Math. Z. 200, 355–378 (1989)

Mathematische Zeitschrift

Symplectic Topology and Hamiltonian Dynamics

I. Ekeland[1]* and H. Hofer[2]*,**

[1] Ceremade, Université de Paris-Dauphine, Place du M. de Lattre de Tassigny, F-75575 Paris Cedex 16, France
[2] Department of Mathematics, Rutgers University, New Brunswick, New Jersey, USA

Contents

I. Symplectic Capacities 355
II. Construction of a Symplectic Capacity 358
III. Embedding Problems 371
IV. Rigidity Problems 372

I. Symplectic Capacities

Denote by (V, ω) a symplectic vectorspace, i.e. ω is a skewsymmetric and non-degenerate bilinear form. We call a linear map $T \in L(V)$ symplectic if it preserves ω, that is

$$\omega(x, y) = \omega(Tx, Ty)$$

for every x, $y \in V$. The set of all linear symplectic maps will be denoted by $\mathrm{Sp}(V)$.

We then have the notion of symplectic for a smooth nonlinear map simply by looking at the derivative. A map $f: U \to V$, where $U \subset V$ is open, is called symplectic if $f'(x) \in \mathrm{Sp}(V)$ for every $x \in U$. With other words f is symplectic if it preserves ω,

$$f^* \omega = \omega.$$

In the following we denote by $D(V)$ the group of symplectic diffeomorphisms of V. By $D_c(V)$ we denote the subgroup of those having compact support, i.e. $f(x) = x$ outside a compact set.

We are interested in studying nonlinear symplectic maps. In the course of this paper we provide answers to such questions as: given two sets S and T in V does there exist a symplectic diffeomorphism of V mapping S into T (mapping problem)?

---

* The results were obtained while the first author held a Distinguished Visiting Professorship at Rutgers and the second author was a member of the Institute for Advanced Study

** Research partially supported by NSF Grant DMS-8603149, the Alfred Sloan Foundation, and a Rutgers Trustees Research Fellowship grant

What is the uniform limit of a sequence of symplectic maps (rigidity problem)? Is there a notion of symplecticity for maps which are only C^0?

Since $f^*\omega = \omega$ implies $f^*\omega^n = \omega^n$ an injective symplectic map preserves the measure associated to ω^n. So the measure is an interesting symplectic invariant. The following discussion however shows that better invariants are needed.

Consider a connected compact smooth hypersurface Δ in V. We define the so-called characteristic distribution $\mathscr{L}_\Delta \to \Delta$ by

$$\mathscr{L}_\Delta = (T\Delta)^\omega = \{(x,\xi) \in T\Delta \mid \xi \perp_\omega T_x \Delta\}.$$

Here $\xi \perp_\omega T_x\Delta$ means ξ is ω-orthogonal to $T_x\Delta$.

Clearly $\mathscr{L}_\Delta \subset T\Delta$ is a one dimensional and therefore integrable distribution and in fact orientable since $V\setminus\Delta$ has exactly two components by Alexander duality. We denote the bounded component by B_Δ. Let $H\colon V \to \mathbb{R}$ be a smooth map having Δ as a regular energy surface so that $\Delta = H^{-1}(1)$, $\nabla H(x) \neq 0$ on $x \in \Delta$, and $\inf H(B_\Delta) < \sup H(V\setminus B_\Delta)$. Then the Hamiltonian vectorfield X_H defined by

$$dH = \omega(X_H, \cdot)$$

induces a nonzero section of $\mathscr{L}_\Delta \to \Delta$ giving $\mathscr{L}_\Delta$ a preferred orientation. A closed integral curve for $\mathscr{L}_\Delta$ is called a closed characteristic or a periodic Hamiltonian trajectory. Denote by $[\Delta]$ the set of all closed characteristics on Δ. If $P \in [\Delta]$ we have

$$TP = \mathscr{L}_\Delta \mid P$$

giving P an induced orientation. Given $f \in D(V)$ we have the following formulas

$$\text{(1)} \qquad Tf(\mathscr{L}_\Delta) = \mathscr{L}_{f(\Delta)}$$

$$f([\Delta]) = [f(\Delta)]$$

Note that all orientations are preserved, for example $f \mid P\colon P \to f(P)$ is orientation preserving. Now let λ be a 1-form on V such that $d\lambda = \omega$. For $P \in [\Delta]$ we define the action $A(P)$ by

$$A(P) = \int \lambda \mid P$$

Clearly the definition of A does not depend on the choice of λ as long as $d\lambda = \omega$ on V. From (1) we derive the formula

$$\text{(2)} \qquad A(f(P)) = A(P) \text{ for every } P \in [\Delta].$$

Here of course $f \in D(V)$. This gives the following.

Lemma 1. *Denote by U_1, U_2 open bounded subsets of V with smooth boundaries. Denote for $i = 1, 2$ by $[\partial U_i]$ the closed characteristics on the boundary components. Then the following sets are equal*

$$\{A(P) \mid P \in [\partial U_1]\} = \{A(Q) \mid Q \in [\partial U_2]\}$$

provided $f(U_1) = U_2$ for some $f \in D(V)$. □

As an example consider $V = \mathbb{C}^n$ with the symplectic form $\omega = d\lambda$, where $\lambda = \sum_{k=1}^{n} p_k\, dq_k$ and $z = q + ip, q, p \in \mathbb{R}^n$. Given $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $0 < \alpha_1 \leqq$

$\alpha_2 \leqq \ldots \leqq \alpha_n$ we define an ellipsoid $E(\alpha)$ by

$$E(\alpha) = \{z \in \mathbb{C}^n \mid \Sigma \alpha_k |z_k|^2 < 1\}.$$

Assuming that the numbers α_k^{-1} are linearly independent over $\mathbb{Z}$ we see that

$$[\partial E(\alpha)] = \{P_1, \ldots, P_n\}$$

and $P_k = \{z \in \partial E(\alpha) \mid z_j = 0 \text{ for } j \neq k\}$. Moreover

$$A(P_k) = \frac{\pi}{\alpha_k}.$$

If now $f(E(\alpha)) = E(\beta)$ for some $f \in D(V)$, where again $0 < \beta_1 \leqq \beta_2 \leqq \ldots \leqq \beta_n$, then the numbers β_k^{-1} are independent over $\mathbb{Z}$ because otherwise as one easily verifies $\{\partial E(\beta)]$ is an infinite set. By the previous discussion applying Lemma 1 we see that $\alpha = \beta$. Note that the conservation of volume would only give a condition like

$$\alpha_1 \alpha_2 \ldots \alpha_n = \beta_1 \beta_2 \ldots \beta_n$$

which is of course much weaker. So better invariants are needed. Symplectic capacities are such invariants.

Denote by 2^V the power set of V, and by $B^{2n}(r)$ the euclidean ball of radius r in $\mathbb{C}^n$, i.e.

$$B^{2n}(r) = \{z \in \mathbb{C}^n \mid |z| < r\}.$$

Given symplectic vectorspaces (V, ω_v) and (W, ω_w) the product is defined by

$$(V, \omega_v) \times (W, \omega_w) = (V \times W, \omega_v \oplus \omega_w).$$

Definition 1. A symplectic capacity is a map c which associates to a subset S of a symplectic vectorspace V a number $c(S) = c_V(S)$ so that the following axioms hold:

(A1) (Normalization) $c_{\mathbb{C}^n}(B^{2n}(1)) = c_{\mathbb{C}^n}(B^2(1) \times \mathbb{C}^{n-1}) = \pi$

(A2) (Monotonicity) $S \subset T \subset (V, \omega) \Rightarrow c_V(S) \leqq c_V(T)$

(A3) (Conformality) If $\sqrt{\alpha} f \in D(V, W)$ for some $\alpha > 0$ then $\alpha c(f(S)) = c(S)$ for all $S \subset V$.

Here $D(V, W)$ is the set of all symplectic diffeomorphisms $f\colon V \to W$.

Let us point out immediately that it is by no means clear that symplectic capacities exist. Gromov, [12], was the first one to construct a symplectic capacity. He calls it *symplectic width* and defines it for every symplectic manifold (V, ω) as the lower bound of the numbers $a > 0$ such that for every almost complex structure J on V tamed by ω, and for every $x \in V$, there exists a nonconstant properly mapped J-holomorphic curve $f\colon S \to V$ (where S is a Riemannian surface) passing through V, such that the symplectic area of f satisfies

$$\int_S f^* \omega \leqq a.$$

In the next section we will give an alternate construction in symplectic vectorspaces V utilizing Hamiltonian systems. This allows to define a capacity for any subset S of V. In a forthcoming paper our approach also allows to define higher order capacities, which provide us with a tool to prove some optimal embedding theorems, for example for Lagrangian Tori into a ball.

Acknowledgement. We would like to thank Professor M. Gromov for very stimulating discussions.

II. Construction of a Symplectic Capacity

Drawing upon Lemma 1 and the subsequent example, we might be tempted to construct a symplectic capacity in the following way. Given an open bounded connected subset U of V with smooth boundary ∂U, pick some $P_U \in [\partial U]$ and set $c(U) := A(P_u)$.

This fails dismally. First of all, there are no results which guarantee the existence of closed characteristics under smoothness and compactness assumption alone. Finally, even if $[\partial U]$ would be nonempty, where should the monotonicity come from?

The approach will be salvaged by exploiting an idea of Hofer and Zehnder, [17]. Loosely speaking they associate with U a variational problem and construct a critical value thereof, which we call $c(U)$. It turns out that $c(U)$ has the desired properties, and that in the cases when $[\partial U]$ is known to be nonempty, there exists always some $P \in [\partial U]$ such that $c(U) = A(P_U)$ or perhaps $kA(P_U)$ for some $k \in \mathbb{N}$, $k \geqq 1$. Let us give some precise statements. We need the following.

Definition 2. A compact smooth connected hypersurface Δ in V is said to be of restricted contact type if there exists a 1-form λ on V such that $d\lambda = \omega$ and $\lambda(x, \xi) \neq 0$ for every nonzero vector in $\mathscr{L}_\Delta$. We denote the collection of all those hypersurfaces by R.

Note that $\lambda(x, \xi) > 0$ for every nonzero vector in $\mathscr{L}_\Delta$ which is positive for the orientation. It has been proved by C. Viterbo that $[\Delta] \neq \phi$ for $\Delta \in R$; see [17] and [21]. Our main result in this section is the following.

Theorem 1. *There exists a symplectic capacity c such that*

$$c(\Delta) = c(B_\Delta) = k_\Delta A(P_\Delta) \tag{1}$$

for every $\Delta \in R$, where B_Δ is the bounded component of $V \setminus \Delta$. Here $k_\Delta \geqq 1$ is a suitable integer and P_Δ a suitable element in $[\Delta]$.

The proof follows in the first part the pattern in [17].

We begin by introducing the Hilbert space E consisting of all functions $x \in L^2(0, 1; \mathbb{C}^n)$ whose Fourier series

$$x(t) = \sum_{k \in \mathbb{Z}} x_k e^{2\pi i k t}, \qquad x_k \in \mathbb{C}^n$$

satisfies

$$\Sigma |k| |x_k|^2 < \infty .$$

The inner product in E is defined by

$$(x,y) = \langle x_0, y_0 \rangle + 2\pi \Sigma |k| \langle x_k, y_k \rangle.$$

Here $\langle \xi, \eta \rangle = \mathrm{Re}\left(\sum_{j=1}^{n} \xi_j \bar{\eta}_j \right)$ and we put as usual

$$\|x\|^2 = (x,x)$$

We denote by P^-, P^0, P^+ the orthogonal projections on the subspaces

$$\begin{aligned} E^- &:= \{x \mid x_k = 0 \quad \text{for } k \geqq 0\} \\ E^0 &:= \{x \mid x_k = 0 \quad \text{for } k \neq 0\} = \mathbb{C}^n \\ E^+ &:= \{x \mid x_k = 0 \quad \text{for } k \leqq 0\}. \end{aligned}$$

For a smooth loop $x\colon R/\mathbb{Z} \to \mathbb{C}^n$ we define by

$$a(x) = \frac{1}{2} \int_0^1 \langle -i\dot{x}, x \rangle \, dt$$

the action of x. a extends to a continuous quadratic form on E which we denote again by a. Then we have

(2) $$a(x) = -\tfrac{1}{2} \| P^- x \|^2 + \tfrac{1}{2} \| P^+ x \|^2.$$

Next we introduce a subgroup Γ of the group of all homeomorphisms of E.

Definition 3. A homeomorphism $h\colon E \to E$ belongs to Γ iff it can be written in the form

$$\begin{aligned} h(x) &= e^{\gamma^+(x)} P^+ x + P^0 x + e^{\gamma^-(x)} P^- x + k(x) \\ &=: u(x)x + k(x) \end{aligned}$$

with $u(x) = e^{\gamma^+(x)} P^+ x + P^0 x + e^{\gamma^-(x)} P^- x$. Here $\gamma^+, \gamma^-\colon E \to \mathbb{R}$ are continuous and map bounded sets into bounded sets, and $k\colon E \to E$ is continuous and maps bounded sets into compact sets. In addition there exists a $\rho > 0$ such that

(3) $\quad a(x) \leqq 0 \quad$ or $\quad \|x\| \geqq \rho$ implies that

$$\gamma^+(x) = \gamma^-(x) = 0 \quad \text{and} \quad k(x) = 0.$$

We leave the easy verification that Γ is a group to the reader.

Now we need an infinite dimensional intersection result. Introduce the unit-sphere in E^+ by

$$S^+ = \{x \in E^+ \mid \|x\| = 1\}$$

and the function $e \in E^+$ defined by

$$e(t) = (e^{2\pi i t}, 0, 0, \ldots, 0) \in \mathbb{C}^n.$$

Proposition 1. *For $h \in \Gamma$ we have*

$$h(S^+) \cap (E^- \oplus E^0 \oplus \mathbb{R}e) \neq \phi.$$

Proof. We shall show that $h(S^+) \cap (E^- \oplus E^0 \oplus [0, +\infty)e) \neq \phi$. Since Γ is a group this is equivalent to

$$S^+ \cap h(E^- \oplus E^0 \oplus [0, +\infty)e) \neq \phi.$$

We embed h in a homotopy h_t, $t \in [0, 1]$, defined by

$$h_t(x) = e^{t\gamma^+(x)} P^+ x + p^0 x + e^{t\gamma^-(x)} P^- x + tk(x)$$

so that $h_0 = \mathrm{Id}$ and $h_1 = h$. We claim

$$S^+ \cap h_t(E^- \oplus E^0 \oplus [0, +\infty)e) \neq \phi \tag{4}$$

for all $t \in [0, 1]$. If $x \in E^- \oplus E^0$ we have $a(x) \leqq 0$ so that $h_t(x) = x$ by condition (3). Hence

$$h_t(E^- \oplus E^0) \cap S^+ = \phi \,\forall\, t \in [0, 1]. \tag{5}$$

Now consider the vectorspace $F = E^- \oplus E^0 \oplus \mathbb{R}e$ and in that space the open cylinder $\Omega = \Omega_b$ defined by

$$\Omega = \{x \in F \mid x = y + \xi e,\ y \in E^- \oplus E^0$$
$$\|y\| < b, 0, < \xi < b\}.$$

Choose $b > \max\{\rho, 10\}$ where ρ is the number associated to h by condition (3). If $x = y + \xi e \in \partial\Omega$ where we take the boundary of Ω in F, then either $\xi = 0$ so that $x \in E^- \oplus E^0$, which gives $h_t(x) \notin S^+$ by (5), or $\xi \neq 0$ in which case $\|x\| \geqq b$ which gives $h_t(x) = x$ by (3), so that $h_t(x) \notin S^+$ again. Hence

$$h_t(\partial\Omega) \cap S^+ = \phi \qquad \text{for } t \in [0, 1]. \tag{6}$$

We can restate problem (4) in the form

$$(P^0 + e^{-t\gamma^-(x)} P^-) h_t(x) = 0$$
$$1 - \|h_t(x)\| = 0$$

to be solved for $x \in \Omega$. This is equivalent to

$$x + T_t(x) = 0, \qquad x \in \Omega$$

with

$$T_t(x) = t(P^0 + e^{-t\gamma^-(x)} P^-) k(x) - (\xi + 1 - \|h_t(x)\|)e$$
$$x = y + \xi e, \quad y \in E^- \oplus E^0.$$

We have just seen that

$$x + T_t(x) \neq 0 \qquad \text{for all } t \in [0, 1] \quad \text{and} \quad x \in \partial\Omega. \tag{7}$$

For $t = 0$, we have

$$x + T_0(x) = x - (\xi + 1 - \|x\|)e$$
$$= y - (1 - \|x\|)e.$$

Since the map $(t, x) \to T_t(x)$ is compact, we can apply Leray-Schauder-Degree theory. Using (7) we see that

$$\deg(I + T_1, \Omega, 0) = \deg(I + T_0, \Omega, 0).$$

On the other hand, $I + T_0$ is homotopic to the translation $I - e$ through maps of the form identity + compact having no zero on $\partial\Omega$ (just take a convex combination of $I - e$ and $I + T_0$). Since $e \in \Omega$ we infer

$$\begin{aligned}\deg(I + T_0, \Omega, 0) &= \deg(I - e, \Omega, 0)\\ &= \deg(I, \Omega, e)\\ &= 1.\end{aligned}$$

This completes the proof. □

Now we set up the variational problems. Denote by $\mathscr{F}$ the family of all smooth maps $H: \mathbb{C}^n \to \mathbb{R}$ such that

(B1) There exists a nonempty open set U of $\mathbb{C}^n$ such that $H = 0$ on U. Moreover $H(z) \geqq 0$ for all $z \in \mathbb{C}^n$.

(B2) There exists $\rho > 0$ and a number $\hat{a} > \pi$ with $\hat{a} \notin \mathbb{N}\pi$ such that $H(z) = \hat{a}|z|^2$ for $|z| \geqq \rho$.

To $H \in \mathscr{F}$ we associate $\Phi_H \in C^\infty(E, \mathbb{R})$ defined by

$$\Phi_H(x) = a(x) - \int_0^1 H(x(t))\,dt. \tag{8}$$

Proposition 2. *Let $H \in \mathscr{F}$ and set*

$$c_H := \sup_{h \in \Gamma} \inf_{x \in S^+} \Phi_H(h(x)). \tag{9}$$

Then

$$0 < c_H < +\infty$$

and c_H is a critical of Φ_H, i.e. *there exists $x_H \in E$ with*

$$\Phi_H'(x_H) = 0 \quad \textit{and} \quad \Phi_H(x_H) = c_H.$$

Proof. Given $\varepsilon > 0$ and $x_0 \in E^0$ we find $h \in \Gamma$ such that

$$h(S^+) = x_0 + \varepsilon S^+.$$

We may assume that H is zero in a neighborhood of x_0. Since $H(x_0) = 0$, $H'(x_0) = 0$ and $H''(x_0) = 0$ we see that

$$\Phi_H(x_0) = 0,\ \Phi_H'(x_0) = 0 \quad \text{and} \quad \Phi_H''(x_0) = -P^- + P^+.$$

Hence, if $\varepsilon > 0$ is small enough

$$\inf \Phi_H(h(S^+)) > 0. \tag{10}$$

From Propostion 1 we know that $h(S^+) \cap (E^- \oplus E^0 \oplus \mathbb{R}e) \neq \phi$ which implies that

$$\inf \Phi_H(h(S^+)) \leqq \sup \Phi_H(E^- \oplus E^0 \oplus \mathbb{R}e).$$

It follows from (B2) that the right-hand side is finite ($\hat{a} > \pi$). Summing up, we have shown that

$$0 < c_H < +\infty.$$

Arguing as in [17] or [10], we see that the Palais-Smale condition holds. Assuming c_H is not a critical level we shall derive a contradiction. Since we have the

Palais-Smale condition we can find $\varepsilon_0 \in (0, \min\{1, c_H\})$ such that

$$|\Phi_H(x) - c_H| \leqq \varepsilon_0 \Rightarrow \|\Phi'_H(x_0)\| \geqq \varepsilon_0. \tag{11}$$

Then we pick $h_0 \in \Gamma$ with

$$c_H \leqq \inf \Phi_H(h_0(S^+)) + \tfrac{1}{2}\varepsilon_0.$$

Now introduce three locally Lipschitz continuous maps $\varphi_1, \varphi_2, \varphi_3: E \to [0,1]$ such that

$$\begin{array}{lll} & \varphi_1(x) = 0 & \text{if } \Phi_H(x) \notin [c_H - \varepsilon_0, c_H + \varepsilon_0] \\ & \varphi_1(x) = 1 & \text{if } \Phi_H(x) \in [c_H - \frac{1}{2}\varepsilon_0 + \frac{1}{2}\varepsilon_0] \\ & \varphi_2(x) = 1 & \text{if } \|\Phi'_H(x)\| \leqq 1 \\ (12) & \varphi_2(x) = \|\Phi'_H(x)\|^{-1} & \text{if } \|\Phi'_H(x)\| > 1 \\ & \varphi_3(x) = 1 & \text{if } \|x\| \leqq (\sup \|h_0(S^+)\|) + \dfrac{2}{\varepsilon_0} \\ & \varphi_3(x) = 0 & \text{if } \|x\| \text{ is large} \end{array}$$

and consider the differential equation

$$\dot{x} = \varphi_1(x)\varphi_2(x)\varphi_3(x)\Phi'_H(x).$$

Solutions are defined globally since the right-hand side is bounded in norm by one. Denote by $\mathbb{R} \times E \to E: (t,x) \to x*t$ the associated flow. If $\|x\|$ is large or $a(x) \leqq 0$ we have $h(x) = x$ where $h(x) = x*\dfrac{1}{\varepsilon_0}$. (In the first case $\varphi_3(x) = 0$; in the second case $\Phi_H(x) \leqq 0 < c_H - \varepsilon_0$ so that $\varphi_1(x) = 0$.) Combining (11) and (12) gives, whenever $|\Phi_H(x) - c_H| \leqq \frac{1}{2}\varepsilon_0$ and $\|x\| \leqq \sup \|h_0(S^+)\|$,

$$\frac{d}{dt}\Phi_H(x*t) \geqq \varepsilon_0^2$$

as long as $|\Phi_H(x*t) - c_H| \leqq \dfrac{1}{2}\varepsilon_0$ and $t \in \left[0, \dfrac{1}{\varepsilon_0}\right]$. Hence

$$\Phi_H\left(x * \frac{1}{\varepsilon_0}\right) \geqq c_H + \frac{1}{2}\varepsilon_0.$$

This implies the inequality

$$\inf \Phi_H(h \circ h_0(S^+)) \geqq c_H + \frac{1}{2}\varepsilon_0.$$

If we can show that $h \in \Gamma$ then we have obtained the contradiction $c_H > c_H$. This would prove that c_H is a critical level. We define a map $d: \mathbb{R} \times E \to \mathbb{R}$ by

$$d(t,x) = \varphi_1(x*t)\,\varphi_2(x*t)\,\varphi_3(x*t).$$

We put $\tau = \frac{1}{\varepsilon_0}$ and define

$$\tilde{\gamma}^+(s,x) = \int_s^\tau d(\sigma,x)\,d\sigma$$

$$\tilde{\gamma}^-(s,x) = \int_s^\tau -d(\sigma,x),d\sigma$$

and

$$k(x) = \int_0^\tau [e^{\tilde{\gamma}^+(s,x)}P^+x + P^0x + e^{\tilde{\gamma}^-(s,x)}P^-]\,d(s,x)h(x*s)\,ds$$

where b is the gradient in E of $x \to \int_0^1 H(x)(t)\,dt$. Finally we define

$$\gamma^+(x) = \tilde{\gamma}^+(0,x)$$

$$\gamma^-(x) = \tilde{\gamma}^-(0,x).$$

This gives the desired decompositions of h. See for example [18] for more details. □

Next we show that c_H is a symplectic invariant.

Proposition 3. *Let $H \in \mathscr{F}$ and let $s \to \Psi_s$ be a smooth homotopy of the identity in $D_c(\mathbb{C}^n)$. Define $H_s \in \mathscr{F}$* by *$H_s = H \circ \Psi_s$. Then the map*

$$s \to c_{H_s}$$

is constant

Proof. The critical levels of $\Phi_s := \Phi_{H_s}$ are independent of s, since $\Phi_{H\circ\varphi}(\varphi^{-1}(x)) = \Phi_H(x)$ for φ exact symplectic and since $\varphi^{-1}(x)$ is a solution for the Hamiltonian vectorfield associated to $H \circ \varphi$ iff x is a solution for the Hamiltonian vectorfield associated to H. Further, there is an estimate of the form

$$|H_s(x) - H_t(x)| \leqq \delta(|t-s|)$$

for all $x \in \mathbb{C}^n$ with $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$. Hence

$$|\Phi_s(x) - \Phi_t(x)| \leqq \delta(|t-s|).$$

From the definition we obtain in obvious notation

$$|c_s - c_t| \leqq \delta(|s-t|).$$

From this estimate it follows that whenever c_H is an isolated critical level of Φ_H the map $s \to c_s$ has to be constant. Since $H(z) = \hat{a}|z|^2$ for $|z| \geqq \rho$ with $\hat{a} \geqq \pi$ and $\hat{a} \notin \mathbb{N}\pi$, there can be no critical point with L^∞-norm exceeding ρ. Using standard genericity results, as in [20], we can find a sequence $(H_n) \subset C^\infty(\mathbb{C}^n, \mathbb{R})$, $H_n \to H$, coinciding with H for $|z| \geqq \rho$ so that $H_n \in \mathscr{F}$ and all positive critical levels are isolated. So $s \to c_{H_n} \circ \Psi_s$ is constant and passing to the limit $n \to \infty$ we obtain the desired result. □

That for example $c_{H_n} \to c_H$ is clear since $|H_n(x) - H(x)| \leqq \varepsilon_n$ for every $x \in \mathbb{C}^n$ where $\varepsilon_n \to 0$. From this it follows immediately that $|c_{H_n} - c_H| \leqq \varepsilon_n$.

Next we need

Proposition 4. *Let $H \in \mathscr{F}$ and define for $\lambda > 0$ the Hamiltonian $H_\lambda \in \mathscr{F}$ by*

$$H_\lambda(z) = \lambda^{-2} H(\lambda z).$$

Then we have the equality

$$\lambda^2 c_\lambda = c_H$$

for $\lambda > 0$ where $c_\lambda := c_{H_\lambda}$.

Proof. Since for large $|z|$ we have $H(z) = \hat{a}|z|^2$ we see that $H_\lambda(z) = \hat{a}|z|^2$ for large $|z|$, too. Since $\hat{a} > \pi$ and $\hat{a} \notin \pi \mathbb{N}$ again the Palais-Smale condition holds, [17]. If x is a 1-periodic solution of $\dot{x} = X_H(x)$ then x_λ defined by

$$x_\lambda(t) = \lambda^{-2} x(t)$$

is a 1-periodic solution of $\dot{y} = X_{H_\lambda}(y)$ and

$$\Phi_{H_\lambda}(x_\lambda) = \lambda^{-2} \Phi_H(x).$$

In other words if $\Sigma \subset \mathbb{R}$ is the set of critical levels of Φ_H, then $\Sigma_\lambda := \lambda^{-2}\Sigma$ is the set of critical levels of Φ_{H_λ}. Arguing as in Proposition 3 we see that the map $\lambda \to \lambda^2 c_\lambda$ is continuous. If c_H is an isolated critical level the map $\lambda \to \lambda^2 c_\lambda$ has to be constant. Finally, using a genericity argument we conclude that in all cases

$$\lambda^2 c_\lambda = c_H$$

proving the desired result. □

Assume now S is a bounded subset of the symplectic vectorspace $(\mathbb{C}^n, \omega)$, ω as introduced before. Denote by $\mathscr{F}(S)$ the collection of all Hamiltonians H in $\mathscr{F}$ such that H vanishes on an open neighborhood U of $cl(S)$. We define

$$c(S) = \inf\{c_H \mid H \in \mathscr{F}(S)\}. \tag{13}$$

Proposition 5. *Given $\Psi \in \mathrm{Diff}(\mathbb{C}^n)$ such that for some $\alpha > 0$ $\Psi^*\omega = \alpha\omega$, we have*

$$c(\Psi(S)) = \alpha c(S)$$

where S is a bounded subset.

Proof. Assume first $\alpha = 1$. Then we find a smooth arc $s \to \tilde{\Psi}_s$ in $D_c(\mathbb{C}^n)$ such that

$$\tilde{\Psi}_1(S) = \Psi(S).$$

By Proposition 3 we obtain

$$c_{H \circ \tilde{\Psi}_1^{-1}} = c_H \quad \text{for } H \in \mathscr{F}(S)$$

and therefore

$$c(\Psi(S)) = c(\tilde{\Psi}_1(S)) = c(S).$$

Next assume $\alpha > 0$ is different from one. Then $\alpha^{-\frac{1}{2}}\Psi \in D(\mathbb{C}^n)$ and by the previous step

$$c(\alpha^{-\frac{1}{2}}\Psi(S)) = c(S).$$

Next we employ Proposition 4 in order to obtain

$$c(\alpha^{-\frac{1}{2}}\Psi(S)) = \alpha^{-1}(\Psi(S)).$$

Summing up we find

$$c(\Psi(S)) = \alpha c(S)$$

as required. □

Next we define for any subset S of $\mathbb{C}^n$ a number $c(S) \subset [0, +\infty]$ by

$$c(S) = \sup\{c(T) \mid T \text{ is bounded, } T \subset S\}. \tag{14}$$

As a corollary of Proposition 5 we find

Corollary 1. *For any subset S of $\mathbb{C}^n$ and any diffeomorphism Ψ of $\mathbb{C}^n$ such that $\Psi^*\omega = \alpha\omega$ for some $\alpha > 0$ we have*

$$c(\Psi(S)) = \alpha c(S). \tag{15}$$

Moreover, if $S \subset T$ then $c(S) \leqq c(T)$.

Proof. The first part is evident by Proposition 5. For the second part we just note that if $S \subset T$, then $\mathscr{F}(T) \subset \mathscr{F}(S)$. □

In view of (15) we can now define for any subset S of a symplectic vectorspace (V, ω_v) a number $c_v(S)$ by

$$c_v(S) := c(\Psi(S)) \tag{16}$$

where $\Psi: (V, \omega_v) \to (\mathbb{C}^n, \omega_{\mathbb{C}^n})$ is a linear symplectic diffeomorphism. Clearly $c_v(S)$ is well defined. We shall show that c is the desired capacity. From our definition it is clear that we may assume in the following that $(V, \omega_v) = (\mathbb{C}^n, \omega)$. From our previous discussion we know that (A2) and (A3) hold. So it remains to prove (A1) (Normalization). In fact, we shall show even more, namely the properties claimed in Theorem 1.

Proposition 6. *Given $\Delta \in R$ we have*

$$c(\Delta) = c(B_\Delta) = k_\Delta \int \lambda \mid P_\Delta$$

for some positive integer k_Δ and a suitable $P_\Delta \in [\Delta]$.

Proof. Following [10, 17, 21, 22] we can foliate a bounded neighborhood of Δ by conformally symplectic images of Δ. More precisely we pick a 1-form on $\mathbb{C}^n$ such $d\lambda = \omega$ and $\lambda(x, \xi) \neq 0$ for nonzero $(x, \xi) \in \mathscr{L}_\Delta$. We define a vectorfield η on V by $\lambda = \omega(\eta, \cdot)$, Then $L_\eta \omega = \omega$ since $d\omega = 0$. Consequently η generates a symplectic dilation; see [22]. We may also assume that η has linear growth as one easily verifies. So η generates a global flow (Ψ_ε). We define $\Delta_\varepsilon := \Psi_\varepsilon(\Delta)$. Note that η is transversal to Δ. Hence for $\varepsilon_0 > 0$ small enough, to simplify notation say $\varepsilon_0 = 1$, we can write a small neighborhood of Δ in $\mathbb{C}^n$ as the disjoint union of the Δ_ε for $\varepsilon \in (-1, 1)$. Clearly $[\Delta_\varepsilon] = \Psi_\varepsilon([\Delta])$ and $\Psi_\varepsilon^* \omega = e^\varepsilon \omega$. Define $r_0 = \operatorname{diam}\left(\bigcup_{\varepsilon \in [-\frac{1}{2}, \frac{1}{2}]} \Delta_\varepsilon\right)$ and fix for $k \in \mathbb{N}, k \geqq 1$, a number $b > (k + \frac{1}{2})\pi r_0^2$. Pick a smooth map g: $\mathbb{R} \to \mathbb{R}$ such that

$$\begin{aligned} g(s) &= b && \text{for } s \leqq r_0 \\ g(s) &= (k+\tfrac{1}{2})\pi s^2 && \text{for } s \text{ large} \\ g(s) &\geqq (k+\tfrac{1}{2})\pi s^2 && \text{for } s \geqq r_0 \\ 0 < g'(s) &\leqq (2k+1)\pi s && \text{for } s > r_0. \end{aligned}$$

Moreover, let φ: $\mathbb{R} \to \mathbb{R}$ be a smooth map such that

$$\begin{aligned} \varphi(s) &= 0 && \text{for } s < \beta_1 \\ \varphi'(s) &> 0 && \text{for } \beta_1 < s < \beta_2 \\ \varphi(s) &= b && \text{for } \beta_2 \leqq s \end{aligned} \tag{17}$$

for suitable $0 < \beta_1 < \beta_2 < \frac{1}{2}$. Define a Hamiltonian $H \in \mathscr{F}(B_\Delta)$ by

$$H(z) = \begin{bmatrix} 0 & \text{if } z \in B_\Delta \\ \varphi(\varepsilon) & \text{if } z \in \Delta_\varepsilon \\ b & \text{if } z \notin B_{\Delta_{\beta^2}}, |z| \leqq r_0 \\ g(|z|) & \text{if } |z| > r_0. \end{bmatrix} \tag{18}$$

Arguing as in [17] we have the following fact:

If $\Phi_H(x) > 0$ and $\Phi'_H(x) = 0$, then $x([0,1]) \subset \Delta_\varepsilon$ for some $\varepsilon \in (\beta_1, \beta_2)$. We denote by Σ the set $\{kA(P) \mid k \in \mathbb{N}, P \in [\Delta]\}$. If Σ_ε denotes the corresponding set for Δ_ε, we see that we must have

$$\Sigma_\varepsilon = e^\varepsilon \Sigma.$$

Let us first assume that Σ is a discrete set. Then the critical levels of Φ_H are discrete as one easily verifies. (Note that $\Sigma \subset (\gamma, +\infty)$ for some $\gamma > 0$ since $\Delta \in R$.) Moreover, the positive critical levels only depend on the choice of φ but not on the choice of g (assuming g has the stated properties). In particular c_H does not depend on the choice of g, since we can homotope from one g to another so that the critical levels would change continuously, whereas the critical level c_H has to be a priori in a discrete set. Define $\mathscr{F}(B_\Delta)_{b,k}$ by

$$\begin{aligned} &H \in \mathscr{F}(B_\Delta)_{b,k} \quad \text{if } H \in \mathscr{F}(B_\Delta) \quad \text{and} \quad H(z) \leqq b \text{ for } |z| \leqq r_0 \text{ and} \\ &H(z) = (k+\tfrac{1}{2})\pi |z|^2 \qquad \text{for } |z| \text{ large.} \end{aligned} \tag{19}$$

Define

$$c(B_\Delta)_{b,k} = \inf\{ c_H \mid H \in \mathscr{F}(B_\Delta)_{b,k}\}.$$

Now we construct a particular sequence (H_m), where H_m is as in (18) and $\varphi = \varphi_m$ and $g = g_m$. For that we choose the β_1^m and β_2^m in such a way that

$$\varphi'_m(s) \notin \Sigma_s \qquad \text{if } \varphi_m(s) \in [\beta_1^m, b - \beta_2^m]. \tag{20}$$

Now given any $H \in \mathscr{F}(B_\Delta)_{b,k}$ we can pick such an H_m such that $H_m \geqq H$. (Perhaps we have to modify the g_m, but positive critical levels do not depend on g_m.)

Hence

$$c_{H_m} \to c(B_\Delta)_{b,k}.$$

Using the choice of b and the definition of H_m we can argue as in [17] to find critical points x_m such that

$$\Phi'_{H_m}(x_m) = 0, \quad c_{H_m} = \Phi_{H_m}(x_m)$$
$$x_m([0,1]) \subset \Delta_{\varepsilon_m}, \quad \varepsilon_m \in (\beta_1^m, \beta_2^m).$$

Now using (20) we see that

$$\varphi_m(\varepsilon_m) \in [0, \beta_1^m] \quad \text{or} \quad \varphi_m(\varepsilon_m) \in [b - \beta_2^m, b].$$

Since $a(x_m) = \varphi'_m(\varepsilon_m)$ as one easily verifies, we obtain

$$|c_{H_m} - a(x_m)| \leqq \beta_1^m$$
$$\text{or} \qquad |c_{H_m} - a(x_m) + b| \leqq \beta_2^m.$$

Since $\Delta \in R$ we have an estimate of the form

$$\text{length}(x_m) \leqq c\, a(x_m).$$

In fact, $a(x_m) = \int_0^1 \lambda(\dot{x}_m)$, where λ is a 1-form such that $d\lambda = \omega$ and $|\lambda(x,\xi)| \geqq 2c|\varphi|$ for $(x,\xi) \in \mathscr{L}_\Delta$. Since "restricted contact type" is a C^1-open condition we see that $|\lambda(x,\xi)| \geqq c|\varphi|$ for $(x,\xi) \in \mathscr{L}_{\Delta_\varepsilon}$ and $|\varepsilon|$ small. Since $a(x_m) > 0$ we infer that

$$\lambda(\dot{x}_m) \geqq c|\dot{x}_m|$$

provided m is large enough, since $\varepsilon_m \to 0$. Consequently the length can be estimated by the action.

As in [17] changing perhaps the parameterization of the x_m we can use the Ascoli-Arzela-Theorem to find $l_b \in \mathbb{Z}$, $l_b \geqq 1$, and $P_b \in [\Delta]$ such that

$$c(B_\Delta)_{b,k} = l_b A(P) \tag{21}$$
$$\text{or} \qquad c(B_\Delta)_{b,k} = l_b A(P_b) - b.$$

The map $b \to c(B_\Delta)_{b,k}$ is nonincreasing. Hence the map $b \to l_b A(P_b)$ must be nonincreasing in both cases in (21) since it takes values in a discrete set Σ. Since

$$c(B_\Delta)_{b,k} \geqq c(B_\Delta)$$

the second case in (21) is impossible for large b. This shows that

$$c(B_\Delta)_{b,k} = l_b A(P_b).$$

Since $c(B_\Delta)_{b,k} \to c(B_\Delta)$ as $b, k \to +\infty$ and since Σ is discrete, we find $P \in [\Delta]$ and $k \in \mathbb{N}, k \geqq 1$, with

$$c(B_\Delta) = kA(P)$$

as required. So far we assumed that Σ is discrete. If Σ is not discrete we take a generic $\tilde{\Delta}$ being C^∞-close to Δ. This can be done due to results in [20]. Now being of restricted contact type is a C^1-open condition. So we can approximate Δ by

$\tilde{\Delta} \in R$ so that our arguments hold for $\tilde{\Delta}$. We can pick $\tilde{k} \in \mathbb{N}$, $\tilde{k} \geqq 1$, and $\tilde{P} \in [\tilde{\Delta}]$ such that

$$c(B_{\tilde{\Delta}}) = \tilde{k} A(\tilde{P}).$$

If $\tilde{\Delta} \to \Delta$ we find $c(B_{\tilde{\Delta}}) \to c(B_\Delta)$. In fact, we have if $\tilde{\Delta}$ is close enough to Δ,

$$B_{\Delta_{-\varepsilon}} \subset B_{\tilde{\Delta}} \subset B_{\Delta_\varepsilon}$$

(The convergence $\tilde{\Delta} \to \Delta$ is understood in the following sense: $\tilde{\Delta}$ is the graph of a section of the normal bundle of Δ and the section converges in C^∞ to the zerosection.) $(*)$ implies using the monotonicity of c,

$$e^{-\varepsilon} c(B_\Delta) \leqq c(B_{\tilde{\Delta}}) \leqq e^{\varepsilon} c(B_\Delta).$$

Moreover, since length $(\tilde{P}) \leqq cA(\tilde{P})$ if $\tilde{\Delta} \to \Delta$ for some constant $c > 0$ (since $\Delta \in R$), we can use the Ascoli-Arzela-Theorem, as in [17] for a suitable parametrization of the $\tilde{P}$ to find $k \in \mathbb{N}$, $k \geqq 1$, and $P \in [\Delta]$ such that

$$c(B_\Delta) = kA(P).$$

By our definition of $c(S)$ we immediately see that

$$c(\Delta) \leqq c(B_\Delta)$$

We like to show equality. In order to do so we go through our construction which gives (18) again. Define $\gamma_a \colon \mathbb{C}^n \to \mathbb{R}$ as follows. Pick for $a > 1$, $\tau_a \colon \mathbb{R} \to \mathbb{R}$ such that

$$\begin{aligned} &\tau_a(s) = a && \text{for } s \leqq -\frac{1}{a} \\ &\tau_a'(s) < 0 && \text{for } -\frac{1}{a} < s < -\frac{1}{2a} \\ &\tau_a(s) = 0 && \text{for } s \geqq -\frac{1}{2a}. \end{aligned}$$

Now define γ_a by

$$\gamma_a(z) = \begin{bmatrix} a & \text{for } z \in B_{\Delta_{(-\frac{1}{a})}} \\ \tau_a(\varepsilon) & \text{for } z \in \Delta_\varepsilon,\ -\frac{1}{a} < \varepsilon \leqq 0 \\ 0 & \text{for } z \notin B_\Delta. \end{bmatrix}$$

Now consider for any Hamiltonian $H \in \mathscr{F}(B_\Delta)$ the Hamiltonian $H_a \in \mathscr{F}(B_\Delta)$ defined by

$$H_a(z) = H(z) + \gamma_a(z).$$

If x is a nonconstant 1-periodic solution of $\dot{x} = X_{H_a}(x)$ and $x(0) \in B_\Delta$, we see that $x(0) \in \Delta_\varepsilon$ for some $\varepsilon \in (-1, 0)$ and moreover

$$\Phi_{H_a}(x) = \tau_a'(\varepsilon) - \tau_a(\varepsilon) < 0.$$

Hence the positive critical levels of Φ_{H_a} and Φ_H are the same. By a genericity argument the map $(0, +\infty) \to (0, +\infty)$: $s \to c_{H+s\gamma_a}$ has to be constant since

$$c_{H+s\gamma_a} = \Phi_{H+s\gamma_a}(x_{H+s\gamma_a}) = \Phi_H(x_{H+s\gamma_a})$$

and the map $s \to c_{H+s\gamma_a}$ has to be continuous. Since for every $\tilde{H} \in \mathscr{F}(\Delta)$ there exists a $H \in \mathscr{F}(B_\Delta)$ and a γ_a as just described so that

$$\tilde{H} \leqq H + \gamma_a$$

we infer

$$c_{\tilde{H}} \geqq c_{H+\gamma_a} = c_H \geqq c(B_\Delta).$$

This implies

$$c(\Delta) \geqq c(B_\Delta)$$

proving our desired result since we already know that $c(\Delta) \leqq c(B_\Delta)$. □

Finally we have to show the normalization property

Lemma 2. *Define for $\beta \in (0,1)$ the set B_β by*

$$B_\beta = \{z \in \mathbb{C}^n \mid |z_1|^2 + \beta(|z_2|^2 + \ldots + |z_n|^2) < 1\}.$$

Then

$$c(B_\beta) = \pi.$$

Proof. Pick a smooth map φ: $\mathbb{R} \to \mathbb{R}$ such that for given $\delta \in (0,1)$

$$\begin{aligned}
&\varphi(s) = 0 && \text{for } s \leqq 1+\delta \\
&\varphi''(s) > 0 && \text{for } s > 1+\delta \\
&\frac{\varphi'(s_0)}{s_0} = 2\pi && \text{for a unique } s_0 \in (1+\delta, 1+2\delta) \\
&\varphi'(s) \leqq 3\pi s && \text{for all } s \in \mathbb{R}.
\end{aligned} \tag{22}$$

We find $H \in \mathscr{F}(B_\beta)$ such that $H(z) \geqq \varphi(|z_1|)$ where $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$. Using the intersection result we estimate

$$\begin{aligned}
&\inf \Phi_H(h(S^+)) \\
&\leqq \sup \Phi_H(E^- \oplus E^0 \oplus \mathbb{R}e) \\
&\leqq \sup \Phi_\varphi(E^- \oplus E^0 \oplus \mathbb{R}e)
\end{aligned} \tag{23}$$

where $\Phi_\varphi(x) = a(x) - \int_0^1 \varphi(|x_1(t))|)dt$ where $x = (x_1, \ldots, x_n) \in E$. The last expression in (23) can be explicitly computed. One easily sees that there exists $x \in (E^- \oplus E^0 \oplus \mathbb{R}e)$ such that $x = (x_1, 0, \ldots, 0)$ and

$$\begin{aligned}
&-i\dot{x} = \frac{\varphi'(|x_1|)}{|x_1|} x \\
&x(0) = x(1).
\end{aligned} \tag{24}$$

This implies immediately that

$$\frac{\varphi'(|x_1(t)|)}{|x_1(t)|} = 2\pi \qquad \text{for all } t \in \mathbb{R}$$

which shows without loss of generality that we may assume that

$$\begin{aligned} x(t) &= (s_0 e^{2\pi i t}, 0, \ldots, 0) \\ &= s_0 e(t). \end{aligned}$$

Hence

$$\begin{aligned} \Phi_\varphi(x) &= \pi s_0^2 - \varphi(s_0) \\ &\leqq \pi s_0^2 \\ &\leqq \pi(1 + 2\delta)^2. \end{aligned}$$

Summing up this implies that

$$c(B_\beta) \leqq c_H \leqq \pi(1 + 2\delta)^2.$$

Since $\delta > 0$ was arbitrary we obtain

$$c(B_\beta) \leqq \pi. \tag{25}$$

On the other hand, we have by monotonicity

$$c(B^{2n}(1)) \leqq c(B_\beta) \leqq \pi. \tag{26}$$

By Proposition 6 we infer that

$$c(B^{2n}(1)) = c(S^{2n}(1)) = k\pi \tag{27}$$

for some positive integer k. Therefore combining (25)–(27) we find that

$$c(B_\beta) = \pi.$$

Now consider $\Sigma(r) = B^2(r) \times \mathbb{C}^{n-1}$. Since $B_\beta \subset \Sigma(1)$ we find that

$$c(\Sigma(1)) \geqq \pi.$$

Next assume S is a bounded set. For given $\delta \in (0, 1)$ we find $\beta \in (0, 1)$ such that

$$\delta S \subset B_\beta.$$

Hence, using the already established properties of c,

$$\delta^2 c(S) \leqq c(B_\beta) = \pi,$$

which gives, since $\delta \in (0, 1)$ was arbitrary:

For every bounded subset S of $\Sigma(1)$ we have $c(S) \leqq \pi$.

By definition

$$\begin{aligned} c(\Sigma(1)) &= \sup\{c(S) \mid S c \Sigma(1), S \text{ is bounded}\} \\ &\leqq \pi. \end{aligned}$$

Since we know already that $c(\Sigma(1)) \geqq \pi$ we must have

$$c(B^2(1) \times \mathbb{C}^{n-1}) = \pi. \qquad \square$$

Remark. To give some more examples of capacities of certain sets consider a bounded open set Ω and a linear subspace W of codimension 2, then, as proved in Lemma 3,

$$c(\Omega + W) = +\infty \quad \text{if } W^{\omega} \subset W$$

$$c(\Omega + W) < +\infty \quad \text{if } W^{\omega} \not\subset W.$$

For example, if $W = \{z \in \mathbb{C}^n \mid p_1 = p_2 = 0\}$, then $W^{\omega} \subset W$, if $W = \{z \in \mathbb{C}^n \mid z_1 = 0\}$, then $W^{\omega} \not\subset W$.

III. Applications to Embedding Results

Assume $\Psi\colon B^{2n}(1) \hookrightarrow \mathbb{C}^n$ is symplectic embedding. Recall the folklore "extension after restriction principle" which says that for given $\delta \in (0,1)$ there exists a $f_{\delta} \in D(\mathbb{C}^n)$ (one can take even $f_{\delta} \in D_c(\mathbb{C}^n)$ such that

$$\Psi \mid B^{2n}(\delta) = f_{\delta} \mid B^{2n}(\delta).$$

This result is not difficult to prove. We may assume that $\Psi(0) = 0$. Then define $\Psi_t(z) = \frac{1}{t}\Psi(tz)$ for $t \in (0,1)$ and $\Psi_0(z) = \Psi'(0)z$. Then $t \to \Psi_t$ is a continuous arc in the space of symplectic embeddings of $B(1)$ into $\mathbb{C}^n$. We can smoothen the arc and take a Hamiltonian $H\colon U \to \mathbb{R}$, where $U \subset [0,1] \times \mathbb{C}^n$ is open in $[0,1] \times \mathbb{C}^n$, which generates the smoothened arc. Then we restrict H to a suitable smaller set and extend the restriction smoothly to a Hamiltonian of compact support. If $\tilde{f}$ is the time 1-map for the associated Hamiltonian system then $f = \tilde{f} \circ \Psi_0$ is the desired global symplectic diffeomorphism. Using this simple remark we can use the capacity c to prove a celebrated result of Gromov [12].

Theorem 2. *Assume $\Psi\colon B^{2n}(r) \hookrightarrow \Sigma(r')\colon = B^2(r') \times \mathbb{C}^{n-1}$ is a symplectic embedding. Then $r \leqq r'$.*

Remark. Note that the cylinder $\hat{\Sigma}(r) = \{z \in \mathbb{C}^n \mid p_1^2 + p_2^2 < r^2\}$ has an infinite capacity for every $r > 0$ as a consequence of the remark before Section III. For $\delta > 0$ sufficiently small consider the symplectic linear map defined by

$$T_{\delta}(z_1, \ldots, z_n) = \left(\frac{1}{\delta}q_1 + i\delta p_1, \frac{1}{\delta}q_2 + i\delta p_2, z_3, \ldots, z_n\right).$$

Then $T_{\delta}(B(r') \subset \hat{\Sigma}(r))$ provided $\delta > 0$ is small enough.

Proof. Let $\delta \in (0,1)$ be given and let $\Psi_{\delta} \in D(\mathbb{C}^n)$ such that

$$\Psi_{\delta} | B(\delta r) = \Psi | B(\delta r).$$

Hence

$$\Psi_{\delta}(B(\delta r) \subset \Sigma(r')$$

which implies

$$\delta^2 \pi r^2 = \delta^2 r^2 c(B(1)) = c(\delta B(r)) \leqq c(\Sigma(r')) = \pi (r')^2.$$

Hence for every $\delta \in (0,1)$ we find $\delta^2 r^2 \leqq (r')^2$ proving that

$$r \leqq r'. \qquad \square$$

This result is remarkable since $\Sigma(r')$ has infinite volume and yet, for $r < r'$, does not have leave room enough to fit a symplectic copy of $B(r)$ which has finite volume. One can also use Theorem 2 as a starting point for constructing symplectic capacities.

For $F \subset \mathbb{C}^n$ define

$$\text{(1)} \qquad \bar{c}(F) = \inf\{\pi r^2 \mid \exists \Psi \in D(\mathbb{C}^n) \quad \text{with } \Psi(F) \subset \Sigma(r)\}$$

$$\underline{c}(F) = \sup\{\pi r^2 \mid \exists \Psi \in D(\mathbb{C}^n) \quad \text{with } \Psi(B(r) \subset F\}.$$

Here $r \in [0, \infty) \cup \{+\infty\}$. From the definition of $\underline{c}$ and $\bar{c}$ it is clear that

Proposition 7. *$\underline{c}$ and $\bar{c}$ are symplectic capacities. Every symplectic capacity satisfies*

$$\underline{c}(F) \leqq c(F) \leqq \bar{c}(F)$$

for all $F \subset \mathbb{C}^n$.

This raises several questions:

(a) For what sets F do we have $\underline{c}(F) = \bar{c}(F)$?

(b) For what sets F is the infimum or the supremum attained in formula (1)?

IV. Rigidity Problems

First we study linear maps which preserve a capacity c. We will then deduce that nonlinear C^1-maps which preserve a symplectic capacity are either symplectic or antisymplectic. C^0-rigidity will be an easy consequence. Define a set $E \subset \mathbb{C}^n$ to be a bounded ellipsoid if there is a positive definite quadratic form q such that

$$E = \{z \in \mathbb{C}^n \mid q(z) < 1\}.$$

In the following we shall consider $\mathbb{C}^n$ exclusively as a $2n$-dimensional real vectorspace. If we speak about a linear subspace of $\mathbb{C}^n$ we mean a linear subspace of the $\mathbb{R}$-vectorspace $\mathbb{C}^n$. Similarly for linear maps.

Theorem 3. *Assume $\Psi: \mathbb{C}^n \to \mathbb{C}^n$ is a linear map such that for every bounded ellipsoid E we have*

$$c(E) = c(\Psi(E)).$$

Then Ψ is symplectic or antisymplectic:

$$\Psi^* \omega = \omega \quad \text{or} \quad \Psi^* \omega = -\omega.$$

The proof depends on a simple lemma. We have to fix some notation. A 2-dimensional subspace W of $\mathbb{C}^n$ is said to be null iff $\omega \mid W = 0$. This is clearly equivalent to $W \subset W^\omega$, where $W^\omega = \{z \in \mathbb{C}^n | z \perp_\omega W\}$ is called the ω-orthogonal of W. Here $z \perp_\omega W$ means $\omega(z, w) = 0$ for every $w \in W$. The ω-orthogonal is of course defined for every subspace of $\mathbb{C}^n$.

Lemma 3. *Let Ω be a bounded open subset of $\mathbb{C}^n$ and W a linear subspace of codimension 2. Consider the cylinder $\Omega + W$. Then*

$$c(\Omega + W) = +\infty \text{ if } W^\omega \text{ is null}$$

$$0 < c(\Omega + W) < +\infty \text{ if } W^\omega \text{ is not null.}$$

Proof. We assume without loss of generality that Ω contains the origin. Using a linear symplectic change of coordinates we can bring ourselves to the situation where:

$$W = [z \in \mathbb{C}^n | \operatorname{Re}(z_1) = \operatorname{Re}(z_2) = 0\} \text{ in the first case}$$

$$W = \{z \in \mathbb{C}^n | z_1 = 0\} \text{ in the second case.}$$

In the first case we take $\varepsilon > 0$ so small that

$$(\operatorname{Re}(z_1))^2 + (Re(z_2))^2 \leqq \varepsilon^2 \Rightarrow z \in \Omega + W.$$

Now observe that for any $N > 0$ the ellipsoid

$$\left\{z \,|\, \frac{1}{\varepsilon^2} \operatorname{Re}(z_1)^2 + \frac{1}{N^2} \operatorname{Im}(z_1)^2 + \frac{1}{\varepsilon^2} \operatorname{Re}(z_2)^2 + \frac{1}{N^2} \operatorname{Im}(z_2)^2 + \sum_{k=3}^{n} \frac{1}{N^2} |z_k|^2 < 1\right\}$$

is contained in $\Omega + W$. This ellipsoid, say E_N, is symplectically equivalent to the ellipsoid $\tilde{E}_N = \left\{z \in \mathbb{C}^n | \frac{1}{\varepsilon N}(|z_1|^2 + |z_2|^2) + \frac{1}{N^2} \sum_{k=3}^{n} |z_k|^2 < 1\right\}$

If N is large enough then

$$c(E_N) = c(\tilde{E}_N) = \pi \varepsilon N.$$

Hence for every N large

$$c(\Omega + W) \geqq \pi \varepsilon N,$$

which proves our first assertion. For the second statement we take N so large so that

$$z \in \Omega + W \Rightarrow |z_1|^2 < N^2.$$

Hence for every $z \in \Omega + W$ we have

$$z \in \Sigma(N)$$

which prove

$$c(\Omega + W) \leqq c(\Sigma(N)) = \pi N^2 < \infty. \qquad \square$$

Given a linear map Ψ: $\mathbb{C}^n \to \mathbb{C}^n$ we denote by Ψ^* the ω-transpose defined by

$$\omega(\Psi x, y) = \omega(x, \Psi^* y).$$

For any linear subspace W of $\mathbb{C}^n$ we have the duality formula

(1) $$\Psi(W^\omega) = ((\Psi^*)^{-1}(W))^\omega$$

Proof of Theorem 3. First assume $\Psi(\mathbb{C}^n) \neq \mathbb{C}^n$. Then $\Psi(\mathbb{C}^n)$ is contained in a hyperplane through zero. Composing if necessary Ψ with a symplectic map we may assume

$$\Psi(\mathbb{C}^n) \subset \{z \in \mathbb{C}^n \mid \mathrm{Re}(z_1) = 0\} =: V.$$

If now S is a bounded ellipsoid, then $\Psi(S)$ is a bounded subset of V. It is now easy to construct bounded ellipsoid with arbitrarily small capacity containing $\Psi(S)$. This would lead to the contradiction

$$0 < c(S) = c(\Psi(S)) = 0.$$

Hence Ψ is invertible. To complete the proof we proceed in two steps.

Step 1. Let $W \subset \mathbb{C}^n$ be of co-dimension two. If W^ω is null so is $\Psi(W)^\omega$. Assume W^ω is null and Ω is a bounded open nonempty subset of $\mathbb{C}^n$. By Lemma 3 we have $c(\Omega + W) = +\infty$. We can exhaust $\Omega + W$ by bounded ellipsoids with arbitrarily high capacity. Hence $c(\Psi(\Omega) + \Psi(W) = +\infty$ as well. By Lemma 3 again $\Psi(W)^\omega$ must be null.

Using the duality formula we compute for a null 2-plane W in view of step 1

$$\begin{aligned}(\Psi^*)^{-1}(W) &= (\Psi(W^\omega))^\omega \\ &\subset \Psi(W^\omega) \\ &= ((\Psi^*)^{-1}(W))^\omega.\end{aligned}$$

Hence we have proved

(2) $(\Psi^*)^{-1}$ maps null 2-planes to null 2-planes.

Step 2. The linear map $\Phi = (\Psi^*)^{-1}$ is symplectic or antisymplectic. Let $e_1, \ldots, e_n \in \mathbb{R}^n \subset \mathbb{C}^n$ be the standard basis of $\mathbb{R}^n$. Then $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ is a symplectic basis for $\mathbb{C}^n$. Using (2) we have

$$\omega(\Phi(e_k), \Phi(e_l)) = 0 \quad \text{if} \quad \omega(e_k, e_l) = 0$$

where $e_{k+n} := ie_k$ for $k = 1, \ldots, n$. Hence

(3) $$\omega(\Phi e_k),(\Phi e_{k+n})) =: d_k, \qquad k = 1, \ldots, n$$

is nonzero, since ω is nondegenerate, Φ is onto and $\omega(\Phi(e_k), \Phi(e_j)) = 0$ for $j \neq k+n$. Since $\omega(e_k - e_j, e_{k+n} + e_{j+n}) = 0$ for $1 \leqq j < k \leqq n$ we see that $\omega(\Phi(e_k) - \Phi(e_j), \Phi(e_{k+n}) + \Phi(e_{j+n})) = 0$ which implies that $d_k = d_j$. So we denote the constant d_k defined by (3) simply by d. If $d > 0$ we define $\tilde{\Phi} = \dfrac{1}{\sqrt{d}} \Phi$ and if $d < 0$ we define $\tilde{\Phi} = \dfrac{1}{\sqrt{-d}} \Gamma \circ \Phi$, where $\Gamma: (\mathbb{C}^n, \omega) \to (\mathbb{C}^n, -\omega)$ is a symplectic isomorphism. It is immediate that $\tilde{\Phi}$ is symplectic. For a symplectic map we have the identity

$(\tilde{\Phi}^*)^{-1} = \tilde{\Phi}$. Hence

$$\tilde{\Phi} = \sqrt{d}\, \Psi \text{ in the first case}$$
$$\tilde{\Phi} = 1 - \sqrt{-d}\, \Gamma \circ \Psi \text{ in the second case.}$$

Observing that Γ is capacity preserving we find for every bounded ellipsoid S

$$c(S) = c(\Phi(S)) = |d|\, c(\Psi(S))$$
$$= |d|\, c(S).$$

Hence $d = 1$ or $d = -1$. So Ψ is either symplectic or antisymplectic as required. □

Here is an interesting consequence of the proof.

Corollary 2. *If $\Psi: \mathbb{C}^n \to \mathbb{C}^n$ is a linear map and there exists an open nonempty subset S such that $c(\Psi(S)) > 0$, then Ψ is invertible.* □

Next we go over to the nonlinear case

Theorem 4. *Let $\Psi: \mathbb{C}^n \to \mathbb{C}^n$ be a C^1-map such that for every bounded ellipsoid S we have*

$$c(\Phi(S)) = c(S).$$

Then Φ is symplectic or antisymplectic. Moreover if in addition $\mathrm{Id} \times \Phi: \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$ *is c-preserving on ellipsoids then Φ is symplectic.*

This result is remarkable because it allows us to extend the notion of symplectic to the C^0-category. We define a C^0-map Φ to be symplectic if Φ and $\mathrm{Id}_{\mathbb{C}} \times \Phi$ both preserve capacity. In view of Theorem 4 this coincides in the C^1-case with the standard definition. Theorem 4 is a consequence of

Theorem 5. *Let (Φ_k) be a sequence of continuous maps of the open unit ball $B(1)$ into $\mathbb{C}^n$ converging uniformly to Φ. Assume all the Φ_k preserve the capacity of bounded ellipsoids*

$$c(\Phi_k(S)) = c(S) \qquad \forall k \in \mathbb{N}$$

for all bounded ellipsoid $S \subset B(1)$. If Φ is differentiable at $0 \in B(1)$ then $\Phi'(0)$ is symplectic or antisympletic.

Proof. Let S be a bounded ellipsoid so that $cl(S) \subset B(1)$. Using the definition of the maximal capacity $\bar{c}$ we find for a given $\varepsilon > 0$ a bounded ellipsoid $\tilde{S}$ and a $\Psi \in D(\mathbb{C}^n)$ such that

$$cl(\Phi(S)) \subset \Psi(\tilde{S})$$
$$\bar{c}(\Psi(\tilde{S})) = \bar{c}(\tilde{S}) \leqq \varepsilon + \bar{c}(\Phi(S)).$$

For k large enough we must have

$$\Phi_k(S) \subset \Psi(\tilde{S})$$

which gives

$$\begin{aligned}\bar{c}(S) &= c(S) \\ &= c(\Phi_k(S)) \\ &\leqq \bar{c}(\Phi_k(S)) \\ &\leqq \bar{c}(\Psi(\tilde{S})) \\ &= \bar{c}(\tilde{S}) \\ &\leqq \varepsilon + \bar{c}(\Phi(S)).\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary we infer that

$$\bar{c}(S) \leqq \bar{c}(\Phi(S)) \tag{4}$$

for every bounded ellipsoid S with closure in $B(1)$. Hence for $t \in (0,1)$

$$\begin{aligned}\bar{c}(S) &= \frac{1}{t^2}\bar{c}(tS) \leqq \frac{1}{t^2}\bar{c}(\Phi(tS)) \\ &= \bar{c}\left(\frac{1}{t}\Phi(tS)\right).\end{aligned} \tag{5}$$

If now $t \downarrow 0$, $\frac{1}{t}\Phi(t\cdot)$ converges uniformly to $\Phi'(0)$ and we conclude by the same argument as above

$$\lim_{t \downarrow 0} \bar{c}\left(\frac{1}{t}\Phi(tS)\right) \leqq \bar{c}(\Phi'(0)S).$$

Consequently

$$\bar{c}(S) \leqq \bar{c}(\Phi'(0)S). \tag{6}$$

for every bounded ellipsoid S with closure in $B(1)$. From Corollary 2 we see that $\Phi'(0)$ is invertible. Next we have to show the reversed inequality in (6). Since h is differentiable at 0 we find a continuous increasing map ε: $(0, 1) \to (0, \infty)$ such that $\varepsilon(s) \to 0$ as $s \to 0$ and assuming $\Phi(0) = 0$

$$|\Phi(x) - \Phi'(0)x| \leqq \varepsilon(|x|)\,|x|.$$

Given $\delta \in (0, 1)$ we find $k(\delta)$ such that for $k \geqq k(\delta)$

$$|\Phi_k(x) - \Phi'(0)x| \leqq \varepsilon(|x|)\,|x| + \delta.$$

Pick a linear ellipsoid S with closure in $B(1)$ and let $t \in (0,1)$. Pick $\gamma > 0$; then, if τ is small enough $(1+\gamma)\tau < 1$. Consider the equation

$$t\Phi_k((1+\gamma)\tau x) + (1-t)\Phi'(0)(1+\gamma)\tau x) = \Phi'(0)(\tau z) \tag{7}$$

for given $z \in S$ and solvable for some $\tau \in [0, 1]$ and $x \in \partial S$. Pick a constant $d > 0$ such that

$$|\Phi'(0)x| \geqq d|x|.$$

Rearranging (7) we estimate using $x \in \partial S$ and $z \in S$ for k large (k depending on τ)

$$\begin{aligned} 2\,\varepsilon((1+\gamma)\tau)(1+\gamma)\tau & \\ &\geqq |\Phi'(0)(\tau z-(1+\gamma)\tau x| \\ &\geqq d|\tau z-(1+\gamma)\tau x| \\ &\geqq d\tau\, d_1(S) \end{aligned}$$

where $d_1(S) > 0$ is some "shape constant" depending on S and γ only. Since $\varepsilon((1+\gamma)\tau) \to 0$ as $\tau \to 0$ we see that this inequality cannot hold for a solution $x \in \partial S$ of (7) (for arbitrary $t \in [0,1]$ provided τ is small and k large. This implies via Brouwer-degree that for τ small and k large

$$\Phi_k((1+\gamma)\tau S) \supset \Phi'(\tau S).$$

This gives

$$\begin{aligned} &\bar{c}(\Phi'(0)(\tau S)) \\ &= c(\Phi'(0)(\tau S)) \\ &\leqq (\Phi_k((1+\gamma)\tau S)) \\ &\leqq \bar{c}((1+\gamma)\tau S). \end{aligned}$$

Hence

$$\bar{c}(\Phi'(0)S) \leqq (1+\gamma)^2\,\bar{c}(S) \tag{7}$$

for every bounded ellipsoid and every $\gamma > 0$. Combining (6) and (7) yields therefore

$$\bar{c}(\Phi'(0)S) = \bar{c}(S).$$

From Theorem 3 the derived result follows. □

A simple corollary of Theorem 5 is

Theorem 6. *Assume $\Phi_k\colon B(1) \to \mathbb{C}^n$ is a sequence of symplectic embeddings converging unifrmly to a continuous map Φ: $B(1) \to \mathbb{C}^n$, which is differentiable at* 0 *with derivative $\Phi'(0)$. Then $\Phi'(0) \in \mathrm{Sp}(\mathbb{C}^n)$.*

Proof. Given a subset S of $B(1)$ with $cl(S) \subset B(1)$ we must have $c(\Phi_k(S)) = c(S)$ for any capacity as the consequence of the extension after restriction principle. From Theorem 5 we conclude that $\Phi'(0)$ is symplectic or antisymplectic. If n is odd a antisymplectic is orientation reversing. Since Φ_k preserves the orientation the same has to be true for $\Phi'(0)$. Hence if n is odd $\Phi'(0)$ can be only symplectic. So assume n is even. Then $\mathrm{Id}_{\mathbb{C}} \times \Phi_k \to \mathrm{Id}_{\mathbb{C}} \times \Phi$ uniformly on $B^{2(n+1)}(\delta)$ for some small $\delta \in (0,1)$. Since $\mathrm{Id}_{\mathbb{C}} \times \Phi_k$ is symplectic we conclude from the previous argument that $\mathrm{Id}_{\mathbb{C}} \times \Phi'(0)$ is symplectic which implies that $\Phi'(0)$ is symplectic. □

Finally we obtain as a simple corollary of Theorem 6 a celebrated result by Eliashberg and Gromov.

Theorem 7. Diff(M, ω) *is closed for the compact open* C^0 *-topology in* Diff(M), *where* (M, ω) *is a symplectic manifold.*

Proof. Using Darboux-charts we can localize and apply Theorem 5. □

Bibliography

1. Amann, H., Zehnder, E.: Nontrivial solutions for a class of nonresonance problems and applications to nonlinear differential equations. Ann. Sc. Norm. Super Pisa **7**, 593–603 (1980)
2. Arnold, V.I.: Sur une propriété topologique des applications canoniques de la mécanique classique. C.R. Acad. Sciences Paris **261**, 3719–3722 (1965)
3. Arnold, V.I.: Mathematical methods of classical mechanics. Berlin Heidelberg New York: Springer 1978
4. Bennequin, D.: Problèmes ellptiques, surfaces de Riemann et structures symplectiques. Sém. Bourbaki Fevr 1986, 1–23
5. Chaperon, M.: Quellques questions de géométrie symplectique (d'après autres Poincaré, Arnold, Conley-Zehnder), S2m Bourbaki, n° 610. Astérisque 105–106, 231–249 (1983)
6. Clarke, F.: Periodic solutions to Hamiltonian inclusions. J. Differ. Equations **40**, 1–6 (1981)
7. Clarke, F., Ekeland, I.: Hamiltonian trajectories having prescribed minimal period, Commun. Pure Appl. Math. **33**, 103–116 (1980)
8. Conley, C., Zehnder, E.: The Birkhoff-Lewis Fixed point theorem and a conjecture by V.I. Arnold. Invent Math. **73**, 33–49 (1983)
9. Ekeland, I., Hofer, H.: Convex Hamiltonian energysurfaces and their periodic trajectories. Commun Math. Phys. **113**, 419–469 (1987)
10. Ekeland, I., Hofer, H.: Two symplectic fixed point theorems. To appear
11. Floer, A.: Morse-theory for Lagrangian intersection. To appear in J. Differ. Geom.
12. Gromov, M.: Pseudo-holomorphic curves in symplectic manifolds. Invent. Math. **82**, 307–347 (1985)
13. Gromov, M.: Partial differential relations. Berlin Heidelberg New York: Springer, 1986
14. Gromov, M.: Symplectic geometry hard and soft. ICM Talk, Berkeley, 1986. To appear
15. Herman, M.: Une remarque, manuscript, Paris, 1983
16. Hofer, H.: Lagrangian embeddings and critical point theory. Ann. Just. Henri Poincaré, **2**, 407–462 (1985)
17. Hofer, H., Zehnder, E.: Periodic solutions on hypersurfaces and a result by C. Viterbo. Invent. Math. **90**, 1–9 (1987)
18. Hofer, H.: On strongly indefinite functionals with application. TAMS **275** (11), 185–214
19. MacDuff, D.: Examples of symplectic structures. Invent. Math. **89**, 13–36 (1987)
20. Takens, F.: Hamiltonian systems: generic properties of closed orbits and local perturbations. Math. Ann. **188**, 304–312 (1970)
21. Viterbo, C.: A proof of the Weinstein conjecture in $\mathbb{R}^{2n}$. Ann. Just. Henri Poincaré, Analyse nonlinéaire **4**, 337–357 (1987)
22. Weinstein, A.: On the hypotheses of Rabinowitz's periodic orbit theorems. J. Differ. Equations **33**, 353–358 (1979)
23. Weinstein, A.: Lagrangian submanifolds and Hamiltonian systems. Ann. Math. **98**, 377–410 (1973)

Received April 25, 1988