zehnder (see e.g. [**FFGW16**] and the references therein).

For the problem at hand, however, it is not necessary to reinvent differential calculus. The situation is saved by the fact that, due to elliptic regularity, elements of $\bar{\partial}_J^{-1}(0)$ are much nicer than arbitrary maps of class $W^{k,p,\delta}$, and moreover, convergence of elements of $\bar{\partial}_J^{-1}(0)$ implies something much stronger than just $W^{k,p,\delta}$-convergence.

LEMMA 7.29. *For the smooth G-invariant neighborhood $\mathcal{V} \subset \bar{\partial}_J^{-1}(0)$ of (j_0, u_0) in Lemma 7.28, the action of G on $\mathcal{V}$ is smooth.*

SKETCH OF THE PROOF. To prove that the action is of class C^1, it suffices to show that for every $(\varphi, (j, u)) \in G \times \mathcal{V}$ and any choice of smooth charts for G near φ and $\mathcal{V}$ near (j, u), all partial derivatives of the map $(\varphi', (j', u')) \mapsto ((\varphi')^*j', u' \circ \varphi')$ are well defined and continuous on a neighborhood of $(\varphi, (j, u))$. For the derivatives in directions tangent to $\bar{\partial}_J^{-1}(0)$ this is already clear, as the action of each individual element $\varphi \in G$ on the infinite-dimensional manifold $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$ was smooth in the first place. It therefore suffices to consider derivatives in directions tangent to G, and for this purpose we can focus on the case $\varphi = \mathrm{Id}$ since any smooth path in G through φ can be written as the composition of φ (which acts smoothly) with a path through Id. Let us therefore choose a smooth 1-parameter family $\varphi_\rho \in G$ with $\varphi_0 = \mathrm{Id}$ and write $X := \partial_\rho \varphi_\rho|_{\rho=0} \in \mathfrak{g} \subset \Gamma(T\Sigma)$. We have

$$\partial_\rho(\varphi_\rho^* j, u \circ \varphi_\rho)\big|_{\rho=0} = (\mathbf{D}_{\mathrm{Id}}(jX), du(X)) \in \ker D\bar{\partial}_J(j, u) = T_{(j,u)}\bar{\partial}_J^{-1}(0),$$

where the fact that the right hand side really belongs to $T_{(j,u)}\bar{\partial}_J^{-1}(0)$ depends on the fact that, by elliptic regularity, u is smooth. To show that this expression really is the derivative of a differentiable path $\rho \mapsto (\varphi_\rho^* j, u \circ \varphi_\rho) \in \mathcal{T} \times \mathcal{B}^{k,p,\delta}$ with respect to the smooth structure of $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$, one needs to choose a Banach manifold chart for $\mathcal{B}^{k,p,\delta}$ and write down the maps $u_\rho \in \mathcal{B}^{k,p,\delta}$ in this chart. Following the prescription of [**Elĭ67**], we can assume that the chart in question has the form

$$\mathcal{B}^{k,p,\delta} \ni u_\rho = \exp_f \eta_\rho \mapsto \eta_\rho \in W^{k,p,\delta}(f^*T\widehat{W}) \oplus V_\Gamma,$$

where $f : \dot{\Sigma} \to \widehat{W}$ is a smooth map that matches trivial cylinders near each of the punctures, and $V_\Gamma \subset \Gamma(f^*T\widehat{W})$ is a finite-dimensional space of sections that are "constant" near infinity as described in §7.2. The path $\rho \mapsto u_\rho \in \mathcal{B}^{k,p,\delta}$ is then differentiable if and only if the difference quotients

$$D^h\eta_\rho\big|_{\rho=0} := \frac{\eta_h - \eta_0}{h} \in W^{k,p,\delta}(f^*T\widehat{W}) \oplus V_\Gamma$$

converge in the topology of $W^{k,p,\delta}(f^*T\widehat{W}) \oplus V_\Gamma$ to $\partial_\rho\eta_\rho|_{\rho=0}$ as $h \to 0$. On this front, the first important observation is that since u is smooth and φ_ρ depends smoothly on ρ, the map $(\rho, z) \mapsto \eta_\rho(z)$ is smooth with respect to ρ as well as z, thus by standard results about difference quotients (see §A.3), the convergence $D^h\eta_\rho|_{\rho=0} \to \partial_\rho\eta_\rho|_{\rho=0}$ as $h \to 0$ is in $C^\infty_{\mathrm{loc}}(\dot{\Sigma})$. One can then apply asymptotic regularity arguments as in §6.5 to achieve exponentially weighted convergence on each cylindrical end.

One can next apply both local and asymptotic regularity in a similar way to show that the tangent vector $du(X) \in T_u\mathcal{B}^{k,p,\delta}$ depends continuously on both $X \in \mathfrak{g}$ and $u \in \mathcal{B}^{k,p,\delta}$ so long as (j, u) belongs to $\bar{\partial}_J^{-1}(0)$, the key point being that under that condition, the $W^{k,p,\delta}$-convergence of u also implies convergence in a much stronger topology, so that the apparent loss of a derivative in $du(X)$ makes no difference. This proves that the group action is of class C^1.

Finally, the following observation gives rise to an inductive trick for going from C^1 to C^∞: there is a natural choice of almost complex structure J' on the manifold $T\widehat{W}$ such that if $\mathcal{M} := \bar{\partial}_J^{-1}(0)$ is a smooth manifold, then its tangent bundle $T\mathcal{M}$ can be identified with a space of J'-holomorphic curves in $\widehat{W}$, and differentiating the G-action on $\mathcal{M}$ gives an action of TG on $T\mathcal{M}$. The same regularity arguments then prove that the action of TG on $T\mathcal{M}$ is of class C^1, thus the original action is of class C^2, and so forth. $\square$

PROOF OF THEOREM 7.1. Combining all of the results above shows that every curve in $\mathcal{M}^{\mathrm{reg}}(J)$ has a neighborhood that can be identified with the quotient of a smooth finite-dimensional manifold $\bar{\partial}_J^{-1}(0) \subset \mathcal{T} \times \mathcal{B}^{k,p,\delta}$ by the smooth action of a Lie group G with finite isotropy such that $\dim \bar{\partial}_J^{-1}(0)/G = \text{vir-dim}\, \mathcal{M}(J)$. The moduli space $\mathcal{M}(J)$ thus inherits orbifold charts from the manifold charts that $\bar{\partial}_J^{-1}(0)$ obtains from applying the implicit function theorem to the section $\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$. There is still a bit of work to be done in showing that transition maps relating any two overlapping charts that arise in this way are smooth. The hard part is actually to show that two charts constructed near a given curve (j_0, u_0) via two different choices of Teichmüller slice are smoothly compatible. This is yet another case where, as in Lemma 7.29, the smoothness cannot be seen at the infinite-dimensional level, but only works on the finite-dimensional submanifolds as a consequence of elliptic regularity. For details, see the proof of Theorem 4.3.6 in [**Wend**]. $\square$

EXERCISE 7.30. Take a deep breath.

It will be useful also to take note of how the analytical picture of $\mathcal{M}(J)$ near a Fredholm regular curve $u_0 \in \mathcal{M}(J)$ characterizes the tangent space $T_{u_0}\mathcal{M}(J)$. The group $G = \mathrm{Aut}(\Sigma, j_0, \Gamma \cup \Theta)$ may be either finite or (if $\chi(\dot{\Sigma}\backslash\Theta) \geqslant 0$) a positive-dimensional Lie group, in which case we shall denote its Lie algebra by

$$\mathfrak{g} := \mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta) = T_{\mathrm{Id}}G.$$

Recall that by Remark 6.31, we only need to consider nontrivial $\mathfrak{g}$ in cases where $u_0 : \dot{\Sigma} \to \widehat{W}$ is not a constant map. Now if u_0 is Fredholm regular, so that $\bar{\partial}_J^{-1}(0)$ near (j_0, u_0) is a manifold with $T_{(j_0,u_0)}\bar{\partial}_J^{-1}(0) = \ker D\bar{\partial}_J(j_0, u_0)$, we can linearize the group action (7.10) with respect to G and obtain a map

$$\mathfrak{g} \to \ker D\bar{\partial}_J(j_0, u_0) : X \mapsto (\mathbf{D}_{\mathrm{Id}}(j_0 X), du_0(X))$$

which is injective since X and du_0 can each only have isolated zeroes. It is therefore natural to regard $\mathfrak{g}$ as a linear subspace of $\ker D\bar{\partial}_J(j_0, u_0)$, and linearizing Proposition 7.23 in the regular case then leads to the following statement:

PROPOSITION 7.31. *If $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)] \in \mathcal{M}(J)$ is Fredholm regular, then the tangent space $T_{u_0}\mathcal{M}(J)$ admits a natural isomorphism*

$$T_{u_0}\mathcal{M}(J) = \ker D\bar{\partial}_J(j_0, u_0) \Big/ \mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta).$$

□

7.5. Evaluation and forgetful maps

In addition to the smoothness of $\mathcal{M}(J)$, we sometimes need to know that certain canonically defined maps on $\mathcal{M}(J)$ are smooth. Considering curves with m marked points and $k_\pm$ positive/negative punctures and writing

$$\ell := m + k_+ + k_-,$$

the first of these is the **forgetful map**

$$\Phi : \mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-) \to \mathcal{M}_{g,\ell}, \tag{7.11}$$

which sends an equivalence class of curves $(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)$ to the equivalence class of its underlying Riemann surface $(\Sigma, j, \Gamma \cup \Theta)$ by forgetting the map $u : \dot{\Sigma} \to \widehat{W}$. Recall that since $\mathcal{T}(\Sigma, \Gamma \cup \Theta)$ is a smooth manifold and $\mathcal{M}_{g,\ell}$ is its quotient by the proper action of $M(\Sigma, \Theta)$ with finite isotropy, $\mathcal{M}_{g,\ell}$ carries a natural smooth orbifold structure.

PROPOSITION 7.32. *The forgetful map (7.11) is smooth in some neighborhood of any Fredholm regular curve.*

PROOF. Suppose $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)] \in \mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ is regular, denote

$$G = \mathrm{Aut}(\Sigma, j_0, \Gamma \cup \Theta), \qquad G_0 := G \cap \mathrm{Diff}_0(\Sigma, j_0, \Gamma \cup \Theta) \quad \text{and} \quad H := G/G_0,$$

and assume $\mathcal{T} \subset \mathcal{J}(\Sigma)$ is a G-invariant Teichmüller slice as provided by Proposition 7.19. By Lemma 7.21, the natural map

$$\mathcal{T}/H \to \mathcal{M}_{g,\ell} : [j] \mapsto [(\Sigma, j, \Gamma \cup \Theta)]$$

is then a homeomorphism between open neighborhoods of $[j_0]$ and $[(\Sigma, j_0, \Gamma \cup \Theta)]$, and in fact one can use this map to define the smooth orbifold structure of $\mathcal{M}_{g,\ell}$ and thus call it a local diffeomorphism. Combining this local picture of $\mathcal{M}_{g,\ell}$ with Proposition 7.23, the forgetful map is expressed locally as

$$\bar{\partial}_J^{-1}(0)/G \to \mathcal{T}/H : [(j, u)] \mapsto [j],$$

where Fredholm regularity implies via the implicit function theorem that $\bar{\partial}_J^{-1}(0)$ is a smooth submanifold of $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$. This map between orbifolds is smooth because it is induced by the smooth map $\bar{\partial}_J^{-1}(0) \to \mathcal{T} : (j, u) \mapsto j$, which is the composition of the smooth inclusion $\bar{\partial}_J^{-1}(0) \hookrightarrow \mathcal{T} \times \mathcal{B}^{k,p,\delta}$ with the manifestly smooth projection map $\mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{T}$. □

In the presence of marked points, we can consider the **evaluation map**

$$\mathrm{ev} : \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) \to \widehat{W}^{\times m}, \tag{7.12}$$

sending the equivalence class of a curve $(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)$ to the m-tuple of points $(u(\zeta_1), \ldots, u(\zeta_m))$ in its image, where $\Theta = (\zeta_1, \ldots, \zeta_m)$.

PROPOSITION 7.33. *The evaluation map (7.12) is smooth on the set of Fredholm regular curves in $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$.*

PROOF. In the neighborhood of a regular curve $u_0 : (\dot{\Sigma} = \Sigma \setminus (\Gamma^+ \cup \Gamma^-), j_0) \to (\widehat{W}, J)$ with marked points $\Theta = (\zeta_1, \ldots, \zeta_m) \in \dot{\Sigma}^{\times m}$, ev takes the form

$$\bar{\partial}_J^{-1}(0) / \operatorname{Aut}(\Sigma, j_0, \Gamma \cup \Theta) \to \widehat{W}^{\times m} : [(j, u)] \mapsto (u(\zeta_1), \ldots, u(\zeta_m)).$$

This lifts to a map $\bar{\partial}_J^{-1}(0) \to \widehat{W}^{\times m}$, which is the composition of the smooth inclusion of the submanifold $\bar{\partial}_J^{-1}(0) \hookrightarrow \mathcal{T} \times \mathcal{B}^{k,p,\delta}$ with the smooth projection $\mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{B}^{k,p,\delta}$ and the map

$$\mathcal{B}^{k,p,\delta} \to \widehat{W}^{\times m} : u \mapsto (u(\zeta_1), \ldots, u(\zeta_m)). \tag{7.13}$$

The latter is smooth for the following reason: by the prescription in [Elĭ67], the Banach manifold structure on $\mathcal{B}^{k,p,\delta}$ is defined via charts of the form $\exp_f \eta \mapsto \eta \in W^{k,p,\delta}(f^*T\widehat{W}) \oplus V_\Gamma$ for smooth reference maps $f : \dot{\Sigma} \to \widehat{W}$ that are cylindrical near infinity, where V_Γ is a finite-dimensional space of smooth vector fields along f. The map (7.13) with respect to a chart thus takes the form

$$W^{k,p,\delta}(f^*T\widehat{W}) \oplus V_\Gamma \supset \mathcal{O} \to \widehat{W} : \eta \mapsto (\exp_{f(\zeta_1)} \eta(\zeta_1), \ldots, \exp_{f(\eta_m)} \eta(\zeta_m))$$

for a suitable neighborhood $\mathcal{O} \subset W^{k,p,\delta}(f^*T\widehat{W}) \oplus V_\Gamma$ of 0. Each factor in this map is just the composition of the smooth map $\exp : T\widehat{W} \to \widehat{W}$ (defined on a neighborhood of the zero-section) with a map of the form $W^{k,p,\delta}(f^*T\widehat{W}) \oplus V_\Gamma \to T_{f(\zeta_i)}\widehat{W} : \eta \mapsto \eta(\zeta_i)$. The latter map is linear, and crucially, it is also continuous (and therefore smooth) since k and p are always chosen for the Sobolev embedding theorem to hold. □

There are also other types of "forgetful" maps defined by forgetting marked points instead of the map u: for instance, the map

$$\pi_m : \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) \to \mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-), \tag{7.14}$$

modifies $[(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)]$ by replacing Θ with the empty set. We would like to show next that this map is a smooth submersion over the set of Fredholm regular curves. From a certain perspective, this statement is very easy to believe, though the proof turns out to inolve a few subtleties. First, suppose $u_0 : (\dot{\Sigma} = \Sigma \setminus (\Gamma^+ \cup \Gamma^-), j_0) \to (\widehat{W}, J)$ represents a regular curve in $\mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$, write $G := \operatorname{Aut}(\Sigma, j_0, \Gamma)$, and choose a suitable Teichmüller slice $\mathcal{T}$ through j_0 so that a neighborhood $\mathcal{U} \subset \mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ of u_0 is identified with $\mathcal{V}/G$ for a suitable G-invariant neighborhood

$$\mathcal{V} \subset \bar{\partial}_J^{-1}(0)$$

of (j_0, u_0) that is a smooth submanifold of $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$. If $\Delta \subset \dot{\Sigma}^{\times m}$ denotes the closed subset consisting of m-tuples in which at least two entries are equal, then there is now an obvious homeomorphism

$$\begin{aligned} \left(\mathcal{V} \times (\dot{\Sigma}^{\times m} \backslash \Delta)\right) / G &\to \pi_m^{-1}(\mathcal{U}), \\ [(j, u, \Theta)] &\mapsto [(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)], \end{aligned} \tag{7.15}$$

where automorphisms $\varphi \in G$ act on (j, u) in the usual way and send $\Theta = (\zeta_1, \ldots, \zeta_m) \in \dot{\Sigma}^{\times m} \backslash \Delta$ to

$$\varphi^* \Theta := (\varphi^{-1}(\zeta_1), \ldots, \varphi^{-1}(\zeta_m)).$$

The isotropy subgroup for each $(j, u, \Theta) \in \mathcal{V} \times (\dot{\Sigma}^{\times m} \backslash \Delta)$ under the G-action is simply the finite automorphism group of u with its marked points Θ, so this picture identifies $\pi_m^{-1}(\mathcal{U})$ with a smooth orbifold having dimension equal to the virtual dimension of $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$. Notice that while $u_0 \in \mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ in this discussion was assumed to be Fredholm regular, we have not considered so far whether any given element of $\pi_m^{-1}(u_0) \subset \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ is Fredholm regular. It would be tempting at this point to circumvent that question and just use (7.15) to *define* the smooth structure on $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$. This would make it obvious that π_m is a smooth submersion, as π_m in this picture looks like the map

$$\left(\mathcal{V} \times (\dot{\Sigma}^{\times m} \backslash \Delta)\right) / G \to \mathcal{V}/G$$

induced by the G-equivariant (and manifestly smooth) projection map $\mathcal{V} \times (\dot{\Sigma}^{\times m} \backslash \Delta) \to \mathcal{V}$. But there is a technical problem: the evaluation map in this picture looks like

$$\begin{aligned} \left(\mathcal{V} \times (\dot{\Sigma}^{\times m} \backslash \Delta)\right) / G &\to \widehat{W}^{\times m}, \\ [(j, u, (\zeta_1, \ldots, \zeta_m)] &\mapsto (u(\zeta_1), \ldots, u(\zeta_m)), \end{aligned}$$

and it is not obvious whether this map is smooth. The most natural way to try to prove it would be to present its lift $\mathcal{V} \times (\dot{\Sigma}^{\times m} \backslash \Delta) \to \widehat{W}^{\times m}$ as the composition of the smooth inclusion

$$\bar{\partial}_J^{-1}(0) \times (\dot{\Sigma}^{\times m} \backslash \Delta) \hookrightarrow \mathcal{T} \times \mathcal{B}^{k,p,\delta} \times (\dot{\Sigma}^{\times m} \backslash \Delta)$$

with the map

$$\mathcal{T} \times \mathcal{B}^{k,p,\delta} \times (\dot{\Sigma}^{\times m} \backslash \Delta) \to \widehat{W}^{\times m} : (j, u, (\zeta_1, \ldots, \zeta_m)) \mapsto (u(\zeta_1), \ldots, u(\zeta_m)).$$

But the latter is definitely *not* a smooth map, as arbitrary elements $u \in \mathcal{B}^{k,p,\delta}$ can only be assumed to have finitely many derivatives. Our intuition says that this should not pose a problem since, by elliptic regularity, u is indeed smooth whenever $(j, u) \in \bar{\partial}_J^{-1}(0)$, but turning this intuition into a rigorous argument is not easy. This difficulty did not arise in Proposition 7.33 because we were only considering parametrizations $u : \dot{\Sigma} \to \widehat{W}$ for which the locations of the marked points were fixed in advanced and not allowed to change—in this case the lack of deriatives of maps $u \in \mathcal{B}^{k,p,\delta}$ makes no difference. But that argument required $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ to be endowed with the particular smooth structure that comes from applying the implicit function theorem to Fredholm regular curves *with marked points*. It means in particular that instead of using (7.15), a neighborhood of $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)] \in \mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ must be identified with a subset of $\bar{\partial}_J^{-1}(0)/\operatorname{Aut}(\Sigma, j_0, \Gamma \cup \Theta)$, with the definition of $\bar{\partial}_J$ now requiring a Teichmüller slice of larger dimension in order to account for the marked points. That perspective has some disadvantages in comparison with (7.15), notably that it makes it much harder to see why the map π_m is a smooth submersion or what its fibers look like. With this as motivation, the next lemma says that it is safe after all to go back and forth between these two perspectives on $\mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$.

LEMMA 7.34. *If $u_0 : (\dot{\Sigma} = \Sigma \backslash (\Gamma^+ \cup \Gamma^-), j_0) \to (\widehat{W}, J)$ represents a Fredholm regular element of the moduli space $\mathcal{M}_{g,0}(J, A, \gamma^+, \gamma^-)$ without marked points, then every element of $\pi_m^{-1}(u_0) \subset \mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ is also Fredholm regular. Moreover, for a sufficiently small neighborhood $\mathcal{U} \subset \mathcal{M}_{g,0}(J, A, \gamma^+, \gamma^-)$ of u_0, there exists a Teichmüller slice $\mathcal{T}$ through (j_0, Γ) and a neighborhood $\mathcal{V} \subset \bar{\partial}_J^{-1}(0) \subset \mathcal{T} \times \mathcal{B}^{k,p,\delta}$ of (j_0, u_0) such that the map (7.15) is a diffeomorphism.*

PROOF. By an inductive argument, it will suffice to consider cases where u_0 already has some marked points and one more is added. Let

$$\pi : \mathcal{M}_{g,m+1}(J, A, \gamma^+, \gamma^-) \to \mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$$

denote the canonical map defined by forgetting the last marked point, suppose $u_0 : (\dot{\Sigma} = \Sigma \backslash (\Gamma^+ \cup \Gamma^-), j_0) \to (\widehat{W}, J)$ with marked points $\Theta = (\zeta_1, \ldots, \zeta_m)$ represents a Fredholm regular element in $\mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$, write $G = \operatorname{Aut}(\Sigma, j_0, \Gamma \cup \Theta)$, and choose a Teichmüller slice $\mathcal{T}$ through $(j_0, \Gamma \cup \Theta)$ and a suitable neighborhood $\mathcal{U} \subset \mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$ of u_0 that is identified with $\mathcal{V}/G$ for some smooth G-invariant neighborhood $\mathcal{V}$ of (j_0, u_0) in the zero set of $\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$. There is then a natural homeomorphism

$$\begin{aligned} \left(\mathcal{V} \times (\dot{\Sigma} \backslash \Theta)\right)/G &\to \pi^{-1}(\mathcal{U}), \\ [(j, u, \zeta)] &\mapsto [(\Sigma, j, \Gamma^+, \Gamma^-, \Theta_\zeta, u)] \end{aligned} \tag{7.16}$$

where

$$\Theta_\zeta := (\zeta_1, \ldots, \zeta_m, \zeta).$$

We claim that every element in $\pi^{-1}(u_0)$ is Fredholm regular, hence a neighborhood of this set in $\mathcal{M}_{g,m+1}(J, A, \gamma^+, \gamma^-)$ has a natural smooth structure given by Theorem 7.1, and moreover, that the map (7.16) is a diffeomorphism with respect to this smooth structure.

Slightly different arguments are required depending on whether $(\Sigma, j_0, \Gamma \cup \Theta)$ is or is not stable, so let us first consider the stable case. Fix $\zeta_{m+1} \in \dot{\Sigma} \backslash \Theta$ and write $\Theta' = (\zeta_1, \ldots, \zeta_{m+1})$. Let $\mathcal{T}$ and $\mathcal{T}'$ denote the pair of Teichmüller slices through $(j_0, \Gamma \cup \Theta)$ and $(j_0, \Gamma \cup \Theta')$ respectively provided by Lemma 7.22, so in particular,

$$\mathcal{T}' = \left\{\varphi_\zeta^* j \in \mathcal{J}(\Sigma)\right\}_{(j,\zeta) \in \mathcal{T} \times \mathcal{U}}$$

for an arbitrarily small neighborhood $\mathcal{U} \subset \dot{\Sigma}$ of ζ_{m+1} and a smooth family of diffeomorphisms $\{\varphi_\zeta : \Sigma \to \Sigma\}_{\zeta \in \mathcal{U}}$ that are supported in a slightly larger neighborhood of

ζ_{m+1} and satisfy $\varphi_{\zeta_{m+1}} = \mathrm{Id}$ and $\varphi_\zeta(\zeta_{m+1}) = \zeta$ for every $\zeta \in \mathcal{U}$. Since $\mathcal{T} \subset \mathcal{T}'$, the operator

$$D\bar{\partial}_J(j_0, u_0) : T_{j_0}\mathcal{T}' \oplus T_{u_0}\mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0)} \tag{7.17}$$

that needs to be surjective in order for $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta', u_0)] \in \mathcal{M}_{g,m+1}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ to be Fredholm regular is simply an extension of the operator

$$D\bar{\partial}_J(j_0, u_0) : T_{j_0}\mathcal{T} \oplus T_{u_0}\mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0)} \tag{7.18}$$

to a larger domain, and the latter is surjective by assumption, so the regularity claim holds. Next, observe that in some neighborhood of (j_0, u_0), $\bar{\partial}_J^{-1}(0) \subset \mathcal{T} \times \mathcal{B}^{k,p,\delta}$ is a smooth submanifold with dimension $I := \mathrm{ind}(u_0) + 2m$, and the set

$$\left\{(\varphi_\zeta^* j, u \circ \varphi_\zeta) \in \mathcal{T}' \times \mathcal{B}^{k,p,\delta} \;\middle|\; (j, u) \in \mathcal{T} \times \mathcal{B}^{k,p,\delta},\ \bar{\partial}_J(j, u) = 0 \text{ and } \zeta \in \mathcal{U}\right\} \tag{7.19}$$

then forms a smooth manifold of dimension $I + 2$ living inside $\bar{\partial}_J^{-1}(0) \subset \mathcal{T}' \times \mathcal{B}^{k,p,\delta}$. But the linearization (7.17) of $\bar{\partial}_J$ on $\mathcal{T}' \times \mathcal{B}^{k,p,\delta}$ is a surjective Fredholm operator of index $I+2$ since the related operator (7.18) has index I and $\dim T_{j_0}\mathcal{T}' = \dim T_{j_0}\mathcal{T} + 2$, so the implicit function theorem implies that (7.19) characterizes an entire neighborhood of $[(j_0, u_0)]$ in $\bar{\partial}_J^{-1}(0) \subset \mathcal{T}' \times \mathcal{B}^{k,p,\delta}$. With this understood, the map (7.16) can now be expressed in terms of a map from a neighborhood of (j_0, u_0, ζ_{m+1}) in $\bar{\partial}_J^{-1}(0) \times (\dot{\Sigma}\backslash\Theta) \subset \mathcal{T} \times \mathcal{B}^{k,p,\delta} \times (\dot{\Sigma}\backslash\Theta)$ to a neighborhood of (j_0, u_0) in $\bar{\partial}_J^{-1}(0) \subset \mathcal{T}' \times \mathcal{B}^{k,p,\delta}$ taking the form

$$(j, u, \zeta) \mapsto (\varphi_\zeta^* j, u \circ \varphi_\zeta).$$

By the same argument as in Lemma 7.29, this map is smooth, and it has nonsingular derivative, so it is a diffeomorphism between the corresponding neighborhoods.

If $(\Sigma, j_0, \Gamma \cup \Theta)$ is not stable, then the treatment of Teichmüller slices in this story simplifies: Lemma 7.22 implies that we can fix a single G-invariant Teichmüller slice through both (j_0, Θ) and (j_0, Θ'), which is then also G'-invariant for $G' := \mathrm{Aut}(\Sigma, j_0, \Gamma \cup \Theta')$. The matter of Fredholm regularity is thus trivial, i.e. u_0 is regular if and only if every element of $\pi^{-1}(u_0)$ is regular. On the other hand, G is no longer finite, but is a positive-dimensional Lie group with G' as a Lie subgroup of codimension 2. The important observation is then that there is a well-defined smooth map

$$G/G' \to \dot{\Sigma} : [\varphi] \mapsto \varphi(\zeta_{m+1})$$

which takes a neighborhood of $[\mathrm{Id}] \in G/G'$ diffeomorphically to a neighborhood of ζ_{m+1} in $\dot{\Sigma}$. We leave the remaining details as an exercise. $\square$

Corollary 7.35. *The map from $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ to $\mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ defined by forgetting all marked points is a smooth submersion on the preimage of the set of Fredholm regular curves in $\mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$.* $\square$

LECTURE 8

Transversality

Contents

The previous lecture proved that the open set of Fredholm regular J-holomorphic curves in $\mathcal{M}(J)$ has a natural smooth finite-dimensional orbifold structure. In practice, however, Fredholm regularity is a technical condition that can rarely be directly checked. To remedy this, the present lecture will be devoted to *genericity* results, showing that Fredholm regularity will always hold if we are willing to make small perturbations of J. We need to prove two slightly different versions of this statement: one for curves in a completed symplectic cobordism (§8.2), and another for curves in symplectizations (§8.3), which presents distinctive problems because the space of allowable perturbations is smaller. The caveat in both cases is that these genericity arguments only work for *somewhere injective* curves, i.e. they fail for multiple covers. This has to do with the fundamental incompatibility between transversality and symmetry, and it is generally not a solvable problem within the framework developed in this book, though there certainly do exist interesting ideas for solving it (see [**FFGW16**]). For all applications that we will discuss rigorously in this book, rigor is possible only because geometric conditions can be used to exclude most multiple covers from consideration. This caveat does not fully negate the value of non-rigorous handwaving based on the fiction that transversality for multiple covers is not a problem—we will indulge in such handwaving in Lectures 12 and 13.

8.1. A paradigm for genericity arguments

Before stating the main results of this lecture, let us discuss in general terms what a "genericity" result is and how one can go about proving it. The canonical example of a genericity result is Sard's theorem: for any smooth map $f : M \to N$

between finite-dimensional manifolds, almost every point in N is a regular value. The words "almost every" imply in particular that the set of regular values is dense, so any point $p \in N$ admits an arbitrarily small perturbation to a nearby point $p' \in N$ such that $f^{-1}(p')$ is a smooth submanifold of M. But "almost every" also implies something better than density: the intersection of two dense sets can be empty, but if there is a countable set of conditions that are satisfied by almost every point, then almost every point satisfies *all* of them, as the the union of countably many sets of measure zero still has measure zero.

A second example of a genericity result is the statement that on any smooth vector bundle $E \to B$, every section can be perturbed to one that is transverse to the zero-section, implying (via the implicit function theorem) that its zero set is a submanifold of B. Stated in this way, one obtains a dense subset $\Gamma^{\mathrm{reg}}(E) \subset \Gamma(E)$ such that every $\eta \in \Gamma^{\mathrm{reg}}(E)$ is transverse to the zero-section, but as with Sard's theorem, more than density is actually true. The statement "almost every section is transverse to the zero-section" would unfortunately not make sense, as there is no natural measure on infinite-dimensional spaces like $\Gamma(E)$ with which to define what "almost every" should mean. The following notion serves as a reasonable substitute in infinite-dimensional settings.

DEFINITION 8.1. If X is a topological space, a subset $Y \subset X$ is called **comeager** if it contains a countable intersection of open and dense sets.[1]

If X is a complete metrizable space, then the Baire category theorem implies that comeager subsets are always dense—moreover, any countable intersection of comeager subsets is also comeager and therefore dense. Informally, we often say that a given statement dependent on a choice of auxiliary data (living in a complete metrizable space) is true **generically**, or "for **generic** choices," if it is true whenever the data are chosen from some comeager subset of the space of all possible data.

In order to describe the standard procedure we will follow for proving genericity results, let us sketch the proof that generic sections of a smooth vector bundle $E \to B$ are transverse to the zero-section. For this purpose, the bundle may in general be either finite or infinite dimensional, though in the latter case, we will see that some extra conditions on sections need to be imposed. Given $s \in \Gamma(E)$, let us call a point $x \in s^{-1}(0)$ **regular** if the linearization at that point

$$Ds(x) : T_x B \to E_x$$

is surjective. A section s is then transverse to the zero-section if and only if every point $x \in s^{-1}(0)$ is regular. To show that this is true generically, the first step is to choose a suitable subset

$$X \subset \Gamma(E)$$

in which one would like to perturb s. This set needs to be sufficiently large, in a sense to be specified below, and it also needs to be a manifold—putting both conditions

[1] Elsewhere in the symplectic literature, comeager subsets are sometimes referred to as "sets of second category," which is unfortunately slightly at odds with the standard meaning of "second category," though it is accurate to say that the *complement* of a comeager subset (also known as a "meager" subset) is a set of first category. The term *Baire subset* is also sometimes used as a synonym for "comeager subset".

together, X will usually need to be an infinite-dimensional Banach manifold. One should also choose X so that if $E' \to X \times B$ denotes the pullback of E along the projection $X \times B \to B$, then the section

$$f : X \times B \to E' : (s, x) \mapsto s(x)$$

is smooth. Its zero set

$$\mathcal{M}(X) := f^{-1}(0) = \{(s, x) \in X \times B \mid s(x) = 0\}$$

is called a **universal moduli space**, as it can be regarded as the union over all admissible perturbed sections $s \in X \subset \Gamma(E)$ of the "moduli spaces"

$$\mathcal{M}(s) := s^{-1}(0).$$

The essential step in most genericity proofs is to show that for every $(s, x) \in \mathcal{M}$, the linearization

$$Df(s, x) : T_s X \oplus T_x B \to E'_{(s,x)} : (t, v) \mapsto Ds(x)v + t(x)$$

is surjective. The way to prove this is typically by ignoring the term $Ds(x)$ and proving that the linearization with respect to variations in s

$$D_1 f(s, x) : T_s X \to E'_{(s,x)} : t \mapsto t(x) \tag{8.1}$$

is surjective. Now perhaps you can see what was meant by the words "sufficiently large" above: since zeroes of a given section $s \in X$ can in principle appear anywhere, the only way to make sure (8.1) is surjective is by constructing X so that for every $x \in B$ and $v \in E_x$, X contains a section $t \in X$ with $t(x) = v$. If this holds, then[2] the universal moduli space becomes a smooth manifold, typically infinite dimensional. The bulk of the effort in most genericity arguments goes into establishing this fact.

The rest of the argument follows a standard pattern: one considers the projection map

$$\pi : \mathcal{M} \to X : (s, x) \mapsto s,$$

which is clearly smooth if $\mathcal{M}$ is a smooth submanifold of $X \times B$. In light of the obvious bijection between $\pi^{-1}(s)$ and $\mathcal{M}(s) = s^{-1}(0)$ for each $s \in X$, we are motivated to ask whether *generic* sections $s \in X$ are regular values of this projection, so that $\pi^{-1}(s)$ becomes a manifold. When X is finite dimensional, this question is answered by Sard's theorem, but we also need an answer for cases where $\dim X = \infty$. The following generalization of Sard's theorem was proved by Smale in 1965; for a concise proof using finite-dimensional reduction, see [MS12, §A.5].[3]

The Sard-Smale Theorem ([Sma65]). *Suppose M and N are second countable Banach manifolds of class C^k and $F : M \to N$ is a map of class C^k with $k \geqslant 1$ such that for every $x \in M$, the derivative $dF(x) : T_x M \to T_{F(x)} N$ is a Fredholm*

[2] There is a detail brushed under the rug here if X is infinite dimensional: in order to apply the implicit function theorem and prove that $f^{-1}(0)$ is a manifold, one needs to know that $Df(s, x)$ is not only surjective but also has a bounded right inverse. This will be automatic for the cases we are interested in because $Ds(x)$ in those cases is Fredholm (see Exercise 8.11.

[3] The Sard-Smale theorem is stated in [MS12] for separable Banach spaces, but the proof is easily adapted for second countable Banach manifolds using local charts and the fact that every open cover has a countable subcover.

operator with $k \geqslant \operatorname{ind} dF(x) + 1$. *Then there exists a comeager subset* $N^{\mathrm{reg}} \subset N$ *such that for every* $y \in N^{\mathrm{reg}}$ *and* $x \in F^{-1}(y)$, $dF(x) : T_x M \to T_y N$ *is surjective.* $\square$

The derivative of the projection $\pi : \mathcal{M} \to X$ at $(s, x) \in \mathcal{M}$ is the linear projection map

$$d\pi(s, x) : \ker Df(s, x) \to T_x X : (t, v) \mapsto t.$$

To complete the picture, we need an easy algebraic lemma about linear maps:

LEMMA 8.2. *Suppose* X, Y *and* Z *are vector spaces,* $D : X \to Z$ *and* $A : Y \to Z$ *are linear maps, and* $L : X \oplus Y \to Z : (x, y) \mapsto Dx + Ay$ *is surjective. Then for the projection*

$$\Pi : \ker L \to Y : (x, y) \mapsto y,$$

there are natural isomorphisms $\ker \Pi \cong \ker D$ *and* $\operatorname{coker} \Pi \cong \operatorname{coker} D$.

PROOF. The isomorphism of the kernels is clear: it is just the restriction of the inclusion $X \hookrightarrow X \oplus Y : x \mapsto (x, 0)$ to $\ker D$. We construct an isomorphism $\operatorname{coker} \Pi \to \operatorname{coker} D$ as follows. Observe that $\operatorname{im} \Pi$ is simply the space of all $y \in Y$ such that $Ay = -Dx$ for some $x \in X$, hence $\operatorname{im} \Pi = A^{-1}(\operatorname{im} D)$, and

$$\operatorname{coker} \Pi = Y / \operatorname{im} \Pi = Y / A^{-1}(\operatorname{im} D).$$

Now it is easy to check that the map $A : Y \to \operatorname{im} A$ descends to an isomorphism

$$A : Y / A^{-1}(\operatorname{im} D) \to \operatorname{im} A / (\operatorname{im} D \cap \operatorname{im} A),$$

and similarly, the inclusion $\operatorname{im} A \hookrightarrow Z$ descends to an injective homomorphism

$$\operatorname{im} A / (\operatorname{im} D \cap \operatorname{im} A) \to Z / \operatorname{im} D.$$

Since every $z \in Z$ can be written as $z = Dx + Ay$ by assumption, this map is also surjective. $\square$

To apply this, suppose every section in X has the property that $Ds(x) : T_x B \to E_x$ is Fredholm for every $x \in s^{-1}(0)$; note that this is always true if B and E are finite dimensional, and it is also true for the infinite-dimensional nonlinear Cauchy-Riemann operator $\bar{\partial} : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$ considered in the previous lecture. Since $\ker Ds(x)$ and $\operatorname{coker} Ds(x)$ are finite dimensional, Lemma 8.2 now implies the same for $\ker D\pi(s, x)$ and $\operatorname{coker} D\pi(s, x)$ for every $(s, x) \in \mathcal{M}$, hence $D\pi(s, x)$ is always Fredholm, and moreover, it is surjective if and only if $Ds(x)$ is surjective. The Sard-Smale theorem therefore gives us a comeager subset $X^{\mathrm{reg}} \subset X$ such that for every $s \in X^{\mathrm{reg}}$ and $x \in s^{-1}(0)$, $Ds(x)$ is surjective—in other words, s is transverse to the zero-section.

Our analytical setup for the moduli space of J-holomorphic curves differs from the story described above in the following respects:

(1) Instead of $\mathcal{M}(J)$ being globally the zero set of a section of a vector bundle, it can *locally* be identified with sets of the form $\bar{\partial}_J^{-1}(0)/G$, i.e. the quotient of a zero set of a bundle by a smooth Lie group action with finite isotropy.
(2) The set of admissible perturbations (called X in the discussion above) will not be quite large enough in general: in particular, it will not be true that for every $(j, u) \in \bar{\partial}_J^{-1}(0)$, perturbations of J can be found realizing arbitrary perturbations in the value of $\bar{\partial}_J(j, u)$.

We will see that the first issue is not really a problem, but the second one is. It is a symptom of the equivariance of $\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$ with respect to the action of the automorphism group: transversality would easily be achieved if arbitrary non-equivariant perturbations to $\bar{\partial}_J$ were allowed, but all perturbations that result from changing J are automatically equivariant, which is a serious restriction. This is why multiply covered curves must be excluded from the main results of this lecture.

There is one other complication that also afflicts the example sketched above but was brushed under the rug: the space of allowable perturbations $X \subset \Gamma(E)$ must be defined to have certain properties for technical reasons, but X is typically not the space we actually want to prove a theorem about. If the goal is to prove that generic smooth sections $\eta \in \Gamma(E)$ are transverse to the zero-section, then the natural choice would seem to be $X := \Gamma(E)$, but this is only a Fréchet space, not a Banach manifold: you cannot use it in the implicit function theorem to prove that the universal moduli space is smooth, and you cannot feed it into the Sard-Smale theorem. There are various ways to overcome this difficulty, but they are all somewhat unnatural and add an extra step to the proof—in the holomorphic curve setting, that step will be described in §8.2.4.

8.2. Generic transversality in cobordisms

8.2.1. A theorem for somewhere injective curves. A smooth map $u : \dot{\Sigma} \to \widehat{W}$ is said to have an **injective point** $z \in \dot{\Sigma}$ if

$$du(z) : T_z\dot{\Sigma} \to T_{u(z)}\widehat{W} \text{ is injective} \quad \text{and} \quad u^{-1}(u(z)) = \{z\}.$$

If u is a proper map, then it is easy to see that the set of injective points is open in $\dot{\Sigma}$, though in general it could also be empty; this is the case e.g. for multiply covered J-holomorphic curves. We say u is **somewhere injective** if its set of injective points is nonempty. For asymptotically cylindrical J-holomorphic curves with nondegenerate asymptotic orbits, Theorem 6.34 implies that somewhere injectivity is equivalent to being *simple*, i.e. not multiply covered.

Here is the first of the two main results in this lecture. It is stated specifically for curves in completed cobordisms; an analogue for curves in symplectizations will be the subject of §8.3.

Theorem 8.3. *Assume as in Theorem 7.1 that all the orbits $\gamma_i^\pm$ are nondegenerate, fix an almost complex structure $J^{\mathrm{fix}} \in \mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ and an open subset*

$$\mathcal{U} \subset W^{r_0}.$$

Then there exists a comeager subset

$$\mathcal{J}^{\mathrm{reg}}_{\mathcal{U}} \subset \left\{ J \in \mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-) \;\middle|\; J = J^{\mathrm{fix}} \text{ on } \widehat{W} \backslash \mathcal{U} \right\},$$

such that for every $J \in \mathcal{J}^{\mathrm{reg}}_{\mathcal{U}}$, every curve $u \in \mathcal{M}(J)$ that has an injective point mapped into $\mathcal{U}$ is Fredholm regular. In particular, the curves with this property define an open subset of $\mathcal{M}(J)$ that is a smooth manifold with dimension equal to its virtual dimension.

REMARK 8.4. The theorem is equally true if $\mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ is replaced by the larger space $\mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$. This distinction makes a difference at only one step in the proof, where the compatible case is slightly harder than the tame case because the space of available perturbations is smaller (see Lemma 8.13 and the discussion that precedes it).

REMARK 8.5. Since $\mathcal{U} \subset \widehat{W}$ has compact closure, the set

$$\left\{ J \in \mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-) \;\middle|\; J = J^{\text{fix}} \text{ on } \widehat{W} \backslash \mathcal{U} \right\}$$

with its natural C^∞-topology is a complete metrizable space; in fact it can be given the structure of a Fréchet manifold, though we will not need to use this fact. The important detail is that the Baire category theorem applies to this space, and guarantees that comeager subsets of it are dense.

REMARK 8.6. As mentioned above, the condition that $u : \dot{\Sigma} \to \widehat{W}$ has an injective point mapped into $\mathcal{U}$ is satisfied if and only if u is simple and $u(\dot{\Sigma}) \cap \mathcal{U} \neq \varnothing$. Instead of expressing it this way in Theorem 8.3, we have stated precisely the condition that is needed in the proof—the reason to do it this way is that in some other contexts, statements analogous to Theorem 8.3 are true but there is no straightforward equivalence between simple curves and somewhere injective curves. The simplest example is that for (unpunctured) J-holomorphic disks with totally real boundary conditions, it is not true in general that every curve factors through a curve whose injective points are dense; see [**Laz00**, **Laz11**].

REMARK 8.7. Theorems 7.1 and 8.3 both admit easy extensions to the study of moduli spaces dependent on finitely many parameters. Concretely, suppose P is a smooth finite-dimensional manifold and $\{J_s\}_{s \in P}$ is a smooth family of almost complex structures satisfying the usual conditions. One can then define a **parametric moduli space**

$$\mathcal{M}(\{J_s\}) = \left\{ (s, u) \;\middle|\; s \in P,\; u \in \mathcal{M}(J_s) \right\}$$

and a notion of *parametric regularity* for pairs $(s, u) \in \mathcal{M}(\{J_s\})$, which is again an open condition, such that the space $\mathcal{M}^{\text{reg}}(\{J_s\})$ of parametrically regular elements will be an orbifold of dimension

$$\dim \mathcal{M}^{\text{reg}}(\{J_s\}) = \text{vir-dim}\, \mathcal{M}(J) + \dim P.$$

The proof of this is the same as in Lecture 7, except that the section $\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$ in §7.4 gets replaced by

$$\bar{\partial}_{\{J_s\}} : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \times P \to \mathcal{E}^{k-1,p,\delta} : (j, u, s) \mapsto du + J_s(u) \circ du \circ j,$$

for a bundle with fibers $\mathcal{E}^{k-1,p,\delta}_{(j,u,s)} := W^{k-1,p,\delta}(\overline{\text{Hom}}_{\mathbb{C}}((T\dot{\Sigma}, j), (u^*T\widehat{W}, J_s)))$, and we call $(s, [(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)]) \in \mathcal{M}(\{J_s\})$ **parametrically regular** if the linearization $D\bar{\partial}_{\{J_s\}}(j, u, s)$ is surjective. Notice that since $D\bar{\partial}_{\{J_s\}}(j, u, s)$ is the sum of $D\bar{\partial}_J(j, u)$ with an extra term defined on T_sP, every $(s, u) \in \mathcal{M}(\{J_s\})$ for which $u \in \mathcal{M}(J_s)$ is Fredholm regular is also parametrically regular. The converse however is false, as

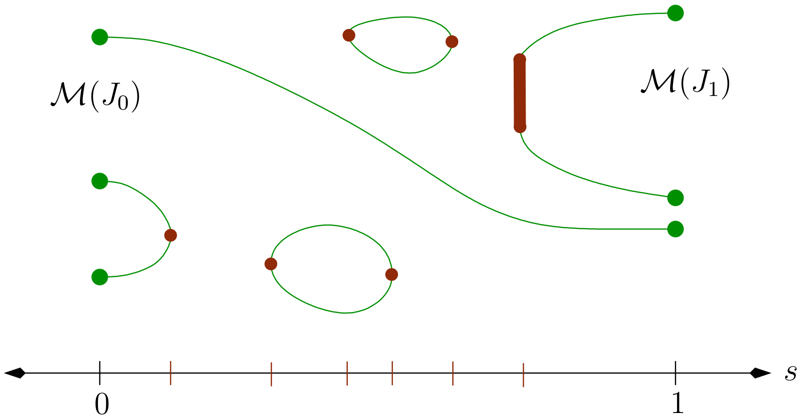

Figure 8.1. The picture shows a smooth parametric moduli space $\mathcal{M}(\{J_s\})$ with $P := [0,1]$ and its projection $\mathcal{M}(\{J_s\}) \to [0,1] : (s,u) \mapsto s$ in a case where vir-dim $\mathcal{M}(J_s) = 0$. The parametric moduli space is 1-dimensional and the spaces $\mathcal{M}(J_s)$ are regular and 0-dimensional for almost every $s \in [0,1]$, but this need not hold when s is a critical value of the projection. In the picture, one such space $\mathcal{M}(J_s)$ contains a 1-dimensional component consisting of non-regular curves, so its dimension differs from its virtual dimension.

$\mathcal{M}^{\text{reg}}(\{J_s\})$ can contain pairs (s,u) for which $u \notin \mathcal{M}^{\text{reg}}(J_s)$; these are precisely the critical points of the map

$$\mathcal{M}^{\text{reg}}(\{J_s\}) \to P : (s,u) \mapsto s. \tag{8.2}$$

The generalization of Theorem 8.3 to the parametric setting states that after generic perturbations of the family $\{J_s\}_{s \in P}$ that are fixed outside of some open subset $\mathcal{U} \subset W^{r_0}$ and fixed everywhere for s outside of some precompact open subset $\mathcal{V} \subset P$, all elements $(s,u) \in \mathcal{M}(\{J_s\})$ for which $s \in \mathcal{V}$ and u has an injective point mapping to $\mathcal{U}$ will be parametrically regular. The proof requires only minor modifications to the proof of Theorem 8.3, so we shall leave it as an exercise. The standard and most important example is $P = [0,1]$ with $\mathcal{V} = (0,1)$, meaning that we consider *generic homotopies* between two fixed almost complex structures J_0 and J_1. Figure 8.1 shows an example in which the moduli spaces $\mathcal{M}(J_s)$ each have virtual dimension zero and $\mathcal{M}(\{J_s\})$ is a 1-manifold. Since the projection (8.2) is not generally a submersion, there can exist critical values $s \in [0,1]$ at which $\mathcal{M}(J_s)$ fails to be a 0-dimensional manifold. In general these cannot be excluded by making generic choices of the homotopy, though in certain cases, one can exclude them using "automatic" transversality results, which give criteria for all J_s to be regular with no need for genericity (see §14.1).

Example 8.8. It is not hard to imagine situations in which transversality *must* fail generically for multiply covered curves. Suppose for instance that (W,ω) is an 8-dimensional symplectic manifold with compatible almost complex structure J_0, and $u_0 : S^2 \to W$ is a simple J_0-holomorphic sphere with no punctures and $[u_0] =$

$A \in H_2(W)$, where $c_1(A) = -1$. This means u_0 represents an element of a moduli space $\mathcal{M}_{0,0}(J_0, A)$ with

$$\text{vir-dim}\, \mathcal{M}_{0,0}(J_0, A) = 2 - 2g + 2c_1(A) = 0.$$

In particular if u_0 is regular and $\{J_s \in \mathcal{J}(\omega)\}_{s \in \mathbb{R}^k}$ is a smooth k-parameter family of compatible almost complex structures including J_0, then Remark 8.7 implies that a neighborhood of $(0, u_0)$ in the parametric moduli space $\mathcal{M}(\{J_s\}) = \{(s, u) \mid s \in \mathbb{R}^k,\ u \in \mathcal{M}_{0,0}(J_s, A)\}$ is a smooth k-dimensional manifold, and this will be true no matter how the family $\{J_s\}$ is chosen. But for each of the elements $(s, u) \in \mathcal{M}(\{J_s\})$ parametrized by a J-holomorphic map $u : (S^2 = \mathbb{C} \cup \{\infty\}, i) \to (W, J_s)$, there is also a double cover

$$u' : S^2 \to W : z \mapsto u(z^2),$$

with $[u'] = 2A$, so $u' \in \mathcal{M}_{0,0}(J_s, 2A)$ and

$$\text{vir-dim}\, \mathcal{M}_{0,0}(J_s, 2A) = 2 - 2g + 2c_1(2A) = -2.$$

Negative virtual dimension means that $\mathcal{M}_{0,0}(J_0, 2A)$ should be empty whenever Fredholm regularity is achieved, but this is clearly impossible, even generically, since elements of $\mathcal{M}_{0,0}(J_s, A)$ always have double covers belonging to $\mathcal{M}_{0,0}(J_s, 2A)$.

REMARK 8.9. The most common way to apply Theorem 8.3 is by setting $\mathcal{U}$ equal to the interior of W^{r_0}, so generic perturbations of J are allowed everywhere except on the regions where it is required to be $\mathbb{R}$-invariant. The theorem then achieves transversality for all simple curves that are not confined to the $\mathbb{R}$-invariant regions. We will show in §8.3 that transversality for all curves of the latter type can also be achieved by generic perturbations within the spaces $\mathcal{J}(\mathcal{H}_\pm)$ of compatible $\mathbb{R}$-invariant almost complex structures on the symplectizations $\mathbb{R} \times M_\pm$, hence generic choices in $\mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ do achieve transversality for all simple curves.

8.2.2. The universal moduli space. Our proof of Theorem 8.3 will roughly follow the paradigm that was described in §8.1, based on the Sard-Smale theorem. The first step is therefore to define a suitable Banach manifold of perturbations of the almost complex structure to incorporate into our functional-analytic setup. All known ways of doing this are in some sense non-ideal, e.g. one could take almost complex structures of class C^k or $W^{k,p}$, but this necessarily introduces non-smooth almost complex structures into the picture, with the consequence that the nonlinear Cauchy-Riemann operator has only finitely many derivatives. That is not the end of the world, and indeed, this is the approach taken in [**MS12**], but I will instead present an approach that was introduced by Floer in [**Flo88b**], in terms of what is now called the "Floer C_ϵ space". The idea is to work with a Banach manifold that continuously embeds into the space of smooth almost complex structures, so that the nonlinear Cauchy-Riemann operator will always be smooth. It's a nice trick, but the catch is that we obtain a space that is strictly smaller than the actual space of smooth almost complex structures we're interested in, and has a much stronger topology. The C_ϵ space should be viewed as a useful tool but not a deeply meaningful object—you might notice that while some of the intermediate results stated below depend on its (somewhat ad hoc) definition, Theorem 8.3 does not. This is due to a general trick described in §8.2.4 below for turning results about C_ϵ into results about C^∞.

As in the statement of Theorem 8.3, assume $\mathcal{U} \subset W^{r_0}$ is open and $J^{\text{fix}} \in \mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$. Let

$$\mathcal{J}_{\mathcal{U}} := \left\{ J \in \mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-) \;\middle|\; J = J^{\text{fix}} \text{ on } \widehat{W} \backslash \mathcal{U} \right\},$$

and choose any almost complex structure

$$J^{\text{ref}} \in \mathcal{J}_{\mathcal{U}}.$$

Informally, we can regard $\mathcal{J}_{\mathcal{U}}$ as an infinite-dimensional manifold[4] with tangent spaces

$$T_{J^{\text{ref}}}\mathcal{J}_{\mathcal{U}} = \left\{ Y \in \Gamma\left(\overline{\text{End}}_{\mathbb{C}}(T\widehat{W}, J^{\text{ref}})\right) \;\middle|\; Y|_{\widehat{W}\backslash\mathcal{U}} \equiv 0 \text{ and } \omega_\psi(\cdot, Y\cdot) + \omega_\psi(Y\cdot, \cdot) \equiv 0 \right\},$$

where the antilinearity of $Y \in T_{J^{\text{ref}}}\mathcal{J}_{\mathcal{U}}$ means that Y is tangent to the space of almost complex structures, and the condition relating it to ω_ψ means that these structures are compatible with ω_ψ. (This condition would be omitted if we had defined $\mathcal{J}_{\mathcal{U}}$ to be a subset of the larger space $\mathcal{J}_\tau(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$ instead of $\mathcal{J}(\omega_\psi, r_0, \mathcal{H}_+, \mathcal{H}_-)$; see Remark 8.4.) One can check that the map

$$Y \mapsto J_Y := \left(\mathbb{1} + \frac{1}{2}J^{\text{ref}}Y\right) J^{\text{ref}} \left(\mathbb{1} + \frac{1}{2}J^{\text{ref}}Y\right)^{-1}$$

sends any sufficiently C^0-small neighborhood of $0 \in T_{J^{\text{ref}}}\mathcal{J}_{\mathcal{U}}$ bijectively to a neighborhood of J^{ref} in $\mathcal{J}_{\mathcal{U}}$. We thus fix a sufficiently small constant $c > 0$ and define the space of "C_ϵ-small perturbations of J^{ref}" by

$$\mathcal{J}_{\mathcal{U}}^{\epsilon} := \left\{ J_Y \in \mathcal{J}_{\mathcal{U}} \;\middle|\; Y \in T_{J^{\text{ref}}}\mathcal{J}_{\mathcal{U}} \text{ with } \sum_{\ell=0}^{\infty} \epsilon_\ell \|Y\|_{C^\ell(\overline{\mathcal{U}})} < c \right\},$$

where $\boldsymbol{\epsilon} := (\epsilon_\ell)_{\ell=0}^{\infty}$ is a fixed sequence of positive numbers with $\epsilon_\ell \to 0$ as $\ell \to \infty$. The sum

$$\|Y\|_{C_\epsilon} := \sum_{\ell=0}^{\infty} \epsilon_\ell \|Y\|_{C^\ell(\overline{\mathcal{U}})}$$

defines a norm, and the space of smooth sections $Y \in T_{J^{\text{ref}}}\mathcal{J}_{\mathcal{U}}$ for which this norm is finite is then a separable Banach space; see Appendix B for a proof of this statement. This makes $\mathcal{J}_{\mathcal{U}}^{\epsilon}$ a separable and metrizable Banach manifold, as the map $J_Y \mapsto Y$ can be viewed as a chart identifying it with an open subset of the aforementioned Banach space. Not every $J \in \mathcal{J}_{\mathcal{U}}$ that is C^∞-close to J^{ref} belongs to $\mathcal{J}_{\mathcal{U}}^{\epsilon}$, but there is a continuous inclusion

$$\mathcal{J}_{\mathcal{U}}^{\epsilon} \hookrightarrow \mathcal{J}_{\mathcal{U}},$$

where the latter carries its usual C^∞-topology and $\mathcal{J}_{\mathcal{U}}^{\epsilon}$ carries the topology induced by the C_ϵ-norm. By a lemma due to Floer, choosing a sequence ϵ_ℓ that decays sufficiently fast makes $\mathcal{J}_{\mathcal{U}}^{\epsilon}$ large enough to contain perturbations in arbitrary directions

[4] Strictly speaking, it is a Fréchet manifold, but not a Banach manifold.

with arbitrarily small support near arbitrary points in $\mathcal{U}$; see Theorem B.6 in Appendix B for a precise version of this statement and its proof. We will assume from now on that a suitably fast decaying sequence has been fixed.

We now define for each $J \in \mathcal{J}_{\mathcal{U}}$ the set

$$\mathcal{M}^*(J) := \left\{ u \in \mathcal{M}(J) \;\middle|\; u \text{ has an injective point with image in } \mathcal{U} \right\},$$

which is an open subset of $\mathcal{M}(J)$ since $\mathcal{U}$ and the condition of being an injective point are both open. The corresponding **universal moduli space** is defined by

$$\mathcal{M}^*(\mathcal{J}^\epsilon_{\mathcal{U}}) := \left\{ (u, J) \;\middle|\; J \in \mathcal{J}^\epsilon_{\mathcal{U}} \text{ and } u \in \mathcal{M}^*(J) \right\}.$$

REMARK 8.10. The notion of convergence in $\mathcal{M}(J)$ defined in §6.4 also makes sense for a sequence of the form $u_\nu \in \mathcal{M}(J_\nu)$ where J_ν is a convergent sequence in $\mathcal{J}^\epsilon_{\mathcal{U}}$. In this way, $\mathcal{M}^*(\mathcal{J}^\epsilon_{\mathcal{U}})$ inherits a natural topology.

The use of the word "universal" is somewhat unfortunate, as $\mathcal{M}^*(\mathcal{J}^\epsilon_{\mathcal{U}})$ depends on many auxiliary choices such as J^{ref} and $(\epsilon_\ell)_{\ell=0}^\infty$. Nonetheless, $\mathcal{M}^*(\mathcal{J}^\epsilon_{\mathcal{U}})$ turns out to have exactly the properties we need for applying the Sard-Smale theorem—in particular, it is a smooth separable Banach manifold. To see this, we can adapt the functional-analytic setup from the proof of Theorem 7.1 and identify $\mathcal{M}^*(\mathcal{J}^\epsilon_{\mathcal{U}})$ locally with a quotient of the zero set of a smooth section of a Banach space bundle. Suppose $J_0 \in \mathcal{J}^\epsilon_{\mathcal{U}}$ and $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \Theta, u_0)] \in \mathcal{M}^*(J_0)$ where $u_0 : \dot{\Sigma} \to \widehat{W}$ has an injective point z_0 with $u_0(z_0) \in \mathcal{U}$. Choose a Teichmüller slice $\mathcal{T}$ through $(j_0, \Gamma \cup \Theta)$ as in Proposition 7.19 and consider the smooth section

$$\bar{\partial} : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \times \mathcal{J}^\epsilon_{\mathcal{U}} \to \mathcal{E}^{k-1,p,\delta} : (j, u, J) \mapsto du + J(u) \circ du \circ j,$$

where $\mathcal{E}^{k-1,p,\delta}$ is the obvious extension of our previous Banach space bundle to a bundle over $\mathcal{T} \times \mathcal{B}^{k,p,\delta} \times \mathcal{J}^\epsilon_{\mathcal{U}}$. We're assuming as before that $k \in \mathbb{N}$, $1 < p < \infty$, $kp > 2$, and $\delta > 0$ is small. A neighborhood of (u_0, J_0) in $\mathcal{M}^*(\mathcal{J}^\epsilon_{\mathcal{U}})$ can then be identified with a neighborhood of $[(j_0, u_0, J_0)]$ in

$$\bar{\partial}^{-1}(0)/G,$$

where $G := \mathrm{Aut}(\Sigma, j_0, \Gamma \cup \Theta)$ acts properly on $\bar{\partial}^{-1}(0)$ by $\varphi \cdot (j, u, J) := (\varphi^* j, u \circ \varphi, J)$. Since u_0 has an injective point, $\mathrm{Aut}(u_0)$ is trivial and the G-action on a neighborhood of (j_0, u_0, J_0) is therefore free. The main task is then to show that $\bar{\partial}^{-1}(0)$ is a smooth Banach manifold on some G-invariant neighborhood of (j_0, u_0, J_0), as elliptic regularity will imply as in Lemma 7.29 that the G-action on this neighborhood is smooth, so its quotient becomes a smooth Banach manifold as well. This will follow from the implicit function theorem if we can show that the linearization

$$D\bar{\partial}(j_0, u_0, J_0) : T_{j_0}\mathcal{T} \oplus T_{u_0}\mathcal{B}^{k,p,\delta} \oplus T_{J_0}\mathcal{J}^\epsilon_{\mathcal{U}} \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0,J_0)}$$

is surjective with a bounded right inverse. Note that $D\bar{\partial}(j_0, u_0, J_0)$ is not a Fredholm operator due to the infinite-dimensional summand $T_{J_0}\mathcal{J}^\epsilon_{\mathcal{U}}$ in its domain, thus in contrast to the situation in Theorem 7.1, it is no longer obvious whether surjectivity implies the existence of a bounded right inverse. However, the restriction of $D\bar{\partial}(j_0, u_0, J_0)$ to the factor $T_{u_0}\mathcal{B}^{k,p,\delta}$ is Fredholm, so the first part of the following exercise shows that surjectivity is sufficient.

EXERCISE 8.11. Given Banach spaces X, Y and Z, a Fredholm operator $\mathbf{T} : X \to Y$ and another bounded linear operator $\mathbf{A} : Z \to Y$, consider the operator $\mathbf{L} : X \oplus Z \to Y : (x, z) \mapsto \mathbf{T}x + \mathbf{A}z$. Prove:

(a) The kernel of $\mathbf{L}$ has a closed complement in $X \oplus Z$. *Hint: Write $X = V \oplus K$ and $Y = W \oplus C$ with $K = \ker \mathbf{T}$ and $C \cong \operatorname{coker} \mathbf{T}$, so that $V \xrightarrow{\mathbf{T}} W$ is an isomorphism. Compute $\ker \mathbf{L}$ in terms of these splittings, and don't forget that C is finite dimensional.*

(b) The image of $\mathbf{L}$ is closed in Y. *Hint: Consider the restriction of $\mathbf{L}$ to the closed complement of $\ker \mathbf{L}$ from part (a).*

The next lemma is now the main technical step in the proof of Theorem 8.3.

LEMMA 8.12. *The operator*

$$W^{k,p,\delta}(u_0^* T\widehat{W}) \oplus T_{J_0}\mathcal{J}^\epsilon_{\mathcal{U}} \xrightarrow{\mathbf{L}} W^{k-1,p,\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^* T\widehat{W}))$$

$$(\eta, Y) \mapsto D\bar{\partial}(j_0, u_0, J_0)(0, \eta, Y) = \mathbf{D}_{u_0}\eta + Y(u_0) \circ du_0 \circ j_0$$

is surjective for every $k \in \mathbb{N}$, $p \in (1, \infty)$ and $\delta > 0$ sufficiently small.

PROOF. Consider first the case $k = 1$,[5] so we are looking at a bounded linear map

$$W^{1,p,\delta}(u_0^* T\widehat{W}) \oplus T_{J_0}\mathcal{J}^\epsilon_{\mathcal{U}} \xrightarrow{\mathbf{L}} L^{p,\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^* T\widehat{W})),$$

and let us fix any value for $\delta > 0$ such that the operator $\mathbf{D}_\delta := \mathbf{D}_{u_0}|_{W^{1,p,\delta}(u_0^* T\widehat{W})} : W^{1,p,\delta}(u_0^* T\widehat{W}) \to L^{p,\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^* T\widehat{W}))$ is Fredholm (cf. Lemma 7.6). Observe that the dual of any space of sections of class $L^{p,\delta}$ can be identified with sections of class $L^{q,-\delta}$ for $\frac{1}{p} + \frac{1}{q} = 1$ (recall Remark 7.4). Indeed, choosing a suitable L^2-pairing defines a bounded bilinear map

$$\langle\ ,\ \rangle_{L^2} : L^{p,\delta} \oplus L^{q,-\delta} \to \mathbb{R}, \tag{8.3}$$

and one can use isomorphisms of the form $L^p \to L^{p,\delta} : \eta \mapsto e^f \eta$ as in the proof of Lemma 7.6 to prove $(L^{p,\delta})^* \cong L^{q,-\delta}$ as a corollary of the standard fact that $(L^p)^* \cong L^q$. With this understood, the fact that $\mathbf{D}_\delta : W^{1,p,\delta} \to L^{p,\delta}$ is Fredholm implies via Exercise 8.11 that $\mathbf{L}$ has closed image. Thus if it is not surjective, the Hahn-Banach theorem provides a nontrivial element $\theta \in L^{q,-\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^* T\widehat{W}))$ that annihilates its image under the pairing (8.3), which amounts to the two conditions

$$\begin{aligned} \langle \mathbf{D}_\delta \eta, \theta \rangle_{L^2} &= 0 \text{ for all } \eta \in W^{1,p,\delta}(u_0^* T\widehat{W}), \\ \langle Y(u_0) \circ du_0 \circ j_0, \theta \rangle_{L^2} &= 0 \text{ for all } Y \in T_{J_0}\mathcal{J}^\epsilon_{\mathcal{U}}. \end{aligned} \tag{8.4}$$

The first relation is valid in particular for all smooth sections η with compact support and thus means that θ is a weak solution to the formal adjoint equation $\mathbf{D}_\delta^* \theta = 0$; applying elliptic regularity and the similarity principle, θ is therefore smooth and has only isolated zeroes. We will see however that this contradicts the second relation as long as there exists an injective point $z_0 \in \dot{\Sigma}$ with $u_0(z_0) \in \mathcal{U}$. Indeed, since the set of injective points with this property is open and zeroes of θ are isolated,

[5] Note that since Lemma 8.12 is a purely linear result, it does not require the assumption $kp > 2$.

let us assume without loss of generality that $\theta(z_0) \neq 0$. Then since $du_0(z_0) \neq 0$, one can apply a standard lemma from symplectic linear algebra (see Lemma 8.13 below) to find a smooth section $Y \in T_{J_0}\mathcal{J}_\mathcal{U}$ whose value at $u_0(z_0)$ is chosen such that $Y(u_0) \circ du_0 \circ j_0 = \theta$ at z_0, implying that the pointwise inner product $\langle Y(u_0) \circ du_0 \circ j_0, \theta\rangle$ is positive in some neighborhood of z_0. By Theorem B.6, one can then multiply a small perturbation of Y by a bump function to produce a section (still denoted by Y) of class C_ϵ so that the pointwise inner product of $Y(u_0) \circ du_0 \circ j_0$ with θ is still positive near z_0 but vanishes everywhere else; note that this requires the assumption $u_0^{-1}(u_0(z_0)) = \{z_0\}$, so that the value of Y near $u_0(z_0)$ affects the value of $Y(u_0) \circ du_0 \circ j_0$ near z_0 but nowhere else. This contradicts the second condition in (8.4) and thus completes the proof for $k = 1$.

In the general case, suppose $\alpha \in W^{k-1,p,\delta}(\overline{\mathrm{Hom}}_\mathbb{C}(T\dot{\Sigma}, u_0^*T\widehat{W}))$. Then α is also of class $L^{p,\delta}$, so surjectivity in the $k = 1$ case implies the existence of $\eta \in W^{1,p,\delta}$ and $Y \in T_{J_0}\mathcal{J}_\mathcal{U}^\epsilon$ with $\mathbf{D}_\delta\eta + Y(u_0) \circ du_0 \circ j_0 = \alpha$. Since $Y(u_0) \circ du_0 \circ j_0$ is smooth with compact support, one can then use elliptic regularity to show $\eta \in W^{k,p,\delta}$, and this proves surjectivity for arbitrary $k \in \mathbb{N}$ and $p \in (1,\infty)$. □

The choice of the bump function $Y \in T_{J_0}\mathcal{J}_\mathcal{U}$ in the proof above required an elementary but slightly non-obvious lemma from linear algebra. This is the only point in the argument where the symplectic structure on W^{r_0} makes any difference, and only if we are requiring perturbed almost complex structures to be *compatible* with ω_ψ (rather than just tame), as this condition meaningfully shrinks the space of available perturbations Y along J_0. But the lemma below shows that this space of perturbations is still large enough. Recall that on any symplectic vector space (V,ω) with compatible complex structure J, one can choose a basis to identify J with i and ω with the standard structure ω_{std}. The linear maps Y that anticommute with i and satisfy $\omega_{\mathrm{std}}(Yv,w) + \omega_{\mathrm{std}}(v,Yw) = 0$ for all $v,w \in V$ are then precisely the *symmetric* matrices that are complex antilinear.

LEMMA 8.13. *For any nonzero vectors $v,w \in \mathbb{R}^{2n}$, there exists a symmetric matrix Y that anticommutes with i and satisfies $Yv = w$.*

PROOF. We borrow the proof directly from [MS12, Lemma 3.2.2] and simply state a formula for Y:

$$Y = \frac{1}{|v|^2}\left(wv^T + vw^T + i\left(wv^T + vw^T\right)i\right) - \frac{1}{|v|^4}\left(\langle w,v\rangle\left(vv^T + ivv^Ti\right) - \langle w,iv\rangle\left(ivv^T - vv^Ti\right)\right),$$

where $\langle\ ,\ \rangle$ denotes the standard real inner product on $\mathbb{R}^{2n} = \mathbb{C}^n$. □

COROLLARY 8.14. *The universal moduli space $\mathcal{M}^*(\mathcal{J}_\mathcal{U}^\epsilon)$ is a smooth, separable and metrizable Banach manifold, and the projection $\mathcal{M}^*(\mathcal{J}_\mathcal{U}^\epsilon) \to \mathcal{J}_\mathcal{U}^\epsilon : (u,J) \mapsto J$ is smooth.*

PROOF. Assume $kp > 2$ and $\delta > 0$ is sufficiently small. The section $\bar{\partial} : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \times \mathcal{J}_\mathcal{U}^\epsilon \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0,J_0)}$ is equivariant with respect to the action of $G :=$

$\mathrm{Aut}(\Sigma, j_0, \Theta)$, and its linearization at (j_0, u_0, J_0) is the sum of the operator considered in Lemma 8.12 with extra terms defined on a finite-dimensional subspace $T_{j_0}\mathcal{T} \oplus V_\Gamma$, and is therefore surjective. In light of equivariance and Exercise 8.11, it follows that the linearization of $\bar{\partial}$ has a bounded right inverse at every point in some G-invariant neighborhood of (j_0, u_0, J_0) in $\bar{\partial}^{-1}(0)$. The implicit function theorem then implies that that neighborhood is a smooth submanifold of $\mathcal{T} \times \mathcal{B}^{k,p,\delta} \times \mathcal{J}^\epsilon_{\mathcal{U}}$, and the G-action on that neighborhood is smooth by elliptic regularity as in Lemma 7.29. The projection map

$$\bar{\partial}^{-1}(0) \to \mathcal{J}^\epsilon_{\mathcal{U}} : (j, u, J) \mapsto J$$

is also smooth on this neighborhood since it is the restriction to a smooth submanifold of the obviously smooth projection map $\mathcal{T} \times \mathcal{B}^{k,p,\delta} \times \mathcal{J}^\epsilon_{\mathcal{U}} \to \mathcal{J}^\epsilon_{\mathcal{U}}$. Since G acts freely and properly on $\bar{\partial}^{-1}(0)$, the neighborhood of $[(j_0, u_0, J_0)]$ in $\bar{\partial}^{-1}(0)/G$ then inherits a smooth Banach manifold structure for which the projection is still smooth, and this quotient is identified locally with $\mathcal{M}^*(\mathcal{J}^\epsilon_{\mathcal{U}})$. Smoothness of transition maps is shown via the same regularity arguments as in the proof of Theorem 7.1. ☐

8.2.3. Applying the Sard-Smale theorem. We claim now that

$$\mathcal{M}^*(\mathcal{J}^\epsilon_{\mathcal{U}}) \to \mathcal{J}^\epsilon_{\mathcal{U}} : (u, J) \mapsto J \tag{8.5}$$

is a smooth nonlinear Fredholm map, i.e. its derivative at every point is a Fredholm operator. Using the local identification of $\mathcal{M}^*(\mathcal{J}^\epsilon_{\mathcal{U}})$ with $\bar{\partial}^{-1}(0)/G$ as in the proof of Corollary 8.14 and lifting the projection to $\bar{\partial}^{-1}(0)$, the derivative of $\bar{\partial}^{-1}(0) \to \mathcal{J}^\epsilon_{\mathcal{U}}$ at (j_0, u_0, J_0) takes the form

$$\ker D\bar{\partial}(j_0, u_0, J_0) \to T_{J_0}\mathcal{J}^\epsilon_{\mathcal{U}} : (y, \eta, Y) \mapsto Y.$$

The Fredholm property for this projection is a consequence of the Fredholm property for $D\bar{\partial}(j_0, u_0)$ via Lemma 8.2, which also implies that the projection is surjective if and only if $D\bar{\partial}(j_0, u_0)$ is surjective, i.e. if u_0 is a Fredholm regular curve. Applying the Sard-Smale theorem to the map (8.5), we conclude:

COROLLARY 8.15. *There exists a comeager subset*

$$\mathcal{J}^{\epsilon,\mathrm{reg}}_{\mathcal{U}} \subset \mathcal{J}^\epsilon_{\mathcal{U}}$$

such that for all $J \in \mathcal{J}^{\epsilon,\mathrm{reg}}_{\mathcal{U}}$, every $u \in \mathcal{M}^*(J)$ is Fredholm regular. ☐

8.2.4. From C_ϵ to C^∞. The arguments above would constitute a proof of Theorem 8.3 if we were allowed to replace the space of smooth almost complex structures $\mathcal{J}_{\mathcal{U}}$ with the space $\mathcal{J}^\epsilon_{\mathcal{U}}$ of C_ϵ-small perturbations of J^{ref}. Let us *define*

$$\mathcal{J}^{\mathrm{reg}}_{\mathcal{U}} \subset \mathcal{J}_{\mathcal{U}}$$

to be the space of all $J \in \mathcal{J}_{\mathcal{U}}$ with the property that all curves in $\mathcal{M}^*(J)$ are Fredholm regular. The theorem claims that this set is comeager in $\mathcal{J}_{\mathcal{U}}$. We can already see at this point that it is dense: indeed, the Baire category theorem implies that $\mathcal{J}^{\epsilon,\mathrm{reg}}_{\mathcal{U}}$ is dense in $\mathcal{J}^\epsilon_{\mathcal{U}}$, so in particular there exists a sequence $J_\nu \in J^{\epsilon,\mathrm{reg}}_{\mathcal{U}}$ that converges to J^{ref} in the C_ϵ-topology and therefore also in the C^∞-topology. The choice of $J^{\mathrm{ref}} \in \mathcal{J}_{\mathcal{U}}$ in this discussion was arbitrary, so this proves density.

To prove that $\mathcal{J}^{\mathrm{reg}}_{\mathcal{U}}$ is not only dense but also contains a countable intersection of *open* and dense sets in $\mathcal{J}_{\mathcal{U}}$, we can adapt an argument originally due to Taubes.

The idea is to present $\mathcal{M}^*(J)$ as a countable union of compact subsets $\mathcal{M}^*_N(J)$ for $N \in \mathbb{N}$, and thus present $\mathcal{J}^{\text{reg}}_{\mathcal{U}}$ as a corresponding countable intersection of spaces $\mathcal{J}^{\text{reg},N}_{\mathcal{U}}$ that achieve regularity only for the elements in $\mathcal{M}^*_N(J)$. The compactness of $\mathcal{M}^*_N(J)$ will then permit us to prove that $\mathcal{J}^{\text{reg},N}_{\mathcal{U}}$ is not only dense but also open.

The definition of $\mathcal{M}^*_N(J)$ is motivated in part by the knowledge that spaces of J-holomorphic curves have natural compactifications. We have not yet discussed the compactification $\overline{\mathcal{M}}(J)$ of $\mathcal{M}(J)$, but we have covered enough of the analytical techniques behind this construction to suffice for most details of the present discussion.

LEMMA 8.16. *For each $J \in \mathcal{J}_{\mathcal{U}}$, there exists a nested sequence of subsets $\mathcal{M}^*_1(J) \subset \mathcal{M}^*_2(J) \subset \ldots \subset \mathcal{M}^*(J)$ such that:*

(1) $\bigcup_{N\in\mathbb{N}} \mathcal{M}^*_N(J) = \mathcal{M}^*(J)$;
(2) For any compact subset $\mathcal{K} \in \mathcal{J}_{\mathcal{U}}$ and every $N \in \mathbb{N}$, the set

$$\mathcal{M}^*_N(\mathcal{K}) := \{(u, J) \mid J \in \mathcal{K} \text{ and } u \in \mathcal{M}^*_N(J)\}$$

with its natural topology (cf. Remark 8.10) is compact.

PROOF. We start by finding a nested sequence in $\mathcal{J}(\Sigma)$ that proves a similar statement about the moduli space of Riemann surface $\mathcal{M}_{g,\ell}$, where $\ell := \#(\Gamma \cup \Theta)$. This depends mainly on the fact that $\mathcal{M}_{g,\ell}$ is a finite-dimensional orbifold, so in particular it is locally compact and second countable. Fix a model surface Σ of genus g along with disjoint sets of punctures $\Gamma \subset \Sigma$ and marked points $\Theta \subset \Sigma$, abbreviating $\Theta' := \Gamma \cup \Theta$. For each $j \in \mathcal{J}(\Sigma)$, choose a Teichmüller slice $\mathcal{T}_j \subset \mathcal{J}(\Sigma)$ through (j, Θ'), and let $\mathcal{V}_j \subset \mathcal{T}_j$ denote a compact neighborhood of j. Then the image of $\mathcal{V}_j$ under the quotient projection $\pi : \mathcal{J}(\Sigma) \to \mathcal{J}(\Sigma)/\operatorname{Diff}(\Sigma, \Theta') = \mathcal{M}_{g,\ell}$ is a compact neighborhood of $[j]$ in $\mathcal{M}_{g,\ell}$. The union of these for all $j \in \mathcal{J}(\Sigma)$ therefore forms an open cover of $\mathcal{M}_{g,\ell}$, which has a countable subcover, i.e.

$$\mathcal{M}_{g,\ell} = \bigcup_{N=1}^{\infty} \pi(\mathcal{V}_{j_N})$$

for some sequence $j_1, j_2, j_3, \ldots \in \mathcal{J}(\Sigma)$. For each $N \in \mathbb{N}$, define

$$\mathcal{J}_N(\Sigma) := \bigcup_{i=1}^{N} \mathcal{V}_{j_i} \subset \mathcal{J}(\Sigma).$$

We now have a nested sequence $\mathcal{J}_1(\Sigma) \subset \mathcal{J}_2(\Sigma) \subset \ldots \subset \mathcal{J}(\Sigma)$ of compact subsets whose union projects surjectively to $\mathcal{M}_{g,\ell}$ under the quotient projection.

With this preparation in place, fix Riemannian metrics on $\widehat{W}$ and $\dot{\Sigma}$ with translation-invariance on the cylindrical ends and use $\operatorname{dist}(\ ,\ )$ to denote the distance functions. For $N \in \mathbb{N}$ and $J \in \mathcal{J}_{\mathcal{U}}$, we define $\mathcal{M}^*_N(J) \subset \mathcal{M}^*(J)$ to be the set of equivalence classes admitting representatives $(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)$ with the following properties:

- $j \in \mathcal{J}_N(\Sigma)$
- $\sup_{z\in\dot{\Sigma}} |du(z)| \leqslant N$;

- There exists $z_0 \in \dot\Sigma$ such that
$$\operatorname{dist}(u(z_0), \widehat{W}\backslash\mathcal{U}) \geqslant \frac{1}{N}, \qquad |du(z_0)| \geqslant \frac{1}{N}, \qquad \operatorname{dist}(z_0, \Gamma) \geqslant \frac{1}{N},$$
and
$$\inf_{z \in \dot\Sigma\backslash\{z_0\}} \frac{\operatorname{dist}(u(z_0), u(z))}{\operatorname{dist}(z_0, z)} \geqslant \frac{1}{N}.$$

Observe that since every $u \in \mathcal{M}^*(J)$ is asymptotically cylindrical, its first derivative is globally bounded and it therefore belongs to $\mathcal{M}^*_N(J)$ for all N sufficiently large. Conversely, every $u \in \mathcal{M}^*_N(J)$ has an injective point lying a fixed distance away from the punctures and whose image lies a fixed distance from the complement of $\mathcal{U}$. Now for any convergent sequence $J_\nu \to J$ in $\mathcal{J}_{\mathcal{U}}$, arbitrary sequences $u_\nu \in \mathcal{M}^*_N(J_\nu)$ have representatives of the form $(\Sigma, j_\nu, \Gamma^+, \Gamma^+, \Theta, u_\nu)$ with $j_\nu \in \mathcal{J}_N(\Sigma)$ and u_ν satisfying the two other two conditions above. Since $\mathcal{J}_N(\Sigma)$ is compact, we can pass to a subsequence and assume $j_\nu \to j \in \mathcal{J}_N(\Sigma)$. The J_ν-holomorphic maps $u_\nu : (\dot\Sigma, j_\nu) \to (\widehat{W}, J_\nu)$ all have intersections with the precompact subset $\mathcal{U} \subset W^{ro}$ on some compact subset of $\dot\Sigma$, and they have uniformly bounded first derivatives, thus they are uniformly C^1-bounded on compact subsets, and therefore also $W^{1,p}_{\text{loc}}$-bounded for $p > 2$. By Corollary 2.25, we can now pass to a further subsequence so that u_ν converges in $C^\infty_{\text{loc}}(\dot\Sigma)$ to a J-holomorphic map $u : (\dot\Sigma, j) \to (\widehat{W}, J)$, which necessarily also satisfies the closed conditions in the definition of the space $\mathcal{M}^*_N(J)$. One can apply further compactness arguments as in Lemma 9.20 to show that in this situation, u is also asymptotically cylindrical and the continuous extensions $\bar u_\nu : \hat\Sigma \to \widehat{W}$ are C^0-convergent to $\bar u : \hat\Sigma \to \widehat{W}$, hence $u \in \mathcal{M}^*_N(J)$ and (u_ν, J_ν) converges to (u, J). □

To complete the proof of Theorem 8.3, define
$$\mathcal{J}^{\text{reg},N}_{\mathcal{U}} \subset \mathcal{J}_{\mathcal{U}}$$
for each $N \in \mathbb{N}$ as the set of all $J \in \mathcal{J}_{\mathcal{U}}$ for which every element of $\mathcal{M}^*_N(J)$ is Fredholm regular.

LEMMA 8.17. *For every $N \in \mathbb{N}$, $\mathcal{J}^{\text{reg},N}_{\mathcal{U}}$ is open and dense in $\mathcal{J}_{\mathcal{U}}$.*

PROOF. Density is immediate, since we've seen already that every $J \in \mathcal{J}_{\mathcal{U}}$ admits a C^∞-small perturbation that achieves regularity for all curves in $\bigcup_{N\in\mathbb{N}} \mathcal{M}^*_N(J)$. For openness, suppose the contrary: then there exists $J_\infty \in \mathcal{J}^{\text{reg},N}_{\mathcal{U}}$ and a sequence $J_\nu \in \mathcal{J}_{\mathcal{U}}\backslash\mathcal{J}^{\text{reg},N}_{\mathcal{U}}$ with $J_\nu \to J_\infty$ in the C^∞-topology. There must also exist a sequence of curves $u_\nu \in \mathcal{M}^*_N(J_\nu)$ that are not Fredholm regular. But then a subsequence of u_ν converges to an element $u_\infty \in \mathcal{M}^*_N(J_\infty)$, which must be Fredholm regular. The latter is an open condition and thus gives a contradiction to the assumption that u_ν is not regular for all ν. □

PROOF OF THEOREM 8.3. Since $\mathcal{M}^*(J) = \bigcup_{N\in\mathbb{N}} \mathcal{M}^*_N(J)$, we have
$$\mathcal{J}^{\text{reg}}_{\mathcal{U}} = \bigcap_{N\in\mathbb{N}} \mathcal{J}^{\text{reg},N}_{\mathcal{U}},$$
which is a countable intersection of open and dense sets. □

8.3. Generic transversality in symplectizations

8.3.1. Main results in the $\mathbb{R}$-invariant setting. If M is a closed manifold with a stable Hamiltonian structure $\mathcal{H}$, one can easily adapt Theorem 8.3 to say that for any symplectic structure of the usual form ω_f on $\mathbb{R} \times M$, generic C^∞-small ω_f-compatible perturbations of any given $J \in \mathcal{J}(\mathcal{H})$ in an open subset $\mathcal{U} \subset \mathbb{R} \times M$ with compact closure will achieve regularity for all curves that map an injective point into $\mathcal{U}$. The trouble with this statement is that after the perturbation, J no longer belongs to $\mathcal{J}(\mathcal{H})$: in particular it cannot be $\mathbb{R}$-invariant, nor can we expect it to preserve the subbundle $\xi = \ker \lambda$ or map ∂_r to the Reeb vector field. Requiring these conditions confines us to a much smaller space of perturbed almost complex structures than in Theorem 8.3, and it is no longer obvious whether this space will be large enough to achieve transversality.

The following statement refers to a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ with induced hyperplane distribution $\xi = \ker \lambda$ and Reeb vector field R, and we denote by

$$\pi_\xi : T(\mathbb{R} \times M) \to \xi$$

the projection along the trivial subbundle generated by ∂_r and R. We will sometimes view λ and $d\lambda$ as forms on $\mathbb{R} \times M$ by pulling them back via the projection $\mathbb{R} \times M \to M$; with this in mind, the **$d\lambda$-complement** of $\xi \subset T(\mathbb{R} \times M)$ at a point $p \in \mathbb{R} \times M$ will be denoted by

$$\xi_p^{\perp d\lambda} := \left\{ X \in T_p(\mathbb{R} \times M) \;\middle|\; d\lambda(X, \cdot)|_{\xi_p} = 0 \right\} \subset T_p(\mathbb{R} \times M).$$

We assume as usual that $\mathcal{M}(J) = \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ denotes a moduli space of asymptotically cylindrical J-holomorphic curves with a fixed genus g and number of marked points m, representing a fixed relative homology class A and asymptotic to fixed sets of Reeb orbits $\gamma_i^\pm$ at its positive and negative punctures.

THEOREM 8.18. *Suppose M is a closed $(2n-1)$-dimensional manifold carrying a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$, $J^{\text{fix}} \in \mathcal{J}(\mathcal{H})$,*

$$\mathcal{U} \subset M$$

is an open subset, and the orbits $\gamma_i^\pm$ in the definition of the moduli space $\mathcal{M}(J) = \mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ are all nondegenerate. Then there exists a comeager subset

$$\mathcal{J}_{\mathcal{U}}^{\text{reg}} \subset \left\{ J \in \mathcal{J}(\mathcal{H}) \;\middle|\; J = J^{\text{fix}} \text{ on } \mathbb{R} \times (M \backslash \mathcal{U}) \right\}$$

such that for every $J \in \mathcal{J}_{\mathcal{U}}^{\text{reg}}$, every curve $u \in \mathcal{M}(J)$ with a representative $u : \dot{\Sigma} \to \mathbb{R} \times M$ that has an injective point $z \in \dot{\Sigma}$ satisfying

(i) $u(z) \in \mathbb{R} \times \mathcal{U}$, *and*
(ii) $\operatorname{im} du(z) \cap \xi_{u(z)}^{\perp d\lambda} = \{0\}$

is Fredholm regular.

This result is applied most frequently with $\mathcal{U} = M$, in which case the condition $u(z) \in \mathbb{R} \times \mathcal{U}$ is vacuous. Since $d\lambda(R, \cdot) \equiv 0$, the second condition on the injective point z can be rephrased by asking for the linear map

$$d\lambda(\pi_\xi \circ Tu(X), \cdot)|_{\xi_{u(z)}} : \xi_{u(z)} \to \mathbb{R}$$

to be nontrivial for every nonzero $X \in T_z\dot{\Sigma}$. If λ is contact, then this is immediate whenever $\pi_\xi \circ Tu(X) \neq 0$ since $d\lambda|_\xi$ is nondegenerate, and the condition $\pi_\xi \circ Tu(X) \neq 0$ is also easy to achieve:

PROPOSITION 8.19. *If $J \in \mathcal{J}(\mathcal{H})$, then for any connected J-holomorphic curve $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$, the section*

$$\pi_\xi \circ du \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(T\dot{\Sigma}, u^*\xi))$$

either is identically zero or has only isolated zeroes.

As you might guess, this result is a consequence of the similarity principle; see §8.3.2 for a proof. Notice that if $\pi_\xi \circ du \equiv 0$, then u is everywhere tangent to the vector fields ∂_r and R, so if it is asymptotically cylindrical, then it can only be a trivial cylinder or a cover thereof.

PROPOSITION 8.20. *All trivial cylinders over nondegenerate Reeb orbits have index 0 and are Fredholm regular.*

PROOF. Let $u_\gamma : \mathbb{R} \times S^1 \to \mathbb{R} \times M$ denote the trivial cylinder over an orbit $\gamma : S^1 \to M$. The virtual dimension formula proved in Lecture 7 gives

$$\begin{aligned}\mathrm{ind}(u_\gamma) &= (n-3)\chi(\mathbb{R} \times S^1) + 2c_1^\tau(u_\gamma^*T(\mathbb{R} \times M)) + \mu_{\mathrm{CZ}}^\tau(\gamma) - \mu_{\mathrm{CZ}}^\tau(\gamma) \\ &= 2c_1^\tau(u_\gamma^*T(\mathbb{R} \times M)) = 0\end{aligned}$$

since the asymptotic trivialization τ has an obvious extension to a global trivialization of $u_\gamma^*\xi$, and $u_\gamma^*T(\mathbb{R} \times M)$ is globally the direct sum of the latter with the trivial line bundle spanned by ∂_r and R. Using this splitting, the linearized Cauchy-Riemann operator $\mathbf{D}_{u_\gamma}$ can be identified with $\bar{\partial} \oplus (\partial_s - \mathbf{A}_\gamma)$, where

$$\bar{\partial} = \partial_s + i\partial_t : W^{k,p,\delta}(\mathbb{R} \times S^1, \mathbb{C}) \oplus V_\Gamma \to W^{k-1,p,\delta}(\mathbb{R} \times S^1, \mathbb{C})$$

and

$$\partial_s - \mathbf{A}_\gamma : W^{k,p,\delta}(u_\gamma^*\xi) \to W^{k-1,p,\delta}(u_\gamma^*\xi).$$

Here we are assuming without loss of generality that V_Γ is a complex 2-dimensional space of smooth sections of the trivial line bundle spanned by ∂_r and R that are constant near infinity, and we are identifying this with a space of smooth complex-valued functions on $\mathbb{R} \times S^1$. Nondegeneracy implies that $\partial_s - \mathbf{A} : W^{k,p} \to W^{k-1,p}$ is an isomorphism, recall Theorem 4.14 in Lecture 4. Using weight functions as in the proof of Lemma 7.6 to define isomorphisms between $W^{k,p,\delta}$ and $W^{k,p}$, one can identify $\partial_s - \mathbf{A}_\gamma : W^{k,p,\delta} \to W^{k-1,p,\delta}$ with a small perturbation of the same operator $W^{k,p} \to W^{k-1,p}$, hence it is also an isomorphism for $\delta > 0$ sufficiently small. To see that $\bar{\partial} : W^{k,p,\delta} \oplus V_\Gamma \to W^{k-1,p,\delta}$ is also surjective, observe first that its index is 2; this follows from our calculation of $\mathrm{ind}(u_\gamma)$ and corresponds to the fact that $\dim \mathrm{Aut}(\mathbb{R} \times S^1, i) = 2$. The kernel of this operator consists of bounded holomorphic $\mathbb{C}$-valued functions on $\mathbb{R} \times S^1$, so it is precisely the real 2-dimensional space of constant functions, implying

$$\dim \mathrm{coker}(\bar{\partial}) = \dim \ker(\bar{\partial}) - \mathrm{ind}(\bar{\partial}) = 2 - 2 = 0,$$

so $\mathbf{D}_{u_\gamma}$ is surjective. ☐

COROLLARY 8.21. *For any contact form α on a closed manifold M, there exists a comeager subset $\mathcal{J}^{\mathrm{reg}}(\alpha) \subset \mathcal{J}(\alpha)$ such that for every $J \in \mathcal{J}^{\mathrm{reg}}(\alpha)$, all somewhere injective asymptotically cylindrical J-holomorphic curves in $\mathbb{R} \times M$ are Fredholm regular.* □

Note that in the setting of Corollary 8.21, a curve that is not a cover of a trivial cylinder always belongs to a smooth 1-parameter family of curves related to each other by $\mathbb{R}$-translation, so that the kernel of the linearized Cauchy-Riemann operator automatically has kernel of dimension at least 1. This precludes Fredholm regularity for curves of index 0, thus:

COROLLARY 8.22. *If α is a contact form and $J \in \mathcal{J}^{\mathrm{reg}}(\alpha)$, then all simple asymptotically cylindrical J-holomorphic curves $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ other than trivial cylinders satisfy*

$$\mathrm{ind}(u) \geqslant 1.$$

□

The following example shows that outside of the contact case, the nonvanishing of $\pi_\xi \circ du$ does not suffice on its own for achieving transversality:

EXAMPLE 8.23 (cf. §6.3.2). Assume (W, Ω) is a closed symplectic manifold of dimension $2n-2$ with a periodic time-dependent Hamiltonian $H : S^1 \times W \to \mathbb{R}$, and $M := S^1 \times W$ is assigned the stable Hamiltonian structure $(\omega, \lambda) := (\Omega + dt \wedge dH, dt)$. A choice of $J \in \mathcal{J}(\mathcal{H})$ is then equivalent to a choice of t-dependent family of Ω-compatible almost complex structures $\{J_t \in \mathcal{J}(W, \Omega)\}_{t \in S^1}$, and for any fixed $t \in S^1$ and $s \in \mathbb{R}$, J_t-holomorphic curves $v : (\Sigma, j) \to (W, J_t)$ give rise to J-holomorphic curves

$$u : (\Sigma, j) \to (\mathbb{R} \times M, J) : z \mapsto (s, t, v(z)).$$

In particular, when $n = 2$ one can consider the example where $W = \Sigma$ is a closed surface, so curves of this form exist for any choice of $J \in \mathcal{J}(\mathcal{H})$, no matter how generic (remember that the domain complex structure j is arbitrary, it is not fixed in advance). If Σ has genus g and the map $v : \Sigma \to \Sigma$ has degree 1, then since u has no punctures and satisfies $c_1([u]) = c_1(u^*T(\mathbb{R} \times S^1 \times \Sigma)) = c_1(T\Sigma) = \chi(\Sigma)$, the index of u is

$$\mathrm{ind}(u) = (n-3)\chi(\Sigma) + 2\chi(\Sigma) = \chi(\Sigma) = 2 - 2g.$$

This shows that u cannot be Fredholm regular unless $g = 0$.

Theorem 8.18 appeared for the first time in the contact case in [Dra04], and alternative proofs have since appeared in the appendix of [Bou06] (for cylinders in the contact case) and in [Wena] (under slightly different assumptions in the stable Hamiltonian setting). What I will describe below is a generalization of Bourgeois's proof.

8.3.2. Injective points of the projected curve. One point of difficulty in proving transversality in $\mathbb{R} \times M$ is that in contrast to the setting of Theorem 8.3, generic perturbations within $\mathcal{J}(\mathcal{H})$ can never be truly local, i.e. if you perturb J near a point $(r, x) \in \mathbb{R} \times M$, then you are also perturbing it in a neighborhood of the entire line $\mathbb{R} \times \{x\}$. We therefore need to know that we can find a point

$z \in \dot{\Sigma}$ that is the *only* point where $u : \dot{\Sigma} \to \mathbb{R} \times M$ passes through such a line; put another way, we need to know that not only $u = (u_{\mathbb{R}}, u_M) : \dot{\Sigma} \to \mathbb{R} \times M$ but also the projected map $u_M : \dot{\Sigma} \to M$ is somewhere injective. The first step in showing this is Proposition 8.19 above, as the zeroes of the section

$$\pi_\xi \circ du \in \Gamma(\mathrm{Hom}_{\mathbb{C}}(T\dot{\Sigma}, u^*\xi))$$

are precisely the non-immersed points of $u_M : \dot{\Sigma} \to M$; everywhere else, u_M is an immersion transverse to the Reeb vector field. To prove Proposition 8.19, we shall use the fact that the vector fields ∂_r and R generate an integrable J-invariant distribution on $\mathbb{R} \times M$. Indeed, the zeroes of $\pi_\xi \circ du$ are the points of tangency with this distribution, hence the result is an immediate consequence of the following statement:

LEMMA 8.24. *Suppose (W, J) is an almost complex manifold, $\Xi \subset TW$ is a smooth integrable J-invariant distribution and $u : (\Sigma, j) \to (W, J)$ is a connected pseudoholomorphic curve whose image is not contained in a leaf of the foliation generated by Ξ. Then all points $z \in \Sigma$ with $\operatorname{im} du(z) \subset \Xi$ are isolated in Σ.*

PROOF. The statement is local, so assume $(\Sigma, j) = (\mathbb{D}, i)$ with coordinates $s+it$, $W = \mathbb{C}^n$, and $u(0) = 0$. Let $2m$ denote the real dimension of Ξ, and observe that since Ξ is integrable, we can change coordinates near 0 and assume without loss of generality that at every point $p \in \mathbb{C}^n$ near 0, $\Xi_p = \mathbb{C}^m \oplus \{0\} \subset \mathbb{C}^n = T_p\mathbb{C}^n$. The J-invariance of Ξ then implies that in coordinates $(w, \zeta) \in \mathbb{C}^m \times \mathbb{C}^{n-m}$, J takes the form

$$J(w, \zeta) = \begin{pmatrix} J_1(w,\zeta) & Y(w,\zeta) \\ 0 & J_2(w,\zeta) \end{pmatrix},$$

where J_1^2 and J_2^2 are both $-\mathbb{1}$, and $J_1 Y + Y J_2 = 0$. Writing $u(z) = (f(z), v(z)) \in \mathbb{C}^m \times \mathbb{C}^{n-m}$, the Cauchy-Riemann equation $\partial_s u + J(u)\partial_t u = 0$ is then equivalent to the two equations

$$\begin{aligned} \partial_s f + J_1(f,v)\, \partial_t f + Y(f,v)\, \partial_t v &= 0, \\ \partial_s v + J_2(f,v)\, \partial_t v &= 0. \end{aligned} \tag{8.6}$$

We have $\operatorname{im} du(z) \subset \Xi$ wherever $\partial_s v = \partial_t v = 0$; notice that it suffices to consider the condition $\partial_s v = 0$ since $\partial_t v = J_2(f,v)\, \partial_s v$. Differentiating the second equation in (8.6) with respect to s gives

$$\partial_s(\partial_s v) + J_2(f,v)\, \partial_t(\partial_s v) + \partial_s \left[J_2(f,v)\right] J_2(f,v)\, \partial_s v = 0,$$

where in the last term we've substituted $J_2(f,v)\, \partial_s v$ for $\partial_t v$. Setting $\bar{J}(z) := J_2(f(z), v(z))$ and $A(z) := \partial_s \left[J_2(f(z), v(z))\right] J_2(f(z), v(z))$, this becomes a linear Cauchy-Riemann type equation $\partial_s(\partial_s v) + \bar{J}\, \partial_t(\partial_s v) + A(\partial_s v) = 0$, so the similarity principle implies that zeroes of $\partial_s v$ are isolated unless it is identically zero. The latter would mean v is constant, so u is contained in a leaf of Ξ. □

LEMMA 8.25. *Suppose $J \in \mathcal{J}(\mathcal{H})$, $\gamma : S^1 \to M$ is a closed Reeb orbit, and $u = (u_{\mathbb{R}}, u_M) : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is an asymptotically cylindrical J-holomorphic curve that is not a cover of a trivial cylinder. Then all intersections of the map $u_M : \dot{\Sigma} \to M$ with the image of the orbit γ are isolated.*

PROOF. The trivial cylinder over γ is a J-holomorphic curve, so the statement follows from the fact that two asymptotically cylindrical J-holomorphic curves can only have isolated intersections unless both are covers of the same simple curve. $\square$

We can now prove the statement we need about somewhere injectivity for $u_M : \dot{\Sigma} \to M$. This result first appeared in [**HWZ99**, Theorem 1.13].

PROPOSITION 8.26. *Suppose $J \in \mathcal{J}(\mathcal{H})$ and*

$$u = (u_{\mathbb{R}}, u_M) : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$$

is a simple asymptotically cylindrical J-holomorphic curve which is not a trivial cylinder and has only nondegenerate asymptotic orbits. Then the set of injective points $z \in \dot{\Sigma}$ of the map $u_M : \dot{\Sigma} \to M$ for which $u_M(z)$ is not contained in any of the asymptotic orbits of u is open and dense.

PROOF. Openness is clear, so our main task is to prove density. The idea is first to show via elementary topological arguments that if the set of injective points is not dense, then $\dot{\Sigma}$ contains two disjoint open sets on which u_M is an embedding with identical images. We will then conclude from this that if u is simple, it must be equivalent to one of its nontrivial $\mathbb{R}$-translations, and the latter is impossible for an asymptotically cylindrical curve.

Step 1: We begin by harmlessly removing some discrete sets of points in $\dot{\Sigma}$ that would make the subsequent arguments more complicated. Let

$$P \subset M$$

denote the union of the images of the asymptotic orbits of u, a finite disjoint union of circles. Lemma 8.25 implies that $u_M^{-1}(P)$ is a discrete subset of $\dot{\Sigma}$. By Proposition 8.19, there is also a discrete set $Z \subset \dot{\Sigma} \backslash u_M^{-1}(P)$ containing all points $z \notin u_M^{-1}(P)$ where $\pi_\xi \circ du(z) = 0$, and we claim that

$$Z' := u_M^{-1}(u_M(Z))$$

is a discrete subset of $\dot{\Sigma} \backslash u_M^{-1}(P)$. Indeed, $u_M(Z)$ is a discrete subset of $M \backslash P$ since the points in Z can only accumulate at infinity,[6] hence accumulation points of $u_M(Z) \subset M$ can occur only in P. For each individual point $p \in u_M(Z)$, the fact that $p \notin P$ implies $u_M^{-1}(p)$ is compact, and it consists of a discrete (and therefore finite) set of points with $\pi_\xi \circ du(z) = 0$, plus possibly some other points where $\pi_\xi \circ du(z) \neq 0$, but u_M is an embedding near each point of the latter type, so that these points of $u_M^{-1}(p)$ must always be isolated and are therefore also finite in number. This proves the claim, and we conclude that

$$\ddot{\Sigma} := \dot{\Sigma} \backslash \left(u_M^{-1}(P) \cup Z' \right)$$

an open and dense subset of $\dot{\Sigma}$, as it is obtained by removing a discrete subset from the open and dense subset $\dot{\Sigma} \backslash u_M^{-1}(P)$. To prove the proposition, it will now suffice to prove that the set of points $z \in \ddot{\Sigma}$ which are injective points of $u_M : \dot{\Sigma} \to M$ is

[6]Actually the asymptotic formula of [**HWZ96**] implies that both Z and $u_M^{-1}(P)$ are always finite for curves that are not covers of trivial cylinders, but we do not need to use that here.

dense in $\ddot{\Sigma}$. We shall argue by contradiction and assume from now on that density fails.

Step 2: We will find two open subsets $\mathcal{U}, \mathcal{V} \subset \dot{\Sigma}$ such that u_M restricts to an embedding on both, but

$$\mathcal{U} \cap \mathcal{V} = \varnothing \quad \text{and} \quad u_M(\mathcal{U}) = u_M(\mathcal{V}).$$

Indeed, assume the set of injective points of u_M lying in $\ddot{\Sigma}$ is not dense in $\ddot{\Sigma}$. Then there exists a point $z_0 \in \ddot{\Sigma}$ with a closed neighborhood $\mathcal{D}(z_0) \subset \ddot{\Sigma}$ such that no $z \in \mathcal{D}(z_0)$ is an injective point. Since $z \in \ddot{\Sigma}$ implies $\pi_\xi \circ du(z) \neq 0$, this means that for every $z \in \mathcal{D}(z_0)$, there exists $\zeta \in \dot{\Sigma} \backslash \{z\}$ with $u_M(z) = u_M(\zeta)$, and the definition of $\ddot{\Sigma}$ implies ζ is also in $\ddot{\Sigma}$, hence $\pi_\xi \circ du(\zeta) \neq 0$ and u_M is a local embedding near ζ. Since $u(z) \notin P$ and u_M maps $\dot{\Sigma} \backslash u_M^{-1}(P)$ properly to $M \backslash P$, we also conclude that $u_M^{-1}(u_M(z))$ is finite. Now suppose $u_M^{-1}(u_M(z_0)) = \{z_0, \zeta_1, \ldots, \zeta_m\}$, and let $\mathcal{D}(\zeta_j) \subset \ddot{\Sigma}$ for $j = 1, \ldots, m$ denote closed neighborhoods on which u_M is an embedding. We claim that after possibly shrinking $\mathcal{D}(z_0)$, we can assume

$$u_M(\mathcal{D}(z_0)) \subset \bigcup_{j=1}^{m} u_M(\mathcal{D}(\zeta_j)).$$

Let us first shrink $\mathcal{D}(z_0)$ so that u_M is an embedding on $\mathcal{D}(z_0)$, which is possible since $\pi_\xi \circ du(z_0) \neq 0$. Then if the claim is false, there exists a sequence $z_\nu \in \mathcal{D}(z_0)$ of noninjective points with $z_\nu \to z_0$, hence there is also a sequence $z'_\nu \in \ddot{\Sigma} \backslash \mathcal{D}(z_0)$ with $u_M(z_\nu) = u_M(z'_\nu)$ but z'_ν not converging to any of $\zeta_1, \ldots, \zeta_m$. But since $u_M(z'_\nu) \to u_M(z_0) \notin P$, the points z'_ν are confined to a compact subset of $\dot{\Sigma}$ and therefore have a subsequence $z'_\nu \to z'_\infty \in \dot{\Sigma}$ with $u_M(z'_\infty) = u_M(z_0)$. The limit cannot be z_0 itself since $z'_\nu \notin \mathcal{D}(z_0)$, thus z'_∞ must be one of the $\zeta_1, \ldots, \zeta_m$, and we have a contradiction. We claim next that at least one of the sets $u_M(\mathcal{D}(z_0)) \cap u_M(\mathcal{D}(\zeta_j))$ has nonempty interior. This is a simple exercise in metric space topology: it can be reduced to the fact that if X is a metric space with closed subsets $V, W \subset X$ that both have empty interior (meaning no open subset of X is contained in V or W), then $V \cup W$ also has empty interior. Since the subsets $u_M(\mathcal{D}(z_0)) \cap u_M(\mathcal{D}(\zeta_j)) \subset u_M(\mathcal{D}(z_0))$ for $j = 1, \ldots, m$ are all closed but their union is $u_M(\mathcal{D}(z_0))$, they cannot all have empty interior. This achieves the goal of Step 2.

Step 3: We show that u is biholomorphically equivalent to one of its $\mathbb{R}$-translations

$$\tau \cdot u := (u_\mathbb{R} + \tau, u_M) : \dot{\Sigma} \to \mathbb{R} \times M$$

for $\tau \in \mathbb{R} \backslash \{0\}$. To see this, note that for $J \in \mathcal{J}(\mathcal{H})$, the nonlinear Cauchy-Riemann equation $Tu \circ j = J \circ Tu$ is equivalent to the two equations

$$\begin{aligned} du_\mathbb{R} &= u_M^* \lambda \circ j, \\ \pi_\xi \circ du_M \circ j &= J(u_M) \circ \pi_\xi \circ du_M. \end{aligned} \tag{8.7}$$

Since $\pi_\xi \circ du_M : T\dot{\Sigma} \to u_M^* \xi$ is fiberwise injective everywhere on the neighborhoods $\mathcal{U}$ and $\mathcal{V}$, the second equation determines j in terms of J on each of these regions; in particular, the identification of $u_M(\mathcal{U})$ with $u_M(\mathcal{V})$ provides a biholomorphic map of $\mathcal{V}$ to $\mathcal{U}$ so that $u|_\mathcal{U}$ and $u|_\mathcal{V}$ may be regarded as two J-holomorphic maps from the same Riemann surface which differ only in the $\mathbb{R}$-factor. But with j and u_M both fixed, the first equation in (8.7) determines $du_{\mathbb{R}}$ and thus determines $u_{\mathbb{R}}$ up to the addition of a constant $\tau \in \mathbb{R}$. If $\tau = 0$, this means u has two disjoint regions on which its images are identical, contradicting the assumption that u is simple. Thus $\tau \neq 0$, and since two distinct simple curves can only intersect each other at isolated points, we conclude $u = \tau \cdot u$ up to parametrization.

Step 4: We now derive a contradiction. The relation $u = \tau \cdot u$ implies that in fact $u = k\tau \cdot u$ for every $k \in \mathbb{Z}$, so we obtain a diverging sequence of $\mathbb{R}$-translations $\tau_k \to \infty$ such that u and $\tau_k \cdot u$ always have identical images in $\mathbb{R} \times M$. It follows that for some point $z \in \dot{\Sigma}$ with $u(z) = (r, x)$ where x is not contained in any of the asymptotic orbits of u, the points $(r - \tau_k, x)$ are all in the image of u as $\tau_k \to \infty$. But this contradicts the asymptotically cylindrical behavior of u. $\square$

8.3.3. Smoothness of the universal moduli space. The overall outline of the proof of Theorem 8.18 is the same as for Theorem 8.3: one needs to define a suitable space $\mathcal{J}_{\mathcal{U}}^{\epsilon}$ of perturbed almost complex structures, giving rise to a universal moduli space $\mathcal{M}^*(\mathcal{J}_{\mathcal{U}}^{\epsilon})$ that is a smooth Banach manifold, and then apply the Sard-Smale theorem to conclude that generic elements of $\mathcal{J}_{\mathcal{U}}^{\epsilon}$ are regular values of the projection $\mathcal{M}^*(\mathcal{J}_{\mathcal{U}}^{\epsilon}) \to \mathcal{J}_{\mathcal{U}}^{\epsilon} : (u, J) \mapsto J$. If $\mathcal{J}_{\mathcal{U}}^{\epsilon}$ is a space of C_{ϵ}-perturbed almost complex structures, then in the final step one can use the Taubes trick as in §8.2.4 to transform the genericity result in $\mathcal{J}_{\mathcal{U}}^{\epsilon}$ into a genericity result within the space $\mathcal{J}(\mathcal{H})$ of smooth almost complex structures. The only step that differs meaningfully from what we've already discussed is the smoothness of the universal moduli space, so let us focus on this detail.

Assume $J^{\mathrm{ref}} \in \mathcal{J}(\mathcal{H})$ with $J^{\mathrm{ref}} = J^{\mathrm{fix}}$ outside $\mathbb{R} \times \mathcal{U}$, and $\mathcal{J}_{\mathcal{U}}^{\epsilon}$ is a Banach manifold of C_{ϵ}-small perturbations of J^{ref} in $\mathcal{J}(\mathcal{H})$ that are also fixed outside of $\mathbb{R} \times \mathcal{U}$. For $J \in \mathcal{J}(\mathcal{H})$, we consider the open subset $\mathcal{M}^*(J) \subset \mathcal{M}(J)$ defined by

$$\mathcal{M}^*(J) := \Big\{ u \in \mathcal{M}(J) \;\Big|\; u = (u_{\mathbb{R}}, u_M) : \dot{\Sigma} \to \mathbb{R} \times M \text{ has an injective point } z \in \dot{\Sigma} \text{ with } u(z) \in \mathbb{R} \times \mathcal{U},\ \pi_{\xi} \circ du(z) \neq 0 \text{ and } \operatorname{im} du_M(z) \cap \xi_{u(z)}^{\perp d\lambda} = \{0\} \Big\}.$$

and the corresponding universal moduli space

$$\mathcal{M}^*(\mathcal{J}_{\mathcal{U}}^{\epsilon}) := \left\{ (u, J) \;\middle|\; J \in \mathcal{J}_{\mathcal{U}}^{\epsilon} \text{ and } u \in \mathcal{M}^*(J) \right\}.$$

The local structure of $\mathcal{M}^*(\mathcal{J}_{\mathcal{U}}^{\epsilon})$ near an element (u_0, J_0) with representative $u_0 : (\dot{\Sigma}, j_0) \to (\mathbb{R} \times M, J_0)$ can again be described via the zero set of a smooth section

$$\bar{\partial} : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \times \mathcal{J}_{\mathcal{U}}^{\epsilon} \to \mathcal{E}^{k-1,p,\delta} : (j, u, J) \mapsto du + J(u) \circ du \circ j,$$

where $\mathcal{T}$ is a Teichmüller slice through $(j_0, \Gamma \cup \Theta)$, and it suffices to show that the linear map

$$\begin{aligned} W^{k,p,\delta}(u_0^* T(\mathbb{R} \times M)) \oplus T_{J_0} \mathcal{J}_{\mathcal{U}}^{\epsilon} &\xrightarrow{\mathbf{L}} W^{k-1,p,\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^* T(\mathbb{R} \times M))) \\ (\eta, Y) &\mapsto \mathbf{D}_{u_0} \eta + Y(u_0) \circ du_0 \circ j_0 \end{aligned}$$

is always surjective. As usual, here we're assuming $k \in \mathbb{N}$, $1 < p < \infty$, and the exponential weight $\delta > 0$ is small but positive so that the map $W^{k,p,\delta} \to W^{k-1,p,\delta} : \eta \mapsto \mathbf{D}_{u_0} \eta$ is Fredholm. The image of $\mathbf{L}$ is then closed, and focusing on the $k = 1$ case, if $\mathbf{L}$ is not surjective then there exists a nontrivial element $\theta \in L^{q,-\delta}(\overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^*T(\mathbb{R} \times M)))$ such that

$$\begin{aligned} \langle \mathbf{D}_{u_0}\eta, \theta \rangle_{L^2} &= 0 \text{ for all } \eta \in W^{1,p,\delta}(u_0^*T(\mathbb{R} \times M)), \\ \langle Y(u_0) \circ du_0 \circ j_0, \theta \rangle_{L^2} &= 0 \text{ for all } Y \in T_{J_0}\mathcal{J}^{\epsilon}_{\mathcal{U}}. \end{aligned} \tag{8.8}$$

The first condition implies via elliptic regularity and the similarity principle that θ is smooth and has only isolated zeroes. So far this is all the same as in the proof of Theorem 8.3, but the next step is trickier: since perturbing J_0 within $\mathcal{J}(\mathcal{H})$ only changes the action of the almost complex structure on ξ but not on the trivial subbundle generated by ∂_r and R, it is not clear whether the range of values allowed for Y is large enough to force $\langle Y(u_0) \circ du_0 \circ j_0, \theta \rangle_{L^2} > 0$.

To overcome this, let us decompose everything in this picture with respect to the natural splitting

$$T(\mathbb{R} \times M) = \varepsilon \oplus \xi,$$

where ε denotes the trivial line bundle spanned by ∂_r and R. In particular, the domain and target bundles of the Cauchy-Riemann type operator $\mathbf{D}_{u_0}$ now split as

$$\begin{aligned} u_0^*T(\mathbb{R} \times M) &= u_0^*\varepsilon \oplus u_0^*\xi, \\ \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^*T(\mathbb{R} \times M)) &= \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^*\varepsilon) \oplus \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^*\xi), \end{aligned}$$

and we shall write $\eta = (\eta^\varepsilon, \eta^\xi)$ and $\theta = (\theta^\varepsilon, \theta^\xi)$ accordingly. This gives a block decomposition of $\mathbf{D}_{u_0}$ as

$$\mathbf{D}_{u_0}\eta = \begin{pmatrix} (\mathbf{D}_{u_0}\eta)^\varepsilon \\ (\mathbf{D}_{u_0}\eta)^\xi \end{pmatrix} = \begin{pmatrix} \mathbf{D}^\varepsilon_{u_0} & \mathbf{D}^{\varepsilon\xi}_{u_0} \\ \mathbf{D}^{\xi\varepsilon}_{u_0} & \mathbf{D}^\xi_{u_0} \end{pmatrix} \begin{pmatrix} \eta^\varepsilon \\ \eta^\xi \end{pmatrix}.$$

By Exercise 6.23, $\mathbf{D}^\varepsilon_{u_0}$ and $\mathbf{D}^\xi_{u_0}$ are each Cauchy-Riemann type operators on $u_0^*\varepsilon$ and $u_0^*\xi$ respectively, while the off-diagonal terms are both tensorial, i.e. zeroth-order operators. Since perturbations of J_0 in $\mathcal{J}(\mathcal{H})$ only change its action on ξ, $Y \in T_{J_0}\mathcal{J}^{\epsilon}_{\mathcal{U}}$ now takes the block form

$$Y = \begin{pmatrix} 0 & 0 \\ 0 & Y^\xi \end{pmatrix},$$

where Y^ξ is a C_ϵ-small section of the bundle $\overline{\mathrm{End}}_{\mathbb{C}}(\xi, J_0)$ over M. Assuming the L^2-pairings are defined so as to respect these splittings, the second condition in (8.8) now becomes

$$\langle Y^\xi(u_0) \circ \pi_\xi \circ du_0 \circ j_0, \theta^\xi \rangle_{L^2} = 0,$$

and given any injective point $z_0 \in \dot{\Sigma}$ of $(u_0)_M : \dot{\Sigma} \to M$ satisfying $u_0(z_0) \in \mathbb{R} \times \mathcal{U}$, we have enough freedom to choose Y^ξ near $\mathbb{R} \times \{u_0(z_0)\}$ such that this pairing becomes positive unless

$$\theta^\xi = 0 \quad \text{near } z_0.$$

Assuming this holds, it remains to show that θ^ε also vanishes near z_0, which will contradict the fact that θ only has isolated zeroes. To this end, notice that the first condition in (8.8) implies via separate choices of the components η^ε and η^ξ with support near z_0 that

$$\langle \mathbf{D}^\varepsilon_{u_0}\eta^\varepsilon, \theta^\varepsilon\rangle_{L^2} = 0 \text{ for all } \eta^\varepsilon \text{ supported near } z_0,$$
$$\langle \mathbf{D}^{\varepsilon\xi}_{u_0}\eta^\xi, \theta^\varepsilon\rangle_{L^2} = 0 \text{ for all } \eta^\xi \text{ supported near } z_0. \tag{8.9}$$

The first of these two conditions gives no new information, since we already know that $\theta = (\theta^\varepsilon, 0)$ solves an anti-Cauchy-Riemann equation. To get some information out of the second condition, we will need an explicit formula for $\mathbf{D}^{\varepsilon\xi}_{u_0}$.

LEMMA 8.27. *The tensorial operator* $\mathbf{D}^{\varepsilon\xi}_{u_0} : u_0^*\xi \to \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^*\varepsilon)$ *takes the form*

$$\mathbf{D}^{\varepsilon\xi}_{u_0}\eta^\xi = \left[-d\lambda\big(\eta^\xi, du \circ j(\cdot)\big)\right]\partial_r + \left[d\lambda\big(\eta^\xi, du(\cdot)\big)\right]R.$$

PROOF. As a preliminary step, notice that $-dr \circ J = \lambda$ for any $J \in \mathcal{J}(\mathcal{H})$; indeed, the conditions $J(\xi) = \xi \subset \ker dr$ and $J\partial_r = R$ imply that these two 1-forms have matching values on ∂_r, R and ξ. As a consequence, $\lambda \circ J_0 = dr$, so in particular $\lambda \circ J_0$ is closed.

Choosing local holomorphic coordinates (s,t) in an arbitrary neighborhood in $\dot{\Sigma}$, we have

$$(\mathbf{D}^{\varepsilon\xi}_{u_0}\eta^\xi)\partial_s = dr\big((\mathbf{D}_{u_0}\eta^\xi)\partial_s\big)\,\partial_r + \lambda\big((\mathbf{D}_{u_0}\eta^\xi)\partial_s\big)\,R.$$

Extend $u_0 : \dot{\Sigma} \to \mathbb{R} \times M$ to a smooth 1-parameter family of maps $\{u_\rho : \dot{\Sigma} \to \mathbb{R} \times M\}_{\rho\in\mathbb{R}}$ with $\partial_\rho u_\rho|_{\rho=0} = \eta^\xi \in \Gamma(u_0^*\xi)$. Then by the definition of the linearized Cauchy-Riemann operator,

$$(\mathbf{D}_{u_0}\eta^\xi)\partial_s = \nabla_\rho\left(\partial_s u_\rho + J_0(u_\rho)\partial_t u_\rho\right)|_{\rho=0},$$

for any choice of connection ∇ on $\mathbb{R} \times M$. Since $\partial_s u_0 + J_0(u_0)\partial_t u_0 = 0$, we find

$$\begin{aligned}
\lambda\big((\mathbf{D}_{u_0}\eta^\xi)\partial_s\big) &= \lambda\big(\nabla_\rho\left(\partial_s u_\rho + J_0(u_\rho)\partial_t u_\rho\right)|_{\rho=0}\big) = \partial_\rho\left[\lambda(\partial_s u_\rho + J_0(u_\rho)\partial_t u_\rho)\right]|_{\rho=0} \\
&= \partial_\rho\left[\lambda(\partial_s u_\rho)\right]|_{\rho=0} + \partial_\rho\left[(\lambda \circ J_0)(\partial_t u_\rho)\right]|_{\rho=0} \\
&= d\lambda(\eta^\xi, \partial_s u) + d(\lambda \circ J_0)(\eta^\xi, \partial_t u) \\
&= d\lambda(\eta^\xi, \partial_s u),
\end{aligned}$$

where we've used the formula

$$d\lambda(X,Y) = \mathcal{L}_X\left[\lambda(Y)\right] - \mathcal{L}_Y\left[\lambda(X)\right] - \lambda([X,Y])$$

and eliminated several terms using the fact that $\lambda(\eta^\xi) = \lambda(J_0\eta^\xi) = 0$ since η^ξ is valued in ξ, plus $d(\lambda \circ J_0) = 0$. A similar computation gives

$$dr\big((\mathbf{D}_{u_0}\eta^\xi)\partial_s\big) = -d\lambda(\eta^\xi, \partial_t u),$$

so removing the local coordinates from the picture produces the stated formula. ☐

LEMMA 8.28. *If* $\operatorname{im} du(z_0) \cap \xi^{\perp d\lambda}_{u(z_0)} = \{0\}$*, then* $\mathbf{D}^{\varepsilon\xi}_{u_0} : u_0^*\xi \to \overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u_0^*\varepsilon)$ *is fiberwise surjective on a neighborhood of* z_0.

PROOF. Choose holomorphic coordinates $s + it$ near z_0 and use Lemma 8.27 to write

$$(\mathbf{D}^{\varepsilon\xi}_{u_0}\eta^\xi)\partial_s = \left[-d\lambda(\eta^\xi, \partial_t u)\right]\partial_r + \left[d\lambda(\eta^\xi, \partial_s u)\right]R.$$

This describes a bundle map $u_0^*\xi \to u_0^*\varepsilon$ near z_0 that is surjective at every point z where $d\lambda(\cdot, \partial_t u)|_{\xi_{u(z)}}$ and $d\lambda(\cdot, \partial_s u)|_{\xi_{u(z)}}$ are linearly independent in $\mathrm{Hom}(\xi_{u(z)}, \mathbb{R})$. The latter is true if and only if for every pair of coefficients $a, b \in \mathbb{R}$ that do not both vanish,

$$0 \neq a\, d\lambda(\cdot, \partial_t u)|_{\xi_{u(z)}} + b\, d\lambda(\cdot, \partial_s u)|_{\xi_{u(z)}} = d\lambda(\cdot, a\partial_t u + b\partial_s u)|_{\xi_{u(u)}},$$

meaning no nontrivial vector in $\operatorname{im} du$ is in $\xi^{\perp d\lambda}$. $\square$

Now assuming that our injective point z_0 also satisfies the condition in Lemma 8.28, we can choose η^ξ with support near z_0 to satisfy

$$\langle \mathbf{D}_{u_0}\eta^\xi, \theta^\varepsilon \rangle_{L^2} > 0$$

unless θ^ε vanishes identically near z_0. The two conditions in (8.9) therefore imply $\theta^\varepsilon \equiv 0$ near z_0 and thus $\theta \equiv 0$, which is a contradiction.

We've proved that the universal moduli space is smooth as claimed. Since the rest of the proof of Theorem 8.18 is the same as in the non-$\mathbb{R}$-invariant case, we leave those details to the reader.

REMARK 8.29. You may have noticed that in both Theorems 8.3 and 8.18, our proof that the universal moduli space is smooth relied on a surjectivity result that was actually stronger than needed: in both cases, we needed to prove that an operator of the form

$$T_{j_0}\mathcal{T} \oplus T_{u_0}\mathcal{B}^{k,p,\delta} \oplus T_{J_0}\mathcal{J}_{\mathcal{U}}^{\epsilon} \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0,J_0)}$$

was surjective, but we ended up proving that its restriction to the smaller domain $T_{u_0}\mathcal{B}^{k,p,\delta} \oplus T_{J_0}\mathcal{J}_{\mathcal{U}}^{\epsilon}$ is already surjective. This technical detail hints at a stronger result that can be proved using these methods: one can show that not only is $\mathcal{M}^*(\mathcal{J}_{\mathcal{U}}^{\epsilon})$ smooth but also the **forgetful map**

$$\mathcal{M}^*(\mathcal{J}_{\mathcal{U}}^{\epsilon}) \to \mathcal{M}_{g,k_+ + k_- + m}$$
$$([(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)], J) \mapsto [(\Sigma, j, \Gamma \cup \Theta)]$$

sending a J-holomorphic curve to its underlying domain in the moduli space of Riemann surfaces is a submersion, cf. the blog post [**Wenb**] and its sequel. One can use this to prove generic transversality results for spaces of J-holomorphic curves whose domains are constrained within the moduli space of Riemann surfaces, which gives rise to more elaborate algebraic structures on SFT, e.g. this idea plays a prominent role in the study of Gromov-Witten invariants.

LECTURE 9

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 lecture 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 lecture is the notion of *bubbling*, of which we will see several examples. Bubbling arises in a natural way from elliptic regularity: recall that in Lecture 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 2.24). 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 9.1. *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

$$
\begin{aligned}
u^*\omega(\partial_s, \partial_t) = \omega(\partial_s u, \partial_t u) &= \frac{1}{2}\left[\omega(\partial_s u, J\partial_s u) + \omega(\partial_t u, J\partial_t u)\right] \\
&= \frac{1}{2}\left(|\partial_s u|_g^2 + |\partial_t u|_g^2\right),
\end{aligned}
\tag{9.1}
$$

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 9.1, 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.

9.1. Removal of singularities

As an important tool for use in the rest of this lecture, we begin with the following result from [**Gro85**]:

THEOREM 9.2 (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 (§9.1.1) establishes that u has a *continuous* extension, and after that (§9.1.2) we will use elliptic regularity to show that the continuous extension is actually smooth.

9.1.1. The continuous extension. We will use as a black box the following additional result from [**Gro85**], which is closely related to a standard result about minimal surfaces:

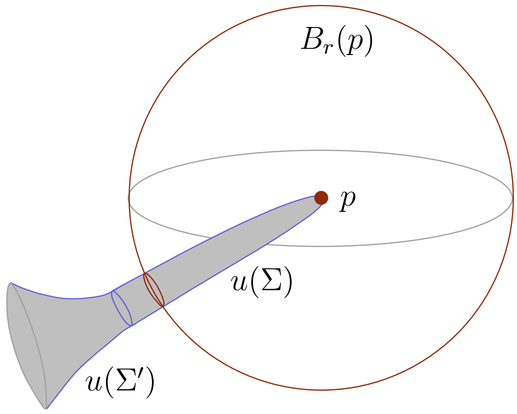

FIGURE 9.1. The intersection of a J-holomorphic curve u with an open ball $B_r(p)$ defines a proper map $\Sigma \to B_r(p)$. The monotonicity lemma prevents this map from having arbitrarily small area if it passes through p.

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 9.1.

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

$$u(s, \cdot) \to p \qquad \text{in} \qquad C^\infty(S^1, W) \text{ as } s \to \infty. \tag{9.2}$$

Fix the obvious flat metric on Z_+ and any Riemannian metric on W in order to define norms such as $|du(s, t)|$ for $(s, t) \in Z_+$.

LEMMA 9.3. *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 $z_k = (s_k, t_k) \in Z_+$ with $|du(z_k)| =: R_k \to \infty$. Choose a sequence of positive numbers $\epsilon_k > 0$ that converge to zero but not too fast, so that $\epsilon_k R_k \to \infty$. We then consider the sequence of reparametrized maps

$$v_k : \mathbb{D}_{\epsilon_k R_k} \to W : z \mapsto u(z_k + z/R_k).$$

These are also J-holomorphic since $z \mapsto z_k + z/R_k$ is holomorphic, and the values of v_k depend only on the values of u over the ϵ_k-disk about z_k. Notice that since $s_k \to \infty$ and $\epsilon_k \to 0$, we are free to assume that all of these ϵ_k-disks are disjoint; moreover, tameness of J implies $u^*\omega \geqslant 0$ and $v_k^*\omega \geqslant 0$, thus

$$\sum_k \int_{\mathbb{D}_{\epsilon_k R_k}} v_k^*\omega = \sum_k \int_{\mathbb{D}_{\epsilon_k}(z_k)} u^*\omega \leqslant \int_{Z_+} u^*\omega < \infty,$$

implying

$$\int_{\mathbb{D}_{\epsilon_k R_k}} v_k^*\omega \to 0 \quad \text{as} \quad k \to \infty. \tag{9.3}$$

We would now like to say something about a limit of the maps v_k as $k \to \infty$, 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_k 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. 9.1) would give a C^∞_{loc} convergent subsequence on the union of all the domains $\mathbb{D}_{\epsilon_k R_k}$, i.e. on the entire plane. We have

$$dv_k(z) = \frac{1}{R_k} du(z_k + z/R_k),$$

hence $|dv_k(0)| = 1$, but we will need to know more about $|du|$ on the rest of $\mathbb{D}_{\epsilon_k}(z_k)$ in order to deduce a C^1-bound for v_k on all of $\mathbb{D}_{\epsilon_k R_k}$. 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:

LEMMA 9.4 (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. $\square$

PROOF OF LEMMA 9.3, PART 2. Applying Lemma 9.4 to $X = Z_+$ with $g(z) = |du(z)|$, we can replace the original sequences ϵ_k and z_k with new sequences for which all the previously stated properties still hold, but additionally,

$$|du(z)| \leqslant 2|du(z_k)| \quad \text{for all} \quad z \in \mathbb{D}_{\epsilon_k}(z_k).$$

Our sequence of reparametrizations v_k then satisfies

$$|dv_k(z)| \leqslant 2 \quad \text{for all} \quad z \in \mathbb{D}_{\epsilon_k R_k},$$

so by elliptic regularity, v_k 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 $|dv_\infty(0)| = \lim_{k\to\infty} |dv_k(0)| = 1$. Informally, we say that the blow-up of the derivatives at z_k has caused a plane to "bubble off". However, (9.3) implies that for every $R > 0$, one can write $\epsilon_k R_k \geqslant R$ for k sufficiently large and thus

$$\int_{\mathbb{D}_R} v_\infty^*\omega = \lim_{k\to\infty} \int_{\mathbb{D}_R} v_k^*\omega \leqslant \lim_{k\to\infty} \int_{\mathbb{D}_{\epsilon_k R_k}} v_k^*\omega = 0,$$

implying $\int_{\mathbb{C}} v_\infty^*\omega = 0$. It follows that v_∞ must be constant, so we have a contradiction. □

To obtain the uniform limit of $u(s,\cdot)$ as $s \to \infty$, we now pick any sequence of nonnegative numbers $s_k \to \infty$ and consider the sequence of J-holomorphic half-cylinders

$$u_k : [-s_k, \infty) \times S^1 \to W : (s,t) \mapsto u(s + s_k, t).$$

By Lemma 9.3, 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 $-s_k/2 \leqslant -c$ for sufficiently large k and thus compute

$$\begin{aligned} \int_{[-c,c]\times S^1} u_\infty^*\omega &= \lim_{k\to\infty} \int_{[-c,c]\times S^1} u_k^*\omega \leqslant \lim_{k\to\infty} \int_{[-s_k/2,\infty)\times S^1} u_k^*\omega \\ &= \lim_{k\to\infty} \int_{[s_k/2,\infty)\times S^1} u^*\omega = 0 \end{aligned}$$

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_k with a subsequence,

$$u(s_k, \cdot) = u_k(0, \cdot) \to p \quad \text{in} \quad C^\infty(S^1, W) \text{ as } k \to \infty.$$

To finish the proof of (9.2), we need to show that one cannot find two sequences $s_k \to \infty$ and $s_k' \to \infty$ such that $u(s_k,\cdot) \to p$ and $u(s_k',\cdot) \to p'$ 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 $s_k'' \to \infty$ for which the loops $u(s_k'',\cdot)$ 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 $\mathcal{U}_k \subset Z_+$ such that each

$$u|_{\mathcal{U}_k} : \mathcal{U}_k \to B_r(q_k)$$

is a proper map passing through some point $q_k \in \partial B_{2r}(p)$. The monotonicity lemma then implies

$$\int_{Z_+} u^*\omega \geqslant \sum_k \int_{\mathcal{U}_k} u^*\omega \geqslant \sum_k cr^2 = \infty,$$

a contradiction.

EXERCISE 9.5. 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 $\int_{\mathbb{C}} f^*\omega = \infty$.

9.1.2. From $C^0 \cap H^1$ to C^∞. To complete the proof of Theorem 9.2, we need to show that a continuous map $u : \mathbb{D} \to W$ that is smooth and J-holomorphic on $\mathbb{D}\backslash\{0\}$ and satisfies $\int_{\mathbb{D}\backslash\{0\}} u^*\omega < \infty$ is also smooth at $0 \in \mathbb{D}$. By (9.1), the first derivative of u on $\mathbb{D}\backslash\{0\}$ is in $L^2(\mathbb{D}\backslash\{0\})$, so the next exercise implies $u \in H^1(\mathbb{D})$.

EXERCISE 9.6. 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.

If we could say $u \in W^{1,p}(\mathbb{D})$ for some $p > 2$ instead of $p = 2$, then the smoothness of u would now follow from the local nonlinear regularity results in Lecture 2 (see Corollary 2.25). The following addendum to those regularity results therefore completes the proof; note that in this statement, the hypothesis $u \in C^0(\mathbb{D})$ is not redundant since we do not assume $p > 2$. Our proof of the lemma is adapted from an argument due to Sikorav, cf. [Sik94, Prop. 2.3.6(i)].

LEMMA 9.7. *Assume $1 < p < \infty$ and J is a continuous almost complex structure on $\mathbb{C}^n$. If $u : \mathbb{D} \to \mathbb{C}^n$ is a J-holomorphic map in $C^0(\mathbb{D}) \cap W^{1,p}(\mathring{\mathbb{D}})$, then for every $q > p$, u is also of class $W^{1,q}$ on all compact subsets of $\mathring{\mathbb{D}}$.*

PROOF. Assume $u : \mathbb{D} \to \mathbb{C}^n$ 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 2.24, we then write $Q := i - J : \mathbb{C}^n \to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n)$ and consider rescaled functions of the form

$$\begin{aligned} \widehat{J} : \mathbb{C}^n &\to \mathcal{J}(\mathbb{C}^n), & \widehat{J}(x) &:= J(x/R), \\ \widehat{Q} : \mathbb{C}^n &\to \mathrm{End}_{\mathbb{R}}(\mathbb{C}^n), & \widehat{Q}(x) &:= Q(x/R) = i - \widehat{J}(x), \\ \hat{u} : \mathbb{D} &\to \mathbb{C}^n, & \hat{u}(z) &:= Ru(z_0 + \epsilon z), \end{aligned} \tag{9.4}$$

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

$$\bar{\partial}\hat{u} - \widehat{Q}(\hat{u})\partial_t\hat{u} = 0. \tag{9.5}$$

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 (9.5), 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 $\beta \in C_0^\infty(\mathring{\mathbb{D}})$ 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 with a fundamental solution of the $\bar{\partial}$-equation. Since $L^q(\mathring{\mathbb{D}}) \subset L^p(\mathring{\mathbb{D}})$ for $q > p$, the same 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 $\beta\hat{u} \in W_0^{1,p}(\mathring{\mathbb{D}})$ (cf. Exercise 2.14). 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 these last two claims, then they imply $\beta\hat{u} \in W^{1,q}(\mathring{\mathbb{D}})$ 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,

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

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. ☐

9.2. Finite energy and asymptotics

As further preparation for the compactness discussion, we now prove the long-awaited converse of the fact that asymptotically cylindrical curves have finite energy. We work in the setting described in §6.2.2: (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, 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 $J \in \mathcal{J}_\tau(\omega_h, r_0, \mathcal{H}_+, \mathcal{H}_-)$ 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

$$E(u) := \sup_{f \in \mathcal{T}(h, r_0)} \int_{\dot\Sigma} u^*\omega_f,$$

where

$$\mathcal{T}(h, r_0) := \left\{ f \in C^\infty(\mathbb{R}, (-\epsilon, \epsilon)) \;\middle|\; f' > 0 \text{ and } f \equiv h \text{ near } [-r_0, r_0] \right\}.$$

The constant $\epsilon > 0$ should always be assumed sufficiently small so that if $J_\pm \in \mathcal{J}(\mathcal{H}_\pm)$ and $X \in \xi_\pm$,

$$(9.6) \qquad (\omega_\pm + \kappa\, d\lambda_\pm)(X, J_\pm X) > 0 \quad \text{whenever} \quad X \neq 0 \text{ and } \kappa \in (-2\epsilon, 2\epsilon).$$

This condition implies that every $J \in \mathcal{J}_\tau(\omega_h, r_0, \mathcal{H}_+, \mathcal{H}_-)$ is tamed by ω_f for every $f \in \mathcal{T}(h, r_0)$; cf. Proposition 6.19. It follows that all J-holomorphic curves satisfy $E(u) \geqslant 0$, with equality if and only if u is constant.

THEOREM 9.8. *Assume all closed Reeb orbits in* $(M_+, \mathcal{H}_+)$ *and* $(M_-, \mathcal{H}_-)$ *are nondegenerate,* $J \in \mathcal{J}_\tau(\omega_h, r_0, \mathcal{H}_+, \mathcal{H}_-)$*,* (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 9.9. 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 9.8 given below.

Like removal of singularities, Theorem 9.8 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

$$\begin{aligned} \varphi_+ &: Z_+ := [0, \infty) \times S^1 \to \dot{\mathbb{D}} : (s,t) \mapsto e^{-2\pi(s+it)} \\ \varphi_- &: Z_- := (-\infty, 0] \times S^1 \to \dot{\mathbb{D}} : (s,t) \mapsto e^{2\pi(s+it)}. \end{aligned} \tag{9.7}$$

THEOREM 9.10. *Suppose* $J \in \mathcal{J}_\tau(\omega_h, r_0, \mathcal{H}_+, \mathcal{H}_-)$ *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* $s_k \to \pm\infty$*, one can restrict to a subsequence such that*

$$u_M(s_k, \cdot) \to \gamma(T \cdot) \quad \text{in} \quad C^\infty(S^1, M_\pm) \text{ as } k \to \infty$$

for some T-periodic Reeb orbit $\gamma : \mathbb{R}/T\mathbb{Z} \to M_\pm$. If γ is nondegenerate or Morse-Bott,[1] *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 $s_k, s_k' \to \pm\infty$ are two distinct sequences, then we have no guarantee in general that the two Reeb orbits obtained as limits of subsequences of $u_M(s_k, \cdot)$ and $u_M(s_k', \cdot)$ 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 §6.5 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 9.10 will reuse most of the techniques that we already saw in our proof of removal of singularities in §9.1. 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 $s_k \to \pm\infty$, the holomorphic maps defined by

$$u_k(s, t) = u \circ \varphi_\pm(s + s_k, t)$$

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 9.2) that the puncture is removable, and the second implies asymptotic convergence to a closed Reeb orbit.

One major difference between the proof of Theorem 9.10 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 $|du_k(s, t)| \leqslant C$ and in the analysis of the sequence u_k itself. In such cases, one can use the $\mathbb{R}$-translation action

$$\tau_c : \mathbb{R} \times M_\pm \to \mathbb{R} \times M_\pm : (r, x) \mapsto (r + c, x) \quad \text{for} \quad c \in \mathbb{R} \tag{9.8}$$

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

[1] The Morse-Bott condition is a standard generalization of nondegeneracy: the T-periodic orbit γ is called **Morse-Bott** if it belongs to a smooth k-dimensional family of T-periodic orbits for some $k \geqslant 0$ such that $\dim \ker \mathbf{A}_\gamma = k$.

$\mathbb{R}$-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 9.11. *Suppose $J \in \mathcal{J}(\mathcal{H})$ for some stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ on an odd-dimensional 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 $u_\gamma \circ \widetilde{\Phi} = u$. Relabeling 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

$$(9.9) \qquad \int_{\dot{\Sigma}} u^*\omega_f = \int_{\dot{\Sigma}} \Phi^* u_\gamma^* \omega_f = \int_{\dot{\Sigma}} \Phi^* d\left(f(s)\, dt\right) = \int_{\dot{\Sigma}} \Phi^* \left(f'(s)\, ds \wedge dt\right)$$

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

$$
\begin{aligned}
\Psi^* \left(f'(s)\,ds \wedge dt\right) &= -\frac{1}{2i}\Psi^* \left(f'(s)\,dz \wedge d\bar{z}\right) = -\frac{1}{2i} f'(\operatorname{Re}\Psi)\,d\Psi \wedge d\bar{\Psi} \\
&= -\frac{1}{2i} f'(\operatorname{Re}\Psi)\,\frac{\partial \Psi}{\partial z}\,dz \wedge \frac{\partial \bar{\Psi}}{\partial \bar{z}}\,d\bar{z} \\
&= -\frac{1}{2i} f'(s/|z|^2)\left(-\frac{1}{z^2}\right)\left(-\frac{1}{\bar{z}^2}\right)\,dz \wedge d\bar{z} \\
&= \frac{f'(s/|z|^2)}{|z|^4}\,ds \wedge dt \quad \text{for} \quad z = s + it \in \mathbb{C}\backslash\{0\}.
\end{aligned}
\tag{9.10}
$$

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

$$
f(\pm r) = \pm\left(\epsilon - \frac{\epsilon}{2r}\right) \quad \text{for} \quad r \geqslant 1 \tag{9.11}
$$

and extend f arbitrarily to $[-1, 1]$ such that $f' > 0$. We can then find a constant $c > 0$ such that f' satisfies

$$
f'(r) \geqslant \min\left\{c, \frac{\epsilon}{2r^2}\right\} \quad \text{for all} \quad r \in \mathbb{R}.
$$

Plugging this into (9.10) 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 9.2 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 9.5), then since $\Phi^* \left(f'(s)\,ds \wedge dt\right) \geqslant \Phi^*\Omega$, (9.9) 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 (9.10) gives

$$u_\gamma^* \omega_f = \frac{T^2}{4\pi^2} \frac{f'\left(\frac{T}{2\pi} \log |z|\right)}{|z|^2}\, dx \wedge dy \quad \text{for} \quad z = x + iy \in \mathbb{C}^*.$$

Now suppose $f \in \mathcal{T}$ is chosen as in (9.11). 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 9.2, 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. $\square$

REMARK 9.12. It may be useful for some applications to observe that Lemma 9.11 does not require M to be compact. In contrast, the compactness arguments in this lecture almost always depend on the assumption that W and $M_\pm$ are compact—without this, one would need add some explicit assumption to guarantee local C^0-bounds on sequences of holomorphic curves, e.g. the assumption in Theorem 9.2 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 stated above (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 9.13. *Given a stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ on M, a sufficiently small constant $\epsilon > 0$ as in (9.6), 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 9.9, 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 $\kappa \in [-\epsilon, \epsilon]$,

$$\frac{1}{c}(\omega + \kappa\, d\lambda)(X, JX) \leqslant \omega(X, JX) \leqslant c(\omega + \kappa\, d\lambda)(X, JX). \tag{9.12}$$

This uses (9.6) 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 (9.12) 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

$$\begin{aligned}
&c^2 \int_{\dot{\Sigma}} u^* \left(\omega + \tilde{f}(r)\, d\lambda\right) + \frac{2\epsilon}{\delta} \int_{\dot{\Sigma}} u^* \left(\tilde{f}'(r)\, dr \wedge \lambda\right) \\
&\qquad = c^2 \int_{\dot{\Sigma}} u^* \left(\omega + \tilde{f}(r)\, d\lambda + \tilde{f}'(r)\, dr \wedge \lambda\right) = c^2 \int_{\dot{\Sigma}} u^*\omega_{\tilde{f}} \leqslant c^2 E_0(u).
\end{aligned}$$

On the other hand if $c^2 < \frac{2\epsilon}{b-a}$, we can set $\delta := b - a$ and write

$$\begin{aligned} c^2 \int_{\dot{\Sigma}} u^* \left(\omega + \widetilde{f}(r)\, d\lambda \right) + \frac{2\epsilon}{\delta} \int_{\dot{\Sigma}} u^* \left(\widetilde{f}'(r)\, dr \wedge \lambda \right) \\ \leqslant \frac{2\epsilon}{b-a} \int_{\dot{\Sigma}} u^* \left(\omega + \widetilde{f}(r)\, d\lambda + \widetilde{f}'(r)\, dr \wedge \lambda \right) \\ = \frac{2\epsilon}{b-a} \int_{\dot{\Sigma}} u^* \omega_{\widetilde{f}} \leqslant \frac{2\epsilon}{b-a} E_0(u). \end{aligned}$$

□

With this preparation out of the way, we now begin in earnest with the proof of Theorem 9.10. 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 (9.7), 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.

$$E(u_\pm) = \sup_{f \in \mathcal{T}(h,r_0)} \int_{Z_\pm} (u \circ \varphi_\pm)^* \omega_f = \sup_{f \in \mathcal{T}(h,r_0)} \int_{\dot{\mathbb{D}}} u^* \omega_f = E(u).$$

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 9.14. *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 9.3. Suppose the contrary, so there exists a sequence $z_k = (s_k, t_k) \in Z_+$ with $R_k := |du_+(z_k)| \to \infty$. Choose a sequence $\epsilon_k > 0$ with $\epsilon_k \to 0$ but $\epsilon_k R_k \to \infty$, and using Lemma 9.4, assume without loss of generality that

$$|du_+(z)| \leqslant 2R_k \quad \text{for all} \quad z \in \mathbb{D}_{\epsilon_k}(z_k).$$

Define a rescaled sequence of J-holomorphic disks by

$$v_k : \mathbb{D}_{\epsilon_k R_k} \to \widehat{W} : z \mapsto u \circ \varphi_+(z_k + z/R_k).$$

These satisfy $|dv_k| \leqslant 2$ 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: $v_k(0)$ has a bounded subsequence.

Then the corresponding subsequence of $v_k : \mathbb{D}_{\epsilon_k R_k} \to \widehat{W}$ is uniformly C^1-bounded on every compact subset and thus (by Proposition 9.1) has a further subsequence convergent in $C^\infty_{\mathrm{loc}}(\mathbb{C})$ to a J-holomorphic plane

$$v_\infty : \mathbb{C} \to \widehat{W}$$

with $|dv_\infty(0)| = \lim_{k\to\infty} |dv_k(0)| = 1$. But by the same argument we used in the proof of Lemma 9.3, the fact that $\int_{Z_+} u_+^*\omega_f < \infty$ for any choice of $f \in \mathcal{T}(h, r_0)$ implies

$$\int_{\mathbb{C}} v_\infty^*\omega_f = 0,$$

hence v_∞ is constant, and this is a contradiction.

Case 2: $v_k(0)$ has a subsequence diverging to $\{+\infty\} \times M_+$. Restricting to this subsequence, suppose

$$v_k(0) \in \{r_k\} \times M_+,$$

so $r_k \to \infty$, and assume without loss of generality that $r_k > r_0$ for all k. Let $\tilde{R}_k \in (0, \epsilon_k R_k]$ for each k denote the largest radius such that $v_k(\mathbb{D}_{\tilde{R}_k}) \subset (r_0, \infty) \times M_+$. Then $\tilde{R}_k \to \infty$ since $|dv_k|$ is bounded. Now using the $\mathbb{R}$-translation maps τ_r defined in (9.8), define

$$\tilde{v}_k := \tau_{-r_k} \circ v_k|_{\mathbb{D}_{\tilde{R}_k}} : \mathbb{D}_{\tilde{R}_k} \to \mathbb{R} \times M_+.$$

Since we're using a translation-invariant metric on $[r_0, \infty) \times M_+$, $\tilde{v}_k$ is now a uniformly C^1_{loc}-bounded sequence of maps into $\mathbb{R} \times M_+$. Proposition 9.1 thus provides a subsequence convergent in $C^\infty_{\mathrm{loc}}(\mathbb{C})$ 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,

$$E(v_\infty) < \infty \quad \text{and} \quad \int_{\mathbb{C}} v_\infty^*\omega_+ = 0, \tag{9.13}$$

where $E(v_\infty)$ is now defined as in Remark 9.9. By Lemma 9.13, 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 $R > 0$,

$$\int_{\mathbb{D}_R} v_\infty^*\Omega_f^+ = \lim_{k\to\infty} \int_{\mathbb{D}_R} v_k^*\tau_{-r_k}^*\Omega_f^+ = \lim_{k\to\infty} \int_{\mathbb{D}_R} v_k^*\Omega_{f_k}^+,$$

where $f_k(r) := f(r - r_k)$. Notice that the dependence of the last integral on the function f_k is limited to the interval $(r_0, \infty) \subset \mathbb{R}$ in its domain since $v_k(\mathbb{D}_R) \subset (r_0, \infty) \times M_+$. Then since $f > a > h(r_0)$ by assumption, there exists for each k a function $h_k \in \mathcal{T}(h, r_0)$ that matches f_k outside some neighborhood of $(-\infty, r_0]$ and thus satisfies

$$\int_{\mathbb{D}_R} v_k^*\Omega_{f_k}^+ = \int_{\mathbb{D}_R} v_k^*\omega_{h_k} \leqslant \int_{\mathbb{D}_{\epsilon_k R_k}} v_k^*\omega_{h_k} = \int_{\mathbb{D}_{\epsilon_k}(z_k)} u_+^*\omega_{h_k} \leqslant \int_{Z_+} u_+^*\omega_{h_k} \leqslant E(u).$$

This is true for every $R > 0$ and thus proves the first part of (9.13). To establish the second part, fix $R > 0$ again and pick any $f \in \mathcal{T}(h, r_0)$. Observe that since we can assume (after perhaps passing to a subsequence) the disks $\mathbb{D}_{\epsilon_k}(z_k)$ are all disjoint,

$$\begin{aligned} 0 &= \lim_{k\to\infty} \int_{\mathbb{D}_{\epsilon_k}(z_k)} u_+^* \omega_f = \lim_{k\to\infty} \int_{\mathbb{D}_{\epsilon_k R_k}} v_k^* \omega_f = \lim_{k\to\infty} \int_{\mathbb{D}_{\epsilon_k R_k}} \tilde{v}_k^* \tau_{r_k}^* \omega_f \\ &\geqslant \lim_{k\to\infty} \int_{\mathbb{D}_R} \tilde{v}_k^* \tau_{r_k}^* \omega_f = \lim_{k\to\infty} \int_{\mathbb{D}_R} \tilde{v}_k^* \Omega_{f_k}^+, \end{aligned}$$

where now $f_k(r) := f(r + r_k)$. Writing $\Omega_{f_k}^+ = \omega_+ + d\big(f_k(r)\,\lambda_+\big) = \omega_+ + f_k(r)\,d\lambda_+ + f_k'(r)\,dr \wedge \lambda_+$, we can choose f such that $f'(r) = f'(r + r_k) \to 0$ as $k \to \infty$, so the third term contributes nothing to the integral. For the second term, let $f_+ := \lim_{k\to\infty} f_k(r) = \lim_{r\to\infty} f(r)$, 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 (9.6) 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 9.11 and conclude that v_∞ is constant, contradicting the fact that $|dv_\infty(0)| = 1$.

Case 3: $v_k(0)$ 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 9.11. □

Consider now a sequence $s_k \to \infty$ and construct the J-holomorphic half-cylinders

$$u_k : [-s_k, \infty) \times S^1 \to \widehat{W} : (s,t) \mapsto u_+(s + s_k, t).$$

The derivatives $|du_k|$ are uniformly bounded due to Lemma 9.14, though again, u_k might fail to be uniformly bounded in C^0. We distinguish three cases.

Case 1: $u_k(0,0)$ has a bounded subsequence.
Then the corresponding subsequence of u_k 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 $f \in \mathcal{T}(h, r_0)$ and any $c > 0$, we have

$$\begin{aligned} \int_{[-c,c]\times S^1} u_\infty^* \omega_f &= \lim_{k\to\infty} \int_{[-c,c]\times S^1} u_k^* \omega_f \leqslant \lim_{k\to\infty} \int_{[-s_k/2,\infty)\times S^1} u_k^* \omega_f \\ &= \lim_{k\to\infty} \int_{[s_k/2,\infty)\times S^1} u_+^* \omega_f = 0 \end{aligned} \tag{9.14}$$

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_k,

$$u_+(s_k, \cdot) \to p \quad \text{in } C^\infty(S^1, \widehat{W}) \qquad \text{as } k \to \infty.$$

Case 2: $u_k(0,0)$ has a subsequence diverging to $\{+\infty\} \times M_+$. Passing to the corresponding subsequence of u_k, suppose

$$u_k(0,0) \in \{r_k\} \times M_+,$$

so $r_k \to \infty$. Since the derivatives $|du_k|$ are uniformly bounded, we can then find a sequence of intervals $[-R_k^-, R_k^+] \subset [-s_k, \infty)$ such that

$$u_k([-R_k^-, R_k^+] \times S^1) \subset [r_0, \infty) \times M_+ \quad \text{and} \quad R_k^\pm \to \infty.$$

Now the translated sequence

$$\tau_{-r_k} \circ u_k|_{[-R_k^-, R_k^+]\times S^1} : [-R_k^-, R_k^+] \times S^1 \to \mathbb{R} \times M_+$$

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 (9.13) and (9.14) above, so I will leave it as such. Lemma 9.11 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

$$u_\infty(s,t) = (\pm Ts + a, \gamma(\pm Tt)) \tag{9.15}$$

for some constant $a \in \mathbb{R}$ and a choice of signs to be determined below (see Lemma 9.18).

Case 3: $u_k(0,0)$ 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 $u_k(0,0) \in \{-r_k\} \times M_-$ with $r_k \to \infty$ and obtain a subsequence for which $\tau_{r_k} \circ u_k$ 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 (9.15) 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 9.15. *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*

$$\mathrm{dist}_{C^0}(u_+(s, \cdot), \ell) < \delta,$$

where $\ell : S^1 \to \widehat{W}$ either is constant or is a loop of the form $\ell(t) = (r, \gamma(\pm Tt))$ 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 $T > 0$.

PROOF. If not, then there exists a sequence $s_k \to \infty$ such that each of the loops $u_+(s_k, \cdot)$ 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 $u(s_k, \cdot)$ becomes arbitrarily C^∞-close to such a loop, so this is a contradiction. □

LEMMA 9.16. *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 $(s_k, t_k) \in Z_+$ such that $u_+(s_k, t_k)$ diverges to $\{+\infty\} \times M_+$. This implies $s_k \to \infty$, 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 $(s_k', t_k') \in Z_+$ with $s_k' \to \infty$ such that $u_+(s_k', t_k') \notin (R, \infty) \times M_+$ for every k. By continuity, we are free to suppose $u_+(s_k', t_k') \in \{R\} \times M_+$ for all k since Lemma 9.15 implies $u_+(\{s_k\} \times S^1) \subset (2R, \infty) \times M_+$ for k sufficiently large. Using Lemma 9.15 again, we also have

$$u_+(\{s_k'\} \times S^1) \subset (R - 1, R + 1) \times M_+$$

for all k large. Assuming $2R > R+2$ without loss of generality, we can therefore find infinitely many pairwise disjoint annuli of the form $[s_k', s_j] \times S^1 \subset Z_+$ containing open sets that u maps properly to small balls centered at points in $\{R + 2\} \times M_+$. Choosing any $f \in \mathcal{T}(h, r_0)$, the monotonicity lemma implies that each of these contributes at least some fixed amount to $\int_{Z_+} u_+^* \omega_f$, contradicting the assumption that $E(u) < \infty$.[2]

A similar argument works if $u_+(s_k, t_k)$ diverges to $\{-\infty\} \times M_-$, proving that for every $R \geqslant r_0$, there exists $s_0 \geqslant 0$ with

$$u_+((s_0, \infty) \times S^1) \subset (-\infty, -R) \times M_-.$$

□

If u is bounded, then the singularity at 0 is removable by Theorem 9.2. If not, then Lemma 9.16 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 9.17. *If the puncture is positive/negative, then the limit*

$$Q := \lim_{s \to \infty} \int_{S^1} u_+(s, \cdot)^* \lambda_\pm \in \mathbb{R}$$

[2]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 $W^{2R} \subset \widehat{W}$.

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

$$\int_{[s_0,\infty)\times S^1} \left|u_+^* d\lambda_+\right| < \infty. \tag{9.16}$$

We claim first that $\int_{[s_0,\infty)\times S^1} u_+^*\omega_+ < \infty$. Indeed, for any $s > s_0$ and $f \in \mathcal{T}(h, r_0)$, 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 (9.16).

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 9.10 as the **charge** of the puncture.

LEMMA 9.18. *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 $s_k \to \infty$ with $u_+(s_k, 0) \in \{\pm r_k\} \times M_\pm$, one can find a sequence $R_k \in [0, s_k]$ with $R_k \to \infty$ such that u_+ maps $[s_k - R_k, \infty) \times S^1$ into the positive/negative cylindrical end for every k, and the sequence of half-cylinders*

$$u_k : [-R_k,\infty) \times S^1 \to \mathbb{R} \times M_+ \qquad \text{or} \qquad u_k : (-\infty, R_k] \times S^1 \to \mathbb{R} \times M_-$$

defined by $u_k(s,t) = \tau_{\mp r_k} \circ u_\pm(s \pm s_k, t)$ has a subsequence convergent in $C^\infty_{\mathrm{loc}}(\mathbb{R} \times S^1)$ 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 9.15, we showed that the sequence $u'(s,t) := \tau_{\mp r_k} \circ u_+(s + s_k, t)$ defined on $[-R_k, \infty) \times S^1$ 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

$$u'_\infty(s,t) = (\sigma T s + a, \gamma(\sigma T t)) \tag{9.17}$$

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

$$0 \neq Q = \lim_{s\to\infty} \int_{S^1} u_+(s,\cdot)^*\lambda_\pm = \lim_{k\to\infty} \int_{S^1} u_k'(0,\cdot)^*\lambda_\pm = \int_{S^1} u_\infty'(0,\cdot)^*\lambda_\pm,$$

so u_∞' cannot be constant, and from (9.17) 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 $s_k \to \infty$ admits a subsequence for which

$$\partial_s u_\mathbb{R}(s_k,\cdot) \to Q \quad \text{in} \quad C^\infty(S^1,\mathbb{R}),$$

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 9.19. *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 $s_k \to \infty$, there exists a point $p \in M_+$ such that $u_+(s_k, 0) \in \{r_k\} \times M_+$ with $r_k \to \infty$ and

$$\tau_{-r_k} \circ u_+(s_k,\cdot) \to (0,p) \in \mathbb{R} \times M_+ \quad \text{in} \quad C^\infty(S^1, \mathbb{R} \times M_+) \text{ as } k \to \infty.$$

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

$$\begin{aligned}
\int_{[s_0,s_1]\times S^1} v_\delta^*\Omega_f^+ &= \int_{[s_0,s_1]\times S^1} v_\delta^*\omega_+ + \int_{[s_0,s_1]\times S^1} v_\delta^* d\,(f(r)\,\lambda_+) \\
&= \int_{[s_0,s_1]\times S^1} v_\delta^*\omega_+ + \int_{S^1} \left[v_\delta(s_1,\cdot)^*\,(f(r)\,\lambda_+) - v_\delta(s_0,\cdot)^*\,(f(r)\,\lambda_+)\right].
\end{aligned}$$

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 9.10 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 9.20. *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 $s_k \to \infty$ is a sequence and $\gamma : \mathbb{R}/T\mathbb{Z} \to M_\pm$ is a Reeb orbit such that

$$u_M(s_k, \cdot) \to \gamma(T\cdot) \quad \text{in} \quad C^\infty(S^1, M_\pm).$$

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 $s_k' \to \infty$ with $u_M(s_k, \cdot) \notin \mathcal{U}$ for all k. We can nonetheless restrict to a subsequence for which $u_M(s_k', \cdot)$ 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 $s_k'' \to \infty$ for which each $u_M(s_k'', 0)$ lies in the region $\mathcal{V}$ some fixed distance away from $\operatorname{im}\gamma$. There must then be a subsequence for which $u_M(s_k'', \cdot)$ converges to another T-periodic orbit, but this is impossible since no such orbits exist in $\mathcal{V}\setminus \operatorname{im}\gamma$. $\square$

9.3. 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 $J_k \to J \in \mathcal{J}_\tau(\omega_h, r_0, \mathcal{H}_+, \mathcal{H}_-)$ 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 to the cylindrical ends by

$$J_+ := J|_{[r_0,\infty)\times M_+} \in \mathcal{J}(\mathcal{H}_+), \qquad J_- := J|_{(-\infty,-r_0]\times M_-} \in \mathcal{J}(\mathcal{H}_-).$$

Suppose

$$u_k := [(\Sigma_k, j_k, \Gamma_k^+, \Gamma_k^-, \Theta_k, u_k)] \in \mathcal{M}_{g,m}(J_k, A_k, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$$

is a sequence of J_k-holomorphic curves in $\widehat{W}$ with fixed genus $g \geqslant 0$ and $m \geqslant 0$ marked points, varying relative homology classes $A_k \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ and fixed collections of asymptotic orbits $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \dots, \gamma_{m_\pm}^\pm)$. Observe that the energies $E(u_k)$ depend only on the orbits $\boldsymbol{\gamma}^\pm$ and relative homology classes A_k, so in particular, $E(u_k)$ is uniformly bounded whenever the relative homology class is also fixed. The fundamental question of this section is:

QUESTION. *If $E(u_k)$ is uniformly bounded and no subsequence of u_k 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?*

9.3.1. Bubbling. Suppose the pointed Riemann surfaces $(\Sigma_k, j_k, \Gamma_k^+ \cup \Gamma_k^- \cup \Theta_k)$ form a convergent sequence, meaning we can assume after biholomorphic reparametrization that the surfaces $\Sigma_k = \Sigma$ are all identical with identical sets of punctures $\Gamma_k^\pm = \Gamma^\pm$ and marked points $\Theta_k = \Theta$, while their complex structures are C^∞-convergent

$$j_k \to j \in \mathcal{J}(\Sigma).$$

Suppose additionally that there exists a point $\zeta_0 \in \dot{\Sigma}$ such that $u_k(\zeta_0) \in \widehat{W}$ is contained in a compact subset for all k, and that for some choice of Riemannian metrics on $\dot{\Sigma} = \Sigma\setminus\Gamma$ and $\widehat{W}$ that are translation-invariant on the cylindrical ends of both, the maps $u_k : \dot{\Sigma} \to \widehat{W}$ are locally C^1-bounded outside some finite subset

$$\Gamma' = \{\zeta_1, \dots, \zeta_N\} \subset \dot{\Sigma},$$

i.e. for every compact set $K \subset \dot{\Sigma}\setminus\Gamma'$, there exists a constant $C_K > 0$ independent of k such that

$$|du_k| \leqslant C_K \quad \text{on } K.$$

Then Proposition 9.1 gives a subsequence that converges in $C^\infty_{\mathrm{loc}}(\dot{\Sigma}\setminus\Gamma')$ to a J-holomorphic curve

$$u_\infty : \dot{\Sigma}\setminus\Gamma' \to \widehat{W}$$

with $E(u_\infty) \leqslant \limsup E(u_k) < \infty$, 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_k, but let's assume this for now (§9.3.2 below discusses 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_k 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 $z_k^i \to \zeta_i$ such that $|du_k(z_k^i)| =: R_k \to \infty$, along with a sequence $\epsilon_k \to 0$ with $\epsilon_k R_k \to \infty$, and using Lemma 9.4, assume without loss of generality that $|du_k(z)| \leqslant 2R_k$ for all z in the ϵ_k-ball about z_k^i. 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 $z_k^i \to 0$ in these coordinates, and assume without loss of generality that they identify our chosen metric near ζ_i with the Euclidean metric. Now setting

$$v_k^i(z) = u_k(z_k^i + z/R_k) \quad \text{for} \quad z \in \mathbb{D}_{\epsilon_k R_k}$$

gives a sequence of J_k-holomorphic maps $v_k^i : \mathbb{D}_{\epsilon_k R_k} \to \widehat{W}$ whose energies and first derivatives are both uniformly bounded. As in the arguments of §2, we now have three possibilities:

- If $u_k(z_k^i)$ has a bounded subsequence, then the corresponding subsequence of v_k^i converges in $C^\infty_{\mathrm{loc}}(\mathbb{C})$ to a J-holomorphic plane $v_\infty^i : \mathbb{C} \to \widehat{W}$ with finite energy.
- If $u_k(z_k^i)$ has a subsequence diverging to $\{\pm\infty\} \times M_\pm$, then translating v_k^i 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 $S^2 \setminus \{\infty\}$, 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_k has "bubbled off a sphere" at ζ_i. Alternatively, v_∞^i may be positively or negatively asymptotic to a Reeb orbit at ∞.

Figure 9.2 shows two scenarios that could occur for a sequence in which $|du_k|$ 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 $u_k(z_k^1)$ 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_k near ζ_1 to show that it must be true, i.e. "bubbles connect". At ζ_3, we have $u_k(z_k^3) \to \{-\infty\} \times M_-$ 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 $u_k(z_k^2) \to \{+\infty\} \times M_+$, 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, $u_k(z_k^2)$ 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_k near ζ_i. One good place to read about the analysis of bubble trees is [**HWZ03**, §4].

EXERCISE 9.21. Fill in the gaps previously left in the proof of Proposition 7.9 regarding the properness of the action of $\mathrm{Diff}(\Sigma, \Theta)$ on $\mathcal{J}(\Sigma)$. Concretely, consider a closed connected surface Σ with two C^∞-convergent sequences of complex structures $j_\nu \to j$ and $j'_\nu \to j'$, along with a sequence of biholomorphic maps $\varphi_\nu : (\Sigma, j'_\nu) \to (\Sigma, j_\nu)$.

(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 $\Sigma = S^2$, then bubbling can occur, but 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 $C^\infty_{\mathrm{loc}}(S^2 \backslash \{\zeta\})$ to a constant, and derive a contradiction from this if there is a set of at least three points $\Theta \subset S^2$ that are fixed by every φ_ν. *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 $\nu \to \infty$.*

9.3.2. Breaking. Figure 9.2 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 9.3 and 9.4 show three such scenarios, where we assume again that $j_k \to j$ and $\dot{\Sigma} = \Sigma \backslash \Gamma$ and $\widehat{W}$ carry Riemannian metrics that are translation-invariant on the cylindrical ends such that

$$|du_k| \leqslant C \quad \text{everywhere on } \dot{\Sigma}$$

for some constant $C > 0$ independent of k. This is a stronger condition than we had in §9.3.1, and if there exists a point $\zeta_0 \in \dot{\Sigma}$ such that $u_k(\zeta_0) \in \widehat{W}$ is bounded, it implies that u_k has a subsequence converging in $C^\infty_{\mathrm{loc}}(\dot{\Sigma})$ to a J-holomorphic map

$$u_\infty : \dot{\Sigma} \to \widehat{W}$$

with $E(u_\infty) \leqslant \limsup E(u_k) < \infty$. Convergence in C^∞_{loc} is, however, not very strong: there may in general be no relation between the asymptotic behavior of u_∞ and u_k at corresponding punctures, e.g. the top scenario in Figure 9.3 shows a case in which a negative puncture of u_k becomes a removable singularity of u_∞. Whenever

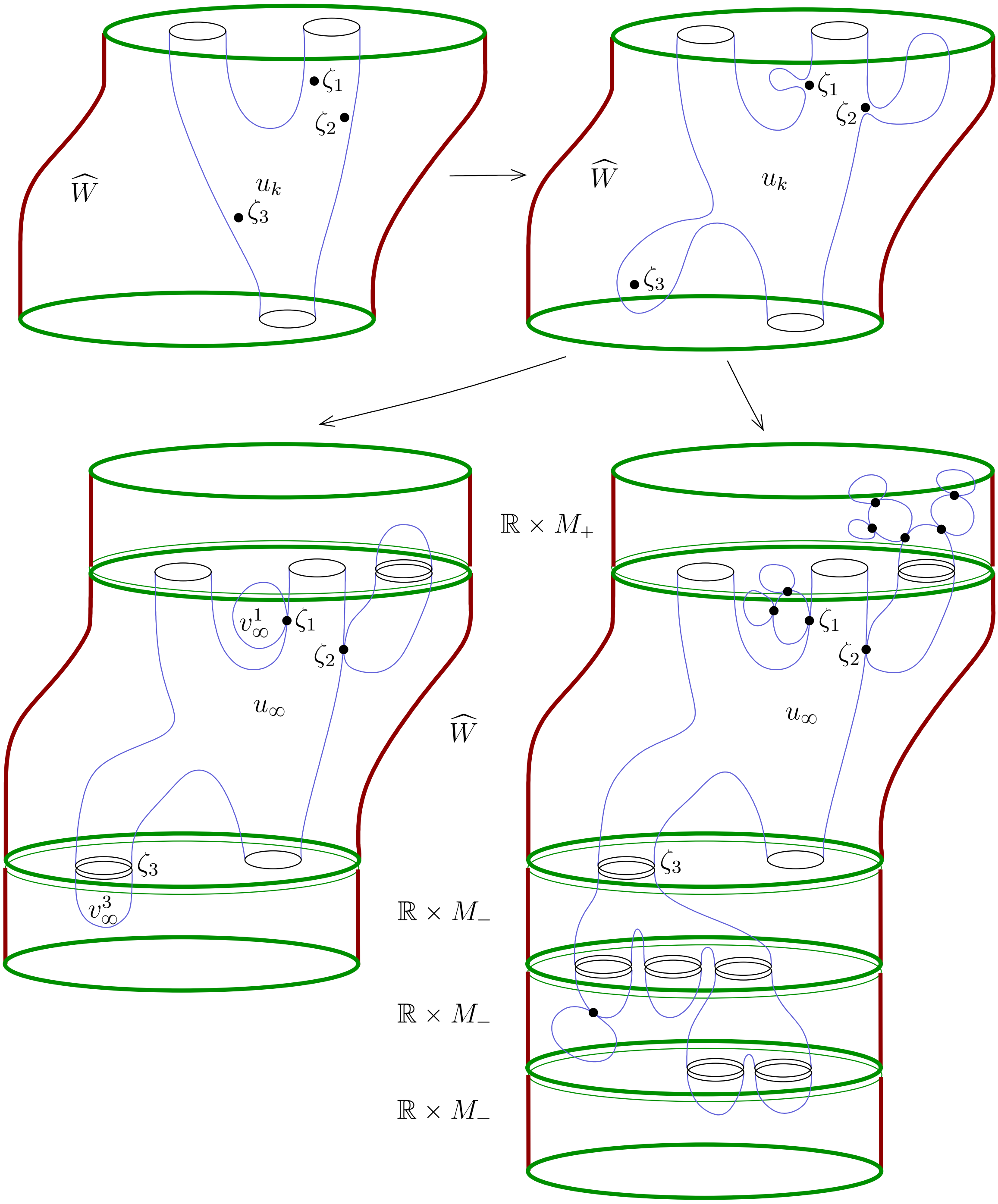

FIGURE 9.2. Two possible pictures of spheres and/or planes that can bubble off when the first derivative blows up near three points.

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_k and find a sequence $s_k \to \infty$ such that the sequence of half-cylinders

$$(-\infty, s_k] \times S^1 \to \widehat{W} : (s,t) \mapsto u_k(s - s_k, t)$$

is uniformly C^1-bounded and thus converges in $C^\infty_{\mathrm{loc}}(\mathbb{R} \times S^1)$ 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_k.

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_k, we can find a sequence $s_k \to \infty$ such that the half-cylinders

$$[-s_k, \infty) \times S^1 \to \widehat{W} : (s,t) \mapsto u_k(s + s_k, t)$$

are uniformly C^1-bounded and converge in $C^\infty_{\mathrm{loc}}(\mathbb{R} \times S^1)$ 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 $s_k \to \infty$, 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.

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_k 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 9.4, we see a scenario in which u_k satisfies the same conditions as above, except that instead of $u_k(\zeta_0)$ being bounded, it diverges to $\{+\infty\} \times M_+$. It follows that after applying suitable $\mathbb{R}$-translations, a subsequence converges in $C^\infty_{\mathrm{loc}}(\dot{\Sigma})$ 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_k. Now observe that since u_k has a negative cylindrical end $(-\infty, 0] \times S^1 \subset \dot{\Sigma}$, one can necessarily find a sequence $s_k \to \infty$ such that $u_k(-s_k, 0)$ is bounded, and the sequence of half-cylinders

$$(-\infty, s_k] \times S^1 \to \widehat{W} : (s,t) \mapsto u_k(s - s_k, t)$$

is then uniformly C^1-bounded and thus has a subsequence convergent in $C^\infty_{\mathrm{loc}}(\mathbb{R} \times S^1)$ to a finite-energy J-holomorphic cylinder $v_0 : \mathbb{R} \times S^1 \to \widehat{W}$. In the picture, v_0 has

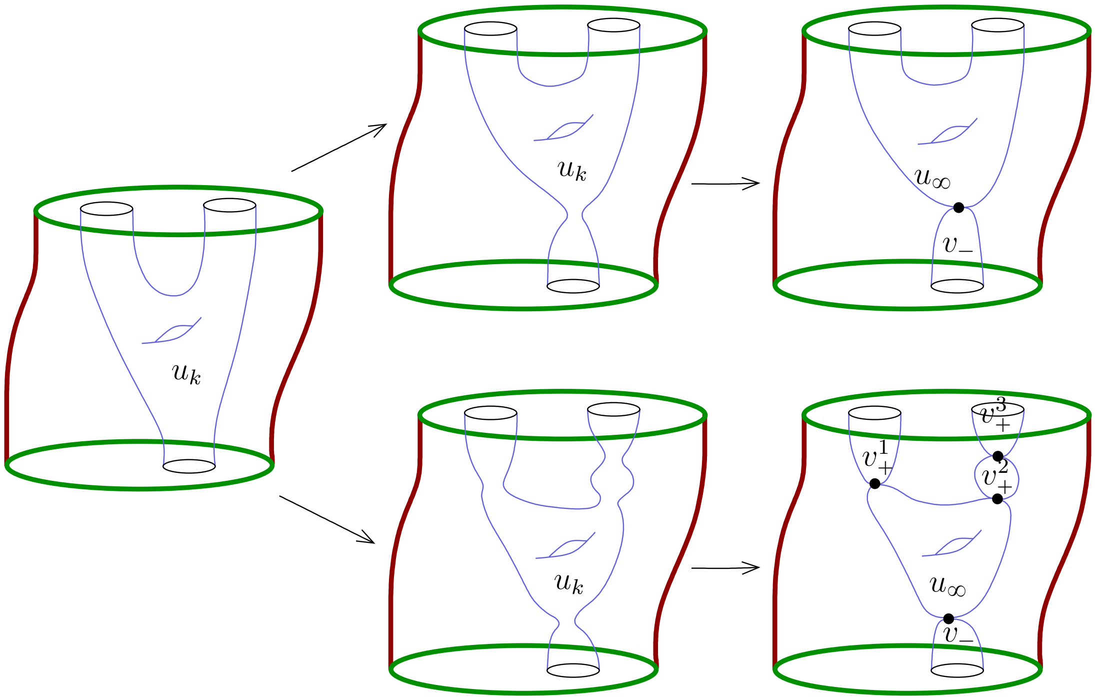

FIGURE 9.3. Even with fixed conformal structures on the domains and without bubbling, a sequence of punctured holomorphic curves in $\widehat{W}$ can break to produce multiple curves in $\widehat{W}$ with extra removable punctures. The picture shows two such scenarios.

both a positive and a negative puncture, but its negative end again approaches a different Reeb orbit from the negative ends of u_k, so one can deduce that there must be still more happening near $-\infty$: there exists another sequence $s_k' \to \infty$ with $s_k' - s_k \to \infty$ such that suitable $\mathbb{R}$-translations of the half-cylinders

$$(-\infty, s_k] \times S^1 \to (-\infty, -r_0] \times M_- : (s,t) \mapsto u_k(s - s_k', t)$$

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_k 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 $s_k \to \infty$ such that $u_k(s_k, 0) \in \{r_k\} \times M_+$ for some $r_k \to \infty$, and suitable $\mathbb{R}$-translations of

$$[-s_k, \infty) \times S^1 \to [r_0, \infty) \times M_+ : (s,t) \mapsto u_k(s + s_k, t)$$

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

$$v_+^2 : \mathbb{R} \times S^1 \to \mathbb{R} \times M_+$$

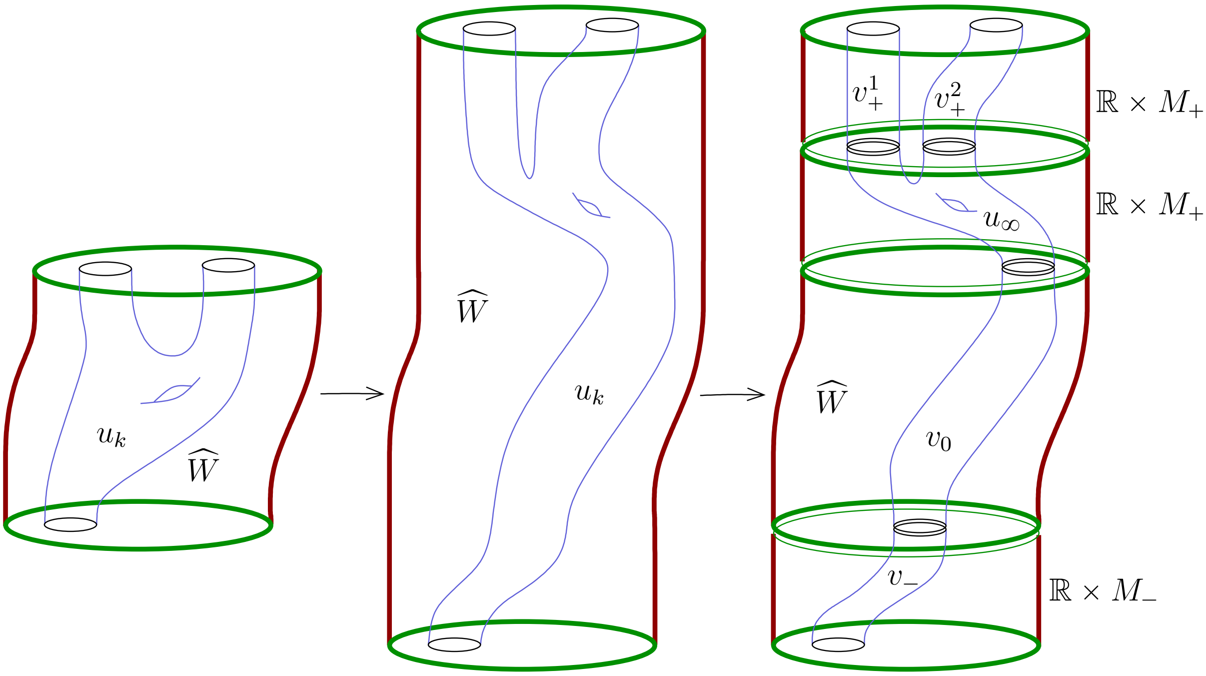

Figure 9.4. Different portions of a breaking sequence of curves may also become infinitely far apart in the limit, so that some live in $\widehat{W}$ while others live in the symplectization of M_+ or M_-.

that connects the errant asymptotic orbit of u_∞ to the corresponding orbit of u_k. One can now perform the same trick at the other positive end of $\dot{\Sigma}$, as there necessarily also exists a sequence $s'_k \to \infty$ in this end such that $u_k(s'_k, 0) \in \{r_k\} \times M_+$ for the same sequence $r_k \to \infty$ 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 9.4 happens whenever the sequence u_k 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_k as $k \to \infty$: the middle picture in Figure 9.4 deals with this by expanding the scale of the cylindrical ends so that the convergence to upper and lower levels becomes visible.

9.3.3. The Deligne-Mumford space of Riemann surfaces. We next need to relax the assumption that the pointed Riemann surfaces $(\Sigma_k, j_k, \Gamma_k^+ \cup \Gamma_k^- \cup \Theta_k)$ converge. Recall that for integers $g \geqslant 0$ and $\ell \geqslant 0$, the moduli space of pointed Riemann surfaces is the space of equivalence classes

$$\mathcal{M}_{g,\ell} = \{(\Sigma, j, \Theta)\} \big/ \sim,$$

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. $\mathcal{M}_{0,\ell}$ 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, one can show that every pointed Riemann surface has a finite automorphism group (see Proposition 7.9), and $\mathcal{M}_{g,\ell}$ 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**. We shall now give a sketch of this construction from the perspective of hyperbolic geometry; for more details, see [**Hum97**, **SS92**].

We recall first the following standard result.

THEOREM 9.22 (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* pointed Riemann surfaces are thus quotients of $(\mathbb{H}, i)$.

PROPOSITION 9.23. *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

$$\begin{aligned} \operatorname{Aut}(\mathbb{H}, i) &= \left\{ \varphi(z) = \frac{az + b}{cz + d} \;\middle|\; a, b, c, d \in \mathbb{R}, \quad ad - bc = 1 \right\} \Big/ \{\pm 1\} \\ &= \mathrm{SL}(2, \mathbb{R}) / \{\pm 1\} =: \mathrm{PSL}(2, \mathbb{R}), \end{aligned}$$

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. For more details on all of this, the book by Hummel [**Hum97**] is highly recommended. $\square$

By lifting to universal covers, this implies the following.

COROLLARY 9.24. *For every pointed Riemann surface* (Σ, j, Θ) *with* $\chi(\Sigma\backslash\Theta) < 0$*, the punctured Riemann surface* $(\Sigma\backslash\Theta, j)$ *admits a complete Riemannian metric* g_j *of constant curvature* -1 *that defines the same conformal structure as* j*, and has the property that 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**, and it is uniquely determined by j. 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 9.24 in the case $\Theta = \varnothing$ without assuming Theorem 9.22, see [**Tro92**].

Every nontrivial class in $\pi_1(\dot{\Sigma})$ contains a unique geodesic for g_j. Now suppose $C \subset \dot{\Sigma}$ is a union of disjoint embedded geodesics such that each connected component of $\dot{\Sigma}\backslash C$ has the homotopy type of a disk with two holes. The components are then called **singular pairs of pants**, and the result is called a **pair-of-pants decomposition** of $(\dot{\Sigma}, j)$. Two examples for the case $g = 1$ and $\ell = 3$ are shown in Figure 9.5.

A pair-of-pants decomposition for (Σ, j, Θ) gives rise to a local parametrization of $\mathcal{M}_{g,\ell}$ near $[(\Sigma, j, \Theta)]$, known as the *Fenchel-Nielsen coordinates*. These consist of two parameters that can be associated to each of the geodesics $\gamma \subset \Sigma$ in the decomposition, namely the length $\ell(\gamma) > 0$ of the geodesic and a *twist* parameter $\theta(\gamma) \in S^1$, which describes how the two neighboring pairs of pants are glued together along γ. Note that by computing Euler characteristics, there are always exactly $-\chi(\Sigma\backslash\Theta) = 2g - 2 + \ell$ pairs of pants in a decomposition, so that the total number of geodesics involved is $[3(2g - 2 + \ell) - \ell]/2 = 3g - 3 + \ell$, thus one can read off the formula $\dim \mathcal{M}_{g,\ell} = 6g - 6 + 2\ell$ from this geometric picture.

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 that the latter possibility is an illusion, but one may need to switch to a different pair-of-pants decomposition to see why:

THEOREM. *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 pointed Riemann surfaces to degenerate is when some of the bounding geodesics in a pair-of-pants decomposition shrink to length zero. Figure 9.6 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 $\overline{\mathcal{M}}_{1,3}$.

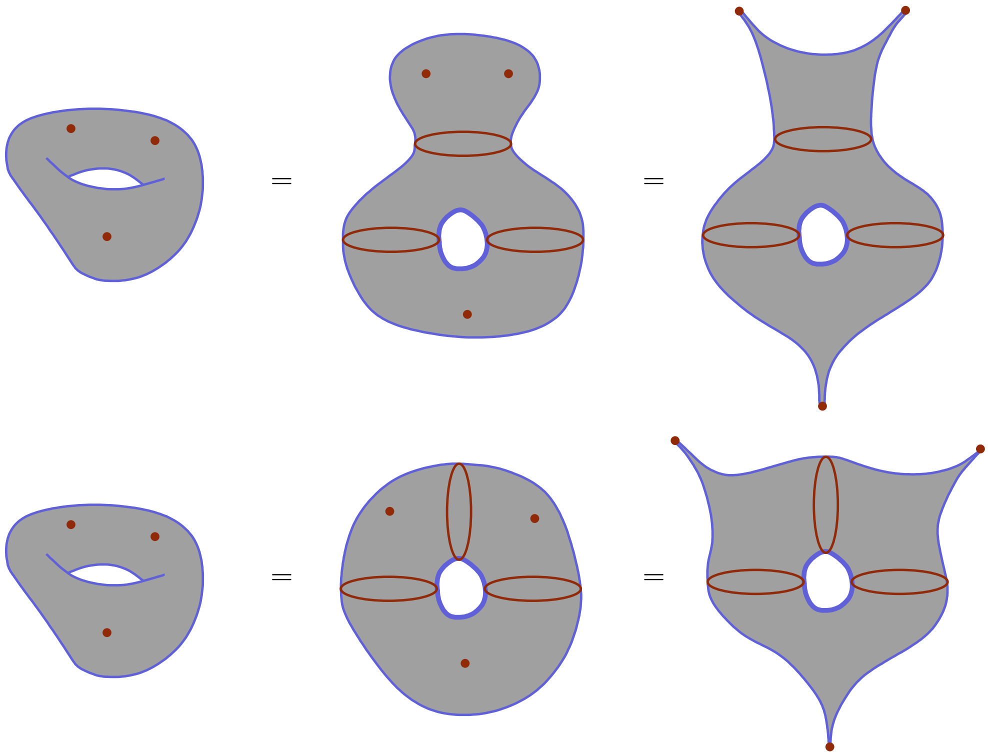

FIGURE 9.5. Two distinct pair-of-pants decompositions for the same genus 1 Riemann surface with three marked points. The decompositions are shown from two perspectives: the pictures at the right are meant to give a more accurate impression of the Poincaré metric, which becomes singular and forms a cusp at each marked point.

DEFINITION 9.25. A **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 $\Theta \subset S$;
- An unordered set of $2N$ points $\Delta \subset S \backslash \Theta$ equipped with an involution $\sigma : \Delta \to \Delta$. 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 $\varphi : (S, j) \to (S', j')$ 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 9.6. We can think of the *stable* nodal surfaces as precisely those which admit (possibly singular) pair-of-pants decompositions. All nodal Riemann

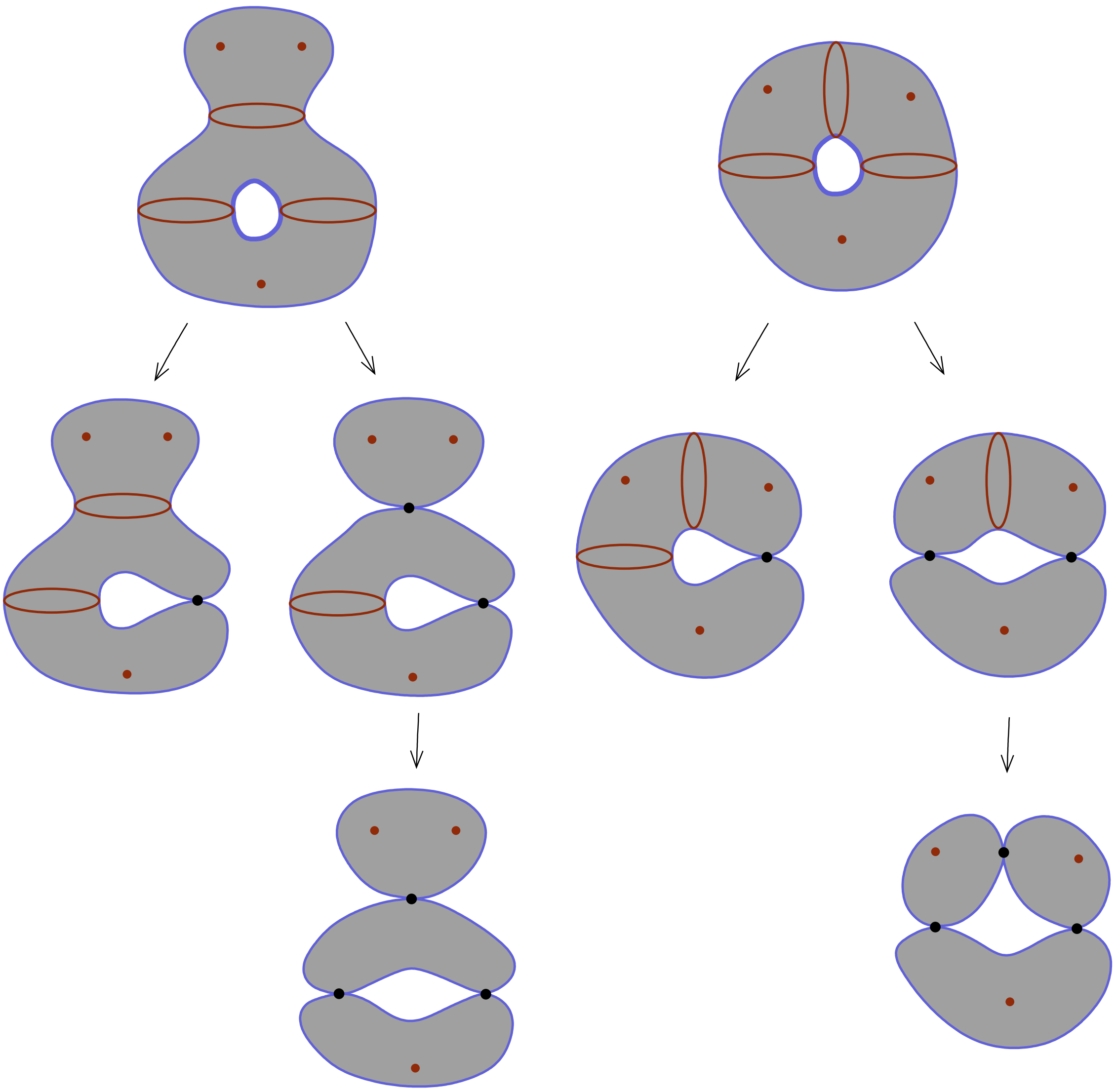

FIGURE 9.6. Starting from each of the pair-of-pants decompositions for the $g = 1$ and $\ell = 3$ case from Figure 9.5, shrinking geodesic lengths to zero produces various examples of stable nodal Riemann surfaces belonging to $\overline{\mathcal{M}}_{1,3}$.

surfaces we consider will be assumed connected in the sense defined above unless otherwise noted; note that S itself can nonetheless be disconnected, as is the case in four out of the six nodal surfaces shown in Figure 9.6.

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 9.25. We shall denote the collection of special circles in $\widehat{S}_\Phi$ where boundray 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 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 $2g + \ell \geqslant 3$, define $\overline{\mathcal{M}}_{g,\ell}$ as the set 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 pointed 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 subset. Rather than formulating all of this in precise terms, let us state the main corollary that is important to know in practice.

THEOREM 9.26. *Fix $g \geqslant 0$ and $\ell \geqslant 0$ with $2g + \ell \geqslant 3$. Then for any sequence $[(\Sigma_k, j_k, \Theta_k)] \in \mathcal{M}_{g,\ell}$, there exists a stable nodal Riemann surface $[(S, j, \Theta, \Delta)] \in \overline{\mathcal{M}}_{g,\ell}$ such that after restricting to a subsequence,*

$$[(\Sigma_k, j_k, \Theta_k)] \to [(S, j, \Theta, \Delta)]$$

in the following sense: (S, j, Θ, Δ) admits a decoration Φ such that for sufficiently large k, there are homeomorphisms

$$\varphi : \widehat{S}_\Phi \to \Sigma_k,$$

smooth outside of C_Φ, which map Θ to Θ_k preserving the ordering and satisfy

$$\varphi^* j_k \to j \quad \text{in} \quad C^\infty_{\mathrm{loc}}(\widehat{S}_\Phi \backslash C_\Delta).$$

As one might gather from the above statement, one could just as well define a compact metrizable topology on the space of equivalence classes of *decorated* nodal Riemann surfaces and then characterize the topology of $\overline{\mathcal{M}}_{g,\ell}$ via the natural projection that forgets the decorations.

EXERCISE 9.27. 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 4-pointed Riemann sphere $(S^2, j, (z_1, z_2, z_3, z_4))$ 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?

9.4. The SFT compactness theorem

We now introduce the natural compactification of $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$.

9.4.1. 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 nodal Riemann surface, with $\Gamma = \Gamma^+ \cup \Gamma^-$ 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 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 and only 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 nodal Riemann surface $(S, j, \Gamma \cup \Theta, \Delta)$ need not be stable in general.

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 $J_k \to J$ and a sequence $[(\Sigma_k, j_k, \Theta_k, u_k)] \in \mathcal{M}_{g,m}(J_k, A_k)$ 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 $[(\Sigma_k, j_k, \Theta_k)] \in \mathcal{M}_{g,m}$ converges to an element of the Deligne-Mumford space $[(S, j, \Theta, \Delta)] \in \overline{\mathcal{M}}_{g,m}$. Concretely, this means that for large k, our sequence in $\mathcal{M}_{g,m}(J_k, A_k)$ admits representatives $(\Sigma, j_k', \Theta, u_k')$, 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

$$j_k' \to j \quad \text{in} \quad C^\infty_{\text{loc}}(\Sigma \backslash C)$$

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 $u_k' : \Sigma \to W$ 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 $[(\Sigma_k, j_k, \Theta_k, u_k)] \in \mathcal{M}_{g,m}(J_k, A_k)$ 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_k' 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 9.3, producing more non-stable domain components—but again, the limiting map on these components will never be constant.

9.4.2. Holomorphic buildings. Only a small subset of the phenomena observed in §9.3 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.

Given integers $g, m, N_+, N_- \geqslant 0$, a **holomorphic building of height** $N_-|1|N_+$ with arithmetic genus g and m marked points is a tuple

$$\mathbf{u} = (S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\mathrm{nd}}, \Delta^{\mathrm{br}}, L, \Phi, u),$$

with the various data defined as follows:

- The **domain** $(S, j, \Gamma \cup \Theta, \Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}})$ is a connected but not necessarily stable nodal Riemann surface of arithmetic genus g, where $\Gamma = \Gamma^+ \cup \Gamma^-$, $\#\Theta = m$, and the involution on $\Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}}$ 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
$$L : S \to \{-N_-, \dots, -1, 0, 1, \dots, N_+\}$$
that attains every value in $\{-N_-, \dots, N_+\}$ except possibly 0, and satisfies:
  (1) $L(z^+) = L(z^-)$ for each node $\{z^+, z^-\} \subset \Delta^{\mathrm{nd}}$;
  (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, 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 $\mathbf{u}$ 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 $N = 0$. 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 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 $[\mathbf{u}] \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$.

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 $N \in \{-N_-, \dots, N_+\}$ 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 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

$$u_1^N \circ \varphi = u_0^N \text{ up to } \mathbb{R}\text{-translation} \quad \text{for each } N \neq 0.$$

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 $A \neq 0$ because they are not constant.

9.4.3. Convergence. For a general definition of the topology of the compactified moduli space $\overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ and the proof that it is both compact and metrizable, we refer to [**BEH**$^+$**03**] or the more comprehensive treatment in [**Abb14**]. We will refer to this topology as **the SFT topology**; in the literature it is sometimes also called the *Gromov-Hofer topology*. The following statement contains all the details about it that one usually needs to know in practice (see Figure 9.7).

THEOREM 9.28. *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 $J_k \to J$ in $\mathcal{J}_\tau(\omega_h, r_0, \mathcal{H}_+, \mathcal{H}_-)$. Then for any sequence*

$$[(\Sigma_k, j_k, \Gamma_k^+, \Gamma_k^-, \Theta_k, u_k)] \in \mathcal{M}_{g,m}(J_k, A_k, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$$

of nonconstant J_k-holomorphic curves in $\widehat{W}$ with uniformly bounded energy $E(u_k)$, there exists a stable holomorphic building

$$[\mathbf{u}_\infty] = [(S, j, \Gamma^+, \Gamma^-, \Theta, \Delta^{\mathrm{nd}}, \Delta^{\mathrm{br}}, L, \Phi, u_\infty)] \in \overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$$

such that after restricting to a subsequence, $[(\Sigma_k, j_k, \Gamma_k^+, \Gamma_k^-, \Theta_k, u_k)] \to [\mathbf{u}_\infty]$ in the following sense. The decorations Φ at Δ^{br} can be extended to decorations at Δ^{nd} so that if $\widehat{S}_\Phi$ denotes the closed oriented topological 2-manifold obtained from $S \backslash (\Delta^{\mathrm{nd}} \cup \Delta^{\mathrm{br}})$ by gluing circle compactifications along Φ, then for k sufficiently large, there exist homeomorphisms

$$\varphi_k : \widehat{S}_\Phi \to \Sigma_k$$

that are smooth outside of C_Φ*, map* $\Gamma^+ \cup \Gamma^- \cup \Theta$ *to* $\Gamma_k^+ \cup \Gamma_k^- \cup \Theta_k$ *with the ordering preserved, and satisfy*

$$\varphi_k^* j_k \to j \quad \text{in} \quad C^\infty_{\text{loc}}(\widehat{S}_\Phi \setminus C_\Phi).$$

Moreover for $N = \{-N_-, \dots, 0, \dots, N\}$*, let*

$$v_k^N := u_k \circ \varphi_k|_{\ddot{S}_N} : \ddot{S}_N \to \widehat{W},$$

with $\ddot{S}_N := \big(S \setminus (\Gamma \cup \Delta^{\text{nd}} \cup \Delta^{\text{br}})\big) \cap L^{-1}(N)$ *regarded as a subset of* $\widehat{S}_\Phi \setminus C_\Phi$*. Then:*

(1) $v_k^0 \to u_\infty^N$ *in* $C^\infty_{\text{loc}}(\ddot{S}_N, \widehat{W})$*;*
(2) For each $\pm N > 0$*,* v_k^N *has image in the positive/negative cylindrical end for all* k *sufficiently large, and there exists a sequence* $r_k^N \to \pm\infty$ *such that the resulting* $\mathbb{R}$*-translations converge:*

$$\tau_{-r_k^N} \circ v_k^N \to u_\infty^N \quad \text{in} \quad C^\infty_{\text{loc}}(\ddot{S}_N, \mathbb{R} \times M_\pm).$$

The rates of divergence of the sequences $r_k^N \to \pm\infty$ *are related by*

$$r_k^{N+1} - r_k^N \to +\infty \quad \text{for all } N < N_+.$$

Finally, let $\overline{S}_\Phi$ *denote the compact topological surface with boundary defined as the circle compactification of* $\widehat{S}_\Phi \setminus \Gamma$*, and let* $\bar{\Sigma}_k$ *denote the circle compactification of* $\dot{\Sigma}_k := \Sigma_k \setminus (\Gamma_k^+ \cup \Gamma_k^-)$*. Then for all* k *large,* φ_k *extends to a continuous map*

$$\bar{\varphi}_k : \overline{S}_\Phi \to \bar{\Sigma}_k$$

such that

$$\bar{u}_k \circ \bar{\varphi}_k \to \bar{u}_\infty \quad \text{in} \quad C^0(\overline{S}_\Phi, W).$$

REMARK 9.29. 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 9.30. Stability of the limit in Theorem 9.28 is guaranteed for the same reasons as in our discussion of Gromov compactness in §9.4.1: 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. Moreover, stability guarantees the *uniqueness* of the limiting building for any convergent sequence, i.e. it is the reason why $\overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ is a Hausdorff space. Indeed, if u_k converges to a stable building $\mathbf{u}_\infty$, then under the notion of convergence described in the theorem, it will also converge to a building $\mathbf{u}'_\infty$ constructed out of $\mathbf{u}_\infty$ by adding to S an extra spherical component, attaching it to the rest by a single node and extending the map u_∞ to be constant on the extra component. One can also insert extra levels into $\mathbf{u}_\infty$ that consist only of trivial cylinders, and u_k will still converge to the resulting building. But these modifications produce buildings that are not stable and thus are not elements of $\overline{\mathcal{M}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$.

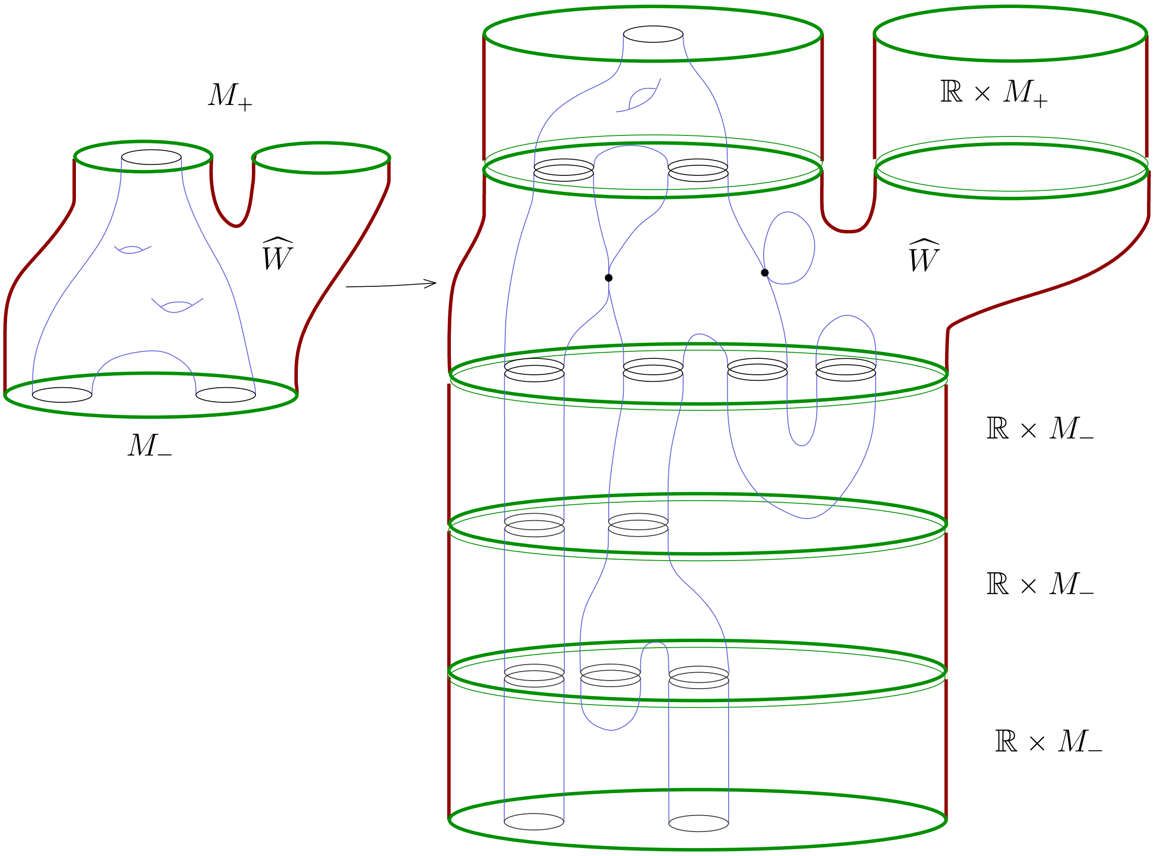

Figure 9.7. Convergence to a building with arithmetic genus 2, one upper level and three lower levels.

9.4.4. Symplectizations, stretching and so forth. 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 §9.4.2, 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 $L : S \to \{1, \ldots, N\}$ for some $N \in \mathbb{N}$, 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 Theorem 9.28.

To complete the picture, we should mention one more type of compactness theorem that appears in [BEH+03], which is colloquially described as *stretching the neck*. The geometric idea is as follows: suppose (W, Ω) is a closed symplectic manifold and $M \subset W$ is a stable hypersurface that separates W into two pieces

$W = W_- \cup_M W_+$, with an induced stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ that orients M as the boundary of W_-.[3] A neighborhood of M in (W, Ω) can then be identified symplectically with

$$(\mathcal{N}_\epsilon, \omega_\epsilon) := ((-\epsilon, \epsilon) \times M, d(r\lambda) + \omega)$$

for sufficiently small $\epsilon > 0$. The idea now is to replace $\mathcal{N}_\epsilon$ with larger collars of the form

$$(\mathcal{N}_T, \omega_T) := ((-T, T) \times M, d\,(f(r)\lambda) + \omega)\,,$$

with C^0-small functions f chosen with $f' > 0$ so that the collar can be glued in smoothly to replace $(\mathcal{N}_\epsilon, \omega_\epsilon)$. These functions are equivalent to choices of diffeomorphisms $(-T, T) \to (-\epsilon, \epsilon)$, which then give rise to symplectomorphisms

$$(\mathcal{N}_T, \omega_T) \xrightarrow{\cong} (\mathcal{N}_\epsilon, \omega_\epsilon) \subset (W, \Omega), \tag{9.18}$$

and we consider pseudoholomorphic curves in the family of closed symplectic manifolds (W_T, Ω_T) obtained by deleting $(\mathcal{N}_\epsilon, \omega_\epsilon)$ from (W, Ω) and replacing it by $(\mathcal{N}_T, \omega_T)$. While (9.18) implies that these manifolds are all symplectomorphic, they admit natural families of tame almost complex structures that exhibit distinctive behavior as $T \to \infty$. Indeed, if we fix $J \in \mathcal{J}_\tau(W, \Omega)$ such that its restriction to $(\mathcal{N}_\epsilon, \omega_\epsilon)$ belongs to $\mathcal{J}(\mathcal{H})$, then for every $T \geqslant \epsilon$, the same recipe gives rise to a tame almost complex structure $J_T \in \mathcal{J}_\tau(W_T, \Omega_T)$ that matches J on the complement of $\mathcal{N}_T$ and also belongs to $\mathcal{J}(\mathcal{H})$ on $\mathcal{N}_T$. This family degenerates as $T \to \infty$, i.e. pushing each J_T forward through the symplectomorphism $(W_T, \Omega_T) \cong (W, \Omega)$ produces a family in $\mathcal{J}_\tau(W, \Omega)$ that has no well-defined limit as $T \to \infty$. Now given a sequence $T_k \to \infty$ and a corresponding degenerating sequence $J_k \in \mathcal{J}_\tau(W_{T_k}, \Omega_{T_k})$ as described above, a sequence u_k of J_k-holomorphic curves with bounded energy has a subsequence convergent to yet another form of holomorphic building, this time involving a bottom level in $\widehat{W}_- := W_- \cup_M ([0, \infty) \times M)$ with positive punctures approaching orbits in M, some finite number of middle levels that live in the symplectization of M, and a top level that lives in $\widehat{W}_+ := ((-\infty, 0] \times M) \cup_M W_+$ with negative punctures approaching M.

A very popular example for applications 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].

[3] The assumption that $M \subset W$ separates W is inessential, but makes certain details in this discussion more convenient.

LECTURE 10

Cylindrical contact homology and the tight 3-tori

Contents

We've now developed enough of the technical machinery of holomorphic curves to be able to give a rigorous construction of the most basic version of SFT and apply it to a problem in contact topology.

10.1. Contact structures on $\mathbb{T}^3$ and Giroux torsion

As a motivating goal in this lecture, we will prove a result about the classification of contact structures on $\mathbb{T}^3 = S^1 \times S^1 \times S^1$. Denote the three global coordinates on $\mathbb{T}^3$ valued in $S^1 = \mathbb{R}/\mathbb{Z}$ by (ρ, ϕ, θ), and for any $k \in \mathbb{N}$, consider the contact structure

$$\xi_k := \ker \alpha_k, \qquad \text{where} \qquad \alpha_k := \cos(2\pi k \rho)\, d\theta + \sin(2\pi k\rho)\, d\phi.$$

It is an easy exercise to verify that these all satisfy the contact condition $\alpha_k \wedge d\alpha_k > 0$; see Figure 10.1 for a visual representation. The following result is originally due to Giroux [**Gir94**] and Kanda [**Kan97**].

THEOREM 10.1. *For each pair of positive integers $k \neq \ell$, the contact manifolds $(\mathbb{T}^3, \xi_k)$ and $(\mathbb{T}^3, \xi_\ell)$ are not contactomorphic.*

One of the reasons this result is interesting is that it cannot be proved using any so-called "classical" invariants, i.e. invariants coming from algebraic topology. An example of a classical invariant would be the Euler class of the oriented vector bundle $\xi_k \to \mathbb{T}^3$, or anything else that depends only on the isomorphism class of this

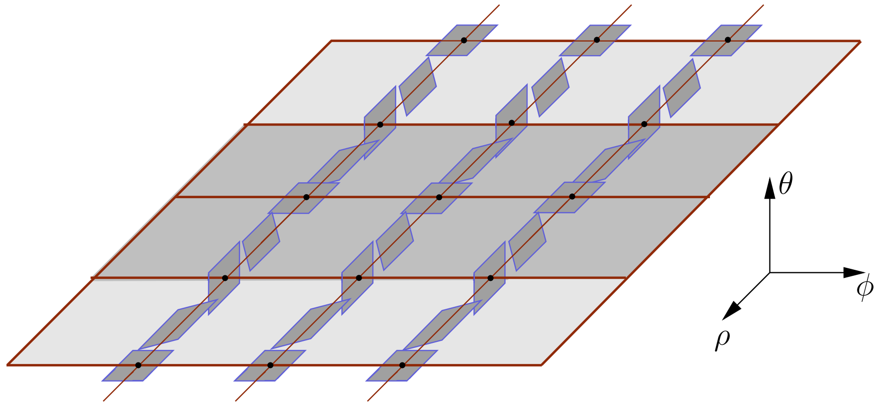

FIGURE 10.1. The contact structures ξ_k on $\mathbb{T}^3$ can be constructed by gluing k copies of the same model $[0,1] \times \mathbb{T}^2$ to each other cyclically.

bundle. The following observation shows that such invariants will never distinguish ξ_k from ξ_ℓ.

PROPOSITION 10.2. *For every $k, \ell \in \mathbb{N}$, ξ_k and ξ_ℓ are homotopic through a smooth family of oriented 2-plane fields on $\mathbb{T}^3$.*

PROOF. In fact, all the ξ_k can be deformed smoothly to $\ker d\rho$, via the homotopy

$$\ker\left[(1-s)\,\alpha_k + s\,d\rho\right], \qquad s \in [0,1].$$

$\square$

REMARK 10.3. One can check in fact that the 1-form in the homotopy given above is contact for every $s \in [0,1)$, so Gray's stability theorem implies that every ξ_k is isotopic to an arbitrarily small perturbation of the foliation $\ker d\rho$. In [**Gir94**], Giroux used this observation to show that all of them are what we now call *weakly symplectically fillable.* If $\ker d\rho$ were also contact, then Gray's theorem would imply that ξ_k and ξ_ℓ are always isotopic. Thus Theorem 10.1 indicates the impossibility of modifying a homotopy from ξ_k to ξ_ℓ into one that passes only through contact structures.

Let us place this discussion in a larger context. Using the coordinates (ρ, ϕ, θ) on $\mathbb{R} \times \mathbb{T}^2$, a pair of smooth functions $f, g : \mathbb{R} \to \mathbb{R}$ gives rise to a contact form

$$\alpha = f(\rho)\,d\theta + g(\rho)\,d\phi$$

whenever the function $D(\rho) := f(\rho)g'(\rho) - f'(\rho)g(\rho)$ is everywhere positive. Indeed, we have $\alpha \wedge d\alpha = D(\rho)\,d\rho \wedge d\phi \wedge d\theta$, and one easily derives a similar formula for the Reeb vector field,

$$R_\alpha = \frac{1}{D(\rho)}\left[g'(\rho)\,\partial_\theta - f'(\rho)\,\partial_\phi\right].$$

The condition $D > 0$ means geometrically that the path $(f, g) : \mathbb{R} \to \mathbb{R}^2$ winds counterclockwise around the origin with its angular coordinate strictly increasing. The simplest special case is the contact form

$$\alpha_{\mathrm{GT}} := \cos(2\pi\rho)\,d\theta + \sin(2\pi\rho)\,d\phi,$$

which matches the formula for α_1 on $\mathbb{T}^3$ given above. Let $\xi_{\mathrm{GT}} := \ker \alpha_{\mathrm{GT}}$ on $\mathbb{R} \times \mathbb{T}^2$.

DEFINITION 10.4. The **Giroux torsion** $\mathrm{GT}(M, \xi) \in \mathbb{N} \cup \{0, \infty\}$ of a contact 3-manifold (M, ξ) is the supremum of the set of positive integers k such that there exists a contact embedding

$$\left([0, k] \times \mathbb{T}^2, \xi_{\mathrm{GT}}\right) \hookrightarrow (M, \xi).$$

We write $\mathrm{GT}(M, \xi) = 0$ if no such embedding exists for any k, and $\mathrm{GT}(M, \xi) = \infty$ if it exists for all k.

EXAMPLE 10.5. The tori $(\mathbb{T}^3, \xi_k)$ for $k \in \mathbb{N}$ are contactomorphic to $(\mathbb{R}\times\mathbb{T}^2, \xi_{\mathrm{GT}})/k\mathbb{Z}$, with $k\mathbb{Z}$ acting by translation of the ρ-coordinate. Thus $\mathrm{GT}(\mathbb{T}^3, \xi_k) \geqslant k - 1$.

A 2-torus $T \subset (M, \xi)$ embedded in a contact 3-manifold is called **pre-Lagrangian** if a neighborhood of T in (M, ξ) admits a contactomorphism to a neighborhood of $\{0\} \times \mathbb{T}^2$ in $(\mathbb{R} \times \mathbb{T}^2, \xi_{\mathrm{GT}})$, identifying T with $\{0\} \times \mathbb{T}^2$. The neighborhood in $\mathbb{R} \times \mathbb{T}^2$ can be arbitrarily small, thus the existence of a pre-Lagrangian torus does not imply $\mathrm{GT}(M, \xi) > 0$; in fact, pre-Lagrangian tori always exist in abundance, e.g. as boundaries of neighborhoods of transverse knots (using the contact model provided by the transverse neighborhood theorem). But given any pre-Lagrangian torus $T \subset (M, \xi)$, one can make a local modification of ξ near T to produce a new contact structure (up to isotopy) with positive Giroux torsion. Define (M', ξ') from (M, ξ) by replacing the small neighborhood $((-\epsilon, \epsilon) \times \mathbb{T}^2, \xi_{\mathrm{GT}})$ with $((-\epsilon, 1 + \epsilon) \times \mathbb{T}^2, \xi_{\mathrm{GT}})$, then identify M' with M by a choice of compactly supported diffeomorphism $(-\epsilon, 1+\epsilon) \to (-\epsilon, \epsilon)$. There is now an obvious contact embedding of $([0, 1] \times \mathbb{T}^2, \xi_{\mathrm{GT}})$ into (M, ξ'), hence $\mathrm{GT}(M, \xi') \geqslant 1$. Moreover, one can adapt the proof of Prop. 10.2 above to show that ξ' is homotopic to ξ through a smooth family of oriented 2-plane fields. The operation changing ξ to ξ' is known as a **Lutz twist** along T. In this language, we see that for each $k \in \mathbb{N}$, $(\mathbb{T}^3, \xi_{k+1})$ is obtained from $(\mathbb{T}^3, \xi_k)$ by performing a Lutz twist along $\{0\} \times \mathbb{T}^2$.

The invariant $\mathrm{GT}(M, \xi)$ is easy to define, but hard to compute in general. The natural guess,

$$\mathrm{GT}(\mathbb{T}^3, \xi_k) = k - 1,$$

turns out to be correct, as was shown in [Gir00], so this is one way to prove Theorem 10.1, but not the approach we will take. The following example shows that one must in any case be careful with such guesses.

EXAMPLE 10.6. For each $k \in \mathbb{N}$, define a model of $S^1 \times S^2$ by

$$S^1 \times S^2 \cong \left([0, k + 1/2] \times \mathbb{T}^2\right) / \sim$$

where the equivalence relation identifies $(\rho, \phi, \theta) \sim (\rho, \phi', \theta)$ for $\rho \in \{0, k + 1/2\}$ and every $\theta, \phi, \phi' \in S^1$. Near $\rho = 0$ and $\rho = k + 1/2$, this means thinking of (ρ, ϕ) as polar coordinates, so the two subsets $\{\rho = 0\}$ and $\{\rho = k+1/2\}$ become circles of the form $S^1 \times \{\text{const}\}$ embedded in $S^1 \times S^2$. Since the ϕ-coordinate is singular at these two circles, the contact form α_{GT} needs to be modified slightly in this region before it will descend to a smooth contact form on $S^1 \times S^2$: this can be done by a C^0-small modification of the form $f(\rho)\,d\theta + g(\rho)\,d\phi$, and the resulting contact structure is then uniquely determined up to isotopy. We shall call this contact manifold

$$(S^1 \times S^2, \xi_k).$$

Now observe that for each $k \in \mathbb{N}$, $(S^1 \times S^2, \xi_{k+1})$ is obtained from $(S^1 \times S^2, \xi_k)$ by a Lutz twist. However, both contact manifolds are also **overtwisted**: a contact 3-manifold (M, ξ) is overtwisted whenever it contains an embedded closed 2-disk $\mathcal{D} \subset M$ such that $T(\partial\mathcal{D}) \subset \xi$ but $T\mathcal{D}|_{\partial\mathcal{D}} \pitchfork \xi$. (Exercise: find a disk with this property in $(S^1 \times S^2, \xi_k)$!) Eliashberg's flexibility theorem for overtwisted contact structures [**Eli89**] implies that whenever ξ and ξ' are two contact structures on a closed 3-manifold that are both overtwisted and are homotopic as oriented 2-plane fields, they are actually isotopic. As a consequence, the contact structures ξ_k on $S^1 \times S^2$ defined above for every $k \in \mathbb{N}$ are all isotopic to each other. As tends to be the case with most interesting h-principles, the isotopy is very hard to see concretely, but it must exist.

EXERCISE 10.7. Show that if (M, ξ) is a closed overtwisted contact 3-manifold, then $\mathrm{GT}(M, \xi) = \infty$.

In contrast to the $S^1 \times S^2$ example above, the contact manifolds $(\mathbb{T}^3, \xi_k)$ are not overtwisted, they are **tight**—in fact, the classification of contact structures on $\mathbb{T}^3$ by Giroux [**Gir94**, **Gir99**, **Gir00**] and Kanda [**Kan97**] states that these are *all* of the tight contact structures on $\mathbb{T}^3$ up to contactomorphism. We will use cylindrical contact homology to show that they are not contactomorphic to each other. The reader should keep Example 10.6 in mind and try to spot the reason why the same argument cannot work for $(S^1 \times S^2, \xi_k)$.

REMARK 10.8. It has been conjectured that the converse of Exercise 10.7 might also hold, so every closed tight contact 3-manifold would have finite Giroux torsion. This conjecture is wide open.

10.2. Definition of cylindrical contact homology

10.2.1. Preliminary remarks. Cylindrical contact homology is the natural "first attempt" at using holomorphic curves in symplectizations to define a Floer-type invariant of contact manifolds (M, ξ). The idea is to define a chain complex generated by Reeb orbits in M and a differential ∂ that counts holomorphic cylinders in $\mathbb{R} \times M$. We saw already in §1.5 some pretty good reasons why this idea cannot work in general: in order to prove $\partial^2 = 0$, we need to be able to identify the space of rigid "broken" holomorphic cylinders (these are what is counted by ∂^2) with the boundary of the compactified 1-dimensional space of index 2 cylinders (up to $\mathbb{R}$-translation). But this compactified boundary has more than just broken cylinders in it, see Figure 10.2. In order to define cylindrical contact homology, one must therefore restrict to situations in which complicated pictures like Figure 10.2 cannot occur. The first useful remark in this direction is that since we are working with a stable Hamiltonian structure of the form $(d\alpha, \alpha)$ for a contact form α, a certain subset of the scenarios allowed by the SFT compactness theorem can be excluded immediately. Indeed:

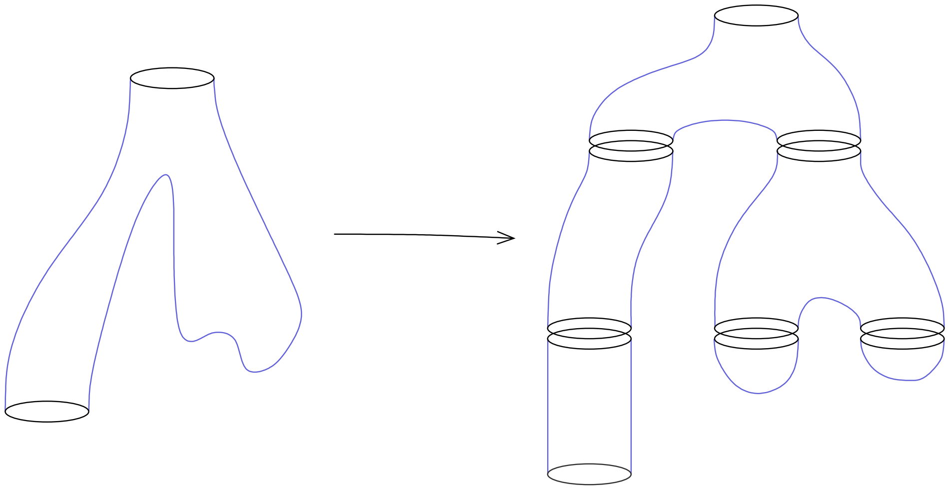

Figure 10.2. A family of holomorphic cylinders can converge in the SFT topology to buildings that include more complicated curves than cylinders—this is why cylindrical contact homology is not well defined for all contact manifolds.

Proposition 10.9. *If $J \in \mathcal{J}(\alpha)$ and $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is a nonconstant asymptotically cylindrical J-holomorphic curve, then u has at least one positive puncture.*

Let us give two proofs of this result, since both contain useful ideas. As preparation for the first proof, recall the definition of energy for curves in symplectizations of contact manifolds that we wrote down in Lecture 1:

$$E(u) := \sup_{f \in \mathcal{T}} \int_{\dot{\Sigma}} u^* d(e^{f(r)} \alpha),$$

where

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

This formula is not identical to the definition of energy used in Lecture 9, but it is equivalent in the sense that any uniform bounds on one imply similar uniform bounds on the other.

First proof of Proposition 10.9. Denote the positive and negative punctures of $u : \dot{\Sigma} \to \mathbb{R} \times M$ by Γ^+ and Γ^- respectively, and suppose u is asymptotic at $z \in \Gamma^\pm$ to the orbit γ_z with period $T_z > 0$. Choose any $f \in \mathcal{T}$ and denote $f_\pm := \lim_{r \to \pm\infty} f(r) \in [-1, 1]$. Since $d(e^{f(r)} \alpha)$ tames $J \in \mathcal{J}(\alpha)$ and u is not constant, Stokes' theorem gives

$$0 < E(u) = e^{f_+} \sum_{z \in \Gamma^+} T_z - e^{f_-} \sum_{z \in \Gamma^-} T_z, \tag{10.1}$$

hence Γ^+ cannot be empty. $\square$

Remark 10.10. The proof via Stokes' theorem works just as well if instead of $\mathbb{R} \times M$, u lives in the completion of an exact symplectic cobordism (W, ω) with concave boundary $(M_-, \xi_- = \ker \alpha_-)$ and convex boundary $(M_+, \xi_+ = \ker \alpha_+)$.

Recall that this means $\partial W = -M_- \coprod M_+$, and $\omega = d\lambda$ for a globally-defined 1-form $\lambda \in \Omega^1(W)$ that restricts to positive contact forms $\lambda|_{TM_\pm} = \alpha_\pm$. As in Lecture 1, we will write

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

for the space of almost complex structures J on $\widehat{W} := ((-\infty, 0] \times M_-) \cup_{M_-} W \cup_{M_+} ([0, \infty) \times M_+)$ that are compatible with ω on W and belong to $\mathcal{J}(\alpha_\pm)$ on the cylindrical ends. The energy of a J-holomorphic curve $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ is then

$$E(u) := \sup_{f \in \mathcal{T}} \int_{\dot{\Sigma}} u^* d\lambda_f,$$

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

$$\lambda_f := \begin{cases} e^{f(r)}\alpha_+ & \text{on } [0, \infty) \times M_+, \\ \lambda & \text{on } W, \\ e^{f(r)}\alpha_- & \text{on } (-\infty, 0] \times M_-. \end{cases}$$

The above proof now generalizes verbatim to show that u must always have a positive puncture. Notice that in both settings, the argument also gives a uniform bound for the energy in terms of the periods of the positive asymptotic orbits.

REMARK 10.11. We can also prove Prop. 10.9 using the fact that $u^* d\alpha \geqslant 0$ for any $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ with $J \in \mathcal{J}(\alpha)$. Indeed, Stokes' theorem then gives

$$0 \leqslant \int_{\dot{\Sigma}} u^* d\alpha = \sum_{z \in \Gamma^+} T_z - \sum_{z \in \Gamma^-} T_z. \tag{10.2}$$

The quantity $\int_{\dot{\Sigma}} u^* d\alpha$ is sometimes called the **contact area** of u. This version of the argument however does not easily generalize to arbitrary exact cobordisms.

The second proof is based on the maximum principle for subharmonic functions.

PROPOSITION 10.12. *Suppose $J \in \mathcal{J}(\alpha)$ and $u = (u_{\mathbb{R}}, u_M) : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is J-holomorphic, where $\dot{\Sigma}$ has no boundary. Then $u_{\mathbb{R}} : \dot{\Sigma} \to \mathbb{R}$ has no local maxima.*

PROOF. In any local holomorphic coordinates (s, t) on a region in $\dot{\Sigma}$, the nonlinear Cauchy-Riemann equation for u is equivalent to the system of equations

$$\begin{aligned} \partial_s u_{\mathbb{R}} - \alpha(\partial_t u_M) &= 0, \\ \partial_t u_{\mathbb{R}} + \alpha(\partial_s u_M) &= 0, \\ \pi_\xi\, \partial_s u_M + J \pi_\xi\, \partial_t u_M &= 0, \end{aligned}$$

where $\pi_\xi : TM \to \xi$ denotes the projection along the Reeb vector field. This gives

$$\begin{aligned} -\Delta u_{\mathbb{R}} &= -\partial_s^2 u_{\mathbb{R}} - \partial_t^2 u_{\mathbb{R}} = -\partial_s \left[\alpha(\partial_t u_M)\right] + \partial_t \left[\alpha(\partial_s u_M)\right] \\ &= -d\alpha(\partial_s u_M, \partial_t u_M) = -d\alpha(\pi_\xi \partial_s u_M, J\pi_\xi \partial_s u_M) \leqslant 0 \end{aligned}$$

since $J|_\xi$ is tamed by $d\alpha|_\xi$, hence $u_{\mathbb{R}}$ is subharmonic. The result thus follows from the maximum principle, see e.g. [Eva98]. □

SECOND PROOF OF PROPOSITION 10.9. If $u = (u_{\mathbb{R}}, u_M) : \dot{\Sigma} \to \mathbb{R} \times M$ has no positive puncture then $u_{\mathbb{R}} : \dot{\Sigma} \to \mathbb{R}$ is a proper function bounded above, and therefore has a local maximum, contradicting Proposition 10.12. □

REMARK 10.13. The proof via the maximum principle does not generalize to arbitrary exact cobordisms $(W, d\lambda)$, but it does work in *Stein* cobordisms, i.e. if λ_f and J are related by $\lambda_f = -dF \circ J$ for some plurisubharmonic function $F : \widehat{W} \to \mathbb{R}$, then $F \circ u : \dot{\Sigma} \to \mathbb{R}$ is subharmonic (cf. [CE12]).

With these preliminaries understood, the next two exercises reveal one natural setting in which breaking of cylinders can be kept under control. Both exercises are essentially combinatorial.

EXERCISE 10.14. Suppose $\mathbf{u}$ is a stable J-holomorphic building in a completed symplectic cobordism $\widehat{W}$ with the following properties:

(1) $\mathbf{u}$ has arithmetic genus 0 and exactly one positive puncture;
(2) Every connected component of $\mathbf{u}$ has at least one positive puncture.

Show that $\mathbf{u}$ has no nodes, and all of its connected components have *exactly* one positive puncture.

EXERCISE 10.15. Suppose that in addition to the conditions of Exercise 10.14, $\mathbf{u}$ has exactly one negative puncture and no connected component of $\mathbf{u}$ is a plane. Show that every level of $\mathbf{u}$ then consists of a single cylinder with one positive and one negative end.

Exercise 10.15 makes it reasonable to define a Floer-type theory counting only cylinders in any setting where planes can be excluded, for instance because the Reeb vector field has no contractible orbits. This is not always possible, e.g. Hofer [Hof93] proved that on overtwisted contact manifolds, there is *always* a plane (which is why the Weinstein conjecture holds). So the invariant we construct will not be defined in such settings, but it happens to be ideally suited to the study of $(\mathbb{T}^3, \xi_k)$.

10.2.2. A compactness result for cylinders. Fix a closed contact manifold (M, ξ) of dimension $2n-1$ and a primitive homotopy class of loops $h \in [S^1, M]$. By **primitive**, we mean that h is not equal to Nh' for any $h' \in [S^1, M]$ and an integer $N > 1$, and this assumption will be crucial for technical reasons in the following.[1] Given a contact form α for ξ, let

$$\mathcal{P}_h(\alpha)$$

denote the set of closed Reeb orbits homotopic to h, where two Reeb orbits are identified if they differ only by parametrization.

DEFINITION 10.16. Given a contact manifold (M, ξ) and a homotopy class $h \in [S^1, M]$, we will say that a contact form α for ξ is **h-admissible** if:

[1]It is to be expected that cylindrical contact homology can be defined also for non-primitive homotopy classes, but this would require more sophisticated methods to address transversality problems. The assumption that h is primitive allows us to assume that all holomorphic curves in the discussion are somewhere injective, hence they are always regular if J is generic.

(1) All orbits in $\mathcal{P}_h(\alpha)$ are nondegenerate;
(2) There are no contractible closed Reeb orbits.

Similarly, we will say that (M, ξ) is **h-admissible** if a contact form with the above properties exists.

DEFINITION 10.17. Given $h \in [S^1, M]$ and an h-admissible contact form α on (M, ξ), we will say that an almost complex structure $J \in \mathcal{J}(\alpha)$ is **h-regular** if every J-holomorphic cylinder in $\mathbb{R} \times M$ with a positive and a negative end both asymptotic to orbits in $\mathcal{P}_h(\alpha)$ is Fredholm regular.

PROPOSITION 10.18. *If $h \in [S^1, M]$ is a primitive homotopy class of loops and α is h-admissible on (M, ξ), then the space of h-regular almost complex structures is comeager in $\mathcal{J}(\alpha)$.*

PROOF. Since h is primitive, the asymptotic orbits for the relevant holomorphic cylinders cannot be multiply covered, hence all of these cylinders are somewhere injective. The result therefore follows from the standard transversality results proved in Lecture 8 for somewhere injective curves in symplectizations. □

PROPOSITION 10.19. *Given an h-admissible contact form α, an h-regular almost complex structure $J \in \mathcal{J}(\alpha)$ and an orbit $\gamma \in \mathcal{P}_h(\alpha)$, suppose u_k is a sequence of J-holomorphic cylinders in $\mathbb{R} \times M$ with one positive puncture at γ and one negative puncture. Then u_k has a subsequence convergent in the SFT topology to a broken J-holomorphic cylinder, i.e. a stable building $\mathbf{u}_\infty$ whose levels $u_\infty^1, \ldots, u_\infty^{N_+}$ are each cylinders with one positive and one negative puncture. Moreover, each level satisfies $\operatorname{ind}(u_\infty^N) \geqslant 1$, thus for large k in the convergent subsequence,*

$$\operatorname{ind}(u_k) = \sum_{N=1}^{N_+} \operatorname{ind}(u_\infty^N) \geqslant N_+.$$

PROOF. Let's start with some bad news: the standard SFT compactness theorem is not applicable in this situation, because we have not assumed that α is nondegenerate, nor even Morse Bott—there is no assumption at all about Reeb orbits in homotopy classes other than h and 0. This fairly loose set of hypotheses is very convenient in applications, as nondegeneracy of a contact form is generally a quite difficult condition to check. The price we pay is that we will have to prove compactness manually instead of applying the big theorem (see Remark 10.20). Fortunately, it is not that hard: the crucial point is that in the situation at hand, there can be no bubbling at all.

Indeed, we claim that the given sequence $u_k : (\mathbb{R} \times S^1, i) \to (\mathbb{R} \times M, J)$ must satisfy a uniform bound

$$|du_k| \leqslant C$$

with respect to any translation-invariant Riemannian metrics on $\mathbb{R} \times S^1$ and $\mathbb{R} \times M$. To see this, note first that since all the u_k have the same positive asymptotic orbit γ, their energies are uniformly bounded via (10.1). Thus if $|du_k(z_k)| \to \infty$ for some sequence $z_k \in \mathbb{R} \times S^1$, we can perform the usual rescaling trick from Lecture 9 and deduce the existence of a nonconstant finite-energy plane $v_\infty : \mathbb{C} \to \mathbb{R} \times M$. Its singularity at ∞ cannot be removable since this would produce a nonconstant J-holomorphic sphere, violating Proposition 10.9. It follows that v_∞ is asymptotic to a Reeb orbit at ∞, but this is also impossible since α does not admit any contractible orbits, and the claim is thus proved.

Suppose now that γ has period $T_+ > 0$, and observe that by nondegeneracy and the Arzelà-Ascoli theorem, the set

$$\mathcal{P}_h(\alpha, T_+) := \left\{\gamma \in \mathcal{P}_h(\alpha) \;\middle|\; \gamma \text{ has period at most } T_+\right\}$$

is finite. Let

$$\mathcal{A}_h(\alpha), \mathcal{A}_h(\alpha, T_+) \subset (0, \infty)$$

denote the set of all periods of orbits in $\mathcal{P}_h(\alpha)$ and $\mathcal{P}_h(\alpha, T_+)$ respectively. By (10.2), the negative asymptotic orbit of each u_k is in $\mathcal{P}_h(\alpha, T_+)$, so we can take a subsequence and assume that these are all the same orbit; call it $\gamma_- \in \mathcal{P}_h(\alpha, T_+)$ and its period $T_- \in \mathcal{A}_h(\alpha, T_+)$. If $T_- = T_+$ then (10.2) implies $u_k^* d\alpha \equiv 0$ for all k, and it follows that u_k is the trivial cylinder over γ, so the sequence trivially converges. Assume therefore $T_- < T_+$. Then since $u_k^* d\alpha \geqslant 0$, Stokes' theorem implies that for each k, the function

$$\mathbb{R} \to \mathbb{R} : s \mapsto \int_{S^1} u_k(s, \cdot)^* \alpha$$

is increasing and is a surjective map onto (T_-, T_+). The uniform bound on the derivatives implies that for any sequences $s_k, r_k \in \mathbb{R}$ with $u_k(s_k, 0) \in \{r_k\} \times M$, the sequence[2]

$$v_k : \mathbb{R} \times S^1 \to \mathbb{R} \times M : (s, t) \mapsto \tau_{-r_k} \circ u_k(s + s_k, t)$$

has a subsequence convergent in $C^\infty_{\mathrm{loc}}(\mathbb{R} \times S^1)$ to some finite-energy J-holomorphic cylinder

$$v_\infty : \mathbb{R} \times S^1 \to \mathbb{R} \times M,$$

which necessarily satisfies

$$\int_{S^1} v_\infty(s, \cdot)^* \alpha = \lim_{k \to \infty} \int_{S^1} u_k(s + s_k, \cdot)^* \alpha \in [T_-, T_+]$$

for every $s \in \mathbb{R}$. This proves that v_∞ is nonconstant, with a positive puncture at $s = \infty$ and negative puncture at $s = -\infty$, and both of its asymptotic orbits are in $\mathcal{P}_h(\alpha, T_+)$.[3] If v_∞ is not a trivial cylinder, then it therefore satisfies

$$\int_{\mathbb{R} \times S^1} v_\infty^* d\alpha \geqslant \delta,$$

where δ is any positive number less than the smallest distance between neighboring elements of $\mathcal{A}_h(\alpha, T_+)$.

Let us call a sequence $s_k \in \mathbb{R}$ *nontrivial* whenever the limiting cylinder v_∞ obtained by the above procedure is not a trivial cylinder, and call two such sequences

---

[2]Recall from Lecture 9 that we denote the $\mathbb{R}$-translation action on $\mathbb{R} \times M$ by $\tau_c(r, x) := (r + c, x)$.

[3]For an alternative argument that v_∞ must have a positive puncture at $s = \infty$ and negative at $s = -\infty$, see Figure 10.3.

s_k and s'_k *compatible* if $s_k - s'_k$ is not bounded. We claim now that if $s_k^1, \dots, s_k^m$ is a collection of nontrivial sequences that are all compatible with each other, then

$$m < \frac{2(T_+ - T_-)}{\delta}.$$

Indeed, we can assume after ordering our collection appropriately and restricting to a subsequence that $s_k^{N+1} - s_k^N \to \infty$ for each $N = 1, \dots, m-1$, and let $v_\infty^N : \mathbb{R} \times S^1 \to \mathbb{R} \times M$ denote the limits of the corresponding convergent subsequences. Then we can find $R > 0$ such that

$$\int_{[-R,R]\times S^1} (v_\infty^N)^* d\alpha > \frac{\delta}{2}$$

and thus

$$\int_{[s_k^N - R, s_k^N + R]\times S^1} u_k^* d\alpha > \frac{\delta}{2}$$

for each $N = 1, \dots, m$ for sufficiently large k. But these domains are also all disjoint for sufficiently large k, implying

$$T_+ - T_- = \int_{\mathbb{R}\times S^1} u_k^* d\alpha \geqslant \sum_{N=1}^m \int_{[s_k^N - R, s_k^N + R]\times S^1} u_k^* d\alpha > \frac{\delta m}{2}.$$

We've shown that there exists a maximal collection of nontrivial sequences $s_k^1, \dots, s_k^{N_+} \in \mathbb{R}$ satisfying $s_k^{N+1} - s_k^N \to \infty$ for each N, such that if $u_k(s_k^N, 0) \in \{r_k^N\} \times M$, then after restricting to a subsequence, the cylinders

$$v_k^N(s,t) := \tau_{-r_k^N} \circ u_k(s + s_k^N, t)$$

each converge in $C^\infty_{\text{loc}}(\mathbb{R} \times S^1)$ as $k \to \infty$ to a nontrivial J-holomorphic cylinder $u_\infty^N : \mathbb{R} \times S^1 \to \mathbb{R} \times M$. Let $\gamma_N^\pm$ denote the asymptotic orbit of u_∞^N at $s = \pm\infty$. We claim,

$$\gamma_N^+ = \gamma_{N+1}^- \quad \text{for each } N = 1, \dots, N_+ - 1.$$

If $\gamma_N^+ \neq \gamma_{N+1}^-$ for some N, choose a neighborhood $\mathcal{U} \subset M$ of the image of γ_N^+ that does not intersect any other orbit in $\mathcal{P}_h(\alpha, T_+)$. Then since each u_k is continuous, there must exist a sequence $s'_k \in \mathbb{R}$ with

$$s'_k - s_k^N \to \infty \quad \text{and} \quad s_k^{N+1} - s'_k \to \infty$$

such that $u_k(s'_k, 0)$ lies in $\mathcal{U}$ for all k but stays a positive distance away from the image of γ_N^+. A subsequence of $(s,t) \mapsto u_k(s + s'_k, t)$ then converges after suitable $\mathbb{R}$-translations to a cylinder $u'_\infty : \mathbb{R} \times S^1 \to \mathbb{R} \times M$ that cannot be trivial since $u'_\infty(0,0)$ is not contained in any orbit in $\mathcal{P}_h(\alpha, T_+)$. This contradicts the assumption that our collection $s_k^1, \dots, s_k^{N_+}$ is maximal. A similar argument shows

$$\gamma_1^- = \gamma^- \quad \text{and} \quad \gamma_{N_+}^+ = \gamma,$$

so the curves $u_\infty^1, \dots, u_\infty^{N_+}$ form the levels of a stable holomorphic building $\mathbf{u}_\infty$. A similar argument by contradiction also shows that the sequence u_k must converge in the SFT topology to $\mathbf{u}_\infty$.

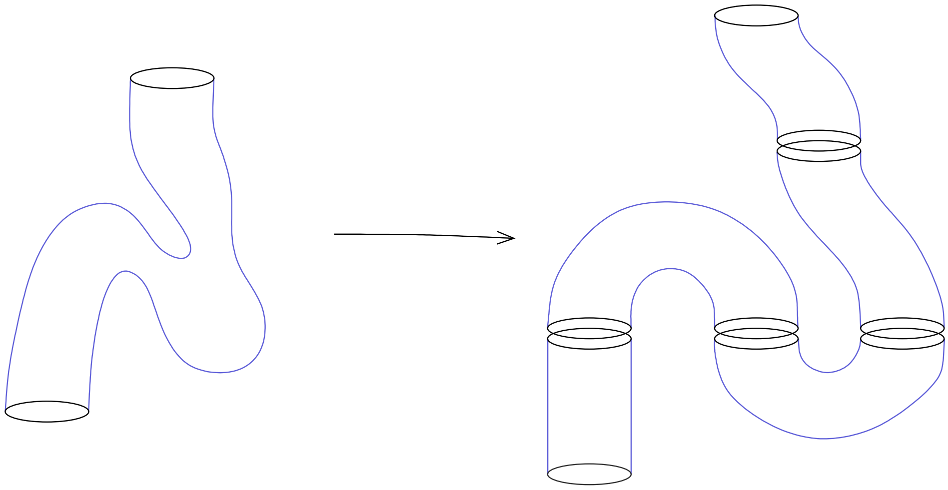

Figure 10.3. A degenerating sequence of holomorphic cylinders $u_k : \mathbb{R} \times S^1 \to \mathbb{R} \times M$ cannot have a limiting level with a puncture of the "wrong" sign unless u_k violates the maximum principle for large k.

Finally, note that since all the breaking orbits in $\mathbf{u}_\infty$ are homotopic to h and J is h-regular, the levels u_∞^N are Fredholm regular. Since all of them also come in 1-parameter families of distinct curves related by the $\mathbb{R}$-action, this implies $\operatorname{ind}(u_\infty^N) \geqslant 1$ for each $N = 1, \dots, N_+$. □

Remark 10.20. Nondegeneracy or Morse-Bott conditions are required for several reasons in the proof of SFT compactness, and indeed, the theorem is not true in general without some such assumption. One can see this by considering what happens to a sequence u_k of J_k-holomorphic curves where $J_k \to J_\infty$ is compatible with a sequence of nondegenerate contact forms α_k converging to one that is only Morse-Bott. A compactness theorem for this scenario is proved in [**Bou02**], but it requires more general limiting objects than holomorphic buildings. On the other hand, it is useful for certain kinds of applications to know when one can do without nondegeneracy assumptions and prove compactness anyway. There are two main advantages to knowing that all Reeb orbits are nondegenerate or belong to Morse-Bott families:

(1) It implies that the set of all periods of closed orbits, the so-called **action spectrum** of α, is a *discrete* subset of $(0, \infty)$; in fact, for any $T > 0$, the set of all periods less than T is finite. Using the relations (10.1) and (10.2), this implies lower bounds on the possible energies of limiting components and thus helps show that only finitely many such components can arise.
(2) Curves asymptotic to nondegenerate or Morse-Bott orbits also satisfy exponential convergence estimates such as in §6.5 and [**HWZ96**, **HWZ01**, **HWZ96**, **Bou02**]. Similar asymptotic estimates yield a result about "long cylinders with small area" (see [**HWZ02**] and [**BEH+03**, Prop. 5.7]) which helps in proving that neighboring levels connect to each other along breaking orbits.

Our situation in Proposition 10.19 was simple enough to avoid using the "long cylinder" lemma, and we did use the discreteness of the action spectrum, but only needed it for orbits in $\mathcal{P}_h(\alpha)$ since we were able to rule out bubbling in the first step. An alternative would have been to assume that all orbits (in all homotopy classes) with period up to the period of γ are nondegenerate: then (10.2) implies that degenerate orbits never play any role in the main arguments of [BEH+03], so the big theorem becomes safe to use.

10.2.3. The chain complex. We now define a $\mathbb{Z}_2$-graded chain complex with coefficients in $\mathbb{Z}_2$ and generators $\langle\gamma\rangle$ for $\gamma \in \mathcal{P}_h(\alpha)$, i.e.

$$CC_*^h(M,\alpha) := \bigoplus_{\gamma \in \mathcal{P}_h(\alpha)} \mathbb{Z}_2.$$

The degree of each generator $\langle\gamma\rangle \in CC_*^h(M,\alpha)$ is defined by

$$|\langle\gamma\rangle| = n - 3 + \mu_{\mathrm{CZ}}(\gamma) \in \mathbb{Z}_2,$$

where $\mu_{\mathrm{CZ}}(\gamma) \in \mathbb{Z}_2$ denotes the parity of the Conley-Zehnder index with respect to any choice of trivialization. The choice to write $n-3$ in front of this is a convention that will make no difference at all in this lecture, but it is consistent with a $\mathbb{Z}$-grading that we will be able to define under suitable assumptions in Lecture 12. To define the differential on $CC_*^h(M,\alpha)$, choose an h-regular almost complex structure $J \in \mathcal{J}(\alpha)$. Given Reeb orbits $\gamma^+, \gamma^- \in \mathcal{P}_h(\alpha)$ and a number $I \in \mathbb{Z}$, let

$$\mathcal{M}^I(J,\gamma^+,\gamma^-)$$

denote the space of all $\mathbb{R}$-equivalence classes of index I holomorphic cylinders in $(\mathbb{R}\times M, J)$ asymptotic to $\gamma^\pm$ at $\pm\infty$, i.e. the union of the spaces $\mathcal{M}_{0,0}(J,A,\gamma^+,\gamma^-)/\mathbb{R}$ for all relative homology classes A such that vir-dim $\mathcal{M}_{0,0}(J,A,\gamma^+,\gamma^-) = I$. Since J is h-regular, all the curves in $\mathcal{M}^I(J,\gamma^+,\gamma^-)$ are Fredholm regular, so if $I \geqslant 1$, $\mathcal{M}^I(J,\gamma^+,\gamma^-)$ is a smooth manifold with

$$\dim \mathcal{M}^I(J,\gamma^+,\gamma^-) = I - 1.$$

Similarly, $\mathcal{M}^0(J,\gamma^+,\gamma^-)$ only contains trivial cylinders and is thus empty unless $\gamma^+ = \gamma^-$, and $\mathcal{M}^I(J,\gamma^+,\gamma^-)$ is always empty for $I < 0$. In particular, $\mathcal{M}^1(J,\gamma^+,\gamma^-)$ is a discrete set whenever $\gamma^+ \neq \gamma^-$, and by Proposition 10.19, it is also compact, hence finite. We can therefore define

$$\partial\langle\gamma\rangle = \sum_{\gamma' \in \mathcal{P}_h(\alpha)} \#_2\mathcal{M}^1(J,\gamma,\gamma')\langle\gamma'\rangle,$$

where for any set X, we denote by $\#_2 X$ the cardinality of X modulo 2. The operator ∂ has odd degree with respect to the grading since every index 1 holomorphic cylinder u with asymptotic orbits γ^+ and γ^- satisfies

$$\mathrm{ind}(u) = 1 = \mu_{\mathrm{CZ}}^\tau(\gamma^+) - \mu_{\mathrm{CZ}}^\tau(\gamma^-)$$

for suitable choices of the trivialization τ.

10.2.4. The homology. Following the standard Floer theoretic prescription, the relation $\partial^2 = 0$ should arise by viewing the compactification $\overline{\mathcal{M}}^2(J, \gamma^+, \gamma^-)$ for each $\gamma^+, \gamma^- \in \mathcal{P}_h(\alpha)$ as a compact 1-manifold whose boundary is identified with the set of rigid broken cylinders, as these are what is counted by ∂^2. Here $\overline{\mathcal{M}}^2(J, \gamma^+, \gamma^-)$ is defined as the closure of $\mathcal{M}^2(J, \gamma^+, \gamma^-)$ in the space of all J-holomorphic buildings in $\mathbb{R} \times M$ modulo $\mathbb{R}$-translation. Proposition 10.19 gives a natural inclusion

$$\overline{\mathcal{M}}^2(J, \gamma^+, \gamma^-) \backslash \mathcal{M}^2(J, \gamma^+, \gamma^-) \subset \coprod_{\gamma_0 \in \mathcal{P}_h(\alpha)} \mathcal{M}^1(J, \gamma^+, \gamma_0) \times \mathcal{M}^1(J, \gamma_0, \gamma^-).$$

We therefore need an inclusion in the other direction, and for this we need to say a word about gluing. We have not had time to discuss gluing in earnest in this book, and we will not do so now either, but the basic idea should be familiar from Floer homology: given $u_+ \in \mathcal{M}^1(J, \gamma^+, \gamma_0)$ and $u_- \in \mathcal{M}^1(J, \gamma_0, \gamma^-)$, one would like to show that there exists a unique (up to $\mathbb{R}$-translation) one-parameter family $\{u_R \in \mathcal{M}^2(J, \gamma^+, \gamma^-)\}_{R \in [R_0, \infty)}$ such that u_R converges as $R \to \infty$ to the building $\mathbf{u}_\infty$ with bottom level u_- and top level u_+. One starts by constructing a family of *preglued* maps

$$\tilde{u}_R : \mathbb{R} \times S^1 \to \mathbb{R} \times M,$$

meaning a smooth family of maps which converge in the SFT topology as $R \to \infty$ to $\mathbf{u}_\infty$ but are only *approximately* J-holomorphic. More precisely, fix parametrizations of u_- and u_+ and a parametrization of the orbit $\gamma_0 : \mathbb{R}/T\mathbb{Z} \to M$ such that

$$\begin{aligned} u_+(s,t) &= \exp_{(Ts, \gamma_0(Tt))} h_+(s,t) && \text{for } s \ll 0, \\ u_-(s,t) &= \exp_{(Ts, \gamma_0(Tt))} h_-(s,t) && \text{for } s \gg 0, \end{aligned}$$

where $h_\pm$ are vector fields along the trivial cylinder satisfying $\lim_{s \to \mp\infty} h_\pm(s,t) = 0$. By interpolating between suitable reparametrizations of h_+ and h_-, one can now define $\tilde{u}_R$ such that

$$\begin{aligned} \tilde{u}_R(s,t) &= \tau_{2RT} \circ u_+(s - 2R, t) && \text{for } s \geqslant R, \\ \tilde{u}_R(s,t) &\approx (Ts, \gamma_0(Tt)) && \text{for } s \in [-R, R], \\ \tilde{u}_R(s,t) &= \tau_{-2RT} \circ u_-(s + 2R, t) && \text{for } s \leqslant -R, \\ \bar{\partial}_J \tilde{u}_R &\to 0 && \text{as } R \to \infty. \end{aligned}$$

Given regularity of u_+ and u_-, one can now use a quantitative version of the implicit function theorem (cf. [**MS12**, §3.5]) to show that a distinguished J-holomorphic cylinder u_R close to $\tilde{u}_R$ exists for all R sufficiently large. For a more detailed synopsis of the analysis involved, see [**Nel13**, Chapter 7], and [**AD14**, Chapters 9 and 13] for the analogous story in Floer homology. The result is:

PROPOSITION 10.21. *For an h-admissible α, an h-regular $J \in \mathcal{J}(\alpha)$ and any two orbits $\gamma^+, \gamma^- \in \mathcal{P}_h(\alpha)$, the space $\overline{\mathcal{M}}^2(J, \gamma^+, \gamma^-)$ admits the structure of a compact 1-dimensional manifold with boundary, where its boundary points can be identified naturally with $\coprod_{\gamma_0 \in \mathcal{P}_h(\alpha)} \mathcal{M}^1(J, \gamma^+, \gamma_0) \times \mathcal{M}^1(J, \gamma_0, \gamma^-)$.* □

COROLLARY 10.22. *The homomorphism $\partial : CC_*^h(M, \alpha) \to CC_{*-1}^h(M, \alpha)$ satisfies $\partial^2 = 0$.* □

We shall denote the homology of this chain complex by

$$HC_*^h(M, \alpha, J) := H_*\big(CC_*^h(M, \alpha), \partial\big).$$

The goal of the rest of this section is to prove that up to natural isomorphisms, $HC_*^h(M, \alpha, J)$ depends on (M, ξ) and h but not on the auxiliary data α and J.

10.2.5. Chain maps. For any constant $c > 0$, there is an obvious bijection between the generators of $CC_*^h(M, \alpha)$ and $CC_*^h(M, c\alpha)$, as the rescaling changes periods of orbits but not the set of closed orbits itself. Moreover, if $J \in \mathcal{J}(\alpha)$ and $J_c \in \mathcal{J}(c\alpha)$ are defined to match on ξ, then there is a biholomorphic diffeomorphism

$$(\mathbb{R} \times M, J) \to (\mathbb{R} \times M, J_c) : (r, x) \mapsto (cr, x),$$

thus giving a bijective correspondence between the moduli spaces of J-holomorphic and J_c-holomorphic curves. It follows that our bijection of chain complexes is also a chain map and therefore defines a canonical isomorphism

$$HC_*^h(M, \alpha, J) = HC_*^h(M, c\alpha, J_c). \tag{10.3}$$

Next suppose α_- and α_+ are two distinct contact forms for ξ, hence

$$\alpha_\pm = e^{f_\pm}\alpha$$

for some fixed contact form α and a pair of smooth functions $f_\pm : M \to \mathbb{R}$. After rescaling α_+ by a constant, we are free to assume $f_+ > f_-$ everywhere. Fix h-regular almost complex structures $J_\pm \in \mathcal{J}(\alpha_\pm)$ and let

$$\partial_\pm : CC_*^h(M, \alpha_\pm) \to CC_{*-1}^h(M, \alpha_\pm)$$

denote the resulting differentials on the two chain complexes. The region

$$W := \big\{(r, x) \in \mathbb{R} \times M \bigm| f_-(x) \leqslant r \leqslant f_+(x)\big\}$$

now defines an exact symplectic cobordism from (M, ξ) to itself: more precisely, setting

$$M_\pm := \big\{(f_\pm(x), x) \in W \bigm| x \in M\big\}$$

gives $\partial W = -M_- \coprod M_+$, and the Liouville form $\lambda := e^r\alpha$ satisfies $\lambda|_{TM_\pm} = \alpha_\pm$. Choose a generic $d\lambda$-compatible almost complex structure J on the completion $\widehat{W}$ that restricts to $J_\pm$ on the cylindrical ends. Now given $\gamma^+ \in \mathcal{P}_h(\alpha_+)$ and $\gamma^- \in \mathcal{P}_h(\alpha_-)$ and a number $I \in \mathbb{Z}$, we shall denote by

$$\mathcal{M}^I(J, \gamma^+, \gamma^-)$$

the union of the spaces $\mathcal{M}_{0,0}(J, A, \gamma^+, \gamma^-)$ for all relative homology classes A such that $\text{vir-dim}\,\mathcal{M}_{0,0}(J, A, \gamma^+, \gamma^-) = I$. Note that we are not dividing by any $\mathbb{R}$-action here since J need not be $\mathbb{R}$-invariant. Since $\gamma^\pm$ are still guaranteed to be simply covered, curves in $\mathcal{M}^I(J, \gamma^+, \gamma^-)$ are again always somewhere injective and therefore regular, hence $\mathcal{M}^I(J, \gamma^+, \gamma^-)$ is a smooth manifold with

$$\dim \mathcal{M}^I(J, \gamma^+, \gamma^-) = I$$

if $I \geqslant 0$, and $\mathcal{M}^I(J, \gamma^+, \gamma^-) = \varnothing$ for $I < 0$. The compactification $\overline{\mathcal{M}}^I(J, \gamma^+, \gamma^-)$ is described via the following straightforward generalization of Proposition 10.19:

PROPOSITION 10.23. *For J as described above, suppose u_k is a sequence of J-holomorphic cylinders in $\widehat{W}$ with one positive puncture at an orbit $\gamma \in \mathcal{P}_h(\alpha_+)$ and one negative puncture. Then u_k has a subsequence convergent in the SFT topology to a broken J-holomorphic cylinder, i.e. a stable building $\mathbf{u}_\infty$ whose levels u_∞^N for $N = -N_-, \dots, -1, 0, 1, \dots, N_+$ are each cylinders with one positive and one negative puncture, living in $\mathbb{R} \times M^\pm$ for $\pm N > 0$ and $\widehat{W}$ for $N = 0$. Moreover, the levels satisfy $\operatorname{ind}(u_\infty^0) \geqslant 0$ and $\operatorname{ind}(u_\infty^N) \geqslant 1$ for $N \neq 0$, thus for large k in the convergent subsequence,*

$$\operatorname{ind}(u_k) = \sum_{N=-N_-}^{N_+} \operatorname{ind}(u_\infty^N) \geqslant N_- + N_+.$$

□

It follows that the set $\mathcal{M}^0(J, \gamma^+, \gamma^-)$ is always finite, and we use this to define a map

$$\Phi_J : CC_*^h(M, \alpha_+) \to CC_*^h(M, \alpha_-) : \langle \gamma \rangle \mapsto \sum_{\gamma' \in \mathcal{P}_h(\alpha_-)} \#_2 \mathcal{M}^0(J, \gamma, \gamma') \langle \gamma' \rangle.$$

This map preserves degrees since it counts index 0 curves, and we claim that it is a chain map:

$$\Phi_J \circ \partial_+ = \partial_- \circ \Phi_J.$$

This follows from the fact that by Proposition 10.23 (in conjunction with a corresponding gluing theorem), $\overline{\mathcal{M}}^1(J, \gamma^+, \gamma^-)$ is a compact 1-manifold whose boundary consists of two types of broken cylinders, depending whether the index 1 curve appears in an upper or lower level:

$$\begin{aligned} \partial \overline{\mathcal{M}}^1(J, \gamma^+, \gamma^-) = & \coprod_{\gamma_0 \in \mathcal{P}_h(\alpha_+)} \left(\mathcal{M}^1(J_+, \gamma^+, \gamma_0) \times \mathcal{M}^0(J, \gamma_0, \gamma^-) \right) \\ & \cup \coprod_{\gamma_0 \in \mathcal{P}_h(\alpha_-)} \left(\mathcal{M}^0(J, \gamma^+, \gamma_0) \times \mathcal{M}^1(J_-, \gamma_0, \gamma^-) \right). \end{aligned}$$

Counting broken cylinders of the first type gives the coefficient in front of $\langle \gamma^- \rangle$ in $\Phi_J \circ \partial_+(\langle \gamma^+ \rangle)$, and the second type gives $\partial_- \circ \Phi_J(\langle \gamma^+ \rangle)$.

It follows that Φ_J descends to a homomorphism

$$\Phi_J : HC_*^h(M, \alpha_+, J_+) \to HC_*^h(M, \alpha_-, J_-). \tag{10.4}$$

10.2.6. Chain homotopies. We claim that the map Φ_J in (10.4) does not depend on J. To see this, suppose J_0 and J_1 are two generic choices of compatible almost complex structures on $\widehat{W}$ that both match $J_\pm$ on the cylindrical ends. The space of almost complex structures with these properties is contractible, so we can find a smooth path

$$\{J_s\}_{s \in [0,1]}$$

connecting them. For $I \in \mathbb{Z}$, consider the parametric moduli space

$$\mathcal{M}^I(\{J_s\}, \gamma^+, \gamma^-) := \left\{ (s, u) \;\middle|\; s \in [0,1],\ u \in \mathcal{M}^I(J_s, \gamma^+, \gamma^-) \right\}.$$

As we observed in Remark 8.7, a generic choice of the homotopy $\{J_s\}$ makes $\mathcal{M}^I(\{J_s\})$ a smooth manifold with

$$\dim \mathcal{M}^I(\{J_s\}, \gamma^+, \gamma^-) = I + 1$$

whenever $I \geqslant -1$, and $\mathcal{M}^I(\{J_s\}, \gamma^+, \gamma^-) = \varnothing$ when $I < -1$. Adapting Proposition 10.23 to allow for a converging sequence of almost complex structures, it implies that $\mathcal{M}^{-1}(\{J_s\}, \gamma^+, \gamma^-)$ is compact and thus finite, so we can use it to define a homomorphism of odd degree by

$$H : CC^h_*(M, \alpha_+) \to CC^h_{*+1}(M, \alpha_-) : \langle\gamma\rangle \mapsto \sum_{\gamma' \in \mathcal{P}_h(\alpha_-)} \#\mathcal{M}^{-1}(\{J_s\}, \gamma, \gamma')\langle\gamma'\rangle.$$

We claim that this is a chain homotopy between Φ_{J_0} and Φ_{J_1}, i.e.

$$\Phi_{J_1} - \Phi_{J_0} = \partial_- \circ H + H \circ \partial_+.$$

This follows by looking at the boundary of the compactified 1-dimensional space $\overline{\mathcal{M}}^0(\{J_s\}, \gamma^+, \gamma^-)$, which consists of four types of objects:

(1) Pairs $(0, u)$ with $u \in \mathcal{M}^0(J_0, \gamma^+, \gamma^-)$, which are counted by Φ_{J_0}.
(2) Pairs $(1, u)$ with $u \in \mathcal{M}^0(J_1, \gamma^+, \gamma^-)$, which are counted by Φ_{J_1}.
(3) Pairs $(s, \mathbf{u})$ with $\mathbf{u}$ a broken cylinder with upper level $u_+ \in \mathcal{M}^1(J_+, \gamma^+, \gamma_0)$ and main level $u_0 \in \mathcal{M}^{-1}(J_s, \gamma_0, \gamma^-)$ for some $s \in (0, 1)$; these are counted by $H \circ \partial_+$.
(4) Pairs $(s, \mathbf{u})$ with $\mathbf{u}$ a broken cylinder with lower level $u_- \in \mathcal{M}^1(J_-, \gamma_0, \gamma^-)$ and main level $u_0 \in \mathcal{M}^{-1}(J_s, \gamma^+, \gamma_0)$ for some $s \in (0, 1)$; these are counted by $\partial_- \circ H$.

The sum $\Phi_{J_0} + \Phi_{J_1} + \partial_- \circ H + H \circ \partial_+$ therefore counts (modulo 2) the boundary points of a compact 1-manifold, so it vanishes.

Since the action of Φ_J on homology no longer depends on J, we will denote it from now on by

$$\Phi : HC^h_*(M, \alpha_+, J_+) \to HC^h_*(M, \alpha_-, J_-).$$

It is well defined for any pair of h-admissible contact forms $\alpha_\pm$ and h-regular $J_\pm \in \mathcal{J}(\alpha_\pm)$ since one can first rescale α_+ to assume $\alpha_\pm = e^{f_\pm}\alpha$ with $f_+ > f_-$, using the canonical isomorphism (10.3).

10.2.7. Proof of invariance. We claim that for any h-admissible α and h-regular $J \in \mathcal{J}(\alpha)$, the cobordism map

$$\Phi : HC^h_*(M, \alpha, J) \to HC^h_*(M, \alpha, J)$$

is the identity. Indeed, the literal meaning of this statement is that for any $c > 1$, the composition of the canonical isomorphism (10.3) with the map

$$\Phi : HC^h_*(M, c\alpha, J_c) \to HC^h_*(M, \alpha, J)$$

defined by counting index 0 cylinders in a trivial cobordism from (M, α, J) to $(M, c\alpha, J_c)$ is the identity. Writing $c = e^a$ for $a > 0$, the Liouville cobordism in question is simply

$$(W, d\lambda) = ([0, a] \times M, d(e^r\alpha)),$$

and one can choose a compatible almost complex structure on this which matches J and J_c on ξ while taking ∂_r to $g(r)R_\alpha$ for a suitable function g with $g(r) = 1$ near $r = 0$ and $g(r) = 1/c$ near $r = a$. The resulting almost complex manifold is biholomorphically diffeomorphic to the usual symplectization $(\mathbb{R} \times M, J)$, so our count of index 0 cylinders is equivalent to the count of such cylinders in $(\mathbb{R} \times M, J)$. The latter are simply the trivial cylinders, all of which are Fredholm regular, so counting these defines the identity map on the chain complex.

Finally, we need to show that for any three h-admissible pairs (α_i, J_i) with $i = 0, 1, 2$, the cobordism maps $\Phi_{ij} : HC^h_*(M, \alpha_j, J_j) \to HC^h_*(M, \alpha_i, J_i)$ satisfy

$$\Phi_{21} \circ \Phi_{10} = \Phi_{20}. \tag{10.5}$$

We will only sketch this part: the idea is to use a stretching construction. After rescaling, suppose without loss of generality that $\alpha_i = e^{f_i}\alpha$ with $f_2 > f_1 > f_0$. Then the cobordism

$$W_{20} := \left\{ (r, x) \;\middle|\; f_0(x) \leqslant r \leqslant f_2(x) \right\}$$

contains a contact-type hypersurface

$$M_1 := \left\{ (f_1(x), x) \;\middle|\; x \in M \right\} \subset W_{20}.$$

As described in §9.4.4, one can now choose a sequence of compatible almost complex structures $\{J^N_{20}\}_{N \in \mathbb{N}}$ on $\widehat{W}_{20}$ that are fixed outside a neighborhood of M_1 but degenerate in this neighborhood as $N \to \infty$, equivalent to replacing a small tubular neighborhood of M_1 with increasingly large collars $[-N, N] \times M$ in which J^N_{20} belongs to $\mathcal{J}(\alpha_1)$. The resulting chain maps

$$\Phi_{J^N_{20}} : CC^h_*(M, \alpha_2, J_2) \to CC^h_*(M, \alpha_0, J_0)$$

are chain homotopic for all N, but as $N \to \infty$, the index 0 cylinders counted by these maps converge to buildings with two levels, the top one an index 0 cylinder in the completion of a cobordism from (M, α_1, J_1) to (M, α_2, J_2), while the bottom one also has index 0 and lives in a cobordism from (M, α_0, J_0) to (M, α_1, J_1). The composition $\Phi_{21} \circ \Phi_{10}$ counts these broken cylinders, so this proves (10.5).

In particular, we conclude now that each of the cobordism maps

$$\Phi : HC^h_*(M, \alpha_+, J_+) \to HC^h_*(M, \alpha_-, J_-)$$

is an isomorphism, since composing it with a cobordism map in the opposite direction must give the identity. The isomorphism class of $HC^h_*(M, \alpha, J)$ is therefore independent of the auxiliary data (α, J), and will be denoted by

$$HC^h_*(M, \xi).$$

This is the **cylindrical contact homology** of (M, ξ) in the homotopy class h. It is defined for any primitive homotopy class $h \in [S^1, M]$ and closed contact manifold that is h-admissible in the sense of Definition 10.16. It is also invariant under contactomorphisms in the following sense:

PROPOSITION 10.24. *Suppose $\varphi : (M_0, \xi_0) \to (M_1, \xi_1)$ is a contactomorphism with $\varphi_* h_0 = h_1$, where $h_0 \in [S^1, M]$ is a primitive homotopy class of loops, and (M_1, ξ_1) is h_1-admissible. Then (M_0, ξ_0) is h_0-admissible, and $HC^{h_0}_*(M_0, \xi_0) \cong HC^{h_1}_*(M_1, \xi_1)$.*

PROOF. Given an h_1-admissible contact form α_1 on (M_1, ξ_1) and an h_1-regular $J_1 \in \mathcal{J}(\alpha_1)$, the contact form $\alpha_0 := \varphi^*\alpha_1$ on M_0 is h_0-admissible since φ defines a bijection from $\mathcal{P}_{h_0}(\alpha_0)$ to $\mathcal{P}_{h_1}(\alpha_1)$ and also a bijection between the sets of contractible Reeb orbits for α_0 and α_1. Since $\varphi_*\xi_0 = \xi_1$, α_0 is a contact form for (M_0, ξ_0), hence the latter is h_0-admissible. The diffeomorphism $\widetilde{\varphi} := \mathrm{Id} \times \varphi : \mathbb{R} \times M_0 \to \mathbb{R} \times M_1$ then maps ∂_r to ∂_r, R_{α_0} to R_{α_1} and ξ_0 to ξ_1, thus $J_0 := \widetilde{\varphi}^* J_1 \in \mathcal{J}(\alpha_0)$, so $\widetilde{\varphi}$ defines a biholomorphic map $(\mathbb{R} \times M_0, J_0) \to (\mathbb{R} \times M_1, J_1)$ and thus a bijection between the sets of holomorphic cylinders in each. It follows that J_0 is h_0-regular, and the bijection $\mathcal{P}_{h_0}(\alpha_0) \to \mathcal{P}_{h_1}(\alpha_1)$ defines an isomorphism between the chain complexes defining $HC_*^{h_0}(M_0, \alpha_0, J_0)$ and $HC_*^{h_1}(M_1, \alpha_1, J_1)$. $\square$

10.3. Computing $HC_*(\mathbb{T}^3, \xi_k)$

10.3.1. The Morse-Bott setup. The contact form α_k on $\mathbb{T}^3$ defined at the beginning of this lecture has Reeb vector field

$$R_k(\rho, \phi, \theta) = \cos(2\pi k\rho)\, \partial_\theta + \sin(2\pi k\rho)\, \partial_\phi.$$

Its Reeb orbits therefore preserve each of the tori $\{\rho\} \times \mathbb{T}^2$ and define linear foliations on them. In particular, none of the closed orbits are contractible, though all of them are also degenerate, as they all come in S^1-parametrized families foliating $\{\text{const}\} \times \mathbb{T}^2$. For certain homotopy classes $h \in [S^1, \mathbb{T}^3]$, this yields a very easy computation of $HC_*^h(\mathbb{T}^3, \xi_k)$, namely whenever h contains no periodic orbits:

THEOREM 10.25. *Suppose $h \in [S^1, \mathbb{T}^3]$ is any primitive homotopy class of loops such that the projection $p : \mathbb{T}^3 \to S^1 : (\rho, \phi, \theta) \mapsto \rho$ satisfies $p_* h \neq 0 \in [S^1, S^1]$. Then α_k is h-admissible and the resulting contact homology $HC_*^h(\mathbb{T}^3, \xi_k)$ is trivial.* $\square$

Now for the interesting part. Every primitive class $h \in [S^1, \mathbb{T}^3]$ not covered by Theorem 10.25 contains closed orbits of R_k, all of them degenerate since they come in S^1-parametrized families foliating the tori $\{\text{const}\} \times \mathbb{T}^2$. This makes it not immediately clear whether $(\mathbb{T}^3, \xi_k)$ is h-admissible, though the following observation in conjunction with Proposition 10.24 shows that if $HC_*^h(\mathbb{T}^3, \xi_k)$ can be defined, it will be the same for all the homotopy classes under consideration.

LEMMA 10.26. *Suppose $h_0, h_1 \in [S^1, \mathbb{T}^3]$ are primitive homotopy classes that are both mapped to the trivial class under the projection $\mathbb{T}^3 \to S^1 : (\rho, \phi, \theta) \mapsto \rho$. Then there exists a contactomorphism $\varphi : (\mathbb{T}^3, \xi_k) \to (\mathbb{T}^3, \xi_k)$ satisfying $\varphi_* h_0 = h_1$.*

PROOF. We can represent h_i for $i = 0, 1$ by loops of the form $\gamma_i(t) = (0, \beta_i(t)) \in S^1 \times \mathbb{T}^2$, where the loops $\beta_i : S^1 \to \mathbb{T}^2$ are embedded and thus represent generators of $\pi_1(\mathbb{T}^2) = \mathbb{Z}^2$. One can thus find a matrix $\begin{pmatrix} m & n \\ p & q \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z})$ such that the diffeomorphism

$$\varphi : \mathbb{T}^3 \to \mathbb{T}^3 : (\rho, \phi, \theta) \mapsto (\rho, m\phi + n\theta, p\phi + q\theta)$$

satisfies $\varphi_* h_0 = h_1$. We have

$$\begin{aligned} \varphi^*\alpha_k &= [q\cos(2\pi k\rho) + n\sin(2\pi k\rho)]\, d\theta + [p\cos(2\pi k\rho) + m\sin(2\pi k\rho)]\, d\phi \\ &=: F(\rho)\, d\theta + G(\rho)\, d\phi. \end{aligned}$$

The loop $(F,G) : S^1 \to \mathbb{R}^2$ satisfies

$$\begin{pmatrix} F(\rho) \\ G(\rho) \end{pmatrix} = \begin{pmatrix} q & n \\ p & m \end{pmatrix} \begin{pmatrix} \cos(2\pi k\rho) \\ \sin(2\pi k\rho) \end{pmatrix},$$

where $\begin{pmatrix} q & n \\ p & m \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$, thus (F,G) winds k times about the origin. Any choice of homotopy from (F,G) to $(\cos(2\pi k\rho), \sin(2\pi k\rho))$ through loops $(F_s, G_s) : S^1 \to \mathbb{R}^2$ winding k times about the origin with positive rotational velocity then gives rise to a homotopy from $\varphi^*\alpha_k$ to α_k through contact forms $F_s(\rho)\, d\theta + G_s(\rho)\, d\phi$. Gray's stability theorem therefore yields a contactomorphism $\psi : (\mathbb{T}^3, \xi_k) \to (\mathbb{T}^3, \ker \varphi^*\alpha_k)$ with ψ smoothly isotopic to the identity. The map $\varphi \circ \psi$ is thus a contactomorphism of $(\mathbb{T}^3, \xi_k)$ with $(\varphi \circ \psi)_* h_0 = \varphi_* \psi_* h_0 = \varphi_* h_0 = h_1$. $\square$

In light of the lemma, we are free from now on to restrict our attention to the particular homotopy class

$$h := [t \mapsto (0,0,t)],$$

which is the homotopy class of the 1-periodic orbits foliating the k tori

$$T_m := \{m/k\} \times \mathbb{T}^2, \qquad m = 0, \dots, k-1$$

since $R_k(m/k, \phi, \theta) = \partial_\theta$. Though the orbits on these tori are degenerate, it is not hard to show that they all satisfy the Morse-Bott condition; in fact, α_k is a Morse-Bott contact form. We will explain a self-contained computation of $HC_*^h(\mathbb{T}^3, \xi_k)$ in the next two sections without using the Morse-Bott condition—but first, it seems worthwhile to sketch how one can guess the answer using Morse-Bott data.

Bourgeois's thesis [**Bou02**] gives a prescription for calculating contact homology in Morse-Bott settings, i.e. for deducing what orbits and what holomorphic curves will appear under certain standard ways of perturbing the Morse-Bott contact form to make it nondegenerate. Notice first that the only orbits in $\mathcal{P}_h(\alpha_k)$ are the ones that foliate the k tori $T_0, \dots, T_{k-1}$, and they all have period 1. By (10.2), it follows that for any $J \in \mathcal{J}(\alpha_k)$, there can be no nontrivial J-holomorphic cylinders connecting two orbits in $\mathcal{P}_h(\alpha_k)$. This makes the calculation of $HC_*^h(\mathbb{T}^3, \xi_k)$ sound trivial, but of course there is more to the story since α_k is not admissible; indeed, the chain complex $CC_*(\mathbb{T}^3, \alpha_k)$ is not even well defined. The prescription in [**Bou02**] now gives the following. Each of the families of orbits in $T_0, \dots, T_{k-1}$ is parametrized by S^1, and by a standard perturbation technique, any choice of a Morse function $f_m : S^1 \to \mathbb{R}$ for $m = 0, \dots, k-1$ yields a contact form α_k' that is C^∞-close to α_k, matches it outside a neighborhood of T_m, but has a nondegenerate Reeb orbit on T_m for each critical point of f_m, while every other closed orbit in the perturbed region can be assumed to have arbitrarily large period. Moreover, there is a corresponding perturbation from $J \in \mathcal{J}(\alpha_k)$ to $J' \in \mathcal{J}(\alpha_k')$ such that every gradient flow line of the function $f_m : S^1 \to \mathbb{R}$ gives rise to a J'-holomorphic cylinder in $\mathbb{R} \times \mathbb{T}^3$ connecting the corresponding nondegenerate Reeb orbits along T_m. In the present situation, since no J-holomorphic cylinders of the relevant type exist before the perturbation, the only ones after the perturbation are those that come from gradient flow lines.

Now imagine performing a similar perturbation near every $T_0, \dots, T_{k-1}$, using Morse functions $f_0, \dots, f_{k-1} : S^1 \to \mathbb{R}$ that each have exactly two critical points.

For the perturbed contact form α'_k, $\mathcal{P}_h(\alpha'_k)$ now consists of exactly $2k$ orbits

$$\gamma_0^{\pm}, \ldots, \gamma_{k-1}^{\pm} \in \mathcal{P}_h(\alpha'_k),$$

where we denote by γ_m^+ and γ_m^- the orbits on T_m corresponding to the maximum and minimum of f_m respectively. For the obvious choice of trivialization τ for the contact bundle along $\gamma_m^{\pm}$, one can relate the Conley-Zehnder indices to the Morse indices of the corresponding critical points, giving

$$\mu_{\mathrm{CZ}}^{\tau}(\gamma_m^+) = 0, \qquad \mu_{\mathrm{CZ}}^{\tau}(\gamma_m^-) = 1, \qquad m = 0, \ldots, k-1.$$

Moreover, the two gradient flow lines connecting maximum and minimum for each f_m give rise two exactly two holomorphic cylinders in $\mathcal{M}^1(J', \gamma_m^-, \gamma_m^+)$ for each $m = 0, \ldots, k-1$, and these are all the curves that are counted for the differential on $CC_*^h(\mathbb{T}^3, \alpha'_k, J')$. Counting modulo 2, we thus have

$$\partial\langle\gamma_m^{\pm}\rangle = 0 \quad \text{for all} \quad m = 0, \ldots, k-1,$$

implying

$$HC_*^h(\mathbb{T}^3, \alpha'_k, J') = \begin{cases} \mathbb{Z}_2^k & * = \text{odd}, \\ \mathbb{Z}_2^k & * = \text{even}. \end{cases}$$

Let us state this as a theorem.

THEOREM 10.27. *Suppose $h \in [S^1, \mathbb{T}^3]$ is a primitive homotopy class that maps to the trivial class under the projection $\mathbb{T}^3 \to S^1 : (\rho, \phi, \theta) \mapsto \rho$. Then $(\mathbb{T}^3, \xi_k)$ is h-admissible and*

$$HC_*^h(\mathbb{T}^3, \xi_k) \cong \begin{cases} \mathbb{Z}_2^k & * = \textit{odd}, \\ \mathbb{Z}_2^k & * = \textit{even}. \end{cases}$$

Theorem 10.1 is an immediate corollary of this: indeed, if $\varphi : (\mathbb{T}^3, \xi_k) \to (\mathbb{T}^3, \xi_\ell)$ is a contactomorphism, choose any $h \in [S^1, \mathbb{T}^3]$ for which Theorem 10.27 applies, and let $h_0 := \varphi^* h \in [S^1, \mathbb{T}^3]$. Then $HC_*^h(\mathbb{T}^3, \xi_\ell) \cong \mathbb{Z}_2^{2\ell}$ implies via Proposition 10.24 that $HC_*^{h_0}(\mathbb{T}^3, \xi_k) \cong \mathbb{Z}_2^{2\ell}$. But Theorems 10.25 and 10.27 imply that the latter is also either 0 or $\mathbb{Z}_2^{2k}$, hence $k = \ell$.

10.3.2. A digression on the Floer equation. In preparation for giving a self-contained proof of Theorem 10.27, we now explain a general procedure for relating holomorphic cylinders in a symplectization to solutions of the Floer equation. This idea is loosely inspired by arguments in [**EKP06**].

To motivate what follows, notice that on a neighborhood of $T_0 = \{0\} \times \mathbb{T}^2 \subset (\mathbb{T}^3, \xi_k)$, we can write

$$\alpha_k = \cos(2\pi k\rho)\,(d\theta + \beta)\,,$$

where $\beta := \tan(2\pi k\rho)\, d\phi$ defines a Liouville form on the annulus $\mathbb{A}_k := [-1/8k, 1/8k] \times S^1$ with coordinates (ρ, ϕ). This makes the neighborhood $\mathbb{A}_k \times S^1 \subset (\mathbb{T}^3, \xi_k)$ a special case of the following general construction.

Definition 10.28. Suppose V is a $2n$-dimensional manifold with an exact symplectic form $d\beta$. The contact manifold $(V \times S^1, \ker(d\theta + \beta))$ is then called the **contactization** of (V, β).[4] Here θ denotes the coordinate on the S^1 factor.

It's easy to check that $d\theta + \beta$ is indeed a contact form on $V \times S^1$ whenever $d\beta$ is symplectic on V: the latter means $(d\beta)^n > 0$ on V, so

$$(d\theta + \beta) \wedge [d(d\theta + \beta)]^n = (d\theta + \beta) \wedge (d\beta)^n = d\theta \wedge (d\beta)^n > 0.$$

Now here's a cute trick one can play with contactizations. For the rest of this subsection, assume

$$(V, d\beta)$$

is an arbitrary compact $2n$-dimensional exact symplectic manifold with boundary. Fix a smooth function

$$H : V \times S^1 \to \mathbb{R},$$

which we shall think of in the following as a time-dependent Hamiltonian $H_\theta := H(\cdot, \theta) : V \to \mathbb{R}$ on $(V, d\beta)$. The 2-form on $V \times S^1$ defined by

$$\omega = d\beta + d\theta \wedge dH = d(\beta - H\, d\theta)$$

is then *fiberwise symplectic*, meaning its restriction to each of the fibers of the projection map $V \times S^1 \to S^1$ is symplectic. We claim that for every $\epsilon > 0$ sufficiently small,

$$\lambda_\epsilon := d\theta + \epsilon(\beta - H\, d\theta)$$

defines a contact form on $V \times S^1$. This is a variation on the construction that was used by Thurston and Winkelnkemper [**TW75**] to define contact forms out of open book decompositions, and the proof is simple enough: since $d\lambda_\epsilon = \epsilon\omega$, we just need to check that $\lambda_\epsilon \wedge \omega^n > 0$ for $\epsilon > 0$ sufficiently small, and indeed,

$$\lambda_\epsilon \wedge \omega^n = d\theta \wedge (d\beta)^n + \epsilon(\beta - H\, d\theta) \wedge \omega^n > 0$$

since the first term is a volume form and ϵ is small. To see the relation between λ_ϵ and the contactization, we can write

$$\lambda_\epsilon = (1 - \epsilon H)\, d\theta + \epsilon\beta = (1 - \epsilon H)\left(d\theta + \frac{\epsilon}{1 - \epsilon H}\beta\right)$$

and observe that $\frac{\epsilon}{1-\epsilon H}\beta$ is also a Liouville form on V whenever H is θ-independent and $\epsilon > 0$ is sufficiently small.

The Reeb vector fields R_ϵ for λ_ϵ vary with ϵ, but their directions do not, since $d\lambda_\epsilon = \epsilon\omega$ has the same kernel for every ϵ. Moreover, while λ_ϵ ceases to be a contact form when $\epsilon \to 0$, the Reeb vector fields still have a well-defined limit: they converge as $\epsilon \to 0$ to the unique vector field R_0 satisfying

$$d\theta(R_0) \equiv 1 \quad \text{and} \quad \omega(R_0, \cdot) \equiv 0.$$

The latter can be written more explicitly as

$$R_0 = \partial_\theta + X_\theta,$$

[4] Elsewhere in the literature, the contactization is also often defined as $V \times \mathbb{R}$ instead of $V \times S^1$. The usage here is consistent with [**MNW13**].

where X_θ is the time-dependent Hamiltonian vector field determined by H_θ, i.e. via the condition

$$d\beta(X_\theta, \cdot) = -dH_\theta.$$

As one can easily compute, the reason for this nice behavior as $\epsilon \to 0$ is that the R_ϵ are also the Reeb vector fields for a smooth family of stable Hamiltonian structures:

PROPOSITION 10.29. *The pairs $\mathcal{H}_\epsilon := (\omega, \lambda_\epsilon)$ for $\epsilon \geqslant 0$ sufficiently small define a smooth family of stable Hamiltonian structures whose Reeb vector fields are R_ϵ.* □

We shall write the hyperplane distributions induced by $\mathcal{H}_\epsilon$ as

$$\Xi_\epsilon := \ker \lambda_\epsilon \subset T(V \times S^1).$$

These are contact structures for $\epsilon > 0$ small, and the space $\mathcal{J}(\mathcal{H}_\epsilon)$ of $\mathbb{R}$-invariant almost complex structures on $\mathbb{R} \times (V \times S^1)$ compatible with $\mathcal{H}_\epsilon$ is then identical to $\mathcal{J}(\lambda_\epsilon)$. On the other hand for $\epsilon = 0$, $\Xi_0 = \ker d\theta$ is a foliation, namely it is the vertical subbundle of the trivial fibration $V \times S^1 \to S^1$. We have seen $\mathcal{H}_0$ before: it is a variation on the "Floer-type" stable Hamiltonian structure that was considered in §6.3.2. Its closed Reeb orbits in the homotopy class of $\gamma : S^1 \to V \times S^1 : t \mapsto (\text{const}, t)$ are all of the form $\gamma(t) = (x(t), t)$ where $x : S^1 \to V$ is a contractible 1-periodic orbit of X_θ. Moreover, suppose $J \in \mathcal{J}(\mathcal{H}_0)$, which is equivalent to a choice of compatible complex structure on the symplectic bundle $(\Xi_0, \omega|_{\Xi_0})$, or in other words, an S^1-parametrized family of $d\beta$-compatible almost complex structures $\{J_\theta\}_{\theta \in S^1}$ on V. Then if

$$u = (f, v, g) : \mathbb{R} \times S^1 \to \mathbb{R} \times (V \times S^1)$$

is a J-holomorphic cylinder asymptotic at $\{\pm\infty\} \times S^1$ to two orbits of the form described above, the nonlinear Cauchy-Riemann equation for u turns out to imply that $(f, g) : \mathbb{R} \times S^1 \to \mathbb{R} \times S^1$ is a holomorphic map with degree 1 sending $\{\pm\infty\} \times S^1$ to $\{\pm\infty\} \times S^1$, and we can therefore choose a unique biholomorphic reparametrization of u so that (f, g) becomes the identity map. Having done this, the equation satisfied by $v : \mathbb{R} \times S^1 \to V$ is now

$$\partial_s v + J_t(v)(\partial_t v - X_t(v)) = 0,$$

in other words, the Floer equation for the data $\{J_\theta\}_{\theta \in S^1}$ and $\{H_\theta\}_{\theta \in S^1}$.

To complete the analogy, notice that since ω is exact, we can write down a natural symplectic action functional with respect to each $\mathcal{H}_\epsilon$ as

$$\mathcal{A}_\epsilon : C^\infty(S^1, V \times S^1) \to \mathbb{R} : \gamma \mapsto \int_{S^1} \gamma^*(\beta - H\, d\theta).$$

For loops of the form $\gamma(t) = (x(t), t)$ with $x : S^1 \to V$ contractible, this reduces (give or take a sign—see Remark 10.31) to the usual formula for the Floer action functional

$$\mathcal{A}_H(\gamma) = \int_{S^1} x^*\beta - \int_{S^1} H(x(t))\, dt = \int_{\mathbb{D}} \bar{x}^* d\beta - \int_{S^1} H(x(t))\, dt, \tag{10.6}$$

where $\bar{x} : \mathbb{D} \to V$ is any map satisfying $\bar{x}|_{\partial\mathbb{D}} = x$. Stokes' theorem gives an easy relation between the action and the so-called *ω-energy* if $u : \mathbb{R} \times S^1 \to \mathbb{R} \times (V \times S^1)$

is a J-holomorphic curve for $J \in \mathcal{J}(\mathcal{H}_\epsilon)$ and is positively/negatively asymptotic to orbits $\gamma^\pm : S^1 \to V \times S^1$ at $s = \pm\infty$: we have

$$0 \leqslant \int_{\mathbb{R}\times S^1} u^*\omega = \mathcal{A}_\epsilon(\gamma^+) - \mathcal{A}_\epsilon(\gamma^-).$$

If $u(s,t) = (s, v(s,t), t)$, then the left hand side is identical to the definition of energy in Floer homology, namely

$$E_H(v) := \int_{\mathbb{R}\times S^1} d\beta(\partial_s v, \partial_t v - X_t(v))\, ds \wedge dt = \int_{\mathbb{R}\times S^1} d\beta(\partial_s v, J_t(v)\partial_s v)\, ds \wedge dt,$$

thus giving the familiar relation

$$E_H(v) = \mathcal{A}_H(\gamma^+) - \mathcal{A}_H(\gamma^-). \tag{10.7}$$

To relate this to the usual notion of energy with respect to a stable Hamiltonian structure, we write the usual formula

$$E_\epsilon(u) := \sup_{\varphi\in\mathcal{T}} \int_{\dot\Sigma} u^* \left[d\big(\varphi(r)\lambda_\epsilon\big) + \omega\right],$$

with $\mathcal{T} := \left\{\varphi \in C^\infty(\mathbb{R}, (-\epsilon_0, \epsilon_0)) \;\middle|\; \varphi' > 0\right\}$ for some constant $\epsilon_0 > 0$ sufficiently small. Notice first that for any fixed ϵ, Stokes' theorem gives a bound for $E_\epsilon(u)$ in terms of the asymptotic orbits of u since ω is exact. Finally, in the case $\epsilon = 0$ with $u(s,t) = (s, v(s,t), t)$, we find

$$E_0(u) = \sup_{\varphi\in\mathcal{T}} \int_{\mathbb{R}\times S^1} \varphi'(s)\, ds \wedge dt + \int_{\mathbb{R}\times S^1} u^*\omega = 2\epsilon_0 + E_H(v),$$

so bounds on $E_0(u)$ are equivalent to bounds on the Floer homological energy $E_H(v)$. The basic fact that Floer trajectories $v : \mathbb{R}\times S^1 \to V$ with $E_H(v) < \infty$ are asymptotic to contractible 1-periodic Hamiltonian orbits can now be regarded as a corollary of our Theorem 9.8.

The above discussion gives a one-to-one correspondence between a certain moduli space of unparametrized J-holomorphic cylinders in $\mathbb{R} \times (V \times S^1)$ and the moduli space of Floer trajectories between contractible 1-periodic orbits in $(V, d\beta)$ with Hamiltonian function H. If we can adequately understand the moduli space of Floer trajectories—in particular if we can classify them and prove that they are regular—then the idea will be to extend this classification via the implicit function theorem to any $J_\epsilon \in \mathcal{J}(\lambda_\epsilon)$ sufficiently close to J for $\epsilon > 0$ small. As the reader may be aware, classifying Floer trajectories is also not easy in general, but it does become easy under certain conditions. Simple examples of contractible 1-periodic Hamiltonian orbits are furnished by the constant loops $\gamma(t) = x$ at critical points $x \in \mathrm{Crit}(H)$, and for each such orbit, $\gamma^*\Xi_0$ has a canonical homotopy class of unitary trivializations, the so-called **constant trivialization**. The following fundamental result is commonly used in proving the isomorphism from Hamiltonian Floer homology to singular homology.

THEOREM 10.30. *Suppose $H : V \to \mathbb{R}$ is a smooth Morse function with no critical points on the boundary, J is a fixed $d\beta$-compatible almost complex structure on V, and the gradient flow of H with respect to the metric $d\beta(\cdot, J\cdot)$ is Morse-Smale* *and transverse to ∂V. Given $c > 0$, let $H^c := cH : V \to \mathbb{R}$, with Hamiltonian vector field $X_{H^c} = cX_H$, and consider the stable Hamiltonian structure*

$$\mathcal{H}_0^c := (d\beta + d\theta \wedge dH^c, d\theta)$$

on $V \times S^1$ with induced Reeb vector field $R_0^c = \partial_\theta + X_{H^c}$. Then for all $c > 0$ sufficiently small, the following statements hold.

(1) The 1-periodic R_0^c-orbit $\gamma_x : S^1 \to V \times S^1 : t \mapsto (x,t)$ arising from any critical point $x \in \mathrm{Crit}(H)$ is nondegenerate, and its Conley-Zehnder index relative to the constant trivialization τ is related to the Morse index $\mathrm{ind}(x) \in \{0,\dots,2n\}$ by

$$\mu_{CZ}^\tau(\gamma_x) = n - \mathrm{ind}(x). \tag{10.8}$$

(2) Any trajectory $\gamma : \mathbb{R} \to V$ satisfying the negative gradient flow equation $\dot\gamma = -\nabla H^c(\gamma)$ gives rise to a Fredholm regular solution $v : \mathbb{R} \times S^1 \to V : (s,t) \mapsto \gamma(s)$ of the time-independent Floer equation

$$\partial_s v + J(v)(\partial_t v - X_{H^c}(v)) = 0, \tag{10.9}$$

and the virtual dimensions of the spaces of Floer trajectories near v and gradient flow trajectories near γ are the same.

(3) Every 1-periodic orbit of X_{H^c} in $\mathring{V}$ is a constant loop at a critical point of H.

(4) Every finite-energy solution $v : \mathbb{R} \times S^1 \to \mathring{V}$ of (10.9) is of the form $v(s,t) = \gamma(s)$ for some negative gradient flow trajectory $\gamma : \mathbb{R} \to V$.

PROOF. The following proof is based on arguments in [SZ92], see in particular Theorem 7.3.

For the first statement, let $\gamma(t) = (x,t)$ for $x \in \mathrm{Crit}(H)$ and recall from Lecture 3 the formula for the asymptotic operator of a 1-periodic orbit,

$$\mathbf{A}_\gamma : \Gamma(\gamma^*\Xi_0) \to \Gamma(\gamma^*\Xi_0) : \eta \mapsto -J\left(\nabla_t \eta - \nabla_\eta R_0^c\right),$$

where ∇ is any symmetric connection on $V \times S^1$. Identifying $\Gamma(\gamma^*\Xi_0)$ in the natural way with $C^\infty(S^1, T_xV)$, using the trivial connection and writing $R_0^c(z,\theta) = \partial_\theta + X_{H^c}(z) = \partial_\theta + cJ(z)\nabla H(z)$, $\mathbf{A}_\gamma$ becomes the operator

$$\mathbf{A}_\gamma = -J\partial_t - c\nabla^2 H(x)$$

on $C^\infty(S^1, T_xV)$, where $\nabla^2 H(x) : T_xV \to T_xV$ denotes the Hessian of H at x. Choosing a unitary basis for T_xV identifies this with $-J_0\partial_t - cS$ for some symmetric $2n$-by-$2n$ matrix S and the standard complex structure $J_0 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$, so $\ker \mathbf{A}_\gamma$ corresponds to the space of 1-periodic solutions to $\dot\eta = B\eta$ for $B := cJ_0S$. The Morse condition implies that S is nonsingular, thus so is B, but it is also small since c is small. We claim that if c is small enough to satisfy

$$\|B\| < \frac{1}{2\pi},$$

then all 1-periodic solutions η to $\dot{\eta} = B\eta$ are constant. To see this, let us complexify and regard η as a loop in $\mathbb{C}^{2n}$ and B as a complex-linear map on $\mathbb{C}^{2n}$. We can then write the smooth loops $\dot{\eta}, \ddot{\eta} : S^1 \to \mathbb{C}^{2n}$ as Fourier series

$$\dot{\eta}(t) = \sum_{k\in\mathbb{Z}} e^{2\pi ikt} a_k, \quad \ddot{\eta}(t) = \sum_{k\in\mathbb{Z}} e^{2\pi ikt} \cdot 2\pi ik a_k, \qquad \text{where} \quad a_k = \int_{S^1} \dot{\eta}(t) e^{-2\pi ikt} \in \mathbb{C}^{2n}.$$

Since $a_0 = \int_{S^1} \dot{\eta}(t)\, dt = 0$, Parseval's identity then implies

$$\|\dot{\eta}\|_{L^2}^2 = \sum_{k\in\mathbb{Z}} |a_k|^2 = \frac{1}{(2\pi)^2} \sum_{k\neq 0} (2\pi)^2 |a_k|^2 \leqslant \frac{1}{(2\pi)^2} \sum_{k\in\mathbb{Z}} |2\pi ik a_k|^2 = \frac{1}{(2\pi)^2} \|\ddot{\eta}\|_{L^2}^2,$$

thus plugging in the derivative of the equation $\dot{\eta} = B\eta$ gives

$$\|\dot{\eta}\|_{L^2} \leqslant \frac{1}{2\pi} \|\ddot{\eta}\|_{L^2} = \frac{1}{2\pi} \|B\dot{\eta}\|_{L^2} \leqslant \frac{\|B\|}{2\pi} \|\dot{\eta}\|_{L^2}.$$

If $\|B\| < 2\pi$, then this inequality gives a contradiction unless $\dot{\eta} \equiv 0$, proving the claim.[5] Since S is nonsingular, it follows that $\mathbf{A}_\gamma$ has only the trivial eigenfunction, hence γ is nondegenerate.

To calculate $\mu_{\mathrm{CZ}}^\tau(\gamma)$, suppose $S_\pm$ denotes a pair of nonsingular symmetric matrices defining asymptotic operators $\mathbf{A}_\pm = -J_0\partial_t - cS_\pm$, and choose a path $\{S_s\}_{s\in[-1,1]}$ of symmetric matrices connecting $S_{\pm 1} := S_\pm$. For $c > 0$ sufficiently small, the claim in the previous paragraph identifies the kernel of $\mathbf{A}_s := -J_0\partial_t - cS_s$ for each $s \in [-1,1]$ with $\ker S_s$ as a space of constant solutions. Similarly, $\lambda \in \mathbb{R}$ is an eigenvalue of $\mathbf{A}_\gamma$ if and only if the kernel of $-J_0\partial_t - (cS + \lambda)$ is nontrivial, which for $|\lambda|$ sufficiently small holds if and only if $-\lambda$ is an eigenvalue of cS. This implies that eigenvalues of the family $\{\mathbf{A}_s\}_{s\in[-1,1]}$ change signs precisely when eigenvalues of the family $\{S_s\}_{s\in[-1,1]}$ change signs in the opposite direction, giving a relation between spectral flows

$$\mu^{\mathrm{spec}}(\mathbf{A}_-, \mathbf{A}_+) = -\mu^{\mathrm{spec}}(S_-, S_+)$$

as long as $c > 0$ is sufficiently small. Denoting the maximal negative-definite subspace of $S_\pm$ by $E^-(S_\pm)$, this relation implies

$$\dim E^-(S_+) - \dim E^-(S_-) = \mu_{\mathrm{CZ}}(\mathbf{A}_-) - \mu_{\mathrm{CZ}}(\mathbf{A}_+).$$

Now suppose S_+ is a coordinate expression for the Hessian $\nabla^2 H(x)$, hence $\dim E^-(S_+) = \operatorname{ind}(x)$ and $\mu_{\mathrm{CZ}}(\mathbf{A}_+) = \mu_{\mathrm{CZ}}^\tau(\gamma)$. Choosing $S_- = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}$ then gives $\dim E^-(S_-) = n$ and $\mu_{\mathrm{CZ}}(\mathbf{A}_-) = 0$ by definition, so $\mu_{\mathrm{CZ}}^\tau(\gamma) = n - \operatorname{ind}(x)$ follows.

The second statement follows in a similar manner by writing down and comparing the linearized operators for the Floer equation and the negative gradient flow equation. Let's leave this as an exercise.

For the third statement, suppose we have a sequence $c_k \to 0$ and a sequence of loops $x_k : S^1 \to \mathring{V}$ satisfying $\dot{x}_k = X_{H^{c_k}}(x_k) = c_k X_H(x_k)$. Pick a number $c > 0$ small

[5] The claim is a linear version of a more general dynamical phenomenon: for an autonomous dynamical system defined via a C^1-small vector field on a compact manifold, all 1-periodic solutions are constant. This fact can be cited as naive motivation for the Arnol′d conjecture on symplectic fixed points; see for instance [**HZ94**, Chapter 5, Prop. 17 and P. 200].

enough for part (1) of the theorem to hold, choose a sequence of integers $N_k \in \mathbb{N}$ such that

$$N_k c_k \to c,$$

and consider the loops $y_k : S^1 \to \mathring{V} : t \mapsto x_k(N_k t)$. These satisfy

$$\dot{y}_k = N_k c_k X_H(y_k),$$

and since X_H is C^∞-bounded on V and $N_k c_k$ is also bounded, the Arzelà-Ascoli theorem provides a subsequence with

$$y_k \to y_\infty \quad \text{in} \quad C^\infty(S^1, V),$$

where $y_\infty : S^1 \to V$ satisfies $\dot{y}_\infty = X_{H^c}(y_\infty)$ for $H^c := cH : V \to \mathbb{R}$. But y_∞ is also constant: indeed, since $y_k(t + 1/N_k) = y_k(t)$ and $N_k \to \infty$, we can find for any $t \in S^1$ a sequence $q_k \in \mathbb{Z}$ satisfying $q_k/N_k \to t$, so

$$y_\infty(t) = \lim_{k\to\infty} y_k(q_k/N_k) = \lim_{k\to\infty} y_k(0) = y_\infty(0). \tag{10.10}$$

Since the constant orbit y_∞ is nondegenerate by part (1) of the theorem, there can only be one sequence of solutions to $\dot{y}_k = X_{H^{N_k c_k}}(y_k)$ converging to y_∞, and we conclude that y_k is also constant for all k sufficiently large.

We will now use a similar trick to prove the fourth statement in the theorem. We shall work under the additional assumption that

$$|\operatorname{ind}(x) - \operatorname{ind}(y)| \leqslant 1 \quad \text{for all pairs} \quad x, y \in \operatorname{Crit}(H), \tag{10.11}$$

which suffices for the application in §10.3.3 below.[6]

Suppose to the contrary that there exists a sequence of positive numbers $c_k \to 0$ with finite-energy solutions $v_k : \mathbb{R} \times S^1 \to \mathring{V}$ of the equation $\partial_s v_k + J(v_k)(\partial_t v_k - X_{H^{c_k}}(v_k)) = 0$, where each $v_k(s,t)$ is not t-independent. By part (3) of the theorem, we can restrict to a subsequence and assume each v_k for large k is asymptotic to a fixed pair of critical points $x_\pm = \lim_{s\to\pm\infty} v_k(s,\cdot) \in \operatorname{Crit}(H)$, and $x_+ \neq x_-$ since v_k would otherwise by constant and therefore t-independent. Choose a sequence $N_k \in \mathbb{N}$ with

$$N_k \to \infty \quad \text{and} \quad N_k c_k \to c,$$

where $c > 0$ is chosen sufficiently small for the first three statements in the theorem to hold. Define $w_k : \mathbb{R} \times S^1 \to V$ by

$$w_k(s,t) = v_k(N_k s, N_k t).$$

Then w_k satisfies another time-independent Floer equation,

$$\partial_s w_k + J(w_k)\left(\partial_t w_k - X_{H^{N_k c_k}}(w_k)\right) = 0, \tag{10.12}$$

where the Hamiltonian functions $H^{N_k c_k}$ converge to H^c. The standard compactness theorem for Floer trajectories should now imply that a subsequence of w_k converges to a broken Floer trajectory whose levels will be t-independent. Since the setting may seem a bit nonstandard, here are some details.

The sequence w_k is uniformly C^0-bounded since V is compact. We claim that it is also C^1-bounded. If not, then there is a sequence $z_k = (s_k, t_k) \in \mathbb{R} \times S^1$ with

[6] Lifting this assumption requires gluing, whereas we shall only need the usual implicit function theorem for Fredholm regular solutions of the Floer equation.

$|dw_k(z_k)| =: R_k \to \infty$, and we can use the usual rescaling trick from Lecture 9 to define a sequence

$$f_k : \mathbb{D}_{\epsilon_k R_k} \to V : z \mapsto w_k(z_k + z/R_k)$$

for a suitable sequence $\epsilon_k \to 0$ with $\epsilon_k R_k \to \infty$ and $|dw_k(z)| \leqslant 2R_k$ for all $z \in \mathbb{D}_{\epsilon_k}(z_k)$. The latter implies that f_k satisfies a local C^1-bound independent of k, and since

$$\partial_s f_k + J(f_k)\partial_t f_k = \frac{1}{R_k} J(f_k) X_{H^{N_k c_k}}(f_k) \to 0 \quad \text{as} \quad k \to \infty,$$

elliptic regularity (Theorem 2.24) and the Arzelà-Ascoli theorem provide a subsequence for which f_k converges in $C^\infty_{\mathrm{loc}}(\mathbb{C}, V)$ to a J-holomorphic plane $f_\infty : \mathbb{C} \to V$, which is nonconstant since

$$|df_\infty(0)| = \lim_{k\to\infty} |df_k(0)| = 1.$$

Since v_k and therefore w_k are all asymptotic to fixed constant orbits $x_\pm$, we have a uniform bound on the Floer energies of w_k,

$$E_{H^{N_k c_k}}(w_k) = \mathcal{A}_{H^{N_k c_k}}(x_+) - \mathcal{A}_{H^{N_k c_k}}(x_-) = N_k c_k \left[H(x_-) - H(x_+)\right], \tag{10.13}$$

where the right hand side is bounded since $N_k c_k \to c$. Using change of variables and the fact that $d\beta(\partial_s f_k, J(f_k)\, \partial_s f_k) \geqslant 0$, this implies a uniform bound

$$\begin{aligned}\int_{\mathbb{D}_{\epsilon_k R_k}} d\beta(\partial_s f_k, J(f_k)\, \partial_s f_k)\, ds \wedge dt &= \int_{\mathbb{D}_{\epsilon_k}(z_k)} d\beta(\partial_s w_k, J(w_k)\, \partial_s w_k)\, ds \wedge dt \\ &\leqslant \int_{\mathbb{R}\times S^1} d\beta(\partial_s w_k, J(w_k)\, \partial_s w_k)\, ds \wedge dt = E_{H^{N_k c_k}}(w_k) \leqslant C,\end{aligned}$$

thus

$$\int_{\mathbb{C}} f_\infty^* d\beta = \int_{\mathbb{C}} d\beta(\partial_s f_\infty, \partial_t f_\infty)\, ds \wedge dt = \int_{\mathbb{C}} d\beta(\partial_s f_\infty, J(f_\infty)\, \partial_s f_\infty)\, ds \wedge dt < \infty.$$

The removable singularity theorem now extends f_∞ to a nonconstant J-holomorphic sphere $f_\infty : S^2 \to V$, but this violates Stokes' theorem since J is tamed by an exact symplectic form.

We've now shown that the sequence $w_k : \mathbb{R} \times S^1 \to V$ is uniformly C^1-bounded, and it has bounded energy due to (10.13). Pick any sequence $s_k \in \mathbb{R}$ and consider the sequence of translated Floer trajectories

$$\widetilde{w}_k(s,t) := w_k(s + s_k, t).$$

These are also uniformly C^1-bounded, so by elliptic regularity again, a subsequence converges in $C^\infty_{\mathrm{loc}}(\mathbb{R} \times S^1)$ to a map $w_\infty : \mathbb{R} \times S^1 \to V$ satisfying

$$\partial_s w_\infty + J(w_\infty)\,(\partial_t w_\infty - X_{H^c}(w_\infty)) = 0,$$

and it has finite energy $E_{H^c}(w_\infty) < \infty$ due to (10.13), implying that w_∞ is asymptotic to a pair of 1-periodic orbits of X_{H^c} as $s \to \pm\infty$. By the same argument used in (10.10) above, w_∞ is also t-independent. It follows that $w_\infty(s,t) = \gamma_\infty(s)$ for some nonconstant gradient flow trajectory $\gamma_\infty : \mathbb{R} \to \mathring{V}$. Depending on the choice of sequence s_k, this trajectory may or may not be constant, but we can always choose s_k to guarantee that γ_∞ is not constant: indeed, since each w_k is asymptotic to two separate critical points at $\pm\infty$, $s_k \in \mathbb{R}$ can be chosen such that $w_k(s_k, 0)$ stays a fixed distance away from every critical point of H, and then

$$w_\infty(0,0) = \lim_{k\to\infty} w_k(s_k, 0) \notin \mathrm{Crit}(H^c).$$

One can now adapt the argument of Proposition 10.19 to find various sequences $s_k \in \mathbb{R}$ that yield potentially separate limiting trajectories forming the levels of a broken trajectory, which is the limit of w_k in the Floer topology. But since all the levels are t-independent and the gradient flow of H^c is Morse-Smale, condition (10.11) implies that the most complicated (and therefore the only) limit possible involves a single level $w_\infty(s,t) = \gamma(s)$, which is a gradient flow trajectory between critical points whose Morse indices differ by 1. This trajectory is Fredholm regular and has index 1 due to part (2) of the theorem, thus by the implicit function theorem, the only solutions to (10.12) that can converge to w_∞ are the obvious reparametrizations of γ, i.e. they are also t-independent. This is a contradiction. □

REMARK 10.31 (sign conventions). You may notice with some horror that (10.8) differs by a sign from what is stated in [**SZ92**]. The mundane reason for this is that the Conley-Zehnder index defined in [**SZ92**] also differs from ours by a sign (see Remark 3.69). On a deeper level, one can relate this discrepancy to the fact that while Floer homology is traditionally defined in terms of a negative gradient flow for the action functional, SFT is based on a *positive* gradient flow—this is also why the action functional in (10.6) differs by a sign from what we saw in Lecture 1, and why the Floer homological formula for asymptotic operators (Exercise 3.3) lacks the initial minus sign that appears in its SFT analogue (see Remark 3.7).

Returning now to the family $\mathcal{H}_\epsilon$, choose $c > 0$ sufficiently small for Theorem 10.30 to hold and define a modified family of stable Hamiltonian structures on $V \times S^1$ by

$$\mathcal{H}^c_\epsilon = (\omega^c, \lambda^c_\epsilon),$$

where

$$\omega^c := d\beta + d\theta \wedge dH^c \quad \text{and} \quad \lambda^c_\epsilon := d\theta + \epsilon(\beta - H^c\, d\theta).$$

Denote the induced hyperplane distributions and Reeb vector fields by Ξ^c_ϵ and R^c_ϵ respectively. We have only changed the Hamiltonian H by rescaling, so all previous statements about $\mathcal{H}_\epsilon$ also apply to $\mathcal{H}^c_\epsilon$, in particular λ^c_ϵ is contact and $\mathcal{J}(\mathcal{H}^c_\epsilon) = \mathcal{J}(\lambda^c_\epsilon)$ for all $\epsilon > 0$ sufficiently small, though the upper bound for the allowed range of ϵ may now depend on c. Once $c > 0$ is fixed by the requirements of Theorem 10.30, we are still free to take $\epsilon > 0$ as small as we like.

THEOREM 10.32. *Assume the same hypotheses as in Theorem 10.30, including (10.11), and denote the unique extension of J to an $\mathbb{R}$-invariant almost complex structure in $\mathcal{J}(\mathcal{H}^c_0)$ by J_0. Given $c > 0$ sufficiently small and any smooth family of compatible $\mathbb{R}$-invariant almost complex structures $J_\epsilon \in \mathcal{J}(\mathcal{H}^c_\epsilon)$ matching J_0 at $\epsilon = 0$, there exists $\epsilon_0 > 0$ such that every critical point $x \in \mathrm{Crit}(H)$ gives rise to a smooth family of nondegenerate closed R^c_ϵ-orbits*

$$x^\epsilon : S^1 \to V \times S^1 \qquad \epsilon \in [0, \epsilon_0]$$

with $x^0(t) = (x, t)$, and every gradient flow trajectory $\gamma : \mathbb{R} \to V$ for H gives rise to a smooth family of Fredholm regular J_ϵ-holomorphic cylinders

$$u^\epsilon_\gamma : \mathbb{R} \times S^1 \to \mathbb{R} \times (V \times S^1) \qquad \epsilon \in [0, \epsilon_0]$$

with $u^0_\gamma(s, t) = (s, \gamma(cs), t)$. Moreover, for all $\epsilon \in [0, \epsilon_0]$, every closed R^c_ϵ-orbit homotopic to $t \mapsto (\mathrm{const}, t)$ belongs to one of the families x^ϵ up to parametrization, and every J_ϵ-holomorphic cylinder with a positive and a negative end asymptotic to orbits of this type belongs to one of the families u^ϵ_γ, up to biholomorphic parametrization.

PROOF. The first part is immediate from the implicit function theorem since the orbits $x^0(t) = (x, t)$ are nondegenerate and the curves $u^0_\gamma(s, t) = (s, \gamma(cs), t)$ are Fredholm regular by Theorem 10.30. For the uniqueness statement, observe that if $\epsilon_k \to 0$ and γ_k is a sequence of $R^c_{\epsilon_k}$-orbits in the relevant homotopy class, then their periods are uniformly bounded, so Arzelà-Ascoli gives a subsequence convergent to a closed R^c_0-orbit, which is a nondegenerate orbit of the form $x^0(t) = (x, t)$ for $x \in \mathrm{Crit}(H)$ by Theorem 10.30, thus sequences converging to this orbit are unique by the implicit function theorem. A similar argument proves uniqueness of J_ϵ-holomorphic cylinders: if $\epsilon_k \to 0$ and u_k is a J_{ϵ_k}-holomorphic sequence, then first by the uniqueness of the orbits, we can extract a subsequence for which all u_k are asymptotic at both ends to orbits in fixed families $x^{\epsilon_k}_\pm$ converging to $x^0_\pm(t) = (x_\pm, t)$ as $k \to \infty$. Since ω is exact, Stokes' theorem then gives a uniform bound on the energies $E_{\epsilon_k}(u_k)$. Since all R^c_0-orbits in the relevant homotopy class are nondegenerate and none are contractible, one can now prove as in Proposition 10.19 that u_k has a subsequence convergent to a finite-energy stable J_0-holomorphic building $\mathbf{u}_\infty$ consisting only of cylinders. Its levels are asymptotic to orbits of the form $x(t) = (x, t)$ for $x \in \mathrm{Crit}(H)$, thus they can be parametrized as $(s, t) \mapsto (s, v(s, t), t)$ for $v : \mathbb{R} \times S^1 \to V$ satisfying the H^c-Floer equation, hence $v(s, t) = \gamma(cs)$ by Theorem 10.30. Now since ∇H is Morse-Smale and indices of critical points can only differ by at most 1, the building $\mathbf{u}_\infty$ can have at most one nontrivial level $u_\infty(s, t) = (s, \gamma(cs), t)$, implying $u_k \to u_\infty$. Since u_∞ is Fredholm regular, the implicit function theorem does the rest. □

10.3.3. Admissible data for $(\mathbb{T}^3, \xi_k)$. We now complete the computation of the cylindrical contact homology $HC^h_*(\mathbb{T}^3, \xi_k)$. We can assume via Lemma 10.26 that h is the homotopy class of the orbits in the special set of tori

$$T_m = \{m/k\} \times \mathbb{T}^2 \subset \mathbb{T}^3, \qquad m = 0, \ldots, k-1.$$

Let's focus for now on the case $k = 1$, as the general case will simply be a k-fold cover of this. Thanks to the Morse-Bott discussion in §10.3.1, we know what we're looking for: we want an h-admissible contact form α for $(\mathbb{T}^3, \xi_1)$ such that $\mathcal{P}_h(\alpha)$ contains exactly two orbits, both in $T_0 \subset \mathbb{T}^3$, along with an h-regular $J \in \mathcal{J}(\alpha)$ such that the differential on $CC^h_*(\mathbb{T}^3, \alpha)$ counts exactly two J-holomorphic cylinders that connect the two orbits in T_0. Let $\mathbb{A}$ denote the annulus

$$\mathbb{A} = [-1, 1] \times S^1 = [-1, 1] \times (\mathbb{R}/\mathbb{Z})$$

with coordinates (ρ, ϕ). This will play the role of the Liouville manifold $(V, d\beta)$ from the previous section, and we set

$$\beta := \rho \, d\phi.$$

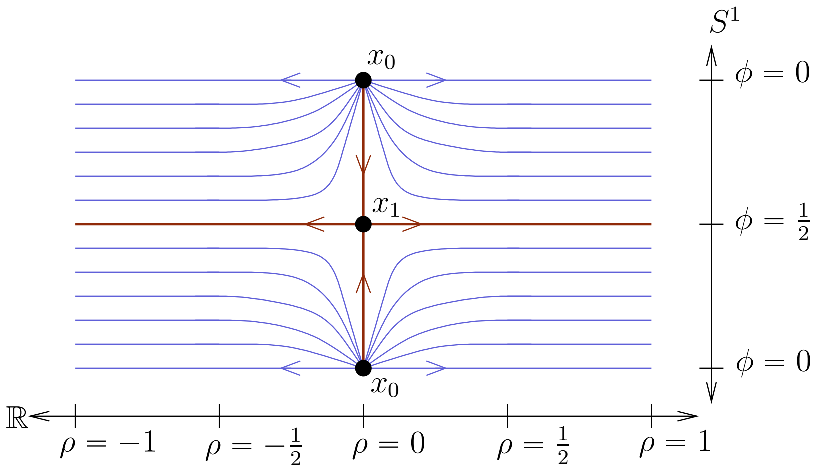

FIGURE 10.4. The gradient flow of the Hamiltonian $H : \mathbb{A} \to \mathbb{R}$

For the Hamiltonian $H : \mathbb{A} \to \mathbb{R}$, choose a Morse function with the following properties (see Figure 10.4):

(1) H has a minimum at $x_0 = (0, 0)$, an index 1 critical point at $x_1 = (0, 1/2)$, and no other critical points;
(2) $H(\rho, \phi) = |\rho|$ for $1/2 \leqslant |\rho| \leqslant 1$;
(3) The gradient flow of H with respect to the standard Euclidean metric on $[-1, 1] \times S^1$ is Morse-Smale.

Fix a number $c > 0$ sufficiently small so that Theorem 10.30 applies for Floer trajectories of $H^c := cH$ in $\mathbb{A}$, and since it will turn out to be useful in Lemma 10.33 below, assume without loss of generality

$$c \in \mathbb{Q}.$$

Then following the prescription described above, we consider the family of stable Hamiltonian structures $\mathcal{H}_\epsilon^c = (\omega^c, \lambda_\epsilon^c)$ on $\mathbb{A} \times S^1$ for $\epsilon \geqslant 0$ small, where

$$\lambda_\epsilon^c = (1 - \epsilon cH)\, d\theta + \epsilon\rho\, d\phi, \qquad \omega^c = d\rho \wedge d\phi + c\, d\theta \wedge dH,$$

with induced Reeb vector fields R_ϵ^c and hyperplane distributions $\Xi_\epsilon^c := \ker \lambda_\epsilon^c$. Choose $J_\epsilon \in \mathcal{J}(\mathcal{H}_\epsilon^c)$ to be any smooth family such that $J_0|_{\Xi_0^c}$ matches the standard complex structure on $\mathbb{A}$ defined by $J_0 \partial_\rho = \partial_\phi$. Then for all $\epsilon > 0$ sufficiently small, Theorems 10.30 and 10.32 give a complete classification of all closed R_ϵ^c-orbits in $\mathbb{A} \times S^1$ homotopic to $t \mapsto (0, 0, t)$, as well as a classification of all J_ϵ-holomorphic cylinders asymptotic to them. Up to parametrization, there are exactly two such orbits,

$$\gamma_i^\epsilon : S^1 \to \mathbb{A} \times S^1, \qquad i = 0, 1,$$

which correspond to the Morse critical points x_0 and x_1 and thus by (10.8) have Conley-Zehnder indices

$$\mu_{\mathrm{CZ}}^\tau(\gamma_i^\epsilon) = 1 - \operatorname{ind}(x_i) = 1 - i \in \{0, 1\}$$

relative to the constant trivialization τ. There are also exactly two J_ϵ-holomorphic cylinders

$$u^\epsilon_\pm : \mathbb{R} \times S^1 \to \mathbb{R} \times (\mathbb{A} \times S^1),$$

corresponding to the two negative gradient flow lines that descend from x_1 to x_0, thus the $u^\epsilon_\pm$ are index 1 curves with a negative end approaching γ^ϵ_1 and a positive end approaching γ^ϵ_0. If we can suitably embed this model into $(\mathbb{T}^3, \xi_1)$ and show that all the orbits and curves needing to be counted are contained in the model, then we will have a complete description of $HC^h_*(\mathbb{T}^3, \xi_1)$, with two generators $\langle \gamma^\epsilon_0 \rangle$ and $\langle \gamma^\epsilon_1 \rangle$, of even and odd degree respectively, satisfying

$$\partial \langle \gamma^\epsilon_0 \rangle = 2\langle \gamma^\epsilon_1 \rangle = 0 \quad \text{and} \quad \partial \langle \gamma^\epsilon_1 \rangle = 0$$

since the former counts two curves and the latter counts none.

LEMMA 10.33. *For any $\epsilon > 0$ sufficiently small, there exists a contact embedding of*

$$(\mathbb{A} \times S^1, \ker \lambda^c_\epsilon) \hookrightarrow (\mathbb{T}^3, \xi_1)$$

identifying the homotopy class of the loops $t \mapsto (0, 0, t)$ in $\mathbb{A} \times S^1$ with h. Moreover, the contact form λ^c_ϵ and almost complex structure $J_\epsilon \in \mathcal{J}(\mathcal{H}^c_\epsilon)$ can then be extended to an h-admissible contact form α on $(\mathbb{T}^3, \xi_1)$ and an h-regular almost complex structure $J \in \mathcal{J}(\alpha)$ such that γ^ϵ_0 and γ^ϵ_1 are the only orbits in $\mathcal{P}_h(\alpha)$, and all J-holomorphic cylinders with a positive and a negative end asymptotic to either of these orbits are contained in the interior of $\mathbb{A} \times S^1$.

PROOF. We've chosen β and H so that in the region $1/2 \leqslant |\rho| \leqslant 1$,

$$\alpha := \lambda^c_\epsilon = (1 - \epsilon c|\rho|)\, d\theta + \epsilon\rho\, d\phi =: f(\rho)\, d\theta + g(\rho)\, d\phi,$$

so the Reeb vector field on this region has the form $\frac{1}{D(\rho)}(g'(\rho)\, \partial_\theta - f'(\rho)\, \partial_\phi)$ with $D := fg' - f'g$. Notice that

$$\frac{f'(\rho)}{g'(\rho)} = \mp\frac{\epsilon c}{\epsilon} = \mp c,$$

and we assumed $c \in \mathbb{Q}$, so the Reeb orbits in this region are all periodic. Next, pick a large number $N \gg 1$ and extend α to a contact form on $[-N, N] \times S^1 \times S^1$ via the same formula. Now extend the path $(f, g) : [1/2, N] \to \mathbb{R}^2$ to $\mathbb{R}$ such that it has period $2N + 2$, satisfies $(f(-\rho), g(-\rho)) = (f(\rho), -g(\rho))$, and winds once around the origin over the interval $[-N - 1, N + 1]$, with positive angular velocity (see Figure 10.5). This gives rise to a contact form α on

$$\mathbb{T}^3_N := \left(\mathbb{R} \big/ (2N + 2)\mathbb{Z}\right) \times S^1 \times S^1$$

which matches λ^c_ϵ in the region $|\rho| \leqslant 1$ and takes the form $f(\rho)\, d\theta + g(\rho)\, d\phi$ outside of that region. We claim in fact that α is also globally homotopic to $f(\rho)\, d\theta + g(\rho)\, d\phi$ through a family of contact forms. To see this, one need only homotop H in the region $|\rho| \leqslant 1/2$ to a Morse-Bott function that depends only on the ρ-coordinate;

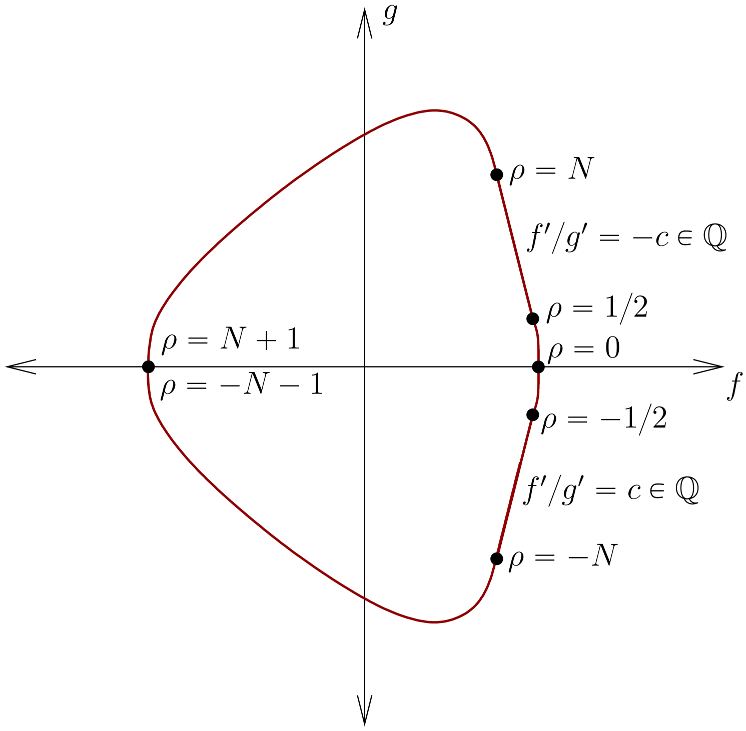

FIGURE 10.5. The periodic path $(f, g) : [-N-1, N+1] \to \mathbb{R}^2$ in the proof of Lemma 10.33.

the contact condition holds for all Hamiltonians in this homotopy as long as $\epsilon > 0$ is sufficiently small. With this understood, the obvious diffeomorphism

$$\mathbb{T}^3_N \to \mathbb{T}^3 : (\rho, \phi, \theta) \mapsto \left(\frac{\rho}{2N+2}, \phi, \theta \right)$$

pushes $\ker \alpha$ forward to a contact structure isotopic to one of the form $F(\rho)\, d\theta + G(\rho)\, d\phi$ for a loop $(F, G) : S^1 \to \mathbb{R}^2$ winding once around the origin, so taking a homotopy of this loop to $(\cos(2\pi\rho), \sin(2\pi\rho))$ and applying Gray's stability theorem produces a contactomorphism

$$(\mathbb{T}^3_N, \ker \alpha) \to (\mathbb{T}^3, \xi_1)$$

that is isotopic to the above diffeomorphism.

The construction clearly guarantees that no closed Reeb orbit of α outside $\mathbb{A} \times S^1$ is homotopic to the preferred class h, and there are also no contractible orbits, so α is an h-admissible contact form on $\mathbb{T}^3_N$. Choose any extension of J_ϵ to some $J \in \mathcal{J}(\alpha)$ on $\mathbb{T}^3_N$. We claim now that if N is chosen sufficiently large, then no J-holomorphic cylinder in $\mathbb{R} \times \mathbb{T}^3_N$ with one positive end at either of the orbits γ_i^ϵ can ever venture outside the region $\mathbb{R} \times (-1/2, 1/2) \times \mathbb{T}^2$. Suppose in particular that u is such a curve and its image intersects $\mathbb{R} \times \{1/2\} \times \mathbb{T}^2$. Since the entire region $[1/2, N] \times \mathbb{T}^2$ is foliated by closed Reeb orbits, we can define Υ to be the set of Reeb orbits γ in that region for which the image of u intersects $\mathbb{R} \times \gamma$. This is a closed subset of the connected topological space of all Reeb orbits in $[1/2, N] \times \mathbb{T}^2$:

indeed, if $\gamma_k \in \Upsilon$ is a sequence converging to some orbit γ_∞, then $u(z_k) \in \mathbb{R} \times \gamma_k$ for some sequence $z_k \in \mathbb{R} \times S^1$, which must be contained in a compact subset since the asymptotic orbits of u lie outside of $[1/2, N] \times \mathbb{T}^2$, hence z_k has a convergent subsequence $z_k \to z_\infty \in \mathbb{R} \times S^1$ with $u(z_\infty) \in \mathbb{R} \times \gamma_\infty$, proving $\gamma_\infty \in \Upsilon$. We claim that Υ is also an open subset of the space of orbits in $[1/2, N] \times \mathbb{T}^2$. This follows from positivity of intersections, as every $\mathbb{R} \times \gamma$ is also a J-holomorphic curve: if $u(z) \in \mathbb{R} \times \gamma$, then for every other closed orbit γ' close enough to γ, there is a point $z' \in \mathbb{R} \times S^1$ near z with $u(z') \in \mathbb{R} \times \gamma'$. This proves that, in fact, u passes through $\mathbb{R} \times \gamma$ for *every* orbit γ in the region $[1/2, N] \times \mathbb{T}^2$. We will now use this to show that if N is sufficiently large, the contact area of u will be larger than is allowed by Stokes' theorem.

Let us write

$$u(s,t) = (r(s,t), \rho(s,t), \phi(s,t), \theta(s,t)) \in \mathbb{R} \times \big(\mathbb{R}/(2N+2)\mathbb{Z}\big) \times S^1 \times S^1$$

and choose two points $\rho_1 \in [1/2, 1]$ and $\rho_2 \in [N-1, N]$ which are both regular values of the function $\rho : \mathbb{R} \times S^1 \to \mathbb{R}/(2N+2)\mathbb{Z}$. The intersections of u with the orbits in $[1/2, N] \times \mathbb{T}^2$ imply that the function $\rho(s,t)$ attains every value in $[1/2, N]$, and since the asymptotic limits of u lie outside this region,

$$\mathcal{U} := \rho^{-1}([\rho_1, \rho_2]) \subset \mathbb{R} \times S^1$$

is then a nonempty and compact smooth submanifold with boundary

$$\partial \mathcal{U} = -C_1 \coprod C_2,$$

where $C_i := \rho^{-1}(\rho_i)$ for $i = 1, 2$. Restricting u to the multicurves C_i then gives a pair of smooth maps

$$w_i : C_i \to \mathbb{T}^2 : (s,t) \mapsto (\phi(s,t), \theta(s,t)), \qquad i = 1, 2,$$

which are homologous to each other. Denote the generators of $H_1(\mathbb{T}^2)$ corresponding to the ϕ- and θ-coordinates by ℓ_ϕ and ℓ_θ respectively, and suppose $[w_i] = m\ell_\phi + n\ell_\theta$ for $m, n \in \mathbb{Z}$. The key observation now is that the restriction of α to each of the tori $\{\rho_i\} \times \mathbb{T}^2$ is a closed 1-form, thus for each $i = 1, 2$, $\int_{C_i} u^*\alpha$ depends only on the homology class $m\ell_\phi + n\ell_\theta \in H_1(\mathbb{T}^2)$ and not any further on the maps w_i. In particular,

$$\int_{C_i} u^*\alpha = f(\rho_i)n + g(\rho_i)m$$

for $i = 1, 2$. We now compute,

$$\begin{aligned}\int_{\mathcal{U}} u^* d\alpha &= \int_{C_2} u^*\alpha - \int_{C_1} u^*\alpha = n[f(\rho_2) - f(\rho_1)] + m[g(\rho_2) - g(\rho_1)] \\ &= n[(1 - \epsilon c\rho_2) - (1 - \epsilon c\rho_1)] + m[\epsilon\rho_2 - \epsilon\rho_1] \\ &= \epsilon(\rho_2 - \rho_1)(m - nc)\end{aligned}$$

This integral has to be positive since $u^*d\alpha \geqslant 0$ and u is not a trivial cylinder, thus $m - nc > 0$. Moreover, c was assumed rational, so if $c = p/q$ for some $p, q \in \mathbb{N}$, we have

$$m - nc = \frac{1}{q}(mq - np) \geqslant \frac{1}{q},$$

implying
$$\int_{\mathbb{R}\times S^1} u^* d\alpha \geqslant \int_{\mathcal{U}} u^* d\alpha \geqslant \frac{\epsilon}{q}(\rho_2 - \rho_1) \geqslant \frac{\epsilon(N-2)}{q}.$$
Having chosen c (which determines q) and ϵ in advance, we are free to make N as large as we like. But by (10.2), $\int_{\mathbb{R}\times S^1} u^* d\alpha$ cannot be any larger than the period of its positive asymptotic orbit, which does not depend on N. So this gives a contradiction, proving that u cannot touch the region $\{\rho \geqslant 1/2\}$. The mirror image of this argument shows that u also cannot touch the region $\{\rho \leqslant -1/2\}$. □

With Lemma 10.33 in hand, the calculation of $HC_*^h(\mathbb{T}^3_N, \alpha, J)$ for sufficiently large N is straightforward: there is one odd generator and one even generator, with a trivial differential, giving
$$HC_*^h(\mathbb{T}^3, \xi_1) \cong \begin{cases} \mathbb{Z}_2 & * = \text{odd}, \\ \mathbb{Z}_2 & * = \text{even}. \end{cases}$$
This calculation can now be extended to $(\mathbb{T}^3, \xi_k)$ by a cheap trick: using the contactomorphism $(\mathbb{T}^3_N, \ker\alpha) \to (\mathbb{T}^3, \xi_1)$, let us identify $\mathbb{T}^3_N$ with $\mathbb{T}^3$ and write $\alpha = F\alpha_1$ for some function $F : \mathbb{T}^3 \to (0, \infty)$. Then the k-fold covering map
$$\Phi_k : \mathbb{T}^3 \to \mathbb{T}^3 : (\rho, \phi, \theta) \mapsto (k\rho, \phi, \theta)$$
maps the homotopy class h to itself and pulls back ξ_1 to ξ_k, so $\Phi_k^*\alpha$ is a contact form for ξ_k. It is also h-admissible: indeed, $\Phi_k^*\alpha$ admits no contractible orbits since they would project down to contractible orbits on $(\mathbb{T}^3, \alpha)$, and every orbit in $\mathcal{P}_h(\Phi_k^*\alpha)$ projects to one in $\mathcal{P}_h(\alpha)$, hence they are all nondegenerate. The almost complex structure $\Phi_k^* J \in \mathcal{J}(\Phi_k^*\alpha)$ then makes the map $\mathrm{Id} \times \Phi_k : (\mathbb{R} \times \mathbb{T}^3, \Phi_k^* J) \to (\mathbb{R} \times \mathbb{T}^3, J)$ holomorphic, so every $\Phi_k^* J$-holomorphic cylinder counted by $HC_*^h(\mathbb{T}^3, \Phi_k^*\alpha, \Phi_k^* J)$ projects to a J-holomorphic cylinder counted by $HC_*^h(\mathbb{T}^3, \alpha, J)$, and conversely, each orbit in $\mathcal{P}_h(\alpha)$ and each J-holomorphic cylinder has exactly k lifts to the cover. The generators of $CC_*^h(\mathbb{T}^3, \Phi_k^*\alpha)$ thus consist of $2k$ orbits, k odd and k even, with $2k$ connecting $\Phi_k^* J$-holomorphic cylinders that cancel each other in pairs, giving a trivial differential. In summary:
$$HC_*^h(\mathbb{T}^3, \xi_k) \cong \begin{cases} \mathbb{Z}_2^k & * = \text{odd}, \\ \mathbb{Z}_2^k & * = \text{even}. \end{cases}$$

REMARK 10.34. The method of this lecture for proving existence and uniqueness of the holomorphic curves in the chain complex for $HC_*^h(\mathbb{T}^3, \xi_k)$ is not the only approach possible. The Morse-Bott technique from [**Bou02**] was mentioned already in §10.3.1 and is used quite often in practice; a higher-dimensional analogue of the same computation using the Morse-Bott method may be found in [**MNW13**, Theorem 9.10(4)]. Alternatively, one can use the same nondegenerate data as in our computation but simplify the uniqueness proof by using intersection theory; we will take this approach to compute other SFT invariants of contact 3-manifolds in Lecture 16. For a higher-dimensional computation that combines intersection theory with the Floer-theoretic approach of §10.3.2, see [**Mora**, **Morb**].

LECTURE 11

Coherent orientations

Contents

11.1. Gluing maps and coherence

This lecture will be concerned with orienting the moduli spaces

$$\mathcal{M}(J) := \mathcal{M}_{g,m}(J, A, \gamma^+, \gamma^-)$$

of J-holomorphic curves in a completed symplectic cobordism $\widehat{W}$, in cases where they are smooth. We assume as usual that all Reeb orbits are nondegenerate so that the usual linearized Cauchy-Riemann operators are Fredholm.

For SFT and other Floer-type theories, it is not enough to know that each component of $\mathcal{M}(J)$ is orientable—relations like $\partial^2 = 0$ rely on having certain compatibility conditions between the orientations on different components. The point is that whenever a space of broken curves is meant to be interpreted as the boundary of some other compactified moduli space, we need to make sure that it carries the boundary orientation. This compatibility is what is known as *coherence*, and in order to define it properly, we need to return to the subject of gluing.

Our discussion of gluing in Lecture 10 was fairly simple because it was limited to somewhere injective holomorphic cylinders that could only break along simply covered Reeb orbits. Recall however that more general holomorphic buildings carry a certain amount of extra structure that was not relevant in that simple case. Even in a building $\mathbf{u}$ that has only two nontrivial levels u_- and u_+, the breaking punctures carry *decorations*: i.e. if $\{z^+, z^-\}$ is a breaking pair in $\mathbf{u}$, then the decoration defines an orientation-reversing orthogonal map

$$\delta_{z^+} \xrightarrow{\Phi} \delta_{z^-}$$

between the two "circles at infinity" $\delta_{z^\pm}$ associated to the punctures $z^\pm$ (see §9.3.3). This extra information is uniquely determined if the breaking orbit is simply covered, but at a multiply covered breaking orbit there is ambiguity, and the decoration cannot be deduced from knowledge of u_- and u_+ alone. We therefore need to consider moduli spaces of curves with a bit of extra structure.

For each Reeb orbit γ in M_+ or M_-, choose a point on its image

$$p_\gamma \in \operatorname{im} \gamma \subset M_\pm.$$

For a J-holomorphic curve $u : (\dot{\Sigma} = \Sigma \backslash (\Gamma^+ \cup \Gamma^-), j) \to (\widehat{W}, J)$ with a puncture $z \in \Gamma^\pm$ asymptotic to γ, an **asymptotic marker** is a choice of a ray $\ell \subset T_z\Sigma$ such that

$$\lim_{t \to 0^+} u(c(t)) = (\pm\infty, p_\gamma)$$

for any smooth path $c(t) \in \Sigma$ with $c(0) = z$ and $0 \neq \dot{c}(0) \in \ell$. If γ has covering multiplicity $m \in \mathbb{N}$, then there are exactly m choices of asymptotic markers at z, related to each other by the action on $T_z\Sigma$ by the mth roots of unity. We shall denote

$$\mathcal{M}^\$(J) := \mathcal{M}^\$_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) := \big\{(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u, \ell)\big\} \big/ \sim,$$

where $(\Sigma, j, \Gamma^+, \Gamma^-, \Theta, u)$ represents an element of $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$, ℓ denotes an assignment of asymptotic markers to every puncture $z \in \Gamma^\pm$, and

$$(\Sigma_0, j_0, \Gamma_0^+, \Gamma_0^-, \Theta_0, u_0, \ell_0) \sim (\Sigma_1, j_1, \Gamma_1^+, \Gamma_1^-, \Theta_1, u_1, \ell_1)$$

means the existence of a biholomorphic map $\psi : (\Sigma_0, j_0) \to (\Sigma_1, j_1)$ which defines an equivalence of $(\Sigma_0, j_0, \Gamma_0^+, \Gamma_0^-, \Theta_0, u_0)$ with $(\Sigma_1, j_1, \Gamma_1^+, \Gamma_1^-, \Theta_1, u_1)$ and satisfies $\psi_* \ell_0 = \ell_1$. There is a natural surjection

$$\mathcal{M}^\$(J) \to \mathcal{M}(J)$$

defined by forgetting the markers. We will say that an element $u \in \mathcal{M}^\$(J)$ is Fredholm regular whenever its image under the map to $\mathcal{M}(J)$ is regular. Let

$$\mathcal{M}^{\$,\mathrm{reg}}(J) = \mathcal{M}^{\$,\mathrm{reg}}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) \subset \mathcal{M}^\$(J)$$

denote the open subset consisting of Fredholm regular curves with asymptotic markers. Note that components of $\mathcal{M}(J)$ and $\mathcal{M}^\$(J)$ consisting of closed curves are identical spaces; components with punctures have the following simple relationship to each other.

PROPOSITION 11.1. *Each component of $\mathcal{M}^{\$,\mathrm{reg}}(J)$ consisting of curves with at least one puncture admits the structure of a smooth manifold, whose dimension on each connected component matches that of $\mathcal{M}^{\mathrm{reg}}(J)$. Moreover, the natural map*

$$\mathcal{M}^{\$,\mathrm{reg}}(J) \to \mathcal{M}^{\mathrm{reg}}(J)$$

is smooth, and the preimage of a curve $u \in \mathcal{M}^{\mathrm{reg}}(J)$ with asymptotic orbits $\{\gamma_z\}_{z\in\Gamma}$ of covering multiplicities $\{\kappa_z\}_{z\in\Gamma}$ contains exactly

$$\frac{\prod_{z\in\Gamma} \kappa_z}{|\operatorname{Aut}(u)|}$$

distinct elements.

PROOF. The smooth structure of $\mathcal{M}^{\$,\mathrm{reg}}(J)$ arises from a small modification of the same argument we used in Lecture 7 for $\mathcal{M}^{\mathrm{reg}}(J)$. Recall that in order to describe a neighborhood of a curve in $\mathcal{M}^{\mathrm{reg}}(J)$ represented by a map $u_0 : (\dot{\Sigma}, j_0) \to (\widehat{W}, J)$ whose domain has punctures $\Gamma \subset \Sigma$, marked points $\Theta \subset \dot{\Sigma} := \Sigma \backslash \Gamma$ and automorphism group $G := \mathrm{Aut}(\Sigma, j_0, \Gamma \cup \Theta)$, we chose a G-invariant Teichmüller slice $\mathcal{T} \subset \mathcal{J}(\Sigma)$ through j_0 and defined the smooth section $\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$ whose zero-set consists of the pairs (j, u) such that $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$ is J-holomorphic. Fredholm regularity implies that a neighborhood of (j_0, u_0) in $\bar{\partial}_J^{-1}(0)$ is a smooth finite-dimensional submanifold of $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$, and G acts smoothly and properly (but possibly not freely) on $\bar{\partial}_J^{-1}(0)$ by $\psi \cdot (j, u) := (\psi^* j, u \circ \psi)$ for $\psi \in G$, so that the natural map $\bar{\partial}_J^{-1}(0)/G \to \mathcal{M}(J)$ is a local homeomorphism near u_0. To include asymptotic markers in this picture, we observe that there is a natural finite covering map

$$\widetilde{\mathcal{M}} \to \bar{\partial}_J^{-1}(0),$$

where the elements of $\widetilde{\mathcal{M}}$ are tuples (j, u, ℓ) where $(j, u) \in \bar{\partial}_J^{-1}(0)$ and ℓ denotes an assignment of asymptotic markers to each puncture of $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$. We can therefore regard $\widetilde{\mathcal{M}}$ as a smooth finite-dimensional manifold by pulling back the smooth structure of $\bar{\partial}_J^{-1}(0)$, and G then also acts smoothly and properly on $\widetilde{\mathcal{M}}$ by

$$\psi \cdot (j, u, \ell) := (\psi^* j, u \circ \psi, (\psi^{-1})_* \ell).$$

We claim however that this action is also free. Indeed, if $\psi \cdot (j, u, \ell) = (j, u, \ell)$, then $\psi \in \mathrm{Aut}(u)$, and the finiteness of $\mathrm{Aut}(u)$ then implies that ψ is a biholomorphic map fixing Γ and satisfying $\psi^k \equiv \mathrm{Id}$ for some $k \in \mathbb{N}$. Choosing a G-invariant neighborhood $\mathcal{U} \subset \Sigma$ of some puncture $\zeta \in \Gamma$, the Riemann mapping theorem permits us to identify $\mathcal{U}$ biholomorphically with the unit disk in $\mathbb{C}$ so that ζ becomes the origin, and ψ on this neighborhood must then be a rotation, specifically a map of the form $z \mapsto e^{2\pi i m/k}$ for some $m \in \{1, \ldots, k\}$. The cases $m \neq k$ are impossible since ψ would then change the asymptotic marker at ζ, thus ψ can only be the identity map on $\mathcal{U}$. Unique continuation then implies $\psi \equiv \mathrm{Id}$, proving the claim. The quotient $\widetilde{\mathcal{M}}/G$ is therefore a smooth manifold, with the same dimension as the orbifold $\bar{\partial}_J^{-1}(0)/G$.

Finally, if $u : (\dot{\Sigma} = \Sigma \backslash \Gamma, j) \to (\widehat{W}, J)$ represents an element of $\mathcal{M}(J)$ with asymptotic orbits $\{\gamma_z\}_{z \in \Gamma}$, then the number of possible choices of asymptotic markers is precisely $\prod_{z \in \Gamma} \kappa_z$. However, not all of these produce inequivalent elements of $\mathcal{M}^{\$}(J)$: indeed, the previous paragraph shows that $\mathrm{Aut}(u)$ acts freely on the set of all choices of markers, so that the total number of inequivalent choices is as stated. $\square$

Suppose u_+ and u_- are two (possibly disconnected and/or nodal) holomorphic curves, with asymptotic markers, such that the number of negative punctures of u_+ equals the number of positive punctures of u_-, and the asymptotic orbit of u_+ at its ith negative puncture matches that of u_- at its ith positive puncture for every i. Then the pair (u_-, u_+) naturally determines a holomorphic building: indeed, the breaking punctures admit unique decorations determined by identifying the markers on u_+ with the markers at corresponding punctures of u_-.

Let us now consider a concrete example of a gluing scenario. Figure 11.1 shows the degeneration of a sequence of curves in $\mathcal{M}_{3,4}(J, A_k, (\gamma_4, \gamma_5), \boldsymbol{\gamma}^-)$ to a building $\mathbf{u} \in \overline{\mathcal{M}}_{3,4}(J, A + B + C, (\gamma_4, \gamma_5), \boldsymbol{\gamma}^-)$ with one main level and one upper level. The main level is a connected curve $u_A \in \mathcal{M}_{1,2}(J, A, (\gamma_1, \gamma_2, \gamma_3), \boldsymbol{\gamma}^-)$, and the upper level consists of two connected curves

$$u_B \in \mathcal{M}_{1,1}(J_+, B, \gamma_4, (\gamma_1, \gamma_2)), \qquad u_C \in \mathcal{M}_{0,1}(J_+, C, \gamma_5, \gamma_3).$$

One can endow each of these curves with asymptotic markers compatible with the decoration of $\mathbf{u}$; this is a non-unique choice, but e.g. if one chooses markers for u_A arbitrarily, then the markers at the negative punctures of u_B and u_C are uniquely determined. Now if all three curves are Fredholm regular, then a substantial generalization of the gluing procedure outlined in Lecture 10 provides open neighborhoods $\mathcal{U}_A^\$$ and $\mathcal{U}_{BC}^\$$,

$$\begin{aligned} u_A \in \mathcal{U}_A^\$ &\subset \mathcal{M}_{1,2}^\$(J, A, (\gamma_1, \gamma_2, \gamma_3), \boldsymbol{\gamma}^-), \\ [(u_B, u_C)] \in \mathcal{U}_{BC}^\$ &\subset \left(\mathcal{M}_{1,1}^\$(J_+, B, \gamma_4, (\gamma_1, \gamma_2)) \times \mathcal{M}_{0,1}^\$(J_+, C, \gamma_5, \gamma_3)\right) \Big/ \mathbb{R} \end{aligned}$$

which are smooth manifolds of dimensions

$$\begin{aligned} \dim \mathcal{U}_A^\$ &= \operatorname{vir-dim} \mathcal{M}_{1,2}(J, A, (\gamma_1, \gamma_2, \gamma_3), \boldsymbol{\gamma}^-), \\ \dim \mathcal{U}_{BC}^\$ &= \operatorname{vir-dim} \mathcal{M}_{1,1}(J_+, B, \gamma_4, (\gamma_1, \gamma_2)) + \operatorname{vir-dim} \mathcal{M}_{0,1}(J_+, C, \gamma_5, \gamma_3) - 1, \end{aligned}$$

along with a smooth embedding

$$\Psi : [R_0, \infty) \times \mathcal{U}_A^\$ \times \mathcal{U}_{BC}^\$ \hookrightarrow \mathcal{M}_{3,4}^\$(J, A + B + C, (\gamma_4, \gamma_5), \boldsymbol{\gamma}^-), \tag{11.1}$$

defined for $R_0 \gg 1$. This is an example of a **gluing map**: it has the property that for any $u \in \mathcal{U}_A^\$$ and $v \in \mathcal{U}_{BC}^\$$, $\Psi(R, u, v)$ converges in the SFT topology as $R \to \infty$ to the unique building (with asymptotic markers) having main level u and upper level v, and moreover, every sequence of smooth curves degenerating in this way is eventually in the image of Ψ.

In analogous ways one can define gluing maps for buildings with a main level and a lower level, or more than two levels, or multiple levels in a symplectization (always dividing symplectization levels by the $\mathbb{R}$-action). It's important to notice that in all such scenarios, the domain and target of the gluing map have the same dimension, e.g. the dimension of both sides of (11.1) is the sum of the virtual dimensions of the three moduli spaces concerned.

DEFINITION 11.2. A set of orientations for the connected components of $\mathcal{M}^\$(J)$ and $\mathcal{M}^\$(J_\pm)$ is called **coherent** if all gluing maps are orientation preserving.

Stated in this way, this definition is based on the pretense that we never have to worry about non-regular curves in any components of $\mathcal{M}^\$(J)$, and that is of course false—sometimes regularity cannot be achieved, in particular for multiply covered curves. As we'll see though in §11.4, the question of orientations can be reframed in a way that completely disjoins it from the question of regularity, thus we will later be able to state a more general version of the above definition that is independent of regularity (see Definition 11.17). The main result whose proof we will outline near the end of this lecture is then:

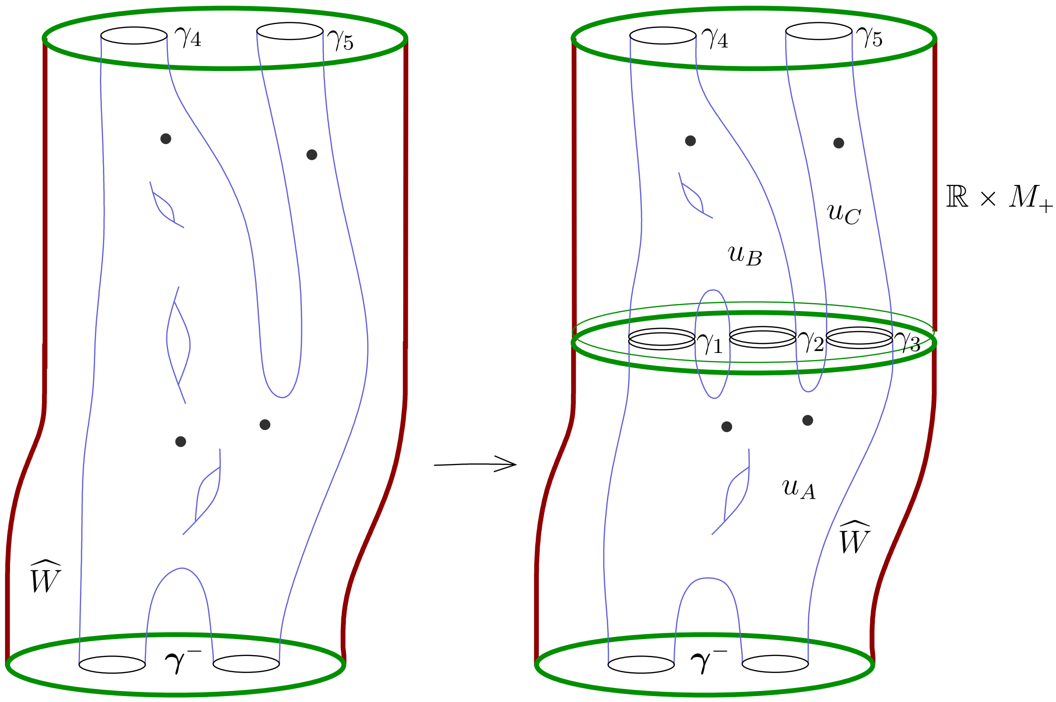

FIGURE 11.1. The degeneration scenario behind the gluing map (11.1)

THEOREM 11.3. *Coherent orientations exist.*

But there is also some bad news. The space $\mathcal{M}^\$(J)$ with asymptotic markers is not actually the space we want to orient. In fact, even the usual moduli space $\mathcal{M}(J)$ has a certain amount of extra information in it that we'd rather not keep track of when we don't have to, for instance the ordering of the punctures. Can we forget this information without forgetting the orientation of the moduli space? Not always:

PROPOSITION 11.4. *Suppose $\hat{\boldsymbol{\gamma}}^+ = (\gamma_1^+, \ldots, \gamma_{k_+}^+)$, and $\check{\boldsymbol{\gamma}}^+$ is a similar ordered list of Reeb orbits obtained from $\hat{\boldsymbol{\gamma}}^+$ by exchanging γ_j^+ with γ_k^+ for some $1 \leqslant j < k \leqslant k_+$. Then for any choice of coherent orientations, the natural map*

$$\mathcal{M}^\$_{g,m}(J, A, \hat{\boldsymbol{\gamma}}^+, \boldsymbol{\gamma}^-) \to \mathcal{M}^\$_{g,m}(J, A, \check{\boldsymbol{\gamma}}^+, \boldsymbol{\gamma}^-)$$

defined by permuting the corresponding punctures $z_j^+, z_k^+ \in \Gamma^+$ along with their asymptotic markers is orientation reversing if and only if the numbers

$$n - 3 + \mu_{CZ}(\gamma_i^+)$$

for $i = j, k$ are both odd. A similar statement holds for permutations of negative punctures.

This result is the reason for the super-commutative algebra that we will see in the next lecture. What about forgetting the markers? It turns out that we can sometimes do that as well, but again not always.

PROPOSITION 11.5. *Suppose* $\mathcal{M}^{\$}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) \to \mathcal{M}^{\$}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ *is the map defined by multiplying the asymptotic marker by* $e^{2\pi i/m}$ *at one of the punctures for which the asymptotic orbit is an* m*-fold cover* γ^m *of a simple orbit* γ*. For any choice of coherent orientations, this map reverses orientation if and only if* m *is even and* $\mu_{CZ}(\gamma^m) - \mu_{CZ}(\gamma)$ *is odd.*

Note that in both of the above propositions, only the odd/even parity of the Conley-Zehnder indices matters, so there is no need to choose trivializations. Proposition 11.5 motivates one of the more mysterious technical definitions in SFT.

DEFINITION 11.6. A closed nondegenerate Reeb orbit γ is called a **bad** orbit if it is an m-fold cover of some simple orbit γ' where m is even and $\mu_{\mathrm{CZ}}(\gamma) - \mu_{\mathrm{CZ}}(\gamma')$ is odd. Orbits that are not bad are called **good**.

The upshot is that coherent orientations can be defined on the union of all components $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ for which all of the orbits in the lists $\boldsymbol{\gamma}^+$ and $\boldsymbol{\gamma}^-$ are good. This does not mean that moduli spaces involving bad orbits cannot be dealt with—in fact, such moduli spaces have the convenient property that the number of distinct choices of asymptotic markers is always even, and every such choice can be cancelled by an alternative choice that induces the opposite orientation. For this reason, while bad orbits certainly can appear in breaking of holomorphic curves, we will see that they do not need to serve as generators of SFT.

EXAMPLE 11.7. The following thought-experiment produces at first glance an apparent counterexample to Proposition 11.5. Working in a symplectization $(\mathbb{R} \times M, J)$ with $J \in \mathcal{J}(\mathcal{H})$ for some stable Hamiltonian structure $\mathcal{H}$ on a $(2n-1)$-manifold M, suppose $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ represents a Fredholm regular curve in a moduli space $\mathcal{M}^{\$}_{g,0}(J, A, \gamma^2, \varnothing)$ of virtual dimension 1, where u has exactly one puncture, asymptotic to a bad orbit γ^2, which is necessarily a double cover of some other orbit γ. Suppose also that u itself is a double cover of some curve $v : (\dot{\Sigma}', j') \to (\mathbb{R} \times M, J)$ asymptotic to γ. Since u has only one puncture, it cannot be a cover of a trivial cylinder, thus it is not invariant under the $\mathbb{R}$-action. Meanwhile the regularity assumption and $\mathrm{ind}(u) = 1$ imply via the implicit function theorem from Lecture 7 that a neighborhood of u in $\mathcal{M}^{\$}_{g,0}(J, A, \gamma^2, \varnothing)$ is a smooth 1-manifold, so it follows that the component of $\mathcal{M}^{\$}_{g,0}(J, A, \gamma^2, \varnothing)$ containing u consists only of the $\mathbb{R}$-translations of u, and it is diffeomorphic to $\mathbb{R}$. But since u is a double cover of v, it also has a nontrivial automorphism, which must necessarily act by a 180-degree rotation on a neighborhood of the puncture, implying that it changes the asymptotic marker. In other words, the two possible choices of asymptotic marker for u at its unique puncture are related to each other by a biholomorphic map, so they define equivalent elements of $\mathcal{M}^{\$}_{g,0}(J, A, \gamma^2, \varnothing)$. It follows that the map $\mathcal{M}^{\$}_{g,0}(J, A, \gamma^2, \varnothing) \to \mathcal{M}^{\$}_{g,0}(J, A, \gamma^2, \varnothing)$ considered in Proposition 11.5 acts as *the identity map* on the component containing u. Clearly, the identity map is never orientation reversing.

Where did we go wrong? It turns out that the particular set of assumptions we adopted in this example can never actually be satisfied all at once—in particular, this is one situation in which transversality for multiply covered curves really is impossible. To see this, note first that if u is Fredholm regular, then the curve v that it covers must also be regular—one can prove this by splitting the linearized Cauchy-Riemann operator for u into symmetric and antisymmetric factors determined by the automorphism group $\mathrm{Aut}(u) = \mathbb{Z}_2$, where the symmetric factor is equivalent to the linearized operator for v (cf. [**Wen**]). Since v also cannot be invariant under the $\mathbb{R}$-action, the orbifold in which it lives contains a 1-dimensional family of $\mathbb{R}$-translations of v, implying $\mathrm{ind}(v) \geqslant 1$. But compare the indices of v and u: for some choice of trivialization τ along γ, we have

$$\begin{aligned}\mathrm{ind}(v) &= (n-3)\chi(\dot{\Sigma}') + 2c_1^\tau(v^*\xi) + \mu_{\mathrm{CZ}}^\tau(\gamma), \qquad \text{and}\\ \mathrm{ind}(u) &= (n-3)\chi(\dot{\Sigma}) + 2c_1^\tau(u^*\xi) + \mu_{\mathrm{CZ}}^\tau(\gamma^2) = 1.\end{aligned}$$

Since $\dot{\Sigma}$ and $\dot{\Sigma}'$ each have exactly one puncture, their Euler characteristics are both odd, but the assumption that γ^2 is a bad orbit implies that $\mu_{\mathrm{CZ}}^\tau(\gamma^2) - \mu_{\mathrm{CZ}}^\tau(\gamma)$ is also odd, thus $\mathrm{ind}(v)$ must be even and therefore satisfies $\mathrm{ind}(v) \geqslant 2$. If this holds, then there exists at least a smooth 2-parameter family of curves near v, and one can then take double covers of these curves to find a smooth 2-parameter family of curves near u, contradicting the conclusion that $\mathcal{M}_{g,0}^{\$}(J, A, \gamma^2, \varnothing)$ is a 1-manifold near u.

The only way out of this paradox is to remember that since u is multiply covered, the results of Lecture 8 provide no guarantee that it can ever be assumed Fredholm regular, no matter how generically $J \in \mathcal{J}(\mathcal{H})$ is chosen. The impossibility of transversality for the curve in this example does not mean that such curves can generally be ignored, but rather that more sophisticated methods are needed for understanding how to count them (or avoid them) in SFT. We will discuss this issue further in §12.4.

11.2. Permutations of punctures and bad orbits

Before addressing the actual construction of coherent orientations, we can already give heuristic proofs of Propositions 11.4 and 11.5. They are not fully rigorous because they are based on the same pretense as Definition 11.2, namely that all curves we ever have to worry about (including multiple covers) are regular. But we will be able to turn these into precise arguments in §11.7, after discussing the determinant line bundle.

HEURISTIC PROOF OF PROPOSITION 11.4. To simplify the notation, suppose $\hat{\boldsymbol{\gamma}}^+$ consists of only two orbits, so $\hat{\boldsymbol{\gamma}}^+ = (\gamma_1, \gamma_2)$ and $\check{\boldsymbol{\gamma}}^+ = (\gamma_2, \gamma_1)$. Consider the gluing scenario shown in Figure 11.2, where $u \in \mathcal{M}_{g,m}^{\$}(J, A, (\gamma_1, \gamma_2), \boldsymbol{\gamma}^-)$ needs to be glued to a disjoint union of two planes

$$u_B \in \mathcal{M}_{0,0}^{\$}(J_+, B, \varnothing, \gamma_1), \qquad u_C \in \mathcal{M}_{0,0}^{\$}(J_+, C, \varnothing, \gamma_2).$$

You might object that there's no guarantee that such planes must exist in $\mathbb{R} \times M_+$, e.g. the orbits γ_1 and γ_2 might not even be contractible. This concern is valid so far as it goes, but it misses the point: since we're talking about gluing rather than compactness, we do not need any seriously global information about $\widehat{W}$ and M_+, as the gluing process doesn't depend on anything outside a small neighborhood of the curves we're considering. Thus we are free to change the global structure of

M_+ elsewhere so that the planes u_B and u_C will exist.[1] If you still can't imagine how one might do this, try not to worry about it and just think of Figure 11.2 as a thought-experiment: it's a situation that certainly does sometimes happen, so when it does, let's see what it implies about orientations.

Assuming all three curves in the picture are regular, there will be smooth open neighborhoods

$$u \in \mathcal{U}_{12} \subset \mathcal{M}^{\$}_{g,m}(J, A, (\gamma_1, \gamma_2), \boldsymbol{\gamma}^-)$$

$$[(u_B, u_C)] \in \mathcal{U}_{BC} \subset \left(\mathcal{M}^{\$}_{0,0}(J_+, B, \varnothing, \gamma_1) \times \mathcal{M}^{\$}_{0,0}(J_+, C, \varnothing, \gamma_2)\right) \Big/ \mathbb{R}$$

and a gluing map

$$\Psi_{BC} : [R_0, \infty) \times \mathcal{U}_{12} \times \mathcal{U}_{BC} \hookrightarrow \mathcal{M}^{\$}_{g,m}(J, A + B + C, \varnothing, \boldsymbol{\gamma}^-),$$

which must be orientation preserving by assumption. But reversing the order of the product $\mathcal{M}^{\$}_{0,0}(J_+, B, \varnothing, \gamma_1) \times \mathcal{M}^{\$}_{0,0}(J_+, C, \varnothing, \gamma_2)$ and letting $u' \in \mathcal{M}^{\$}_{g,m}(J, A, (\gamma_2, \gamma_1), \boldsymbol{\gamma}^-)$ denote the image of u under the map that switches the order of its positive punctures, there are also smooth open neighborhoods

$$u' \in \mathcal{U}_{21} \subset \mathcal{M}^{\$}_{g,m}(J, A, (\gamma_2, \gamma_1), \boldsymbol{\gamma}^-)$$

$$[(u_C, u_B)] \in \mathcal{U}_{CB} \subset \left(\mathcal{M}^{\$}_{0,0}(J_+, C, \varnothing, \gamma_2) \times \mathcal{M}^{\$}_{0,0}(J_+, B, \varnothing, \gamma_1)\right) \Big/ \mathbb{R}$$

and a gluing map

$$\Psi_{CB} : [R_0, \infty) \times \mathcal{U}_{21} \times \mathcal{U}_{CB} \hookrightarrow \mathcal{M}^{\$}_{g,m}(J, A + B + C, \varnothing, \boldsymbol{\gamma}^-).$$

If both of these gluing maps preserve orientation, then the effect on orientations of the map from $\mathcal{M}^{\$}_{g,m}(J, A, (\gamma_1, \gamma_2), \boldsymbol{\gamma}^-)$ to $\mathcal{M}^{\$}_{g,m}(J, A, (\gamma_2, \gamma_1), \boldsymbol{\gamma}^-)$ defined by interchanging the positive punctures must be the same as that of the map

$$\mathcal{M}^{\$}_{0,0}(J_+, B, \varnothing, \gamma_1) \times \mathcal{M}^{\$}_{0,0}(J_+, C, \varnothing, \gamma_2) \to \mathcal{M}^{\$}_{0,0}(J_+, C, \varnothing, \gamma_2) \times \mathcal{M}^{\$}_{0,0}(J_+, B, \varnothing, \gamma_1)$$
$$(u_B, u_C) \mapsto (u_C, u_B).$$

The latter is orientation reversing if and only if both moduli spaces of planes are odd dimensional, which means $n - 3 + \mu_{CZ}(\gamma_i)$ is odd for $i = 1, 2$. $\square$

HEURISTIC PROOF OF PROPOSITION 11.5. Let us reuse the thought-experiment of Figure 11.2, but with different details in focus. Suppose γ_1 in the picture is an m-fold covered orbit γ^m, where γ is simply covered, and suppose that u_B is also an m-fold cover, taking the form

$$u_B(z) = v(z^m)$$

for a somewhere injective plane $v \in \mathcal{M}_{0,0}(J_+, B_0, \varnothing, \gamma)$. We're going to assume again that all curves in the discussion are regular, including the multiple cover u_B; while this doesn't sound very plausible, we will see once the determinant line bundle

[1] Of course by the maximum principle, planes with only negative ends will not exist in $\mathbb{R} \times M_+$ if this is the symplectization of a contact manifold. But we could also change the contact data to a stable Hamiltonian structure for which such planes are allowed.

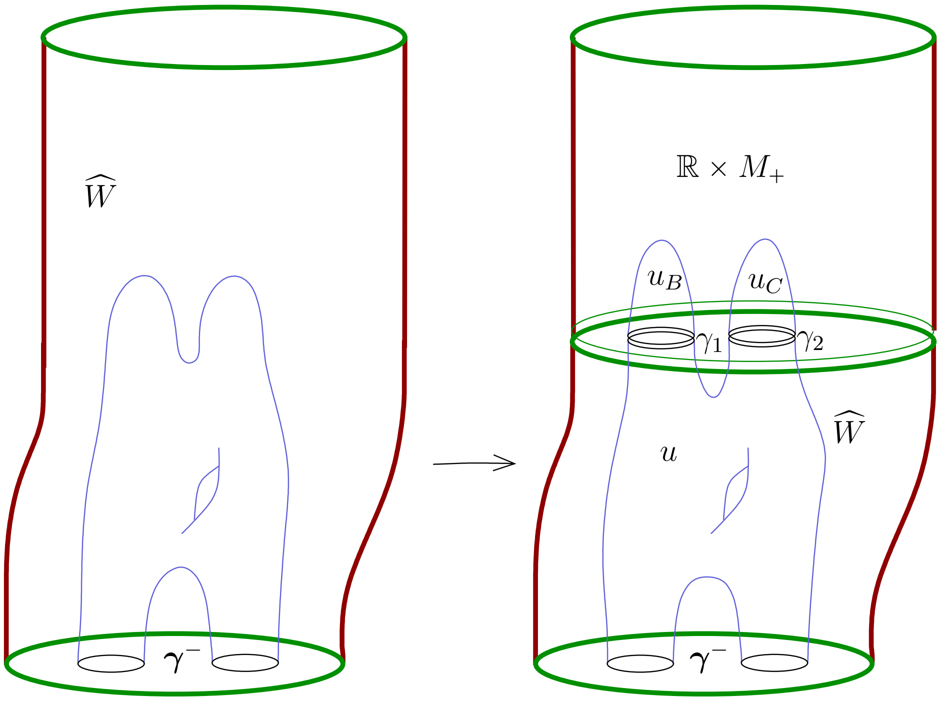

FIGURE 11.2. The gluing thought-experiment used for proving Propositions 11.4 and 11.5.

enters the picture in §11.4 that it is an irrelevant detail. Now, u_B has a cyclic automorphism group

$$\mathrm{Aut}(u_B) = \mathbb{Z}_m \subset \mathrm{U}(1)$$

which acts freely on the set of m choices of asymptotic marker for u_B. Then if we act with the same element of $\mathbb{Z}_m$ on u_B and on the corresponding asymptotic marker for u, the building is unchanged, as it has the same decoration. Coherence therefore implies that the effect on orientations of the map from $\mathcal{M}^{\$}_{g,m}(J, A, (\gamma_1, \gamma_2), \boldsymbol{\gamma}^-)$ to itself defined by acting with the canonical generator of $\mathbb{Z}_m \subset \mathrm{U}(1)$ on the marker at γ_1 is the same as the effect of the map $\mathcal{M}^{\$}_{0,0}(J_+, mB_0, \varnothing, \gamma^m) \to \mathcal{M}^{\$}_{0,0}(J_+, mB_0, \varnothing, \gamma^m)$ defined by composing $u_B : \mathbb{C} \to \mathbb{R} \times M_+$ with $\psi(z) := e^{2\pi i/m} z$.

The derivative of this map from $\mathcal{M}^{\$}_{0,0}(J_+, mB_0, \varnothing, \gamma^m)$ to itself at u_B defines a linear self-map

$$\Psi : T_{u_B}\mathcal{M}_{0,0}(J_+, mB_0, \varnothing, \gamma^m) \to T_{u_B}\mathcal{M}_{0,0}(J_+, mB_0, \varnothing, \gamma^m)$$

with $\Psi^m = \mathbb{1}$. The latter implies that Ψ cannot reverse orientation if m is odd. If m is even, observe that the representation theory of $\mathbb{Z}_m$ gives a decomposition

$$T_{u_B}\mathcal{M}_{0,0}(J_+, mB_0, \varnothing, \gamma^m) = V_1 \oplus V_{-1} \oplus V_{\mathrm{rot}},$$

where Ψ acts on $V_{\pm 1}$ as $\pm\mathbb{1}$, and V_{rot} is a direct sum of real 2-dimensional subspaces on which Ψ acts by rotations (and therefore preserves orientations). Thus Ψ reverses the orientation of $T_{u_B}\mathcal{M}_{0,0}(J_+, mB_0, \varnothing, \gamma^m)$ if and only if $\dim V_{-1}$ is odd. As we will review in the next section, $T_{u_B}\mathcal{M}_{0,0}(J_+, mB_0, \varnothing, \gamma^m)$ is a space of holomorphic sections of $u_B^* T(\mathbb{R} \times M_+)$ modulo a subspace defined via the linearized automorphisms of $\mathbb{C}$, so V_1 consists of precisely those sections η that satisfy $\eta = \eta \circ \psi$, meaning they are m-fold covers of sections of $v^* T(\mathbb{R} \times M_+)$. This defines a bijective correspondence between V_1 and $T_v \mathcal{M}_{0,0}(J_+, B_0, \varnothing, \gamma)$, so

$$\dim V_{-1} = \dim \mathcal{M}_{0,0}(J_+, mB_0, \varnothing, \gamma^m) - \dim \mathcal{M}_{0,0}(J_+, B_0, \varnothing, \gamma) \qquad (\text{mod } 2).$$

The result then comes from plugging in the dimension formulas for these two moduli spaces. ☐

11.3. Orienting moduli spaces in general

We now discuss concretely what is involved in orienting a moduli space of J-holomorphic curves.

Recall from Lecture 7 that whenever a curve $u_0 : (\dot{\Sigma} = \Sigma \backslash \Gamma, j_0) \to (\widehat{W}, J)$ with marked points $\Theta \subset \dot{\Sigma}$ is Fredholm regular, a neighborhood of u_0 in $\mathcal{M}(J)$ can be identified with

$$\bar{\partial}_J^{-1}(0)/G,$$

where $G = \mathrm{Aut}(\Sigma, j_0, \Gamma \cup \Theta)$ and $\bar{\partial}_J$ is the smooth Fredholm section

$$\mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta} : (j, u) \mapsto du + J(u) \circ du \circ j,$$

defined on the product of a G-invariant Teichmüller slice $\mathcal{T}$ through $(j_0, \Gamma \cup \Theta)$ with a Banach manifold $\mathcal{B}^{k,p,\delta}$ of $W^{k,p}$-smooth maps $\dot{\Sigma} \to \widehat{W}$ satisfying an exponential decay condition at the cylindrical ends. Here G acts on $\bar{\partial}_J^{-1}(0)$ by

$$G \times \bar{\partial}_J^{-1}(0) \to \bar{\partial}_J^{-1}(0) : (\varphi, (j, u)) \mapsto (\varphi^* j, u \circ \varphi). \qquad (11.2)$$

Regularity means that the linearization $D\bar{\partial}_J(j_0, u_0) : T_{j_0}\mathcal{T} \oplus T_{u_0}\mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0)}$ is surjective, and the implicit function theorem then gives a natural identification

$$T_{u_0}\mathcal{M}(J) = \ker D\bar{\partial}_J(j_0, u_0) \big/ \mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta),$$

where $\mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta)$ denotes the Lie algebra of G, which acts on $\ker D\bar{\partial}_J(j_0, u_0)$ by differentiating (11.2).[2] This action actually defines an *inclusion* of $\mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta)$ into $\ker D\bar{\partial}_J(j_0, u_0)$ whenever u_0 is not constant, thus we can regard $\mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta)$ as a subspace of $\ker D\bar{\partial}_J(j_0, u_0)$.

As outlined in the proof of Proposition 11.1, the space $\mathcal{M}^\$(J)$ with asymptotic markers admits a similar local description: here one only needs to replace $\bar{\partial}_J^{-1}(0)$ with a finite cover that includes information about asymptotic markers, and the action of G on the cover becomes free.

We now make a useful observation about the spaces $\mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta)$ and $T_{j_0}\mathcal{T}$: namely, they both carry natural complex structures and are thus canonically oriented. This follows from the fact that both the automorphism group G and the Teichmüller space $\mathcal{T}(\Sigma, \Gamma \cup \Theta) = \mathcal{J}(\Sigma) \big/ \mathrm{Diff}_0(\Sigma, \Gamma \cup \Theta)$ are naturally complex manifolds. On the linearized level, one way to see it is via the fact—discussed previously in §7.3.1—that $\mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta)$ and $T_{[j_0]}\mathcal{T}(\Sigma, \Gamma \cup \Theta)$ have natural identifications

[2] The presence of $\mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta)$ in this discussion is only relevant in the finite set of "non-stable" cases where $\chi(\dot{\Sigma} \backslash \Theta) \geqslant 0$, since otherwise G is finite and thus $\mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta)$ is trivial.

with the kernel and cokernel respectively of the canonical linear Cauchy-Riemann type operator on (Σ, j_0),

$$\mathbf{D}_{\mathrm{Id}} : W^{k,p}_{\Gamma\cup\Theta}(T\Sigma) \to W^{k-1,p}(\overline{\mathrm{End}}_{\mathbb{C}}(T\Sigma)), \tag{11.3}$$

which is the linearization at $\mathrm{Id} : \Sigma \to \Sigma$ of the nonlinear operator that detects holomorphic maps $(\Sigma, j_0) \to (\Sigma, j_0)$. This operator is equivalent to the operator that defines the holomorphic structure of $T\Sigma$, thus it is complex linear. To handle the punctures and marked points, one needs to restrict the nonlinear operator to the space of $W^{k,p}$-smooth maps $\Sigma \to \Sigma$ that fix every point in $\Gamma \cup \Theta$, thus the domain of the linearization becomes the finite-codimensional subspace

$$W^{k,p}_{\Gamma\cup\Theta}(T\Sigma) := \left\{ X \in W^{k,p}(T\Sigma) \;\middle|\; X|_{\Gamma\cup\Theta} = 0 \right\}.$$

This subspace is still complex, thus so is (11.3), and its kernel and cokernel inherit natural complex structures.

The complex structure on $\mathfrak{aut}(\Sigma, j_0, \Gamma \cup \Theta)$ means that defining an orientation on the tangent space $T_{u_0}\mathcal{M}^{\$}(J)$ is equivalent to defining one on $\ker D\bar\partial_J(j_0, u_0)$. The latter operator takes the form

$$D\bar\partial_J(j_0, u_0) : T_{j_0}\mathcal{T} \oplus T_{u_0}\mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0)} : (y, \eta) \mapsto J(u_0) \circ du_0 \circ y + \mathbf{D}_{u_0}\eta,$$

where $\mathbf{D}_{u_0} : W^{k,p,\delta}(u_0^*T\widehat{W}) \oplus V_\Gamma \to W^{k-1,p,\delta}(\overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot\Sigma, u_0^*T\widehat{W}))$ is the usual linearized Cauchy-Riemann operator at u_0, with V_Γ denoting a complex $(\#\Gamma)$-dimensional space of smooth sections that are constant near infinity. The remarks above and the fact that u_0 is J-holomorphic imply that the first term in this operator,

$$T_{j_0}\mathcal{T} \to \mathcal{E}^{k-1,p,\delta}_{(j_0,u_0)} : y \mapsto J(u_0) \circ du_0 \circ y$$

is a complex-linear map. Now if $\mathbf{D}_{u_0}$ happens also to be a complex-linear map, then we are done, because $\ker D\bar\partial_J(j_0, u_0)$ will then be a complex vector space and inherit a natural orientation.

In general, $\mathbf{D}_{u_0}$ is not complex linear, though it does have a *complex-linear part*,

$$\mathbf{D}^{\mathbb{C}}_{u_0}\eta := \frac{1}{2}\left(\mathbf{D}_{u_0}\eta - J\mathbf{D}_{u_0}(J\eta)\right),$$

which is also a Cauchy-Riemann type operator. The space of all Cauchy-Riemann type operators on a fixed vector bundle is affine, so one can interpolate from $\mathbf{D}_{u_0}$ to $\mathbf{D}^{\mathbb{C}}_{u_0}$ through a path of Cauchy-Riemann type operators, though they may not all be Fredholm—this depends on the asymptotic operators at the punctures. In the special case however where there are no punctures, one can easily imagine making use of this idea: if $\dot\Sigma = \Sigma$ is a closed surface, then the obvious homotopy from $\mathbf{D}_{u_0}$ to its complex-linear part yields a homotopy from $D\bar\partial_J(j_0, u_0)$ to its complex-linear part, and if every operator along this homotopy happens to be surjective, then the canonical orientation defined on the kernel of the complex-linear operator determines an orientation on $\ker D\bar\partial_J(j_0, u_0)$.

There are two obvious problems with the above discussion:

(1) We have no way to ensure that every operator in the homotopy from $D\bar\partial_J(j_0, u_0)$ to its complex-linear part is surjective;

(2) If there are punctures, then we cannot even expect every operator in this homotopy to be Fredholm.

The first problem motivates the desire to define a notion of orientations for a Fredholm operator $\mathbf{T}$ that does not require $\mathbf{T}$ to be surjective but reduces to the usual notion of orienting $\ker \mathbf{T}$ whenever it is. The solution to this problem is the *determinant line bundle*, which we will discuss in the next section. With this object in hand, the above discussion for the case of closed curves can be made rigorous, so that all smooth moduli spaces of closed J-holomorphic curves inherit canonical orientations. One of the advantages of using the determinant line bundle is that the question of orientations becomes entirely disjoined from the question of transversality: if one can orient the determinant line bundle then moduli spaces of regular curves inherit orientations, but orienting the determinant bundle does not require knowing in advance whether the curves are regular.

The second problem is obviously significant because in the punctured case, moduli spaces of J-holomorphic curves sometimes have *odd* real dimension, making it clearly impossible to homotop $D\bar{\partial}_J(j_0, u_0)$ through Fredholm operators to one that is complex linear. The solution in this case will be to define orientations algorithmically via the coherence condition, and we will describe a suitable algorithm for this in §11.6.

11.4. The determinant line bundle

Fix real Banach spaces X and Y and let $\mathrm{Fred}_{\mathbb{R}}(X, Y)$ denote the space of real-linear Fredholm operators, viewed as an open subset of the Banach space $\mathscr{L}_{\mathbb{R}}(X, Y)$ of all bounded linear operators. We'll use the following notation throughout: if V is an n-dimensional real vector space, then the top-dimensional exterior power of V is denoted by

$$\Lambda^{\max} V := \Lambda^n V.$$

This 1-dimensional real vector space is spanned by any wedge product of the form $v_1 \wedge \ldots \wedge v_n$ where $(v_1, \ldots, v_n)$ is a basis of V. Denoting the dual space of V by V^*, note that there is a canonical isomorphism $(\Lambda^{\max} V)^* = \Lambda^{\max} V^*$. If $\dim V = 0$, then we adopt the convention $\Lambda^{\max} V = \mathbb{R}$.

DEFINITION 11.8. Given $\mathbf{T} \in \mathrm{Fred}_{\mathbb{R}}(X, Y)$, the **determinant line** of $\mathbf{T}$ is the real 1-dimensional vector space

$$\det(\mathbf{T}) = (\Lambda^{\max} \ker \mathbf{T}) \otimes (\Lambda^{\max} \operatorname{coker} \mathbf{T})^* .$$

The union of these vector spaces defines the set

$$\det(X, Y) := \{(\mathbf{T}, v) \mid \mathbf{T} \in \mathrm{Fred}_{\mathbb{R}}(X, Y),\ v \in \det(\mathbf{T})\},$$

and for a subset $\mathcal{U} \subset \mathrm{Fred}_{\mathbb{R}}(X, Y)$, we will denote by $\det(X, Y)|_{\mathcal{U}} \subset \det(X, Y)$ the preimage of $\mathcal{U}$ under the projection $\det(X, Y) \to \mathrm{Fred}_{\mathbb{R}}(X, Y) : (\mathbf{T}, v) \mapsto \mathbf{T}$.

Recall that the set $\mathcal{U} \subset \mathrm{Fred}_{\mathbb{R}}(X, Y)$ of surjective Fredholm operators $X \to Y$ is open, and $\{(\mathbf{T}, v) \in \mathcal{U} \times X \mid v \in \ker \mathbf{T}\}$ forms a smooth finite-rank subbundle of the trivial Banach space bundle $\mathcal{U} \times X$ (see Proposition 3.23). For $\mathbf{T} \in \mathcal{U}$, we have $\det(\mathbf{T}) = \Lambda^{\max} \ker \mathbf{T}$, so it follows that $\det(X, Y)|_{\mathcal{U}}$ has a natural smooth vector bundle structure. Our main goal in this section is to prove that the latter is true for the space of *all* Fredholm operators, not just the surjective ones:

THEOREM 11.9. *The natural vector bundle structure of* $\det(X, Y)$ *over the space of surjective Fredholm operators extends over* $\mathrm{Fred}_{\mathbb{R}}(X, Y)$ *to define a smooth vector bundle* $\det(X, Y) \xrightarrow{\pi} \mathrm{Fred}_{\mathbb{R}}(X, Y)$ *of real rank* 1 *such that* $\pi^{-1}(\mathbf{T}) = \det(\mathbf{T})$ *for each* $\mathbf{T} \in \mathrm{Fred}_{\mathbb{R}}(X, Y)$.

If $\mathbf{T} \in \mathrm{Fred}_{\mathbb{R}}(X, Y)$ is surjective, an orientation of $\det(\mathbf{T})$ is equivalent to an orientation of $\ker \mathbf{T}$. More generally, an orientation of $\det(\mathbf{T})$ is equivalent to an orientation of $\ker \mathbf{T} \oplus \operatorname{coker} \mathbf{T}$. If $\mathbf{T}$ is an isomorphism, then $\det(\mathbf{T})$ is simply $\mathbb{R}$, in which case an orientation of $\det(\mathbf{T})$ can be called **positive** or **negative** depending on whether it matches the canonical orientation of $\mathbb{R}$.

To construct local trivializations of $\det(X, Y) \to \mathrm{Fred}_{\mathbb{R}}(X, Y)$, we start with the special case where X and Y are both finite dimensional. Here *every* linear map is Fredholm, including the zero map, and its determinant is simply $\Lambda^{\max} X \otimes (\Lambda^{\max} Y)^*$. It will be convenient to use the following notation: if V is a real 1-dimensional vector space, we define the (nonlinear) bijection

$$V \backslash \{0\} \to V^* \backslash \{0\} : v \mapsto v^* \quad \text{such that} \quad v^*(v) = 1.$$

LEMMA 11.10. *Suppose* X *and* Y *are real finite-dimensional vector spaces. Then for every* $\mathbf{T} \in \mathscr{L}_{\mathbb{R}}(X, Y)$, *there exists a canonical isomorphism*

$$\Psi_{\mathbf{T}} : (\Lambda^{\max} \ker \mathbf{T}) \otimes (\Lambda^{\max} \operatorname{coker} \mathbf{T})^* \xrightarrow{\cong} (\Lambda^{\max} X) \otimes (\Lambda^{\max} Y)^*,$$

uniquely characterized by the condition that for any nonzero elements $\mathbf{k} \in \Lambda^{\max} \ker \mathbf{T}$, $\mathbf{c} \in \Lambda^{\max} \operatorname{coker} \mathbf{T}$ *and any linearly independent set* $v_1, \ldots, v_q \in X$ *spanning a subspace complementary to* $\ker \mathbf{T} \subset X$,

$$\Psi_{\mathbf{T}}(\mathbf{k} \otimes \mathbf{c}^*) = (\mathbf{k} \wedge v_1 \wedge \ldots \wedge v_q) \otimes (\mathbf{c} \wedge \mathbf{T}v_1 \wedge \ldots \wedge \mathbf{T}v_q)^*. \tag{11.4}$$

Here the product $\mathbf{c} \wedge \mathbf{T}v_1 \wedge \ldots \wedge \mathbf{T}v_q \in \Lambda^{\max} Y$ *is defined by choosing any complement* $C \subset Y$ *of* $\operatorname{im} \mathbf{T}$ *and identifying* $\mathbf{c} \in \Lambda^{\max} \operatorname{coker} \mathbf{T}$ *with an element of* $\Lambda^{\max} C \subset \Lambda^{\dim C} Y$ *via the natural isomorphism* $C \xrightarrow{\cong} \operatorname{coker} \mathbf{T}$ *arising from the restriction of the quotient projection* $Y \to Y / \operatorname{im} \mathbf{T}$, *and this product is independent of choices.*

PROOF. Assume $\dim X = n$, $\dim Y = m$, $\dim \ker \mathbf{T} = k$ and $\dim \operatorname{coker} \mathbf{T} = \ell$, so $\operatorname{ind}(\mathbf{T}) = k - \ell = n - m$, thus $n - k = m - \ell$, and the latter is also q in (11.4). The vectors $\mathbf{T}v_1, \ldots, \mathbf{T}v_q$ thus form a basis of $\operatorname{im} \mathbf{T}$. We show first that the product $\mathbf{c} \wedge \mathbf{T}v_1 \wedge \ldots \wedge \mathbf{T}v_q \in \Lambda^{\max} Y$ is independent of the choice of complement $C \subset Y$ for $\operatorname{im} \mathbf{T}$. Indeed, choose a basis $c_1, \ldots, c_\ell$ of $\operatorname{coker} \mathbf{T}$ and let $\tilde{c}_1, \ldots, \tilde{c}_\ell \in C$ denote the image of this basis under the natural isomorphism of $\operatorname{coker} \mathbf{T}$ to some choice of complement $C \subset Y$. Changing the complement would then change each $\tilde{c}_i$ by vectors in $\operatorname{im} \mathbf{T}$, and since $\mathbf{T}v_1, \ldots, \mathbf{T}v_q$ is a basis of $\operatorname{im} \mathbf{T}$, this does not change $\tilde{c}_1 \wedge \ldots \tilde{c}_\ell \wedge \mathbf{T}v_1 \wedge \ldots \wedge \mathbf{T}v_q$.

For the same reason, the product $\mathbf{k} \wedge v_1 \wedge \ldots \wedge v_q \in \Lambda^{\max} X$ depends only on $\mathbf{k} \in \Lambda^{\max} \ker \mathbf{T}$ and $\mathbf{v} := [v_1] \wedge \ldots \wedge [v_q] \in \Lambda^{\max}(X / \ker \mathbf{T})$. To see that the stated formula for $\Psi_{\mathbf{T}}(\mathbf{k} \otimes \mathbf{c}^*)$ is independent of the choice of $\mathbf{v} \neq 0 \in \Lambda^{\max}(X / \ker \mathbf{T})$, observe that replacing $\mathbf{v}$ by $\lambda \mathbf{v}$ for some $\lambda \in \mathbb{R} \backslash \{0\}$ replaces $\mathbf{T}\mathbf{v} := \mathbf{T}v_1 \wedge \ldots \wedge \mathbf{T}v_q \in$

$\Lambda^{\max}(\operatorname{im}\mathbf{T})$ by $\lambda\mathbf{T}\mathbf{v}$, so $\mathbf{c}\wedge\mathbf{T}\mathbf{v}$ becomes $\lambda\mathbf{c}\wedge\mathbf{T}\mathbf{v}$ and $(\mathbf{c}\wedge\mathbf{T}\mathbf{v})^*$ therefore becomes $\frac{1}{\lambda}(\mathbf{c}\wedge\mathbf{T}\mathbf{v})^*$, giving

$$(\mathbf{k}\wedge\lambda\mathbf{v})\otimes(\mathbf{c}\wedge\mathbf{T}(\lambda\mathbf{v}))^* = \lambda(\mathbf{k}\wedge\mathbf{v})\otimes\frac{1}{\lambda}(\mathbf{c}\wedge\mathbf{T}\mathbf{v})^* = (\mathbf{k}\wedge\mathbf{v})\otimes(\mathbf{c}\wedge\mathbf{T}\mathbf{v})^*.$$

This proves that the formula (11.4) is independent of choices. It is also bilinear with respect to $\mathbf{k}$ and $\mathbf{c}^*$ and is nontrivial, thus it is an isomorphism. $\square$

EXERCISE 11.11. This exercise establishes a naturality property for the isomorphism $\Psi_{\mathbf{T}}$ in Lemma 11.10. Fix Banach space isomorphisms $\varphi\in\mathscr{L}_{\mathbb{R}}(X,X')$ and $\psi\in\mathscr{L}_{\mathbb{R}}(Y',Y)$ and, for a given $\mathbf{T}'\in\operatorname{Fred}_{\mathbb{R}}(X',Y')$, let $\mathbf{T}:=\psi\mathbf{T}'\varphi\in\operatorname{Fred}_{\mathbb{R}}(X,Y)$. Then φ and ψ also define isomorphisms

$$\ker\mathbf{T}\xrightarrow{\varphi}\ker\mathbf{T}', \qquad \operatorname{im}\mathbf{T}'\xrightarrow{\psi}\operatorname{im}\mathbf{T}, \qquad \operatorname{coker}\mathbf{T}'\xrightarrow{\psi}\operatorname{coker}\mathbf{T},$$

which determine corresponding isomorphisms $\varphi_*:\Lambda^{\max}\ker\mathbf{T}\to\Lambda^{\max}\ker\mathbf{T}'$ and $\psi^*:\Lambda^{\max}(\operatorname{coker}\mathbf{T})^*\to\Lambda^{\max}(\operatorname{coker}\mathbf{T}')^*$. Prove that if the Banach spaces in this discussion are finite dimensional, then the following diagram commutes:

$$\begin{tikzcd}
\det(\mathbf{T}) \arrow[r, "\Psi_{\mathbf{T}}"] \arrow[d, "\varphi_*\otimes\psi^*"] & \Lambda^{\max}X\otimes(\Lambda^{\max}Y)^* \arrow[d, "\varphi_*\otimes\psi^*"] \\
\det(\mathbf{T}') \arrow[r, "\Psi_{\mathbf{T}'}"] & \Lambda^{\max}X'\otimes(\Lambda^{\max}Y')^*
\end{tikzcd}$$

A second naturality property of the isomorphism in Lemma 11.10 concerns direct sums. If $\mathbf{T}_i\in\operatorname{Fred}_{\mathbb{R}}(X_i,Y_i)$ for $i=1,2$, then $\mathbf{T}_1\oplus\mathbf{T}_2$ is in $\operatorname{Fred}_{\mathbb{R}}(X_1\oplus X_2,Y_1\oplus Y_2)$, and the canonical isomorphisms

$$\begin{aligned}
\ker(\mathbf{T}_1\oplus\mathbf{T}_2) &= \ker\mathbf{T}_1\oplus\ker\mathbf{T}_2,\\
(\operatorname{coker}(\mathbf{T}_1\oplus\mathbf{T}_2))^* &= (\operatorname{coker}\mathbf{T}_1)^*\oplus(\operatorname{coker}\mathbf{T}_2)^*
\end{aligned}$$

give rise to a canonical isomorphism

$$\begin{aligned}
&\det(\mathbf{T}_1)\otimes\det(\mathbf{T}_2)\to\det(\mathbf{T}_1\oplus\mathbf{T}_2)\\
&(\mathbf{k}_1\otimes\mathbf{c}_1^*)\otimes(\mathbf{k}_2\otimes\mathbf{c}_2^*)\mapsto(-1)^{(\dim\ker\mathbf{T}_2)(\dim\operatorname{coker}\mathbf{T}_1)}(\mathbf{k}_1\wedge\mathbf{k}_2)\otimes(\mathbf{c}_1^*\wedge\mathbf{c}_2^*).
\end{aligned}\tag{11.5}$$

The sign in this formula is related to the reordering of $\mathbf{k}_2\in\Lambda^{\max}\ker\mathbf{T}_2$ and $\mathbf{c}_1^*\in(\Lambda^{\max}\operatorname{coker}\mathbf{T}_1)^*$, and is necessary in order for the result of the following exercise to hold.

EXERCISE 11.12. Show that if X_1,X_2,Y_1,Y_2 are finite dimensional, then the diagram

$$\begin{tikzcd}
\det(\mathbf{T}_1)\otimes\det(\mathbf{T}_2) \arrow[r, "\Psi_{\mathbf{T}_1}\otimes\Psi_{\mathbf{T}_2}"] \arrow[d] & \Lambda^{\max}X_1\otimes(\Lambda^{\max}Y_1)^*\otimes\Lambda^{\max}X_2\otimes(\Lambda^{\max}Y_2)^* \arrow[d] \\
\det(\mathbf{T}_1\oplus\mathbf{T}_2) \arrow[r, "\Psi_{\mathbf{T}_1\oplus\mathbf{T}_2}"] & \Lambda^{\max}(X_1\oplus X_2)\otimes(\Lambda^{\max}(Y_1\oplus Y_2))^*
\end{tikzcd}$$

commutes, where the vertical arrows represent the canonical map (11.5), the one at the right defined in terms of the Fredholm operators $X_1\xrightarrow{0}Y_1$ and $X_2\xrightarrow{0}Y_2$.

The bundle structure of $\det(X,Y)$ can now be understood via the following construction from §3.3.1. Given $\mathbf{T}_0 \in \mathrm{Fred}_{\mathbb{R}}(X,Y)$, write $X = V_{\mathbf{T}_0} \oplus K_{\mathbf{T}_0}$ and $Y = W_{\mathbf{T}_0} \oplus C_{\mathbf{T}_0}$ where $K_{\mathbf{T}_0} = \ker \mathbf{T}_0$, $C_{\mathbf{T}_0} \cong \operatorname{coker} \mathbf{T}_0$, $W_{\mathbf{T}_0} = \operatorname{im} \mathbf{T}_0$ and $\mathbf{T}_0|_{V_{\mathbf{T}_0}} : V_{\mathbf{T}_0} \to W_{\mathbf{T}_0}$ is an isomorphism. We use these splittings to write any other operator $\mathbf{T} \in \mathrm{Fred}_{\mathbb{R}}(X,Y)$ as

$$\mathbf{T} = \begin{pmatrix} \mathbf{A}_{\mathbf{T}} & \mathbf{B}_{\mathbf{T}} \\ \mathbf{C}_{\mathbf{T}} & \mathbf{D}_{\mathbf{T}} \end{pmatrix} : V_{\mathbf{T}_0} \oplus K_{\mathbf{T}_0} \to W_{\mathbf{T}_0} \oplus C_{\mathbf{T}_0},$$

and let $\mathcal{U} \subset \mathrm{Fred}_{\mathbb{R}}(X,Y)$ denote an open neighborhood of $\mathbf{T}_0$ such that the block $\mathbf{A}_{\mathbf{T}} : V \to W$ is invertible for all $\mathbf{T} \in \mathcal{U}$. This gives rise to a pair of smooth maps

$$\mathcal{U} \to \mathscr{L}_{\mathbb{R}}(V_{\mathbf{T}_0} \oplus K_{\mathbf{T}_0}) = \mathscr{L}_{\mathbb{R}}(X) : \mathbf{T} \mapsto F_{\mathbf{T}} := \begin{pmatrix} \mathbb{1} & -\mathbf{A}_{\mathbf{T}}^{-1}\mathbf{B}_{\mathbf{T}} \\ 0 & \mathbb{1} \end{pmatrix},$$

and

$$\mathcal{U} \to \mathscr{L}_{\mathbb{R}}(W_{\mathbf{T}_0} \oplus C_{\mathbf{T}_0}) = \mathscr{L}_{\mathbb{R}}(Y) : \mathbf{T} \mapsto G_{\mathbf{T}} := \begin{pmatrix} \mathbb{1} & 0 \\ -\mathbf{C}_{\mathbf{T}}\mathbf{A}_{\mathbf{T}}^{-1} & \mathbb{1} \end{pmatrix}$$

such that $F_{\mathbf{T}}$ and $G_{\mathbf{T}}$ are invertible for every $\mathbf{T} \in \mathcal{U}$ and satisfy

$$G_{\mathbf{T}}\mathbf{T}F_{\mathbf{T}} = \begin{pmatrix} \mathbf{A}_{\mathbf{T}} & 0 \\ 0 & \Phi_{\mathbf{T}} \end{pmatrix} \in \mathrm{Fred}_{\mathbb{R}}(V_{\mathbf{T}_0} \oplus K_{\mathbf{T}_0}, W_{\mathbf{T}_0} \oplus C_{\mathbf{T}_0}), \tag{11.6}$$

where we write

$$\Phi_{\mathbf{T}} := \mathbf{D}_{\mathbf{T}} - \mathbf{C}_{\mathbf{T}}\mathbf{A}_{\mathbf{T}}^{-1}\mathbf{B}_{\mathbf{T}} \in \mathscr{L}_{\mathbb{R}}(K_{\mathbf{T}_0}, C_{\mathbf{T}_0}).$$

This shows in particular that $F_{\mathbf{T}}$ maps $\ker \Phi_{\mathbf{T}} \subset K_{\mathbf{T}_0}$ isomorphically to $\ker \mathbf{T}$, while $G_{\mathbf{T}}$ maps $\operatorname{im} \mathbf{T}$ isomorphically to $W_{\mathbf{T}_0} \oplus \operatorname{im} \Phi_{\mathbf{T}}$ and thus descends to an isomorphism of $\operatorname{coker} \mathbf{T}$ to $\operatorname{coker} \Phi_{\mathbf{T}}$. Using this construction, we can define splittings of the trivial Banach space bundles $\mathcal{U} \times X$ and $\mathcal{U} \times Y$ into direct sums of smooth subbundles $V \oplus K$ and $W \oplus C$ respectively, with fibers

$$V_{\mathbf{T}} := F_{\mathbf{T}}(V_{\mathbf{T}_0}), \quad K_{\mathbf{T}} := F_{\mathbf{T}}(K_{\mathbf{T}_0}), \quad W_{\mathbf{T}} := G_{\mathbf{T}}^{-1}(W_{\mathbf{T}_0}), \quad C_{\mathbf{T}} := G_{\mathbf{T}}^{-1}(C_{\mathbf{T}_0}),$$

such that for all $\mathbf{T} \in \mathcal{U}$,

$$\ker \mathbf{T} \subset K_{\mathbf{T}} \quad \operatorname{im} \mathbf{T} \supset W_{\mathbf{T}}, \qquad \mathbf{T}(V_{\mathbf{T}}) \subset W_{\mathbf{T}} \text{ and } \quad \mathbf{T}(K_{\mathbf{T}}) \subset C_{\mathbf{T}}.$$

We will refer to any collection $(\mathcal{U}, V, K, W, C)$ consisting of a neighborhood $\mathcal{U} \subset \mathrm{Fred}_{\mathbb{R}}(X,Y)$ of $\mathbf{T}_0$ and subbundles $V, K \subset \mathcal{U} \times X$ and $W, C \subset \mathcal{U} \times Y$ constructed in the above manner as a **reduction** of $\mathrm{Fred}_{\mathbb{R}}(X,Y)$ **near** $\mathbf{T}_0$. Note that a reduction is uniquely determined on some neighborhood of $\mathbf{T}_0$ by the choice of complements $V_{\mathbf{T}_0} \subset X$ for $\ker \mathbf{T}_0$ and $C_{\mathbf{T}_0} \subset Y$ for $\operatorname{im} \mathbf{T}_0$.

A reduction $(\mathcal{U}, V, K, W, C)$ near $\mathbf{T}_0$ determines a smooth vector bundle structure on $\det(X,Y)|_{\mathcal{U}}$ as follows. Let us write $\mathbf{T}|_{V_{\mathbf{T}}} = \mathbf{T}^\infty \in \mathrm{Fred}_{\mathbb{R}}(V_{\mathbf{T}}, W_{\mathbf{T}})$ and $\mathbf{T}|_{K_{\mathbf{T}}} = \mathbf{T}^{\mathrm{red}} \in \mathrm{Fred}_{\mathbb{R}}(K_{\mathbf{T}}, C_{\mathbf{T}})$, noting that by (11.6), $\mathbf{T}^\infty$ is conjugate to the isomorphism $\mathbf{A}_{\mathbf{T}} : V_{\mathbf{T}_0} \to W_{\mathbf{T}_0}$, so that $\det(\mathbf{T}^\infty) = \mathbb{R}$, and this isomorphism is canonical, i.e. it does not depend on any choices. The natural isomorphisms in (11.4) and (11.5) therefore give rise to a chain of natural isomorphisms

$$\begin{aligned} \det(\mathbf{T}) = \det(\mathbf{T}^\infty \oplus \mathbf{T}^{\mathrm{red}}) &\xleftarrow{\cong} \det(\mathbf{T}^\infty) \otimes \det(\mathbf{T}^{\mathrm{red}}) = \mathbb{R} \otimes \det(\mathbf{T}^{\mathrm{red}}) \\ &= \det(\mathbf{T}^{\mathrm{red}}) \xrightarrow{\Psi_{\mathbf{T}^{\mathrm{red}}}} \Lambda^{\max} K_{\mathbf{T}} \otimes (\Lambda^{\max} C_{\mathbf{T}})^*. \end{aligned}$$

Since K and C are smooth subbundles of the trivial bundles $\mathcal{U} \times X$ and $\mathcal{U} \times Y$ respectively, the right hand side is now the fiber of a smooth vector bundle $\Lambda^{\max} K \otimes (\Lambda^{\max} C)^*$ over $\mathcal{U}$, so this family of isomorphisms determines a vector bundle structure on $\det(X, Y)|_{\mathcal{U}}$. One can use the naturality properties to show that any two reductions determine smoothly isomorphic bundle structures on the region where they overlap.

EXERCISE 11.13. Show that if X and Y are complex Banach spaces, then the restriction of $\det(X, Y)$ to the subspace of complex-linear Fredholm operators $\mathrm{Fred}_{\mathbb{C}}(X, Y) \subset \mathrm{Fred}_{\mathbb{R}}(X, Y)$ admits a canonical orientation compatible with the complex structures of $\ker \mathbf{T}$ and $\operatorname{coker} \mathbf{T}$ for each $\mathbf{T} \in \mathrm{Fred}_{\mathbb{C}}(X, Y)$. Show also that whenever $\mathbf{T} \in \mathrm{Fred}_{\mathbb{C}}(X, Y)$ is an isomorphism, the resulting orientation of $\det(\mathbf{T}) = \mathbb{R}$ is positive.

The orientation of $\det(\mathbf{T})$ for $\mathbf{T} \in \mathrm{Fred}_{\mathbb{C}}(X, Y)$ described in Exercise 11.13 is called the **complex orientation**.

11.5. Determinant bundles of moduli spaces

Combining ideas from the previous two sections, let

$$\det(J) \to \mathcal{M}^{\$}(J)$$

denote the topological line bundle that associates to any $u \in \mathcal{M}^{\$}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ the determinant line of the Fredholm operator

$$\mathbf{D}_u : W^{k,p,\delta}(u^*T\widehat{W}) \oplus V_\Gamma \to W^{k-1,p,\delta}(\overline{\mathrm{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u^*T\widehat{W})).$$

One can construct local trivializations for this bundle using Theorem 11.9 and any choice of local trivializations for the Banach space bundles $T\mathcal{B}^{k,p,\delta}$ and $\mathcal{E}^{k-1,p,\delta}$.

PROPOSITION 11.14. *Any orientation of* $\det(J) \to \mathcal{M}^{\$}(J)$ *canonically determines an orientation of* $\mathcal{M}^{\$,\mathrm{reg}}(J)$.

PROOF. As explained in §11.3, an orientation of $\mathcal{M}^{\$,\mathrm{reg}}(J)$ near a particular curve $u_0 : (\dot{\Sigma}, j_0) \to (\widehat{W}, J)$ is equivalent to a continuously varying choice of orientations for the kernels

$$\ker D\bar{\partial}_J(j, u) \subset T_j\mathcal{T} \oplus T_u\mathcal{B}^{k,p,\delta}$$

for all $(j, u) \in \bar{\partial}_J^{-1}(0)$, where $\mathcal{T}$ is a Teichmüller slice through $(j_0, \Gamma \cup \Theta)$. The operator $D\bar{\partial}_J(j, u)$ is of the form

$$\mathbf{L}(y, \eta) := J(u) \circ du \circ y + \mathbf{D}_u \eta$$

and thus is homotopic through Fredholm operators to

$$\mathbf{L}^0(y, \eta) := \mathbf{D}_u \eta,$$

namely via the homotopy $\mathbf{L}^s(y, \eta) := sJ(u) \circ du \circ y + \mathbf{D}_u \eta$ for $s \in [0, 1]$. The kernel and cokernel of $\mathbf{L}^0$ are $T_j\mathcal{T} \oplus \ker \mathbf{D}_u$ and $\operatorname{coker} \mathbf{D}_u$ respectively, and since $T_j\mathcal{T}$ carries a complex structure, the orientation of $\det(\mathbf{D}_u)$ naturally determines an orientation of $\det(\mathbf{L}^0)$. Using the homotopy $\mathbf{L}^s$, this determines orientations of $\det(D\bar{\partial}_J(j, u))$ and thus orientations of $\ker D\bar{\partial}_J(j, u)$ for all (j, u) near (j_0, u_0), and this orientation does not depend on the choice of Teichmüller slice since the operators $\mathbf{D}_u$ also do not. $\square$

From now on, when we speak of an **orientation of** $\mathcal{M}^\$(J)$, we will actually mean an orientation of the bundle $\det(J) \to \mathcal{M}^\$(J)$. The above proposition implies that this is equivalent to what we want in applications, but one advantage of talking about $\det(J)$ is that there is no need to limit the discussion to curves that are regular, i.e. the notion of an orientation of $\mathcal{M}^\$(J)$ now makes sense even though $\mathcal{M}^\$(J)$ is not globally a smooth object.

PROPOSITION 11.15. *Suppose all Reeb orbits in* $\boldsymbol{\gamma}^\pm$ *are nondegenerate and have the property that their asymptotic operators are complex linear. Then* $\mathcal{M}^\$_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ *admits a natural orientation, known as the* ***complex orientation****.*

PROOF. Having complex-linear asymptotic operators implies that the obvious homotopy from each Cauchy-Riemann operator $\mathbf{D}_u$ to its complex-linear part does not change the asymptotic operators and is therefore a homotopy through Fredholm operators. We therefore have a continuously varying homotopy of each of the relevant fibers of $\det(J)$ to the determinant bundle over a family of complex-linear operators, which inherit the complex orientation described in Exercise 11.13. $\square$

Proposition 11.15 applies in particular to all moduli spaces of closed J-holomorphic curves, and thus solves the orientation problem in that case.

11.6. An algorithm for coherent orientations

We now briefly describe the construction of coherent orientations due to Bourgeois and Mohnke [BM04]. A slightly different construction is sketched in [EGH00], but the Bourgeois-Mohnke construction has become the standard.

Recall from Lecture 4 the notion of an *asymptotically Hermitian* vector bundle (E, J) over a punctured Riemann surface $(\dot{\Sigma}, j)$. Here $(\dot{\Sigma}, j)$ is endowed with the extra structure of fixed cylindrical ends $(\dot{\mathcal{U}}_z, j) \cong (Z_\pm, i)$ for each puncture $z \in \Gamma^\pm$, which determines a choice of asymptotic markers. Likewise, the bundle E comes with an asymptotic bundle $(E_z, J_z, \omega_z) \to S^1$ associated to each puncture, carrying compatible complex and symplectic structures. We shall now endow E with a bit more structure that is always naturally present in the case $E = u^*T\widehat{W}$: namely, assume each of the asymptotic bundles comes with a splitting

$$(E_z, J_z, \omega_z) = (\mathbb{C} \oplus \widehat{E}_z, i \oplus \widehat{J}_z, \omega_0 \oplus \widehat{\omega}_z), \tag{11.7}$$

where ω_0 is the standard symplectic structure on the trivial complex line bundle $(\mathbb{C}, i)$ over S^1, and $(\widehat{E}_z, \widehat{J}_z, \widehat{\omega}_z) \to S^1$ is another Hermitian bundle. Fix a choice $\{\mathbf{A}_z\}_{z\in\Gamma}$ of nondegenerate asymptotic operators on each of the bundles $(\widehat{E}_z, \widehat{J}_z, \widehat{\omega}_z)$, and define the topological space

$$\mathcal{CR}(E, \{\mathbf{A}_z\}_{z\in\Gamma})$$

to consist of all Cauchy-Riemann type operators on E that are asymptotic at the punctures $z \in \Gamma$ to the asymptotic operators

$$(-i\partial_t) \oplus \mathbf{A}_z : \Gamma(\mathbb{C} \oplus \widehat{E}_z) \to \Gamma(\mathbb{C} \oplus \widehat{E}_z).$$

This is an affine space, so it is contractible, and if $\delta > 0$ is sufficiently small and $V_\Gamma \subset \Gamma(E)$ denotes a complex $(\#\Gamma)$-dimensional space of smooth sections that take constant values in $\mathbb{C} \oplus \{0\} \subset E_z$ near each puncture z, then every $\mathbf{D} \in \mathcal{CR}(E, \{\mathbf{A}_z\}_{z\in\Gamma})$ determines a Fredholm operator

$$\mathbf{D} : W^{k,p,\delta}(E) \oplus V_\Gamma \to W^{k-1,p,\delta}(\overline{\mathrm{Hom}}_{\mathbb{C}}(T\Sigma, E)).$$

It follows that a choice of orientation of the determinant line for any one of these operators determines an orientation for all of them. The point of this construction is that every $u \in \mathcal{M}^{\$}(J)$ determines an operator $\mathbf{D}_u$ belonging to a space of this form.

We now construct a gluing operation for Cauchy-Riemann operators that linearizes the gluing maps described in §11.1. Suppose $(E^i, J^i) \to (\dot{\Sigma}_i = \Sigma_i \backslash \Gamma_i, j_i)$ for $i = 0, 1$ is a pair of asymptotically Hermitian bundles of the same rank, endowed with asymptotic splittings as in (11.7) and asymptotic operators $\{\mathbf{A}_z\}_{z\in\Gamma_i}$, and that there exists a pair of punctures $z_0 \in \Gamma_0^+$ and $z_1 \in \Gamma_1^-$ such that some unitary bundle isomorphism

$$\widehat{E}^1_{z_1} \xrightarrow{\cong} \widehat{E}^0_{z_0}$$

identifies $\mathbf{A}_{z_1}$ with $\mathbf{A}_{z_0}$. Note that such an isomorphism is uniquely determined up to homotopy whenever it exists. For $R > 0$, we can define a family of glued Riemann surfaces

$$(\dot{\Sigma}_R = \Sigma_R \backslash \Gamma_R, j_R)$$

by cutting off the ends $(R, \infty) \times S^1 \subset \dot{\mathcal{U}}_{z_0}$ and $(-\infty, -R) \times S^1 \subset \dot{\mathcal{U}}_{z_1}$ and gluing $\{R\} \times S^1 \subset \dot{\Sigma}_0$ to $\{-R\} \times S^1 \subset \dot{\Sigma}_1$. The glued Riemann surface contains an annulus biholomorphic to $([-R, R] \times S^1, i)$ in place of the infinite cylindrical ends at the punctures z_0 and z_1. The unitary isomorphism $\widehat{E}^1_{z_1} \to \widehat{E}^0_{z_0}$ then determines an isomorphism $E^1_{z_1} \to E^0_{z_0}$ via the splitting (11.7) and hence an asymptotically Hermitian bundle

$$(E^R, J^R) \to (\dot{\Sigma}_R, J_R).$$

Using cutoff functions in the neck $[-R, R] \times S^1$, any Cauchy-Riemann operators $\mathbf{D}_i \in \mathcal{CR}(E^i, \{\mathbf{A}_z\}_{z\in\Gamma_i})$ for $i = 0, 1$ now determine a family of operators

$$\mathbf{D}_R \in \mathcal{CR}(E^R, \{\mathbf{A}_z\}_{z\in\Gamma_R})$$

uniquely up to homotopy. Analogously to the gluing maps in §11.1, one can arrange this construction so that the operators $\mathbf{D}_R$ converge in some sense to the pair $(\mathbf{D}_0, \mathbf{D}_1)$ as $R \to \infty$, which has the following consequence:

LEMMA 11.16 ([BM04, Corollary 7]). *For $R > 0$ sufficiently large, there is a natural isomorphism*

$$\det(\mathbf{D}_0) \otimes \det(\mathbf{D}_1) \to \det(\mathbf{D}_R)$$

that is defined up to homotopy. □

Up to some additional direct sums and quotients by finite-dimensional complex vector spaces, this isomorphism should be understood as the linearization of a gluing map between moduli spaces, generalized to a setting in which the holomorphic curves involved need not be regular. To orient $\mathcal{M}^{\$}(J)$ coherently, it now suffices to choose orientations for the operators in $\mathcal{CR}(E, \{\mathbf{A}_z\}_{z\in\Gamma})$ that vary continuously under deformations of j and E and are preserved by the isomorphisms of Lemma 11.16. This motivates the following generalization of Definition 11.2.

Definition 11.17. A system of **coherent orientations** is an assignment to each asymptotically Hermitian bundle $(E, J) \to (\dot\Sigma, j)$ with asymptotic splittings as in (11.7) and asymptotic operators $\{\mathbf{A}_z\}_{z\in\Gamma}$ of an orientation for the determinant line of each $\mathbf{D} \in \mathcal{CR}(E, \{\mathbf{A}_z\})$, such that these orientations vary continuously with $\mathbf{D}$ as well as the data j and J, and such that the isomorphisms in Lemma 11.16 are always orientation preserving.

The prescription of [**BM04**] to construct such systems is now as follows.

(1) For any trivial bundle E over $\dot\Sigma = \mathbb{C}$ with ∞ as a negative puncture and any asymptotic operator $\mathbf{A}_\infty$, choose an arbitrary continuous family of orientations for the operators in $\mathcal{CR}(E, \{\mathbf{A}_\infty\})$, subject only to the requirement that these should match the complex orientation whenever $\mathbf{A}_\infty$ is complex linear.
(2) For any trivial bundle E_- over $\dot\Sigma = \mathbb{C}$ with ∞ as a positive puncture, any asymptotic operator $\mathbf{A}_\infty$ and any $\mathbf{D}_- \in \mathcal{CR}(E_-, \{\mathbf{A}_\infty\})$, let E_+ denote the trivial bundle over $\mathbb{C}$ with a negative puncture as in step (1), choose any $\mathbf{D}_+ \in \mathcal{CR}(E_+, \{\mathbf{A}_\infty\})$ and construct the resulting family of glued operators
$$\mathbf{D}_R \in \mathcal{CR}(E^R),$$
where the E^R are trivial bundles over S^2. Since S^2 has no punctures, $\mathbf{D}_R$ has a natural complex orientation, so define the orientation of $\mathbf{D}_-$ to be the one that is compatible via Lemma 11.16 with this and the orientation chosen for $\mathbf{D}_+$ in step (1).
(3) For an arbitrary $(E, J) \to (\dot\Sigma, j)$, glue positive and negative planes to $\dot\Sigma$ to produce a bundle over a closed surface $\widehat\Sigma$, and define the orientation of any $\mathbf{D} \in \mathcal{CR}(E, \{\mathbf{A}_z\}_{z\in\Gamma})$ to be compatible via Lemma 11.16 with the choices in steps (1) and (2) and the complex orientation for operators over $\widehat\Sigma$.

It should be easy to convince yourself that if we now vary the bundle $(E, J) \to (\dot\Sigma, j)$ or the operators on this bundle (but *not* the asymptotic operators!) continuously, the capping procedure described in step (3) above produces a continuous family of Cauchy-Riemann type operators on bundles over closed Riemann surfaces. Since these all carry the complex orientation, the resulting orientations of the original operators vary continuously. It is similarly clear from the construction that any Cauchy-Riemann operator whose asymptotic operators are all complex linear will end up with the complex orientation. Bourgeois and Mohnke use this fact to prove that any system of orientations constructed in this way is compatible with *all* possible linear gluing maps arising from Lemma 11.16. The idea is to reduce it to the complex-linear case by gluing cylinders to the ends of any asymptotically Hermitian bundle so that the asymptotic operators can be changed at will; see [**BM04**, Proposition 8].

11.7. Permutations and bad orbits revisited

The heuristic proofs in §11.2 can now be made precise in the following way.

Suppose $\mathbf{D} \in \mathcal{CR}(E, \{\mathbf{A}_z\}_{z\in\Gamma})$, and $\mathbf{D}'$ is the same operator after interchanging two of the punctures in Γ. Imagine gluing $(E, J) \to (\dot{\Sigma}, j)$ to trivial bundles E^1 and E^2 over planes in order to cap off the two punctures that are being interchanged, and choose Cauchy-Riemann operators $\mathbf{D}_1$ and $\mathbf{D}_2$ on these planes to form a glued operator on the capped surface. This capping procedure is done one plane at a time, and the order of the two punctures determines which plane is glued first. Compatibility with the isomorphisms of Lemma 11.16 then dictates that the orientations of $\det(\mathbf{D})$ and $\det(\mathbf{D}')$ match if and only if the orientations of $\det(\mathbf{D}_1) \otimes \det(\mathbf{D}_2)$ and $\det(\mathbf{D}_2) \otimes \det(\mathbf{D}_1)$ match. Since orientations of $\det(\mathbf{D}_i)$ for $i = 1, 2$ are equivalent to orientations of $\ker \mathbf{D}_i \oplus \operatorname{coker} \mathbf{D}_i$, reversing the order of the tensor product changes orientations if and only if both of these direct sums are odd dimensional, which means $\operatorname{ind}(\mathbf{D}_1)$ and $\operatorname{ind}(\mathbf{D}_2)$ are both odd. If the bundles have complex rank n and the asymptotic operators are $\mathbf{A}_i$ for $k = 1, 2$, we have

$$\operatorname{ind}(\mathbf{D}_i) = n\chi(\mathbb{C}) \pm \mu_{\mathrm{CZ}}((-i\partial_t \oplus \mathbf{A}_i) \pm \delta) = n - 1 \pm \mu_{\mathrm{CZ}}(\mathbf{A}_i),$$

which matches $n - 3 + \mu_{\mathrm{CZ}}(\mathbf{A}_i)$ modulo 2. This proves Proposition 11.4.

Similarly for Proposition 11.5, we consider the action of the generator $\psi \in \mathbb{Z}^m$ on $\det(\mathbf{D})$ where ψ rotates the cylindrical end by $1/m$ at some puncture where the trivialized asymptotic operator $\mathbf{A}$ is of the form $-i\partial_t - S(mt)$ for a loop of symmetric matrices $S(t)$. Capping off this puncture with a plane carrying a Cauchy-Riemann operator $\mathbf{D}_\infty$, coherence dictates that the same transformation must act the same way on the orientation of $\det(\mathbf{D}_\infty)$. Since $\psi^m = 1$, ψ cannot reverse this orientation if m is odd. To understand the case of m even, note first that we are free to choose $\mathbf{D}_\infty$ so that it is an m-fold cover, meaning it is related to the branched cover $\varphi : \mathbb{C} \to \mathbb{C} : z \mapsto z^m$ by

$$\mathbf{D}_\infty(\eta \circ \varphi) = \varphi^* \widehat{\mathbf{D}}_\infty \eta$$

for some other Cauchy-Riemann operator $\widehat{\mathbf{D}}_\infty$, which is asymptotic to $\widehat{\mathbf{A}} := -i\partial_t - S(t)$. Now the group $\mathbb{Z}_m$ generated by ψ acts on $\ker \mathbf{D}_\infty$ and $\operatorname{coker} \mathbf{D}_\infty$, so representation theory tells us

$$\ker \mathbf{D}_\infty = V_1 \oplus V_{-1} \oplus V_{\mathrm{rot}}$$
$$\operatorname{coker} \mathbf{D}_\infty = W_1 \oplus W_{-1} \oplus W_{\mathrm{rot}},$$

where ψ acts on $V_{\pm 1}$ and $W_{\pm 1}$ as $\pm \mathbb{1}$ and acts as orientation-preserving rotations on V_{rot} and W_{rot}. It follows that ψ reverses the orientation of $\ker \mathbf{D}_\infty \oplus \operatorname{coker} \mathbf{D}_\infty$ if and only if $\dim V_{-1} - \dim W_{-1}$ is odd. Now observe that there are natural isomorphisms

$$V_1 = \ker \widehat{\mathbf{D}}_\infty, \qquad W_1 = \operatorname{coker} \widehat{\mathbf{D}}_\infty,$$

hence

$$\dim V_{-1} - \dim W_{-1} = \operatorname{ind}(\mathbf{D}_\infty) - \operatorname{ind}(\widehat{\mathbf{D}}_\infty) \qquad (\mathrm{mod}\ 2).$$

This difference in Fredholm indices is precisely $\mu_{\mathrm{CZ}}(\mathbf{A}) - \mu_{\mathrm{CZ}}(\widehat{\mathbf{A}})$ up to a sign, and this completes the proof of Proposition 11.5.

LECTURE 12

The generating function of SFT

Contents

It is time to begin deriving algebraic consequences from the analytical results of the previous lectures. We saw the simplest possible example of this in Lecture 10, where the behavior of holomorphic cylinders in symplectizations of contact manifolds without contractible Reeb orbits led to a rudimentary version of cylindrical contact homology $HC_*(M, \xi)$ with $\mathbb{Z}_2$ coefficients. Unfortunately, the condition on contractible orbits means that this version of $HC_*(M, \xi)$ cannot always be defined, and even when it can, it only counts cylinders—we would only expect it to capture a small fragment of the information contained in more general moduli spaces of holomorphic curves. Extracting information from these general moduli spaces will require enlarging our algebraic notion of what a Floer-type theory can look like.

12.1. Some important caveats on transversality

For most of this and the next lecture, we fix the following fantastically optimistic assumption:

Assumption 12.1 (science fiction). *One can choose suitably compatible almost complex structures so that all pseudoholomorphic curves are Fredholm regular.*

This assumption held in Lecture 10 for the curves we were interested in, because they were all guaranteed for topological reasons to be somewhere injective. It can also be shown to hold under some very restrictive conditions on Conley-Zehnder indices in dimension three, see [**Nel15**, **Nel20**]. Both of those are very lucky situations, and as we've discussed before, the assumption cannot generally be achieved merely by perturbing J generically—it *must* sometimes fail for curves that are multiply covered, and such curves always exist (see §12.4 for more on this). The only way in reality to ensure something like Assumption 12.1 is to perturb the nonlinear Cauchy-Riemann equation more abstractly, e.g. by replacing $\bar{\partial}_J u = 0$ with an inhomogeneous equation of the form

$$\bar{\partial}_J u = \nu$$

for a generic perturbation ν. This is the standard technique in certain versions of Gromov-Witten theory (see e.g. [**RT95**, **RT97**]), and we will outline the main ideas in §12.4.3. Alternatively, one can allow J to depend generically on points in the domain rather than just points in the target, as in [**MS12**, §7.3]. Both approaches eliminate the initial problem with multiple covers, but they both also run into serious and subtle difficulties concerning the relationship between $\mathcal{M}(J)$ and the strata of its compactification $\overline{\mathcal{M}}(J)$. As observed in [**Sal99**, §5], the presence of natural symmetries in $\overline{\mathcal{M}}(J)$ makes it necessary for any sufficiently general abstract perturbation scheme to involve *multivalued* perturbations, and it is important for these perturbations to be "coherent" in a sense analogous to our discussion of orientations in the previous lecture. These notions have not yet all been developed in a sufficiently consistent and general way to give a rigorous definition of SFT, though there has been much progress: this is the main objective of the long-running *polyfold* project initiated by Hofer-Wysocki-Zehnder [**Hof06**]. More recently, a quite different and much more topological approach was introduced by Pardon [**Par19**] specifically for contact homology, and an approach to general SFT based on Kuranishi structures has been proposed by Ishikawa [**Ish**].

For most of this lecture we will ignore these subtleties and simply adopt Assumption 12.1 as a convenient fiction, thus pretending that all components of $\mathcal{M}(J)$ are smooth orbifolds of the correct dimension and all gluing maps are smooth. All "theorems" stated under this assumption should be read with the caveat that they are only true in a fictional world in which the assumption holds. Even if it is a fiction, one can get quite far with this point of view: it is still possible not only to deduce the essential structure of what we assume will someday be a rigorously defined polyfold-based SFT, but also to infer the existence of certain contact invariants that have interesting rigorous applications requiring only well-established techniques, e.g. the cobordism obstructions discovered in [**LW11**].

12.2. Auxiliary data, grading and supercommutativity

The goal is to define an invariant of closed $(2n-1)$-dimensional contact manifolds (M, ξ) with closed nondegenerate Reeb orbits as generators and a Floer-type differential counting J-holomorphic curves in the symplectization $(\mathbb{R} \times M, d(e^r \alpha))$. The auxiliary data we choose must obviously therefore include a nondegenerate contact form α and a generic $J \in \mathcal{J}(\alpha)$, for which we shall suppose Assumption 12.1 holds. For convenience, we will also assume throughout most of this lecture:

ASSUMPTION 12.2. *$H_1(M)$ is torsion free.*

This is needed mainly in order to be able to define an integer grading, though without this assumption, it is still always possible to define a $\mathbb{Z}_2$-grading—see §12.7.1 for more on what to do when Assumption 12.2 does not hold. We now supplement the auxiliary data (α, J) with the following additional choices:

(1) Coherent orientations as in Lecture 11 for the moduli spaces $\mathcal{M}^\$(J)$ with asymptotic markers.
(2) A collection of **reference curves**
$$S^1 \cong C_1, \ldots, C_r \subset M$$
whose homology classes form a basis of $H_1(M)$.
(3) A unitary trivialization of ξ along each of the reference curves $C_1, \ldots, C_r$, denoted collectively by τ.
(4) A **spanning surface** C_γ for each periodic Reeb orbit γ: this is a smooth map of a compact and oriented surface with boundary into M such that
$$\partial C_\gamma = \sum_i m_i[C_i] - [\gamma]$$
in the sense of singular 2-chains, where $m_i \in \mathbb{Z}$ are the unique coefficients with $[\gamma] = \sum_i m_i[C_i] \in H_1(M)$.

These choices determine the following. To any collections of Reeb orbits $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \ldots, \gamma_{k_\pm}^\pm)$ and any relative homology class $A \in H_2(M, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ with $\partial A = \sum_i[\gamma_i^+] - \sum_j[\gamma_j^-]$, we can now associate a cycle in absolute homology,
$$A + \sum_i C_{\gamma_i^+} - \sum_j C_{\gamma_j^-} \in H_2(M).$$
Indeed, the boundary of this real 2-chain is a sum of linear combinations of the reference curves C_i, which add up to zero because $\sum_i[\gamma_i^+]$ and $\sum_j[\gamma_j^-]$ are homologous. We shall abuse notation and use this correspondence to associate the absolute homology class
$$[u] \in H_2(M)$$
to any asymptotically cylindrical holomorphic curve u in $\mathbb{R} \times M$. Adapting the previous notation,
$$\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$$
for $A \in H_2(M)$ will now denote a moduli space of curves whose relative homology classes glue to the chosen capping surfaces to form A.

Secondly, the chosen trivializations τ along the reference curves can be pulled back and extended over every capping surface C_γ, giving trivializations of ξ along every orbit γ uniquely up to homotopy. We shall define
$$\mu_{\mathrm{CZ}}(\gamma) \in \mathbb{Z}$$
from now on to mean the Conley-Zehnder index of γ relative to this trivialization.

EXERCISE 12.3. Show that if $H_1(M)$ has no torsion and $u : \dot{\Sigma} \to \mathbb{R} \times M$ is asymptotically cylindrical, then its relative first Chern number with respect to the trivializations τ described above satisfies
$$c_1^\tau(u^*T(\mathbb{R} \times M)) = c_1([u]),$$

where $c_1([u])$ denotes the evaluation of $c_1(\xi) \in H^2(M)$ on $[u] \in H_2(M)$.

By Exercise 12.3, the index of a curve $u : (\dot{\Sigma} = \Sigma \backslash \Gamma, j) \to (\mathbb{R} \times M, J)$ with $[u] = A \in H_2(M)$ and asymptotic orbits $\{\gamma_z\}_{z \in \Gamma^\pm}$ can now be written as

$$\operatorname{ind}(u) = -\chi(\dot{\Sigma}) + 2c_1(A) + \sum_{z \in \Gamma^+} \mu_{\mathrm{CZ}}(\gamma_z) - \sum_{z \in \Gamma^-} \mu_{\mathrm{CZ}}(\gamma_z). \tag{12.1}$$

In order to keep track of homology classes of holomorphic curves algebraically, we can define our theory to have coefficients in the group ring $\mathbb{Q}[H_2(M)]$, or more generally,

$$R := \mathbb{Q}[H_2(M)/G]$$

for a given subgroup $G \subset H_2(M)$. Elements of R will be written as finite sums

$$\sum_i c_i e^{A_i} \in R, \qquad c_i \in \mathbb{Q}, \; A_i \in H_2(M)/G,$$

where the multiplicative structure of the group ring is derived from the additive structure of $H_2(M)/G$ by $e^A e^B := e^{A+B}$. The most common examples of G are $H_2(M)$ and the trivial subgroup, giving $R = \mathbb{Q}$ or $R = \mathbb{Q}[H_2(M)]$ respectively. We will see a geometrically meaningful example in between these two extremes in the next lecture.

Finally, we define certain formal variables which have degrees in $\mathbb{Z}$ or $\mathbb{Z}_{2N}$ for some $N \in \mathbb{N}$, and will serve as generators in our graded algebra. To each closed Reeb orbit γ we associate two variables, q_γ, p_γ, whose integer-valued degrees are

$$|q_\gamma| = n - 3 + \mu_{\mathrm{CZ}}(\gamma), \qquad |p_\gamma| = n - 3 - \mu_{\mathrm{CZ}}(\gamma).$$

To remember these numbers, think of the index of a J-holomorphic plane u positively or negatively asymptotic to γ, with $[u] = 0$.

We also assign an integer grading to the group ring $\mathbb{Q}[H_2(M)]$ such that rational numbers have degree 0 and

$$|e^A| = -2c_1(A), \qquad \text{for} \qquad A \in H_2(M).$$

If $c_1(A) = 0$ for every $A \in G$, in particular if $c_1(\xi) = 0$, then this descends to an integer grading on the ring $R = \mathbb{Q}[H_2(M)/G]$. Otherwise, R inherits a $\mathbb{Z}_{2N}$-grading, where

$$N := \min \left\{ c_1(A) > 0 \;\middle|\; A \in G \right\}.$$

A $\mathbb{Z}_2$-grading is well defined in every case.

The algebra will include one additional formal variable $\hbar$, which is defined to have degree

$$|\hbar| = 2(n-3).$$

The degrees of $\hbar$ and the p_γ and q_γ variables should all be interpreted modulo $2N$ if $c_1(\xi)|_G \neq 0$.

The algebra of SFT uses monomials in the variables p_γ and q_γ respectively to encode sets of positive and negative asymptotic orbits of holomorphic curves, while the group ring $R = \mathbb{Q}[H_2(M)/G]$ is used to keep track of the homology classes of such curves, and powers of $\hbar$ are used to keep track of their genus. More precisely, given $g \geqslant 0$, $A \in H_2(M)$ and ordered lists of Reeb orbits $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \dots, \gamma_{k_\pm}^\pm)$, we encode the moduli space $\mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ formally via the product

$$e^A \hbar^{g-1} q^{\boldsymbol{\gamma}^-} p^{\boldsymbol{\gamma}^+} := e^A \hbar^{g-1} q_{\gamma_1^-} \dots q_{\gamma_{k_-}^-} p_{\gamma_1^+} \dots p_{\gamma_{k_+}^+}, \tag{12.2}$$

where we are abusing notation by identifying A with its equivalence class in $H_2(M)/G$ if G is nontrivial. Notice that according to the above definitions, this expression has degree

(12.3)

$$\begin{aligned} |e^A \hbar^{g-1} q^{\boldsymbol{\gamma}^-} p^{\boldsymbol{\gamma}^+}| = |e^A| + (g-1)|\hbar| + \sum_{i=1}^{k_-} \left[(n-3) + \mu_{\mathrm{CZ}}(\gamma_i^-)\right] \\ + \sum_{i=1}^{k_+} \left[(n-3) - \mu_{\mathrm{CZ}}(\gamma_i^+)\right] \\ = -2c_1(A) + (2g - 2 + k_+ + k_-)(n-3) - \sum_{i=1}^{k_+} \mu_{\mathrm{CZ}}(\gamma_i^+) + \sum_{i=1}^{k_-} \mu_{\mathrm{CZ}}(\gamma_i^-) \\ = -\operatorname{vir-dim} \mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-), \end{aligned}$$

interpreted modulo $2N$ if $c_1(\xi)|_G \neq 0$. The orientation results in Lecture 11 suggest introducing a *supercommutativity* relation for the variables q_γ and p_γ: defining the graded commutator bracket by

$$[F, G] := FG - (-1)^{|F||G|} GF, \tag{12.4}$$

we define a relation on the set of all monomials of the form $q^{\boldsymbol{\gamma}^-} p^{\boldsymbol{\gamma}^+}$ by setting

$$[q_{\gamma_1}, q_{\gamma_2}] = [p_{\gamma_1}, p_{\gamma_2}] = 0 \tag{12.5}$$

for all pairs of orbits γ_1 and γ_2. As a consequence, permuting the orbits in the lists $\boldsymbol{\gamma}^\pm$ changes the sign of the monomial (12.2) if and only if it changes the orientation of the corresponding moduli space. In particular, any product that includes multiple copies of an odd generator q_γ or p_γ is identified with 0. This accounts for the fact that any rigid moduli space $\mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ with two copies of γ among its positive or negative asymptotic orbits contains zero curves when counted with the correct signs: every curve is cancelled by a curve that looks identical except for a permutation of two of its punctures.

12.3. The definition of H and commutators

To write down the SFT generating function, let

$$\mathcal{M}^\sigma(J) := \mathcal{M}(J)/\sim$$

denote the space of equivalence classes where two curves are considered equivalent if they have parametrizations that differ only in the ordering of the punctures. This space is in some sense more geometrically natural than $\mathcal{M}(J)$ or $\mathcal{M}^\$(J)$, but due to the orientation results in the previous lecture, less convenient for technical reasons.

Given $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ representing a nonconstant element of $\mathcal{M}^\sigma(J)$ with no marked points, it is natural to define

$$\mathrm{Aut}^\sigma(u) \subset \mathrm{Aut}(\Sigma, j)$$

as the (necessarily finite) group of biholomorphic transformations $\varphi : (\Sigma, j) \to (\Sigma, j)$ satisfying $u = u \circ \varphi$; in particular, elements of $\mathrm{Aut}^\sigma(u)$ are allowed to permute the punctures, so $\mathrm{Aut}^\sigma(u)$ is generally a larger group than the usual $\mathrm{Aut}(u)$. For $k \in \mathbb{Z}$, let

$$\mathcal{M}^\sigma_k(J) \subset \mathcal{M}^\sigma(J)$$

denote the subset consisting of index k curves that have no marked points and whose asymptotic orbits are all *good* (see Definition 11.6 in Lecture 11).

We now define the **SFT generating function** as a formal power series

$$\mathbf{H} = \sum_{u \in \mathcal{M}^\sigma_1(J)/\mathbb{R}} \frac{\epsilon(u)}{|\mathrm{Aut}^\sigma(u)|} \hbar^{g-1} e^A q^{\boldsymbol{\gamma}^-} p^{\boldsymbol{\gamma}^+}, \tag{12.6}$$

where the terms of each monomial are determined by $u \in \mathcal{M}^\sigma_1(J)$ as follows:

- g is the genus of u;
- A is the equivalence class of $[u] \in H_2(M)$ in $H_2(M)/G$;
- $\boldsymbol{\gamma}^\pm = (\gamma^\pm_1, \dots, \gamma^\pm_{k_\pm})$ are the asymptotic orbits of u after arbitrarily fixing orderings of its positive and negative punctures;
- $\epsilon(u) \in \{1, -1\}$ is determined by the chosen coherent orientations on $\mathcal{M}^\$(J)$. Specifically, given the chosen ordering of the punctures and an arbitrary choice of asymptotic markers at each puncture, u determines a 1-dimensional connected component of $\mathcal{M}^\$(J)$, and we define $\epsilon(u) = +1$ if and only if the coherent orientation of $\mathcal{M}^\$(J)$ matches its tautological orientation determined by the $\mathbb{R}$-action.

Note that while both $\epsilon(u)$ and the corresponding monomial $q^{\boldsymbol{\gamma}^-} p^{\boldsymbol{\gamma}^+}$ depend on a choice of orderings of the punctures, their product does not depend on this choice. Moreover, $\epsilon(u)$ does not depend on the choice of asymptotic markers since curves with bad asymptotic orbits are excluded from $\mathcal{M}^\sigma_1(J)$. Since every monomial in $\mathbf{H}$ corresponds to a holomorphic curve of index 1, (12.3) implies

$$|\mathbf{H}| = -1.$$

There are various combinatorially more elaborate ways to rewrite $\mathbf{H}$. For any Reeb orbit γ, let

$$\kappa_\gamma := \mathrm{cov}(\gamma) \in \mathbb{N}$$

denote its covering multiplicity, and for a finite list of orbits $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_k)$, let

$$\kappa_{\boldsymbol{\gamma}} := \prod_{i=1}^k \kappa_{\gamma_i}.$$

Given $u \in \mathcal{M}^\sigma(J)$ with $k_\pm \geqslant 0$ positive/negative punctures asymptotic to the set of orbits $\boldsymbol{\gamma}^\pm = (\gamma^1_\pm, \dots, \gamma^{k_\pm}_\pm)$, there are $k_+!k_-!\kappa_{\boldsymbol{\gamma}^+}\kappa_{\boldsymbol{\gamma}^-}$ ways to order the punctures and choose asymptotic markers, but some of them are equivalent since (by an easy variation on Proposition 11.1) the finite group $\mathrm{Aut}^\sigma(u)$ acts freely on this set of choices. As a result, (12.6) is the same as

$$\mathbf{H} = \sum_{u \in \mathcal{M}_1^\$(J)/\mathbb{R}} \frac{\epsilon(u)}{k_+!k_-!\kappa_{\gamma^+}\kappa_{\gamma^-}} \hbar^{g-1} e^A q^{\gamma^-} p^{\gamma^+}, \tag{12.7}$$

where $\mathcal{M}_1^\$(J)$ denotes the space of all (see Remark 12.5 below for some important commentary about the word "all") index 1 curves without marked points in $\mathcal{M}^\$(J)$, and the rest of the mononomial is determined by the condition that u belongs to $\mathcal{M}_{g,0}^\$(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$, with no need for any arbitrary choices. Another way of writing this is

$$\mathbf{H} = \sum_{g,A,\boldsymbol{\gamma}^+,\boldsymbol{\gamma}^-} \frac{\#\left(\mathcal{M}_{g,0}^\$(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)/\mathbb{R}\right)}{k_+!k_-!\kappa_{\gamma^+}\kappa_{\gamma^-}} \hbar^{g-1} e^A q^{\gamma^-} p^{\gamma^+}, \tag{12.8}$$

where the sum ranges over all integers $g \geqslant 0$, homology classes $A \in H_2(M)$ and ordered tuples of Reeb orbits $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \dots, \gamma_{k_\pm}^\pm)$ for which $\text{vir-dim}\, \mathcal{M}_{g,0}^\$(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) = 1$, and $\#\left(\mathcal{M}_{g,0}^\$(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)/\mathbb{R}\right)$ is the signed count of connected components in $\mathcal{M}_{g,0}^\$(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$. For fixed g and $\boldsymbol{\gamma}^\pm$, the union of these spaces for all $A \in H_2(M)$ is finite due to SFT compactness, as the energy of curves in $(\mathbb{R} \times M, d(e^t\alpha))$ is computed by integrating exact symplectic forms and thus (by Stokes) admits a uniform upper bound in terms of $\boldsymbol{\gamma}^+$. For this reason, (12.8) defines a formal power series in the p variables and in $\hbar$, with coefficients that are *polynomials* in the q variables and the group ring R.

REMARK 12.4. The formulas (12.6) and (12.7) for $\mathbf{H}$ are predicated on the explicit assumption that all J-holomorphic curves are Fredholm regular, and some modification is required for the reality in which this assumption cannot always hold. We will sketch in §12.4.3 how (12.8) can be defined without this assumption, but the count $\#\left(\mathcal{M}_{g,0}^\$(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)/\mathbb{R}\right)$ appearing in this formula will require a more general interpretation in which it may sometimes be a rational number, not an integer.

REMARK 12.5. We played a slightly sneaky trick in writing down (12.7) and (12.8): these summations do not exclude bad orbits, whereas (12.6) was a sum over curves u that are not asymptotic to any bad orbits—a necessary exclusion in that case because $\epsilon(u)$ would otherwise depend on choices of asymptotic markers. The reason bad orbits are allowed in (12.8) is that their total contribution adds up to zero. Indeed, bad orbits are always multiple covers with even multiplicity, so whenever $u \in \mathcal{M}^\$(J)$ has a puncture approaching a bad orbit with multiplicity $2m$, there are $2m - 1$ other elements of $\mathcal{M}^\$(J)$ that differ only by adjustment of the marker at that one puncture, and by Proposition 11.5, half of these cancel out the other half in the signed count.[1] For similar reasons related to Proposition 11.4, (12.8) will not

[1]There is one caveat: if $\mathrm{Aut}(u)$ is nontrivial, then not all of these $2m$ curves are inequivalent elements of $\mathcal{M}^\$(J)$, which is fine if the equivalence classes still cancel in pairs, but one can imagine scenarios as in Example 11.7, where two choices of asymptotic markers defining different signs $\epsilon(u)$ give the same element of $\mathcal{M}^\$(J)$. This is nonsense, but it takes some effort to see why. The

contain any terms with multiple factors of any odd generator q_γ or p_γ, even though curves with multiple ends asymptotic to γ sometimes exist.

REMARK 12.6. Readers famliar with Floer homology may see a resemblance between the group ring $R = \mathbb{Q}[H_2(M)/G]$ and the Novikov rings that often appear in Floer homology, though R is not a Novikov ring since it only allows finite sums. In Floer homology, the Novikov ring sometimes must be included because counts of curves may fail to be finite, though they only do so if the energies of those curves blow up. The situation above is somewhat different: since the symplectization is an exact symplectic manifold, Stokes' theorem implies that energy cannot blow up if the positive asymptotic orbits are fixed, and one therefore obtains well-defined curve counts no matter the choice of the coefficient ring R. The use of the group ring is convenient however for two reasons: first, without it one cannot always define an integer grading, and second, different choices of coefficients can sometimes be used to detect different geometric phenomena via SFT. We will see an example of the latter in Lecture 13.

The compactness and gluing theory of SFT is encoded algebraically by viewing $\mathbf{H}$ as an element on a noncommutative operator algebra determined by the commutator relations

$$\begin{aligned} [p_\gamma, q_\gamma] &= \kappa_\gamma \hbar \\ [p_\gamma, q_{\gamma'}] &= 0 \quad \text{if } \gamma \neq \gamma'. \end{aligned} \tag{12.9}$$

Here $[\ ,\ ]$ again denotes the graded commutator (12.4), so "commuting" generators actually anticommute whenever they are both odd. The rest of the multiplicative structure of this algebra is determined by requiring all elements of R and powers of $\hbar$ (all of which are even generators) to commute with everything, meaning all operators are $R[[\hbar]]$-linear.

One concrete representation of this operator algebra is as follows: let $\mathcal{A}$ denote the graded supercommutative unital algebra over R generated by the set

$$\{q_\gamma \mid \gamma \text{ a good Reeb orbit}\}.$$

The ring of formal power series $\mathcal{A}[[\hbar]]$ is then an $R[[\hbar]]$-module. Define each of the generators q_γ to be $R[[\hbar]]$-linear operators on $\mathcal{A}[[\hbar]]$ via multiplication from the left, and define $p_\gamma : \mathcal{A}[[\hbar]] \to \mathcal{A}[[\hbar]]$ by

$$p_\gamma = \kappa_\gamma \hbar \frac{\partial}{\partial q_\gamma}. \tag{12.10}$$

Here the $R[[\hbar]]$-linear partial derivative operator is defined via

$$\frac{\partial}{\partial q_\gamma} q_\gamma = 1, \qquad \frac{\partial}{\partial q_\gamma} q_{\gamma'} = 0 \quad \text{for } \gamma \neq \gamma'$$

main reason is the observation in Example 11.7 that scenarios of this type are incompatible with the assumption that all curves are Fredholm regular. This may seem an unconvincing explanation in light of the unreasonable optimism of our transversality assumption, but as we will outline in §12.4.3, regularity really *can* be achieved for all curves using multivalued inhomogeneous perturbations, so in that context, a similar argument rules out unwanted automorphisms without having to rely on science fiction.

and the graded Leibniz rule

$$\frac{\partial}{\partial q_\gamma}(FG) = \frac{\partial F}{\partial q_\gamma} G + (-1)^{|q_\gamma||F|} F \frac{\partial G}{\partial q_\gamma}$$

for all homogeneous elements $F, G \in \mathcal{A}[[\hbar]]$.

EXERCISE 12.7. Check that the operator $p_\gamma : \mathcal{A}[[\hbar]] \to \mathcal{A}[[\hbar]]$ defined above has the correct degree and satisfies the commutation relations (12.5) and (12.9).

Notice that while $\mathbf{H}$ contains terms of order -1 in $\hbar$, every term also contains at least one p_γ variable since all index 1 holomorphic curves in $(\mathbb{R} \times M, d(e^t \alpha))$ have at least one positive puncture. The substitution (12.10) thus produces a differential operator in which every term contains a nonnegative power of $\hbar$, giving a well-defined $R[[\hbar]]$-linear operator

$$\mathbf{D}_{\mathrm{SFT}} : \mathcal{A}[[\hbar]] \xrightarrow{\mathbf{H}} \mathcal{A}[[\hbar]].$$

The following may be regarded as the fundamental theorem of SFT.

THEOREM 12.8. $\mathbf{H}^2 = 0$.

We will discuss in §12.6 how this relation follows from the compactness and gluing theory of punctured holomorphic curves, and we will use it in Lecture 13 to define various Floer-type contact invariants. The first and most obvious of these is the homology

$$H_*^{\mathrm{SFT}}(M, \xi) := H_*(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}}),$$

which will turn out to be an invariant of (M, ξ) in the sense that any two choices of α, J and the other auxiliary data described in §12.2 gives rise to a functorial isomorphism between the two graded homology groups. Notice that while $\mathcal{A}[[\hbar]]$ is an algebra, its product structure does not descend to $H_*^{\mathrm{SFT}}(M, \xi)$ since $\mathbf{D}_{\mathrm{SFT}}$ is not a derivation—indeed, it is a formal sum of differential operators of all orders, not just order one. In the next lecture we will discuss various ways to produce homological invariants out of $\mathbf{H}$ with nicer algebraic structures.

On the other hand, it is fairly easy to understand the geometric meaning of the complex $(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$ in Floer-theoretic terms. Each individual curve $u \in \mathcal{M}_1^\sigma(J)$ with genus g, homology class $A \in H_2(M)$ and asymptotic orbits $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \dots, \gamma_{k_\pm}^\pm)$ contributes to $\mathbf{D}_{\mathrm{SFT}}$ the differential operator

$$\frac{\epsilon(u)}{|\operatorname{Aut}^\sigma(u)|} \kappa_{\boldsymbol{\gamma}^+} \hbar^{g+k_+-1} e^A q_{\gamma_1^-} \dots q_{\gamma_{k_-}^-} \frac{\partial}{\partial q_{\gamma_1^+}} \cdots \frac{\partial}{\partial q_{\gamma_{k_+}^+}}.$$

Applying this operator to a monomial $q_{\gamma_1} \dots q_{\gamma_m} \in \mathcal{A}[[\hbar]]$ that does not contain all of the generators $q_{\gamma_1^+}, \dots, q_{\gamma_{k_+}^+}$ will produce zero, and its effect on a product that does contain all of these generators will be to eliminate them and multiply $q_{\gamma_1^-} \dots q_{\gamma_{k_-}^-}$ by whatever remains, plus some combinatorial factors and signs that may arise from differentiating by the same q_γ more than once. Ignoring the combinatorics and signs for the moment, this operation on $q_{\gamma_1} \dots q_{\gamma_m}$ has a geometric interpretation: it counts all *potentially disconnected* J-holomorphic curves of index 1 (i.e. disjoint unions of u with trivial cylinders) that have $\gamma_1, \dots, \gamma_m$ as their positive asymptotic

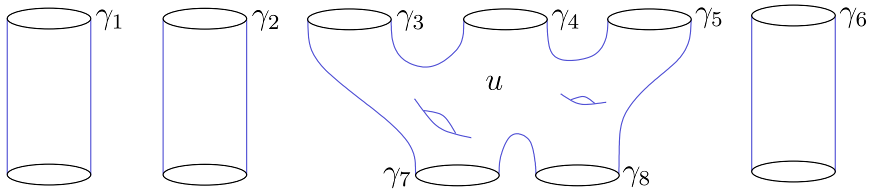

FIGURE 12.1. Counting disjoint unions of index 1 curves $u \in \mathcal{M}_{2,0}(J, A, (\gamma_3, \gamma_4, \gamma_5), (\gamma_7, \gamma_8))$ with some trivial cylinders contributes a multiple of $\hbar^4 e^A q_{\gamma_1} q_{\gamma_2} q_{\gamma_7} q_{\gamma_8} q_{\gamma_6}$ to $\mathbf{D}_{\mathrm{SFT}}(q_{\gamma_1} q_{\gamma_2} q_{\gamma_3} q_{\gamma_4} q_{\gamma_5} q_{\gamma_6})$.

orbits; see Figure 12.1. In other words, the action of $\mathbf{D}_{\mathrm{SFT}}$ on each monomial $q^{\boldsymbol{\gamma}}$ for $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_m)$ is determined by a formula of the form

$$\mathbf{D}_{\mathrm{SFT}} q^{\boldsymbol{\gamma}} = \sum_{g=0}^{\infty} \sum_{A \in H_2(M)} \sum_{\boldsymbol{\gamma}'} \sum_{k=1}^{m} \hbar^{g+k-1} e^A n_g(\boldsymbol{\gamma}, \boldsymbol{\gamma}', k) q^{\boldsymbol{\gamma}'}, \tag{12.11}$$

where $n_g(\boldsymbol{\gamma}, \boldsymbol{\gamma}', k)$ is a product of some combinatorial factors with a signed count of generally disconnected index 1 holomorphic curves of genus g and homology class A with positive ends at $\boldsymbol{\gamma}$ and negative ends at $\boldsymbol{\gamma}'$, such that the nontrivial connected component has exactly k positive ends. The presence of the combinatorial factors hidden in $n_g(\boldsymbol{\gamma}, \boldsymbol{\gamma}', k)$ is a slightly subtle point which we will try to clarify in the following sections.

12.4. Interlude: Orbifolds and branched manifolds

12.4.1. How to count zeroes in an orbifold. As in all versions of Floer theory, the proof that $\mathbf{H}^2 = 0$ is based on the idea that certain moduli spaces are compact oriented 1-dimensional manifolds with boundary, and the signed count of their boundary points is therefore zero. We must be careful of course because, strictly speaking, $\overline{\mathcal{M}}(J)$ is not a manifold, even when Assumption 12.1 holds—it is an orbifold, with the possibility of singularities at multiply covered curves with nontrivial automorphism groups. On the other hand, if we exclude curves with bad asymptotic orbits, then it is an *oriented* orbifold, and oriented 1-dimensional orbifolds happen to be very simple objects: since smooth finite group actions on $\mathbb{R}$ cannot be nontrivial without reversing orientation, all oriented 1-dimensional orbifolds are actually manifolds, so that the standard classification of compact 1-manifolds should apply and lead to the simple formula

$$\text{“}\#\partial\overline{\mathcal{M}}_1(J) = 0.\text{”}$$

I have placed this formula in quotation marks for a reason. The reality of the situation is somewhat more complicated.

This is in fact where it becomes important to remember that Assumption 12.1 in its stated form is really not just science fiction, but *fantasy*: transversality is sometimes impossible to achieve for multiple covers, and we must therefore at least have a sensible back-up plan for such cases. To see the problem, remember that our local structure theorem for $\mathcal{M}(J)$ was proved by identifying it in a neighborhood of any curve $u_0 : (\dot{\Sigma}, j_0) \to (\mathbb{R} \times M, J)$ with a set of the form

$$\bar{\partial}_J^{-1}(0)/G,$$

where $\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$ is a smooth section of a Banach space bundle $\mathcal{E}^{k-1,p,\delta}$ over the product of a Teichmüller slice $\mathcal{T}$ with a Banach manifold $\mathcal{B}^{k,p,\delta}$ of maps $\dot{\Sigma} \to \mathbb{R} \times M$, and G is the group of automorphisms of the domain of u_0, whose action on the base[2]

$$G \times (\mathcal{T} \times \mathcal{B}^{k,p,\delta}) \to \mathcal{T} \times \mathcal{B}^{k,p,\delta} : (\psi, (j, u)) \mapsto (\psi^* j, u \circ \psi)$$

preserves $\bar{\partial}_J^{-1}(0)$. In fact, the action of G on $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$ is covered by a natural action on the bundle $\mathcal{E}^{k-1,p,\delta}$, and the reason for it preserving the zero set is that $\bar{\partial}_J$ is an equivariant section,

$$\bar{\partial}_J(\psi^* j, u \circ \psi) = \psi^* \bar{\partial}_J(j, u).$$

If G is finite, then another way to say this is that $\bar{\partial}_J$ is a smooth Fredholm section of the infinite-dimensional **orbibundle** $\mathcal{E}^{k-1,p,\delta}/G$ over the orbifold $(\mathcal{T} \times \mathcal{B}^{k,p,\delta})/G$, whose isotropy group at (j_0, u_0) is $\mathrm{Aut}(u_0)$. This section is transverse to the zero-section if and only if the usual regularity condition holds, making $\bar{\partial}_J^{-1}(0)/G$ a suborbifold of $(\mathcal{T} \times \mathcal{B}^{k,p,\delta})/G$ whose isotropy group at (j_0, u_0) is some quotient of $\mathrm{Aut}(u_0)$.

REMARK 12.9. Most sensible definitions of the term "orbifold" (cf. [**ALR07**, **Dav**, **FO99**]) require local models of the form $\mathcal{U}/G$, where $\mathcal{U}$ is a G-invariant open subset of a vector space on which the finite group G acts smoothly and *effectively*—the latter condition is necessary in order to have isotropy groups that are well-defined up to isomorphism at every point. In the above example, G acts effectively on $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$ but might have a nontrivial subgroup $H \subset G$ of transformations that fix every element of $\bar{\partial}_J^{-1}(0)$, in which case the G-action on $\bar{\partial}_J^{-1}(0)$ can be replaced by an effective action of G/H. The isotropy group of $(j_0, u_0) \in \bar{\partial}_J^{-1}(0)/G$ is then $\mathrm{Aut}(u_0)/(\mathrm{Aut}(u_0) \cap H)$.

Now to see just how unreasonably optimistic Assumption 12.1 is, notice that it's easy to think up examples of smooth orbibundles in which zeroes of sections can *never* be regular if they have nontrivial isotropy:

EXAMPLE 12.10. Let M denote the real 2-dimensional orbifold $\mathbb{C}/\mathbb{Z}_2$, with $\mathbb{Z}_2$ acting on $\mathbb{C}$ via the involution $z \mapsto \bar{z}$, and define an orbibundle $E \to M$ as the quotient of the trivial complex line bundle $\mathbb{C} \times \mathbb{C} \to \mathbb{C}$ by the $\mathbb{Z}_2$-action generated by the involution $(z, v) \mapsto (\bar{z}, -v)$. A smooth function $f : \mathbb{C} \to \mathbb{C}$ then represents a section $f \in \Gamma(E)$ if and only if it satisfies $f(\bar{z}) = -f(z)$ for all z. The zero set of f therefore is guaranteed to contain the 1-dimensional submanifold $\mathbb{R} \subset M$, so that $\partial_x f(x + iy)|_{y=0} = 0$ and the zeroes along $\mathbb{R}$ can therefore never be regular.

[2]If G is a Lie group of positive dimension, then this discussion should be taken with a grain of salt since, as mentioned in §7.4 (see the discussion preceding Lemma 7.29), the action of G on $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$ is not differentiable. There is no problem however if G is finite, e.g. if the underlying Riemann surface is stable, which is true outside of finitely many cases.

EXAMPLE 12.11. Let $M = \mathbb{C}/\mathbb{Z}_2$ with $\mathbb{Z}_2$ acting as the antipodal map, and consider the trivial complex line bundle $E = M \times \mathbb{C} = (\mathbb{C} \times \mathbb{C})/\mathbb{Z}_2$, where in this case the $\mathbb{Z}_2$ action identifies (z, v) with $(-z, v)$. A smooth function $f : \mathbb{C} \to \mathbb{C}$ then represents a section of E if and only if f is even, implying $df(0) = 0$, thus for any choice of function f that has a zero at the origin, this zero cannot be regular. In contrast to Example 12.10, it is possible in this case to perturb f generically to a section that is transverse to the zero-section, but the zero set of such a perturbation can never contain $0 \in M$.

Continuing with this example for a moment, let us observe that while a zero at $0 \in M$ cannot be regular, it may still be isolated, and in this case we do know how to assign it a $\mathbb{Z}$-valued order. For instance, $f(z) = z^2$ defines a section of $E \to M$ with a zero of order 2 at 0. Notice however that if we perturb this to $f_\epsilon(z) = z^2 + \epsilon$ for $\epsilon > 0$ small, then f_ϵ has two simple zeroes at points near the origin, but they are actually *the same point* in $\mathbb{C}/\mathbb{Z}_2$, giving a count of only 1 zero. This means that if we give the zero of f at the origin its full weight, then we are counting wrongly—the resulting count will not be homotopy invariant. The correct algebraic count of zeroes is evidently

$$\#f^{-1}(0) := \sum_{z \in f^{-1}(0) \subset M} \frac{\operatorname{ord}(f; z)}{\kappa_z} \in \mathbb{Q}, \tag{12.12}$$

where $\operatorname{ord}(f; z) \in \mathbb{Z}$ is the order of the zero (computed in the usual way as a winding number, or in higher dimensions as the degree of a map of spheres, cf. [Mil97]), and $\kappa_z \in \mathbb{N}$ denotes the order of the isotropy group at z.

EXERCISE 12.12. Convince yourself that for any smooth oriented orbibundle $E \to M$ of real rank m over a compact, smooth and oriented m-dimensional orbifold M without boundary, the count (12.12) gives the same result for any section with isolated zeroes.[3] *Hint: The space of sections of an orbibundle is still a vector space, so any two are homotopic. Since M and $[0, 1]$ are both compact, it suffices to focus on small perturbations of a single section in a single orbifold chart.*

For a slightly different perspective on (12.12), consider the special case of a closed orbifold that is the quotient of a closed manifold $\widetilde{M}$ by an effective orientation-preserving finite group action,

$$M = \widetilde{M}/G.$$

Suppose $\widetilde{E} \to \widetilde{M}$ is an oriented vector bundle with rank equal to $\dim M$, and G also acts on $\widetilde{E}$ by orientation-preserving linear bundle maps that cover its action on $\widetilde{M}$, so the quotient

$$E = \widetilde{E}/G \to M$$

is an orbibundle. A section $f : M \to E$ is then equivalent to a G-equivariant section $\widetilde{f} : \widetilde{M} \to \widetilde{E}$, and the signed count of zeroes

$$\#\widetilde{f}^{-1}(0) = \sum_{z \in \widetilde{f}^{-1}(0) \subset \widetilde{M}} \operatorname{ord}(\widetilde{f}; z) \in \mathbb{Z}$$

[3] If you're still not sure what an orbibundle is, a definition can be found in [FO99, Chapter 1].

is of course the same for any section that has only isolated zeroes. It can also be expressed in terms of f since any $z \in f^{-1}(0) \subset M$ has exactly $|G|/\kappa_z$ lifts to points in $\widetilde{f}^{-1}(0) \subset \widetilde{M}$, implying

$$\#\widetilde{f}^{-1}(0) = \sum_{z \in f^{-1}(0) \subset M} \frac{|G|}{\kappa_z} \operatorname{ord}(f; z),$$

and thus $\#f^{-1}(0) = \frac{1}{|G|}\#\widetilde{f}^{-1}(0)$. The invariance of (12.12) is now an immediate consequence of the invariance of $\#\widetilde{f}^{-1}(0)$, which follows from the standard argument as in [Mil97].

12.4.2. Multivalued perturbations. If you enjoyed reading [Mil97] as much as I did, then it may seem tempting to try proving invariance of (12.12) in general by choosing a generic homotopy $H : [0, 1] \times M \to E$ between two generic sections f_0 and f_1 and showing that $H^{-1}(0) \subset [0, 1] \times M$ is a compact oriented 1-dimensional orbifold with boundary. As we observed at the beginning of this section, $H^{-1}(0)$ is then actually a manifold, so the signed count of its boundary points should be zero. But this would give the wrong result: it would suggest that $\sum_{z \in f^{-1}(0) \subset M} \operatorname{ord}(f; z)$ should be homotopy invariant, without the rational weights, and we've already seen that this is not true. What is going on here? The answer is that the homotopy H cannot in general be made transverse to the zero-section, now matter how generically we perturb it! It is an illustration of the fundamental conflict between the notions of *genericity* and *equivariance*.

EXAMPLE 12.13. Let $M = \mathbb{C}/\mathbb{Z}_2$ as in Example 12.11, but define the complex orbibundle $E \to M$ by

$$E = (\mathbb{C} \times \mathbb{C})\big/(z, v) \sim (-z, -v),$$

i.e. the $\mathbb{Z}_2$-action also acts antipodally on fibers. Now a smooth function $f : \mathbb{C} \to \mathbb{C}$ defines a section of E if and only if it is odd, hence *all* such sections have a zero at the origin. Compare the two sections

$$f_0(x + iy) = x + iy, \qquad f_1(x + iy) = x^3 - x + iy.$$

They have qualitatively the same behavior near infinity, meaning in particular that they are homotopic through a family of sections whose zeroes are confined to some compact subset, thus we expect the algebraic count of zeroes to be the same for both. This is true if the count is defined by (12.12): we have $\#f_0^{-1}(0) = \#f_1^{-1}(0) = \frac{1}{2}$, in particular the negative zero of f_1 at the origin counts for $-1/2$ while the positive zero at $(1, 0) \sim (-1, 0)$ counts for 1. We see that the inclusion of the rational weights $\frac{1}{\kappa_z}$ is crucial for this result. Notice that if $H : [0, 1] \times M \to E$ is a homotopy of sections $f_\tau := H(\tau, \cdot)$ from f_0 to f_1, then $f_\tau(0) = 0$ for all τ, thus $\partial_\tau H(\tau, 0)$ vanishes and

$$D_2 H(\tau, 0) = df_\tau(0).$$

But $df_\tau(0)$ cannot be an isomorphism for all $\tau \in (0, 1)$ since $df_0(0)$ preserves orientation while $df_1(0)$ reverses it: there must exist a parameter $\tau_0 \in (0, 1)$ for which $df_{\tau_0}(0)$ is singular (see Figure 12.2), making $(\tau_0, 0)$ a critical zero of H. This is not

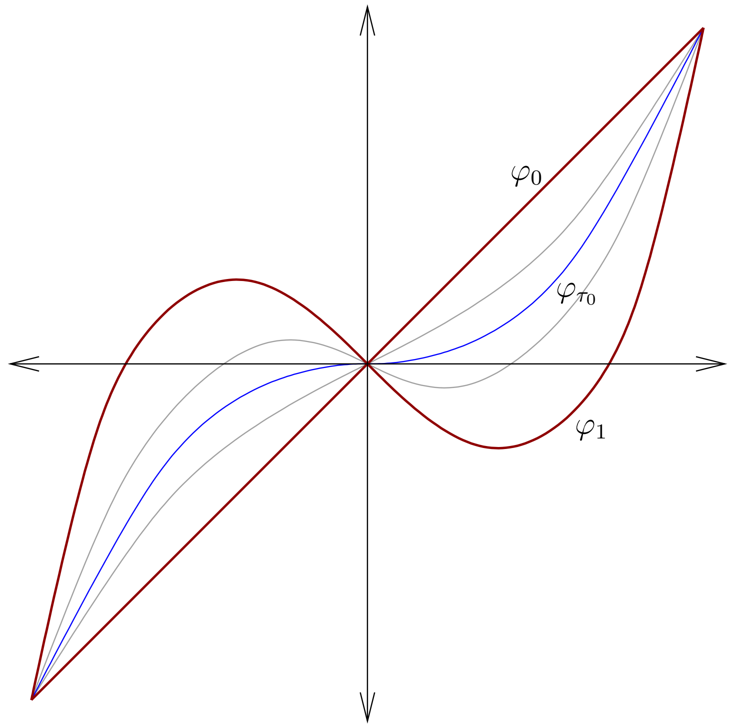

FIGURE 12.2. A family of odd functions $\varphi_\tau : \mathbb{R} \to \mathbb{R}$ defining a possible homotopy between the sections $f_0, f_1 \in \Gamma(E)$ in Example 12.13 via $f_\tau(x + iy) := \varphi_\tau(x) + iy$.

a problem that can be fixed by making H more generic—the homotopy will never be transverse to the zero-section, no matter what we do.

The need to address issues of the type raised by Examples 12.10 and 12.13 leads naturally to the notion of **multisections** as outlined in [**Sal99**, §5] and [**FO99**], and this is a major feature of the analysis under development by Fish-Hofer-Wysocki-Zehnder (see for example [**HWZ10**, **FH**]). In Example 12.13 for instance, one can consider functions

$$g : \mathbb{C} \to \mathrm{Sym}_2(\mathbb{C}) := (\mathbb{C} \times \mathbb{C})/(z_1, z_2) \sim (z_2, z_1),$$

which can be regarded as doubly-valued sections of $E \to M$ if g is $\mathbb{Z}_2$-equivariant for the antipodal action of $\mathbb{Z}_2$ on the symmetric product $\mathrm{Sym}_2(\mathbb{C})$. Such a section is considered single-valued at any point z where $g(z)$ is of the form $[(v, v)]$, so one can now imagine homotopies from $g_0(z) := [(f_0(z), f_0(z))]$ to $g_1(z) := [(f_1(z), f_1(z))]$ through doubly-valued sections $g_\tau : \mathbb{C} \to \mathrm{Sym}_2(\mathbb{C})$, the advantage being that g_τ is now allowed to take nonzero values of the form $[(v, -v)]$ at the origin. Indeed, if $f : \mathbb{C} \to \mathbb{C}$ is an odd function, then for any constant $c \in \mathbb{C}$, $g(z) := [(f(z) + c, f(z) - c)]$ defines an odd function $\mathbb{C} \to \mathrm{Sym}_2(\mathbb{C})$ and therefore also a well-defined multisection of E.

We can now construct a homotopy of multisections with better properties than the homotopy $f_\tau = H(\tau, \cdot)$ in Example 12.13. Assume the homotopy in that example is as depicted in Figure 12.2, and there is a subinterval $[\tau_-, \tau_+] \subset (0, 1)$ such that

$f_\tau = f_0$ for $\tau \leqslant \tau_-$ and $f_\tau = f_1$ for $\tau \geqslant \tau_+$. Choosing a smooth function $\beta : [0,1] \to [0,1]$ with compact support in $(0,1)$ such that $\beta|_{[\tau_-,\tau_+]} \equiv 1$, we define $g_\tau : \mathbb{C} \to \mathrm{Sym}_2(\mathbb{C})$ by

$$g_\tau(z) := [(f_\tau(z) + \beta(\tau), f_\tau(z) - \beta(\tau))].$$

The zero set of a multisection is defined to be the set of points at which *any* of its branches vanishes, e.g. for g_τ, we write $g_\tau^{-1}(0) := f_\tau^{-1}(\beta(\tau)) \cup f_\tau^{-1}(-\beta(\tau))$, and the zero set of the homotopy is

$$\widetilde{Z} := \left\{(\tau, z) \in [0,1] \times \mathbb{C} \;\middle|\; f_\tau(z) = \beta(\tau) \text{ or } f_\tau(z) = -\beta(\tau)\right\}.$$

We also associate to each point in $g_\tau^{-1}(0)$ a rational *weight*, in this example namely $1/2$ if the multisection has two branches but only one of them vanishes, and 1 if both do. This produces the picture in Figure 12.3 (left) of $\widetilde{Z}$ as a subset of $[0,1] \times \mathbb{C}$. What we really want however is to view the zero set as a subset of the orbifold $[0,1] \times M = [0,1] \times (\mathbb{C}/\mathbb{Z}_2)$, which gives the picture of $Z := \widetilde{Z}/\mathbb{Z}_2 \subset [0,1] \times M$ in Figure 12.3 (right). Here it is appropriate to adjust the rational weights by dividing by κ_z at any point with a nontrivial isotropy group, as e.g. some of the branching phenomena in $\widetilde{Z}$ no longer look like branching in Z, but are merely points at which the local isotropy of the ambient orbifold changes. Finally, we observe that the orbifold and orbibundle in this example both have natural orientations, thus so do the zero sets; these orientations are indicated with arrows in Figure 12.3, and they give rise to signs at each boundary point, which are shown multiplied by the corresponding rational weights. You will notice that the sum of these signed rational weights at ∂Z is zero.

What we have constructed in this example is known as a **weighted branched 1-manifold with boundary**. (For precise definitions of this notion, see [Sal99, McD06].) It leads to a general proof of the invariance of $\#f^{-1}(0) \in \mathbb{Q}$ for sections f of an oriented orbibundle over a closed oriented orbifold: given any two sections f_0 and f_1 with isolated zeroes, the zero set of a generic multivalued homotopy between them defines a compact oriented weighted branched 1-manifold that can be understood as an oriented cobordism between $f_0^{-1}(0)$ and $f_1^{-1}(0)$. One must then appeal to a certain straightforward generalization (and corollary) of the classification of 1-manifolds: if Z is a compact oriented weighted branched 1-manifold with boundary, then the algebraic count (with signs and rational weights) of points in ∂Z is zero. A complete proof of this statement may be found e.g. in [Sal99, Lemma 5.11]. Note that this depends crucially on orientations—in contrast to ordinary 1-manifolds, branched 1-manifolds need not be orientable in general, and there is no "mod 2 version" of the statement about the number of boundary points if orientations are ignored. For this reason, the coherent orientations from Lecture 11 are an indispensable ingredient in any version of SFT for which curves with nontrivial automorphisms cannot be excluded. The precise meaning of the formula

$$\#\partial\overline{\mathcal{M}}_1(J) = 0$$

should now be understood in these terms: after applying a sufficiently generic multivalued perturbation of the nonlinear Cauchy-Riemann equation, $\overline{\mathcal{M}}_1(J)$ carries the

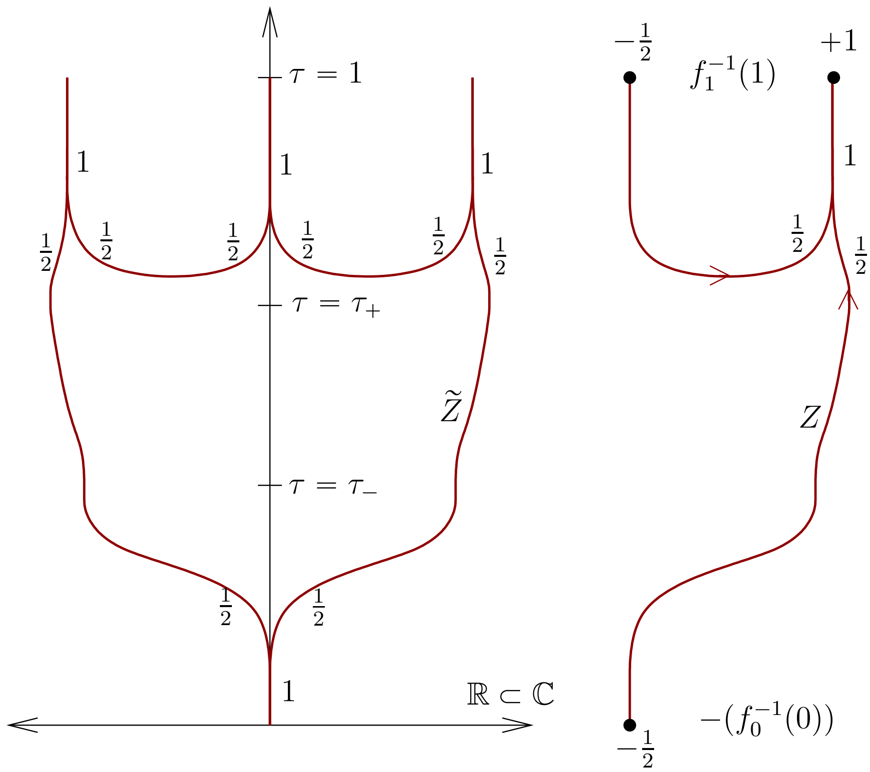

FIGURE 12.3. The zero set of a multivalued homotopy between two sections of the orbibundle in Example 12.13. The picture at the left shows $\widetilde{Z} \subset [0,1] \times \mathbb{C}$, while the picture at the right shows its quotient $Z \subset [0,1] \times M$, together with orientations and rational weights.

structure of a compact oriented weighted branched 1-manifold with boundary, hence the rational count of points in its boundary must vanish.

One can similarly use multivalued perturbations to define $\#f^{-1}(0)$ for sections of oriented orbibundles E over closed oriented orbifolds M on which transversality for single-valued sections is impossible. If $\operatorname{rank} E = \dim M$, the zero set of a generic multisection $f \in \Gamma(E)$ is an oriented weighted branched 0-manifold, which means a discrete set of points with signs and rational weights attached to them. As Figure 12.3 suggests, the sum of these signed rational numbers is the correct homotopy-invariant definition of $\#f^{-1}(0)$ and thus of the Euler number of such an orbibundle. In general it is a rational number, not an integer.

12.4.3. The inhomogeneous Cauchy-Riemann equation. Let us now briefly sketch how the rational count of index 1 curves

$$\# \left(\mathcal{M}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)/\mathbb{R} \right) \in \mathbb{Q} \tag{12.13}$$

appearing in the formula (12.8) for the SFT generating function $\mathbf{H}$ can be defined without making unrealistic assumptions about transversality. As of this writing, no definition with complete details has yet appeared in writing, though there exists a considerable body of literature on polyfolds meant to prepare the way for this definition; see [FH] and the references therein. For our discussion, we will avoid the difficult global aspects of this story and give only a local picture.

The functional-analytic setup from Lecture 7 identifies a neighborhood of any curve $[(\Sigma, j_0, \Gamma^+, \Gamma^-, \varnothing, u_0)] \in \mathcal{M}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ with a neighborhood of (j_0, u_0) in $\bar{\partial}_J^{-1}(0)/G$ for the nonlinear Cauchy-Riemann operator $\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$, where $\mathcal{B}^{k,p,\delta}$ is a Banach manifold of asymptotically cylindrical maps $\dot{\Sigma} \to \widehat{W}$ and $\mathcal{T} \subset \mathcal{J}(\Sigma)$ is a G-invariant Teichmüller slice through (j_0, Γ), with $G := \mathrm{Aut}(\Sigma, j_0, \Gamma)$ acting on $\mathcal{T} \times \mathcal{B}^{k,p,\delta}$ by $\psi \cdot (j, u) := (\psi^* j, u \circ \psi)$. In Lecture 8, we showed that taking generic perturbations of J makes $\bar{\partial}_J$ transverse to the zero-section along the open set of $(j, u) \in \bar{\partial}_J^{-1}(0)$ such that u is somewhere injective. The set of perturbations achieved in this way is, in fact, much smaller than the set of all possible perturbations of the section $\bar{\partial}_J : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$; in particular, all of them are G-equivariant. The latter is a condition we need to keep if we want a moduli space that is geometrically meaningful, but it might still make life easier if we perturb the section $\bar{\partial}_J$ directly instead of perturbing J. This leads to the idea of an *inhomogeneous* perturbation.

Ignoring equivariance for the moment, let us consider what happens if we replace the section $\bar{\partial}_J$ with

$$\bar{\partial}_{J,\nu} := \bar{\partial}_J - \nu$$

for a generic section $\nu : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$, so that instead of looking for pairs (j, u) such that $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is J-holomorphic, we are looking for solutions (j, u) of the inhomogeneous nonlinear Cauchy-Riemann equation

$$du + J(u) \circ du \circ j = \nu(j, u).$$

It is clear that $\bar{\partial}_J - \nu$ will be transverse to the zero-section for generic choices of the section ν, but in practice, we also need ν to satisfy a few more conditions which make this transversality result less obvious. One of those is that the linearization of $\bar{\partial}_J - \nu$ should be Fredholm on its zero set and have the same index as the linearization of the unperturbed section $\bar{\partial}_J$, so that the implicit function theorem can still be used to prove that $\bar{\partial}_J^{-1}(\nu)$ is a smooth manifold of the correct dimension. This will be true for instance if $D\bar{\partial}_{J,\nu}(j, u)$ differs from $D\bar{\partial}_J(j, u)$ only be a zeroth-order operator that decays to zero at infinity. To achieve this, let $\Xi \to \mathcal{T} \times \dot{\Sigma} \times \widehat{W}$ denote the vector bundle with fibers

$$\Xi_{(j,z,p)} := \overline{\mathrm{Hom}}_{\mathbb{C}}\big((T_z\dot{\Sigma}, j), (T_p\widehat{W}, J)\big),$$

choose a section $K \in \Gamma(\Xi)$ that is compactly supported with respect to $z \in \dot{\Sigma}$ and define the section $\nu : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$ by

$$\nu(j, u)(z) := K(j, z, u(z)).$$

The linearization of $\bar{\partial}_{J,\nu}$ at $(j, u) \in \bar{\partial}_{J,\nu}^{-1}(0)$ now takes the form

$$D\bar{\partial}_{J,\nu}(j, u)(y, \eta) = [J(u) \circ du \circ y - D_1 K(j, z, u)y] + \mathbf{D}_u^\nu \eta,$$

where

$$\mathbf{D}_u^\nu := \nabla\eta + J(u) \circ \nabla\eta \circ j + \nabla_\eta J \circ du \circ j - \nabla_\eta K$$

for any symmetric connection ∇ on $\widehat{W}$. The second term here is a Cauchy-Riemann type operator on $u^*T\widehat{W}$, and the compact support of K implies that $\mathbf{D}_u^\nu$ has the same asymptotic behavior as in the unperturbed case, so that the Fredholm condition still holds and the index is unchanged. If we also allow K to vary in a suitable Banach manifold of sections of Ξ, say the space $\Gamma_{\boldsymbol{\epsilon}}(\Xi)$ of $C_{\boldsymbol{\epsilon}}$-small perturbations of some chosen $K^{\text{ref}} \in \Gamma(\Xi)$, and differentiate it in some direction $(\eta, K') \in T_u\mathcal{B}^{k,p,\delta} \oplus T_K\Gamma_{\boldsymbol{\epsilon}}(\Xi)$, we obtain the operator

$$W^{k,p,\delta}(u^*T\widehat{W}) \oplus T_K\Gamma_{\boldsymbol{\epsilon}}(\Xi) \to W^{k-1,p,\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u^*T\widehat{W}))$$
$$(\eta, K') \mapsto \mathbf{D}_u^\nu\eta - K'(j, z, u),$$

which is easily shown to be surjective by an argument similar to the proof of Lemma 8.12. Indeed, for $k = 1$, there must otherwise exist a nontrivial $\theta \in L^{q,-\delta}(\overline{\operatorname{Hom}}_{\mathbb{C}}(T\dot{\Sigma}, u^*T\widehat{W}))$ for $\frac{1}{p} + \frac{1}{q} = 1$ such that

$$\langle \mathbf{D}_u^\nu\eta, \theta\rangle_{L^2} = 0 \text{ for all } \eta \in W^{1,p,\delta}(u^*T\widehat{W}),$$
$$\langle K'(j, z, u), \theta\rangle_{L^2} = 0 \text{ for all } K' \in T_K\Gamma_{\boldsymbol{\epsilon}}(\Xi).$$

The first relation implies that θ is a weak solution to a linear Cauchy-Riemann type equation, so it is smooth and its zeroes are isolated. But it is then easy to find a perturbation $K' \in T_K\Gamma_{\boldsymbol{\epsilon}}(\Xi)$ for which the pointwise inner product of $K'(j, z, u)$ with θ is positive in a neighborhood of some point z_0 where $\theta(z_0) \neq 0$, and vanishes outside such a neighborhood, giving a contradiction. The implicit function theorem now implies the smoothness of a certain universal moduli space to which one can apply the Sard-Smale theorem as in Lecture 8 and conclude that for generic $K \in \Gamma(\Xi)$, the section $\bar{\partial}_{J,\nu}$ is transverse to the zero-section everywhere.

The inhomogeneous perturbation gives us the freedom to avoid assuming that u is somewhere injective, but we still cheated in the above argument by ignoring the question of equivariance. For a generic $K \in \Gamma(\Xi)$, the resulting section $\nu : \mathcal{T} \times \mathcal{B}^{k,p,\delta} \to \mathcal{E}^{k-1,p,\delta}$ is typically *not* G-equivariant unless K satisfies the additional condition

$$K = \varphi^*K \text{ for all } \varphi \in G, \qquad \text{where} \qquad (\varphi^*K)(j, z, x) := K(\varphi_*j, \varphi(z), x) \circ d\varphi(z).$$

Requiring this in the transversality argument above would have killed the argument: it takes away our ability to choose K' at will near a specific point $z_0 \in \dot{\Sigma}$ without also changing the inner product of $K'(j, z, u)$ with θ in other regions. This should not be surprising, as it is just another example of the standard conflict between transversality and equivariance.

But in this form, the problem is easy to solve with a multivalued perturbation. Indeed, if G is finite,[4] we can for any given $K \in \Gamma(\Xi)$ define a finite set $\{K_1, \ldots, K_N\} \subset \Gamma(\Xi)$ as the G-orbit of K, i.e. the set of sections of the form $\varphi^*K \in \Gamma(\Xi)$ for $\varphi \in G$, and let $\{\nu_1, \ldots, \nu_N\}$ denote the corresponding set of sections of the Banach space bundle $\mathcal{E}^{k-1,p,\delta} \to \mathcal{T} \times \mathcal{B}^{k,p,\delta}$. After a generic perturbation of K, we can assume each of the perturbations $\bar{\partial}_{J,\nu_i}$ for $i = 1, \ldots, N$ is transverse to

[4]For an idea of what to do when $\dim G > 0$, see [Sal99, §5.3].

the zero-section, and while the individual sections $\bar\partial_{J,\nu_i}$ are not G-equivariant, they are equivariant as a collection, so that in particular their collective zero set

$$\bigcup_{i=1}^{N} \bar\partial_{J,\nu_i}^{-1}(0) \subset \mathcal{T} \times \mathcal{B}^{k,p,\delta}$$

is G-invariant. Associating to each element (j,u) in this set the weight $\lambda(j,u) := q/N$ where q is the number of elements $i \in \{1,\dots,N\}$ such that $\bar\partial_{J,\nu_i}(j,u) = 0$, the quotient of this zero set by G now acquires the structure of a finite-dimensional weighted branched orbifold.

We can incorporate asymptotic markers into this picture the same way as in Proposition 11.1: replace each of the finite-dimensional manifolds $\mathcal{M}_i := \bar\partial_{J,\nu_i}^{-1}(0)$ with a finite cover $\widetilde{\mathcal{M}}_i$ consisting of triples (j,u,ℓ), where $(j,u) \in \bar\partial_{J,\nu_i}^{-1}(0)$ and ℓ is a choice of asymptotic markers at every puncture. Fixing coherent orientations for the determinant line bundles of the Cauchy-Riemann type operators $\mathbf{D}_u^\nu$ now determines orientations of the manifolds $\widetilde{\mathcal{M}}_i$, so that the quotient of $\bigcup_{i=1}^N \widetilde{\mathcal{M}}_i$ by G becomes an oriented weighted branched orbifold. For the definition of the SFT generating function $\mathbf{H}$, we are only interested in spaces of virtual dimension 1 living in a symplectization $\mathbb{R} \times M$, in which case the orientation prevents orbifold singularities from appearing, and we are left with an oriented weighted branched 1-manifold, whose quotient by the $\mathbb{R}$-action is then an oriented weighted branched 0-manifold. The count (12.13) is now defined as the sum of the signed rational weights of all points in this space.

We have oversimplified several aspects of this discussion. One point worth emphasizing is that what we've given above is a purely local description of a moduli space of solutions to a perturbed equation—it depends on several choices such as a Teichmüller slice through j_0 and the parametrization of some curve $u_0 : (\dot\Sigma, j_0) \to (\widehat{W}, J)$. We have not given a *global* definition of this moduli space, which would be necessary in order to give a proper definition of the count of perturbed solutions. There are many subtle technical issues involved in formulating these notions globally so that the result can be shown to be independent of choices. For an introduction to this topic, see [**FH**].

REMARK 12.14. Multivalued perturbations are not the only possible approach to transversality problems in SFT, and easier methods are possible in some special situations. One well-known example is the rational Gromov-Witten invariants of semipositive symplectic manifolds: if one considers only closed holomorphic curves of genus $g = 0$ with $m \geqslant 3$ marked points, then equivariance can be eliminated from the picture because $\mathrm{Aut}(S^2, i, \Theta)$ is trivial whenever $\#\Theta \geqslant 3$. It then suffices to take single-valued inhomogeneous perturbations (as in [**RT95**]) or domain-dependent almost complex structures (as in [**MS12**]) to achieve transversality for all moduli spaces, and the resulting curve counts are not just rational numbers, they are integers. The same technique can also be used without the assumption $m \geqslant 3$, by identifying the invariants for $m \leqslant 2$ with invariants that count curves with 3 marked points, $3 - m$ of which are required to satisfy an incidence condition in which the evaluation map intersects a suitable cycle in the image. The curve count obtained in this way is, however, generally an overcount, as there may be multiple inequivalent ways to introduce auxiliary marked points satisfying the incidence condition; the count therefore needs to be divided by a combinatorial factor, producing a result that is generally in $\mathbb{Q}$ rather than $\mathbb{Z}$. The same thing happens in the more sophisticated approach by Cieliebak-Mohnke [CM07] (see also Gerstenberger [Ger13] for the higher-genus case), in which Donaldson hypersurfaces are used to define an incidence condition satisfied by a generally very large number of auxiliary marked points. This trick kills equivariance and thus eliminates the need for multivalued perturbations, but the invariants are still in $\mathbb{Q}$ due to the combinatorics of auxiliary marked points (see §12.7.3 for more on this). The Donaldson hypersurface technique has also been applied (with single-valued inhomogeneous perturbations) for punctured holomorphic curves in [CM18].

We will not discuss multisections or weighted branched manifolds any further, but the main takeaway from this discussion should be that defining an algebraic count of objects that respect symmetries can often be done, even when achieving transversality while respecting these symmetries is impossible, but the count one obtains in this way is generally rational, not an integer. For the rest of this lecture, we will return to the convenient fiction that transversality is always satisfied, but you may want to keep in mind that the objects we describe as 0- or 1-dimensional oriented manifolds are in general actually weighted branched manifolds.

12.5. Cylindrical contact homology revisited

Under an extra assumption on the complex $(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$, we can recover from it a more general version of the cylindrical contact homology we saw in Lecture 10. Suppose in particular that there are no index 1 holomorphic planes in $\mathbb{R} \times M$, so every term in $\hbar\mathbf{H}$ has at least one factor of either $\hbar$ or one of the q_γ variables. Then

$$\mathbf{D}_{\mathrm{SFT}} = \sum_{\gamma,\gamma',A} \kappa_\gamma \left(\sum_{u \in \mathcal{M}_{0,0}(J,A,\gamma,\gamma')/\mathbb{R}} \frac{\epsilon(u)}{|\operatorname{Aut}(u)|} e^A q_{\gamma'} \frac{\partial}{\partial q_\gamma} \right) + \ldots,$$

where the first sum is over all pairs of good Reeb orbits γ and γ', and the ellipsis is a sum of terms that all include at least a positive power of $\hbar$ or two q_γ variables or two partial derivatives. Let us abbreviate the spaces $\mathcal{M}_{0,0}(J,A,\gamma,\gamma')/\mathbb{R}$ of $\mathbb{R}$-equivalence classes of J-holomorphic cylinders by $\mathcal{M}_A(\gamma,\gamma')$, and notice that for any $u \in \mathcal{M}_A(\gamma,\gamma')$, the automorphism group is a cyclic group of order equal to the covering multiplicity

$$|\operatorname{Aut}(u)| = \kappa_u := \operatorname{cov}(u) \in \mathbb{N}.$$

Thus for any single generator q_γ, we have

$$\mathbf{D}_{\mathrm{SFT}} q_\gamma = \partial_{\mathrm{CCH}} q_\gamma + O(|q|^2, \hbar),$$

where

$$\partial_{\mathrm{CCH}} q_\gamma := \kappa_\gamma \sum_{\gamma',A} \left(\sum_{u \in \mathcal{M}_A(\gamma,\gamma')} \frac{\epsilon(u)}{\kappa_u} \right) e^A q_{\gamma'}. \tag{12.14}$$

The fact that $\mathbf{D}_{\mathrm{SFT}}^2 = 0$ thus implies

$$\partial_{\mathrm{CCH}}^2 = 0,$$

and the homology of the graded R-module generated by $\{q_\gamma \mid \gamma \text{ good}\}$ with differential ∂_{CCH} is an obvious generalization of the cylindrical contact homology from Lecture 10. What we saw there was a special case of this where the combinatorial factor κ_γ/κ_u did not appear because we were restricting to a homotopy class in which all orbits were simply covered, and all holomorphic cylinders were thus somewhere injective.

The presence of the factor κ_γ/κ_u deserves further comment. According to the above formula, we have

$$\partial_{\mathrm{CCH}}^2 q_\gamma = \sum_{\gamma',\gamma'',A,A'} \sum_{u \in \mathcal{M}_A(\gamma,\gamma')} \sum_{v \in \mathcal{M}_{A'}(\gamma',\gamma'')} e^{A+A'} \frac{\kappa_\gamma \kappa_{\gamma'} \epsilon(u)\epsilon(v)}{\kappa_u \kappa_v} q_{\gamma''},$$

hence $\partial_{\mathrm{CCH}}^2 = 0$ holds if and only if for all $A \in H_2(M)$ and all pairs of good orbits γ_+, γ_-,

$$\sum_{\gamma_0} \sum_{B+C=A} \left(\sum_{(u,v) \in \mathcal{M}_B(\gamma_+,\gamma_0) \times \mathcal{M}_C(\gamma_0,\gamma_-)} \frac{\kappa_{\gamma_0}}{\kappa_u \kappa_v} \epsilon(u)\epsilon(v) \right) = 0. \tag{12.15}$$

If γ_+ and γ_- happen to be simply covered orbits, then u and v in this expression always have trivial automorphism groups and it is clear what this sum means: every such pair $(u, v) \in \mathcal{M}_B(\gamma_+, \gamma_0) \times \mathcal{M}_C(\gamma_0, \gamma_-)$ corresponds to exactly κ_{γ_0} distinct holomorphic buildings obtained by different choices of decoration, so (12.15) is the count of boundary points of the compactified 1-dimensional manifold of index 2 cylinders $\mathcal{M}_A(\gamma_+, \gamma_-)/\mathbb{R}$. This sum skips over all bad orbits γ_0, but this is fine because whenever the breaking orbit is bad, there are evenly many choices of decoration such that half of these choices cancel the other half when counted with the correct signs (cf. Remark 12.5).

To understand why this formula is still correct in the presence of automorphisms, let us outline two equivalent approaches.

The easiest option is to instead consider moduli spaces with asymptotic markers, which never have automorphisms: removing unnecessary factors of κ_{γ_+} and κ_{γ_-} then transforms (12.15) into

$$\sum_{\gamma_0} \sum_{B+C=A} \frac{1}{\kappa_{\gamma_0}} \#\mathcal{M}_B^\$(\gamma_+, \gamma_0) \cdot \#\mathcal{M}_C^\$(\gamma_0, \gamma_-) = 0.$$

Now since each pair $(u, v) \in \mathcal{M}_B^\$(\gamma_+, \gamma_0) \times \mathcal{M}_C^\$(\gamma_0, \gamma_-)$ carries a canonical decoration and thus determines a holomorphic building, the division by κ_{γ_0} accounts for the fact that $\#\mathcal{M}_B^\$(\gamma_+, \gamma_0) \cdot \#\mathcal{M}_C^\$(\gamma_0, \gamma_-)$ overcounts the set of broken cylinders from γ_+ to γ_- with asymptotic markers at $\gamma_\pm$ by precisely this factor, as a simultaneous adjustment of the marker at γ_0 in both $u \in \mathcal{M}_B^\$(\gamma_+, \gamma_0)$ and $v \in \mathcal{M}_C^\$(\gamma_0, \gamma_-)$ produces the same decoration and therefore the same building.

The following alternative perspective will be more useful when we generalize beyond cylinders in the next section. We can directly count points in $\partial\overline{\mathcal{M}}_A(\gamma_+, \gamma_-)$, though as we saw in §12.4, rational weights should be included in the count whenever there is isotropy. Let us write

$$\mathcal{M}_A(\gamma_+,\gamma_-) = \mathcal{M}^\$_A(\gamma_+,\gamma_-)/G,$$

where $G \cong \mathbb{Z}_{\kappa_{\gamma_+}} \times \mathbb{Z}_{\kappa_{\gamma_-}}$ is a finite group acting by adjustment of the asymptotic markers. Since $\overline{\mathcal{M}}^\$_A(\gamma_+,\gamma_-)$ is a compact oriented 1-manifold with boundary under Assumption 12.1, the signed count of its boundary points is 0. We can ignore buildings broken along bad orbits in this count, since (by Remark 12.5) these always come in cancelling pairs. Let us now transform this into a count of buildings $(u|\Phi|v) \in \partial\overline{\mathcal{M}}_A(\gamma_+,\gamma_-)$ broken along good orbits γ_0: here $u \in \mathcal{M}_B(\gamma_+,\gamma_0)$ and $v \in \mathcal{M}_C(\gamma_0,\gamma_-)$ for some homology classes with $B + C = A$, and Φ is a decoration which describes how to glue the ends of u and v at γ_0. The automorphism group of such a building is the subgroup

$$\mathrm{Aut}(u|\Phi|v) \subset \mathrm{Aut}(u) \times \mathrm{Aut}(v)$$

consisting of all pairs $(\varphi,\psi) \in \mathrm{Aut}(u) \times \mathrm{Aut}(v)$ that define the same rotation at the two punctures asymptotic to γ_0; note that this group does not actually depend on the decoration Φ. Since we're talking about cylinders, we can be much more specific: we have $\mathrm{Aut}(u) = \mathbb{Z}_{\kappa_u}$ and $\mathrm{Aut}(v) = \mathbb{Z}_{\kappa_v}$, and if both are regarded as subgroups of $\mathrm{U}(1)$,

$$\mathrm{Aut}(u|\Phi|v) = \mathbb{Z}_{\kappa_u} \cap \mathbb{Z}_{\kappa_v} = \mathbb{Z}_{\gcd(\kappa_u,\kappa_v)},$$

which is injected into $\mathrm{Aut}(u) \times \mathrm{Aut}(v)$ by $\psi \mapsto (\psi,\psi)$. The boundary of $\overline{\mathcal{M}}^\$_A(\gamma_+,\gamma_-)$ can be understood likewise as a space of equivalence classes

$$[(u,v)] \in \left(\mathcal{M}^\$_B(\gamma_+,\gamma_0) \times \mathcal{M}^\$_C(\gamma_0,\gamma_-)\right) / \sim,$$

where two such pairs are equivalent if their asymptotic markers at the ends asymptotic to γ_0 determine the same decoration. Now observe that the group $G \cong \mathbb{Z}_{\kappa_{\gamma_+}} \times \mathbb{Z}_{\kappa_{\gamma_-}}$ also acts on buildings in $\partial\overline{\mathcal{M}}^\$_A(\gamma_+,\gamma_-)$, and the stabilizer of this action at (u,v) is $\mathrm{Aut}(u|\Phi|v)$, hence each $(u|\Phi|v) \in \partial\overline{\mathcal{M}}_A(\gamma_+,\gamma_-)$ gives rise to $\frac{|G|}{\gcd(\kappa_u,\kappa_v)}$ terms in the count of $\partial\overline{\mathcal{M}}^\$_A(\gamma_+,\gamma_-)$, implying

$$\sum_{(u|\Phi|v) \in \partial\overline{\mathcal{M}}_A(\gamma_+,\gamma_-)} \frac{\epsilon(u)\epsilon(v)}{\gcd(\kappa_u,\kappa_v)} = 0. \tag{12.16}$$

Finally, notice that while each pair $(u,v) \in \mathcal{M}_B(\gamma_+,\gamma_0) \times \mathcal{M}_C(\gamma_0,\gamma_-)$ determines buildings with κ_{γ_0} distinct choices of decoration, some of these buildings may be equivalent: every pair of automorphisms $(\varphi,\psi) \in \mathrm{Aut}(u) \times \mathrm{Aut}(v)$ transforms a building $(u|\Phi|v)$ by potentially changing the decoration Φ, thus producing an equivalent building. This action on buildings is trivial if and only if $(\varphi,\psi) \in \mathrm{Aut}(u|\Phi|v)$, hence every pair $(u,v) \in \mathcal{M}_B(\gamma_+,\gamma_0) \times \mathcal{M}_C(\gamma_0,\gamma_-)$ gives rise to exactly

$$\frac{\kappa_{\gamma_0}}{\left|(\mathrm{Aut}(u) \times \mathrm{Aut}(v)) / \mathrm{Aut}(u|\Phi|v)\right|} = \frac{\kappa_{\gamma_0} \gcd(\kappa_u,\kappa_v)}{\kappa_u \kappa_v}$$

elements of $\partial\overline{\mathcal{M}}_A(\gamma_+,\gamma_-)$, so that (12.16) becomes

$$\sum_{\gamma_0}\sum_{B+C=A}\left(\sum_{(u,v)\in\mathcal{M}_B(\gamma_+,\gamma_0)\times\mathcal{M}_C(\gamma_0,\gamma_-)}\frac{\epsilon(u)\epsilon(v)}{\gcd(\kappa_u,\kappa_v)}\frac{\kappa_{\gamma_0}\gcd(\kappa_u,\kappa_v)}{\kappa_u\kappa_v}\right)$$
$$=\sum_{\gamma_0}\sum_{B+C=A}\left(\sum_{(u,v)\in\mathcal{M}_B(\gamma_+,\gamma_0)\times\mathcal{M}_C(\gamma_0,\gamma_-)}\frac{\epsilon(u)\epsilon(v)\kappa_{\gamma_0}}{\kappa_u\kappa_v}\right)=0,$$

reproducing (12.15).

12.6. Combinatorics of gluing

Now let's try to justify the formula $\mathbf{H}^2=0$. The product of $\mathbf{H}$ with itself is the formal sum over all pairs of index 1 curves $u,v\in\mathcal{M}_1^\sigma(J)/\mathbb{R}$ of certain monomials: in particular if these two curves respectively have genus g_u and g_v, homology classes A_u and A_v, and asymptotic orbits $\boldsymbol{\gamma}_u^\pm$ and $\boldsymbol{\gamma}_v^\pm$, then the corresponding term in $\mathbf{H}^2$ is

$$\frac{\epsilon(u)\epsilon(v)}{|\operatorname{Aut}^\sigma(u)||\operatorname{Aut}^\sigma(v)|}\hbar^{g_u+g_v-2}e^{A_u+A_v}q^{\boldsymbol{\gamma}_u^-}p^{\boldsymbol{\gamma}_u^+}q^{\boldsymbol{\gamma}_v^-}p^{\boldsymbol{\gamma}_v^+}.$$

Before we can add up all monomials of this form, we need to put all the q and p variables in the same order: within each of the products $q^{\boldsymbol{\gamma}_u^-}$, $p^{\boldsymbol{\gamma}_u^+}$ and so forth this is simply a matter of permuting the variables and changing signs as appropriate, but the interesting part is the product $p^{\boldsymbol{\gamma}_u^+}q^{\boldsymbol{\gamma}_v^-}$, for which we can apply the commutation relations (12.9) to put all q variables before all p variables. Before discussing how this works in general, let us consider a more specific example.

Assume γ_i for $i=1,2$ are two specific orbits with $n-3+\mu_{\mathrm{CZ}}(\gamma_i)$ even, so the corresponding q and p variables have even degree, and suppose

$$\boldsymbol{\gamma}_u^+=(\gamma_1,\gamma_1,\gamma_2),\qquad \boldsymbol{\gamma}_v^-=(\gamma_1,\gamma_1).$$

After applying the relation $p_{\gamma_1}q_{\gamma_1}=q_{\gamma_1}p_{\gamma_1}+\kappa_{\gamma_1}\hbar$ a total of five times, one obtains the expansion

$$p_{\gamma_1}p_{\gamma_1}p_{\gamma_2}q_{\gamma_1}q_{\gamma_1}=q_{\gamma_1}^2p_{\gamma_1}^2p_{\gamma_2}+4\kappa_{\gamma_1}\hbar q_{\gamma_1}p_{\gamma_1}p_{\gamma_2}+2\kappa_{\gamma_1}^2\hbar^2p_{\gamma_2},$$

thus contributing a total of three terms to $\mathbf{H}^2$, namely the products of the factor $\frac{\epsilon(u)\epsilon(v)}{|\operatorname{Aut}(u)||\operatorname{Aut}(v)|}e^{A_u+A_v}$ with each of the expressions

$$\hbar^{g_u+g_v-2}q^{\boldsymbol{\gamma}_u^-}q_{\gamma_1}^2p_{\gamma_1}^2p_{\gamma_2}p^{\boldsymbol{\gamma}_v^+},\tag{12.17}$$

$$4\kappa_{\gamma_1}\hbar^{g_u+g_v-1}q^{\boldsymbol{\gamma}_u^-}q_{\gamma_1}p_{\gamma_1}p_{\gamma_2}p^{\boldsymbol{\gamma}_v^+},\tag{12.18}$$

$$2\kappa_{\gamma_1}^2\hbar^{g_u+g_v}q^{\boldsymbol{\gamma}_u^-}p_{\gamma_2}p^{\boldsymbol{\gamma}_v^+}.\tag{12.19}$$

As shown in Figure 12.4, this sum of three terms can be interpreted as the count of all possible holomorphic buildings obtained by gluing v on top of u together with a collection of trivial cylinders. Indeed, since $\boldsymbol{\gamma}_u^+$ and $\boldsymbol{\gamma}_v^-$ include two matching orbits (which also happen to be the same one), there are several choices to be made:

(1) The top-right picture shows what we might call the "stupid gluing," in which no ends of u are matched with any ends of v, but all are instead glued to trivial cylinders, thus producing a disconnected building. This possibility is encoded by (12.17), and we will see that in the total sum forming $\mathbf{H}^2$, this term gets cancelled out by a similar term for the stupid gluing of u on top of v.
(2) The lower-left picture shows the building obtained by gluing one end of u to an end of v along the matching orbit γ_1. This option is encoded by (12.18), where the factor $4\kappa_{\gamma_1}$ appears because there are precisely $4\kappa_{\gamma_1}$ distinct buildings of this type: indeed, there are four choices of which end of u should be glued to which end of v, and for each of these, a further κ_{γ_1} choices of the decoration. The arithmetic genus of the resulting building is $g_u + g_v$, as represented by the factor $\hbar^{g_u+g_v-1}$.
(3) The lower-right picture is encoded by (12.19): here there are two choices of bijections between the two pairs of punctures asymptotic to γ_1, and taking the choices of decoration at each breaking orbit into account, we obtain the combinatorial factor $2\kappa_{\gamma_1}^2$. The presence of two nontrivial breaking orbits increases the arithmetic genus to g_u+g_v+1, as encoded in the factor $\hbar^{g_u+g_v}$.

You may now be able to extrapolate from the above example why the commutator algebra we've defined encodes gluing of holomorphic curves in the symplectization and thus leads to the relation $\mathbf{H}^2 = 0$. Think of the algorithm by which you change $q^{\gamma_u^-}p^{\gamma_u^+}q^{\gamma_v^-}p^{\gamma_u^+}$ into a sum of products with all q's appearing before p's: for the first q you see appearing after a p, move it past each p for different orbits (changing signs as necessary) until it encounters a p for the *same* orbit. Now you replace $p_\gamma q_\gamma$ with $(-1)^{|p_\gamma||q_\gamma|}q_\gamma p_\gamma + \kappa_\gamma\hbar$, turning one product into a sum of two. This represents a choice between two options: either you move q_γ past p_γ and apply the usual sign change, or you eliminate them both but replace them with the combinatorial factor κ_γ and an extra $\hbar$. Then you continue this process until all q's appear before all p's.

The key point is that the process of gluing v on top of u in all possible ways is governed by *exactly the same algorithm*: first consider the disjoint union of the two curves as a single disconnected curve, with its punctures ordered in the same way in which their orbits appear in the monomial. Now reorder negative punctures of v and positive punctures of u, changing orientations as appropriate, until you see two such punctures next to each other approaching the same orbit γ. Here you have two options: either glue them together, or don't glue them but exchange their order. If you exchange the order, then you may again have to change orientations (depending on the parity of $n - 3 + \mu_{\mathrm{CZ}}(\gamma)$), but if you glue, then you have κ_γ distinct choices of decoration and will also increase the arithmetic genus of the eventual building by 1. In this way, every individual term in the final expansion of $q^{\gamma_u^-}p^{\gamma_u^+}q^{\gamma_v^-}p^{\gamma_u^+}$ represents a particular choice of which positive of ends of u should or should not be glued to which negative ends of v. Additional factors of $\hbar$ appear to keep track of the increase in arithmetic genus, and covering multiplicities of the breaking orbits also appear due to distinct choices of decorations. At the end these must still be divided by orders of automorphism groups in order to avoid counting equivalent

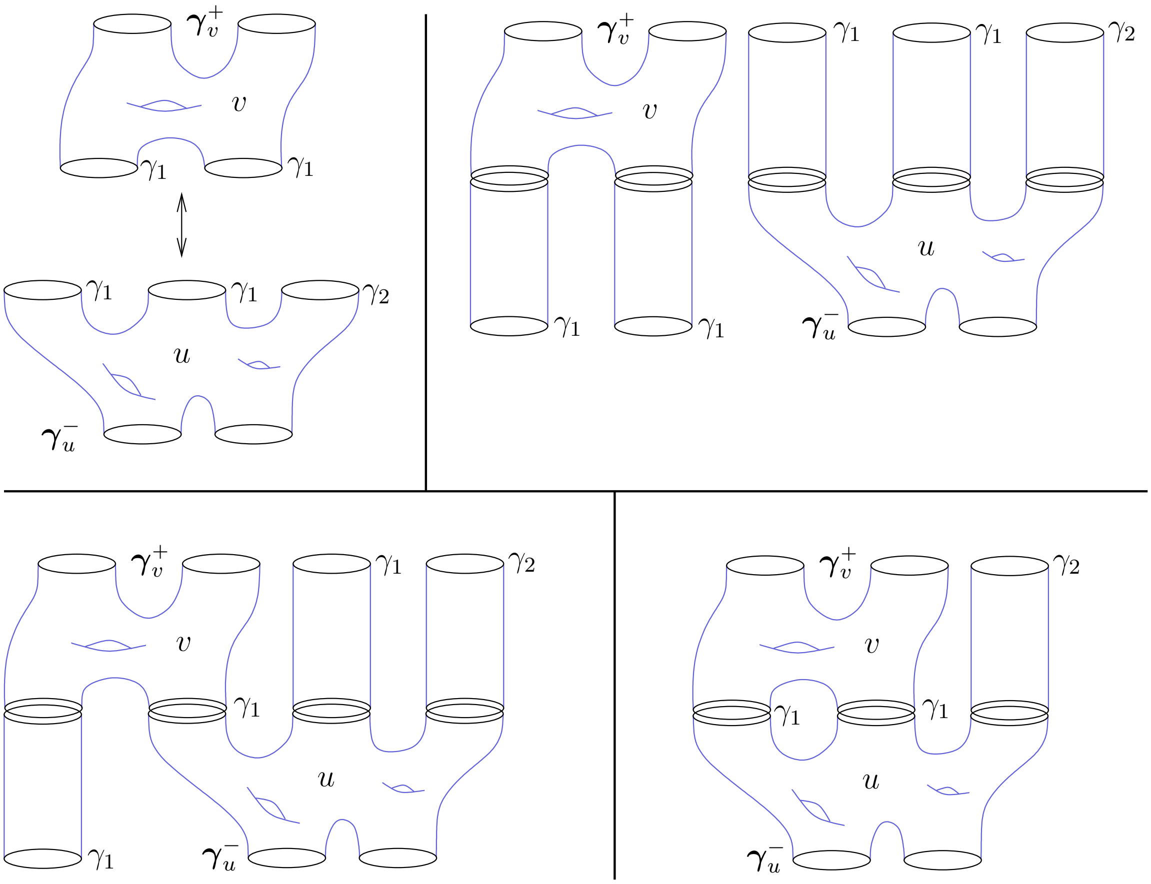

FIGURE 12.4. Three possible ways of gluing the curves u and v along with trivial cylinders to form index 2 curves.

buildings separately. Fleshing out these details leads to the following explanation for the relation $\mathbf{H}^2 = 0$:

PROPOSITION 12.15. *Let $\partial\overline{\mathcal{M}}_2^\sigma(J)$ denote the space of two-level holomorphic buildings*[5] *in $\overline{\mathcal{M}}(J)$ that have total index 2 and no bad asymptotic or breaking orbits, divided by the equivalence relation that forgets the order of the punctures. Then*

$$\mathbf{H}^2 = \sum_{\mathbf{u} \in \partial\overline{\mathcal{M}}_2^\sigma(J)} \frac{\epsilon(\mathbf{u})}{|\operatorname{Aut}^\sigma(\mathbf{u})|} \hbar^{g-1} e^A q^{\gamma^-} p^{\gamma^+},$$

where the terms in each monomial are determined by $\mathbf{u} \in \partial\overline{\mathcal{M}}_2^\sigma(J)$ as follows:

(1) *g is the arithmetic genus of $\mathbf{u}$;*
(2) *A is the equivalence class of $[\mathbf{u}] \in H_2(M)$ in $H_2(M)/G$;*
(3) *$\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \dots, \gamma_{k_\pm}^\pm)$ are the asymptotic orbits of $\mathbf{u}$ after arbitrarily fixing orderings of its positive and negative punctures;*

[5] Recall from §9.4.4 that for a holomorphic building in a symplectization, each level is defined only up to $\mathbb{R}$-translation. Under our present transversality assumptions, both levels of each building in $\partial\overline{\mathcal{M}}_2^\sigma(J)$ must have index 1, meaning each has a single index 1 component and possibly also some trivial cylinders, thus $\partial\overline{\mathcal{M}}_2^\sigma(J)$ is a discrete (and therefore finite) space.

(4) $\epsilon(\mathbf{u}) \in \{1, -1\}$ *is the boundary orientation at* $\mathbf{u}$ *determined by the chosen coherent orientations on* $\mathcal{M}^{\$}(J)$. *Specifically, given the chosen ordering of the punctures and an arbitrary choice of asymptotic markers at each puncture,* $\mathbf{u}$ *determines a boundary point of a* 1*-dimensional connected component of* $\overline{\mathcal{M}}^{\$}(J)$*, and we define* $\epsilon(\mathbf{u}) = +1$ *if and only if the orientation of* $\overline{\mathcal{M}}^{\$}(J)$ *at this point is outward.*

Once again $\epsilon(\mathbf{u})$ and $q^{\gamma^-}p^{\gamma^+}$ change signs in the same way under any reordering of the punctures, so their product is well defined, and there is no dependence on choices of markers since bad orbits have been excluded.

PROOF OF PROPOSITION 12.15. Our original formula for $\mathbf{H}$ gives rise to an expansion

$$\mathbf{H}^2 = \sum_{(u,v)\in\mathcal{M}_1^\sigma(J)/\mathbb{R}\times\mathcal{M}_1^\sigma(J)/\mathbb{R}} \frac{\epsilon(u)\epsilon(v)}{|\operatorname{Aut}^\sigma(u)||\operatorname{Aut}^\sigma(v)|}\hbar^{g_u+g_v-2}e^{A_u+A_v}q^{\gamma_u^-}p^{\gamma_u^+}q^{\gamma_v^-}p^{\gamma_v^+}.$$

As explained in the previous paragraph, the process of reordering $p^{\gamma_u^+}q^{\gamma_v^-}$ to put all q's before p's produces an expansion, each term of which can be identified with a specific choice of which positive punctures of u should be glued to which negative punctures of v. If k punctures are glued, then the resulting power of $\hbar$ is $g_u + g_v - 2 + k$, corresponding to the fact that the resulting building has arithmetic genus $g_u + g_v + k - 1$. We claim that the term for $k = 0$ is cancelled out by the corresponding term of $\mathbf{H}^2$ that has the roles of u and v reversed. To see this, imagine first the case where u and v have no asymptotic orbits in common, hence no nontrivial gluings are possible and all the q and p variables in the expression supercommute with each other. Then since both curves have index 1, the monomials $q^{\gamma_u^-}p^{\gamma_u^+}$ and $q^{\gamma_v^-}p^{\gamma_v^+}$ must both have odd degree, implying

$$q^{\gamma_u^-}p^{\gamma_u^+}q^{\gamma_v^-}p^{\gamma_v^+} = -q^{\gamma_v^-}p^{\gamma_v^+}q^{\gamma_u^-}p^{\gamma_u^+}$$

and thus the desired cancelation. If u and v do have orbits in common, then the result for the $k = 0$ terms is still not any different from this: all signs still change in the same way when applying $[p_\gamma, q_\gamma] = \kappa_\gamma\hbar$ to change $p_\gamma q_\gamma$ into $q_\gamma p_\gamma$, we simply ignore the extra term $\kappa_\gamma\hbar$ since it is only relevant for gluings with $k > 0$. This proves the claim, and consequently, that the expansion resulting from the curves u and v has no term containing $\hbar^{g_u+g_v-2}$.

The combinatorial factors can be explained as follows. The commutator expansion for $p^{\gamma_u^+}q^{\gamma_v^-}$ automatically produces combinatorial factors that count the different possible gluings, but if u and v have automorphisms, then not all of these give inequivalent buildings. This part of the discussion is a straightforward extension of what we did for cylindrical contact homology at the end of §12.5. Indeed, the actual set of inequivalent buildings is the quotient of this larger set by an action of

$$(\operatorname{Aut}^\sigma(u) \times \operatorname{Aut}^\sigma(v)) / \operatorname{Aut}^\sigma(\mathbf{u}),$$

where for a building $\mathbf{u}$ formed by endowing the pair (u, v) with decorations, $\operatorname{Aut}^\sigma(\mathbf{u})$ denotes the subgroup consisting of pairs $(\varphi, \psi) \in \operatorname{Aut}^\sigma(u) \times \operatorname{Aut}^\sigma(v)$ that preserve pairs of breaking punctures along with their decorations. This is what changes the factor $\frac{1}{|\operatorname{Aut}^\sigma(u)||\operatorname{Aut}^\sigma(v)|}$ into $\frac{1}{|\operatorname{Aut}^\sigma(\mathbf{u})|}$ as in the statement of the proposition. $\square$

The theorem that $\mathbf{H}^2 = 0$ now follows once you believe the propaganda from §12.4, arguing that $\sum_{\mathbf{u}\in\partial\overline{\mathcal{M}}_2^\sigma(J)} \frac{\epsilon(\mathbf{u})}{|\operatorname{Aut}^\sigma(\mathbf{u})|}$ is the correct way to count the boundary points of $\overline{\mathcal{M}}_2^\sigma(J)$. As we did with cylindrical contact homology, one can also use the obvious projection $\overline{\mathcal{M}}^\$(J) \to \overline{\mathcal{M}}^\sigma(J)$ to reduce this to the fact that if the 1-dimensional components of $\overline{\mathcal{M}}^\$(J)$ are manifolds (which is true if Assumption 12.1 holds), then the integer-valued signed count of their boundary points vanishes.

12.7. Some remarks on torsion, coefficients, and conventions

12.7.1. What if $H_1(M)$ has torsion? The main consequence for SFT if $H_1(M)$ has torsion is that one cannot define an integer grading, though there is always a canonical $\mathbb{Z}_2$-grading.[6] The setup in §12.2 must now be modified as follows. The **reference curves**

$$C_1, \dots, C_r \subset M$$

are required to form a basis of $H_1(M)$/torsion, so for every integral homology class $[\gamma]$, there is a unique collection of integers $m_1, \dots, m_r$ such that $[\gamma] = \sum_i m_i[C_i] \in H_1(M;\mathbb{Q})$. Instead of spanning surfaces for each orbit, one can define **spanning chains** C_γ, which are singular 2-chains with rational coefficients satisfying

$$\partial C_\gamma = \sum_i m_i[C_i] - [\gamma]$$

for the aforementioned set of integers $m_i \in \mathbb{Z}$. Note that C_γ must in general have non-integral coefficients since $\sum_i m_i[C_i]$ and $[\gamma]$ might not be homologous in $H_1(M;\mathbb{Z})$, so C_γ cannot always be represented by a smooth map of a surface. One consequence of this is that the absolute homology class associated to an asymptotically cylindrical holomorphic curve $u : \dot{\Sigma} \to \mathbb{R} \times M$ will now be rational,

$$[u] \in H_2(M;\mathbb{Q}),$$

and we must therefore take G to be a linear subspace

$$G \subset H_2(M;\mathbb{Q}).$$

Another consequence is that we cannot use capping chains to transfer trivializations from the reference curves to the orbits, so there is no natural way to define $\mu_{\mathrm{CZ}}(\gamma)$ as an integer. The easiest thing to do instead is to take the mod 2 Conley-Zehnder index

$$\mu_{\mathrm{CZ}}(\gamma) \in \mathbb{Z}_2$$

and define all degrees of generators as either even or odd with no further distinction. In particular, we now have

$$|q_\gamma| = n - 3 + \mu_{\mathrm{CZ}}(\gamma) \in \mathbb{Z}_2, \qquad |p_\gamma| = n - 3 - \mu_{\mathrm{CZ}}(\gamma) \in \mathbb{Z}_2,$$

[6] In fact there is a bit more than a $\mathbb{Z}_2$-grading, see [EGH00, §2.9.1].

while $\hbar$ and all elements of $R = \mathbb{Q}[H_2(M;\mathbb{Q})/G]$ are even. With these modifications, the rest of the discussion also becomes valid for the case where $H_1(M)$ has torsion, and leads to $\mathbb{Z}_2$-graded contact invariants.

12.7.2. Combinatorial conventions. The combinatorial factors appearing in our definition of $\mathbf{H}$ may at first look slightly different from what appears elsewhere in the literature. Actually, most papers seem to agree on this detail, but various subtle differences and ambiguities in notation mean that it sometimes requires intense concentration to recognize this fact.

The original propaganda paper [EGH00] expresses everything in terms of moduli spaces with asymptotic markers, and the formula for $\mathbf{H}$ in §2.2.3 of that paper (which is expressed in a slightly more general form involving marked points) agrees with our (12.8).

Cieliebak and Latschev [CL09, §2] write down the same formula in terms of moduli spaces that have no asymptotic markers but remember the order of the punctures, thus it includes some factorials that do not appear in (12.6) but is missing the κ_γ terms of (12.8). The notation $n_g(\Gamma^-, \Gamma^+)$ used in [CL09] for curve counts must be understood implicitly to include rational weights arising from automorphisms (or multivalued perturbations, as the case may be).

My paper with Latschev [LW11] uses moduli spaces with asymptotic markers and attempts to write down the same formula as in [EGH00, CL09], but gets it slightly wrong due to some missing κ_γ terms that should appear in front of each $\frac{\partial}{\partial q_\gamma}$. Mea culpa.

For cylindrical contact homology, the combinatorial factors in §12.5 also agree with what appears in [Bou03]. As observed by Nelson [Nel13, Remark 8.3], there are other conventions for ∂_{CCH} that appear in the literature and lead to equivalent theories: in particular it is possible to replace (12.14) with

$$\partial_{\mathrm{CCH}} q_\gamma := \sum_{\gamma', A} \kappa_{\gamma'} \left(\sum_{u \in \mathcal{M}_A(\gamma,\gamma')} \frac{\epsilon(u)}{\kappa_u} \right) e^A q_{\gamma'}.$$

One can derive this from the same definition of $\mathbf{H}$ by applying a "change of coordinates" to the algebra $\mathcal{A}[[\hbar]]$, or equivalently, by choosing a slightly different representation of the operator algebra defined by the p_γ and q_γ variables. To avoid confusion, let us write the generators of $\mathcal{A}$ as x_γ instead of q_γ, and then define the operators q_γ and p_γ on $\mathcal{A}[[\hbar]]$ by

$$q_\gamma = \kappa_\gamma x_\gamma, \qquad p_\gamma = \hbar \frac{\partial}{\partial x_\gamma}.$$

These operators still satisfy $[p_\gamma, q_\gamma] = \kappa_\gamma \hbar$ and thus define an equivalent theory, but the resulting differential operator $\mathbf{D}_{\mathrm{SFT}}$ on $\mathcal{A}[[\hbar]]$ now includes extra factors of κ_γ for the negative punctures instead of the positive punctures.

12.7.3. Coefficients: $\mathbb{Q}$, $\mathbb{Z}$ or $\mathbb{Z}_2$? While we were able to use $\mathbb{Z}_2$ coefficients for cylindrical contact homology in a primitive homotopy class in Lecture 10, a quick glance at any version of the formula for $\mathbf{H}$ should make the reader very skeptical about doing this for more general versions of SFT. The existence of curves with automorphisms means that $\mathbf{H}$ always contains terms with rational (but nonintegral) coefficients. And this is only what is true in the fictional world of Assumption 12.1: in the general version of the theory, we expect to have to replace expressions like $\sum_u \frac{\epsilon(u)}{|\mathrm{Aut}(u)|}$ with counts of 0-dimensional branched manifolds with rational weights, arising as zero sets of generic multisections.

As mentioned in Remark 12.14, there are various tricks available for avoiding multivalued perturbations, but these typically also produce rational counts for combinatorial reasons. For instance, in the approach of Cieliebak-Mohnke [**CM07**] for the rational Gromov-Witten invariants of a closed symplectic manifold (W^{2n}, ω) with $[\omega] \in H^2(W; \mathbb{Q})$, the invariants are defined by replacing the usual moduli space $\mathcal{M}_{0,m}(J, A)$ by a space $\mathcal{M}_{0,m+N}(J, A; Y)$ consisting of J-holomorphic spheres $u : S^2 \to W$ with some large number of auxiliary marked points $\zeta_1, \ldots, \zeta_N$ required to satisfy the condition

$$u(\zeta_i) \in Y, \qquad i = 1, \ldots, N.$$

Here $Y^{2n-2} \subset W^{2n}$ is a J-holomorphic hypersurface with $[Y] = D \cdot \mathrm{PD}([\omega]) \in H_{2n-2}(W)$ for some degree $D \in \mathbb{N}$, and the number of extra marked points is determined by

$$N = A \cdot [Y] = D\langle [\omega], A \rangle,$$

so positivity of intersections implies that u *only* intersects Y at the auxiliary marked points. These auxiliary points are convenient for technical reasons involving transversality—their role is vaguely analogous to the way that asymptotic markers get rid of isotropy in SFT—but they are not geometrically meaningful, as we'd actually prefer to count curves in $\mathcal{M}_{0,m}(J, A)$. Every such curve has N intersections with Y, so accounting for permutations, it lifts to $N!$ distinct elements of $\mathcal{M}_{0,m+N}(J, A; Y)$, and the correct count is therefore obtained as an integer count of curves in the latter space divided by $N!$. Perturbing to achieve transversality breaks the symmetry, however, so there is no guarantee that counting curves in $\mathcal{M}_{0,m+N}(J, A; Y)$ will produce a multiple of $N!$, and moreover, N could be arbitrarily large since one needs to take hypersurfaces of arbitrarily large degree in order to show that the invariants don't depend on this choice. For these reasons, the resulting Gromov-Witten invariants are rational numbers rather than integers, and their denominators cannot be predicted or bounded.

The upshot of this discussion is that there is probably no hope of defining SFT with integer coefficients in general, much less with $\mathbb{Z}_2$ coefficients, and the inclusion of orientations in the picture is unavoidable.

The good news however is that whenever formulas like $\sum_u \frac{\epsilon(u)}{|\mathrm{Aut}(u)|}$ can be taken literally as a count of curves, the chain complex $(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$ can in fact be defined with $\mathbb{Z}$ coefficients, and one can even reduce to a $\mathbb{Z}_2$ version in order to ignore signs. A special case of this was observed for cylindrical contact homology in [**Nel15**, Remark 1.5], and you may notice it already when you look at the formula (12.14) for ∂_{CCH}: the factor κ_γ/κ_u is always an integer since the multiplicity of a holomorphic cylinder always divides the covering multiplicity of both its asymptotic orbits. Surprisingly, something similar turns out to be true for the much larger chain complex of SFT. The following result is only a statement about a chain complex and makes no claims about any chain maps between complexes defined via different choices of α and J—one should therefore not expect it to imply anything interesting about actual contact invariants, but it is included here mainly for the author's amusement.

PROPOSITION 12.16. *If Assumption 12.1 holds then the rational coefficients $n_g(\boldsymbol{\gamma}, \boldsymbol{\gamma}', k)$ in the formula (12.11) for $\mathbf{D}_{\mathrm{SFT}} q^{\boldsymbol{\gamma}}$ are all integers.*

COROLLARY 12.17. *Under Assumption 12.1, there exist well-defined chain complexes*

$$(\mathcal{A}_{\mathbb{Z}}[[\hbar]], \mathbf{D}_{\mathrm{SFT}}) \quad \text{and} \quad (\mathcal{A}_{\mathbb{Z}_2}[[\hbar]], \mathbf{D}_{\mathrm{SFT}}),$$

where for a general commutative ring $\mathscr{R}$, $\mathcal{A}_{\mathscr{R}}$ denotes the graded supercommutative unital algebra over $\mathscr{R}[H_2(M)/G]$ generated by the q_γ variables for good Reeb orbits γ. The differentials $\mathbf{D}_{\mathrm{SFT}}$ on $\mathcal{A}_{\mathbb{Z}}[[\hbar]]$ and $\mathcal{A}_{\mathbb{Z}_2}[[\hbar]]$ are defined by the same formula as on $\mathcal{A}[[\hbar]]$, where in the $\mathbb{Z}_2$ case we are free to set all signs $\epsilon(u)$ equal to 1.

PROOF OF PROPOSITION 12.16. We need to show that expressions of the form

$$\frac{\kappa_{\boldsymbol{\gamma}^+}}{|\operatorname{Aut}^\sigma(u)|} \frac{\partial}{\partial q_{\gamma_1^+}} \cdots \frac{\partial}{\partial q_{\gamma_{k_+}^+}} q^{\boldsymbol{\gamma}}$$

produce integer coefficients for every holomorphic curve u with asymptotic orbits $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \ldots, \gamma_{k_\pm}^\pm)$ and every tuple $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_m)$. It suffices to consider the special case $\boldsymbol{\gamma} = \boldsymbol{\gamma}^+$, as the derivative in question is only nonzero on monomials that are divisible by $q^{\boldsymbol{\gamma}^+}$. Up to a sign change, we can reorder the orbits and write $\boldsymbol{\gamma}^+$ in the form

$$\boldsymbol{\gamma}^+ = (\underbrace{\gamma_1, \ldots, \gamma_1}_{m_1}, \ldots, \underbrace{\gamma_N, \ldots, \gamma_N}_{m_N})$$

for some finite set of distinct orbits $\gamma_1, \ldots, \gamma_N$ and numbers $m_i \in \mathbb{N}$, $i = 1, \ldots, N$. We then have

(12.20)
$$\begin{aligned}\frac{\kappa_{\boldsymbol{\gamma}^+}}{|\operatorname{Aut}^\sigma(u)|} \frac{\partial}{\partial q_{\gamma_1^+}} \cdots \frac{\partial}{\partial q_{\gamma_{k_+}^+}} q^{\boldsymbol{\gamma}^+} &= \frac{\kappa_{\gamma_1}^{m_1} \cdots \kappa_{\gamma_N}^{m_N}}{|\operatorname{Aut}^\sigma(u)|} \left(\frac{\partial}{\partial q_{\gamma_1}}\right)^{m_1} \cdots \left(\frac{\partial}{\partial q_{\gamma_N}}\right)^{m_N} \left(q_{\gamma_1}^{m_1} \cdots q_{\gamma_N}^{m_N}\right) \\ &= \pm \frac{\kappa_{\gamma_1}^{m_1} \cdots \kappa_{\gamma_N}^{m_N} m_1! \ldots m_N!}{|\operatorname{Aut}^\sigma(u)|}.\end{aligned}$$

We claim that this number is always an integer. Indeed, if $\operatorname{Aut}^\sigma(u)$ is nontrivial, then $u : \dot{\Sigma} \to \mathbb{R} \times M$ is a multiple cover $u = v \circ \varphi$ for some holomorphic branched cover $\varphi : (\Sigma, j) \to (\Sigma', j')$ and somewhere injective curve $v : (\dot{\Sigma}' = \Sigma' \backslash \Gamma', j') \to (\mathbb{R} \times M, J)$. Automorphisms $\psi \in \operatorname{Aut}^\sigma(u)$ thus define biholomorphic maps on (Σ, j) that permute each of the sets of punctures asymptotic to the same orbit. Given any puncture $z \in \Gamma$ where u is asymptotic to γ_i, the $\operatorname{Aut}^\sigma(u)$-orbit of z consists of $\ell \leqslant m_i$ other punctures also asymptotic to γ_i, and its stabilizer is a cyclic subgroup of order $k = |\operatorname{Aut}^\sigma(u)|/\ell$, acting on a neighborhood of z by biholomorphic rotations. It follows that κ_{γ_i} is divisible by k, hence

$$\frac{\kappa_{\gamma_i} \ell}{|\operatorname{Aut}^\sigma(u)|} \in \mathbb{N},$$

and (12.20) is a multiple of this. □

REMARK 12.18. Since $1 = -1$ in $\mathcal{A}_{\mathbb{Z}_2}$, anticommuting elements of $\mathcal{A}_{\mathbb{Z}_2}[[\hbar]]$ actually commute, so unless one imposes extra algebraic conditions in the case of $\mathbb{Z}_2$ coefficients, higher powers of odd generators p_γ and q_γ do not vanish. Nonetheless, these powers still do not appear in $\mathbf{H}$, so the complex $(\mathcal{A}_{\mathbb{Z}_2}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$ ignores curves with multiple ends approaching an orbit of odd degree (and also bad orbits, for that matter).

LECTURE 13

Contact invariants

Contents

In the previous lecture, we introduced an operator algebra defined via the supercommutators $[p_\gamma, q_\gamma] = \kappa_\gamma \hbar$, then we defined the SFT generating function

$$\mathbf{H} = \sum_{u \in \mathcal{M}_1^\sigma(J)/\mathbb{R}} \frac{\epsilon(u)}{|\operatorname{Aut}^\sigma(u)|} \hbar^{g-1} e^A q^{\gamma^-} p^{\gamma^+}$$

and proved (modulo transversality) that $\mathbf{H}^2 = 0$. The generating function is a formal power series whose coefficients are rational counts of holomorphic curves, and these counts are strongly dependent on the choices of contact form α, almost complex structure $J \in \mathcal{J}(\alpha)$ and further auxiliary data such as coherent orientations. Thus in contrast to Gromov-Witten theory, the generating function does not define an invariant, but one can follow the standard prescription of Floer-type theories and define invariants via homology. We saw that for the natural representation $\mathcal{A}[[\hbar]]$ of the operator algebra defined by setting $p_\gamma = \kappa_\gamma \hbar \frac{\partial}{\partial q_\gamma}$, $\mathbf{H}$ defines a differential operator $\mathbf{D}_{\mathrm{SFT}} : \mathcal{A}[[\hbar]] \to \mathcal{A}[[\hbar]]$ with $\mathbf{D}_{\mathrm{SFT}}^2 = 0$. One of our goals in this lecture will be to explain (again modulo transversality) why the resulting homology

$$H_*^{\mathrm{SFT}}(M, \xi; R) = H_*(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$$

is an invariant of the contact structure. We will then use it to define simpler numerical invariants that detect symplectic fillability properties of contact manifolds.

But first, $\mathcal{A}[[\hbar]]$ is not the only possible representation of the operator algebra of SFT: other choices lead to different invariants with different algebraic structures. Let's begin by describing the original hierarchy of contact invariants that were outlined in [EGH00].

REMARK 13.1. We will continue in this lecture under the convenient fiction of Assumption 12.1 that choosing J generically suffices to achieve transversality for all holomorphic curves, with only occasional remarks on what needs to be modified in the cold hard reality where multiple covers cannot be ignored. All theorems stated in this lecture should therefore be understood with the caveat that they refer to objects whose complete definitions remain work in progress (see e.g. [FH]).

We will also continue to assume for simplicity that $H_1(M)$ has no torsion, and the same assumption is made about cobordisms in §13.2.2. Only minor changes are necessary if this condition is lifted, e.g. one could then replace all instances of $H_1(M)$ with $H_1(M;\mathbb{Q})$ and assume always that the grading is $\mathbb{Z}_2$; see §12.7.1.

13.1. The Eliashberg-Givental-Hofer package

In the following, (M,ξ) is a $(2n-1)$-dimensional closed contact manifold with a contact form α and almost complex structure $J \in \mathcal{J}(\alpha)$ for which the usual optimistic transversality condition (Assumption 12.1) is assumed to hold. We fix also the auxiliary data described in §12.2, plus a choice of subgroup $G \subset H_2(M)$ which determines the coefficient ring

$$R = \mathbb{Q}[H_2(M)/G].$$

Each of the differential graded algebras described below then carries the same grading that was described in that lecture, i.e. there is always at least a $\mathbb{Z}_2$-grading, and it lifts to $\mathbb{Z}$ if $H_1(M)$ is torsion free and $c_1(\xi)|_G = 0$, or possibly $\mathbb{Z}_{2N}$ if $N \in \mathbb{N}$ is the smallest possible value for $c_1(A)$ with $A \in G$.

13.1.1. Full SFT as a Weyl superalgebra.

We start with some seemingly trivial algebraic observations. First, the relation $\mathbf{H}^2 = 0$ is equivalent to

$$[\mathbf{H},\mathbf{H}] = 0.$$

Remember that $[\ ,\ ]$ is a *super*-commutator, so $[\mathbf{F},\mathbf{F}] = 0$ holds automatically for operators $\mathbf{F}$ with even degree, but $\mathbf{H}$ is odd, and for odd operators the commutator is defined by $[\mathbf{F},\mathbf{G}] = \mathbf{FG} + \mathbf{GF}$, hence $[\mathbf{H},\mathbf{H}] = 2\mathbf{H}^2$. Formally speaking $[\ ,\ ]$ is a **super Lie bracket** and thus satisfies the "super Jacobi identity":

$$\big[\mathbf{F},[\mathbf{G},\mathbf{K}]\big] + (-1)^{|\mathbf{F}||\mathbf{G}|+|\mathbf{F}||\mathbf{K}|}\big[\mathbf{G},[\mathbf{K},\mathbf{F}]\big] + (-1)^{|\mathbf{F}||\mathbf{K}|+|\mathbf{G}||\mathbf{K}|}\big[\mathbf{K},[\mathbf{F},\mathbf{G}]\big] = 0. \tag{13.1}$$

In order to create a homology theory out of $\mathbf{H}$, we don't absolutely need to find a representation of the entire operator algebra: it suffices to find a representation of the induced super Lie algebra. Indeed, suppose V is a graded $R[[\hbar]]$-module and L is a linear grading-preserving map that associates to operators $\mathbf{F}$ (expressed as power series functions of p's, q's and $\hbar$ with coefficients in R) an $R[[\hbar]]$-linear map

$$L_{\mathbf{F}} : V \to V$$

such that

$$L_{[\mathbf{F},\mathbf{G}]} = L_{\mathbf{F}}L_{\mathbf{G}} - (-1)^{|\mathbf{F}||\mathbf{G}|}L_{\mathbf{G}}L_{\mathbf{F}}$$

for every pair of operators $\mathbf{F},\mathbf{G}$. Then the $R[[\hbar]]$-linear map $L_{\mathbf{H}} : V \to V$ satisfies

$$L_{\mathbf{H}}^2 = \frac{1}{2}[L_{\mathbf{H}},L_{\mathbf{H}}] = \frac{1}{2}L_{[\mathbf{H},\mathbf{H}]} = 0,$$

hence $(V, L_{\mathbf{H}})$ is a chain complex. The complex $(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$ was a special case of this, in which we represented the super Lie algebra via a faithful representation of the whole operator algebra.

EXERCISE 13.2. Verify (13.1).

REMARK 13.3 (supersymmetric sign rules). To see where the signs in (13.1) come from, it suffices to know the following basic rule of superalgebra: for any pair of $\mathbb{Z}_2$-graded vector spaces V and W, the natural "commutation" isomorphism $c : V \otimes W \to W \otimes V$ is defined on homogeneous elements by

$$c(v \otimes w) = (-1)^{|v||w|} w \otimes v.$$

For any permutation of a finite tuple of $\mathbb{Z}_2$-graded vector spaces, one can derive the appropriate isomorphism from this: in particular the cyclic permutation isomorphism $\sigma : X \otimes Y \otimes Z \to Y \otimes Z \otimes X$ takes the form

$$\sigma = (\mathbb{1} \otimes c_{23}) \circ (c_{12} \otimes \mathbb{1}) : x \otimes y \otimes z \mapsto (-1)^{|x||y|+|x||z|} y \otimes z \otimes x.$$

Writing the Jacobi identity as $[\cdot, [\cdot, \cdot]] \circ (\mathbb{1} + \sigma + \sigma^2) = 0$ then produces (13.1). In this sense, it only differs from the usual Jacobi identity in being based on a different definition of the commutation isomorphism $V \otimes W \to W \otimes V$. For more on this perspective, see [**Var04**, §3.1].

Now here is a different kind of example, where the representation does not respect the product structure of the operator algebra but does respect its Lie bracket. Let $\mathfrak{W}$ denote the graded unital algebra consisting of formal power series

$$\sum_{\boldsymbol{\gamma}, k} f_{\boldsymbol{\gamma}, k}(q) \hbar^k p^{\boldsymbol{\gamma}},$$

where the sum ranges over all integers $k \geqslant 0$ and all ordered sets $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_m)$ of good Reeb orbits for $m \geqslant 0$, and the $f_{\boldsymbol{\gamma}, k}$ are polynomial functions of the q_γ variables, with coefficients in R. Note that the case of the empty set of orbits is included here, which means $p^{\boldsymbol{\gamma}} = 1$. The multiplicative structure of $\mathfrak{W}$ is defined via the usual (super)commutation relations, and its elements can be interpreted as operators. If we now associate to each $\mathbf{F} \in \mathfrak{W}$ the $R[[\hbar]]$-linear map

$$D_{\mathbf{F}} : \mathfrak{W} \to \mathfrak{W} : \mathbf{G} \mapsto [\mathbf{F}, \mathbf{G}],$$

then the Jacobi identity (13.1) implies

$$D_{[\mathbf{F}, \mathbf{G}]} = D_{\mathbf{F}} D_{\mathbf{G}} - (-1)^{|\mathbf{F}||\mathbf{G}|} D_{\mathbf{G}} D_{\mathbf{F}}.$$

This is just the graded version of the standard *adjoint representation* of a Lie algebra. The only problem in applying this idea to define a differential

$$D_{\mathbf{H}} : \mathfrak{W} \to \mathfrak{W} : \mathbf{F} \mapsto [\mathbf{H}, \mathbf{F}] \tag{13.2}$$

is that $\mathbf{H}$ is not technically an element of $\mathfrak{W}$: indeed, $\mathbf{H}$ contains terms of order -1 in $\hbar$, thus

$$\mathbf{H} \in \frac{1}{\hbar} \mathfrak{W}.$$

On the other hand, the failure of supercommutativity in $\mathfrak{W}$ is a "phenomenon of order $\hbar$," i.e. since every nontrivial commutator contains a factor of $\hbar$, we have

$$[\mathbf{F}, \mathbf{G}] = \mathcal{O}(\hbar) \quad \text{for all} \quad \mathbf{F}, \mathbf{G} \in \mathfrak{W}.$$

Here and in the following we use the symbol

$$\mathcal{O}(\hbar^k)$$

to denote any element of the form $\hbar^k \mathbf{F}$ for $\mathbf{F} \in \mathfrak{W}$. As a consequence, $[\mathbf{H}, \mathbf{F}] \in \mathfrak{W}$ whenever $\mathbf{F} \in \mathfrak{W}$, hence (13.2) is well defined, and the Jacobi identity now implies

$$D_{\mathbf{H}}^2 = 0.$$

The homology of the resulting chain complex gives another version of what is often called **full SFT**,

$$H_*^{\mathfrak{W}}(M, \xi; R) := H_*(\mathfrak{W}, D_{\mathbf{H}}).$$

A proof (modulo transversality) that this defines a contact invariant is outlined in [**EGH00**, §2], but it is algebraically somewhat more involved than for $H_*^{\mathrm{SFT}}(M, \xi; R)$, so I will skip it since I don't have any applications of $H_*^{\mathfrak{W}}(M, \xi; R)$ in mind. As far as I am aware, no contact-topological applications of this invariant or computations of it (outside the trivial case—see §13.1.4 below) have yet appeared in the literature. This is a pity, because $H_*^{\mathfrak{W}}(M, \xi; R)$ actually has much more algebraic structure than $H_*^{\mathrm{SFT}}(M, \xi; R)$. Indeed, using the identities

$$\begin{aligned} [\mathbf{F}, \mathbf{G}\mathbf{K}] &= [\mathbf{F}, \mathbf{G}]\mathbf{K} + (-1)^{|\mathbf{F}||\mathbf{G}|}\mathbf{G}[\mathbf{F}, \mathbf{K}], \\ [\mathbf{F}\mathbf{G}, \mathbf{K}] &= \mathbf{F}[\mathbf{G}, \mathbf{K}] + (-1)^{|\mathbf{G}||\mathbf{K}|}[\mathbf{F}, \mathbf{K}]\mathbf{G}, \end{aligned} \tag{13.3}$$

one sees that $D_{\mathbf{H}} : \mathfrak{W} \to \mathfrak{W}$ satisfies a graded Leibniz rule,

$$D_{\mathbf{H}}(\mathbf{F}\mathbf{G}) = (D_{\mathbf{H}}\mathbf{F})\mathbf{G} + (-1)^{|\mathbf{F}|}\mathbf{F}\, D_{\mathbf{H}}\mathbf{G}.$$

It follows that $D_{\mathbf{H}} : \mathfrak{W} \to \mathfrak{W}$ is also a derivation with respect to the bracket structure on $\mathfrak{W}$, i.e.

$$D_{\mathbf{H}}[\mathbf{F}, \mathbf{G}] = [D_{\mathbf{H}}\mathbf{F}, \mathbf{G}] + (-1)^{|\mathbf{F}|}[\mathbf{F}, D_{\mathbf{H}}\mathbf{G}]$$

for all $\mathbf{F}, \mathbf{G} \in \mathfrak{W}$. As a consequence, the product and bracket structures on $\mathfrak{W}$ descend to $H_*^{\mathfrak{W}}(M, \xi; R)$, giving it the structure of a *Weyl superalgebra*.

As a matter of interest, we observe that $(\mathfrak{W}, D_{\mathbf{H}})$, as with $(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$ in the previous lecture, can be defined with $\mathbb{Z}$ or $\mathbb{Z}_2$ coefficients whenever the transversality results are good enough to take the usual expression $\sum_u \frac{\epsilon(u)}{|\mathrm{Aut}^\sigma(u)|}$ literally as a count of holomorphic curves. This result is of limited interest since it cannot hold in general cases where transversality for multiple covers is impossible without multivalued perturbations—nonetheless I find it amusing.[1]

PROPOSITION 13.4. *If Assumption 12.1 in Lecture 12 holds, then $D_{\mathbf{H}}$ is also well defined if the ring $R = \mathbb{Q}[H_2(M)/G]$ is replaced by $\mathbb{Z}[H_2(M)/G]$ or $\mathbb{Z}_2[H_2(M)/G]$.*

[1]The same arguments used to define SFT chain complexes over the integers can also be applied to the chain maps involved in the proof of invariance (see §13.3.1), so the SFT invariants *should* be defined over the integers if transversality can be achieved for multiple covers. There are known situations however in which this cannot hold: even if the chain complexes are well defined over $\mathbb{Z}$, invariance may hold only over $\mathbb{Q}$, due to the failure of transversality in cobordisms. See [**Hut**].

PROOF. Since $D_{\mathbf{H}}$ is a derivation, it suffices to check that for every good Reeb orbit γ, $D_{\mathbf{H}}q_\gamma$ and $D_{\mathbf{H}}p_\gamma$ are each sums of monomials of the form $ce^A\hbar^k q^{\gamma^-}p^{\gamma^+}$ with coefficients $c \in \mathbb{Z}$. Suppose $u \in \mathcal{M}_1(J)$ is an index 1 holomorphic curve with positive and/or negative asymptotic orbits

$$\boldsymbol{\gamma}^\pm = (\underbrace{\gamma_1^\pm, \dots, \gamma_1^\pm}_{m_1^\pm}, \dots, \underbrace{\gamma_{k_\pm}^\pm, \dots, \gamma_{k_\pm}^\pm}_{m_{k_\pm}^\pm}),$$

where $\gamma_i^\pm \neq \gamma_j^\pm$ for $i \neq j$. We can assume all the orbits $\gamma_i^\pm$ are good and that $m_i^\pm = 1$ whenever $n - 3 + \mu_{\mathrm{CZ}}(\gamma_i^\pm)$ is odd. Up to a sign and factors of e^A and $\hbar$ which are not relevant to this discussion, u then contributes a monomial

$$\mathbf{H}_u := \frac{1}{|\operatorname{Aut}^\sigma(u)|} q_{\gamma_1^-}^{m_1^-} \dots q_{\gamma_{k_-}^-}^{m_{k_-}^-} p_{\gamma_1^+}^{m_1^+} \dots p_{\gamma_{k_+}^+}^{m_{k_+}^+}$$

to $\mathbf{H}$. The commutator $[\mathbf{H}_u, q_\gamma]$ vanishes unless γ is one of the orbits $\gamma_1^+, \dots, \gamma_{k_+}^+$, so suppose $\gamma = \gamma_{k_+}^+$. If $n - 3 + \mu_{\mathrm{CZ}}(\gamma)$ is odd, then $m := m_{k_+}^+ = 1$, and (13.3) with $[p_\gamma, q_\gamma] = \kappa_\gamma \hbar$ implies

$$\begin{aligned}
[\mathbf{H}_u, q_\gamma] &= \frac{1}{|\operatorname{Aut}^\sigma(u)|} \left[q_{\gamma_1^-}^{m_1^-} \dots q_{\gamma_{k_-}^-}^{m_{k_-}^-} p_{\gamma_1^+}^{m_1^+} \dots p_{\gamma_{k_+-1}^+}^{m_{k_+-1}^+} p_\gamma, q_\gamma \right] \\
&= \frac{\kappa_\gamma}{|\operatorname{Aut}^\sigma(u)|} \hbar q_{\gamma_1^-}^{m_1^-} \dots q_{\gamma_{k_-}^-}^{m_{k_-}^-} p_{\gamma_1^+}^{m_1^+} \dots p_{\gamma_{k_+-1}^+}^{m_{k_+-1}^+}.
\end{aligned}$$

The fraction in front of this expression is an integer since u can have only one end asymptotic to γ, and κ_γ is thus divisible by the covering multiplicity of u. If $n - 3 + \mu_{\mathrm{CZ}}(\gamma)$ is even, then we generalize this calculation by using (13.3) to write

$$[p_\gamma^m, q_\gamma] = m\kappa_\gamma \hbar p_\gamma^{m-1},$$

so then,

$$\begin{aligned}
[\mathbf{H}_u, q_\gamma] &= \frac{1}{|\operatorname{Aut}^\sigma(u)|} \left[q_{\gamma_1^-}^{m_1^-} \dots q_{\gamma_{k_-}^-}^{m_{k_-}^-} p_{\gamma_1^+}^{m_1^+} \dots p_{\gamma_{k_+-1}^+}^{m_{k_+-1}^+} p_\gamma^m, q_\gamma \right] \\
&= \frac{\kappa_\gamma m}{|\operatorname{Aut}^\sigma(u)|} \hbar q_{\gamma_1^-}^{m_1^-} \dots q_{\gamma_{k_-}^-}^{m_{k_-}^-} p_{\gamma_1^+}^{m_1^+} \dots p_{\gamma_{k_+-1}^+}^{m_{k_+-1}^+} p_\gamma^{m-1}.
\end{aligned}$$

To see that $\frac{\kappa_\gamma m}{|\operatorname{Aut}^\sigma(u)|}$ is always an integer, recall from our proof of Prop. 12.16 in the previous lecture that transformations in $\operatorname{Aut}^\sigma(u)$ permute each of the sets of punctures that are asymptotic to the same Reeb orbit. Suppose the set of positive punctures of u asymptotic to γ is partitioned by the $\operatorname{Aut}^\sigma(u)$-action into N subsets, each consisting of $\ell_1, \dots, \ell_N$ punctures, where $\ell_1 + \dots + \ell_N = m$. If z is a puncture in the ith of these subsets, then its stabilizer is a cyclic subgroup of order k_i acting on a neighborhood of z by biholomorphic rotations, where $k_i\ell_i = |\operatorname{Aut}^\sigma(u)|$. Each of these orders k_i necessarily divides the multiplicity κ_γ, so we can write $k_i a_i = \kappa_\gamma$ for some $a_i \in \mathbb{N}$. Putting all this together, we have

$$\kappa_\gamma m = \sum_{i=1}^N \kappa_\gamma \ell_i = \sum_{i=1}^N k_i a_i \ell_i = |\operatorname{Aut}^\sigma(u)| \sum_{i=1}^N a_i.$$

Following this same procedure, you should now be able to verify on your own that the coefficient appearing in $[\mathbf{H}_u, p_\gamma]$ is also always an integer. The existence of a chain complex with $\mathbb{Z}_2$ coefficients follows from this simply by projecting $\mathbb{Z}$ to $\mathbb{Z}_2$. $\square$

13.1.2. The semiclassical limit: rational SFT. The idea of rational symplectic field theory (RSFT) is to extract as much information as possible from genus zero holomorphic curves but ignore curves of higher genus. The algebra of SFT provides a fairly obvious mechanism for this: RSFT should be what SFT becomes in the "limit as $\hbar \to 0$," i.e. the classical approximation to a quantum theory. Let

$$\mathfrak{P} := \mathfrak{W}/\hbar\mathfrak{W},$$

so $\mathfrak{P}$ is a graded unital algebra generated by the p_γ and q_γ variables and the coefficient ring R, but it does not include $\hbar$ as a generator. Since all commutators in $\mathfrak{W}$ are in $\hbar\mathfrak{W}$, the product structure of $\mathfrak{P}$ is supercommutative. Let us use the distinction between capital and lowercase letters to denote the quotient projection

$$\mathfrak{W} \to \mathfrak{P} : \mathbf{F} \mapsto \mathbf{f}.$$

We will make an exception for the letter "H": recall that $\mathbf{H}$ is not an element of $\mathfrak{W}$ since its genus zero terms have order -1 in $\hbar$, but $\hbar\mathbf{H} \in \mathfrak{W}$, so we will define

$$\mathbf{h} = \sum_u \frac{\epsilon(u)}{|\operatorname{Aut}^\sigma(u)|} e^A q^{\gamma^-} p^{\gamma^+} \in \mathfrak{P}$$

to be the image of $\hbar\mathbf{H}$ under the projection. The sum in this expression ranges over all $\mathbb{R}$-equivalence classes of index 1 curves with genus zero and good asymptotic orbits, so $\mathbf{h}$ will serve as the generating function of RSFT. To encode gluing of genus zero terms, note first that the commutator operation would not be appropriate since it prodcues terms for *every* possible gluing of two curves, including those which glue genus zero curves along more than one breaking orbit to produce buildings with positive arithmetic genus. We need instead to have an algebraic operation on $\mathfrak{P}$ that encodes gluing along only one breaking orbit at a time.

You already know what to expect if you've ever taken a quantum mechanics course: in the "classical limit," commutators become Poisson brackets. To express this properly, we need to make a distinction between differential operators operating from the left or the right: let

$$\frac{\overrightarrow{\partial}}{\partial q_\gamma} : \mathfrak{W} \to \mathfrak{W}$$

denote the usual operator $\frac{\partial}{\partial q_\gamma}$, which was previously defined on $\mathcal{A}[[\hbar]]$ but has an obvious extension to $\mathfrak{W}$ such that $\frac{\overrightarrow{\partial}}{\partial q_\gamma} p_{\gamma'} = 0$ for all γ'. This operator satisfies the graded Leibniz rule

$$\frac{\overrightarrow{\partial}}{\partial q_\gamma}(\mathbf{F}\mathbf{G}) = \left(\frac{\overrightarrow{\partial}}{\partial q_\gamma}\mathbf{F}\right)\mathbf{G} + (-1)^{|q_\gamma||\mathbf{F}|}\mathbf{F}\left(\frac{\overrightarrow{\partial}}{\partial q_\gamma}\mathbf{G}\right).$$

The related operator

$$\frac{\overleftarrow{\partial}}{\partial q_\gamma} : \mathfrak{W} \to \mathfrak{W} : \mathbf{F} \mapsto \mathbf{F}\frac{\overleftarrow{\partial}}{\partial q_\gamma}$$

is defined exactly the same way on individual variables p_γ and q_γ, but satisfies a slightly different Leibniz rule,

$$(\mathbf{F}\mathbf{G})\frac{\overleftarrow{\partial}}{\partial q_\gamma} = \mathbf{F}\left(\mathbf{G}\frac{\overleftarrow{\partial}}{\partial q_\gamma}\right) + (-1)^{|q_\gamma||\mathbf{G}|}\left(\mathbf{F}\frac{\overleftarrow{\partial}}{\partial q_\gamma}\right)\mathbf{G}.$$

The point of writing $\frac{\overleftarrow{\partial}}{\partial q_\gamma}$ so that it acts from the right is to obey the usual conventions of superalgebra: signs change whenever the order of two odd elements (or operators) is interchanged. Partial derivatives with respect to p_γ can be defined analogously on $\mathfrak{W}$. With this notation in hand, the **graded Poisson bracket** on $\mathfrak{W}$ is defined by

$$\{\mathbf{F},\mathbf{G}\} = \sum_\gamma \kappa_\gamma \left(\mathbf{F}\frac{\overleftarrow{\partial}}{\partial p_\gamma}\frac{\overrightarrow{\partial}}{\partial q_\gamma}\mathbf{G} - (-1)^{|\mathbf{F}||\mathbf{G}|}\mathbf{G}\frac{\overleftarrow{\partial}}{\partial p_\gamma}\frac{\overrightarrow{\partial}}{\partial q_\gamma}\mathbf{F}\right), \tag{13.4}$$

where the sum ranges over all good Reeb orbits. In the same manner, the differential operators and the bracket $\{\ ,\ \}$ can also be defined on $\mathfrak{P}$.

It is easy to check that $\{\ ,\ \}$ on $\mathfrak{W}$ *almost* satisfies a version of (13.3): we have

$$\begin{aligned}\{\mathbf{F},\mathbf{G}\mathbf{K}\} &= \{\mathbf{F},\mathbf{G}\}\mathbf{K} + (-1)^{|\mathbf{F}||\mathbf{G}|}\mathbf{G}\{\mathbf{F},\mathbf{K}\} + \mathcal{O}(\hbar),\\ \{\mathbf{F}\mathbf{G},\mathbf{K}\} &= \mathbf{F}\{\mathbf{G},\mathbf{K}\} + (-1)^{|\mathbf{G}||\mathbf{K}|}\{\mathbf{F},\mathbf{K}\}\mathbf{G} + \mathcal{O}(\hbar)\end{aligned} \tag{13.5}$$

for all $\mathbf{F},\mathbf{G},\mathbf{K} \in \mathfrak{W}$. The extra terms denoted by $\mathcal{O}(\hbar)$ arise from the fact that in proving (13.5), we must sometimes reorder products $\mathbf{F}\mathbf{G}$ by writing them as $(-1)^{|\mathbf{F}||\mathbf{G}|}\mathbf{G}\mathbf{F} + [\mathbf{F},\mathbf{G}]$, where $[\mathbf{F},\mathbf{G}] = \mathcal{O}(\hbar)$. Since the terms with $\hbar$ disappear in $\mathfrak{P}$, the relations become exact in $\mathfrak{P}$:

$$\begin{aligned}\{\mathbf{f},\mathbf{g}\mathbf{k}\} &= \{\mathbf{f},\mathbf{g}\}\mathbf{k} + (-1)^{|\mathbf{f}||\mathbf{g}|}\mathbf{g}\{\mathbf{f},\mathbf{k}\},\\ \{\mathbf{f}\mathbf{g},\mathbf{k}\} &= \mathbf{f}\{\mathbf{g},\mathbf{k}\} + (-1)^{|\mathbf{g}||\mathbf{k}|}\{\mathbf{f},\mathbf{k}\}\mathbf{g}\end{aligned} \tag{13.6}$$

for all $\mathbf{f},\mathbf{g},\mathbf{k} \in \mathfrak{P}$.

PROPOSITION 13.5. *For all* $\mathbf{F},\mathbf{G} \in \mathfrak{W}$,

$$[\mathbf{F},\mathbf{G}] = \hbar\{\mathbf{f},\mathbf{g}\} + \mathcal{O}(\hbar^2),$$

and $\{\ ,\ \}$ *satisfies the conditions of a super Lie bracket on* $\mathfrak{P}$.

REMARK 13.6. In formulas like the one in the above proposition, we interpret $\{\mathbf{f},\mathbf{g}\} \in \mathfrak{P}$ as an element of $\mathfrak{W}$ via any choice of R-linear inclusion $\mathfrak{P} \hookrightarrow \mathfrak{W}$ that acts as the identity on the generators p_γ, q_γ. There is ambiguity in this choice due to the noncommutativity of $\mathfrak{W}$, but the ambiguity is in $\hbar\mathfrak{W}$ and thus makes no difference to the formula.

Proof of Proposition 13.5. The formula is easily checked when $\mathbf{F}$ and $\mathbf{G}$ are individual variables of the form p_γ or q_γ; in fact the extra term $\mathcal{O}(\hbar^2)$ can be omitted in these cases. The case where $\mathbf{F}$ and $\mathbf{G}$ are general monomials follows from this via (13.3) and (13.5) using induction on the number of variables in the product. This implies the general case via bilinearity.

Given the formula, the condition $\{\mathbf{f},\mathbf{g}\} + (-1)^{|\mathbf{f}||\mathbf{g}|}\{\mathbf{g},\mathbf{f}\} = 0$ and the Poisson version of the super Jacobi identity (13.1) follow from the corresponding properties of $[\ ,\ ]$. $\square$

The proposition implies that our genus zero generating function $\mathbf{h} \in \mathfrak{P}$ satisfies $0 = \hbar^2[\mathbf{H},\mathbf{H}] = [\hbar\mathbf{H},\hbar\mathbf{H}] = \hbar\{\mathbf{h},\mathbf{h}\} + \mathcal{O}(\hbar^2)$, thus

$$\{\mathbf{h},\mathbf{h}\} = 0.$$

This relation can be interpreted as the count of boundary points of all 1-dimensional moduli spaces of genus zero curves: indeed, any pair of genus zero curves $u, v \in \mathcal{M}_1^\sigma(J)/\mathbb{R}$ contributes to $\{\mathbf{h},\mathbf{h}\}$ a term of the form

$$\sum_\gamma \frac{\kappa_\gamma}{|\operatorname{Aut}^\sigma(u)||\operatorname{Aut}^\sigma(v)|} e^{A_u + A_v} q^{\gamma_u^-} \left(p^{\gamma_u^+} \overleftarrow{\frac{\partial}{\partial p_\gamma}} \right) \left(\overrightarrow{\frac{\partial}{\partial q_\gamma}} q^{\gamma_v^-} \right) p^{\gamma_v^+},$$

plus a corresponding term with the roles of u and v reversed. This sums all the monomials that one can construct by cancelling one p_γ variable from u with a matching q_γ variable from v, in other words, constructing a building by gluing v on top of u along one matching Reeb orbit.

The graded Jacobi identity will again imply that any representation of the super Lie algebra $(\mathfrak{P}, \{\ ,\ \})$ gives rise to a chain complex with $\mathbf{h}$ as its differential. For example we can take the adjoint representation,

$$\mathfrak{P} \to \operatorname{End}_R(\mathfrak{P}) : \mathbf{f} \mapsto d_{\mathbf{f}}, \qquad d_{\mathbf{f}}\mathbf{g} := \{\mathbf{f},\mathbf{g}\},$$

which satisfies $d_{\{\mathbf{f},\mathbf{g}\}} = d_{\mathbf{f}} d_{\mathbf{g}} - (-1)^{|\mathbf{f}||\mathbf{g}|} d_{\mathbf{g}} d_{\mathbf{f}}$ due to the Jacobi identity. Then $d_{\mathbf{h}}^2 = 0$ since $\mathbf{h}$ has odd degree and $\{\mathbf{h},\mathbf{h}\} = 0$, and the homology of **rational SFT** is defined as

$$H_*^{\mathrm{RSFT}}(M,\xi;R) := H_*(\mathfrak{P}, d_{\mathbf{h}}).$$

We again refer to [EGH00] for an argument that $H_*^{\mathrm{RSFT}}(M,\xi;R)$ is an invariant of the contact structure. Notice that Proposition 13.5 yields a simple relationship between the chain complexes $(\mathfrak{W}, D_{\mathbf{H}})$ and $(\mathfrak{P}, d_{\mathbf{h}})$, namely

$$D_{\mathbf{H}}\mathbf{F} = d_{\mathbf{h}}\mathbf{f} + \mathcal{O}(\hbar), \tag{13.7}$$

where $d_{\mathbf{h}}\mathbf{f}$ is interpreted as an element of $\mathfrak{W}$ via Remark 13.6. In other words, the projection $\mathfrak{W} \to \mathfrak{P} : \mathbf{F} \to \mathbf{f}$ is a chain map. Moreover, $d_{\mathbf{H}}$ is a derivation on $\mathfrak{P}$ with respect to both the product and the Poisson bracket: this follows via Proposition 13.5 and (13.7) from the fact that $D_{\mathbf{H}}$ satisfies the corresponding properties on $\mathfrak{W}$. We conclude that $H_*^{\mathrm{RSFT}}(M,\xi;R)$ inherits the structure of a Poisson superalgebra, and the map

$$H_*^{\mathfrak{W}}(M,\xi;R) \to H_*^{\mathrm{RSFT}}(M,\xi;R)$$

induced by the chain map $(\mathfrak{W}, D_{\mathbf{H}}) \to (\mathfrak{P}, d_{\mathbf{h}})$ is both an algebra homomorphism and a homomorphism of graded super Lie algebras.

13.1.3. The contact homology algebra. Contact homology is the most popular tool in the SFT package and was probably the first to be understood beyond the more straightforward cylindrical theory. In situations where cylindrical contact homology cannot be defined due to bubbling of holomorphic planes, the next simplest thing one can do is to define a theory that counts genus zero curves with one positive end but *arbitrary* numbers of negative ends (cf. Exercise 10.14 in Lecture 10).

The proper algebraic setting for such a theory turns out to be the algebra $\mathcal{A}$ generated by the q_γ variables, and it can be derived from RSFT by setting all p_γ variables to zero. Using the obvious inclusion $\mathcal{A} \hookrightarrow \mathfrak{P}$, define $\partial_{\mathrm{CH}} : \mathcal{A} \to \mathcal{A}$ by

$$\partial_{\mathrm{CH}}\mathbf{f} = d_{\mathbf{h}}\mathbf{f}|_{p=0}.$$

We can thus write $d_{\mathbf{h}}\mathbf{f} = \partial_{\mathrm{CH}}\mathbf{f} + \mathcal{O}(p)$, where

$$\mathcal{O}(p^k)$$

will be used generally to denote any formal sum consisting exclusively of terms of the form $p_{\gamma_1} \dots p_{\gamma_k}\mathbf{f}$ for $\mathbf{f} \in \mathfrak{P}$. Now observe that for any good orbit γ,

$$d_{\mathbf{h}}p_\gamma = \{\mathbf{h}, p_\gamma\} = -(-1)^{|p_\gamma|}\sum_{\gamma'}\left(p_\gamma\frac{\overleftarrow{\partial}}{\partial p_{\gamma'}}\right)\left(\frac{\overrightarrow{\partial}}{\partial q_{\gamma'}}\mathbf{h}\right) = -(-1)^{|p_\gamma|}\frac{\partial \mathbf{h}}{\partial q_\gamma} = \mathcal{O}(p)$$

since every term in $\mathbf{h}$ has at least one p variable. It follows that $d_{\mathbf{h}}(\mathcal{O}(p)) = \mathcal{O}(p)$, so the fact that $d_{\mathbf{h}}^2 = 0$ implies $\partial_{\mathrm{CH}}^2 = 0$, and **contact homology** is defined as

$$HC_*(M, \xi; R) := H_*(\mathcal{A}, \partial_{\mathrm{CH}}).$$

Since $d_{\mathbf{h}}$ is a derivation on $\mathfrak{P}$, the formula $d_{\mathbf{h}}\mathbf{f} = \partial_{\mathrm{CH}}\mathbf{f} + \mathcal{O}(p)$ implies that ∂_{CH} is likewise a derivation on $\mathcal{A}$, so $HC_*(M, \xi; R)$ has the structure of a graded supercommutative algebra with unit. Moreover, the projection $\mathfrak{P} \to \mathcal{A} : \mathbf{f} \mapsto \mathbf{f}|_{p=0}$ is a chain map, giving rise to an algebra homomorphism

$$H_*^{\mathrm{RSFT}}(M, \xi; R) \to HC_*(M, \xi; R).$$

The invariance of $HC_*(M, \xi; R)$ will follow from the invariance of $H_*^{\mathrm{SFT}}(M, \xi; R)$, to be discussed in §13.3.1 below.

To interpret ∂_{CH}, we can separate the part of $\mathbf{h}$ that is linear in p variables, writing

$$\mathbf{h} = \sum_\gamma \mathbf{h}_\gamma(q)p_\gamma + \mathcal{O}(p^2),$$

where for each good Reeb orbit γ, $\mathbf{h}_\gamma(q)$ denotes a polynomial in q variables with coefficients in R. Since elements $\mathbf{f} \in \mathcal{A}$ have no dependence on p variables, we then have

$$d_{\mathbf{h}}\mathbf{f} = \{\mathbf{h}, \mathbf{f}\} = \sum_\gamma \kappa_\gamma\left(\mathbf{h}\frac{\overleftarrow{\partial}}{\partial p_\gamma}\right)\left(\frac{\overrightarrow{\partial}}{\partial q_\gamma}\mathbf{f}\right) = \sum_\gamma \kappa_\gamma \mathbf{h}_\gamma \frac{\partial \mathbf{f}}{\partial q_\gamma} + \mathcal{O}(p),$$

hence

$$\partial_{\mathrm{CH}}\mathbf{f} = \sum_\gamma \kappa_\gamma \mathbf{h}_\gamma \frac{\partial \mathbf{f}}{\partial q_\gamma}.$$

In particular, ∂_{CH} acts on each generator $q_\gamma \in \mathcal{A}$ as

$$\partial_{\mathrm{CH}} q_\gamma = \kappa_\gamma \mathbf{h}_\gamma = \sum_u \frac{\epsilon(u)\kappa_\gamma}{\mathrm{Aut}^\sigma(u)} e^A q^{\boldsymbol{\gamma}^-},$$

where the sum is over all $\mathbb{R}$-equivalence classes of index 1 J-holomorphic curves u with genus zero, one positive end at γ, negative ends at good orbits $\boldsymbol{\gamma}^-$, and homology class $A \in H_2(M)/G$.

13.1.4. Algebraic overtwistedness. Even the simplest of the three differential graded algebras described above is too large to compute in most cases. The major exception is the case of overtwisted contact manifolds.

THEOREM 13.7. *If (M, ξ) is overtwisted, then $HC_*(M, \xi; R) = 0$ for all choices of the coefficient ring R.*

REMARK 13.8. If X is an algebra with unit, then saying $X = 0$ is equivalent to saying that $1 = 0$ in X.

The notion of overtwisted contact structures in dimension three was introduced by Eliashberg in [**Eli89**], who proved that they are *flexible* in the sense that their classification up to isotopy reduces to the purely obstruction-theoretic classification of oriented 2-plane fields up to homotopy. This means in effect that an overtwisted contact structure carries no distinctly contact geometric information, so it should not be surprising when "interesting" contact invariants such as $HC_*(M, \xi)$ vanish. The three-dimensional case of Theorem 13.7 seems to have been among the earliest insights about SFT: its first appearance in the literature was in [**Eli98**], and a proof later appeared in a paper by Mei-Lin Yau [**Yau06**], which includes a brief appendix sketching Eliashberg's original proof. We will discuss Eliashberg's proof in detail in Lecture 16.

The definitive higher-dimensional notion of overtwistedness was introduced much more recently by Borman-Eliashberg-Murphy [**BEM15**], following earlier steps in this direction by Niederkrüger [**Nie06**] and others. There are now two known proofs of Theorem 13.7 in higher dimensions: the first uses the fact that since overtwisted contact manifolds are flexible, they always admit an embedding of a *plastikstufe*, which implies vanishing of contact homology by an unpublished result of Bourgeois and Niederkrüger (see [**Bou09**, Theorem 4.10] for a sketch of the argument). The second argument appeals to an even more recent result of Casals-Murphy-Presas [**CMP19**] showing that (M, ξ) is overtwisted if and only if it is supported by a negatively stabilized open book, in which case $HC_*(M, \xi) = 0$ was proven by Bourgeois and van Koert [**BvK10**].

It is not known whether the vanishing of contact homology *characterizes* overtwistedness, i.e. there are not yet any examples of tight contact manifolds with $HC_*(M, \xi) = 0$. I will go out on a limb and say that such examples seem unlikely to exist in dimension three but are much more likely in higher dimensions; in fact various candidates are known [**MNW13**, **CDvK16**], but we do not yet have adequate methods to prove that any of them are tight. The lack of known counterexamples has nonetheless given rise to the following definition.

DEFINITION 13.9. A closed contact manifold (M, ξ) is **algebraically overtwisted** if $HC_*(M, \xi; R) = 0$ for every choice of the coefficient ring R.

REMARK 13.10. The coefficient ring is not always mentioned in statements of the above definition, but it should be. We will see in §13.3.2 below that this detail makes a difference to issues like symplectic filling obstructions. Note that for any nested pair of subgroups $G \subset G' \subset H_2(M)$, the natural projection $H_2(M)/G' \to H_2(M)/G$ induces an algebra homomorphism

$$HC_*(M, \xi; \mathbb{Q}[H_2(M)/G']) \to HC_*(M, \xi; \mathbb{Q}[H_2(M)/G]).$$

Since algebra homomorphisms necessarily map $1 \mapsto 1$ and $0 \mapsto 0$, the target of this map must vanish whenever its domain does, so for checking Definition 13.9, it suffices to check the case $R = \mathbb{Q}[H_2(M)]$.

We've seen above that there exist algebra homomorphisms

$$H_*^{\mathfrak{W}}(M, \xi; R) \to H_*^{\mathrm{RSFT}}(M, \xi; R) \to HC_*(M, \xi; R), \tag{13.8}$$

thus the vanishing of either of the algebras $H_*^{\mathfrak{W}}(M, \xi; R)$ or $H_*^{\mathrm{RSFT}}(M, \xi; R)$ with all coefficient rings R is another sufficient condition for algebraic overtwistedness. Bourgeois and Niederkrüger observed that, in fact, these conditions are also necessary:

THEOREM 13.11 ([BN10]). *For any coefficient ring R, the following conditions are equivalent:*

(1) $HC_*(M, \xi; R) = 0$,
(2) $H_*^{\mathrm{RSFT}}(M, \xi; R) = 0$,
(3) $H_*^{\mathrm{SFT}}(M, \xi; R) = 0$.

PROOF. The implications (3) ⇒ (2) ⇒ (1) are immediate from the algebra homomorphisms (13.8), thus it will suffice to prove (1) ⇒ (3). Suppose $1 = 0 \in HC_*(M, \xi; R)$, which means $\partial_{\mathrm{CH}} \mathbf{f} = 1$ for some $\mathbf{f} \in \mathcal{A}$. Using the obvious inclusion $\mathcal{A} \hookrightarrow \mathfrak{W}$, this means

$$D_{\mathbf{H}} \mathbf{f} = 1 - \mathbf{G},$$

where $\mathbf{G} = \mathcal{O}(p, \hbar)$, i.e. $\mathbf{G}$ is a sum of terms that all contain at least one p_γ variable or a power of $\hbar$. It follows that $\mathbf{G}^k = \mathcal{O}(p^k, \hbar^k)$ for all $k \in \mathbb{N}$, and the infinite sum

$$\sum_{k=0}^{\infty} \mathbf{G}^k$$

is therefore an element of $\mathfrak{W}$, as the coefficient in front of any fixed monomial $\hbar^k p^\gamma$ in this sum is a polynomial function of the q variables. This sum is then a multiplicative inverse of $1 - \mathbf{G}$, and since

$$0 = D_{\mathbf{H}}^2 \mathbf{f} = 0 = -D_{\mathbf{H}} \mathbf{G},$$

it also satisfies $D_{\mathbf{H}} \left((1 - \mathbf{G})^{-1} \right) = 0$. Using the fact that $D_{\mathbf{H}}$ is a derivation, we therefore have

$$D_{\mathbf{H}} \left((1 - \mathbf{G})^{-1} \mathbf{f} \right) = (1 - \mathbf{G})^{-1} (1 - \mathbf{G}) = 1,$$

implying $1 = 0 \in H_*^{\mathrm{SFT}}(M, \xi; R)$. □

13.2. SFT generating functions for cobordisms

All invariance proofs in SFT are based on a generating function analogous to **H** that counts index 0 holomorphic curves in symplectic cobordisms. The basic definition is a straightforward extension of what we saw in Lecture 12, but there is an added wrinkle due to the fact that, in general, one must include *disconnected* curves in the count.

13.2.1. Weak, strong and stable cobordisms. First some remarks about the category we are working in. Since the stated purpose of SFT is to define invariants of contact structures, we have been working since Lecture 12 with symplectizations of contact manifolds rather than more general stable Hamiltonian structures. We've made use of this restriction on several occasions, namely so that we can assume:

(1) All nontrivial holomorphic curves in $\mathbb{R} \times M$ have at least one positive puncture;
(2) The energy of a holomorphic curve in $\mathbb{R} \times M$ can be bounded in terms of its positive asymptotic orbits.

It will be useful however for certain applications to permit a slightly wider class of stable Hamiltonian structure. Recall that a hypersurface V in an almost complex manifold (W, J) is called **pseudoconvex** if the maximal complex subbundle

$$\xi := TV \cap J(TV) \subset TV$$

defines a contact structure on V whose canonical conformal symplectic bundle structure tames $J|_\xi$. For example, if α is a contact form on M and $J \in \mathcal{J}(\alpha)$, then each of the hypersurfaces $\{\text{const}\} \times M$ is pseudoconvex in $(\mathbb{R} \times M, J)$. The contact structure ξ induces an orientation on the hypersurface V; if V comes with its own orientation (e.g. as a boundary component of W), then we call it *pseudoconvex* if ξ is a positive contact structure with respect to this orientation, and **pseudoconcave** otherwise. For example, if (W, ω) is a symplectic cobordism from (M_-, ξ_-) to (M_+, ξ_+) and $J \in \mathcal{J}(W, \omega, \alpha_+, \alpha_-)$, then M_+ is pseudoconvex and M_- is pseudoconcave.

DEFINITION 13.12. Given an odd-dimensional manifold M, we will say that an almost complex structure J on $\mathbb{R} \times M$ is **pseudoconvex** if $\{r\} \times M$ is a pseudoconvex hypersurface in $(\mathbb{R} \times M, J)$ for every $r \in \mathbb{R}$, with the induced orientation such that ∂_r and $\{r\} \times M$ are positively transverse.

If $\mathcal{H} = (\omega, \lambda)$ is a stable Hamiltonian structure on M, then pseudoconvexity of $J \in \mathcal{J}(\mathcal{H})$ imposes conditions on $\mathcal{H}$, in particular λ must be a contact form. It also requires $J|_\xi$ to be tamed by $d\lambda|_\xi$, but unlike the case when $J \in \mathcal{J}(\lambda)$, $J|_\xi$ need not be *compatible* with it, i.e. the positive bilinear form $d\lambda(\cdot, J\cdot)|_\xi$ need not be symmetric. As always, $J|_\xi$ must be compatible with $\omega|_\xi$, but ω need not be an *exact* form for this to hold—the freedom to change $[\omega] \in H^2_{\text{dR}}(M)$ will be the main benefit of this generalization, particularly when we discuss weak symplectic fillings below.

PROPOSITION 13.13. *Suppose $\mathcal{H} = (\omega, \lambda)$ is a stable Hamiltonian structure on a closed manifold M and $J \in \mathcal{J}(\mathcal{H})$ is pseudoconvex. Then all nonconstant finite-energy J-holomorphic curves in $\mathbb{R} \times M$ have at least one positive puncture, and* *their energies satisfy a uniform upper bound in terms of the periods of their positive asymptotic orbits.*

PROOF. It is straightforward to check that either of the two proofs of Proposition 10.9 given in Lecture 10 generalizes to any J on $\mathbb{R} \times M$ that is pseudoconvex. In particular, pseudoconvexity implies that if $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$ is a J-holomorphic curve, then $u^* d\lambda \geqslant 0$, with equality only at points where u is tangent to ∂_r and the Reeb vector field. Stokes' theorem thus gives

$$0 \leqslant \int_{\dot{\Sigma}} u^* d\lambda = \sum_{z \in \Gamma^+} T_z - \sum_{z \in \Gamma^-} T_z, \tag{13.9}$$

where $T_z > 0$ denotes the period of the asymptotic orbit at each positive/negative puncture $z \in \Gamma^\pm$. Since $J|_\xi$ is also tamed by $\omega|_\xi$ and ω annihilates the Reeb vector field, we similarly have $u^*\omega \geqslant 0$, with the same condition for equality, and the compactness of M then implies an estimate of the form

$$0 \leqslant u^*\omega \leqslant c u^* d\lambda$$

for every J-holomorphic curve $u : (\dot{\Sigma}, j) \to (\mathbb{R} \times M, J)$, with a constant $c > 0$ that depends only on M, $\mathcal{H}$ and J. In light of (13.9), this implies an upper bound on $\int_{\dot{\Sigma}} u^*\omega$ in terms of the periods T_z for $z \in \Gamma^+$. Writing $\omega_\varphi = \omega + d(\varphi(r)\lambda)$ for suitable C^0-small increasing functions $\varphi : \mathbb{R} \to \mathbb{R}$, we can then apply Stokes' theorem to the second term in

$$E(u) = \sup_\varphi \int_{\dot{\Sigma}} u^*\omega_\varphi = \int_{\dot{\Sigma}} u^*\omega + \sup_\varphi \int_{\dot{\Sigma}} u^* d(\varphi(r)\lambda),$$

implying a similar upper bound for $E(u)$. □

COROLLARY 13.14. *For any stable Hamiltonian structure $\mathcal{H} = (\omega, \lambda)$ with a nondegenerate Reeb vector field $R_{\mathcal{H}}$ and a pseudoconvex $J \in \mathcal{J}(\mathcal{H})$, one can use closed $R_{\mathcal{H}}$-orbits and count J-holomorphic curves in $\mathbb{R} \times M$ to define the chain complexes $(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$, $(\mathfrak{W}, D_{\mathbf{H}})$, $(\mathfrak{P}, d_{\mathbf{h}})$ and $(\mathcal{A}, \partial_{\mathrm{CH}})$.* □

We shall denote the homologies of the above chain complexes with coefficients in $R = \mathbb{Q}[H_2(M)/G]$ by

$$H_*^{\mathrm{SFT}}(M, \mathcal{H}, J; R), \quad H_*^{\mathfrak{W}}(M, \mathcal{H}, J; R), \quad H_*^{\mathrm{RSFT}}(M, \mathcal{H}, J; R), \quad HC_*(M, \mathcal{H}, J; R).$$

We make no claim at this point about these homologies being invariant. For the examples that we actually care about, this will turn out to be an irrelevant question due to Proposition 13.16 and Exercise 13.36 below.

EXAMPLE 13.15. Suppose α is a contact form on (M, ξ) and $\mathcal{H} = (\Omega, \alpha)$ is a stable Hamiltonian structure. Then for all constants $c > 0$ sufficiently large, $\mathcal{H}_c := (\Omega + c\, d\alpha, \alpha)$ is also a stable Hamiltonian structure and there exists a pseudoconvex $J_c \in \mathcal{J}(\mathcal{H}_c)$. To see the latter, notice that $\mathcal{H}'_c := \left(\frac{1}{c}\Omega + d\alpha, \alpha\right)$ is another family of stable Hamiltonian structures, with $\mathcal{J}(\mathcal{H}'_c) = \mathcal{J}(\mathcal{H}_c)$ for all c, and $\mathcal{H}'_c \to (d\alpha, \alpha)$ as $c \to \infty$. Thus one can select $J_c \in \mathcal{J}(\mathcal{H}_c)$ converging to some $J_\infty \in \mathcal{J}(\alpha)$ as $c \to \infty$, and these are pseudoconvex for $c > 0$ sufficiently large since J_∞ is.

PROPOSITION 13.16. *In the setting of Example 13.15, assume α is nondegenerate and $J_\infty \in \mathcal{J}(\alpha)$ satisfies Assumption 12.1. If $HC_*(M, \xi; R)$ vanishes, then $HC_*(M, \mathcal{H}_c, J_c; R)$ also vanishes for all $c > 0$ sufficiently large.*

PROOF. We will assume in the following that the usual (unrealistic) transversality assumptions hold, but the essential idea of the argument would not change in the presence of abstract perturbations.

Let $(\mathcal{A}, \partial^\infty_{\mathrm{CH}})$ denote the contact homology chain complex generated by closed R_α-orbits, with $\partial^\infty_{\mathrm{CH}}$ counting J_∞-holomorphic curves in $\mathbb{R} \times M$. The assumption $HC_*(M, \xi; R) = 0$ means there exists an element $\mathbf{f} \in \mathcal{A}$ with $\partial^\infty_{\mathrm{CH}}\mathbf{f} = 1$. Here $\mathbf{f}$ is a polynomial function of the q_γ variables, and $\partial^\infty_{\mathrm{CH}}\mathbf{f}$ counts a specific finite set of Fredholm regular index 1 curves in $(\mathbb{R} \times M, J_\infty)$. Now let $(\mathcal{A}, \partial^c_{\mathrm{CH}})$ denote the chain complex for $HC_*(M, \mathcal{H}_c, J_c; R)$, and notice that since α is contact, the stable Hamiltonian structures $(d\alpha, \alpha)$ and $\mathcal{H}_c = (\Omega + c\, d\alpha, \alpha)$ define matching Reeb vector fields, so the set of generators is unchanged. There is also no change to this complex if we replace $\mathcal{H}_c = (\Omega + c\, d\alpha, \alpha)$ by $\mathcal{H}'_c = \left(\frac{1}{c}\Omega + d\alpha, \alpha\right)$: this changes the energies of individual J_c-holomorphic curves, but the sets of finite-energy curves are still the same in both cases. We can assume $J_c \to J_\infty$ in C^∞ as $c \to \infty$. The implicit function theorem then extends each of the finitely many J_∞-holomorphic curves counted by $\partial^\infty\mathbf{f}$ uniquely to a smooth 1-parameter family of J_c-holomorphic curves for $c > 0$ sufficiently large.[2] We claim that these are the only curves counted by $\partial^c_{\mathrm{CH}}\mathbf{f}$ when $c > 0$ is large. Indeed, there would otherwise exist a sequence $c_k \to \infty$ for which additional J_{c_k}-holomorphic index 1 curves u_k contribute to $\partial^{c_k}_{\mathrm{CH}}\mathbf{f}$, and since $\mathbf{f}$ has only finitely many terms representing possible positive asymptotic orbits, we can find a subsequence for which all the u_k have the same positive asymptotic orbits. A further subsequence then has all the same negative asymptotic orbits as well since the Reeb flow is nondegenerate and the total period of the negative orbits is bounded by the total period of the positive orbits. Finally, since the sequence of stable Hamiltonian structures $\mathcal{H}'_{c_k}$ converges to $(d\alpha, \alpha)$, the curves u_k have uniformly bounded energy with respect to $\mathcal{H}'_{c_k}$, so that SFT compactness yields a subsequence converging to a J_∞-holomorphic building of index 1, which can only be one of the curves counted by $\partial^\infty_{\mathrm{CH}}\mathbf{f}$. This contradicts the uniqueness in the implicit function theorem and thus proves the claim. We conclude that for all $c > 0$ sufficiently large, $\partial^c_{\mathrm{CH}}\mathbf{f} = 1$. □

DEFINITION 13.17. Assume (W, ω) is a symplectic cobordism with stable boundary $\partial W = -M_- \coprod M_+$, with induced stable Hamiltonian structures $\mathcal{H}_\pm = (\omega_\pm, \lambda_\pm)$ at $M_\pm$, and suppose J is an almost complex structure on the completion $\widehat{W}$ that is ω-tame on W and belongs to $\mathcal{J}(\mathcal{H}_\pm)$ on the cylindrical ends. We will say that J is **pseudoconvex near infinity**[3] if the $\mathbb{R}$-invariant almost complex structures $J_\pm$ defined by restricting J to $[0, \infty) \times M_+$ and $(-\infty, 0] \times M_-$ are both pseudoconvex.

[2] In case you are concerned about the parametric moduli space being an orbifold instead of a manifold, just add asymptotic markers so that there is no isotropy, and divide by the appropriate combinatorial factors to count.

[3] If I were being hypercorrect about use of language, I might insist on saying that J is "pseudoconvex near $+\infty$ and *pseudoconcave* near $-\infty$," as the orientation reversal at the negative boundary makes M_- technically a pseudoconcave hypersurface in $(\widehat{W}, J)$, not pseudoconvex. But this definition will only be useful to us in cases where $M_- = \emptyset$, so my linguistic guilt is limited.

Note that the condition on J in the above definition can only be satisfied if $\lambda_\pm$ are both positive contact forms on $M_\pm$, but the 2-forms $\omega_\pm$ need not be exact.

Proving contact invariance of SFT requires counting curves in trivial exact symplectic cobordisms, but it is also natural to try to say things about non-exact **strong** symplectic cobordisms using SFT.[4] These fit naturally into our previously established picture since every strong cobordism has collar neighborhoods near the boundary in which it matches the symplectization of a contact manifold. The following more general notion of cobordism is also natural from a contact topological perspective, but fits less easily into the SFT picture.

DEFINITION 13.18 ([MNW13]). Given closed contact manifolds (M_+, ξ_+) and (M_-, ξ_-) of dimension $2n-1$, a **weak symplectic cobordism** from (M_-, ξ_-) to (M_+, ξ_+) is a compact symplectic manifold (W, ω) with $\partial W = -M_- \coprod M_+$ admitting an ω-tame almost complex structure J for which the almost complex manifold (W, J) is pseudoconvex at M_+ and pseudoconcave at M_-, with

$$\xi_\pm = TM_\pm \cap J(TM_\pm).$$

Weak cobordisms are characterized by the existence of a tame almost complex structure J whose restriction to $\xi_\pm$ is tamed by *two* symplectic bundle structures, $\omega|_{\xi_\pm}$ and $d\alpha_\pm|_{\xi_\pm}$ (for any choices of contact forms $\alpha_\pm$ defining $\xi_\pm$). Notice that in dimension 4, the second condition is mostly vacuous, and the weak cobordism condition just reduces to

$$\omega|_{\xi_\pm} > 0.$$

In this form, the low-dimensional case of Definition 13.18 has been around since the late 1980's, and there are many interesting results about it, e.g. examples of contact 3-manifolds that are weakly but not strongly fillable [Gir94, Eli96]. We will see in §13.3.2 that this distinction is detectable via SFT. Higher-dimensional examples of this phenomenon were found in [MNW13].

One major difference between weak and strong cobordisms is that the latter are always exact near the boundary, as the Liouville vector field is dual to a primitive of ω. It turns out that up to deformation, weak fillings that are exact at the boundary are the same thing as strong fillings—this was first observed by Eliashberg in dimension three [Eli91, Prop. 3.1], and was extended to higher dimensions in [MNW13]:

PROPOSITION 13.19. *Suppose (W, ω) is a weak filling of a $(2n-1)$-dimensional contact manifold (M, ξ) such that $\omega|_{TM}$ is exact. Then after a homotopy of ω through a family of symplectic forms that vary only in a collar neighborhood of ∂W and define weak fillings of (M, ξ), (W, ω) is a strong filling of (M, ξ).*

PROOF. Choose any contact form α for ξ, and let $\Omega = \omega|_{TM}$. By Proposition 6.9 and Remark 6.10, we can identify a collar neighborhood of ∂W in W with $(-\epsilon, 0] \times M$, with the coordinate on $(-\epsilon, 0]$ denoted by r, such that $\omega = \Omega + d(r\alpha)$ on the collar.

[4] By *strong cobordism*, we mean the usual notion of a compact symplectic manifold with convex and/or concave boundary components (see §1.4). The word "strong" is included in order to contrast this notion with its weaker cousin described in Definition 13.18.

By assumption, $\Omega = d\eta$ for some 1-form η on M, and since (W, ω) is a weak filling of $(M, \xi = \ker \alpha)$, we can choose a complex structure J_ξ on ξ that is tamed by both $d\alpha|_\xi$ and $d\eta|_\xi$. Now choose a smooth cutoff function $\beta : [0, \infty) \to [0, 1]$ that has compact support and equals 1 near 0. We claim that

$$\omega := d(\beta(r)\eta) + d(r\alpha)$$

is a symplectic form on $[0, \infty) \times M$ if $|\beta'|$ is sufficiently small. Indeed, writing $\omega = dr \wedge (\alpha + \beta'(r)\,\eta) + [\beta(r)\,d\eta + r\,d\alpha]$, we have

$$\omega^n = n\,dr \wedge \alpha \wedge [\beta(r)\,d\eta + r\,d\alpha]^{n-1} + n\beta'(r)\,dr \wedge \eta \wedge [\beta(r)\,d\eta + r\,d\alpha]^{n-1}.$$

The first term is positive and bounded away from zero since $d\eta|_\xi$ and $d\alpha|_\xi$ both tame J_ξ, hence so does $\beta\,d\eta + r\,d\alpha|_\xi$. The second term is then harmless if $|\beta'|$ is sufficiently small, proving $\omega^n > 0$.

This defines an extension of the original weak filling to a symplectic completion $\widehat{W} = W \cup_M ([0, \infty) \times M)$, and for each $r_0 \geqslant 0$, the compact subdomains defined by $r \leqslant r_0$ define weak fillings of $(\{r_0\} \times M, \xi)$ since $\omega|_\xi = (\beta(r_0)\,d\eta + r_0\,d\alpha)|_\xi$ also tames J_ξ. Notice that for r_0 sufficiently large, the $d\eta$ term disappears, so ω has a primitive that restricts to $\{r_0\} \times M$ as a contact form for ξ, meaning we have a *strong* filling of this hypersurface. The desired deformation of ω can therefore be defined by pulling back via a smooth family of diffeomorphisms $(-\epsilon, 0] \to (-\epsilon, r_0]$, where r_0 varies from 0 to a sufficiently large constant. □

Unlike strong cobordisms, being a weak cobordism is an open condition: if (W, ω) is a weak cobordism, then so is $(W, \omega + \epsilon\sigma)$ for any $\epsilon > 0$ sufficiently small and a closed 2-form σ, which need not be exact at ∂W. As a consequence, the cylindrical ends of a completed weak cobordism cannot always be deformed to look like the symplectization of a contact manifold. This is where Definition 13.17 comes in useful. The proof of the next lemma is very much analogous to Proposition 13.19.

LEMMA 13.20 ([**MNW13**, Lemma 2.10]). *Suppose (W, ω) is a weak filling of a $(2n-1)$-dimensional contact manifold (M, ξ), α is a contact form for ξ and Ω is a closed 2-form on M with $[\Omega] = [\omega|_{TM}] \in H^2_{\mathrm{dR}}(M)$. Then for any constant $c > 0$ sufficiently large, after a homotopy of ω through a family of symplectic forms that vary only in a collar neighborhood of ∂W and define weak fillings of (M, ξ), $\omega|_{TM} = \Omega + c\,d\alpha$.* □

The following result then provides a suitable model that can be used as Ω in the above lemma when $\omega|_{TM}$ is nonexact. The statement below is restricted to the case where $[\omega|_{TM}]$ is a rational cohomology class; the reason for this is that it relies on a Donaldson-type existence result for contact submanifolds obtained as zero sets of approximately holomorphic sections, due to Ibort, Martiínez-Torres and Presas [**IMTP00**]. It seems likely that the rationality condition could be lifted with more work, and in dimension three this is known to be true; see [**NW11**, Prop. 2.6].

LEMMA 13.21 ([**CV15**, Prop. 2.18]). *For any rational cohomology class $\eta \in H^2(M; \mathbb{Q})$ on a closed $(2n-1)$-dimensional contact manifold (M, ξ), there exists a closed 2-form Ω and a nondegenerate contact form α for ξ such that (Ω, α) is a stable Hamiltonian structure.* □

Combining all of the above results (including Example 13.15) proves:

Proposition 13.22. *Suppose (W, ω) is a weak filling of a $(2n-1)$-dimensional contact manifold (M, ξ) such that $[\omega|_{TM}] \in H^2_{\mathrm{dR}}(M)$ is rational or $n = 2$. Fix a nondegenerate contact form α for ξ. Then there exists a closed 2-form Ω cohomologous to $\omega|_{TM}$ such that $\mathcal{H} := (\Omega, \alpha)$ is a stable Hamiltonian structure, and for all $c > 0$ sufficiently large, ω can be deformed in a collar neighborhood of ∂W, through a family of symplectic forms defining weak fillings of (M, ξ), to a new weak filling for which ∂W is also stable and inherits the stable Hamiltonian structure $\mathcal{H}_c := (\Omega + c\, d\alpha, \alpha)$. In particular, after this deformation, the completed stable filling admits a tame almost complex structure that is pseudoconvex near infinity and may be assumed C^∞-close to any given $J \in \mathcal{J}(\alpha)$.* □

We will use this in §13.3.2 to define obstructions to weak fillability via SFT.

Remark 13.23. There is apparently no analogue of Propositions 13.19 and 13.22 for negative boundary components of weak cobordisms, and this is one of a few reasons why they are not often discussed. For example, if L is a Lagrangian torus in the standard symplectic 4-ball $\mathbb{D}^4$, then the complement of a neighborhood of L in B^4 defines a strong cobordism from the standard contact $\mathbb{T}^3$ to S^3. The symplectic form on this cobordism is obviously exact, but if any result analogous to Proposition 13.19 were to hold at the concave boundary, then we could deform it to a Liouville cobordism. No such Liouville cobordism exists—it would imply that the Lagrangian $L \subset B^4$ is exact, thus violating Gromov's famous theorem [**Gro85**] on exact Lagrangians.

13.2.2. Counting disconnected index 0 curves. Fix a symplectic cobordism (W, ω) with stable boundary $\partial W = -M_- \coprod M_+$ carrying stable Hamiltonian structures $\mathcal{H}_\pm = (\omega_\pm, \lambda_\pm)$, along with a generic almost complex structure J that is ω-tame on W, belongs to $\mathcal{J}(\mathcal{H}_\pm)$ on the cylindrical ends, and is pseudoconvex near infinity. This implies that the stabilizing 1-forms $\lambda_\pm$ are both contact forms. Let us also assume that the $\lambda_\pm$ are both nondegenerate, and that both J and the induced $\mathbb{R}$-invariant almost complex structures $J_\pm \in \mathcal{J}(\mathcal{H}_\pm)$ satisfy Assumption 12.1. These assumptions mean that all the usual SFT chain complexes are well defined for $(M_\pm, \mathcal{H}_\pm, J_\pm; R_\pm)$ with any choice of coefficient ring $R_\pm = \mathbb{Q}[H_2(M_\pm)/G_\pm]$. Denote the corresponding SFT generating functions by $\mathbf{H}_\pm$.

Recall from Lecture 12 that the auxiliary data on M_+ and M_- includes a choice of capping surface C_γ for each closed Reeb orbit γ (or a capping *chain* with rational coefficients if $H_1(M_\pm)$ has torsion). These surfaces satisfy

$$\partial C_\gamma = \sum_i m_i [C_i^\pm] - [\gamma],$$

where the m_i are integers and $C_i^\pm \subset M_\pm$ are fixed curves forming a basis of $H_1(M_\pm)$. Assume $H_1(W)$ and $H_1(M_\pm)$ are all torsion free. (Only minor modifications are needed if this assumption fails to hold, see Remark 13.1.) We can then fix the following additional auxiliary data:

(1) A collection of **reference curves**

$$S^1 \cong C_1, \ldots, C_r \subset W$$

whose homology classes from a basis of $H_1(W)$.

(2) A unitary trivialization of TW along each of the reference curves $C_1, \dots, C_r$, denoted collectively by τ.

(3) A **spanning surface** $S_i^{\pm}$ for each of the positive/negative reference curves $C_i^{\pm} \subset M_{\pm}$, i.e. a smooth map of a compact and oriented surface with boundary into W such that

$$\partial S_i^{\pm} = \sum_j m_{ji}[C_j] - [C_i^{\pm}]$$

in the sense of singular 2-chains, where $m_{ji} \in \mathbb{Z}$ are the unique coefficients with $[C_i^{\pm}] = \sum_j m_{ji}[C_j] \in H_1(W)$.

Now to any collections of orbits $\boldsymbol{\gamma}^{\pm} = (\gamma_1^{\pm}, \dots, \gamma_{k_{\pm}}^{\pm})$ in $M_{\pm}$ and a relative homology class $A \in H_2(W, \bar{\boldsymbol{\gamma}}^+ \cup \bar{\boldsymbol{\gamma}}^-)$ with $\partial A = \sum_i [\gamma_i^+] - \sum_j [\gamma_j^-]$, we can associate an absolute homology class in two steps: first add A to suitable sums of the capping surfaces $C_{\gamma_i^{\pm}}$ producing a 2-chain whose boundary is a linear combination of positive and negative reference curves, then add a suitable linear combination of the $S_i^{\pm}$ so that the boundary becomes the *trivial* linear combination of $C_1, \dots, C_r$. With this understood, we can now associate an absolute homology class

$$[u] \in H_2(W)$$

to any asymptotically cylindrical J-holomorphic curve $u : (\dot{\Sigma}, j) \to (\widehat{W}, J)$, and this defines the notation $\mathcal{M}_{g,m}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ with $A \in H_2(W)$. We now require the trivializations of $\xi_{\pm}$ along each $C_i^{\pm}$ to be compatible with τ in the sense that they extend to trivializations of TW along the capping surfaces $S_i^{\pm}$. With this convention, the Fredholm index formula takes the expected form

$$\operatorname{ind}(u) = (n-3)\chi(\dot{\Sigma}) + 2c_1([u]) + \sum_{i=1}^{k_+} \mu_{\mathrm{CZ}}(\gamma_i) - \sum_{j=1}^{k_-} \mu_{\mathrm{CZ}}(\gamma_j).$$

If $H_1(W)$ has torsion, then this whole discussion can be adapted as in §12.7.1 by replacing integral homology with rational homology and capping surfaces with capping chains, and the Conley-Zehnder indices can be defined modulo 2.

We will also need to impose a compatibility condition relating the coefficient rings $R_{\pm} = \mathbb{Q}[H_2(M_{\pm})/G_{\pm}]$ to a corresponding choice on the cobordism W. Choose a subgroup $G \subset H_2(W)$ such that

$$\langle [\omega], A \rangle = 0 \quad \text{for all} \quad A \in G, \tag{13.10}$$

and such that the maps $H_2(M_{\pm}) \to H_2(W)$ induced by the inclusions $M_{\pm} \hookrightarrow W$ send $G_{\pm}$ into G. If $[\omega] \neq 0 \in H^2_{\mathrm{dR}}(W)$, then we will have to deal with noncompact sequences of J-holomorphic curves that have unbounded energy, so it becomes necessary to "complete" R to a **Novikov ring** $\bar{R}$, which contains R but also includes infinite formal sums

$$\sum_{i=1}^{\infty} c_i e^{A_i} \quad \text{such that} \quad \langle [\omega], A_i \rangle \to +\infty \text{ as } i \to \infty.$$

Note that the evaluation $\langle [\omega], A\rangle \in \mathbb{R}$ is well defined for $A \in H_2(W)/G$ due to (13.10).

Analogously to our definition of $\mathbf{H}$ in Lecture 12, the generating function for index 0 curves in $\widehat{W}$ is defined as a formal power series in the variables $\hbar$, q_γ (for orbits in M_-), and p_γ (for orbits in M_+), with coefficients in $\bar{R}$:

$$\mathbf{F} = \sum_{u \in \mathcal{M}_0^\sigma(J)} \frac{\epsilon(u)}{|\operatorname{Aut}^\sigma(u)|} \hbar^{g-1} e^A q^{\gamma^-} p^{\gamma^+}, \tag{13.11}$$

where $\mathcal{M}_0^\sigma(J)$ denotes the moduli space of stable connected J-holomorphic curves u in $\widehat{W}$ with $\operatorname{ind}(u) = 0$ and only good asymptotic orbits, modulo permutations of the punctures, and for each u:

- g is the genus of u;
- A is the equivalence class of $[u] \in H_2(W)$ in $H_2(W)/G$;
- $\boldsymbol{\gamma}^\pm = (\gamma_1^\pm, \dots, \gamma_{k_\pm}^\pm)$ are the asymptotic orbits of u after arbitrarily fixing orderings of its positive and negative punctures;
- $\epsilon(u) \in \{1, -1\}$ is the sign of u as a point in the 0-dimensional component of $\mathcal{M}^\$(J)$ (after choosing an ordering of the punctures and asymptotic markers), relative to a choice of coherent orientations on $\mathcal{M}^\$(J)$.

As usual, the product $\epsilon(u) q^{\gamma^-} p^{\gamma^+}$ is independent of choices. Similarly to our discussion of $\mathbf{H}$ in §12.3, writing $\mathbf{F}$ in the form of (13.11) only makes sense under the unrealistic assumption that all index 0 curves in $\widehat{W}$ (including multiple covers) are regular, but under the same assumption, the formula is equivalent to

$$\mathbf{F} = \sum_{g, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-} \frac{\#\mathcal{M}_{g,0}^\$(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)}{k_+! k_-! \kappa_{\boldsymbol{\gamma}^+} \kappa_{\boldsymbol{\gamma}^-}} \hbar^{g-1} e^A q^{\gamma^-} p^{\gamma^+}, \tag{13.12}$$

in which the sum ranges over all integers $g \geqslant 0$, equivalence classes $A \in H_2(M)/G$, and ordered tuples of Reeb orbits $\boldsymbol{\gamma}^+ = (\gamma_1^+, \dots, \gamma_{k_+}^+)$ in M_+ and $\boldsymbol{\gamma}^- = (\gamma_1^-, \dots, \gamma_{k_-}^-)$ for which $\operatorname{vir-dim} \mathcal{M}_{g,0}^\$(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) = 0$. In this form, the transversality issues can be dealt with as in §12.4.3 by defining $\#\mathcal{M}_{g,0}^\$(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-) \in \mathbb{Q}$ as a count of a compact weighted branched 0-manifold obtained as the zero set of a generic multivalued inhomogeneous perturbation of the nonlinear Cauchy-Riemann operator.

REMARK 13.24. The word "stable" in the above definition deserves further comment. Recall from Remark 6.31 that a smooth connected J-holomorphic curve without marked points is called **stable** if either it is nonconstant or its domain has negative Euler characteristic; equivalently, u is stable if and only if $\operatorname{Aut}(u)$ is finite. This condition was not relevant in our definition of $\mathbf{H}$ because constant curves never have index 1, but they *can* have index 0, thus the definition of $\mathbf{F}$ explicitly excludes closed J-holomorphic spheres and tori with $[u] = 0$, and $\#\mathcal{M}_{g,0}^\$(J, 0, \emptyset, \emptyset)$ in (13.12) is defined to be 0 for $g = 0, 1$. This is appropriate in light of the SFT compactness theorem, in which smooth non-stable curves can never appear as components in stable holomorphic buildings. What may seem a bit confusing at first is that if $\dim W = 6$, then constant curves of genus $g \geqslant 2$ have index 0 and are stable, so (13.12) does count them. But constant curves can never be isolated, so this is yet another case where Assumption 12.1 on transversality definitely cannot be satisfied. The contributions $\#\mathcal{M}^{\$}_{g,0}(J, 0, \varnothing, \varnothing)$ for $g \geqslant 2$ can be defined in this case using multivalued inhomogeneous perturbations as in §12.4.3, meaning that we count (with signs and rational weights) nullhomologous solutions $u : \Sigma_g \to \widehat{W}$ to equations of the form $\bar{\partial}_J u = \nu$.

We shall regard $\mathbf{F}$ as an element in an enlarged operator algebra that includes q and p variables for good orbits in both M_+ and M_-, related to each other by the supercommutation relations

$$[p_{\gamma_-}, q_{\gamma_+}] = [p_{\gamma_+}, q_{\gamma_-}] = [q_{\gamma_-}, q_{\gamma_+}] = [p_{\gamma_-}, p_{\gamma_+}] = 0$$

whenever γ_- is an orbit in M_- and γ_+ is an orbit in M_+. Since all curves counted by $\mathbf{F}$ have index 0, $\mathbf{F}$ is homogeneous with degree

$$|\mathbf{F}| = 0.$$

Notice that for any fixed monomial $q^{\gamma^-} p^{\gamma^+}$, the corresponding set of curves in $\mathcal{M}^{\sigma}_0(J)$ may be infinite if ω is nonexact, but SFT compactness implies that the set of such curves with any given bound on $\langle [\omega], [u] \rangle$ is bounded. As a consequence, the coefficient of $\hbar^{g-1} q^{\gamma^-} p^{\gamma^+}$ in $\mathbf{F}$ belongs to the Novikov ring $\bar{R}$.

REMARK 13.25. In the case $\partial M_+ = \partial M_- = \varnothing$, there are no q or p variables, so $\mathbf{F}$ is an element of $\frac{1}{\hbar}\bar{R}[[\hbar]]$ recording the counts of closed J-holomorphic curves of all genera in all homology classes in the closed symplectic manifold (W, ω). These are, in other words, Gromov-Witten invariants of (W, ω), and $\mathbf{F}$ is a simple version of the so-called **Gromov-Witten potential**.[5]

REMARK 13.26. In many standard presentations of Gromov-Witten theory, the moduli space $\mathcal{M}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ in (13.12) is replaced with the compactification $\overline{\mathcal{M}}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$. This distinction is academic in our presentation, because if Assumption 12.1 on transversality holds, then $\overline{\mathcal{M}}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ does not contain any nodal curves, for dimensional reasons, so the two spaces are the same finite set. It is important to keep in mind however that in the real world where transversality cannot be achieved merely by perturbing J, there exist situations (see Example 13.27 below) where $\mathcal{M}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ is empty but $\overline{\mathcal{M}}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ is not, in which case $\#\mathcal{M}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ may be nonzero. The latter can happen because under generic inhomogeneous perturbations, there may exist solutions u of the perturbed equation $\bar{\partial}_J u = \nu$ that converge to nodal curves or buildings (but not to smooth curves) as the perturbation ν is turned off. The moral is: if you want to deduce a computation of $\#\mathcal{M}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$ without actually carrying out abstract perturbations, then you generally need to understand the compactified space $\overline{\mathcal{M}}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$, not just $\mathcal{M}^{\$}_{g,0}(J, A, \boldsymbol{\gamma}^+, \boldsymbol{\gamma}^-)$, even if the latter is finite or empty.

[5]The more elaborate standard version of the Gromov-Witten potential also involves moduli spaces with marked points, and includes extra generators to keep track of intersection numbers of the evaluation map with homology cycles in the target manifold. One can similarly build this type of information into the algebraic formalism of SFT, making it a direct generalization of Gromov-Witten theory—see [**EGH00**, §2.2–2.3] for details.

EXAMPLE 13.27. Consider the space of closed genus $g \geqslant 1$ curves of degree 1 in (S^2, i): $\mathcal{M}_{g,0}(i, [S^2], \varnothing, \varnothing)$ is empty since a holomorphic map $(\Sigma_g, j) \to (S^2, i)$ of degree 1 would have to be a diffeomorphism, but $\overline{\mathcal{M}}_{g,0}(i, [S^2], \varnothing, \varnothing)$ contains a nodal curve u with one spherical component on which u is the identity map, attached by a node to a genus g component on which u is constant. Nodal curves of this type are known to make nontrivial contributions to Gromov-Witten invariants; see e.g. [**MS12**, Example 8.6.12].

Consider next the series

$$\exp(\mathbf{F}) := \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{F}^k.$$

We will be able to view this as a formal power series in q and p variables and a formal Laurent series in $\hbar$ with coefficients in $\bar{R}$, though it is not obvious at first glance whether its coefficients are in any sense finite. We will deduce this after interpreting it as a count of *disconnected* index 0 curves: first, write

$$\exp(\mathbf{F}) = \sum_{k=0}^{\infty} \frac{1}{k!} \Bigg(\sum_{(u_1,\dots,u_k)\in(\mathcal{M}_0^\sigma(J))^k} \frac{\epsilon(u_1)\dots\epsilon(u_k)}{|\operatorname{Aut}^\sigma(u_1)|\dots|\operatorname{Aut}^\sigma(u_k)|} \hbar^{g_1+\dots+g_k-k} e^{A_1+\dots A_k} \cdot q^{\gamma_1^-} p^{\gamma_1^+} \dots q^{\gamma_k^-} p^{\gamma_k^+} \Bigg).$$

Observe that since each of the curves $u_i \in \mathcal{M}_0^\sigma(J)$ in this expansion has index 0, the monomials $q^{\gamma_i^-} p^{\gamma_i^+}$ all have even degree and thus the order in which they are written does not matter. Now for a given collection of distinct curves $v_1, \dots, v_N$ and integers $k_1, \dots, k_N \in \mathbb{N}$ with $k_1 + \dots + k_N = k$, the various permutations of

$$(u_1, \dots, u_k) := (\underbrace{v_1, \dots, v_1}_{k_1}, \dots, \underbrace{v_N, \dots, v_N}_{k_N}) \in (\mathcal{M}_0^\sigma(J))^{\times k}$$

occur $\frac{k!}{k_1!\dots k_N!}$ times in the above sum, so if we forget the ordering, then the contribution of this particular k-tuple of curves to $\exp(\mathbf{F})$ is

$$\frac{\epsilon(u_1)\dots\epsilon(u_k)}{k_1!\dots k_N!|\operatorname{Aut}^\sigma(u_1)|\dots|\operatorname{Aut}^\sigma(u_k)|} \hbar^{g_1+\dots+g_k-k} e^{A_1+\dots+A_k} q^{\gamma_1^-} p^{\gamma_1^+} \dots q^{\gamma_k^-} p^{\gamma_k^+}.$$

Notice next that the denominator $k_1!\dots k_N!|\operatorname{Aut}^\sigma(u_1)|\dots|\operatorname{Aut}^\sigma(u_k)|$ is the order of the automorphism group of the *disconnected* curve formed by the disjoint union of $u_1, \dots, u_k$: the extra factors $k_i!$ come from automorphisms that permute connected components of the domain. Thus $\exp(\mathbf{F})$ can also be written as in (13.11), but with $\mathcal{M}_0^\sigma(J)$ replaced by the moduli space of *potentially disconnected* index 0 curves with unordered punctures, and $g-1$ generalized to $g_1 + \dots + g_k - k$ for any curve that has k connected components of genera $g_1, \dots, g_k$. One subtlety that was glossed over in the above discussion: the sum also includes the unique curve with *zero* components, i.e. the "empty" J-holomorphic curve, which appears as the initial 1 in the series expansion of $\exp(\mathbf{F})$.

With this interpretation of $\exp(\mathbf{F})$ understood, we can now address the possibility that the infinite sum defining $\exp(\mathbf{F})$ might include infinitely many terms for a given monomial $\hbar^m q^{\gamma^-} p^{\gamma^+}$, i.e. that there are infinitely many disconnected index 0 curves with fixed asymptotic orbits and a fixed sum of the genera minus the number of connected components. We claim that this can indeed happen, but only if the curves belong to a sequence of homology classes $A_i \in H_2(M)/G$ with $\langle [\omega], A_i \rangle \to \infty$, hence the coefficient of $\hbar^m q^{\gamma^-} p^{\gamma^+}$ in $\exp(\mathbf{F})$ belongs to the Novikov ring $\bar{R}$. The danger here comes only from *closed* curves, since a disjoint union of two curves with punctures always has strictly more punctures. Notice also that for any given tuples of orbits $\boldsymbol{\gamma}^{\pm}$, there exists a number $c \in \mathbb{R}$ depending only on these orbits and the chosen capping surfaces such that every (possibly disconnected) J-holomorphic curve $u : \dot{\Sigma} \to \widehat{W}$ asymptotic to $\boldsymbol{\gamma}^{\pm}$ satisfies

$$\langle [\omega], [u] \rangle \geqslant c.$$

This follows from the fact that the integral of ω over the relative homology class of u always has a nonnegative integrand.

LEMMA 13.28. *Given constants $C \in \mathbb{R}$ and $k \in \mathbb{Z}$, there exists a number $N \in \mathbb{N}$ such that if $u : (\Sigma, j) \to (\widehat{W}, J)$ is a closed J-holomorphic curve satisfying $\langle [\omega], [u] \rangle \leqslant C$, with m connected components, all stable, of genera $g_1, \ldots, g_m$ satisfying $g_1 + \ldots + g_m - m = k$, then $m \leqslant N$.*

PROOF. Note first that for each integer $g \geqslant 0$, there is an **energy threshold**, i.e. a constant $c_g > 0$ such that every nonconstant closed and connected J-holomorphic curve $u : \Sigma \to \widehat{W}$ of genus g has

$$\langle [\omega], [u] \rangle \geqslant c_g.$$

This is an easy consequence of Gromov's compactness theorem, which is just the closed case of SFT compactness. Indeed, if there were no such constant, then we would find a sequence $u_k : \Sigma \to \widehat{W}$ of connected closed curves with genus g such that

$$E(u_k) = \sup_{\varphi} \int_{\Sigma} u_k^* \omega_{\varphi} = \sup_{\varphi} \langle [\omega_{\varphi}], [u_k] \rangle = \langle [\omega], [u_k] \rangle \to 0,$$

where we have used the fact that the cohomology class $[\omega_{\varphi}] \in H^2_{\mathrm{dR}}(\widehat{W})$ is independent of the choice of auxiliary function $\varphi : \mathbb{R} \to (-\epsilon, \epsilon)$ and matches $[\omega]$ under the isomorphism $H^2_{\mathrm{dR}}(\widehat{W}) = H^2_{\mathrm{dR}}(W)$ defined via a deformation retraction of the cylindrical ends. There is then a subsequence of u_k that converges to a stable holomorphic building in which every component has zero energy—in other words, the limit is a nodal curve whose components are all constant, and its total homology class is therefore zero. But the latter cannot happen unless $[u_k] = 0 \in H_2(\widehat{W})$ for all k sufficiently large, and since $E(u_k)$ for closed curves depends only on the homology class, this would imply $E(u_k) = 0$ as well, so that the curves u_k are also constant, giving a contradiction.

Now if u is a disconnected curve satisfying the stated conditions, the bound on $\langle [\omega], [u] \rangle$ combines with the energy threshold to give bounds on the number of connected components of u with genus 0 or 1, as stability requires these to be nonconstant. All other components contribute positively to the left hand side of the relation $\sum_{i=1}^{m}(g_i - 1) = k$, so this implies a universal bound on m. $\square$

COROLLARY 13.29. *Fix constants $C \in \mathbb{R}$ and $k \in \mathbb{Z}$, and tuples of Reeb orbits $\boldsymbol{\gamma}^{\pm}$, and assume that the usual transversality conditions hold. Then there exist at most finitely many potentially disconnected J-holomorphic curves $u : \dot{\Sigma} \to \widehat{W}$ with index 0 such that the number of connected components m and the genera $g_1, \dots, g_m$ of its components satisfy $g_1 + \dots + g_m - m = k$.* $\square$

COROLLARY 13.30. *The expression $\exp(\mathbf{F})$ is a formal power series in q and p variables and a formal Laurent series in $\hbar$, with coefficients in the Novikov ring $\bar{R}$.* $\square$

For a compact symplectic manifold (W, ω) with no boundary, the coefficients in $\mathbf{F}$ and $\exp(\mathbf{F})$ are symplectic invariants; in particular they are independent of the choice of almost complex structure J, and depend in fact only on the deformation class of the symplectic structure ω. One proves this by choosing a generic homotopy $\{J_s\}_{s \in [0,1]}$ of compatible almost complex structures associated to a deformation of ω and viewing the resulting parametric moduli space as an oriented cobordism between the moduli spaces for J_0 and J_1. In the absence of transversality problems, the cobordism is compact due to a dimension counting argument: the formation of nodal curves is a codimension two phenomenon, so in a 1-dimensional parametric moduli space, it never happens. But when $\partial W \neq \varnothing$, breaking can also occur, giving rise to holomorphic buildings with multiple levels that form codimension one strata of the compatification. This kills the dimension counting argument, with the consequence that $\mathbf{F}$ and $\exp(\mathbf{F})$ are not independent of the choice of J. Instead, the compactness and gluing theory give rise to an algebraic relation between $\mathbf{F}$ and $\mathbf{H}_{\pm}$.

Consider the 1-dimensional moduli space of connected index 1 curves in $\widehat{W}$ with genus g. The boundary points of the compactification of this space consist of two types of buildings:

TYPE 1 : A main level of index 0 and an upper level of index 1;
TYPE 2 : A main level of index 0 and a lower level of index 1.

This is clear under the usual transversality assumptions since regular curves in $\widehat{W}$ must have index at least 0, while regular curves in the symplectizations $\mathbb{R} \times M_{\pm}$ have index at least 1 unless they are trivial cylinders. The building must also be connected and have arithmetic genus g, but there is nothing to guarantee that each individual level is connected. In fact, we already saw this issue in Lecture 12 when proving $\mathbf{H}^2 = 0$, but it was simpler to deal with there, because disconnected regular curves of index 1 in a symplectization always have a unique nontrivial component, while the rest are trivial cylinders. In the cobordism $\widehat{W}$, on the other hand, a disconnected index 0 curve can be formed by any disjoint union of index 0 curves, all of which are nontrivial. The resulting algebraic relation therefore involves $\exp(\mathbf{F})$ instead of $\mathbf{F}$. Since the union of all buildings of types 1 and 2 described above forms the boundary of a compact oriented 1-manifold,[6] the count of these buildings is zero,

[6]Or a weighted branched 1-manifold, if transversality problems are solved via multivalued perturbations.

and this fact is encoded in the so-called **master equation**

$$\mathbf{H}_- \exp(\mathbf{F})|_{p_-=0} - \exp(\mathbf{F})\mathbf{H}_+|_{q_+=0} = 0, \tag{13.13}$$

where the expressions "$p_- = 0$" and "$q_+ = 0$" mean that we discard all terms in $\mathbf{H}_- \exp(\mathbf{F}) - \exp(\mathbf{F})\mathbf{H}_+$ containing any variables p_γ for orbits in M_- or q_γ for orbits in M_+. The resulting expression is a formal power series in q variables for orbits in M_- and p variables for orbits in M_+, representing a count of generally disconnected index 1 holomorphic buildings in $\widehat{W}$ with the specified asymptotics. The various ways to form such buildings by choices of gluings is again encoded by the commutator algebra. The master equation (13.13) can be used to prove the chain map property for counts of curves in cobordisms, thus it is an essential piece of the invariance proof for each of the homology theories introduced above.

EXERCISE 13.31. Fill in the details of the proof of (13.13), modulo transversality.

13.3. Full SFT as a BV_∞-algebra

In this section we discuss the specific theory $H_*^{\mathrm{SFT}}(M,\xi;R)$, defined as the homology of the chain complex $(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$. The case $G = H_2(M)$ with trivial group ring coefficients $\mathbb{Q}[H_2(M)/G] = \mathbb{Q}$ will be abbreviated as

$$H_*^{\mathrm{SFT}}(M,\xi) := H_*^{\mathrm{SFT}}(M,\xi;\mathbb{Q}).$$

As we defined it, $\mathbf{D}_{\mathrm{SFT}}$ acts on $\mathcal{A}[[\hbar]]$ by treating the generating function $\mathbf{H}$ as a differential operator via the substitution

$$p_\gamma = \kappa_\gamma \hbar \frac{\partial}{\partial q_\gamma}. \tag{13.14}$$

According to [CL09], this makes $(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}})$ into a BV_∞-algebra; we'll have no particular need to discuss here what that means, but one convenient feature is the expansion

$$\mathbf{D}_{\mathrm{SFT}} = \frac{1}{\hbar}\sum_{k=1}^{\infty} \mathbf{D}_{\mathrm{SFT}}^{(k)} \hbar^k, \tag{13.15}$$

in which each $\mathbf{D}_{\mathrm{SFT}}^{(k)} : \mathcal{A} \to \mathcal{A}$ is a differential operator *of order* $\leqslant k$ (see [CL09, §5]). For each $k \in \mathbb{N}$, $\mathbf{D}_{\mathrm{SFT}}^{(k)}$ is a count of all index 1 holomorphic curves that have genus $g \geqslant 0$ and $m \geqslant 1$ positive punctures such that $g + m = k$. In particular, $\mathbf{D}_{\mathrm{SFT}}^{(1)}$ is simply the contact homology differential ∂_{CH}, and the expansion (13.15) implies together with $\mathbf{D}_{\mathrm{SFT}}^2 = 0$ that $(\mathbf{D}_{\mathrm{SFT}}^{(1)})^2 = 0$, hence we again see the chain complex for contact homology hidden inside a version of the "full" SFT complex.

13.3.1. Cobordism maps and invariance. One can use the master equation (13.13) to prove invariance of $H_*^{\mathrm{SFT}}(M,\xi;R)$ by a straightforward generalization of the usual Floer-theoretic argument. Suppose $(W, d\lambda)$ is an exact symplectic cobordism from (M_-,ξ_-) to (M_+,ξ_+) with $\lambda|_{TM_\pm} = \alpha_\pm$, and choose a generic almost complex structure J on $\widehat{W}$ that is $d\lambda$-compatible on W and restricts to the cylindrical ends as generic elements $J_\pm \in \mathcal{J}(\alpha_\pm)$. Let $(\mathcal{A}^\pm[[\hbar]], \mathbf{D}_{\mathrm{SFT}}^\pm)$ denote the chain complexes associated to the data $(\alpha_\pm, J_\pm)$, and for simplicity in this initial discussion, choose the trivial coefficient ring $R = \mathbb{Q}$ for both. We then define a map

$$\mathbf{\Phi} : \mathcal{A}^+[[\hbar]] \to \mathcal{A}^-[[\hbar]] : \mathbf{f} \mapsto \exp(\mathbf{F})\mathbf{f}|_{q_+=0},$$

where the generating function $\exp(\mathbf{F})$ is regarded as a differential operator via the substitution (13.14), with $e^A := 1$ for all $A \in H_2(W)$ since we are using trivial coefficients, and "$q_+ = 0$" means that after applying $\exp(\mathbf{F})$ to change $\mathbf{f}$ into a function of q variables for orbits in both M_+ and M_-, we discard all terms that involve orbits in M_+. The exactness of the cobordism implies that negative powers of $\hbar$ do not appear in $\mathbf{\Phi f}$, thus producing an element of $\mathcal{A}^-[[\hbar]]$: indeed, since there are no holomorphic curves in $\widehat{W}$ without positive punctures, every term in $\mathbf{F}$ contains at least one p variable, so that negative powers of $\hbar$ do not appear in $\exp(\mathbf{F})$ after applying (13.14).

The master equation for $\mathbf{F}$ now translates into the fact that $\mathbf{\Phi}$ is a chain map,

$$\mathbf{D}^-_{\mathrm{SFT}} \circ \mathbf{\Phi} - \mathbf{\Phi} \circ \mathbf{D}^+_{\mathrm{SFT}},$$

thus it descends to homology. The geometric meaning of $\mathbf{\Phi}$ is straightforward to describe: analogous to (12.11) in Lecture 12, we can write

$$\mathbf{\Phi} q^\gamma = \sum_{g=0}^{\infty} \sum_{\gamma'} \hbar^{g+k-1} n_g(\boldsymbol{\gamma}, \boldsymbol{\gamma}', k) q^{\gamma'}, \tag{13.16}$$

where $n_g(\boldsymbol{\gamma}, \boldsymbol{\gamma}', k)$ is a product of some combinatorial factors with a signed count of disconnected index 0 holomorphic curves with connected components of genera $g_1, \ldots, g_m$ satisfying $g_1 + \ldots + g_m - m = g - 1$, and with positive ends at $\boldsymbol{\gamma}$ and negative ends at $\boldsymbol{\gamma}'$, where k is the number of positive ends.

Let's discuss two applications of the cobordism map $\mathbf{\Phi}$. First, note that if W is a *trivial* symplectic cobordism $[0, 1] \times M$, then the above discussion can easily be generalized with $(\mathcal{A}^\pm, \mathbf{D}^\pm_{\mathrm{SFT}})$ both defined over the same group ring $R = \mathbb{Q}[H_2(M)/G]$ for any choice of $G \subset H_2(M)$. There is no need to consider a Novikov ring in defining $\mathbf{F}$ here since the cobordism is exact. We therefore obtain a chain map with arbitrary group ring coefficients, and extending this discussion along standard Floer-theoretic principles will imply that the chain map is an isomorphism: this can be used in particular to prove that $H^{\mathrm{SFT}}_*(M, \xi; R)$ does not depend on the choices of contact form and almost complex structure. There are two additional steps involved in this argument: first, one needs to use a chain homotopy to prove that $\mathbf{\Phi}$ does not depend on the choice of almost complex structure J on $\widehat{W}$. Given a generic homotopy $\{J_s\}_{s\in[0,1]}$, the chain homotopy map

$$\mathbf{\Psi} : \mathcal{A}^+[[\hbar]] \to \mathcal{A}^-[[\hbar]]$$

is defined as a differential operator in the same manner as $\mathbf{\Phi}$, but counting pairs (s, u) where $s \in [0, 1]$ is a parameter value for which J_s is nongeneric and u is a disconnected J_s-holomorphic curve in $\widehat{W}$ with index -1. We saw how this works for cylindrical contact homology in Lecture 10, but there is a new subtlety now that should be mentioned: in principle, a *disconnected* index -1 curve in $\widehat{W}$ could have arbitrarily many components, including perhaps many with index -1 and others with arbitrarily large index. Even worse, the compactified 1-dimensional space of pairs (s, u) for J_s-holomorphic curves u of index 0 may include buildings that have symplectization levels of index greater than 1, balanced by disjoint unions of many index -1 curves in the main level. This sounds horrible, but it can actually be ignored, for the following reason: first, since there are only finitely many pairs (s, u) where u is a *connected* J_s-holomorphic curve with index -1, one can (if transversality is achievable at all) use a genericity argument to assume without loss of generality that for any given $s \in [0, 1]$, at most *one* connected index -1 curve exists. This means that in any building that has multiple index -1 components, those components are just multiple copies of the same curve. Now, since that curve has odd index, it is represented by a monomial $q^{\gamma^-}p^{\gamma^+}$ that contains an odd number of odd generators, and any nontrivial product of such generators therefore *disappears* in $\mathcal{A}$ since odd generators anticommute with themselves. This algebraic miracle encodes a convenient fact about coherent orientations: whenever one of the horrible buildings described above appears, one can reorder two of the index -1 components to produce from it a different building that lives in a moduli space with the opposite orientation. Gluing this building back together then produces a continuation of the 1-dimensional moduli space, so that the horrible building can actually be interpreted as an "interior" point of the 1-dimensional space, rather than boundary. The actual count of boundary points is then exactly what we want it to be: it is represented algebraically by the chain homotopy relation!

Finally, compositions of cobordism maps can be understood via a stretching argument that is not substantially different from the case of cylindrical contact homology. Since the trivial cobordism with $\mathbb{R}$-invariant data gives a cobordism map that just counts trivial cylinders and is therefore the identity, it follows that cobordism maps relating different pairs of data $(\alpha_\pm, J_\pm)$ are always invertible, and this proves the invariance of $H_*^{\mathrm{SFT}}(M, \xi; R)$.

The second application concerns nontrivial exact cobordisms, and it is immediate from the fact that $\mathbf{\Phi}$ is a chain map:

THEOREM 13.32. *Any exact cobordism $(W, d\lambda)$ from (M_-, ξ_-) to (M_+, ξ_+) gives rise to a $\mathbb{Q}[[\hbar]]$-linear map*

$$H_*^{\mathrm{SFT}}(M_+, \xi_+) \to H_*^{\mathrm{SFT}}(M_-, \xi_-).$$

□

It is much more complicated to say what happens in the event of a nonexact cobordism, but slightly easier if we restrict our attention to fillings, i.e. the case with $M_- = \varnothing$. Assume (W, ω) is a compact symplectic manifold with stable boundary M, inheriting a stable Hamiltonian structure $\mathcal{H} = (\Omega, \alpha)$ for which α is a nondegenerate contact form, and assume also that the completion $\widehat{W}$ admits an almost complex structure J that is ω-tame on W and has a pseudoconvex restriction $J_+ \in \mathcal{J}(\mathcal{H})$ to the cylindrical end. We saw in Proposition 13.22 that these conditions can always be achieved for a weak filling after deforming the symplectic structure. Let

$$G := \ker[\omega] := \{A \in H_2(W) \mid \langle[\omega], A\rangle = 0\},$$

and choose $G_+ \subset H_2(M)$ to be any subgroup such that the map $H_2(M) \to H_2(W)$ induced by the inclusion $M \hookrightarrow W$ sends G_+ into G. In other words, G_+ can be any subgroup of $\ker[\Omega] \subset H_2(M)$. Define the group rings

$$R_+ = \mathbb{Q}[H_2(M)/G_+], \qquad R = \mathbb{Q}[H_2(W)/\ker[\omega]],$$

with the Novikov completion of R denoted by $\bar{R}$. The map $H_2(M)/G_+ \to H_2(W)/G$ induced by $M \hookrightarrow W$ then gives a natural ring homomorphism

$$R_+ \to \bar{R}. \tag{13.17}$$

If ω is not exact, then it may no longer be true that every term in $\mathbf{F}$ has at least one p variable. Let us write

$$\mathbf{F} = \mathbf{F}_0 + \mathbf{F}_1,$$

where $\mathbf{F}_0$ contains no p variables and $\mathbf{F}_1 = \mathcal{O}(p)$, i.e. $\mathbf{F}_0$ counts all closed curves in $\widehat{W}$, and $\mathbf{F}_1$ counts everything else. Since $\mathbf{F}_0$ and $\mathbf{F}_1$ have even degree, they commute, and thus

$$\exp(\mathbf{F}) = \exp(\mathbf{F}_0)\exp(\mathbf{F}_1).$$

where $\exp(\mathbf{F}_0)$ is an invertible element of $\bar{R}[[\hbar, \hbar^{-1}]]$ since $\exp(-\mathbf{F}_0)\exp(\mathbf{F}_0) = 1$. By the master equation,

$$\exp(\mathbf{F}_0)\exp(\mathbf{F}_1)\mathbf{H} = \mathcal{O}(q),$$

hence $\exp(\mathbf{F}_1)\mathbf{H} = \exp(-\mathbf{F}_0)\mathcal{O}(q) = \mathcal{O}(q)$ since $\exp(-\mathbf{F}_0)$ contains no p variables. Using the substitution (13.14), and using (13.17) to map coefficients in R_+ to $\bar{R}$, it follows that $\exp(\mathbf{F}_1)$ gives rise to a differential operator

$$\mathbf{\Phi} : \mathcal{A}[[\hbar]] \to \bar{R}[[\hbar]] : \mathbf{f} \mapsto \exp(\mathbf{F}_1)\mathbf{f}|_{q=0},$$

which is a chain map to the SFT of the empty set with Novikov coefficients, meaning

$$\mathbf{\Phi} \circ \mathbf{D}_{\mathrm{SFT}} = 0.$$

This chain map counts the disconnected index 0 curves in $\widehat{W}$ whose connected components all have at least one positive puncture.

THEOREM 13.33. *Suppose (W, ω) is a compact symplectic manifold with stable boundary $(M, \mathcal{H} = (\Omega, \alpha))$, where α is a nondegenerate contact form, and its completion $\widehat{W}$ admits an almost complex structure that is ω-tame on W and has a generic and pseudoconvex restriction $J_+ \in \mathcal{J}(\mathcal{H})$ to the cylindrical end. Let $\bar{R}$ denote the Novikov completion of $\mathbb{Q}[H_2(W)/\ker[\omega]]$, and let $R_+ = \mathbb{Q}[H_2(M)/G_+]$, where $G_+ \subset H_2(M)$ is any subgroup on which the evaluation of $[\Omega] \in H^2_{\mathrm{dR}}(M)$ vanishes. Then there exists an $\bar{R}[[\hbar]]$-linear map $H_*^{\mathrm{SFT}}(M, \mathcal{H}, J_+; R_+) \to \bar{R}[[\hbar]]$.* □

13.3.2. Algebraic torsion. We can now generalize the notion of algebraic overtwistedness. Notice that since every term in $\mathbf{D}_{\mathrm{SFT}}$ is a differential operator of order at least 1,

$$\mathbf{D}_{\mathrm{SFT}}\mathbf{f} = 0 \quad \text{for all} \quad \mathbf{f} \in R[[\hbar]],$$

hence every element of the extended coefficient ring $R[[\hbar]]$ represents an element of $H_*^{\mathrm{SFT}}(M, \xi; R)$ that may or may not be trivial. Since $\mathbf{D}_{\mathrm{SFT}}$ commutes with all elements of $R[[\hbar]]$, the subset consisting of elements that are trivial in homology forms an ideal. The following definition originates in [LW11].

DEFINITION 13.34. We say that a closed contact manifold (M, ξ) has **algebraic torsion of order** k (or k-*torsion* for short) with coefficients in R if

$$[\hbar^k] = 0 \in H_*^{\mathrm{SFT}}(M, \xi; R).$$

The numerical invariant

$$\mathrm{AT}(M, \xi; R) \in \mathbb{N} \cup \{0, \infty\}$$

is defined to be the smallest integer k such that (M, ξ) has algebraic k-torsion but no $(k-1)$-torsion, or ∞ if there is no algebraic torsion of any order.

Several consequences of algebraic torsion can be read off quickly from the properties of SFT cobordism maps. Consider first the case of trivial coefficients $R = \mathbb{Q}$, which we shall refer to as **untwisted** algebraic torsion and abbreviate

$$\mathrm{AT}(M, \xi) := \mathrm{AT}(M, \xi; \mathbb{Q}).$$

If (W, ω) is a strong filling of (M, ξ), then the hypotheses of Theorem 13.33 are fulfilled even with $G_+ = H_2(M)$ since ω is exact at the boundary, thus we obtain a $\mathbb{Q}[[\hbar]]$-linear map $H_*^{\mathrm{SFT}}(M, \xi) \to \bar{R}[[\hbar]]$, with $\bar{R}$ denoting the Novikov completion of $\mathbb{Q}[H_2(W)/\ker[\omega]]$. If $[\hbar^k] = 0 \in H_*^{\mathrm{SFT}}(M, \xi)$, then the cobordism map implies a contradiction since $\hbar^k$ does not equal 0 in $\bar{R}[[\hbar]]$. Similarly, if $(W, d\lambda)$ is an exact cobordism from (M_-, ξ_-) to (M_+, ξ_+), then the cobordism map $H_*^{\mathrm{SFT}}(M_+, \xi_+) \to H_*^{\mathrm{SFT}}(M_-, \xi_-)$ of Theorem 13.32 is also $\mathbb{Q}[[\hbar]]$-linear, and thus any algebraic k-torsion in (M_+, ξ_+) is inherited by (M_-, ξ_-). This proves:

THEOREM 13.35. *Contact manifolds with* $\mathrm{AT}(M, \xi) < \infty$ *are not strongly fillable. Moreover, if there exists an exact symplectic cobordism from* (M_-, ξ_-) *to* (M_+, ξ_+)*, then* $\mathrm{AT}(M_-, \xi_-) \leqslant \mathrm{AT}(M_+, \xi_+)$*.* □

It is known (see [**Wen13b**]) that the second part of the above theorem does not hold for strong symplectic cobordisms in general, so exactness of cobordisms is a meaningful symplectic topological condition, not just a technical hypothesis. It is also known (see [**Ghi05**]) that strong and exact fillability are not equivalent conditions.

There are many known examples of contact manifolds that have untwisted algebraic torsion but are weakly fillable. The simplest are the tight tori $(\mathbb{T}^3, \xi_k)$ for $k \geqslant 2$, for which weak fillings were first constructed by Giroux [**Gir94**], but Eliashberg [**Eli96**] showed that strong fillings do not exist, and we will see in Lecture 16 that $\mathrm{AT}(\mathbb{T}^3, \xi_k) = 1$. The weak/strong distinction can often be detected via the choice of coefficients in SFT. We saw in §13.2.1 that a weak filling of a contact manifold (M, ξ) can always be deformed so as to have stable boundary with data $(\mathcal{H} = (\Omega, \alpha), J_+)$ for which α is a nondegenerate contact form and J_+ is C^∞-close to any given element of $\mathcal{J}(\alpha)$. Proposition 13.16 showed that if (M, ξ) is algebraically overtwisted, then the contact homology for the stable Hamiltonian data $(\mathcal{H}, J_+)$ can also be made to vanish.

EXERCISE 13.36. Generalize the proof of Prop. 13.16 to show that if (M, ξ) has algebraic k-torsion with coefficients in R, then also $[\hbar^k] = 0 \in H_*^{\mathrm{SFT}}(M, \mathcal{H}_c, J_c; R)$ for sufficiently large $c > 0$.

It then follows using Theorem 13.33 that algebraic torsion with suitably twisted coefficients also gives an obstruction to weak filling. Let us say that (M, ξ) has **fully twisted** algebraic k-torsion whenever $[\hbar^k] = 0 \in H_*^{\mathrm{SFT}}(M, \xi; \mathbb{Q}[H_2(M)])$. Note that in parallel with Remark 13.10, any nested pair of subgroups $G \subset G' \subset H_2(M)$ gives rise to a map

$$H_*^{\mathrm{SFT}}(M, \xi; \mathbb{Q}[H_2(M)/G']) \to H_*^{\mathrm{SFT}}(M, \xi; \mathbb{Q}[H_2(M)/G]),$$

which is a morphism in the sense that it maps the unit and all powers of $\hbar$ to themselves. This implies that (M, ξ) has fully twisted k-torsion if and only if it has k-torsion for every choice of coefficients.

THEOREM 13.37. *If (M, ξ) is a closed contact manifold with a finite order of algebraic torsion with coefficients in $R = \mathbb{Q}[H_2(M)/G]$ for some subgroup G, then (M, ξ) does not admit any weak symplectic filling (W, ω) for which $[\omega|_{TM}] \in H^2_{\mathrm{dR}}(M)$ is rational and annihilates all elements of G. In particular, if (M, ξ) has fully twisted algebraic torsion of some finite order, then it is not weakly fillable.* □

REMARK 13.38. The rationality condition in Theorem 13.37 can probably be lifted, and is known to be unnecessary at least in dimension three. It is clear in any case that if (M, ξ) admits a weak filling (W, ω), then one can always make a small perturbation of ω to produce a weak filling for which $[\omega|_{TM}] \in H^2(M; \mathbb{Q})$.

We will see some concrete examples of algebraic torsion computations in Lecture 16. Let us conclude this discussion for now with the observation that algebraic torsion of order *zero* is a notion we've seen before:

PROPOSITION 13.39. *For any closed contact manifold (M, ξ) and group ring $R = \mathbb{Q}[H_2(M)/G]$, the following conditions are equivalent:*

(1) (M, ξ) has algebraic 0-torsion (with coefficients in R);
(2) (M, ξ) is algebraically overtwisted (with coefficients in R);
(3) $H_*^{\mathrm{SFT}}(M, \xi; R) = 0$.

PROOF. It is obvious that (3) implies (1). Since $\mathbf{D}_{\mathrm{SFT}}\mathbf{f} = \partial_{\mathrm{CH}}\mathbf{f} + \mathcal{O}(\hbar)$ for $\mathbf{f} \in \mathcal{A}$, the linear map

$$\mathcal{A}[[\hbar]] \to \mathcal{A} : \mathbf{F} \mapsto \mathbf{F}|_{\hbar=0}$$

defines a chain map $(\mathcal{A}[[\hbar]], \mathbf{D}_{\mathrm{SFT}}) \to (\mathcal{A}, \partial_{\mathrm{CH}})$ and thus descends to a linear map $H_*^{\mathrm{SFT}}(M, \xi; R) \to HC_*(M, \xi; R)$ which sends $1 \in \mathcal{A}[[\hbar]]$ to $1 \in \mathcal{A}$. The existence of this map proves that (1) implies (2).

To prove that (2) implies (3), recall first that if there exists $\mathbf{f} \in \mathcal{A}$ with $\partial_{\mathrm{CH}}\mathbf{f} = 1$, then the fact that $HC_*(M, \xi; R) = 0$ follows easily since for any $\mathbf{g} \in \mathcal{A}$ with $\partial_{\mathrm{CH}}\mathbf{g} = 0$, the graded Leibniz rule implies $\partial_{\mathrm{CH}}(\mathbf{fg}) = (\partial_{\mathrm{CH}}\mathbf{f})\mathbf{g} - \mathbf{f}(\partial_{\mathrm{CH}}\mathbf{g}) = \mathbf{g}$. This works because ∂_{CH} is a derivation—but $\mathbf{D}_{\mathrm{SFT}}$ is not one, so the same trick will not quite work for $\mathbf{D}_{\mathrm{SFT}}$. The trick in proving $H_*^{\mathrm{SFT}}(M, \xi; R) = 0$ will be to quantify the failure of $\mathbf{D}_{\mathrm{SFT}}$ to be a derivation. For our purposes, it suffices to know that

$$\mathbf{D}_{\mathrm{SFT}}(\mathbf{FG}) = (\mathbf{D}_{\mathrm{SFT}}\mathbf{F})\mathbf{G} + (-1)^{|\mathbf{F}|}\mathbf{F}(\mathbf{D}_{\mathrm{SFT}}\mathbf{G}) + \mathcal{O}(\hbar) \tag{13.18}$$

holds for all $\mathbf{F}, \mathbf{G} \in \mathcal{A}[[\hbar]]$, which follows from the fact that ∂_{CH} is a derivation.

With this remark out of the way, suppose $\mathbf{f} \in \mathcal{A}$ satisfies $\partial_{\mathrm{CH}}\mathbf{f} = 1$, in which case

$$\mathbf{D}_{\mathrm{SFT}}\mathbf{f} = 1 + \hbar\mathbf{G} \tag{13.19}$$

for some $\mathbf{G} \in \mathcal{A}[[\hbar]]$. We claim then that for any $\mathbf{Q} \in \mathcal{A}[[\hbar]]$ with $\mathbf{D}_{\mathrm{SFT}}\mathbf{Q} = 0$, there exists $\mathbf{Q}_1 \in \mathcal{A}[[\hbar]]$ with

$$\mathbf{D}_{\mathrm{SFT}}(\mathbf{f}\mathbf{Q}) = \mathbf{Q} + \hbar\mathbf{Q}_1 \tag{13.20}$$

and $\mathbf{D}_{\mathrm{SFT}}\mathbf{Q}_1 = 0$. Indeed, (13.20) follows from (13.18) and (13.19) since $\mathbf{D}_{\mathrm{SFT}}\mathbf{Q} = 0$, and $\mathbf{D}_{\mathrm{SFT}}\mathbf{Q}_1 = 0$ then follows by applying $\mathbf{D}_{\mathrm{SFT}}$ to (13.20) and using $\mathbf{D}_{\mathrm{SFT}}^2 = 0$. Fixing $\mathbf{Q}_0 := \mathbf{Q} \in \mathcal{A}[[\hbar]]$, we can now define a sequence $\mathbf{Q}_k \in \mathcal{A}[[\hbar]]$ satisfying $\mathbf{D}_{\mathrm{SFT}}\mathbf{Q}_k = 0$ for all integers $k \geqslant 0$ via the inductive condition

$$\mathbf{D}_{\mathrm{SFT}}(\mathbf{f}\mathbf{Q}_k) = \mathbf{Q}_k + \hbar\mathbf{Q}_{k+1}.$$

Then $\sum_{k=0}^{\infty}(-1)^k\hbar^k\mathbf{Q}_k \in \mathcal{A}[[\hbar]]$, and

$$\mathbf{D}_{\mathrm{SFT}}\left(\mathbf{f}\sum_{k=0}^{\infty}(-1)^k\hbar^k\mathbf{Q}_k\right) = \mathbf{Q}.$$

□

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 2.6, 2.7 and 2.10 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 $f_1, \dots, f_m$,

$$\text{(A.1)} \qquad \left\| \prod_{i=1}^m |f_i| \right\|_{L^p} \leqslant \prod_{i=1}^m \|f_i\|_{L^{p_i}} \quad \text{for } 1 \leqslant p \leqslant p_1, \dots, p_m \leqslant \infty \text{ with } \frac{1}{p} = \sum_{i=1}^m \frac{1}{p_i}.$$

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.2 below). Let

$$W_{\mathrm{loc}}^{k,p}(\mathcal{U}) := \big\{ \text{functions } f \text{ on } \mathcal{U} \;\big|\; f \in W^{k,p}(\mathcal{V}) \text{ for all open subsets } \mathcal{V} \subset \mathcal{U} \text{ with compact closure } \overline{\mathcal{V}} \subset \mathcal{U} \big\},$$

and say that a sequence $f_j \in W_{\mathrm{loc}}^{k,p}(\mathcal{U})$ converges in $W_{\mathrm{loc}}^{k,p}$ to $f \in W_{\mathrm{loc}}^{k,p}(\mathcal{U})$ 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 ([**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.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.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.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.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.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.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.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 $W_0^{k,p}(\mathcal{U})$ 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.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.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.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.9. Derive Corollary A.8 from Theorem A.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.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

$$W^{k+d,p}(\mathcal{U}) \hookrightarrow C^{d,\alpha}(\mathcal{U}) \quad \text{for } \alpha \in (0, k-n/p),$$
$$W^{k+d,p}(\mathcal{U}) \hookrightarrow W^{d,q}(\mathcal{U}) \quad \text{for } q \in [p, \infty)$$

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

$$W^{k+d,p}(\mathcal{U}) \hookrightarrow W^{d,q}(\mathcal{U}) \quad \text{for } q \in [p, p^*)$$

are compact.

In particular, the continuous inclusion $W^{k,p}(\mathcal{U}) \hookrightarrow W^{m,q}(\mathcal{U})$ in Corollary A.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.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.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 2.6, 2.7 and 2.10 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 $W_0^{k,p}$.

We begin by generalizing Prop. 2.6, 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.7, and at least as many derivatives as functions of class $W^{m,q}$.

THEOREM A.13. *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.6 and Hölder's inequality.

LEMMA A.14. *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*

$p_1, \dots, p_m \geqslant 1$ and integers $k_1, \dots, k_m \geqslant 0$. Let $I := \{ i \in \{1, \dots, m\} \mid k_i p_i \leqslant n \}$. 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

$$W^{k_1,p_1}(\mathcal{U}) \times \dots \times W^{k_m,p_m}(\mathcal{U}) \to L^q(\mathcal{U}).$$

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 $q_1, \dots, q_m \in [q, \infty]$ satisfying $1/q = 1/q_1 + \dots + 1/q_m$ for which Theorem A.6 provides continuous inclusions

$$W^{k_i,p_i}(\mathcal{U}) \hookrightarrow L^{q_i}(\mathcal{U})$$

for each $i = 1, \dots, m$. 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. $\square$

PROOF OF THEOREM A.13. 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.14, 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.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 $a + b \leqslant m$. $\square$

REMARK A.15. 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.13, 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.13 provides a means of resolving this question whenever f and g belong to suitable Sobolev spaces.

PROPOSITION A.16. *Suppose k, m, p, q and $\mathcal{U} \subset \mathbb{R}^n$ satisfy the same conditions as in Theorem A.13, 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.17. A slight simplification of the same argument as in Proposition A.16 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 2.7 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_0 and

$$(x, v) \in \mathcal{V} \quad \Rightarrow \quad (x, tv + (1-t) f_0(x)) \in \mathcal{V} \text{ for all } t \in [0, 1].$$

THEOREM A.18. *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 $\mathcal{O}^k(\overline{\mathcal{V}}, \mathbb{R}^N) \subset C^k(\overline{\mathcal{V}}, \mathbb{R}^N)$ is a subset such that all $\Psi \in \mathcal{O}^k(\overline{\mathcal{V}}, \mathbb{R}^n)$ 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

$$\mathcal{O}^k(\overline{\mathcal{V}}, \mathbb{R}^N) \times \mathcal{O}^{k,p}(\mathcal{U}; \mathcal{V}) \to W^{k,p}(\mathcal{U}, \mathbb{R}^N) : (\Psi, f) \mapsto \Psi \circ (\mathrm{Id} \times f),$$

[1]Both of the conditions on $\Psi \in \mathcal{O}^k(\overline{\mathcal{V}}, \mathbb{R}^n)$ are vacuous if $\mathcal{U} \subset \mathbb{R}^n$ is bounded.

and for each $\Psi \in \mathcal{O}^k(\overline{\mathcal{V}}, \mathbb{R}^N)$ *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 $\Psi \in \mathcal{O}^k(\overline{\mathcal{V}}, \mathbb{R}^N)$. 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

$$\|\Psi \circ (\mathrm{Id} \times f) - \Psi \circ (\mathrm{Id} \times f_0)\|_{L^p} \leqslant \|\Psi\|_{C^1(\mathcal{V})} \cdot \|f - f_0\|_{L^p},$$

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

$$(D_1^\gamma D_2^\ell \Psi \circ (\mathrm{Id} \times f))(\partial^{\beta_1} f, \ldots, \partial^{\beta_\ell} f) : \mathcal{U} \to \mathbb{R}^N, \tag{A.7}$$

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

$$\left\| (D_1^\gamma D_2^\ell \Psi \circ (\mathrm{Id} \times f))(\partial^{\beta_1} f, \ldots, \partial^{\beta_\ell} f) \right\|_{L^p(\mathcal{U})} \leqslant \|D_1^\gamma D_2^\ell \Psi\|_{C^0(\mathcal{V})} \cdot \left\| \prod_{j=1}^{\ell} |\partial^{\beta_j} f| \right\|_{L^p(\mathcal{U})}$$

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 $m_1, \ldots, m_\ell \geqslant 0$ are integers satisfying $m_1 + \ldots + m_\ell \leqslant k$. 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.14 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 $m_1 + \ldots + m_\ell \leqslant k$, we find

$$\begin{aligned}\sum_{j=1}^{\ell} \left(\frac{1}{p} - \frac{k - m_j}{n} \right) &= \ell \left(\frac{1}{p} - \frac{k}{n} \right) + \frac{m_1 + \ldots + m_\ell}{n} \\ &\leqslant \frac{1}{p} + (\ell - 1) \left(\frac{1}{p} - \frac{k}{n} \right) < \frac{1}{p}.\end{aligned}$$

This proves that $\Psi \circ (\mathrm{Id} \times f) \in W^{k,p}(\mathcal{U}, \mathbb{R}^N)$.

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 $\Psi_i \in \mathcal{O}^k(\overline{\mathcal{V}}, \mathbb{R}^N)$ converges to $\Psi \in \mathcal{O}^k(\overline{\mathcal{V}}, \mathbb{R}^N)$ 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.7) 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.19 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 $(\Psi, f) \mapsto \Psi \circ (\mathrm{Id} \times f)$. $\square$

EXERCISE A.19. 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.18, 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.20 ([**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 2.10. 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.21. *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.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.6 gives a continuous inclusion

$$W^{k-m,p}(\mathring{\mathbb{D}}^n) \hookrightarrow L^q(\mathring{\mathbb{D}}^n) \tag{A.8}$$

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.8) 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

$$|h| < \mathrm{dist}(\mathcal{V}, \mathbb{R}^n\backslash\mathcal{U}) := \inf\left\{|x - y| \;\middle|\; x \in \mathcal{V} \text{ and } y \in \mathbb{R}^n\backslash\mathcal{U}\right\}.$$

The following result about difference quotients is useful for proving local regularity of solutions to PDEs, as in §2.4.

THEOREM A.22. *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.22.

LEMMA A.23. *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.24. *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 $f \in W^{1,p}(\mathcal{U})$,
$$\|D_i^h f\|_{L^p(\mathcal{V})} \leqslant \|\partial_i f\|_{L^p(\mathcal{U})}, \qquad i = 1, \ldots, n \tag{A.9}$$
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
$$\begin{aligned}\left|D_i^h f(x)\right| = \left|\frac{f(x + he_i) - f(x)}{h}\right| &= \left|\frac{1}{h} \int_0^1 \frac{d}{dt} f(x + the_i)\, dt\right| \\ &= \left|\int_0^1 \partial_i f(x + the_i)\, dt\right| \leqslant \int_0^1 |\partial_i f(x + the_i)|\, dt.\end{aligned}$$
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

$$\begin{aligned}\|D_i^h f\|^p_{L^p(\mathcal{V})} &= \int_{\mathcal{V}} \left|D_i^h f(x)\right|^p \, d\mu(x) \leqslant \int_{\mathcal{V}} \left(\int_0^1 |\partial_i f(x + the_i)| \, dt \right)^p d\mu(x) \\ &\leqslant \int_{\mathcal{V}} \int_0^1 |\partial_i f(x + the_i)|^p \, dt \, d\mu(x) = \int_0^1 \int_{\mathcal{V}} |\partial_i f(x + the_i)|^p \, d\mu(x) \, dt \\ &\leqslant \int_0^1 \|\partial_i f\|^p_{L^p(\mathcal{U})} \, dt = \|\partial_i f\|^p_{L^p(\mathcal{U})}.\end{aligned}$$

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.23, $D_i^h f_\epsilon \to \partial_i f_\epsilon$ 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.9),

$$\|D_i^h f_\epsilon - D_i^h f\|_{L^p(\mathcal{V})} \leqslant \|\partial_i f_\epsilon - \partial_i f\|_{L^p(\mathcal{U})} \leqslant \|f_\epsilon - f\|_{W^{1,p}(\mathcal{U})} < \epsilon/3,$$

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.25 (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.26. 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.25 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.27. *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.23, $D_i^{-h_j}\varphi \to \partial_i \varphi$ uniformly on $\mathcal{V}$ and thus also in $L^q(\mathcal{V})$, so taking the limit of the integrals, we've shown

$$\int_{\mathcal{V}} g\varphi = -\int_{\mathcal{V}} f \partial_i \varphi \quad \text{for all } \varphi \in C_0^\infty(\mathcal{V}),$$

or in other words, $\partial_i f = g \in L^p(\mathcal{V})$. □

PROOF OF THEOREM A.22. The two statements in the theorem follow by applying Propositions A.24 and A.27 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 $\mathbb{R}^n_+ := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_n \geqslant 0\}$;
(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

$$g_{\beta\alpha} : \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to \mathrm{GL}(m, \mathbb{F}) \quad \text{such that} \quad \begin{array}{l} \Phi_\beta \circ \Phi_\alpha^{-1}(x, v) = (x, g_{\beta\alpha}(x)v) \\ \text{for } x \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta,\ v \in \mathbb{F}^m \end{array}$$

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

$$W^{k,p}_{\mathrm{loc}}(E) := \big\{ \eta : M \to E \ \big|\ \text{sections such that } \eta^\alpha \in W^{k,p}_{\mathrm{loc}}(\mathring{\Omega}_\alpha, \mathbb{F}^m) \\ \text{for all } \alpha \in \mathcal{A}(\pi) \big\},$$

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.28. 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.29. *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.28 satisfies $\|\eta_j - \eta\|_{W^{k,p}(E)} \to 0$.*

The proposition is an immediate consequence of the following.

LEMMA A.30. *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.20 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

$$\begin{aligned}
\|\eta^\beta\|_{W^{k,p}(\mathring{\Omega}_\beta)} &= \left\| \sum_{\alpha\in J} (\rho_\alpha\eta)^\beta \right\|_{W^{k,p}(\mathring{\Omega}_\beta)} \leqslant \sum_{\alpha\in J} \|(\rho_\alpha\eta)^\beta\|_{W^{k,p}(\mathring{\Omega}_{\beta\alpha})} \\
&= \sum_{\alpha\in J} \|(\rho_\alpha \circ \varphi_\beta^{-1})(g_{\beta\alpha} \circ \varphi_\beta^{-1})(\eta^\alpha \circ \varphi_{\alpha\beta})\|_{W^{k,p}(\mathring{\Omega}_{\beta\alpha})} \\
&\leqslant c \sum_{\alpha\in J} \|\eta^\alpha\|_{W^{k,p}(\mathring{\Omega}_{\alpha\beta})} \leqslant c \sum_{\alpha\in J} \|\eta^\alpha\|_{W^{k,p}(\mathring{\Omega}_\alpha)}.
\end{aligned}$$

□

COROLLARY A.31. *If M is compact, then the norm on $W^{k,p}(E)$ given by Definition A.28 is independent of all auxiliary choices up to equivalence of norms.* □

THEOREM A.32. *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.30 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.33. 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.34. For $\mathcal{U} \subset \mathbb{R}^n$ an open subset, the space $W^{k,p}_{\text{loc}}(\mathcal{U})$ 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 $W^{k,p}_{\text{loc}}(\mathcal{U})$ are equivalent.

EXERCISE A.35. 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.35, 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 $W_0^{k,p}(E)$ is a closed subspace of $W^{k,p}(E)$.

The partition of unity argument in Theorem A.32 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.36. *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*
$$W^{k,p}(\mathcal{V}) := \left\{\eta \in W^{k,p}(E) \;\middle|\; \eta(M) \subset \mathcal{V}\right\}$$
and
$$C^k_M(\mathcal{V}, F) := \left\{\Phi : \mathcal{V} \to F \mid \text{fiber-preserving maps of class } C^k\right\},$$
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.37. 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.13 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 $M = \dot{\Sigma} = \Sigma \setminus \Gamma$ 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 Lecture 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 $W^{k,p}_{\mathrm{loc}}(E)$ 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.28 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.36 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.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, $W^{k,p}(E) = W^{k,p}_0(E)$.

THEOREM A.38. *Given an asymptotically Hermitian bundle E over a punctured Riemann surface $\dot{\Sigma}$, the space $C^\infty_0(E)$ 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.28 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 provides for each $\alpha \in I$ a sequence $\eta^\alpha_j \in W^{k,p}(\Omega_\alpha)$ of smooth functions with bounded support such that $\eta^\alpha_j \to \eta^\alpha$ 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

$$\eta_j := \sum_{\alpha \in I} \rho_\alpha (\eta^\alpha_j \circ \varphi_\alpha) \in W^{k,p}(E).$$

These sections are smooth and have compact support since the η^α_j have bounded support in Ω_α, and they converge in $W^{k,p}$ to η. □

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

$$C_{\boldsymbol{\epsilon}}(E) = \left\{ \eta \in \Gamma(E) \;\middle|\; \|\eta\|_{C_\epsilon} < \infty \right\},$$

where the C_ϵ-norm is defined by

$$\|\eta\|_{C_\epsilon} = \sum_{k=0}^{\infty} \epsilon_k \|\eta\|_{C^k}. \tag{B.1}$$

The norm for $C_\epsilon(E)$ 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 $C_\epsilon(E)$ 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

$$C_\epsilon(E;\mathcal{U}) = \left\{ \eta \in C_\epsilon(E) \;\middle|\; \eta|_{M\setminus\mathcal{U}} \equiv 0 \right\}.$$

Remark B.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 $C_\epsilon(E;\mathcal{U})$ can be defined as a closed subspace of $C_\epsilon(E|_{M_0})$ 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.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 $j \geqslant N$. $\square$

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.3. Adapt the proof of Theorem B.2 to show that X_ϵ is also a Banach space.

LEMMA B.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.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.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.2 and B.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.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 $C_\epsilon(E;\mathcal{U})$, as

$$\|\eta_j\|_{C_\epsilon} < \sum_{k=0}^{j-1} \epsilon_k \|\eta_j\|_{C^k} + \sum_{k=j}^{\infty} \frac{1}{2^k} < \infty.$$

The second property is essentially local, so it can be deduced from Lemma B.7 below. $\square$

LEMMA B.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 3.35, which was needed for our definition of spectral flow in §3.3. The proof combines some ideas from that section with the technique used in Lecture 8 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 3.35, you might want to skip this for now and come back after you've read as far as Lecture 8.

Recalling the notation from Lecture 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 $S \in L^\infty(S^1, \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n}))$, 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 $\mathbf{T}_{\mathrm{ref}}$,

$$\mathrm{Fred}_{\mathbb{R}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) = \left\{ \mathbf{T}_{\mathrm{ref}} + \mathbf{K} : \mathcal{D} \to \mathcal{H} \;\middle|\; \mathbf{K} \in \mathscr{L}_{\mathbb{R}}^{\mathrm{sym}}(\mathcal{H}) \right\},$$

which we regard as a smooth Banach manifold via its obvious identification with the space $\mathscr{L}_{\mathbb{R}}^{\mathrm{sym}}(\mathcal{H})$ of symmetric bounded linear operators on $\mathcal{H}$. For $k \in \mathbb{N}$, we denote by

$$\mathrm{Fred}_{\mathbb{R}}^{\mathrm{sym},k}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}}) \subset \mathrm{Fred}_{\mathbb{R}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$$

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 3.35 again.

Lemma. *Fix a smooth path $[-1, 1] \to L^\infty(\mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})) : s \mapsto S_s$ 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 $B : [-1, 1] \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ 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

$$\mathcal{A}_{\boldsymbol{\epsilon}} := \left\{ B \in C^\infty([-1,1]\times S^1, \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n})) \;\middle|\; \|B\|_{C_{\boldsymbol{\epsilon}}} < \infty \text{ and } B(\pm 1,\cdot) \equiv 0 \right\},$$

and assume via Theorem B.6 that $\mathcal{A}_{\boldsymbol{\epsilon}}$ is dense in the Banach space of continuous functions $[-1,1]\times S^1 \to \mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n})$ 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 3.20 and 3.21 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 $L^\infty(\mathrm{End}^{\mathrm{sym}}_{\mathbb{R}}(\mathbb{R}^{2n}))$. 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 §3.3 that there exist decompositions

$$\mathcal{D} = V \oplus K, \qquad \mathcal{H} = W \oplus K$$

where $K = \ker\left(\mathbf{A}^B_{s_0} - \lambda_0\right)$, $W = \operatorname{im}\left(\mathbf{A}^B_{s_0} - \lambda_0\right)$ 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 Lecture 8: 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.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 $B : S^1 \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ 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 3.16. 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 $\widehat{B}(t) \in \mathrm{End}(\mathbb{R}^{2n})$ 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

$$\Psi(s, \lambda, B) := \Psi_B(s, \lambda) \in \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(K),$$

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 $K = \ker\left(\mathbf{A}_{s_0}^{B_0} - \lambda_0\right)$. The partial derivative of Ψ with respect to the third variable at (s_0, λ_0, B_0) is then a linear map

$$\mathbf{L} := D_3\Psi(s_0, \lambda_0, B_0) : \mathcal{A}_\epsilon \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(K)$$

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 $\Upsilon \in \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(K)$, Lemma C.1 provides a continuous loop $C_0 : S^1 \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ 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 $C : [-1,1] \times S^1 \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\mathbb{R}^{2n})$ 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 $C_{\boldsymbol{\epsilon}}$, but since it can be approximated arbitrarily well in the C^0-norm by functions in $\mathcal{A}_{\boldsymbol{\epsilon}}$, we conclude that the image of $\mathbf{L}$ is dense in $\mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(K)$. 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 $(-1,1) \times \mathbb{R} \times \mathcal{A}_{\boldsymbol{\epsilon}}$, and it is separable since the latter is also separable. Consider the projection

$$\pi : \mathcal{V}^k \to \mathcal{A}_{\boldsymbol{\epsilon}} : (s, \lambda, B) \mapsto B, \tag{C.2}$$

which is a smooth map of separable Banach manifolds whose fibers $\pi^{-1}(B)$ are the spaces $\mathcal{V}^k(B)$. Using Lemma 8.2, 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

$$\mathcal{A}_{\boldsymbol{\epsilon}}^{\mathrm{reg},k} \subset \mathcal{A}_{\boldsymbol{\epsilon}}.$$

The intersection

$$\mathcal{A}_{\boldsymbol{\epsilon}}^{\mathrm{reg}} := \bigcap_{k \in \mathbb{N}} \mathcal{A}_{\boldsymbol{\epsilon}}^{\mathrm{reg},k}$$

is then another comeager subset of $\mathcal{A}_{\boldsymbol{\epsilon}}$, with the property that for each $B \in \mathcal{A}_{\boldsymbol{\epsilon}}^{\mathrm{reg}}$ and every $k \in \mathbb{N}$ and $(s, \lambda) \in \mathcal{V}^k(B)$, $d\Psi_B(s, \lambda)$ is (by Lemma 8.2) 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 $B \in \mathcal{A}_{\boldsymbol{\epsilon}}$ 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

$$\mathcal{V}^0 = \left\{ (s, B) \in (-1,1) \times \mathcal{A}_{\boldsymbol{\epsilon}} \;\middle|\; s \in \mathcal{V}^0(B) \right\}.$$

A neighborhood of any (s_0, B_0) in $\mathcal{V}^0$ is then the zero set of a smooth map of the form

$$\Psi(s, B) = \Phi(\mathbf{A}_s^B) \in \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\ker \mathbf{A}_{s_0}^{B_0}),$$

defined for all $(s, B) \in (-1,1) \times \mathcal{A}_{\boldsymbol{\epsilon}}$ close enough to (s_0, B_0). For a fixed $B \in \mathcal{A}_{\boldsymbol{\epsilon}}$ 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 $s \mapsto \mathbf{A}_s^B \in \mathrm{Fred}_{\mathbb{R}}^{\mathrm{sym}}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ with $\mathrm{Fred}_{\mathbb{R}}^{\mathrm{sym},1}(\mathcal{D}, \mathcal{H}, \mathbf{T}_{\mathrm{ref}})$ at $s = s_1$ is transverse if and only if

$$d\Psi_B(s_1) : \mathbb{R} \to \mathrm{End}_{\mathbb{R}}^{\mathrm{sym}}(\ker \mathbf{A}_{s_0}^{B_0})$$

is surjective. At (s_0, B_0), the partial derivative of Ψ with respect to B is again the same operator

$$\mathbf{L} = D_2\Psi(s_0, B_0) : \mathcal{A}_{\boldsymbol{\epsilon}} \to \operatorname{End}_{\mathbb{R}}^{\mathrm{sym}}(\ker \mathbf{A}_{s_0}^{B_0})$$

as in (C.1), which we've already seen is surjective due to Lemma C.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 3.35.

Bibliography

[Abb11] C. Abbas, *Holomorphic open book decompositions*, Duke Math. J. **158** (2011), no. 1, 29–82.

[Abb14] ———, *An introduction to compactness results in symplectic field theory*, Springer, Heidelberg, 2014. MR3157146

[ACH05] C. Abbas, K. Cieliebak, and H. Hofer, *The Weinstein conjecture for planar contact structures in dimension three*, Comment. Math. Helv. **80** (2005), no. 4, 771–793.

[AS] A. Abbondandolo and F. Schlenk, *Floer homologies, with applications*. Preprint arXiv:1709.00297.

[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.

[ALR07] A. Adem, J. Leida, and Y. Ruan, *Orbifolds and stringy topology*, Cambridge Tracts in Mathematics, vol. 171, Cambridge University Press, Cambridge, 2007.

[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.

[Ale23] J. W. Alexander, *A lemma on systems of knotted curves*, Nat. Acad. Proc. **9** (1923), 93–95.

[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.

[Avd] R. Avdek, *Liouville hypersurfaces and connect sum cobordisms*. Preprint arXiv:1204.3145.

[Bao15] E. Bao, *On J-holomorphic curves in almost complex manifolds with asymptotically cylindrical ends*, Pacific J. Math. **278** (2015), no. 2, 291–324.

[BEM15] M. S. Borman, Y. Eliashberg, and E. Murphy, *Existence and classification of overtwisted contact structures in all dimensions*, Acta Math. **215** (2015), no. 2, 281–361.

[Bou02] F. Bourgeois, *A Morse-Bott approach to contact homology*, Ph.D. Thesis, Stanford University, 2002.

[Bou03] ———, *Introduction to Contact Homology* (2003). Lecture notes from the summer school "Courbes holomorphes et Topologie de contact" in Berder, available at `http://www.math.u-psud.fr/~bourgeois/papers/berder.html`.

[Bou06] ———, *Contact homology and homotopy groups of the space of contact structures*, Math. Res. Lett. **13** (2006), no. 1, 71–85.

[Bou09] ———, *A survey of contact homology*, New perspectives and challenges in symplectic field theory, CRM Proc. Lecture Notes, vol. 49, Amer. Math. Soc., Providence, RI, 2009, pp. 45–71.

[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.

[BM04] F. Bourgeois and K. Mohnke, *Coherent orientations in symplectic field theory*, Math. Z. **248** (2004), no. 1, 123–146.

[BN10] F. Bourgeois and K. Niederkrüger, *Towards a good definition of algebraically overtwisted*, Expo. Math. **28** (2010), no. 1, 85–100.

[BvK10] F. Bourgeois and O. van Koert, *Contact homology of left-handed stabilizations and plumbing of open books*, Commun. Contemp. Math. **12** (2010), no. 2, 223–263.

[Bra15] B. Bramham, *Periodic approximations of irrational pseudo-rotations using pseudoholomorphic curves*, Ann. of Math. (2) **181** (2015), no. 3, 1033–1086.

[CMP19] R. Casals, E. Murphy, and F. Presas, *Geometric criteria for overtwistedness*, J. Amer. Math. Soc. **32** (2019), no. 2, 563–604.

[CDvK16] R. Chiang, F. Ding, and O. van Koert, *Non-fillable invariant contact structures on principal circle bundles and left-handed twists*, Internat. J. Math. **27** (2016), no. 3, 1650024, 55.

[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.

[CL09] K. Cieliebak and J. Latschev, *The role of string topology in symplectic field theory*, New perspectives and challenges in symplectic field theory, CRM Proc. Lecture Notes, vol. 49, Amer. Math. Soc., Providence, RI, 2009, pp. 113–146.

[CM05] K. Cieliebak and K. Mohnke, *Compactness for punctured holomorphic curves*, J. Symplectic Geom. **3** (2005), no. 4, 589–654. Conference on Symplectic Topology.

[CM07] K. Cieliebak and K. Mohnke, *Symplectic hypersurfaces and transversality in Gromov-Witten theory*, J. Symplectic Geom. **5** (2007), no. 3, 281–356.

[CM18] K. Cieliebak and K. Mohnke, *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.

[CE] J. Conway and J. B. Etnyre, *Contact surgery and symplectic caps*. Preprint arXiv:1811.00387, to appear in Bull. London Math. Soc.

[Dav] M. W. Davis, *Lectures on orbifolds and reflection groups*. Available at `https://people.math.osu.edu/davis.12/papers/lectures%20on%20orbifolds.pdf`.

[DFa] A. Doicu and U. Fuchs, *Asymptotic behaviour for $\mathcal{H}$-holomorphic cylinders of small area*. Preprint arXiv:1802.05573.

[DFb] ———, *A compactness result for $\mathcal{H}$-holomorphic curves in symplectizations*. Preprint arXiv:1802.08991.

[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.

[Dra04] D. L. Dragnev, *Fredholm theory and transversality for noncompact pseudoholomorphic maps in symplectizations*, Comm. Pure Appl. Math. **57** (2004), no. 6, 726–763.

[Eli89] Y. Eliashberg, *Classification of overtwisted contact structures on* 3*-manifolds*, Invent. Math. **98** (1989), no. 3, 623–637.

[Eli91] Y. Eliashberg, *On symplectic manifolds with some contact properties*, J. Differential Geom. **33** (1991), no. 1, 233–238.

[Eli96] Y. Eliashberg, *Unique holomorphically fillable contact structure on the* 3*-torus*, Internat. Math. Res. Notices **2** (1996), 77–82.

[Eli98] Y. Eliashberg, *Invariants in contact topology*, Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998), 1998, pp. 327–338.

[Eli07] ———, *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.

[EKP06] Y. Eliashberg, S. S. Kim, and L. Polterovich, *Geometry of contact transformations and domains: orderability versus squeezing*, Geom. Topol. **10** (2006), 1635–1747.

[ET98] Y. M. Eliashberg and W. P. Thurston, *Confoliations*, University Lecture Series, vol. 13, American Mathematical Society, Providence, RI, 1998.

[Elĭ67] H. I. Elĭasson, *Geometry of manifolds of maps*, J. Differential Geometry **1** (1967), 169–194.

[Etn06] J. B. Etnyre, *Lectures on open book decompositions and contact structures*, Floer homology, gauge theory, and low-dimensional topology, Clay Math. Proc., vol. 5, Amer. Math. Soc., Providence, RI, 2006, pp. 103–141.

[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.

[EL15] J. D. Evans and Y. Lekili, *Floer cohomology of the Chiang Lagrangian*, Selecta Math. (N.S.) **21** (2015), no. 4, 1361–1404.

[FFGW16] O. Fabert, J. W. Fish, R. Golovko, and K. Wehrheim, *Polyfolds: a first and second look*, EMS Surv. Math. Sci. **3** (2016), no. 2, 131–208.

[FH] J. W. Fish and H. Hofer, *Lectures on Polyfolds and Symplectic Field Theory*. Preprint arXiv:1808.07147.

[FS] J. W. Fish and R. Siefring, *Connected sums and finite energy foliations I: Contact connected sums*. Preprint arXiv:1311.4221.

[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.

[FO99] K. Fukaya and K. Ono, *Arnold conjecture and Gromov-Witten invariant*, Topology **38** (1999), no. 5, 933–1048.

[Gei08] H. Geiges, *An introduction to contact topology*, Cambridge Studies in Advanced Mathematics, vol. 109, Cambridge University Press, Cambridge, 2008.

[GW] C. Gerig and C. Wendl, *Generic transversality for unbranched covers of closed pseudoholomorphic curves*. To appear in Comm. Pure Appl. Math., Preprint arXiv:1407.0678.

[Ger13] A. Gerstenberger, *Universal moduli spaces in Gromov-Witten theory*, Ph.D. Thesis, Ludwig-Maximilians-Universität Munich, 2013.

[Ghi05] P. Ghiggini, *Strongly fillable contact 3-manifolds without Stein fillings*, Geom. Topol. **9** (2005), 1677–1687.

[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).

[Gir99] ———, *Une infinité de structures de contact tendues sur une infinité de variétés*, Invent. Math. **135** (1999), no. 3, 789–802 (French).

[Gir00] ———, *Structures de contact en dimension trois et bifurcations des feuilletages de surfaces*, Invent. Math. **141** (2000), no. 3, 615–689 (French).

[Gir02] ———, *Géométrie de contact: de la dimension trois vers les dimensions supérieures*, Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002), 2002, pp. 405–414.

[Gro85] M. Gromov, *Pseudoholomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), no. 2, 307–347.

[Gro57] A. Grothendieck, *Sur la classification des fibrés holomorphes sur la sphère de Riemann*, Amer. J. Math. **79** (1957), 121–138 (French).

[Hin00] R. Hind, *Holomorphic filling of* $\mathbf{R}\mathrm{P}^3$, Commun. Contemp. Math. **2** (2000), no. 3, 349–363.

[Hin03] R. Hind, *Stein fillings of lens spaces*, Commun. Contemp. Math. **5** (2003), no. 6, 967–982.

[HL15] R. Hind and S. Lisi, *Symplectic embeddings of polydisks*, Selecta Math. (N.S.) **21** (2015), no. 3, 1099–1120.

[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.

[Hof00] ———, *Holomorphic curves and real three-dimensional dynamics*, Geom. Funct. Anal. **Special Volume** (2000), 674–704. GAFA 2000 (Tel Aviv, 1999).

[Hof06] H. Hofer, *A general Fredholm theory and applications*, Current developments in mathematics, 2004, Int. Press, Somerville, MA, 2006, pp. 1–71.

[HLS97] H. Hofer, V. Lizan, and J.-C. Sikorav, *On genericity for holomorphic curves in four-dimensional almost-complex manifolds*, J. Geom. Anal. **7** (1997), no. 1, 149–159.

[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.

[HLS15] U. Hryniewicz, J. E. Licata, and P. A. S. Salomão, *A dynamical characterization of universally tight lens spaces*, Proc. Lond. Math. Soc. (3) **110** (2015), no. 1, 213–269.

[HMS15] U. Hryniewicz, A. Momin, and P. A. S. Salomão, *A Poincaré-Birkhoff theorem for tight Reeb flows on* S^3, Invent. Math. **199** (2015), no. 2, 333–422.

[HWZ95a] 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.

[HWZ95b] ———, *A characterisation of the tight three-sphere*, Duke Math. J. **81** (1995), no. 1, 159–226.

[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.

[HWZ98] ———, *The dynamics on three-dimensional strictly convex energy surfaces*, Ann. of Math. (2) **148** (1998), no. 1, 197–289.

[HWZ99] ———, *Properties of pseudoholomorphic curves in symplectizations. III. Fredholm theory*, Topics in nonlinear analysis, 1999, pp. 381–475.

[HWZ02] ———, *Finite energy cylinders of small area*, Ergodic Theory Dynam. Systems **22** (2002), no. 5, 1451–1486.

[HWZ03] ———, *Finite energy foliations of tight three-spheres and Hamiltonian dynamics*, Ann. of Math. (2) **157** (2003), no. 1, 125–255.

[HWZ10] ———, *Integration theory on the zero sets of polyfold Fredholm sections*, Math. Ann. **346** (2010), no. 1, 139–198.

[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.

[Hut02] M. Hutchings, *An index inequality for embedded pseudoholomorphic curves in symplectizations*, J. Eur. Math. Soc. (JEMS) **4** (2002), no. 4, 313–361.

[Hut14] ———, *Lecture Notes on Embedded Contact Homology*, Contact and Symplectic Topology, Bolyai Society Mathematical Studies, vol. 26, Springer, 2014, pp. 389–484.

[Hut] ———, *Is cylindrical contact homology defined with integer coefficients?*. Blog post, available at https://floerhomology.wordpress.com/2014/12/15/.

[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.

[IMTP00] A. Ibort, D. Martínez-Torres, and F. Presas, *On the construction of contact submanifolds with prescribed topology*, J. Differential Geom. **56** (2000), no. 2, 235–283.

[Ish] S. Ishikawa, *Construction of general symplectic field theory*. Preprint arXiv:1807.09455.

[IS99] S. Ivashkovich and V. Shevchishin, *Structure of the moduli space in a neighborhood of a cusp-curve and meromorphic hulls*, Invent. Math. **136** (1999), no. 3, 571–602.

[Kan97] Y. Kanda, *The classification of tight contact structures on the 3-torus*, Comm. Anal. Geom. **5** (1997), no. 3, 413–438.

[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.

[Laz] O. Lazarev, *Maximal contact and symplectic structures*. Preprint arXiv:1810.11728, to appear in J. Topol.

[Laz00] L. Lazzarini, *Existence of a somewhere injective pseudo-holomorphic disc*, Geom. Funct. Anal. **10** (2000), no. 4, 829–862.

[Laz11] L. Lazzarini, *Relative frames on J-holomorphic curves*, J. Fixed Point Theory Appl. **9** (2011), no. 2, 213–256.

[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.

[MNW13] P. Massot, K. Niederkrüger, and C. Wendl, *Weak and strong fillability of higher dimensional contact manifolds*, Invent. Math. **192** (2013), no. 2, 287–373.

[McD06] D. McDuff, *Groupoids, branched manifolds and multisections*, J. Symplectic Geom. **4** (2006), no. 3, 259–315.

[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.

[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.

[Mor18] A. Moreno, Ph.D. Thesis, Humboldt-Universität zu Berlin, 2018.

[Mora] A. Moreno, *SFT computations and intersection theory in higher-dimensional contact manifolds*. Preprint arXiv:1903.11884.

[Morb] ______, *Algebraic and Giroux torsion in higher-dimensional contact manifolds*. Preprint arXiv:1903.11880.

[Nel13] J. Nelson, *Applications of automatic transversality in contact homology*, Ph.D. Thesis, University of Wisconsin at Madison, 2013.

[Nel15] ______, *Automatic transversality in contact homology I: regularity*, Abh. Math. Semin. Univ. Hambg. **85** (2015), no. 2, 125–179.

[Nel20] ______, *Automatic transversality in contact homology II: filtrations and computations*, Proc. Lond. Math. Soc. (3) **120** (2020), no. 6, 853–917.

[Nie06] K. Niederkrüger, *The plastikstufe—a generalization of the overtwisted disk to higher dimensions*, Algebr. Geom. Topol. **6** (2006), 2473–2508.

[NW11] K. Niederkrüger and C. Wendl, *Weak symplectic fillings and holomorphic curves*, Ann. Sci. École Norm. Sup. (4) **44** (2011), no. 5, 801–853.

[Pal68] R. S. Palais, *Foundations of global non-linear analysis*, W. A. Benjamin, Inc., New York-Amsterdam, 1968.

[Par19] J. Pardon, *Contact homology and virtual fundamental cycles*, J. Amer. Math. Soc. **32** (2019), no. 3, 825–919.

[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.

[RS95] J. Robbin and D. Salamon, *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.

[Sat56] I. Satake, *On a generalization of the notion of manifold*, Proc. Nat. Acad. Sci. U.S.A. **42** (1956), 359–363.

[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.

[Sie11] ———, *Intersection theory of punctured pseudoholomorphic curves*, Geom. Topol. **15** (2011), 2351–2457.

[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.

[Sma65] S. Smale, *An infinite dimensional version of Sard's theorem*, Amer. J. Math. **87** (1965), 861–866.

[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.

[Thu76] W. P. Thurston, *Some simple examples of symplectic manifolds*, Proc. Amer. Math. Soc. **55** (1976), no. 2, 467–468.

[TW75] W. P. Thurston and H. E. Winkelnkemper, *On the existence of contact forms*, Proc. Amer. Math. Soc. **52** (1975), 345–347.

[Tro92] A. J. Tromba, *Teichmüller theory in Riemannian geometry*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1992.

[Var04] V. S. Varadarajan, *Supersymmetry for mathematicians: an introduction*, Courant Lecture Notes in Mathematics, vol. 11, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2004.

[vBa] J. von Bergmann, *Compactness results for H-holomorphic maps*. Preprint arXiv:0904.1603.

[vBb] ———, *Embedded H-holomorphic maps and open book decompositions*. Preprint arXiv:0907.3939.

[Wei78] A. Weinstein, *Periodic orbits for convex Hamiltonian systems*, Ann. of Math. (2) **108** (1978), no. 3, 507–518.

[Wen08] C. Wendl, *Finite energy foliations on overtwisted contact manifolds*, Geom. Topol. **12** (2008), 531–616.

[Wen10a] ———, *Automatic transversality and orbifolds of punctured holomorphic curves in dimension four*, Comment. Math. Helv. **85** (2010), no. 2, 347–407.

[Wen10b] ———, *Strongly fillable contact manifolds and J-holomorphic foliations*, Duke Math. J. **151** (2010), no. 3, 337–384.

[Wen10c] ———, *Open book decompositions and stable Hamiltonian structures*, Expos. Math. **28** (2010), no. 2, 187–199.

[Wen13a] ———, *A hierarchy of local symplectic filling obstructions for contact* 3*-manifolds*, Duke Math. J. **162** (2013), no. 12, 2197–2283.

[Wen13b] ———, *Non-exact symplectic cobordisms between contact* 3*-manifolds*, J. Differential Geom. **95** (2013), no. 1, 121–182.

[Wena] ———, *Generic transversality in symplectizations.* Two-part blog post, available at https://symplecticfieldtheorist.wordpress.com/2014/11/27/.

[Wenb] ———, *Some bad news about the forgetful map in SFT.* Blog post, available at https://symplecticfieldtheorist.wordpress.com/2015/07/16/.

[Wenc] ———, *Signs (or how to annoy a symplectic topologist).* Blog post, available at https://symplecticfieldtheorist.wordpress.com/2015/08/23/.

[Wend] ———, *Lectures on holomorphic curves in symplectic and contact geometry.* Version 3.3, http://www.math.hu-berlin.de/~wendl/pub/jhol_bookv33.pdf.

[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.

[Wen] ———, *Transversality and super-rigidity for multiply covered holomorphic curves.* Preprint arXiv:1609.09867.

[Wit82] E. Witten, *Supersymmetry and Morse theory*, J. Differential Geom. **17** (1982), no. 4, 661–692 (1983).

[Yau06] M.-L. Yau, *Vanishing of the contact homology of overtwisted contact 3-manifolds*, Bull. Inst. Math. Acad. Sin. (N.S.) **1** (2006), no. 2, 211–229. With an appendix by Yakov Eliashberg.