Treball final de màster

MÀSTER DE MATEMÀTICA AVANÇADA

Facultat de Matemàtiques
Universitat de Barcelona

Gromov compactness theorem for pseudoholomorphic curves

Autor: Carlos Sáez Calvo

Director: Dr. Ignasi Mundet i Riera
Realitzat a: Departament de Àlgebra i Geometria

Barcelona, 14 de setembre de 2014

Contents

Introduction

The main goal of this master thesis is to give a self-contained proof of the Gromov compactness theorem for pseudoholomorphic curves and the non-squeezing theorem in symplectic topology. Pseudoholomorphic curves are smooth maps from a Riemann surface into an almost complex manifold that respect the almost complex structures. If the target manifold is a complex manifold, we recover the notion of holomorphic maps, so pseudoholomorphic maps can be seen as the generalization of holomorphic maps to the almost complex setting. Pseudoholomorphic curves were introduced by Gromov in a ground-breaking paper published in 1985, [Gro]. Since then, they have become one of the main tools in the field of symplectic topology.

Any symplectic manifold can be endowed with an almost complex structure that satisfies a condition of compatibility with the symplectic form. In this setting, one can consider moduli spaces of pseudoholomorphic curves, which in most situations are manifolds of finite dimension. However, in general these moduli spaces are not compact, and the problem of seeking a natural compactification for these moduli spaces arises. This problem was first solved by Gromov in his paper [Gro], and a complete picture of the compactified moduli space was finally given by Kontsevich in terms of stable maps. Essentially, the Gromov compactness theorem asserts that the failure of compactness of the moduli spaces of pseudoholomorphic curves can happen only in a very controlled way.

Compactness of the moduli spaces is crucial for applications in symplectic topology. One of the important results of Gromov in [Gro] was a proof of the non-squeezing theorem, asserting that a ball can be symplectically embedded in a symplectic cylinder if and only if the radius of the ball is smaller than the radius of the cylinder. Its importance lies in the fact that this leads to the definition of non-trivial global symplectic invariants, while it has been known for a long time that there are no local symplectic invariants. Moreover, compactness of the moduli spaces also play a crucial role for further developments of the theory, such as the definition of Gromov-Witten invariants or Floer homology.

The main novelty of this thesis, compared to the proofs of the compactness theorem that can be found in the literature, lies in the fact that we use exclusively the L^2 theory of Sobolev spaces, while the usual proofs use either L^p theory of Sobolev spaces (for a general p) or isoperimetric inequalities.

The thesis is divided into several chapters. In the first one we introduce all the elements of symplectic geometry necessary for the rest of the work. In particular, we introduce symplectic manifolds, complex manifolds, almost complex manifolds and Kaehler manifolds, studying its properties and the relations between them.

In the second chapter, we introduce our main object of interest, pseudoholomorphic curves. After studying its basic properties, we organize them in a topological space, the moduli space of pseudoholomorphic curves. An important property of the moduli spaces of pseudoholomorphic curves is that, under suitable conditions, they are finite-dimensional smooth manifolds. However, this result needs some analytic machinery that is beyond the scope of this work. Therefore, we only give a brief idea of how this can be proved. In the last section, we see that these moduli spaces are not, in general, compact. Therefore, the need for a geometrically meaningful compactification of the moduli spaces arises.

In the third chapter, we introduce the objects that form the compactification of the moduli spaces, namely, the stable curves. We introduce the moduli spaces of stable curves and give them a suitable topology. In the last section of the chapter, we give the statement of the Gromov compactness theorem. For the notations and definitions for stable curves, we follow closely the reference [M-S1].

The fourth chapter constitutes the main contribution of this thesis. In this chapter of a technical nature, we develop estimates for pseudoholomorphic curves needed for the proof of the compactness theorem using only L^2 methods. We use the obtained estimates to give a self-contained proof of the removal of singularities theorem, which is an analogue of the Riemann extension theorem for holomorphic functions (asserting that any bounded holomorphic function defined in a punctured disk extends uniquely to a holomorphic function in the whole disk) in the pseudoholomorphic setting.

The fifth chapter uses the results obtained in the previous chapter in order to give a complete proof of the Gromov compactness theorem. The proof also follows closely [M-S1].

In the last chapter, we give a proof of the Gromov's non-squeezing theorem and discuss its importance. In particular, we use the theorem to define symplectic invariants. Our proof is essentially the same given by Gromov in [Gro], but with more detail. Our exposition follows closely that of [Hum].

Finally, I want to thank my advisor Ignasi Mundet i Riera for his guidance and ideas.

Chapter 1

Introduction to symplectic and almost complex manifolds

This chapter is a short introduction to symplectic geometry, where we cover the basics and the topics we need for our exposition of pseudoholomorphic curves and the Gromov compactness theorem. In particular, we discuss symplectic manifolds, almost complex manifolds, complex manifolds and how all this concepts relate to each other. In the last section, we give a very brief overview of Kähler manifolds, which are a big source of examples of symplectic manifolds.

1.1 Symplectic vector spaces

We start our exposition of symplectic geometry by studying the linear case, that is, a vector space with a non degenerate alternating bilinear form. We will see later that most of the results obtained in this linear setting extend to symplectic manifolds.

Definition 1.1.1. *Let V be a finite dimensional vector space over $\mathbb{R}$. A bilinear form ω satisfying the two following conditions:*

1. *(Alternating) $\omega(v,w) = -\omega(w,v)$ for all $v, w \in V$*
2. *(Non degenerate) If $\omega(v,w) = 0$ for all $w \in V$, then $v = 0$*

is called a symplectic form, and the pair (V, ω) is called a symplectic vector space.

Observe that an alternating bilinear form in V is the same as a 2-form in V.
We introduce now our main examples of symplectic vector spaces.

Example 1.1.2. Consider the vector space $\mathbb{R}^{2n}$ (of dimension $2n$), and write the standard basis as $e_1, ..., e_n, f_1, ..., f_n$. We define

$$\omega_0 = \sum_{i=1}^{n} e_i \wedge f_i$$

Clearly ω is an alternating bilinear form, and it is easy to see that it is non degenerate. Indeed, observe that if $v = a_1e_1 + ... + a_ne_n + b_1f_1 + ... + b_nf_n$ with $a_n, b_n \in \mathbb{R}$, then $\omega_0(v, e_i) = -b_i$

and $\omega_0(v, f_i) = a_i$. Therefore, if $\omega_0(v, w) = 0$ for all $w \in V$, then $0 = \omega_0(v, e_i) = -b_i$ and $0 = \omega_0(v, f_i) = a_i$, so $v = 0$.

Therefore $(\mathbb{R}^{2n}, \omega_0)$ is a symplectic vector space for each n.

Proposition 1.1.3. *Let V be a vector space and let ω be a bilinear alternating form in V (maybe degenerate). Then, there exists a basis of V,$u_1, ..., u_n, v_1, ..., v_n$, $w_1, ..., w_k$, such that $\omega = \sum_{i=1}^n u_i \wedge v_i$.*

Proof. Let us define $\ker(\omega) = \{v \in V : \omega(v, \cdot) = 0\}$. Observe that $\ker(\omega)$ is a subspace of V. Let $w_1, ..., w_k$ be a basis for $\ker(\omega)$. Choose now a complementary subspace to $\ker(\omega)$, W, so that $V = ker(\omega) \oplus W$. Observe that $\omega|_W$ is non degenerate. Take $u_1 \in W$ such that $u_1 \neq 0$. Since $\omega|_W$ is non degenerate, there exists a vector $v_1 \in W$ such that $\omega(u_1, v_1) = 1$. Define now $W_1 =< u_1, v_1 >$ and $W_1^\omega = \{v \in V : \omega(v, w) = 0 \text{ for all } w \in W_1\}$.

We claim that $W_1 \cap W_1^\omega = 0$. Indeed, if $v \in W_1 \cap W_1^\omega$, we have $v = au_1 + bv_1$. But $0 = \omega(v, u_1) = a\omega(u_1, u_1) + b\omega(v_1, u_1) = -b$ and $0 = \omega(v, v_1) = a\omega(u_1, v_1) + b\omega(v_1, v_1) = a$, so $v = 0$ as wanted.

Moreover, $W = W_1 \oplus W_1^\omega$. To see this, let $v \in W$ and let $a = \omega(v, u_1)$, $b = \omega(v, v_1)$. Then, $v = (-au_1 + bv_1) + (v + au_1 - bv_1)$. It is clear that $-au_1 + bv_1 \in W_1$ and $v + au_1 - bv_1 \in W_1^\omega$ by seeing that $\omega(v + au_1 - bv_1, u_1) = \omega(v + au_1 - bv_1, v_1) = 0$.

Observe now that $\omega|_{W_1^\omega}$ is non degenerate, because if $v \in W_1^\omega$ satisfies $\omega(v, w) = 0$ for all $w \in W_1^\omega$, since $\omega(v, u_1) = \omega(v, v_1) = 0$ by the definition of W_1^ω, we have that $\omega(v, w) = 0$ for all $w \in W_1$, so $v = 0$. Therefore, we can repeat the process and finish the proof by induction, since $\dim V < \infty$.

For the last statement, observe that by construction, $\omega(u_i, v_j) = \delta_{ij}$, $\omega(u_i, u_j) = \omega(v_i, v_j) = 0$ and $\omega(u_i, w_j) = \omega(v_i, w_j) = 0$. Therefore, we have $\omega = \sum_{i=1}^n u_i \wedge v_i$. □

This proposition gives us immediately the following corollary.

Corollary 1.1.4. *Let (V, ω) be a symplectic vector space. Then $\dim V$ is even.*

Moreover, this proposition tells us that any symplectic vector space is essentially one of the symplectic vector spaces from example 1.1.2.

Definition 1.1.5. *Let (V, ω), (V', ω') be two symplectic vector spaces. A map $\phi : V \longrightarrow V'$ is called a linear symplectomorphism if it is a linear isomorphism that satisfies $\phi^*\omega' = \omega$ (that is, $\omega'(\phi(v), \phi(w)) = \omega(v, w)$ for all $v, w \in V$). If there exists a linear symplectomorphism between two symplectic vector spaces, we say that the vector spaces are symplectomorphic.*

Corollary 1.1.6. *Let (V, ω) be a symplectic vector space of dimension $2n$. Then, V is symplectomorphic to $(\mathbb{R}^{2n}, \omega_0)$.*

Proof. Consider the basis $u_1, ..., u_n, v_1, ..., v_n$ of V given by the previous proposition (observe that since ω is non degenerate, we have $k = 0$). Define a linear symplectomorphism $\phi : V \longrightarrow \mathbb{R}^{2n}$ by $\phi(u_i) = e_i$, $\phi(v_i) = f_i$. Since ϕ takes a basis of V to a basis of $\mathbb{R}^{2n}$, ϕ is an isomorphism. Since $\omega = \sum_{i=1}^n u_i \wedge v_i$ and $\omega_0 = \sum_{i=1}^n e_i \wedge f_i$, it is immediate to check that $\phi^*(\omega_0) = \omega$. □

As in the case of a vector space with an inner product, the symplectic form gives us a natural isomorphism between V and its dual. In fact, this is true for any vector space with a non degenerate bilinear form.

Proposition 1.1.7. *Let V be a finite dimensional vector space and ω a non degenerate bilinear form in V. Then, the map $\tilde{\omega} : V \longrightarrow V^*$ defined by $\tilde{\omega}(u)(v) = \omega(u, v)$ is an isomorphism.*

Proof. $\tilde{\omega}$ is linear because ω is bilinear. Suppose that $\tilde{\omega}(u) = 0$. Then, $0 = \tilde{\omega}(u)(v) = \omega(u, v)$ for all $v \in V$. Therefore, since ω is non degenerate, we have $u = 0$. So $\tilde{\omega}$ is injective. Then, since V and V^* are vector spaces of the same dimension, $\tilde{\omega}$ is an isomorphism. □

In the symplectic setting, as in the inner product case, we can define the orthogonal of a subspace.

Definition 1.1.8. *Let (V, ω) be a symplectic vector space. If W is a subset of V, we define the symplectic complement of W as*

$$W^\omega = \{v \in V : \omega(v, w) = 0 \text{ for all } w \in W\}$$

However, the behaviour of the symplectic complement is very different from the orthogonal in the inner product case. For instance, recall that if g is an inner product on a vector space V, we have that $W^\perp \cap W = \{0\}$.

Example 1.1.9. Consider $(\mathbb{R}^{2n}, \omega_0)$.

1. Let $W =< e_1, ..., e_j >$ for some $j \leq n$. Then, $W^\omega =< e_1, ..., e_n, f_{j+1}, ..., f_n >$ and $W \subset W^\omega$ ($W = W^\omega$ if $j = n$).
2. Let $W =< e_1, ..., e_n, f_1, ..., f_j >$ for some $j \leq n$. Then, $W^\omega =< e_{j+1}, ..., e_n >$ and $W^\omega \subset W$.
3. Let $W =< e_1, f_1 >$. Then, $W^\omega =< e_2, ..., e_n, f_2, ..., f_n >$ and $W \cap W^\omega = \{0\}$.

This example illustrates the most important types of subspaces in a symplectic vector space. Each of these types of subspaces receive a name.

Definition 1.1.10. *Let (V, ω) be a symplectic vector space, and let $W \subset V$ be a subspace.*

1. *W is an isotropic subspace if $W \subset W^\omega$.*
2. *W is a coisotropic subspace if $W^\omega \subset W$.*
3. *W is a lagrangian subspace if $W = W^\omega$.*
4. *W is a symplectic subspace if $W \cap W^\omega = 0$.*

The previous example, together with the fact that any symplectic vector space is symplectomorphic to some $(\mathbb{R}^{2n}, \omega_0)$, shows that any symplectic vector space has subspaces of all kinds described in the definition. Note also that there exist in general subspaces that are of none of the above types.

However, an important feature shared by both the orthogonal and the symplectic complement, is the following.

Proposition 1.1.11. *Let (V, ω) be a symplectic vector space. Let $W \subset V$ be a subspace. Then*

$$\dim W + \dim W^{\omega} = \dim V$$

Proof. Consider the map $\Phi : V \longrightarrow W^*$ defined by $\Phi(u) = \omega(u, \cdot)|_W$. We have $\ker \Phi = \{u \in V : \omega(u, v) = 0 \text{ for all } v \in W\} = W^{\omega}$. Moreover, Φ is exhaustive, since we have seen in proposition 1.1.7 that $\tilde{\omega}$ is exhaustive. Since $\dim W^* = \dim W$, we have $\dim V = \dim \ker \Phi + \dim \Phi(V) = \dim W^{\omega} + \dim W$. □

In particular, it follows that if W is a lagrangian subspace, then $\dim W = \frac{\dim V}{2}$, and conversely, that a isotropic subspace is lagrangian if its dimension is $\frac{\dim V}{2}$.
To finish this section, we observe that any symplectic vector space comes equipped with a natural choice of orientation.

Proposition 1.1.12. *Let V be a vector space, and let ω be an alternating bilinear form. Then ω is non degenerate if and only if $\omega^n := \omega \wedge \overset{n)}{\ldots} \wedge \omega \neq 0$. In particular, if ω is a symplectic form, V has a natural choice of orientation.*

Proof. Suppose first that ω is degenerate. Then, there is some $0 \neq v \in V$ such that $\omega(v, \cdot) = 0$. Pick a basis $v, u_2, ..., u_k$ of V. Then $\omega^n(v, u_2, ..., u_k) = 0$, so $\omega^n = 0$. Conversely, suppose ω is non degenerate. Then, by 1.1.3, V has a basis $u_1, ..., u_n, v_1, ..., v_n$ such that $\omega = \sum_{i=1}^{n} u_i \wedge v_i$. Then, $\omega^n = \frac{1}{n!} u_1 \wedge v_1 \wedge ... \wedge u_n \wedge v_n \neq 0$. The last assertion follows from the fact that a choose non zero $2n$-form is equivalent to a choose of orientation of V. □

1.2 Symplectic manifolds

Definition 1.2.1. *Let X be a smooth manifold, and let ω be a 2-form in X. ω is called a symplectic form in X if it is non degenerate (i.e., ω_p is a non degenerate 2-form in T_pX for every $p \in X$) and ω is closed (i.e., $d\omega = 0$).*

A smooth manifold X endowed with a symplectic form ω is called a symplectic manifold.

From the work done in the previous section, we can draw immediately some facts about symplectic manifold.

Proposition 1.2.2. *Let (X, ω) be a symplectic manifold. Then X has even dimension and it is an orientable manifold.*

Proof. Since $\dim X = \dim T_pX$ for any $p \in X$, and (T_pM, ω_p) is a symplectic vector space, we have that $\dim X$ is even. Moreover, if $\dim X = 2n$, $\omega^n = \omega \wedge \overset{n)}{\ldots} \wedge \omega$ is a $2n$-form on X which is non zero at every point of X. Therefore ω^n is a volume form, and X is orientable. □

It follows from this proposition that not all manifolds admit a symplectic form. In the case of compact manifolds, the fact that ω is closed implies further topological constraints on X.

Proposition 1.2.3. *Let (X, ω) be a compact symplectic manifold. Then, $H^2_{deR}(X) \neq 0$, and $0 \neq [\omega] \in H^2_{deR}(X)$*

Proof. Since ω is a closed 2-form, it represents a second de Rham cohomology class. It is enough to see that $[\omega] \neq 0$, that is, that ω cannot be exact. Suppose to the contrary that ω is exact. Then, $[\omega^n] = 0$, so $\int_X \omega^n = 0$. But this is a contradiction with the fact that ω^n is a volume form in X. □

Therefore, if X is a compact manifold admitting a symplectic form, it has non zero second de Rham cohomology group. In particular, no S^n with $n \neq 2$ admits a symplectic structure. We see now that S^2 does admit a symplectic form. In fact, much more is true:

Proposition 1.2.4. *Let X be an orientable surface. Then, X admits a symplectic form.*

Proof. Since X is orientable, X has an area form dA, that is, a nowhere vanishing 2-form. We claim that $\omega = dA$ is a symplectic form. Indeed, dA_p is non-degenerate by definition of area form, and it is closed because $d\omega \in \Omega^3(X) = 0$, since X is a surface. □

The main example of a non compact symplectic manifold is the following:

Example 1.2.5. Consider $X = \mathbb{R}^{2n}$ with coordinates $x_1, ..., x_n, y_1, ..., y_n$. Define the 2-form $\omega_0 = \sum_{i=1}^n dx_i \wedge dy_i$. Then ω_0 is clearly a symplectic form. We will call ω_0 the standard symplectic form on $\mathbb{R}^{2n}$.

We will now state a deep fact that completely determines the local structure of symplectic manifolds.

Definition 1.2.6. *Let (X, ω) and (X', ω') be symplectic manifolds. A map $\phi : X \longrightarrow X'$ is a symplectomorphism if it is a diffeomorphism and $\phi^*(\omega') = \omega$ (that is, $\omega_p(u, v) = \omega_{\phi(p)}(d\phi_p(u), d\phi_p(v))$ for all $u, v \in T_pX$).*

Theorem 1.2.7 (Darboux)**.** *Let (X, ω) be a symplectic manifold of dimension $2n$. For any $p \in X$ there exists a neighbourhood U of p such that $(U, \omega|_U)$ is symplectomorphic to an open subset of $\mathbb{R}^{2n}$ with its standard symplectic form. In particular, there are local coordinates $x_1, ..., x_n, y_1, ..., y_n$ near p such that $\omega = \sum_{i=1}^n dx_i \wedge dy_i$.*

The proof is not difficult, using the formula of Cartan, $\mathcal{L}_X = \iota_X d + d\iota_X$. For a proof, see any book on symplectic geometry, like [M-S2] or [Sil].

Observe that Darboux's theorem tells us that all symplectic manifolds of the same dimension are locally the same, that is, there are no local symplectic invariants. This is a sharp contrast with the analogous case of Riemannian manifolds, since there we have the curvature as a local invariant. In consequence, all symplectic invariants must be global. Towards the end of this thesis, we will define a symplectic invariant, but in order to show that the definition makes sense, we will have to use the machinery of pseudoholomorphic curves, which are global objects defined in a symplectic manifold.

We finish this section by defining some important classes of submanifolds of a symplectic manifold. They correspond with the special types of subspace of a symplectic vector space defined in the previous section.

Definition 1.2.8. *Let (X, ω) be a symplectic manifold, and let $Y \subset X$ be an embedded submanifold, with $i : Y \longrightarrow X$ the inclusion map. Then Y is a lagrangian (resp. isotropic, coisotropic, symplectic) submanifold, if $T_pY \subset T_pX$ is a lagrangian (resp. isotropic, coisotropic, symplectic) subspace, for each $p \in Y$.*

1.3 Complex manifolds and almost complex structures

Definition 1.3.1. *Let V be a finite dimensional vector space over $\mathbb{R}$. A complex structure on V is an automorphism $J : V \longrightarrow V$ such that $J^2 = -Id$.*

A first observation is that if V has a complex structure, then V must have even dimension. Indeed, if $n = \dim V$, from $J^2 = -Id$ we obtain $(\det J)^2 = (-1)^n$. Since $\det J$ is a real number, this is possible only if n is even.

Observe also that if V is a real vector space with $\dim V = 2n$ and V has a complex structure, then it can be seen as a complex vector space of (complex) dimension n in the following way: $(x + iy)v = xv + yJv$. Hence the name of complex structure.

Conversely, if V is a real vector space of dimension $2n$, then there exists a complex structure on V. Indeed, fix a basis $e_1, ..., e_n, f_1, ..., f_n$ and consider the linear map $J : V \longrightarrow V$ such that $J(e_j) = f_j$, $J(f_j) = -e_j$
Then, it is immediate to check that J is an automorphism and $J^2 = -Id$.

We are mainly interested in vector spaces with a symplectic structure. In this setting, we would want a complex structure in V to be compatible in some sense with the symplectic form. In this respect, there are two relevant definitions.

Definition 1.3.2. *Let (V, ω) be a symplectic vector space, and let J be a complex structure on V. We say that J is ω-tame if we have $\omega(v, Jv) > 0$ for all $v \neq 0$ in V. We say that J is ω-compatible if it is ω-tame, and in addition $\omega(Jv, Jw) = \omega(v, w)$ for all $v, w \in V$.*

We have seen before that every vector space of even dimension admits a complex structure. Refining our argument, it is easy to see that every symplectic vector space admits an ω-compatible complex structure. Just take as a basis for V that given by proposition 1.1.3 and define again $J : V \longrightarrow V$ by $J(e_j) = f_j$, $J(f_j) = -e_j$. Then, it is immediate to check that J is ω-compatible.

However, we will present another construction that gives us lots of different ω-compatible complex structures on J and that will be useful when we deal with almost complex structures on manifolds.

Proposition 1.3.3. *Let (V, ω) be a symplectic vector space. Then, there exist ω-compatible complex structures J on V.*

Proof. We start by choosing any inner product g on V. Then, observe that since g and ω are non degenerate, there is a linear map $A : V \longrightarrow V$ such that $\omega(u, v) = g(Au, v)$ for all $u, v \in V$.

This map A satisfies the following properties:

1. A is skew-symmetric with respect to g: $g(A^*u, v) = g(u, Av) = g(Av, u) = \omega(v, u) = -\omega(u, v) = g(-Au, v)$, so $A^* = -A$.
2. AA^* is symmetric: $(AA^*)^* = AA^*$
3. AA^* is positive: $g(AA^*u, u) = g(A^*u, A^*u) > 0$ if $u \neq 0$.

Since AA^* is symmetric and positive, ir diagonalizes by the spectral theorem:
$AA^* = B \operatorname{diag}(\lambda_1, ..., \lambda_{2n})B^{-1}$

Therefore, $\sqrt{AA^*}$ exists: $\sqrt{AA^*} = B\ \mathrm{diag}(\sqrt{\lambda_1}, ..., \sqrt{\lambda_{2n}})B^{-1}$. Moreover, observe that $\sqrt{AA^*}$ does not depend on B nor on the order of the eigenvuales λ_i. Indeed, we can characterize $\sqrt{AA^*}$ as the linear map that acts as $\sqrt{AA^*}(u) = \sqrt{\lambda_i}u$ if $u \in V_{\lambda_i}$ where V_{λ_i} is the eigenspace of eigenvalue λ_i of AA^*.

Defining $J = (\sqrt{AA^*})^{-1}A$, we obtain the polar decomposition of A: $A = \sqrt{AA^*}J$.

Since A^* commutes with A (because $A^* = -A$), we have that AA^* commutes with A. Therefore $\sqrt{AA^*}$ commutes also with A: if $u \in V_{\lambda_i}$ then $\sqrt{AA^*}Au = \lambda_i\sqrt{AA^*}u = \lambda_i\sqrt{\lambda_i}u$ and $A\sqrt{AA^*}u = \sqrt{\lambda_i}Au = \sqrt{\lambda_i}\lambda_i u$. Hence, A commutes with J.

Finally, observe that J is orthogonal

$$JJ^* = (\sqrt{AA^*})^{-1}AA^*(\sqrt{AA^*})^{-1} = (\sqrt{AA^*})^{-2}AA^* = Id$$

and skew-adjoint

$$\begin{aligned} J^* &= J^{-1} = A^{-1}\sqrt{AA^*} = A^{-1}(\sqrt{AA^*})^{-1}AA^* = (\sqrt{AA^*})^{-1}A^* \\ &= -(\sqrt{AA^*})^{-1}A = -J \end{aligned}$$

With these properties, we can check that J is our wanted ω-compatible complex structure:

1. $J^2 = -J(-J) = -JJ^* = -Id$
2. $\omega(u, Ju) = g(Au, Ju) = g(-JAu, u) = g(\sqrt{AA^*}u, u) > 0$
3. $\omega(Ju, Jv) = g(AJu, Jv) = g(JAu, Jv) = g(Au, J^*Jv) = g(Au, v) = \omega(u, v)$

This finishes the proof.

□

A very useful remark about the previous proof is that the almost complex structure J is obtained in a canonical way once we fix the inner product g.

To finish with the linear theory, observe that if we have a symplectic vector space with a complex structure, then we get an inner product.

Proposition 1.3.4. *Let (V, ω, J) be a symplectic vector space together with an ω-compatible complex structure. Then, $g(u, v) = \omega(Ju, v)$ is an inner product in V.*

Proof. That g is a bilinear map is clear since ω is bilinear and J is linear. g is positive definite because $g(u, u) = \omega(Ju, u) > 0$ by the ω-tameness. Finally, g is symmetric since $g(u, v) = \omega(Ju, v) = -\omega(v, Ju) = -\omega(-JJv, Ju) = \omega(JJv, Ju) = \omega(Jv, u) = g(v, u)$, where we have used the ω-compatibility. □

We will extend now our results on complex structures to manifolds. The manifold analog of a complex structure is just a collection of complex structures in each tangent space which varies smoothly from point to point.

Definition 1.3.5. *Let X be a smooth manifold. An almost complex structure in X is an automorphism of the tangent bundle of X, $J : TX \longrightarrow TX$ such that $J^2 = -Id$.*

A smooth manifold X together with an almost complex structure J is called an almost complex manifold.

Observe that then, an almost complex structure J in X induce a complex structure in each tangent space.

Let us study how almost complex structures behave under diffeomorphisms.

Definition 1.3.6. *Let (X, J) be an almost complex manifold. Let Y be a smooth manifold, and let $\psi : X \longrightarrow Y$ a diffeomorphism. Then we define $\psi_*(J) = d\psi \circ J \circ (d\psi)^{-1}$.*

Since $d\psi$ is an isomorphism of tangent bundles, it is clear that $\psi_*(J)$ is an almost complex structure in Y. Observe that we can use this to express an almost complex structure J on a manifold X in local coordinates: if (U, ξ) is a chart of X, then $\xi_*(J|_U)$ is an almost complex structure on the open set $\xi(U) \subset \mathbb{R}^{2n}$.

As before, if (X, ω) is a symplectic manifold, then we are interested in the almost complex structures satisfying some kind of compatibility condition with the symplectic structure.

Definition 1.3.7. *Let (X, ω) be a symplectic manifold, and let J be an almost complex structure on X. We say that J is ω-tame if we have $\omega_p(v, Jv) > 0$ for all $v \neq 0$ in T_pX and for all $p \in X$. We say that J is ω-compatible if it is ω-tame, and in addition $\omega(Jv, Jw) = \omega_p(v, w)$ for all $v, w \in T_pX$ and all $p \in X$.*

We will show now that any symplectic manifold X admits ω-compatible almost complex structures.

Proposition 1.3.8. *Let (X, ω) be a symplectic manifold. Then, there exist ω-compatible almost complex structures on X.*

Proof. Choose a Riemannian metric g on X. Now, we can apply at each tangent space T_pX the construction from proposition 1.3.3 to the inner product g_p in order to obtain an ω_p-compatible complex structure J_p on T_pX. Define J as $J(p) = J_p$. The only thing that is left is to check that J is smooth. But this follows from the fact that g is smooth and the construction from 1.3.3 is canonical. Following the proof of that result, we can therefore check that J_p depends smoothly on p. □

As in the linear case, a symplectic manifold together with an ω-compatible almost complex structure determines a Riemannian metric on X.

Proposition 1.3.9. *Let (X, ω, J) be a symplectic manifold together with an ω-compatible almost complex structure. Then, g defined by $g_p(u, v) = \omega_p(J_pu, v)$ for $u, v \in T_pX$ is a Riemannian metric on X.*

Proof. The proof is the same as in the linear case, except that here we have to check the smoothness of g. But this is immediate since both ω and J are smooth. □

We will denote the metric in this proposition by g_J.

We will now introduce the concept of a complex manifold and study its relations with almost complex structures.

Definition 1.3.10. *Let X be a topological manifold. A complex structure on X is a collection of charts $(U_i, \xi_i)_{i\in I}$, where U_i are open sets in X and $\xi_i : U_i \longrightarrow V \subset \mathbb{C}^n$ are homeomorphisms from U_i onto an open subset of $\mathbb{C}^n$, such that $\bigcup_{i\in I} U_i = X$ and for all $i, j \in I$ such that $U_i \cap U_j \neq \emptyset$ we have that $\xi_i \circ \xi_j^{-1}$ is a holomorphic map.*

A manifold X endowed with such a complex structure is called a complex manifold of (complex) dimension n.

Observe that in particular, identifying $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ we have that a complex manifold of dimension n is a smooth manifold of dimension $2n$. Indeed, under this identification we have for a chart $\xi(p) = (z_1(p), ..., z_n(p)) = (x_1(p), y_1(p), ..., x_n(p), y_n(p))$, where $z_i : U \longrightarrow \mathbb{C}$ are complex functions on the domain of the chart, and $x_i, y_i : U \longrightarrow \mathbb{R}$ are real functions. As in the real case, $(z_1, ..., z_n)$ are called local coordinates near p if p is in the domain of the chart. The relation between z_i and the pair x_i, y_i is given by $z_i = x_i + iy_i$.

In the same way as smooth manifolds can be considered as the appropriate setting for doing calculus in the large, complex manifolds are the appropriate setting for doing complex analysis in the large. In particular, we can extend the notion of holomorphic map to maps between complex manifolds.

Definition 1.3.11. *Let X, Y be two complex manifolds. A map $f : X \longrightarrow Y$ is said to be holomorphic at $p \in X$ if there are charts (U, ξ) near p and (V, ψ) near $f(p)$ such that $\psi \circ f \circ \xi^{-1}$ is holomorphic at $\xi(p)$. f is said to be holomorphic if it is holomorphic at every point of X.*

Observe that the previous definition is independent of the chosen charts because of the fact that the transition maps in complex manifolds are holomorphic.

In a complex manifold we can define several notions of tangent space at a point. Fix $p \in X$, and fix local coordinates $(z_1, ..., z_n)$ near p.

First of all, we can consider the tangent space of X at p as a $2n$-dimensional smooth manifold, which we denote by $T_{\mathbb{R},p}X$. This is a real vector space of dimension $2n$ with basis $\{\frac{\partial}{\partial x_1}, \frac{\partial}{\partial y_1}, ..., \frac{\partial}{\partial x_n}, \frac{\partial}{\partial y_n}\}$.

Secondly, we can consider the complexified tangent space $T_{\mathbb{C},p}X = T_{\mathbb{R},p}X$, which is a complex vector space of complex dimension $2n$, with basis
$\{\frac{\partial}{\partial x_1}, \frac{\partial}{\partial y_1}, ..., \frac{\partial}{\partial x_n}, \frac{\partial}{\partial y_n}\}$.
Equivalently, we can consider the basis $\{\frac{\partial}{\partial z_1}, \frac{\partial}{\partial \overline{z_1}}, ..., \frac{\partial}{\partial z_n}, \frac{\partial}{\partial \overline{z_n}}\}$.

Finally, we can consider the holomorphic tangent space T'_pX, which is a complex vector space of complex dimension n, with basis $\{\frac{\partial}{\partial z_1}, ..., \frac{\partial}{\partial z_n}\}$. This is the natural tangent space to consider when we consider X as a complex manifold.

A basic topological property of complex manifolds is that, like symplectic manifolds, they are always orientable.

Proposition 1.3.12. *Let X be a complex manifold. Then X is orientable.*

Proof. It suffices to check that if $(U, \xi), (V, \xi')$ are two complex charts, then the jacobian of the change $\xi' \circ \xi^{-1}$ is always positive. Observe that, since $\xi' \circ \xi^{-1}$ is a holomorphic function, it satisfies the Cauchy-Riemann equations. Using the basis $\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n}, \frac{\partial}{\partial y_1}, ..., \frac{\partial}{\partial y_n}$, writing $\xi' \circ \xi^{-1} = (\psi_1, ..., \psi_n) = (u_1, ..., u_n, v_1, ..., v_n)$, where $\psi_i = u_i + iv_i$, and using the Cauchy-Riemann equations $\frac{\partial v_i}{\partial y_j} = \frac{\partial u_i}{\partial x_j}$ and $\frac{\partial v_i}{\partial x_j} = -\frac{\partial u_i}{\partial y_j}$ to express everything in terms of x, we get that the jacobian of the change of variables is a sum of squares, hence positive. □

The following proposition shows that every complex manifold carries a natural almost complex structure.

Proposition 1.3.13. *Let X be a complex manifold. Then there exists a canonical choice of almost complex structure J on X.*

Proof. Fix some chart $U \subset X$ with local coordinates $(z_1, ..., z_n)$ in X, and put $z_i = x_i + iy_i$. Then, for $p \in U$ we can consider the almost complex structure defined by $J(p)(\frac{\partial}{\partial x_i}) = \frac{\partial}{\partial y_i}$ and $I(p)(\frac{\partial}{\partial y_i}) = -\frac{\partial}{\partial x_i}$. We will check that $J(p)$ is independent of the choice of local coordinates near p. Indeed, let $(w_1, ..., w_n)$ be another set of local coordinates and put $w_i = u_i + iv_i$. Let $h = f + ig$ the change of coordinates from the w_i's to the z_i's. Recall that, since X is a complex manifold, h is holomorphic. Then we have that $\frac{\partial}{\partial u_i} = \sum_j \frac{\partial f_j}{\partial u_i}\frac{\partial}{\partial x_j} + \frac{\partial g_j}{\partial u_i}\frac{\partial}{\partial y_j}$ and $\frac{\partial}{\partial v_i} = \sum_j \frac{\partial f_j}{\partial v_i}\frac{\partial}{\partial x_j} + \frac{\partial g_j}{\partial v_i}\frac{\partial}{\partial y_j}$. Since h is holomorphic, it satisfies the Cauchy-Riemann equations, hence $\frac{\partial f_j}{\partial u_i} = \frac{\partial g_j}{\partial v_i}$ and $\frac{\partial f_j}{\partial v_i} = -\frac{\partial g_j}{\partial u_i}$.

So finally we have:

$$\frac{\partial}{\partial u_i} = \sum_j \frac{\partial f_j}{\partial u_i}\frac{\partial}{\partial x_j} - \frac{\partial f_j}{\partial v_i}\frac{\partial}{\partial y_j}$$

$$\frac{\partial}{\partial v_i} = \sum_j \frac{\partial f_j}{\partial v_i}\frac{\partial}{\partial x_j} + \frac{\partial f_j}{\partial u_i}\frac{\partial}{\partial y_j}$$

Therefore $J(p)((\frac{\partial}{\partial u_i}) = \sum_j \frac{\partial f_j}{\partial u_i} J(p)(\frac{\partial}{\partial x_j}) - \frac{\partial f_j}{\partial v_i} J(p)(\frac{\partial}{\partial y_j}) = \sum_j \frac{\partial f_j}{\partial u_i}\frac{\partial}{\partial y_j} + \frac{\partial f_j}{\partial v_i}\frac{\partial}{\partial x_j} = \frac{\partial}{\partial v_i}$ and similarly $J(p)(\frac{\partial}{\partial v_i}) = -\frac{\partial}{\partial u_i}$.

This shows that the definition of J is independent of the local coordinates, and since it is clearly smooth in each chart, this shows that J is an almost complex structure on X. □

Note that the canonical almost structure of a complex manifold X acts as a rotation of $\pi/2$ in each plane $\mathbb{R}\{\frac{\partial}{\partial x_i}, \frac{\partial}{\partial y_i}\}$, or equivalently as the multiplication by i in this planes. Hence, the name almost complex structure comes because it gives a notion of multiplication by i in a smooth manifold.

Now we can consider the inverse problem, that is, given an almost complex structure J on a smooth manifold X, there exists a complex structure on X such that J is the canonical almost complex structure induced by that complex structure?

Definition 1.3.14. *Let X be a smooth manifold and let J be an almost complex structure on X. J is said to be integrable if it is the canonical almost complex structure for some complex structure on X.*

In general, not all almost complex structures J are integrable. In fact, necessary and sufficient conditions for integrability were given by Newlander and Nirenberg.

Definition 1.3.15. *Given a map $A : TX \longrightarrow TX$, we define the Nijenhuis tensor as*

$$N_A(X,Y) = -A^2[X,Y] + A([AX,Y] + [X,AY]) - [AX,AY]$$

One can easily check that N_A is indeed a tensor.

Proposition 1.3.16 (Newlander-Nirenberg)**.** *An almost complex structure $J : TX \longrightarrow TX$ is integrable if and only if $N_J = 0$.*

Therefore we see that integrability is equivalent to the fact that J satisfies a certain differential equation. The proof of this result can be found in theorem 11.8 of [Dem].

If we assume this result we can easily show that all orientable surfaces admit a complex structure.

Proposition 1.3.17. *Let X be an orientable smooth surface. Then X admits a complex structure.*

Proof. First of all, observe that X admits an almost complex structure J. For instance, we have seen before that X admits a symplectic form, and that any symplectic manifold admits ω-compatible almost complex structures.

Now we only have to check that $N_J = 0$. Since N is a tensor, it is enough to check that $N_J(a,b) = 0$ for some basis a, b of T_pX. Let $0 \neq a \in T_pX$. Then $\{a, Ja\}$ is a basis of T_pX (otherwise, $Ja = \lambda a$ and $\omega(Ja, a) = 0$, a contradiction with the ω-tameness of J). But then we have $N_J(a, Ja) = [a, Ja] + J([Ja, Ja] + [a, JJa]) - [Ja, JJa] = [a, Ja] + J(-[a, a]) + [Ja, a] = [a, Ja] - [a, Ja] = 0$. So by the Newlander-Nirenberg theorem, J is integrable. This shows that X admits a complex structure. □

We can now give several examples of complex manifolds. The most important ones are the manifolds of complex dimension 1.

Definition 1.3.18. *A connected complex manifold of dimension 1 is called a Riemann surface.*

Example 1.3.19. $\mathbb{C}^n$ is a complex manifold of dimension n, with a complex structure given by the only chart $(\mathbb{C}^n, id)$. In particular, $\mathbb{C}$ is a Riemann surface.

Example 1.3.20. Let Σ be a compact Riemann surface. Then, since Σ is an orientable smooth manifold of dimension 2, that is, a compact orientable smooth surface, we know by the classification theorem for surfaces that it is diffeomorphic to a sphere S^2 or to a connected sum of tori, and therefore that topologically (and smoothly) they are classified according to their genus g. In fact, as we have seen in the previous proposition, one can construct Riemann surfaces of every genus, hence all compact orientable smooth surfaces admit a complex structure.

The most important examples of compact complex manifolds are the complex projective spaces, which we now describe.

Example 1.3.21. Let $\mathbb{C}P^n$ be the quotient space of $\mathbb{C}^{n+1} - \{0\}$ by the equivalence relation $(z_0, z_1, ..., z_n) \sim (w_0, w_1, .., w_n)$ if and only if there exists a $\lambda \in \mathbb{C}$, $\lambda \neq 0$, such that $z_i = \lambda w_i$ for all $i = 1, ..., n$. Alternatively, we can see it as the space of all lines in $\mathbb{C}^{n+1}$. We can give so-called homogeneous coordinates to $\mathbb{C}P^n$ by assigning to a point $p = [(z_0, ..., z_n)]$ its coordinates $[z_0, ..., z_n]$. Note that then a point p has infinitely many homogeneous coordinates, and $[z_0, ..., z_n]$ and $[w_0, ..., w_n]$ represent the same point if and only if $z_i = \lambda w_i$ for all i and some $\lambda \neq 0$.

As it is well-known from general topology, $\mathbb{C}P^n$ is a Hausdorff compact topological space. We will see now that it has the structure of a complex manifold of dimension n. First of all, define $U_i = \{p = [z_0, ..., z_n] \in \mathbb{C}P^n : z_i \neq 0\}$, and observe that it is well-defined (i.e., the condition $z_i \neq 0$ is independent of the homogeneous coordinates that we choose for p).

Clearly we have $\mathbb{C}P^n = \bigcup_{i=0}^n U_i$. We give maps $\xi_i : U_i \longrightarrow \mathbb{C}^n$ given by $\xi_i([z_0, ..., z_n]) = (\frac{z_0}{z_i}, ..., \frac{z_{i-1}}{z_i}, \frac{z_{i+1}}{z_i}, ..., \frac{z_n}{z_i})$, which are seen to be homeomorphisms. Finally, the transition maps $\xi_{ji} := \xi_j \circ \xi_i^{-1} : \xi_i(U_i \cap U_j) \longrightarrow \xi_j(U_i \cap U_j)$ are given by $\xi_{ji}(z_1, ..., z_n) = \left(\frac{z_1}{z_j}, ..., \frac{z_{i-1}}{z_j}, \frac{1}{z_j}, \frac{z_{i+1}}{z_j}, ..., \frac{z_n}{z_j}\right)$, where the i-th coordinate has been omited, which is a holomorphic map.

Therefore, $\mathbb{C}P^n$ is a complex manifold of dimension n.

Definition 1.3.22. *Let (V, ω) be a symplectic vector space. We define $\mathcal{J}(V, \omega)$ as the set of all ω-compatible complex structures on V, topologized as a subspace of $End(V)$.*

Let (X, ω) be a symplectic manifold. We denote by $\mathcal{J}(\omega)$ (or simply by $\mathcal{J}$ if ω is understood) the set of all the ω-compatible almost complex structures on X. Observe that an element of $\mathcal{J}$ is a section of the fibre bundle $\mathcal{J} \longrightarrow X$ with fibre over p, $\mathcal{J}(T_pX, \omega_p)$. Therefore, we consider $\mathcal{J}(\omega)$ as a topological space as the space of sections of the fibre bundle $\mathcal{J} \longrightarrow X$.

For the applications it will be very important the fact that $\mathcal{J}$ is path-connected. In fact, much more is true: $\mathcal{J}$ is contractible. There are several proofs of this fact. The one we give emphasizes the relation between lagrangian subspaces, almost complex structures and inner products.

Proposition 1.3.23. *$\mathcal{J}(\omega)$ is contractible.*

Proof. It is enough to check that the space $\mathcal{J}(V, \omega)$ of ω-compatible complex structures on a symplectic vector space (V, ω), given the topology as a subspace of $End(V)$, is contractible. This is because then $\mathcal{J}(\omega)$ is the space of sections of a fibre bundle with contractible fibers, hence it is also contractible.

Fix a lagrangian subspace L_0 of (V, ω). Let $\mathcal{L}(V, \omega, L_0)$ be the space of all lagrangian subspaces of (V, ω) which intersect L_0 transversally (that is, $V = L_0 + L$). Let $\mathcal{G}(L_0)$ be the space of all positive inner products on L_0, and consider the map

$$\Psi : \mathcal{J}(V, \omega) \longrightarrow \mathcal{L}(V, \omega, L_0) \times \mathcal{G}(L_0)$$

defined by $\Psi(J) = (JL_0, G_J|_{L_0})$.
First, note that Ψ is well-defined. That is, $JL_0 \in \mathcal{L}(V, \omega, L_0)$ and $G_J|_{L_0} \in \mathcal{G}(L_0)$. The second fact is immediate, since G_J is an inner product on V and the restriction of an inner product to any subspace is also an inner product. For the first fact, we need to see that JL_0 is a lagrangian subspace of V and that it intersects L_0 transversally. Observe that if $v \in L_0$ and $w \in L_0 = L_0^\omega$ we have $\omega(Jv, Jw) = \omega(v, w) = 0$, so $(JL_0)^\omega \subset JL_0$. Since J is an automorphism of V, we have

$\dim JL_0 = \dim L_0 = n/2$, so JL_0 is indeed a lagrangian subspace of V. Let us see that it intersects L_0 transversally. Since $\dim L_0 = \dim JL_0 = n/2$ it is enough to see that $L_0 \cap JL_0 = \{0\}$. Let $u \in L_0 \cap JL_0$. Then, $u = Jv$ for some $v \in L_0$. So $g_J(u,u) = \omega(u, Ju) = \omega(u, JJv) = -\omega(u,v) = 0$ because $L_0 = L_0^\omega$. Since g_J is an inner product, it follows that $u = 0$.

Now, we see that Ψ is bijective. Let $(L, G) \in \mathcal{L}(V, \omega, L_0) \times \mathcal{G}(L_0)$. We define J as follows. For $v \in L_0$, $v^\perp = \{u \in L_0 : G(u,v) = 0\}$ is an $n-1$-dimensional subspace of L_0. Its symplectic complement, $(v^\perp)^\omega$ is therefore $(n+1)$-dimensional. Observe that $(v^\perp)^\omega \cap L$ is 1-dimensional. Indeed, $L \cap L_0 = \{0\}$ and $L_0 \subset (v^\perp)^\omega$ (because $L_0 = L_0^\omega$ and $v^\perp \subset L_0$), which together with the fact that $(v^\perp)^\omega$ is $n+1$-dimensional, implies that $(v^\perp)^\omega \cap L$ is 1-dimensional. Let $Jv \in (v^\perp)^\omega \cap L$ be the unique vector satisfying $\omega(v, Jv) = 1$. Take now $v_1, ..., v_n$ a G-orthonormal basis of L_0, and let $Jv_1, ..., Jv_n$ constructed as above. Note that $Jv_1, ..., Jv_n$ are linearly independent, as follows from the fact that $\omega(v_i, Jv_j) = \delta_{ij}$.

We claim that this defines the wanted J such that $\Psi(J) = (L, G)$. Since $Jv_i \in L$ and $L \cap L_0 = \{0\}$, $v_1, ..., v_n, Jv_1, ..., Jv_n$ is a basis of V and J is completely determined by defining $J(Jv_i) = -v_i$. It is clear in this way that J is an automorphism of V satisfying $J^2 = -1$, that is, J is an almost complex structure on V. We have $\omega(v_i, Jv_i) = 1$ and $\omega(Jv_i, JJv_i) = \omega(Jv_i, -v_i) = \omega(v_i, Jv_i) = 1$, so J is ω-tame. Moreover, $\omega(Jv_i, Jv_j) = 0 = \omega(v_i, v_j)$ (because $Jv_i, Jv_j \in L = L^\omega$), $\omega(JJv_i, Jv_j) = -\omega(v_i, Jv_j) = -\delta_{ij} = \omega(Jv_i, v_j)$, and $\omega(JJv_i, JJv_j) = \omega(v_i, v_j) = \omega(Jv_i, Jv_j)$, so J is in fact ω-compatible.

Since $L \cap L_0 = \{0\}$ and using the fact that $v_1, ..., v_n$ is a basis of L_0 and $Jv_1, ..., Jv_n$ a basis of L, it follows that $L = JL_0$. Finally, we must check that $G_J|_{L_0} = G$. Indeed, $G(v_i, v_j) = \delta_{ij}$, while $G_J|_{L_0}(v_i, v_j) = \omega(v_i, Jv_j) = \delta_{ij}$. So Ψ is exhaustive.

In order to see that Ψ is injective, suppose that $\Psi(J) = \Psi(J')$. Then, $JL_0 = J'L_0 =: L$ and $G_J|_{L_0} = G_{J'}|_{L_0} =: G$. Take $v_1, ...v_n$ an orthonormal basis of L_0 with respect to G. Then, using the same notation as before, we have $v_i^\perp =< v_1, ..., v_{i-1}, v_{i+1}, ..., v_n >$ and $(v_i^\perp)^\omega =< v_1, ..., v_{i-1}, v_{i+1}, ..., v_n, u_i >$ with $u_i \in L$. Since $Jv_i \in L$, $\omega(v_i, Jv_i) = 1$ and $G(v_j, Jv_i) = \omega(v_j, JJv_i) = -\omega(v_j, v_i) = 0$ for all $j \neq i$, it follows that Jv_i is the only vector in $(v_i^\perp)^\omega \cap L$ satisfying $\omega(v_i, Jv_i) = 1$, so $J = J'$.

Observe that $\mathcal{G}(L_0)$ is contractible (in fact, it is convex since $tG_1 + (1-t)G_2$ is still an inner product on L_0 for $t \in [0,1]$).
$\mathcal{L}(V, \omega, L_0)$ is also contractible. In order to see this, we observe that $\mathcal{L}(V, \omega, L_0)$ can be identified with the vector space of all symmetric $n \times n$ matrices, which is convex and hence contractible. Fix an ω-compatible almost complex structure J. Observe that if L is an n-dimensional subspace of V transversal to L_0, we have that L is the graph of a linear map $S : JL_0 \longrightarrow L_0$ where $L_0 =< v_1, ..., v_n >$ and $L =< Jv_1 + SJv_1, ..., Jv_n + SJv_n >$. Observe that in these basis, the linear map S is expressed as a symmetric $n \times n$ matrix A. Indeed, we have $A_{ij} = \omega(SJv_i, Jv_j)$. Then, using that L_0 and L are Lagrangian, we have that $0 = \omega(Jv_i + SJv_i, Jv_j + SJv_j) = \omega(Jv_i, Jv_j) + \omega(Jv_i, SJv_j) + \omega(SJv_i, Jv_j) + \omega(SJv_i + SJv_j) = -\omega(SJv_j, Jv_i) + \omega(SJv_i, Jv_j)$, so $A_{ij} = \omega(SJv_i, Jv_j) = \omega(SJv_j, v_i) = A_{ji}$ and A is symmetric. Conversely, we can see similarly that if A is symmetric, then L is Lagrangian.

To finish, just observe that since Ψ is a bijection and we have just seen that $\mathcal{L}(V, \omega, L_0) \times \mathcal{G}(L_0)$ is contractible, we have that $\mathcal{J}(V, \omega)$ is contractible, as wanted.

$\square$

1.4 Kaehler manifolds

The most important class of symplectic manifolds is the class of Kähler manifolds. In this section we will define them and give some examples of Kähler manifolds and of non-Kähler symplectic manifolds.

Definition 1.4.1. *A symplectic manifold (X, ω, J) with an ω-compatible almost complex structure J is said to be a Kähler manifold if J is integrable. In this situation, the symplectic form ω is called a Kähler form.*

There is an equivalent definition coming from complex geometry as a complex manifold equipped with an Hermitian metric.

Definition 1.4.2. *Let X be a complex manifold. An Hermitian metric h in X is a map $h : TX \times TX \longrightarrow TX$ such that h_p is a nondegenerate Hermitian form in T_pX, that is, if we write in local coordinates $h = \sum_{i,j} h_{ij} dz^i \otimes d\overline{z}^j$, h_{ij} is a positive definite Hermitian matrix.*

A complex manifold X endowed with an Hermitian metric is called a Hermitian manifold.

Observe that in an Hermitian manifold, we have in a natural way a 2-form and a Riemannian metric. Indeed, the Hermitian form is a 2-form defined as $\omega := \frac{i}{2}(h - \overline{h})$, and we have the Riemannian metric $g = \frac{1}{2}(h + \overline{h})$. In this way, we can express the Hermitian metric as $h = g - i\omega$.

Definition 1.4.3. *A complex manifold X together with a Hermitian metric H is called a Kähler manifold if its Hermitian form ω is closed.*

Proposition 1.4.4. *The two definitions of Kähler manifold coincide.*

Proof. (Sketch). We only indicate the relation between the Kähler form and the hermitian metric.

It can be shown that if h is an Hermitian metric on X its Hermitian form ω is symplectic, and if J is the canonical almost complex structure, J is ω-compatible.

Conversely, if J is an integrable almost complex structure in X and ω is a symplectic form such that J is ω-compatible, then it is can be checked that $h = g_J + i\omega$ is an Hermitian metric, and obviously it has a closed Hermitian form. □

A very usual way to give a Kähler structure on a manifold is by means of a Kähler potential. Let us describe it.

Definition 1.4.5. *Let X be a complex manifold. A function $\rho : X \longrightarrow \mathbb{R}$ is strictly plurisubharmonic (spsh) if, on each complex chart U with local coordinates $z_1, ..., z_n$, the matrix $\left(\frac{\partial^2 \rho}{\partial z_i \partial \overline{z}_j}(p)\right)_{ij}$ is positive definite for all $p \in U$.*

Proposition 1.4.6. *Let X be a complex manifold and $\rho : X \longrightarrow \mathbb{R}$ a spsh function. Then*

$$\omega = \frac{i}{2} \partial \overline{\partial} \rho$$

is a Kähler form.

Such a function ρ is then called a Kähler potential.

Locally, we also have the converse. That is, in a Kähler manifold, the Kähler form always derive locally from a Kähler potential.

Theorem 1.4.7. *Let ω be a Kähler form of a Kähler manifold X and let $p \in X$. Then there exist a neighbourhood U of p such that, on U,*

$$\omega = \frac{i}{2}\partial\overline{\partial}\rho$$

With this fact, we can prove that complex submanifolds of a Kähler manifold are also Kähler manifolds.

Proposition 1.4.8. *Let X be a complex manifold, $\rho : X \longrightarrow \mathbb{R}$ spsh and M a complex submanifold of X (that is, the inclusion $i : M \longrightarrow X$ is a holomorphic embedding). Then $i^*\rho$ is spsh.*

Proof. Take a chart $(U, z_1, ..., z_n)$ centered at p and adapted to M, that is, such that $U \cap M$ is given by $z_1 = ... = z_m = 0$ (where $m = \dim M$. Then, we have that $i^*\rho$ is spsh if and only if the matrix $\frac{\partial^2\rho}{\partial z_i \partial \overline{z}_j}(0, ..., 0, z_{m+1}, ..., z_n)_{i=m+1,...,n;j=m+1,...,n}$ is positive definite, but this is a minor of the matrix $\frac{\partial^2\rho}{\partial z_i \partial \overline{z}_j}(p))_{ij}$, which is positive definite by hypothesis, so it is also positive definite. $\square$

Corollary 1.4.9. *Any complex submanifold of a Kähler manifold is also Kähler.*

Now we can give lots of examples of Kähler (and hence, symplectic) manifolds.

Example 1.4.10. Consider in $\mathbb{C}^n$ the function $\rho(z) = |z|^2 = z\overline{z}$. This function is real valued and it is spsh, since $\frac{\partial^2\rho}{\partial z_i \partial \overline{z}_j}(z)_{ij} = \delta_{ij}$. Therefore, it is a Kähler potential and $\mathbb{C}^n$ is a Kähler manifold. By the last corollary, all complex submanifolds of $\mathbb{C}^n$ are also Kähler manifolds.

Example 1.4.11. Let us return to $\mathbb{C}P^n$. We have seen that it is a complex manifold. We claim that it is also a Kähler manifold. Indeed, we can define a Kähler potential in $\mathbb{C}^n$ by $\rho_{FS} := log(|z|^2 + 1)$. Then, putting $\omega_{FS}^{\mathbb{C}^n} = \frac{i}{2}\partial\overline{\partial}\rho$, we can write $\omega_{FS}|_{U_i} = \phi_i^*(\omega_{FS}^{\mathbb{C}^n})$ where $\phi_i : U_i \longrightarrow \mathbb{C}^n$ is the usual chart map. One can check that ω_{FS} is then a well-defined Kähler form in $\mathbb{C}P^n$ (that is, $\phi_i^*(\omega_{FS}^{\mathbb{C}^n}) = \phi_j^*(\omega_{FS}^{\mathbb{C}^n})$, called the Fubini-Study Kähler form. Therefore, $\mathbb{C}P^n$ is a Kähler manifold, as are all its complex submanifolds.

Example 1.4.12. [Smooth projective varieties]
All smooth projective algebraic varieties (the locus of zeros of a set of homogeneous polynomials in $\mathbb{C}P^n$) are complex submanifolds of $\mathbb{C}P^n$ (as can be easily seen by an application of the holomorphic inverse function theorem), and hence, a Kähler manifold. In particular, smooth projective curves are Kähler manifolds. Curves have complex dimension 1, and hence they are Riemann surfaces. Since we already know that Riemann surfaces are symplectic manifolds, this gives another proof of the fact that smooth projective curves are symplectic manifolds.

Kähler manifolds have been studied intensively, and there are plenty of interesting results about them. Here we will only quote one result, restricting the class of compact manifolds that admit a Kähler structure, which can be proved by means of Hodge theory. Recall that the Betti numbers of X are defined as

$$b^k(X) := \dim H^k_{deR}(X)$$

Theorem 1.4.13. *If X is a compact Kähler manifold, then the odd Betti numbers b^{2k+1} are even.*

1.5 Counterexamples

In this chapter we have introduced several structures on manifolds: we have seen symplectic, almost complex, complex and Kähler manifolds. We have pointed out several relations between them, that we summarize in the following proposition

Proposition 1.5.1. *We have the following*

(1) Any Kähler manifold is a complex manifold.

(2) Any Kähler manifold is a symplectic manifold.

(3) Any complex manifold is an almost complex manifold.

(4) Any symplectic manifold can be given the structure of an almost complex manifold.

We will give now examples (without proof that they satisfy the desired properties) showing that neither of the above implications can be reversed.

The following example is due to Thurston.

Example 1.5.2. [A symplectic and complex manifold that is not Kähler] Consider $\mathbb{R}^4$ with the usual symplectic form $\omega_0 = dx_1 \wedge dy_1 + dx_2 \wedge dy_2$. Let Γ be the discrete group of symplectomorphisms generated by the following symplectomorphisms:

$$\gamma_1(x_1, x_2, y_1, y_2) = (x_1, x_2 + 1, y_1, y_2)$$

$$\gamma_1(x_1, x_2, y_1, y_2) = (x_1, x_2, y_1, y_2 + 1)$$

$$\gamma_1(x_1, x_2, y_1, y_2) = (x_1 + 1, x_2, y_1, y_2)$$

$$\gamma_1(x_1, x_2, y_1, y_2) = (x_1, x_2 + y_2, y_1 + 1, y_2)$$

Consider now the compact manifold $X = \mathbb{R}^4/\Gamma$. It is symplectic with the symplectic form induced from ω_0, and it can be shown that it admits a complex structure. However, it cannot be Kähler, since $\pi_1(X) = \Gamma$, so $H^1(X; \mathbb{Z}) = \Gamma/[\Gamma, \Gamma]$ which has rank 3. Therefore, $b_1 = 3$, which implies that it cannot be Kähler as seen in theorem 1.4.13.

The following example is due to Hopf.

Example 1.5.3. [A complex manifold that does not admit any symplectic structure] Consider the compact manifold $S^1 \times S^3$, called the Hopf surface. It is not symplectic, since $H^2(S^1 \times S^3) = 0$. However, it is complex, since $S^1 \times S^3 \cong (\mathbb{C}^2 - \{0\})/\Gamma$, where $\Gamma = \{2^n id : n \in \mathbb{Z}\}$ is a discrete group of holomorphic transformations.

Example 1.5.4. [There are almost complex manifolds that are neither complex nor symplectic] Consider $\mathbb{C}P^2 \sharp \mathbb{C}P^2 \sharp \mathbb{C}P^2$. It can be shown that this manifold admits an almost complex structure, but it is neither complex nor symplectic (proved by Taubes).

There are also examples of symplectic manifold that does not admit any complex structure. Fernández-Gotay-Gray proved the existence of such manifolds. Their examples are circle bundles over circle bundles over a 2-torus. However we will not discuss them further here.

Chapter 2

Moduli spaces of pseudoholomorphic curves

In this chapter we introduce the main object of this thesis: pseudoholomorphic curve. After defining them, we prove some of its basic properties. In the second section, we show how they form moduli spaces and explain briefly how these moduli spaces have the structure of a finite dimensional manifold, for most situations. In the last section, we show by means of an example that these moduli spaces fail to be compact in general, and we will see how this failure of compactness occurs, introducing the phenomenon of bubbling.

2.1 Pseudoholomorphic curves

In all this section, let (X, ω) be a symplectic manifold, and let J be an ω-compatible almost complex structure on X. Moreover, denote by Σ a Riemann surface, by j the canonical almost complex structure on Σ and by $[\Sigma] \in H_2(\Sigma)$ the canonical choice of orientation induced by the canonical orientation of $\mathbb{C}$ via any holomorphic chart (since change of coordinates between holomorphic charts are orientation preserving, this orientation is well-defined). We also equip Σ with a volume form $dvol_\Sigma$. Then, the Riemann surface Σ carries a natural metric determined by j and $dvol_\Sigma$ as follows. Pick local coordinates $z = s + it$ near a point $p \in \Sigma$. Then, $\frac{\partial}{\partial s}$ and $\frac{\partial}{\partial t} = j\left(\frac{\partial}{\partial s}\right)$ form a local frame of $T\Sigma$. By rescaling the coordinates (using λz instead of z for $\lambda \in \mathbb{C} - \{0\}$), we may assume that $dvol_\Sigma = ds \wedge dt$. We then define a metric g on Σ by specifying that $\frac{\partial}{\partial s}$ and $\frac{\partial}{\partial t}$ form an orthonormal local frame. This gives as a well-defined metric, since using the fact that the change of coordinates is holomorphic one can check that the definition is independent of the chosen coordinates.

Definition 2.1.1. *A smooth map $u : \Sigma \longrightarrow X$ is a J-holomorphic curve (or pseudoholomorphic curve if J is understood) if it satisfies $du \circ j = J \circ du$.*

Observe that $du \circ j = J \circ du$ is equivalent to

$$du + J \circ du \circ j = 0$$

as is easily seen composing with J and using $J^2 = -id$.

In order to handle pseudoholomorphic curves it will be very useful to introduce the following two operators, adapted at the almost complex structure J.

Definition 2.1.2. *We define the operators ∂_J and $\bar{\partial}_J$ by:*

$$\partial_J u = \frac{1}{2}(du - J \circ du \circ j)$$

$$\bar{\partial}_J u = \frac{1}{2}(du + J \circ du \circ j)$$

Therefore, the pseudoholomorphicity condition can be expressed as $\bar{\partial}_J u = 0$. Observe also that for each u, $\partial_J u$ and $\bar{\partial}_J u$ are 1-forms.

Now we are going to express the pseudoholomorphicity condition as a differential equation in terms of local coordinates. In order to get a good expression, we will use only holomorphic charts for Σ, that is, charts in the complex structure of Σ. Recall that if $z = s + it$ are such local coordinates, we have that $j(\frac{\partial}{\partial s}) = \frac{\partial}{\partial t}$ and $j(\frac{\partial}{\partial t}) = -\frac{\partial}{\partial s}$. Such local coordinates are also called conformal.

Proposition 2.1.3. *Let (U, ϕ_α) be a chart in Σ with $z = s + it$ as conformal local coordinates and put $u_\alpha = u \circ \phi_\alpha$. In this coordinates, the condition $du \circ j = J \circ du$ can be expressed as*

$$\partial_s u_\alpha + J(u_\alpha)\partial_t u_\alpha = 0$$

where as usual we still call J the expression of the almost complex structure on X in a given chart.

Proof. Using the local coordinates, and recalling that in these coordinates $j(\frac{\partial}{\partial s}) = \frac{\partial}{\partial t}$ and $j(\frac{\partial}{\partial t}) = -\frac{\partial}{\partial s}$ we have:

$$du\left(\frac{\partial}{\partial s}\right) + Jdu\left(j\left(\frac{\partial}{\partial s}\right)\right) = \partial_s u_\alpha + (J \circ u_\alpha)\partial_t u_\alpha$$

$$du\left(\frac{\partial}{\partial t}\right) + Jdu\left(j\left(\frac{\partial}{\partial t}\right)\right) = \partial_t u_\alpha - (J \circ u_\alpha)\partial_s u_\alpha$$

Therefore, in such a chart, the condition to be pseudoholomorphic reduces to the following two differential equations:

$$\partial_s u_\alpha + (J \circ u_\alpha)\partial_t u_\alpha = 0$$

$$\partial_t u_\alpha - (J \circ u_\alpha)\partial_s u_\alpha = 0$$

Now just observe that they are in fact the same equation, as seen composing one of them with $J \circ U_\alpha$, so the result follows. □

With this, we can see where the name pseudoholomorphic comes from. Assume that X is a complex manifold and J is its canonical almost complex structure. In this situation, choosing conformal coordinates for X also, so that $J = J_0$ in these coordinates (where J_0 is the canonical almost complex structure on $\mathbb{C}^n$), the condition of pseudoholomorphicity in local coordinates reduce to the usual Cauchy-Riemann equations, and thus say that the map is holomorphic.

The above proof also shows

Proposition 2.1.4. *Let (U, ϕ_α) be a chart in Σ with $z = s + it$ as conformal local coordinates and put $u_\alpha = u \circ \phi_\alpha$. In this coordinates*

$$\bar{\partial}_J u = \frac{1}{2}(\partial_s u_\alpha + J(u_\alpha)\partial_t u_\alpha)du + \frac{1}{2}(\partial_t u_\alpha - J(u_\alpha)\partial_s u_\alpha)dt$$

In general, the equations of the previous proposition are called non linear Cauchy-Riemann equations. They are non linear but elliptic, since its linearization are clearly the usual Cauchy-Riemann equations, which are well-known to be elliptic.

An essential property of a pseudoholomorphic curve is its energy, which we now define and give some properties.

Definition 2.1.5. *Let $u : \Sigma \longrightarrow X$ be a smooth map. Its energy is*

$$E(u) = \frac{1}{2}\int_\Sigma |du|_J^2 dvol_\Sigma$$

where $|du|_J$ means the norm of du with respect to the metric g_J in X.

We will also use the notation $E(u; A) := \frac{1}{2}\int_A |du|_J^2 dvol_\Sigma$, for $A \subset \Sigma$. We will call energy density to $|du|_J$.

Observe then that, by definition, the energy is just one half of the square of the L^2 norm of du with respect to the metric g_J.

A trivial remark that we will use several times is that u is constant if and only if $E(u) = 0$.

We will see now that the energy of a pseudoholomorphic curve is a topological invariant if J is ω-compatible.

Definition 2.1.6. *Let $u : \Sigma \longrightarrow X$ be a pseudoholomorphic curve, and let $[A] \in H_2(X)$ be a homology class. We say that u represents $[A]$ if $u_*([\Sigma]) = [A]$, where $u_* : H_2(\Sigma) \longrightarrow H_2(X)$ is the induced map in the homology.*

The main proposition about the energy is the following identity.

Proposition 2.1.7. *Let $u : \Sigma \longrightarrow X$ be a smooth map, and J an ω-compatible almost complex structure on the symplectic manifold (X, ω). Then*

$$E(u) = \int_\Sigma |\bar{\partial}_J u|_J^2 dvol_\Sigma + \int_\Sigma u^*\omega$$

where $u^*\omega$ represents the 2-form on Σ that is the pullback by the map u of the symplectic form ω.

Proof. We choose conformal coordinates $z = s + it$ on Σ, and assume without loss of generality that Σ is an open subset of $\mathbb{C}$. We have, in this case

$$\begin{aligned}
\frac{1}{2}|du|_J^2 dvol_\Sigma &= \frac{1}{2}(|\partial_s u|_J^2 + |\partial_t u|_J^2)ds \wedge dt \\
&= \frac{1}{2}|\partial_s u + J\partial_t u|_J^2 ds \wedge dt - < \partial_s u, J\partial_t u >_J ds \wedge dt \\
&= |\bar{\partial}_J u|_J^2 dvol_\Sigma - < \partial_s u, J\partial_t u >_J ds \wedge dt
\end{aligned}$$

where we have used in the third equality the ω-compatibility of J.

Observe that we have $< \partial_s u, J\partial_t u >_J = -\omega(\partial_s u, \partial_t u) = -u^*\omega$.

Therefore, we finally get the formula from the statement.

$$E(u) = \int_\Sigma |\overline{\partial}_J u|_J^2 dvol_\Sigma + \int_\Sigma u^*\omega$$

□

Corollary 2.1.8. *If $u : \Sigma \longrightarrow X$ is a pseudoholomorphic curve, its energy is*

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

In particular, the energy of a pseudoholomorphic curve depends only on the homology class it represents.

Proof. The first part follows immediately from the proposition, since being pseudoholomorphic implies that $\overline{\partial}_J u = 0$.

For the second part, observe that, since ω is closed it defines a cohomology class $[\omega]$. Then, we have

$$E(u) = \int_\Sigma u^*\omega =< u^*[\omega], [\Sigma] >=< [\omega], [A] >$$

where $[A]$ is the homology class represented by u, and $< -, - >$ represents the Kronecker pairing. □

Observe also that it follows immediately from the formula in proposition 2.1.7 the following very important fact about pseudoholomorphic curves.

Corollary 2.1.9. *Let u be a pseudoholomorphic map representing the homology class A. Then, the energy of u is minimal with respect to the energies of all smooth maps representing A.*

Another very important property of the energy of a pseudoholomorphic curve is that it is conformal invariant, that is, it only depends on the almost complex structure j of Σ, but not on the specific metric on Σ, determined by $dvol_\Sigma$. However, the energy density $|du|_J^2$ does depend on the metric of Σ in general. This is indeed immediate from what we have done, since in fact we have seen that the energy of a pseudoholomorphic map only depends on topological data. Therefore, if we reparametrize a pseudoholomorphic map $u : \Sigma \longrightarrow X$ by composing it with an automorphism of Σ, $\psi \in Aut(\Sigma)$, we get $E(u) = E(u \circ \psi)$. This will be used extensively in the proof of the Gromov compactness theorem.

2.2 Moduli spaces of pseudoholomorphic curves: a quick overview

In this section we will describe the topology of the moduli spaces of pseudoholomorphic curves. Since the construction of moduli spaces and the proof of its properties (such as the fact that they are finite-dimensional manifolds) require advanced analytic tools such as Fredholm theory on Banach manifods, which are beyond the scope of this work, we will content ourselves to give a brief overview of the theory and the construction of moduli spaces. This overview will be enough for our purposes, and its specific details will not be very important for what follows, with the exception of the proof of the non-squeezing theorem, where we will make non-trivial use of some of the theorems stated here. For simplicity, we restrict ourselves to the genus 0 case.

We begin with a quick overview of Fredholm theory. First of all let us introduce the notion of Banach manifolds, which are essentially manifolds modelled on a Banach space instead of $\mathbb{R}^n$. In particular, this allows us to speak about infinite-dimensional manifolds. Before introducing it, we recall the definition of differentiability of a map between Banach spaces.

Definition 2.2.1. *Let V, W be Banach spaces. Let $\Omega \subset V$ be an open set, and let $f : \Omega \longrightarrow W$ be a map between them. We say that F is differentiable at a point $x \in V$ if there exists a linear map $T_x : V \longrightarrow W$ such that*

$$\lim_{||h|| \to 0} \frac{||f(x+h) - f(x) - T_x h||}{||h||} = 0$$

That is, if the map T_x is a good linear aproximation near x. In this situation, T_x is called the differential of f at x.

We say that f is differentiable if it is differentiable at each $x \in V$. In this situation we say that $T : V \longrightarrow W$ defined as $T(x) := T_x$ is its differential. If moreover T is continuous, we say that f is of class C^1. Similarly, if f is differentiable k times with continuity, we say that f is of class C^k (where $k = \infty$ means that f can be differentiated any number of times).

Definition 2.2.2. *Let E be a Banach space. A C^k E-atlas on a topological space X consists on charts (U_i, ξ_i) such that the U_i form an open cover of X, $\xi_i : U_i \longrightarrow \xi_i(U_i)$ is a homeomorphism from U_i onto an open subset $\xi_i(U_i) \subset E$ such that $\xi_j \circ \xi_i^{-1} : \xi_i(U_i \cap U_j) \longrightarrow \xi_j(U_i \cap U_j)$ is a C^k function.*

Definition 2.2.3. *A Banach C^k-manifold $\mathcal{X}$ is a Hausdorff space having a C^k E-atlas, for some Banach space E.*

A very important class of operators between Banach spaces is that of the Fredholm operators.

Definition 2.2.4. *Let E, V be Banach spaces and $F : X \longrightarrow Y$ a continuous operator between them. We say that F is a Fredholm map if both $kerF$ and $cokerF$ are finite-dimensional. In this case, we say that the index of F is $indexF := dimkerF + dimcokerF$.*

As in the case of finite-dimensional manifolds, we can define the tangent space at a point of a Banach manifold.

Definition 2.2.5. *Let X be a Banach manifold of class C^k, with $k \geq 1$, and let $p \in X$. Consider triples (U, φ, v) where (U, φ) is a chart centred at p and v an element of the model Banach space. We define an equivalence relation on the set of all these triples specifying that (U, φ, v)*

and (U', φ', v') are equivalent if the derivative of $\varphi' \circ \varphi^{-1}$ at $\varphi(p)$ sends v to v'.

We define the tangent space T_pX of X at p as the set of all such equivalence classes. Each chart (U, φ) centred at p gives a bijection of T_pX with a Banach space, via (U, φ, v) is sent to v. This bijection gives T_pX the structure of a Banach space.

Likewise, we have a definition for the differential of a map between Banach manifolds.

Definition 2.2.6. *Let $f : X \longrightarrow Y$ be a map between Banach manifolds. Then, the differential $df_p : T_pX \longrightarrow T_{f(p)}Y$ is the unique linear map satisfying that for (U, φ) a chart centred at $p \in X$ and (V, ψ) a chart centred at $f(p) \in Y$ with $f(U) \subset V$ and $\tilde{v} \in T_pX$ is represented in the chart (U, φ) by v, then $df_p(\tilde{v})$ is represented in the chart (V, ψ) by the vector $D(\psi \circ f \circ \varphi^{-1})(\varphi(p))v$.*

We can extend the notion of a Fredholm operator to maps between Banach manifolds. This will be the main maps we will be interested in, since they are the maps for which most of the theorems from finite-dimensional differential topology extend to the infinite-dimensional setting.

Definition 2.2.7. *A C^1-map $\phi : X \longrightarrow Y$ between Banach manifolds is called a Fredholm map if at each point $p \in X$, its differential $d_p\phi : T_pX \longrightarrow T_pY$ is a Fredholm operator.*

We now consider analogues of finite dimensional theorems that hold for Banach manifolds and Fredholm maps between them. Recall that a subset of a topological space is called residual if it contains a countable intersection of open dense sets. In the case of a complete metric space we have that any residual set is dense, by the Baire category theorem.

Definition 2.2.8. *Let $F : M \longrightarrow N$ be a map between Banach manifolds. We say that $y \in N$ is a regular value of F if $dF_x : T_xM \longrightarrow T_yN$ is surjective for all $x \in F^{-1}(y)$. Observe that if $F^{-1}(y) = \emptyset$, y is a regular value of F.*

Theorem 2.2.9 (Sard-Smale)**.** *The set of regular values of a C^∞ Fredholm map $F : M \longrightarrow N$ is residual in N.*

Theorem 2.2.10 (Transversality theorem)**.** *Let $F : M \longrightarrow N$ be a smooth Fredholm map and $G : Y \longrightarrow N$ be an embedding of a finite-dimensional manifold Y into N. Suppose that F is transverse to G on a closed subset of $W \subset Y$. Then, there exists a map $G' : Y \longrightarrow N$ arbitrarily close to G in the C^1-topology such that G' is transverse to F and $G'|_W = G$.*

We are now ready to introduce the moduli spaces of pseudoholomorphic curves, give them a topology, and give a hint at the proof that they are finite-dimensional manifolds.

Definition 2.2.11. *Let (X, ω) be a compact manifold, J an ω-compatible almost complex structure on X and $A \in H_2(X)$ a second homology class of X. We define*

$$\mathcal{M}(A, J) = \{u : S^2 \longrightarrow X : u_*[S^2] = A, df \circ j = J \circ df\}$$

That is, $\mathcal{M}(A, J)$ is the set of all J-holomorphic curves that represent A.

In order to have good properties on the moduli spaces (that is, in order to prove that they are finite-dimensional manifolds), we need to restrict ourselves to a special class of pseudoholomorphic curves, that we now introduce.

Definition 2.2.12. *Let $u : \Sigma \longrightarrow X$ be a J-holomorphic curve. We say that u is multiply covered if there exists a Riemann surface Σ' and a holomorphic branched covering $\phi : \Sigma \longrightarrow \Sigma'$ such that $u = u' \circ \phi$, with $\deg(\phi) > 1$. If a pseudoholomorphic curve is not multiply covered, it is said to be simple.*

Definition 2.2.13. *Let (X, ω) be a compact manifold, J an ω-compatible almost complex structure on X and $A \in H_2(X)$ a second homology class of X. We define*

$$\mathcal{M}^*(A, J) = \{u : S^2 \longrightarrow X : u_*[S^2] = A, df \circ j = J \circ df, u \text{ is simple}\}$$

That is, $\mathcal{M}^*(A, J)$ is the set of all simple J-holomorphic curves that represent A.

Let us give a topology to the moduli spaces. Observe first that we can assume $\mathcal{M}(A, I) \subset C^\infty(S^2, X)$. Now, $C^\infty(S^2, X)$ has a natural topology as a Fréchet space. Namely, we say that a sequence f_n converge to f if for all α $D^\alpha f_n \to D^\alpha f$ uniformly on S^2. In this case we say that f_n converges to f in the C^∞ topology. Then, we define the C^∞ topology on S^2 by saying that a set $A \subset C^\infty(S^2, X)$ is closed if every convergent sequence in A converges to a function in A. One can check that this defines indeed a Fréchet topology on $C^\infty(S^2, X)$. Since $\mathcal{M}(A, J) \subset C^\infty(S^2, X)$, $\mathcal{M}(A, J)$ is a topological space with the C^∞ topology inherited from $C^\infty(S^2, X)$. Observe however, that $C^\infty(S^2, X)$ is not a Banach manifold.

Using the above techniques, it can be seen that $\mathcal{M}^*(A, J)$ with this topology is in fact a finite dimensional manifold for a dense set of $J \in \mathcal{J}$. Such J's are called regular, and we denote the space of all regular almost complex structures by $\mathcal{J}_{reg}$. This is done by expressing $\mathcal{M}^*(A, J)$ as $\bar{\partial}_J^{-1}(0)$ (where we interpret $\bar{\partial}$ as a Fredholm map between certain Banach manifolds, and therefore we have to consider an embedding of $\mathcal{M}^*(A, J)$ in a more general space than $C^\infty(S^2, X)$), and using that for $J \in \mathcal{J}_{reg}$ 0 is a regular value of $\bar{\partial}_J$, hence $\mathcal{M}^*(A, J)$ is a finite dimensional manifold for such J's.

For the applications, where we will need to consider several almost complex structures, it will be useful to introduce also the universal moduli space.

Definition 2.2.14. *Let $A \in H_2(X)$ a homology class. We define the universal moduli space of pseudoholomorphic curves representing A as*

$$\mathcal{M}(A, \mathcal{J}) := \{(u, J) \in C^\infty(S^2, X) \times \mathcal{J} : \bar{\partial}_J(u) = 0, u_*([S^2]) = A\}$$

We denote by $P : \mathcal{M}(A, \mathcal{J}) \longrightarrow \mathcal{J}$ the natural projection.

Observe that $P^{-1}(J) = \mathcal{M}(A, J) \times \{J\}$, which is essentially the moduli space of J-holomorphic curves.

It turns out that $\mathcal{M}(A, \mathcal{J})$ is a Banach manifold, the projection P is a Fredholm map and the set of regular values of P is $\mathcal{J}_{reg}$.

Usually, however, we will not be interested just in pseudoholomorphic curves, but in unparametrized pseudoholomorphic curves.

Recall that the group of conformal automorphisms of the Riemann sphere S^2 is the group of Moebius transformations, $PSL(2, \mathbb{C})$, that from now on will be denoted just by G. We recall that its elements are of the form

$$\sigma(z) = \frac{az+b}{cz+d}$$

with $ad - bc \neq 0$.
Recall also that this group acts transitively in the set of all three distinct points of S^2. That is, given distinct points $z_1, z_2, z_3 \in S^2$ and distinct points $w_1, w_2, w_3 \in S^2$, there is a unique map $\sigma \in Aut(S^2)$ satisfying $\sigma(z_1) = w_1$, $\sigma(z_2) = w_2$, and $\sigma(z_3) = w_3$.

If $u : S^2 \longrightarrow X$ is a pseudoholomorphic curve, we can consider the curves $u \circ \sigma$ for all conformal automorphisms $\psi \in Aut(S^2)$. This gives us an action of the group $Aut(S^2)$ on the moduli spaces $\mathcal{M}(A, J)$. We will be interested in the moduli spaces $\mathcal{M}(A, J)/G$, quotient of $\mathcal{M}(A, J)$ by the action of this group. In the case that $\mathcal{M}(A, J)$ is a finite dimensional manifold, this quotient can be seen to be also a (finite dimensional) manifold.

Finally, we introduce the evaluation map, which is crucial for applications and further development of the theory, like the definition of Gromov-Witten invariants.

We have a natural evaluation map $\tilde{ev} : \mathcal{M}(A, \mathcal{J}) \times S^2 \longrightarrow V$ defined by $\tilde{ev}(f, z) = f(z)$. Consider now the action of G on $\mathcal{M}(A, \mathcal{J}) \times S^2$ given by $\sigma \cdot (f, z) = (f \circ \sigma^{-1}, \sigma(z))$. It is clear that the map $\tilde{ev}$ factors through this action. If we consider the restriction to the moduli space $\mathcal{M}(A, J)$ for some almost complex structure J, we arrive to the next definition.

Definition 2.2.15. *We call the evaluation map to the map $ev_J : (\mathcal{M}(A, J) \times S^2)/G \longrightarrow V$ defined by $ev_J([f, z]) = f(z)$.*

Now, fix a regular J. Then, $\mathcal{M}(A, J)$ is a finite-dimensional manifold. It can be seen that in this situation, $(\mathcal{M}(A, J) \times S^2)/G$ is also a manifold.

2.3 Failure of compactness

In this section we will see a geometrically interesting example of moduli space of pseudoholomorphic curves where compactness fails.
Considering $S^2 = \mathbb{C}P^1$ and $\mathbb{C}P^2$ as Kähler manifolds, with their standard complex structures (denoted by j and J respectively), we have that $f : \mathbb{C}P^1 \longrightarrow \mathbb{C}P^2$ is a pseudoholomorphic curve if it is a holomorphic curve, in the sense of complex geometry.

Consider the following sequence of maps $u_n : \mathbb{C}P^1 \longrightarrow \mathbb{C}P^2$ given by $u_n([x : y]) = [x^2 : y^2 : nxy]$.

First of all, it is clear that u_n are holomorphic curves, hence they are J-holomorphic curves. We see that they all represent the same homology class. Indeed, they are all homotopic to the map u_0 defined by $u_0([x : y]) = [x^2 : y^2 : 0]$ via the homotopy $H_n : \mathbb{C}P^1 \times I \longrightarrow \mathbb{C}P^2$ given by

$$H_n([x : y], t) = [x^2 : y^2 : (1 - t)nxy]$$

We only have to check that H_n is well-defined and continuous. Indeed, it is well-defined since $[x^2 : y^2 : (1 - t)nxy] \in \mathbb{C}P^2$ for all $[x : y] \in \mathbb{C}P^1$ and $H_n([\lambda x : \lambda y], t) = H_n([x : y], t)$ for all $\lambda \in \mathbb{C} - \{0\}$. The continuity is clear. Since the energy is a topological invariant, it follows that all the u_n represent the same homology class, say $A \in H_2(\mathbb{C}P^2)$. Therefore, all of them are curves in the moduli space $\mathcal{M}(A, J)$, and represent elements in the moduli space $\mathcal{M}(A, J)/G$.

We claim that there is no reparametrization $u'_n := u_n \circ \phi_n$ with $\phi_n \in G$ such that u'_n has a convergent subsequence in the C^∞ topology. Therefore the moduli space $\mathcal{M}(A, J)/G$ is non compact.

In order to see this, observe that the image of the u_n is the family of smooth algebraic curves of degree 2 (conics) defined by $XY = \frac{Z^2}{n}$. Observe that this family degenerates as $n \to \infty$ to the pair of lines $XY = 0$. Since the reparametrization of the sequence u_n does not change its image, we have that there is no convergent subsequence of any reparametrized sequence. For otherwise, its limit would have image $XY = 0$, while there is no smooth parametrization $f : \mathbb{C}P^1 \longrightarrow \mathbb{C}P^2$ of this curve.

However, $XY = 0$ is the union of two lines, hence of two smooth holomorphic curves $\mathbb{C}P^1 \longrightarrow \mathbb{C}P^2$. Let us see that each of these lines arise as the limit of a suitable reparametrization of our original sequence (except for the point $[0 : 0 : 1]$). Take $\phi_n^1([x : y]) = [x : ny]$ and observe that $\phi_n^1 \in G$. Indeed, we have that $\phi_n^1(z) = nz$. Then, $u_n^1 := u_n \circ \phi_n^1$ is defined by $u_n^1([x : y]) = [x^2 : n^2y^2 : n^2xy] = [x^2/n^2 : y^2 : xy]$ which converges to $u^1([x : y]) = [0 : y^2 : xy] = [0 : y : x]$ which has as image the line $X = 0$. Similarly, if we take $\phi_n^2([x : y]) = [nx : y]$ then we have $u_n^2 := u_n \circ \phi_n^2$ is defined by $u_n^2([x : y]) = [n^2x^2 : y^2 : n^2xy] = [x^2 : y^2/n^2 : xy]$ which has as limit the map $u^2([x : y]) = [x^2 : 0 : xy] = [x : 0 : y]$ which has as image the line $Y = 0$.

We see here a phenomenon characteristic of pseudoholomorphic curves. We have seen that the original sequence has no convergent subsequence (in the C^∞ sense), but suitable reparametrizations of the sequence have subsequences converging to pseudoholomorphic spheres, in such a way that the limit curve of the sequence is the union of this pseudoholomorphic spheres. A pseudoholomorphic sphere that develops in the limit by reparametrizing the sequence is called a bubble. We will formalize all this in the next section, and we will see that the only way in which compactness can fail is bubbling (formation of bubbles at certain points). This is the content of the Gromov compactness theorem.

Chapter 3

The Gromov compactness theorem

In this chapter we state the Gromov compactness theorem in terms of Kontsevich's stable maps. We will consider only pseudoholomorphic maps of genus 0, which are easier to describe.

In this chapter we follow closely [M-S1].

3.1 Moduli space of stable curves

As we have seen, the moduli spaces $\mathcal{M}(A,J)/G$ of pseudoholomorphic curves $u : S^2 \longrightarrow X$ are not compact in general. What we would like now is to compactify it in a geometrically meaningful way. The wanted compactification of $\mathcal{M}(A,J)/G$ will be denoted by $\overline{\mathcal{M}}(A,J)$. In this section we will describe this compactification by first presenting the objects that we will consider, and then describing the topology we will consider in this set.

As a set, the points of $\overline{\mathcal{M}}(A,J)$ will be equivalence classes of stable maps. Our first task is to define them.

As we have noted in the last chapter, we will see that the only way in which compactness can fail is that a sequence of pseudoholomorphic maps develop bubbles at some points. Since a bubble is a pseudoholomorphic sphere, we will model our limiting objects as maps defined over some set of spheres connected between themselves. Considering each sphere as a vertex of a graph, and an edge between vertices as indicating that the two spheres represented by the vertices are connected, it turns out that the adequate model will be that of a tree.

Definition 3.1.1. *A graph $G = (V,E)$ consists on a set V of vertices and a relation E (edges) in V which is symmetric and such that $x \not E x$ for all $x \in V$.*

Let $x, y \in V$. A path in G starting at x and ending at y is a finite sequence of vertices $(v_1, .., v_n)$ such that $v_1 = x$, $v_n = y$ and $v_i E v_{i+1}$ for $i = 1, ..., n-1$.

A graph G is said to be connected if for every pair of vertices $x, y \in G$ there exists a path in G starting at x and ending at y.

A cycle in G is a finite sequence of vertices $(v_1, ..., v_n)$ such that $n \geq 3$, $v_i \neq v_j$ for $i \neq j$, $v_i E v_{i+1}$ for $i = 0, ..., n-1$, and $v_1 = v_n$.

A finite graph G is said to be a tree if it is connected and has no cycles.

Given a tree T and two vertices $\alpha, \beta \in T$ with $\alpha E \beta$, we will denote by $T_{\alpha\beta}$ the set of all vertices γ in T such that there exists a path from α to *gamma* passing trough β.

We will also need the notion of tree homomorphism.

Definition 3.1.2. *Let T, T' be trees. A map $f : T \longrightarrow T'$ is said to be a tree homomorphism if $f^{-1}(\gamma)$ is a tree for each vertex $\gamma \in T'$, and for all $\alpha, \beta \in T$ with $\alpha E \beta$, either $f(\alpha) E' f(\beta)$ or $f(\alpha) = f(\beta)$.*

If moreover f is bijective and f^{-1} is also a tree homomorphism, we say that f is a tree isomorphism.

Our stable maps will be modelled on trees. We define them now.

Definition 3.1.3. *Let (X, ω) be a compact symplectic manifold and let J be an ω-compatible almost complex structure on X.*
A stable J-holomorphic map of genus zero into X, modelled over the tree T is a tuple

$$(u, z) = (\{u^\alpha\}_{\alpha \in T}, \{z_{\alpha\beta}\}_{\alpha E \beta})$$

where each $u^\alpha : S^2 \longrightarrow X$ is a J-holomorphic curve and $z_{\alpha\beta}$ are points in S^2 (called nodal points) labelled with the oriented edges $\alpha E \beta$ of the tree T, such that the following conditions are satisfied:

1. *For all $\alpha, \beta \in T$ such that $\alpha E \beta$, we have that $u^\alpha(z_{\alpha\beta}) = u^\beta(z_{\beta\alpha})$.*
2. *For every $\alpha \in T$, we have $z_{\alpha\beta} \neq z_{\alpha\beta'}$ if $\beta \neq \beta'$.*
   We put $Z_\alpha = \{z_{\alpha\beta} : \alpha E \beta\}$.
3. *If u^α is a constant map, then $\#Z_\alpha \geq 3$.*

Observe that this definition correspond with the intuitive image of the limit of a pseudoholomorphic curves, since as we have say, convergence can only fail via bubbling. Therefore, it is reasonable to expect that the limit is a map defined on a bunch of spheres, one connected to others only through a point (the point where bubble develops) in such a way that there are no cycles, and such that the map restricted to each sphere is a pseudoholomorphic map, and the only point that two contiguous spheres have in common have the same image.

We have also an evident notion of energy for an stable map: just sum the energy of all the bubbles.

Definition 3.1.4. *Let (u, z) be a stable map. Then we define its energy as*

$$E(u) := \sum_{\alpha \in T} E(u^\alpha)$$

We also define

$$m_{\alpha\beta}(u) := \sum_{\gamma \in T_{\alpha\beta}} E(u^\gamma)$$

In the similar way, stable maps represent second homology classes of X.

Definition 3.1.5. *Let (u, z) be a stable map and $A \in H_2(X)$. We say that (u, z) represents the class A if*

$$A = \sum_{\alpha \in T} u^\alpha_*([S^2])$$

We define the moduli space of stable curves representing the homology class A.

Definition 3.1.6. *The moduli space of J-stable curves representing the homology class A is defined by*

$$SC(A, J) = \{(u, z) : (u, z) \text{ is a } J\text{-stable curve representing } A\}$$

We have a natural notion of equivalence of stable curves by reparametrization, that we now describe.

Definition 3.1.7. *Two stable maps (u, z), (u', z'), modelled over trees T, T' respectively, are called equivalent if there exists a tree isomorphism $f : T \longrightarrow T'$ and a function $\phi : T \longrightarrow G$ (where we denote $\phi(\alpha)$ by ϕ_α) such that for all α*

1. $u'_{f(\alpha)} \circ \phi_\alpha = u_\alpha$
2. $z'_{f(\alpha)f(\beta)} = \phi_\alpha(z_{\alpha\beta})$

We define now the desired moduli spaces of pseudoholomorphic curves.

Definition 3.1.8. *Let $A \in H_2(X)$, and J an almost complex structure on X. We define the moduli space of J-stable curves representing A, and denote it by $\overline{\mathcal{M}}(A, J)$, as the quotient of $SC(A, J)$ by the equivalence relation defined above (that is, equivalence of stable trees).*

We have to define a topology on the moduli space $\overline{\mathcal{M}}(A)$. We will define the topology by defining first a notion of convergence of a sequence of stable maps.

Definition 3.1.9. *A sequence of stable maps*

$$(u_n, z_n) = (\{u^\alpha_n\}_{\alpha \in T_n}, \{z^n_{\alpha\beta}\}_{\alpha E_n \beta})$$

is said to Gromov converge to a stable map

$$(u, z) = \{u^\alpha\}_{\alpha \in T}, \{z_{\alpha\beta}\}_{\alpha E \beta})$$

if there exists a n_0 such that for all $n \geq n_0$ there exists a surjective tree homomorphism

$$f_n : T \longrightarrow T_n$$

and a collection of Moebius transformations $\{\phi^\alpha_n\}_{\alpha \in T}$, such that the following holds:

1. *For every $\alpha \in T$ the sequence of maps $u^{f_n(\alpha)}_n \circ \phi^\alpha_n : S^2 \longrightarrow X$ converges uniformly on compact subsets of $S^2 - Z_\alpha$.*

2. *If $\alpha E \beta$, then:*

$$m_{\alpha\beta}(u) = \lim_{\epsilon \to 0} \lim_{n \to \infty} E_{f_n(\alpha)}(u_n; \phi_n^\alpha(B_\epsilon(z_{\alpha\beta})))$$

3. *If $\alpha, \beta \in T$ are such that $\alpha E \beta$ and n_j is a subsequence such that $f_{n_j}(\alpha) = f_{n_j}(\beta)$ then $\phi_{n_j}^{\alpha\beta} := (\phi_{n_j}^\alpha)^{-1} \circ \phi_{n_j}^\beta$ converges to $z_{\alpha\beta}$ uniformly on compact subsets of $S^2 - \{z_{\alpha\beta}$.*

4. *If $\alpha, \beta \in T$ are such that $\alpha E \beta$ and n_j is a subsequence such that $f_{n_j}(\alpha) = f_{n_j}(\beta)$ then $z_{\alpha\beta} = \lim_{n\to\infty} (\phi_{n_j}^\alpha)^{-1}(z^{n_j}_{f_{n_j}(\alpha), f_{n_j}(\beta)})$.*

Observe that this definition induces a notion of convergence in $\overline{\mathcal{M}}(A, J)$.

Before going on, let us explain what this notion of convergence means. In order to do so, it will be helpful to write the definition of Gromov convergence in the more usual situation where the sequence of stable maps consist in pseudoholomorphic spheres. That is, the definition of convergence of a sequence of pseudoholomorphic spheres to a stable map.

Definition 3.1.10. *A sequence of pseudoholomorphic maps*

$$u_n : S^2 \longrightarrow X$$

is said to Gromov converge to a stable map

$$(u, z) = \{u_\alpha\}_{\alpha \in T}, \{z_{\alpha\beta}\}_{\alpha E \beta})$$

if there exists a collection of Moebius transformations $\{\phi_n^\alpha\}_{\alpha \in T}$, such that the following holds:

(Map) *For every $\alpha \in T$ the sequence of maps $u_n^\alpha := u_n \circ \phi_n^\alpha : S^2 \longrightarrow X$ converges to u^α uniformly on compact subsets of $S^2 - Z_\alpha$.*

(Energy) *If $\alpha E \beta$, then:*

$$m_{\alpha\beta}(u) = \lim_{\epsilon \to 0} \lim_{\nu \to \infty} E(u_n; \phi_n^\alpha(B_\epsilon(z_{\alpha\beta})))$$

(Rescaling) *If $\alpha, \beta \in T$ are such that $\alpha E \beta$ then the sequence $\phi_n^{\alpha\beta} := (\phi_n^\alpha)^{-1} \circ \phi_n^\beta$ converges to $z_{\alpha\beta}$ in the C^∞ sense on $S^2 - \{z_{\beta\alpha}\}$.*

Observe that in the first condition the sequence u_n does not enter directly, but only composed with Moebius transformations. This is because we are only interested in the limit of the sequence modulo the action of the reparametrization group G. Therefore, we do not ask the sequence to converge, but only a suitable reparametrization of the sequence. Taking this into account, the first condition says that a suitable reparametrized sequence converge in the C^∞ sense everywhere on S^2 except at the points of Z_α, which are interpreted as the points where bubbling occurs in the limit.

The second condition asserts that there is no energy loss in the limit. Indeed, it says that the energy of the bubble tree bubbling at the point $z_{\alpha\beta}$ is all the energy concentrating at small neighbourhoods of the point $z_{\alpha\beta}$. The fact that there is no energy loss in the limit implies that this is the correct notion of convergence, since we do not loss geometric information about the sequence when passing to the limit.

Now we can finally define the topology on the set of stable maps as follows.

Definition 3.1.11. *A set $C \subset \overline{\mathcal{M}}(A, J)$ is said to be Gromov closed if every Gromov convergent sequence of stable curves in C has a limit in C. A set $U \subset \overline{\mathcal{M}}(A, J)$ is said to be Gromov open if its complement is Gromov closed.*

It can be checked that this defines a topology on $\overline{\mathcal{M}}(A, J)$, called the Gromov topology.

3.2 Statement of the compactness theorem

In this section we will state and discuss the theorem we want to prove. As we have seen in the previous chapter, the moduli spaces $\mathcal{M}(A, J)/G$ are not compact in general. Since the compactness of a moduli space is a very desirable property (it is crucial for some applications, as we will see in the last chapter), we seek a compactification of the moduli space. Moreover, we want a compactification that is geometrically meaningful, so that the points we add to the moduli space gives us some insight about pseudoholomorphic curves. As said above, the new points we need to add to the moduli space are the so-called stable maps. They are geometrically meaningful, since they arise in a natural way as the limit of reparametrized sequences of pseudoholomorphic curves.

We state now a first version of the Gromov compactness theorem that tells us that any sequence of pseudoholomorphic maps converge to a stable map.

Theorem 3.2.1 (Gromov compactness theorem)**.** *Let (X, ω) be a symplectic manifold and let J_n be ω-compatible almost complex structures converging in the C^∞ sense to an almost complex structure J. Let $u_n : S^2 \longrightarrow X$ be a sequence of J_n-holomorphic curves with bounded energy. Then, there is a subsequence of the u_n that Gromov converges to a J-stable map (u, z).*

We can improve this theorem in order to prove that every sequence of stable maps of bounded energy has a subsequence that Gromov converges to a stable map.

Theorem 3.2.2 (Gromov compactness theorem for stable maps)**.** *Let (X, ω) be a symplectic manifold and let J_n be ω-compatible almost complex structures converging in the C^∞ sense to an almost complex structure J. Let $(u_n, <_n)$ be a sequence of J_n-stable maps with bounded energy. Then, there is a subsequence of the u_n that Gromov converges to a J-stable map (u, z).*

In particular, we can apply this last result to $\overline{\mathcal{M}}(A, J)$, and it tells us that any sequence of stable maps from $\overline{\mathcal{M}}(A, J)$ has a convergent subsequence (observe that convergence in the topology of the moduli space of stable maps coincides with Gromov convergence of stable maps). However, this is not enough to see that $\overline{\mathcal{M}}(A, J)$ is compact, unless we know that the topology satisfy some condition like having a countable basis of open sets. In fact, one can prove the following theorem:

Theorem 3.2.3. *The moduli spaces $\overline{\mathcal{M}}(A, J)$ are compact metrizable spaces.*

However, proving the metrizability of the moduli spaces would take us too far, and we will not do it in this work. For a proof, see chapter 3 of [M-S1].

Chapter 4

Local estimates

In this chapter we will study in detail the local behaviour of pseudoholomorphic curves. In particular, we will study the behaviour of pseudoholomorphic curves defined on disks and pseudoholomorphic curves defined on cylinders (which are equivalent to pseudoholomorphic curves defined on annuli), focusing on the distribution of the energy density along a sufficiently long cylinder.

After that, we will put to use the results we obtain for pseudoholomorphic cylinders in order to prove the removal of singularities theorem, which states that a pseudoholomorphic map defined on a punctured disk with bounded energy extends to a pseudoholomorphic map defined on the whole disk, and so it can be thought of as a pseudoholomorphic analogue of the Riemann extension theorem of complex analysis, which says that any bounded holomorphic function defined on a punctured disk extends holomorphically to the whole disk. We give here an elementary proof of this result, assuming only knwoledge of the usual L^2 theory of linear elliptic operators.

The local estimates obtained in this chapter (and especially the removal of singularities theorem) constitute the core of our proof of the Gromov compactness theorem, that we will develop in the following chapter.

In all this chapter we denote by X a compact manifold endowed with an almost complex structure J. We consider X as embedded in some euclidean space $\mathbb{R}^N$, and take as metric in X the pullback of the euclidean metric in $\mathbb{R}^n$. Observe that since X is compact, all the metrics in X induce equivalent norms.

We also assume in the following that all pseudoholomorphic curves are smooth.

4.1 Review of Sobolev spaces and elliptic regularity

In this section we collect the relevant definitions and theorems from L^2 Sobolev theory that we need in the rest of the chapter. We will not present proofs, and refer the reader to Chapter 6 of [War].

We will consider functions $\varphi : \mathbb{R}^{\ltimes} \longrightarrow \mathbb{C}^m$. Moreover, in order to simplify things, and since we are only interested in the local behaviour of functions, we will assume that they are periodic of period 2π in each variable. This is not a restriction since we can always extend a function

defined in a neighbourhood of a point to a periodic function.

If we consider smooth periodic functions, we can develop them in Fourier series:

$$\varphi(x) = \sum_{\xi} \varphi_\xi e^{ix\cdot\xi}$$

where $\xi = (\xi_1, ..., \xi_n)$ is an n-tuple of integers, and the sum is extended over $\mathbb{Z}^n$. Recall also that the Fourier coefficients are defined by

$$\varphi_\xi = \frac{1}{(2\pi)^n} \int_Q \varphi(x) e^{-ix\cdot\xi}$$

where Q is a cube of side 2π.

We will identify a periodic function with the set of corresponding Fourier coefficients. This allows us to define Sobolev spaces of periodic functions. We denote by $\mathcal{S}$ the complex vector space of sequences of complex vectors in $\mathbb{C}^m$ indexed by n-tuples of integers ξ.

Definition 4.1.1. *For each integer k we define the k-Sobolev space as*

$$L^2_k = \{u \in \mathcal{S} : \sum_{\xi} (1 + |\xi|^2)^k |u_\xi|^2 < \infty\}$$

and we endow it with the inner product defined by

$$< u, v >_k = \sum_{\xi} (1 + |\xi|^2)^k u_\xi \cdot v_\xi$$

It can be seen that L^2_k are Hilbert spaces.

Observe that, for $s \geq 0$, each $u \in L^2_s$ represents an L^2 function defined by

$$\phi(x) = \sum_{\xi} u_\xi e^{ix\cdot\xi}$$

Proposition 4.1.2. *The space of all smooth periodic functions is dense in every Sobolev space L^2_k.*

For a proof, see Theorem 6.18(c) of [War].

An interesting result, that gives some insight into the meaning of Sobolev spaces, is the following:

Proposition 4.1.3. *Let k be a non-negative integer, and let φ be a smooth periodic function. Then, the Sobolev norm in L^2_k is equivalent to the norm given by*

$$||\varphi|| = \sum_{[\alpha]=0}^{k} ||D^\alpha \varphi||$$

where we use multi-index notation ($\alpha = (\alpha_1, ..., \alpha_n)$ and $[\alpha] = \alpha_1 + ... + \alpha_n$), and $||D^\alpha \varphi|| = \sup_{x\in Q} |D^\alpha \varphi(x)|$.

For a proof, see Theorem 6.18(a) of [War].

Therefore, for smooth functions, the k-Sobolev norm takes into account the values of the function and the values of all the derivatives up to order k.

A key property of the Sobolev spaces is the fact that, if some function is in a sufficiently high Sobolev space, then it is sufficiently derivable. More precisely,

Theorem 4.1.4 (Sobolev lemma)**.** *If $t > [n/2] + m + 1$ (where n is the dimension of the space) and $u \in L^2_t$, then the series $D^\alpha u = \sum_\xi \xi^\alpha u_\xi e^{ix\cdot\xi}$ converges uniformly for $[\alpha] \leq m$. Therefore, each $u \in L^2_t$ for this range of t corresponds to a function of class C^m.*

For a proof, see Lemma 6.22 of [War].

The other important theorem about Sobolev spaces is the following.

Theorem 4.1.5 (Rellich lemma)**.** *Let u^n be a sequence of elements of L^2_t with $||u^n||_t \leq 1$. If $s < t$, then there is a subsequence of u^n which converges in L^2_s. In other words, the inclusion $i : L^2_t \longrightarrow L^2_s$ is compact.*

For a proof, see Lemma 6.33 of [War].
For the proof of the removal of singularities theorem, we will need the following lemma.

Definition 4.1.6. *Let $u \in L^2_k$ and $0 \neq h \in \mathbb{R}^n$. We define the difference quotient of u determined by h as*

$$u^h = \frac{T_h(u) - u}{|h|}$$

where $T_h : L^2_k \longrightarrow L^2_k$ is the operator defined (using the Fourier coefficients) by $(T_h(u))_\xi := e^{ih\cdot\xi} u_\xi$.

Lemma 4.1.7. *Let $u \in L^2_k$. Then, for any $0 \neq h \in \mathbb{R}^n$, we have that $||u^h||_{L^2_k} \leq C||u||_{L^2_{k+1}}$, where C does not depend on h. Conversely, if $||u^h||_{L^2_k} \leq K$ for some constant K independent of h, then $u \in L^2_{k+1}$.*

For a proof see 6.19 and 6.20 of [War].

Definition 4.1.8. *A linear differential operator L of order l on the $\mathbb{C}^m$-valued C^∞ functions on $\mathbb{R}^n$ consists of an $m \times m$ matrix (L_{ij}) in which*

$$L_{ij} = \sum_{[\alpha]=0}^{l} a^\alpha_{ij} D^\alpha$$

where $a^\alpha_{ij} : \mathbb{R}^n \longrightarrow \mathbb{C}^m$ are C^∞ and at least one $a^\alpha_{ij} \neq 0$ for some i, j and some α with $[\alpha] = l$.

In principle, a differential operator is defined only for sufficiently differentiable functions. However, since smooth functions are dense in all Sobolev spaces, it extends by density to an operator $L : L^2_{k+l} \longrightarrow L^2_k$, where l is the order of L.

Definition 4.1.9. *Let L be a differential operator of order l. We write L as*

$$L = P_l(D) + ... + P_0(D)$$

where $P_j(D)$ is an $m \times m$ matrix where each entry is a homogeneous differential operator or order j, $\sum_{[\alpha]=j} a_\alpha D^\alpha$.

We say that L is elliptic at $x \in \mathbb{R}^n$ if $P_l(\xi)$ is non-singular at x for every $0 \neq \xi \in \mathbb{R}^n$, where $P_l(\xi)$ represents the matrix obtained from $P_l(D)$ by replacing each occurrence of D^α with ξ^α.

We say that L is elliptic if L is elliptic at every $x \in \mathbb{R}^n$.

We have the following characterization of elliptic operators.

Proposition 4.1.10. *L is elliptic if and only if $L(\phi^l u)(x) \neq 0$ for each $\mathbb{C}^m$-valued function u such that $u(x) \neq 0$ and each smooth real-valued ϕ such that $\phi(x) = 0$ but $d\phi(x) \neq 0$.*

See 6.28 of [War] for a proof.

We will be mainly concerned with the Cauchy-Riemann operator, which for a smooth function $f : \mathbb{R}^2 \longrightarrow \mathbb{C}^m$ is defined as

$$\bar{\partial} f = \partial_x f + i\partial_y f$$

We have that

Proposition 4.1.11. *The Cauchy-Riemann operator $\bar{\partial}$ is elliptic.*

Proof. Simply observe that $\xi_1 + i\xi_2$ is never 0 for $\xi = (\xi_1, \xi_2) \neq (0, 0)$. □

The following is the main inequality for elliptic operators. We will use it extensively in what follows

Proposition 4.1.12 (Gårding's inequality)**.** *Let L be an elliptic operator of order l, and let s be an integer. There is a constant $C > 0$ such that*

$$||u||_{L^2_{s+l}} \leq C(||Lu||_{L^2_s} + ||u||_{L^2_s}$$

for all $u \in L^2_{s+l}$.

See 6.29 of [War] for a proof.

Definition 4.1.13. *Let L be a differential operator of order l, and $v \in L^2_s$ for some s. Any solution $u \in L^2_k$ for some k of an equation $Lu = v$ is called a weak solution of the differential equation. Such a weak solution is called a strong solution if $u \in C^l$.*

The main theorem about elliptic operators is the following.

Theorem 4.1.14 (Elliptic regularity)**.** *Let L be a periodic elliptic operator of order l. Assume that $u \in L^2_k$ for some $k \in \mathbb{Z}$, $v \in L^2_t$ and*

$$Lu = v$$

Then, $u \in L^2_{t+l}$.

For a proof, see 6.30 of [War].

4.2 Estimates on disks

We start by proving some estimates for pseudoholomorphic maps defined on disks.

The following lemma tells us that the L^2_k norms of the derivative of a pseudoholomorphic map on a small disk is controlled by the C^0 norm of the derivative on a larger disk, provided the C^0 norm of the derivative is small enough.

Lemma 4.2.1. *Let J be an almost complex structure in $\mathbb{C}^n$ satisfying $J(0) = J_0$ (where J_0 represents the canonical complex structure on the vector space $\mathbb{C}^n$). For all $k \in \mathbb{N}$, there exists an $\epsilon > 0$ such that if $u : D_1 \longrightarrow \mathbb{C}^n$ is a J-holomorphic curve (i.e., it satisfy $\overline{\partial}_J(u) = 0$), satisfying $|du|_{C^0(D_1)} < \epsilon$ and $|u(0)| < \epsilon$, then $||u||_{L^2_k(D_{1/2})} < C_k|du|_{C^0(D_{3/4})}$, for some constant $C_k > 0$ independent of u.*

Proof. First, observe that since u is pseudoholomorphic, it is smooth. Therefore, $u \in L^2_k(D_1)$ for all k. We prove the theorem by induction on k.

For $k = 0$, we have that $||u||_{L^2(D_{1/2})} \leq const|u|_{C^0(D_{1/2})}$.
But $|u|_{C^0(D_{1/2})} < 2\epsilon$ if $|du|_{C^0(D_1)} < \epsilon$ and $|u(0)| < \epsilon$, since $|u(z) - u(0)| \leq \int_{[0,z]} |du| \leq \frac{1}{2}\epsilon < \epsilon$ and $|u(z)| \leq |u(0)| + |u(z) - u(0)| < 2\epsilon$, if $|z| \leq \frac{1}{2}$.
Therefore, $||u||_{L^2(D_{1/2})} \leq const|du|_{C^0(D_{3/4})}$, where the constant is independent of u.

Fix a $k > 0$ and suppose the lemma is true for this k.

By Gårding's inequality,

$$||u||_{L^2_{k+1}(D_{1/2})} \leq const(||\overline{\partial}_{J_0}u||_{L^2_k(D_{1/2})} + ||u||_{L^2_k(D_{1/2})})$$

Since $0 = \overline{\partial}_J u = \overline{\partial}_{J_0} + (J(u) - J_0)\partial_y u$ we have $\overline{\partial}_{J_0}u = -(J(u) - J_0)\partial_y u$.

Therefore:

$$||u||_{L^2_{k+1}(D_{1/2})} \leq const(||(J(u) - J_0)\partial_y u||_{L^2_k(D_{1/2})} + ||u||_{L^2_k(D_{1/2})})$$

We now want to estimate $||(J(u)-J_0)\partial_y u||_{L^2_k(D_{1/2})}$. In order to do that, observe that $\nabla^r((J(u)-J_0)\partial_y u) = \sum_{p+q=r} \nabla^p(J(u) - J_0)\nabla^q\partial_y u$. Since both J and u are smooth, so is $J(u) - J_0$.

Observe that for $z \in D_{1/2}$ we have $|u(z) - u(0)| \leq \int_{[0,z]} |du| < \frac{\epsilon}{2}$ if $|du|_{C^0(D_{3/4})} < \epsilon$. Then, $|u(z)| \leq |u(0)| + \epsilon/2 < \epsilon + \epsilon < 2\epsilon$. Therefore, we can ensure that $u(D_{1/2}) \subset B_{2\epsilon}$ if we have $|du| < \epsilon$ and $|u(0)| < \epsilon$, choosing $\epsilon > 0$ small enough.
Let $\delta > 0$, an choose $\epsilon > 0$ so that it also satisfies $||J||_{C^k(B_{2\epsilon})} < \delta$ (such a ball $B_{2\epsilon}$ exists since J is smooth).

For $p, q < k$, we have, by induction hypothesis,

$$\begin{aligned}
||\nabla^p(J(u)-J_0)\nabla^q\partial_x u||_{L^2(D_{1/2})} &= ||(\nabla^p(J-J_0))\circ u\nabla^p(u)\nabla^q\partial_x u||_{L^2(D_{1/2})} \\
&\leq ||(\nabla^p(J-J_0))\circ u||_{C^0(D_{1/2})}||\nabla^p(u)\nabla^{q+1}(u)||_{L^2(D_{1/2})} \\
&\leq \delta||\nabla^p(u)\nabla^{q+1}(u)||_{L^2(D_{1/2})}
\end{aligned}$$

Observe now that, by the Sobolev embedding theorem, we have that if $||u||_{L^2(D_{1/2})} < \epsilon$ then $|u|_{C^{k-2}} \leq ||u||_{L^2(D_{1/2})}$, if $k \geq 2$.

Since $p+q=r\leq k$, either $p<k-1$ or $q+1<k-1$, if $k>3$. In this case, if $p<k-1$, we obtain

$$\begin{aligned}
||\nabla^p(u)\nabla^q(u)||_{L^2(D_{1/2})} &\leq |u|_{C^p(D_{1/2})}||\nabla^q(u)||_{L^2(D_{1/2})} \\
&\leq |u|_{C^p(D_{1/2})}||u||_{L^2_{k+1}(D_{1/2})} \leq C||u||_{L^2_{k+1}(D_{1/2})}
\end{aligned}$$

where in the last step we have used $|u|_{C^{k-2}} \leq ||u||_{L^2(D_{1/2})}$, the induction hypothesis, and we assume that $|du|_{C^0(D_{1/2})} < 1$. We get the same result if $q+1<k-1$. In the cases $k=1,2,3$, observe that we always have $p\leq 1$ or $q\leq 1$. Then, we can use the fact that we may assume $|du|_{C^0}<1$ and $|u|_{C^0}\leq 2\epsilon$, to obtain in all the cases $||\nabla^p(u)\nabla^q(u)||_{L^2(D_{1/2})} \leq C||u||_{L^2_{k+1}(D_{1/2})}$, where C is a constant independent of u.

Therefore, we conclude that $||\nabla^r((J(u)-J_0)\partial_x u)||_{L^2(D_{1/2})} \leq const\delta||u||_{L^2_{k+1}(D_{1/2})}$, if $r<k$, where the constant is independent of u.

Putting all together, we obtain $||R(u)||_{L^2_k(D_{1/2})} = ||(J(u)-J_0)\partial_x u||_{L^2_k(D_{1/2})} \leq C(\delta)||u||_{L^2_{k+1}(D_{1/2})}$, where $C(\delta)$ is a constant depending only on δ. Looking at the definition of δ, we can take ϵ small enough such that if $|du|_{C^0(D_1)}<\epsilon$, then $const.C(\delta)<1/2$.

Then, returning at the Gårding's inequality, we get

$$||u||_{L^2_{k+1}(D_{1/2})} \leq 1/2||u||_{L^2_{k+1}(D_{1/2})} + const||u||_{L^2_k(D_{1/2})}$$

So, $||u||_{L^2_{k+1}(D_{1/2})} \leq 2const||u||_{L^2_k(D_{1/2})}$. And, by induction hypothesis,

$$||u||_{L^2_k(D_{1/2})} \leq C_k|du|_{C^0(D_1)}$$

if $|du|_{C^0(D_1)}$ is small enough. Finally, we conclude that there exists an $\epsilon>0$ and a constant $C_k>0$ such that if $|du|_{C^0(D_1)}<\epsilon$, $||u||_{L^2_{k+1}(D_{1/2})} \leq C_k|du|_{C^0(D_1)}$, as wanted. □

We can now prove that the L^2_k norms of du are controlled by the C^0 norm of du.

Corollary 4.2.2. *Let $u: D_1 \longrightarrow X$ be a pseudoholomorphic map. Then, for each $k\in\mathbb{N}$ there is some $\epsilon>0$ and a constant $C_k>0$, such that if $|du|_{C^0(D_1)}<\epsilon$, then $||du||_{L^2_k(D_{1/2})} \leq C_k|du|_{C^0(D_{3/4})}$, where ϵ and C_k are independent of the map u.*

Proof. Fix a k. Let $\delta>0$ be the number ϵ corresponding to k in the preceding proposition. By the compactness of X, we can find a number $\epsilon>0$, a finite open cover of X, $U_1,...,U_n$, and charts $\xi_i: U_i \longrightarrow D_1$ such that for any pseudoholomorphic map $u: D_1 \longrightarrow X$ with $|du|_{C^0(D_1)}<\epsilon$

there exists an i with $u(D_1) \subset \xi_i^{-1}(D_{1/2})$ and, if we define $u' = \xi_i \circ u$, then $|du'|_{C^0(D_1)} < \delta$ and $|u'(0)| < \delta$.

Then, by the previous lemma, $||u'||_{L^2_{k+1}(D_{1/2})} \leq C|du|_{C^0(D_{3/4})}$. In particular, $||du'||_{L^2_k(D_{1/2})} \leq ||u'||_{L^2_{k+1}(D_{1/2})} \leq C|du'|_{C^0(D_{3/4})}$.

Now, observe that the norm induced by the norm of X in $\xi_i^{-1}(\overline{D}_{1/2})$, and the norm induced by the pullback of the standard metric of D_1 by ξ_i in $\xi_i^{-1}(\overline{D}_{1/2})$ are equivalent. Since this is true for any i and there are a finite number of such i's, there exist a constant $C > 0$ (independent of u) such that if u satisfies the above conditions, then $\frac{1}{C}|du(z)| \leq |du'(z)| \leq C|du'(z)|$. Therefore, the result for u follows from that of u'. □

With this result at our disposal, we can prove a compactness result for sequences of pseudoholomorphic curves with bounded energy density.

Proposition 4.2.3. *Let $u_n : \Sigma \longrightarrow X$ be a sequence of J_n-holomorphic curves, and suppose that the almost complex structures J_n converge to an almost complex structure J in the C^∞ sense. Let $K \subset \mathring{K'} \subset K' \subset \Sigma$, where K, K' are two compact sets. Assume also that $\sup_n \sup_{x \in K'} |du_n(x)| < C < \infty$, for some $C > 0$. Then, there exists a subsequence $(u_{n_k})_k$ that converges to a J-holomorphic curve u in $C^\infty(K)$.*

Proof. We start by showing that there is a subsequence that converges in $C^k(K)$, for every k. By the Sobolev embedding theorem, it is enough to check that there is a convergent subsequence in $L^2_k(K)$ for any natural k.

Fix a natural k. We want to find a convergent subsequence of u_n in $L^2_k(K)$. Let ϵ_{k+1} be the ϵ coming from lemma 4.2.1 for $k+1$. Choose, for each $x \in K$, an open set $U_x \subset K'$ that is the domain of a chart $\chi_x : U_x \longrightarrow \mathbb{C}$ centered at x.

For each open set U_x, put $u_n^x = u_n|_{U_x} \circ \chi_x$. Without loss of generality, we may assume that $|du_n^x|_{C^0} < \epsilon_{k+1}$ (if not, replace χ_x by $\chi_x \circ \psi$, where $\psi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is defined by $\psi(z) = \frac{\epsilon_{k+1}}{C} z$).

Since X is compact, passing to a subsequence we may assume that the sequence $(u_n^i(0))_n$ converges to a point $x \in X$. Pick a chart (V, ξ) of X centred at x and such that $J(0) = J_0$ in this chart. For n big enough we have that $\xi(u_n^i(0)) \in D_{\epsilon_{k+1}}$. Since also $|du_n^i|_{C^0(D_1)} < \epsilon_{k+1}$, by the lemma above, we get that $||u_n^i||_{L^2_{k+1}(D_{1/2})} \leq A\epsilon_{k+1}$, if we can obtain an ϵ_{k+1} that works for all almost complex structures J_n, for n big enough.

Indeed, fixed a $\delta > 0$, for n big enough, we have $||J(x) - J_n(x)||_{L^2_k(D_r)} < \delta$ (by the Sobolev embedding theorem together with the fact that J_n converges to J in the C^∞ sense). Then, for n big enough and any $\epsilon > 0$, if $||J - J_0||_{L^2_k(D_r)} < \epsilon - \delta$, we have $||J_n - J_0||_{L^2_k(D_r)} < \delta + ||J - J_0||_{L^2_k(D_r)} < \epsilon$ for n big enough. By looking at the proof of lemma 4.2.1, we see that this is enough to see that the lemma applies to the present situation and we can take ϵ_{k+1} to be the same for all the almost complex structures J_n.
Hence, by Rellich's lemma, it has a convergent subsequence in L^2_k.

Therefore, for any U_x we have a subsequence of u_n that converges to a L^2_k function u_x on U_x. Since K is compact and the $\{U_x\}_{x \in X}$ is an open covering of K, we can find a finite subcover $U_1, ..., U_n$ of K. Then, we can obtain a subsequence of u_n that converges in L^2_k to a function u in

all of K. Indeed, we can extract a convergent subsequence u_n^1 that converges in U_1. Then from this sequence we can extract a subsequence u_n^2 that converges in U_2, and continue until we get a subsequence that converges in $L_k^2(K)$. Finally, we repeat this process with respect to k, that is, we choose a subsequence u_n^1 of the original sequence that converges in $L_1^2(K)$, extract from this another subsequence u_n^2 that converges in $L_2^2(K)$ and proceed recursively for each k. Finally consider the diagonal sequence $(u_n^n)_n$. By construction, this is a subsequence of our original sequence that converges in $L_k^2(K)$ for every k. By the Sobolev embedding theorem, we have that this sequence in fact converges in $C^n(K)$ for every n, or in other words, that it converges in $C^\infty(K)$. Call u the limit of this sequence. Then, $u \in C^\infty(K)$. Moreover, u satisfies $\overline{\partial}_J u = 0$. Indeed, we have that $\overline{\partial}_{J_n} u_n^n = 0$ for each n, and u_n^n converges to u uniformly, so taking the limit as n goes to ∞, we obtain that u is a J-holomorphic curve, as wanted.

□

To finish, we prove a converse of corollary 4.2.2, namely that we can control the C^0 norm of du on some disk if we have control over the L^2 norm on some larger disk, provided the energy of the curve is small enough. But first, we need the following lemma.

Lemma 4.2.4. *Let $u : D_1 \longrightarrow X$ be a J-holomorphic map. Then there exists an $\epsilon > 0$ such that if $||du||_{L^2(D_1)} < \epsilon$ then $|du(0)| \leq K||du||_{L^2(D_1)}$ for some constant $K > 0$ independent of u.*

Proof. Suppose that the conclusion of the statement is not true. In this case, we can find a sequence of J-holomorphic maps $u_j : D_1 \longrightarrow X$ such that $||du_j||_{L^2(D_1)} \to 0$ and $\lim_j \frac{|du_j(0)|}{||du_j||_{L^2(D_1)}} = \infty$.

We will split the proof in two cases: that $|du_j(0)| \to 0$, and that there is some $\epsilon > 0$ such that $|du_j(0)| \geq \epsilon$ for all j.

Case 1: $|du_j(0)| \to 0$.

We will see that we can reduce this case to the second one. Indeed, passing to a subsequence, we may assume that the images of all the u_j are contained in one coordinate chart of X, by taking the ϵ of the statement small enough. Let $\xi : U \longrightarrow X$ be this chart, and we denote by J again the almost complex structure $\xi_*(J)$ in $\mathbb{R}^{2n}$ induced by J via ξ. Consider $u_j' = \xi \circ u_j$. Then, also $|du_j'(0)| \to 0$ and $\lim_j \frac{|du_j'(0)|}{||du_j'||_{L^2(D_1)}} = \infty$.

Let, for each j, $\psi_j(z) = \epsilon|du_j'(0)|^{-1}u_j(z)$. Then, $|d\psi_j(0)| = \epsilon$ for all j, but $||d\psi_j||_{L^2(D_1)} = \epsilon|du_j'(0)|^{-1}||du_j'||_{L^2(D_1)} \to 0$. Moreover, ψ_j satisfies $\overline{\partial}_{J_j}\psi_j = 0$, where J_j is the almost complex structure in $\mathbb{R}^{2n}$ induced by J via the map $x \mapsto \epsilon|du_j'(0)|^{-1}\xi(x)$. Observe that, since $|du_j'(0)| \to 0$, the almost complex structures J_j converge to J_0 in the C^∞ sense.

Case 2: There is some $\epsilon > 0$ such that $|du_j(0)| \geq \epsilon$ for all j.

First of all, we prove that if $f : D_1 \longrightarrow \mathbb{R}_{\geq 0}$ is a smooth function satisfying $f(0) > 0$, then there exists some integer $k \geq 1$ and a point $x \in D_1$ such that $f(x) \geq 3^{k-1}f(0)$, $D_{3^{-k}}(x) \subset D_1$, and $sup_{D_{3^{-k}}(x)} < 3f(x)$.
Indeed, we can define in a recursive way a sequence of points $x_0, x_1, ... \in D_1$ satisfying $|x_j - x_{j-1}| \leq 3^{-j}$ for any j, as follows. Let $x_0 = 0$ and suppose $x_0, x_1, ..., x_{i-1}$ have been defined. By hypothesis,

$|x_{i-1} - x_0| \leq |x_{i-1} - x_{i-2}| + |x_{i-2} - x_{i-3}| + ... + |x_1 - x_0| \leq 3^{-(i-1)} + 3^{-(i-2)} + ... + 3^{-1} < \sum_{j\geq 1} 3^{-j} = 2/3$
so $\overline{D_{3^{-i}}(x_{i-1})} \subset D_1$. Choose $x_i \in \overline{D_{3^{-i}}(x_{i-1})}$ in such a way that $f(x_i) = sup\{f(y)|y \in \overline{D_{3^{-i}}(x_{i-1})}\}$.
For some $j \geq 1$ we must have $f(x_j) < 3f(x_{j-1})$. Indeed, if this is not the case, we would have $f(x_j) \geq 3^j f(0)$ for every j (using induction). Since $f(0) > 0$ and we have seen that $|x_j| \leq 2/3$ for every j, we have a contradiction with the fact that $f|_{\overline{D_{2/3}}}$ is bounded. Let k be the smallest $j \geq 1$ such that $f(x_j) < 3f(x_{j-1})$, and put $x = x_j$. These x and k satisfy the required properties.

With this, we can show that the assumption leads to a contradiction. In order to handle the case 1, we allow the maps u_j to be J_j-holomorphic maps where J_j are different almost complex structures that converge to an almost complex structure J in the C^∞ sense.

Suppose that the statement is false. Then, there is a sequence of pseudoholomorphic maps $u_j : D_1 \longrightarrow X$ such that $||du_j||_{L^2(D_1)} \to 0$ and $lim_j \frac{|du_j(0)|}{||du_j||_{L^2(D_1)}} = \infty$.

We prove first that for any $\epsilon > 0$ there is a $C > 0$ such that if $u : D_1 \longrightarrow X$ is pseudoholomorphic and $||du||_{L^2(D_1)} < C$ then $|du(0)| < \epsilon$.
By the result just proved, for each j there are $x_j \in D_1$ and $k_j \in \mathbb{N}$ such that:

1. $|du_j(x_j)| \geq 3^{k_j}|du_j(0)| \geq 3^{k_j} j||du_j||_{L^2(D_1)}$
2. there is an inclusion $D_{3^{-j}}(x_j) \subset D_1$
3. the number $\mu_j = sup_{D_{3^{-k_j}}(x_j)}|du_j|$ satisfies $\mu_j \leq 3|du_j(x_j)|$

Define, for each j, $\rho_j = min\{3^{-k_j}, \mu_j^{-1}\delta\}$, where δ is the number C coming from 4.2.3.
Define now maps $\psi_j : D_1 \longrightarrow X$, by $\psi_j(z) = j||du_j||_{L^2(D_1)}u(x_j + \rho_j z)$, which make sense because $\rho_j \leq 3^{-k_j}$ and then $x_j + \rho_j z \in D_{3^{-k_j}} \subset D_1$.

We have:

$$\overline{\partial}_I \psi_j = 0$$

$$sup_{D_1}|d\psi_j| \leq \rho_j \sup_{D_{3^{-k_j}}(x_j)} |du_j| = \rho_j \mu_j \leq \delta$$

$$||d\psi_j||_{L^2(D_1)} \leq ||du_j||_{L^2(D_{3^{-k_j}}(x_j))} \leq ||du_j||_{L^2(D_1)} \to 0$$

Then,

$$|du_j(0)| = \rho_j|du_j(x_j)| = min\{3^{-k_j}, \mu_j^{-1}\delta\}|du_j(x_j)| = min\{3^{-k_j}|du_j(x_j)|, \mu_j^{-1}\delta|du_j(x_j)|\}$$

We can estimate $3^{-k_j}|du_j(x_j)| \geq \epsilon$ and $\mu_j^{-1}\delta|du_j(x_j)| \geq 3^{-1}\delta$ which implies that $|d\psi_j(0)| \geq min\{\epsilon, 3^{-1}\delta\}$.

By proposition 4.2.3 there exists a subsequence ψ_{j_k} that converges in the C^1 sense uniformly on $D_{1/2}$ to a map $\psi : D_{1/2} \longrightarrow X$. Then, by the above estimates, we have that $||d\psi||_{L^2(D_{1/2})} = 0$,

so $d\psi$ is identically 0. But we have seen also that $|d\psi_j(0)| \geq min\{\epsilon, 3^{-1}\delta\}$ which implies that $|d\psi(0)| \neq 0$, a contradiction.

□

Corollary 4.2.5. *Let $u : D_1 \longrightarrow X$ be a J-holomorphic map. Then there exists an $\epsilon > 0$ such that if $||du||_{L^2(D_1)} < \epsilon$ then $|du|_{C^0(D_{1/2})} \leq K||du||_{L^2(D_1)}$ for some constant $K > 0$ independent of u.*

Proof. Consider, for each $w \in D_{1/2}$, the map $u_w : D_1 \longrightarrow X$ defined by $u_w(z) = u(w + z/4)$. This is well defined because $|w + z/4| \leq 1/2 + 1/4 < 1$. Moreover, $|du_w(0)| = |du(w)|/4$, and $||du_w||_{L^2(D_1)} \leq ||du||_{L^2(D_1)} < \epsilon$.

We can therefore apply the previous lemma to each of the maps u_w in order to obtain $|du(w)|/4 = |du_w(0)| < K||du_w||_{L^2(D_1)} \leq K/4||du||_{L^2(D_1)}$, and therefore, $|du(w)| < K||du||_{L^2(D_1)}$ for all $w \in D_{1/2}$, as wanted. □

4.3 Estimates on cylinders

We now study the behaviour of pseudoholomorphic curves defined on cylinders.
The goal of this section is to prove that in a finite cylinder, the energy density decays exponentially when we move towards the center of the cylinder, where these estimates are given by constants which are independent of the length of the cylinder. This estimate will be essential in our proof of the removal of singularities theorem (see theorem 4.4.1), and also in the analysis of bubbling that we will undertake in the following chapter.

For the next lemma, let $E_i(a) := E(u; [a(i-1), ai] \times S^1) = \frac{1}{2}\int_{[a(i-1),ai]\times S^1} |du(t,\theta)|^2$.

Lemma 4.3.1. *Let $a > 0$ be a real number, and let $u : S^1 \times [0, 3a] \longrightarrow \mathbb{C}^n$ be holomorphic (where we consider the standard complex structure on the cylinder). Then, we have $E_2(a) \leq \frac{\gamma}{2}(E_1(a) + E_3(a))$ for some $\gamma < 1$. Moreover, any $\cosh^{-1}(2a) < \gamma < 1$ works.*

Proof. First, suppose that $n = 1$. Then, $u : S^1 \times [0, 3a] \longrightarrow \mathbb{C}$ is holomorphic. We can develop u as a Fourier series in the θ variable: $u(\theta, t) = \sum_{k\in\mathbb{Z}} u_k(t)e^{ik\theta}$. Since u is holomorphic, it must satisfy the Cauchy-Riemann equations: $\partial_\theta u = i\partial_t u$. If we compute it with the Fourier series expression, we obtain: $\sum_{k\in\mathbb{Z}} u_k(t)ike^{ik\theta} = i\sum_{k\in\mathbb{Z}} u'_k(t)$. We must then have $u'_k = ku_k$ for all $k \in \mathbb{Z}$. Solving this differential equation, we see that $u_k(t) = C_k e^{kt}$, and we have $u(\theta, t) = \sum_{k\in\mathbb{Z}} C_k e^{kt} e^{ik\theta}$.

Now, $\partial_t u(\theta, t) = \sum_{k\in\mathbb{Z}} C_k k e^{kt} e^{ik\theta}$. From the Cauchy-Riemann equations, we have $|du|^2 = |\partial_t u|^2 + |\partial_\theta u|^2 = 2|\partial_t u|^2$. If we define the energy at t as $E(t) = \frac{1}{2}\int_0^{2\pi} |du(\theta, t)|^2 d\theta$, we get

$$E(t) = \int_0^{2\pi} \sum_{k,k'\in\mathbb{Z}} C_k C^*_{k'} kk' e^{(k+k')t} e^{i(k-k')\theta} = 2\pi \sum_{k\in\mathbb{Z}} |C_k|^2 k^2 e^{2kt}$$

Finally, we have

$$\begin{aligned} E_i(a) &= \int_{(i-1)a}^{ia} E(t)dt = 2\pi \sum_{k\in\mathbb{Z}} |C_k|^2 k^2 \int_{(i-1)a}^{ia} e^{2kt}dt \\ &= \pi \sum_{k\in\mathbb{Z},k\neq 0} |C_k|^2 k(e^{2kai} - e^{2ka(i-1)}) \end{aligned}$$

Then, there exists a γ with $0 < \gamma < 1$ that satisfies $E_2(a) \leq \frac{\gamma}{2}(E_1(a) + E_3(a))$ if and only if $e^{4ka} - e^{2ka} \leq \frac{\gamma}{2}(e^{2ka} - 1 + e^{6ka} - e^{4ka}) = \frac{\gamma}{2}(e^{4ka} - e^{2ka})(e^{2ka} - e^{-2ka})$ for all k, if and only if $1 \leq \gamma \cosh(2ka)$ for all $k \neq 0$, if and only if $\gamma \geq \cosh^{-1}(2ka)$ for all $k \neq 0$. Clearly, the worst case is $k = 1$, so it is enough to take γ with $\cosh^{-1}(2a) < \gamma < 1$.

The general case follows easily from the n=1 case. Indeed, let $u : S^1 \times [0, 3a] \longrightarrow \mathbb{C}^n$. We can write then $u = (u_1, ..., u_n)$, where $u_j : S^1 \times [0, 3a] \longrightarrow \mathbb{C}$ are holomorphic for $j = 1, ..., n$. Therefore, $E_i(a) = \frac{1}{2}\int_{a(i-1)}^{ai}\int_0^{2\pi} |du(\theta, t)|^2 dtd\theta = \frac{1}{2}\sum_{j=0}^n \int_{a(i-1)}^{ai}\int_0^{2\pi} |du_j(\theta, t)|^2 dtd\theta = \sum_{j=0}^n E_i^j(a)$. Hence, $E_2(a) = \sum_{j=0}^n E_2^j(a) \leq \frac{\gamma}{2}\sum_{j=0}^n (E_1^j(a) + E_3^j(a)) = \frac{\gamma}{2}(E_1(a) + E_3(a))$, as wanted. □

Proposition 4.3.2. *Let $a > 0$, $\eta > 0$ and let $u : S^1 \times [-\eta, 3a+\eta] \longrightarrow X$ be a pseudoholomorphic curve. Then, for any $\gamma \in (\cosh^{-1}(2a), 1)$, there exists an $\epsilon > 0$ such that $E_2(a) \leq \frac{\gamma}{2}(E_1(a) + E_3(a))$ whenever $|du|_{C^0} < \epsilon$.*

Proof. Suppose that this is not the case. Then, there exists a sequence of pseudoholomorphic maps $u_j : S^1 \times [-\eta, 3a + \eta] \longrightarrow X$ with $|du_j|_{C^0} \to 0$ and $E_2^j(a) > \frac{\gamma}{2}(E_1^j(a) + E_3^j(a))$, where $E_i^j(a) := E(u_j; [(i-1)a, ia] \times S^1)$.

Since $|du_j|_{C^0} \to 0$, we can assume without loss of generality (maybe passing to a subsequence) that the image of the u_j's is contained in a coordinate chart of X. In other words, we can assume that in fact $u_j : S^1 \times [-\eta, 3a + \eta] \longrightarrow \mathbb{R}^{2n}$, where we have in $\mathbb{R}^{2n}$ the almost complex structure J induced from the manifold X (that will not coincide, in general, with the standard complex structure J_0 of $\mathbb{R}^{2n}$). Without loss of generality we can assume that $J(0) = J_0$.

Consider now maps u'_j defined by $u'_j(z) = u_j(z)/|du_j|_{C^0}$. Then, $|du'_j|_{C^0} = 1$, so the energy density of the sequence remains bounded. Moreover, this maps are pseudoholomorphic maps with respect to almost complex structures J_j satisfying $J_j(z) = J(|du_j|_{C^0} z)$. Since $|du_j|_{C^0} \to 0$, $J(0) = J_0$, and the smoothness of the almost complex structure, we obtain that in fact J_j converges to J_0 in the C^∞ sense.
Finally, note that the maps u'_j satisfy $E_2^j(a) > \frac{\gamma}{2}(E_1^j(a) + E_3^j(a))$.

Now, by 4.2.3, we can extract a subsequence of u'_j that converge to a map $u : S^1 \times [0, 3] \longrightarrow \mathbb{C}^n$ which is holomorphic (which is the same as J_0-holomorphic). By the above remarks, it satisfy $E_2(a) \geq \frac{\gamma}{2}(E_1(a)+E_3(a)) > \frac{\gamma c}{2}(E_1(a)+E_3(a))$. But this is a contradiction with lemma 4.3.1. □

The following lemma will give us the link between the convexity relations we have found in the above proposition and the exponential decay behaviour of the energy density.

Lemma 4.3.3. *Let $e_0, e_1, ..., e_N$ nonnegative real numbers. Assume that there is $\gamma \in (0, 1)$ such that $e_j \leq \frac{\gamma}{2}(e_{j-1} + e_{j+1})$ for all $k = 1, ..., N - 1$. Then, there exists a $\sigma > 0$ and a constant C (independent of N) such that $e_j \leq C(e_0 + e_N)\exp(-\sigma \min\{j, N - j\})$, for all $j = 1, ..., N - 1$.*

Proof. First, suppose that we have $e_j = \frac{\gamma}{2}(e_{j-1} + e_{j+1})$ for all $k = 1, ..., N-1$. In this case, we can solve explicitly the recurrence equation, and we can obtain an expression for e_j in terms of e_0 and e_N. The solution is

$$e_j = \frac{e_0 r_+^N - e_n}{r_+^N - r_-^N} r_-^j + \frac{e_n - e_0 r_-^N}{r_+^N - r_-^N} r_+^j$$

where $r_+ = \frac{1}{\gamma}(1 + \sqrt{1-\gamma^2})$ and $r_- = \frac{1}{\gamma}(1 - \sqrt{1-\gamma^2})$. Observe that, since $\gamma < 1$, $r_+ > 1$ and $r_- < 1$, and note also that $r_-^{-1} = r_+$.

Now,

$$\frac{e_0 r_+^N - e_n}{r_+^N - r_-^N} = \frac{e_0 - e_n r_+^{-N}}{1 - (r_-/r_+)^N} \leq \frac{1}{1 - (r_-/r_+)}(e_0 + e_N)$$

On the other hand,

$$\frac{e_n - e_0 r_-^N}{r_+^N - r_-^N} = \frac{e_n - e_0 r_-^N}{1 - (r_-/r_+)^N} r_+^{-N} \leq \frac{1}{1 - (r_-/r_+)}(e_0 + e_N) r_+^{-N}$$

Therefore, putting $C = \frac{1}{1-(r_-/r_+)}$, we get

$$\begin{aligned} e_j &\leq C(e_0 + e_N)(r_-^j + r_+^{-(N-j)}) = C(e_0 + e_N)(r_+^{-j} + r_+^{-(N-j)}) \\ &\leq 2C(e_0 + e_N)\exp(-\sigma \min\{j, N-j\}) \end{aligned}$$

where $\sigma = log(r_+)$.

For the general case, given e_0, e_N we put $\tilde{e}_j$ the solution of $e_j = \frac{\gamma}{2}(e_{j-1} + e_{j+1})$ for all $k = 1, ..., N-1$. We define $r_j = \tilde{e}_j - e_j$. It is enough to see that $r_j \geq 0$ for all j. This is true for $j = 0, N$ (since $e_0 = \tilde{e}_0$ by hypothesis, and similarly with N, so $r_0 = r_N = 0$).

Observe that the r_j's satisfy $r_j \geq \frac{\gamma}{2}(r_{j-1} + r_{j+1})$ for $j = 1, ..., N-1$.

Suppose that for some $j > 0$ we have $r_j < 0$, and let j be the first one for which this happens. We will prove that in this case, we have $r_k < 0$ for all $k \geq j$, which is a contradiction with the fact that $r_N = 0$.
Since $0 > r_j$, we have $0 > r_j \geq \frac{\gamma}{2}(r_{j-1} + r_{j+1}) \geq \frac{\gamma}{2} r_{j+1}$, so we have $0 > r_j \geq \frac{\gamma}{2} r_{j+1}$, and in particular, $r_{j+1} < 0$. We prove by induction on n that $0 > r_{j+n} \geq \frac{(n+1)\gamma}{n+2} r_{j+n+1}$. We have just proved the case $n = 0$. Suppose now that this holds for n. Then,

$$r_{j+n+1} \geq \frac{\gamma}{2} r_{j+n} + \frac{\gamma}{2} r_{j+n+2} \geq \frac{(n+1)\gamma^2}{2(n+2)} r_{j+n+1} + \frac{\gamma}{2} r_{j+n+2}$$

So we have $\left(1 - \frac{(n+1)\gamma^2}{2(n+2)}\right) r_{j+n+1} \geq \frac{\gamma}{2} r_{j+n+2}$.
But observe that since $\gamma < 1$, we have

$$\left(1 - \frac{(n+1)\gamma^2}{2(n+2)}\right) \geq 1 - \frac{n+1}{2(n+2)} = \frac{n+3}{2(n+2)}$$

and since $r_{j+n+1} < 0$,

$$\left(1-\frac{(n+1)\gamma^2}{2(n+2)}\right)r_{n+j+1} \leq \frac{n+3}{2(n+2)}r_{n+j+1}$$

Finally, we get

$$\frac{n+3}{2(n+2)}r_{n+j+1} \geq \frac{\gamma}{2}r_{j+n+2}$$

and

$$r_{n+j+1} \geq \frac{(n+2)\gamma}{n+3}r_{j+n+2}$$

Since $0 > r_{n+j+1}$, it follows that $0 > r_{j+n+2}$ and the inductive hypothesis holds.

This finishes the proof of the lemma.

□

This last result, together with 4.3.2 implies that if $u : C \longrightarrow X$ is a pseudoholomorphic curve, then for every k has an exponential decay with t for t large enough. This is the content of the next proposition.

Proposition 4.3.4. *Let $u : C \longrightarrow X$ be a pseudoholomorphic curve with domain a semiinfinite cylinder $C = [0, +\infty) \times \mathbb{R}$, and suppose that u has finite energy. Then, there is some $T > 0$ such that $|\nabla^k u(t, \theta)| \leq C_k \exp(-\sigma t)$ for any $t > T$, where we can take, for any $\epsilon > 0$, $\sigma < 1 - \epsilon$.*

Proof. Fix $\epsilon > 0$. By proposition 4.3.2 and lemma 4.3.3, we obtain that, for any $a > 0$, $E_j(a) \leq C \exp(-\sigma' j)$ for j big enough, where $\sigma' = \frac{1}{\gamma}(1 + \sqrt{1-\gamma^2})$, and γ any real number satisfying $\frac{1}{\cosh(2a)} < \gamma < 1$.

Observe that for $a > 0$ small enough, we have $\frac{1}{\cosh(2a)} < e^{-2a(1-\epsilon)}$. Taking such an a, we can take a γ such that $\frac{1}{\cosh(2a)} < \gamma < e^{-2a(1-\epsilon)}$. Then we have $\sigma' > log\frac{1}{\gamma} > 2a(1-\epsilon)$.

Note now that $E_j(a) = ||du_{[(j-1)a,ja]}||^2_{L^2}$, so in fact we obtain $||du_{[(j-1)a,ja]}||^2_{L^2} \leq C \exp(-\frac{\sigma'}{a} ja) \leq C \exp(-2(1-\epsilon)ja)$ and finally $||du_{[(j-1)a,ja]}||_{L^2} \leq C' \exp(-(1-\epsilon)t)$, for any $t \in [(j-1)a, ja]$ and j big enough. This means that we can take $\sigma > 1 - \epsilon$.

Fix a $\delta > 0$. Since the energy of u is finite, there is some $T > 0$ such that the energy of $u|_{[T,+\infty)\times S^1}$ is less than $\delta > 0$.

Now we want to apply corollary 4.2.5 in order to obtain a bound

$$|du|_{C^0([(j-1)a,ja])} < C' \exp(-\sigma' j)$$

Indeed, we can cover $[-\eta + (j-1)a, ja + \eta] \times S^1$ by finitely many open sets $U_1, ..., U_n$ in the cylinder and charts $\xi_i : U_i \longrightarrow D_1$ such that $\xi_i^{-1}(D_{1/2})$ also cover the cylinder.

Define now $u_j : D_1 \longrightarrow X$ by $u_j = u \circ \xi_i$. Choosing appropriately the charts ξ_i, we can assume that $||du_j||_{L^2(D_1)} < \delta$. By 4.2.5, we have then that $|du_j|_{C^0(D_{1/2})} \leq C||du_j||_{L^2(D_1)}$.

So we obtain $|du|_{C_0(\xi_j^{-1}(D_{1/2}))} \leq C_j||du||_{L^2(\xi_j^{-1}(D_{1/2}))}$, where in general the C_j is different for each j, since it depends on the C^0 norm of $d\xi_j$. Since this is true for every j, choosing $C = \max\{C_1, ..., C_n\}$ we finally obtain our desired bound

$$|du|_{C^0([(j-1)a,ja])} \leq C||du||_{L^2([(j-1)a,ja])}$$

Therefore for any $t \in [(j-1)a, ja]$ we have

$$|du(t,\theta)| \leq C' \exp(-\frac{\sigma'}{a}t) \leq C' \exp(-\sigma t)$$

where $\sigma > 1 - \epsilon$ for $t > T > 0$.

We now show the analogous C^0 bound for $\nabla^k u$. Repeating the same argument as above, we see by corollary 4.2.2, that, for any k,

$$||du||_{L^2_k([(j-1)a,ja])} < C'' \exp(-\sigma t)$$

Now, applying Sobolev lemma, we get our desired conclusion,

$$|du|_{C^r([(j-1)a,ja])} < C'' \exp(-\sigma t)$$

In particular, we get the result of the statement.

□

We finish this section by proving a result about annuli that we will need in the proof of the Gromov compactness theorem. Recall that an annulus is conformally equivalent to a finite cylinder, therefore we can apply the methods of this section.

We define $A(r,R) := \overline{B}_R - B_r$, with $r < R$. Observe that $A(r,R)$ is conformally equivalent to the cylinder $C_{r,R} := [0, \log R - \log r] \times S^1$ via the conformal map $\psi : C_{r,R} \longrightarrow A(r,R)$ defined by $\psi(t,\theta) = e^{t+\log r + i\theta}$.

Lemma 4.3.5. *For every $\sigma < 1$ there exist constants $\epsilon > 0$ and $C > 0$ such that for every pseudoholomorphic map $u : A(r,R) \longrightarrow X$ with $E(u) < \epsilon$ we have, for all T such that $e^T r < e^{-T} R$:*

$$E(u; A(e^T r, e^{-T} R)) \leq C e^{2\sigma T} E(u)$$

Proof. Put $u' : C \longrightarrow X$ the map $u' := u \circ \psi$. Since the energy is conformally invariant, and we have $\psi(\log r + T, \theta) = e^T e^{i\theta}$ and $\psi(\log R - \log r - T, \theta) = Re^{-T} e^{i\theta}$, we obtain that

$$E(u; A(e^T r, e^{-T} R) = E(u'; [T, \log R - \log r - T] \times S^1)$$

Now, pick $a > 0$ so that $T = ka$, $(k'-1)a \leq \log R - \log r - T < k'a$, with $k' > k$, and $(N-1)a \leq \log R - \log r < Na$. Then, using lemma 4.3.3, we get:

$$E(u'; [T, \log R - \log r - T] \times S^1) \leq \sum_{j=k}^{k'-1} E_j(a)$$

$$\leq \sum_{j=k}^{k'-1} C(E_1(a) + E_N(a)) \exp(-2\sigma a \min\{j, N-j\})$$

if $E(u) < \epsilon$, by taking $\epsilon > 0$ as in proposition 4.3.2, and $a > 0$ small enough as in the proof of proposition 4.3.4 to get the inequality with the desired σ.

Observe that $E_1(a) + E_N(a) \leq E(u') = E(u)$. Observe also that $k = \frac{T}{a}$ and $k' = \frac{\log R - \log r - T}{a} = N - \frac{T}{a}$, so $\min\{j, N-j\} \leq \frac{T}{a}$ for all $k \leq j \leq k'$.

So we conclude

$$E(u'; [T, \log R - \log r - T] \times S^1) \leq C' E(u) e^{2\sigma T}$$

as wanted. □

4.4 Removal of singularities

In this section we will prove the theorem of removal of singularities, which is a key theorem in the proof of the Gromov compactness theorem.

In this section we put $D = D_1$ and $D^* = D - \{0\}$.

Theorem 4.4.1 (Removal of singularities)**.** *Let $u_0 : D^* \longrightarrow X$ be a pseudoholomorphic map with finite energy. Then, u extends to a pseudoholomorphic map $u : D \longrightarrow X$.*

Recall that $C = [0, +\infty) \times S^1$ and D^* are conformally equivalent. A conformal map between them is $h : C \longrightarrow D^*$ given by $h(t, \theta) = e^{-t+i\theta}$.

Proposition 4.4.2. *Let $u_0 : D^* \longrightarrow X$ be a pseudoholomorphic curve with finite energy. Then, u_0 extends to a continuous map $u : D \longrightarrow X$.*

Proof. Since D^* is conformally equivalent to $C = [0, +\infty) \times S^1$ and the energy is a conformal invariant, we can see u_0 as a pseudoholomorphic curve $u_0' : C \longrightarrow X$ with finite energy (defined by $u' = u \circ h$). Applying proposition 4.3.4, we have that $|du_0'(t, \theta)| < C \exp(-\sigma t)$, for some $\sigma > 0$.

From this, we get that u_0 has a continuous extension to all of D. Indeed, suppose that x_k is a sequence of points in D such that $x_k \to 0$. Write $(t_k, \theta_k) = h^{-1}(x_k)$. Then, we have

$$\begin{aligned} |u_0(x_p) - u_0(x_q)| &= |u_0'(t_p, \theta_p) - u_0'(t_q, \theta_q)| \leq \int_\gamma |du'(t, \theta)| \\ &\leq C \int_\gamma \exp(-\sigma t) = C \int_{t_p}^{t_q} \exp(-\sigma t) < \epsilon \end{aligned}$$

if $p, q > N$ for some N, because $\int_0^\infty \exp(-\sigma t) dt$ is convergent. This shows that the sequence x_k is Cauchy, and using the fact that X is complete (because it is compact), we obtain that $u_0(x_k)$ is convergent to a point $x \in X$. Moreover, the argument also shows that the limit is independent of the sequence (if this were not the case, we could choose a sequence $x_k \to 0$ that is not Cauchy). Therefore, we can define $u : D \longrightarrow X$ extending u_0 continuously by putting $u(0) = x$. □

Now that we have that u_0 has a continuous extension to D, we need to show that this extension is in fact C^∞. We will do it by proving a regularity theorem. But in order to put that regularity theorem to use, we need to start with a $L^2_3(D)$ map. Therefore, we will show first that the continuous extension we have just obtained is in $L^2_3(D)$.

The following useful result will allow us to take charts of X where the almost complex structure has a nice expression.

Lemma 4.4.3. *Let X be a manifold and J an almost complex structure on X. Let $p \in X$. Then, there exists a chart (U, χ) centered at p, with $\chi : U \longrightarrow \chi(U) \subset \mathbb{R}^{2n}$, such that $\chi_*(J) : \mathbb{R}^{2n} \longrightarrow End(\mathbb{R}^{2n})$ (which we also denote by J) satisfy:*

$$(i) J(0) = J_0$$

$$(ii) DJ(0) = 0$$

In particular, $|J(q) - J(0)| \leq C|q|^2$.

Proof. We can always take a chart centered at p in which $J(0) = J_0$. Indeed, it is enough to take a chart centered at p and compose it with a linear isomorphism of $\mathbb{R}^{2n}$ that takes $J(0)$ to J_0.

Now, observe that if $J : \mathbb{R}^{2n} \longrightarrow End(\mathbb{R}^{2n})$ is such that $J^2 = -Id$ and $J(0) = J_0$, and $\psi : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$ is a diffeomorphism satisfying $\psi(0) = 0$ and $D\psi(0) = Id$, then $\psi_* J$ (which recall is defined by $\psi_* J(x) = D\psi(x)^{-1} J(\psi(x)) D\psi(x)$) still satisfy $\psi_* J(0) = J_0$. Therefore, it is enough to prove that there exists such a ψ that moreover satisfy $D(\psi_* J)(0) = 0$.
We impose the condition and get:

$$\begin{aligned} 0 &= D_x(\psi_* J(x))|_{x=0} = D\psi(0)^{-1} J(\psi(0))(D_x(D\psi)(0)) \\ &+ D\psi(0)^{-1}(D_x J(\psi(0)) D\psi(0)) D\psi(0) \\ &- D\psi(0)^{-1} D_x(D\psi(0)) D\psi(0)^{-1} J(\psi(0)) D\psi(0) \end{aligned}$$

where we have used that, since $D\psi(0)^{-1} D\psi(0) = Id$, we have

$$D_x(D\psi^{-1})(0) D\psi(0) + D\psi(0)^{-1} D_x D\psi(0) = 0$$

Taking into account that $D\psi(0) = Id$ and $\psi(0) = 0$, this simplify to:

$$0 = J(0)(D_x(D\psi)(0)) + D_x J(0) - D_x(D\psi)(0) J(0)$$

That is, $0 = [J(0), D_x D\psi(0)] + D_x J(0)$.

Now, we must see that we can choose ψ satisfying these equations. Observe that this is a condition only on the second derivatives of ψ at 0. The only restriction we have on the second derivatives is that $D_{x_k} D_{x_l} \psi = D_{x_l} D_{x_k} \psi$ for all k, l. Indeed, given numbers a_{kl} satisfying $a_{kl} = a_{lk}$, we can consider $\psi(x) = \sum_k x_k + \frac{1}{2} \sum_{k,l} a_{kl} x_k x_l$. Then, $D\psi(0) = Id$ and $D_{x_k} D_{x_l} \psi(0) = a_{kl}$. The implicit function theorem assures us that ψ is a diffeomorphism in some neighbourhood of 0, and this is all we need.

We can see that there are numbers $D_{x_k}D_{x_l}\psi(0)$ satisfying $0 = [J(0), D_xD\psi(0)] + D_xJ(0)$ by seeing that these conditions together with $D_{x_k}D_{x_l}\psi(0)$ is an underdetermined linear system with $D_{x_k}D_{x_l}\psi(0)$ as unknowns.
Indeed, we have $4n^2$ unknowns. Let us count now the number of equations. We have $2n(2n-1)/2 = 2n^2 - n$ equations from the conditions of the symmetry of the partial derivatives. Let us now count the number of equations coming from $[J(0), D_xD\psi(0)] = -D_xJ(0)$. For each $1 \leq k \leq 2n$ we have one equation $[J(0), D_{x_k}D\psi(0)] = -D_{x_k}J(0)$. We can write

$$D_{x_k}J(0) = \begin{pmatrix} a_k & b_k \\ c_k & d_k \end{pmatrix}$$

Observe that since $J^2 = -1$ we have $D_{x_k}J(0)J(0) + J(0)D_{x_k}J(0) = 0$. Recalling that $J(0) = J_0$, this implies that $d_k = -a_k$ and $c_k = b_k$. Therefore,

$$D_{x_k}J(0) = \begin{pmatrix} a_k & b_k \\ b_k & -a_k \end{pmatrix}$$

If we write

$$D_{x_k}D\psi(0) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

we see that we can write our equation as

$$\begin{pmatrix} -B-C & A-D \\ A-D & B+C \end{pmatrix} = -\begin{pmatrix} a_k & b_k \\ b_k & -a_k \end{pmatrix}$$

Therefore, we have in fact only n distinct equations for each k. Since we have $2n$ variables, we get a total of $2n^2$ equations. Taking into account also the symmetry equations, we have a total of $2n^2 + 2n^2 - n = 4n^2 - n$. Hence, we have less equations than unknowns, and the system has a solution. This finishes our proof.
The final assertion follows from the Taylor expansion of J around 0, taking into account that $DJ(0) = 0$. □

Let $u' : D \longrightarrow X$ be a continuous map such that $u'|_{D^*}$ is smooth and satisfies $\bar{\partial}_J u' = 0$. From now on, we choose a chart (U, χ) as in the previous lemma centred at the point $u'(0)$, and consider the map $u := \chi \circ u$. Then, in order to prove the removal of singularities theorem, it will be enough to prove that any J-holomorphic map $u : D \longrightarrow \mathbb{C}^n$ is smooth at 0.

Observe that the pseudoholomorphic condition $\overline{\partial}_J u = 0$ can be written as $\overline{\partial}_{J_0} u = g$, where $g = -(J(u) - J_0)\partial_y u$.

We can now prove that $g \in L^2_2(D_{1/2})$.

Proposition 4.4.4. *Let $u : D \longrightarrow \mathbb{C}^n$ be a continuous map that is pseudoholomorphic in D^*. Then, $g = -(J(u) - J_0)\partial_y u \in L^2_2(D_{1/2})$.*

Proof. Put $u_0 : D^* \longrightarrow \mathbb{C}^n$ the restriction of u to D^*. Then, since D^* and $C = [0, +\infty) \times S^1$ are conformally equivalent, we can see u_0 as a map $u_0' : C \longrightarrow \mathbb{C}^n$ defined by $u_0'(t, \theta) = u_0(e^{-t}, \theta)$. By proposition 4.3.4, we have that $|\nabla^k u(t, \theta)| < Cexp(-\sigma t)$ for some $\sigma > 1 - \epsilon$ and t big enough. Now, observe that $(\partial_i u_0)(e^{-t}, \theta) = r^{-1}(\partial_i u_0')(t, \theta)$. Therefore, $|du_0(r, \theta)| \leq Cr^{-1}r^{\sigma} \leq Cr^{-\epsilon}$. Similarly, one see that $|\nabla^2 u_0(r, \theta)| \leq Cr^{-1-\epsilon}$ and $|\nabla^3 u_0(r, \theta) \leq Cr^{-2-\epsilon}$ for r small enough.

Now, we use lemma 4.4.3 to obtain a chart ξ centered at $u(0)$ with the property that $|J(p)-J_0| \leq C|p|^2$, where we call J again to the expression of the almost complex structure J in this chart. Then, consider the expression of u in this local coordinates, that is, $\xi \circ u$, which by simplicity will be called also u in what follows.

Since $|du_0(r,\theta)| \leq r^{-\epsilon}$ for r small enough, we have that

$$|u(z) - u(0)| \leq \int_{[0,z]} |du(z)| \leq C|z|^{1-\epsilon}$$

So $|u(z)| \leq C|z|^{1-\epsilon}$ for $|z|$ small enough.

If $z \in D$ is such that $|u(z)| \leq |z|^{1-\epsilon}$ we have that $|J(u(z)) - J_0| \leq C|u(z)|^2 \leq C|z|^{2-2\epsilon}$. Therefore, since by our previous estimates we have $|\partial_y u(z)| \leq |z|^{-\epsilon}$, we get $|g(z)| \leq C|z|^{2-3\epsilon}$. Since $|g|^2$ is integrable in a neighbourhood of 0, we obtain that $g \in L^2$.

We can see similarly that $\nabla g \in L^2$ and $\nabla^2 g \in L^2$.

Indeed,

$$\nabla g = \nabla J(u)\nabla(u)\partial_y u + (J(u) - J_0)\nabla\partial_y u$$

and

$$\begin{aligned} \nabla^2 g &= \nabla^2 J(u)(\nabla u)^2\partial_y u + \nabla J(u)\nabla^2 u\partial_y u + \nabla J(u)\nabla u\nabla\partial_y u \\ &+ \nabla J(u)\nabla u\nabla\partial_y u + (J(u) - J_0)\nabla^2\partial_y u \end{aligned}$$

Taking into account the previous estimates for $||\nabla u||_{L^2}$ and $||\nabla^2 u||_{L^2}$, and the fact that by our choice of chart, we have $|\nabla J(z)| \leq C|z|$ and $|\nabla^2 J(z)| \leq C'$, we obtain the bounds $|\nabla g(z)| \leq C|z|^{1-3\epsilon}$ and $|\nabla^2 g(z)| \leq C|z|^{-3\epsilon}$. Therefore, if we take ϵ small enough, we get $\nabla g \in L^2$ and $\nabla^2 g \in L^2$. That is, $g \in L^2_2$. □

Lemma 4.4.5. *u is a weak solution of $\overline{\partial}_{J_0} u = g$ in D.*

Proof. Observe first that the formal adjoint of the operator $\overline{\partial}_{J_0} = \partial_x + J_0\partial_y$ is $-\partial_x + J_0\partial_y = -\partial_{J_0}$

Therefore, in order to check that u is a weak solution of the equation, we only need to check that for any $\psi \in C^\infty(D)$ we have $-\int_D u\overline{\partial_{J_0}\psi} = \int_D g\overline{\psi}$.

Fix a small disk of radius $\epsilon > 0$ around the origin. Since we know that u is a strong solution of the equation in D^*, we have that $-\int_{D-D_\epsilon} u\overline{\partial_{I_0}\psi} = \int_{D-D_\epsilon} g\overline{\psi}$. Therefore,

$$|-\int_D u\overline{\partial_{I_0}\psi} - \int_D g\overline{\psi}| = |\int_{D_\epsilon} u\overline{\partial_{I_0}\psi} + \int_{D_\epsilon} g\overline{\psi}| \leq \int_{D_\epsilon} |u\overline{\partial_{I_0}\psi}| + \int_{D_\epsilon} |g\overline{\psi}|$$

But the first integral is bounded by $|u\overline{\partial_{J_0}\psi}|_{C^0} A(D_\epsilon)$ while the second one is as small as we want because the integrand is in $L^2(D)$. Hence, when ϵ goes to 0, also does $-\int_D u\overline{\partial_{J_0}\psi} = \int_D g\overline{\psi}$. □

Putting together the last two results, we obtain:

Corollary 4.4.6. *u is in $L^2_3(D)$.*

Proof. The first statement follows from proposition 4.4.4, lemma 4.4.5 and the standard elliptic regularity theorem, since $\overline{\partial}_{J_0} u = g$, and $g \in L^2_2$. □

Now we can show that in fact our weak solution u is a strong solution in all of D, using a bootstrapping method. In order to prove this, we will need the following embedding lemma.

Lemma 4.4.7. *There is an embedding $L^2_1([0,1]^2) \longrightarrow L^4([0,1]^2)$. In particular, multiplication of functions gives a map $L^2_1([0,1]^2) \times L^2_1([0,1]^2) \longrightarrow L^2([0,1]^2)$, that is, if $f, g \in L^2_1$, then $fg \in L^2$.*

Proof. This is a special case of a much more general embedding theorem. However, since we only need this special case, and according to our philosophy of using only L^2 techniques, we give an elementary proof for this result.

First observe that the second part of the lemma follows from the first, since if we have such an embedding, $\int |fg|^2 \leq (\int |f|^4)^{1/2} (\int |g|^4)^{1/2}$, so if $f, g \in L^2_1$, then $fg \in L^2$ (where we have used the Cauchy-Schwarz inequality).

Now, let us prove the existence of such an embedding. In fact, we will prove that $||f||_{L^4} \leq 6^{1/4} ||f||_{L^2_1}$. For that, it suffices to prove that $\int_{[0,1]\times[0,1]} |f|^4 dxdt \leq 6$ for all $f : [0,1] \times [0,1] \longrightarrow \mathbb{C}$ satisfying $f \in L^2_1$.

Fix such a function f and define $f_t : [0,1] \longrightarrow \mathbb{C}$ for any $t \in [0,1]$ by $f_t(x) = f(t,x)$. Denoting by $< g, h >= g\bar{h}$ (the usual Hermitian product), we have:

$$\begin{aligned}
\left| \frac{\partial}{\partial t} \int_0^1 |f_t(x)|^2 dx \right| &= \left| \int_0^1 2Re \left\langle f_t(x), \frac{\partial f_t(x)}{\partial t} \right\rangle dx \right| \\
&\leq 2 \left(\int_0^1 |f_t(x)|^2 dx \right)^{1/2} \left(\int_0^1 \left| \frac{\partial f_t(x)}{\partial t} \right|^2 dx \right)^{1/2} \\
&\leq \left(\int_0^1 |f_t(x)|^2 dx \right) + \left(\int_0^1 \left| \frac{\partial f_t(x)}{\partial t} \right|^2 dx \right)
\end{aligned}$$

where we have used successively the Cauchy-Schwartz inequality and the Arithmetic-Geometric inequality.
Integrating, we get

$$\begin{aligned}
|||f_{\tau'}|||^2_{L^2} - ||f_\tau||^2_{L^2}| &= \left| \int_\tau^{\tau'} \frac{\partial}{\partial t} \int_0^1 |f_t(x)|^2 dx \right| \leq \left| \int_\tau^{\tau'} \int_0^1 (|f_t(x)|^2 + |\frac{\partial f_t(x)}{\partial t}|^2) dxdt \right| \\
&\leq ||f||^2_{L^2_1} = 1
\end{aligned}$$

for any $0 \leq \tau \leq \tau' \leq 1$.

Since

$$\int_0^1 ||f_t||^2_{L^2} dt = ||f||^2_{L^2} \leq 1$$

we must have $||f_{\tau_0}||_{L^2}^2 \leq 1$ for some $\tau_0 \in [0,1]$. Then it follows from the last estimate that $||f_t||_{L^2}^2 \leq 2$ for all $t \in [0,1]$.

Now, since the roles of x and t are symmetric, we can interchange them and repeat all the argument to get $\int_0^1 |f_t(x)|^2 dx \leq 2$ for all $x \in [0,1]$.

Define, for any $t, x \in [0,1]$

$$F_t(x) := \int_0^x |f_t(y)|^2 dy$$

Then, we have

$$F_t(x) \leq ||f_t||_{L_2}^2 \leq 2$$

Writing now $|f|^4 = |f|^2|f|^2$ and using integration by parts in the x variable, we get

$$\begin{aligned}
\int_0^1 \int_0^1 |f(t,x)|^4 dxdt &= \int_0^1 \left(\int_0^1 |f_t(x)|^4 dx \right) dt \\
&= \int_0^1 \left(|f_t|^2 F_t|_{x=0}^1 - \int_0^1 \frac{\partial |f_t(x)|^2}{\partial x} F_t(x) dx \right) dt \\
&= \int_0^1 \left(|f_t(1)|^2 F_t(1) - \int_0^1 \frac{\partial |f_t(x)|^2}{\partial x} F_t(x) dx \right) dt \\
&= \int_0^1 |f_t(1)|^2 F_t(1) dt - \int_0^1 \int_0^1 \frac{\partial |f_t(x)|^2}{\partial x} F_t(x) dxdt
\end{aligned}$$

We estimate the two terms separately.

For the first term, we have

$$\int_0^1 |f_t(1)|^2 F_t(1) dt \leq 2 \int_0^1 |f_t(1)|^2 dt \leq 2 \cdot 2 = 4$$

And for the second term, using again Cauchy-Schwartz and the Arithmetic-Geometric inequality, we have

$$\begin{aligned}
\left| \int_0^1 \int_0^1 \frac{\partial |f_t(x)|^2}{\partial x} F_t(x) dxdt \right| &\leq 2 \int_0^1 \int_0^1 \left| \frac{\partial |f_t(x)|^2}{\partial x} \right| dxdt \\
&= 2 \int_0^1 \int_0^1 2Re \left| \left\langle f_t(x), \frac{\partial f_t(x)}{\partial x} \right\rangle \right| dxdt \\
&\leq 4 \left(\int_0^1 \int_0^1 |f_t(x)|^2 dxdt \right)^{1/2} \left(\int_0^1 \int_0^1 \left| \frac{\partial f_t(x)}{\partial x} \right|^2 dxdt \right)^{1/2} \\
&\leq 2 \int_0^1 \int_0^1 |f_t(x)|^2 dxdt + \int_0^1 \int_0^1 \left| \frac{\partial f_t(x)}{\partial x} \right|^2 dxdt \\
&= 2||f||_{L_1^2}^2 = 2
\end{aligned}$$

Putting all together, we finally obtain

$$||f||_{L^4}^4 \leq 6$$

as wanted. □

Proposition 4.4.8. *Let $u : D \longrightarrow \mathbb{C}^n$ be a map such that $u \in L_3^2(D)$, u is a weak solution of the equation $\partial_{J_0} u = g$, and $u|_{D^*}$ is a pseudoholomorphic curve from the punctured disk into $\mathbb{C}^n$. Then, u is a pseudoholomorphic curve from the disk into $\mathbb{C}^n$.*

Proof. First of all, observe that $u \in L_3^2(D)$ implies that $u \in C^1(D)$.

We will show that $u \in C^\infty(D)$. In order to do that, by the Sobolev lemma, it is enough to prove that $u \in L_k^2(D)$ for all $k > 0$. Since we know already that $u|_{D^*}$ is smooth, in order to prove that $u \in L_k^2(D)$, it is enough to prove that $u \in L_k^2(D_\epsilon)$ for some $\epsilon > 0$.

We proceed by induction on k. By hypothesis, $u \in L_3^2(D)$. Suppose now that $u \in L_k^2(D)$. In order to show that $u \in L_{k+1}^2(D_\epsilon)$ for some $\epsilon > 0$, it suffices to show that for all $h \in \mathbb{R}^2$ (or equivalently, for all h such that $|h|$ is small enough), $||u^h||_{L_k^2} < C$ for some constant $C > 0$ independent of h. (Recall that $u^h(x) = \frac{u(x+h)-u(x)}{|h|}$).

By Gårding's inequality, we have

$$||u^h||_{L_k^2(D)} \leq C(||\overline{\partial}_{J_0}(u^h)||_{L_{k-1}^2(D)} + ||u^h||_{L_{k-1}^2})$$

Since $u \in L_k^2(D)$, we have $||u^h||_{L_{k-1}^2} \leq ||u||_{L_k^2}$, which is independent of h. Let us examine the other term, $||\overline{\partial}_{J_0}(u^h)||_{L_{k-1}^2(D_\epsilon)}$.

Since $\overline{\partial}_{J_0}$ is a differential operator with constant coefficients, we have

$$\overline{\partial}_{J_0}(u^h) = (\overline{\partial}_{J_0}(u))^h = ((J(u) - J_0)\partial_x u)^h$$

Now,

$$\begin{aligned}
& ((J(u) - J_0)\partial_x u)^h(x) \\
= \; & \frac{1}{|h|}((J(u(x+h)) - J_0)\partial_x u(x+h) - (J(u(x)) - J_0)\partial_x u(x)) \\
= \; & \frac{1}{|h|}((J(u(x+h)) - J_0)\partial_x u(x+h) - (J(u(x)) - J_0)\partial_x u(x+h)) \\
+ \; & \frac{1}{|h|}((J(u(x)) - J_0)\partial_x u(x+h) - (J(u(x)) - J_0)\partial_x u(x)) \\
= \; & (J(u) - J_0)^h \tau_h \partial_x u + (J(u) - J_0)(\partial_x u)^h
\end{aligned}$$

Therefore,

$$\begin{aligned}
||\overline{\partial}_{J_0}(u^h)||_{L_{k-1}^2(D_\epsilon)}|| &\leq ||(J(u) - J_0)^h \tau_h \partial_x u||_{L_{k-1}^2(D_\epsilon)} + ||(J(u) - J_0)(\partial_x u)^h||_{L_{k-1}^2(D_\epsilon)} \\
&\leq C||(J(u) - J_0)^h \partial_x u||_{L_{k-1}^2(D_\epsilon)} + ||(J(u) - J_0)(\partial_x u)^h||_{L_{k-1}^2(D_\epsilon)}
\end{aligned}$$

where in the last step we have used that $|\tau_h \partial_x u(z)| \leq C|\partial_x u(z)|$ (using that the derivative is continuous).

Now we can do a similar analysis as in lemma 4.2.1.

We start by analysing the base case, $k = 3$. In this case, we have that $||(J(u) - J_0)^h \partial_x u||_{L^2} \leq |\partial_x u|_{C^0} ||(J(u) - J_0)^h||_{L^2} \leq C||\nabla(J(u) - J_0)||_{L^2}$, where we have used that $\partial_x u$ is continuous.

Similarly, $||\nabla((J(u) - J_0)^h)\partial_x u||_{L^2}$ and $||(J(u) - J_0)^h \nabla(\partial_x u)||_{L^2}$ are bounded independently of h, because

$$||\nabla((J(u) - J_0)^h)\partial_x u||_{L^2} \leq |\partial_x u|_{C^0} ||\nabla^2(J(u) - J_0)||_{L^2}$$

and

$$||(J(u) - J_0)^h \nabla(\partial_x u)||_{L^2} \leq C||\nabla(J(u) - J_0)\nabla(\partial_x u)||_{L^2} \leq |\nabla(J(u) - J_0)|_{C^0} ||\partial_x u||_{L^2_1}$$

Finally, $||\nabla^2(J(u) - J_0)^h \partial_x u||_{L^2}$, $||(J(u) - J_0)^h \nabla^2(\partial_x u)||_{L^2}$ and $||\nabla(J(u) - J_0)^h)\nabla(\partial_x u)||_{L^2}$ are bounded independently of h, the first two with similar arguments as above (using that $\partial_x u \in L^2_2$ and $J(u) - J_0 \in L^2_2$, hence they are continuous), and the last one, using that $||\nabla(J(u) - J_0)^h)\nabla(\partial_x u)||_{L^2} \leq C||\nabla^2(J(u) - J_0)\nabla(\partial_x u)||_{L^2}$, and applying lemma 4.4.7 taking into account that $\nabla^2(J(u) - J_0), \nabla(\partial_x u) \in L^2_1$.

Therefore, $||(J(u) - J_0)(\partial_x u)^h||_{L^2_2}$ is bounded independently of h.

Now, the same kind of analysis allows us to write

$$||(J(u) - J_0)(\partial_x u)^h||_{L^2_2} \leq C + |J(u) - J_0|_{C^0} ||u^h||_{L^2_3}$$

Indeed, we can bound all the terms $||\nabla^p(J(u) - J_0)\nabla^q((\partial_x u)^h)||^2_L$ for $p+q \leq 2$ as before, except for the term $p = 0, q = 2$, for which we get the bound $|J(u) - J_0|_{C^0} ||\partial_x(u^h)||_{L^2_2} \leq |J(u) - J_0|_{C^0} ||u^h||_{L^2_3}$. Summing up, we have:

$$||u^h||_{L^2_3} \leq C + C'|J(u) - J_0|_{C^0} ||u^h||_{L^2_3}$$

Taking the radius of the disk, ϵ, small enough, we can ensure that $C'|J(u) - J_0|_{C^0} < 1/2$. Then, we get:

$$||u^h||_{L^2_3(D_\epsilon)} \leq 2C$$

Hence, since $||u^h||_{L^2_3(D_\epsilon)}$ is bounded independently of h, we obtain that $u \in L^2_4(D_\epsilon)$.

We can now consider the general case, $k > 3$. We proceed in the same way as before, only that in this case we don't need to use lemma 4.4.7. The analogous analysis as in the previous case (taking into account that if $u \in L^2_k$ then $u \in C^{k-2}$ and $I(u) - I_0 \in C^{k-2}$) allows us to conclude that $u \in L^2_{k+1}(D_\epsilon)$, for a suitable ϵ.

Therefore, we obtain that in fact $u \in L^2_k(D_{1/2})$ for all k, and by the Sobolev embedding theorem, $u \in C^\infty(D)$ (since our argument shows that it is smooth at the origin, and it is smooth outside the origin by hypothesis). That is, u is a C^∞ solution to $\overline{\partial}_J u = 0$ in D.

□

With this, we finish the proof of the removal of singularities theorem.

Chapter 5

Proof of the Gromov compactness theorem

In this chapter we finish the proof of the Gromov compactness theorem, using the results obtained in the last chapter. In the first section we analyse how bubbling occurs and prove a first rough compactness theorem for pseudoholomorphic curves. In the second section we study in more depth the phenomenon of bubbling, and in the final section we put all together and interpret bubbling in terms of Gromov convergence, giving a proof of the Gromov compactness theorem.

In this chapter we follow closely [M-S1].

5.1 Bubbling

In this section we will see how bubbling appears as a consequence of the conformal invariance of the energy. Recall that we have seen a manifestation of this phenomenon in section 2.3. Moreover, as a first step towards the Gromov compactness theorem, we will prove the following compactness theorem.

Proposition 5.1.1. *Let J_n be ω-compatible almost complex structures on X converging in the C^∞ sense to an ω-compatible almost complex structure J. Let $u_n : S^2 \longrightarrow X$ be a sequence of J_n-holomorphic curves with $\sup_n E(u_n) < \infty$. Then there exists a finite set of points $F = \{z^1, ..., z^n\} \subset S^2$, a subsequence of u_n (which we again denote by u_n for simplicity) and a J-holomorphic curve $u : S^2 \longrightarrow X$ such that:*

1. *u_n converge uniformly to u with all derivatives in compact subsets of $S^2 - F$.*
2. *For every j and every $\epsilon > 0$ such that $\overline{D}_\epsilon(z^j) \cap F = \{z^j\}$, the limit*

$$m_\epsilon(z^j) := \lim_n E(u_n; \overline{D}_\epsilon(z_j))$$

exists and is a continuous function of ϵ, and

$$m(z^j) := \lim_{\epsilon \to 0} m_\epsilon(z^j) \geq \hbar$$

3. For every compact subset $K \subset S^2$ with $F \subset \circ K$,

$$E(u;K) + \sum_{j=1}^{l} m(z^j) = \lim_n E(u_n;K)$$

First of all, we prove the following important proposition, telling us that a non-constant pseudoholomorphic sphere has a lower bound to its energy depending only on the target manifold.

Proposition 5.1.2. *There exists a constant $\hbar > 0$, depending only on the symplectic manifold (X,ω) and the ω-compatible almost complex structure J, such that if $u : S^2 \longrightarrow X$ is a non-constant pseudoholomorphic map, then $E(u) \geq \hbar$.*

Moreover, if we denote $\hbar(J)$ the constant corresponding to the almost complex structure J, and J_n is a sequence of ω-compatible almost complex structures converging to J in the C^∞ sense, then $\hbar(J_n)$ converges to $\hbar(J)$.

Proof. Assume that the proposition is false. Then, for each $j \in \mathbb{N}$ there exists a non-constant pseudoholomorphic map $u_j : S^2 \longrightarrow X$ such that $E(u_j) < \frac{1}{j}$.

Let $\epsilon > 0$. Cover S^2 by finitely many open charts with chart maps $\xi_i : U_i \longrightarrow D_1$ such that the sets $\xi_i^{-1}(D_{1/2})$ also cover S^2. Then, taking j large enough we have that $|d(u_j \circ \xi_i^{-1})|_{C^0(D_{1/2})}$ is as small as we want for all i, by corollary 4.2.5. Therefore, we have $|du_j|_{C^0} < \epsilon$ for j large enough. In this case, we have that $\operatorname{diam}(u_j(S^2)) < 2\pi\epsilon$. If we take ϵ small enough so that any set with diameter less than $2\pi\epsilon$ is contained in a coordinate chart of M homeomorphic to the open ball of radius 1, B_1 (this always exists since X is compact, so an open covering with coordinate charts homeomorphic to B_1 has a Lebesgue number), then $u_j(S^2)$ is contained in such a coordinate chart of X. Since these coordinate charts of X are contractible, we have that u_j is homotopic to a constant map. Therefore, $E(u_j) = 0$ (because energy is a topological invariant of a pseudoholomorphic curve), which is a contradiction with the fact that u_j is not constant.

For the second part, observe that the only place where the almost complex structure enters is in the definition of ϵ necessary to apply corollary 4.2.5. But this number depends continuously on the almost complex structure J, since by looking at the proof of corollary 4.2.5, we see that it is used in order to apply proposition 4.2.3, which can be applied to any sequence of J_n-holomorphic maps, where J_n converge to J in the C^∞ sense. □

Observe that in this proposition we make full use of the condition that J is ω-compatible, since this is what allows us to say that the energy depends only on the class of homology that the curve represents. Because of this proposition, one often says that the energy of a pseudoholomorphic sphere is quantized.

Now we can give a proof of proposition 5.1.1.

Suppose that u_n is a sequence of J_n-holomorphic curves $u_n : S^2 \longrightarrow X$ such that $\sup_n E(u_n) < \infty$. If we have $\sup_n |du_n|_{C^0(S^2)} < \infty$, we can apply proposition 4.2.3 and conclude that a subsequence converges in all of S^2. Therefore the only failure of compactness occur when we have $\sup_n |du_n|_{C^0(S^2)} = \infty$. Let $z_n \in S^2$ be such that $c_n := |du_n(z_n)| = |du_n|_{C^0(S^2)}$. Passing to a subsequence we can assume that z_n converges to a point $z_0 \in S^2$ and c_n diverges to infinity.

Consider a map $\phi : U \longrightarrow \phi(U) \subset S^2$ where $U \subset \mathbb{C}$, $\phi(0) = z_0$ and ϕ^{-1} is a complex chart in S^2 such that $\phi(0) = z_0$. We can assume without loss of generality that $\frac{1}{2} \leq |d\phi(z)| \leq 2$ for all $z \in U$ and $|d\phi(0)| = 1$. Moreover, we can also assume without loss of generality (maybe passing to a subsequence) that $z_n \in \phi(U)$ for all n. Put $u'_n = u_n \circ \phi$ and $z'_n = \phi^{-1}(z_n)$. Then, $|du'_n(z'_n)| = |du(z_n)||d\phi(z'_n)|$, and therefore $\frac{1}{2}c_n \leq |du'_n(z'_n)| \leq c_n$ and $\lim_n z'_n = 0$.

Now we pick $\epsilon > 0$ such that $\overline{D}_\epsilon(z_n) \subset U$ and consider the reparametrized sequence $v_n : \overline{D}_{\epsilon c_n} \longrightarrow X$ defined as $v_n(z) = u'_n(z_n + z/c_n)$. Then, $|dv_n(0)| = |du'_n(z_n)|\frac{1}{c_n} \geq \frac{1}{2}$, and $|dv_n|_{C^0} \leq |du_n|_{C_0}\frac{1}{c_n} \leq 2$. Moreover, we have $E(v_n; \overline{D}_{\epsilon c_n}) = E(u'_n; \overline{D}_\epsilon(z_n)) \leq E(u'_n) \leq E(u_n)$, where we have used the conformal invariance of energy twice, since both ϕ and $z \mapsto z_n + z/c_n$ are conformal maps.

We can now apply proposition 4.2.3 to the sequence v_n (since we have seen that it has uniformly bounded derivative) to get a J-holomorphic curve $v : \mathbb{C} \longrightarrow X$. Moreover, by the previous bounds we have $|dv(0)| \geq \frac{1}{2}$ (in particular, v is not constant, so $E(v) > 0$), and $E(v) \leq \sup_i E(u_i) < \infty$, so v has finite energy. Consider the curve $\tilde{v} : \mathbb{C} - \{0\} \longrightarrow X$ defined by $\tilde{v}(z) = v(1/z)$. By the conformal invariance of energy, $\tilde{v}$ has finite energy. Therefore, by the removal of singularities theorem, $\tilde{v}$ extends to a J-holomorphic curve defined in the whole $\mathbb{C}$. Therefore, we can see v as a J-holomorphic curve $v : S^2 \longrightarrow X$. Moreover, as we have noted above, v is not constant, so by proposition 5.1.2 we have that $E(v) > \hbar$. Observe also that for any $R > 0$ and every $\epsilon > 0$ we have

$$E(v; \overline{D}_R) = \lim_n E(v_n; \overline{D}_R) = \lim_n E(u'_n; \overline{D}_{R/c_n}(z'_n)) \leq \lim_n E(u'_n; \overline{D}_\epsilon(z'_n))$$

where in the last step we have used that c_n diverges to infinity. Taking the limit $R \to \infty$, we finally obtain $E(v) \leq \liminf_n E(u'_n; \overline{D}_\epsilon(z'_n))$ for every $\epsilon > 0$. Therefore, also for any $\epsilon > 0$ we have $\hbar \leq E(v) \leq \liminf_n E(u_n; \overline{D}_\epsilon(z_n))$, where we have used that ϕ is conformal, so $E(u'_n; \overline{D}_\epsilon(z'_n)) = E(u_n; \overline{D}_\epsilon(z_n))$.

The maps $v : S^2 \longrightarrow X$ arising in this way are called bubbles, and it is said that v bubbles off in the limit. Observe that we have just seen that any sequence of points in S^2 where the derivatives du_n diverge gives rise to a non constant bubble in the limit. Moreover, the energy $E(v)$ comes from arbitrarily small neighbourhoods of the points z_n after a suitable conformal rescaling of the sequence. Since any bubble has energy at least $\hbar > 0$ and the energy of the sequence is bounded, this implies that there can only be a finite number of bubbles that bubble off in the limit.

Let us formalize this argument in order to have a proper proof of proposition 5.1.1.

Proof of Proposition 5.1.1. We start with a sequence u_n of J_n-holomorphic curves with bounded energy, that is, $\sup_n E(u_n) < \infty$. Suppose $\sup_n |du_n|_{C^0} = \infty$. Then we can apply the previous construction and passing to a subsequence u^1_n find a sequence of points z_n of S^2 converging to $z \in S^2$ and such that $|du_n(z_n)| \to \infty$, so that a bubble bubbles off at z in the limit. Now, if it exists, we can take a subsequence u^2_n and points z^1_n such that $z^1_n \to z^1 \neq z$ and $|du^2_n(z^1_n)| \to \infty$, so that by the previous argument a bubble bubbles off at z^1 in the limit. We continue this process until it stops.

Now we will show that this process must necessarily stop in a finite number of steps. Indeed, it must stop in at most $\frac{\sup_n E(u_n)}{\hbar}$ steps. For suppose we have a subsequence $(u^N_n)_n$ and points $z^1, ..., z^N \in S^2$ with $N > \frac{\sup_n E(u_n)}{\hbar}$ such that there are sequences $(z^j_n)_n \subset S^2$ for

$j = 1, ..., N$ satisfying $z_n^j \to z^j$ and $|du_n^N(z_n^j)| \to \infty$ for each $j = 1, ..., N$. Choose $\epsilon > 0$ such that the balls $\overline{D}_\epsilon(z^1), ..., \overline{D}_\epsilon(z^N)$ are all disjoint. By the above estimates we have that $\hbar \leq \liminf_n E(u_n; \overline{D}_\epsilon(z^j))$ for every $j = 1, ..., N$. So, since the balls are disjoint,

$$\sup_n E(u_n) \geq \liminf_n E(u_n) \geq \sum_{j=1}^N \liminf_n E(u_n; \overline{D}_\epsilon(z^j)) \geq N\hbar > \sup_n E(u_n)$$

a contradiction.

Put $F = \{z^1, ..., z^N\}$, and for simplicity call u_n to the subsequence u_n^N obtained at the end of the process. Consider a countable family of compact sets K_i such that $K_i \subset \overset{\circ}{K}_{i+1} \subset K_{i+1}$ for each i and such that $\bigcup_i K_i = S^2 - F$. By construction we have $|du_n|_{C^0(K_i)} < \infty$ for each i. Therefore, we can apply proposition 4.2.3 and conclude that there exists a subsequence of the u_n that converges to a J-holomorphic curve $u^i : K_i \longrightarrow X$ in the C^∞ sense in K_i for each i. Finally, by using a diagonal argument we obtain a subsequence of the u_n that converges to a J-holomorphic curve $u^\infty : S^2 - F \longrightarrow X$ in the C^∞ sense (that is, it converges uniformly in compact with all its derivatives). Now just note that $E(u^\infty) < \sup_n E(u_n) < \infty$, so that it has finite energy and then by N applications of the removal of singularities theorem (at each of the points $z_1, ..., z_N$) we obtain a pseudoholomorphic curve $u : S^2 \longrightarrow X$ satisfying the condition (i).

Now we take a further subsequence such that $m_\epsilon(z^j) = \lim_n E(u_n; \overline{D}_\epsilon(z^j))$ exist for all j. This subsequence fulfills all the conditions of the proposition. It is clear that it still satisfies (i). For (ii), take $0 < \epsilon' < \epsilon$. Then, since u_i converge uniformly in the annulus $\overline{D}_\epsilon(z^j) - D_{\epsilon'}(z^j)$, we have that $m_\epsilon(z^j) = m_{\epsilon'}(z^j) + E(u; \overline{D}_\epsilon(z^j) - D_{\epsilon'}(z^j))$, so clearly $m_{\epsilon'}(z^j)$ is defined for all $\epsilon' < \epsilon$ and is continous in ϵ'. Since as we have seen before, we have $m_\epsilon(z^j) \geq \hbar$ for all $\epsilon > 0$, we have $m(z^j) \geq \hbar$ for all j.

Finally, for (iii), since the $\overline{D}_\epsilon(z^j)$ are disjoint we have that

$$E(u; S^2 - \bigcup_{j=1}^n D_\epsilon(z^j)) = \lim_i E(u_i; K) - \sum_{j=1}^n \lim_i E(u_i; D_\epsilon(z^j)) = \lim_i E(u_i; K) - \sum_{j=1}^n m_\epsilon(z^j)$$

and taking the limit when $\epsilon \to 0$, we obtain

$$E(u; S^2 - F) = \lim_i E(u_i; K) - \sum_{j=1}^n m(z^j)$$

□

Observe that we have now a compactness theorem, but we have not said anything about the behaviour of the limit at the singular points in F. By the above argument we know that under a suitable rescaling of the sequence a bubble appears at each of this points. But this is not the end of the story, since multiple bubbling or nested bubbling can appear at a given point, so that if we just consider the bubble given by the argument above (which is a single bubble) we are losing energy in the process. This happens because we have not taken into account the behaviour of the sequence in small annuli around the singular points. In the next section we take a closer look at this and use the results obtained to give a full proof of the Gromov compactness theorem.

5.2 Soft rescaling

In this section we will improve the results we have just obtained by analyzing in detail the way in which bubbling occurs. In particular, we will see that by refining our bubbling argument, choosing apropriately the rescaling of the u_n's, we can obtain a bubble that connects to the limit of u_n obtained in proposition 5.1.1. In order to prove this refined compactness theorem, we will have to study carefully the behaviour of the sequence in annuli around a singular point. This annuli can be thought as the necks that join the principal bubble with a developing bubble, so that in the limit the necks contract to a single point and the resulting limit is singular.

The main result of this section is the following:

Theorem 5.2.1. *Let J_n be ω-compatible almost complex structures on X that converge in the C^∞ sense to an ω-compatible almost complex structure J. Let $z_0 \in \mathbb{C}$ and $r > 0$. Let $u_n : \overline{D}_r(z_0) \longrightarrow X$ be a sequence of J_n-holomorphic curves and $u : \overline{D}_r(z_0) \longrightarrow X$ a J-holomorphic map such that the following holds*

(a) u_n converges to u in the C^∞ sense on $\overline{D}_r(z_0) - \{z_0\}$.

(b) $m_0 := \lim_{\epsilon \to 0} \lim_n E(u_n; \overline{D}_\epsilon(z_0))$ exists and $m_0 > 0$.

Then, there exist a subsequence (which we denote also by u_n for simplicity), a sequence of Moebius transformations $\psi_n \in G$, a J-holomorphic sphere $v : S^2 \longrightarrow X$ and finitely many distinct points $z_1, ..., z_l, z_\infty \in S^2$ such that the following holds

(i) ψ_n converges to z_0 in the C^∞ sense on compact subsets of $S^2 - \{z_\infty\}$.

(ii) The sequence $v_n := u_n \circ \psi_n$ converges to v in the C^∞ sense on all compact subsets of $S^2 - \{z_1, ..., z_l, z_\infty\}$, and the limits

$$m_j := \lim_{\epsilon \to 0} \lim_n E(v_n; \overline{D}_\epsilon(z_j))$$

exist and satisfy $m_j > 0$ for all $j = 1, ..., l$.

(iii) $E(v) + \sum_{j=1}^{l} m_j = m_0$

(iv) If v is constant, then $l \geq 2$.

Proof. Since the proof is quite long, we will split it in 4 steps.

Step 1: We may assume without loss of generality that $z_0 = 0$ and $|du_n(z)|$ reaches its maximum at $z = 0$, for all n.

Since by (a), u and its derivative converge uniformly away from z_0, the sequence $|du_n|$ is uniformly bounded in the annulus $\overline{D}_r(z_0) - D_\epsilon(z_0)$ for every $\epsilon > 0$. We have also, for every $\epsilon > 0$, that

$$\lim_n \sup_{\overline{D}_\epsilon(z_0)} |du_n| = \infty$$

for otherwise, we would had $E(u_n; \overline{D}_\epsilon(z_0)) = \int_{\overline{D}_\epsilon(z_0)} |du_n|^2 \leq C\epsilon$ for some $0 < C < \infty$ independent of n, and hence $m_0 \leq \lim_{\epsilon \to 0} C\epsilon = 0$, a contradiction with (b).

Therefore, every sequence $z_n \in \overline{D}_R(z_0)$ with $|du_n(z_n)| = \sup_{\overline{D}_R(z_0)} |du_n|$ converges to z_0 (otherwise, the sequence stays infinitely often in some annulus $\overline{D}_r(z_0) - D_\epsilon(z_0)$ where the energy density is bounded, contradicting the hypothesis that the energy density reaches its maximum at the points of the sequence). We may then assume $z_n \in \overline{D}_{r/2}(z_0)$ for all i (by passing to a subsequence of the u_n).

Consider the maps $\tilde{u}_n : \overline{D}_{R/2} \longrightarrow X$ defined by $\tilde{u}_n(z) := u_n(z + z_n)$ (observe that the map $z \mapsto z + z_n$ is in G). These new maps clearly satisfy the hypothesis (a) and (b), with r replaced by $r/2$ and z_0 replaced by 0. Moreover, if they satisfy the conclusion for some sequence $\psi_n \in G$, so do the old maps with the sequence $z \mapsto \psi_n(z) + z_n$. This finishes step 1.

Step 2: Definition of the sequence $\psi_n \in G$ satisfying condition (ii) and condition (i) with $z_0 = 0$ and $z_\infty = \infty$.

Choose $0 < \delta < \hbar$ be such that lemma 4.3.5 holds with the ϵ of the lemma equal to δ. By passing to a subsequence, we may choose a sequence of numbers $\delta_n > 0$ such that

$$E(u_n; \overline{D}_{\delta_n}) = m_0 - \frac{\delta}{2}$$

Let us see that such a choice (for n large enough) is possible. Observe that, by passing to a subsequence, for a fixed n, we have that $E(u_n; \overline{D}_\epsilon)$ is strictly positive and decreases continuously with $\epsilon \to 0$ with limit 0. On the other hand, for a fixed $\epsilon > 0$, we have that $m_\epsilon = \lim_n E(u_n; \overline{D}_\epsilon) \geq m_0$ and is decreasing with ϵ. Therefore, for n large enough, we have $E(u_n; \overline{D}_\epsilon) > m_0 - \frac{\delta}{2}$ and it follows that there is some δ_n with $E(u_n; \overline{D}_{\delta_n}) = m_0 - \frac{\delta}{2}$.

We see now that $\delta_n \to 0$. Indeed, suppose this is not the case. Then there is an $\epsilon_0 > 0$ such that for any n_0 there exists $n > n_0$ such that $\delta_n > \epsilon_0$. So, passing to a subsequence we may assume that $\delta_n > \epsilon_0$ for all n. But this is a contradiction, since we would have:

$$m_0 = \lim_{\epsilon \to 0} \lim_n E(u_n; \overline{D}_\epsilon) \leq \lim_n E(u_n; \overline{D}_{\delta_n}) < m_0$$

Let us explain intuitively what this condition means. When we take $\overline{D}_{\delta_n}$, we are taking neighbourhoods of z_0 such that all bubbling near z_0 occur inside this neighbourhoods. This is because m_0 is the energy of all the bubbles occurring at z_0 in the limit, while outside this neighbourhood we have energy at most $m_0 - \hbar/2$ and any non-constant bubble has energy at least $\hbar$. Therefore, by considering these neighbourhoods we are considering all the zone near z_0 relevant to bubbling.

Consider now the sequences $\psi_n \in G$ and $v_n : \overline{D}_{R/\delta_n} \longrightarrow X$ defined by $\psi_n(z) := \delta_n z$ and $v_n(z) := u_n(\delta_i z)$. Since $\delta_n \to 0$, we have that ψ_n satisfies (i) with $z_\infty = \infty$. Moreover,

$$|dv_n(0)| = \sup_{\overline{D}_r} |dv_n| \tag{5.1}$$

if $r < R/\delta_n$ (recall that $|du_n(z)|$ reaches its maximum at $z = 0$). The energy of v_n in $\overline{D}_r$ is bounded by the energy of u_n in $\overline{D}_R$, for any fixed r and n large enough. Then, we can apply proposition 5.1.1 to the v_n and obtain a subsequence (again called v_n) and points $z_1, ..., z_l \in S^2$ such that (ii) holds. Let $F = \{z_1, ..., z_l\} \subset S^2$. Suppose that $F \neq \emptyset$. Then, equation 5.1 implies that $0 \in F$, by our discussion on bubbling in the previous section. This finishes the proof of step 2.

Now we have already obtained the desired subsequence of the u_n and the desired maps ψ_n. It only remains to show that conditions (iii) ad (iv) are fulfilled. In order to do that, we will need the following identity.

Step 3: $\lim_{R\to\infty} \lim_n E(u_n; \overline{D}_{R\delta_n}) = m_0$

The proof is by contradiction. We assume that it is false. Then, there exists $\rho > 0$ and a subsequence (also called u_n) such that for every $R \geq 1$,

$$\lim_n E(u_n; \overline{D}_{R\delta_n}) \leq m_0 - \rho$$

Indeed, this follows from the fact that, for a fixed $R \geq 1$, we have

$$\lim_n E(u_n; \overline{D}_{R\delta_n}) \leq m_0$$

because for any $\epsilon > 0$ there is an n_0 such that $R\delta_n < \epsilon$, hence for all $n \geq n_0$, $E(u_n, \overline{D}_{R\delta_n}) \leq E(u_n; \overline{D}_\epsilon)$, so $\lim_n E(u_n; \overline{D}_{R\delta_n}) \leq \lim_n E(u_n; \overline{D}_\epsilon)$, and therefore $\lim_n E(u_n; \overline{D}_{R\delta_n}) \leq \lim_{\epsilon\to 0} \lim_n E(u_n; \overline{D}_\epsilon) = m_0$.

So, using the definition of δ_n, we have that for every $R \geq 1$,

$$\lim_n E(u_n; A(\delta_n, R\delta_n)) = \lim_n E(u_n; \overline{D}_{R\delta_n}) - \lim_n E(u_n; \overline{D}_{\delta_n}) \leq \frac{\delta}{2} - \rho \tag{5.2}$$

Now we show that there is a sequence of numbers $\epsilon_n > 0$ such that $\lim_n \epsilon_n = 0$ and $\lim_n E(u_n; \overline{D}_{\epsilon_n}) = m_0$. Indeed, by the definition of m_0, for each $l \in \mathbb{N}$ there exists $\epsilon_l \in (0, 1/l)$ and $n_l \in \mathbb{N}$ such that $|E(u_n; \epsilon_l) - m_0| \leq 1/l$ for all $n > n_l$. Without loss of generality we can assume also that $\epsilon_{l+1} < \epsilon_l$ and $n_{l+1} > n_l$ for all l. Then, we can define $\epsilon_n := \epsilon_l$ for $n_l \leq n < n_{l+1}$.

Since $R\epsilon_n \to 0$ for each fixed $R \geq 1$, we also have for $R \geq 1$

$$\lim_n E(u_n; \overline{D}_{R\epsilon_n}) = m_0 \tag{5.3}$$

where we use that, as with the δ_n, we have $\lim_n E(u_n; \overline{D}_{R\epsilon_n}) \leq m_0$ and $\lim_n E(u_n; \overline{D}_{\epsilon_n}) \leq \lim_n E(u_n; \overline{D}_{R\epsilon_n})$.

Observe that the ratios ϵ_n/δ_n are not bounded on any subsequence. Because otherwise, by passing to a subsequence we get $\epsilon_n < C\delta_n$ for some $C \geq 1$ and that would imply $m_0 = \lim_n E(u_n; \overline{D}_{\epsilon_n}) \leq \lim_n E(u_n; \overline{D}_{C\delta_n}) \leq m_0 - \rho$, a contradiction.

So, by the definition of ϵ_n and δ_n, we have

$$\lim_n E(u_n; A(\delta_n, \epsilon_n)) = \delta/2 \tag{5.4}$$

and

$$\lim_n \frac{\delta_n}{\epsilon_n} = 0$$

Therefore, for n large enough we have $\delta_n < \epsilon_n$ and we can consider the annuli $A(\delta_n, \epsilon_n)$. Now we will examine the behaviour of the sequence on these annuli. For each $T > 0$ (such that it makes

sense) we split each of these annuli in three annuli $A(\delta_n, e^T\delta_n)$, $A(e^T\delta_n, e^{-T}\epsilon_n)$ and $A(e^{-T}\epsilon_n, \epsilon_n)$.

We will prove

$$\lim_{T\to\infty}\lim_n E(u_n, A(\delta_n, e^T\delta_n)) = \delta/2 \tag{5.5}$$

which leads to a contradiction with equation 5.2 and establishes step 3.

To prove it, consider the sequence $w_n : \overline{D}_{1/\epsilon_n} \longrightarrow X$ defined by $w_n(z) := u_n(\epsilon_n z)$. By passing to a subsequence, we may assume that w_n converges in the C^∞ sense to a pseudoholomorphic map w outside a finite set of points. Using that $\delta_n/\epsilon_n \to 0$, we have for each $T \geq 1$ that

$$\lim_n E(w_n; A(e^{-T}, e^T)) = \lim_n E(u_n; A(e^{-T}\epsilon_n, e^T\epsilon_n)) \leq \lim_n E(u_n; A(\delta_n, e^T\epsilon_n)) \leq \frac{\hbar}{2}$$

where in the last inequality we have used equation 5.3 and the definition of δ_n. This implies that w_n converges to a constant in the C^∞ sense on $\mathbb{C} - \{0\}$, because the bound imply that there is no bubbling there, hence for any compact $K \subset \mathbb{C} - \{0\}$ w_n converges to w in the C^∞ sense on K and in particular $|dw|_{C^0(K)} = \lim_n |dw_n|_{C^0(K)} = \lim_n \epsilon_n |du_n|_{C^0(K/\epsilon_n)} = 0$, where we have used that u_n converges in the C^∞ sense away from 0.

Therefore, for all $T \geq 0$,

$$\lim_n E(u_n; A(e^{-T}\epsilon_n, \epsilon_n)) = \lim_n E(w_n; A(e^{-T}, 1)) = E(w; A(e^{-T}, 1)) = 0$$

Now, by equation 5.4, the definition of δ and lemma 4.3.5 with $\sigma = 1/2$, we have that there exists a constant $C > 0$ such that

$$\lim_n E(u_n; A(e^T\delta_n, e^{-T}\epsilon_n)) \leq Ce^{-T}\frac{\delta}{2}$$

Since $\lim_n E(u_n; A(\delta_n, \epsilon_n)) = \delta/2$ and $\lim_n E(u_n; A(e^{-T}\epsilon_n, \epsilon_n)) = 0$, we have that

$$\lim_n E(u_n; A(\delta_n, e^T\delta_n)) \geq (1 - Ce^{-T})\frac{\delta}{2}$$

Taking the limit when $T \to \infty$, and taking into account that equation 5.4 shows that $\lim_n E(u_n; A(\delta_n, e^T\delta_n)) \leq \frac{\delta}{2}$, this proves 5.5, and thus establishes step 3.

Step 4: Finishing the proof.

Let us prove (iii). Since $\lim_{R\to\infty}\lim_n E(u_n; \overline{D}_{R\delta_n}) = m_0$, and recalling the definition of δ_n, we have that

$$\lim_{R\to\infty}\lim_n E(v_n; \overline{D}_R) = \lim_{R\to\infty}\lim_n E(u_n; \overline{D}_{R\delta_n}) = m_0$$

and for every n,

$$E(v_n; \overline{D}_1) = E(u_n; \overline{D}_{\delta_n}) = m_0 - \frac{\delta}{2} \geq m_0 - \frac{\hbar}{2}$$

Therefore, all the bubble points $z_1, ..., z_l$ of the sequence v_i lie in the disk $\overline{D}_1$. Fix a number $s > 1$. Then, we have

$$\begin{aligned}
m_0 &= \lim_{R\to\infty}\lim_n E(v_n;\overline{D}_R)\\
&= \lim_{R\to\infty}\lim_n E(v_n;\overline{D}_R - D_s) + \lim_{R\to\infty}\lim_n E(v_n;D_s)\\
&= \lim_{R\to\infty} E(v;\overline{D}_R - D_s) + \lim_n E(v_n;D_s)\\
&= E(v;\mathbb{C}-D_s) + \lim_{\epsilon\to 0}\lim_n E\left(v_n; D_s - \bigcup_{j=1}^{l} D_\epsilon(z_j)\right) + \sum_{j=1}^{l} m_j\\
&= E(v) + \sum_{j=1}^{l} m_j
\end{aligned}$$

where we have used the identity of step 3. This proves (iii).

Let us prove (iv). Since all bubble points of v_n lie in the closed unit disk, if v is constant and $1 < R_1 < R_2$, we have

$$\lim_n E(v_n; A(R_1,R_2)) = E(v; A(R_1,R_2)) = 0$$

Therefore the limit $\lim_n E(v_n;\overline{D}_R) = \lim_n E(u_n;\overline{D}_{R\delta_n})$ is independent of $R > 1$. So by the identity of step 3, $\lim_n E(v_n;\overline{D}_R) = m_0$ for every $R > 1$. Since $E(v_n;\overline{D}_1) = m_0 - \frac{\delta}{2}$, we have $\lim_n E(v_n;\overline{D}_1) < m_0$. This is only possible if there is a bubble point z of the sequence v_n lying on the boundary of the unit disk (that is, with $|z| = 1$). But we have seen at step 2 that if the set of singular points F for v is not empty, then also $0 \in F$, so $l = \#F \geq 2$, as wanted.

This finishes the proof. □

Now we prove a lemma telling us that if when a bubble bubbles off in the limit there is no loss of energy, then the bubble connects with the principal component, in the sense that the point where the bubble develops in the principal component and that of the bubble have the same image. We will prove this first in a special situation (the following lemma) and then proving that the general situation can be reduced to the situation of the lemma. This, together with the last result, makes plausible that the limit of a sequence of pseudoholomorphic maps converge to a stable map. We will prove this in the next section.

Lemma 5.2.2. *Let J_n be a sequence of ω-compatible almost complex structures converging in the C^∞ sense to an ω-compatible almost complex structure J. Let $r > 0$ and a positive sequence $\delta_n \to 0$. Let $u_n : A(\delta_n/r, r) \longrightarrow X$ be a sequence of $J-n$-holomorphic curves satisfying*

(i) u_n converge to a J-holomorphic curve $u : \overline{D}_r \longrightarrow X$ in the C^∞ sense on $\overline{D}_r - \{0\}$.

(ii) $u_n(\delta_n\cdot)$ converge to a J-holomorphic curve $v : S^2 - \overline{D}_{1/r} \longrightarrow X$ in the C^∞ sense on $\mathbb{C} - \overline{D}_{1/r}$ (where here we think of $\mathbb{C}$ as a subset of S^2).

(iii) $\lim_{\rho\to 0}\lim_n E(u_n; A(\delta_n/\rho, \rho)) = 0$

Then, we have

$$u(0) = v(\infty)$$

Proof. Observe that the annuli $A(r,R)$ is conformally equivalent to a finite cylinder, so the total energy remains invariant. Fix δ as the ϵ in lemma 4.3.5. By (iii), we have that there exists $\rho > 0$ (which we take such that $2\rho < r$) and n_0 such that for all $n > n_0$ we have $E(u_n; A(\delta_n/(2\rho), 2\rho)) < \delta$. Then by lemma 4.3.5 with $\sigma = 1/2$ and $T = \log 2$ we obtain that $E(u_n; A(\delta_n/\rho e^T, e^{-T}\rho) \leq 2CE(u_n)$. Now, by using corollaries 4.2.5 and 4.2.2 applied to obtain bounds over all the annulus (using the standard argument of covering it by suitable small open sets, applying there the corollaries and using the compactness of the annulus), we get

$$|du_n|_{C^0(A(\delta_n/\rho,\rho))} \leq C'||du_n||_{L^2(A(\delta_n/\rho,\rho))} \leq C''||du_n||_{L^2(A(\delta_n/(2\rho),2\rho))}$$

and

$$d(u_n(z), u_n(z')) \leq \int_{[z,z']} |du_n| \leq r|du_n|_{C^0(A(\delta_n/\rho,\rho))} \leq C||du_n||_{L^2(A(\delta_n/(2\rho),2\rho))}$$

for $z, z' \in A(\delta_n/\rho, \rho)$.

From this, taking the limit when $n \to \infty$

$$d(u(\rho), v(1/\rho)) = \lim_n d(u_n(\rho), u_n(\delta_n/\rho)) \leq C \lim_n ||du_n||_{L^2(A(\delta_n/(2\rho),2\rho))}$$

To finsh, just observe that $||du_n||_{L^2(A(\delta_n/(2\rho),2\rho))} = \sqrt{E(u_n; A(\delta_n/(2\rho), 2\rho))}$ and that, by (iii), $\lim_{\rho\to 0} \lim_n E(u_n; A(\delta_n/(2\rho), 2\rho)) = 0$, so taking the limit $\rho \to 0$, we get

$$d(u(0), v(\infty)) = \lim_{\rho\to 0} d(u(\rho), v(1/\rho)) \leq \lim_{\rho\to 0} \lim_n E(u_n; A(\delta_n/(2\rho), 2\rho)) = 0$$

That is, $u(0) = v(\infty)$, as wanted. □

Proposition 5.2.3. *Let J_n be a sequence of ω-compatible almost complex structures converging in the C^∞ sense to an ω-compatible almost complex structure J. Let $z_0 \in \mathbb{C}$ and $u, u_n : \overline{D}_r(z_0) \longrightarrow X$ a J-holomorphic and J_n-holomorphic (respectively) curves satisfying the following conditions*

(a) u_n converges to u in the C^∞ sense on $\overline{D}_r(z_0) - \{z_0\}$

(b) The limit $m_0 := \lim_{\epsilon\to 0} \lim_n E(u_n; \overline{D}_\epsilon(z_0))$ exists and $m_0 > 0$.

Suppose moreover that $\psi_n \in G$, $v : S^2 \longrightarrow X$ and $z_1, ..., z_n, z_\infty \in S^2$ satisfy

(i) ψ_n converges in the C^∞ sense to z_0 on every compact subset of $S^2 - \{z_\infty\}$.

(ii) The sequence $v_n := u_n \circ \psi_n$ converges to v in the C^∞ sense on $S^2 - \{z_1, ..., z_l, z_\infty\}$ and the limits

$$m_j := \lim_{\epsilon\to 0} \lim_n E(v_n; \overline{D}_\epsilon(z_j))$$

exist and satisfy $m_j > 0$ for all $j = 1, ..., l$.

(iii) $E(v) + \sum_{j=1}^{l} m_j = m_0$

Then, $u(z_0) = v(z_\infty)$.

Proof. We will reduce this situation to the particular case of the previous lemma.

First of all, we will show that we can assume without loss of generality that $z_0 = 0$ and $z_\infty = \infty$. For this, choose isometries $\phi_0, \phi_\infty \in G$ such that $\phi_0(z_0) = 0$ and $\phi_\infty(z_\infty) = \infty$. Now, replacing u_n, v_n, ψ_n, u and v with $\tilde{u}_n = u_n \circ \phi_0^{-1}$, $\tilde{v}_n = v_n \circ \phi_\infty^{-1}$, $\tilde{\psi}_n = \phi_0 \circ \psi_n \circ \phi_\infty^{-1}$, $\tilde{u} = u \circ \phi_0^{-1}$ and $\tilde{v} = v \circ \phi_\infty^{-1}$ respectively, we have that they still satisfy the conditions of the statement (we use that elements of G preserve the energy of the curve and that the group of isometries of S^2 is compact, so composing with ϕ_0, ϕ_∞ does not spoil convergence in condition (i)) and they satisfy $\tilde{u}(0) = u(\phi_0^{-1}(0)) = u(z_0) = v(z_\infty) = v(\phi_\infty^{-1}(\infty)) = \tilde{v}(\infty)$. Moreover, it is clear that if the new maps satisfy the conclusion of the proposition ($\tilde{u}(0) = \tilde{v}(\infty)$) then also the old maps satisfy it.

Now we show that again without loss of generality we can assume that the maps ψ_n are of the form $\psi_n(z) = \delta_n z$, where $\delta_i \to 0$. Observe that by hypothesis we have that $\psi_n(0) \to 0$ and $\psi_n^{-1}(\infty) \to \infty$. The last assertion follows because otherwise there exists a subsequence ψ_j and a sequence of points $z_j \in S^2$ with $z_j \in K \subset S^2 - \{\infty\}$ (K compact) for all j such that $\psi_j(z_j) = \infty$, contradicting the fact that ψ_n converges to 0 uniformly on compact sets of $S^2 - \{\infty\}$. Then, we can find sequences $\rho_n^0, \rho_n^\infty \in G$ that converge uniformly to the identity such that $\rho_n^0(\psi_n(0)) = 0$, $\rho_n^0(\infty) = \infty$, $\rho_n^\infty(\psi_n^{-1}(\infty)) = \infty$, $\rho_n^\infty(0) = 0$, and for n large enough, $\rho_n^\infty(1) = 1$ and $\rho_n^0(\psi_n(1)) =: \delta_n \in \mathbb{R}_+$. Indeed, recall that an element of G is completely determined by its image in three distinct points of S^2 and that G acts transitively in the set of all three distinct points of S^2, so ρ_n^0 and ρ_n^∞ are completely determined by this conditions, and they converge to the identity, since in the limit they fix three points of S^2. Now replace u_n, v_n and ψ_n by $\tilde{u}_n := u_n \circ (\rho_i^0)^{-1}$, $\tilde{v}_n := v_n \circ (\rho_n^\infty)^{-1}$ and $\tilde{\psi}_n := \rho_n^0 \circ \psi_n \circ (\rho_n^\infty)^{-1}$. Then, we have $\tilde{\psi}_n(1) = \rho_n^0(\psi_n(1)) = \delta_n$, $\tilde{\psi}_n(0) = \rho_n^0(\psi_n(0)) = 0 = \delta_n 0$ and $\tilde{\psi}_n(\infty) = \rho_n^0(\infty) = \infty = \delta_n \infty$, so $\tilde{\psi}_n(z) = \delta_n z$ for all $z \in S^2$ (since both are elements of G and coincide in three points). Moreover, if the conclusion of the theorem holds for the new maps, then it holds for the old maps also.

We want to prove now that $E(\rho) := \lim_n E(u_n; A(\delta_n/\rho, \rho))$ converges to 0 as $\rho \to 0$. Indeed, by using the conformal invariance of energy, we obtain

$$\begin{aligned} E(\rho) &= \lim_n E(u_n; \overline{D}_\rho) - \lim_n E(v_n; \overline{D}_{1/\rho}) \\ &= m_0 + E(u; \overline{D}_\rho) - E(v; \overline{D}_{1/\rho}) - \sum_{j=1}^{l} m_j \\ &= E(u; \overline{D}_\rho) + E(v; \mathbb{C} - \overline{D}_{1/\rho}) \end{aligned}$$

where in the second equality we have used that $\lim_n E(u_n; \overline{D}_\rho) = E(u; \overline{D}_\rho) + m_0$ (by definition of m_0) and $\lim_n E(v_n; \overline{D}_\rho) = E(v; \overline{D}_\rho) + \sum_{j=1}^{l} m_j$ (by definition of m_j), and in the third equality we have condition (iii) from the statement.

To finish, note that we have just proved that we are in the conditions of the previous lemma. So its application gives us $u(0) = v(\infty)$, as wanted. □

5.3 Gromov compactness

Now that we have all the hard work done, we will put everything together and interpret the obtained results in terms of the notion of Gromov convergence.

Theorem 5.3.1 (Gromov compactness)**.** *Let J_n be ω-compatible almost complex structures converging, in the C^∞ sense, to an ω-compatible almost complex structure J. Let $u_n : S^2 \longrightarrow X$ be a sequence of J_n-holomorphic curves such that $\sup_n E(u_i) < \infty$. Then, u_n has a Gromov convergent subsequence.*

Proof. Without loss of generality (by passing to a subsequence) we may assume that $E(u_n)$ converges. We put

$$E = \lim_n E(u_n)$$

Before we actually start the proof, we will introduce some useful notation. We will describe a tree T with N vertices as a vector $\mathbf{j} = (j_2, ..., j_N)$ with $j_k < k$ for all k. Let us explain how to do it. We can enumerate the vertices of the tree $v_1, ..., v_N$ in such a way that $T_i := \{v_1, ..v_i\}$ is a tree for any i. Indeed, start with any vertex v_1 an put $v_2, ..., v_{i_1}$ a maximal path starting at v_1. If there are no more vertices in T we are done. Otherwise, choose a vertex $v_k \in \{v_1, v_2, ..., v_{i_1}$ that is related to a vertex not yet considered, and choose again a maximal path starting with v_k. We continue in this way until we have all the vertices of T (the process stops because T is finite). It is clear that with the vertices chosen in this way, T_i is a tree for all i. Observe that by definition of the ordering, for each i there is some $j_i < i$ such that $v_i E v_{j_i}$. Therefore, the vector $\mathbf{j} = (j_2, ..., j_N)$ with $j_k < k$ for all k determines completely the tree T, since it encodes the relation E.

We will construct by induction

(a) a $(2N-2)$-tuple $\mathbf{u} = (u^1, u^2, ..., u^N; j_2, ..., j_N; z_2, ..., z_N)$, where $u_n : S^2 \longrightarrow X$ are pseudo-holomorphic spheres, $j_k < k$ for $k \geq 2$ are positive integers, and $z_k \in \mathbb{C}$ with $|z_k| \leq 1$.

(b) finite subsets $F_k \subset \overline{D}_1$ for $k = 1, ..., N$.

(c) sequences of Moebius transformations $\{\phi_n^k\}_{n\in\mathbb{N}}$ for $k = 1, ..., N$

such that some subsequence satisfies the following conditions

(i) $u_n \circ \phi_n^1$ converges to u^1 in the C^∞ sense on $S^2 - F_1$. For $k = 2, ..., N$, $u_n \circ \phi_n^k$ converges to u^k in the C^∞ sense on $\mathbb{C} - F_k$. Moreover, $F_1 \subset \{0\}$ and $Z_N = \emptyset$.

(ii) If u^k is constant, then $\#F_k \geq 2$.

(iii) The limit

$$m_k(z) := \lim_{\epsilon\to 0} \lim_n E(u_n \circ \phi_n^k; \overline{D}_\epsilon(z))$$

exists and satisfies $m_k(z) > 0$ for all $z \in F_k$. If $F_1 = \{0\}$ then $E = E(u^1) + m_1(0)$. If $k \geq 2$ then $z_k \in F_{j_k}$ and

$$m_{j_k}(z_k) = E(u^k) + \sum_{z\in Z_k} m_k(z)$$

(iv) If $j_k = j_{k'}$ then $z_k \neq z_{k'}$.

(v) For $k = 2, ..., N$, $u^{j_k}(z_k) = u^k(\infty)$

(vi) For $k = 2, ..., N$, $(\phi_n^{j_k})^{-1} \circ \phi_n^k$ converges to z_k in the C^∞ sense on $\mathbb{C} = S^2 - \{\infty\}$.

(vii) For $k = 2, ..., N$, $F_k = \{z_i : k < i \leq N, j_i = k\}$.

Let us explain briefly what these conditions mean. (i) and (iii) imply that the sequence $u_n \circ \phi_n^k$ bubbles exactly at the points of F_k. Therefore, the condition $F_1 \subset \{0\}$ of (i) means that $u_n \circ \phi_n^1$ has at most one bubble point (at 0), that is, the bubble u^1 is an endpoint of the tree T. (v) means that the bubble u^j is connected to the bubble u^{j_k} at the point z_k. (iv) means that there are no two bubbles bubbling at the same point. (iii) implies also that there is no loss of energy, that is, that the energy of u^{j_k} at the point z_k equals the energy of the bubble u^k together with all the bubbles accessible from u^k. (vii) will be satisfied only at the end of the induction, since it says that we have captured all the bubble points at each bubble. Finally, observe that (ii) is the stability condition asserting that if u^k is constant, then it is connected to at least three bubbles. Indeed, one point is ∞ by construction and then we have the additional points at F_k.

We think of the u_n as maps $u_n : \mathbb{C} \longrightarrow X$. We denote by $|du_n(z)|$ the norm of the differential where we use the usual norm in $\mathbb{C}$. We consider in S^2 the Fubini-Study metric (identifying S^2 with $\mathbb{C}P^1$), which in the open set $U_1 \cong \mathbb{C}$ is given by $|du_n(z)|_{FS} = |du_n(z)|(1 + |z|^2)$. Since moreover $|du_n(\infty)|_{FS}$ is finite, we must have $|du_n(z)| \to 0$ when $|z| \to \infty$. Therefore, $\sup_{z\in\mathbb{C}} |du_n(z)|$ assumes its maximum at some point of $\mathbb{C}$. Therefore, there is a sequence $z_n \in \mathbb{C}$ such that

$$|du_n(z_n)| = \sup_{z\in\mathbb{C}} |du_n(z)| =: c_n$$

Put $v_n(z) = u_n(z_n + z/c_n)$ and observe that $\sup_{z\in\mathbb{C}} |dv_n(z)| = 1 = |dv_n(0)|$. Therefore, some subsequence converges in the C^∞ sense on $\mathbb{C} = S^2 - \{\infty\}$ to a nonconstant pseudoholomorphic curve v. Passing to a further subsequence we may also assume that $E(v_n; \mathbb{C} - \overline{D}_1)$ converges. Then, the functions $u^1(z) = v(1/z)$ and $\phi_n^1(z) = z_n + \frac{1}{c_n z}$ satisfy $(i) - (iii)$ with $F_1 = \emptyset$ or $F_1 = \{0\}$. Conditions $(iv) - (vi)$ are void for $N = 1$. If $F_1 = \emptyset$ then $u_n \circ \phi_n^1$ converges to u^1 in the C^∞ sense on all of S^2 and the theorem is proved. So we will assume $F_1 = \{0\}$.

Let $l \geq 1$ and suppose by induction that we have constructed u^k, j_k, z_k, F_k and $\{\phi_n^k\}_{n\in\mathbb{N}}$ for $k \leq l$ in such a way that they satisfy $(i) - (vi)$ but not (vii) with l instead of N. Then, there is a $j \leq l$ such that $F_j \neq \{z_k : j < k \leq l, j_k = j\}$. We put

$$F_{j;l} := F_j - \{z_k | j < k \leq l, j_k = j\}$$

Choose any element $z_{l+1} \in F_{j;l}$ and apply proposition 5.2.1 to the sequence $u_n \circ \phi_n^j$ and the point $z_0 = z_{l+1}$. Calling the obtained subsequence again by $u_n \circ \phi_n^j$, we get that there exist maps $\psi_n \in G$ satisfying the conclusion of 5.2.1. Hence, $u_n \circ \phi_n^j \circ \psi_n$ converges to a pseudoholomorphic sphere $u^{l+1} : S^2 \longrightarrow X$ in the C^∞ sense on $S^2 - F$, where $F \subset \overline{D}_1$ is a finite set. We also have that ψ_n converges to z_{l+1} in the C^∞ sense on compact subsets of $\mathbb{C} = S^2 - \{\infty\}$. Moreover by proposition 5.2.3 we have that $u^{l+1}(\infty) = u^j(z_{l+1})$. This shows that the conditions $(i) - (vi)$ with $N = l + 1$ are satisfied for $k \leq l + 1$ with $j_{l+1} := j$, $F_{l+1} = F$ and $\phi_n^{l+1} = \phi_n^j \circ \psi_n$. This finishes with the induction.

Now we have to see that the induction terminates after finitely many steps. Consider at the l-th step of the induction the tree $T_l = (j_2, ..., j_l)$ and the numbers

$$m(j;l) = E(u^j) + \sum_{z \in F_{j;l}} m_j(z)$$

for $j = 1, ..., l$ where $F_{j;l}$ is as before. This quantity is the energy of u^j plus the energy of its bubble points that have not yet been considered. We have $m(j;l) = 0$ only when u^j is constant and all its bubble points have been considered. By (ii) this implies that j is a stable component of T_l (recall that T_l is the tree formed by the vertices $v_1, ..., v_l$). Note that by proposition 5.1.2, if $m(j;l) \neq 0$, then we have $m(j;l) \geq \hbar$, where $\hbar > 0$ is taken as the infimum of the numbers $\hbar(J_n) > 0$, where we use that $\hbar(J_n)$ converges to $\hbar(J)$ and hence its infimum is greater than 0 after passing to a subsequence of the u_n. We claim that

$$\sum_{j=1}^{l} m(j;l) = E \tag{5.6}$$

If $l = 1$, we have $E = E(u^1) + m_1(0) = m(1;1)$. Suppose that we have proved the identity for some integer $l \geq 1$. Note that $m(j;l+1) = m(j;l)$ whenever $j \neq j_{l+1}$ and $j \leq l$. But

$$m(j_{l+1}; l+1) = m(j_{l+1}; l) - m_{j_{l+1}}(z_{l+1})$$

and by (iii) with $k = l+1$,

$$m(l+1; l+1) = m_{j_{l+1}}(z_{l+1})$$

Hence the identity holds for $l+1$. Thus we have proved the identity for all $l \geq 1$. This means that the total energy at each step is E.

Let us see that the number of edges of T_l is bounded above by $2E/\hbar$. Since any bubble that is not constant must have energy at least $\hbar$, there are at most $E/\hbar$ nonconstant bubbles. Fix a vertex v_i of the tree (that is, some bubble). Observe that, since by the stability condition each constant bubble is connected to at least three bubbles, and for each edge of a constant bubble there are (connected to the constant bubble through some path containing that edge) at least two non-constant bubbles. Therefore, there can be at most $E/(2\hbar)$ non-constant bubbles, which means that we have at most $3E/(2\hbar)$ bubbles. Since at a tree with N vertices we have $N-1$ edges, we obtain that we have at most $3E/(2\hbar) - 1 < 2E/\hbar$ edges. Since at each step of the induction we add an edge to the tree, this shows that the induction terminates after a finite number of steps.

So we have proved the existence of u^k, j_k, z_k and $\{\phi_n^k\}_{n\in\mathbb{N}}$ for $k = 1, ..., N$ such that a suitable subsequence satisfies $(i) - (vii)$. We claim now that the subsequence u_n Gromov converges to the stable map $\mathbf{u}$. First of all we see that $\mathbf{u}$ is stable. In order to see this, observe that the vertices k and j with $k > j$ are connected by an edge if and only if $j = j_k$. Moreover, using the notation of the definition of stable maps, $z_{j_k k} = z_k$ and $z_{k j_k} = \infty$. So the singular set of the vertex $j \in \{2, ..., N\}$ is $F_j \cup \{\infty\}$ with F_j as in (vii), while the singular set of the vertex 1 is F_1. But if u^k is constant, necessarily $k \geq 2$ and $\#F_k \geq 2$. So $\mathbf{u}$ is stable.

To finish the proof we must verify the conditions in the definition of Gromov convergence. The (Map) and (Rescaling) conditions follow from (i) and (vi). It remains to verify the (Energy) condition. Denote by T_k the set of vertices k' which can be reached from j_k by a chain of edges, starting with the one from j_k to k. These sets can be defined recursively for $k = N, N-1, ..., 1$ by

$$T_k := \{k\} \cup \bigcup_{j_i=n} T_i$$

In particular $T_k = \{k\}$ whenever there is no $i > k$ such that $j_i = k$. We can prove by induction over the number of elements in T_k that

$$m_{j_k}(z_k) = \sum_{k' \in T_k} E(u^{k'}) \tag{5.7}$$

This follows from (iii) if $Z_k = \emptyset$ and so $T_k = \{k\}$. Assume that we have proved the formula for $\#T_k \leq m$ and suppose $\#T_k = m+1$. Put $\{k_1, ..., k_s\} := \{i > n : j_i = n\}$.
Then, the set $\{k_1, ..., k_s\} \cup \{j_k\}$ is the set of vertices adjacent to k and $\#T_{k_r} \leq m$ for $r = 1, ..., s$. Hence, by induction hypothesis,

$$m_k(z_k) = E(u^k) + \sum_{r=1}^{s} m_k(z_{k_r})$$

for $r = 1, ..., s$.
It follows from (iii) that

$$m_{j_k}(z_k) = E(u^k) + \sum_{r=1}^{s} m_k(z_{k_r})$$

Thus, combining the previous formulas and using the fact that $T_k = \{k\} \cup \bigcup_r T_{k_r}$ we obtain our claim 5.7. This is just (Energy) for $\alpha = j_k$ and $\beta = k$. With $\alpha = k$ and $\beta = j_k$ we have $T_{\alpha\beta} = \{1, ..., N\} - T_k$ and $z_{\alpha\beta} = \infty$ so that (Energy) has the form

$$m_k(\infty) := \lim_{R\to\infty} \lim_n E(u_n \circ \phi_n^k; \mathbb{C} - \overline{D}_R) = \sum_{k' \notin T_k} E(u^{k'}) \tag{5.8}$$

We can see this as follows

$$E - m_k(\infty) = \lim_{R\to\infty} \lim_n E(u_n \circ \phi_n^k; \overline{D}_r) = E(u^k) + \sum_{z \in F_k} m_k(z) = m_{j_k}(z_k) = \sum_{k' \in T_k} E(u^{k'})$$

where we have used the definitions of E and $m_k(\infty)$ in the first equality, the definition of $m_k(z)$ in the second, (iii) in the third and 5.7 in the fourth. Moreover, by 5.6 with $l = N$ we have $E = \sum_{k=1}^{N} E(u^k)$, which together with the last identity implies 5.8.

This finishes the proof of the theorem. □

To finish the chapter, we prove the Gromov compactness theorem for stable maps, which is an easy corollary of the version we have just proved.

Theorem 5.3.2 (Gromov compactness for stable maps)**.** *Let (u_n, z_n) be a sequence of J_n-stable maps such that $\sup_n E(u_n) < \infty$. Then, (u_n, z_n) has a convergent subsequence.*

Proof. Put $c := \sup_n E(u_n, z_n)$. Observe that, since the number of edges of the trees are bounded above by $\frac{2E}{\hbar}$ (as noted in the proof of the previous theorem), there appear only finitely many trees in the stable maps of the sequence, and passing to a subsequence we may assume that all trees are the same. We call it T.

By the previous theorem, for each $\alpha \in T$ the sequence u_n^α Gromov converges to a stable map (u^α, z^α). Now just connect the limit stable maps for different edges by adding an additional edge for each pair $\alpha E \beta$. $\square$

Chapter 6

The non-squeezing theorem

In this section we will give a complete proof (modulo the analytical details about moduli spaces of pseudoholomorphic curves and a result about minimal surfaces) of the celebrated Gromov non-squeezing theorem.

The proof we present in this chapter is essentially the same proof of Gromov, given in [Gro]. The exposition of this chapter closely follows that of [Hum].

6.1 The theorem and its consequences

We will denote by B_r^{2m} and $\overline{B}_r^{2m}$ the open and closed balls of radius r centered at the origin in $\mathbb{R}^{2m}$. We define also the symplectic cylinders

$$Z_r^{2m} := \{(x_1, y_1, ..., x_n, y_n) : x_1^2 + y_1^2 < r^2\} = B_r^2 \times \mathbb{R}^{2m-2} \subset \mathbb{R}^{2m}$$

With these notations we can state the theorem.

Theorem 6.1.1 (Non-squeezing theorem for cylinders)**.** *Suppose that there exists a symplectic embedding $i : \overline{B}_r^{2m} \longrightarrow Z_R^{2m}$. Then, $r < R$.*

We will prove a more general result from which this follows.

Theorem 6.1.2 (Non-squeezing theorem)**.** *Let M be a closed symplectic manifold of dimension $2m-2$ and assume that $\pi_2(M) = 0$ (where $\pi_2(M)$ represents the second homotopy group of M). Suppose also that there exists a symplectic embedding $i : \overline{B}_r^{2m} \longrightarrow B_R^2 \times M$. Then, $r < R$.*

We will see that the non-squeezing theorem for cylinders follows from this version. Since $i(\overline{B}_r^{2m}) \subset \mathbb{R}^{2m}$ is compact, we can assume that it lies inside $B_R^2 \times C$ for some cube $[a_1, b_1] \times ... \times [a_{2m-2}, b_{2m-2}]$. Consider the action of the group $\Gamma = \{(b_1 - a_1)n_1 + ... + (b_{2m-2} - a_{2m-2})n_{2m-2}\}$ on $\mathbb{R}^{2m-2}$ by translations. Then, $\mathbb{R}^{2m-2}/\Gamma \simeq T^{2m-2}$, the $2m-2$-dimensional torus. Let us see that this torus inherits a symplectic structure from the symplectic structure in $\mathbb{R}^{2m-2}$. Indeed, put $\pi : \mathbb{R}^{2m-2} \longrightarrow T^{2m-2}$ and define $\omega^T = \pi^*\omega$. Since π is a local diffeomorphism and ω is a symplectic form, it follows that ω^T is also a symplectic form. Therefore, $\pi \circ i : \overline{B}_r^{2m} \longrightarrow \overline{B}_R^2 \times T^{2m-2}$ is also a symplectic embedding (since by choice of the group Γ, π

is an embedding on $i(\overline{B}_r^{2m})$. Moreover, $\pi_2(T^{2m-2}) \simeq \pi_2(S^1) \times ... \times \pi_2(S^1) = 0$ (since as it is well-known, $\pi_n(S^1) = 0$ for $n > 1$). So, by the second version of the non-squeezing theorem, we have that $r < R$, as wanted.

To realize the importance of the non-squeezing theorem, observe that we have the following

Theorem 6.1.3. *Let M be a compact and oriented manifold. Suppose σ and σ' are two volume forms on M of the same total volume, that is,*

$$\int_M \sigma = \int_M \sigma'$$

Then, there exists a volume and orientation preserving diffeomorphism $\psi : N \longrightarrow N$, that is $\psi^*\sigma' = \sigma$.

The proof of this theorem is similar to Darboux's theorem 1.2.7. Details can be found in [H-Z].

Observe that, since a symplectic form determines a volume form in a symplectic manifold, any symplectomorphism between compact symplectic manifolds must be volume preserving. Taking this into account, we see that the non-squeezing theorem gives us a fundamental difference between symplectomorphisms and volume preserving diffeomorphisms, because any ball can be embedded in a volume preserving way into a cylinder, regardless of their radii. This gives a hint that there are non-trivial symplectic invariants, that are in some sense 2-dimensional.

We now define these non-trivial symplectic invariants, the symplectic capacities.

Definition 6.1.4. *A symplectic capacity c is a function that assigns to each symplectic manifold (X, ω) a number in $[0, \infty]$ satisfying*

1. *If there is a symplectic embedding $i : (X, \omega) \longrightarrow (X', \omega')$ and $\dim X = \dim X'$ then $c(X, \omega) \leq c(X', \omega')$.*
2. *$c(X, \lambda\omega) = c(X, \omega)$*
3. *$c(B^{2n}(1), \omega_0) > 0$ and $c(Z^{2n}(1), \omega_0) < \infty$.*

Observe that the first axiom implies that two symplectomorphic manifolds have the same symplectic capacity, hence it is a real symplectic invariant. Observe also that the volume cannot be a symplectic capacity, since the third axiom implies that $(Z^{2n}(1), \omega_0)$ has finite capacity, while it has infinite volume. Symplectic capacities must be thought of as measuring a two-dimensional quantities.

It is not trivial to show that there exist symplectic capacities. However, with the non-squeezing theorem, we can prove that they exist.

Definition 6.1.5. *The Gromov width of a symplectic manifold (X, ω) is defined by*

$$w_G(X, \omega) = w_G(X) = \sup\{\pi r^2 : B^{2n}(r) \text{ embeds symplectically in } X\}$$

Proposition 6.1.6. *The Gromov width is a symplectic capacity.*

Proof. Indeed, the Gromov width of a symplectic manifold is a symplectic capacity with the desired additional condition. That it satisfies the first and second axioms is immediate from the definition. The fact that it satisfies the third axiom follows from the non-squeezing theorem. □

6.2 Proof of the theorem

In this section we will give the proof of the non-squeezing theorem.
Let (M, ω) be a closed symplectic manifold with $\pi_2(M) = 0$. Consider the symplectic product

$$(V, \Omega) := (S^2 \times M, \sigma \oplus \omega)$$

where σ is any symplectic structure on S^2.

We denote by $A \in H_2(V)$ the homology class represented by the inclusion $i_{S^2} : S^2 \longrightarrow V$ defined by $i_{S^2}(z) = (z, q_0)$.

The main ingredient is the following result telling us that there exist a lot of pseudoholomorphic curves.

Theorem 6.2.1. *For each Ω-compatible almost complex structure J_V on V and any given point $p_0 \in V$ there exists an J_V-holomorphic curve $u : S^2 \longrightarrow V$ in the homology class A containing p_0 in its image.*

Recall that $P : \mathcal{M}(A, \mathcal{J}) \longrightarrow \mathcal{J}$ is the canonical projection. Note that if J_M is an almost complex structure on M, then, putting $\sigma_q : S^2 \longrightarrow S^2 \times M$ defined by $\sigma_q(z) = (\sigma(z), q)$ for $\sigma \in G$, we have that $P^{-1}(j \oplus J_M) = \{\sigma_q : \sigma \in G, q \in M\}$. In order to prove this, observe that if u is such that $P(u) = j \oplus J_M$ (so that u is a $j \oplus J_M$-holomorphic map) we can consider $\pi \circ u : S^2 \longrightarrow M$ where $\pi : S^2 \times M \longrightarrow M$ is the canonical projection, which is an J_M-holomorphic curve. Since $\pi_2(M) = 0$, we have that $\pi \circ u$ is homotopic to a constant map. Since pseudoholomorphic maps minimize the energy in its homology class, it follows that $\pi \circ u$ must be constant. On the other hand, $\pi' \circ u : S^2 \longrightarrow S^2$, where $\pi' : S^2 \times M \longrightarrow S^2$ is the canonical projection is a j-holomorphic map, it follows that $\pi' \circ u =: \sigma \in G$, as wanted.

We will use the following fact without proof.

Lemma 6.2.2. *Suppose J_M is integrable on an open set $U_0 \subset M$. Then, for each $\sigma \in G$ and each $q \in U_0$, the $j \oplus J_M$-holomorphic curve σ_q is a regular point of P.*

The next theorem, which is of crucial importance in our proof, is a corollary of the Gromov compactness theorem. In particular, it tells us that the spaces $\mathcal{M}(A, J) \times S^2/G$ are already compact, so there is no need to consider stable curves.

Proposition 6.2.3. *Suppose that $\{[u_n, z_n]\}_{n\geq 1}$ is a sequence in $\mathcal{M}(A, \mathcal{J}) \times S^2/G$ and that some subsequence $\{P(u_n)\}_{n\geq 1}$ converges in $\mathcal{J}$. Then, $\{[u_n, z_n]\}_{n\geq 1}$ has a convergent subsequence.*

Proof. Put $J = \lim_n P(u_n)$. Under the assumption of the theorem we have then that u_n is a sequence of J_n-holomorphic maps, where J_n converge to J.

Since each u_n represents A and J_n is Ω-compatible, we have that the energy of u_n is given by $< \Omega, A >$, so the sequence has bounded energy. By the compactness theorem, we have that u_n Gromov converges to a stable curve $\mathbf{u}$ of energy $E(\mathbf{u}) =< \Omega, A >$. Let us see that the stable curve $\mathbf{u}$ is actually a pseudoholomorphic sphere. Indeed, suppose it has at least two non-constant components, u_1 and u_2. Then we have

$$0 << \Omega, [u_1] >= E(u_1) < E(u) =< [\Omega], A >$$

Therefore there is a homology class with energy positive but strictly smaller than that of A. This leads to a contradiction as follows. Since $\pi_2(M) = 0$, we have that $\pi_2(V) = \pi_2(S^2 \times M) \simeq \pi_2(S^2) \simeq \mathbb{Z}$, where the last isomorphism is via the degree of the map. Suppose now that $g : S^2 \longrightarrow V$ is a pseudoholomorphic curve and that $\pi \circ g : S^2 \longrightarrow S^2$ is a map of degree n. Then, $g_*[S^2] = (deg(g))A$ (since A is a generator of $H_2(V)$), which implies that $E(g) =< [\Omega], g_*[S^2] >= (deg(g)) < [\Omega], A >$ cannot lie strictly between 0 and $< [\Omega], A >$.

Hence, $\mathbf{u}$ is actually a J-holomorphic sphere. The assertion that u_n Gromov converges to u means that there exist Moebius transformations $\psi_n \in G$ such that $u_n \circ \psi_n$ converge to u in the C^∞ sense. Since we are in $\mathcal{M}(A, \mathcal{J}) \times S^2/G$, we have that $[u_n, z_n] = [u_n \circ \psi_n, \psi_n^{-1}(z_n)]$. By passing to a subsequence if necessary, we may assume that $\psi_n^{-1}(z_n)$ converges to some $z \in S^2$ (by compacity of S^2). Therefore, we have that $[u_n, z_n]$ converge to $[u, z]$, as wanted. □

Using this proposition, we will give a proof of theorem 6.2.1. Before we go into the details, let us give a brief outline of the proof. First of all, note that if the almost complex structure J_V is of the form $j \oplus J_M$, where j is the usual almost complex structure on S^2 and J_M is a regular almost complex structure in M, then the theorem is trivial. This is because, in this case, putting $p_0 = (r_0, q_0)$ it suffices to consider a pseudoholomorphic map $u : S^2 \longrightarrow V = S^2 \times M$ of the form $u(z) = (\sigma(z), q_0)$ for some $\sigma \in G$. Now, for a general almost complex structure J_V we will show by using a cobordism argument with the almost complex structure $j \oplus J_M$ that there are J_b-holomorphic curves passing through p_b for J_b regular arbitrarily close to J_V and p_b arbitrarily close to p_0. Then, choosing a sequence of such maps and by applying proposition 6.2.3, we will show that we have a convergent subsequence and the limit pseudoholomorphic map has the desired properties.

Let us carry out this program in detail. First of all, observe that if J_M is an almost complex structure on M and j is the usual almost complex structure in S^2, we have

$$P^{-1}(j \oplus J_M) = \{\sigma_q : \sigma \in G, q \in M\}$$

where $\sigma_q : S^2 \longrightarrow V$ is defined by $\sigma_q(z) = (\sigma(z), q)$. It is clear that every σ_q is a $j \oplus J_M$-holomorphic map. Conversely, if u is a $j \oplus J_M$-holomorphic map, consider $\pi_M \circ u : S^2 \longrightarrow M$ where $\pi_M : S^2 \times M \longrightarrow M$ is the canonical projection. Then, $\pi_M \circ f$ is an J_M-holomorphic map in M. But since $\pi_2(M) = 0$, $\pi_M \circ f$ is homotopic to a constant map. It follows that $\pi_M \circ u$ has the zero energy, hence it is constant. On the other hand, $\pi_{S^2} \circ u : S^2 \longrightarrow S^2$ is a j-holomorphic map, hence a Moebius transfomation, as wanted.

Lemma 6.2.4. *Given neighbourhoods $\mathcal{U}$ of $j \oplus J_M$ in $\mathcal{J}$ and U of p_0 in V, there exist an almost complex structure $J_a \in \mathcal{U} \cap \mathcal{J}_{reg}$ and a point $p_a \in U$ such that p_a is a regular value of ev_{J_a} and $ev_{J_a}^{-1}(p_a)$ has exactly one point.*

Proof. Consider the following map

$$\Phi : \mathcal{M}(A, \mathcal{J}) \times S^2 \longrightarrow \mathcal{J} \times V$$

defined by $\Phi(f, z) = (P(f), f(z))$. Since Φ is obtained from P and the evaluation map, we see that it is a Fredholm map. From the previous description of $P^{-1}(j \oplus J_M)$ and lemma 6.2.2, we can see that $(j \oplus I_m, p_0)$ is a regular value of Φ. Therefore, there is an open connected

neighbourhood $\mathcal{V}$ of $(id_{q_0}, z_0) = (j \oplus J_M, p_0)$ in $\mathcal{J} \times V$ such that $\Phi|_{\mathcal{V}} : \mathcal{V} \longrightarrow \mathcal{W}$ is a submersion. Then, by a form of the inverse function theorems for Banach manifolds, there is a diffeomorphism $\xi : \mathcal{W} \times B \longrightarrow \mathcal{V}$. By shrinking the neighbourhoods, we may assume without loss of generality that $B = \Phi|_{\mathcal{V}}^{-1}(j \oplus J_M, p_0)$ is a connected open ball in some $\mathbb{R}^n$ such that

$$\Phi \circ \xi : \mathcal{W} \times B \longrightarrow \mathcal{W}$$

is the projection onto $\mathcal{W}$.
Observe that $\dim B = \dim G$, because there is only one $j \oplus J_M$-holomorphic curve passing through p_0 modulo the action of G. Choose $(w, b) \in \mathcal{W} \times B$. Then the set of points in $\{w\} \times B$ equivalent to (w, b) by the action of G form an open set in $\{w\} \times B$ because $\dim B = \dim G$, and moreover it is closed, so by connectedness, it is all of $\{w\} \times B$. Therefore, given $(J, p) \in \mathcal{W}$, we have that there is a unique J-holomorphic curve in $\mathcal{V}$ passing through p modulo G.

Now the Sard-Smale theorem tells us that there exists a sequence (J_n) in $\mathcal{J}_{reg}$ converging to $j \oplus J_M$ and elements $(u_n, z_n) \in \mathcal{M}(A, J_n) \times S^2$ converging to (id_{q_0}, z_0) such that $[u_n, z_n]$ are regular points of ev_{J_n}.

Suppose now that the lemma is false, that is, that there exist neighbourhoods of $\mathcal{U}$ and of p_0 such that for any almost complex structure J and point p in the corresponding neighbourhoods, p is not a regular value of ev_J. Then, from some n on, we have that there are elements $(u'_n, z'_n) \in \mathcal{M}(A, I_n) \times S^2$ with $ev_{J_n}([u'_n, z'_n]) = ev_{J_n}([u_n, z_n])$ and $[u'_n, z'_n]$ a singular point of ev_{J_n}. Using proposition 6.2.3 and the compactness of S^2, we may assume that (u'_n, z'_n) converge to some $(u,', z') \in \mathcal{M}(A, | \oplus \mathcal{J}_{\mathcal{M}}) \times S^2$. We have that $u'(z') = \lim_n u'_n(z'_n) = \lim_n u_n(z_n) = p_0$ and using that there is only one $j \oplus J_M$-holomorphic curve through p_0, we may assume (after a reparametrization) that $(u', z') = (id_{q_0}, z_0)$. Hence, $(u'_n, z'_n) \in \mathcal{V}$ for n large enough, and then $(u'_n, z'_n) = (u_n, z_n)$, which is a contradiction since $[u'_n, z'_n]$ is singular for ev_{J_n} while $[u_n, z_n]$ is regular. This finishes the proof of the first assertion.

The second assertion is now trivial, since we have already seen that there is a unique J_a-holomorphic curve passing through p_a, modulo reparametrization by G. □

Recall that the set of regular almost complex structures is open in $\mathcal{J}$, so we can pick a regular almost complex structure J_b in any neighbourhood of J_V. By the Sard-Smale theorem, the set of regular values for ev_{J_b} is also dense, hence we can pick such a regular value p_b in any neighbourhood of p_0. Now we have

Lemma 6.2.5. *Let J_b and p_b as before. Then there exists an J_b-holomorphic curve representing A and passing through p_b.*

Proof. We choose p_a and J_a in a neighbourhood of $(p_0, j \oplus J_M)$ satisfying the previous lemma. By the path-connectedness of $\mathcal{J}$ and the transversality theorem, there is a smooth path $\alpha : [0, 1] \longrightarrow \mathcal{J}$ joining J_a to J_b transverse to the P. Then, $\mathcal{M}(A, \alpha) := P^{-1}(\alpha([0, 1]))$ is a finite-dimensional smooth manifold with boundary, and its boundary is $\partial\mathcal{M}(A, \alpha) = \mathcal{M}(A, J_a) \cup \mathcal{M}(A, J_b)$, where the union is disjoint. Observe that G acts in $\mathcal{M}(A, \alpha) \times S^2$ as always, by $g \cdot (u, z) = (u \circ g^{-1}, g(z))$, and this action gives rise to a manifold $\mathcal{M}(A, \alpha) \times S^2/G$. We can consider again an evaluation map

$$ev_\alpha : \mathcal{M}(A, \alpha) \times S^2/G \longrightarrow V \times [0, 1]$$

by $ev_\alpha([u,z]) = (u(z),t)$ is f is an $\alpha(t)$-holomorphic curve. This is well-defined since α is transversal to P. Then, the points $(p_a,0)$ and $(p_b,1)$ are regular values of ev_α, since p_a and p_b are regular values of J_a and J_b respectively. Since $V \times S^2$ is connected, we can choose a path $\beta : [0,1] \longrightarrow V \times [0,1]$ from p_a to p_b transverse to ev_α and transverse to the boundary $V \times \{0\} \cup V \times \{1\}$.

Put now $N := ev_\alpha^{-1}(\beta([0,1]))$. Then, N is a smooth submanifold of $\mathcal{M}(A,\alpha)$ and has boundary $\partial N = N \cap \partial\mathcal{M}(A,\alpha) = (N \cap (\mathcal{M}(A,J_a) \times S^2)/G) \cup (N \cap (\mathcal{M}(A,J_b) \times S^2)/G)$. Observe that by applying proposition 6.2.3, we obtain that any subsequence of $\mathcal{M}(A,\alpha)$ contains a convergent subsequence, so $\mathcal{M}(A,\alpha)$ is compact. Then N, being a closed submanifold of $\mathcal{M}(A,\alpha)$ is compact too. By the previous lemma, the component $N \cap (\mathcal{M}(A,I_a) \times S^2)/G) = ev_{J_a}^{-1}(p_a)$ of the boundary of N consists on exactly one point. Therefore, N must be one-dimensional and its boundary is a disjoint union of an even number of points. This implies that $(N \cap (\mathcal{M}(A,J_b) \times S^2)/G \neq \emptyset$. That is, there is an J_b-holomorphic curve passing through p_b, as wanted. □

Now we can easily finish the proof of theorem 6.2.1.

Proof of theorem 6.2.1. The previous lemma implies that there is a sequence (J_n,p_n) in $\mathcal{J} \times V$ converging to (J_V,p_0) and a sequence u_n of J_n-holomorphic curves representing the homology class A passing through p_n. Let z_n such that $u_n(z_n) = p_n$. Use proposition 6.2.3 to extract a subsequence of the (u_n,z_n) such that there exists $\psi_n \in G$ satisfying that $u_n \circ \psi_n$ converges to a J_V-holomorphic curve u in the C^∞ sense and $\psi^{-1}(z_n)$ converges to a point $z \in S^2$. Then, $u(z) = \lim_n (u_n \circ \psi_n)(\psi_n^{-1}(z_n)) = \lim_n p_n = p_0$. So u is the wanted pseudoholomorphic curve. □

Now that we know that there exist plenty of pseudoholomorphic curves, we are ready to prove theorem 6.1.2. We will use the following result from the theory of minimal surfaces, which we quote without proof. Recall that a surface in $\mathbb{R}^{2n}$ is proper if each closed ball in $\mathbb{R}^{2n}$ contains a compact portion of the surface with respect to its intrinsic topology.

Theorem 6.2.6 (Monotonicity theorem)**.** *Any proper minimal surface passing through the origin in $B^{2n}(R) \subset \mathbb{C}^n$ has its area greater or equal than πR^2.*

For a proof, see [M-P], Theorem 2.6.2.
Observe that for a minimal surface $S \subset \mathbb{R}^{2n}$, its area is given by

$$A(S) = \int_S \omega_0$$

where ω_0 is the restriction to S of the standard symplectic form of $\mathbb{R}^{2n}$.

Suppose we have a holomorphic map $u : S \longrightarrow \mathbb{C}^n$. Then, we have:

$$A(S) = \int_S \omega_0 = \int_S u^*\omega_0 = E(u)$$

This tells us that the area of a surface which is the domain of a holomorphic curve equals the energy of the curve. Moreover, since holomorphic curves (which are just J_0-holomorphic curves) are energy minimizing in their homology class, this means that S is an area minimizing surface, that is, a minimal surface. We will use this identity in the following proof.

Proof of theorem 6.1.2. Assume that there is a symplectic embedding $i : B^{2m}(r) \longrightarrow B^2(R)\times M$. Choose $\epsilon > 0$ and a symplectic structure σ on S^2 such that $A(S^2, \sigma) = \pi R^2 + \epsilon$. Note that this is possible simply by rescaling the symplectic form (which in the case of a surface is the same as an area form). Hence, we have $A(B^2(R)) = \pi R^2 < A(S^2, \sigma)$, so we can embed symplectically (which in this case is just the same as volume-preserving) $B^2(R)$ into S^2. In this way, we will view $B^2(R)$ as a subset of (S^2, σ). So composing the two embeddings, we have an embedding (still denoted by i) $i : B^{2m}(r) \longrightarrow (V, \Omega) = (S^2 \times M, \sigma \oplus \omega)$.

We will choose now an Ω-compatible almost complex structure J on V in such a way that $J_0|_{B^{2m}(r-\epsilon)} = i^*(J|_{i(B^{2m}(r-\epsilon)})$. Let us see how to do this. Pick, for each point $x \in i(\overline{B}^{2m}(r-\epsilon))$, a coordinate chart (U_x, ξ_x) and a partition of unity $\{\rho_x\}_x$ subordinated to the $\{U_x\}_x$. Define a metric g_x in each of the U_x with the condition that $i^*(g_x|_{U_x\cap i(B^{2m}(r)}) = g_0|_{\xi_x(U_x)\cap B^{2m}(r)}$ and put $g = \sum_x \rho_x g_x$. Then, we have obtained a metric g on V such that $i^*(g|_{B^{2m}(r-\epsilon))}) = g_0|_{B^{2m}(r-\epsilon)}$. Now apply the construction in the proof of proposition 1.3.3 to obtain an Ω-compatible almost complex structure J from g. Since starting from ω_0 and g_0 one obtains the usual almost complex structure J_0, we see that our J satisfy the desired conditions.

Observe now that $i(B^{2m}(r-\epsilon))$ is holomorphically isometric to the Euclidean ball $B^{2m}(r-\epsilon)$, where we are considering in $i(B^{2m}(r-\epsilon))$ the metric g_J induced by the almost complex structure J. Indeed, the map i is an isometry because if $u, v \in TB^{2m}(r-\epsilon)$ we have

$$\begin{aligned} i^*(g_J(u,v)) &= g_J(di(u), di(v)) = \Omega(di(u), Jdi(v)) \\ &= \Omega(di(u), di(J_0(v))) = \omega_0(u, J_0(v)) = g_0(u,v) \end{aligned}$$

where we have used that i is symplectic and that $Jdi = diJ_0$. The holomorphicity follows from the fact that $i^*(I) = J_0$, where J_0 is the usual almost complex structure in $\mathbb{R}^{2m}$.

Now we apply theorem 6.2.1 to obtain a J-holomorphic curve $u : S^2 \longrightarrow V$ representing the homology class A and passing through $i(0)$. Observe that, since $i(\overline{B}^{2m}(r-\epsilon))$ is a $2m$-dimensional submanifold with boundary of V, $S := u^{-1}(i(\overline{B}^{2m}(r-\epsilon)))$ is a proper 2-dimensional submanifold with boundary of S^2 (it cannot be the whole S^2, in other words, the image of u is not completely contained in $i(\overline{B}^{2m}(r))$, because if this were the case u would be constant, being a pseudoholomorphic curve homotopic to a constant map). Now consider $u|_S : S \longrightarrow \overline{B}^{2m}(r-\epsilon)$, where we have identified $i(\overline{B}^{2m}(r-\epsilon))$ with the euclidean ball $\overline{B}^{2m}(r-\epsilon)$ via the holomorphic isometry i^{-1}. Then, $u|_S$ is a J_0-holomorphic curve, and since J_0 is the usual almost complex structure in $\mathbb{R}^{2m}$, we have that $u|_S$ is in fact a holomorphic curve. Moreover, since holomorphic maps minimize energy in their homology class, and in the case of holomorphic maps, energy coincides with the area, we have that $u|_S$ is area minimizing among all the smooth maps $\phi : S \longrightarrow \overline{B}^{2m}(r-\epsilon)$ which satisfy $\phi|_{\partial S} = u|_{\partial S}$. This means that S is a minimal surface.

Therefore, we can apply the monotonicity lemma for minimal surfaces, to conclude that

$$\pi(r-\epsilon)^2 \leq A(S) = E(u|_S)$$

From here, we get the following chain of inequalities:

$$\pi(r-\epsilon)^2 \leq E(u|_S) < E(u) =< [\Omega], A >= \int_{S^2} incl^*\Omega = \pi R^2 + \epsilon.$$

Therefore, since this is valid for any ϵ, it follows that $r \leq R$, as claimed. $\square$

Bibliography

[Dem] Demailly, J-P.; *Complex analytic and differential geometry.* Online text. 2012.

[Gro] Gromov, M.; *Pseudoholomorphic curves in symplectic manifolds.* Invent. math. 82, 307-347, 1985.

[Hum] Hummel, C.; *Gromov's compactness theorem for pseudoholomorphic curves.* Springer Basel AG, 1997.

[H-Z] Hofer, H., Zehnder, E.; *Symplectic invariants and Hamiltonian Dynamics.* Birkhäuser Advanced Text, Birkhäuser, 1994.

[M-S1] McDuff, D.; Salamon, D. *J-Holomorphic curves and symplectic topology.* Colloquium Publications, Vol. 52. American Mathematical Society, 2004.

[M-S2] McDuff, D.; Salamon, D. *Introduction to symplectic topology.* Oxford University Press, 1998.

[M-P] Meeks, W., Pérez, J.; *A survey on classical minimal surface theory.* Online text.

[Sil] da Silva, C.; *Lectures on symplectic geometry.* Lecture notes in mathematics, book 1764. Springer, 2008.

[War] Warner, F. *Foundations of differentiable manifolds and Lie groups.* Graduate texts in mathematics, No. 94. Springer-Verlag, 1983.