THE NONSQUEEZING THEOREM. A JOURNEY THROUGH SYMPLECTIC GEOMETRY

GABRIEL AGUILAR PEREZ

Abstract. We develop the fundamental notions of symplectic geometry, including symplectic vector spaces and symplectic manifolds, in order to show some fundamental results in the dynamics of Hamiltonian systems and symplectic manifolds. With this background, we finally introduce the set of invariants called symplectic capacities, particularly the Hofer-Zehnder capacity, and finally prove Gromov's Nonsqueezing theorem.

Contents

1. Introduction

Symplectic geometry is born as a grand mathematical generalization of classical mechanics (in particular, it is born from the Hamiltonian formulation of mechanics), and in this way becomes its underlying mathematical formalism. More particularly, the fundamental tool of symplectic geometry, the *symplectic form*, together with the Hamiltonian function of a system, control the underlying physics of any given phase space (in the classical sense). The generalization of this tools has helped develop many important mathematical results, such as Gromov's *Nonsqueezing theorem*. We begin by defining the properties of a symplectic form, in order to develop the notion of symplectic vector spaces.

2. Fundamental notions

Definition 2.1. Let V be a real vector space such that dim $V = m$, and let $\omega : V \times V \to \mathbb{R}$ be a bilinear form. If $\omega(v, u) = -\omega(u, v)$ for all $v, u \in V$ we say ω is **skew-symmetric**.

Date: AUGUST 28, 2021.

Note that a consequence of this definition is that $\omega(v,v) = -\omega(v,v)$, which means that $\omega(v,v) = 0$ for all $v \in V$.

Theorem 2.2. ***(Standard Form for skew-symmetric bilinear forms)*** *If ω is a skew-symmetric bilinear form on $\mathbb{R}$, then there exists a basis $x_1, \ldots, x_k, e_1, \ldots, e_n, f_1, \ldots, f_n$ in V such that*

$$\omega(x_i, v) = 0 \text{ for all } i \text{ and all } v \in V,$$
$$\omega(e_i, e_j) = \omega(f_i, f_j) = 0 \text{ for all } i \text{ and } j,$$
$$\omega(e_i, f_j) = \delta_{ij} \text{ for all } i \text{ and } j \text{ (and } \delta_{ij} \text{ is the Kronecker delta).}$$

Proof. Let $U := \{u \in V \,:\, \omega(u,v) = 0 \;\forall\; v \in V\}$. Also, let W be a subspace of V such that $V = U \oplus W$. We can choose $e_1 \in W$, $e_1 \neq 0$, and by definition of W we can find $f_1 \in W$ such that $\omega(e_1, f_1) \neq 0$. By taking care of the necessary constants, we can safely assume that $\omega(e_1, f_1) = 1$. Let us define

$$W_1 := \text{span}\{e_1, f_1\}$$

and

$$W_1^\omega := \{w \in W \,:\, \omega(w,v) = 0 \;\forall\; v \in W_1\}.$$

If we consider some $v \in W_1 \cap W_1^\omega$, we know that $v = ae_1 + bf_1$, and we find that

$$\begin{cases} \omega(v, e_1) = \omega(ae_1 + bf_1, e_1) = a \cdot \omega(e_1, e_1) + b \cdot \omega(f_1, e_1) = -b, \\ \omega(v, f_1) = \omega(ae_1 + bf_1, f_1) = a \cdot \omega(e_1, f_1) + b \cdot \omega(f_1, f_1) = a \end{cases}$$

since $v \in W_1^\omega$, it follows that $a = b = 0$, and therefore $v = 0$. This means that $W_1 \cap W_1^\omega = \{0\}$. Hence, for some $v \in W$ assume $\omega(v, e_1) = c$ and $\omega(v, f_1) = d$. Thus, $(-cf_1 + de_1) \in W_1$ and $(v + cf_1 - de_1) \in W_1^\omega$, and we can write

$$v = (-cf_1 + de_1) + (v + cf_1 - de_1).$$

Therefore, $W = W_1 \oplus W_1^\omega$. Similarly, we can find a vector $e_2 \in W_1^\omega$ such that $e_2 \neq 0$, and there exists $f_2 \in W_1^\omega$ such that $\omega(e_2, f_2) \neq 0$. Call

$$W_2 := \text{span}\{e_2, f_2\},$$

and again

$$W_2^\omega := \{v \in W_1^\omega \,:\, \omega(v,w) = 0 \;\forall\; w \in W_2\}.$$

We can continue this process as in the first step by further decomposing the subspaces, as it will eventually end since V is finite-dimensional. Finally, we find that

$$V = U \oplus W_1 \oplus W_2 \oplus \cdots \oplus W_n,$$

where each W_i has a basis e_i, f_i, and $\omega(e_i, f_i) = 1$. □

Note that this result is independent of the choice of basis of the subspace U defined above. Moreover, from the proof we see that dim $V = k + 2n$, where $2n$ is the rank of ω and $k :=$dim U.

Definition 2.3. The linear map $\iota : V \to V^*$ is defined by $(\iota(v))(w) = \omega(v,w)$, and is usually called the **interior product**.

Note that the subspace U defined in the proof of Theorem 2.2 is the kernel of the interior product. We can now state the following, fundamental definition.

Definition 2.4. A **symplectic form** on a finite-dimensional real vector space V is a skew-symmetric bilinear form $\omega : V \times V \to \mathbb{R}$ such that the interior product ι is an isomorphism; that is, $U = \{0\}$. In other words, a symplectic form is a skew-symmetric bilinear form which is **non-degenerate**: for any $v \in V$, $v \neq 0$, there exists $w \in V$ such that $\omega(v, w) \neq 0$.

In this way, a **symplectic vector space** is a pair (V, ω), where V is a infinite-dimensional real vector space and ω is the symplectic form we just defined. Because ω is non-degenerate, it follows that $\dim U = k = 0$, which means that $\dim V = 2n$. Therefore, any symplectic vector space V is necessarily even-dimensional. In this spirit, a basis in the symplectic vector space as defined in Theorem 2.2 $(e_1, \ldots, e_n, f_1, \ldots, f_n)$ is called a **symplectic basis**.

3. Symplectic geometry vs. Euclidean geometry

Let us denote by ω_0 the **standard symplectic form** in $\mathbb{R}^{2n}$, such that

$$\omega_0 = \sum_{j=1}^{n} dx_j \wedge dy_j.$$

Recall that in euclidean geometry there exists a natural bilinear form often called the *dot product*, defined in the euclidean space $\mathbb{R}^n$ by $v \cdot w = |v| \cdot |w| \cos\theta$, for $v, w \in \mathbb{R}^n$ and θ the angle between v and w. Generally, this product is naturally associated to the length of a vector when projected onto a certain direction. In particular, $v \cdot v = |v|^2$, where $|v|$ is the length (or norm) of the vector itself.

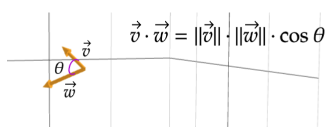

Figure 1.

In contrast, since we saw that $\omega_0(v, v) = 0$ for any $v \in V = \mathbb{R}^{2n}$, it is clear that contrary to the dot product, the symplectic form cannot measure lengths. However, because it is skew-symmetric and nondegenerate, it is natural to associate it to the measurement of a **signed-area**.

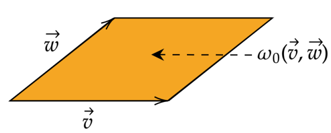

Figure 2.

For example, in $\mathbb{R}^2$, $\omega_0 = dx \wedge dy$, where indeed $dx \wedge dy$ is an area form. More particularly, we can see that $\omega_0(v, w) = \det(v, w)$, as in Figure 2.

Likewise, in $\mathbb{R}^4$, with coordinates (x_1, x_2, y_1, y_2), $\omega_0 = dx_1 \wedge dy_1 + dx_2 \wedge dy_2$, which is a sum of area forms. In this sense, for some surface S embedded in $\mathbb{R}^4$, $\omega_0(S)$ is the sum of the areas of its projections onto the $x_1 - y_1$ plane and $x_2 - y_2$

plane, as in Figure 3. Moreover, a surface in the $x_1 - x_2$ plane only projects a line onto the $x_1 - y_1$ and $x_2 - y_2$ planes, and thus $\omega_0(S) = 0$.

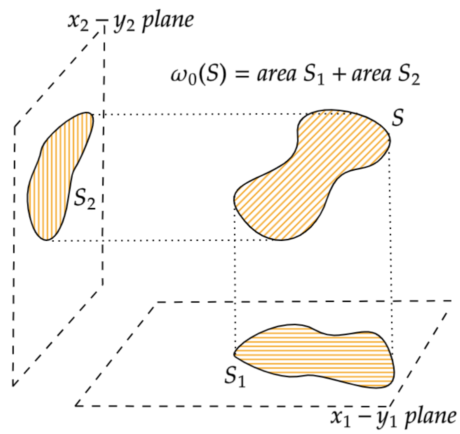

FIGURE 3. Surface in $\mathbb{R}^4$ as measured by ω_0.

4. Linear symplectic geometry and the affine Nonsqueezing theorem

Definition 4.1. A **linear symplectomorphism** of a symplectic vector space (V, ω) is an isomorphism $\varphi : V \to V$ such that

$$\varphi^* \omega = \omega,$$

where φ^* is the pullback of φ. That is, φ preserves the symplectic structure.

The set of linear symplectomorphisms of (V, ω) forms a group which is usually denoted by $\mathrm{Sp}(V, \omega)$. In particular, we write $\mathrm{Sp}(2n) = \mathrm{Sp}(\mathbb{R}^{2n}, \omega_0)$.

In terms of matrices, we define

$$\mathrm{Sp}(2g, \mathbb{R}) := \{A \in \mathrm{GL}(2g, \mathbb{R}) : A^* \omega = \omega\}.$$

Definition 4.2. A map $\psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ such that

$$\psi(x) = \Psi x + x_0,$$

where $\Psi \in \mathrm{Sp}(2n)$ and $x_0 \in \mathbb{R}^{2n}$, is called an **affine symplectomorphism**. We denote by $\mathrm{ASp}(2n)$ the group of affine symplectomorphisms.

Given the standard basis $e_1, \ldots, e_n, f_1, \ldots, f_n$ in $\mathbb{R}^{2n}$, we denote the **symplectic cylinder** with radius $R > 0$ by

$$\mathbf{Z^{2n}} := \mathbf{B^2(R)} \times \mathbb{R}^{\mathbf{2n-2}} = \{\mathbf{x} \in \mathbb{R}^{\mathbf{2n}} : \langle \mathbf{x}, \mathbf{e_1} \rangle^{\mathbf{2}} + \langle \mathbf{x}, \mathbf{f_1} \rangle^{\mathbf{2}} \leq \mathbf{R^2}\}, \tag{4.3}$$

where B^2 is the standard ball in $\mathbb{R}^2$, and $\langle \cdot, \cdot \rangle$ denotes the standard inner product in $\mathbb{R}^{2n}$. Now we can state the affine version of the main theorem in the paper.

Theorem 4.4. ***(Affine Nonsqueezing theorem)*** *Let $\psi \in ASp(2n)$. If ψ is such that $\psi(B^{2n}(r)) \subset Z^{2n}(R)$, then $r \leq R$.*

Proof. Consider an affine symplectomorphism $\psi(z)$ as it was just defined. Without loss of generality, let us assume $r = 1$. Given a symplectic basis $e_1, \dots, e_n, f_1, \dots, f_n$ in $\mathbb{R}^{2n}$, we shall consider the supremum of the radii of the balls B^{2n} obtained from the symplectic splitting of a cylinder Z^{2n} as we did in (4.3):

$$\begin{aligned}
\sup_{|x|=1} \left(\langle e_1, \psi(x)\rangle^2 + \langle f_1, \psi(x)\rangle^2\right) &= \sup_{|x|=1} \left(\langle e_1, \Psi x + x_0\rangle^2 + \langle f_1, \Psi x + x_0\rangle^2\right) \\
&= \sup_{|x|=1} \Big((\langle e_1, \Psi x\rangle + \langle e_1, x_0\rangle)^2 + (\langle f_1, \Psi x\rangle + \langle f_1, x_0\rangle)^2 \Big) \\
&= \sup_{|x|=1} \Big(\big(\langle \Psi^T e_1, x\rangle + a\big)^2 + \big(\langle \Psi^T f_1, x\rangle + b\big)^2 \Big) \\
&= \sup_{|x|=1} \Big((\langle v, x\rangle + a)^2 + (\langle w, x\rangle + b)^2 \Big) \leq R^2
\end{aligned}$$

where we let $v := \Psi^T e_1$, $w := \Psi^T f_1$, $a := \langle e_1, x_0\rangle$, and $b := \langle f_1, x_0\rangle$. Given such v and w, we know that $\Psi, \Psi^T \in \mathrm{Sp}(2n)$, and therefore $\omega_0(v, w) = 1$. Moreover, $1 = \omega_0(v, w) \leq |v| \cdot |w|$, which implies that either v or w must have norm greater than or equal to 1. Without loss of generality, let us assume that $|v| \geq 1$, and a natural choice for x is $\pm v/|v|$, with the sign depending on a. Hence,

$$\sup_{|x|=1} \Big((\langle v, x\rangle + a)^2 + (\langle w, x\rangle + b)^2 \Big) = \sup_{|x|=1} \Big((\pm|v| + a)^2 + (\langle w, x\rangle + b)^2 \Big) \leq R^2.$$

This means that $(\pm|v| + a)^2 \geq 1$, and thus $R \geq 1$. □

5. Symplectic manifolds. Where does symplectic geometry come from?

First, let us introduce introduce the notion of a manifold alongside some of its key concepts. Put in a few words, a manifold is a space which is locally euclidean.

Definition 5.1. Let M be a topological space. Then M is a **(topological) manifold of dimension** n if it satisfies:

- for any two distinct points p, q in M, there exist disjoint open neighborhoods $U, V \subset M$ with $p \in U$ and $q \in V$; that is, M is a **Hausdorff space**,
- there exists a countable basis for the topology of M; that is, M is **second-countable**,
- for any point in M, there exists an open neighborhood of each point that is homeomorphic to some subset $V \subset \mathbb{R}^n$; we say that M is **locally euclidean**.

In this sense, the fundamental property of a manifold M is the third: for any point $p \in M$, there exists an open neighborhood U of p and a subset $V \subset \mathbb{R}^n$, such that there exists a homeomorphism $\phi : U \to V$.

Definition 5.2. The pair (U, φ) is called a **coordinate chart**, and in general, we call the homeomorphism φ a **chart**. Moreover, a collection of charts whose domains cover M is called an **atlas** (for M).

Definition 5.3. Given two subsets $U, V \subset M$ with charts $\varphi : U \to \hat{U} \subset \mathbb{R}^n$ and $\psi : V \to \hat{V} \subset \mathbb{R}^n$, we call the map $\psi \circ \varphi^{-1} : \hat{U} \to \hat{V}$ a **transition map**. It is a composition of homeomorphisms and is therefore a homeomorphism.

Intuitively, a transition map acts like a function that allows you to "read page-to-page" the sections of such manifold through the atlas. See Figure 5.

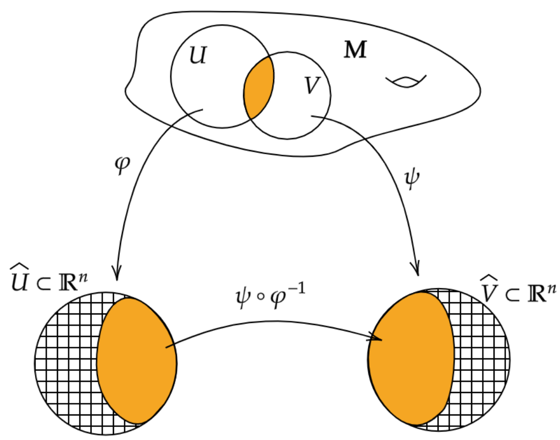

FIGURE 4.

Given two charts (U, φ) and (V, ψ), these are said to be **smoothly compatible** if either $U \cap V = \emptyset$ or the transition map $\psi \circ \varphi^{-1}$ is a diffeomorphism. In this spirit, we say that any atlas $\mathcal{A}$ is a **smooth atlas** if any given two charts contained in $\mathcal{A}$ are smoothly compatible. Furthermore, if any such atlas is **maximal** (that is, for any chart that is smoothly compatible with every other chart in $\mathcal{A}$, such chart is already in $\mathcal{A}$), we say that this atlas is a **smooth structure** on the manifold M.

Definition 5.4. The pair $(M, \mathcal{A})$, where M is a manifold and $\mathcal{A}$ is a smooth structure on M, is called a **smooth manifold**.

Finally, to produce some of the results later in the paper, we will briefly use the notion of a Riemannian manifold, so it is encouraged to see [2] for a brief introduction.

In the spirit of the "linearization of a manifold", we proceed to briefly introduce the notion of a tangent space. For the notion of *derivations*, see [2] pg. 52.

Definition 5.5. Given a point $p \in M$, the set of all derivations in $C^\infty(M)$ at p is a vector space called the **tangent space to M at** p, and is denoted by T_pM. Any element in T_pM is called a **tangent vector at** p.

To help develop a geometric intuition, the tangent space is usually associated to the set of velocity vectors of a curve $t \mapsto c(t)$, and the velocity vector of c at t_0, given by $dc/dt|_{t_0}$ is defined to be the vector $dc_{t_0} \in T_pM$.

For the rest of this section, let us assume that M is a connected C^∞-smooth manifold.

Definition 5.6. A **symplectic structure** on a smooth manifold M is a 2-form $\omega \in \Omega^2(M)$

$$\omega : T_pM \times T_pM \to \mathbb{R},$$

(where T_pM is the tangent space of M at a point $p \in M$) such that ω is nondegenerate and closed, that is, $d\omega = 0$.

Nondegeneracy implies that (T_pM, ω) is a symplectic vector space. The pair (M, ω) is the called a symplectic manifold.

Now, recall that from Classical Mechanics we learn that given a *Hamiltonian function* (a function that measures the energy in a given system) $H(x, y) : \mathbb{R}^{2n} \to \mathbb{R}$, there are two equations (Hamilton's equations) that uniquely determine the equations of motion of the system (given the proper initial conditions). Here, the symplectic manifold is identified with the phase space $\mathbb{R}^{2n}$. These equations can be written as:

$$\dot{x} = \frac{\partial H}{\partial y}, \;\; \dot{y} = -\frac{\partial H}{\partial x}.$$

Compactly, these equations give rise to a **Hamiltonian vector field**

$$X_H := -J_0 \nabla H(z) = -\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H}{\partial x} \\ \frac{\partial H}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial y} \\ -\frac{\partial H}{\partial x} \end{pmatrix} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix},$$

where J_0 is called a **complex structure** because $J_0^2 = -I$; and it geometrically corresponds to a $\pi/2$ rotation. As we will introduce later, the solution operator to this system is denoted by ϕ_H^{t,t_0} if given the solution $z(t)$ of the system we know that $z(t_0) = z_0$. In this case, $\phi_H^t(z_0) = z(t)$, and ϕ_H^{t,t_0} is called the **Hamiltonian flow** of H since given the Hamiltonian vector field, its solution resembles a fluid flow.

Note that given Hamilton's equations, we can write

$$\omega = \sum_i \partial x_i \wedge \partial y_i,$$

which naturally gives rise to the definition of a symplectic form and its properties (most easily, notice the skew-symmetric relation between x_i and y_i in Hamilton's equations), which in turn give back Hamilton's equations in the form of a Hamiltonian vector field. In other words: ω and H control the underlying physics of the system!

In this spirit, the significance of symplectomorphisms is the idea that whenever ω is preserved, the physics should remain the same.

6. Periodic solutions on Hamiltonian trajectories

Throughout this and the following sections, we closely follow the proofs presented in [1]. Consider some symplectic manifold (M, ω), and let us define the set

$$\mathcal{H}(M) := \left\{ H \in C_0^\infty(M) : H \geq 0, \; \mathrm{supp}(H) \subset M, \; H|_U = \sup H \text{ for some open set } U \right\},$$

that is, the set of nonnegative Hamiltonian smooth functions defined on M with compact support and which attain their maximum on some open set. For any $H \in \mathcal{H}(M)$ let us denote by ϕ_H^t the solution operator or *flow* of the Hamiltonian vector field X_H. If a solution $x(t) = \phi_H^t(x_0)$ is such that $x(t+T) = x(t)$ for some period T and for very $t \in \mathbb{R}$, we call it a **T-periodic orbit.**

Definition 6.1. The **symplectic action functional** of a loop z in $\mathbb{R}^{2n}$ with period T is defined by

$$A(z) = \int_0^1 \left(\langle y, \dot{x} \rangle - H(x, y) \right),$$

where $z(t) = (x(t), y(t))$.

Let us denote by $H_0 : Z^{2n}(1) \to \mathbb{R}$ a Hamiltonian function satisfying:

I $\pi < \sup H_0 < \infty$ and $0 \leq H_0(z) \leq \sup H_0 \;\forall\; z \in Z^{2n}(1)$, II There is a compact set $K \subset \operatorname{int} Z^{2n}(1)$ such that for $z \notin K$, $H_0(z) = \sup H_0$,

III There exists an open set $U \subset K$ such that $H_0(z) = 0$ for $z \in U$.

Definition 6.2. We say a function H_0 has **quadratic growth** if

$$\left\|d^2 H(z)\right\| \leq c$$

for some constant $c > 0$, for all $z \in \mathbb{R}^{2n}$.

Throughout the rest of the section we will also denote by $\operatorname{Per}(H)$ the set of $1-$periodic solutions of $\dot{z} = X_H(z)$, and by $\operatorname{Fix}(\phi_H^1)$ the set of fixed points of the *time-1-map* of the Hamiltonian flow of H. Note that $\operatorname{Per}(H) \cong \operatorname{Fix}(\phi_H^1)$.

Lemma 6.3. *For a function $H_0 : Z^{2n} \to \mathbb{R}$ satisfying (I), (II), and (III), there exists a corresponding Hamiltonian $H : \mathbb{R}^{2n} \to \mathbb{R}$ satisfying:*
(i) There exists $R > 0$ such that $K \subset B^{2n}(R)$ and $H(z) = H_0(z)$ whenever $z \in Z^{2n}(1) \cap B^{2n}(R)$,
(ii) H has quadratic growth,
(iii) There is some constant $c > 0$ such that if $|z| > c$, then $\left|\phi_H^1 - z\right| \geq 1$,
(iv) If $z \in$Per(H) and $A(z) > 0$, then $z(t)$ is not a constant function and $z(t) \in K$ for any t.

Proof. Consider some $\varepsilon > 0$ such that $\varepsilon < \pi/2$ and $M > \pi + \varepsilon$. We can find a smooth function $f : [0, \infty) \to \mathbb{R}$ such that

$$\begin{cases} f(s) = M, \ \ 0 \leq s \leq 1, \\ f(s) = (\pi + \varepsilon)s, \ \ s \geq M, \\ f(s) \geq (\pi + \varepsilon)s, \ \ \forall\ s, \\ \pi + \varepsilon \geq f'(s) \geq 0, \ \ \forall\ s. \end{cases}$$

Also, we can find $R \geq 1$ such that for $z \notin B^{2n}(R)$, $H_0(z) = M$. Consider the smooth function $g : [0, \infty) \to \mathbb{R}$

$$\begin{cases} g(s) = 0, \ \ 0 \leq s \leq R^2, \\ g(s) = \pi s/2, \ \ s \geq 3R^2, \\ 0 \leq g'(s) \leq \pi, \ \ \forall s. \end{cases}$$

Finally, consider $z = (z_1, w) \in \mathbb{R}^{2n}$, with $z_1 \in \mathbb{R}^2$ and $w = (z_2, \dots, z_n) \in \mathbb{R}^{2n-2}$. We can write $Z^{2n}(1) := \{|z_1| \leq 1\}$. The hamiltonian H we look for is now given by

$$H(z) = \begin{cases} H_0(z) \ \ z \in Z^{2n}(1) \cap B^{2n}(R), \\ f(|z_1|^2 + g(|w|^2), \ \ \text{otherwise}. \end{cases}$$

Therefore, $H(z)$ has quadratic growth because f and g have quadratic growth, and it also satisfies (i) by its definition. To show that $H(z)$ also satisfies (iii), note that we can split the Hamiltonian vector field $\dot{z} = -J_0 \nabla H(z)$ into the system

$$\begin{cases} \dot{z}_1 & = -2if'\left(|z_1|^2\right) z_1, \\ \dot{w} & = -2ig'\left(|w|^2\right) w, \end{cases}$$

whenever $|z_1| \geq 1$ or $|w| \geq R$. We can solve it and thus obtain the flow ϕ_H^t of the system. By setting $t = 1$ to obtain the time-1 map, we find that in the domain

$\{\max\{|z_1|, |w|/R\} \geq 1\}$,

$$\phi_H^1(z_1, w) = \left(e^{-2if'(|z_1|^2)}z_1, e^{-2ig'(|w|^2)}w\right).$$

So $|z|^2 = |z_1|^2 + |w|^2 \geq c^2 \geq 2\max\left\{M, 3R^2, \left|1 - e^{-2i\varepsilon}\right|^{-2}\right\}$, and we have two cases:

- $|z|_1^2 \geq c^2/2 \geq M$, and hence $e^{-2if'(|z_1|^2)} = e^{-2i\varepsilon} \neq 1$,
- $|w|^2 \geq c^2/2 \geq 3R^2$ and thus $e^{-2ig'(|w|^2)} = -1$.

For the first case $\left|\phi_H^1 - z\right| \geq \left|1 - e^{-2i\varepsilon}\right| \cdot |z_1| \geq \left|1 - e^{-2i\varepsilon}\right| \cdot c/\sqrt{2} \geq 1$. For the second case, we obtain that $\left|\phi_H^1 - z\right| \geq 2|w| \geq 1$. This proves (iii).

To prove (iv), notice that the symplectic action of a constant solution $z(t) \equiv z_0$ is $A(z) = \int_0^1 0 - H(z) = -H(z_0) \leq 0$, which is a contradiction. Thus $z(t)$ must be nonconstant. Similarly, if we assume that the conditions in (iv) hold, and we also assume that $z(t) \notin K$, we obtain that $A(z) \leq 0$, which is once more a contradiction. This proves (iv). Now the proof is complete. □

Lemma 6.4. *Consider a Hamiltonian $H : \mathbb{R}^{2n} \to \mathbb{R}$ that has quadratic growth. There exists sufficiently small $\epsilon > 0$ such that there exists a function $V_\epsilon : \mathbb{R}^{2n} \to \mathbb{R}$ satisfying*
(i) $(x_1, y_1) = \phi_H^\epsilon(x_0, y_0)$ if and only if

$$\frac{x_1 - x_0}{\epsilon} = \frac{\partial V_\epsilon}{\partial y}(x_1, y_0), \quad \frac{y_1 - y_0}{\epsilon} = -\frac{\partial V_\epsilon}{\partial x}(y_1, y_0).$$

(ii) V_ϵ converges to H in the $C^\infty-$topology as $\epsilon \to 0$. Furthermore, V_ϵ has quadratic growth and there exists some $c > 0$ such that

$$\sup_{z\in\mathbb{R}^{2n}} |V_\epsilon(z) - H(z)| \leq c\epsilon\left(|z|^2 + 1\right).$$

(iii) $Crit(H) = Crit(V_\epsilon) = Fix(\phi_H^t)$ and thus $H(z) = V_\epsilon$ for every critical point $z \in Crit(H)$.
(iv) There exists a unique solution $z(t) = z(t; x_1, y_0, \epsilon)$ to the boundary value problem

$$\dot{z} = X_H(z), \;\; x(\epsilon) = x_1, \;\; y(0) = y_0,$$

for all $x_1, y_0 \in \mathbb{R}^n$. The action $A(z)$ of this solution on $[0, \epsilon]$ is given by

$$A_H^{[0,\epsilon]}(z) = \langle y_0, x_1 - x_0\rangle - \epsilon \cdot V_\epsilon(x_1, y_0).$$

*Any function V_ϵ with these properties is called a **generating function of type** V_ϵ.*

Proof. See [1] (pg. 493) for a proof. □

Now, let us choose some $\epsilon = 1/N$ for large enough $N \in \mathbb{N}$. Let us denote by $\mathbb{R}^{2nN}$ the space of discrete loops $\mathbf{z} = (z_j)_{j\in\mathbb{Z}}$, where $z_{j+N} = z_j$ for all j. We define the **discrete symplectic action** $A_H^\epsilon : \mathbb{R}^{2nN} \to \mathbb{R}$ by

$$A_H^\epsilon(z) = \sum_{j=0}^{N-1} \left(\langle y_j, x_{j+1} - x_j\rangle - \epsilon V_\epsilon(x_{j+1}, y_j)\right), \tag{6.5}$$

where V_ϵ is the generating function defined in Lemma 6.4. In this discrete sense, the critical points of A_H^ϵ are the sequences $z_j \in \mathbb{R}^{2n}$ such that

$$z_{j+1} = \phi_H^\epsilon(z_j), \;\; z_{j+N} = z_j.$$

Corollary 6.6. *The map* $C^\infty(\mathbb{R}/\mathbb{Z}, \mathbb{R}^{2n}) \to \mathbb{R}^{2nN} : z \mapsto \{z(j/N)\}_{j\in\mathbb{Z}}$*, identifies the critical points of* A_H *with those of* A_H^ϵ *for* $\epsilon = 1/N$*. Furthermore, for every* $z \in Per(H)$*,*

$$A_H(z) = A_H^\epsilon(\{z(j/N)\}_{j\in\mathbb{Z}}).$$

Proof. This is a direct consequence of Lemma 6.4. □

Now, consider a Riemannian manifold X. Let $\Phi : X \to X$ be a smooth function such that the vector field $\nabla\Phi : X \to TX$ is complete, in the sense that for any $x_0 \in X$, the flow $\phi_t : X \to X$ (with $x(t) = \phi_t(x_0)$ whenever $x(0) = x_0$) is the unique solution the system

$$\dot{x} = -\nabla\Phi(x).$$

Definition 6.7. We say that Φ satisfies the **Palais-Smale** condition if any sequence $x_\nu \in X$ satisfying

$$\lim_{\nu\to\infty} \|\nabla\Phi(x_\nu)\| = 0$$

has a convergent subsequence. This implies that the set $\mathrm{Crit}(\Phi)$ is compact.

Lemma 6.8. *Consider a function* Φ *satisfying the Palais-Smale condition. Let* $c \in \mathbb{R}$ *be a regular (non-critical) value of* Φ*. Then, for any* $T > 0$ *there exists some* $\delta > 0$ *such that*

$$\Phi(x) \le c + \delta \ \Rightarrow \Phi(\phi_T(x)) \le c - \delta.$$

Proof. We can argue by contradiction. That is, there exists a sequence $x_\nu \in X$ such that $\Phi(x_\nu) \to c$ and $\Phi(\phi_T(x_\nu)) \to c$, which is the same as having
$\Phi(x_\nu) - \Phi(\phi_T(x_\nu)) \to 0$.

Therefore, since

$$\frac{d}{dt}\Phi(\phi_t(x_\nu)) = -\|\nabla\Phi(\phi_t(x_\nu))\|^2,$$

it follows that

$$\int_0^T \|\nabla\Phi(\phi_t(x_\nu))\|^2 dt = \Phi(x_\nu) - \Phi(\phi_T(x_\nu)$$

which means that $\int_0^T \|\nabla\Phi(\phi_t(x_\nu))\|^2 dt \to 0$. Hence, there must exist some sequence $t_\nu \in [0, T]$ such that

$$\lim_{t_\nu\to\infty} \nabla\Phi(\phi_{t_\nu}(x_\nu)) = 0.$$

Since Φ satisfies the Palais-Smale condition, the sequence $\phi_{t_\nu}(x_\nu)$ has a convergent subsequence and the limit point x of this subsequence satisfies $\nabla\Phi(x) = 0$ and $\Phi(x) = c$. This is a contradiction because c is a regular point. The proof is complete. □

We can now show that A_H^ϵ satisfies the Palais-Smale condition.

Lemma 6.9. *Consider a Hamiltonian that has quadratic growth and assume there exist constants* $c, \delta > 0$ *such that*

$$|z| > c \ \Rightarrow \left|\phi_H^1 - z\right| \ge \delta.$$

Then, for any sufficiently small $\epsilon > 0$*,* $A_H^\epsilon : \mathbb{R}^{2nN} \to \mathbb{R}$ *satisfies the Palais-Smale condition.*

Proof. See [1] (pg. 497) for a proof. □

We need to prove that the discrete symplectic action has a critical point $\mathbf{z}$ such that $A_H^\epsilon(\mathbf{z}) > 0$. Notice that A_H^ϵ is a perturbation of the functional

$$A^\epsilon(z) = \sum_{j=0}^{N-1} \langle y_j, x_{j+1} - x_j \rangle = \frac{1}{2}\langle z, L^\epsilon z\rangle_\epsilon,$$

that is, when $H = 0$ in (6.5), which implies that $V_\epsilon \to 0$. L^ϵ denotes the Hessian of A^ϵ, which is a self-adjoint operator on $\mathbb{R}^{2nN}$ that maps a path $\zeta = \{\zeta_j\}_{j\in\mathbb{Z}}$ to the path (or sequence) $\zeta' = \{\zeta'_j\}_{j\in\mathbb{Z}}$, where we denote the variations in a path by $\zeta = (\xi_0, \dots, \xi_N, \eta_0, \dots, \eta_N)$, and

$$\xi'_{j+1} = \frac{\eta_j - \eta_{j+1}}{\epsilon}, \quad \eta'_j = \frac{\xi_{j+1} - \xi_j}{\epsilon}.$$

Lemma 6.10. *The smallest positive eigenvalue of L^ϵ is less than or equal to 2π, and for each eigenvalue there exists a corresponding eigenvector $\zeta = \{\zeta_j\}_{j\in\mathbb{Z}}$ such that $\zeta_j \in \mathbb{R}^2 \times 0 \subset \mathbb{R}^{2n}$. Furthermore,*

$$\lambda \in \sigma(L^\epsilon) \iff -\lambda \in \sigma(L^\epsilon),$$

where $\sigma(L^\epsilon)$ denotes the spectrum of L^ϵ.

Proof. Consider the eigenvalues of L^ϵ. Let us notice that $L^\epsilon(\xi, \eta) = \lambda(\xi, \eta)$ is equivalent to

$$\lambda\xi_{j+1} = \frac{\eta_j - \eta_{j+1}}{\epsilon}, \quad \lambda\eta_j = \frac{\xi_{j+1} - \xi_j}{\epsilon}.$$

We can combine these expressions into

$$\begin{pmatrix} \xi_{j+1} \\ \eta_{j+1} \end{pmatrix} = \begin{pmatrix} 1 & \lambda/N \\ -\lambda/N & 1-(\lambda/N)^2 \end{pmatrix} \begin{pmatrix} \xi_j \\ \eta_j \end{pmatrix}.$$

Then, the eigenvalues of L^ϵ are given by

$$\det(I - A_N(\lambda)^N) = 0, \quad A_N(\lambda) = \begin{pmatrix} 1 & \lambda/N \\ -\lambda/N & 1-(\lambda/N)^2 \end{pmatrix}$$

and the eigenvectors have the form $\zeta = \{\zeta_j\}_{j\in\mathbb{Z}}$, where $\zeta_j = (\zeta_{j1}, \dots, \zeta_{jn}) \in \mathbb{R}^{2n}$. Notice that $\det(A_N(\lambda)) = 1$ and $\mathrm{trace}(A_N(\lambda)) = 2 - \frac{\lambda^2}{N^2} \in [-2, 2]$ if $\lambda^2/N^2 \leq 4$. If $|\lambda|/N > 2$, then the eigenvalues would not be coherent with the determinant of the matrix. Thus, it must be the case that $|\lambda|/N < 2$. It follows that the eigenvalues are given by $e^{\pm i\beta_N(\lambda)}$, where $\beta_N : [-2N, 2N] \to \mathbb{R}$ is a continuous function such that $\beta_N(0) = 0$ and

$$\cos\beta_N(\lambda) = 1 - \frac{\lambda^2}{2N^2}, \quad -2N \leq \lambda \leq 2N.$$

Taking its derivative, we see that

$$N\beta'_N(\lambda) = \frac{1}{\sqrt{1 - \lambda^2/4N^2}} \geq 1,$$

which means that $N\beta_N(\lambda) \geq \lambda \geq 0$. Therefore, the smallest positive eigenvalue $\lambda_N > 0$ of L^ϵ is such that $N\beta_N(\lambda_N) = 2\pi$. This shows that $\lambda_N \leq 2\pi$. Finally, we can check the last assertion by noticing that the following two are equivalent:

$$L^\epsilon(\xi, \eta) = \lambda(\xi, \eta) \iff L^\epsilon(\xi, -\eta) = -\lambda(\xi, -\eta).$$

The proof is complete. □

Having described the eigenvalues of L^ϵ, consider the eigenspace decomposition

$$\mathbb{R}^{2nN} = E^+ \cup E^0 \cup E^-$$

such that $E^0 = \ker L^\epsilon$, the eigenvalues of $L^\epsilon|_{E^+}$ are positive, and the eigenvalues of $L^\epsilon|_{E^-}$ are negative. Notice that E^0 consists of the constant paths or sequences. We can finally state and prove the lemma on which the proof of the Nonsqueezing theorem rests.

Lemma 6.11. *Consider $H_0 : Z^{2n}(1) \to \mathbb{R}$ satisfying (I), (II), and (III) from definition 6.1. Then the flow of H_0 has a nonconstant 1-periodic solution.*

Proof. Let us assume without loss of generality that the open set U from (III) is a neighbourhood of 0. Let H be the extension of H_0 found in Lemma 6.3. Let V_ϵ from Lemma 6.4 be the corresponding generating function V_ϵ of H, for sufficiently small $\epsilon = 1/N$. We wish to prove that A^ϵ_H has a critical point with $A^\epsilon_H > 0$.

Let us define $\Gamma := \{z \in E^+ : \|z\|_\epsilon = \alpha\}$. In this way, if $\alpha > 0$ is sufficiently small, then $\inf_\Gamma A^\epsilon_H > 0$. By Lemma 6.4, V_ϵ vanishes near zero. Thus A^ϵ_H becomes the unperturbed functional A^ϵ near zero. By Lemma 6.10, there exists some eigenvector $\zeta \in E^+$ of L^ϵ with a corresponding eigenvalue $\lambda \leq 2\pi$. We define

$$\Sigma = \{z + s\zeta : z \in E^- \oplus E^0,\ \|z\|_\epsilon \leq T,\ 0 \leq s \leq T\}.$$

In this way, if T is sufficiently large, then $\sup_{\partial\Sigma} A^\epsilon_H \leq 0$. By the definition of H, there exists some $C > 0$ such that

$$H(z_1, w) \geq (\pi + \varepsilon)|z_1|^2 + \frac{\pi}{2}|w|^2 - C.$$

Once more by Lemma 6.4 (ii), we see that for small enough ϵ,

$$V_\epsilon \geq (\pi + \frac{\varepsilon}{2})|z_1|^2 + \frac{\pi}{4}|w|^2 - 2C.$$

Let $z = z^- + z^0 \in E^- \oplus E^0$ and $s \in \mathbb{R}$. Thus,

$$\begin{aligned} A^\epsilon_H(z + s\zeta) &= A^\epsilon(z^-) + s^2 A^\epsilon(\zeta) - \frac{1}{N}\sum_{j=0}^{N-1} V_\epsilon(x_{j+1} + s\xi_{j+1}, y_j + s\eta_j) \\ &\leq \frac{\lambda s^2}{2}\|\zeta\|_\epsilon^2 - \left(\pi + \frac{\varepsilon}{2}\right)\|z_1 + s\zeta\|_\epsilon^2 - \frac{\pi}{4}\|z\|_\epsilon^2 + 2C \\ &\leq -\frac{\varepsilon s^2}{2}\|\zeta\|_\epsilon^2 - \frac{\pi}{4}\|z\|_\epsilon^2 + 2C, \end{aligned}$$

where we have used the fact that $A^\epsilon(z^-) \leq 0$ and $A^\epsilon(\zeta) = \frac{\lambda}{2}\|\zeta\|_\epsilon^2$ (as one can check from the definition), and that $\lambda \leq 2\pi$. Therefore, $\partial\Sigma$ consists of those points of the form $z + s\zeta \in \Sigma$ that satisfy one of the following:
(i) $\|z_\epsilon\| = T$, $0 \leq s \leq T$,
(ii) $\|z_\epsilon\| \leq T$, $s = T$,
(iii) $\|z_\epsilon\| \leq T$, $s = 0$.

Next, we show that $A^\epsilon_H : \mathbb{R}^{2nN} \to \mathbb{R}$ has a positive critical value. We define the **linking number** as the intersection number of any ball in the sphere $\partial\Sigma$ with Γ.

In this way, $\partial\Sigma$ and Γ have linking number 1. Let us denote by $\Sigma_t := \phi_t(\Sigma)$ the image of Σ under the gradient flow of A_H^ϵ for $t \geq 0$. Also, let us define

$$c := \inf_{t\geq 0} \sup_{\Sigma_t} A_H^\epsilon,$$

and we wish to show that $c > 0$ and that it is a critical value of A_H^ϵ. From the estimate obtained above it follows that

$$\sup_{\partial\Sigma_t} A_H^\epsilon \leq 0, \ \forall\, t \geq 0.$$

This implies that $\partial\Sigma_t \cap \Gamma = \emptyset$ for every $t \geq 0$ and thus the linking number remains 1. For every $t \geq 0$, $\Sigma_t \cap \Gamma \neq \emptyset$, and therefore

$$\sup_{\Sigma_t} A_H^\epsilon \geq \inf_{\Gamma_t} A_H^\epsilon > 0,$$

which means that $c > 0$. Finally, assume c is not a critical value of A_H^ϵ. By Lemma 6.3, H satisfies the condition (iii) in the Lemma. Therefore, by Lemma 6.9, A_H^ϵ satisfies the Palais-Smale condition. Furthermore, by Lemma 6.8, there exists $\delta > 0$ with $T = 1$ such that the assertion of the same lemma holds. Hence, we can choose $t^* \geq 0$ such that

$$\sup_{\Sigma_{t^*}} A_H^\epsilon \leq c + \delta,$$

and therefore, since we assume c is a regular value,

$$\sup_{\Sigma_{t^*+1}} A_H^\epsilon \leq c - \delta.$$

However, this is a contradiction to the definition of c. Therefore c is positive and a critical value of A_H^ϵ. By Corollary 6.6, this gives a $1-$periodic solution $z \in$Per(H) with $A_H(z) > 0$. By Lemma 6.3, this is not a constant function, and it is also a periodic solution of the flow of H_0. Now the proof is complete. □

7. Symplectic capacities and Gromov's Non-squeezing theorem

One of the most important invariants in symplectic geometry is that of **capacity**. We understand capacities as the 2-dimensional *size* of symplectic manifolds, and in its development lies a proof for the invariant (coordinate-free) version of the Nonsqueezing theorem. Before going further, let us state a final theorem before the Nonsqueezing theorem.

Recall from Section 6 that we defined the set $\mathcal{H}(M)$ as the set of nonnegative Hamiltonian smooth functions defined on M with compact support and which attain their maximum on some open set. This gives rise to the following definition.

Definition 7.1. A function $H \in \mathcal{H}(M)$ is called **admissible** if every nonconstant periodic orbit of its corresponding flow ϕ_H^t has period $T > 1$. We denote the set of admissible Hamiltonian functions by $\mathcal{H}_{\text{ad}}(M, \omega)$.

Theorem 7.2. *Consider a Hamiltonian $H \in \mathcal{H}(Z^{2n}(1))$ with $\sup H > \pi$. The Hamiltonian flow of H has a nonconstant periodic orbit of period 1.*

Proof. Given such H, notice that the function $H_0 = (\sup H) - H$ satisfies conditions I, II, and III from definition 6.1. Hence, by Lemma 6.11, the flow of H_0 has a nonconstant $1-$periodic solution, and therefore so does the flow of H (recall the proof of the Lemma). □

Now we proceed to define capacities and show their full power.

Definition 7.3. A **symplectic capacity** is a map that assigns to every symplectic manifold (M, ω) a number $c(M, \omega)$ such that

- (monotonicity) If there exists an embedding $(M_1, \omega_1) \hookrightarrow (M_1, \omega_1)$, then $c(M_2, \omega_2) \leq c(M_2, \omega_2)$. In particular, if two symplectic manifolds are symplectomorphic, then $c(M_1, \omega_1) = c(M_2, \omega_2)$.
- (conformality) Let a be a scalar. Then $c(M, a\omega) = a \cdot c(M, \omega)$
- (nontriviality) $c(B^{2n}(1), \omega_0) > 0$ and $c(Z^{2n}(1), \omega_0) < \infty$.

For the purpose of this paper, we shall restrict the study of capacities to $\mathbb{R}^{2n}$. Now, a **symplectic capacity** is a map that assigns a number $c(\Omega)$ to some arbitrary set $\Omega \subset \mathbb{R}^{2n}$ such that:

- (monotonicity) If there exists an embedding $\psi : \Omega \to \mathbb{R}^{2n}$ such that $\psi(\Omega) \subset B$, then $c(\Omega) \leq c(B)$,
- (conformality) Let λ be a scalar. The set $c(\lambda \cdot \Omega)$ consists of all the points $(\lambda\Omega x, \lambda\Omega y)$ with $(x, y) \in \Omega$. Then, $c(\lambda\Omega) = \lambda^2 \cdot c(\Omega)$.
- (nontriviality) $c(B^{2n}(1)) > 0$ and $c(Z^{2n}(1)) < \infty$.

Definition 7.4. The **Hofer norm** on the Lie algebra of a given Hamiltonian $H \in C_0^\infty(M)$ is defined by

$$\|H\| := \max_M H - \min_M H,$$

with M being a noncompact symplectic manifold. If M is compact, any constant function has norm zero and so the Hofer norm defines a norm on the space of Hamiltonian vector fields.

We now introduce the **Hofer-Zehnder** capacity in order to prove the Nonzqueezing theorem.

Definition 7.5. We define **Hofer-Zehnder capacity** by

$$c_{HZ}(M, \omega) := \sup_{H \in \mathcal{H}_{ad}(M,\omega)} \|H\|,$$

where $\|.\|$ is the Hofer norm.

Theorem 7.6. *The capacity c_{HZ} is a symplectic capacity and*

$$c_{HZ}(Z^{2n}(1)) = c_{HZ}(B^{2n}(1)) = \pi.$$

More generally, $c_{HZ}(Z^{2n}(r)) = c_{HZ}(B^{2n}(r)) = \pi r^2$.

Proof. We want to show that c_{HZ} satisfies the properties of a capacity. Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds of dimension $2n$ and let $\psi : M_1 \to M_2$ be an embedding. Given some Hamiltonian $H_1 \in C_0(M_1, \mathbb{R})$, there exists a unique $H_2 \in C_0(M_2, \mathbb{R})$ such that $H_2 \equiv 0$ on $M_2 \backslash \psi(M_1)$ and $H_1 = H_2 \circ \psi$. Due to ψ, there is a one-to-one correspondence between the nonconstant periodic orbits of the flows H_1 and H_2. Therefore,

$$\begin{aligned} c_{HZ}(M_1, \omega_1) &= \sup_{H_1 \in \mathcal{H}_{ad}(M_1,\omega_1)} \|H_1\| \\ &= \sup_{H_2 \in \mathcal{H}_{ad}(M_2,\omega_2)} \|H_2\| \\ &\leq c_{HZ}(M_2, \omega_2). \end{aligned}$$

This proves monotonicity. To prove conformality, notice that $\mathcal{H}_{ad}(M, \lambda\omega) = \{\lambda H : H \in \mathcal{H}_{ad}(M, \omega)\}$.
Now, because the Hamiltonian(s) $H \in \mathcal{H}_{ad}$ considered by c_{HZ} are admissible, it follows that the flows have nonconstant periodic orbits with period greater than 1. Hence, by Theorem 7.2, $c_{HZ}(Z^{2n}(1) \leq \pi$. By constructing a smooth function in a similar way as was done for the proof of Lemma 6.3, we obtain that $c_{HZ}(B^{2n}(1)) \geq \pi$ (see [1] (pg. 484) for the complete proof). Therefore, having proved the monotonicity axiom, we obtain that $c_{HZ}(B^{2n}(1)) = c_{HZ}(Z^{2n}(1)) = \pi$. The proof is complete. □

For a grand finale, we can now show Gromov's nonsqueezing theorem.

Theorem 7.7. ***Nonsqueezing theorem*** *If there exists a symplectic embedding* $(B^{2n}(r), \omega_0) \hookrightarrow (Z^{2n}(R), \omega_0)$, *then* $r \leq R$.

Proof. Assuming there exists a symplectic embedding $\psi : B^{2n}(r) \hookrightarrow Z^{2n}(R)$, we obtain that

$$\pi r^2 = c_{HZ}(B^{2n}(r)) = c_{HZ}(\psi\left(B^{2n}(r)\right)) \leq c_{HZ}(Z^{2n}(R)) = \pi R^2.$$

This shows that $r \leq R$. □

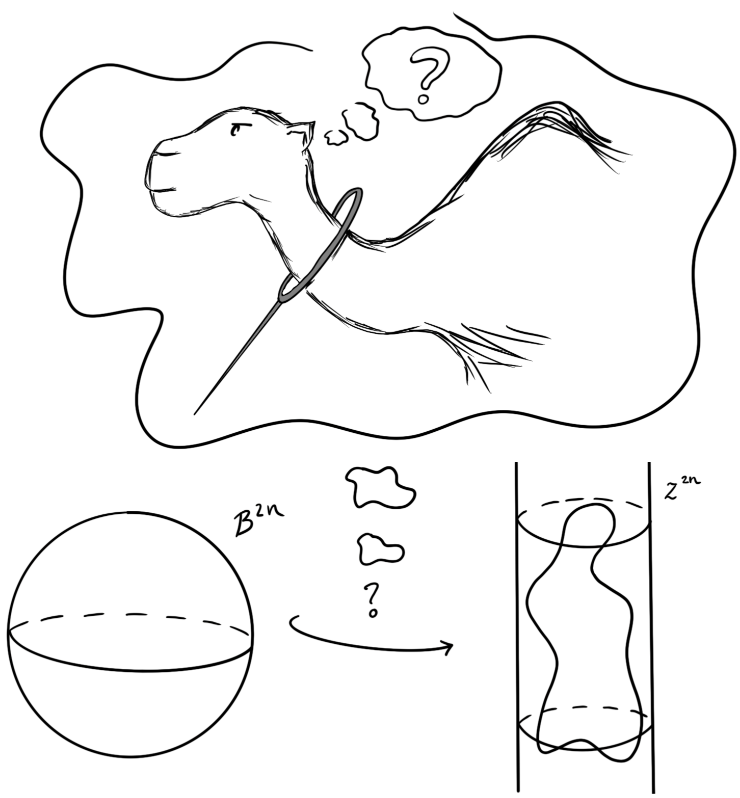

FIGURE 5. A symplectic camel through the eye of a symplectic needle.

Acknowledgments

I feel honored to thank my mentor, DeVon Ingram, for helping me open my way onto many worlds of math I didn't even know existed. I wouldn't have been able to learn anywhere near as much math as I did in one summer without his guidance. I also want to thank Professor May for organizing such an amazing program full of an insane amount of awesome mathematical ideas, and Professor Rudenko for the great lectures and fun problems from the apprentice program.

References

[1] McDuff, Dusa Salamon. Introduction to Symplectic Topology. Oxford University Press. 2017.
[2] John M. Lee. Introduction to Smooth Manifolds. Springer Graduate Texts in Mathematics. 2013.
[3] Loring W. Tu. An Introduction to Manifolds. Springer. 2010.
[4] Ana Cannas da Silva. Lectures on Symplectic Geometry. Springer. 2008.
[5] B. Aebischer, et.al. Symplectic Geometry. An Introduction based on the Seminar in Bern, 1992. Springer. 1994.
[6] Kathleen T. Alligood, et.al. Chaos. An Introduction to Dynamical Systems. Springer. 1997.
[7] Richard J. Brown. A modern introduction to Dynamical Systems. Oxford University Press. 2018.
[8] Sergei Treil. Linear Algebra Done Wrong. 2017.
[9] Jorge V. José, Eugene J. Saletan. Classical Dynamics. A contemporary approach. Cambridge University Press. 1998.