

## ELLIPTIC METHODS IN SYMPLECTIC GEOMETRY

DUSA MCDUFF

The past few years have seen several exciting developments in the field of symplectic geometry, and a beginning has been made towards solving many important and hitherto inaccessible problems. The new techniques which have made this possible have come both from the calculus of variations and from the theory of elliptic partial differential operators. This paper describes some of the results that Gromov obtained using elliptic methods, and then shows how Floer applied these elliptic techniques to develop a new approach to Morse theory, which has important applications in the theory of 3- and 4-manifolds as well as in symplectic geometry. To give some idea of the context of their results, we begin with a section on symplectic geometry, which concentrates on questions about symplectic diffeomorphisms. For more general recent surveys of the field, see for example [A2], [E2], [G1], [G3], [H2], [V1], and [V2].

The contents of this paper are:

### CONTENTS

#### §1. Symplectic geometry

- (1.1) Basic notions
- (1.2) Symplectic images of balls
- (1.3) Fixed point theorems

#### §2 $J$ -holomorphic spheres

- (2.1) Basic definitions
- (2.2) Compatible metrics and conformality
- (2.3) The analytic setup

---

Received by the editors March 1, 1990.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 53C15, 58E05, 58F05.

*Key words and phrases.* Symplectic geometry, almost complex manifold, elliptic PDE, Hamiltonian systems, pseudoholomorphic curves.

Partially supported by NSF grant no. DMS-8803056.

©1990 American Mathematical Society  
0273-0979/90 \$1.00 + \$.25 per page

- (2.4) Compactness
- (2.5) Applications

### §3 Floer homology

- (3.1) The variational setup
- (3.2) Morse theory
- (3.3) Floer's proof of the Arnold conjectures
- (3.4) Some details of the analysis
- (3.5) Other applications of Floer's theory

The first three sections of §3 are independent of §2 apart from some elementary definitions. In particular, (3.2) gives a self-contained description of Floer's approach to Morse theory. We treat this here only in the context of symplectic geometry. However, because the Chern-Simons functional has analytic properties very similar to those of the symplectic action functional, there is another important application of these ideas to the study of 3- and 4-dimensional manifolds. This is beautifully described in Atiyah's survey article [At], and the details may be found in [F4]. See also [T] and [W2] which give a very general context for the theory described here.

I wish to thank Floer, Gromov, Hofer, Oh, Salamon, and Weinstein for many useful comments and discussions, and Floer for giving me some of his unpublished lecture notes which form the basis for my treatment of the analysis.

## §1. SYMPLECTIC GEOMETRY

(1.1) **Basic notions.** We will assume throughout that  $V$  is a smooth compact manifold without boundary and of dimension  $2n$ . A symplectic form (or symplectic structure) on  $V$  is a closed 2-form  $\omega$  on  $V$ , which is nondegenerate in the sense that its highest power  $\omega \wedge \dots \wedge \omega = \omega^n$  never vanishes. The simplest example is the standard form

$$\omega_0 = dx_1 \wedge dx_2 + dx_3 \wedge dx_4 + \dots + dx_{2n-1} \wedge dx_{2n}$$

on euclidean space  $\mathbf{R}^{2n}$ . The first theorem in symplectic geometry is:

**Darboux's theorem.** *Every symplectic form is locally diffeomorphic to the standard form  $\omega_0$  on  $\mathbf{R}^{2n}$ .*

(For a proof see [A1] or [Wn1, §4]). Thus symplectic manifolds all have the same local structure. It follows that all symplectic

invariants are global in nature, which gives the subject a very different flavor from Riemannian or Kählerian geometry.

(1.2) **Symplectic images of balls.** A diffeomorphism  $\varphi$  is said to be symplectic if  $\varphi^*(\omega) = \omega$ . Since any such diffeomorphism preserves the volume form  $\omega^n$  on  $V$ , the group  $G_{\text{Symp}}(V)$  of symplectic diffeomorphisms is contained in the group  $G_{\text{Vol}}(V)$  of volume-preserving diffeomorphisms of  $V$ . Volume-preserving diffeomorphisms do not seem to have any special geometric properties (apart from the fact that they preserve volume). For example, it follows from Moser's stability theorem for volume forms (see [M] and [Kr]) that if  $D$  is any subset of  $\mathbf{R}^{2n}$  which is diffeomorphic to the closed unit ball  $B = B^{2n}(1)$  and has the same volume as  $B$ , then there is a volume-preserving diffeomorphism  $g$  from  $B$  onto  $D$ . A similar result easily follows for subsets diffeomorphic to the open unit ball: see [GS]. However, it is not hard to see that such a set  $D$  need not be symplectically diffeomorphic to  $B$ , since the restrictions of  $\omega$  to the boundaries of  $D$  and  $B$  need not be equivalent. It is not so easy to rule out the possibility that there is a symplectic diffeomorphism between the interiors of  $B$  and  $D$ , or that the volume-preserving diffeomorphism  $g : B \rightarrow D$  may be uniformly approximated by symplectic embeddings of  $B$  into  $\mathbf{R}^{2n}$ .

Gromov's celebrated squeezing theorem is the first major result in this connection. (A proof will be sketched in §2. See also (3.5).)

**Theorem 1.2.1.** *If there is a symplectic embedding  $g$  of the ball  $B^{2n}(r) \subset \mathbf{R}^{2n}$  of radius  $r$  into the product  $B^2(R) \times \mathbf{R}^{2n-2} \subset \mathbf{R}^2 \times \mathbf{R}^{2n-2} = \mathbf{R}^{2n}$ , then  $r \leq R$ .*

This theorem has many consequences. First observe that, because there are volume-preserving embeddings of  $B^{2n}(r)$  into  $B^2(R) \times \mathbf{R}^{2n-2}$  for any values of  $r$  and  $R$ , it implies that a volume-preserving embedding  $g : B \rightarrow \mathbf{R}^{2n}$  cannot always be uniformly approximated by symplectic embeddings. By using Darboux's theorem together with the isotopy extension theorem for volume-preserving embeddings [Kr], it is not hard to deduce that, for any compact symplectic manifold  $(V, \omega)$ , the group  $G_{\text{Symp}}(V)$  of symplectic diffeomorphisms is not  $C^0$ -dense in the group  $G_{\text{Vol}}(V)$  of volume-preserving diffeomorphisms. But it is a consequence of Gromov's implicit function theorem that, at least if  $H^1(V; \mathbf{R}) = 0$ , there are no subgroups of  $G_{\text{Vol}}(V)$  lying strictly between

$G_{\text{Symp}}(V)$  and  $G_{\text{Vol}}(V)$ , except perhaps for  $G_{\pm \text{Symp}}(V)$  which consists of all volume-preserving diffeomorphisms  $h$  such that  $h^*(\omega) = \pm\omega$ . (Note that  $G_{\pm \text{Symp}}(V) = G_{\text{Symp}}(V)$ , unless  $n$  is even.) When  $H^1(V; \mathbf{R}) \neq 0$ , the above statement may not be quite true because of problems arising from the flux homomorphism. However, in any case, one can easily deduce:

**Corollary 1.2.2.** (See [G2, 3.4.4 (H)].)  $G_{\text{Symp}}(V)$  is  $C^0$ -closed in the group of all diffeomorphisms of  $V$ .

A much simpler proof of this corollary was given by Eliashberg in [E1] and rediscovered in [EH]. They observe that if  $\{g_k\}$  is a sequence of symplectic embeddings of  $(B, 0)$  into  $(\mathbf{R}^{2n}, 0)$  which converges uniformly to a differentiable map  $g_0$ , then Theorem 1.2.1 implies that the derivative  $dg_0(0)$  of  $g_0$  at 0 cannot map the unit ball  $B$  into any set of the form  $L(B^2(R) \times \mathbf{R}^{2n-2})$ , where  $R < 1$  and  $L \in \text{Sp}(2n, \mathbf{R})$  is a linear symplectic map. (This follows because  $dg_0(0)$  is approximated by suitable rescalings  $tg_k(x/t)$  of the  $g_k$ .) It remains to show that this fact implies that  $dg_0(0)$  itself belongs to  $\text{Sp}(2n, \mathbf{R})$ , which is a question of linear algebra.

Another consequence of Theorem 1.2.1 is that it provides us with ways to measure the symplectic size of subsets  $U$  of  $\mathbf{R}^{2n}$ . For example, let  $\text{proj}$  denote the projection of  $\mathbf{R}^{2n}$  onto the  $(x_1, x_2)$ -plane and define  $c(U)$  by:

$$c(U) = \inf\{\text{area}(\text{proj } g(U)) : g \in G_{\text{Symp}}(\mathbf{R}^{2n})\}.$$

Clearly,  $c(U)$  is a symplectic invariant of  $U$  which is monotone (i.e.  $c(U_1) \leq c(U_2)$  if  $U_1 \subset U_2$ ). Further, because there are no nontrivial invariants of closed regions of the plane under area-preserving diffeomorphisms, it is not hard to see that

$$c(B^{2n}(1)) = c(B^2(1) \times \mathbf{R}^{2n-2}) = \text{area } B^2(1) = \pi.$$

This, together with the homogeneity property  $c(\lambda U) = \lambda^2 c(U)$ , shows that, in contrast to the volume,  $c$  is a 2-dimensional invariant. It is one of the capacity functions considered by Ekeland and Hofer in [EH]. As remarked by Weinstein, its existence is closely related to the uncertainty principle, since it gives one a way to measure a quantity  $c(U)$  which can reasonably be thought of as the ‘‘uncertainty’’ involved in predicting the pair  $(x_1, x_2)$  for points  $(x_1, \dots, x_{2n})$  in the subset  $U$  of  $\mathbf{R}^{2n}$ . (Here one identifies  $\mathbf{R}^{2n}$  with phase space  $T^*(\mathbf{R}^n)$ , taking  $x_1$  as the first spacial coordinate and  $x_2$  as its conjugate momentum.)

There are many possible definitions of these symplectic capacities: for example, Gromov defines a symplectic radius in [G1, 0.3.A], and in [EH] Ekeland and Hofer give a definition connected with the behavior of the periodic orbits of the Hamiltonian flow on the boundary of smooth regions  $U$ . However, the definitions all seem to agree at least on ellipsoids. The properties of these functions have been fruitfully explored by Viterbo, Hofer, Ekeland, and Zehnder. Using variational methods which exploit the special structure of  $(\mathbf{R}^{2n}, \omega_0)$ , these authors developed a powerful theory which is described in the survey articles [H2], [V1], and [V2]. As shown in [H3], this theory also has important applications in the study of symplectic diffeomorphisms of  $\mathbf{R}^{2n}$ . We will see in (3.5) that one way to extend this theory to general symplectic manifolds is by means of Floer's elliptic techniques.

(1.3) **Fixed point theorems.** Another important group of questions concern the number of fixed points of a symplectic diffeomorphism. The prototypical result is:

(1.3.1) **Birkhoff's twist theorem.** *Let  $\varphi$  be an area-preserving diffeomorphism of the annulus  $\{(x, y), \in \mathbf{R}^2 : a \leq x^2 + y^2 \leq b\}$  which rotates the inner and outer boundaries in opposite directions. Then  $\varphi$  has at least two distinct fixed points.*

Note that this theorem is false without the twist condition since a rotation has no fixed points on the annulus. It also clearly fails if  $\varphi$  does not preserve area. In dimension 2, the symplectic and volume-preserving cases coincide, and so when one tries to generalize a result such as this to higher dimensions, one can look either in the symplectic or in the volume-preserving categories. However, just as in (1.2) the volume-preserving condition is too weak to provide interesting results. In fact, using Gromov's method of convex integration one can construct a nonvanishing volume-preserving (i.e. divergence free) vector field  $\xi$  on any manifold of dimension  $\geq 3$  which has zero Euler class: see [G2, 2.4.3]. One can even make  $\xi$  irrotational (which is the appropriate version of the twist condition). Hence, there is no volume-preserving analogue of Birkhoff's twist theorem. On the other hand, there is a generalization to symplectic diffeomorphisms which are  $C^1$ -close to the identity. To explain this, we must recall some facts about Hamiltonian flows.

(1.3.2) *Hamiltonian flows and exact diffeomorphisms.* Just as in the case of an inner product (i.e. nondegenerate symplectic bilinear form), a nondegenerate skew-symmetric bilinear form  $\omega$  on a manifold  $V$  gives rise to a linear isomorphism from the tangent bundle  $TV$  to its dual  $T^*V$  which takes the vector field  $\xi$  to the 1-form  $\beta = i(\xi)\omega$  given by contracting  $\omega$  with  $\xi$ . Thus,  $\beta(\zeta) = \omega(\xi, \zeta)$ . It is not hard to check that the flow  $\varphi_t$  associated with  $\xi$  preserves  $\omega$  if and only if the corresponding 1-form  $\beta$  is closed. (This follows from the fact that the Lie derivative  $\mathcal{L}_\xi(\omega)$  of  $\omega$  along  $\xi$  is given by  $i(\xi)d\omega + d(i(\xi)\omega) = d\beta$ .) In particular, every smooth real-valued function  $H$  on  $V$  gives rise to a closed 1-form  $\beta = dH$  and hence to an associated symplectic flow  $\Psi_t$ . Such a flow  $\Psi_t : V \rightarrow V$  is called a *Hamiltonian flow*, and the function  $H$ , although it is just an ordinary function, is called a *Hamiltonian function*. The basic example is the time flow of classical mechanical system, which is generated by the “energy” function  $H : \mathbf{R}^{2n} \rightarrow \mathbf{R}$ .

It follows from the above remarks that every 1-parameter subgroup  $\{\psi_t\}$  of  $G_{\text{Symp}}(V)$  corresponds to a closed 1-form  $\beta$ . Clearly, the fixed points of the diffeomorphisms  $\psi_t$  correspond exactly to the zeros of  $\beta$  when  $t$  is sufficiently small. In particular, if  $\beta$  is exact, it has at least as many zeros as a smooth function on  $V$  has critical points; while if  $\beta$  is not exact, it need have no zeros. Hence, the largest group of diffeomorphisms which one might hope to have “many fixed points” is the closed subgroup of  $G_{\text{Symp}}(V)$  generated by those diffeomorphisms which belong to some Hamiltonian flow. The elements of this subgroup are said to be *exact*. It is not hard to show that the exact diffeomorphisms are precisely those which may be joined to the identity by a smooth path  $\varphi_t$ ,  $0 \leq t \leq 1$ , which is generated by a time-dependent Hamiltonian function  $H_t$ . In other words, the path  $\varphi_t$  is tangent to a family  $\xi_t$  of vector fields such that  $i(\xi_t)\omega = dH_t$  for each  $t$ . Equivalently, the path  $\varphi_t$  has zero flux through all 1-cycles  $\gamma$ , in the sense that  $\omega$  has zero integral over the cylinders  $\varphi_t(\gamma)$ ,  $0 \leq t \leq 1$ . Thus, the lift  $(\varphi_1, \{\varphi_t\})$  of  $\varphi_1$  to the universal cover  $\overline{G}$  of  $G_{\text{Symp}}(V)$  is in the kernel of the flux homomorphism  $\Phi$  from  $\overline{G}$  onto  $H^1(V, \mathbf{R})$ : see [MD1]. As an example, observe that a symplectic diffeomorphism of the standard torus  $T^{2n}$  which is isotopic to the identity is exact, if and only if it is covered by a symplectic diffeomorphism of  $\mathbf{R}^{2n}$  which moves the fundamental

domain  $\Delta$  of  $T^{2n}$  to a set  $\Delta'$  which has the same center of gravity as  $\Delta$ . (See [CZ, Appendix], where this is proved for  $n = 1$ . The argument in fact works for all  $n$ .)

(1.3.3) *Arnold's conjectures.* Arnold conjectured that the number of distinct fixed points of an exact symplectic diffeomorphism  $\varphi$  of  $V$  should be at least the number of critical points of some function  $f(\varphi)$  on  $V$ . (See [A1, Appendix 9].) It is easy to see that this statement holds for exact diffeomorphisms which are sufficiently  $C^1$ -close to the identity since, by [Wn1, §6] for example, these may be identified via their graphs with exact 1-forms  $df$  in such a way that their fixed points correspond to the critical points of  $f$ . But it is not at all clear that it holds globally. The conjectures have actually not been proved in quite the form in which Arnold stated them, since it is hard to work with estimates of the number of critical points of a function. Therefore, we will follow Floer in [F5] and make the following definitions.

**Definition 1.3.4.** A diffeomorphism  $\varphi$  satisfies the cup-length estimate with respect to a ring  $R$  if the number  $\#F(\varphi)$  of its fixed points is greater than or equal to the cup-length  $\text{cl}(V, R)$  of  $V$ , i.e., to the largest integer  $k$  such that there is a nonzero cup-product in  $H^*(V; R)$  of length  $k - 1$ .

Since, by Ljusternik–Schnirelman theory (see [B]), the number of critical points of a function on  $V$  is at least  $\text{cl}(V; \mathbf{Z})$ , any exact diffeomorphism of  $(V, \omega)$  which is sufficiently  $C^1$ -close to the identity satisfies the cup-length estimate with respect to any ring.

**Definition 1.3.5.** If all fixed points of  $\varphi$  are nondegenerate, then the fixed point set  $F(\varphi)$  of  $\varphi$  is finite, and we will say that  $\varphi$  satisfies the ungraded Morse inequalities with respect to  $R$  if there exists a homomorphism  $\partial : F_* \rightarrow F_*$  of the free  $R$ -module  $F_*$  with generators  $F(\varphi)$ , such that  $\partial^2 = 0$  and  $\text{Ker } \partial / \text{Im } \partial$  is isomorphic to the direct sum  $H_*(V; R)$  of the homology groups  $H_k(V; R)$ ,  $0 \leq k \leq \dim V$ .

In particular, if this holds for some field  $\mathcal{K}$ , we have  $\#F(\varphi) \geq \sum_k \beta_k$ , where  $\beta_k$  is the Betti number  $\dim H_k(V, \mathcal{K})$ . In the classical case of a Morse function  $f$  on a finite-dimensional manifold, the generators  $F(f)$  correspond to the fixed points of  $f$  and their index provides a natural grading of  $F_*$  which is compatible with

that on  $H_*(V; R)$ . It follows easily that the existence of a boundary operator  $\partial$  of degree  $-1$  corresponds to the classical (graded) Morse inequalities

$$N_k - N_{k-1} + N_{k-2} - \dots \pm N_0 \geq \beta_k - \beta_{k-1} + \beta_{k-2} - \dots \pm \beta_0,$$

for  $0 \leq k \leq \dim V$ , with equality when  $k = \dim V$ , where  $N_k = \dim F_k$ . Again, any exact diffeomorphism of  $(V, \omega)$  with nondegenerate fixed points which is sufficiently  $C^1$ -close to the identity satisfies both versions of the Morse inequalities.

We will say that the *Arnold conjectures hold on*  $(V, \omega)$  with respect to a coefficient ring  $R$  if every exact diffeomorphism satisfies the cup-length estimates for  $R$ , and (provided it is nondegenerate) the ungraded Morse inequalities. For example, Arnold's conjectures say that every exact diffeomorphism of the standard torus  $T^{2n}$  should have at least  $2n + 1$  distinct fixed points, and at least  $2^{2n}$  if they are all nondegenerate. By way of contrast, because  $\pi_1(\mathbb{C}P^n) = 0$ , all symplectic diffeomorphisms of the Kähler manifold  $\mathbb{C}P^n$  are exact, and Arnold's conjectures predict that they all have at least  $n + 1$  distinct fixed points. In the nondegenerate case, this follows from the Lefschetz fixed point formula, but in the degenerate case Arnold's conjecture requires  $n + 1$  distinct fixed points while Lefschetz's formula would be satisfied if there were just one fixed point of multiplicity  $n + 1$ .

The first major contribution to the solution of this problem was by Conley and Zehnder in 1982, who established the validity of these conjectures for the standard torus: see [CZ]. To do this, they reformulated the problem so that it became a question of finding critical points for some functional on the infinite-dimensional space  $\Omega\mathbb{R}^{2n} = C^\infty(S^1, \mathbb{R}^{2n})$  of loops on  $\mathbb{R}^{2n}$ . This approach is explained in (3.1.1) below. They then reduced the problem to finite dimensions by using Fourier analysis, and counted the critical points by means of an index theory developed by Conley.

These variational techniques were very influential in the development of the theory of capacities mentioned in (1.2) above. They have also been refined by several authors in an attempt to extend the above result to other manifolds. In particular, Weinstein [Wn2] succeeded in showing that the Arnold conjectures hold for all exact diffeomorphisms (on any symplectic manifold) which are isotopic to the identity through a  $C^0$ -small isotopy. However, no one has so far found purely variational methods which work for an arbitrary exact diffeomorphism on an arbitrary symplectic manifold.



Instead, Floer took a different tack and, influenced by ideas of Witten, Gromov, and Conley, defined a “Conley index” by elliptic methods. This Morse-theoretic invariant is described in (3.2) below. It is easiest to work with when the symplectic form vanishes on all elements of  $\pi_2(V)$ . However, even if this condition is not satisfied, one can still manage if  $(V, \omega)$  is *monotone*, that is, if there is a constant  $k > 0$  such that the cohomology class  $[\omega] - kc_1(V)$  is zero on each 2-sphere in  $V$ . (As we shall see in §2, the tangent bundle of any symplectic manifold has a complex structure which is well defined up to homotopy and so it has Chern classes  $c_i(V)$ .) Using this, Floer established the following result.

**Theorem 1.3.6.** [F4, Theorem 1] and [F7]. *Let  $(V, \omega)$  be a compact symplectic manifold. If both  $[\omega]$  and  $c_1(V)$  vanish on  $\pi_2(V)$ , then Arnold’s conjectures hold for every coefficient ring  $R$ . Moreover, if  $(V, \omega)$  is monotone, then the Morse inequalities are satisfied with respect to any ring  $R$  by any exact diffeomorphism with nondegenerate fixed points.*

In §3, we sketch the proof of the latter statement in the case  $R = \mathbf{Z}/2\mathbf{Z}$ . The paper [F5] also contains various partial results on the degenerate case for certain special monotone manifolds, as well as a nice history of the problem. An alternative and simpler treatment of the degenerate case (again for  $\mathbf{Z}/2\mathbf{Z}$  coefficients) is given by Hofer in [H1].

§2 *J*-HOLOMORPHIC SPHERES

(2.1) **Basic definitions.** Because  $\text{Sp}(2n, \mathbf{R})$  deformation retracts onto  $U(n, \mathbf{R})$ , the tangent bundle of any symplectic manifold has a complex structure which is well defined up to homotopy. It is usually described by the automorphism  $J : TV \rightarrow TV$  which corresponds to multiplication by  $i$ . This automorphism, which must satisfy the equation  $J^2 = -Id$ , is called an almost complex structure on  $V$ . Of course, in general  $J$  is not integrable, i.e. one cannot (even locally) choose complex coordinates on  $V$  which induce this map  $J$  on  $TV$ . However, as we will see below, one can always choose  $J$  so that it is compatible with  $\omega$  in the sense that

$$(2.1.1)(i) \quad \omega(\xi, J\xi) > 0 \quad \text{for all nonzero } \xi \in TV,$$

and

$$(2.1.1)(ii) \quad \omega(J\xi, J\xi) = \omega(\xi, \xi) \quad \text{for all } \xi, \zeta \in TV.$$

Then the bilinear form  $\mu_J$  defined by setting  $\mu_J(\xi, \zeta) = \omega(\xi, J\zeta)$  is symmetric by (ii) and positive by (i), so that it is a Riemannian metric on  $V$ . Such a metric is said to be compatible with  $\omega$ . An example is a Kähler manifold  $V$  with its integrable complex structure  $J$ , its Kähler form  $\omega$ , and its Kähler metric.

A map  $f$  from a Riemann surface  $(\Sigma, J_0)$  to an almost complex manifold  $(V, J)$  is said to be  $J$ -holomorphic (or pseudo-holomorphic) if the derivative  $df : T\Sigma \rightarrow TV$  is complex linear, that is, if  $df \circ J_0 = J \circ df$ . The image of  $f$  is often called a “ $J$ -holomorphic curve”, or, when  $\Sigma = S^2$ , a  $J$ -holomorphic sphere. When  $J$  is integrable, the space of such curves tells one a great deal about the structure of  $V$ . This was well known in the case of complex surfaces, but has only recently been used to analyze 3-folds: see [Ko].

Gromov realized that a considerable part of the theory goes through even when  $J$  is not integrable. One reason for this is that there are no local obstructions to the existence of  $J$ -holomorphic curves, because every almost complex structure on a 2-manifold is integrable. Further, the ellipticity of the Cauchy-Riemann equation  $df + J \circ df \circ J_0 = 0$  implies that, for generic  $J$ , its space of solutions  $\mathcal{M}(J)$  is an oriented finite-dimensional manifold. When  $J$  is compatible with  $\omega$ , one can even compactify many components of the moduli space  $\mathcal{M}(J)/G$  of unparametrized  $J$ -holomorphic curves by adding to them certain well-understood families of degenerate curves. (Here  $G$  is the reparametrization group of holomorphic transformations of  $\Sigma$ .) As  $J$  varies in the space of  $\omega$ -compatible  $J$ , the manifold  $\mathcal{M}(J)/G$  changes at most by a cobordism, so that (provided that it can be compactified) any invariants of its cobordism class are actually invariants of the symplectic form  $\omega$ . (Recall that compact-oriented manifolds  $M_0$  and  $M_1$  are said to be cobordant if there is a compact-oriented manifold  $W$  whose boundary is the disjoint union of  $M_1$  with  $-M_0$ .) In the next sections, we shall outline this theory, using Floer’s formulation of the analysis. For the sake of simplicity we will restrict ourselves to the case when  $\Sigma$  is the Riemann sphere  $S^2$ .

(2.2) **Compatible metrics and conformality.** To see that compatible almost complex structures exist, observe that if  $\mu$  is any Riemannian metric on  $V$  there is a unique automorphism  $A$  of  $TV$  such that  $\omega(\xi, \zeta) = \mu(A\xi, \zeta)$ , for all  $\xi, \zeta \in TV$ . Using the skew symmetry of  $\omega$ , one easily sees that  $-A^2$  is positive definite with

respect to  $\mu$ . Further, one can check that the metric  $\mu'$  defined by  $\mu'(\xi, \zeta) = \mu(\xi, (-A^2)^{1/2}\zeta)$  satisfies (2.1.1) with  $J = (-A^2)^{-1/2}A$  and so is compatible with  $\omega$ . It follows that the set of all metrics deformation retracts onto the set of compatible metrics, so that the latter set is contractible and nonempty, as is the set of  $\omega$ -compatible  $J$ .

Now suppose that  $\mu_0$  is the standard metric on  $S^2$  and let  $\mu_J$  be any compatible metric on  $V$ . Then, because a  $J$ -holomorphic map  $f : S^2 \rightarrow V$  preserves right angles, it is conformal with respect to  $\mu_0$  and  $\mu_J$ , i.e. there is a positive function  $\rho$  on  $S^2$  such that  $f^*(\mu_J) = \rho\mu_0$ . Therefore, if  $|df_z|$  denotes the norm of the linear map  $df_z : T_z S^2 \rightarrow T_{f(z)}V$  with respect to these metrics,  $|df_z| = \rho(z)$ , and so the  $L^2$ -norm  $\|df\|_2$  of  $df$  is given by

$$\|df\|_2^2 = \int \rho(z)^2 \mu_0 = \mu_J\text{-area of } (\text{Im } f) = \int f^*(\omega).$$

Thus, it depends only on the homology class represented by  $f$ , and is not affected by conformal reparametrizations of  $f$ . The quantity  $A(f) = \|df\|_2^2$  is called the *area* of  $f$ , or sometimes its *harmonic energy*. It is this boundedness of the area functional which underlies the proof of compactness, just as the boundedness of the  $L^2$ -norm of the curvature underlies the analysis of the compactness properties of Yang-Mills connections: see [Wo, §2].

Note also that any  $J$ -holomorphic map  $f : S^2 \rightarrow (V, \mu_J)$  is absolutely area minimizing in its homology class. As in the Kähler case described in [L, Theorem 2.2], this follows from Wirtinger's inequality, which in this context states that  $\omega(\tau_1, \tau_2) \leq 1$  for any pair of orthonormal vectors  $\tau_1$  and  $\tau_2$ , with equality if and only if  $\tau_2 = J\tau_1$ . In particular, any  $J$ -holomorphic sphere in  $(V, \mu_J)$  satisfies the minimal surface equation.

(2.3) **The analytic setup.** Our analytic setup is an adaptation of that in [F1] and [F5] and is nicely described in [FHV]. For  $p > 2$ , let  $W_{1,p}(S^2, \mathbf{R}^N)$  denote the Sobolev space of all maps  $f : S^2 \rightarrow \mathbf{R}^N$  whose derivative  $df$  is  $L^p$ , with the corresponding Sobolev norm. By the Sobolev embedding theorems, each such  $f$  is continuous. Hence, if we embed the manifold  $V$  in  $\mathbf{R}^N$ , we may define  $W_A = W_{1,p}(S^2, V)_A$  to be the set of all such functions with values in  $V$  which represent the homology class  $A \in H_2(V; \mathbf{Z})$ . As in [K, Chapter 1.2], one can prove that  $W_A$  is a smooth Banach manifold. Further, the inclusion  $W_{1,p}(S^2, V)_A \rightarrow C^0(S^2, V)$  is

compact, so that every bounded sequence in  $W_{1,p}(S^2, V)_A$  has a subsequence which converges uniformly to some continuous maps.

For each  $f \in W_A$  and  $p > 2$ , let  $\Lambda^p(f)$  be the space of  $L^p$ -sections of the bundle  $\text{Hom}(TS^2, f^*(TV))$ , and let  $\overline{\Lambda}^p(f)$  be the subbundle of all antiholomorphic sections; i.e.

$$\overline{\Lambda}^p(f) = \{\gamma \in \Lambda^p(f) : \gamma - J \circ \gamma \circ J_0 = 0\}.$$

Again, one can show that these Banach spaces fit together to form Banach bundles  $\mathcal{L}$  and  $\overline{\mathcal{L}}$  over  $W_A$ .

Now, let  $\mathcal{F}$  be the Fréchet space of all  $C^\infty$ -smooth  $\omega$ -compatible almost complex structures on  $V$ . For each  $J \in \mathcal{F}$  and  $f \in W_A$ , the operator

$$\overline{\partial}_J f = df + J \circ df \circ J_0$$

defines a smooth section  $\overline{\partial}_J$  of  $\overline{\mathcal{L}}$  over  $W_A$  whose zeros are the  $J$ -holomorphic maps. Let us write this out in local coordinates, choosing a local holomorphic coordinate  $z = x + iy$  near some point  $p \in S^2$ , and identifying a neighborhood of  $f(p)$  with a neighborhood of  $0 \in \mathbf{R}^{2n}$  in such a way that the induced almost complex structure  $J$  on  $\mathbf{R}^{2n}$  equals the standard one  $J_0$  at the origin. Then the homomorphism  $\overline{\partial}_J f$  takes the vector field  $\partial/\partial x$  on  $S^2$  to the vector field on  $\mathbf{R}^{2n}$  whose  $\alpha$ th component is:

(2.3.1)

$$\begin{aligned} [\overline{\partial}_J f(\partial/\partial x)]^\alpha &= \partial f^\alpha/\partial x + [J(f(z))]_{\alpha\beta} \partial f^\beta/\partial y \\ &= \partial f^\alpha/\partial x + [J_0]_{\alpha\beta} \partial f^\beta/\partial y + [A]_{\alpha\beta} \partial f^\beta/\partial y \end{aligned}$$

where the entries of the matrix  $A$  are functions of  $z = x + iy$  which vanish at  $p$ . Hence the first-order terms of the linearization of the operator  $f \mapsto \overline{\partial}_J f(\partial/\partial x)$  at  $p$  make up the usual  $\overline{\partial}$ -operator for functions  $\xi : \mathbf{C} \rightarrow \mathbf{C}^n = \mathbf{R}^{2n}$ . Thus this linearization is elliptic. Note also that the homomorphism  $\overline{\partial}_J f$  is antiholomorphic and so is determined by its value on  $\partial/\partial x$ , i.e.  $\overline{\partial}_J f(\partial/\partial y) = \overline{\partial}_J f(J_0(\partial/\partial x)) = -J \overline{\partial}_J f(\partial/\partial x)$ .

One says that a section  $\sigma$  of the bundle  $\pi : \overline{\mathcal{L}} \rightarrow W_A$  is *elliptic* near its zero section if, for every  $f \in W_A$  for which  $\sigma(f) = 0$ , there is a trivialization  $\tau : \pi^{-1}(U) \rightarrow U \times \overline{\Lambda}^p(f)$  over some neighborhood of  $f$  such that the composite map  $U \rightarrow \overline{\Lambda}^p(f)$  given by  $g \mapsto \tau \circ \sigma(g)$  is an elliptic partial differential operator. The remarks in the previous paragraph show that  $\overline{\partial}_J$  is elliptic.

Hence, a generic perturbation  $\sigma$  of  $\bar{\partial}_J$  is regular, i.e. the derivative  $d(\tau \circ \sigma)(f)$  is surjective at all points  $f$  such that  $\sigma(f) = 0$ . This implies that the inverse image  $\sigma^{-1}(0)$  of the zero section of  $\bar{\mathcal{L}}$  is a manifold of dimension equal to the (finite) index of  $\bar{\partial}_J$ .

In fact, one can always find regular sections of the form  $\bar{\partial}_J$  provided that one allows  $J$  to vary in a sufficiently large subset of  $\mathcal{J}$ , and provided that one stays away from multiply covered spheres, i.e. spheres of the form  $f \circ \eta$  where  $\eta$  is a  $J_0$ -holomorphic map  $S^2 \rightarrow S^2$  of degree  $> 1$ . In order to avoid this latter difficulty, we will assume from now on that  $A$  is not a multiple class, i.e. that  $A \neq kB$  for any  $B \in H_2(V, \mathbf{Z})$ . In [MD2, §4] we showed that one can always perturb  $J$  within a space of  $W_{k,p}$ -smooth almost complex structures to make  $\bar{\partial}_J$  regular. However, it is often convenient to be able to restrict to  $C^\infty$ -almost complex structures  $J$ . This may be done by Floer’s argument in [F1, §5]: given any  $J \in \mathcal{J}$ , one can define a Hilbert space  $\mathcal{F}' = \mathcal{F}'(J)$  of  $C^\infty$ -perturbations of  $J$  which is so large that its closure in  $\mathcal{J}$  with respect to the  $L^2$ -norm contains an open neighborhood of  $J$  in  $\mathcal{J}$ . The proof of [MD2, Proposition 4.1] then shows:

**Lemma 2.3.2.** *Suppose that  $A$  is not a multiple class. Then the space  $\mathcal{M}_A = \{(f, J) : \bar{\partial}_J f = 0\}$  is a Banach submanifold of  $W_A \times \mathcal{F}'$ , and the projection  $P_A : \mathcal{M}_A \rightarrow \mathcal{F}'$  is Fredholm.*

By elliptic regularity theory (see Theorem 2.4.1 below), the elements  $f$  which occur in  $\mathcal{M}_A$  are  $C^\infty$ -smooth and do not depend on the value of  $p > 2$  used to define  $W_A = W_{1,p}(S^2, V)_A$ .

We will denote the space  $P_A^{-1}(J)$  by  $\mathcal{M}_A(J)$  and will call its elements  $J$ -holomorphic  $A$ -spheres. The index of  $P_A$  may be calculated by the Atiyah-Singer index theorem in terms of the dimension,  $2n$ , of  $V$  and the value,  $c(A)$ , of the first Chern class  $c = c_1(J)$  of  $(V, J)$  on  $A$ . We find:

**Theorem 2.3.3.** *There is a dense set  $\mathcal{J}_{\text{reg}}$  in  $\mathcal{J}$  such that for all  $J \in \mathcal{J}_{\text{reg}}$ , the space  $\mathcal{M}_A(J)$  of  $J$ -holomorphic  $A$ -spheres is an oriented manifold of dimension  $2(c(A) + n)$ . Further, if  $J_1$  and  $J_2$  are two elements of  $\mathcal{J}_{\text{reg}}$ , any path in  $\mathcal{J}$  which joins them may be slightly perturbed (relative to its endpoints) in such a way that its inverse image  $P_A^{-1}(\alpha)$  forms a (noncompact) oriented cobordism between  $\mathcal{M}_A(J_1)$  and  $\mathcal{M}_A(J_2)$ .*

*Proof.* As usual, an element  $J$  of  $\mathcal{F}'$  is said to be *regular* if the derivative (or linearization)  $dP_A$  of  $P_A$  at  $(f, J)$  is surjective for all  $(f, J) \in P_A^{-1}(J)$ . The implicit function theorem for Banach spaces implies that, for regular  $J \in \mathcal{F}'$ , the space  $P_A^{-1}(J)$  is a smooth manifold of dimension  $\text{Int } P_A = 2(c(A) + n)$ . Further, by Proposition 4.3 of [MD2], the manifold  $P_A^{-1}(J)$  has a canonical orientation. (In fact, it has a well-defined stable almost complex structure.)

By standard Fredholm theory (see [Sm]), the set of regular points in  $\mathcal{F}'$  is “of second category” in  $\mathcal{F}'$  (i.e. it contains a countable intersection of open, dense sets). Hence, the set  $\mathcal{F}_{\text{reg}}$  of all elements in  $\mathcal{F}$  which are regular in  $\mathcal{F}'(J)$  for some  $J \in \mathcal{F}$  is dense in  $\mathcal{F}$ . Observe that, although  $\mathcal{F}_{\text{reg}}$  may not have second category in  $\mathcal{F}$ , it has all the essential features of a set of second category. For example, the intersection of two such sets, one defined for the class  $A$  and another for the class  $B$ , is always nonempty.

The second statement may be proved in a similar way, by defining for each path  $\alpha$  a suitable subset  $\mathcal{F}'(\alpha)$  of  $\mathcal{F}$  which is a Banach submanifold and contains  $\alpha$ .  $\square$

When the complex structure  $J$  is integrable, one can often prove that it is a regular value for  $P_A$  by using the following lemma. Recall that Grothendieck proved in [Gr] that any holomorphic bundle  $\eta$  over  $S^2 = CP^1$  is holomorphically equivalent to a sum of holomorphic line bundles. Moreover, this splitting is unique apart from the order in which the summands are arranged, so that  $\eta = L_1 \oplus \dots \oplus L_n$  is completely characterized by the set of Chern classes  $c_1(L_1), \dots, c_1(L_n)$ . Since this set is a topological invariant, it must be the same for all bundles  $f^*(TV)$ , where  $f : S^2 \rightarrow V$  varies in a connected component of the space of  $J$ -holomorphic  $A$ -spheres.

**Lemma 2.3.4.** *Suppose that  $J$  is integrable, and that, for all  $J$ -holomorphic  $A$ -spheres  $f$ , every summand of  $f^*(TV)$  has  $c_1 > -2$ . Then  $J$  is regular for  $P_A$ .*

*Proof.* When  $J$  is integrable, it follows from [MD2, Proposition 4.2] that  $dP_A$  is essentially equal to the usual  $\bar{\partial}$ -operator on sections of  $f^*(TV)$ . In particular,  $dP_A$  is surjective if and only if  $\bar{\partial}$  is surjective, which happens if and only if the Dolbeault cohomology group  $H^{0,1}(CP^1, f^*(TV)) = H^{0,1}(CP^1, L_1) \oplus \dots \oplus H^{0,1}(CP^1, L_n)$  vanishes. But, for any holomorphic line bundle

$L, H^{0,1}(\mathbb{C}P^1, L) \cong H^1(\mathbb{C}P^1, \mathcal{O}(L)) \cong H^0(\mathbb{C}P^1, \mathcal{O}(L^* \otimes K))$  by Kodaira–Serre duality (see [GH, Chapter 1, §2]), and so vanishes if and only if  $c_1(L^* \otimes K) < 0$ . But  $c_1(L^* \otimes K) = c_1(L) - 2$ , and so we need  $c_1(L) > -2$ .  $\square$

**Example 2.3.5.** Let  $V$  be  $S^2 \times S^2 \times \dots \times S^2$  with symplectic form  $\tau = \lambda_1 \sigma_1 \oplus \lambda_2 \sigma_2 \oplus \dots \oplus \lambda_n \sigma_n$ , where the  $\sigma_k$  are area forms on the different factors of  $S^2$  with total area  $\pi$ , and where the numbers  $\lambda_k$  are positive. Let  $A$  be the class represented by  $S^2 \times pt \times \dots \times pt$ . Then  $c(A) = 2$ , and  $\mathcal{M}_A(J)$  has dimension  $2n + 4$ . Further, by Lemma 2.3.4, the usual complex structure  $J_0$  on  $S^2 \times S^2 \times \dots \times S^2$  is a regular value for the projection operator  $P_A$ . Note that the set of  $J_0$ -holomorphic maps from  $S^2$  to  $S^2 \times S^2 \times \dots \times S^2$  is the  $(2n + 4)$ -parameter family  $z \mapsto (\gamma(z), w_2, \dots, w_n)$ , where  $\gamma$  runs through the 6-dimensional reparametrization group  $G = \text{PSL}(2, \mathbb{C})$ .

(2.4) **Compactness.** Theorem 2.3.3 is useless unless one can establish some kind of compactness. For, any manifold  $M$  is cobordant to the empty manifold via the noncompact cobordism  $M \times [0, 1)$ . Now, the manifold  $\mathcal{M}_A(J)$  itself cannot be compact (unless it is empty) since the noncompact group  $G = \text{PSL}(2, \mathbb{C})$  of holomorphic self-maps of  $S^2$  acts on it by reparametrization. However, in many cases the space  $\mathcal{M}_A(J)/G$  of unparametrized spheres is compact.

The key to proving this is the fact, established in (2.2), that the area functional (i.e.  $L^2$ -norm of  $df$ ) is uniformly bounded on  $\mathcal{M}_A(J)$ . This is the “borderline” case of the Sobolev theory. Indeed, as remarked in (2.3), any sequence  $\{f_\alpha\}$  in  $\mathcal{M}_A(J) \subset \mathcal{W}_{1,p}(S^2, V)_A$  which is bounded in the  $L^p_1$ -norm, for some  $p > 2$ , has a subsequence which converges uniformly, and by the usual bootstrapping procedure, one can show that it must converge in the  $L^p_k$ -norm, for any  $k \geq 1$ . On the other hand, if the derivatives  $df_\alpha$  are bounded in  $L^2$  but not in  $L^p$ , a simple geometric argument allows one to construct a  $J$ -holomorphic map  $g : \mathbb{C} \rightarrow V$  with finite area, which, by “removal of singularities”, may be completed to a map with domain  $S^2 = \mathbb{C} \cup \{\infty\}$ . This is the phenomenon of “bubbling off of spheres”. Sometimes, by choosing the class  $A$  carefully, one can show that this cannot happen. In this case, the spaces  $\mathcal{M}_A(J)/G$  are compact. In other cases, bubbling off can occur, and one must proceed with more care.

The following paragraphs contain more details of this argument, again following the approach of Floer. Wolfson in [Wo] has a slightly different viewpoint which emphasizes the affinities between this and other geometric problems concerning harmonic maps of surfaces, Yang–Mills connections, and so on. Gromov’s approach was very different: he argued geometrically, using isoperimetric inequalities and the Schwartz lemma for conformal maps. Details of his proofs have been written up by Pansu [P], but unfortunately have not yet been published. The flavor of his arguments may be sampled in the proof of Theorem 2.4.7 (removal of singularities) which is sketched below.

*Elliptic regularity.* Choose  $p > 2$ , and  $k \geq 1$ , and let  $W_{k,p}(S^2, V)$  be the Sobolev space of all maps  $f: S^2 \rightarrow V$  whose  $k$ th derivatives lie in  $L^p$ . We denote the corresponding Sobolev norm by  $\|\cdot\|_{k,p}$ . By the Sobolev embedding theorems, the inclusion of  $W_{k+1,p}(S^2, V)$  into  $W_{k,p}(S^2, V)$  is compact for all  $k$ . Further, if  $k - 2/p > m + \alpha$  where  $0 < \alpha < 1$ ,  $W_{k,p}(S^2, V)$  embeds compactly into the Hölder space  $C^{m+\alpha}(S^2, V)$ . (Note that the “2” in the conditions  $p > 2$  and  $k - 2/p > n$  denotes the dimension of the domain  $S^2$ .) Because of this, the main elliptic regularity theorem includes a statement on compactness.

**Theorem 2.4.1.** *Let  $k \geq 1$  and  $p > 2$ . If  $f \in W_{k,p}(S^2, V)$  and  $\bar{\partial}_J f = 0$ , then  $f \in C^\infty(S^2, V)$ . Further, every subset of  $\bar{\partial}_J^{-1}(0)$  which is bounded in  $W_{k,p}(S^2, V)$  has compact closure in  $C^m(S^2, V)$ , for all positive integers  $m$ .*

*Proof.* By the above remarks, it clearly suffices to show that if  $S$  is a subset of  $W_{k,p}(S^2, V)$  such that both  $\|f\|_{k,p}$  and  $\|\bar{\partial}_J f\|_{k,p}$  are bounded by the constant  $M$  as  $f$  ranges over  $S$ , then, for all  $f \in S$ ,

$$(*) \quad f \in W_{k+1,p}(S^2, V) \text{ and } \|f\|_{k+1,p} \leq C,$$

where the constant  $C$  depends only on  $k \geq 1$ ,  $M$  and  $J$ .

We first show how to reduce this to a local problem. Because  $W_{1,p}(S^2, V)$  embeds compactly into  $C^\alpha(S^2, V)$ , the elements of  $S$  have a uniform modulus of continuity. Therefore, given an atlas on  $V$ , the elements of  $S$  will take any sufficiently small ball in  $S^2$  into one of the coordinate charts of this atlas. Hence, using partitions of unity on  $S^2$  and  $V$ , we can reduce to the case when



$S$  is a set of maps from the disc  $D_\varepsilon$  of radius  $\varepsilon$  in  $\mathbf{C}$  to  $\mathbf{C}^n$ , such that both  $\|f\|_{k,p}$  and  $\|\bar{\partial}_J f\|_{k,p}$  are bounded for  $f \in S$ , and where  $J$  is an almost complex structure on  $\mathbf{C}^n$  which is standard at  $\{0\}$ .

The statement about the smoothness of  $f$  now follows by the standard trick. Given  $f \in W_{k,p}(D_\varepsilon, \mathbf{C}^n)$  such that  $\bar{\partial}_J f = 0$ , one considers the linear elliptic operator

$$L(u) = du + Jf(z) \circ du \circ i \quad \text{for } u \in W_{1,p}(D_\varepsilon, \mathbf{C}^n).$$

This has coefficients in  $W_{k,p}(D_\varepsilon, \mathbf{C}^n) \subset C^{m+\alpha}(D_\varepsilon, \mathbf{C}^n)$ , where  $k - 2/p > m + \alpha$  as above. Hence, by standard linear theory, all its solutions lie in  $C^{m+1+\alpha}(D_\varepsilon, \mathbf{C}^n)$ . Hence it follows that  $f \in C^{m+1+\alpha}(D_\varepsilon, \mathbf{C}^n)$ . Repeating this, we find that  $f \in C^{m+1+\alpha}(D_\varepsilon, \mathbf{C}^n)$  for all  $m$ , as claimed.

Thus we may suppose that the set  $S$  consists of all  $f \in C^\infty(D_\varepsilon, \mathbf{C}^n)$  with  $\|f\|_{k;\varepsilon} \leq M$ , where  $\|\cdot\|_{k;\varepsilon}$  denotes the norm on  $W_{k,p}(D_\varepsilon, \mathbf{C}^n)$ . (For simplicity, we suppress  $p$  from now on.) Clearly, the a priori estimate (\*) will follow if we show that for each  $M$  there is an  $\varepsilon > 0$  and a constant  $C = C(\varepsilon, J, M)$  such that

$$(**) \quad \|f\|_{k+1;\varepsilon/2} \leq C(\|\bar{\partial}_J f\|_{k;\varepsilon} + \|f\|_{k;\varepsilon}),$$

for all  $f \in S$ . In fact, since  $|f(0)|$  is uniformly bounded for  $f \in S$ , we need only establish (\*\*) for  $f \in S_0 = \{f \in S : f(0) = 0\}$ .

The existence of such a constant  $C$  is well known for the standard  $\bar{\partial}$ -operator since this is linear. Indeed, one can prove this explicitly by expressing  $f$  as  $Sf + T\bar{\partial}f$ , where  $S$  is the Cauchy kernel

$$Sf(z) = \frac{1}{2\pi i} \oint_{\partial D_\varepsilon} \frac{f(\zeta)}{\zeta - z} d\zeta$$

and  $T$  is the integral operator

$$Tu(z) = \frac{1}{2\pi i} \iint_{D_\varepsilon} \frac{u(\zeta)}{\zeta - z} d\zeta \wedge d\bar{\zeta}.$$

Basic properties of  $T$ , plus the Hölder estimates corresponding to (\*\*), are worked out in [NW]. For example, they show that both  $\sup|Tu(z)|$  and  $\sup|dT_u(z)|$  for  $z \in D_\varepsilon$  are bounded by  $\sup|u(z)|$ . On the other hand, it is easy to see that one can only estimate  $Sf(z)$  and  $dSf(z)$  in terms of  $f|\partial D_\varepsilon$  for  $z$  in some compact subset of  $\text{Int } D_\varepsilon$ . This is why the norm  $\|f\|_{k;\varepsilon/2}$  appears on the left of (\*\*).

Following [F1] Lemma 2.3, we now show how to extend this result to  $\bar{\partial}_J$ . We will use the fact that, because  $Sf = 0$  when  $f$  has compact support, (\*\*) implies that there is a constant  $C_1$  such that

$$(\#) \quad \|u\|_{k+1:\varepsilon} \leq C_1 \|\bar{\partial}u\|_{k:\varepsilon}$$

for all  $C^\infty$  functions  $u$  with compact support in  $\text{Int } D_1$  and any  $\varepsilon \leq 1$ . Let  $\beta : \text{Int } D_1 \rightarrow [0, 1]$  be a smooth bump function of compact support, which equals 1 on  $D_{1/2}$ , and set  $\beta_\varepsilon(x) = \beta(x/\varepsilon)$ . Then  $\beta_\varepsilon f$  has compact support in  $\text{Int } D_\varepsilon$  and clearly there is a constant  $C_2 = C_2(\varepsilon)$  such that

$$(\#\#) \quad \|f\|_{m:\varepsilon/2} \leq \|\beta_\varepsilon f\|_{m:\varepsilon} \leq C_2 \|f\|_{m:\varepsilon} \quad \text{for } m = k, k + 1.$$

Thus, there are constants  $C_1$  and  $C_3$ , which are independent of  $\varepsilon \leq 1$ , and a constant  $C_\varepsilon$  which depends on  $\varepsilon$ , such that

$$\begin{aligned} \|\beta_\varepsilon f\|_{k+1:\varepsilon} &\leq C_1 \|\bar{\partial} \beta_\varepsilon f\|_{k:\varepsilon} \quad \text{by } (\#), \\ &\leq C_1 \{C_2 \|\bar{\partial}_J f\|_{k:\varepsilon} + \|\beta_\varepsilon (J\beta_\varepsilon f(z) - J_0)df \circ i\|_{k:\varepsilon} + C_\varepsilon \|f\|_{k:\varepsilon}\} \\ &\leq C_1 \{C_1 \|\bar{\partial}_J f\|_{k:\varepsilon} + \|J\beta_\varepsilon f(z) - J_0\|_{\text{sup}:\varepsilon} \cdot \|\beta_\varepsilon f\|_{k+1:\varepsilon} + \\ &\quad C_3 \|J\beta_\varepsilon f(z) - J_0\|_{k:\varepsilon} \cdot \|\beta_\varepsilon f\|_{k:\varepsilon} + C_\varepsilon \|f\|_{k:\varepsilon}\}, \end{aligned}$$

where  $\|\cdot\|_{\text{sup}:\varepsilon}$  is the sup norm on  $D_\varepsilon$ . Because the elements of  $S_0$  have a uniform modulus of continuity and because  $f(0) = 0$  and  $0 \leq \beta_\varepsilon \leq 1$ , there is a constant  $K = K(M, J)$  such that  $\|J\beta_\varepsilon f(z) - J_0\|_{\text{sup}:\varepsilon} \leq \varepsilon K$ . Thus, if we choose  $\varepsilon$  so small that  $C_1 \varepsilon K \leq 1/2$  we may take the term involving  $\|\beta_\varepsilon f\|_{k+1:\varepsilon}$  to the left of the above inequality, thereby estimating  $\|\beta_\varepsilon f\|_{k+1:\varepsilon}$ . Inequality (\*\*) now follows easily from (##).  $\square$

*Bubbles.* We now prove the simplest version of Gromov’s compactness theorem, which gives a criterion for the moduli space  $\mathcal{M}_A(J)/G$  to be compact. If this criterion is not satisfied, it is often possible to compactify  $\mathcal{M}_A(J)/G$  by adding suitable “cusp-curves”, but we will not discuss this here. Throughout, we assume that  $A$  is not a multiple class.

**Theorem 2.4.2.** *If  $\mathcal{M}_A(J)/G$  is not compact, there is a continuous map  $g : S^2 \rightarrow V$  which represents a homology class  $B$  such that  $0 < \omega(B) < \omega(A)$ .*

*Proof.* Suppose that  $\{f_\alpha\}$  is a sequence in  $\mathcal{M}_A(J)$  whose image in  $\mathcal{M}_A(J)/G$  has no convergent subsequence. By (2.2) the sequence  $\{f_\alpha\}$  is bounded in the  $W_{1,2}$ -norm. Further, because each

$f_\alpha$  is a smooth function on the compact manifold  $S^2$ , there is a point  $z_\alpha$  at which  $|df(z_\alpha)|$  attains its maximum. If the sequence  $\{|df_\alpha(z_\alpha)|\}$  is bounded, then  $\{f_\alpha\}$  is bounded in the  $W_{1,p}$ -norm for  $p > 2$  and so converges in  $\mathcal{M}_A(J)$  by Theorem 2.4.1. So let us suppose that  $|df_\alpha(z_\alpha)| = r_\alpha \rightarrow \infty$ . We now rescale the  $f_\alpha$  conformally, defining maps  $g_\alpha$  from the disc  $B(0, r_\alpha)$  of radius  $r_\alpha$  about  $0 \in \mathbb{C}$  to  $V$  by setting

$$(2.4.3) \quad g_\alpha(z) = f_\alpha(\psi_\alpha(z)),$$

where  $\psi_\alpha$  is a conformal map of  $B(0, r_\alpha)$  into  $S^2$  which takes 0 to  $z_\alpha$  and is such that  $|d\psi_\alpha(0)| = \sup |d\psi_\alpha(z)| = 1/r_\alpha$ . Then the  $g_\alpha$  are  $J$ -holomorphic, and

$$(2.4.4)(i) \quad \|dg_\alpha\|_2^2 \leq \omega(A) \quad \text{by (2.2);}$$

$$(2.4.4)(ii) \quad |dg_\alpha(0)| = \sup |dg_\alpha(z)| = 1.$$

Hence, for each  $R$ , the restrictions of the  $g_\alpha$  to  $B(0, 2R)$  are uniformly bounded in the  $W_{1,p}$ -norm, and so, by Theorem 2.4.1, have a subsequence which converges in the  $C^1$ -topology to a holomorphic map  $g_R : B(0, R) \rightarrow V$ . Repeating this for all positive integers  $R$  and choosing subsequences so that  $g_{R+1}$  extends  $g_R$ , we obtain a  $J$ -holomorphic map  $g : \mathbb{C} \rightarrow V$  such that

$$(2.4.5)(i) \quad \|dg\|_2^2 \leq \omega(A)$$

and

$$(2.4.5)(ii) \quad |dg(0)| = 1.$$

By Theorem 2.4.7 below, this map  $g$  extends to a continuous map of  $S^2 = \mathbb{C} \cup \{\infty\}$  to  $V$ , which cannot be constant because of (2.4.5)(ii).

By (2.2),  $\|dg\|_2^2 = \omega(B)$ , where  $B$  is the class represented by  $g$ . Hence if  $\|dg\|_2^2 < \omega(A)$ , the proof is complete. On the other hand, if  $\|dg\|_2^2 = \omega(A)$ , we claim that the  $f_\alpha$  may be reparametrized so that they converge to  $g$ . To see this, let  $\gamma_\alpha \in G$  be the reparametrization of  $f_\alpha$  which corresponds to the rescaling

in (2.4.3), and let  $w \in S^2$  be the point corresponding to  $\{\infty\}$ . Then  $f'_\alpha = f_\alpha \circ \gamma_\alpha$  converges to  $g$  in  $W_{1,p,\text{loc}}(S^2 - \{w\}, V)$ . If the norms  $\|df'_\alpha\|_{p; B(w,1)}$  are bounded, then the  $f'_\alpha$  converge in  $W_{1,p}(S^2, V)$  by Theorem 2.4.1, and the limit has to be  $g$ . But if these norms are unbounded, we may repeat the argument at the beginning of this proof and find a subsequence which, after appropriate rescaling near  $w$ , converges to a nonconstant continuous map  $h : S^2 \rightarrow V$ . Because  $g$  and  $h$  are limits of “disjoint pieces” of the  $f'_\alpha$ , we have  $\|dg\|_2^2 + \|dh\|_2^2 \leq \omega(A)$ . Therefore our hypothesis  $\|dg\|_2^2 = \omega(A)$  implies that  $\|dh\|_2^2 = 0$  which is impossible if  $h$  is nonconstant.  $\square$

*Note 2.4.6.* For the sake of simplicity, Theorem 2.4.2 was stated for a fixed  $J$ . However, the same methods show that if  $\{f_\alpha\}$  is a sequence of  $J_\alpha$ -holomorphic  $A$ -spheres where the  $J_\alpha$  tend to  $J_0$  in the  $C^\infty$ -topology, then either the  $f_\alpha$  may be reparametrized so that they converge to a  $J_0$ -holomorphic  $A$ -sphere, or there is a continuous map  $g : S^2 \rightarrow V$  which represents a homology class  $B$  such that  $0 < \omega(B) < \omega(A)$ .

**Theorem 2.4.7** (Removal of singularities). *If  $g$  is a  $J$ -holomorphic map with finite area of the punctured disc  $D - \{0\} \subset \mathbb{C}$  to  $V$ , then  $g$  extends to a continuous map  $D \rightarrow V$ .*

*Proof.* As pointed out by Pansu in [P, #37], one can prove this by using the monotonicity theorem for minimal surfaces. This states that there are constants  $c > 0$  and  $\varepsilon_0 > 0$  (which depend on  $V$  and the metric  $\mu_j$ ) such that for every minimal surface  $S$  in  $(V, \mu_j)$  which goes through the point  $x$ ,  $\mu_j$ -area  $(S \cap B_\varepsilon(x)) \geq c\varepsilon^2$  for all  $\varepsilon < \varepsilon_0$ . See [L, 3.15]. To apply this, recall from (2.2) that  $\text{Im } g$  is minimal and suppose that  $g(z)$  has two limit points  $p$  and  $q$  as  $z \rightarrow 0$ . If  $\delta$  is chosen to be less than  $d(p, q)/3$ , the monotonicity theorem implies that each connected component of  $g^{-1}(B_\delta(p))$  which meets  $g^{-1}(B_{\delta/2}(p))$  has area  $\geq c\delta^2/4$ . Therefore, because  $\text{Im } g$  is minimal and has finite area  $A(g)$  by (2.2), there can only be a finite number of such components. Similar remarks apply to  $q$ . Hence there exists an  $r_0 > 0$  such that, for any  $r < r_0$ , the image  $\gamma_r$  of the circle  $\{z : |z| = r\}$  under  $g$  meets both  $B_{\delta/2}(p)$  and  $B_{\delta/2}(q)$ , and so must have length  $\ell(\gamma_r) > \delta$ . But then, the conformality of  $g$

implies that  $|dg| = (1/r)|\partial g/\partial\theta|$ , and we find that

$$\begin{aligned} A(g) &= \int_{(0,1] \times S^1} \left( \frac{|\partial g/\partial\theta|^2}{r^2} \right) r dr \wedge d\theta \\ &\geq \int_{(0,1]} \left[ \int_{S^1} |\partial g/\partial\theta| d\theta \right]^2 \frac{1}{2\pi r} dr \\ &= \int_{(0,1]} \frac{\ell(\gamma_r)^2}{2\pi r} dr \geq \int_{(0,r_0]} \frac{\delta^2}{2\pi r} dr, \end{aligned}$$

which is impossible because  $A(g)$  is finite.  $\square$

*Note 2.4.8.* In fact, one can always extend  $g$  smoothly over the singular point. However, it is more difficult to prove this. By the elliptic regularity theorem (Theorem 2.4.1), it suffices to show that  $dg$  is bounded in the  $L^p$ -norm for some  $p > 2$ , which may be done by identifying  $D - \{0\}$  with the cylinder  $(-\infty, 0] \times S^1$ , and establishing some exponential decay estimates as in [FU, Appendix D]. (Compare [F1, Lemma 4.3.]). For another approach see [O2].

**Example 2.4.9.** Let us return to Example 2.3.5. If the weights  $\lambda_k$  are all positive integers with  $\lambda_1 = 1$ , then  $\omega(A) = \pi$  is the smallest positive value taken by  $[\omega]$  on  $\pi_2(V)$  and so  $\mathcal{M}_A(J)/G$  is compact by Theorem 2.4.2. If the weights are arbitrary, then  $\mathcal{M}_A(J)/G$  is not always compact. For example, if  $n = 2$  and  $\lambda_1 > \lambda_2$ , then the antidiagonal  $\{(z, \iota(z)) : z \in S^2\}$  (where  $\iota$  is the antipodal map) may be realized as a  $J_1$ -holomorphic sphere for some  $J_1$  which is compatible with  $\omega$ . (In fact, one can take  $J_1$  to be Hirzebruch’s complex structure on  $S^2 \times S^2$  which is obtained by identifying  $S^2 \times S^2$  with the projectivization of the rank 2 complex vector bundle over  $S^2 = CP^1$  which has first Chern class equal to 2: see [MD2, §3].) Since the antidiagonal represents the class  $A - B$ , where  $B = [pt \times S^2]$ , we may write  $A$  as a sum of two classes  $A - B$  and  $B$ , both of which may be represented by  $J_1$ -holomorphic spheres. Therefore, according to Theorem 2.4.2, it is possible that the moduli space  $\mathcal{M}_A(J_1)/G$  is not compact, and indeed one can check that it is not. However, when  $n = 2$ ,  $\mathcal{M}_A(J)/G$  is always compact for generic  $J$ : see [MD3, §4].

(2.5) **Applications.** Gromov thought of many interesting ways to use holomorphic spheres in symplectic geometry. For example, he showed that there is an exotic symplectic structure  $\rho$  on  $\mathbf{R}^{2n}$  which does not embed symplectically in  $\mathbf{R}^{2n}$  with its standard

structure. A simple and explicit formula for such  $\rho$  may be found in [BP]. Gromov also proved that the group of compactly supported symplectic diffeomorphisms of  $\mathbf{R}^4$  is contractible, and that (up to a multiple) there is only one symplectic form on the complex projective plane  $CP^2$  which admits a symplectically embedded 2-sphere in the class of  $CP^1$ . All these results are surveyed in [G3] and [Be]. More recently, Eliashberg developed these ideas further, using families of holomorphic discs (rather than 2-spheres) to prove, for example, that if  $W$  is a symplectic 4-manifold whose boundary  $\partial W$  is a 3-sphere of contact type, then the contact structure on  $\partial W$  must be the standard one: see [E3].

Here, we will content ourselves with proving Gromov's squeezing theorem (Theorem 1.2.1). We must show that it is impossible to embed the closed ball  $B^{2n}(r)$  symplectically inside the product  $B^2(R) \times \mathbf{R}^{2n-2}$  unless  $r \leq R$ . To this end, suppose that  $g : B^{2n}(r) \rightarrow B^2(R) \times \mathbf{R}^{2n-2}$  is a symplectic embedding, and choose any  $\varepsilon > 0$  and integer  $K > 0$  such that

$$\text{Im } g \subset \text{Int } B^2(R + \varepsilon) \times B^2(K(R + \varepsilon)) \times \dots \times B^2(K(R + \varepsilon)).$$

Then, by compactifying this product of 2-balls, we obtain a symplectic embedding  $g$  of  $B^{2n}(r)$  into  $(S^2 \times \dots \times S^2, \tau)$  where  $\tau$  is the symplectic form  $(R + \varepsilon)^2[\sigma_1 \oplus K^2\sigma_2 \oplus \dots \oplus K^2\sigma_n]$ .

**Lemma 2.5.1.** *The evaluation map*

$$e_A(J) : \mathcal{M}_A(J) \times_G S^2 \rightarrow S^2 \times \dots \times S^2 : (f, z) \mapsto f(z)$$

*is surjective for every  $\tau$ -compactible  $J$ .*

*Proof.* By Lemma 2.3.4, the usual complex structure  $J_0$  on  $S^2 \times \dots \times S^2$  is regular. Therefore, by Theorem 2.3.3 and the remarks in Example 2.4.9, the evaluation map  $e_A(J_0)$  is compactly bordant to  $e_A(J)$  for every regular  $J$ . It follows from Example 2.3.5 that the manifolds  $\mathcal{M}_A(J) \times_G S^2$  and  $S^2 \times \dots \times S^2$  have the same dimension and that the map  $e_A(J_0)$  has degree 1. Since degree is a cobordism invariant, the map  $e_A(J)$  has degree 1 whenever  $J$  is regular. Thus  $e_A(J)$  is surjective for all  $J$  in a dense subset of  $\mathcal{J}$ . Hence, by Note 2.4.6, it is surjective for all  $J$ .  $\square$

(2.5.2) *Proof of the Squeezing Theorem.* Let  $J_0$  be the usual complex structure on  $\mathbf{R}^{2n}$ . Because the set  $\mathcal{J}$  of  $\tau$ -compatible  $J$  forms a fiber bundle over  $V = S^2 \times \dots \times S^2$  with contractible fibers, it is possible to extend  $g_*(J_0)$  to an element  $J \in \mathcal{J}$ . By

Lemma 2.5.1, there is a  $J$ -holomorphic  $A$ -sphere through  $g(0)$ . Let  $S$  be the pull-back by  $g$  of the intersection of  $\text{Im } g$  with this sphere. Then  $S$  is a surface in  $B^{2n}(r)$  which goes through the origin and has boundary on  $\partial B^{2n}(r)$ . Moreover, it is minimal with respect to the euclidean metric  $\mu_0$  because it is  $J_0$ -holomorphic. Therefore, the monotonicity theorem mentioned in the proof of Theorem 2.4.7 implies that its  $\mu_0$ -area is at least  $\pi r^2$ . (When  $(V, \mu_j)$  is euclidean space, the constant  $c$  may be taken to be 1: see [L] or [V1, Appendix].) But by (2.2),

$$\text{area } S = \int_S \omega_0 = \int_{g(S)} \tau < \tau(A) = \pi(R + \varepsilon)^2,$$

and so  $r \leq R$  as claimed.  $\square$

### §3 FLOER HOMOLOGY

In this section, we first describe the variational setup used in the proof of Arnold’s conjectures and then outline Floer’s way of doing Morse theory on infinite-dimensional spaces. In (3.3) we construct the Floer complex and sketch the proof of the nondegenerate case of Arnold’s conjectures. (3.4) contains some details of the analysis, and (3.5) outlines some other applications of Floer homology. The article [S] provides another introduction to these ideas with more emphasis on the details of the Morse theory, and contains many useful references and technical details.

(3.1) **The variational setup.** Let  $\varphi$  be an exact diffeomorphism of the compact symplectic manifold  $(V, \omega)$ . By (1.3.3) this means that there is a time-dependent Hamiltonian  $H_t$  which generates a family  $\varphi_t$  of diffeomorphisms with  $\varphi_0 = \text{id}$  and  $\varphi_1 = \varphi$ . By reparametrizing  $H_t$  with respect to time, we may suppose that  $H_t$  depends smoothly on  $t \in S^1 = \mathbf{R}/\mathbf{Z}$ . If we look at the set of paths  $\{\varphi_t(x), 0 \leq t \leq 1 : x \in V\}$ , then the fixed points of  $\varphi$  correspond bijectively to the subset of these paths which close up to loops. However, there is no good way of distinguishing this subset from the full path space. Therefore, Conley and Zehnder in [CZ] looked instead at the space  $\Omega V$  of all smooth and contractible loops on  $V$  and then analyzed the subset consisting of all loops which have the form  $\varphi_t(x)$ ,  $0 \leq t \leq 1$ . We show below that these loops are precisely the critical points of the *action functional*  $a_H$ , which means that we can count them using Morse theory. Note that because we restrict to the component of contractible

loops on  $V$ , we are ignoring all fixed points which correspond to noncontractible loops.

A loop  $z : S^1 \rightarrow V$  has the form  $z(t) = \varphi_t(x)$  if and only if

$$(3.1.1) \quad \dot{z}(t) = \xi_t(z(t)), \quad 0 \leq t \leq 1,$$

where  $\xi_t$  is the Hamiltonian vector field of  $H_t$  and  $\dot{z}$  is the time derivative  $dz/dt$ . Suppose for the moment that  $\omega$  vanishes on  $\pi_2(V)$  and, for  $z \in \Omega V$ , define

$$(3.1.2) \quad a_H(z) = - \int_{D^2} \tilde{z}^*(\omega) - \int_{S^1} H_t(z(t)) dt$$

where  $\tilde{z} : D^2 \rightarrow V$  restricts to  $z$  on  $\partial D^2$ . Because  $\omega$  vanishes on  $\pi_2(V)$ , this is independent of the choice of  $\tilde{z}$ .

Recall that the tangent space to  $\Omega V$  at  $z$  consists of sections  $\zeta(t)$  of the bundle  $z^*(TV)$ . (We will be more careful later with questions of smoothness.) It is not hard to check that the value of the derivative  $da_H(z)$  in the direction of  $\zeta$  is given by the formula:

$$(3.1.3) \quad da_H(z)(\zeta) = \int_{S^1} \{ \omega(\dot{z}(t), \zeta(t)) - dH_t(z(t))\zeta(t) \} dt.$$

Because  $\omega(\dot{z}, \cdot) = i(\dot{z})\omega$  and  $dH_t = i(\xi_t)\omega$  by (1.3.3), this vanishes for all  $\zeta$  if and only if  $z(t)$  satisfies the equation (3.1.1). Thus the solutions of (3.1.1) are exactly the critical points of the function  $a_H$ . (Note that our signs differ from Floer's in [F5] since his metric  $g$  satisfies the identity  $g(\cdot, \cdot) = \omega(J\cdot, \cdot)$  rather than (2.1.1).)

If  $\omega$  does not vanish on  $\pi_2(V)$ , then we must pass to an appropriate covering of  $\Omega V$  in order to define  $a_H$ . However, its derivative  $da_H$  is defined for  $z \in \Omega V$  as before.

Let us now consider the gradient  $\mathcal{J}_{J,H}$  of  $a_H$  with respect to the inner product

$$\langle \xi, \zeta \rangle = \int_{S^1} \mu_J(\xi(t), \zeta(t)) dt,$$

on  $\Omega V$ , where  $\mu_J(J\cdot, \cdot) = \omega(\cdot, \cdot)$  as in (2.1). If  $\nabla H_t$  denotes the  $\mu_J$ -gradient of the function  $H_t$  on  $V$ , then (3.1.3) implies that  $\mathcal{J}_{J,H}(z)$  is given by the formula:

$$(3.1.4) \quad \mathcal{J}_{J,H}(z)(t) = J(z(t))\dot{z}(t) - \nabla H_t(z),$$

where  $J(z(t))$  is the automorphism of  $T_{z(t)}V$  induced by  $J$ . Hence if  $z_\tau \in \Omega V$  is a trajectory of the flow  $\partial z_\tau / \partial \tau = -\mathcal{J}_{J,H}(z_\tau)$ ,



the associated map  $u : \mathbf{R} \times S^1 \rightarrow V$  given by  $u(\tau, t) = z_\tau(t)$  satisfies the equation:

$$(3.1.5) \quad \partial u / \partial \tau + J(u) \partial u / \partial t - \nabla H_t(u) = 0.$$

Observe that the first-order terms in this equation make up the Cauchy-Riemann operator  $\bar{\partial}_J$  of (2.3.1).

(3.2) **Morse theory.** The usual way to study the critical points of a function  $\Psi$  on a finite-dimensional manifold  $X$  is to look at the change in topology of the sets  $X_\lambda = \Psi^{-1}((-\infty, \lambda])$  as  $\lambda$  passes through a critical value of  $\Psi$ : see [B]. Recall that if the critical point  $w$  is isolated and nondegenerate, then there is a well-defined nondegenerate quadratic form  $\text{Hess}(\Psi)$  on  $T_w X$  which is called the Hessian. In local coordinates, it is given by the matrix of second derivatives  $(\partial^2 \Psi / \partial x_i \partial x_j)$ . Further, if one defines the index  $\text{Int}(w)$  of  $w$  to be the dimension of the largest subspace on which the Hessian is negative, then one can show that  $X_{a+\varepsilon}$  is obtained from  $X_{a-\varepsilon}$  by adding a handle whose dimension equals the index of the relevant critical point.

This method can still be used when  $X$  is infinite dimensional provided that  $\Psi$  has a well-behaved gradient flow with respect to a suitable metric on  $X$ , and all the critical points of  $\Psi$  (or of  $-\Psi$ ) have finite index. This is the case for some important functionals, for example, the Yang–Mills functional on the quotient space of connections by the gauge group (see [AB]), but, by the results of [CZ], does not hold for  $a_H$  on the loop space  $\Omega V$ .

However, there is an alternative approach to Morse theory on finite-dimensional manifolds which was recently emphasized by Witten in [W1] and which does generalize to our situation. This arises by considering the trajectories of the negative  $-\nabla \Psi$  of the gradient vector field of  $\Psi$  with respect to a generic metric  $\mu$  on  $X$ . For, if  $\mu$  is generic, one can assume that for every pair of critical points  $x, y$ , the unstable manifold of  $x$  (i.e. set of trajectories going away from  $x$ ) intersects the stable manifold of  $y$  (i.e. set of trajectories going towards  $y$ ) transversally. It follows that if  $\text{Int}(x) - \text{Int}(y) = k$ , the manifold of trajectories of  $\Psi$  going from  $x$  to  $y$  has dimension  $k$  (if it is nonempty). Since a trajectory  $u : \tau \mapsto u(\tau)$  has a 1-dimensional family of reparametrizations  $u_\sigma : \tau \mapsto u(\tau + \sigma)$ , this means that there is a  $(k - 1)$ -dimensional manifold of unparametrized trajectories between  $x$  and  $y$ . In particular, if  $\text{Int}(x) - \text{Int}(y) = 1$ , there is a finite set of un-

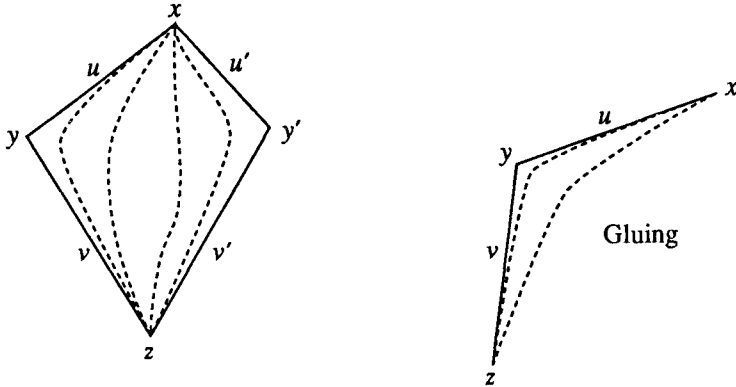


FIGURE 1

parametrized trajectories from  $x$  to  $y$ . We will call these isolated trajectories.

Now consider the space of unparametrized trajectories between critical points  $x$  and  $z$  with  $\text{Int}(x) - \text{Int}(z) = 2$ . Since  $\mu$  is generic, this space consists of a finite number of 1-dimensional components. If a component is noncompact, each of its two ends must converge (in an obvious sense) to a pair  $(u, v)$  of trajectories, where  $u$  is an isolated trajectory from  $x$  to a critical point  $y$ , and  $v$  is an isolated trajectory from  $y$  to  $z$ . (See Figure 1.) Conversely, given such a pair of isolated trajectories, one can “glue them together” at  $y$  and construct a 1-parameter family of trajectories from  $x$  to  $z$  which abuts on  $(u, v)$ .

Using this information, one defines a complex  $(F_*(X), \partial)$  as follows. For  $0 \leq k \leq \dim X$ , let  $F_k(X)$  be the free  $\mathbf{Z}/2\mathbf{Z}$ -module with one generator for each critical point  $x$  of index  $k$ , and define the boundary operator  $\partial : F_k(X) \rightarrow F_{k-1}(X)$  by  $\partial(x) = \sum \langle x, y \rangle y$ , where  $\langle x, y \rangle$  is the number (counted mod 2) of unparametrized trajectories from  $x$  to  $y$ . Using the above remarks about the 1-dimensional families of trajectories between points  $x$  and  $z$  with  $\text{Int}(x) - \text{Int}(z) = 2$ , one easily checks that  $\partial^2 = 0$ . (See Lemma 3.3.4 below.)

The chain groups  $F_k(X)$  are just those of the usual cell complex (over  $\mathbf{Z}/2\mathbf{Z}$ ) which one builds from the Morse function  $\Psi$ . It is not hard to see that the boundary maps are also the same. Hence, standard Morse theory implies that the homology of the complex  $(F_*(X), \partial)$  is isomorphic to  $H_*(X; \mathbf{Z}/2\mathbf{Z})$ . Observe, also, that it is possible to define  $(F_*(X), \partial)$  over the integers: one just has

to find a consistent way of assigning an orientation (i.e. + or - sign) to each isolated trajectory from  $x$  to  $y$ . For more details, see [W1], [F6], and [S].

The complex  $(F_*(X), \partial)$  involves only those trajectories of the gradient flow which go from one critical point to another. Floer realized that there is therefore no need to have a globally defined gradient flow: all one needs is to have well-behaved spaces of such trajectories. (He points out in [F5] that this is an extension of Conley's idea of an "isolating block".) In particular, in order to build a Morse theory for  $\alpha_H$  on  $\Omega V$ , one can use the  $L^2$ -gradient  $\mathcal{J}_{J,H}$  whose trajectories satisfy the elliptic equation (3.1.5). Floer calls the homology of the complex  $(F_*, \partial)$  so constructed the *Conley index*  $I(J, H)$ , but it is also known as the *Floer homology*. To compute it, one must establish some kind of homotopy invariance, and then make a calculation for a special choice of  $J$  and  $H$ .

*Note.* There is no way of completing  $\Omega V$  so that  $\mathcal{J}_{J,H}$  integrates to a globally defined flow: for example, if  $H$  is real analytic, all solutions of (3.1.5) are also real analytic so that there is a trajectory through  $z$  only if  $z$  is real analytic as well. However, when  $V = \mathbf{R}^{2n}$ , one can complete  $\Omega V$  to the Sobolev space  $H^{1/2}(S^1, \mathbf{R}^{2n})$  which consists of all functions whose "derivative of order 1/2" is  $L^2$ . (This is best understood in terms of the Fourier transform.) Again, this is a borderline case for Sobolev theory: the functions in  $H^{1/2}(S^1, \mathbf{R}^{2n})$  are not continuous, and one cannot substitute an arbitrary manifold  $V$  for  $\mathbf{R}^{2n}$ . On the other hand, one can check that the gradient of  $\alpha_H$  with respect to the Hilbert inner product on  $H^{1/2}(S^1, \mathbf{R}^{2n})$  integrates to give a flow which is everywhere defined. This is the flow considered in [CZ], [V1], [V2], [EH], etc. It has the advantage of being globally defined, but its trajectories no longer satisfy an elliptic equation such as (3.1.5).

(3.3) **Floer's proof of the Arnold conjectures.** In this section we describe Floer's complex  $(F_*, \partial)$ , and outline his proof of the following result.

**Theorem 3.3.1.** *Exact nondegenerate diffeomorphisms on compact monotone manifolds  $(V, \omega)$  satisfy the Morse inequalities over  $\mathbf{Z}/2\mathbf{Z}$ .*

We will assume throughout that  $(V, \omega)$  is monotone, i.e. that there is a constant  $k > 0$  such that the cohomology class  $[\omega] - k.c_1(V)$  is zero on all 2-spheres. This hypothesis is crucial to the

proof of compactness: see Proposition 3.4.11.

Given critical points  $x$  and  $y$  for the action  $a_H$  on  $\Omega V$ , and an  $\omega$ -compatible  $J$ , we define the space  $\mathcal{M}(x, y)$  of trajectories from  $x$  and  $y$  to be the space of all solutions  $u : \mathbf{R} \times S^1 \rightarrow V$  of the equation

$$(3.3.2) \quad \bar{\partial}_{J,H}(u) = \partial u / \partial \tau + J(u) \partial u / \partial t - \nabla H_t(u) = 0$$

such that  $\lim_{\tau \rightarrow -\infty} u(\tau, t) = x(t)$  and  $\lim_{\tau \rightarrow \infty} u(\tau, t) = y(t)$ . By the remarks in (3.1) these are precisely the trajectories of the gradient flow  $\partial z_\tau / \partial \tau = -\varphi_{J,H}(z_\tau)$ , which go from  $x$  to  $y$ . (For our notation to be consistent with that of §2, we should really mention  $J$  and  $H$  in the name  $\mathcal{M}(x, y)$ . However, following Floer, we will not do this.) The main results needed about the  $\mathcal{M}(x, y)$  are contained in the following proposition which is a slightly simplified version of [F5, Proposition 1b]. Its proof is discussed in (3.4). As before, we write  $\mathcal{F}$  for the Fréchet space of all  $C^\infty$   $\omega$ -compatible almost complex structures on  $V$ , and  $\mathcal{H}$  for the Fréchet space  $C^\infty(V \times S^1, \mathbf{R})$  of Hamiltonians.

**Theorem 3.3.3.** (i) *There is a dense set  $(\mathcal{F} \times \mathcal{H})_{\text{reg}}$  in  $\mathcal{F} \times \mathcal{H}$  such that for every element  $(J, H) \in (\mathcal{F} \times \mathcal{H})_{\text{reg}}$  the action functional  $a_H$  has a finite set  $Z$  of critical points, and the trajectory spaces  $\mathcal{M}(x, y)$ ,  $x, y \in Z$ , are smooth manifolds. Further, if  $\Gamma$  denotes the subgroup of  $\mathbf{Z}$  generated by the values taken by the first Chern class of  $(V, J)$  on the elements of  $\pi_2(V)$ , there is a  $\text{mod } 2\Gamma$  index function  $\text{Int} : Z \rightarrow \mathbf{Z}/2\Gamma$  such that  $\dim \mathcal{M}(x, y) \equiv \text{Int}(x) - \text{Int}(y) \pmod{2\Gamma}$ .*

(ii) *Let  $\widehat{\mathcal{M}}(x, y) = \mathcal{M}(x, y) / \mathbf{R}$  denote the quotient by the translational symmetry. Then, for any triple  $x_1, x_2, x_3 \in Z$ , there is a local diffeomorphism  $\sharp$  from an open subset  $\mathcal{O}$  of  $\widehat{\mathcal{M}}(x_1, x_2) \times \widehat{\mathcal{M}}(x_2, x_3) \times \mathbf{R}$  into  $\widehat{\mathcal{M}}(x_1, x_3)$  such that:*

- (T1) *for each compact set  $K$  in  $\widehat{\mathcal{M}}(x_1, x_2) \times \widehat{\mathcal{M}}(x_2, x_3)$ , there is a number  $\rho(K) > 0$  such that  $K \times [\rho(K), \infty) \subset \mathcal{O}$ ;*
- (T2) *for  $i = 1, 2$  there is a lifting  $\sharp_i : \mathcal{O} \rightarrow \mathcal{M}(x_1, x_3)$  such that for each  $(u_1, u_2) \in \mathcal{M}(x_1, x_2) \times \mathcal{M}(x_2, x_3)$ , the family of maps  $\sharp_i(u_1, u_2, \rho)$  converges to  $u_i$  in the local  $C^\infty$ -topology as  $\rho \rightarrow \infty$ ;*
- (T3) *the 0- and 1-dimensional part of  $\widehat{\mathcal{M}} = \bigcup_{x,y \in Z} \widehat{\mathcal{M}}(x, y)$  is compact up to the image of  $\sharp$ .*

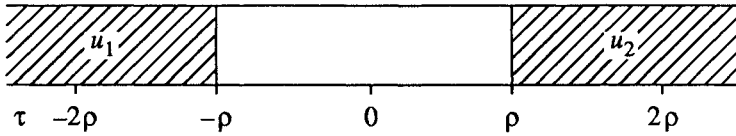


FIGURE 2

Intuitively speaking, the map  $\#$  glues a trajectory  $u_1$  from  $x_1$  to  $x_2$  to a trajectory  $u_2$  from  $x_2$  to  $x_3$  to make one trajectory  $u$  from  $x_1$  to  $x_3$ , with a gluing parameter  $\rho$  which measures how far apart one places  $u_1$  and  $u_2$ . More precisely, consider the following diagram, in which we illustrate an approximate trajectory which we will call  $u_1\#u_2\#\rho$ . Here

$$\begin{aligned} u_1\#u_2\#\rho(\tau, t) &= u_1(\tau + 2\rho, t) \quad \text{for } \tau \leq -\rho, \\ &= u_2(\tau - 2\rho, t) \quad \text{for } \tau \geq \rho, \end{aligned}$$

and is very small elsewhere (Figure 2). Using a technique developed by Taubes, one proves that, when  $\rho$  is sufficiently large, there is a deformation retraction  $\mathbf{D}$  from approximate trajectories such as  $u_1\#u_2\#\rho$  into the space of actual trajectories. Hence the maps  $\#_i$  of (T2) are given by:

$$\#_1(u_1, u_2, \rho)(\tau, t) = \mathbf{D}(u_1\#u_2\#\rho)(\tau + 2\rho, t),$$

and

$$\#_2(u_1, u_2, \rho)(\tau, t) = \mathbf{D}(u_1\#u_2\#\rho)(\tau - 2\rho, t).$$

(Remember that a sequence converges in the local  $C^\infty$ -topology if its restriction to each truncated cylinder  $[-R, R] \times S^1$  converges in the  $C^\infty$ -topology.)

It is clear from this construction that, if  $\{\rho_N\}$  is any sequence of real numbers, the only possible limits of subsequences of  $\{\mathbf{D}(u_1\#u_2\#\rho_N)(\cdot + \rho_N)\}$  are reparametrizations of  $u_1$  and  $u_2$ , or are  $\tau$ -constant trajectories at  $x_1$ ,  $x_2$ , and  $x_3$ . Under these circumstances, we say that the sequence  $\{\mathbf{D}(u_1\#u_2\#\rho_N)\}$  converges weakly to the pair  $u_1, u_2$ . See [F5, 3a].

Next observe that, because  $\#$  is a local diffeomorphism, the image of  $\#$  lies in a component of  $\widehat{\mathcal{M}}(x_1, x_3)$  which has dimension  $1 + \dim \widehat{\mathcal{M}}(x_2, x_2) + \dim \widehat{\mathcal{M}}(x_2, x_3)$ . In particular, if  $u_1$  and  $u_2$  are isolated, they glue together to form part of a 1-dimensional component of  $\widehat{\mathcal{M}}$ . By (T2), this part must in fact be a neighborhood of one of the ends of this component. (T3) states that this is the only noncompactness found in the 0- and 1-dimensional

components of  $\widehat{\mathcal{M}}$ . Thus there are only finitely many isolated trajectories and 1-dimensional components of  $\widehat{\mathcal{M}}$ . Moreover, any 1-dimensional component of  $\widehat{\mathcal{M}}(x, y)$  either is a circle or is an open interval, each end of which lies in the image of  $\sharp$  and so converges weakly to a pair of isolated trajectories. Note, finally, that because there is a  $\text{mod } 2\Gamma$  index, there are isolated trajectories between  $x$  and  $y$  only if  $\text{Int } x + \text{Int } y \equiv 1 \pmod{2\Gamma}$ , and there are 1-parameter components in  $\widehat{\mathcal{M}}(x, y)$  only if  $\text{Int } x + \text{Int } y \equiv 0 \pmod{2\Gamma}$ .

Now consider the Floer complex  $(F_*, \partial)$  built from these data. Its vertices are the elements  $x$  of  $Z$  and we define the boundary  $\partial$  by  $\partial(x) = \sum \langle x, y \rangle y$ , where  $\langle x, y \rangle$  is the number of isolated trajectories from  $x$  to  $y$ , counted  $\text{mod } 2$ .

**Lemma 3.3.4.**  $\partial^2 = 0$ .

*Proof.* If  $\partial^2 x \neq 0$ , there is, for some  $z \in Z$ , an odd number of pairs of isolated trajectories  $u, v$  where  $u$  goes from  $x$  to some  $y$  and  $v$  goes from  $y$  to  $z$ . But, by (T1), for each pair we may construct an arc  $\alpha(u, v)$  which lies in a 1-dimensional component of  $\widehat{\mathcal{M}}$  and has one end which converges weakly to the pair  $u, v$ . By the remarks above, the other end of the component containing  $\alpha(u, v)$  must converge to another pair of trajectories between  $x$  and  $z$ . Therefore, the number of such pairs is even.  $\square$

This completes the construction of the Floer complex  $(F_*, \partial)$ . *Invariance of  $(F_*, \partial)$  under “continuations.”* We now investigate the dependence of  $(F_*, \partial)$  on the choice of  $(J, H)$ . Because  $(F_*, \partial)$  is finite, one can easily show that it changes by an isomorphism if  $(J, H)$  is slightly perturbed. In Floer’s first version of his theory in [F3] (which applied to the Lagrangian intersection problem), he dealt with large perturbations by considering 1-parameter families of  $(J, H)$ , and investigating in detail what happens when one passes through a generic codimension 1 singularity. However, he realized that this argument was unnecessarily complicated and in [F5] used the elegant argument which we now present.

The appropriate notion of homotopy is called by Floer a “continuation”. In our context (which is slightly simpler than Floer’s), this is just a smooth family  $\{J_\lambda, H_\lambda\}_{\lambda \in \mathbb{R}}$  of pairs, where  $J_\lambda$  is an almost complex structure which is compatible with  $\omega$  and  $H_{t, \lambda}$  is a time-dependent Hamiltonian. We assume further that these are independent of  $\lambda$  when  $\lambda$  is outside some interval  $[-R, R]$ .

Consider the equation

$$(3.3.5) \quad \bar{\partial}_{J,H,\lambda}(u) = \partial u / \partial \tau + J_\tau(u) \partial u / \partial t - \nabla H_{t,\tau}(u) = 0.$$

This is almost identical to equation (3.3.2) except that the coefficients  $J$  and  $H$  depend on  $\tau$ . It is not hard to check that the analysis goes through as before, except that the solution set  $\mathcal{M}_\lambda$  is no longer translationally invariant. In particular, if we write  $x, y, \mathcal{M}$ , etc. for objects related to the pair  $(J, H) = (J_{-R}, H_{-R})$  and  $x', y', \mathcal{M}'$  for objects related to the pair  $(J', H') = (J_R, H_R)$ , it follows from Lemma 3.4.2 that every solution  $u$  of (3.3.5) on which the action  $a_H$  is finite converges to some element  $x \in Z$  as  $\tau \rightarrow -\infty$  and to some element  $x' \in Z'$  as  $\tau \rightarrow +\infty$ . Further, if  $(u_1, u_2)$  belongs to  $\mathcal{M}_\lambda(x, x') \times \mathcal{M}'(x', y')$  or to  $\mathcal{M}(x, y) \times \mathcal{M}_\lambda(y, y')$ , one can construct approximate trajectories  $u_1 \# u_2 \# \rho$  as above. Using this, it follows as before that, if the continuation  $\{J_\lambda, H_\lambda\}$  is regular (i.e. generic), there are gluing maps from open subsets in  $\mathcal{M}_\lambda(x, x') \times \widehat{\mathcal{M}}(x', y') \times \mathbf{R}$  and  $\widehat{\mathcal{M}}(x, y) \times \mathcal{M}_\lambda(y, y') \times \mathbf{R}$  to  $\mathcal{M}_\lambda(x, y')$ . Moreover, the proof of Proposition 3.4.11 shows that any noncompact end of a 1-dimensional component in  $\mathcal{M}_\lambda$  converges weakly to either the composite of an isolated trajectory in  $\widehat{\mathcal{M}}$  with one in  $\mathcal{M}_\lambda$  or the composite of an isolated trajectory in  $\mathcal{M}_\lambda$  with one in  $\widehat{\mathcal{M}}$ . Hence these gluing maps account for all the noncompactness in the 0- and 1-dimensional part of  $\mathcal{M}_\lambda$ . In particular, there are only finitely many isolated trajectories and arcs in  $\mathcal{M}_\lambda$ .

Thus, a regular continuation  $\{J_\lambda, H_\lambda\}$  defines a map  $h : F_* \rightarrow F'_*$  by

$$(*) \quad h(x) = \sum \langle x, x' \rangle x',$$

where  $\langle x, x' \rangle$  is the number of isolated trajectories in  $\mathcal{M}_\lambda$  from  $x$  to  $x'$ , counted mod 2.

**Proposition 3.3.6.** (i)  $h$  is a chain map, i.e.  $h\partial = \partial'h$ .

(ii)  $h$  induces an isomorphism on homology.

*Proof.* (i) is proved by the arguments of Lemma 3.3.4 and the above remarks. To prove (ii), note first that one can compose continuations. For if  $\{J_\lambda, H_\lambda\}$  is a continuation from  $(J, H)$  to  $(J', H')$ , and if  $\{J'_\lambda, H'_\lambda\}$  is a continuation from  $(J', H')$  to

$(J'', H'')$ , we may define  $\{J_\lambda^\rho, H_\lambda^\rho\}$  for large  $\rho$  by setting

$$\begin{aligned} J_\lambda^\rho &= J_{\lambda+2\rho} & \text{for } \lambda < 0, \\ &= J' & \text{for } \lambda = 0, \quad \text{and} \\ &= J_{\lambda-2\rho} & \text{for } \lambda > 0, \end{aligned}$$

and similarly for  $H_\lambda^\rho$ . (Here we assume that  $\rho$  is so large that both  $\{J_\lambda, H_\lambda\}$  and  $\{J'_\lambda, H'_\lambda\}$  are constant for  $|\lambda| \geq \rho$ .) A slight perturbation of  $\{J_\lambda^\rho, H_\lambda^\rho\}$  will be regular, and will define a chain map from  $(F_*, \partial)$  to  $(F''_*, \partial'')$ . (As remarked above,  $(F_*, \partial)$  does not change under small perturbations of  $(J, H)$ , and so it does not matter if we move the ends of  $\{J_\lambda^\rho, H_\lambda^\rho\}$  a little bit.) That this map (for large  $\rho$ ) is the composite of  $h$  with  $h'$  follows from the fact that the  $\{J_\lambda^\rho, H_\lambda^\rho\}$ -trajectories are in 1-to-1 correspondence with pairs  $u, u'$  of trajectories for  $\{J_\lambda, H_\lambda\}$  and  $\{J'_\lambda, H'_\lambda\}$ . (This may be proved by the same methods which establish statements (T1) and (T2) of Theorem 3.3.3.)

We next claim that the chain homotopy class of  $h$  depends only on the homotopy class of  $\{J_\lambda, H_\lambda\}$  relative to its ends. To see this, suppose that  $\{J_\lambda(\nu), H_\lambda(\nu)\}$ ,  $0 \leq \nu \leq 1$ , is a family of continuations with fixed ends, and consider the set

$$\begin{aligned} \overline{\mathcal{M}}(x, x') &= \{(\nu, u) : \nu \in [0, 1], u \text{ satisfies} \\ &\text{equation (3.3.5) for } \{J_\lambda(\nu), H_\lambda(\nu)\}\}. \end{aligned}$$

The extra parameter  $\nu$  can be easily incorporated into the analytic setup of (3.4), and one can show that  $\overline{\mathcal{M}}(x, x')$  is the zero set of a Fredholm section of a Banach bundle over the space  $[0, 1] \times P_{1,\rho}(x, x')$ , and that it is a manifold for generic families  $\{J_\lambda(\nu), H_\lambda(\nu)\}$ .

Define a map  $S : F_* \rightarrow F'_*$  by the formula (\*), this time using the isolated trajectories in  $\overline{\mathcal{M}}(x, x')$ . We can think of this map as having degree  $-1 \pmod{\Gamma}$ , because the extra parameter  $\nu$  makes the relative index  $\mu(u, \nu)$  of a pair of critical points  $x \in Z$  and  $x' \in Z'$  with respect to the family  $\{J_\lambda(\nu), H_\lambda(\nu)\}$ ,  $0 \leq \nu \leq 1$ , one more than their relative index  $\mu(u)$  (see (3.4)) with respect to any individual continuation  $\{J_\lambda(\nu), H_\lambda(\nu)\}$  in this family. Note also that if  $(u, \nu)$  is an isolated element of  $\overline{\mathcal{M}}(x, x')$ , the continuation  $\{J_\lambda(\nu), H_\lambda(\nu)\}$  cannot be regular, since if it were, it would admit no trajectories  $u$  of index  $\mu(u) = -1$ . In particular,  $\nu$  cannot equal 0 or 1.



We claim that  $S$  is a chain homotopy between  $h_0$  and  $h_1$ , i.e.  $h_1 - h_0 = S\delta - \delta S$ . This follows because of the structure of the 1-dimensional components of  $\overline{\mathcal{M}}(x, x')$ . A typical component will lie over a path  $\alpha$  in  $[0, 1]$  with endpoints  $\alpha(0), \alpha(1)$ . If either  $\alpha(0)$  or  $\alpha(1)$  equals 0 or 1, the component must end at a  $\{J_\lambda(\nu), H_\lambda(\nu)\}$ -trajectory from  $x$  to  $x'$ , where  $\nu = 0$  or 1, as appropriate. Otherwise, the component abuts in a pair of trajectories which are either of the form  $(S, \delta) \in \mathcal{M}_\lambda((\alpha(i), x, z') \times \overline{\mathcal{M}}(z', x')$  or of the form  $(\delta, S) \in \overline{\mathcal{M}}(x, y) \times \mathcal{M}_\lambda(\alpha(i), y, y')$ . It follows that there are only finitely many 1-dimensional components, and that they are one of the three kinds shown in Figure 3.

By using the fact that any pair  $(S, \delta)$  or  $(\delta, S)$  may be glued together to form one end of a 1-dimensional component in  $\overline{\mathcal{M}}(x, x')$ , it is now not hard to see that  $s$  is indeed a chain homotopy.

Since the composition of  $\{J_\lambda, H_\lambda\}$  with its “inverse”  $\{J_{-\lambda}, H_{-\lambda}\}$  is homotopic to the constant map  $\{J, H\}$ , the chain map  $h$  must induce an isomorphism on homology.  $\square$

By the remarks in Definition 1.3.5, Theorem 3.3.1 will follow if we prove:

**Proposition 3.3.7.** *The homology of the complex  $(F_*, \partial)$  is isomorphic to  $H_*(V; \mathbb{Z}/2\mathbb{Z})$ .*

*Proof.* By Proposition 3.3.6 it suffices to calculate  $(F_*, \partial)$  for any generic  $(J, H)$ . Because  $(V, \omega)$  is monotone, we may apply Proposition 3.4.10 below and hence assume that  $H$  is time independent and so small that the only critical points of  $a_H$  on  $\Omega V$  are the constant loops. Then equation (3.3.2) has a whole class

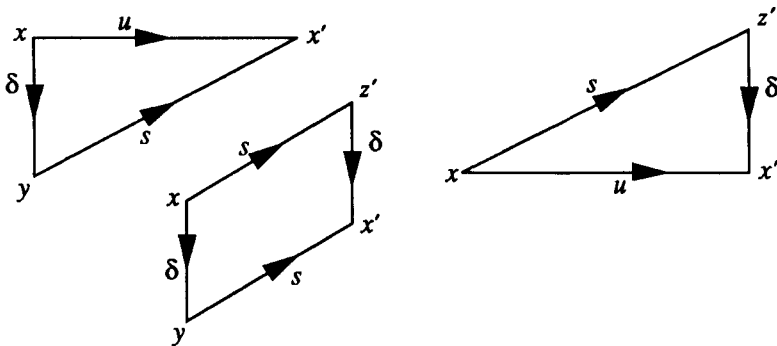


FIGURE 3

of solutions which are also independent of time, and so consist of trajectories of the  $\mu_J$ -gradient flow of  $H$  on  $V$ . Since this flow is generic in the sense of (3.2), the remarks in (3.2) imply that the Floer complex made out of these trajectories has homology equal to  $H_*(V; \mathbb{Z}/2\mathbb{Z})$ . The desired result now follows because, by Proposition 3.4.10, no other elements of  $\mathcal{M}(x, y)$  contribute to  $(F_*, \partial)$ .  $\square$

**(3.4) Some details of the analysis.** The proof of Theorem 3.3.3 is enormously complicated, and it is impossible to deal with all its complexities here. At its basis lies the analysis of elliptic operators on the cylinder  $\mathbb{R} \times S^1$ , and a gluing procedure first developed by Taubes. I have chosen to say rather little about these topics, since Floer’s discussion of them is quite adequate, and instead have tried to spell out some of the more superficial details involved in putting everything together.

*The trajectory spaces  $\mathcal{M}(x, y)$ .* We begin with a useful result concerning solutions of the equation

$$(3.3.2) \quad \bar{\partial}_{J,H}(u) = \partial u / \partial \tau + J \partial u / \partial t - \nabla H_t(u) = 0.$$

For the time being, we will work in the space  $W_{1,p,\text{loc}}(\mathbb{R} \times S^1, V)$ , where  $p > 2$  as in (2.3), and where “loc” means that a neighborhood basis is given by the  $W_{1,p}$ -norms of the restrictions of the maps  $u$  to the compact subsets  $[-\rho, \rho] \times S^1$  of  $\mathbb{R} \times S^1$ . The length  $\ell(u)$  of  $u \in W_{1,p,\text{loc}}(\mathbb{R} \times S^1, V)$  is defined to be:

$$(3.4.1) \quad \ell^2(u) = \int_{\mathbb{R} \times S^1} \left| \frac{\partial u}{\partial \tau} \right|^2 d\tau \wedge dt.$$

Further, given critical points  $x$  and  $y$  for the action  $a_H$  on  $\Omega V$ , we define the trajectory space  $\mathcal{M}(x, y)$  to be the subspace of  $W_{1,p,\text{loc}}(\mathbb{R} \times S^1, V)$  consisting of all solutions to (3.3.2) such that  $\lim_{\tau \rightarrow -\infty} u(\tau, t) = x(t)$  and  $\lim_{\tau \rightarrow \infty} u(\tau, t) = y(t)$ . Note that solutions of (3.3.2) are  $C^\infty$ -smooth by elliptic regularity theory.

**Lemma 3.4.2**(c.f. [F1, Theorem 2]). *Let  $u \in W_{1,p,\text{loc}}(\mathbb{R} \times S^1, V)$  be a solution of  $\bar{\partial}_{J,H}u = 0$ . Then  $u$  belongs to a trajectory space  $\mathcal{M}(x, y)$  if and only if it has finite length.*

*Proof.* If  $\bar{\partial}_{J,H}u = 0$ , then  $\partial u/\partial \tau = -\varphi_{J,H}(u(\tau))$  by (3.1) and

$$\begin{aligned} \frac{\partial a_H(u(\tau))}{\partial \tau} &= da_H(u(\tau)) \left( \frac{\partial u}{\partial \tau} \right) \\ &= -\langle \varphi_{J,H}(u(\tau)), \varphi_{J,H}(u(\tau)) \rangle \\ &= -\int_{S^1} \left| \frac{\partial u(\tau)}{\partial \tau} \right|^2 dt. \end{aligned}$$

Hence,

$$\ell^2(u) = -\int_R \frac{\partial}{\partial \tau} a_H(u(\tau)) d\tau.$$

Thus, if  $u \in \mathcal{M}(x, y)$ ,  $\ell^2(u)$  is the change in  $a_H(u(\tau))$ , as  $\tau$  goes from  $-\infty$  to  $+\infty$  and so is finite.

Conversely, let  $u \in W_{1,p,\text{loc}}$  be a solution of  $\bar{\partial}_{J,H}u = 0$  of finite length, and write  $u_R$  for its restriction to  $[-R, R] \times S^1$ . Then, by (3.1.3),

$$\ell^2(u_R) = A(u_R) + \int_{S^1} [H(u(R, t)) - H(u(-R, t))] dt,$$

where the area  $A(u_R)$  is defined in (2.2). Since  $\ell(u)$  is finite and  $H$  is bounded, it follows easily that  $A(u)$  is finite. Let  $v_\alpha$  be the restriction of  $u(\cdot + \rho_\alpha)$  to  $[-2, 2] \times S^1$ , where  $\rho_\alpha$  is any increasing sequence in  $\mathbf{R}$  which tends to  $\infty$ . Then, the sequence  $\{v_\alpha\}$  converges to 0 in the  $W_{1,2}$ -norm, since  $A(u)$  is finite. If it is not uniformly bounded in the  $W_{1,p}$ -norm for some  $p > 0$ , we will see below that a sphere bubbles off. Since this sphere must have positive area, this contradicts the fact that  $A(v_\alpha) \rightarrow 0$ . Hence  $\{v_\alpha\}$  is uniformly bounded in the  $W_{1,p}$ -norm and so, as in Theorem 2.4.1, has a subsequence whose restriction to  $[-1, 1] \times S^1$  converges in the  $W_{1,p}$ -norm. Clearly, the limit  $v$  has zero length, and so must be independent of  $\tau$ . Since  $\bar{\partial}_{J,H}v = 0$ , this means that  $v(\tau, t) = x(t)$  where  $x$  is a critical point of  $a_H$ . Further, because this holds for every reparametrization sequence  $\rho_\alpha$ , it is easy to see that the elements  $u(\tau)$  of  $\Omega V$  must converge to  $x$  in the  $C^0$ -topology on  $\Omega V$  as  $\tau \rightarrow -\infty$ . A similar argument shows that  $u(\tau)$  converges as  $\tau \rightarrow +\infty$ . Hence  $u \in \mathcal{M}(x, y)$ .

We will prove the statement about the sphere bubbling off by an argument due to Hofer. Observe first that if we had sequences  $z_\alpha \in [-2, 2] \times S^1$  and  $\varepsilon_\alpha \in (0, 1]$  such that

$$(3.4.3)(i) \quad r_\alpha = \varepsilon_\alpha \cdot |dv_\alpha(z_\alpha)| \rightarrow \infty,$$

and

$$(3.4.3)(ii) \quad |dv_\alpha(z)| \leq 2|dv_\alpha(z_\alpha)| \quad \text{for all } z \in B(z_\alpha, \varepsilon_\alpha),$$

then the argument of Theorem 2.4.2 would go through. For, if  $g_\alpha$  were the rescaling of  $v_\alpha$  defined as in (2.4.3), but with a map  $\psi_\alpha$  of  $B(0, r_\alpha)$  onto  $B(z_\alpha, \varepsilon_\alpha)$  such that  $|d\psi_\alpha(z)| \leq 1/|dv_\alpha(z_\alpha)|$ , then  $\sup |dg_\alpha(z)| \leq 2|dg_\alpha(0)| = 2$  by (3.4.3)(ii). Therefore, just as before, a subsequence of the  $g_\alpha$  would converge to a map  $g : \mathbf{C} \rightarrow V$ . Note that the limit  $g$  would be  $J$ -holomorphic, since, in the rescaling, the contribution from  $H$  in the equation  $\bar{\partial}_{J,H}u = 0$  becomes vanishingly small. Hence  $g$  would extend to a continuous map  $S^2 \rightarrow V$  as claimed.

To find suitable  $z_\alpha$  and  $\varepsilon_\alpha$ , we start from any sequences which satisfy (3.4.3)(i). Suppose, inductively, that (3.4.3)(ii) holds for  $\alpha < k$  and consider the  $k$ th term. We will find a point  $z'_k \in B(z_k, 2\varepsilon_k)$  and  $\varepsilon'_k > 0$  such that (3.4.3)(ii) holds and  $\varepsilon'_k \cdot |dv_k(z'_k)| \geq \varepsilon_k \cdot |dv_k(z_k)|$ . Start with  $(w_1, \delta_1) = (z_k, \varepsilon_k)$ . If this does not satisfy (3.4.3)(ii), choose  $w_2 \in B(w_1, \delta_1)$  so that  $|dv_k(w_2)| > 2|dv_k(w_1)|$  and put  $\delta_2 = \delta_1/2$ . Then  $(w_2, \delta_2)$  clearly satisfies (3.4.3)(i). If it does not satisfy (3.4.3)(ii), then repeat the above construction using  $(w_2, \delta_2)$  instead of  $(w_1, \delta_1)$  to get a new point  $(w_3, \delta_3)$  with  $\delta_3 = \delta_2/2$ , and so on. Either this process stops after  $N$  steps and one can take  $(z'_k, \varepsilon'_k) = (w_N, \delta_N)$ , or one obtains a sequence  $w_1, w_2, \dots$  which converges to some point  $w$  in  $B(z_k, 2\varepsilon_k)$ . But then  $v_k$  cannot be differentiable at  $w$ , which contradicts the regularity theorem (Theorem 2.4.1). Hence the process must stop, which completes the proof of the claim.  $\square$

*Note.* Another way to prove that the  $v_\alpha$  are uniformly bounded in the  $W_{1,p}$ -norm is to prove directly that if  $u$  is any solution of  $\bar{\partial}_{J,H}u = 0$  with finite length then  $|du(z)|$  is bounded over  $\mathbf{R} \times S^1$ . This follows from an *a priori* estimate for the derivative  $|df(0)|$  of a solution  $f : B(0, 1) \rightarrow V$  of  $\bar{\partial}_{J,H}f = 0$  in terms of the area of its image, which holds provided that this area is sufficiently small. (Here we identify the disc  $B(0, 1)$  with the corresponding subset in  $\mathbf{R} \times S^1 = \mathbf{C}/\mathbf{Z}$ .) Gromov (see [P, # 8, # 12]) and Wolfson [Wo, Theorem 4.1] prove such an estimate when  $H = 0$ . For the general case, see Salamon [S, Lemma 5.1]. (In fact, the general case can be reduced to the case  $H = 0$  by Gromov's trick in [G1, 1.4.C'].)

*Fredholm theory.* For each pair  $x, y$  of critical points of  $a_H$ , Floer defines a subspace  $P_{1,p}(x, y)$  of  $W_{1,p,\text{loc}}(\mathbf{R} \times S^1, V)$  which consists of functions which decay exponentially to  $x$  and  $y$  at the ends of the cylinder  $\mathbf{R} \times S^1$ . He then shows that, provided that  $x$  and  $y$  are nondegenerate (i.e. that the Hessians of  $a_H$  at  $x$  and  $y$  have no kernel), one can set up a Fredholm theory for the operator  $\bar{\partial}_{J,H}$  on  $P_{1,p}(x, y)$  much as in (2.3). Thus, he defines a Banach bundle  $\bar{\mathcal{L}}$  over  $P_{1,p}(x, y)$  and a Fredholm section  $\bar{\partial}_{J,H}$  of  $\bar{\mathcal{L}}$  whose zeros are solutions of the equation (3.3.2). (The details, which are rather complicated, may be found in [F5, 2a, b]. The Hessians mentioned here are defined in (3.4.4) below, where they are called  $A_x$  and  $A_y$ .) Further, the argument of [F1, Theorem 4a] shows that each element of  $\mathcal{M}(x, y)$  does have exponential decay at its ends, and so belongs to  $P_{1,p}(x, y)$ . Hence  $\mathcal{M}(x, y)$  is precisely the zero set of  $\bar{\partial}_{J,H}$ .

*The index  $\mu(u)$ .* In the finite-dimensional case, all the components of  $\mathcal{M}(x, y)$  have the same dimension  $\text{Int}(x) - \text{Int}(y)$ . In our case, the critical points  $x$  of  $a_H$  on  $\Omega V$  do not have a naturally defined integer-valued index. However, the difference  $\mu(u) = \text{Int}(x) - \text{Int}(y)$  can be measured along each trajectory  $u \in \mathcal{M}(x, y)$ . For, according to (3.2),  $\mu(u)$  should be a measure of the difference in dimension between the maximal negative subspaces of the Hessians of  $a_H$  at  $x$  and  $y$ . Suppose that, for each  $u$ , we can define a family of self-adjoint operators  $A_\tau$ , where  $\tau \in [0, 1]$ , on a smoothly varying family  $L_\tau$  of Hilbert spaces, such that  $A_0$  is the Hessian at  $x$  and  $A_1$  is the Hessian at  $y$ . Suppose further that all the eigenvalues of  $A_\tau$  have multiplicity 1. Then there are smooth families  $a(\tau)$  of eigenvalues of  $A_\tau$ , and we can measure  $\mu(u)$  by the spectral flow of the  $A_\tau$ , that is by counting the number of eigenvalues of the  $A_\tau$  which cross 0 as  $\tau$  goes from 0 to 1. More precisely,

$$\mu(u) = \#\{a : a(0) < 0 < a(1)\} - \#\{a : a(0) > 0 > a(1)\}.$$

To define these operators, let  $L_z$  be the tangent space  $L^2(z^*TV)$  to  $\Omega V$  at  $z$ , and define  $A_z$  on a dense subspace of  $L_z$  by setting it equal to the covariant derivative of the gradient vector field  $\mathcal{J}_{J,H}$ . If the domain of  $A_z$  is extended to the Sobolev space  $W_{1,2}(z^*TV)$ , it is not hard to check that  $A_z$  is self-adjoint. Further, if  $z$  is a critical point of  $a_H$ ,  $A_z$  is the Hessian of  $a_H$  at

$z$ , and one can check that it is given by the formula:

(3.4.4)

$$A_z(\xi) = D_\xi(J\dot{z}) - D_\xi(\nabla H_t(z)) = J(D_t\xi) + (D_\xi J)\dot{z} - D_\xi(\nabla H_t(z)),$$

where  $D$  denotes the covariant derivative with respect to the Levi-Civita connection of  $\mu_J$ . (In [F2] (2.10) Floer uses  $\nabla$  instead of  $D$ , and denotes  $D_t(\xi)$  by  $\nabla_z(\xi)$ .) Floer shows in [F2] Prop. 2.1 that, provided that the critical points  $x$  and  $y$  are nondegenerate, we may take the family  $A_\tau$  to be a slight perturbation of the family  $A_{u(\tau)}$  (suitably reparametrized). Moreover, the spectral flow of  $A_\tau$  is independent of the perturbation chosen. Since  $A_z$  is defined for all  $z \in \Omega V$ , it follows that  $\mu(u)$  is defined for every path  $u$  between  $x$  and  $y$ , and depends only on the homotopy class of this path relative to its endpoints. He also shows that  $\mu(u)$  is the index of the relevant Fredholm operator, so that it is the dimension of the component of  $\mathcal{M}(x, y)$  which contains  $u$ .

Notice that because  $\Omega V$  consists of contractible loops, two paths  $u$  and  $v$  in  $\Omega V$  from  $x$  to  $y$  differ homotopically by a 2-sphere, i.e.  $v$  is homotopic (rel  $x, y$ ) to the connected sum of  $u$  with the image of a 2-sphere in  $V$ . If the free homotopy class of this 2-sphere is  $A$ , we will write  $v = u\#A$ .

**Lemma 3.4.5.**  $\mu(u\#A) = \mu(u) + 2c_1(A)$ .

*Proof.* See Theorem 4 in [F2]. This result can also be proved using the Maslov index of [SZ].  $\square$

This lemma shows that one can define a mod  $2\Gamma$  index  $\text{Int}(x)$  for the critical points  $x$  of  $a_H$ . We now come to a very important application of the monotonicity hypothesis.

**Lemma 3.4.6.** *When  $(V, \omega)$  is monotone, there is for each  $K$  a number  $L_K$  such that  $\ell(u) \leq L_K$  whenever  $\mu(u) \leq K$ .*

*Proof.* As we saw in Lemma 3.4.2,

$$\ell^2(u) = A(u) + \int_{S^1} [H(x(t)) - H(y(t))] dt.$$

Hence  $\ell^2(u\#A) = \ell^2(u) + \omega(A) = \ell^2(u) + kc_1(A)$ , since, by monotonicity,  $\omega(A) = kc_1(A)$ . The result now follows from Lemma 3.4.5 and the fact that  $a_H$  has only finitely many critical points.  $\square$

*The set  $(\mathcal{F} \times \mathcal{H})_{\text{reg}}$ .* The remaining step in the proof of part (i) of Theorem 3.3.3 is to show that  $\mathcal{F} \times \mathcal{H}$  contains a dense set of regular elements  $(\mathcal{F} \times \mathcal{H})_{\text{reg}}$ . It is not hard to check that

the critical points of  $a_H$  are nondegenerate for generic  $(J, H)$ . (This point is discussed further in Lemma 3.4.9 below.) Hence the Fredholm theory may be set up for generic  $(J, H)$  and as in (2.3), it suffices to show that one can obtain a generic perturbation of  $\bar{\partial}_{J,H}$  by varying the pair  $(J, H)$  within a Banach manifold  $\mathcal{F}' \times \mathcal{H}'$ , where  $\mathcal{F}'$  is as in (2.3) and  $\mathcal{H}'$  is a similarly defined Banach subspace of  $\mathcal{H}$ . This is easy to do, since we have such freedom in the choice of  $H$ : see [F5, §2c]. Thus, just as in Theorem 2.3.3, we can show:

**Proposition 3.4.7.** *There is a dense set  $(\mathcal{F} \times \mathcal{H})_{\text{reg}}$  in  $\mathcal{F} \times \mathcal{H}$  such that for every element  $(J, H) \in \mathcal{F} \times \mathcal{H}$  the action functional  $a_H$  has a finite set  $Z$  of nondegenerate critical points, and the trajectory spaces  $\mathcal{M}(x, y)$ ,  $x, y \in Z$ , are smooth manifolds. Further, the dimension of the component of  $\mathcal{M}(x, y)$  containing the trajectory  $u$  equals  $\mu(u)$ .*

This “transversality” result is sufficient for many purposes. However, when we calculated the Floer complex  $(F_*, \partial)$  in Proposition 3.3.7, it was convenient to use a time-independent  $H$ . Since the property of being time independent is not generic in  $\mathcal{H}$ , we must justify this by further arguments. Observe that when  $H$  is time independent, there may be multiply covered trajectories, i.e. trajectories  $v$  such that  $v(\tau, t) = u(m\tau, mt)$  for some  $(J, H/m)$ -trajectory  $u$  and some  $m > 1$ . These do not occur for generic elements of  $\mathcal{H}$  since the corresponding critical points of  $a_H$  are not multiply covered. However, because the Hamiltonian flows of time-independent functions have fixed points, they do occur for generic time-independent  $H$ . We will say that a trajectory  $u$  is *simple* if it is not multiply covered. Similarly, the pair  $(u, J)$  is simple if  $u$  is.

**Lemma 3.4.8.** *If  $H$  is any element of  $\mathcal{H}$ , and  $x, y$  are distinct nondegenerate critical points of  $a_H$ , then the subset of  $P_{1,p}(x, y) \times \mathcal{F}'$  formed by all simple pairs  $(u, J)$  such that  $\bar{\partial}_{J,H}u = 0$  is a Banach manifold.*

*Sketch of proof.* We will indicate the changes in [MD2, Proposition 4.1] needed to deal with the extra term in the equation, suppressing various details which are caused by the complications of analysis on  $\mathbf{R} \times S^1$ . (For these, see [F1, §5].) To be consistent with the notation of [MD2], we will write  $f$  instead of  $u$ .

The setup is just the same as before, except that  $S^2$  should be replaced everywhere by  $\mathbf{R} \times S^1$  and the operator  $\Phi$  has the formula

$$\Phi(f, J) = \Psi(f, J)[df + J \circ df \circ i + h].$$

Here  $h$  is the section of  $\Lambda_J^{0,1}(E_f)$  given by  $h(\partial/\partial\tau) = -\nabla H_t = JdH_t$ , and  $h(\partial/\partial t) = J\nabla H_t = -dH_t$ . Let  $\alpha$  be the section of  $\text{Hom}(T(\mathbf{R} \times S^1), E_f)$  given by  $\alpha(\partial/\partial\tau) = dH_t$ , and  $\alpha(\partial/\partial t) = 0$ . Then  $h = J\alpha - \alpha \circ i$ , and it is easy to check that the linearization  $L\Phi$  of  $\Phi$  at  $(f, J)$  satisfies

$$L\Phi(0, v) = \Psi(f, J)[v \circ df \circ i + v \circ \alpha].$$

Further  $v \circ df \circ i + v \circ \alpha$  is anti- $J$ -holomorphic, i.e. belongs to  $\Lambda_J^{0,1}(E_f)$ , because  $\Phi(f, J) = 0$ .

The calculation of  $L\Phi(w, 0)$  is much as before, except that there are some new terms of zero order coming from  $h$ . Because the endpoints  $x$  and  $y$  are nondegenerate, this operator is Fredholm, and so has closed image with finite-dimensional cokernel. Hence, it suffices to show that the image of  $L\Phi$  is dense.

Observe that equation (b) in [MD2] now reads  $\langle \xi, v \circ df \circ i + v \circ \alpha \rangle_2 = 0$ . There is an open set on which  $df \circ i + \alpha \neq 0$ , because the value of this homomorphism on  $\partial/\partial t$  is  $\partial f/\partial\tau$ , and this must be nonzero on an open set since  $x \neq y$  and  $f$  is smooth. Further, one can show as in [MD2, Lemma 4.4(i)] that the hypothesis that  $f$  is simple implies that  $df \circ i + \alpha \neq 0$  on some open subset of  $\mathbf{R} \times S^1$  on which  $f$  is injective. (This statement is the analogue of [F1, Lemma 5.3]. Note that the local properties of solutions of the equation  $\bar{\partial}_J f = g(f)$  are just like those of solutions of  $\bar{\partial}_J f = 0$ : indeed, by [G1, 1.4.C'], the graphs of solutions of the former equation are  $J_g$ -holomorphic for suitable  $J_g$ .) Hence the argument that  $L\Phi$  has dense image goes through without essential change.  $\square$

Now consider the case when  $H$  is time independent. Let  $x \in \Omega V$  be the constant loop at a critical point  $x_0$  of  $H$  on  $V$  and let  $S_J : T \rightarrow T$  denote its Hessian with respect to the metric  $\mu_J$ , where  $T$  is the tangent space to  $V$  at the point  $x_0$ . It is easy to check that  $\mu_J(S_J(X), Y) = XY(H) = -\omega(JS_J(X), Y)$ , for all vector fields  $X, Y$  on  $V$ . Hence the operator  $JS_J$  is independent of the choice of  $J$ , which implies that the condition on  $H$  in the lemma below is also independent of the choice of  $J$ .



**Lemma 3.4.9.** *Let  $H$  be time independent. Then  $x$  is a nondegenerate critical point of  $a_H$  if and only if there is no integer  $k$  such that  $2\pi ik$  is in the spectrum of  $JS_J$ . In particular,  $x_0$  must be a nondegenerate critical point of  $H$ .*

*Proof.* Since  $\dot{x} = 0$ , the formula (3.4.4) for the Hessian  $A_x$  of  $a_H$  at  $x$  simplifies to  $A_x(\xi(t)) = J(D_t\xi(t)) - S_J(\xi(t))$ , where  $\xi(t) \in W_{1,2}(S^1, T)$ . Therefore,  $x$  is nondegenerate if and only if the equation  $A_x(\xi) = 0$  has no nonzero solutions  $\xi(t)$  with  $\xi(0) = \xi(1)$ . Since its solutions have the form  $\xi(t) = \exp(-tJS_J)\xi(0)$ , this is equivalent to requiring that the spectrum of  $JS_J$  contains no element  $2\pi ik$ ,  $k \in \mathbf{Z}$ .  $\square$

**Proposition 3.4.10.** *Let  $H$  be a generic time-independent function on  $V$ . When  $(V, \omega)$  is monotone, and  $\lambda$  is sufficiently small, there is an element  $(J, \lambda H) \in (\mathcal{F} \times \mathcal{H})_{\text{reg}}$  such that the only components of the  $(J, \lambda H)$ -trajectory spaces  $\mathcal{M}(x, y)$  which contribute to  $(F_*, \partial)$  consist of time-independent trajectories.*

*Proof.* Let  $H$  be a generic function on  $V$  and choose  $\lambda > 0$  small enough so that:

- (i)  $\sup(\lambda H) - \inf(\lambda H) \leq k$  (where  $[\omega] = kc_1$  by monotonicity), and
- (ii)  $\lambda H$  has no nonconstant periodic orbits of period  $\leq 1$ .

Condition (ii) may be achieved because, for each  $K > 0$ , the Hamiltonian flow on  $(V, \omega)$  of a generic function  $H$  has finitely many periodic orbits of period  $\leq K$ . It follows that the only critical points of  $a_{\lambda H}$  are constant loops at the critical points of  $H$ . Hence, because  $H$  is generic, we can assume that the nondegeneracy condition of Lemma 3.4.9 is satisfied. Further, the image  $\text{Im } u$  of each trajectory is a 2-sphere in  $V$ , so that its area  $A(u)$  is just the value  $\omega([u])$  of  $\omega$  on this sphere. By Lemma 3.4.2,  $\ell^2(u) = \omega([u]) + \lambda H(x) - \lambda H(y)$  and is positive. Hence (i) implies that  $\omega([u]) > -k$ . But this means that  $\omega([u]) \geq 0$  since, by the integrality of  $c_1$ , all negative values of  $[\omega] = kc_1$  are  $\leq -k$ . Therefore,  $c_1([u]) \geq 0$ , and so, by Lemma 3.4.5, the index  $\mu(u^m)$  of the multiple trajectory  $u^m = u(m \cdot, m \cdot)$  is  $\geq \mu(u)$ .

By Lemma 3.4.8, we may choose  $J \in \mathcal{F}$  so that, for all  $m \geq 1$ , the pair  $(J, \lambda H/m)$  is regular for all simple trajectories, i.e. each simple  $(J, \lambda H/m)$ -trajectory  $u$  has a neighborhood in  $\mathcal{M}(x, y)$  which is diffeomorphic to  $\mathbf{R}^{u(u)}$ . Then, although the  $(J, \lambda H)$ -trajectory space  $\mathcal{M}(x, y)$  could conceivably have singularities

near its multiple trajectories, each such trajectory has the form  $u^m$  where  $u$  is a regular simple  $(J, \lambda H/m)$ -trajectory.

The crucial point now is that any solution  $u$  of  $\bar{\partial}_{J, \lambda H/m} u = 0$  which is not constant with respect to  $t$  has an extra rotational symmetry. Hence it has a 2-parameter family of reparametrizations, and so, provided that it is regular so that its index equals the local dimension of  $\mathcal{M}(x, y)$ , it has index  $\mu(u) \geq 2$ . Hence its multiple coverings  $u^m$  also have index  $\geq 2$ . It follows that every  $(J, \lambda H)$ -trajectory which depends on time lies in a component of  $\mathcal{M}(x, y)$  of dimension  $\geq 2$ . Since the boundary operator  $\partial$  is defined by the components of dimension 1, such trajectories do not contribute to  $(F_*, \partial)$ .  $\square$

*Compactness.* The situation here is very similar to that in (2.4), except that now there is another source of noncompactness: besides the possibility of  $J$ -holomorphic spheres bubbling off, a family of trajectories from  $x$  to  $y$  can split up into a  $k$ -trajectory from  $x$  to  $y$ , which is a sequence of  $k$ -trajectories (with  $k > 1$ ) going from  $x$  to  $z_1$ , from  $z_1$  to  $z_2$ , and so on until it reaches  $z_k = y$ . The following proposition proves part (T3) of Theorem 3.3.3.

**Proposition 3.4.11.** *When  $V$  is monotone and  $(J, H)$  is regular, the components of  $\widehat{\mathcal{M}}(x, y)$  of dimension 0 are compact (and hence finite) and those of dimension 1 are compact except for sequences of trajectories which split up into two.*

*Proof.* Consider a sequence  $\{u_\alpha\}$  in  $\mathcal{M}(x, y)$  with  $\mu(u_\alpha) = m$ . By Lemma 3.4.6 the lengths of the  $u_\alpha$  are bounded, which means that the  $u_\alpha$  are uniformly bounded in  $W_{1,2}(\mathbf{R} \times S^1)$ . Let us suppose first that the  $u_\alpha$  are also uniformly bounded in  $W_{1,p,\text{loc}}(\mathbf{R} \times S^1)$  for some  $p > 2$ , i.e. that for each  $\rho > 0$ , the restrictions of the  $u_\alpha$  to  $[-\rho, \rho] \times S^1$  are  $W_{1,p}$ -bounded. Then the argument of Theorem 2.4.1 shows that there is a subsequence (also called  $u_\alpha$ ) which converges in the local  $C^n$  topology. (The extra term in  $\bar{\partial}_{J,H}$  involving  $H$  makes no essential difference to the analysis.) We claim that this implies that  $\{u_\alpha\}$  splits up into a  $k$ -trajectory. To see this, observe that any sequence  $u_\alpha \circ \sigma_\alpha = u_\alpha(\cdot + \sigma_\alpha)$  which is obtained from  $u_\alpha$  by reparametrization also has a subsequence which converges in the local  $C^n$ -topology. Its limit will be a solution  $v$  of the equation  $\bar{\partial}_{J,H} v = 0$  in  $C^n(\mathbf{R} \times S^1)$  with bounded length. By Lemma 3.4.2,  $v$  belongs to some trajectory space

$\mathcal{M}(x', y')$ , and is constant with respect to  $\tau$  if and only if it has zero length.

It is now easy to see that as  $\{\sigma_\alpha\}$  ranges over all possible reparametrizations, there are only a finite number of possible limits  $v$ , and that they may be arranged in a sequence  $v_1, \dots, v_k$  where  $v_1$  goes from  $x = z_0$  to  $z_1$ ,  $v_2$  goes from  $z_1$  to  $z_2$ , and so on. Clearly, the index of the  $u_\alpha$  is the sum of the indices of the  $v_i$ . But because  $(J, H)$  is regular there are no components of  $\mathcal{M}(x, y)$  with dimension  $< 1$ . (Because there is a 1-dimensional reparametrization group, no component of  $\mathcal{M}(x, y)$  can have dimension 0.) Hence this splitting cannot occur if  $u$  has index 1 and, if the  $u_\alpha$  have index 2, they can only split into a pair of isolated trajectories.

Now, consider the case when the  $u_\alpha$  are not uniformly bounded in  $W_{1,p,\text{loc}}(\mathbf{R} \times S^1)$  for any  $p > 2$ . Then, as in Theorem 2.4.2 and Lemma 3.4.2, one can find a subsequence from which a  $J$ -holomorphic sphere bubbles off. In fact, there may be a finite number of spheres bubbling off, and the trajectories  $u_\alpha$  may also converge to a  $k$ -trajectory. But, in any case, there must be a subsequence which converges to the union of a  $k$ -trajectory  $v$  with a finite number of  $J$ -holomorphic spheres of classes  $A_1, \dots, A_p$ . By Lemma 3.4.5 we must have  $\mu(v) + 2c_1(A_1) + \dots + 2c_1(A_p) = \mu(u_\alpha) = m$ , where  $\mu(v) \geq 1$  if  $v$  is not constant because  $(J, H)$  is regular. Further,  $\omega(A_i) > 0$  for each  $i$  since  $A_i$  has a  $J$ -holomorphic representative. Hence  $c_1(A_i) > 0$  by monotonicity, and so this cannot occur when  $m = \mu(u_\alpha) \leq 2$ . It is also possible that  $x = y$ ,  $v$  is constant,  $p = 1$  and  $c(A_1) = 1$ . But in this case, the loop  $x$  must be a single point lying on the bubble. This cannot happen for generic  $J$  because, by (2.3.3), when  $c(A) = 1$  the set of points which lie on an  $A$ -sphere has codimension  $\geq 2$ .  $\square$

The remaining parts of Theorem 3.3.3 involve the technique of gluing. Floer’s discussion of this in [F5, 2d] is very clear, and there is nothing I can usefully add to it.

**(3.5) Other applications of Floer’s theory.**

(3.5.1) *Lagrangian intersection theory.* Because the graph of a symplectic diffeomorphism is Lagrangian, one can generalize Arnold’s conjectures to questions about the intersection of Lagrangian submanifolds. Instead of counting fixed points of exact diffeomorphisms, one counts the points of intersection of a Lagrangian submanifold  $L$  with its image  $L_1$  under an exact de-

formation. Floer's construction applies to this case, and many of his papers on the Arnold conjectures are in this language. His results are proved for Lagrangian submanifolds  $L$  of  $(V, \omega)$  such that  $\omega$  vanishes on all elements of  $\pi_2(V, L)$ . Recently Oh [O1] generalized this to the case of symmetric Lagrangian submanifolds in monotone  $(V, \omega)$ , i.e. to Lagrangian submanifolds  $L$  which form the fixed point set of some antisymplectic involution. Givental in [Gi] has a completely different and rather more geometric approach to the problem which works for the symmetric submanifold  $L = \mathbf{R}P^n$  in  $CP^n$ , and it would be interesting to know how far these methods can be combined and/or generalized.

(3.5.2) *Capacities.* Roughly speaking, Ekeland and Hofer define the capacity  $c(H)$  of a function  $H$  on  $\mathbf{R}^{2n}$  to be the action  $a_H(z)$  of some particular periodic orbit  $z$  of its Hamiltonian flow. The capacity of a set  $S$  is then defined to be the infimum of  $c(H)$  taken over all  $H$  which vanish on  $S$  and satisfy some growth condition at infinity. The main difficulty here is to find a natural way to describe the special periodic orbit used to define  $c(H)$ . The variational methods used for this purpose by Hofer and Viterbo are rather indirect. Moreover, as one sees from [HV2], even to transfer them from  $\mathbf{R}^{2n}$  to the class of cotangent bundles requires a huge amount of work. In contrast, Floer's elliptic techniques work in any symplectic manifold, provided that the  $J$ -holomorphic spheres in that manifold can be understood. Further, as we shall see in (3.5.3) below, they do allow one to get a handle on the periodic orbits since these appear as the vertices of the Floer complex. Therefore, one may be able to extend significantly the range and power of capacity theory by using Floer's elliptic techniques.

Floer and Hofer have made a beginning in this definition, by working out a definition of Ekeland and Hofer's sequence of capacities for subsets of  $\mathbf{R}^{2n}$  which uses these techniques instead of critical point theory. (This work has not been written up yet.) However, their definition still applies only to  $\mathbf{R}^{2n}$ . Together with Viterbo, they have also tried to extend the theory of capacities to more general symplectic manifolds. Their first attempt may be found in [FHV], where they consider manifolds of the form  $V \times \mathbf{R}^{2n}$  where  $\omega$  vanishes on  $\pi_2(V)$ . The argument there, although interesting, is still somewhat clumsy. However, it is refined by Hofer and Viterbo in [HV1] (see also [V1]), where, among

other things, one can find a simple and elegant proof of Gromov's squeezing theorem. To illustrate the kind of ideas involved, we will end this article by summarizing their argument.

(3.5.3) *A Floer-type proof of Gromov's squeezing theorem.* First of all, Hofer and Viterbo reformulate the problem so that it becomes a question about the existence of periodic orbits for a certain Hamiltonian flow. They observe, as before, that it suffices to consider a symplectic embedding  $g$  of  $B^{2n}(r)$  into  $(S^2 \times V, \tau)$  where  $V$  is a product of 2-spheres and  $\tau$  is the form  $(R + \varepsilon)^2[\sigma_1 \oplus K^2\sigma_2 \oplus \dots \oplus K^2\sigma_n]$ . Pick points  $x$  in  $\text{Im } g$  and  $y$  outside  $\text{Im } g$ , and let  $H$  be any smooth function  $S^2 \times V \rightarrow [0, 1]$  which is 0 near  $x$  and 1 near  $y$  and has no critical values other than 0 and 1. (For example,  $H$  could be constant except in a small collar neighborhood of the  $(2n - 1)$ -sphere  $g(\partial B^{2n}(r))$ .) Then they prove that the Hamiltonian flow of such a function  $H$  has a nonconstant periodic solution of period  $T < \pi(R + \varepsilon)^2$ . By considering the corresponding functions  $H \circ g$  on  $B^{2n}(r)$ , one deduces that the Hamiltonian flow of a function  $\phi(|x|^2)$  has periodic orbits of period  $< \pi(R + \varepsilon)^2$ , whenever  $\phi$  is a nondecreasing function from  $[0, \infty)$  onto  $[0, 1]$  such that  $\phi(s) = 1$  for  $s^2 \geq r^2$ . But an easy calculation shows that the period of these orbits is  $\pi/\phi'(s)$ . Hence, choosing suitable  $\phi$ , one sees that  $r \leq R$ , as required.

In terms of Floer's theory, we are just asserting the existence of a nontrivial critical point of  $a_{\lambda H}$  for some  $\lambda < \pi(R + \varepsilon)^2$ , since these critical points correspond to 1-periodic solutions of the Hamiltonian flow of  $\lambda H$ , and hence to  $\lambda$ -periodic solutions of the Hamiltonian flow of  $H$ . Consider the family of action functionals  $a_{\lambda H}$  for  $\lambda \geq 0$ , and let  $x, y$  be as above. By composing  $g$  with a symplectic diffeomorphism of  $S^2 \times V$ , we may suppose that  $x, y$  both lie on the 2-sphere  $S^2 \times \{p\}$ , for some  $p \in V$ . Let  $J$  be the standard complex structure on  $S^2 \times V$ . When  $\lambda = 0$ , the elements of  $\mathcal{M}_0(x, y)$  are  $J$ -holomorphic spheres (reparametrized as maps with domain  $\mathbf{R} \times S^1$ ), and so these are completely understood. In particular, there is a compact 1-dimensional component of  $\mathcal{M}_0(x, y)$  corresponding to the holomorphic sphere  $S^2 \times \{p\}$ . However, when  $\lambda \geq \pi(R + \varepsilon)^2 = \tau(S^2 \times \{p\})$ , there is no trajectory  $u$  in  $\mathcal{M}_\lambda(x, y)$  in this homology class, because by the formula in Lemma 3.4.2,  $\ell^2(u)$  would have to be  $\leq 0$ . Therefore, this component must disappear as  $\lambda$  increases from 0 to  $\pi(R + \varepsilon)^2$ . Since

the minimum value of  $\omega$  on the  $J$ -holomorphic spheres occurs for the sphere  $S^2 \times \{p\}$ , bubbling off of spheres can be ruled out, and hence the only way that this component can disappear is for a trajectory in  $\mathcal{M}_0(x, y)$  to split up into a pair of trajectories, one from  $x$  to  $z$  and the other from  $z$  to  $y$ . The critical point  $z$  is the desired periodic orbit.

## REFERENCES

- [A1] V. I. Arnold, *Mathematical methods of classical mechanics*, Springer-Verlag, Berlin, Heidelberg, New York, 1978.
- [A2] —, *First steps in symplectic topology*, Russian Math. Surveys **41** (6), (1986), 1–21.
- [At] M. Atiyah, *New invariants of 3 and 4 dimensional manifolds*, Proc. Sympos. Pure Math., vol. 48, Amer. Math. Soc., Providence, R.I., 1988, 285–299.
- [AB] M. Atiyah and R. Bott, *The Yang–Mills equations over Riemann surfaces*, Proc. Roy. Soc. London Ser. A **308** (1982), 523–615.
- [BP] L. Bates and G. Peschke, *A remarkable symplectic structure*, preprint, Calgary, 1989.
- [Be] D. Bennequin, *Problèmes elliptiques, surfaces de Riemann, et structures symplectiques*, Sémin. Bourbaki 1985–86 #657, Astérisque **145–146** (1987), 4, 111–136.
- [B] R. Bott, *Lectures on Morse theory, old and new*, Bull. Amer. Math. Soc. **7** (1982), 331–358.
- [CZ] C. Conley and E. Zehnder, *The Birkhoff–Lewis fixed point theorem and a conjecture of V. I. Arnold*, Invent. Math. **73** (1983), 33–49.
- [E1] Ya. Eliashberg, *A theorem on the structure of wave fronts*, Functional Anal. Appl. **21** (1987), 65–72.
- [E2] —, Three lectures on symplectic topology, Conference at Cala Gonone, Italy, 1988.
- [E3] —, *On symplectic manifolds which are bounded by standard contact spheres*, J. Differential Geom. (to appear).
- [EH] I. Ekeland and H. Hofer, *Symplectic topology and Hamiltonian dynamics I*, Math. Z. **200** (1989), 355–378; *Part II*: Math. Z. (to appear).
- [F1] A. Floer, *The unregularized gradient flow of the symplectic action*, Comm. Pure Appl. Math. **XLI** (1988), 775–813.
- [F2] —, *A relative Morse index for the symplectic action*, Comm. Pure Appl. Math. **XLI** (1988), 393–407.
- [F3] —, *Morse theory for Lagrangian intersections*, J. Differential Geom. **28** (1988), 513–547.
- [F4] —, *An instanton invariant for 3-manifolds*, Comm. Math. Phys. **118** (1988), 215–240.
- [F5] —, *Symplectic fixed points and holomorphic spheres*, Comm. Math. Phys. **120** (1989), 575–611.
- [F6] —, *Witten’s complex and infinite dimensional Morse theory*, J. Differential Geom. **30** (1989), 207–221.

- [F7] —, *Cuplength estimates on Lagrangian intersections*, Comm. Pure Appl. Math. **XLII** (1989), 335–357.
- [FHV] A. Floer, H. Hofer and C. Viterbo, *The Weinstein conjecture in  $P \times \mathbf{C}^1$* , Math Zeitschrift **203** (1990), 469–482.
- [FU] D. Freed and K. Uhlenbeck, *Instantons and four-manifolds*, Publ. MSRI, vol. 1, Springer-Verlag, New York, 1984.
- [Gi] A. Givental, *Nonlinear Maslov index and Lagrangian intersections*, preprint, 1990.
- [GS] R. Greene and K. Shiohama, *Diffeomorphisms and volume-preserving embeddings of non-compact manifolds*, Trans. Amer. Math. Soc. **255** (1979), 403–414.
- [GH] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Wiley, New York, 1978.
- [G1] M. Gromov, *Pseudo-holomorphic curves on almost complex manifolds*, Invent. Math. **82** (1985), 307–347.
- [G2] —, *Partial differential relations*, Ergebnisse der Math., Springer-Verlag, Berlin, Heidelberg, New York, 1986.
- [G3] —, *Soft and hard symplectic geometry*, Proceedings of the ICM at Berkeley, 1986, 1987, vol. 1, Amer. Math. Soc., Providence, R.I., pp. 81–98.
- [Gr] A. Grothendieck, *Sur las classification des fibrés holomorphes sur la sphère de Riemann*, Amer. J. Math. **76** (1957), 121–138.
- [H1] H. Hofer, *On Ljusternik-Schnirelman theory for Lagrangian intersections*, Ann. Inst. H. Poincaré Anal. NonLinéaire (1988).
- [H2] —, *Recent progress in symplectic geometry*, Nonlinear functional analysis, edited by P. S. Milojevic, Marcel Dekker, New York (to appear).
- [H3] —, *On the topological properties of symplectic maps*, Proc. Roy. Soc. Edinburgh (special volume on the occasion of J. Hale's 60th birthday) (to appear).
- [HV1] H. Hofer and C. Viterbo, *The Weinstein conjecture for compact manifolds and related results*, preprint, 1989.
- [HV2] —, *The Weinstein conjecture for cotangent bundles and related results*, Ann. Scuola Norm. Sup. Pisa Cl. Sci., vol. 15, 1988, 411–446.
- [HZ] H. Hofer and E. Zehnder, *A new capacity for symplectic manifolds* (to appear in the Proceedings of the Conference on the occasion of J. Moser's 60th birthday).
- [K] W. Klingenberg, *Lectures on closed geodesics*, Grundlehren Math. Wiss., vol. 230, Springer-Verlag, Berlin, Heidelberg, New York, 1978.
- [Ko] J. Kollar, *The structure of algebraic threefolds, an introduction to Mori's program*, Bull. Amer. Math. Soc. **17** (1987), 211–277.
- [Kr] A. Krygin, *Continuation of diffeomorphisms preserving volume*, Functional Anal. Appl. **5** (1971), 147–150.
- [L] H. B. Lawson, *Minimal varieties in real and complex geometry*, Sem. Math. Sup., vol. 57, Presses Univ. Montreal, Montreal, 1974.
- [MD1] D. McDuff, *Symplectic diffeomorphisms and the flux homomorphism*, Invent. Math. **77** (1984), 353–366.
- [MD2] —, *Examples of symplectic structures*, Invent. Math. **89** (1987), 13–36.
- [MD3] —, *Rational and ruled symplectic 4-manifolds*, J. Amer. Math. Soc. (to appear).

- [M] J. Moser, *On the volume elements on manifolds*, Trans. Amer. Math. Soc. **120** (1965), 280–296.
- [O1] Y.-G. Oh, *Floer cohomology of symplectic Lagrangian submanifolds*, preprint, (NYU) 1989.
- [O2] —, *Removal of boundary singularities of pseudo-holomorphic curves with Lagrangian boundary conditions*, preprint NYU, 1990.
- [P] P. Pansu, *Pseudo-holomorphic curves in symplectic manifolds*, preprint, Ecole Polytechnique, Palaiseau, 1986.
- [S] D. Salamon, *Morse theory, the Conley index and Floer homology*, Bull. London Math. Soc. **22** (1990), 113–140.
- [SZ] D. Salamon and E. Zehnder, *Floer homology, the Maslov index and periodic orbits of Hamiltonian equations*, Analysis Et Cetera (P. H. Rabinowitz and E. Zender, eds.), Academic Press, 1990, 573–600.
- [Sm] S. Smale, *An infinite dimensional version of Sard's theorem*, Amer. J. Math. **87** (1973), 213–221.
- [T] C. Taubes, *Moduli spaces and homotopy theory*, Proc. Sympos. Pure Math., vol. 48, Amer. Math. Soc., Providence, R.I., 1988, 301–315.
- [V1] C. Viterbo, *Capacités symplectiques et applications*, Sem. Bourb. #714, June 1989.
- [V2] —, *Recent progress in periodic orbits of autonomous Hamiltonian systems*, Non-Linear Functional Analysis (P. S. Milojevic, Marcel Dekker, eds.) (to appear).
- [Wn1] A. Weinstein, *Symplectic manifolds and their Lagrangian submanifolds*, Adv. in Math. **6** (1971), 329–346.
- [Wn2] —, *On extending the Conley-Zehnder fixed point theorem to other manifolds*, Proc. Sympos. Pure Math., vol. 45, Amer. Math. Soc., Providence, R.I., 1986, 451–454.
- [W1] E. Witten, *Supersymmetry and Morse theory*, J. Differential Geom. **17** (1982), 661–692.
- [W2] —, *Geometry and quantum field theory*, Amer. Math. Soc. Lecture, summer meeting, 1988.
- [Wo] J. G. Wolfson, *Gromov's compactness of pseudo-holomorphic curves and symplectic geometry*, J. Differential Geom. **28** (1988), 383–405.

STATE UNIVERSITY OF NEW YORK, STONY BROOK, NEW YORK

E-mail: DUSA@MATH.SUNYSB.EDU