J. DIFFERENTIAL GEOMETRY
88 (2011) 519-532

THE HOFER CONJECTURE ON EMBEDDING SYMPLECTIC ELLIPSOIDS

Dusa McDuff

Abstract

In this note we show that one open 4-dimensional ellipsoid embeds symplectically into another if and only if the ECH capacities of the first are no larger than those of the second. This proves a conjecture due to Hofer. The argument uses the equivalence of the ellipsoidal embedding problem with a ball embedding problem that was recently established by McDuff. Its method is inspired by Hutchings' recent results on embedded contact homology (ECH) capacities but does not use them.

1. Introduction

Consider the ellipsoid $E(a,b) := \{z \in \mathbb{C}^2 : \frac{|z_1|^2}{a} + \frac{|z_2|^2}{b} \le 1\}$ with the symplectic structure induced from the standard structure on Euclidean space. Define $\mathcal{N}(a,b)$ to be the sequence of numbers formed by arranging all the positive integer combinations $ma+nb, m, n \ge 0$ in nondecreasing order (with repetitions). We say that $\mathcal{N}(a,b)$ is less than or equal to $\mathcal{N}(c,d)$ (written $\mathcal{N}(a,b) \preccurlyeq \mathcal{N}(c,d)$) if, for all $k \ge 0$, the kth entry of $\mathcal{N}(a,b)$ is at most equal to the kth entry in $\mathcal{N}(c,d)$. Hofer's conjecture evolved as earlier guesses, such as those by Cieliebak, Hofer, Latschev, and Schlenk in [**3**], proved inadequate. Finally, in private conversation, he conjectured that the numbers $\mathcal{N}(a,b)$ should detect precisely when these embeddings exist.

We show in this note that this is indeed the case.

Theorem 1.1. *There is a symplectic embedding* $\operatorname{int} E(a,b) \overset{s}{\hookrightarrow} E(c,d)$ *exactly if*

$$\mathcal{N}(a,b) \preccurlyeq \mathcal{N}(c,d).$$

Using embedded contact homology (ECH), Hutchings showed in [**6**] that the condition $\mathcal{N}(a,b) \preccurlyeq \mathcal{N}(c,d)$ is necessary. In fact, the main result of his paper is that there are quantities called ECH capacities, defined for any closed bounded subset of $\mathbb{R}^4$, that are monotone under symplectic embeddings. The application to embedding ellipsoids then follows because the ECH capacities of $E(a,b)$ are just the sequence

Received 8/11/2010.

$\mathcal{N}(a,b)$. Hutchings also shows that his ECH capacities give sharp obstructions to the problem of embedding a union of disjoint balls into a ball.

In the case of embedding an ellipsoid into a ball, McDuff–Schlenk [**14**, thm. 1.1.3] calculated exactly when the embedding exists, and concluded that in this case the condition $\mathcal{N}(a,b) \preccurlyeq \mathcal{N}(c,d)$ is sufficient. Combining these results, we see that Hofer's conjecture holds when the target is a ball. Below we prove the result in general by a much shorter argument that uses none of the geometric results in ECH. Instead it uses some elementary combinatorics that develop some of Hutchings' ideas, as well as the result from McDuff [**12**] that reduces the ellipsoidal embedding problem to a ball embedding problem. See Hutchings [**7**] for a survey that gives more of the background.

The higher dimensional analog of Theorem 1.1 is completely open; there is even no good guess of what the answer should be. However, the analog of the Hofer conjecture does not hold. The first counterexamples are due to Guth [**4**] who showed that there are constants a, b, c such that $E(1, R, R)$ embeds symplectically in $E(a, b, cR^2)$ for all $R > 0$, with similar results in higher dimensions. In [**5**] Hind–Kerman improved Guth's embedding method to show that $E(1, R, \dots, R)$ embeds in $E(a, a, R^2, \dots, R^2)$ whenever $a > 3$, but found an obstruction when $a < 3$.

Example 1.2. (i) The sequence $\mathcal{N}(a,a)$ is

$$\mathcal{N}(a,a) = (0, a, a, 2a, 2a, 2a, 3a, 3a, 3a, 3a, \dots).$$

In other words, for each d there are $d+1$ entries of da occurring as the terms $\mathcal{N}_k(a,a)$ for $\frac{1}{2}(d^2+d) \le k \le \frac{1}{2}(d^2+3d)$. Thus $\mathcal{N}(a,a)$ is the maximal sequence with $\mathcal{N}_k = da$ for $k = \frac{1}{2}(d^2+3d)$.

(ii) When $k = \frac{1}{2}(d^2+3d)$, the sequence

$$\mathcal{N}(1,4) = (0, 1, 2, 3, 4, 4, 5, 5, 6, 6, 7, 7, 8, 8, 8, \dots)$$

has $\mathcal{N}_k(1,4) = 2d = \mathcal{N}_k(2,2)$. Hence the maximality of $\mathcal{N}(2,2)$ implies that $\mathcal{N}(1,4) \preccurlyeq \mathcal{N}(2,2)$. Thus $\operatorname{int} E(1,4) \overset{s}{\hookrightarrow} B(2)$. (Here, and elsewhere, we write $X \overset{s}{\hookrightarrow} Y$ to mean that X embeds symplectically in Y.)

The first construction of an embedding of this kind is due to Opshtein [**16**].[1] The paper [**12**] develops a general method of embedding ellipsoids, which in most cases is not very geometric. However, as is shown in [**12**, §1], in the special case of $E(1,4)$ the argument can be made rather explicit. One still cannot see the geometry of the image as clearly as in Opshtein because one uses symplectic inflation to increase

[1] In fact, he constructed an explicit embedding from $\operatorname{int} E(1,4)$ into projective space by using properties of neighborhoods of curves of degree 2, but one can easily arrange that the embedding avoids a line so that there is a corresponding embedding into a ball. Cf. also theorem 4 in [**17**].

the size of the image of the ellipsoid, i.e. rather than embedding larger and larger ellipsoids into a fixed ball, one embeds a small ellipsoid E into the ball B and then increases the relative size of E by distorting the symplectic form on $B \setminus E$.

Remark 1.3. (i) We phrase all our results in terms of embedding the interior of X into Y (or, equivalently, embedding $\operatorname{int} X$ into $\operatorname{int} Y$), while Hutchings talks about embedding X into the interior of Y. But these amount to the same when X is the disjoint union of ellipsoids, since in this case there is a symplectic embedding $\operatorname{int} X \overset{s}{\hookrightarrow} Y$ exactly if λX embeds symplectically into $\operatorname{int} Y$ for all $\lambda < 1$; cf. [**12**, cor. 1.5].

(ii) The current methods extend to give a simple numerical criterion for embedding disjoint unions of ellipsoids into an ellipsoid. See Proposition 3.5 for a precise result. Also, all the methods used here extend to the case when the target manifold is a polydisc, i.e. a product of two discs with a product form, cf. Müller [**15**], or a blowup of a rational or ruled surface. However, just as in [**10**], one gets no information when the target is a closed 4-manifold with $b_2^+ > 1$ such as T^4 or a surface of general type.

See Bauer [**1**] for more information on the numerical properties of the sequences $\mathcal{N}(a, b)$.

Acknowledgments. I warmly thank Michael Hutchings for his patient explanations of the index calculations in ECH and for many illuminating discussions; Felix Schlenk for some useful comments on an earlier version of this note; and also FIM at ETH, Zürich, for providing a very stimulating environment at Edifest, 2010. This work was partially supported by NSF grant DMS 0905191.

2. Combinatorics

A standard continuity argument implies that it suffices to prove Theorem 1.1 when the ratios b/a and d/c are rational. One of the main results of [**12**] states that for each integral ellipsoid $E(a, b)$ there is a sequence of integers $\mathbf{W}(a, b) := (W_1, \ldots, W_M)$ called the (normalized) **weight sequence** of a, b, such that $\operatorname{int} E(a, b)$ embeds in $B(\mu)$ exactly if the disjoint union $\sqcup \operatorname{int} B(\mathbf{W}) := \sqcup \operatorname{int} B(W_i)$ embeds in $B(\mu)$. This section shows how $\mathcal{N}(a, b)$ may be calculated in terms of $\mathbf{W}(a, b)$.

We begin with some definitions. They are basically taken from [**12**], but are modified as in [**14**]. Given positive integers p, q with $q \le p$, we denote by $\mathbf{W}(p, q) = \mathbf{W}(q, p)$ the normalized weight sequence of p/q. Thus $\mathbf{W}(p, q) = (W_1, W_2, \ldots, W_M)$ is a finite sequence of positive integers defined recursively by the following rules:[2]

[2]The proof that the weight expansion $\mathbf{W}(p, q)$ described above agrees with that used in [**12**] is given in the appendix to [**14**].

- $W_1 = q$ and $W_n \geq W_{n+1} > 0$ for all n;
- if $W_i > W_{i+1} = \cdots = W_n$ (where we set $W_0 := p$), then

$$W_{n+1} = \min\bigl\{W_n,\ W_i - (n-i)W_{i+1}\bigr\};$$

- the sequence stops at W_M if the above formula gives $W_{M+1} = 0$.

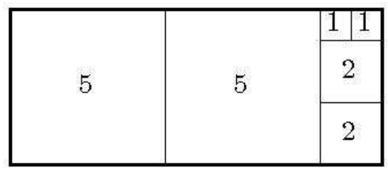

Figure 2.1. One obtains the weights by cutting a rectangle into squares: $W(5,12) = (5,5,2,2,1,1) =: (5^{\times 2}, 2^{\times 2}, 1^{\times 2})$.

It is often convenient to write $\mathbf{W}(p,q)$ as

$$\mathbf{W}(p,q) = \bigl(X_0^{\times \ell_0}, X_1^{\times \ell_1}, \dots, X_K^{\times \ell_K}\bigr), \tag{2.1}$$

where $X_0 > X_1 > \cdots > X_K > 0$ and $\ell_K \geq 2$. Thus the ℓ_i are the multiplicities of the entries in $\mathbf{W}(p,q)$ and, as is well known, give the continued fraction expansion of p/q: namely,

$$\frac{p}{q} = \ell_0 + \frac{1}{\ell_1 + \frac{1}{\ell_2 + \dots \frac{1}{\ell_K}}} =: [\ell_0; \ell_1, \dots, \ell_K]. \tag{2.2}$$

In this notation, the defining formulas for the terms in $\mathbf{W}(p,q)$ become:

$$X_{-1} := p, \quad X_0 = q, \quad X_{i+1} = X_{i-1} - \ell_i X_i, \ i \geq 0.$$

In particular, because $X_1 = p - \ell_0 q$,

$$W(p,q) = \bigl(q^{\times \ell_0}, X_1^{\times \ell_1}, X_2^{\times \ell_2}, \dots\bigr) = \bigl(q^{\times \ell_0}, W(q, X_1)\bigr). \tag{2.3}$$

More generally, the following holds.

Lemma 2.1. *Define $p, q, X_i, \ell_i, 0 \leq i \leq K$ as in equations* (2.1) *and* (2.2) *and set $X_{-1} := p, \ell_{-1} := 1$. Then for all $i = 0, \dots, K-1$, we have*

$$\mathbf{W}(X_{i-1}, X_i) = \bigl(X_i^{\times \ell_i}, \mathbf{W}(X_i, X_{i+1})\bigr).$$

Remark 2.2. The weight sequence $\mathbf{W}(p,q)$ does not seem to be mentioned in elementary treatments of continued fractions. Instead, one considers the convergents $p_k/q_k := [\ell_0; \ell_1, \dots, \ell_k]$ of p/q. However, it is well known that the two mirror fractions $p/q := [\ell_0; \ell_1, \dots, \ell_K]$ and $P/Q := [\ell_K; \ell_{K-1}, \dots, \ell_0]$ have the same numerator $p = P$. It follows that when p, q are relatively prime the sequence

$$(X_0, X_1, \dots, X_K),$$

when taken in reverse order, is just the sequence of numerators of the convergents of the mirror P/Q. More precisely, $X_{K-k} = P_{k-1}$ for $1 \le k \le K+1$. Further, the rectangle definition of the W_i (cf. Figure 2.1) easily implies that $\sum_i W_i^2 = pq$. Another, less obvious, quadratic relation for the W_i is discussed in [**14**, §2.2].

We now show that $\mathcal{N}(a,b)$ may be calculated from the weights $\mathbf{W}(a,b)$ using an operation # first considered by Hutchings. Given two nondecreasing sequences $\mathcal{C} = (\mathcal{C}_k)_{k\ge 0}, \mathcal{D} = (\mathcal{D}_k)_{k\ge 0}$ with $\mathcal{C}_0 = \mathcal{D}_0 = 0$, define $\mathcal{C}\#\mathcal{D}$ by

$$(\mathcal{C}\#\mathcal{D})_k := \max_{0\le i\le k}(\mathcal{C}_i + \mathcal{D}_{k-i}).$$

The next result follows immediately from the definition.

Lemma 2.3. (i) *The operation $\mathcal{C},\mathcal{D} \mapsto \mathcal{C}\#\mathcal{D}$ is associative and commutative.*
(ii) *If $\mathcal{C} \preccurlyeq \mathcal{C}'$ then $\mathcal{C}\#\mathcal{D} \preccurlyeq \mathcal{C}'\#\mathcal{D}$ for all sequences $\mathcal{D}$.*

For any sequence $\mathbf{a} = (a_1, \ldots, a_M)$ of positive integers, we define

$$\mathcal{N}(\mathbf{a}) := \mathcal{N}(a_1,a_1)\#\mathcal{N}(a_2,a_2)\#\ldots\#\mathcal{N}(a_M,a_M).$$

If $a := a_1 = \cdots = a_M$, we abbreviate this product as $\#^M\mathcal{N}(a,a)$.

To understand the effect of the operation # on sequences of the form $\mathcal{N}(a,b)$, it is convenient to interpret the numbers $\mathcal{N}_k(a,b)$ in terms of lattice counting as in Hutchings [**6**, §3.3]. For $A > 0$ consider the triangle

$$T^A_{a,b} := \big\{(x,y) \in \mathbb{R}^2 : x,y \ge 0,\ ax+by \le A\big\}.$$

Each integer point $(m,n) \in T^A_{a,b}$ gives rise to an element of the sequence $\mathcal{N}(a,b)$ that is $\le A$. If a/b is irrational, there is for all A at most one integer point on the slant edge of $T^A_{a,b}$. It follows that $\mathcal{N}_k(a,b) = A$, where A is such that $|T^A_{a,b} \cap \mathbb{Z}^2| = k+1$. Since for rational a/b there might be more than one integral point on this slant edge, the general definition is:

$$\mathcal{N}_k(a,b) = \inf\big\{A : |T^A_{a,b} \cap \mathbb{Z}^2| \ge k+1\big\}.$$

Lemma 2.4. *For all $a,b > 0$, we have $\mathcal{N}(a,a)\#\mathcal{N}(a,b) = \mathcal{N}(a,a+b)$. More generally, for all $\ell \ge 1$, we have*

$$\big(\#^\ell\mathcal{N}(a,a)\big)\#\mathcal{N}(a,b) = \mathcal{N}(a,b+\ell a).$$

Proof. By continuity and scaling, it suffices to prove this when $a,b \in \mathbb{Z}$. Suppose that $\mathcal{N}_k(a,a+b) = A$. Then there is at least one integer point (m,n) on the slant edge QP of the triangle $T := T^A_{a,a+b}$; see Figure 2.2 (I). Let $d := \lceil \frac{A}{a+b} \rceil$ be the smallest integer greater than $\frac{A}{a+b}$, and let R be the point where the line joining $Y = (0,d)$ to $X = (d,0)$ meets the slant edge of T. Then because the intersection of the line $x+y = d-1$ with the first quadrant lies entirely in T, all the points in the interior of

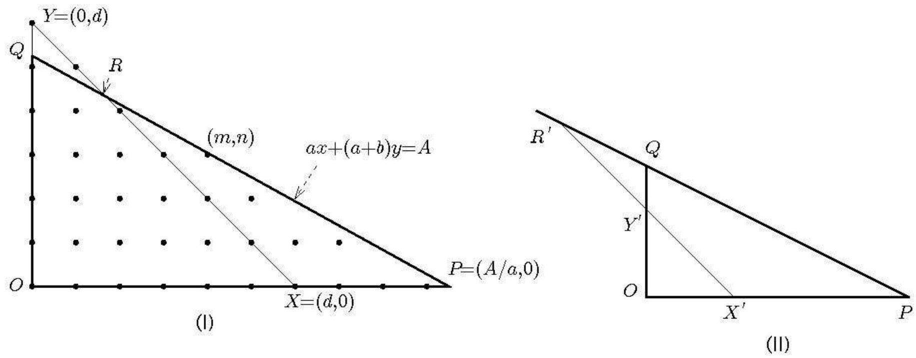

Figure 2.2. Different cuts in Lemma 2.4.

triangle OXY lie in T. Hence $\mathcal{S} := T \cap \mathbb{Z}^2$ divides into two sets $\mathcal{S}_1$ and $\mathcal{S}_2$, where $\mathcal{S}_1$ contains X plus all points (m,n) in T with $m+n<d$, and $\mathcal{S}_2$ consists of all other points in $\mathcal{S}$. Let

$$|\mathcal{S}_1| = k_1 + 1, \quad |\mathcal{S}_2| = k_2.$$

Then $k_1+k_2 = k$, and our remark above about the triangle OXY implies $\mathcal{N}_{k_1}(a,a) = da$.

Let α be the integral affine transformation that fixes Y and translates the x axis by $-d$ so that X goes to the origin O. Then α takes the triangle XRP to the triangle $T^B_{a,b}$, where $B/a = A/a - d$. The set of integral points in $T^B_{a,b}$ is $\alpha(\mathcal{S}_2 \cup \{X\})$. Hence $\mathcal{N}_{k_2}(a,b) = B = A - da$. Thus $\mathcal{N}_{k_1}(a,a) + \mathcal{N}_{k_2}(a,b) = \mathcal{N}_k(a,a+b)$.

We claim that for all other $i \le k$ we have $\mathcal{N}_i(a,a) + \mathcal{N}_{k-i}(a,b) \le \mathcal{N}_k(a,a+b)$. To see this, we slightly modify the above argument as follows. If $i > k_1$, then $\mathcal{N}_i(a,a) = ad'$ for some $d' \ge d$. Decompose T by the line $x+y = d'$ as above. Then choose $\mathcal{S}'_1 \subset \mathcal{S}$ to contain $X' = (d',0)$ together with all points in T with $x+y<d'$ and let $\mathcal{S}'_2 = \mathcal{S} \setminus \mathcal{S}'_1$. Then $k'_1 + 1 := |\mathcal{S}'_1| \le i+1$ (since now there may be some integer points in the interior of triangle $QY'R'$) so that $\mathcal{N}_{k'_1}(a,a) \le ad'$, while, as above, $\mathcal{N}_{k-k'_1}(a,b) = A - ad'$.

If $i < k_1$, then we choose d' as before and again slice T by the line $x+y = d'$. The corresponding triangle $OX'Y'$ is illustrated in Figure 2.2 (II). The line $X'Y'$ now meets the slant edge of T at R' lying beyond Q. Hence, if we partition the integral points in T as before, $k'_2+1 := |S_2|+1$ is at most the number $\ell+1$ of integral points in $X'R'P$ (and may well be strictly smaller). Thus $\mathcal{N}_{k'_2}(a,b) \le \mathcal{N}_\ell(a,b) = A - ad'$. On the other hand, $\mathcal{N}_{k'_1}(a,a) = ad'$ as before.

This completes the proof of the first statement. The second follows immediately by induction. q.e.d.

Corollary 2.5. *Let* $\mathbf{W}(a,b)$ *be the weight sequence for* $(a,b) \in \mathbb{N}^2$. *Then* $\mathcal{N}(\mathbf{W}(a,b)) = \mathcal{N}(a,b)$.

Proof. Without loss of generality, we may suppose that a, b are relatively prime and that $a \le b$. If $a = 1$, then $\mathcal{N}(1,b) = \#^b \mathcal{N}(1,1)$ by the second statement in Lemma 2.4.

For a general pair (a,b), we argue by induction on the length K of the weight expansion $\mathbf{W}(a,b) = \big(X_0^{\times \ell_0}, \dots, X_K^{\ell_K}\big)$. We saw in (2.3) above that $X_1 = b - \ell_0 a$ and that $\mathbf{W}(X_1, a) = \big(X_1^{\times \ell_1}, \dots, X_K^{\ell_K}\big)$. Therefore, we may assume by induction that $\mathcal{N}\big(\mathbf{W}(X_1,a)\big) = \mathcal{N}(X_1,a)$. Thus

$$\begin{aligned}\mathcal{N}\big(\mathbf{W}(a,b)\big) &= \big(\#^{\ell_0}\mathcal{N}(a,a)\big)\#\mathcal{N}\big(\mathbf{W}(X_1,a)\big)\\ &= \big(\#^{\ell_0}\mathcal{N}(a,a)\big)\#\mathcal{N}(X_1,a)\\ &= \mathcal{N}(a, X_1 + \ell_0 a) = \mathcal{N}(a,b),\end{aligned}$$

where the first equality follows from the definition since $\mathbf{W}(a,b) = \big(a^{\times \ell_0}, \mathbf{W}(X_1,a)\big)$, the second holds by the inductive hypothesis, and the third holds by Lemma 2.4. q.e.d.

The first part of the next lemma was independently observed by David Bauer during Edifest 2010.

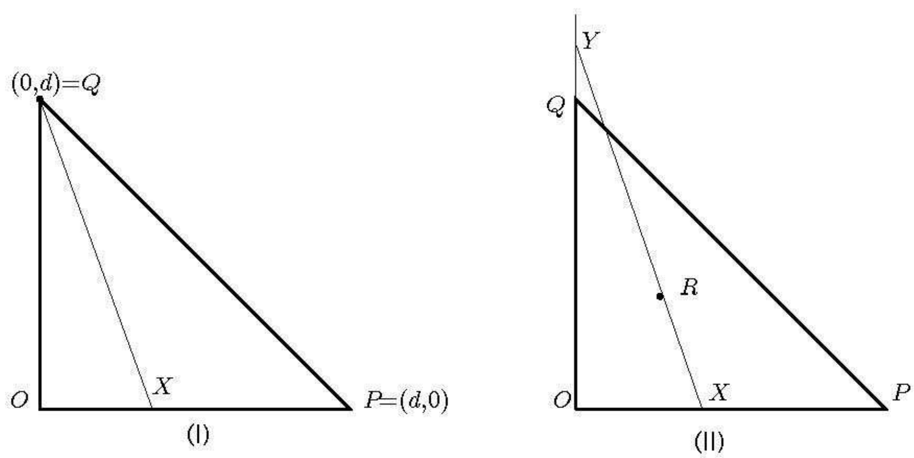

Figure 2.3. Decompositions considered in Lemma 2.6.

Lemma 2.6. (i) *Given integers* $0 < a < b$, *we have*

$$\mathcal{N}(a,b)\#\mathcal{N}(b-a,b) \le \mathcal{N}(b,b).$$

(ii) *For each* $k \ge 1$ *there is* ℓ *such that*

$$\mathcal{N}_{k+\ell}(b,b) = \mathcal{N}_k(a,b) + \mathcal{N}_\ell(b-a,b).$$

Proof. (i) Since $\mathcal{N}_i(b,b)$ is constant and equal to bd on the sets of the form $(d^2+d)/2 \le i \le (d^2+3d)/2$, it suffices to consider the case $i = (d^2+3d)/2$. Then the points counted by $\mathcal{N}_i(b,b)$ are those in the triangle $T = OPQ$ in diagram (I) in Figure 2.3. Cut this triangle by the line QX of slope $-b/a$, and divide the integer points in T into two groups $\mathcal{S}_1, \mathcal{S}_2$ where $\mathcal{S}_1$ consists of Q plus the points to the left of QX and $\mathcal{S}_2$ is the rest. Let $|\mathcal{S}_1| = k_1 + 1$ and $|\mathcal{S}_2| = k_2$ as before. Then, counting the points in triangle $OQX = T^{ad}_{b,a}$, we see that $\mathcal{N}_{k_1} = ad$. Further, if we move the triangle PXQ first by the affine transformation that fixes the x-axis and takes Q horizontally to the point (d,d), and then by the reflection in the vertical line $x = d$, it is a horizontal translate of $T^{d(b-a)}_{b,b-a}$. Hence, counting the points in the triangle PQX, we see that $\mathcal{N}_{k_2} = (b-a)d$. Thus $\mathcal{N}_{k_1} + \mathcal{N}_{k_2} = bd$. This proves (i).

To prove (ii), suppose $\mathcal{N}_k(a,b) = B$ and let R be an integer point on the slant edge of the corresponding triangle $T^B_{b,a}$. Let Y be the point where the slant edge of this triangle meets the y-axis, so that $Y = (0, B/a)$, and let X be where it meets the x-axis, so that $X = (B/b, 0)$. Next, let $Q = (0,d)$ be the integer point with $d = \lfloor B/a \rfloor$, and put $P = (d,0)$ as before. Then no points in the triangle $T^B_{b,a}$ lie above the line PQ since $|YQ| < 1$. We now divide the points in triangle OPQ into two sets as before, with $\mathcal{S}_1$ the union of R with all points to the left of XY and $\mathcal{S}_2$ the rest. Then $|\mathcal{S}_1| = k+1$ by construction. Further, as in (i), the points in $\mathcal{S}_2$ lie in a triangle that is affine equivalent to $T^A_{b,b-a}$, where $A/b = |PX| = d - B/b$. Moreover, $\mathcal{S}_2$ contains P (which corresponds to the origin) but not R, which is a point on the slant edge. Thus $T^A_{b,b-a}$ has $|\mathcal{S}_2| + 1$ integer points, so that if $\ell := |\mathcal{S}_2|$, we have $\mathcal{N}_\ell(b, b-a) = A = db - B$. Therefore, $\mathcal{N}_k(a,b) + \mathcal{N}_\ell(b,b-a) = db$. But $k + \ell = (d^2+3d)/2$ by construction, so that $\mathcal{N}_{k+\ell}(b,b) = db$. The result follows. q.e.d.

Corollary 2.7. *Let $\mathcal{C}$ be a nondecreasing sequence of nonnegative numbers such that $\mathcal{C} \# \mathcal{N}(d-c,d) \le \mathcal{N}(d,d)$. Then $\mathcal{C} \le \mathcal{N}(c,d)$.*

Proof. If not, there is k such that $\mathcal{C}_k > \mathcal{N}_k(c,d)$. But by the previous lemma, there is ℓ such that $\mathcal{N}_{k+\ell}(d,d) = \mathcal{N}_k(c,d) + \mathcal{N}_\ell(d-c,d)$. Then $\mathcal{C}_k + \mathcal{N}_\ell(d-c,d) > \mathcal{N}_{k+\ell}(d,d)$, contradicting the hypothesis. q.e.d.

3. Proof of Theorem 1.1

Our argument is based on the following key results.

Proposition 3.1. [**12**, thm. 3.11] *Let c,d,e,f be any positive integers and $\lambda > 0$. Then there is a symplectic embedding*

$$\Phi_E : \ \operatorname{int} \lambda E(e,f) \overset{s}{\hookrightarrow} \operatorname{int} E(c,d)$$

if and only if there is a symplectic embedding

$$\Phi_B : \ \operatorname{int} \lambda B\big(\mathbf{W}(e,f)\big) \ \sqcup \ \operatorname{int} B\big(\mathbf{W}(d-c,d)\big) \ \stackrel{s}{\hookrightarrow} \ B(d).$$

The proof that the numerical condition is sufficient for an embedding to exist involves a significant use of Taubes–Seiberg–Witten theory in conjunction with the theory of J-holomorphic curves. However, it is easy to see why it is necessary, since the ellipsoid $E(p,q)$ decomposes into a union of balls whose sizes are given by the weights $\mathbf{W}(p,q)$; cf. Figure 3.1. To see this, recall that the moment (or toric) image of the ball is affine equivalent to a standard triangle (a right-angled isosceles triangle), while that of an ellipsoid is an arbitrary right-angled triangle. As the diagram shows, the decomposition of a rectangle into squares given by the weights (as in Figure 2.1) yields a corresponding decomposition of the rectangle into (affine) standard triangles; see [**11**, §2] for more detail.

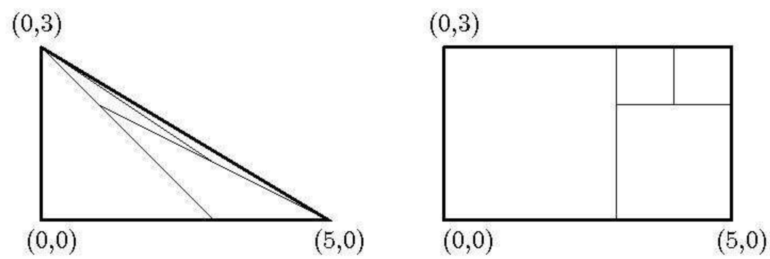

Figure 3.1. Cutting a triangle into standard triangles.

The next result follows from Hutchings' observation in [**6**, remark 1.10] that the ECH capacities $\mathcal{N}(a,b)$ give a sharp obstruction for ball embeddings. We give the proof to make it clear that it does not use any knowledge of ECH, although it does use easier gauge theoretic results.

Proposition 3.2. [**6**] *Let* $\mathbf{a} := (a_1, \dots, a_M)$ *be any sequence of positive numbers. Then there is a symplectic embedding* $\Phi_B : \sqcup_i \operatorname{int} B(a_i) \stackrel{s}{\hookrightarrow} B(\mu)$ *exactly if* $\mathcal{N}(\mathbf{a}) := \#_i \mathcal{N}(a_i, a_i) \preccurlyeq \mathcal{N}(\mu,\mu)$.

Proof. Hutchings showed in [**6**, proposition 1.9] that $\mathcal{N}(\mathbf{a}) \preccurlyeq \mathcal{N}(\mu,\mu)$ exactly if

$$\sum_i m_i a_i \le \mu d, \quad \text{whenever } \sum_i m_i^2 + m_i \le d^2 + 3d, \tag{3.1}$$

where $(d; m_1, \dots, m_M) =: (d; \mathbf{m})$ is any sequence of nonnegative integers. This holds because the kth entry in the sequence $\mathcal{N}(\mathbf{a}) = \#_{i=1}^M \mathcal{N}(a_i, a_i)$ is the maximum of the numbers

$$\sum_i \mathcal{N}_{k_i}(a_i, a_i) = \sum_i m_i a_i,$$

where $\sum_i k_i = k$ and $\frac{1}{2}(m_i^2 + m_i) \le k_i \le \frac{1}{2}(m_i^2 + 3m_i)$.

On the other hand, it was shown by McDuff–Polterovich [**13**] that Φ_B exists exactly if for each $\lambda \in (0,1)$ there is a symplectic representative of the cohomology class $\alpha_\lambda := \mu\ell - \lambda^2 \sum_i a_i e_i$ on the M-fold blowup $X_M := \mathbb{C}P^2 \# M\overline{\mathbb{C}P}^2$ that is deformation equivalent to the small blowup of a form on $\mathbb{C}P^2$ and hence has standard first Chern class. (Here ℓ and $e_i, i = 1, \dots, M$ denote the Poincaré duals of the classes of the line L and the exceptional divisors E_i.) Thus we need α_λ to lie in the symplectic cone $\mathcal{C}_K$ of X_M given by the classes of all symplectic forms with first Chern class Poincaré dual to $-K := 3L - \sum_i E_i$.

After preliminary work by McDuff [**10**] and Biran [**2**] concerning the closure of $\mathcal{C}_K$, Li–Liu in [**9**, theorem 3] used Taubes–Seiberg–Witten theory to describe $\mathcal{C}_K$ in the following terms. Let $\mathcal{E}_K \subset H_2(X_M;\mathbb{Z})$ be the set of classes E with $-K\cdot E = 1$ that can be represented by smoothly embedded spheres of self-intersection -1. Then Li–Liu showed that $\mathcal{C}_K$ is connected and has the following description:

$$\mathcal{C}_K = \{\alpha \in H^2(X_M;\mathbb{R}) : \alpha^2 > 0, \langle \alpha, E\rangle > 0 \; \forall E \in \mathcal{E}_K\}.$$

Notice that if $E = dL - \sum m_i E_i \in \mathcal{E}_K$, then $\sum m_i^2 = d^2 + 1$ while $\sum m_i = 3d - 1$. Therefore, each such E does give rise to an inequality of the type considered in (3.1).[3] It is also easy to check that the inequalities in (3.1) for $d \to \infty$ imply that $\sum a_i^2 \le \mu$, which corresponds to the volume condition $\alpha^2 > 0$; cf. [**6**, remark 3.13]. However, because many tuples $(d;\mathbf{m})$ with $\sum_i m_i^2 + m_i \le d^2 + 3d$ do not correspond to elements in $\mathcal{E}_K$, it seems on the face of it that the conditions in (3.1) are more stringent than the geometric condition $\alpha \in \mathcal{C}_K$. Lemma 3.3 below gives a purely algebraic argument showing that this is not the case. q.e.d.

Let $\mathcal{F}$ be the set of all tuples $(d;\mathbf{m}) := (d; m_1, \dots, m_M)$ such that $\sum_i (m_i^2 + m_i) \le d^2 + 3d$. We denote by $\mathcal{F}^+$ those elements $(d;\mathbf{m}) \in \mathcal{F}$ with $d > 0$ and $m_i \ge 0$ for all i. Further, we define $(\mu;\mathbf{a})\cdot(d;\mathbf{m}) := d\mu - \sum a_i m_i =: d\mu - \mathbf{a}\cdot\mathbf{m}$. Then, if $-K = (3;1,\dots,1)$, we have

$$(d;\mathbf{m}) \in \mathcal{F} \iff (d;\mathbf{m})\cdot\big((d;\mathbf{m}) - K\big) \ge 0.$$

Similarly, we identify the cone $\mathcal{C}_K \subset H^2(X_M;\mathbb{R})$ with the set of tuples $(\mu;\mathbf{a})$ given by the coefficients of the classes $\alpha = \mu\ell - \sum a_i e_i \in \mathcal{C}_K$. In this notation, it suffices to prove the following lemma.

Lemma 3.3. *Let $(\mu;\mathbf{a}) := (\mu; a_1, \dots, a_M)$ be a tuple of nonnegative numbers such that*

(i) $(\mu;\mathbf{a})\cdot(d;\mathbf{m}) \ge 0$ *for all* $(d;\mathbf{m}) \in \mathcal{E}_K$.

(ii) $\|\mathbf{a}\| := \sqrt{\sum_i a_i^2} \le \mu$.

Then $(\mu;\mathbf{a})\cdot(d;\mathbf{m}) \ge 0$ *for all* $(d;\mathbf{m}) \in \mathcal{F}^+$.

[3] The cone $\mathcal{C}_K$ is described by strict inequalities, but when we let $\lambda \to 1$, these correspond to the $\le$ signs in (3.1).

The proof is based on the following elementary result that is part of Li–Li [**8**, lemma 3.4]. We say that a tuple $(\mu;\mathbf{a})$ is **positive** if $\mu \geq 0$ and $a_i \geq 0$ for all i; that it is **ordered** if $a_1 \geq \cdots \geq a_M$; and that a positive, ordered $(\mu;\mathbf{a})$ is **reduced** if $\mu \geq a_1 + a_2 + a_3$.

Lemma 3.4. *Let $(\mu;\mathbf{a})$ be reduced and $(d;\mathbf{m})$ be a positive tuple such that $-K \cdot (d;\mathbf{m}) = 3d - \sum m_i \geq 0$. Then $(\mu;\mathbf{a}) \cdot (d;\mathbf{m}) \geq 0$.*

Proof. Because $\sum_i m_i \leq 3d$, we may partition the list

$$(1^{\times m_1}, 2^{\times m_2}, \ldots, M^{\times m_M})$$

(considered with multiplicities) into sets $I_n, 1 \leq n \leq d$, where each $I_n = \{j_{n1}, j_{n2}, j_{n3}\}$ is a set of at most three distinct numbers chosen so that each element $j \in \{1, \ldots, M\}$ occurs in precisely m_j different sets I_n. Then

$$\sum_{i=1}^{M} a_i m_i = \sum_{n=1}^{d} \Big(\sum_{i \in I_n} a_i\Big).$$

Further, $\sum_{i \in I_n} a_i \leq \mu$ for all I_n because $(\mu;\mathbf{a})$ is reduced. Hence

$$\sum_{i=1}^{M} a_i m_i \leq \sum_{n=1}^{d} \mu = d\mu,$$

as required. q.e.d.

Proof of Lemma 3.3. In the argument below we assume $M \geq 3$. Since we allow the a_i to be 0, we can always reduce to this case by increasing M if necessary. Next observe that it suffices to prove the result for integral tuples $(\mu;\mathbf{a})$ and $(d;\mathbf{m})$. We suppose throughout that $(\mu;\mathbf{a})$ satisfies conditions (i) and (ii). Further, if $(d;\mathbf{m})$ is such that $\|\mathbf{m}\| := \sqrt{\sum m_i^2} \leq d$, then $(d;\mathbf{m}) \cdot (\mu;\mathbf{a}) \geq d\mu - \|\mathbf{m}\|\,\|\mathbf{a}\| \geq 0$ as required. Therefore, we only need consider $(d;\mathbf{m})$ with $(d;\mathbf{m}) \cdot (d;\mathbf{m}) < 0$.

Following Li–Li [**8**], consider the Cremona transformation Cr that acts on tuples by $Cr(d;\mathbf{m}) = (d';\mathbf{m}')$, where $m'_j = m_j$ for $j \geq 4$ and

$$d' = 2d - (m_1 + m_2 + m_3), \quad m'_i = d - (m_j + m_k) \text{ for } \{i,j,k\} = \{1,2,3\}.$$

Then Cr preserves the class K and the intersection product, and hence preserves $\mathcal{F}$. Because Cr, when considered as acting on $H_2(X_M)$, is induced by a diffeomorphism (the reflection in the sphere in class $L - E_1 - E_2 - E_3$), it preserves the set of classes represented by embedded spheres and hence preserves $\mathcal{E}_K$ and $\mathcal{C}_K$.

Now suppose that $(\mu;\mathbf{a}) \in \mathcal{C}_K$, and denote by $\mathrm{Orb}\,(\mu;\mathbf{a})$ its orbit under permutations and Cremona transformations. Since $\mathrm{Orb}\,(\mu;\mathbf{a}) \subset \mathcal{C}_K$, all elements in $\mathrm{Orb}\,(\mu;\mathbf{a})$ are positive. Moreover, if $(\mu;\mathbf{a})$ is ordered, a Cremona move decreases μ unless $(\mu;\mathbf{a})$ is also reduced. Hence, we can transform an ordered $(\mu;\mathbf{a})$ to a reduced element $(\mu';\mathbf{a}') := C(\mu;\mathbf{a})$

by a sequence of k moves $C_1, \dots, C_k$, each consisting of Cr followed by a reordering. Thus $C := C_k \circ \cdots \circ C_1$. Take any $(d; \mathbf{m}) \in \mathcal{F}^+$ with $(d; \mathbf{m}) \cdot (d; \mathbf{m}) < 0$ and denote by $(d'; \mathbf{m}') := C(d; \mathbf{m})$ its image under these moves. Then we must check that $(d'; \mathbf{m}') \cdot (\mu'; \mathbf{a}') \geq 0$.

There are three cases to consider.

Case (i) $(d'; \mathbf{m}')$ *is positive.*

Since $(d'; \mathbf{m}') \cdot (d'; \mathbf{m}') = (d; \mathbf{m}) \cdot (d; \mathbf{m}) < 0$ and $(d'; \mathbf{m}') \in \mathcal{F}$, we must have $-K \cdot (d'; \mathbf{m}') > 0$. Hence, the result follows from Lemma 3.4.

Case (ii) $d' > 0$ *but some* $m_i' < 0$.

In this case, let $(d', \mathbf{m}'')$ be the positive tuple obtained from $(d'; \mathbf{m}')$ by replacing the negative terms m_i' by 0. Because $m^2 + m \geq 0$ for all m, we still have $(d'; \mathbf{m}'') \in \mathcal{F}^+$. Further,

$$(d'; \mathbf{m}') \cdot (\mu'; \mathbf{a}') \geq (d'; \mathbf{m}'') \cdot (\mu'; \mathbf{a}').$$

Therefore, it suffices to show that $(d'; \mathbf{m}'') \cdot (\mu'; \mathbf{a}') \geq 0$. If $(d'; \mathbf{m}'') \cdot (d'; \mathbf{m}'') \geq 0$, then this holds by the argument in the first paragraph of this proof. Otherwise, $(d'; \mathbf{m}'') \cdot (d'; \mathbf{m}'') < 0$, and it holds as in case (i) above.

Case (iii) $d' < 0$.

In this case, we show that $(d'; \mathbf{m}') \cdot (\mu'; \mathbf{a}') \geq 0$ by induction on k, the length of the reducing sequence for $(\mu; \mathbf{a})$. Consider the sequence $(d_\ell; \mathbf{m}_\ell) := C_\ell \circ \cdots \circ C_1(d; \mathbf{m})$ of elements of $\mathcal{F}$ obtained by applying the moves $C_i, i = 1, \dots, k$, to $(d; \mathbf{m})$, and let $(\mu_\ell; \mathbf{a}_\ell)$ be the corresponding elements of $\mathcal{C}_K$. Consider the smallest ℓ for which $(d_\ell; \mathbf{m}_\ell)$ is negative.

Suppose first that $d_\ell < 0$. Then the entries $s_i := m_{\ell-1,i}$ of the previous term $\mathbf{m}_{\ell-1}$ are nonnegative, while if $t := d_{\ell-1}$, we have

$$0 < 2t < s_1 + s_2 + s_3, \quad \sum_{i=1}^{3} s_i^2 + s_i \leq t^2 + 3t.$$

Therefore, if $\sum_{i=1}^3 s_i^2 = \lambda t^2$ with $\lambda > 1$, we have

$$2t < \sum_{i=1}^{3} s_i \leq 3t - (\lambda - 1)t^2,$$

so that $\lambda < 1 + 1/t$. Thus in all cases $\sum_{i=1}^3 s_i^2 \leq t(1+t)$. But the minimum of the expression $x^2 + y^2 + z^2$ subject to the constraints $x, y, z \geq 0, x + y + z = 2$ is assumed when $x = y = z$ and is $\frac{4}{3}$. Therefore $t(1+t) \geq \frac{4}{3}t^2$, which gives $d_{\ell-1} = t \leq 3$. But there is no integral solution for $(d_{\ell-1}; \mathbf{m}_{\ell-1})$ with such a low value for $d_{\ell-1}$. Thus this case does not occur.

Hence the first negative element $(d_\ell; \mathbf{m}_\ell)$ must have $d_\ell > 0$ and some negative entry in $\mathbf{m}_\ell$. But then define $\mathbf{m}_\ell''$ as in Case (ii) by replacing all negative entries in $\mathbf{m}_\ell$ by 0. We saw there that it suffices to show that $(d_\ell; \mathbf{m}''_\ell) \cdot (\mu_\ell; \mathbf{a}_\ell) \geq 0$. If $(d_\ell; \mathbf{m}''_\ell) \cdot (d_\ell; \mathbf{m}''_\ell) \geq 0$ this is automatically true. Otherwise, since $\ell \geq 1$, it holds by the inductive hypothesis.

This completes the proof of Lemma 3.3. □

We are now ready to prove the main result. We denote by $\lambda E(a,b)$ the ellipsoid $\{\lambda x : x \in E(a,b)\}$. Thus $\lambda E(a,b) = E(\lambda^2 a, \lambda^2 b)$ has corresponding sequence $\lambda^2 \mathcal{N}(a,b)$.

Proof of Theorem 1.1. By standard continuity properties as explained in [**12**, cor. 1,5], it suffices to prove this when a,b,c,d are rational. Therefore, we will suppose that $c \leq d$ are mutually prime integers and that $(a,b) = \lambda^2(e,f)$ where $e \leq f$ are also mutually prime integers. We need to show that there is an embedding $\Phi_E : \operatorname{int} \lambda E(e,f) \overset{s}{\hookrightarrow} E(c,d)$ exactly if $\lambda^2 \mathcal{N}(e,f) \preccurlyeq \mathcal{N}(c,d)$.

By Proposition 3.1, it suffices to consider the corresponding ball embedding Φ_B, and by Proposition 3.2, this exists exactly if

$$\mathcal{N}\Big(\lambda B\big(\mathbf{W}(e,f)\big) \sqcup B\big(\mathbf{W}(d,d-c)\big)\Big) \preccurlyeq \mathcal{N}(B(d)) = \mathcal{N}(d,d).$$

Since $\mathcal{N}\big(\lambda B\big(\mathbf{W}(e,f)\big) \sqcup B\big(\mathbf{W}(d,d-c)\big)\big) = \lambda^2 \mathcal{N}(e,f) \# \mathcal{N}(d,d-c)$, this condition is equivalent to

$$(*) \qquad \lambda^2 \mathcal{N}(e,f) \# \mathcal{N}(d,d-c) \preccurlyeq \mathcal{N}(d,d).$$

Thus the theorem will hold if we show that $(*)$ is equivalent to the condition $\lambda^2 \mathcal{N}(e,f) \preccurlyeq \mathcal{N}(c,d)$.

But if $\lambda^2 \mathcal{N}(e,f) \preccurlyeq \mathcal{N}(c,d)$, then we have

$$\lambda^2 \mathcal{N}(e,f) \# \mathcal{N}(d,d-c) \preccurlyeq \mathcal{N}(c,d) \# \mathcal{N}(d,d-c) \preccurlyeq \mathcal{N}(d,d)$$

by Lemma 2.3 (ii) and Lemma 2.6 (i). Conversely, suppose that $(*)$ holds. Then $\lambda^2 \mathcal{N}(e,f) \preccurlyeq \mathcal{N}(c,d)$ by Corollary 2.7. Hence the two conditions are equivalent. □

Proposition 3.5. *The disjoint union of open ellipsoids* $\sqcup_{i=1}^n E(a_i,b_i)$ *embeds symplectically in* $E(c,d)$ *exactly if*

$$\mathcal{N}(a_1,b_1) \# \ldots \# \mathcal{N}(a_n,b_n) \preccurlyeq \mathcal{N}(c,d).$$

Proof. Though this case is not considered in [**12**], the proof of Proposition 3.1 works just as well when the domain is a disjoint union of ellipsoids. Hence, if $c \leq d$, the necessary and sufficient condition for this embedding of unions of ellipsoids to exist is that

$$\mathcal{N}\big(\sqcup_i \mathbf{W}(a_i,b_i)\big) \# \mathcal{N}\big(\mathbf{W}(d-c,d)\big) \preccurlyeq \mathcal{N}(d,d).$$

The proof of Corollary 2.5 adapts to show that

$$\mathcal{N}\big(\sqcup_i \mathbf{W}(a_i,b_i)\big) = \mathcal{N}(a_1,b_1) \# \ldots \# \mathcal{N}(a_n,b_n) =: \mathcal{C}.$$

Now use Lemma 2.6 and Corollary 2.7 as before. q.e.d.

References

[1] D. Bauer, *The generating function of the embedding capacity for* 4-*dimensional symplectic ellipsoids*, arXiv:1102.5630.

[2] P. Biran, *Symplectic packing in dimension* 4, Geometric and Functional Analysis, **7** (1997), 420–37, MR 1466333 Zbl 0892.53022.

[3] K. Cieliebak, H. Hofer, J. Latschev & F. Schlenk, *Quantitative symplectic geometry, Dynamics, ergodic theory, and geometry*, 1–44, Math. Sci. Res. Inst. Publ. **54**, Cambridge Univ. Press, Cambridge, 2007, MR 2369441, Zbl 1143.53341.

[4] L. Guth, *Symplectic embeddings of polydisks*, Invent. Math. **172** (2008), 477–489. MR2393077, Zbl 1153.53060.

[5] R. Hind & E. Kerman, *New obstructions to symplectic embeddings*, arxiv:0906:4296.

[6] M. Hutchings, *Quantitative embedded contact homology*, arXiv:1005.2260, to appear in J. Differential Geometry.

[7] M. Hutchings, *Recent progress on symplectic embedding problems in four dimensions*, to appear in PNAS.

[8] Bang-He Li & T.-J. Li, *Symplectic genus, minimal genus and diffeomorphisms*, Asian J. Math. **6** (2002), 123–44, MR 1902650, Zbl 1008.57024.

[9] Tian-Jun Li & A. K. Liu, *Uniqueness of symplectic canonical class, surface cone and symplectic cone of* 4-*manifolds with* $b^+ = 1$, J. Diff. Geom. **58** (2001), 331–70, MR 1913946, Zbl 1051.57035.

[10] D. McDuff, *From symplectic deformation to isotopy*, Topics in Symplectic 4-manifolds (Irvine CA 1996), ed. Stern, Internat. Press, Cambridge, MA (1998), pp 85–99, MR 1635697, Zbl 0928.57018.

[11] D. McDuff, *Symplectic embeddings and continued fractions: a survey*, arxiv:0908.4387, Journ. Jap. Math. Soc. **4** (2009), 121–139. MR2576029, Zbl pre05675980.

[12] D. McDuff, *Symplectic embeddings of* 4-*dimensional ellipsoids*, J. Topol. **2** (2009), 1–22, MR 2499436, Zbl 1166.53051.

[13] D. McDuff & L. Polterovich, *Symplectic packings and algebraic geometry*, Inventiones Mathematicae **115**, (1994) 405–29, MR 1262938, Zbl 0833.53028.

[14] D. McDuff & F. Schlenk, *The embedding capacity of* 4-*dimensional symplectic ellipsoids*, arXiv:0912.0532 v2.

[15] Dorothee Müller, *Symplectic embeddings of ellipsoids into polydiscs*, PhD thesis, Université de Neuchâtel, in preparation.

[16] E. Opshtein, *Maximal symplectic packings in* $\mathbb{P}^2$, arxiv:0610677, Compos. Math. **143** (2007), 1558–1575. MR2371382, Zbl 1133.53057

[17] E. Opshtein, *Singular polarizations and ellipsoid packings*, arXiv:1011.6358.

[18] L. Traynor, *Symplectic packing constructions*, J. Diff. Geom. **42** (1995), 411–29, MR 1366550, Zbl 0861.52008.

Department of Mathematics
Barnard College
Columbia University
New York, NY 10027-6598, USA
E-mail address: dmcduff@barnard.edu