Symplectic folding and non-isotopic polydisks

R. Hind

October 10, 2012

1 Introduction

1.1 Main results

Our main result demonstrates the existence of different Hamiltonian isotopy classes of symplectically embedded polydisks inside a 4-ball, and by the same argument also in the complex projective plane. Furthermore, we find exactly how large the ball can be before the embeddings become isotopic; the optimal isotopy is a version of symplectic folding.

Before stating the result precisely we fix some notation. We work in $\mathbb{R}^4$ with coordinates (x_1, y_1, x_2, y_2) and standard symplectic form $\omega = \sum_i dx_i \wedge dy_i$. We denote by $B(a)$ the open ball of capacity a, that is

$$B(a) = \{\pi \sum_i (x_i^2 + y_i^2) < a\} \subset \mathbb{R}^4$$

with the restricted symplectic form. The polydisk $P(a, b)$ is defined by

$$P(a, b) = \{\pi(x_1^2 + y_1^2) \leq a, \pi(x_2^2 + y_2^2) \leq b\} \subset \mathbb{R}^4.$$

For the results in this section, when describing a polydisk we will always assume that $a < b$. Sometimes $D(a)$ will be used to denote a closed disk of area a in $\mathbb{R}^2$. The inclusion map then gives a symplectic embedding $P(a, b) \to B(R)$ for any $R > a + b$.

Let $H_t : \mathbb{R}^4 \to \mathbb{R}$, $0 \leq t \leq 1$ be a smooth family of compactly supported functions. The forms dH_t are dual to vectorfields X_{H_t} under the symplectic form, that is, we define X_{H_t} by $X_{H_t} \rfloor \omega = dH_t$. Then the Hamiltonian diffeomorphism generated by $\{H_t\}$ is the time-1 flow of the time-dependent vectorfield X_{H_t}. Hamiltonian diffeomorphisms are symplectic. On the other hand, by a theorem of Gromov, [8], a compactly supported symplectomorphism of $B(a)$ is Hamiltonian.

Here is the first part of our main result.

Theorem 1.1. *Let $S < a + b < R < 2a + b$. Then there does not exist a Hamiltonian diffeomorphism ϕ with support contained in $B(R)$ such that $\phi(P(a, b)) \subset B(S)$.*

In cases when the ratio $\frac{b}{a}$ is sufficiently large there do exist Hamiltonian diffeomorphisms ψ of $\mathbb{R}^4$ such that $\psi(P(a, b)) \subset B(S)$ for $S < a+b$. In fact, as the ratio $\frac{b}{a} \to \infty$ we may take $S \to \sqrt{ab}$, that is, the image of the polydisk can occupy an arbitrarily large proportion of the volume of the ball. For this, see for example [17].

Hence, together with Gromov's theorem mentioned above we can produce examples of the following.

Corollary 1.2. *Suppose that $a + b < R < 2a + b$, so in particular $P(a, b) \subset B(R)$, and $\frac{b}{a}$ is very large. Then there exist symplectic embeddings $P(a, b) \to B(R)$ which do not extend to symplectomorphisms $B(R) \to B(R)$.*

On the other hand, when the ratio $\frac{b}{a}$ is close to 1 there may be no symplectic embeddings $P(a, b) \to B(S)$ for any $S < a + b$. For example, the second Ekeland-Hofer capacity of $P(1, 1)$ is 2 whereas the second Ekeland-Hofer capacity of $B(S)$ is S, see [5], and these capacities are monotonic under symplectic embeddings.

The easiest way to construct an embedding $P(a, b) \to B(S)$ for $S < a+b$ (when such an embedding exists) is via symplectic folding, and we shall see that Theorem 1.1 is sharp in the sense that this method applies whenever $R > 2a + b$.

Theorem 1.3. *Let $R > 2a + b$. Then for any $\epsilon > 0$ there exists a Hamiltonian diffeomorphism ϕ with compact support in $B(R)$ such that $\phi(P(a, b)) \subset B(2a + \frac{b}{2} + \epsilon)$.*

Let us consider symplectic embeddings $f : P(a, b) \to B(R)$ under the equivalence relation that $f_1 \sim f_2$ if there exists a Hamiltonian diffeomorphism ϕ with support in $B(R)$ such that $\phi \circ f_1 = f_2$. Hamiltonian diffeomorphisms with support in $B(R)$ can be generated by functions which all have support in $B(R)$. Thus we call the equivalence classes Hamiltonian isotopy classes. Then combining with Theorem 1.1 we can say the following, making Corollary 1.2 more precise.

Corollary 1.4. *Suppose that* $2a < b$ *and* $a+b < R < 2a+b$*. Then there exists at least two Hamiltonian isotopy classes of polydisks* $P(a, b)$ *inside* $B(R)$*.*

In section 2 we study holomorphic curves in manifolds with cylindrical ends in order to prove Theorem 1.1. The technique of symplectic folding is now fairly well known. It was introduced in [14] and studied extensively in [17], but in section 3 we review the construction and show that it implies Theorem 1.3.

1.2 Related work

1.2.1 Isotopies of polydisks

There are known to be non-isotopic polydisks embedded in larger polydisks. Observe that if $a, b < 1$ then the polydisks $P(a, b)$ and $P(b, a)$ both lie in $P(1, 1)$. We consider these as giving two embeddings $g_1, g_2 : P(a, b) \to P(1, 1)$. For any $R > a + b$, there exists a unitary transformation mapping $P(a, b)$ to $P(b, a)$ generated by a Hamiltonian with support contained in $B(R) \subset P(R, R)$. The following theorem is due to A. Floer, H. Hofer and K. Wysocki.

Theorem 1.5. *(Floer-Hofer-Wysocki [7] Theorem* 4*) When* $a, b < 1$ *and* $a + b > 1$ *there does not exist a Hamiltonian diffeomorphism* ϕ *with support contained in* $P(1, 1)$ *such that* $\phi \circ g_1 = g_2$*.*

In other words, if the ambient polydisk does not contain the support of a rotation moving $P(a, b)$ onto $P(b, a)$ then the polydisks are not isotopic under any Hamiltonian flow. We contrast this with our Theorems 1.1 and 1.3 saying that if the ambient ball is not large enough to support a Hamiltonian generating the symplectic folding construction, then the polydisks are not isotopic under any Hamiltonian flow.

The theorem was established as an application of symplectic homology, although as pointed out by K. Ono it can also be derived from Floer theoretic methods applied to the distinguished boundary $\partial D(a) \times \partial D(b) \subset \partial D(a, b)$ of the polydisk, which is a Lagrangian torus.

1.2.2 Isotopies of ellipsoids

Besides polydisks, another convenient class of domains used to investigate symplectic embedding problems are ellipsoids. The ellipsoid $E(a,b)$ is defined by

$$E(a,b) = \{\frac{\pi(x_1^2 + y_1^2)}{a} + \frac{\pi(x_2^2 + y_2^2)}{b} \leq 1\} \subset \mathbb{R}^4.$$

In particular $E(a,a) = \overline{B}(a)$. The ellipsoid $E(a,b)$ is embedded in the ball $B(R)$ by inclusion whenever $R > \max\{a,b\}$. In contrast to Theorem 1.1 however, according to a theorem of D. McDuff this is the only isotopy class inside the ball. Indeed, this remains the case if we study any embedding from an ellipsoid to an ellipsoid.

Theorem 1.6. *(McDuff [15] Corollary 1.6) The space of symplectic embeddings $E(a,b) \to E^\circ(a',b')$ is path connected whenever it is nonempty.*

2 Restriction on isotopies

In this section we establish our constraint on the Hamiltonian diffeomorphisms of the standard polydisk $P(a,b) \subset B(R)$ for some $a+b < R < 2a+b$. Our techniques apply to smooth domains so first we approximate $P(a,b)$ by a slightly smaller domain W with smooth boundary. We recall that a Hamiltonian diffeomorphism with support in $B(R)$ generates a symplectic isotopy of $P(a,b)$ inside $B(R)$ by choosing generating Hamiltonian functions with support in the ball. As an isotopy of $P(a,b)$ restricts to one of W constraints on isotopies of W imply constraints for the polydisk.

2.1 The smooth domain W and orbits on the boundary

Fix a small irrational $\epsilon > 0$ and a $0 < \delta << \epsilon$ and let $x_0 = \frac{a-\epsilon b}{1-\epsilon^2}$. Then let $f : [0,a] \to [0,b]$ be a smooth convex function with $f(x) = b - \epsilon x$ for $x < x_0 - \delta$ and $f(x) = \frac{1}{\epsilon}(a-x)$ for $x > x_0 + \delta$. Finally set $K(x,y) = y - f(x)$.

It will be convenient to use symplectic polar coordinates on $\mathbb{R}^4$, so we set $R_i = \pi(x_i^2 + y_i^2)$ and $\tan \frac{y_i}{x_i} = \theta_i$. Then we define our approximation W to $P(a,b)$ by

$$W = \{K(R_1, R_2) \leq 0\}.$$

This is a smooth domain with contact-type boundary and Reeb flow R generated by the Hamiltonian vector field of K, namely

$$R = 2\pi \left(K_{R_1} \frac{\partial}{\partial \theta_1} + K_{R_2} \frac{\partial}{\partial \theta_2} \right)$$

where K_{R_i} denotes the partial derivative with respect to R_i as usual.

The Reeb flow preserves the Lagrangian tori $\{(R_1, R_2) = \text{const.}\}$ in ∂W and using the coordinates (θ_1, θ_2) we can identify these tori with a fixed torus T^2 and the integer homology with $H_1(T^2, \mathbb{Z}) = \mathbb{Z}^2$. Then we get closed orbits in a homology class $(m, n) \in H_1(T^2)$ exactly when $\frac{K_{R_1}}{K_{R_2}} = \frac{m}{n}$. By the definition of K, in our case $K_{R_2} \equiv 1$ and K_{R_1} is a function only of R_1 and is increasing from ϵ to $\frac{1}{\epsilon}$ as R_1 moves through x_0. Therefore, given a rational $\frac{m}{n} \in [\epsilon, \frac{1}{\epsilon}]$ we will have a 1-parameter family of closed Reeb orbits in the corresponding homology class lying in the torus over (R_1, R_2) where $K_{R_1} = \frac{m}{n}$. As ϵ is irrational these are the only closed Reeb orbits in ∂W except for covers of $\{R_1 = 0\}$ and $\{R_2 = 0\}$.

Definition 2.1. Given $m, n \in \mathbb{Z}_{\geq 0}$ with $\frac{m}{n} \in (\epsilon, \frac{1}{\epsilon})$ let $\gamma_{m,n}$ denote a Reeb orbit in the corresponding homology class (m, n). Thus if m, n are coprime, $\gamma_{rm,rn}$ denotes an r-fold cover of a primitive Reeb orbit $\gamma_{m,n}$.

Let γ_1^r denote the r-fold cover of $\gamma_1 = \{R_2 = 0\}$ and γ_2^r denote the r-fold cover of $\gamma_2 = \{R_1 = 0\}$.

Now, if we fix a symplectic trivialization of $T\mathbb{R}^4|_\gamma$, the tangent bundle of $\mathbb{R}^4$ restricted to a closed orbit γ of R of period T, then the derivative of the Reeb flow (extended to act trivially normal to ∂W) gives a map $\psi : [0, T] \to \text{Symp}(4, \mathbb{R})$, where $\text{Symp}(4, \mathbb{R})$ is the group of 4×4 symplectic matrices. Associated to such a path is a Conley-Zehnder index $\mu(\gamma)$ defined in this case by Robbin and Salamon in [16].

Lemma 2.2. *With respect to the standard basis of* $\mathbb{R}^4$ *we have* $\mu(\gamma_i^r) = 2r + 1$ *and* $\mu(\gamma_{m,n}) = 2(m + n) + \frac{1}{2}$.

Proof. The computation for the orbits in the coordinate planes is contained for example in [9], Lemma 1.6 and we omit that. Similar computations for the other orbits can be found in [1], see also [13], but we review this anyway.

First of all, consider a different symplectic trivialization consisting of the Reeb vector field R itself, a normal vector n to ∂W, a vector v perpendicular to R and tangent to the Lagrangian torus, and a vector w in ∂W symplectically orthogonal to n and v. The flow preserves R and n (by definition) and so the index is determined by its action on the symplectic complement $< v, w >$. Restricted to this subspace the marix $\psi(t)$ is of the form

$$\psi(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \tag{1}$$

that is, v is preserved (as the flow is linear on the Lagrangian fibers) but the image of w has a component in the direction of v due to the change in direction of R as we vary the fibers. Thus the orbits $\gamma_{m,n}$ are said to be of hyperbolic type and the index with respect to this trivialization is $\frac{1}{2}$.

We wish to work with a different trivialization and so the matrix $\psi(t)$ above must be composed with a symplectic change of basis matrix $\zeta(t)$ taking the basis described above to the standard basis of $\mathbb{R}^4$. If we identify $\mathbb{R}^4$ with $\mathbb{C}^2$ then this change of basis is actually complex and so up to scale we can write $\zeta(t)$ in complex coordinates as

$$\zeta(t) = \begin{pmatrix} ie^{imt} & ie^{imt} \\ ie^{int} & -ie^{int} \end{pmatrix}, 0 \leq t \leq 2\pi. \tag{2}$$

The path $\zeta(t)$ has Maslov index $2(m+n)$ and so by the catenation formula for Conley-Zehnder indices, see [16] Theorem 2.3, with respect to our chosen trivialization we have $\mu(\gamma_{m,n}) = 2(m+n) + \frac{1}{2}$. □

2.2 Holomorphic curves in $\mathbb{C}P^2 \setminus W$

Let $\mathbb{C}P^2(R)$ denote the complex projective plane with its Fubini-Study symplectic form scaled such that lines have area R. There exists a natural symplectic embedding $B(R) \subset \mathbb{C}P^2(R)$ whose image is the complement of $\mathbb{C}P^1(\infty)$, the line at infinity. In order to study holomorphic curves, we can therefore think of W as embedded $W \subset B(R) \subset \mathbb{C}P^2(R)$.

2.2.1 The almost-complex structure on X and finite energy curves

A neighborhood of $\partial W \subset \mathbb{C}P^2(R)$ is symplectomorphic to $\partial W \times (-\epsilon, \epsilon)$ with symplectic form $d(e^t\lambda)$, where λ is a contact form on ∂W and t is the coordinate on $(-\epsilon, \epsilon)$. Indeed, let Z denote the radial vector field

$$Z = 2 \sum R_i \frac{\partial}{\partial R_i}$$

on $\mathbb{R}^4$. This is a Liouville vector field transverse to ∂W, that is, $\mathcal{L}_Z\omega = \omega$. Then $\sigma = Z\rfloor\omega$ is a primitive of ω, and we also have $\mathcal{L}_Z\sigma = \sigma$. Note that $\sigma(Z) = 0$. We identify a neighborhood of $\partial W = \partial W \times \{0\}$ with $\partial W \times (-\epsilon, \epsilon)$ such that Z is given by $\frac{\partial}{\partial t}$, where t is the coordinate on $(-\epsilon, \epsilon)$. Then in these coordinates we can write $\sigma = e^t\lambda$ where λ is the restriction of $\sigma = Z\rfloor\omega$ to ∂W. Let $\xi = \{\lambda = 0\}$ be our corresponding contact structure on ∂W.

Now, following [6] we will study holomorphic curves in

$$X = (\mathbb{C}P^2(R) \setminus W) \cup (\partial W \times (-\infty, 0])$$

where the gluing is defined using the above identification of a neighborhood of ∂W with $\partial W \times (-\epsilon, \epsilon)$. The symplectic form on $\mathbb{C}P^2(R) \setminus W$ extends smoothly to all of X after a small perturbation near ∂W by the formula $\omega = d(\phi(t)\lambda)$ where ϕ is an increasing function with $\phi(t) = e^t$ for t close to ϵ and $\phi(t) \to 1$ as $t \to -\infty$. We associate a tame almost-complex structure J which is translation invariant outside of a compact set and maps the Reeb vector R to $\frac{\partial}{\partial t}$. We also assume that J preserves the contact planes ξ. The symplectic form is chosen so that X is symplectomorphic to $\mathbb{C}P^2(R) \setminus W$, but the description of the complex structure on the cylindrical end is clearer in terms of X.

There is a theory of holomorphic curves mapping Riemann surfaces with punctures into X, which are asymptotic at their punctures to cylinders $\gamma \times (-\infty, 0)$, where γ is a closed Reeb orbit in ∂W. These are sometimes called finite energy curves. The literature is fairly extensive, fundamental works include [9], [10], [11], [12]. Of course, under the symplectomorphism from X to $\mathbb{C}P^2(R) \setminus W$ we could also think of holomorphic curves with image in $\mathbb{C}P^2(R) \setminus W$. In this case, the curves extend to maps from the oriented blow-up of the Riemann surface at its punctures, mapping the boundary circles to closed Reeb orbits on ∂W, see Proposition 5.10 of [3].

2.2.2 Area and index formulas for degree 1 curves

The punctures of a finite energy holomorphic curve are called elliptic if they are asymptotic to a γ_i^r or hyperbolic if they are asymptotic to a $\gamma_{m,n}$. We can define the Chern class $c_1(C)$ of such a curve C to be $3d$, where d is the degree, that is, the intersection number of C with $\mathbb{C}P^1(\infty)$. This is exactly the number of zeros of a section of the determinant line bundle $\Lambda^2(TX, J)|_C$ which agrees with a trivial section of $\mathbb{C}^2$ near the punctures.

Suppose that a curve C of degree d has e_1 punctures asymptotic to orbits $\gamma_1^{r_i}$ for $1 \leq i \leq e_1$ and e_2 punctures asymptotic to orbits $\gamma_2^{s_j}$ for $1 \leq j \leq e_2$. Also, suppose the curve has h hyperbolic punctures asymptotic to γ_{m_k,n_k} for $1 \leq k \leq h$ respectively.

We recall from section 2.1 that δ is a very small parameter chosen to control the approximation of our rounded domain W to $P(a,b)$.

Lemma 2.3. *The symplectic area of C is given by*

$$\int_C \omega = dR - \sum_{i=1}^{e_1} r_i a - \sum_{j=1}^{e_2} s_j b - \sum_{k=1}^{h} (m_k a + n_k b) + \delta(C),$$

where $\delta(C)$ is an error term of order δ.

Proof. A Reeb orbit $\gamma_{m,n} \subset \partial W$ lying over some (R_1, R_2) bounds a disk in W of area $mR_1 + nR_2$. Due to our rounding of the polydisk all hyperbolic orbits lie over radial coordinates which are approximately (a,b). The difference between the actual radial coordinates and (a,b) accounts for the error term in our formula.

Given this, each of the negative terms in the formula correspond roughly to the areas of disks which can be glued to C (thought of now as lying in $\mathbb{C}P^2(R) \setminus W$) to produce a closed cycle in $\mathbb{C}P^2$ of degree d and hence area dR. □

Remark 2.4. We observe immediately from Lemma 2.3 that curves of non-positive degree have negative area and so cannot exist for a tame almost-complex structure.

Under the assumption that $a+b < R < 2a+b$ we now document all genus 0 holomorphic curves of degree 1 satisfying a certain area restriction. The motivation for the restriction is that such curves can appear as boundary components of a certain moduli space of finite energy planes which will be the basis of our proof of Theorem 1.1. We will henceforth also use δ to denote the maximum of the possible error terms in Lemma 2.3 and assume it is chosen very small, in particular $\delta < 2a + b - R$.

Lemma 2.5. *Finite energy curves in X of degree* 1 *and area at most $R - a - b + \delta$ are of one of the following types.*

I. $e_2 = h = 0$, $1 + \frac{b}{a} \leq \sum_1^{e_1} r_i < 2 + \frac{b}{a}$;

II. $e_1 = h = 0$, $e_2 = 1$, $s_1 = 2$;

III. $e_1 = h = 0$, $e_2 = 2$, $s_1 = s_2 = 1$;

IV. $e_1 = e_2 = 1$, $h = 0$, $r_1 = s_1 = 1$;

V. $e_1 = e_2 = 0$, $h = 1$, $m_1 = n_1 = 1$.

Proof. The area formula of Lemma 2.3 implies that for such a curve we must have

$$-\delta < R - \sum_{i=1}^{e_1} r_i a - \sum_{j=1}^{e_2} s_j b - \sum_{k=1}^{h} (m_k a + n_k b) \leq R - a - b + \delta,$$

which, as $R < 2a + b$, implies

$$a + b - \delta \leq \sum_{i=1}^{e_1} r_i a + \sum_{j=1}^{e_2} s_j b + \sum_{k=1}^{h} (m_k a + n_k b) < 2a + b$$

when δ is sufficiently small. Now, recalling that the r_i, s_j, m_k and n_k are all nonnegative integers we can simply check that the only possibilities are those described, at least provided that δ is chosen sufficiently small. $\square$

Remark 2.6. As holomorphic curves have positive area, curves of types II and III are possible only if $b < 2a$.

Moduli spaces of finite energy curves, that is, spaces of such curves of a certain type modulo reparameterizations of the underlying Riemann surface, have a virtual index, and we now compute this for the types of curves appearing in Lemma 2.5.

Lemma 2.7. *The virtual index of the curves described in Lemma 2.5 are as follows.*

I. index $= 4 - 2\sum r_i \leq -2$;

II. index $= 0$;

III. index $= 0$;

IV. index $= 0$;

V. index $= 1$.

Proof. The general index formula is

$$\mathrm{index}(C) = e_1+e_2+h-2+2c_1(C)-\sum_{i=1}^{e_1}\mu(\gamma_1^{r_i})-\sum_{j=1}^{e_2}\mu(\gamma_2^{s_j})-\sum_{k=1}^{h}(\mu(\gamma_{m_k,n_k})-\frac{1}{2}\mathrm{dim}V_k).$$

For this, see [2]. Here $c_1(C)$ is the Chern class which we have normalized to be 3 for degree 1 curves, and $\mathrm{dim}V_k$ is the dimension of the family of hyperbolic orbits containing γ_{m_k,n_k}, namely $\mathrm{dim}V_k = 1$. Substituting the Conley-Zehnder indices from Lemma 2.2 we get

$$\mathrm{index}(C) = 4+e_1+e_2+h-2-\sum_{i=1}^{e_1}(2r_i+1)-\sum_{j=1}^{e_2}(2s_j+1)-\sum_{k=1}^{h}2(m_k+n_k)$$

or equivalently

$$\mathrm{index}(C) = 4+h-2\sum_{i=1}^{e_1}r_i-2\sum_{j=1}^{e_2}s_j-2\sum_{k=1}^{h}(m_k+n_k).$$

The formulas follow readily from here. For the inequality on the index for curves of type I we note from Lemma 2.5 that $\sum r_i \geq 1+\frac{b}{a}$ and so the index satisfies

$$\mathrm{index} = 4-2\sum r_i \leq 4-2(1+\frac{b}{a}) = 2(1-\frac{b}{a}) \leq -2$$

since we always assume $b > a$. □

2.3 Proof of Theorem 1.1

Our method of proof is to construct a moduli space of holomorphic curves $\mathcal{M}_t$ for $0 \leq t \leq 1$ corresponding to each polydisk (or rather copy of W) in a symplectic isotopy; $\mathcal{M}_0$ will relate to the standard polydisk and we shall argue by contradiction and suppose that $\mathcal{M}_1$ corresponds to a symplectically isotopic polydisk in $B(S)$. We will show that $\mathcal{M}_0$ is nonempty (and in fact represents a nontrivial cobordism class) while $\mathcal{M}_1$ is empty. The key to the proof is then that $\cup_t\mathcal{M}_t$ gives a compact cobordism between $\mathcal{M}_0$ and $\mathcal{M}_1$ provided all polydisks lie in $B(R)$ or $\mathbb{C}P^2(R)$ for some $R < 2a+b$. This gives a contradiction. If $R > 2a+b$ then the universal moduli space is noncompact.

2.3.1 Existence of finite energy curves

Here we establish the following. Recall from Lemma 2.5 that curves of type V are finite energy planes (that is, curves with domain the complex plane $\mathbb{C}$) of degree 1 which are asymptotic at infinity to an orbit of type $\gamma_{1,1}$.

Lemma 2.8. *There exists a compatible almost-complex structure J_0 on X such that there exist J_0 holomorphic finite energy curves of type V.*

Proof. We can produce finite energy curves in the manifold X by starting with curves in $\mathbb{C}P^2$ and stretching the neck along ∂W. This means that we start with a tame almost-complex structure restricting to a translation invariant structure on $\partial W \times (-\epsilon, \epsilon) \subset \mathbb{C}P^2$ satisfying $J(R) = \frac{\partial}{\partial t}$ and replace it by a sequence of tame almost-complex structures J_N satisfying $J_N(R) = \frac{1}{N}\frac{\partial}{\partial t}$, but such that $J_N = J$ on the contact planes ξ, which are preserved. For this see section 3.4 of [3]. We can choose embeddings ϕ_N of $\mathbb{C}P^2(R) \setminus W$ into X pushing forward J_N to an almost-complex structure which on increasingly large subsets of the cylindrical end $\partial W \times (-\infty, 0] \subset X$ is translation invariant and satisfies $J(R) = \frac{\partial}{\partial t}$.

Now, there exists a unique pair (R_1, R_2) with $K(R_1, R_2) = 0$ and $K_{R_1} = K_{R_2}$, see the definitions at the start of section 2.1. In the corresponding Lagrangian fiber the Reeb flow is generated by $\frac{\partial}{\partial \theta_1} + \frac{\partial}{\partial \theta_2}$. Let $c \in \mathbb{C}$ satisfy $|c|^2 = \frac{R_2}{R_1}$. Then any complex line $z_2 = cz_1$ (thought of as the affine part of a sphere in $\mathbb{C}P^2(R)$) is tangent to both the radial vector field Z (that is, $\frac{\partial}{\partial t}$ in the local coordinates on the neck region) and to the Reeb vector field on ∂W. Thus, if we choose our initial almost-complex structure J such that these curves are indeed holomorphic, then they will remain holomorphic during our stretching construction. Therefore, in the limit as $N \to \infty$, see [3], such curves converge to a holomorphic building whose component in X is a degree 1 finite energy plane asymptotic to an orbit $\gamma_{1,1}$. Denote by J_0 the limiting almost-complex structure on X. $\square$

2.3.2 Moduli spaces

Let us fix an orbit η of type $\gamma_{1,1}$. Consider the moduli space

$$\mathcal{M}_0 = \{u : \mathbb{C} \to X | \text{degree}(u) = 1, \overline{\partial}_{J_0} u = 0, u \sim \eta\}/G$$

where $u \sim \eta$ means that u is asymptotic at infinity to η, and G is the reparameterization group of $\mathbb{C}$.

Lemma 2.9. *The moduli space $\mathcal{M}_0$ is a non-empty oriented manifold of dimension 0, and all points have positive orientation.*

Proof. The fact that $\mathcal{M}_0$ is non-empty is Lemma 2.8 and its virtual dimension is 0 by Lemma 2.7, part V. The index is 0 here rather than 1 as we are requiring our curves to be asymptotic to a specific orbit η, rather than any orbit of the type $\gamma_{1,1}$. Moduli spaces of finite energy curves can be oriented by following [4]. That all points have the same orientation follows from automatic regularity for finite energy planes, which is due to C. Wendl, [18] Theorem 1. $\square$

Arguing by contradiction, suppose that we have a symplectic isotopy $W_t \subset B(R)$ for $0 \leq t \leq 1$ of the domain W with $W_0 = W$ and $W_1 \subset B(S)$. Then we can form a family of manifolds $X_t = \mathbb{C}P^2(R) \setminus W_t$ as before, equip them with almost-complex structures J_t, and study finite energy curves. The analysis is exactly the same as for X, in particular Lemmas 2.5 and 2.7 hold unchanged. The area formula of Lemma 2.3 remains the same since the isotopy is symplectic. We can now define moduli spaces $\mathcal{M}_t$ for each t and a universal moduli space

$$\mathcal{M} = \{(u,t) | u : \mathbb{C} \to X, \mathrm{degree}(u) = 1, \overline{\partial}_{J_t} u = 0, u \sim \eta, t \in [0,1]\}/G.$$

Assuming that the family J_t is chosen generically, $\mathcal{M}$ is a 1-dimensional manifold with boundary $\mathcal{M}_0 \coprod \mathcal{M}_1$. Next we have the following.

Lemma 2.10. *The moduli space $\mathcal{M}_1$ is empty.*

The proof of this will rely on a monotonicity theorem.

Lemma 2.11. *Let $u : \Sigma \to \mathbb{C}P^2(R) \setminus B(S)$ be a proper holomorphic curve where Σ is a Riemann surface and $\mathbb{C}P^2(R) \setminus B(S)$ is equipped with its integrable complex structure. Then $\int_\Sigma \omega \geq R - S$.*

Proof of Lemma 2.11. First we observe by the maximum principle that $u(\Sigma) \cap \mathbb{C}P^1(\infty) \neq \emptyset$. There exist disjoint balls in $\mathbb{C}P^2(R)$ with its integrable complex structure embedded both holomorphically and symplectically of capacities S and $R - S$ respectively. Using a holomorphic isometry we may assume that the ball of capacity S is the standard one and that of capacity $R - S$ is centered at a point of $u(\Sigma) \cap \mathbb{C}P^1(\infty)$. Then $u(\Sigma)$ passes through the center of a ball of capacity $R - S$ and therefore has area at least $R - S$ from the monotonicity theorem, see [8], section $2.3.E_2'$. $\square$

Proof of Lemma 2.10. First note that by Lemma 2.3 curves in $\mathcal{M}$ have area $R-(a+b)$, at least up to an error of order δ. Next, as $W_1 \subset B(S)$ we can restrict elements of $\mathcal{M}_1$ to $\mathbb{C}P^2(R) \setminus B(S)$ and as holomorphic curves have positive area any such restriction has area bounded above by $R-(a+b)+\delta$, which, by choosing δ sufficiently small, we may assume is (strictly) bounded above by $R-S$ (as $S < a+b$). As W_1 is disjoint from $\mathbb{C}P^2(R) \setminus B(S)$ we may assume that J_1 restricts to the standard integrable structure here, then we have a contradiction to Lemma 2.11. $\square$

Together Lemmas 2.9 and 2.10 imply that the moduli space $\mathcal{M}$ is not compact. In the next section we investigate this using the compactness theorem of [3].

2.3.3 Compactness

According to [3], any sequence $(u_t, t) \in \mathcal{M}$ has a subsequence converging in the sense of holomorphic buildings. Roughly speaking, a holomorphic building is a union of finite energy curves whose domains are components of a nodal Riemann surface minus the nodes. Matching nodes in the nodal Riemann surface correspond to matching asymptotic limits.

More precisely, in our situation the nodal Riemann surface is a degeneration of the complex plane given by contracting circles. The components of the limiting holomorphic building map to either X or the symplectization $\partial W \times \mathbb{R}$ of ∂W. For the components in $\partial W \times \mathbb{R}$, we can distinguish positive and negative punctures in the obvious way. Each positive puncture is matched either to a negative puncture of another component, or to a puncture of a component in X. There will be one unmatched asymptotic limit, corresponding to the original puncture in $\mathbb{C}$ which is asymptotic to η.

It follows from this that the curves in the symplectization layer can be projected to ∂W and glued along their matching asymptotic limits to produce a map from a (possibly disconnected) Riemann surface whose remaining positive limits are the asymptotic limits of components in X and which has a single remaining negative limit asymptotic to η.

The sum of the degrees of the components in X is 1, but as by Remark 2.4 there are no components of nonpositive degree, we conclude that the limit has a single component in X of degree 1. Also, as area is preserved in the limit and curves of $\mathcal{M}$ have area $R-(a+b)$, we have that the component in X is of one of those types listed in Lemma 2.5. Furthermore, assuming that we choose a sufficiently generic family J_t of almost-complex structures, any limiting components appearing should have index at least -1. Hence the limit is of type II, III, IV or V. We deal with these separately.

If the limit is of type V then it is asymptotic to an orbit $\gamma_{1,1}$ and any components in $\partial W \times \mathbb{R}$ have area 0 (as all $\gamma_{1,1}$ orbits have exactly the same area). Thus these components must be trivial cylinders, the $\gamma_{1,1}$ orbit is actually η and (u_t, t) converge to an element of $\mathcal{M}$.

Suppose now that this component is of type IV. Then the components in $\partial W \times \mathbb{R}$ have a total of two positive ends, asymptotic to γ_1 and γ_2 respectively, and a negative end asymptotic to η. As all such components have nonnegative area and both γ_1 and γ_2 have smaller action than η we conclude that there must be a single component in $\partial W \times \mathbb{R}$ with two positive ends and a single negative end. But this is impossible as gluing the positive ends of curves in $\partial W \times \mathbb{R}$ to the negative ends of curves in X we must produce a surface of genus 0, the genus of the u_t. Gluing the limits as described here results in a genus 1 curve.

Next suppose this component is of type III. Then the components in $\partial W \times \mathbb{R}$ again have two positive ends, and we can exclude such curves as for type IV.

Hence if there exists an isotopy with $W_1 \subset B(S)$ we are left with the conclusion that a sequence $(u_t, t) \in \mathcal{M}$ converges to a holomorphic building whose component in X is of type II. In other words, for some $t_0 \in [0, 1]$ we have produced a J_{t_0}-holomorphic finite energy plane asymptotic to γ_2^2.

Remark 2.12. Remark 2.6 observed that this is already a contradiction in the case that $b > 2a$, and given Corollary 1.4 this is precisely the case in which we are especially interested.

Despite the above remark, we proceed to prove Theorem 1.1 in its entirety by showing that the bubbling of type II curves also leads to a contradiction. We repeat the argument above using type II curves instead of type V and for t now in the interval $[t_0, 1]$. To be precise, we define moduli spaces

$$\mathcal{N}_t = \{u : \mathbb{C} \to X | \text{degree}(u) = 1, \overline{\partial}_{J_t} u = 0, u \sim \gamma_2^2\}/G$$

and a universal moduli space

$$\mathcal{N} = \{(u,t) | u : \mathbb{C} \to X, \text{degree}(u) = 1, \overline{\partial}_{J_t} u = 0, u \sim \gamma_2^2, t \in [t_0, 1]\}/G.$$

Again we see that $\mathcal{N}_{t_0}$ consists of a number of points which are all counted positively, and $\mathcal{N}_1$ is empty. However now the 1-dimensional manifold $\mathcal{N}$ is compact. Indeed, if there is a loss of compactness we must see curves of types I, III, IV or V in X. But type I can be still be excluded by genericity, and type III in exactly the same way as above. Finally curves of types IV and V have strictly larger area than curves of type II as $b > a$ and so do not appear either. Hence $\mathcal{N}$ is a compact oriented cobordism from $\mathcal{N}_0$ to the empty set and this is a contradiction.

3 Symplectic folding

Here we establish Theorem 1.3 by following the symplectic folding construction while minimizing the size of the balls in which the support of our various diffeomorphisms lie. It clearly suffices to find a Hamiltonian diffeomorphism with support in an arbitrarily small neighborhood of $B(2a+b)$ mapping $P(a,b)$ arbitrarily close to $B(2a+\frac{b}{a})$.

For convenience we will use symplectic polar coordinates (R,θ) on the (x_2,y_2) plane, where $R = \pi(x_2^2+y_2^2)$ and $\tan\theta = \frac{y_2}{x_2}$. Therefore we can write a disk of area a in this plane as $D(a) = \{R \le a\}$.

We start with a polydisk $P(a,b) = D(a) \times D(b)$ where $D(a)$ and $D(b)$ denote disks in the (x_1,y_1) and (x_2,y_2) planes respectively. This polydisk lies inside $B(a+b)$.

Step 1. We apply a symplectomorphism ψ_1 of the (R,θ) plane mapping $D(b)$ to an arbitrarily small neighborhood of

$$U = \{R \le \frac{b}{2}\} \cup \{\theta = 0, \frac{b}{2} \le R \le a + \frac{b}{2} + \delta\} \cup \{a + \frac{b}{2} + \delta \le R \le a + b + \delta\}.$$

We can take δ arbitrarily small, but for the following steps to apply it must be strictly positive. Figure 1 is a sketch of the image of $D(b)$.

Now for Step 1 we apply the symplectomorphism $\phi_1 = \text{id.} \times \psi_1$ to the polydisk $P(a,b)$. As U lies in a disk $D(a+b+\delta)$ the image of $P(a,b)$ lies in a ball $B(a+b+\delta+a)$ and we can realize ϕ_1 as a Hamiltonian diffeomorphism with support arbitrarily close to $B(2a+b)$. Let $P_1 = \phi_1(P(a,b))$.

Step 2. There exists a Hamiltonian diffeomorphism ψ_2 displacing $D(a)$ from itself and having compact support in a small neighborhood of $D(2a)$. Let $H_2(x_1,y_1)$ be a generating Hamiltonian, that is, ψ_2 is the time-1 flow of the Hamiltonian vector field corresponding to H_2. We may assume that $0 \le H_2 \le a$ and it is also not hard to arrange that for all $0 < \lambda < 1$ the

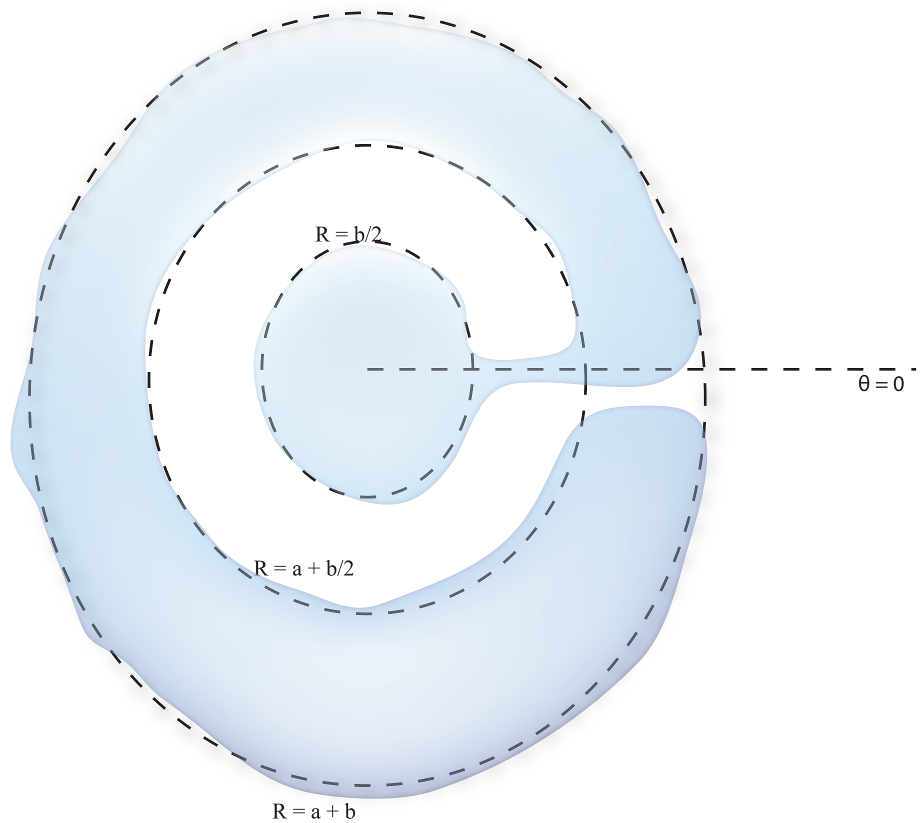

Figure 1: The image of $D(b)$ in the (R, θ) plane.

Hamiltonian diffeomorphism generated by λH_2 maps $D(a)$ into a neighborhood of $D((1+\lambda)a)$.

Define $\chi(R)$ to be a decreasing function equal to 1 when $R \leq \frac{b}{2}$ and 0 when $R \geq a + \frac{b}{2} + \delta$ and having slope bounded by $\frac{1}{a}$.

Our second step is to apply the Hamiltonian diffeomorphism ϕ_2 generated by $\chi(R)H_2(x_1, y_1)$. Set $P_2 = \phi_2(P_1)$. We examine separately the images of points with $\{R \leq \frac{b}{2}\}$, $\{\frac{b}{2} \leq R \leq a + \frac{b}{2} + \delta\}$ and $\{a + \frac{b}{2} + \delta \leq R \leq a + b + \delta\}$ under this flow. As the generating Hamiltonian is independent of θ the flow preserves the R coordinate and in particular these regions. Figure 2 is a sketch of the projection of P_2 to the (R, θ) plane, the figure is justified by the following analysis.

• If $R \leq \frac{b}{2}$ then the flow fixes (R, θ) and the (x_1, y_1) coordinates remain in a neighborhood of a disk $D(2a)$. Thus the flow remains close to a ball

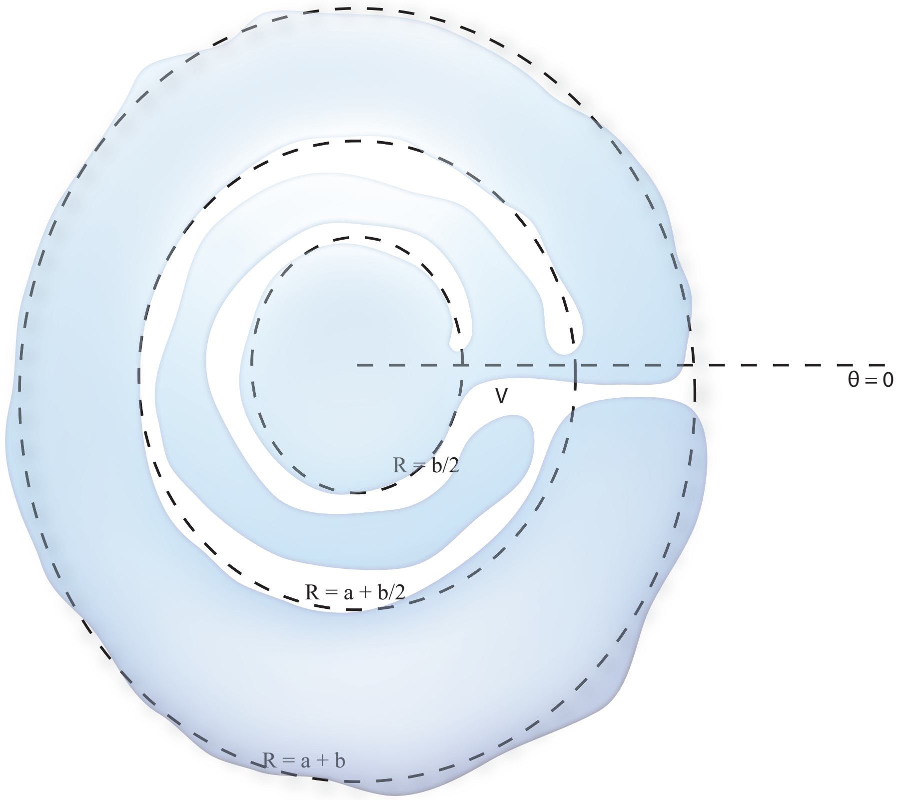

Figure 2: The projection of P_2 to the (R, θ) plane.

$B(2a + \frac{b}{2})$.

• If $\frac{b}{2} \le R \le a + \frac{b}{2} + \delta$ then, as the flow of $\lambda H_2(x_1, y_1)$ maps $D(a)$ into a neighborhood of $D((1+\lambda)a)$, the trace of the (x_1, y_1) coordinates throughout lie in a neighborhood of a disk $D((1+\chi(R))a)$. As χ is roughly $\frac{a+b/2-R}{a}$, this disk is roughly equal to $D(2a + \frac{b}{2} - R)$. Therefore, as the corresponding (x_2, y_2) coordinates lie in $\partial D(R)$, the flow remains close to the ball $B(2a+\frac{b}{2})$. Meanwhile the rate of increase of the θ coordinate is bounded by $2\pi\chi'(R)H_2(x_1, y_1) < 2\pi$. Hence we may assume that the projection of the image to the (R, θ) plane continues to avoid a narrow segment V just below the $\{\theta = 0\}$ axis as shown in Figure 2.

• Finally, the flow is constant in the region $\{a + \frac{b}{2} + \delta \le R \le a + b\}$.

Step 3. Here we apply a symplectomorphism ϕ_3 generated by a Hamiltonian $H_3(R, \theta)$ with support in a neighborhood of $D(a+b+\delta)$ only to points of P_2 which lie in the region $\{a + \frac{b}{2} + \delta \le R \le a + b + \delta\}$. The flow of H_3 will fix points close to the boundary $\{a + \frac{b}{2} + \delta = R\}$ of this region and so can be extended as a constant to the rest of P_2. In fact, we will take H_3 to be identically 0 on the region $\{\frac{b}{2} \le R \le a + \frac{b}{2} + \delta\} \setminus V$.

We choose H_3 such that the corresponding flow moves all of $P_2 \cap \{a + \frac{b}{2} + \delta \le R \le a + b + \delta\}$ to a neighborhood of $\{R \le \frac{b}{2}\} \cup V$. Such an H_3 exists because $\{a + \frac{b}{2} + \delta \le R \le a + b + \delta\}$ has area roughly $\frac{b}{2}$ and we may assume that the projection of P_2 to the (R, θ) plane is simply connected, see Figure 1. Then H_3, thought of as a Hamiltonian on $\mathbb{R}^4$, can be used to define a symplectomorphism ϕ_3 of P_2 once we check that the traces of points moved by the flow of H_3 are disjoint from the points of P_2 where we are defining ϕ_3 to be the identity. This follows since the only possible intersections lie in $\{R \le \frac{b}{2}\}$, but the (x_1, y_1) coordinates of points in $P_2 \cap \{R \le \frac{b}{2}\}$ lie outside $D(a)$ while those of points in $P_2 \cap \{a+\frac{b}{2}+\delta \le R \le a+b+\delta\}$ lie in $D(a)$ (and ϕ_3 fixes our (x_1, y_1) coordinates). Hence we have a well-defined ϕ_3, a symplectic folding map. The only points moved by the flow have $(x_1, y_1) \in D(a)$ and $R \le a+b+\delta$, and so the Hamiltonian isotopy remains in an arbitrarily small neighborhood of $B(2a + b)$.

Finally, we recall that points of $P_2 \cap \{R \le a + \frac{b}{2} + \delta\}$ all lie in a neighborhood of $B(2a + \frac{b}{2})$, and since they fixed by ϕ_3 the image $P_3 = \phi_3(P_2)$ lies in the union of a neighborhood of $B(2a + \frac{b}{2})$ together with $\phi_3(P_2 \cap \{a + \frac{b}{2} + \delta \le R \le a + b + \delta\})$. This second part of the image lies in a neighborhood of $D(a) \times \{R \le a+\frac{b}{2}+\delta\}$ which in turn lies also in a neighborhood of $B(2a+\frac{b}{2})$.

In summary, our isotopy has support in a neighborhood of $B(2a+b)$ (we already approach the boundary of this region in Step 1), and the resulting diffeomorphism has image close to $B(2a+\frac{b}{2})$ as required.

References

[1] F. Bourgeois, A Morse-Bott approach to contact homology, PhD Thesis, Stanford University, 2002.

[2] F. Bourgeois, A Morse-Bott approach to contact homology. Symplectic and contact topology: interactions and perspectives (Toronto, ON/Montreal, QC, 2001), 55–77, Fields Inst. Commun., 35, Amer. Math. Soc., Providence, RI, 2003.

[3] F. Bourgeois, Y. Eliashberg, H. Hofer, K. Wysocki, and E. Zehnder, Compactness results in symplectic field theory, *Geom. Topol.*, **7** (2003), 799–888.

[4] F. Bourgeois and K. Mohnke, Coherent orientations in symplectic field theory, *Math. Z.*, 248 (2004), no. 1, 123–146.

Hamiltonian dynamics, *Math. Z.*, **200** (1989), 355–378.

[5] I. Ekeland and H. Hofer, Symplectic topology and Hamiltonian dynamics II, *Math. Z.*, **203** (1990), 553–567.

[6] Y. Eliashberg, A. Givental and H. Hofer, Introduction to symplectic field theory, GAFA 2000 (Tel Aviv, 1999), *Geom. Funct. Anal.*, 2000, Special Volume, Part II, 560–673.

[7] A. Floer, H. Hofer and K. Wysocki, Applications of symplectic homology. I. *Math. Z.*, **217** (1994), 577-606.

[8] M. Gromov, Pseudo-holomorphic curves in symplectic manifolds, *Inv. Math.*, **82** (1985), 307–347.

[9] H. Hofer, K. Wysocki and E. Zehnder, A characterisation of the tight three-sphere, *Duke Math. J.*, **81** (1995), 159-226.

[10] H. Hofer, K. Wysocki and E. Zehnder, Properties of pseudoholomorphic curves in symplectisations I: Asymptotics, *Ann. Inst. H. Poincaré Anal. Non Lineaire*, **13** (1996) , 337–379.

[11] H. Hofer, K. Wysocki and E. Zehnder, Properties of pseudoholomorphic curves in symplectisations II: Embedding controls and algebraic invariants, *Geom. Funct. Anal.*, **5** (1995), 337–379.

[12] H. Hofer, K. Wysocki and E. Zehnder, Properties of pseudoholomorphic curves in symplectisations III: Fredholm theory, *Topics in nonlinear analysis*, 381-475, Prog. Nonlinear Differential Equations Appl., 35, Birkhäuser, Basel, 1999.

[13] M. Hutchings and M. Sullivan, Rounding corners of polygons and the embedded contact homology of T^3, *Geom. Topol.*, **10** (2006), 169-266.

[14] F. Lalonde and D. McDuff, The geometry of symplectic energy, *Ann. of Math.*, **141** (1995), 349–371.

[15] D. McDuff, Symplectic embeddings of 4-dimensional ellipsoids, *J. Topol.*, **2** (2009), 1-22.

[16] J. Robbin and D. Salamon, The Maslov index for paths, *Topology*, **32** (1993), 827–844.

[17] F. Schlenk, Embedding problems in symplectic geometry De Gruyter Expositions in Mathematics 40. Walter de Gruyter Verlag, Berlin. 2005.

[18] C. Wendl, Automatic transversality and orbifolds of punctured holomorphic curves in dimension four, *Comment. Math. Helv.*, **85** (2010), 347–407.