Lectures on Symplectic Geometry

Fraydoun Rezakhanlou
Departmet of Mathematics, UC Berkeley

December 12, 2015

1 Introduction

Hamiltonian systems appear in conservative problems of mechanics as in celestial mechanics but also in statistical mechanics governing the motion of particles and molecules in fluid. A mechanical system of N planets (particles) is modeled by a Hamiltonian function $H(x)$ where $x = (q, p)$, $q = (q_1, \dots, q_N)$, $p = (p_1, \dots, p_N)$ with $(q_i, p_i) \in \mathbb{R}^d \times \mathbb{R}^d$ being the position and the momentum of the i-th particle. The Hamiltonian's equations of motion are

$$\dot{q} = H_p(q,p), \;\; \dot{p} = -H_q(q,p), \tag{1.1}$$

which is of the form

$$\dot{x} = \bar{J}\nabla_x H(x), \quad \bar{J} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}, \tag{1.2}$$

where $n = dN$ and I_n denotes the $n \times n$ identity matrix. The equation (1.2) is an ODE that possesses a unique solution for every initial data x_0 provided that we make some standard assumptions on H. If we denote such a solution by $\phi_t(x_0) = \phi(t, x_0)$, then ϕ enjoys the group property

$$\phi_t \circ \phi_s = \phi_{t+s}, \quad t, s \in \mathbb{R}.$$

The ODE (1.1) is a system of $2n = 2dN$ unknowns. Such a typically large system can not be solved explicitly. A reduction of such a system is desirable and this can be achieved if we can find some conservation laws associated with our system. To find such conservation laws systematically, let us look at a general ODE of the form

$$\frac{dx}{dt} = b(x), \tag{1.3}$$

with the corresponding flow denoted by ϕ_t, and study $u(x,t) = T_t f(x) = f(\phi_t(x))$. A celebrated theorem of Liouville asserts that the function u satisfies

$$\frac{\partial u}{\partial t} = \mathcal{L}u = b \cdot u_x. \tag{1.4}$$

Recall that a function $f(x)$ is *conserved* if $\frac{d}{dt} f(\phi_t(x)) = 0$. From (1.4) we learn that a function f is conserved if and only if

$$b \cdot f_x = 0. \tag{1.5}$$

In the case of a Hamiltonian system, $b = (H_p, -H_q)$ and the equation (1.5) becomes

$$\{f, H\} := f_q \cdot H_p - f_p \cdot H_q = 0. \tag{1.6}$$

As an obvious choice, we may take $f = H$ in (1.6). In general, we may have other conservation laws that are not so obvious to be found. *Nother's principle* tells us how to find a conservation law using a symmetry of the ODE (1.3). With the aid of the symmetries, we may reduce our system to a simpler one that happens to be another Hamiltonian-type system.

Liouville discovered that for a Hamiltonian system of Nd-degrees of freedom ($2Nd$ unknowns) we only need Nd conserved functions in order to solve the system completely by means of quadratures. Such a system is called *completely integrable* and unfortunately hard to come by. Recently there has been a revival of the theory of completely integrable systems because of several infinite dimensional examples (Korteweg–deVries equation, nonlinear Schrödinger equation, etc.).

As we mentioned before, the conservation laws can be used to simplify a Hamiltonian system by reducing its size. To get more information about the solution trajectories, we may search for other conserved quantities. For example, imagine that we have a flow ϕ_t associated with (1.3) and we may wonder how the volume of $\phi_t(A)$ changes with time for a given measurable set A. For this, imagine that there exists a density function $\rho(x, t)$ such that

$$\int J(\phi_t(x))\rho^0(x)dx = \int J(x)\rho(x,t)dx \tag{1.7}$$

for every bounded continuous function J. This is equivalent to saying that for every nice set A,

$$\int_{\phi_{-t}(A)} \rho^0(x)dx = \int_A \rho(x,t)dx. \tag{1.8}$$

In words, the ρ^0-weighted volume of $\phi_{-t}(A)$ is given by the $\rho(\cdot, t)$-weighted volume of A. Using (1.4), it is not hard to see that in fact ρ satisfies the (dual) Liouville's equation

$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho b) = 0. \tag{1.9}$$

As a result, the measure $\rho^0(x)dx$ is invariant for the flow ϕ_t if and only if

$$\mathrm{div}(\rho^0 b) = 0.$$

In particular, if $\mathrm{div}\, b = 0$, then the Lebesgue measure is invariant. In the case of a Hamiltonian system $b = J\nabla H$, we do have $\mathrm{div}\, b = 0$, and as a consequence,

$$\mathrm{vol}(\phi_t(A)) = \mathrm{vol}(A), \tag{1.10}$$

for every measurable set A.

In our search for other invariance properties, let us now look for vector fields $F : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\frac{d}{dt} \int_{\phi_t(\gamma)} F \cdot dx \equiv 0 \tag{1.11}$$

for every closed curve γ. Such an invariance property is of interest in (for example) fluid mechanics because $\int_\gamma F \cdot dx$ measures the *circulation* of the velocity field F around γ. To calculate the left-hand side of (1.11), observe

$$\int_{\phi_t(\gamma)} F \cdot dx = \int_\gamma F(\phi_t(x)) D\phi_t(x) \cdot dx$$

where $D\phi_t$ denotes the derivative of ϕ_t in x and we regard F as a row vector. Set $F = (F^1, \dots, F^k)$, $\phi = (\phi^1, \dots, \phi^k)$, $u = (u^1, \dots, u^k)$, $u(x,t) = T_t F(x) = F \circ \phi_t(x) D\phi_t(x)$, so that

$$u^j(x,t) = \sum_i F^i(\phi_t(x)) \frac{\partial \phi_t^i}{\partial x_j}(x).$$

To calculate the time derivative, we write

$$\begin{aligned} u(x, t+h) &= F(\phi_{t+h}) D\phi_{t+h} = F(\phi_t \circ \phi_h)\ D\phi_t \circ \phi_h D\ \phi_h \\ &= u(\phi_h(x), t) D\phi_h(x), \end{aligned}$$

so that

$$\begin{aligned} \frac{d}{dh} u^j(x, t+h)\bigg|_{h=0} &= \sum_i \left[(\nabla u^i \cdot b)\delta_{ij} + u^i \frac{\partial b^i}{\partial x_j} \right] \\ &= \nabla u^j \cdot b + \sum_i u^i \frac{\partial b^i}{\partial x_j} \\ &= (u \cdot b)_{x_j} + \sum_i (u^j_{x_i} - u^i_{x_j}) b^i. \end{aligned}$$

In summary,

$$\frac{\partial u}{\partial t} = \nabla(u \cdot b) + \mathcal{C}(u) b \tag{1.12}$$

where $\mathcal{C}(u)$ is the matrix $[u^i_{x_j} - u^j_{x_i}]$. In particular,

$$\frac{d}{dt} \int_{\phi_t(\gamma)} F \cdot dx = \int_\gamma \mathcal{C}(u) b \cdot dx \tag{1.13}$$

for every closed curve γ. Recall that we would like to find vector fields F for which (1.11) is valid. For this it suffices to have $\mathcal{C}(F)b$ a gradient. Indeed if $\mathcal{C}(F)b$ is a gradient, then

$$\begin{aligned}\frac{d}{dt}\int_{\phi_t(\gamma)} F\cdot dx &= \frac{d}{dh}\int_{\phi_{t+h}(\gamma)} F\cdot dx\bigg|_{h=0}\\ &= \frac{d}{dh}\int_{\phi_h(\phi_t(\gamma))} F\cdot dx\bigg|_{h=0}\\ &= \int_{\phi_t(\gamma)} \mathcal{C}(F)b\cdot dx = 0.\end{aligned}$$

Let us examine this for some examples.

Example 1.1. (i) Assume that $k = 2n$ with $x = (q,p) = (q_1,\dots,q_n,p_1,\dots,p_n)$. Let $b(q,p) = (H_p,-H_q)^* = J\nabla H$ for a Hamiltonian $H(q,p)$. Choose $F(q,p) = (p,0)$. We then have $\mathcal{C}(F) = \begin{bmatrix} 0 & I_n \\ -I_n & 0\end{bmatrix} = J$, and $\mathcal{C}(F)b = JJ\nabla H = -\nabla H$. This and (1.13) imply that for a Hamiltonian flow ϕ_t and closed γ,

$$\frac{d}{dt}\int_{\phi_t(\gamma)} p\cdot dq = 0, \tag{1.14}$$

which was discovered by Poincaré originally.

(ii) Assume $k = 2n+1$ with $x = (q,p,t)$ and $b(q,p,t) = (H_p,-H_q,1)^*$ where H is now a time-dependent Hamiltonian function. Define

$$F(q,p,t) = (p,0,-H(q,p,t)).$$

We then have

$$\mathcal{C}(F) = \left[\begin{array}{c|c|c} 0 & I_n & H_q^* \\ \hline -I_n & 0 & H_p^* \\ \hline -H_q & -H_p & 0\end{array}\right] = \left[\begin{array}{c|c} J & \nabla H^t \\ \hline -\nabla H & 0\end{array}\right]. \tag{1.15}$$

Since $\mathcal{C}(F)b = 0$, we deduce that for any closed (q,p,t)-curve γ,

$$\frac{d}{ds}\int_{\phi_s(\gamma)} (q\cdot dp - H(q,p,t)dt) = 0, \tag{1.16}$$

proving a result of *Poincaré* and *Cartan*. Note that if γ has no t-component in (1.16), then (1.16) becomes (1.14).

(iii) Assume $n = 3$. Then $\mathcal{C}(F) = \begin{bmatrix} 0 & -\alpha_3 & \alpha_2 \\ \alpha_3 & 0 & -\alpha_1 \\ -\alpha_2 & \alpha_1 & 0 \end{bmatrix}$ with $(\alpha_1, \alpha_2, \alpha_3) = \nabla \times F$. Now if $b = \nabla \times F$, then $\mathcal{C}(F)b = (\nabla \times F) \times b$ and $\frac{d}{dt}\int_{\phi_t(\gamma)} F \cdot dx = 0$. In words, the F-circulation of a curve moving with velocity field $\nabla \times F$ is preserved with time. □

Example 1.2. A Hamiltonian system (1.2) simplifies if we can find a function $w(q,t)$ such that $p(t) = w(q(t),t)$. If such a function w exists, then $q(t)$ solves

$$\frac{dq}{dt} = H_p(q, w(q,t), t). \tag{1.17}$$

The equation $\dot{p}$ gives us the necessary condition for the function w:

$$\begin{aligned} \dot{p} &= w_q \dot{q} + w_t = w_q \cdot H_p(q, w, t) + w_t, \\ \dot{p} &= -H_q(q, w, t). \end{aligned}$$

Hence $w(q,t)$ must solve,

$$w_t + w_q \cdot H_p(q, w, t) + H_q(q, w, t) = 0. \tag{1.18}$$

For example, if $H(q,p,t) = \frac{1}{2}|p|^2 + V(q,t)$, then (1.18) becomes

$$w_t + w_q w + V_q(q,t) = 0. \tag{1.19}$$

The equation (1.17) simplifies to

$$\frac{dq}{dt} = w(q,t) \tag{1.20}$$

in this case. If the flow of (1.20) is denoted by ψ_t, then $\phi_t(q, w(q,0)) = (\psi_t(q), w(\psi_t(q), t))$. Now (1.14) means that for any closed q-curve η,

$$\frac{d}{dt}\int_{\psi_t(\eta)} w(q,t) \cdot dq = 0.$$

This is the celebrated *Kelvin's circulation theorem.* □

We may use Stokes' theorem to rewrite (1.14) as

$$\frac{d}{dt}\int_{\phi_t(\Gamma)} \bar{\omega} := \frac{d}{dt}\int_{\phi_t(\Gamma)} dp \wedge dq = 0 \tag{1.21}$$

for every two-dimensional surface Γ. In words, the 2-form $\bar{\omega}$ is invariant under the Hamiltonian flow ϕ_t. In summary, we have found various invariance principles for Hamiltonian flows:

- The conserved functions f satisfying (1.6) is an example of an invariance principle for 0-forms.
- The Liouville's theorem (1.10) is an example of an invariance principle of an n-form.
- Poincaré's theorem (1.21) is an instance of an invariance principle involving a 2-form.

In fact (1.21) implies (1.10) because the invariance of $\bar{\omega}$ implies the invariance of the $k = 2n$ form $\bar{\omega}^n = \bar{\omega} \wedge \cdots \wedge \bar{\omega}$ which is a constant multiple of the volume form. More generally, we may take an arbitrary l-form ω and evolve it by the flow ϕ_t of a velocity field b. If we write $\omega(t)$ for $\phi_t^*\omega$:

$$\int_\Gamma \omega(t) = \int_\Gamma \phi_t^*\omega = \int_{\phi_t(\Gamma)} \omega,$$

then by a formula of Cartan,

$$\frac{d\omega}{dt} = \mathcal{L}_b\omega := d(i_b\omega) + i_b(d\omega)$$

where $\mathcal{L}_b\omega$ denotes the *Lie derivative*.

The configuration space of a system with constraints is a manifold. Also, when we use conservation laws to reduce our Hamiltonian system, we obtain a Hamiltonian system on a manifold. If the configuration space is a n-dimensional differentiable manifold N, and $L : TN \to \mathbb{R}$ is a differentiable *Lagrangian function*, then $p = \frac{\partial L}{\partial \dot{q}}$ is a cotangent vector. The cotangent bundle $M = T^*N$ is an example of a *symplectic manifold* because it possesses a natural closed non-degenerate form $\bar{\omega}$ which is simply $\sum_1^n dp_i \wedge dq_i$, in local coordinates. More generally we may study an even dimensional manifold M, equipped with a non-degenerate closed 2-form ω, and construct vector fields X_H associated with scalar functions H such that $i_{X_H}(\omega) = -dH$. The vector field X_H is the analog of $J\nabla H$ in the Euclidean case $M = \mathbb{R}^{2n}$. By the non-degeneracy of ω, such X_H exists for every differentiable Hamiltonian function H.

A celebrated theorem of Darboux asserts that any symplectic manifold is locally equivalent to a Euclidean space with its standard symplectic structure. As a result, the most important questions in symplectic geometry are the global ones.

Consider the Euclidean space $(\mathbb{R}^{2n}, \bar{\omega})$. If the hypersurface $\Gamma = H^{-1}(c)$ is a compact energy level set with $\nabla H \neq 0$ on Γ, then the unparametrized orbits on Γ of the Hamiltonian vector field $X_H = J\nabla H$ are independent of the choice of H. One can therefore wonder what hypersurfaces carry a periodic orbit. P. Rabinowitz showed that every star-like hypersurface carries a periodic orbit. Later, Viterbo showed that the same holds more generally for hypersurfaces of *contact type*, establishing affirmatively a conjecture of A. Weinstein.

Consider two compact connected domains U_1 and U_2 in $\mathbb{R}^n$ with smooth boundaries. U_1 and U_2 are diffeomorphic and volume(U_1) = volume(U_2), then we can find a diffeomorphism between U_1 and U_2 that is also volume preserving (DaCorogna–Moser). We may wonder whether or not there exists a symplectic diffeomorphism between U_1 and U_2. *Gromov's squeezing theorem* shows that the symplectic transformations are more rigid; if there exists a symplectic embedding from the ball

$$B_R(0) = \{(q,p) : |q|^2 + |p|^2 < R^2\}$$

into the cylinder

$$Z_r(0) = \{(q,p) : q_1^2 + p_1^2 < r^2\},$$

then we must have $r \geq R$! Motivated by this, Gromov defines the *symplectic radius* $r(M)$ of a symplectic manifold (M, ω) as the largest r for which there exists a symplectic embedding from $B_r(0)$ into M. The Gromov's radius is an example of a *symplectic capacity* that is a symplectic invariant. Since the discovery of the Gromov radius, new capacities have been discovered. The existence of some of these capacities can be used to prove various global properties of Hamiltonian systems such as Viterbo's existence of periodic orbits.

Another rigidity of symplectic transformation is illustrated in an important result of Eliashberg and Gromov: If $\{f_m\}$ is a sequence of symplectic transformation that converges uniformly to a differentiable function f, then f is also symplectic. The striking aspect of this result is that our definition of a symplectic function f involves the first derivative of f. As a result, we should expect to have a definition of symplicity that does not involve any derivative. This should be compared to the definition of a volume preserving transformation that can be formulated with or without using derivative.

2 Quadratic Hamiltonians and Linear Symplectic Geometry

In this section, we discuss several central concepts and fundamental results of symplectic geometry in linear setting. More specifically, we establish Darboux Theorem for *symplectic vector spaces*, define *symplectic spectrum* for quadratic Hamiltonian functions, construct *linear symplectic capacities* for ellipsoids, establish *symplectic rigidity* for linear symplectic maps, and analyze *complex structures* that are compatible with a symplectic form.

A symplective vector space (V, ω) is a pair of finite dimensional real vector space V and a bilinear form $\omega : V \times V \to \mathbb{R}$ which is antisymmetric and non-degenerate. That is, $\omega(a, b) = -\omega(b, a)$ for all $a, b \in V$, and that $\forall a \in V$ with $a \neq 0$, $\exists b \in V$ such that $\omega(a, b) \neq 0$. The non-degeneracy is equivalent to saying that the transformation $a \mapsto \omega(a, \cdot)$ is a linear isomorphism between V and its dual V^*. Clearly $(\mathbb{R}^k, \bar{\omega})$ is an example of a symplectic vector space when $k = 2n$ and $\omega(a, b) = \bar{J}a \cdot b$, where $\bar{J}$ was defined in (1.2). More generally, given a k by k matrix C, the bilinear form $\omega(a, b) = Ca \cdot b$ is symplectic if C is invertible and skew-symmetric. Note that since

$$\det C = \det C^* = \det(-C) = (-1)^n \det C,$$

necessarily $k = 2n$ is even. Given a symplectic (V, ω), then we say a and b are *ω-orthogonal* and write $a \amalg b$ if $\omega(a, b) = 0$. If W is a linear subspace of V, then

$$W^{\amalg} = \{a \in V : a \amalg W\}.$$

In our first result we state a linear version of Darboux's theorem and some elementary facts about symplectic vector spaces. Darboux's theorem in Euclidean setting asserts that for every invertible skew-symmetric matrix C we can find an invertible matrix T such that $T^* C T = \bar{J}$.

Proposition 2.1 *Let (V, ω) be a symplectic linear space of dimension $k = 2n$ and W be a subspace of V.*

- **(i)** $\dim W + \dim W^{\amalg} = \dim V$.
- **(ii)** $\left(W^{\amalg}\right)^{\amalg} = W$.
- **(iii)** (W, ω) *is symplectic iff* $W \oplus W^{\amalg} = V$.
- **(iv)** *If W is a symplectic subspace, then $W^{\amalg}$ is also symplectic.*
- **(v)** *There exists a basis $e_1, \dots, e_n, f_1, \dots, f_n$ such that $\omega(e_i, e_j) = \omega(f_i, f_j) = 0$ and $\omega(f_i, e_j) = \delta_{ij}$. Equivalently, if $x = (q_1, \dots, q_n, p_1, \dots, p_n)$, $y = (q_1', \dots, q_n', p_1', \dots, p_n')$, $a = \sum_1^n q_j e_j + p_j f_j$, $a' = \sum_1^n q_j' e_j + p_j' f_j$, then $\omega(a, a') = \bar{\omega}(x, y)$.*

Proof. **(i)** Assume that $\dim V = n$ and $\dim W = m$. Choose a basis $\{a_1, \dots, a_m\}$ for W. Then by non-degeneracy $a_1^*, \dots, a_m^*$ are independent where $a_{\mathcal{I}}(b) = \omega(a_j, b)$. Since $W^{\mathrm{II}} = \{a : a_{\mathcal{I}}(a) = 0 \text{ for } j = 1, \dots, m\}$, with $a_{\mathcal{I}}$ independent, we have that $\dim W^{\mathrm{II}} = n - m$.

(ii) Evidently $W \subseteq (W^{\mathrm{II}})^{\mathrm{II}}$. Since $\dim W + \dim W^{\mathrm{II}} = \dim W^{\mathrm{II}} + \dim (W^{\mathrm{II}})^{\mathrm{II}}$, we deduce that $W = (W^{\mathrm{II}})^{\mathrm{II}}$.

(iii) By definition, (W, ω) is symplectic iff $W \cap W^{\mathrm{II}} = \{0\}$. Since $\dim W + \dim W^{\mathrm{II}} = \dim V$, we have that $W \oplus W^{\mathrm{II}} = V$.

(iv) If W is symplectic, then $V = W \oplus W^{\mathrm{II}} = (W^{\mathrm{II}})^{\mathrm{II}} \oplus W^{\mathrm{II}}$, which implies that in fact W^{II} is symplectic.

(v) Evidently $\dim V \geq 2$. Let e_1 be a non-zero vector of V. Since ω is non-degenerate, we can find $f_1 \in V$ such that $\omega(f_1, e_1) = 1$. Clearly f_1, e_1 are linearly independent. Let $V_1 = \operatorname{span}\{e_1, f_1\}$. If $V = V_1$, then we are done. Otherwise $V = V_1 \oplus V_1^{\mathrm{II}}$ with both (V_1, ω), $(V_1^{\mathrm{II}}, \omega)$ symplectic. Now we repeat the previous argument to find f_2, e_2 etc. $\square$

We now turn our attention to quadratic Hamiltonian functions and ellipsoids. By a quadratic Hamiltonian we mean a function $H(x) = \frac{1}{2} Bx \cdot x$ for a symmetric matrix B. We are particularly interested in the case $B \geq 0$. We note that for such quadratic Hamiltonians, the corresponding Hamiltonian vector field $X(x) = JBx$ is linear. Since the flow of

$$\dot{x} = JBx, \tag{2.1}$$

preserves H, we also study the level sets of nonnegative quadratic functions. By an *ellipsoid* we mean a set E of the form

$$E = \{x : H(x) \leq 1\}$$

where $H(x) = \frac{1}{2} Bx \cdot x$ with $B \geq 0$. Note that if $B > 0$, the ellipsoid E is a bounded set. Otherwise, the set E is unbounded and may be also called a *cylinder* or *cylindrical ellipsoid*. Our goal is to show that we can make a change of coordinates to turn the ODE to a simpler Hamiltonian system for which B is a diagonal matrix. Before embarking on this, let us first review some well-known facts about symmetric matrices, which is the symmetric counterpart of what we will discuss for symplectic matrices.

To begin, let us recall that the standard Euclidean inner product is preserved by a matrix A if A is orthogonal. That is

$$Aa \cdot Ab = a \cdot b \ \text{ for all } a, b \in \mathbb{R}^k \Leftrightarrow A^{-1} = A^*.$$

Let us write $O(k)$ for the space of $k \times k$ orthogonal matrices. We also write $S(k)$ for the space of symmetric matrices. A quadratic function $H : \mathbb{R}^k \to \mathbb{R}$ is defined by $H(x) = \frac{1}{2} Bx \cdot x$ with $B \in S(k)$.

Proposition 2.2 *Let H_1 and H_2 be two quadratic functions associated with the symmetric matrices B_1 and B_2. Then there exists $A \in O(k)$ such that $H_1 \circ A = H_2$ if and only if B_1 and B_2 have the same spectrum.*

As a consequence, if $H(x) = \frac{1}{2}Bx \cdot x$ and B has eigenvalues $\boldsymbol{\lambda}(H) = (\lambda_1, \dots, \lambda_k)$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$, then there exists $A \in O(k)$ such that $H(Ax) = \frac{1}{2}\sum_{j=1}^{k} \lambda_j x_j^2$. In particular, if $B \geq 0$, then λ_j's are nonnegative and we may define *radii* $\mathbf{R}(H) = (R_1(H), \dots, R_k(H)) \in (0, \infty]^k$ by $R_i^2 = R_i^2(H) = \frac{2}{\lambda_j}$ so that $0 < R_1(H) \leq R_2(H) \leq \dots \leq R_k(H)$ and

$$H(A(x)) = \sum_{j=1}^{k} \frac{x_j^2}{R_j^2}.$$

If E is the corresponding ellipsoid,

$$E = \{x : H(x) \leq 1\},$$

then we write

$$\mathbf{R}(E) = (R_1(E), \dots, R_k(E))$$

for $\mathbf{R}(H)$ and refer to its coordinates as the radii of E. We now rephrase Proposition 2.1 as

Corollary 2.1 *Let E_1 and E_2 be two ellipsoids. Then there exists $A \in O(k)$ such that $A(E_1) = E_2$ if and only if $\mathbf{R}(E_1) = \mathbf{R}(E_2)$.*

We next discuss the monotonicity of $\mathbf{R}$.

Proposition 2.3 **(i)** *Let H_1 and H_2 be two quadratic functions. Then $H_1 \circ A \leq H_2$ for some $A \in O(k)$ if and only if $\boldsymbol{\lambda}(H_1) \leq \boldsymbol{\lambda}(H_2)$.*

(ii) *Let E_1 and E_2 be two ellipsoids. Then $A(E_2) \subseteq E_1$ for some $A \in O(k)$ if and only if $\mathbf{R}(E_2) \leq \mathbf{R}(E_1)$.*

Proof We note that **(i)** implies **(ii)** because if $E_r = \{x : H_r(x) \leq 1\}$ for $r = 1$ and 2, then

$$E_2 \subseteq A^{-1}E_1 \Leftrightarrow H_1 \circ A \leq H_2.$$

As for the proof of **(i)**, observe that if $\boldsymbol{\lambda} = \boldsymbol{\lambda}(H_1) \leq \boldsymbol{\lambda}' = \boldsymbol{\lambda}(H_2)$, then we can find A_1 and $A_2 \in O(k)$ such that

$$H_1(A_1 x) = \frac{1}{2}\sum_{j=1}^{k} \lambda_j x_j^2 \leq \frac{1}{2}\sum_{j=1}^{k} \lambda'_j x_j^2 = H_2(A_2 x),$$

proving the "if" part of **(i)**. The "only if" is an immediate consequence of Courant–Hilbert Minimax Principle that will be stated in Lemma 2.1 below. $\square$

Lemma 2.1 *(Courant–Hilbert). Let $B \in S(k)$ with eigenvalues $\mu_1 \le \mu_2 \le \cdots \le \mu_k$. Then*

$$\mu_j = \inf_{\dim V = j} \sup_{x \in V \setminus \{0\}} \frac{Bx \cdot x}{|x|^2}, \tag{2.2}$$

$$\mu_j = \sup_{\dim V = j-1} \inf_{x \in V^\perp \setminus \{0\}} \frac{Bx \cdot x}{|x|^2}, \tag{2.3}$$

where V denotes a linear subspace of $\mathbb{R}^k$.

Proof Let us write X for the right-hand side of (2.2). Let $u_1, u_2, \dots, u_k$ be an orthonormal basis with $Bu_j = \mu_j u_j$, $j = 1, \dots, k$. Note that

$$\sup \left\{ \frac{Bx \cdot x}{|x|^2} : x \in \operatorname{span}\{u_1, \dots, u_j\}, x \neq 0 \right\} = \sup_{c_1, \dots, c_j} \frac{\sum_1^j \mu_l c_l^2}{\sum_1^j c_l^2} \le \mu_j,$$

proving $X \le \mu_j$. For $X \ge \mu_j$, pick a linear subspace V of dimension j and choose non-zero $x \in V$ such that $x \perp u_1, \dots, u_{j-1}$. Such x exists because $\dim V = j$ and we are imposing $j-1$ many conditions. Since we can write $x = \sum_{l=j}^k c_l u_l$, we have

$$\frac{Bx \cdot x}{|x|^2} = \frac{\sum_j^k \mu_l c_l^2}{\sum_j^k c_l^2} \ge \mu_j. \tag{2.4}$$

As a result, $X \ge \mu_j$ and this completes the proof of (2.2).

As for (2.3), note that if $x \perp u_1, \dots, u_{j-1}$, $x \neq 0$, then $\frac{Bx \cdot x}{|x|^2} \ge \mu_j$ by (2.4). Hence, if Y denotes the right-hand side of (2.3), then $Y \ge \mu_j$ by choosing $V = \operatorname{span}\{u_1, \dots, u_{j-1}\}$. For $\mu_j \ge Y$, let V be any linear space of dimension $j-1$ and pick a non-zero $x \in \operatorname{span}\{u_1, \dots, u_j\} \cap V^\perp$. For such a vector x we have $x = \sum_1^j c_l u_l$, $(c_1, \dots, c_j) \neq 0$, and $\frac{Bx \cdot x}{|x|^2} \le \mu_j$. This implies that $\mu_j \ge Y$. $\square$

We would like to develop a theory similar to what we have seen in this section but now for the bilinear form $\bar{\omega}$. The following table summarizes our main results:

	Symmetric	Antisymmetric
Form	$a \cdot b$	$\bar{\omega}(a, b) = Ja \cdot b$
Invariant matrix	$A \in O(k) : A^{-1} = A^*$	$T \in Sp(n) : T^{-1} = -\bar{J}T^*\bar{J}$
Vector field	$\nabla H(x) = Bx,\ B \in S(k)$	$\bar{J}\nabla H(x) = \bar{J}Bx;\ \bar{J}B \in \operatorname{Ham}(n)$
Spectral theorem	Proposition 2.2	Weirstrass Theorem (Theorem 2.1)
Monotonicity	Courant–Hilbert Minimax (Lemma 2.1)	Theorem 2.2, Lemma 2.2

We say a matrix T is *symplectic* if $\bar\omega(Ta, Tb) = \bar\omega(a, b)$. Equivalently $T^*\bar JT = \bar J$ or $T^{-1} = -\bar JT^*\bar J$. The set of $2n \times 2n$ symplectic matrices is denoted by $Sp(n)$. We say a matrix C is Hamiltonian if $C = \bar JB$ for a symmetric matrix B. The space of $2n \times 2n$ Hamiltonian matrices is denoted by $\mathrm{Ham}(n)$. We have

$$C \in \mathrm{Ham}(n) \Leftrightarrow \bar JC + C^*\bar J = 0 \Leftrightarrow C^* = \bar JC\bar J.$$

We note that if $H(x) = \frac{1}{2}Bx \cdot x$ with $B \in S(2n)$, then $J\nabla H(x) = \bar JBx$ with $\bar JB \in \mathrm{Ham}(n)$.

We are now ready to state Weirstrass Theorem which allows us to diagonalize a Hamiltonian matrix using a symplectic change of variable.

Theorem 2.1 *Let B be a positive matrix. Then the matrix $C = \bar JB$ has purely imaginary eigenvalues of the form $\pm i\lambda_1, \dots, \pm i\lambda_n$ with $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. Moreover there exists $T \in Sp(n)$ such that the quadratic Hamiltonian function $H(x) = \frac{1}{2}Bx \cdot x$ can be represented as*

$$H \circ T(x) = \sum_{j=1}^{n} \frac{\lambda_j}{2}(q_j^2 + p_j^2),$$

where $x = (q_1, p_1, \dots, q_n, p_n)$.

Proof *Step 1.* Let $\mu + i\lambda$ be an eigenvalue of C associated with the (non-zero) eigenvector $a + ib$. As a result,

$$Ca = \mu a - \lambda b, \quad Cb = \lambda a + \mu b \quad \text{or} \quad Ba = \lambda Jb - \mu Ja, \quad Bb = -\lambda Ja - \mu Jb.$$

Hence

$$Ba \cdot b = \mu\bar\omega(b, a), \quad Bb \cdot a = -\mu\bar\omega(b, a), \quad Ba \cdot a = Bb \cdot b = \lambda\bar\omega(b, a).$$

From this, $B = B^*$, $\mu + i\lambda \neq 0$, and $a + ib \neq 0$, we deduce that $\mu = 0$, $\bar\omega(b, a) \neq 0$, and $Ba \cdot b = 0$. As a result, for every nonzero eigenvalue $i\lambda$, we can find an eigenvector $a + ib$, such that

$$\begin{aligned} &Ba \cdot a = Bb \cdot b = \lambda, \quad Ba \cdot b = 0, \\ &Ca = -\lambda b, \quad Cb = \lambda a. \end{aligned} \tag{2.5}$$

Step 2. If $\lambda_1 = 0$, then all eigenvalues are 0 and there is nothing to prove. If $\lambda_1 \neq 0$, we use Step 1 and (2.5) find a_1 and b_1 such that

$$Ba_1 \cdot a_1 = Bb_1 \cdot b_1 = \lambda, \quad Ba_1 \cdot b_1 = 0, \quad Ca = -\lambda b, \quad Cb = \lambda a.$$

As a result, if $V_1 = \text{span}\{a_1, b_1\}$, then $CV_1 \subseteq V_1$, and for q_1 and $p_1 \in \mathbb{R}$,

$$H(q_1 a_1 + p_1 b_1) = q_1^2 H(a_1) + p_1^2 H(b_1) + q_1 p_1 B a_1 \cdot b_1 = \frac{\lambda_1}{2}(q_1^2 + p_1^2). \tag{2.6}$$

By Proposition 2.1, the spaces V_1 and V_1^{II} are symplectic and $\mathbb{R}^{2n} = V_1 \oplus V_1^{\text{II}}$. We now claim

$$a \in V_1,\ b \in V_1^{\text{II}} \Rightarrow Ba \cdot b = 0, \quad \textit{and} \quad CV_1^{\text{II}} \subseteq V_1^{\text{II}}. \tag{2.7}$$

Indeed if $a \in V_1$, $b \in V_1^{\text{II}}$, then $Ca \in V_1$, and

$$\bar{\omega}(Cb, a) = \bar{J}Cb \cdot a = -Bb \cdot a = -b \cdot Ba = -\bar{J}b \cdot Ca = -\bar{\omega}(b, Ca) = 0,$$

which proves both claims in (2.7) because we just showed that $Cb \in V_1^{\text{II}}$.

Final Step. From (2.7) we learn that if $a \in V_1$ and $b \in V_1^{\text{II}}$, then

$$H(a+b) = H(a) + H(b).$$

Let us look at the restriction of H to the symplectic vector space $(V_1^{\text{II}}, \bar{\omega})$. By Proposition 2.1, this pair is isomorphic with $(\mathbb{R}^{2n-2}, \bar{\omega})$. As a result, we may repeat the above argument to assert that there exits a pair of vectors $a_2, b_2 \in V_1^{\text{II}}$ with $\bar{\omega}(b_2, a_2) = -1$, $H(a_2) = H(b_2) = r_2^{-2}$ and $Ba_2 \cdot b_2 = 0$. Continuing this process would yield a basis $(a_1, b_1, a_2, b_2, \dots, a_n, b_n)$ such that

$$\begin{aligned} &\bar{\omega}(a_i, a_j) = \bar{\omega}(b_i, b_j) = 0, \quad \bar{\omega}(b_i, a_j) = \delta_{ij}, \\ &H(a_j) = H(b_j) = r_j^{-2},\ Ba_j \cdot b_j = 0. \end{aligned}$$

From this we learn that the linear map $T : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ defined by

$$T(q_1, p_1, \dots, q_n, p_n) = \sum_1^n q_j a_j + p_j b_j$$

is symplectic, and

$$H(T(x)) = \sum_1^n \frac{\lambda_j}{2}(q_j^2 + p_j^2).$$

□

Given $H(x) = \frac{1}{2}Bx \cdot x$ with $B \geq 0$, let us write $\frac{1}{2}\lambda_j = \frac{1}{r_j^2}$ so that $r_j = r_j(H) \in (0, \infty]$ satisfy

$$0 < r_1(H) \leq r_2(H) \leq \cdots \leq r_n(H).$$

We also write $\mathbf{r}(H) = (r_1(H), \dots, r_n(H))$ and if E is the corresponding ellipsoid, we write $\mathbf{r}(E)$ for $\mathbf{r}(H)$. We may rephrase Theorem 2.1 as follows:

Corollary 2.2 **(i)** *If H_1 and H_2 are two positive definite quadratic forms, then $\mathbf{r}(H_1) = \mathbf{r}(H_2)$ if and only if $H_2 = H_1 \circ T$ for some $T \in Sp(n)$.*

(ii) *Let E_1 and E_2 be two ellipsoid. Then $T(E_2) = E_1$ for $T \in Sp(n)$ if and only if $\mathbf{r}(E_1) = \mathbf{r}(E_2)$.*

Example 2.1 Let $n = 1$ and $H(q_1, p_1) = \frac{q_1^2}{R_1^2} + \frac{p_1^2}{R_2^2}$ so that $\mathbf{R}(H) = (R_1, R_2)$. Here $H(x) = \frac{1}{2}Bx \cdot x$ with

$$B = \begin{bmatrix} \frac{2}{R_1^2} & 0 \\ 0 & \frac{2}{R_2^2} \end{bmatrix}.$$

We have

$$C = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} B = \begin{bmatrix} 0 & \frac{2}{R_2^2} \\ -\frac{2}{R_1^2} & 0 \end{bmatrix}.$$

The matrix C has eigenvalues $\pm i\frac{2}{R_1 R_2}$. Hence $\mathbf{r}(H) = (r_1(H))$ with $r_1(H) = \sqrt{R_1 R_2}$ and there exists $T \in Sp(n)$ such that $H \circ T(q_1, p_1) = \frac{q_1^2 + p_1^2}{R_1 R_2}$. □

As our next corollary to Theorem 2.1, we solve (2.1) with the aid of a symplectic change of coordinates:

Corollary 2.3 *Let H, T and $(\lambda_1, \dots, \lambda_n)$ be as in Theorem 2.1. Let $x(t)$ be a solution of (2.1) and define $y(t) = T^{-1}x(t)$. Then $\dot{y} = \bar{J}B_0 y$, where B_0 is a diagonal matrix that has the entries*

$$\lambda_1, \dots, \lambda_n, \lambda_1, \dots, \lambda_n,$$

on its main diagonal.

Proof From (2.1), we learn that $T\dot{y} = \bar{J}BTy$. Also, by Theorem 2.1 we know that $(BTx) \cdot (Tx) = B_0 x \cdot x$, which means that $T^*BT = B_0$. As a result

$$\dot{y} = T^{-1}\bar{J}BTy = -\bar{J}T^*\bar{J}\bar{J}BTy = \bar{J}T^*BT = \bar{J}B_0.$$

□

Remark 2.1 If we write $\phi_t(y)$ for the flow of $\dot{y} = \bar{J}B_0 y$ and use the complex notation $y = (z_1, \dots, z_n)$ with $z_j = q_j + ip_j$, then

$$\phi_t(z_1, \dots, z_n) = (e^{-i\lambda_1 t}z_1, \dots, e^{-i\lambda_n t}z_n).$$

In particular, for each $z_j \neq 0$, $\lambda_j \neq 0$, if we set $\hat{z}_j$ for the complex vector that has z_j for the j-coordinate and 0 for the other coordinates, then the orbit $(\phi_t(\hat{z}_j) : t \in \mathbb{R})$ is periodic of period $2\pi/\lambda_j = \pi r_j^2$. □

We now turn to the question of monotonicity.

Theorem 2.2 **(i)** *If H_1 and H_2 are two non-negative definite quadratic forms, then $\mathbf{r}(H_1) \geq \mathbf{r}(H_2)$ if and only if there exists $T \in Sp(n)$ such that $H_1 \circ T \leq H_2$.*

(ii) *Let E_1 and E_2 be two ellipsoids. Then $\mathbf{r}(E_2) \leq \mathbf{r}(E_1)$ if and only if $T(E_2) \subseteq E_1$ for some $T \in Sp(n)$.*

Proof As before **(i)** implies **(ii)**. By approximation, it suffices to establish **(i)** when H_1 and H_2 are positive definite. In this case, **(ii)** is an immediate consequence of a variational formula we obtain for $r_j(H)$ in Lemma 2.2 below. □

Lemma 2.2 *Let H be a positive definite quadratic function of $\mathbb{R}^{2n}$. Then*

$$\frac{1}{2} r_j^2(H) = \inf_{\dim V = 2n+2j} \sup_{[x,y]^* \in V \setminus \{0\}} \frac{\bar{\omega}(x,y)^+}{H(x)+H(y)}. \tag{2.8}$$

Here V is for linear subspace of $\mathbb{R}^{4n}$.

Proof Recall that $\pm i 2 r_j^{-2}$ are the eigenvalues of $C = JB$ where $H(x) = \frac{1}{2} Bx \cdot x$. Hence $\pm \frac{i}{2} r_j^2$ are the eigenvalues of C^{-1}. If $a_j + i b_j$ denotes the corresponding eigenvector, then $C^{-1}(a_j + i b_j) = \frac{i}{2} r_j^2 (a_j + i b_j)$. This means

$$C^{-1} a_j = -\frac{r_j^2}{2} b_j, \; C^{-1} b_j = \frac{r_j^2}{2} a_j. \tag{2.9}$$

This suggests looking at the $4n \times 4n$ matrix

$$D = \begin{bmatrix} 0 & C^{-1} \\ -C^{-1} & 0 \end{bmatrix}.$$

From (2.9) we readily deduce

$$D \begin{bmatrix} a_j \\ b_j \end{bmatrix} = \frac{r_j^2}{2} \begin{bmatrix} a_j \\ b_j \end{bmatrix}, \; D \begin{bmatrix} b_j \\ -a_j \end{bmatrix} = \frac{r_j^2}{2} \begin{bmatrix} b_j \\ -a_j \end{bmatrix},$$
$$D \begin{bmatrix} b_j \\ a_j \end{bmatrix} = -\frac{r_j^2}{2} \begin{bmatrix} b_j \\ a_j \end{bmatrix}, \; D \begin{bmatrix} -a_j \\ b_j \end{bmatrix} = -\frac{r_j^2}{2} \begin{bmatrix} -a_j \\ b_j \end{bmatrix}.$$

Note that since $a_j + i b_j \neq 0$, the vectors $\begin{bmatrix} a_j \\ b_j \end{bmatrix}, \begin{bmatrix} b_j \\ -a_j \end{bmatrix}$ are linearly independent. Hence $\pm \frac{i}{2} r_j^2$ produces eigenvalue $\pm \frac{r_j^2}{2}$ of multiplicity 2 for D. We would like to apply Courant–Hilbert minimax principle to D, except that D is not symmetric with respect to the dot product of $\mathbb{R}^{4n}$. However if we define an inner product

$$\langle [x,y]^*, [x',y']^* \rangle = Bx \cdot x' + By \cdot y',$$

with corresponding norm

$$\|[x,y]^*\| = 2H(x) + 2H(y),$$

then D is $\langle \cdot, \cdot \rangle$-symmetric. Indeed,

$$\begin{aligned} \left\langle D \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x' \\ y' \end{bmatrix} \right\rangle &= \left\langle \begin{bmatrix} C^{-1}y \\ -C^{-1}x \end{bmatrix}, \begin{bmatrix} x' \\ y' \end{bmatrix} \right\rangle \\ &= BC^{-1}y \cdot x' - BC^{-1}x \cdot y' \\ &= \bar{J}x' \cdot y + \bar{J}x \cdot y', \end{aligned}$$

which is symmetric. Moreover,

$$\left\langle D \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle = 2\bar{\omega}(x,y).$$

We are now in a position to apply Lemma 2.1 to obtain (2.8) with $\bar{\omega}$ instead of ω^+ in the numerator. (Note that $\frac{1}{2}r_j^2(H)$ is the $2n+2j$-th eigenvalue of D.) Finally we need to replace $\bar{o}$ with $\bar{\omega}^+$. This is plausible because the left-hand side is positive. $\square$

An immediate consequence of Theorem 2.2 is a linear version of Gromov's non-squeezing theorem. More precisely, if we define

$$B_R = \{x : |x| \le R\}, \quad Z_R = \{x : q_1^2 + p_1^2 \le R^2\},$$

then $\mathbf{r}(B_R) = (R, R, \ldots, R)$ and $\mathbf{r}(Z_R) = (R, \infty, \infty, \ldots, \infty)$. By Theorem 2.2**(ii)** if for some $T \in Sp(n)$, we have $T(B_r) \subseteq Z_R$, then $r \le R$. We now slightly improve this and give a direct proof of it.

Proposition 2.4 *Suppose that for some $T \in Sp(n)$ and $z^0 \in \mathbb{R}^{2n}$, $T(B_r) \subseteq z^0 + Z_R$. Then $r \le R$.*

Proof Write $z^0 = (q_1^0, \ldots, q_n^0, p_1^0, \ldots, p_n^0)$ and let $(s_1, \ldots, s_n, t_1, \ldots, t_n)$ denote the rows of T. By assumption

$$(x \cdot s_1 - q_1^0)^2 + (x \cdot t_1 - p_1^0)^2 \le R^2$$

for x satisfying $|x| \le r$. Hence

$$(x \cdot s_1)^2 + (x \cdot t_1)^2 - 2x \cdot (q_1^0 s_1 + p_1^0 t_1) \le R^2. \tag{2.10}$$

On the other hand, since T^* is symplectic,

$$\bar\omega(s_1,t_1)=\bar\omega(T^*e_1,T^*f_1)=\bar\omega(e_1,f_1)=-1 \tag{2.11}$$

where $\{e_1,\dots,e_n,f_1,\dots,f_n\}$ denote the standard symplectic basis for $\mathbb{R}^{2n}$, i.e., $e_j\cdot x=q_j$ and $f_j\cdot x=p_j$ for $x=(q_1,\dots,q_n,p_1,\dots,p_n)$. From (2.11) we learn that

$$1=|\bar\omega(s_1,t_1)|=|Js_1\cdot t_1|\le|s_1||t_1|.$$

So either $|s_1|\ge 1$ or $|t_1|\ge 1$. Both cases can be treated similarly, so let us assume that for example $|t_1|\ge 1$. We then choose $x=\pm r\frac{t_1}{|t_1|}$ in (2.13). We select $+$ or $-$ for x so that $x\cdot(q_1^0s_1+p_1^0t_1)\le 0$. This would allow us to deduce $r^2\le R^2$ from (2.13), and this completes the proof. □

Remark 2.2 Note that if we consider $Z'_R=\{x:\ q_1^2+q_2^2\le R^2\}$ instead, then $\mathbf{r}(Z'_R)=(\infty,\dots,\infty)$ and we can embed B_r symplectically inside Z'_R no matter how large r is. This is because the map $T(q,p)=(\varepsilon q,\varepsilon^{-1}p)$ is symplectic (use $\omega(q,p,q',p')=p\cdot q'-q\cdot p'$ to check this), and $T(B_r)$ consists of points (q,p) such that $\varepsilon^{-2}|q|^2+\varepsilon^2|p|^2\le r^2$. □

As our next topic, we address the issue of symplectic rigidity for linear transformations. Note that the condition $|\det A|=1$ for a matrix A is equivalent to the claim that the sets E and $A(E)$ have the same Euclidean volume. To be able to establish Eliashberg-Gromov rigidity, we would like to have a similar criterion for symplectic maps. Since a symplectic change of variables does not change symplectic radii, the volume must be replaced with suitable linear capacities that are defined in terms of the symplectic radii. Though as in the case of volume, the orientation could be reversed when a symplectic capacity is preserved. So instead of $T\in Sp(k)$, what we really have is

$$|\bar\omega(Ta,Tb)|=|\bar\omega(a,b)|. \tag{2.12}$$

We set $S'(k)$ to be the set of matrices T for which (2.12) is valid for all $a,b\in\mathbb{R}^k$. We also say that matrix T is *anti-symplectic* if $\bar\omega(Tx,Ty)=-\bar\omega(x,y)$, or equivalently $T^*\bar JT=-\bar J$. It is not hard to show that $T\in S'(k)$ iff T is either symplectic or anti-symplectic (see Exercise 2.1). On the other hand, it is straightforward to check that a linear map T is anti-symplectic iff $T\circ\tau\in S(k)$, where $\tau(q,p)=(p,q)$. From this we learn that indeed in Theorem 2.1 and Corollary 2.2 apply to anti-symplectic transformations as well. In summary,

Proposition 2.5 **(i)** *Let H be a positive definite quadratic function. Then there exists $T\in S'(k)$ such that $H\circ T(x)=\sum_1^n\frac{q_j^2+p_j^2}{r_j^2}$, where $r_j=r_j(H)$.*

(ii) *Let E_1 and E_2 be two ellipsoids. Then $T(E_2)=E_1$ for some $T\in S'(k)$ if and only if $\mathbf{r}(E_2)=\mathbf{r}(E_1)$.*

We are now ready for a converse to Proposition 2.5, which will be used for the proof of Eliashberg's theorem in Section 6.

Theorem 2.3 *Let T be an invertible $2n \times 2n$ matrix. Then $T \in S'(k)$ iff $r_1(E) = r_1(T(E))$ for every ellipsoid E.*

Proof Given a pair of vectors (a, b) with $\bar{\omega}(a, b) \neq 0$, let us define

$$Z(a, b) = \{x : (x \cdot a)^2 + (x \cdot b)^2 \leq 1\},$$

which is a cylinder. We claim that in fact $Z(a, b)$ is a (degenerate) ellipsoid with

$$r_1(Z(a, b)) = |\bar{\omega}(a, b)|^{-1/2}, \quad r_j(Z(a, b)) = \infty, \text{ for } j \geq 2. \tag{2.13}$$

Once we establish this, we are done: If $r_1(Z(a, b)) = r_1(T(Z(a, b))$ for every a and b with $\bar{\omega}(a, b) \neq 0$, then using $Z(a, b) = T\big(Z(T^*a, T^*b)\big)$, we deduce

$$|\bar{\omega}(a, b)| = |\bar{\omega}(T^*a, T^*b)|,$$

whenever $\bar{\omega}(a, b), \bar{\omega}(T^*a, T^*b) \neq 0$. As a result

$$A := \{(a, b) : \ |\bar{\omega}(a, b)| \neq |\bar{\omega}(T^*a, T^*b)|\} \subseteq A' := \{(a, b) : \ \bar{\omega}(a, b)\bar{\omega}(T^*a, T^*b) = 0\}.$$

Since the set A is open and the set A' is the union of two linear sets of codimension 1, we must have $A = \emptyset$, which in turn implies that $T^* \in S'(k)$. From this, we can readily show that $T \in S'(k)$.

It remains to verify (2.13). First observe that if $r = |\bar{\omega}(a, b)|^{-1/2}$ and $(a_1, b_1) = r(a, b)$, then $|\bar{\omega}(a_1, b_1)| = 1$, and

$$Z(a, b) = \left\{x : \ H(x) := \left[(x \cdot a')^2 + (x \cdot b')^2\right] / r^2 \leq 1\right\}.$$

Without loss of generality, let us assume that in fact $\bar{\omega}(a_1, b_1) = -1$. We then build a (symplectic) basis $\{a_1, \dots, a_n, b_1, \dots, b_n\}$ such that

$$\bar{\omega}(b_j, a_i) = \delta_{i,j}, \quad \bar{\omega}(a_i, a_j) = \bar{\omega}(b_i, b_j) = 0,$$

for all i and j. Let us write $\{e_1, \dots, e_n, f_1, \dots, f_n\}$ for the standard basis, in other words, e_i and f_i satisfy $e_i \cdot x = q_i$ and $f_i \cdot x = p_i$. We then choose a map $\hat{T}$ so that $\hat{T}^* a_i = e_i$ and $\hat{T}^* b_i = f_i$. We have

$$H \circ \hat{T}(x) = \left[(\hat{T}x \cdot a_1)^2 + (\hat{T}x \cdot b_1)^2\right] / r^2 = \left[(x \cdot e_1)^2 + (x \cdot f_1)^2\right] / r^2 = (q_1^2 + p_1^2)/r^2.$$

From this we deduce (2.13) by definition. $\square$

As we observed in Remark 2.1, sometimes it is beneficiary to identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ and use complex number. More precisely, we may write $z_j = q_j + ip_j$, so that if $a = (z_1, \dots, z_n)$ and $b = (z'_1, \dots, z'_n)$ are in $\mathbb{C}^n$, then $\bar{J}a = (-iz_1, \dots, -iz_n)$, and

$$a \cdot b = Re \sum_{j=1}^{n} z_j \bar{z}'_j, \qquad \bar{\omega}(a,b) = Im \sum_{j=1}^{n} z_j \bar{z}'_j.$$

More generally, given a symplectic vector space (V, ω), we may try to express ω as

$$\omega(a,b) = g(Ja,b), \tag{2.14}$$

where g is an inner product on V and $J : V \to V$ is a linear map satisfying $J^2 = -I$. When (2.14) is true, we say that the pair (g, J) is *compatible* with ω. Let us write $\mathcal{I}(\omega)$ for the space of compatible pairs (g, J). We also define

$$\mathcal{G}(\omega) := \{g : \ (g,J) \in \mathcal{I}(\omega), \text{ for some } \ J\}.$$

Note that if $(g, J) \in \mathcal{I}(\omega)$, then

$$g(Ja, Jb) = \omega(a, Jb) = -\omega(Jb, a) = g(b,a) = g(a,b),$$

which means that $J^*J = I$, where J^* is the g-adjoint or transpose of J. From this and $J^2 = -I$, we learn that for every $(g, J) \in \mathcal{I}(\omega)$, we have $J^* = -J$. Define $T^*(g)(a,b) = g(Ta, Tb)$ and $T^*\omega(a,b) = \omega(Ta, Tb)$.

Proposition 2.6 *Let $T : V \to V$ be an invertible linear map. Then $T^*\omega = \omega'$ iff $T^*(\mathcal{G}(\omega)) \subseteq \mathcal{G}(\omega')$.*

Proof Set $\hat{T}(J) = T^{-1}JT$. If $(g, J) \in \mathcal{I}(\omega)$, then

$$(T^*\omega)(a,b) = \omega(Ta, Tb) = g(JTa, Tb) = \omega(T\hat{T}(J)a, Tb) = (T^*\omega)(\hat{T}(J)a, b).$$

From this we deduce that if $(g, J) \in \mathcal{I}(\omega)$, then $(T^*g, \hat{T}(J)) \in \mathcal{I}(T^*\omega)$. As a result, $T^*(\mathcal{G}(\omega)) \subseteq \mathcal{G}(T^*\omega)$. Similarly, $(T^{-1})^*\mathcal{G}(T^*\omega) \subseteq \mathcal{G}(\omega)$. Hence $T^*(\mathcal{G}(\omega)) = \mathcal{G}(T^*\omega)$. From this we learn that we only need to show that if $\mathcal{G}(\alpha) \subseteq \mathcal{G}(\beta)$, then $\alpha = \beta$. $\square$

Example 2.2 Identifying a metric $g(a,b) = Ga \cdot b$ with the matrix $G > 0$, one can readily show

$$\mathcal{G}(\bar{\omega}) = \{G : \ G > 0, \ G \in S(2n)\}. \tag{2.15}$$

Indeed if $g(a,b) = Ga \cdot b$ and $(g, J) \in \mathcal{I}(\omega)$, then $\bar{J} = GJ$ so that $G = -\bar{J}J$. This implies

$$G\bar{J}G = \bar{J}J\bar{J}\bar{J}J = \bar{J},$$

which means that G is symplectic. Conversely, if $G > 0$ and $G \in S(2n)$, then set $J = G^{-1}\bar{J}$ and observe that since G^{-1} is also symplectic, then $J^2 = G^{-1}\bar{J}G^{-1}\bar{J} = \bar{J}\bar{J} = -I$. $\square$

When $g \in \mathcal{G}(\omega)$, we know how to calculate the area of the parallelogram associated with two vectors a and b, namely

$$A_g(a,b) = \left(\|a\|_g^2\|b\|_g^2 - g(a,b)^2\right)^{\frac{1}{2}},$$

where $\|a\|_g = g(a,a)^{1/2}$. Of course $\omega(a,b)$ offers the symplectic area of the same parallelogram. In the next proposition, we compare these two areas.

Proposition 2.7 *For every $g \in \mathcal{G}(\omega)$, we have*

$$\omega(a,b) \le A_g(a,b) \le \frac{1}{2}\left(\|a\|_g^2 + \|b\|_g^2\right). \tag{2.16}$$

Moreover we have equality iff $b = Ja$.

Proof The second inequality is obvious and the first inequality is also obvious when $g(a,b) = 0$, because

$$\omega(a,b)^2 = g(Ja,b)^2 \le g(Ja,Jb)g(b,b) = \|a\|_g^2\|b\|_g^2.$$

Given arbitrary a and b, with $a \neq 0$, set $t = -g(a,b)/g(a,a)$ so that $g(a,b') = 0$, for $b' = ta + b$. We certainly have

$$\begin{aligned}\omega(a,b)^2 = \omega(a,b')^2 \le g(Ja,Ja)g(b',b') &= g(a,a)g(b',b') = g(a,a)g(b',b)\\ = g(a,a)g(b,b) + tg(a,a)g(a,b) &= A_g(a,b),\end{aligned}$$

with equality iff $Ja = \theta b'$ for some $\theta \ge 0$. However, if we require $2\omega(a,b)^2 = \|a\|_g^2 + \|b\|_g^2$, then $g(a,b) = 0$, $g(a,a) = g(b,b)$, and $\omega(a,b)^2 = g(a,a)^2$. As a result,

$$\|Ja - b\|_g^2 = \|a\|_g^2 + \|b\|_g^2 - 2g(Ja,b) = 2\|a\|_g^2 - 2\omega(a,b) = 0,$$

as desired. $\square$

Exercise 2.1

(i) Let V be a vector space with $\dim V = 2n$. Then a 2-form $\omega : V \times V \to \mathbb{R}$ is non-degenerate if and only if $\omega^n = \underbrace{\omega \wedge \cdots \wedge \omega}_{n \text{ times}} \neq 0$.

(ii) Use part (i) to deduce that if $T \in Sp(n)$, then $\det T = 1$.

(iii) Recall that an invertible matrix T can be written as $T = PO$ with $P > 0$ symmetric and O orthogonal, and that this decomposition is unique (polar decomposition). Show that if $T \in Sp(n)$, then P and $O \in Sp(n)$.

(iv) Show that if $T \in Sp(n) \cap O(2n)$, then $T = \begin{bmatrix} X & -Y \\ Y & X \end{bmatrix}$ with X, Y two $n \times n$ matrices such that $X + iY$ is a unitary matrix.

(v) Let V be a vector space with $\dim V = 2n + 1$. Assume β is an antisymmetric 2-form on V with

$$\dim\{v \in V : \beta(v, a) = 0 \text{ for all } a \in V) = 1.$$

Then there exists a basis $\{e_1, \ldots, e_n, f_1, \ldots, f_n, \bar{a}\}$ such that $\beta(e_i, e_j) = \beta(f_i, f_j) = 0$, $\beta(f_j, e_i) = \delta_{ij}$, and $\beta(\bar{a}, f_j) = \beta(\bar{a}, e_j) = 0$.

(vi) If $T \in Sp(n)$ and $A \in Ham(n)$, then $T^{-1}AT \in Ham(n)$.

(vii) An invertible T maps the flows of $\frac{dx}{dt} = Ax$ to the flows of $\frac{dx}{dt} = Bx$ iff $B = TAT^{-1}$.

(viii) If $C_1, C_2 \in \mathrm{Ham}(n)$ and $r \in \mathbb{R}$, then $C_1 + C_2, [C_1, C_2] = C_1C_2 - C_2C_1, C_1^t, rC_1 \in \mathrm{Ham}(n)$.

(ix) If $T_1, T_2 \in Sp(n)$, then $T_1^{-1}, T_1^*, T_1T_2 \in Sp(n)$.

(x) Show that if (2.13) is true for all a and b, then T is either symplectic or ani-symplectic.

(xi) If $Z(t_0, t)$ denotes the fundamental solution of $\dot{x} = JB(t)x$ with $B : [t_0, \infty) \to S(2n)$ a C^1-function, then $Z(t_0, t) \in Sp(n)$ for every $t \geq t_0$.

(xii) If $e^{tC} \in Sp(n)$ for every t, then $C \in \mathrm{Ham}(n)$.

(xiii) If $C \in \mathrm{Ham}(n)$, and $p_C(\lambda) = \det(\lambda I - C)$, then $p_C(\lambda) = p_C(-\lambda)$.

(xiv) If $T \in Sp(n)$, then $p_T(\lambda^{-1}) = \lambda^{-2n} p_T(\lambda)$.

(xv) Let T be an invertible matrix and assume that $n \geq 2$. Show that if $r_2(T(E)) = r_2(E)$ for every ellipsoid E, then $T \in S'(2n)$.

$\square$

3 Symplectic Manifolds and Darboux's Theorem

Before discussing symplectic manifolds, let us review some useful facts about our basic example $(\mathbb{R}^{2n},\bar\omega)$ with $\bar\omega(a,b)=\bar J a\cdot b$. We write $Sp(\mathbb{R}^{2n})$ for the space of differentiable functions φ such that $\varphi^*\bar\omega=\omega$. This means

$$\varphi^*\bar\omega(a,b)=\bar\omega(\varphi'(x)a,\varphi'(x)b)=\bar\omega(a,b),$$

for every $a,b,x\in\mathbb{R}^{2n}$. Here $\varphi'(x)$ denotes the derivative of φ. A function $\varphi\in Sp(\mathbb{R}^{2n})$ is called *symplectic*. Note that $\varphi\in Sp(\mathbb{R}^{2n})$ iff $\varphi'(x)\in Sp(2n)$ for every x. Hence for a symplectic transformation φ, we have

$$\varphi'(x)^*\bar J\varphi'(x)=\bar J.$$

Evidently $\bar\omega=\sum_{i=1}^n dp_i\wedge dq_i=d\bar\lambda$ where $\bar\lambda=\sum_1^n p_i dq_i=p\cdot dq$. Let us define

$$A(\gamma)=\int_\gamma\bar\lambda. \tag{3.1}$$

Clearly $\varphi\in Sp(\mathbb{R}^{2n})$ iff $d(\varphi^*\bar\lambda-\bar\lambda)=0$. Hence $\varphi\in Sp(\mathbb{R}^{2n})$ is equivalent to saying

$$A(\varphi\circ\gamma)=A(\gamma) \tag{3.2}$$

for every closed curve γ. It is worth mentioning that if γ is parametrized by $\theta\mapsto x(\theta)$, $\theta\in[0,T]$, then

$$A(\gamma)=\int_0^T p\cdot\dot q d\theta=\frac12\int_0^1(p\cdot\dot q-q\cdot\dot p)d\theta=\frac12\int_0^1(\bar Jx\cdot\dot x)d\theta. \tag{3.3}$$

Given a scalar-valued (0-form) function H, we may use non-degeneracy of $\bar\omega$ to define a vector-field X_H such that

$$\bar\omega(X_H(x),a)=-dH(x)a=-\nabla H(x)\cdot a,$$

which means that $X_H=J\nabla H$. We write $\phi_t=\phi_t^H$ for the corresponding flow:

$$\begin{cases}\frac{d}{dt}\phi_t(x)=X_H(\phi_t(x)),\\ \phi_0(x)=x.\end{cases} \tag{3.4}$$

Our interest in symplectic transformation stems from two important facts. Firstly, $\phi_t\in Sp(\mathbb{R}^{2n})$ if ϕ_t is a Hamiltonian flow. We have seen this in Section 1 and will be proved later in this section for general symplectic manifolds. Secondly a symplectic change of coordinates preserve Hamiltonian structure (see Proposition 3.1 below). More precisely, if $\varphi\in Sp(\mathbb{R}^{2n})$, and ϕ_t is the flow of X_H, then $\psi_t = \varphi^{-1} \circ \phi_t \circ \varphi$ is the flow of a Hamiltonian system. To guess what the Hamiltonian function of ψ_t is, observe

$$\begin{aligned} \frac{d}{dt}\psi_t\Big|_{t=0} &= (\varphi^{-1})' \circ \varphi \; X_H \circ \varphi = (\varphi')^{-1} \; \bar{J}\nabla H \circ \varphi \\ &= -\bar{J}(\varphi')^* \; \bar{J}\bar{J}\nabla H \circ \varphi = \bar{J}(\varphi')^* \; \nabla H \circ \varphi = \bar{J}\nabla(H \circ \varphi). \end{aligned}$$

A pair (M, ω) is called a *symplectic manifold* if M is an even dimensional manifold and ω is a closed non-degenerate 2-form on M. This implies that for each $x \in M$, the pair (T_xM, ω_x) is a symplectic vector space. Also, by Exercise 2.1**(i)** we know that if $dim(M) = 2n$, then the form ω^n is a volume form. Hence M is an orientable manifold. In fact if M is a compact symplectic manifold without boundary, then ω is never exact. This is because if $\omega = d\lambda$, then $\Omega := \omega^n = d(\lambda \wedge \omega^{n-1})$. But by Stokes' theorem $\int_M \Omega = \int_M d(\lambda \wedge \omega^{n-1}) = 0$, which contradicts the non-degeneracy of Ω. Note however that $(\mathbb{R}^{2n}, \bar{\omega})$ is an example of a non-compact symplectic manifold with $\bar{\omega} = d\bar{\lambda}$.

Example 3.1 (i) Any orientable 2-dimensional manifold is symplectic where ω is chosen to be any volume form.

(ii) The sphere S^{2n} with $n > 1$ is not symplectic because any closed 2-form is exact, hence degenerate.

(iii) The classical example $(\mathbb{R}^{2n}, \bar{\omega})$ has a natural generalization that is relavant for models in classical mechanics: Every cotangent bundle T^*M can be equipped with a symplectic $\omega = d\lambda$ where λ is a *standard* 1-form that can be defined in two ways; using local charts and pull-backing $\bar{\lambda}$ to λ, or giving a chart-free description. We start with the former. Let us assume that M is an n-dimensional C^2 manifold and choose an atlas $\mathcal{A}$ of charts (U, h) of M such that U is an open subset of M and $h : U \to h(U) := V \subseteq \mathbb{R}^n$ is a diffeomorphism. This induces a C^1 transformation $dh : TU \to TV = V \times \mathbb{R}^n$. We next define a natural transformation $\bar{h} : T^*U \to T^*V = V \times \mathbb{R}^n$. To construct $\bar{h}$, take the standard basis $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$ and define $\hat{e}_j(q) = (d\phi)_q(e_j)$, where $\phi = h^{-1}$. (We may define T_qM as the equivalence classes of curves $\gamma : (-\delta, \delta) \to M$ with $\gamma(0) = q$, and two such curves γ_1 and γ_2 are equivalent if $(h \circ \gamma_1)'(0) = (h \circ \gamma_2)'(0)$. We may then define $\hat{e}_j(q)$ as the equivalent class of $\gamma^j(\theta) = h^{-1}(h(q)+\theta e_j)$.) Certainly $\{\hat{e}_1(q), \dots, \hat{e}_n(q)\}$ defines a basis for T_qM. We now define a basis for T_q^*L by taking dual vectors $e_1^*, \dots, e_n^*$ that are defined $e_i^*(q)\left(\sum_{j=1}^n v_j \hat{e}_j(q)\right) = v_j$. We now define $\bar{h}$ by

$$\bar{h}\left(q, \sum_{j=1}^n p_j e_j^*(q)\right) = (h(q), (p_1, \dots, p_n)).$$

We finally define λ as the unique 1-form such that for each chart (U, h), the restriction of λ to T^*U is given by $\lambda = \bar{h}^*\bar{\lambda}$. For an alternative chart-free description, observe that if

$\hat{\pi} : T^*\mathbb{R}^n \to \mathbb{R}^n$ is the projection $\hat{\pi}(q,p) = q$, then

$$d\hat{\pi} : T(T^*\mathbb{R}^n) \to T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$$

is simply given by

$$d\hat{\pi}_{(q,p)}(\alpha, \beta) = \alpha.$$

Hence we may write $\bar{\lambda}_{(q,p)}(\alpha,\beta) = p \cdot \alpha = p \cdot (d\hat{\pi})_{(q,p)}(\alpha,\beta)$. Going back to M, let us also define $\pi : T^*M \to M$ to be the projection onto the base point, i.e., $\pi(q,p) = q$ with $q \in M$ and $p \in T_q^*M$. Since the following diagram commutes

$$\begin{array}{ccc} T^*U & \xrightarrow{\pi} & U \\ \bar{h}\big\downarrow & & \big\downarrow h \\ \mathbb{R}^{2n} & \xrightarrow[\hat{\pi}]{} & \mathbb{R}^n \end{array}$$

we also have that their derivatives

$$\begin{array}{ccc} T(T^*U) & \xrightarrow{d\pi} & TU \\ d\bar{h}\big\downarrow & & \big\downarrow dh \\ T\mathbb{R}^{2n} & \xrightarrow[d\hat{\pi}]{} & \mathbb{R}^{2n} \end{array}$$

commute. Hence we also have

$$\lambda_{(q,p)}(a) = p\left(d\pi_{(q,p)}(a)\right),$$

which gives the desired chart-free description of λ. $\square$

A differentiable map $f : (M_1, \omega^1) \to (M_2, \omega^2)$ between two symplectic manifolds is called *symplectic* if $f^*\omega^2 = \omega^1$. This means

$$\omega^2_{f(x)}(df(x)a, df(x)b) = \omega^1_x(a,b) \tag{3.5}$$

for $x \in M_1$ and $a, b \in T_xM_1$. We write $Sp(M_1, M_2)$ for the space of symplectic transformations. When $M_1 = M_2 = M$ and $\omega^1 = \omega^2 = \omega$, we simply write $Sp(M)$ for $Sp(M_1, M_2)$. We note that if $f \in Sp(M_1, M_2)$, then $df(x)$ is injective by (3.5) and non-degeneracy of ω^1_x. Hence, if such f exists, then $\dim M_1 \le \dim M_2$.

Let (M, ω) be a symplectic manifold and assume that X is sufficiently nice vector field for which the ODE $\dot{x} = X(x,t)$ is well defined. The flow of this vector field is denoted by $\phi_t = \phi_t^X$. We wish to find conditions on X to guarantee that $\phi_t^*\omega = \omega$. To prepare for this let us take an arbitrary ℓ-form α and evolve it with the flow; set $\alpha(t) = \phi_t^*\alpha$. We would like to derive an evolution equation for $\alpha(t)$. Let's examine some examples.

Example 3.2 Assume that $(M, \omega) = (\mathbb{R}^{2n}, \bar{\omega})$.

(i) If $\alpha = f$ is a 0-form, then $\alpha(t)$ is a function u that is given by $u(x,t) = f(\phi_t(x))$. By differentiating u at $t = 0$ and using the group property of the flow (see the proof of Proposition 3.1 below), we can readily show

$$u_t = X \cdot u_x.$$

(ii) If $\alpha = m(x)\ dx_1 \dots dx_k$, with $k = 2n$, then $\alpha(t) = m(\phi_t(x)) \det(\phi_t'(x))\ dx_1 \dots dx_k$. In this case, differentiating in t yields,

$$\rho_t = \rho_x \cdot X + \rho\ divX = div(\rho X),$$

because for small t, we have $\det(\phi_t'(x)) = \det(I + tX'(x) + o(t)) = 1 + t divX + o(t)$.

(iii) If $\omega = F \cdot dx$ is a 1-form, then

$$\alpha(t) = \phi_t^* \alpha = F(\phi_t(x)) \cdot \phi_t'(x) dx = \phi_t'(x)^* F(\phi_t(x)) \cdot dx.$$

Hence, if we set $u(x,t) = \phi_t'(x)^* F(\phi_t(x))$, then

$$u_t(x,t) = X_x(x,t)^* u(x,t) + u_x(x,t) X(x,t). \tag{3.6}$$

If $u = (u^1, \dots, u^k)$ and $X = (X^1, \dots, X^k)$, then the i-th component of the right-hand side of (3.6) equals

$$\sum_j \left(X^j_{x_i} u^j + u^i_{x_j} X^j \right) = \left(\sum_j X^j u^j \right)_{x_i} + \sum_j (u^i_{x_j} - u^j_{x_i}) X^j.$$

As a result,

$$u_t = (X \cdot u)_x + \mathcal{C}(u) X. \tag{3.7}$$

□

As Example 3.2**(iii)** indicates, a simple manipulation of the right-hand side of (3.6) leads to the compact expression of the right-hand side of (3.7), that may be recognized as

$$[(X \cdot u)_x + \mathcal{C}(u) X] \cdot dx = d\left(i_X \alpha(t)\right) + i_X d\alpha(t).$$

More generally we have the following useful result of Cartan:

Proposition 3.1 **(i)** *Let X be a vector field with flow ϕ_t and let α be a ℓ-form. Then*

$$\frac{d}{dt} \phi_t^* \alpha = \mathcal{L}_X \phi_t^* \alpha = \phi_t^* \mathcal{L}_X \alpha \tag{3.8}$$

with $\mathcal{L}_X = i_X \circ d + d \circ i_X$.
(ii) *Let $X(\cdot, t)$ be a possibly time dependent vector field and denote is flow by $\phi_{s,t}$. If $\alpha(t) = \phi_t \alpha$ for $\phi_t = \phi_{0,t}$, then*

$$\frac{d}{dt} \alpha(t) = \mathcal{L}_{X(\cdot,t)} \alpha(t).$$

Proof We only establish Part **(i)** as the proof of **(ii)** is similar. Let us define

$$\mathcal{L}\beta = \lim_{h\to 0} \frac{1}{h}(\phi_h^*\beta - \beta)$$

whenever the limit exists. Since

$$(\phi_{t+h}^* - \phi_t^*)\alpha = \phi_t^*(\phi_h^*\alpha - \alpha) = \phi_h^*(\phi_t^*\alpha) - \phi_t^*\alpha,$$

it suffices to show that $\mathcal{L} = \mathcal{L}_X$. Let us study some properties of $\mathcal{L}$. From $\phi_t^*(\alpha \wedge \beta) = \phi_t^*\alpha \wedge \phi_t^*\beta$, we learn

$$\phi_t^*(\alpha \wedge \beta) - \alpha \wedge \beta = (\phi_t^*\alpha - \alpha) \wedge \phi_t^*\beta + \alpha \wedge (\phi_t^*\beta - \beta).$$

From this we deduce

$$\mathcal{L}(\alpha \wedge \beta) = \alpha \wedge \mathcal{L}\beta + \mathcal{L}\alpha \wedge \beta. \tag{3.9}$$

From $\phi_t^* \circ d = d \circ \phi_t^*$, we deduce

$$\mathcal{L} \circ d = d \circ \mathcal{L}. \tag{3.10}$$

We can readily show that $\mathcal{L}_X$ satisfy (3.9) and (3.10) as well. Since locally every form can be built from 0-th forms using the operations $\wedge$ and d, we only need to check that $\mathcal{L} = \mathcal{L}_X$ on 0-forms. That is, if $f : M \to \mathbb{R}$, then $\mathcal{L}f = i_X \circ df = df(X)$. This is trivially verified because $\phi_t(x) = x + tX(x) + o(t)$. $\square$

Armed with (3.8), we can readily find necessary and sufficient conditions on a vector field X such that the flow of X preserves ω. In view of Proposition 3.1,

$$\left(\phi_t^X\right)^* \omega = \omega \quad \text{for all } t \text{ iff} \quad \mathcal{L}_X\omega = d\left(i_X\omega\right) = 0.$$

This leads to two definitions:

Definition 3.1

- We call a vector field X *symplectic* iff $i_X\omega$ is exact.
- Given a differentiable $H : M \to \mathbb{R}$, we can find a unique vector field $X = X_H = X_H^\omega$ such that
$$(i_{X_H}\omega) = \omega(X_H, \cdot) = -dH.$$
(Note that the non-degeneracy of ω guarantees the existence X_H.) The vector field X_H is called *Hamiltonian* and its corresponding flow is denoted by ϕ_t^H.

Example 3.2 If $\omega_x(v_1, v_2) = C(x)v_1 \cdot v_2$ is a symplectic form in $\mathbb{R}^{2n}$, then $X_H = -C^{-1}\nabla H$. □

As in the case of $\bar{\omega}$, a change of coordinates turn a Hamiltonian flow to another Hamiltonian flow as our next Proposition demonstrates.

Proposition 3.2 *Let (M, ω) be a symplectic manifold. If $\varphi : N \to M$ is a diffeomorphism and $H : M \to \mathbb{R}$ is a smooth Hamiltonian, then $X^{\varphi^*\omega}_{H\circ\varphi} = (d\varphi)^{-1} X^{\omega}_H \circ \varphi$. In other words, $\varphi^* X^{\omega}_H = X^{\varphi^*\omega}_{H\circ\varphi}$, and*

$$\varphi^{-1} \circ \phi_t^{X^{\omega}_H} \circ \varphi = \phi_t^{\varphi^* X^{\omega}_H} = \phi_t^{X^{\varphi^*\omega}_{H\circ\varphi}},$$

where ϕ_t^X denotes the flow of the vector field X.

Proof By Lemma 10.2 of Appendix A, ψ_t is the flow of $\varphi^* X^{\omega}_H$. Furthermore, for $X = X^{\omega}_H$ and $\hat{X} = (d\varphi)^{-1} X \circ \varphi$,

$$\begin{aligned}(\varphi^*\omega)_x(\hat{X}(x), v) &= \omega_{\varphi(x)}\left((d\varphi)_x \hat{X}(x), (d\varphi)_x v\right) = \omega_{\varphi(x)}\left(X(\varphi(x)), (d\varphi)_x v\right) \\ &= -(dH)_{\varphi(x)}\left((d\varphi)_x v\right) = -\left(\varphi^*(dH)\right)_x(v) = -d(H \circ \varphi)_x(v),\end{aligned}$$

for every v. Hence $\hat{X} = X^{\varphi^*\omega}_{H\circ\varphi}$, as desired. □

We now turn to the question of the equivalence of two symplectic manifolds or the embedding of one symplectic manifold inside another symplectic manifold. As a warm-up, let us discuss the analogous question for volume forms.

Theorem 3.1 *(Moser) Let M be a connected oriented compact manifold with no boundary. Assume that α and β are two volume forms. Then there exists a diffeomorphism φ such that $\varphi^*\alpha = \beta$ iff $\int_M \alpha = \int_M \beta$.*

Proof Evidently if for some diffeomorphism φ, we have $\varphi^*\alpha = \beta$, then

$$\int_M \beta = \int_M \varphi^*\alpha = \int_{\varphi(M)} \alpha = \int_M \alpha.$$

As for the converse, assume that $c = \int_M \alpha = \int_M \beta$. Without loss of generality, we may assume that $c = 1$, and regard α and β as two mass (or probability) distributions on M. We may interpret φ as a plan of transportation; a unit mass at x is transported to $\phi(x)$ so that after this transportation is performed for all points, the mass distribution changes from α to β. With this interpretation in mind, we may design a route for our transportation so that this change of transportation is carried out in one unit of time. Equivalently, we may search for a (possibly time dependent) vector field X such that if ϕ_t denotes its flow, then $\varphi = \phi_1$. Note that if we set $\alpha(t) = \phi_t^*\alpha$, then $\alpha(0) = \alpha$ and $\alpha(1) = \beta$. This scheme of finding φ has a chance to work only if there exists a path of volume forms connecting α to β. So for achieving our goal, let us first such a path. We now argue that in fact the path $\alpha(t) = t\beta + (1-t)\alpha$ would do the job. Since α and β are volume forms, locally $\alpha = a\ dx_1 \wedge \cdots \wedge dx_k$ and $\beta = b\ dx_1 \cdots \wedge \ldots dx_k$ with a and b non-zero and continuous. Since $\int_M \alpha = \int_M \beta$, the functions a and b must have the same sign. Hence $t\beta + (1-t)\alpha$ is never zero, and as a result, $\alpha(t)$ is never degenerate. We next search for a vector field $X(\cdot, t)$ such that its flow ϕ_t satisfies $\phi_t^*\alpha = t\beta + (1-t)\alpha$. Differentiating both sides with respect to t yields

$$\beta - \alpha = \frac{d}{dt}\alpha(t) = \mathcal{L}_X \alpha(t) = d(i_{X(\cdot,t)}\alpha(t)).$$

by Proposition 3.1. But $\int_M(\beta - \alpha) = 0$ implies that $\beta - \alpha = d\gamma$ for some $k-1$-form γ (See Lemma A1 of the Appendix). The existence of $X(\cdot, t)$ with $i_{X(\cdot,t)}\alpha(t) = \gamma$ follows from the non-degeneracy of $\alpha(t)$. □

Remark 3.1. **(i)** As a consequence of Theorem 3.1, if M and N are two oriented compact closed manifolds with volume forms α and η respectively, then they there exists a diffeomorphism $\phi : N \to M$ with $\phi^*\alpha = \eta$ iff M and N are diffeomorphic and $\int_M \alpha = \int_N \eta$. Indeed if $\psi : M \to N$ is a diffeomorphism, then α and $\beta = \psi^*\eta$ are two volume forms on M for which Theorem 3.1 applies: there exists a diffeomorphism $\varphi : M \to M$, such that $\phi^*\alpha = \psi^*\eta$. We then set $\phi = \phi \circ \psi^{-1}$ to deduce that $\phi^*\alpha = \eta$.

(ii) If $M = \mathbb{T}^k$ and $\alpha = \hat{\alpha}(x)\ dx_1 \wedge \cdots \wedge dx_k$, $\beta = \hat{\beta}(x)\ dx_1 \wedge \cdots \wedge dx_k$, then $\varphi^*\alpha = \beta$ means that $\hat{\alpha}(\varphi(x)) \det \varphi'(x) = \hat{\beta}(x)$. If $\varphi = \nabla w$ for some $w : \mathbb{T}^k \to \mathbb{R}$, then $\hat{\alpha}(\nabla w)\det(D^2 w) = \hat{\beta}$. This is the celebrated Monge–Ampère's equation. However the function φ in Moser's proof is not a gradient. In fact the vector field X in the proof of Theorem 3.1 can be constructed by the following recipe: First find a vector field Y such that $div\ Y = \hat{\beta} - \hat{\alpha}$, so that if $\gamma = \sum_1^k (-1)^{j-1} Y^j dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_k$, then $d\gamma = \beta - \alpha$. Such a vector field Y exists because $\int (\hat{\beta} - \hat{\alpha}) dx = 0$. We then set

$$X(x,t) = -\frac{Y(x)}{t\hat{\beta}(x) + (1-t)\hat{\alpha}(x)}.$$

In fact for the existence of Y, we may try a gradient $Y = \nabla u$ so that the scalar-valued function u satisfies $\Delta u = \hat{\beta} - \hat{\alpha}$. Again this equation has a solution because $\int (\hat{\beta} - \hat{\alpha}) dx = 0$.

(iii) There has been new developments for Moser's theorem. In fact the transformation φ in Theorem 3.1 is by no means unique. However, one may wonder whether or not a "nice" φ exists. More precisely, let (M, g) be a Riemannian manifold. Assume that M is compact, connected with no boundary. Using g we can talk about a Riemannian distance. More precisely, let $d(x, y)$ be the length of the geodesic distance between two points x and y.

We have a natural volume form Ω that is expressed by $(\det[g_{ij}])^{1/2}dx_1 \wedge \cdots \wedge dx_k$ in local coordinates. Consider two forms $\alpha = a\Omega$ and $\beta = b\Omega$ with $a, b > 0$ and $\int_M \alpha = \int_M \beta = 1$. Set

$$S(\alpha, \beta) = \{f : M \to M \text{ with } \varphi^*\alpha = \beta\}.$$

By Moser's theorem, this set is non-empty. Monge's problem searches for a function $f \in S(\alpha, \beta)$ which minimizes the cost function

$$\mathcal{I}(f) = \int_M c(x, f(x))\beta,$$

with $c(x, y)$ a suitable function of $M \times M$. If we choose $c(x, y) = \frac{1}{2}(d(x, y))^2$, then the minimizer φ is of the form $\varphi = \nabla w$ for a *convex* function w. This was shown by Brenier (1987, 1991) in the Euclidean case and by McCann in (2001) in the case of a Riemannian manifold. Brenier observed that such a minimizer can be used to find a *non-linear polar decomposition.* To explain this, let $F : U \to \mathbb{R}^k$ be an invertible integrable function with $\alpha = (F^{-1})^*\beta$ where $\beta = dx_1 \wedge \cdots \wedge dx_k$ and α is a volume form. According to Brenier's theorem, there exists a convex function ψ such that $(\nabla\psi)^*\alpha = \beta$. If we write $\rho = (\nabla\psi)^{-1} \circ F$, then $\rho^*\beta = F^*(\nabla\psi)^{-1*}\beta = F^*\alpha = \beta$. As a consequence, any function F can be decomposed as $F = \nabla\psi \circ \rho$ with ψ convex and ρ volume preserving. It turns out that polar decomposition implies the Hodge decomposition. To see this assume that $F^\epsilon(x) = x + \epsilon f(x)$ with ϵ small. We then expect to have $\psi^\epsilon(x) = \frac{1}{2}|x|^2 + \epsilon\varphi(x) + o(\epsilon)$ and $\rho^\epsilon(x) = x + \epsilon m(x) + o(\epsilon)$. Hence

$$\begin{aligned} x + \epsilon f(x) &= x + \epsilon\nabla\varphi(x + \epsilon m(x)) + \epsilon m(x) + o(\epsilon) \\ &= x + \epsilon(\nabla\varphi(x) + m(x)) + o(\epsilon). \end{aligned} \tag{3.11}$$

On the other hand, since ρ^ϵ is volume preserving,

$$1 = \det(id + \epsilon m') = 1 + \epsilon \operatorname{div}(m) + o(\epsilon).$$

From this and (3.11) we learn

$$f(x) = \nabla\varphi(x) + m(x) \text{ with } \operatorname{div} m = 0.$$

(iv) Theorem 3.1 can be used to show that the total volume is the only invariant of volume preserving diffeomorphisms. More precisely, if

$$\mathcal{V}(M) = \{\alpha : \alpha \text{ is a volume form on } M\}$$

with M compact and closed, and

$$c : \mathcal{V}(M) \to \mathbb{R}$$

is a function such that $c(\alpha) = c(\beta)$ whenever $\varphi^*\alpha = \beta$ for some diffeomorphism φ, then c must be a function of $\int_M \alpha$. $\square$

Given a manifold M with two symplectic forms α and β, we may wonder whether or not for some diffeomorphism φ, we have $\varphi^*\alpha = \beta$. By Theorem 3.1, this would be the case if $dim\ M = 2$ and $\int_M \alpha = \int_M \beta$. However, when $dim\ M = 2n \geq 4$, we may have non-isomorphic symplectic forms α and β with $\int_M \alpha^n = \int_M \beta^n$. To see how the proof of Theorem 3.1 breaks down, we note that we may fail to find a path of symplectic forms $\alpha(t)$ that connects α to β. Even if such a path exists, the equation $\mathcal{L}_X\alpha(t) = d\ i_X\alpha(t) = d\alpha(t)/dt$ requires that $d\alpha(t)/dt$ to be exact for $t \in [0,1]$. Though both of these issues can be handled locally as the next theorem demonstrate.

Theorem 3.2 *(Darboux) Let (M,ω) be a symplectic manifold of dimension $2n$ and take $x^0 \in M$. Then there exists an open set $U \subseteq \mathbb{R}^{2n}$ with $0 \in U$ and a diffeomorphism $\varphi : U \to M$ such that $\varphi(0) = x^0$ and $\varphi^*\omega = \bar{\omega}$.*

Proof Since M is locally diffeomorphic to an open subset of $\mathbb{R}^{2n}$, we may assume that $M = U_0 \subseteq \mathbb{R}^{2n}$ and $x^0 = 0 \in U$. In view of Proposition 2.1, we may also assume that $\omega_0 = \bar{\omega}$. Our goal is finding an open set $U \subseteq U_0$ with $0 \in U$, and a diffeomorphism $\varphi : U \to U$ with $\varphi(0) = 0$ and $\varphi^*\omega = \bar{\omega} = \omega_0$. Indeed if $\omega_x(v_1, v_2) = C(x)v_1 \cdot v_2$ and $\omega(t) = \bar{\omega} + t(\omega - \bar{\omega})$, for $t \in [0,1]$, then $\omega(t)_x(v_1, v_2) = C(x,t)v_1 \cdot v_2$, with $C(x,t) = \bar{J} + t(C(x) - \bar{J})$ and $C(0) = \bar{J}$. Clearly, we can find an open neighborhood $U = B_r(0)$ of 0 such that for $x \in U$ we have that $\|C(x) - \bar{J}\| < \|\bar{J}\| = 1$, which in turn guarantees that $C(x,t)$ is invertible for $(x,t) \in U \times [0,1]$. This means that in U, the form $\omega(t)$ is symplectic for all $t \in [0,1]$. We then search for a time-dependent vector field $X(\cdot, t)$ such that its flow ϕ_t satisfies

$$\phi_t^*\omega = \omega(t) \quad \text{for } t \in [0,1].$$

From differentiating both sides with respect to t and using Proposition 3.1, we learn

$$\omega - \bar{\omega} = \mathcal{L}_{X(\cdot,t)}\omega(t) = d\ i_{X(\cdot,t)}\omega(t). \tag{3.12}$$

In the ball $U = B_r(0)$, we can express $\omega - \bar{\omega} = d\alpha$ for a 1-form α such that $\alpha_0 = 0$. Hence (3.12) would follow if we can find a time-dependent vector field X such that $i_{X(\cdot,t)}\omega(t) = \alpha$. By nondegeneracy of $\omega(t)$, such a vector field X exists with $X(0,t) = 0$ for every $t \in [0,1]$. We are done. $\square$

Remark 3.2 If we write $\omega = u \cdot dx$ and $\bar{\omega} = \bar{u} \cdot dx$ near the origin, then X constructed in the proof has the form

$$X(x,t) = \mathcal{C}\big(tu(x) + (1-t)\bar{u}(x)\big)^{-1}(u(x) - \bar{u}(x)) = C(x,t)^{-1}(u(x) - \bar{u}(x)).$$

$\square$

As an immediate consequence of Darboux's theorem we learn that any symplectic manifold M of dimension $2n$ has a an atlas consisting of pairs $\{(U_j, h_j\}$ with $h_j : U_j \to \mathbb{R}^{2n}$, such that

$$h_i \circ h_j^{-1} : h_j(U_i \cap U_j) \to h_i(U_i \cap U_j)$$

is symplectic for every i and j. The family $\{(U_j, h_j)\}$ is an example of a *symplectic atlas* that always exists by Darboux's theorem.

Exercise 3.1 (i) Show that $\mathcal{L}_X = d \circ i_X + i_X \circ d$ satisfy

$$\mathcal{L}_X \circ d = d \circ \mathcal{L}_X, \quad \mathcal{L}_X(\alpha \wedge \beta) = (\mathcal{L}_X \alpha) \wedge \beta + \alpha \wedge (\mathcal{L}_X \beta).$$

(ii) Let (M, ω) be a $2n$-dimensional symplectic manifold. Assume that M is compact with no boundary. Then for every $j \in \{1, \dots, n\}$, there exists a closed $2j$-form which is not exact.

(iii) Let (M_1, ω^1) and (M_2, ω^2) be two symplectic manifolds. Define $(M_1 \times M_2, \omega^1 \times \omega^2)$ with

$$(\omega^1 \times \omega^2)_{(x,y)}((a_1, a_2), (b_1, b_2)) = \omega^1_x(a_1, b_1) + \omega^2_y(a_2, b_2).$$

Show that $(M_1 \times M_2, \omega^1 \times \omega^2)$ is symplectic.

(iv) Let (M, ω) be a symplectic manifold. Show that if $\varphi^*\omega = f\omega$ for some diffeomorphism $\varphi : M \to M$, and C^1 scalar function f, then either $dim\ M = 2$ or f is a constant function.

(v) Use polar coordinates in $\mathbb{R}^{2n}$ to write $q_i = r_i \cos\theta_i$, $p_i = r_i \sin\theta_i$, and let e_i (respectively f_i) denote the vector for which the q_i-th coordinate (respectively p_i-th coordinate) is 1 and any other coordinate is 0. Set

$$e_i(\theta_i) = (\cos\theta_i)e_i + (\sin\theta_i)f_i, \quad f_i(\theta_i) = -(\sin\theta_i)e_i + (\cos\theta_i)f_i.$$

Given a vector field u, we may write

$$u = \sum_{i=1}^{n} \left(a^i e_i(\theta_i) + b^i f_i(\theta_i)\right).$$

The form $\alpha = u \cdot dx$ can be written as

$$\alpha = \sum_{i=1}^{n} \left(a^i dr_i + r_i b^i d\theta_i\right) =: \sum_{i=1}^{n} \left(a^i dr_i + B^i d\theta_i\right).$$

Assume that all a^is and b^is depend on $r = (r_1, \dots, r_n)$ only. What are the the necessary and sufficient conditions on a and B in order for $\omega = d\alpha$ to be symplectic.

(vi) Let U_1 and $U_2 \subset B_r$ be two planar open sets with smooth boundaries that are diffeomorphic to a disc. Show that if $area(U_1) = U_2$, then there is an area preserving diffeomorphism $\psi : \mathbb{R}^2 \to \mathbb{R}^2$ such that $\psi(U_1) = U_2$ and $\psi(x) = x$ for $x \notin B_r$. Here $B_r = \{x : |x| < r\}$.
Hint: Construct ψ from three diffeomrphisms ψ_1, ψ_2 and ψ_3 with the following properties:

- $\psi_1 : \mathbb{R}^2 \to \mathbb{R}^2$ with $\psi_1(x) = x$, for $x \notin B_r$, $\psi_1(U_1) = U_2$, and ψ_1 is area-preserving on a small neigberhood of ∂U_1.
- $\psi_2 : \bar{U}_1 \to \bar{U}_2$ is area preserving and $\psi_2(x) = \psi_1(x)$ for x near ∂U_1.
- $\psi_3 : B_r \setminus \bar{U}_1 \to B_r \setminus \bar{U}_2$ is area-preserving and $\psi_3(x) = \psi_1(x)$ for x near $\partial U_1 \cup \partial B_r$.

$\square$

4 Contact Manifolds and Weinstein Conjecture

In this section we will give two motivations for studying contact manifolds. As our first motivation we observe that we can construct *exotic* symplectic forms on $\mathbb{R}^4$ from certain contact forms in $\mathbb{R}^3$. Our second motivation is the *Weinstein's conjecture* that predicts every compact nonsingular contact level set of a Hamiltonian function carries at least one periodic orbit of the corresponding Hamiltonian vector field.

We first give a simple recipe for constructing a symplectic form on $\hat{M} = M \times \mathbb{R}$ from certain 1-forms on a manifold M: Given a 1-form α on M, define $\hat{\alpha}_{(x,s)}(v,\tau) = e^s\alpha_x(v)$ for every $x \in M$, $v \in T_xM$, and $s, \tau \in \mathbb{R}$. There is nothing special about e^s and our approach is applicable if we replace e^s with a strictly increasing or decreasing C^1 function of s. The question is whether or not $\hat{\omega} = d\hat{\alpha}$ is symplectic.

Proposition 4.1 *Let M be any manifold of odd dimension and α any 1-form on M. Then the form $d\hat{\alpha}$ is symplectic iff $\ell_x(\alpha) \cap \xi_x(\alpha) = \{0\}$ for every $x \in M$, where*

$$\begin{aligned}
\ell_x &= \ell_x(\alpha) = \{v \in T_xM : \ d\alpha_x(v,w) = 0 \ \ \textit{for every } w \in T_xM\}\,, \\
\xi_x &= \xi_x(\alpha) = \ker \alpha_x = \{v \in T_xM : \ \alpha_x(v) = 0\}\,.
\end{aligned}$$

Proof We certainly have $d\hat{\alpha} = e^s\big(d\alpha - \alpha \wedge ds\big)$. As a result,

$$(d\hat{\alpha})_{(x,s)}\left((v,\tau),(w,\tau')\right) = e^s\big((d\alpha)_x(v,w) - \alpha_x(v)\tau' + \alpha_x(w)\tau\big).$$

From this we can readily show that if $v \in \ell_x \cap \xi_x$, then

$$(d\hat{\alpha})_{(x,s)}\left((v,0),(w,\tau')\right) = 0,$$

for every $(w,\tau') \in T_xM \times \mathbb{R}$.

Conversely, suppose that $\ell_x(\alpha) \cap \xi_x(\alpha) = \{0\}$, and that for some $(v,\tau) \in T_xM \times \mathbb{R}$, we have

$$(d\alpha)_x(v,w) - \alpha_x(v)\tau' + \alpha_x(w)\tau = 0, \tag{4.1}$$

for every $(w,\tau') \in T_xM \times \mathbb{R}$. We wish to show that $(v,\tau) = 0$. To see this, first vary τ' in (4.1), to deduce that $\alpha_x(v) = 0$, or $v \in \xi_x$. Hence we now have

$$(d\alpha)_x(v,w) + \alpha_x(w)\tau = 0, \tag{4.2}$$

for every $w \in T_xM$. If $\tau = 0$, then (4.2) means that $v \in \ell_x$. As a result, $v \in \ell_x \cap \xi_x = \{0\}$ and we are done. On the other hand, if $\tau \neq 0$, choose any non-zero $w \in \ell_x$ in (4.2) to deduce that $\alpha_x(w) = 0$, or $w \in \xi_x$, which contradicts our assumption $\ell_x \cap \xi_x = \{0\}$. (Note that since the dimension of M is odd, $\ell_x \neq \{0\}$.) This completes the proof. □

Note that since $\dim M = 2n-1$ is odd, the dimension of ℓ_x is at least 1. Hence the condition $\ell_x \cap \xi_x = \{0\}$ implies that $\dim \ell_x = 1$ and $\dim \xi_x = 2n-2$. Equivalently,

$$T_xM = \ell_x \oplus \xi_x, \tag{4.3}$$

for every $x \in M$.

Definition 4.1 The pair (M, α) of a manifold M and 1-form α is called a contact manifold if $\ell_x(\alpha) \cap \xi_x(\alpha) = \{0\}$ for every $x \in M$. The *Reeb vector field* $R = R^\alpha$ is the unique $R \in \ell_x$ such that $\alpha(R) = 1$. The R-projection onto ξ is denoted by $\pi = \pi^\alpha$; $\pi_x(v) = v - \alpha_x(v)R(x)$. □

One of the main interest in contact manifold is the following conjecture:

Weinstein Conjecture The Reeb vector field of a closed contact manifold has a closed (periodic) orbit. □

We next discuss the analog of Hamiltonian vector fields for contact manifolds.

Proposition 4.2 *Let (M, α) be a contact manifold, and let $H : M \to \mathbb{R}$ be a C^1 function. Set $\hat{H}(x, s) = e^s H(x)$. If $X_{\hat{H}}^{d\hat{\alpha}} = (Z, V)$, then the vector field Z and the function $V : M \to \mathbb{R}$ depend on x and are uniquely determined by the equations*

$$i_Z d\alpha = dH(R^\alpha)\alpha - dH, \qquad \alpha(Z) = H, \qquad V = -dH(R^\alpha). \tag{4.4}$$

Moreover, $\mathcal{L}_Z \alpha = -V\alpha$.

Proof By definition

$$i_{(Z,V)} d\hat{\alpha} = e^s \left(i_Z d\alpha - \alpha(Z)\ ds + V\alpha \right) = -e^s \left(H\ ds + dH \right).$$

This implies that $\alpha(Z) = H$ and $i_Z d\alpha + V\alpha = -dH$. By evaluating both sides at R^α we learn that $V = -dH(R^\alpha)$, completing the proof of (4.4). The vector field Z is uniquely determined by the first two equations of (4.4) because the restriction of $d\alpha$ to ξ is symplectic, and $\pi(Z)$ satisfies

$$i_{\pi(Z)} d\alpha = i_Z d\alpha = dH(R^\alpha)\alpha - dH := \nu,$$

with $\nu(R) = 0$. Finally,

$$\mathcal{L}_Z \alpha = i_Z d\alpha + d\big(\alpha(Z)\big) = i_Z d\alpha + dH = dH(R^\alpha)\alpha = -V\alpha,$$

by (4.4). □

Definition 4.2 Given a contact manifold (M, α), we say that a vector field Z is an α-*contact*, if $\mathcal{L}_Z \alpha + V\alpha = 0$, for some function $V : M \to \mathbb{R}$. Moreover, given a C^1 Hamiltonian function H, the unique α-contact vector field Z and scalar-valued function V satisfying(4.4) are denoted by $Z_H = Z_H^\alpha$ and $V_H^\alpha = V_H$ respectively. $\square$

Remark 4.1 (i) Note that if the flow of an α-contact vector filed Z is denoted by ϕ_t, then

$$\frac{d}{dt}\phi_t^*\alpha = \phi_t^*\mathcal{L}_Z\alpha = -\phi_t^*(V\alpha) = -(V\circ\phi_t)\phi_t^*\alpha,$$

which implies

$$\left(\phi_t^*\alpha\right)_x = e^{-\int_0^t \phi_\theta(x)d\theta}\alpha_x.$$

(ii) Since $\mathcal{L}_{R^\alpha}\ \alpha = 0$, the Reeb vector field is an example of an α-contact vector field. In fact the flow of R^α preserves α. $\square$

Example 4.1 When $M = \mathbb{R}^3$, the form $\alpha = u\cdot dx$ is contact if and only if $\rho = (\nabla\times u)\cdot u$ is never 0. Indeed,

$$d\alpha(v_1, v_2) = \big((\nabla\times u)\times v_1\big)\cdot v_2 =: \big[(\nabla\times u), v_1, v_2\big]$$

which implies that when $(\nabla\times u)(x) \neq 0$, the set ℓ_x is the line spanned by $(\nabla\times u)(x)$. In this case the Reeb vector field is given by $R = \rho^{-1}(\nabla\times u)$, and

$$\alpha\wedge d\alpha = \rho\ dx_1\wedge dx_2\wedge dx_3, \tag{4.5}$$

is a volume form. Moreover, if we write $\mathcal{L}_Z(u\cdot dx) = \big(\mathcal{L}'_Z u\big)\cdot dx$, then

$$\mathcal{L}'_Z u = \nabla(u\cdot Z) + (\nabla\times u)\times Z. \tag{4.6}$$

The contact vector field associated with H is given by

$$\rho X_H = u\times\nabla H + H(\nabla\times u). \tag{4.7}$$

More generally, if $M = \mathbb{R}^{2n-1}$ with $n \geq 2$, we may express a 1-form α as $\alpha = u\cdot dx$ for a vector field u. Moreover

$$\beta_x(v_1, v_2) = d\alpha_x(v_1, v_2) = C(u)v_1\cdot v_2,$$

where $C(u) = Du - (Du)^*$. The form α is contact iff u never vanishes, and the null space of $C(u)(x)$ is not orthogonal to u. The set ℓ_x is one dimensional and R^α is the unique vector in ℓ_x such that $u(x)\cdot R(x) \equiv 1$. Writing $u^\perp$ and $R^\perp$ for the space of vectors perpendicular to u and R respectively, then $\xi = u^\perp$, and we may define a matrix $C'(u)$ which is not exactly the inverse of $C(u)$ (because $C(u)$ is not invertible), but it is specified uniquely by two requirements:

- (i) $C'(u)$ restricted to $R^{\perp}$ is the inverse of $C(u) : u^{\perp} \to R^{\perp}$.
- (ii) $C'(u)u = 0$.

The α-contact vector field associated with H is given by

$$X_H = -C'(u)\nabla H + HR. \tag{4.8}$$

□

As Example 4.1 and (4.5) indicates, in dimension 3, a form α is contact iff $\alpha \wedge d\alpha$ is a volume form. More generally we have the following elementary result.

Proposition 4.3 *Let α be a 1-form on a manifold M. Then α is contact iff $\alpha \wedge (d\alpha)^{n-1}$ is a volume form.*

Proof. Assume that (M, α) is contact and set $\gamma = (d\alpha)^{n-1}$. Since the restriction of $d\alpha$ to ξ_x is non-degenerate and $\dim \xi_x = 2(n-1)$, we may use Exercise 2.1**(i)** to assert that $(d\alpha)_x^{n-1}$ is a volume form on ξ_x. More specifically, we may choose a symplectic basis $\{e_1, \dots, e_{n-1}, f_1, \dots, f_{n-1}\}$ for ξ_x so that if their duals are denoted by

$$\{e_1^*, \dots, e_{n-1}^*, f_1^*, \dots, f_{n-1}^*\} =: \{dx_1 \dots, dx_{2n-2}\}$$

then $d\alpha_x = \sum_i f_i^* \wedge e_i^*$. From this we deduce

$$\gamma_x = (n-1)!\; dx_1 \wedge \cdots \wedge dx_{2n-2}.$$

To show that $\alpha \wedge \gamma$ is a volume form, observe

$$(\alpha \wedge \gamma)(v_1, \dots, v_{2n-1}) = \sum_{i=1}^{2n-1} (-1)^{i-1} \alpha(v_i)\gamma(v_1, \dots, \hat{v}_i, \dots, v_{2n-1})$$

which means that if dz denotes the dual of the vector R, then

$$\alpha \wedge \gamma = (n-1)!\; dz \wedge dx_1 \wedge \cdots \wedge dx_{2n-2}.$$

The converse is left as an exercise. (See Exercise 4.1**(ii)**.) □

If (N, ω) is a symplectic manifold and M is a closed co-dimension 1 submanifold of N, we may wonder whether or not we can find a contact form α on M, $\delta > 0$, and a diffeomorphism $\psi : \hat{M} = M \times (-\delta, \delta) \to U \subseteq M$, such that $\psi^*\omega = \hat{\omega}$, where $\hat{\omega} = d\hat{\alpha}$, for $\hat{\alpha} = e^s\alpha$. In words, a neighborhood U of N in M is isomorphic to $(\hat{M}, \hat{\omega})$. As we will see below that this is possible whenever M is compact and there is a vector field X that plays the role of $\frac{\partial}{\partial s}$. The point is that if the flow of the vector field $\frac{\partial}{\partial s}$ is denoted by ϕ_t, then $\phi_t(x,s) = (x, s+t)$ and

$$\phi_t^* \hat{\alpha} = e^{s+t}\alpha = e^t \hat{\alpha}.$$

Hence $\phi_t^* \hat{\omega} = e^{s+t} d\alpha = e^t \hat{\omega}$, which means

$$\left(\psi \circ \phi_t \circ \psi^{-1}\right)^* \omega = e^t \omega.$$

As a result, if

$$X = \psi_* \left(\frac{\partial}{\partial s}\right),$$

then

$$\left(\phi_t^X\right)^* \omega = e^t \omega. \tag{4.9}$$

In fact the existence of a vector field X that satisfies (4.9) is exactly what we need for ω to be isomorphic to $\hat{\omega}$ in a neighborhood of M.

Definition 4.3 Let X be a vector field on a symplectic manifold (M, ω). We say X is an *ω-Liouville vector field* if $\mathcal{L}_X \omega = d i_X \omega = \omega$. Equivalently, X is ω-Liouville if (4.9) is valid. $\square$

Remark 4.2 Note that if a Liouville vector field X exists, then we can define a 1-form $\alpha' = i_X \omega$ which satisfies $d\alpha' = \omega$. Moreover, since $\alpha'(X) = \omega(X, X) = 0$, we also have

$$\mathcal{L}_X \alpha' = i_X d\alpha' = i_X \omega = \alpha', \qquad \left(\phi_t^X\right)^* \alpha' = e^t \alpha'. \tag{4.10}$$

$\square$

Theorem 4.1 *Let (N, ω) be a symplectic manifold of dimension $2n$, and let M be compact closed submanifold of M with $\dim M = 2n-1$. The following statements are equivalent:*

- **(i)** *There exists a neighborhood U of M in N and an ω-Liouville vector field X such that X is transverse to M.*
- **(ii)** *There exists a contact form α on M such that $d\alpha$ is the restrictions of ω to M.*
- **(iii)** *There exists a contact form α on M, $\delta > 0$, and a diffeomorphism $\psi : \hat{M} = M \times (-\delta, \delta) \to U \subseteq M$, with $\psi(x, 0) = 0$, such that if $\hat{\alpha} = e^s \alpha$ and $\alpha' = \left(\psi^{-1}\right)^* \hat{\alpha}$, then $d\alpha' = \omega$.*

Proof Suppose that **(i)** is true. In other words, there exists a Liouville vector field X such that $X(x) \notin T_xM$ for every $x \in M$. Set $\alpha' = i_X\omega$ so that $d\alpha' = \omega$. Write α for the restriction of α' to M. Set $\xi_x = \ker \alpha$, and

$$\ell_x := \{v \in T_xM : \ \omega_x(v, w) = 0 \quad \text{for all } w \in T_xM\},$$

for $x \in M$. We wish to show that $\ell_x \cap \xi_x = \{0\}$. Pick any $v \in \ell_x \cap \xi_x$. By definition, $\omega_x(v, w) = 0$ for all $w \in T_xM$. On the other hand $\omega_x(X(x), v) = \alpha_x(v) = 0$. Hence $\omega_x(v, a) = 0$ for all $v \in T_xN$ because X transverses to M. Since ω_x is symplectic, we deduce that $v = 0$. Thus α is contact. In summary **(i)** implies **(ii)**.

The converse is carried out in several steps.

Step1. Starting from a contact form α on M with $d\alpha = \omega_{|M}$, we wish extend α to α' with $d\alpha' = \omega$, and construct a Liouville vector field X in a neighborhood of M. First we construct X on M. Recall that $R = R^\alpha$ denotes the Reeb vector field and $\xi_x = \ker \alpha_x \subset T_xM$ is of dimension $2n - 2$. Since we hope to find X satisfying $\alpha = i_X\omega$ on M, the restriction of such X to M much satisfy $\omega_x(X(x), v) = \alpha(v)$ for every $v \in T_xM$. As a result, for every $x \in M$ and $v \in \xi_x$, we must have

$$\omega_x(X(x), R(x)) = 1, \qquad \omega_x(X(x), v) = 0. \tag{4.11}$$

It is not hard to construct a vector field X that satisfies (4.11). For example, we may use Proposition 4.2 to take an almost complex structure J and a Riemannian metric g so that $\omega(v_1, v_2) = g(Jv_1, v_2)$. We then search for a vector field X of the form

$$X(x) = f(x)J_xR(x) + Y(x), \qquad Y(x) \in \xi_x,$$

with $f : M \to \mathbb{R}$ a scalar-valued function. To satisfy (4.11), we need

$$f = -g(R, R)^{-1}, \qquad i_Y\omega = -f\big(i_{JR}\, \omega\big).$$

The latter equation has a unique solution $Y \in \xi$ because $d\alpha_{|\xi}$ is symplectic. We note that the first requirement in (4.11) implies that $X(x) \notin T_xM$ for every $x \in M$. In summary, we have constructed a vector field X that transverses to M and satisfies $i_X\omega = \alpha$ on M.

Step2. So far our vector field $X(x)$ is defined only on M. Since M is compact and X transverses to M, we can use the g-exponential map to define $\varphi(x, s) = \exp_x(sX(x))$, so that $\varphi : M \times (-\delta, \delta) \to U \subseteq N$ is a diffeomorphism. Here $\varphi(x, s) = \gamma(s)$, where γ is the unique geodesic with $\gamma(0) = x$ and $\dot\gamma(0) = X(x)$. (Alternatively, extend X to a vector field $\hat X$ that is defined on a neighborhood of M, and set $\varphi(x, s) = \phi_s^{\hat X}(x)$.) This map has the desired property on M;

$$\begin{aligned}(\varphi^*\omega)_{(x,0)}\big((v_1,\tau_1),(v_2,\tau_2)\big) &= \omega_x\big(d\varphi_{(x,0)}(v_1,\tau_1), d\varphi_{(x,0)}(v_2,\tau_2)\big)\\ &= \omega_x\big(v_1+\tau_1 X(x), v_2+\tau_2 X(x)\big)\\ &= \omega_x(v_1,v_2)+\tau_1\omega_x(X(x),v_2)+\tau_2\omega_x(v_1,X(x))\\ &= \omega_x(v_1,v_2)+\tau_1\alpha(v_2)-\tau_2\alpha(v_1)\\ &= (\omega+ds\wedge\alpha)_{(x,0)}\big((v_1,\tau_1),(v_2,\tau_2)\big),\end{aligned}$$

because $\varphi_{(x,0)}(v,\tau)=\tau X(x)+v$. Hence, if we define $\hat\alpha=e^s\alpha$ and $\hat\omega=d\hat\alpha$ on $\hat M = M\times(-\delta,\delta)$, then $\varphi^*\omega=\hat\omega$ on $M=M\times\{0\}$. As a result, if we set $\omega'=\varphi^*\omega$, then $\eta:=\omega'-\hat\omega$ is a closed form on $\hat M$ such that $\eta_{|M}=0$. As we will see in Lemma 4.1 below, the form η is necessarily exact. In fact, there exists a 1-form β such that $\beta_{|M}=0$ and $d\beta=\eta$. Hence, $\omega'=d(\hat\alpha+\beta)$ with $(\hat\alpha+\beta)_{|M}=\alpha$. From this we learn that if $\alpha'=\big(\varphi^{-1}\big)^*(\hat\alpha+\beta)$, then $d\alpha'=\omega$, and $\alpha'_{|M}=\alpha$.

Step3. So far we now that there exists a neighborhood U of M, and a 1-form α' that is defined in U and satisfies

$$\alpha'_{|M}=\alpha,\qquad d\alpha'=\omega.$$

Since ω is symplectic, we can find a vector field X in U such that $i_X\omega=\alpha'$. This vector field is a Liouville vector field because $di_X\omega=d\alpha'=\omega$. Moreover, since $\omega_x(X(x),v)=\alpha_x(v)$, for every $x\in M$ and $v\in T_xM$, we learn that $X(x)\notin T_xM$ for every $x\in M$; otherwise $\alpha_x(R(x))=0$ which is impossible. This completes the proof of **(i)**.

It remains to show the equivalence of **(i)** and **(iii)**. If **(i)** is true, simply define $\psi(x,s)=\phi_s^X(x)$. Note that by (4.10), we have $\phi_s^*\alpha'=e^s\alpha'$. Since

$$\hat\phi_t(x,s):=\big(\psi^{-1}\circ\phi_t\circ\psi\big)(x,s)=(x,t+s),$$

we deduce

$$\psi^*X=\frac{\partial}{\partial s}.$$

On the other hand, if $\hat\alpha=\psi^*\alpha'$, then $\big(\hat\phi_s\big)^*\hat\alpha=e^s\hat\alpha$, which implies that $\hat\alpha_{(x,s)}(v,\tau)=e^s\alpha_x(v)$, because $\hat\phi$ is the flow of $\partial/\partial s$. We are done.

Conversely, if **(iii)** is true, then the Liouville vector field is given by

$$X=\big(\psi^{-1}\big)^*\frac{\partial}{\partial s}.$$

□

We continue with the proof of a Poincaré-type lemma that was used in the proof of Theorem 4.1.

Lemma 4.1 *Let η be a closed l-form on $\hat{M} = M \times (-\delta, \delta)$ with $\eta_{|M\times\{0\}} = 0$. Then there exists an $(l-1)$-form β on $\hat{M}$ such that $d\beta = \eta$ and $\beta_{|M\times\{0\}} = 0$.*

Proof Define $\Phi_\theta(x,s) = (x, e^\theta s)$ for $y = (x,s) \in \hat{M}$ and $\theta \in \mathbb{R}$. Note that $\frac{d}{d\theta}\Phi_\theta(y) = Y(\Phi_\theta(y))$, for $Y(x,s) = (0,s)$. Let us simply write M for $M \times \{0\}$. Since $\Phi_0(x,-\infty) = (x,0)$, we have $\Phi_{-\infty}^*\eta = \eta_{|M} = 0$ by our assumption. We now write

$$\begin{aligned}\eta = \Phi_0^*\eta - \Phi_{-\infty}^*\eta &= \int_{-\infty}^0 \frac{d}{d\theta}\Phi_\theta^*\eta \; d\theta = \int_{-\infty}^0 \Phi_\theta^*\mathcal{L}_Y\eta \; d\theta \\ &= d\left[\int_{-\infty}^0 \Phi_\theta^* i_Y\eta \; d\theta\right] =: d\beta.\end{aligned}$$

For example, when $l = 2$,

$$\begin{aligned}\beta_{(x,s)}(v,\tau) &= \int_{-\infty}^0 \eta_{(x,e^\theta s)}\big((0,e^\theta s),(v,e^\theta\tau)\big)\; d\theta \\ &= \int_{-\infty}^0 \eta_{(x,e^\theta s)}\big((0,s),(v,e^\theta\tau)\big)\; e^\theta d\theta \\ &= \int_0^1 \eta_{(x,\theta s)}\big((0,s),(v,\theta\tau)\big)\; d\theta.\end{aligned}$$

From this we learn that β is well-defined and that $\beta_{|M} = 0$. The case of general l can be treated in the same way. □

Definition 4.2 If the assumptions of Theorem 4.1 are satisfied, we say that the submanifold M is of *contact type*. □

Remark 4.3 An important consequence of Theorem 4.1 is that a neighborhood of a submanifold M of contact type is isomorphic to $\big(\hat{M}, \hat{\omega}\big)$. This means that such a neighborhood can be foliated into submanifolds that are, in some sense isomorphic to M. More precisely, if ϕ_t is the flow of the corresponding Liouville vector field X, then the hypersurfaces $\big(M^s = \phi_s(M) : \; s \in (-\delta,\delta)\big)$ are all of contact type. In fact if we write $\ell(M^s)$ for the corresponding line bundle,

$$\ell_x(M^s) = \{v \in T_xM^s : \; \omega_x(v,w) = 0 \text{ for all } w \in T_xM^s\},$$

and $R^s(x)$ for the Reeb vector field associated with $\big(M^s, \alpha'_{|M^s}\big)$, then we have the following identities:

$$(4.12) \qquad (d\phi_s)_x\big(\ell_x(M^0)\big) = \ell_{\phi_s(x)}(M^s), \qquad (d\phi_s)_x(R^0(x)) = e^sR^s(\phi_s(x)).$$

The second equation is a consequence of (4.10):

$$e^s = e^s\alpha_x(R^0(x)) = e^s\alpha'_x(R^0(x)) = \left(\phi_s^*\alpha'\right)_x(R^0(x)) = \alpha'_{\phi_s(x)}\big((d\phi_s)_x(R^0(x))\big).$$

We may also define a Hamiltonian function $K = K_M : U \to (-\delta,\delta)$ with $K(M^s) = e^s$. In other words, if ψ is as in Theorem 4.1**(iii)** and $\hat{H} : \hat{M} \to (-\delta,\delta)$ is defined by $\hat{H}(x,s) = e^s$, then

$$K = \hat{H} \circ \psi^{-1}. \tag{4.13}$$

We can readily show

$$X_{\hat{H}}^{\hat{\omega}}(x,s) = (R(x),0).$$

From this we deduce

$$X_K^{\omega}(\phi_s(x)) = (d\psi)_{(x,s)}(R(x),0) = (d\phi_s)_x(R(x)) = e^s R^s(\phi_s(x)),$$

because by Proposition 3.2 $X_K^{\omega} \circ \psi = (d\psi)X_{\hat{H}}^{\hat{\omega}}$. In summary,

$$(X_K^{\omega})_{|M^s} = e^s R^s. \tag{4.14}$$

Hence the periodic orbits of K coincide with the closed orbits of $\big(R^s : \ s \in (-\delta,\delta)\big)$. □

As we learned from Remark 4.2, the hypersurfaces of contact-type may be regarded as the level sets of a Hamiltonian function for which the corresponding Hamiltonian vector field is closely related to the Reeb vector field. Now imagine that we start with a $2n$ dimensional symplectic manifold (N,ω) and a Hamiltonian vector field X_H^{ω}, and wonder whether or not level sets of H carry periodic orbits. We note that when the Hamiltonian function H is independent of time, then the level sets of H are conserved because

$$\frac{d}{dt}\left(H \circ \phi_t^H\right) = dH\left(X_H^{\omega} \circ \phi_t^H\right) = -\omega\big(X_H^{\omega}, X_H^{\omega}\big) \circ \phi_t^H = 0.$$

Let us write $M(c) = \{x : \ H(x) = c\}$, and set $\ell_x(c) := (T_xM(c))^{\amalg}$. We call $M = M(c)$ *regular* if $dH_x \neq 0$ for all $x \in M(c)$. Since $\dim T_xM(c) = 2n-1$ for a regular $M(c)$, we use Proposition 2.1**(i)** to assert

$$\dim \ell_x(c) = 1.$$

On the other hand, since

$$T_xM(c) = \{w : \ (dH)_x(w) = 0\},$$

we deduce that indeed $X_H(x) \in \ell_x(c)$, for $c = H(x)$ because for every $w \in T_xM(c)$,

$$\omega_x(X_H(x),w) = -(dH)_x(w) = 0.$$

This means that the line l_x is parallel $X_H(x)$. What we learn from this is that the existence of a periodic orbit of X_H on $M = M(c)$ is a property of M and ω and does not depend on H. In other words, if M possesses a periodic orbit of X_H and if $M = \{x : H'(x) = c'\}$ for another regular H', then $X_{H'}$ possesses a periodic orbit as well. Evidently $l_x = (T_xM)^{\text{II}}$ offers an H-independent candidate for the tangent lines to the orbit. More precisely, if M is a closed hypersurface of a symplectic manifold (N, ω), define $\ell_x(M) = (T_xM)^{\text{II}}$ which is a line. We always have $\ell_x(M) \subseteq T_xM$ because if $v \in \ell_x(M)$ but $v \notin T_xM$, then $v \,\text{II}\, \big(\ell_x(M) \oplus T_xM\big) = T_xN$ which implies $v = 0$ because ω is symplectic. As a result, $L_M = \bigsqcup_{x \in M} l_x$ is a line bundle of M, and we can express our question of the existence of periodic orbits purely in terms of this line bundle. The existence of periodic orbits is now reduced to the existence of closed characteristics of the line bundle L_M. It turns out that there are hypersurfaces M and symplectic structures ω such that the corresponding line bundle has no closed characteristics. However Weinstein's conjecture asserts that if M is of contact type, then such a closed characteristic can be found. In view of (4.14), the existence of a closed orbit of the Reeb vector field follows from the existence of a periodic orbit of the vector field X_K where K was defined by (4.13). From the preceding discussion and (4.14) we learn that we only need to find a periodic orbit in a neighborhood of a hypersurface of contact type:

Proposition 4.4 *Let M be a compact hypersurface of contact type of the symplectic manifold (M, ω). Let $K = K_M : U \to \mathbb{R}$ be the corresponding Hamiltonian function that is defined for a neighborhood U of M as in* (4.13). *If the vector field X_K^ω has a periodic orbit in U, then the Reeb vector field of M has a closed orbit.*

Remark 4.4 As we mentioned in the introduction, Weinstein's conjecture has been established for the hypersurfaces of $(\mathbb{R}^{2n}, \bar{\omega})$ by Viterbo. As we will see in Section 6 below, we will use *Hofer-Zehnder* Capacity to give a rather straightforward proof of Viterbo's theorem. The point is that if Hofer-Zehnder Capacity of a neighborhood U of a contact type hypersurface M is positive, then by very definition of this capacity, a periodic orbit would exist for every Hamiltonian vector field with compact support inside U. From this it is not hard to deduce that X_{K_M} also possesses periodic orbits. We then use Proposition 4.4 to conclude that M carries a closed orbit. It is not known how to use the same line of reasoning to settle Weinstein's conjecture for general (N, ω); it is not known how to show the positivity of Hofer-Zehnder Capacity of open sets of general symplectic manifolds. Though the Weinstein's conjecture has now been established for all closed 3-dimensional manifolds by Taubes (2007) employing a variant of *Seiberg-Witten Floer homology.* □

Starting from a symplectic manifold (N, ω) with $\dim N = 2n$, and a Hamiltonian function $H : N \to \mathbb{R}$ for which the level set $M = \{x : \; H(x) = c\}$ is compact and regular, we have seen that the line bundle $\ell_x(c) = (T_xM)^{\text{II}}$ is well-defined. If we already know that $\omega = d\alpha'$ for a 1-form α', in a neighborhood U of M, then using the fact that ω is symplectic, we learn that there exists a unique vector field X such that $i_X\omega = \alpha'$. On the other hand

$$i_{X^\omega_H}\omega = -dH, \quad \text{and} \quad i_X\omega = \alpha' \quad \Rightarrow \quad \alpha'\big(X^\omega_H\big) = dH(X).$$

Hence

$$X(x) \notin T_xM \ \text{ for all } \ x \in M \quad \Leftrightarrow \quad \alpha = \alpha'_{|M} \ \text{ is of contact type.}$$

If $M = M(c)$ is of contact type, then we also have a natural volume form $\alpha\wedge(d\alpha)^{n-1}$ that is defined on M and we expect it to be invariant under the flow of X^ω_H. Indeed, if $(\hat{M}, \hat{\omega})$ with $\hat{\omega} = d\hat{\alpha}$ is as in Theorem 4.1, then

$$\hat{\Omega} := (n!)^{-1}\ \hat{\omega}^n = (n!)^{-1}\ e^{ns}\ ds\wedge\alpha\wedge(d\alpha)^{n-1} = d\hat{H}\wedge m,$$

where $\hat{H} = e^s$, and

$$m = (n!)^{-1}\ e^{(n-1)s}\ \alpha\wedge(d\alpha)^{n-1}.$$

Using ψ as in Theorem 4.1, we can write

$$\Omega := (n!)^{-1}\ \omega^n = dH\wedge\mu, \tag{4.15}$$

where $H = \hat{H}\circ\psi^{-1}$, and $\mu = \big(\psi^{-1}\big)^* m$. In particular, $\mu_{|M}$ is a volume form for M when M is of contact type. We now assert that in general we can always find a $2n-1$- form on M that satisfies (4.15).

Proposition 4.5 *Let (N,ω) be a symplectic manifold with $\dim N = 2n$, and assume that $H : N \to \mathbb{R}$ is a function for which the level set $M = \{x :\ H(x) = c\}$ is compact and regular.*

(i) *There exists a $(2n-1)$-form μ such that the volume form $\Omega = (n!)^{-1}\ \omega^n$ can be expressed as*

$$\Omega = dH\wedge\mu, \tag{4.16}$$

in a neighborhood of M.

(ii) *The form $\mu_{|M}$ is uniquely determined by (4.16): If $\omega = dH\wedge\mu = dH\wedge\mu'$ then $\mu_{|M} = \mu'_{|M}$.*

(ii) $\big(\phi^H_t\big)^*\mu_{|M} = \mu_{|M}$. *In words, μ is invariant for ϕ^H_t, restricted to M.*

Proof **(i)** The existence of μ is an immediate consequence of the non-degeneracy of $dH \neq 0$ on M. (See Exercise 4.1**(vi)**.)

(ii) Write $j : M \to N$ for the inclusion map. We certainly have $dH\wedge(\mu-\mu') = 0$. From Exercise 4.1**(vii)** below we deduce that for some $2n-2$ form γ, $\mu-\mu' = dH\wedge\gamma$. As a result, $\mathcal{I}(\mu-\mu') = \mathcal{I}(dH\wedge\gamma) = (\mathcal{I}dH)\wedge\mathcal{I}\gamma = 0$ because $\mathcal{I}dH = d(H\circ j) = 0$.

(iii) Write ϕ_t for ϕ_t^H. By (4.16),

$$\Omega = \phi_t^*\Omega = (\phi_t^* dH) \wedge \phi_t^*\mu = d(H \circ \phi_t) \wedge \phi_t^*\mu = dH \wedge \phi_t^*\mu.$$

By uniqueness result of **(ii)**, we deduce $\mathcal{I}\phi_t^*\mu = \mathcal{I}\mu$, or $(\phi_t \circ j)^*\mu = \mathcal{I}\mu$. We may write $\phi_t \circ j = j \circ \phi_t$ where the second $\phi_t : M \to M$ is the restriction of ϕ_t to M. Hence $\Phi_t^*\mathcal{I}\mu = \mathcal{I}\mu$, as desired. □

Remark 4.4 When $(N, \omega) = (\mathbb{R}^{2n}, \bar{\omega})$, we can describe the form μ of (4.16) explicitly. To ease the notation, let us write $k = 2n$. Note that if $\{g_j : j = 1, \dots, k\}$ is a collection of smooth functions with $\sum_{j=1}^k g_j H_{x_j}(-1)^{j-1} \equiv 1$, then we may choose for μ,

$$\mu = \sum_{j=1}^{k} g_j dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_k.$$

For example $g_j = (-1)^{j-1} H_{x_j}/|\nabla H|^2$ would do the job.) We now claim

$$\int_\Gamma \alpha = c_0 \int_\Gamma |\nabla H|^{-1} d\sigma, \tag{4.17}$$

for a constant c_0. To see this, let us assume $H_{x_k} \neq 0$ so that locally we can find a function $f(x_1, \dots, x_{k-1})$ with

$$H(x_1, \dots, x_{k-1}, f(x_1, \dots, x_{k-1})) = c.$$

This means that the graph of f gives a parametrization of M and that $f_{x_j} = -\frac{H_{x_j}}{H_{x_k}}$ for $j = 1, \dots, k-1$. On the other we can use the function f to express the volume measure on M that is induced by the standard volume of $\mathbb{R}^k$. More precisely, consider the vector $a = (a_1, \dots, a_k)$ with

$$a_j = \frac{\partial(x_1, \dots, \hat{x}_j, \dots, x_k)}{\partial(x_1, \dots, x_{k-1})} = (-1)^{k-j} \frac{H_{x_j}}{H_{x_k}},.$$

for $1 \le j \le k$. If $\Gamma = \{(x_1, \dots, x_{k-1}, f(x_1, \dots, x_{n-1})) : (x_1, \dots, x_k) \in U\}$, then

$$\int_\Gamma \mu = \int_U \left(\sum_{j=1}^k a_j g_j \right) dx_1, \dots, dx_{k-1}.$$

On the other hand

$$\bar{a} = \left(\sum_{j=1}^k a_j^2 \right)^{1/2} = \frac{|\nabla H|}{|H_{x_k}|}.$$

Since locally either $H_{x_k} > 0$ or $H_{x_k} < 0$, we have

$$\begin{aligned}\int_\Gamma \mu &= \int_U (-1)^{k-1} \frac{1}{H_{x_k}}\ dx_1, \dots, dx_{k-1} \\ &= \pm \int_U \frac{1}{|\nabla H|} \bar{a}\ dx_1, \dots, dx_{k-1} = \pm \int_\Gamma \frac{1}{|\nabla H|}\ d\sigma,\end{aligned}$$

proving (4.17). Note that this is consistent with the coarea formula

$$\int f\ dx_1, \dots, dx_k = \int_{-\infty}^{\infty} \left[\int_{H=c} f \frac{d\sigma}{|\nabla H|} \right] dc.$$

□

We now give examples of contact type hypersurfaces.

Example 4.2(i) Let $N = T^*Q$ for an n-dimensional manifold Q, and let $\omega = d\lambda$ denotes its standard symplectic form as was defined in Example 3.1. The level set M of $H : T^*Q \to \mathbb{R}$ is of contact type if

$$\lambda_{(q,p)}\big(X_H(q,p)\big) = p\left((d\pi)_{(p,q)}\left(X_H(q,p)\right)\right) \neq 0, \tag{4.18}$$

on M. When $Q = \mathbb{R}^n$, then (4.19) means that $p \cdot H_p \neq 0$ on M.

(ii) Let M be as in **(i)** and assume that $H : T^*Q \to R$ is homogeneous of degree $l > 0$ in the p-variable: for every $r > 0$,

$$H(q, rp) = r^l H(q,p). \tag{4.19}$$

We claim that this condition is equivalent to

$$i_{X_H}\lambda = lH. \tag{4.20}$$

The proof of the equivalence of (4.19) and (4.20) is straightforward when $Q = \mathbb{R}^n$: By differentiating both sides of (4.19) we can show that H is homogeneous of degree $l > 0$ iff $p \cdot H_p = lH$, which is exactly (4.20). The proof of general Q is similar. Use the Darboux charts we defined in Example 3.1, namely take $h : U \to \mathbb{R}^n$ as a chart of Q and use this to construct $\bar{h} : T^*U \to \mathbb{R}^{2n}$ so that $\lambda = \bar{h}^*\bar{\lambda}$. Define $\bar{H} : h(U) \times \mathbb{R}^n \to \mathbb{R}$ so that $\bar{H} \circ \bar{h} = H$. By the construction,

$$\bar{h}(q,p) = (h(q), \bar{p}) \quad \Rightarrow \quad \bar{h}(q, rp) = (h(q), r\bar{p}),$$

for every $r > 0$. This implies that H is homogeneous of degree $l > 0$ iff $\bar{H}$ is homogeneous of degree $l > 0$. Similarly,

$$i_{X^{\bar{\omega}}_{\bar{H}}}\lambda = l\bar{H} \quad \Leftrightarrow \quad \bar{h}^*\left(i_{X^{\bar{\omega}}_{\bar{H}}}\lambda\right) = l\bar{H} \circ \bar{h} \quad \Leftrightarrow \quad i_{X_H}\lambda = lH.$$

This completes the proof of the equivalence of (4.19) and (4.20). From this and (4.18) we deduce that $M = \{x : \ H(x) = c\}$ is always of contact type for every homogeneous H and non-zero c. $\square$

Remark 4.5 We note that the Hamiltonian vector field X_H on T^*Q preserves the 1-form λ if and only if $H(q,p)$ is (positively) homogeneous of degree 1 in p. Indeed by Proposition 3.1, the form λ is preserved by the flow of X_H iff

$$0 = \mathcal{L}_{X_H}\lambda = d \circ i_{X_H}\lambda + i_{X_H} \circ d\lambda = d\big(i_{X_H}\lambda - H\big).$$

This is equivalent to assert that $i_{X_H}\lambda - H$ is a constant. Since adding a constant does not change X_H, we may assume that the constant is zero so that the condition now is $i_{X_H}\lambda = H$. In view of the equivalence of (4.19) and (4.20), we deduce that $\mathcal{L}_{X_H}\lambda = 0$ is equivalent to the degree-1 p-homogeneity of H. $\square$

The positions of a conservative system are points in an n-dimensional manifold M that is known as the configuration space. The phase space in Lagrangian formulation is the tangent bundle TQ. The motion is determined by a Lagrangian $L : TQ \to \mathbb{R}$. In the Hamiltonian formulation we use a Hamiltonian function $H : M = T^*Q \to \mathbb{R}$ to determine the motion of the system. In our next example we study the level set of classical Hamiltonians.

Example 4.3 Observe that the Riemannian metric g on a manifold Q, allows us to define an operator $\sharp : T^*Q \to TQ$ that maps 1-forms to vector fields by requiring

$$g_q\left((\sharp p)_q, v\right) = p_q(v),$$

for every vector $v \in T_qM$. This duality also induces a metric on T^*Q by

$$G_q(p,p') = g_q(\sharp p, \sharp p').$$

Given a smooth potential energy $V : Q \to \mathbb{R}$, we define

$$H(q,p) = \frac{1}{2}G_q(p,p) + V(q). \tag{4.21}$$

We now claim

$$\lambda\left(X_H^\omega\right) = G_q(p,p), \tag{4.22}$$

for H given by (4.21). The proof is very similar to the equivalence of (4.19) and (4.20); using h and $\bar{h}$ as in Examples 3.1 and 4.2, we can find a metric $\bar{G}$ on $h(U)$ and a function $\bar{V} : h(U) \to \mathbb{R}$ such that $V = \bar{V} \circ h$, and $H = \bar{H} \circ \bar{h}$, where

$$\bar{H}(q,p) = \frac{1}{2}\bar{G}_q(p,p) + \bar{V}(q).$$

Writing $\bar{G}_q(p,p) = A(q)p \cdot p$ with $A > 0$, we deduce that $X^{\bar{\omega}}_{\bar{H}} = (S(q)p, -\nabla V(q))$. From this we can readily deduce

$$\bar{\lambda}\,(X^{\bar{\omega}}_{\bar{H}}) = \bar{G}_q(p,p).$$

We then apply $\bar{h}^*$ to both sides to arrive at (4.22). Using the elementary Lemma 4.10, we can now readily show that if

$$M_E = \{(q,p) : H(q,p) = E\}$$

is compact and regular, and if $E > \max V$, then M_E is of contact type simply because for such E we always have $p \neq 0$, whenever $(q,p) \in M_E$. It turns out that M_E is of contact type even when $E \le \max V$ (See [HZ]). It is worth mentioning that when $E > \max V$, then we can define a new Riemannian metric

$$\hat{G}_q(p,p) = \frac{G_q(p,p)}{E - V(q)},$$

that is known as the *Jacobi metric*. We then have

$$M_E = \{(q,p) = H(q,p) = E\} = \left\{(q,p) = \frac{1}{2}\hat{G}_q(p,p) = 1\right\}.$$

The Hamiltonian $\hat{H}(q,p) = \frac{1}{2}\hat{G}_q(p,p)$ induces a Hamiltonian vector field $X_{\hat{H}}$. It is simply related to X_H by

$$X_H(q,p) = G_q(p,p)X_{\hat{H}}(q,p), \ (q,p) \in S. \tag{4.23}$$

Hence a periodic orbit exists on M_E for X_H if the same is true for $X_{\hat{H}}$. Geometrically, $X_{\hat{H}}$ generates the geodesic flow defined by the Jacobi metric $\hat{G}$ on $M = T^*Q$. □

We next discuss some examples of non-contact types hypersurfaces.

Example 4.3 (Zehnder [Z])

(i) Let $\hat{M} = \mathbb{T}^3 \times I$ where $\mathbb{T}^3$ is the 3-dimensional torus and I is an open interval. We write $\theta = (\theta_1, \theta_2, \theta_3) \in \mathbb{T}^3$ for points in $\mathbb{T}^3$ and define $H : \hat{M} \to I$ by $H(\theta, s) = s$. The level sets of H are $M_s = \mathbb{T}^3 \times \{s\}$ for $s \in I$. Writing $(x_1, \dots, x_4)$ for $(\theta_1, \theta_2, \theta_3, s)$, any constant skew-symmetric $C = [c_{ij}]$, define a 2-form $\omega = \sum_{i<j} c_{ij} dx_i \wedge dx_j$. Note that each $d\theta_i$ is closed (but not exact), and ds is exact. Hence ω is always closed because C is constant. As a result, ω is symplectic iff C is invertible. Given a vector $\zeta = (\zeta_1, \zeta_2, \zeta_3)$ with $\zeta_3 \neq 0$, we define C by

$$-C^{-1} = \hat{C} := \begin{bmatrix} 0 & 1 & 0 & \zeta_1 \\ -1 & 0 & 0 & \zeta_2 \\ 0 & 0 & 0 & \zeta_3 \\ -\zeta_1 & -\zeta_2 & -\zeta_3 & 0 \end{bmatrix}.$$

We further assume that ζ is *irrational* in the following sense:

$$a \in \mathbb{Z}^3, \ a \cdot \zeta = 0 \quad \Rightarrow \quad a = 0.$$

Evidently, $X_H = X_H^\omega = \hat{C}\nabla H$ that in our case leads to $X_H = (\xi, 0)$. Hence $\phi_t^H(\theta, s) = (\theta + t\zeta, s)$ where $\theta + s\zeta$ is understood as a $(mod\ 1)$ summation; here we identify $\mathbb{T}$ as the interval $[0,1]$ with 0 and 1 regarded as the same point. The irrationality ζ guarantees that X_H has no periodic orbit. Hence $\omega_{|M_s}$ does not induce a contact structure.

(ii) Let us write $x = (q_1, q_2, p_1, p_2)$ for angles in a 4-dimensional torus $\mathbb{T}^4$ that may be defined as $[0,1]^4$ with $0 = 1$. Take a 1-periodic function $f : \mathbb{R} \to \mathbb{R}$ that can be regarded as a function on the circle $\mathbb{T}$. Let $\omega = dp_1 \wedge dq_1 + dp_2 \wedge dq_2 + \zeta dp_1 \wedge dq_2$ for an irrational number ζ. Take $H(x) = f(p_1)$ so that $X_H^\omega(x) = (f'(p_1), \zeta f'(p_1), 0, 0)$. The corresponding flow is

$$\phi_t^H(x) = (q_1 + tf'(p_1), q_2 + \zeta f'(p_1), p_1, p_2),$$

with $(mod\ 1)$ additions. Again since ζ is irrational, $\phi_t^H(x)$ is never periodic whenever $f'(p_1) \neq 0$. $\square$

As we mentioned in the beginning of this section, contact structures in $\mathbb{R}^3$ can be used to produce exotic structures. The process of going from a contact manifold (M, α) to the symplectic manifold $(\hat{M}, \hat{\omega})$ as in Proposition 4.1 is called the *symplectization.* For example we equip

$$\mathbb{R}^{2n-1} = \{(q, p, z) : \ q, p \in \mathbb{R}^{n-1}, \ z \in \mathbb{R}\},$$

with the so-called *standard contact form* $\bar{\alpha} = p \cdot q + dz$, then the corresponding symplectization $(\mathbb{R}^{2n}, \hat{\omega})$ is isomorphic to the standard $(\mathbb{R}^{2n}, \bar{\omega})$. However if we equip $\mathbb{R}^3$ with an *overtwisted* contact form, the symplectization yields a symplectic form on $\mathbb{R}^4$ that is not equivalent to the standard symplectic structure. Before giving the definition of overtwised contact structures, let us observe that if (M, α) is contact, then $\xi = \ker \alpha$ is never integrable. The reason is that the integrability of ξ is equivalent to $\alpha \wedge d\alpha = 0$ (see Lemma A.4 in Appendix), whereas the contact assumption requires $\alpha \wedge (d\alpha)^{n-1}$ to be a volume form by Proposition 4.3. In fact we have a following results for integrals of ξ:

Proposition 4.6 *Let (M, α) be a contact manifold of dimension $2n-1$. If Γ is an l dimensional submanifold of M with $T_x\Gamma \subseteq \xi_x = \ker \alpha_x$ for every $x \in \Gamma$, then $\dim \Gamma \leq n-1$.*

Proof Using Lemma A.3 of the Appendix, we can readily show that if X and Y are two vector fields that are tangent to Γ, then $d\alpha(X, Y) = 0$. Simply because for $X, Y \in \xi$, we have $\alpha(X) = \alpha(Y) = 0$, and from $X, Y \in T\Gamma$ we know that $[X, Y] \in T\Gamma$ by Frobenius' theorem. This means that if we use the symplectic form $\omega = d\alpha_{|\xi}$ for orthogonality, then $T_x\Gamma \subseteq (T_x\Gamma)^{\amalg}$, which by Proposition 2.1 implies

$$2n - 2 = \dim T_x\Gamma + \dim(T_x\Gamma)^{\amalg} \geq 2 \dim T_x\Gamma.$$

This clearly implies that $\dim T_x\Gamma \leq n-1$. $\square$

Definition 4.3 We call a submanifold Γ of a $2n-1$ dimensional contact manifold (M, α) Legendrian if $T_x\Gamma \subset T_xM$ and $\dim \Gamma = n-1$. $\square$

Let us now examine contact forms on $\mathbb{R}^3$. If α is a contact form on $\mathbb{R}^3$ and $\xi = \ker \alpha$ is the corresponding contact structure, we expect that the plan field ξ experience some twists as we move along a hypersurface because ξ is not integrable. Given a 2 dimensional surface Γ, we may define a plane field $\gamma_x = T_x\Gamma \cap \xi_x$. For generic points of Γ, γ_x is a line and the integrals of this line bundle have been used to classify the contact structures in $\mathbb{R}^3$ or even arbitrary 3-dimensional manifolds. The simplest way to produce examples of exotic contact structures in $\mathbb{R}^3$ is using cylindrical coordinates; $q = r\cos\theta, p = r\sin\theta$ so that if $x = (q, p, z)$, then

$$dq = \cos\theta \; dr - \sin\theta \; rd\theta, \qquad dp = \sin\theta \; dr + \cos\theta \; rd\theta.$$

Example 4.4(i) If $\alpha = qdp - pdq - dz = r^2 \; d\theta - dz$, then for $x = (q, p, z)$,

$$\xi_x = \{(\hat{q}, \hat{p}, \hat{z}) : \; q\hat{p} - p\hat{q} = \hat{z}\}.$$

Observe that for $\Gamma = \{z = 0\}$, the set $\gamma_x = T_x\Gamma \cap \xi_x$ is a line only when $x \neq 0$. More precisely,

$$\gamma_x = \{(\hat{q}, \hat{p}, 0) : \; q\hat{p} - p\hat{q} = 0\},$$

is the ray $\{(sq, sp, 0) : \; s \in \mathbb{R}\}$ if $x \neq 0$, whereas $\xi_0 = \Gamma$. So the line bundle $\{\gamma_x : \; x \in \Gamma\}$ has a singularity at 0. For example, on the $\{(s, 0, 0) : \; s > 0\}$, the vector $u = (-p, q, -1)$ twists from $-\pi/2$ to 0 as s goes from 0 to ∞.

(ii) Let $\beta = \sin r \; rd\theta + \cos r \; dz$, so that

$$\xi_x = \{(\hat{q}, \hat{p}, \hat{z}) : \; (r\tan r)\hat{\theta} = \hat{z}\},$$

where $(1 + \tan^2\theta)\hat{\theta} = q^{-2}(qdp - pdq)$. Note that the curve $\{r = \pi, \; z = 0\}$ is Legendrian. If we set

$$\Gamma_\varepsilon = \{(q, p, \varepsilon) : \; r \leq \pi,$$

then the interior Γ_ε has no singular point for $\varepsilon > 0$ and $\partial\Gamma$ is Legendrian. We say a contact form is over twisted if such $\Gamma = \Gamma_\varepsilon$ exists for β. Note that for this particular example, if we move along a ray emanating from the origin, the plane ξ_x make complete turns infinitely many times. $\square$

Exercise 4.1

(i) Verify (4.6)-(4.8).

(ii) Let M be a manifold of dimension $2n-1$ and let α be a 1-form on M. Show that if $\alpha \wedge (d\alpha)^{n-1}$ is a volume form, then α is contact.

(iii) Consider $(\mathbb{R}^{2n}, \bar{\omega})$ and let S be the boundary of a set A that is star-shaped with respect to the origin. Assume that each ray emanating from the origin intersect S at exactly one point. Show that S is of contact type. **Hint:** Show that $X(x) = \frac{1}{2}x$ is a Liouville vector field.

(iv) Consider the contact manifold $(\mathbb{R}^{2n-1}, \bar{\alpha})$. Let $H(q, p, z)$ be a smooth function. Show that the corresponding contact vector field Z_H is given by

$$(H_p, -H_q + pH_z, -p \cdot H_p + H).$$

(v) Let S^{2n-1} be a unit sphere in $\mathbb{R}^{2n}$. The form $\bar{\lambda}$ induces a 1-form $\bar{\mu}$ on S^{2n-1}. Show that $(S^{2n-1}, \bar{\mu})$ is not a contact manifold. Define $\mu = \frac{1}{2}(p \cdot dq - q \cdot dp)$. Show that (S^{2n-1}, μ) is a contact manifold. Find its Reeb flow.

(vi) Let Ω be a volume form and let τ be a non-degenerate 1-form. Show that there exists a form β such that $\Omega = \tau \wedge \beta$.

(vii) Let η be an l-form and τ a 1-form with $\tau \wedge \eta = 0$. Show that $\eta = \tau \wedge \gamma$ for some $(l-1)$-form γ.

(viii) Verify (4.22). □

5 Variational Principle and Convex Hamiltonian

In this section, we use variational techniques to prove the following result of A. Weinstein:

Theorem 5.1 *Assume that the hypersurface $S \subseteq \mathbb{R}^{2n}$ is the smooth boundary of a compact strictly convex region. Then the Reeb's vector field on S has a periodic orbit.*

In fact Theorem 5.1 allows us to define some kind of a *symplectic width* or *capacity* of convex sets K with nonempty interiors:

$$\mathbf{c}_0(K) := \inf\big\{|A(\gamma)|;\ \gamma \text{ is a Reeb characteristic of } \partial K\big\}, \tag{5.1}$$

where the *symplectic action* $A(\gamma)$ of γ was defined by (3.1). Of course Theorem 5.1 guarantees that $\mathbf{c}_0(K) < \infty$ for every bounded convex set K of nonempty interior. Our strategy for proving Theorem 5.1 is by figuring out an alternative *dual* like variational principle of (5.1). Once this dual problem is formulated, we can readily establish its finiteness.

Example 5.1 Assume that U is an ellipsoid with $0 \in U$. We learned in Chapter 2 that there are radii $r_1 \le \dots \le r_n$ and a linear symplectic T such that $T(U) = E$, where $E = \{x : H(x) \le 1\}$, with

$$H(x) = \sum_{j=1}^{n} \frac{q_j^2 + p_j^2}{r_j^2}.$$

The corresponding Hamiltonian flow is $\phi_t(z_1, \dots, z_n) = \big(e^{-i\lambda_1 t} z_1, \dots, e^{-i\lambda_n t} z_n\big)$, where $z_j = q_j + ip_j$, and $\lambda_j = 2/r_j^2$. For $y = (z_1, \dots, z_n)$ and $\gamma : [0, T] \to \mathbb{R}^{2n}$, defined by $\gamma(t) = \phi_t(y)$, we have

$$\begin{aligned} A(\gamma) &= \frac{1}{2}\int_0^T \bar{\omega}(\gamma(s), \dot{\gamma}(s))\ ds = \frac{1}{2}\int_0^T \sum_j Im\left(e^{-i\lambda_j s} z_j\ i\lambda_j e^{i\lambda_j s} \bar{z}_j\right)\ ds \\ &= \frac{1}{2} T \sum_j \lambda_j |z_j|^2 = TH(\gamma(0)). \end{aligned}$$

If this γ lies on ∂E, we must have $H(\gamma(0)) = 1$ which means that $A(\gamma) = T$ for such γ. Now if γ is a periodic orbit of period $T > 0$, then we must have

$$z_j \neq 0 \quad \Rightarrow \quad \lambda_j T = 2k\pi, \qquad \text{for some } k \in \mathbb{N}.$$

This implies $T \ge 2\pi/\lambda_1 = \pi r_1^2$. Hence, $\mathbf{c}_0(E) \ge \pi r_1^2$. On the other hand, if we choose $(z_1, \dots, z_n)$ satisfying $|z_1| = r_1$, and $z_j = 0$ for $j \neq 1$, we have that $A(\gamma) = T = \pi r_1^2$ for the corresponding γ. From all this we deduce that for an ellipsoid E,

$$\mathbf{c}_0(E) = \pi r_1(E)^2. \tag{5.2}$$

$\square$

Before embarking on finding the dual problem associated with (5.1), let us first review the Lagrangian formulation of Hamiltonian systems. As we will see below, this formulation is not well-suited for finding periodic characteristics on the boundary of a convex set. Nonetheless we will learn some fruitful ideas from it that will play essential role later in showing the finiteness of $\mathbf{c}_0(U)$.

To study Newton's equation with constraints, Lagrange initiated variational formulation of conservative mechanical problems. Let $L : TQ \to \mathbb{R}$ be a C^1-function. Given $q^0, q^1 \in Q$, we define $\mathcal{B} : \Gamma_T(q^0, q^1) \to \mathbb{R}$ with

$$\begin{aligned}\Gamma_T(q^0, q^1) &= \{\gamma : [0,T] \to Q \text{ is } C^1 \text{ and } \gamma(0) = q^0,\ \gamma(T) = q^1\},\\ \mathcal{B}(\gamma) &= \int_0^T L(\gamma(t), \dot{\gamma}(t))dt.\end{aligned}$$

Let us denote the argument of L by (q, v) with $v \in T_qQ$ and write L_q and L_v for the partial derivatives of L.

We now claim that if $q(\cdot)$ is a critical point of $\mathcal{B}$, then q solves the Euler–Lagrange–Newton's equation

$$\frac{d}{dt} L_v(q, \dot{q}) = L_q(q, \dot{q}). \tag{5.3}$$

Indeed if we use the L^2-inner product, then the derivative of $\mathcal{B}$ is given by

$$\partial\mathcal{B}(\gamma) = L_q(\gamma, \dot{\gamma}) - \frac{d}{dt} L_v(\gamma, \dot{\gamma}), \tag{5.4}$$

for every $\gamma \in \Gamma_T(q^0, q^1)$. To see this, take any $\eta = \eta(\theta, t)$ with $\theta \in (-\delta, \delta)$ and $t \in [0, T]$, such that $\eta_\theta(0, t) = \tau(t)$, and $\eta(0, \cdot) = \gamma(\cdot)$. Now

$$\begin{aligned}\langle \partial\mathcal{B}(\gamma), \tau\rangle &:= \frac{d}{d\theta}\mathcal{B}(\eta(\theta, \cdot))\bigg|_{\theta=0}\\ &= \int_0^T [L_q(\gamma, \dot{\gamma}) \cdot \tau + L_v(\gamma, \dot{\gamma}) \cdot \dot{\tau}]\ dt\\ &= \int_0^T \left[L_q(\gamma, \dot{\gamma}) - \frac{d}{dt} L_v(\gamma, \dot{\gamma})\right] \cdot \tau\ dt.\end{aligned}$$

Here we used the fact that $\tau(0) = \tau(T) = 0$ which follows from $\eta(\theta, 0) = q^0$, $\eta(\theta, T) = q^1$ for all θ. As an example, let $L(q, v) = \frac{m}{2}|v|^2 - V(q)$ in $\mathbb{R}^n$. Then the equation (5.3) reads as $m\ddot{q} = -\nabla V(q)$.

To explain the connection between Lagrangian formulation with the the Hamiltonian, let us assume that $Q = \mathbb{R}^n$ and that for $p \in \mathbb{R}^n = T_q^*Q$, we can solve the relationship

$$p = L_v(q, v)$$

uniquely for v. Denoting the solution by $v(q, p)$, and setting

$$H(q, p) = p\big(v(p, q)\big) - L(q, v(q, p)),$$

we learn

$$\begin{aligned} H_p &= v + pv_p - v_p L_v = v + pv_p - pv_p = v, \\ H_q &= pv_q - L_q(q, v) - L_v(q, v)v_q = -L_q(q, v). \end{aligned}$$

Hence

$$p = L_v(q, v) \quad \Leftrightarrow \quad H_p(q, p) = v, \tag{5.5}$$

and

$$H_q(q, p) = -L_q(q, v). \tag{5.6}$$

If we set $p(t) = L_v(q(t), \dot{q}(t))$, then by (5.3), (5.5), and (5.6),

$$\dot{q} = H_p(q, p), \quad \dot{p} = -H_q(q, p).$$

The inversion (5.5) is possible if we assume that L is strictly convex in v. In this case H can be constructed from L by *Legendre transform*:

$$H(q, p) = \sup_v (p \cdot v - L(q, v)).$$

More generally, if $L : TQ \to \mathbb{R}$ is convex in v, we may define $H : T^*Q \to \mathbb{R}$ by

$$H(q, p) = \sup_v \big(p(v) - L(q, v)\big).$$

This relationship maybe inverted to yield

$$L(q, v) = \sup_p \big(p(v) - H(q, p)\big).$$

If we assume that L has superlinear growth in variable v as $|v| \to \infty$, then H is finite and the supremum is attained at $v = v(q, p)$ so that (5.5) is valid.

For a v-convex L, we may find solutions to (5.1) by minimizing $\mathcal{B}$. For example, if for some constants $c_1, c_2 > 0$ and $\alpha > 1$,

$$L(q,v) \geq c_1|v|^\alpha - c_2, \tag{5.7}$$

for every $v \in T_qN$ and $q \in N$, and if γ_l is a sequence in $\Gamma_T(q^0, q^1)$ such that $\lim_{l\to\infty} \mathcal{B}(\gamma_l) = A = \inf \mathcal{B}$, then by (5.7) we have the bound

$$\sup_l \int_0^T |\dot{\gamma}_l(t)|^\alpha dt < \infty.$$

This bound allows us to extract a subsequence of γ_l which converges weakly with respect to the topology of Sobolev space $W^{1,\alpha}$. It turns out that $\mathcal{B}$ is lower semicontinuous because L is convex in v. This allows us to deduce that for any limit point $q(\cdot)$ of the sequence $\gamma_l(\cdot)$, we have that $\mathcal{B}(q) \leq A$. Since $A = \inf \mathcal{B}$, we learn that $\mathcal{B}(q) = A$ and that the infimum is achieved.

In spite of the appeal of the above argument, it is not clear how to use it to prove Theorem 5.1. Recall that we are searching for a periodic solution on a given energy surface. Of course we could have chosen $q^0 = q^1$ so that our solution satisfies $q(0) = q(T)$. But to lie on the surface S we need to make sure that $(q(0), p(0)) \in S$. In fact the solution we have found may not even be periodic in (q,p) coordinates because we cannot guarantee $p(0) = p(T)$.

To this end let us define another functional $\mathcal{A}$ of which critical points solve the Hamiltonian system. Set $\mathcal{A} : \Gamma_T \to \mathbb{R}$ to be

$$\begin{aligned}\mathcal{A}(x(\cdot)) &= \int_{x(\cdot)} [p \cdot dq - H(x)]\ dt \\ &= \int_0^T \left[\frac{1}{2}\bar{J}x \cdot \dot{x} - H(x)\right] dt,\end{aligned}$$

where

$$\Gamma_T = \{x : \mathbb{R} \to \mathbb{R}^{2n} \text{ is } C^1 \text{ and } T\text{-periodic}\}.$$

We have

$$\partial\mathcal{A}(x(\cdot)) = -\bar{J}\dot{x} - \nabla H(x), \tag{5.8}$$

because

$$\begin{aligned}\langle \partial\mathcal{A}(x(\cdot)), \tau(\cdot)\rangle &= \frac{d}{d\delta}\mathcal{A}(x+\delta\tau)\bigg|_{\delta=0} \\ &= \int_0^T \left[\frac{1}{2}\bar{J}\tau \cdot \dot{x} + \frac{1}{2}\bar{J}x \cdot \dot{\tau} - \nabla H(x)\cdot\tau\right]\ dt \\ &= -\int_0^T (\bar{J}\dot{x} + \nabla H(x)) \cdot \tau\ dt\end{aligned}$$

for every $x, \tau \in \Gamma_T$. From (5.8) we learn that $\partial \mathcal{A}(x(\cdot)) \equiv 0$ iff x solves

$$\dot{x} = \bar{J} \nabla H(x). \tag{5.9}$$

Note that $\mathcal{A}$ involves H explicitly and no additional assumption such as convexity is needed. However typically the critical points of $\mathcal{A}$ are saddle points and it is helpless to search for (local) maximizers or minimizers. Because of this, finding critical points for $\mathcal{A}$ is far more challenging. Before discovering a remedy for this, let us show that a hypersurface S as in Theorem 5.1 can be realized as a level set of a convex homogeneous Hamiltonian function.

Definition 5.1 Let $U \subseteq \mathbb{R}^{2n}$ be an open convex set with $0 \in U$. Set $K = \bar{U}$ for its closure.

(i) The *gauge function* associated with K is defined by

$$g_K(x) = \|x\|_K = \inf\{r > 0 : \ x/r \in K\} = \inf\{r > 0 : \ x/r \in U\}.$$

(ii) The *polar set* associated with K is defined by

$$K^o := \{x : \ x \cdot y \leq 1 \quad \text{for all } y \in K\}.$$

(iii) The *support function* h_K associated with K is defined by

$$h_K(x) := \sup_y \{x \cdot y : \ y \in K\}. \tag{5.10}$$

□

Proposition 5.1 *Let $U \subseteq \mathbb{R}^{2n}$ be an open convex set with $0 \in U$ and $K = \bar{U}$.*

(i) *The gauge function $\|\cdot\|_K$ is a norm. Its Legendre transform is given by*

$$\sup_y \left(x \cdot y - \|y\|_K\right) = \begin{cases} 0 & \text{if} \quad x \in K^o; \\ \infty & \text{otherwise,} \end{cases}$$

(ii) *We have*

$$K = \{y : \ x \cdot y \leq h_K(x) \quad \text{for all } x\}. \tag{5.11}$$

(iii) *We have $K^{\circ\circ} = K$. Moreover,*

$$\|x\|_{K^\circ} = h_K(x), \qquad \|x\|_K = h_{K^\circ}(x). \tag{5.12}$$

(iv) *If $H_K(x) = \frac{1}{2}\|x\|_K^2$, then its Legendre transform is given by $H_K^* = \frac{1}{2}h_K^2 = H_{K^\circ}$.*

Proof The proof of **(i)** and **(ii)** are left as an exercise. As for **(iii)**, observe

$$\begin{aligned}\|x\|_{K^\circ} &= \inf\{\rho > 0 : \ x/\rho \in K^\circ\} = \inf\{\rho : \ x \cdot y \le \rho \text{ for every } y \in K\} \\ &= \inf\{\rho : \ h_K(x) \le \rho\} = h_K(x).\end{aligned}$$

For the second equality, observe that since $x \cdot y \le \|x\|_K$ for every $y \in K^\circ$, we have

$$h_{K^\circ}(x) = \sup\{x \cdot y : \ y \in K^\circ\} \le \|x\|_K.$$

On the other hand, by (5.11),

$$\begin{aligned}\|x\|_K &= \inf\{\rho : \ x/\rho \in K\} = \inf\{\rho : \ \rho^{-1}(x \cdot z) \le h_K(z) \text{ for every } z\} \\ &= \inf\{\rho : \ (x \cdot z/h_K(z)) \le \rho \text{ for every } z\} = \sup_z (x \cdot z/h_K(z)) \\ &\le \sup_{y \in K^\circ} (x \cdot y) = h_{K^\circ}(x),\end{aligned}$$

because $z/h_K(z) \in K^\circ$ always. This completes the proof of second equality in (5.12). Note that two equalities of (5.12) imply that $\|x\|_K = \|x\|_{K^{\circ\circ}}$. As an immediate consequence we conclude that $K = K^{\circ\circ}$.

We now turn to **(iv)**. We certainly have

$$H_K^*(x) = \sup_y \left(x \cdot y - \frac{1}{2}\|y\|_K^2 \right) = \sup_y \sup_{t \ge 0} \left(tx \cdot y - \frac{t^2}{2}\|y\|_K^2 \right).$$

If $x \cdot y > 0$, then the t-supremum is attained at $t = (x \cdot y)/\|y\|_K^2$; otherwise the t-supremum is 0. As a result,

$$\begin{aligned}2H_K^*(x) &= \sup_y \frac{((x \cdot y)^+)^2}{\|y\|_K^2} = \sup_y \left[\left(x \cdot \frac{y}{\|y\|_K} \right)^+ \right]^2 \\ &= \big[\sup\{x \cdot y : \ \|y\|_K = 1\}\big]^2 = \big[\sup\{x \cdot y : \ \|y\|_K \le 1\}\big]^2 \\ &= \big[\sup\{x \cdot y : \ y \in K\}\big]^2 = h_K(x)^2,\end{aligned}$$

as desired. □

We next address the regularity of the function H_K.

Lemma 5.1 *Let $U \subseteq \mathbb{R}^{2n}$ be an open strictly convex set with $0 \in U$, $S = \partial U$, and $K = \bar{U}$. If S is C^2, then H_K is C^1 and strictly convex.*

Proof Write $F(x) = \|x\|_K$, $2H = 2H_K = F^2$. Note that F is not differentiable at 0. On the other hand,

$$D^2F(x)|_{T_xS} > 0, \qquad D^2F(x)x = 0 \tag{5.13}$$

for $x \in S$. The latter follows from the homogeneity of F; $x \cdot \nabla F = F$ and $D^2F(x)x = 0$. To see (5.13), let us look at the curvatures of S. Note that $n(x) = \frac{\nabla F(x)}{|\nabla F(x)|}$ is the unit normal to S and

$$Dn(x) = |\nabla F(x)|^{-1}D^2F(x) + \nabla F(x) \otimes \nabla(|\nabla F(x)|^{-1}).$$

As a result,

$$Dn(x)a \cdot a = |\nabla F(x)|^{-1}D^2F(x)a \cdot a + (\nabla F(x) \cdot a)(\nabla|\nabla F(x)|^{-1} \cdot a).$$

As a result, if $a \in T_xS$, then

$$Dn(x)a \cdot a = |\nabla F(x)|^{-1}D^2F(x)a \cdot a. \tag{5.14}$$

Since S is strictly convex, we have $Dn(x) > 0$. This and (5.14) imply (5.13).

Evidently H is C^1 and C^2 off the origin. Since H is homogeneous of degree 2,

$$\begin{aligned} \nabla H &= F\nabla F, \\ D^2H &= \nabla F \otimes \nabla F + FD^2F. \end{aligned}$$

Hence,

$$D^2H(x)a \cdot a = (\nabla F(x) \cdot a)^2 + F(x)D^2F(x)a \cdot a.$$

Note that if $a = b+c$ with $b\|x$ and $c \in T_xS$, then $(\nabla F(x)\cdot a)^2 > 0$ if $b \neq 0$ and $D^2F(x)a\cdot a > 0$ if $b = 0$ and $c \neq 0$. Hence $D^2H(x) > 0$ for all $x \neq 0$. □

Let us set $H = H_K$ and study the corresponding functional $\mathcal{A}$. Note that $\mathcal{A} = \mathcal{A}_T$ is defined for T-periodic functions whereas for Theorem 5.1 we need a periodic orbit on $S = \{H = 1/2\}$ of some period. Of course if $x(\cdot)$ is such a periodic orbit of period T, then $y(t) = x\,(Tt)$ is 1-periodic and

$$\dot{y} = T\bar{J}\nabla H(y). \tag{5.15}$$

In view of the form of functional $\mathcal{A}_T$, perhaps we should fix the period to be 1 always and now insist that $y(\cdot)$ solves (5.15) for some T. As a result, we now want to find a critical point of

$$T^{-1}\int_0^1 \left[\frac{1}{2}\bar{J}y \cdot \dot{y} - TH(y)\right]\,dt = \frac{1}{2T}\int_0^1 \bar{J}y \cdot \dot{y}\;dt - \int_0^1 H(y)\;dt$$

with y a 1-periodic path and some $T \in \mathbb{R}$. The scalar $\eta = T^{-1}$ resembles a Lagrange multiplier that now is employed for a functional defined on an infinite dimensional space. Motivated by this, define $\mathcal{C} : \Lambda \to \mathbb{R}$ by

$$\mathcal{C}(x(\cdot)) = \int_0^1 H(x(t))\;dt$$

with

$$\Lambda = \left\{ x : \mathbb{R} \to \mathbb{R}^{2n} \text{ is } C^1, \text{ 1-periodic and } \int_0^1 \bar{J}x \cdot \dot{x}\ dt = 1 \right\}.$$

It is not hard to show that $\Lambda \neq \emptyset$.

Lemma 5.2 **(i)** *Let $y(\cdot)$ be a non-constant critical point of $\mathcal{C} : \Lambda \to \mathbb{R}$. Then either $\nabla H(y) \equiv 0$, or there exists a constant T such that $\dot{y} = T\bar{J}\nabla H(y)$.*

(ii) *If $H = H_K$, then any critical $y(\cdot)$ satisfies $\dot{y} = T\bar{J}\nabla H(y)$ for some $T > 0$. Moreover, $z(t) = \sqrt{T}y(t/T)$ is a T-periodic solution of $\dot{z} = \bar{J}\nabla H(z)$, such that $A\big(z(\cdot)\big) = T/2$ and $2H(z) \equiv 1$.*

Proof First, let us determine the space $T_{y(\cdot)}\Lambda$. Take a path $z : (-\delta, \delta) \to \Lambda$ with $z(\cdot, 0) = y(\cdot)$ and $z_\theta(\cdot, 0) = \tau(\cdot)$. The condition

$$\int_0^1 \bar{J}z(\cdot, \theta) \cdot z_t(\cdot, \theta)\ dt = 1$$

can be differentiated with respect to θ to yield

$$0 = \int_0^1 \left[\bar{J}y \cdot \dot{\tau} + \bar{J}\tau \cdot \dot{y}\right]\ dt = -2\int_0^1 \bar{J}\dot{y} \cdot \tau\ dt.$$

In other words, for $\tau \in T_{y(\cdot)}\Lambda$, we always have

$$\int_0^1 \bar{J}\dot{y} \cdot \tau\ dt = 0. \tag{5.16}$$

The converse is also true; if a 1-periodic function τ satisfies (5.16), then the path $z(t, \theta) = y(t) + \theta\tau(t)$ lies in Λ and satisfies $z_\theta(\cdot, 0) = \tau(\cdot)$.

Now, if $y(\cdot)$ is critical for $\mathcal{C}$, then $\frac{d}{d\theta}\mathcal{C}(z(\cdot, \theta))\big|_{\theta=0} = 0$. This means that $y(\cdot)$ is critical iff

$$\int_0^1 \bar{J}\dot{y} \cdot \tau\ dt = 0 \quad \Rightarrow \quad \int_0^1 \nabla H(y) \cdot \tau\ dt = 0. \tag{5.17}$$

In particular if $y \in \Lambda$ satisfies $\nabla H(y) \equiv 0$, it is critical. Assume that $\nabla H(y)$ is not identically 0. We then use (5.17) to deduce that for every 1-periodic τ satisfying (5.16), and every $T \in \mathbb{R}$, we have

$$\int_0^1 \big(\bar{J}\dot{y} + T\nabla H(y)\big) \cdot \tau\ dt = 0. \tag{5.18}$$

We wish to select $\tau = \bar{J}\dot{y} + T\nabla H(y)$ to deduce that $\bar{J}\dot{y} + T\nabla H(y) = 0$. For the admissibility of such selection, we need

$$\int_0^1 |\dot{y}|^2 \, dt + T \int_0^1 \bar{J}\dot{y} \cdot \nabla H(y) \, dt = 0.$$

This can be solved for T if

$$\int_0^1 \bar{J}\dot{y} \cdot \nabla H(y) \, dt \neq 0. \tag{5.19}$$

In fact if (5.19) were not the case, then we could choose $\tau = \nabla H(y)$ in (5.17) to deduce that $\nabla H(y) = 0$, which contradicts our assumption. Hence y must satisfy $\dot{y} = T\bar{J}\nabla H(y)$.

We now turn to the proof of **(ii)**. For a critical $y \in \Lambda$, we know that $\dot{y} = T\bar{J}\nabla H(y)$. On the other hand, since H is homogeneous of degree 2, we also know that $y \cdot \nabla H(y) = 2H(y)$. Hence

$$1 = \int_0^1 \bar{J}y \cdot \dot{y} \, dt = 2T \int_0^1 H(y) \, dt = 2TH(y), \tag{5.20}$$

because $H(y)$ is a constant function. Since y cannot be identically 0 by the constraint, we deduce that $T > 0$ and that $H(y) = (2T)^{-1}$. Now if $z(t) = \sqrt{T}y(t/T)$, then obviously z is T-periodic with $H(z) = 1/2$ by the homogeneity of H. On the other hand,

$$\dot{z}(t) = T^{-1/2} \, \dot{y}(t/T) = \sqrt{T}\bar{J} \, \nabla H\big(y(t/T)\big) = \bar{J} \, \nabla H\big(\sqrt{T}y(t/T)\big) = \bar{J} \, \nabla H(z(t)).$$

Finally

$$A(z) = \frac{1}{2}\int_0^T \bar{J}z \cdot \dot{z} \, dt = \frac{T}{2}\int_0^T \bar{J}y(t/T) \cdot \dot{y}(t/T)T^{-1} \, dt = \frac{T}{2}\int_0^1 \bar{J}y(t) \cdot \dot{y}(t) \, dt = \frac{T}{2}.$$

This completes the proof of **(ii)**. □

On account of Lemma 5.2, we only need to find critical points for $\mathcal{C}$ on Λ. Since H is convex, we may wonder whether or not a minimum provides us with a critical point. It turns out that $\sup_\Lambda \mathcal{C} = +\infty$ and $\inf_\Lambda \mathcal{C} = 0$. Note that for H as in Lemma 5.2 infimum is not achieved because if $\mathcal{C}(x(\cdot)) = 0$, then $x(\cdot) \equiv 0$ and $\frac{1}{2}\int_0^1 \bar{J}x \cdot \dot{x} \neq 1$ for such $x(\cdot)$. Let us study an example to see why the infimum is not achieved.

Example 5.2 Assume that $n = 1$ and $H : \mathbb{R}^2 \to \mathbb{R}$ is given by $H(x) = \pi|x|^2$. We already know how to solve the corresponding system $\dot{x} = \bar{J}\nabla H(x)$; the solutions are given by $x(t) = (q(t), p(t)) = r(\sin 2\pi t, \cos 2\pi t)$, $t \geq 0$. The set $\{x : H(x) = 1/2\}$ carries the periodic orbit $x(t) = (2\pi)^{-1/2}(\sin 2\pi t, \cos 2\pi t)$. The set Λ consists of 1-periodic y with $\int_0^1 \bar{J}y \cdot \dot{y} \, dt = 1$.

If y is a simple curve with coordinates q and p, then $\int_0^1 \bar{J}y \cdot \dot{y}\, dt$ is twice the area enclosed by y. We can easily construct a sequence $y_l(t)$ in Λ with $\mathcal{C}(y_l) \to \infty$ as $l \to \infty$. For example, set $y_l(t) = \left(2\sqrt{l}\sin 2\pi t, \frac{1}{\pi\sqrt{l}}\cos 2\pi t\right)$, so that $y_l \in \Lambda$ and $\mathcal{C}(y_l) = 2\pi l + \frac{1}{2\pi l}$. This confirms $\sup_\Lambda \mathcal{C} = +\infty$. As for the infimum, let us choose a sequence $z_l \in \Lambda$ of high oscillation, say $z_l(t) = (2\pi l)^{-1/2}(\sin 2\pi lt, \cos 2\pi lt)$. Since $H(z_l) = (2l)^{-1}$, we learn that $\inf_\Lambda \mathcal{C} = 0$. □

From Example 5.2 it is clear that a control on $\mathcal{C}(y_l)$ for $y_l \in \Lambda$ guarantees no control on $\dot{y}_l$ and this results in $\inf_\Lambda \mathcal{C} = 0$. We now follow an idea of Clarke to switch to a new functional $\mathcal{D}$ which involves the derivative. To motivate the definition, observe that if

$$\dot{y} = T\bar{J}\ \nabla H(y), \tag{5.21}$$

for some $T > 0$, then for any constant $c \in \mathbb{R}^{2n}$,

$$\frac{d}{dt}\bar{J}(c - y) = T\nabla H(y).$$

Writing $w = \bar{J}(c - y)$, we deduce that for $H = H_K$ and $G = H_{K^\circ}$,

$$\nabla G(\dot{w}) = T(\bar{J}w + c). \tag{5.22}$$

because ∇H is homogeneous of degree 1 and $\big(\nabla H\big)^{-1} = \nabla G$. Here we are using Proposition 5.1 to assert that $H^* = G$. The condition $\int_0^1 \bar{J}y \cdot \dot{y}\, dt = 1$ becomes

$$1 = \int_0^1 \bar{J}(\bar{J}w + c) \cdot \bar{J}\dot{w}\, dt = \int_0^1 \bar{J}w \cdot \dot{w}\, dt.$$

Motivated by this and (5.21), let us define $\mathcal{D} : \Lambda \to \mathbb{R}$ by

$$\mathcal{D}(y) = \int_0^1 G(\dot{y}(t))\, dt.$$

Lemma 5.3 *If $w(\cdot)$ is a C^1 critical point of $\mathcal{D} : \Lambda \to \mathbb{R}$, then w solves* (5.22) *for some constants $T > 0$ and $c \in \mathbb{R}^{2n}$. Moreover if $y = \bar{J}w + c$, then $y \in \Lambda$, y satisfies* (5.21), *$2TH(y(\cdot)) \equiv 1$, and $\int_0^1 G(\dot{y})\, dt = T/2$.*

Proof As in the proof of Lemma 5.2, we can show that if w is a critical point for $\mathcal{D}$ on Λ, then

$$\int_0^1 \bar{J}\dot{w} \cdot \tau\, dt = 0 \ \Rightarrow\ \int_0^1 \nabla G(\dot{w}) \cdot \dot{\tau}\, dt = 0,$$

for every periodic C^1 path τ. This can be rewritten as

$$\int_0^1 \bar{J}w \cdot \dot{\tau}\ dt = 0 \quad \Rightarrow \quad \int_0^1 \nabla G(\dot{w}) \cdot \dot{\tau}\ dt = 0,$$

Hence

$$\int_0^1 \big(T\bar{J}w + c' - \nabla G(\dot{w})\big) \cdot \dot{\tau}\ dt = 0, \tag{5.23}$$

for every $T \in \mathbb{R}$ and $c' \in \mathbb{R}^{2n}$. We wish to choose τ so that

$$\dot{\tau} = T\bar{J}w + c' - \nabla G(\dot{w}). \tag{5.24}$$

For this, we need to satisfy two conditions:

$$\int_0^1 \big(T\bar{J}w + c' - \nabla G(\dot{w})\big)\ dt = 0, \qquad \int_0^1 \big(T\bar{J}w + c' - \nabla G(\dot{w})\big) \cdot \bar{J}w\ dt = 0.$$

These equations determine c' and T because $w \in \Lambda$ is never a constant. Indeed if we solve the first equation for c' and substitute it in the second equation, we get a linear equation for T that has the coefficient

$$\int_0^1 |Jw|^2\ dt - \left(\int_0^1 Jw\ dt\right)^2,$$

which is never 0 unless w is constant. Now choosing τ satisfying (5.24) in (5.23) implies that $\dot{\tau}$ is identically 0, which in turn implies

$$\nabla G(\dot{w}) = T\bar{J}w + c'. \tag{5.25}$$

Since G is homogeneous of degree 2, we know that $\dot{w} \cdot \nabla G(\dot{w}) = 2G(\dot{w})$. From this and (5.25) we deduce

$$\int_0^1 2G(\dot{w})\ dt = \int_0^1 \dot{w} \cdot \nabla G(\dot{w})\ dt = T \int_0^1 \bar{J}w \cdot \dot{w}\ dt = T.$$

Hence $T > 0$ because $G > 0$ away from 0. We then choose c so that $cT = c'$. Substituting this in the (5.25) yields (5.22). We can readily show that if $y = \bar{J}w + c$, then $y \in \Lambda$ and y satisfies (5.21). Clearly (5.21) implies that $2TH(y) = 1$ by (5.20). We are done. □

We are now ready for the proof of Theorem 5.1.

Proof of Theorem 5.1 First we obtain some useful properties of G. Note that by homogeneity of G,

$$G(x) = |x|^2 G\left(\frac{x}{|x|}\right).$$

As a result, there are positive constants c_1 and c_2, such that

$$c_1\frac{|x|^2}{2} \le G(x) \le c_2\frac{|x|^2}{2} \tag{5.26}$$

for all $x \in \mathbb{R}^{2n}$. From the homogeneity of G, we also know that $\nabla G(a) = |a|\nabla G\left(\frac{a}{|a|}\right)$. This in turn implies that for some constant c_3,

$$|\nabla G(a)| \le c_3|a|. \tag{5.27}$$

We then remark that even though we have worked with the space C^1, our Lemma 5.3 is valid for a larger space $\mathcal{H}^1$. The space $\mathcal{H}^1$ consists of 1-periodic functions $w : \mathbb{R} \to \mathbb{R}^{2n}$ which are weakly differentiable with the weak derivative $\dot{w} \in L^2$. That is, there exists $v \in L^2$ such that for every $\zeta \in C^1$,

$$\int_0^1 w \cdot \dot{\zeta}\ dt = -\int_0^1 v \cdot \zeta.$$

We simply write $\dot{w}$ for the weak derivative v. Note that we can define $\mathcal{D} : \bar{\Lambda} \to \mathbb{R}$ where

$$\bar{\Lambda} = \left\{ w \in \mathcal{H}^1 : \int_0^1 \bar{J}w \cdot \dot{w}\ dt = 1 \right\}.$$

In Lemma 5.3, we may replace Λ with $\bar{\Lambda}$.

On account of our extension of Lemma 5.3, it suffices to find a critical point of $\mathcal{D} : \bar{\Lambda} \to \mathbb{R}$. This can be achieved by showing the existence of a minimizer of $\mathcal{D}$.

Set $a = \inf_{\bar{\Lambda}} \mathcal{D}$ and choose a sequence $w_l \in \bar{\Lambda}$ such that $w_l(0) = 0$ and $\mathcal{D}(w_l) \to a$. In view of (5.24), we certainly have

$$c_4 := \sup_l \int_0^1 |\dot{w}_l|^2\ dt < \infty. \tag{5.28}$$

From this and Exercise 5.1**(ii)** we know

$$|w_l(t) - w_l(s)| \le \sqrt{c_4}|t-s|^{1/2}.$$

This and $w_l(0)$ imply that w_l has a convergent subsequence with respect to the uniform topology. On account of (5.28) we may choose a subsequence such that

$$\begin{aligned} &w_l \to w \text{ uniformly}, \\ &\dot{w}_l \to v \text{ weakly}, \end{aligned}$$

for some $v \in L^2$ and continuous w. We now assert that in fact w is weakly differentiable and its weak derivative is v. Indeed if $\zeta \in C^1$, then

$$\int_0^1 w \cdot \dot{\zeta}\ dt = \lim_{l\to\infty}\int_0^1 w_l \cdot \dot{\zeta}\ dt = -\lim_{l\to\infty}\int_0^1 \dot{w}_l \cdot \zeta\ dt = -\int_0^1 v \cdot \zeta\ dt.$$

It remains to show that $w \in \bar{\Lambda}$ and

$$a = \mathcal{D}(w). \tag{5.29}$$

For the former observe

$$\begin{aligned}\int_0^1 \bar{J}w \cdot \dot{w}\ dt &= \int_0^1 \bar{J}w \cdot (\dot{w} - \dot{w}_l)\ dt + \int_0^1 \bar{J}(w - w_l) \cdot \dot{y}_l\ dt + \int_0^1 \bar{J}w_l \cdot \dot{w}_l\ dt \\ &=: \Omega_1 + \Omega_2 + \Omega_3.\end{aligned}$$

We certainly have that $\lim_{l\to\infty} \Omega_1 = 0$ and $\Omega_3 = 1$. Moreover

$$|\Omega_2| \le \|w - w_l\|_{L^2}\|\dot{w}_l\|_{L^2} \to 0$$

as $l \to \infty$ by (5.28) and uniform convergence of w_l to w. This shows that $w \in \bar{\Lambda}$.

It remains to establish (5.29). Since $a = \inf \mathcal{D}$, it suffices to show that $\mathcal{D}(w) \le a$. This follows from the lower semi-continuity of the functional $\mathcal{D}$ which is a consequence of the convexity of G. Indeed by convexity of G,

$$G(\dot{w}) + \nabla G(\dot{w}) \cdot (\dot{w}_l - \dot{w}) \le G(\dot{w}_l).$$

Hence,

$$\int_0^1 G(\dot{w})\ dt + \int_0^1 \nabla G(\dot{w}) \cdot (\dot{w}_l - \dot{w})\ dt \le \int_0^1 \nabla G(\dot{w}_l)\ dt.$$

We now send $l \to \infty$. Since $\nabla G(\dot{w}) \in L^2$ by (5.17) and $\dot{w} \in L^2$, we know that the second term on the left-hand side goes to 0. As a result,

$$\int_0^1 G(\dot{w})\ dt \le \liminf_{l\to\infty} \int_0^1 G(\dot{w}_l) = a.$$

Using $\mathcal{H}^1$ version of Lemma 5.3, we know that for some $y \in \mathcal{H}^1$, and $T > 0$, we have $\dot{y} = T\bar{J}\nabla H(y)$. Since the right-hand side is continuous, we use Exercise 5.6(i) to deduce that for almost all t,

$$y(t) = y(0) + T\int_0^t \bar{J}\nabla H(y(s))\ ds.$$

From this we deduce that in fact $y \in C^1$. As in Lemma 5.2, we may use y to construct a solution x to $\dot{x} = \bar{J}\nabla H(x)$ with $2H(x) = 1$ to complete the proof. □

As a corollary to the proof of Theorem 5.1, we derive two useful expressions for the symplectic capacity that was defined in (5.1).

Theorem 5.2 *For every compact set K with the origin in its interior, we have $\mathbf{c}_0(K) = \mathbf{c}_0'(K) = \mathbf{c}_0''(K)$, where*

$$(5.30) \qquad \begin{aligned} \mathbf{c}_0'(K) &:= \inf_{w\in\Lambda} \frac{1}{2}\int_0^1 h_K^2(\dot{w})\ dt \\ \mathbf{c}_0''(K) &:= \inf\left\{\frac{T}{2}:\quad \text{There exists a } T\text{-periodic orbit of } X_{H_K} \text{ on } \partial K\right\}. \end{aligned}$$

Proof Let w be a 1-periodic function in Λ that minimizes $\mathcal{D}$. By Lemma 5.3, we can find $T > 0$ and $c \in \mathbb{R}^{2n}$ such that

$$(5.31) \qquad \dot{y} = T\bar{J}\nabla H(y), \quad \int_0^1 \bar{J}y\cdot\dot{y}\ dt = 1, \quad 2TH(y) = 1, \quad \int_0^1 G(\dot{y})\ dt = \frac{T}{2},$$

for $y = \bar{J}w + c$. We then build z out of y by $z(t) = \sqrt{T}y(t/T)$, so that z is T-periodic and

$$(5.32) \qquad \dot{z} = \bar{J}\nabla H(z), \quad A\big(z(\cdot)\big) = \frac{T}{2}, \quad 2H(z) = 1,$$

by Lemma 5.2. Hence $z(\cdot)$ is a periodic orbit of X_H that lies on $S = \partial K$. This immediately implies

$$(5.33) \qquad \mathbf{c}_0(K) \le \mathbf{c}_0'(K), \quad \mathbf{c}_0''(K) \le \mathbf{c}_0'(K).$$

For the reverse inequalities, let z be a periodic orbit of the Reeb vector field. By reversing the orientation if necessary, we may assume that $A(z) > 0$. After a reparametrization, we may assume that z is a T-periodic orbit of the vector field X_H with $H = H_K$, that lies on the level set $H = 1/2$. Note that since H satisfies $z\cdot\nabla H(z) = 2H$, we have

$$\frac{1}{2}\int_0^T \bar{J}z\cdot\dot{z}\ dt = \frac{1}{2}\int_0^T \bar{J}z\cdot\bar{J}\nabla H(z)\ dt = \frac{1}{2}\int_0^T 2H(z)\ dt = \frac{T}{2}.$$

This implies that $\mathbf{c}_0 = \mathbf{c}_0''$. Also this z (a T-periodic orbit of the vector field X_H) satisfies (5.32). Defining $y(t) = T^{-1/2}z(Tt)$ yields a 1-periodic function which satisfies the first three equations in (5.31). We then take any constant $c \in \mathbb{R}^{2n}$ and set $w = \bar{J}(c - y)$ so that (5.22) is valid, which in turn can be used as in the proof of Lemma 5.3 to derive the last equation of (5.31). As in the calculation right after (5.22), we can use (5.32) to assert that $y \in \Lambda$. From this and the last equation of (5.31) we deduce that $\mathbf{c}_0'(K) \le \mathbf{c}_0(K) = \mathbf{c}_0''(K)$. This and (5.33) complete the proof. □

Remark 5.1(i) In (5.1), we defined $\mathbf{c}_0(K)$ for a convex set with smooth boundary. However the equality $\mathbf{c}_0(K) = \mathbf{c}_0'(K)$ gives us an expression that is well defined for arbitrary convex sets with nonempty interior. In fact if w is a minimizer for $c'(K)$ and z is the corresponding orbit as in Lemmas 5.2 and 5.3, then z still solves $\dot{z} = \bar{J}\nabla H(z)$ in some generalized sense. Given any convex function H, we may define its *subdifferential* $\partial H(z)$ as the set of vectors v such that

$$H(a) - H(z) \geq v \cdot (a - z),$$

for all $a \in \mathbb{R}^{2n}$. Now the corresponding Hamiltonian ODE reads as

$$\dot{z}^{\pm} \in J\partial H(z),$$

where $\dot{z}^{\pm}$ denotes the left and right derivatives. Similarly, the line bundle ℓ_x for $x \in \partial K$ consists of lines in the direction Jv, where $v \in \partial g_K(x)$. So a Reeb orbit will be tangent to a single line in the direction of $J\nabla g_K(x)$ if it passes through a point $x \in \partial K$ at which $\nabla g_K(x)$ exists; otherwise there is a cone of directions $\{Jv : \ v \in \partial g_K(x)\}$ that represents ℓ_x.

(ii) When $H = H_K$, then a T-periodic orbit $z(\cdot)$ on ∂K yields periodic orbits on the other level sets of the same period. Indeed if $\lambda > 0$ and $u(t) = \sqrt{2\lambda}z(t)$, then u still solves $\dot{u} = \bar{J}\nabla H(u)$, simply because ∇H is homogeneous of degree 1. What we learn from this is that if T is the smallest period that a periodic orbit of X_H can have on ∂K, then this T is the smallest period that a non-constant periodic orbit can have any where in $\mathbb{R}^{2n}$. Moreover, since $H(u) = 2\lambda H(z) = \lambda$, then

$$\mathbf{c}_0\left(\{H \leq \lambda\}\right) = T\lambda.$$

It is this formula that generalizes to the *Hofer-Zehnder capacity* in Chapter 6. □

Example 5.3 Let V and W be two open convex subsets of $\mathbb{R}^n$ with $0 \in V$ and W. Set $K = A \times B$, where $A = \bar{V}$ and $B = \bar{W}$. This convex set does not have a smooth boundary. Nonetheless we can use $\mathbf{c}'(K)$ to define its symplectic capacity as we discussed in Remark 5.1. We certainly have

$$\partial K = \big(\partial V \times W\big) \cup \big(V \times \partial W\big) \cup \big(\partial V \times \partial W\big).$$

It is the set $\partial U \times \partial V$ that is responsible for the non-smoothness of ∂K. (For example if $n = 1$, then K is a rectangle and $\partial U \times \partial V$ consists of its 4 corners.) To figure out how a Reeb orbit looks like, let us use the Hamiltonian function g_K for which K is the level set $\{g_K = 1\}$. In this case $g_K(q,p) = \max\{g_A(q), g_B(p)\}$, and $\partial g_K(x) = \{\nabla g_K(x)\}$ is singleton only in $\big(\partial V \times W\big) \cup \big(V \times \partial W\big)$:

$$\nabla g_K(q,p) = \begin{cases} \big(0, \nabla g_B(p)\big) & \text{if } (q,p) \in V \times \partial W \\ \big(\nabla g_A(q), 0\big) & \text{if } (q,p) \in \partial V \times W. \end{cases}$$

The corresponding Hamiltonian vector field is given by

$$X_{g_K}(q,p) = \begin{cases} \big(\nabla g_B(p), 0\big) & \text{if } (q,p) \in V \times \partial W \\ \big(0, -\nabla g_A(q)\big) & \text{if } (q,p) \in \partial V \times W. \end{cases}$$

We now describe an orbit $(q(t), p(t))$ of X_{g_K} on ∂K with $(q(0), p(0)) = (q_0, p_0) \in V \times \partial W$:

(i) While in $V \times \partial W$, $p(t)$ stays put and $q(t)$ travels with velocity $\nabla g_B(p_0)$;

(ii) At some time t_0, the position $q(t)$ reaches the boundary ∂V. Then it enters the set $\partial V \times W$ through the point $(q(t_0), p_0) \in \partial V \times \partial W$. While in $\partial V \times W$, $q(t) = q(t_0) = q_1$ stays put and $p(t)$ travels in W with velocity $-\nabla g_A(q(t_0))$.

(iii) At some time t_1, the momentum $p(t)$ reaches the boundary ∂W. Then it enters the set $V \times \partial W$ through the point $(q(t_0), p(t_1)) \in \partial V \times \partial W$. Then **(i)** is repeated with (q_0, p_0) replaced with $(q(t_0), p(t_1))$.

The orbit we described in **(i)-(iii)** is a trajectory of a *Minkowski Billiard*. To see why we have a billiard, let us describe the orbit when $W = B_0(1)$ is the unit ball in $\mathbb{R}^n$: Given $q_0 \in V$ and p_0 of length 1, we have the following orbit:

(i) A particle starts at q_0 and travels with velocity p_0 in V.

(ii) As q reaches the boundary at $q_1 = q(t_0)$, it enters $\{q(t_0)\} \times B_0(1)$. Then $q(t_0)$ stays put while p_0 travels with velocity $-n(q_1) = -\nabla g_A(q_1)$ which is the inward normal to ∂V at q_1.

(iii) As $p(t)$ reaches the boundary of the unit ball at time t_1, we simply have $p_1 = p(t_1) = p_0 - 2(n_1 \cdot p_1)n_1$, where $n_1 = n(q_1)$. So, the relationship between p_0 and p_1 is that of a *specular reflection.*

From the preceding description, it is clear the if we ignore the time spent in $\partial V \times W$, what we have is a billiard trajectory. The part of dynamics spent in $\partial V \times W$ is for transforming the incident velocity $p(t_{2i})$ to the reflecting velocity $p(t_{2i+1})$. If we calculate the symplectic action of an orbit $\int_0^T p \cdot dq$, the part of the orbit inside $\partial V \times W$ does not contribute because $q(t)$ stays put. From this we learn that for the capacity of K, we are dealing with a billiard trajectory that is defined in the following way:

(i) We start from a point $q_0 \in V$ that has a velocity $\nabla g_A(p_0)$ with $p_0 \in \partial W$. The q-point travels according to its velocity until it reaches the boundary.

(ii) At the boundary point $q_1 = q(t_0)$, the velocity changes from $\nabla g_A(p_0)$ to $\nabla g_A(p_1)$, where $p_1 = \eta_B(p_0, q_1)$. The function η_B depends on the incoming velocity and the normal at the boundary point.

Now a periodic Reeb orbit corresponds to a periodic Minkowski billiard trajectory. If we write $q_0, q_1, \dots, q_l \in \partial V$ with $q_l = q_1$, for the hitting locations of a periodic Reeb orbit γ of the boundary, then

$$A(\gamma) = (q_1 - q_0) \cdot p_0 + \cdots + (q_l - q_{l-1}) \cdot p_{l-1},$$

where $p_0, p_1, \dots, p_l$ are the corresponding momenta with $p_l = p_1$, and $p_{i+1} = \eta_B(q_i, p_i)$. In the case of a standard billiard (when $B = B_0(1)$ is the unit ball), we have that $p_i =$

$(q_{i+1}-q_i)/|q_{i+1}-q_i|$, and $A(\gamma)$ is the total length of the billiard trajectory. So, $\mathbf{c}_0\big(A\times B_0(1)\big)$ is simply the length of the shortest periodic billiard trajectory in the convex set A. $\square$

In Example 5.3, we related $\mathbf{c}_0(A \times B)$ to the the shortest (with respect to B) periodic billiard trajectories in A. However in general it is hard to calculate the capacity of a convex body. We now state a conjecture that would allow us to compare the capacity of a set with its volume.

Conjecture 5.1 (Viterbo) If B_{2n} denotes the unit ball in $\mathbb{R}^{2n}$, then

$$\frac{\mathbf{c}_0(K)}{\mathbf{c}_0(B_{2n})} \le \left[\frac{Vol(K)}{Vol(B_{2n})}\right]^{\frac{1}{n}}, \tag{5.34}$$

for every bounded convex set $K \subset \mathbb{R}^{2n}$. $\square$

Of course we know that $\mathbf{c}_0(B_{2n}) = \pi$ and $Vol(B_{2n}) = \pi^n/n!$. Also we have equality in (5.34) if K is a ball.

Recently Artstein-Avidan, Karasev and Ostrover have observed that Viterbo's conjecture implies an old conjecture of Mahler in Convex Geometry that was formulated in 1939.

Conjecture 5.2 (Mahler) For every centrally symmetric convex set A of $\mathbb{R}^n$ with 0 in its interior,

$$Vol(A)\ Vol(A^\circ) \ge 4^n/n!. \tag{5.35}$$

$\square$

To see how (5.34) implies (5.35), observe that if we apply (**??**) to the set $A \times A^\circ$, then (5.34) reads as

$$\frac{c^n(A\times A^\circ)}{\pi^n} \le \frac{n!\ Vol(A)\ Vol(A^\circ)}{\pi^n},$$

which implies (5.34) if we can show that $\mathbf{c}_0(A \times A^\circ) \ge 4$. This inequality was established by Artstein-Avidan et al. in [AKO]. According to [AKO],

$$\mathbf{c}_0(A\times B) = 4\max\{r>0:\ rB^\circ \subset A\}.$$

Remark 5.2(i) Several weaker versions of (5.34) have been already established. Viterbo showed

$$\frac{\mathbf{c}_0(K)}{\mathbf{c}_0(B_{2n})} \le \gamma_n \left[\frac{Vol(K)}{Vol(B_{2n})}\right]^{\frac{1}{n}},$$

where $\gamma_n = 2n$ for a centrally symmetric K and $\gamma_n = 32n$ for general convex K. According to Arestein-Avidan Milman and Ostrove (2007), we can choose the constant γ_n to be independent of n.

(ii) For a centrally symmetric convex set K with the origin in the interior, Gluskin and Ostrover [GO] obtained the following bound

$$\left(\sup_{a,b\in K^\circ} \bar{\omega}(a,b)\right)^{-1} \le \mathbf{c}_0(K) \le 4\left(\sup_{a,b\in K^\circ} \bar{\omega}(a,b)\right)^{-1}.$$

□

Exercise 5.1 (i) Verify Parts **(i)** and **(ii)** of Proposition 5.1.

(ii) Let $w \in \mathcal{H}^1$. Show that $w(t) = \int_0^t \dot{w}(\theta)d\theta$ for almost all t. Also show

$$|w(t) - w(s)| \le \|\dot{w}\|_{L^2}\ |t-s|^{\frac{1}{2}}.$$

for $s, t \in [0,1]$ and $L^2 = L^2[0,1]$. **(iii)** Prove that if $y \in \mathcal{H}^1$ and $\int_0^1 y\ dt = 0$, then

$$\|y\|_{L^2} \le \frac{1}{2\pi}\|\dot{y}\|_{L^2}.$$

(iv) Show that if A is a centrally symmetric convex subset $\mathbb{R}^n$ with 0 in its interior, then $\mathbf{c}_0(A \times A^\circ) \le 4$. (*Hint:* Show that if the Hamiltonian flow starts from $(0, \nabla g_A(q))$ for some $q \in \partial A$ then it reaches the boundary at q and reflect backs on itself to hit ∂A at $-q$ for the next hitting location. Such a periodic orbit has the value 4 for its action.) □

6 Capacities and Their Applications

In this chapter we assume the existence of a capacity for the symplectic manifolds and deduce several properties of Hamiltonian systems. As we mentioned in the introduction the non-squeezing theorem of Gromov inspired the search for symplectic capacities.

In (5.1), we assign a positive number $\mathbf{c}_0(K)$ to every convex set with nonempty interior. By Theorem 5.2, we have an alternative variational expression $\mathbf{c}_0'(K)$ for $\mathbf{c}_0(K)$ that allows us to easily verify two important properties of $\mathbf{c}_0(\cdot)$:

$$K \subseteq K' \;\Rightarrow\; \mathbf{c}_0(K) \le \mathbf{c}_0(K'), \qquad \mathbf{c}_0(\lambda K) = \lambda^2 \mathbf{c}_0(K), \tag{6.1}$$

for every $\lambda > 0$. Moreover, (5.1) allows us to show that $\mathbf{c}_0(K)$ is invariant under a symplectic change of coordinates. More precisely, if ψ is a symplectic diffeomorphism and both K and $\psi(K)$ are convex then

$$\mathbf{c}_0(K) = \mathbf{c}_0(\psi(K)), \tag{6.2}$$

because $(d\psi)_x T_x(\partial K) = T_{\psi(x)}(\partial\psi(K))$, which implies

$$(d\psi)_x\big(\ell_x(\partial K)\big) = \ell_{\psi(x)}(\psi(\partial K)),$$

where the line $\ell_x(A)$ denotes the kernel of $\bar\omega_{|A}$. (See also (4.12) and Proposition 4.4.) Clearly the finiteness of $\mathbf{c}_0(K)$ is exactly Weinstein's conjecture for convex subsets of $\mathbb{R}^{2n}$. We now wonder whether $\mathbf{c}_0(\cdot)$ has an extension to all subsets of $\mathbb{R}^{2n}$. The finiteness of $\mathbf{c}_0(A)$ for a bounded A would lead to the existence of a periodic orbit of the Reeb vector field on ∂A as we discussed in Remark 4.4. As a naive guess we may try $\mathbf{c}_0(K) = \inf |A(\gamma)|$, with infimum over the periodic characteristic of the line bundle $\ell_x(\partial K)$, which is always well-defined because $\bar\omega$ is symplectic. But this cannot be a useful extension for our purposes because many smooth hypersurfaces carry no periodic orbit as we saw in Example 4.3. We will offer shortly several legitimate extensions of (5.1). Let us first formulate our wish-list for what an extension should satisfy.

Definition 6.1 Write $\mathcal{P}\big(\mathbb{R}^{2n}\big)$ for the set of subsets of $\mathbb{R}^{2n}$. Then $\mathbf{c} : \mathcal{P}\big(\mathbb{R}^{2n}\big) \to [0,\infty]$ is called a *weak Euclidean capacity* if the following conditions hold:

(i) If $\psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is a symplectomorphism (i.e., a symplectic diffeomorphism with $\psi(\mathbb{R}^{2n}) = \mathbb{R}^{2n}$) and $\psi(A) \subseteq B$ then $\mathbf{c}(A) \le \mathbf{c}(B)$.

(ii) $\mathbf{c}(\lambda A) = \lambda^2 \mathbf{c}(A)$ for $\lambda > 0$.

(iii) $\mathbf{c}(B^{2n}(1)) > 0$ and $\mathbf{c}(Z^{2n}(1)) < \infty$, where $B^{2n}(1)$ is the Euclidean ball of radius 1, and $Z^{2n}(1)$ is the cylinder $\{(q,p) : q_1^2 + p_1^2 \le 1\}$.

We say c is a *strong Euclidean capacity* if in place of **(iii)**, we require the stronger assumption

$$\mathbf{c}(B^{2n}(1)) = \mathbf{c}(Z^{2n}(1)) = \pi.$$

$\square$

Definition 6.2 Write $\mathcal{SM}$ for the set of all symplectic manifolds. Then $\mathbf{c} : \mathcal{SM} \to [0, \infty]$ is a *weak symplectic capacity* if the following conditions hold:

(i) If $\psi : (M_1, \omega_1) \to (M_2, \omega_2)$ is a symplectic embedding and $\dim M_1 = \dim M_2$, then $\mathbf{c}(M_1, \omega_1) \leq \mathbf{c}(M_2, \omega_2)$.

(ii) $\mathbf{c}(M, \lambda\omega) = \lambda \mathbf{c}(M, \omega)$ for $\lambda > 0$.

(iii) $\mathbf{c}(B^{2n}(1), \bar{\omega}) > 0$, and $\mathbf{c}(Z^{2n}(1), \bar{\omega}) < \infty$.

We say c is a *strong symplectic capacity* if in place of **(iii)**, we require the stronger assumption

$$\mathbf{c}(B^{2n}(1), \bar{\omega}) = \mathbf{c}(Z^{2n}(1), \bar{\omega}) = \pi. \tag{6.3}$$

$\square$

Note that the condition $\mathbf{c}(B^{2n}(1), \bar{\omega}) > 0$, guarantees that $\mathbf{c} \equiv 0$ is not a capacity. The requirement $\mathbf{c}(Z^{2n}(1)) < \infty$, disqualifies $\mathbf{c}(M, \omega) = |\int_M \omega^n|^{1/n}$, $n = dim(M)$, to be a capacity. A strong capacity for symplectic manifolds can be used to define a strong Euclidean capacity by

$$\mathbf{c}(A) = \inf \left\{ \mathbf{c}(U, \bar{\omega}) : A \subseteq U,\ U \text{ open in } \mathbb{R}^{2n} \right\}. \tag{6.4}$$

The proof of this is left to the reader (Exercise 6.1**(i)**).

We now define four capacities.

Definition 6.3(i) (*Gromov Width*)

$$\underline{\mathbf{c}}(M, \omega) = \sup \left\{ \pi r^2 : \text{ There exists an embedding from } (B^{2n}(r), \bar{\omega}) \text{ into } (M, \omega) \right\},$$
$$\bar{\mathbf{c}}(M, \omega) = \inf \left\{ \pi r^2 : \text{ There exists an embedding from } (M, \omega) \text{ into } (Z^{2n}(r), \bar{\omega}) \right\}$$

(ii) (*Hofer-Zehnder Capacity*) Given a symplectic (M, ω), we first set $\mathcal{H}(M)$ to denote the set C^1 functions $H : M \to [0, \infty)$ such that $H = maxH$ outside a compact subset of M and $H = \min H = 0$ on some non-empty open set. We then set

$$\hat{\mathcal{H}}(M, \omega) = \{H \in \mathcal{H}(M) : X_H \text{ has no periodic orbit of period } T \in (0, 1]\}.$$

We now define

$$\mathbf{c}_{HZ}(M,\omega) = \sup_{H\in\hat{\mathcal{H}}(M,\omega)} \max H. \tag{6.5}$$

(iii) (*Displacement Energy*) First we define it for open subsets $U \subseteq \mathbb{R}^{2n}$. Given a time dependent Hamiltonian function $H : \mathbb{R}^{2n} \times [0,1] \to \mathbb{R}$, define

$$\|H\|_{\infty,1} := \int_0^1 \left(\sup_x H(x,t) - \inf_x H(x,t) \right) dt.$$

We then define the *displacement energy* of a set U by

$$\mathbf{e}(U) := \inf \left\{ \|H\|_{\infty,1} : \ \phi_1^H(U) \cap U = \emptyset, \ H \text{ is of compact support} \right\}.$$

As in (6.4), we may also define $\mathbf{e}(A)$ for arbitrary subsets of $\mathbb{R}^{2n}$. More generally, if (M,ω) is symplectic and U is an open subset of M, then

$$\mathbf{e}(U; M,\omega) := \inf \left\{ \|H\|_{\infty,1} : \ \phi_1^{H,\omega}(U) \cap U = \emptyset, \ H \text{ is of compact support} \right\}.$$

□

Remark 6.1(i) The motivation behind the definition is Gromov's non-squeezing theorem. Later in Chapter 10, we give a minimax-type expression for $\underline{c}$ and $\bar{c}$ that is based on Gromov's proof of non-squeezing theorem and involves *pseudo-holomorphic curves*.

(ii) Let us write

$$T^{\min}(H;\omega) = T^{\min}(H) = \inf \left\{ T : \ X_H^\omega \text{ has a non-constant periodic orbit of period } T \right\}.$$

We may then write

$$\mathbf{c}_{HZ}(M,\omega) = \sup_{H\in\mathcal{H}(M)} \left\{ \max H : \ T^{\min}(H;\omega) \geq 1 \right\}.$$

Obviously,

$$\mathbf{c}_{HZ}(M,\omega) \leq \sup_{H\in\mathcal{H}(M)} \left(T^{\min}(H;\omega) \max H \right).$$

Since a T-periodic orbit $x(\cdot)$ of X_H yields a T/λ-periodic orbit $y(t) = x(t\lambda)$ of $X_{\lambda H}$, we learn

$$T^{\min}(\lambda H) = \lambda^{-1} T^{\min}(H), \tag{6.6}$$

for every $\lambda > 0$. As a result, if $H \in \mathcal{H}(M)$ and $\bar{T} = T^{\min}(H)$, then $T^{\min}(\bar{T}H) = 1$ and $\bar{T}H \in \hat{\mathcal{H}}(M,\omega)$, which in turn implies that $\bar{T} \max H \leq \mathbf{c}_{HZ}(M,\omega)$. In summary, for every H of compact support,

$$T^{\min}(H;\omega) \leq \frac{\mathbf{c}_{HZ}(M,\omega)}{\max H}, \quad \text{and} \quad \mathbf{c}_{HZ}(M,\omega) = \sup_{H\in\mathcal{H}(M)} \left(T^{\min}(H;\omega) \max H \right). \tag{6.7}$$

When $M=U$ is a convex open subset of $\mathbb{R}^{2n}$ with $0\in U$ and $K=\bar U$, then we may set $H=\min\{H_K,1/2\}$ so that $H=1/2$ in the complement of K and $H(0)=0$. In some sense H is the Hamiltonian for which the supremum in (6.7) is attained. Since $H\notin\mathcal{H}(U)$, we choose a sequence $H_l\in\mathcal{H}(U)$ so that $H_l\to H$ as $l\to\infty$. This implies that $\mathbf{c}_{HZ}(U)=\mathbf{c}_{HZ}(\bar U)\ge\mathbf{c}(\bar U)$. The details are left to the reader (see Exercise 6.1(**(v)**).

(iii) If $X_H=X_H^{\bar\omega}$ has a periodic orbit x of period T, then $y(t)=x(Tt)$ is 1-periodic and solves $\dot y=T\bar J\nabla H(y)$. By differentiating we obtain $\ddot y=T\bar JD^2H(y)\dot y$, which in turn yields the bound

$$\|\ddot y\|_{L^2}\le aT\|\dot y\|_{L^2}\le\frac{aT}{2\pi}\|\ddot y\|_{L^2},$$

where $a=\max|D^2H|$ and for the last inequality we used Exercise 5.1**(iii)**. From this we learn

$$T^{\min}(H)\ge\frac{2\pi}{\max|D^2H|}. \tag{6.8}$$

However, the capacity $\mathbf{c}_{HZ}(M,\omega)$, provides us with an upper-bound on the period that depends on $\|H\|$ only. Note that if we write $\mathcal{T}(H)=\mathcal{T}(H;\omega)$ for $T^{min}(H;\omega)$, then in some sense, $\mathcal{T}(H)$ measures the size of the function H in some symplectic sense. In fact, for $H\in\mathcal{H}(U)$ and $\lambda>0$,

$$\mathcal{T}(\lambda H)=\lambda\mathcal{T}(H),\qquad \mathbf{c}_{HZ}(U)^{-1}\ \max H\le\mathcal{T}(H)\le(2\pi)^{-1}\ \max|D^2H|.$$

(iii) From comparing (6.8) with (6.7) we deduce the bound

$$\|H\|\le(2\pi)^{-1}\mathbf{c}_{HZ}(U)\max|D^2H|,$$

for every H with compact support in U. In fact it is elementary to show

$$\|H\|\le diam(U)\max|\nabla H|,\qquad \max|\nabla H|\le diam(U)\max|D^2H|,$$

where *diamU* denotes the diameter of U. So, we have maximum-type-principle inequality of the form

$$\|H\|\le diam(U)^2\max|D^2H|.$$

According to *Alexandroff-Bakelman-Pucci (ABP) Maximum Principle*,

$$\max_U u\le\max_{\partial U}u+\frac{diam(U)}{Vol(B^{2n})^{1/n}}\left[\int_U|\det D^2u|\ dx\right]^{\frac{1}{n}}.$$

In Viterbo [V], a symplectic geometry proof of ABP is given by using Hofer's theorem for the displacement energy (Theorem 6.2 **(iv)** below). □

Here are some straightforward properties of capacities.

Proposition 6.1 *Suppose that c is a strong capacity.*

(i) *For every ellipsoid E, we have $\mathbf{c}(E,\bar\omega) = \pi r_1(E)^2$.*

(ii) *For every symplectic manifold (M,ω),*

$$\underline{\mathbf{c}}(M,\omega) \le \mathbf{c}(M,\omega) \le \bar{\mathbf{c}}(M,\omega).$$

Proof By Corollary 2.2, the ellipsoid E is symplectomorphic to the ellipsoid

$$E' = \left\{ x : \sum_1^n r_j^{-2}(q_j^2 + p_j^2) \le 1 \right\}.$$

Hence $\mathbf{c}(E,\bar\omega) = \mathbf{c}(E',\bar\omega)$. On the other hand,

$$B^{2n}(r_1) \subseteq E \subseteq Z^{2n}(r_1).$$

As a result, $\pi r_1^2 \le \mathbf{c}(E',\bar\omega) = \mathbf{c}(E,\bar\omega) \le \pi r_1^2$. This completes the proof of **(i)**. The proof of **(ii)** is left as an exercise. □

In Chapter 5 we had a candidate for the symplectic capacity of a convex set. It is conjectured that all capacities coincide on convex sets.

Conjecture 6.1 (*Viterbo*) For every convex subset K of $\mathbb{R}^{2n}$ with nonempty interior, $\underline{\mathbf{c}}(K) = \bar{\mathbf{c}}(K)$. □

In Chapter 7 we will show that $\mathbf{c}_{HZ}$ is a capacity.

Theorem 6.1 *The Hofer-Zehnder function $\mathbf{c}_{HZ}$ is a strong capacity.*

Remark 6.2 As we will see in Exercise 6.1**(iv)**, it is not hard to show that $\mathbf{c}_{HZ}\big(B^{2n}(1)\big) \le \pi$. Chapter 7 is devoted to the proof of $\mathbf{c}_{HZ}\big(Z^{2n}(1)\big) \le \pi$. Let us explain how the *coarea formula* can be used to prove the latter bound for $n = 1$. In fact we can even show that $\mathbf{c}_{HZ}(U) \le area(U)$ for every connected open set in the plane. According to the coarea formula,

$$\int f|\nabla H|\, dx = \int_{-\infty}^{\infty} \left(\int_{\{H=r\}} f\, d\ell \right) dr,$$

where dx and $d\ell$ denote the area and the length integration (see Proposition 4.5 and Remark 4.4, or [EG] for the coarea formula). Hence if $H \in \mathcal{H}(M)$ with support in an open set U,

$$area\,(U \cap \{\nabla H \ne 0\}) = \int_0^{\max H} \left(\int_{U\cap\{H=r\}} \mathbb{1}(\nabla H \ne 0)\, \frac{d\ell}{|\nabla H|} \right) dr. \tag{6.9}$$

Now if $x(\cdot)$ is a periodic orbit of period T that lies on the level set $\{H = r\}$, then its arc length is given by

$$d\ell = |\dot{x}(t)|dt = |\nabla H(x(t))|dt.$$

Hence

$$\int_{U\cap\{H=r\}} \mathbb{1}(\nabla H \neq 0)\ \frac{d\ell}{|\nabla H|} \geq T \tag{6.10}$$

If $\hat{T}$ is the smallest possible period for the non-constant periodic orbits of X_H, then by (6.9) and (6.10), we have

$$(\max H)\hat{T} \leq area(U).$$

This in turn implies that $\mathbf{c}_{HZ}(U) \leq area(U)$. For the reverse inequality $\mathbf{c}_{HZ}(U) \geq area(U)$, we need to find $H : U \to [0, \infty)$ so that each nonzero level set $\{H = r\}$ consists of exactly one periodic orbit of period very close to $area(U)$. In fact when U is convex with $0 \in U$, then a slight modification of $H_{\bar{U}}$ would do the job as we demonstrated this in Remark 5.1**(ii)**. If U is simply connected, we can find an area preserving φ such that $\varphi(B^{2n}(r)) = U$ with $\pi r^2 = area(U)$. Then a modification of $H(x) = |x|^2$ yields a Hamiltonian in $\mathcal{H}(B^{2n}(r)$, and this will be pushed forward to U by φ to yield the desired Hamiltonian. $\square$

We now discuss four fundamental results in symplectic geometry that will be established in this chapter with the aid of Theorem 6.1.

Theorem 6.2 **(i)** *(Gromov) If there exists a symplectic diffeomorphism $\psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ with $\psi(B^{2n}(r)) \subseteq Z^{2n}(R)$, then $r \geq R$.*

(ii) *(Gromov, Eliashberg) If $\psi_k : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ are symplectomorphisms such that $\psi_k \to \psi$ uniformly with ψ a diffeomorphism, then ψ is symplectic.*

(iii) *(Viterbo) The Reeb vector field of a closed hypersurface S of contact type in $\mathbb{R}^{2n}$ has a periodic orbit.*

(iv) *(Hofer) For every open set U, we have $\mathbf{c}_{HZ}(U) \leq \mathbf{e}(U)$.*

As we will see later, parts **(i)** and **(ii)** are consequences of the existence a strong symplectic capacity. Though **(iii)** and **(iv)** relies on the form of $\mathbf{c}_{HZ}$.

Proof of Theorem 6.2(i) If ψ is a symplectomorphsim with $\psi(B^{2n}(r)) \subseteq Z^{2n}(R)$, and c is a strong capacity, then

$$r^2\pi = \mathbf{c}(rB^{2n}(1)) = \mathbf{c}(B^{2n}(r)) = \mathbf{c}(\psi(B^{2n}(r)) \leq \mathbf{c}(Z^{2n}(R)) = R^2\pi,$$

as desired.

Proof of Theorem 6.2(iii) On account of Proposition 4.4, we only need to show that if $K = K_S$, then the Hamiltonian vector field X_K has a periodic orbit in a bounded neighborhood V of S. Recall that there exists an open set U with $U = \cup_{t\in(-\delta,\delta)} S_t$, and $K(S_s) = e^s$. To turn K to a Hamiltonian $H \in \mathcal{H}(U)$, take a smooth function $g : [e^{-\delta}, e^{\delta}] \to [0,1]$ such that

$$g(r) = 0 \text{ for } r \le e^{-\delta/2}, \quad g(r) = 1 \text{ for } r \ge e^{\delta/2}, \quad g'(r) > 0 \text{ for } e^{-\delta/2} < r < e^{\delta/2}.$$

If we $H = g(K)$, then $H \in \mathcal{H}(U)$. By Theorem 6.1, we know that $\mathbf{c}_{HZ}(U) < \infty$. From this and (6.7), we deduce that X_H has a non-constant periodic orbit in U. This periodic orbit must lie in

$$V = \cup\{S_t : \ t \in (-\delta/2, \delta/2)\}.$$

We are done by Proposition 4.4. □

We now turn to the proof of Theorem 6.2**(ii)**. The key idea of the proof is an equivalent criterion for the symplecticity of a transformation that does not involve any derivative. This should be compared to the notion of a volume preserving transformations. We can say that a transformation is volume preserving if its Jacobian is one, or equivalently it preserves the volume. The latter criterion is more useful when we want to show that a limit of a sequence of measure preserving transformations is again measure preserving.

Theorem 6.3 *(*Ekeland-Hofer*) Let* $\mathbf{c}$ *be a strong capacity and let* $\psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ *be a diffeomorphism. Then the following statements are equivalent:*

(i) *For every small ellipsoid* E*, we have* $\mathbf{c}(\psi(E)) = \mathbf{c}(E)$*.*

(ii) *Either* $\psi^*\bar{\omega} = \bar{\omega}$ *or* $\psi^*\bar{\omega} = -\bar{\omega}$*.*

Proof Evidently if $\psi^*\bar{\omega} = \bar{\omega}$, then (i) is true. If $\psi^*\bar{\omega} = -\bar{\omega}$, then $\hat{\psi}^*\bar{\omega} = \bar{\omega}$ for $\hat{\psi} = \psi \circ \tau$, where τ is given by $\tau(q,p) = (p,q)$. We then have that for an ellipsoid E, $\mathbf{c}(\psi(\tau(E))) = \mathbf{c}(E)$. On the other hand, if E is a *standard ellipsoid* given by

$$\sum_{i=1}^{n} r_i^{-2}(q_i^2 + p_i^2) \le 1,$$

then $\tau(E) = E$. Hence for such ellipsoids, $\mathbf{c}(\psi(E)) = \mathbf{c}(E)$. Since for any symplectic φ we have $(\psi \circ \varphi)^*\bar{\omega} = -\bar{\omega}$, we also have

$$\mathbf{c}(\varphi(E)) = \mathbf{c}(E) = \mathbf{c}(\psi(\varphi(E))),$$

for any standard ellipsoid . Now any ellipsoid can be represented as $\varphi(E)$ for some linear symplectic φ and a standard E. This completes the proof of (ii) $\Rightarrow$ (i).

For the converse, assume (i) is true and define

$$\phi_k(x) = k(\psi(a + k^{-1}x) - \psi(a)).$$

We have that $\phi_k(x) \to \psi'(a)x$ locally uniformly as $k \to \infty$. Note that if E is an ellipsoid and ψ satisfies (i), then

$$\begin{aligned}\mathbf{c}(\phi_k(E)) &= \mathbf{c}(k\psi(a + k^{-1}E) - k\psi(a)) \\ &= k^2\mathbf{c}(\psi(a + k^{-1}E)) \\ &= k^2\mathbf{c}(k^{-1}E) = \mathbf{c}(E).\end{aligned}$$

On the other hand, it follows from Lemma 6.1 below that $\lim \phi_k(\cdot) = \psi'(a)$ also satisfies (i). We finally use Theorem 2.3 to deduce that the matrix $A = \psi'(a)$ satisfies $A^*\bar{\omega} = \bar{\omega}$ or $A^*\bar{\omega} = -\bar{\omega}$. By continuity we have $\psi(a)'^*\bar{\omega} = \bar{\omega}$ for all a or $\psi(a)'^*\bar{\omega} = -\bar{\omega}$ for all a. $\square$

To complete the proof of Theorem 6.3, we till need to show that the property (i) of Theorem 2.3 is preserved under a uniform limit. In Lemma 6.1 we prove a stronger variant that does not assume that the functions $\{\psi_k\}_{k\in\mathbb{N}}$ are diffeomorphism.

Lemma 6.1 *Let $\{\psi_k\}$ be a sequence of continuous functions for which (i) of Theorem 6.3 is valid. If $\psi_k \to \psi$ locally uniformly and ψ is a homeomorphism, then ψ satisfies (i) as well.*

Proof Imagine that we can prove this: For every $\lambda \in (0, 1)$ and every 0-centered ellipsoid E, there exists k_0 such that for $k > k_0$ we have

$$\psi_k(\lambda E) \subseteq \psi(E) \subseteq \psi_k(\lambda^{-1}E). \tag{6.11}$$

Then we are done because

$$\begin{aligned}\mathbf{c}(\psi(E)) &\le \mathbf{c}(\psi_k(\lambda^{-1}E)) = \mathbf{c}(\lambda^{-1}E) = \lambda^{-2}\mathbf{c}(E), \\ \mathbf{c}(\psi(E)) &\ge \mathbf{c}(\psi_k(\lambda E)) = \mathbf{c}(\lambda E) = \lambda^2\mathbf{c}(E)\end{aligned}$$

and yields property (i) for ψ by sending λ to 1.

To establish (6.11), let us write $\varphi_k = \psi^{-1} \circ \psi_k$. Clearly the sequence φ_k converges identity, locally uniformly in large k limit. From this it is clear that for large k,

$$\psi^{-1} \circ \psi_k(\lambda E) \subseteq E,$$

establishing the first inclusion in (6.11).

If ψ_k is a homeomorphism for each k, then the second inclusion in (6.11) can be established in the same way. It remains to show that for large k,

$$E \subseteq \varphi_k(\lambda^{-1}E), \tag{6.12}$$

even when ψ_k's are not homeomorphism. If (6.12) fails, then there exists a sequence $k_l \to \infty$ and $y_l \in E$ such that $y_l \notin \varphi_{k_l}(\lambda^{-1}E)$. This allows us to define

$$F_l(x) = \frac{\varphi_{k_l}(x) - y_l}{|\varphi_{k_l}(x) - y_l|},$$

for $x \in \lambda^{-1}E =: E_\lambda$. It follows from Lemma A.1 of the Appendix that $\deg f_l = 0$ where $f_l : \partial E_\lambda \to S^{2n-1}$ is the restriction of F_l to ∂E_λ. On the other hand, we may define $g_l : \partial E_\lambda \to \partial S^{2n-1}$ by $g_l(x) = \frac{\varphi_{k_l}(x)}{|\varphi_{k_l}(x)|}$. The function g_l is well-defined for large l because φ_{k_l} is uniformly close to identity over the set ∂E_λ. The function g_l has $\deg 1$ simply because g_l is uniformly close to $x \mapsto \frac{x}{|x|}$ which has degree 1. To arrive at a contradiction, it suffices to show that g_l is homotopic to f_k. (By Lemma A.1 homotopic transformations have the same degree.) For homotopy, define

$$\Phi_l(x,t) = \frac{\varphi_{k_l}(x) - ty_l}{|\varphi_{k_l}(x) - ty_l|},$$

for $x \in \partial E_\lambda$, $t \in [0,1]$. Again, since $\varphi_k \to id$ and $ty_l \in E$ the homotopy is well-defined.

□

With Theorem 6.3 and Lemma 6.1 at our disposal, we can now give a straightforward proof for Theorem 6.2(**(iii)**.

Proof of Theorem 6.2(iii) Let ψ_k be a sequence of symplectic transformations such that $\psi_k \to \psi$ locally uniformly. Assume that ψ is a diffeomorphism. By Lemma 6.1, ψ preserves the capacity of ellipsoids and by Theorem 6.3, either $\psi^*\bar{\omega} = \bar{\omega}$ or $\psi^*\bar{\omega} = -\bar{\omega}$. We now need to rule out the second possibility. Indeed if $\psi^*\bar{\omega} = -\bar{\omega}$, and we define

$$\phi_n = \psi_n \times id : \mathbb{R}^{4n} \to \mathbb{R}^{4n}, \qquad \omega = \bar{\omega} \times \bar{\omega},$$

then $\phi_n^*\omega = \omega$, and $\phi_n \to \phi := \psi \times id$ locally uniformly. But $\phi^*\omega = \psi^*\bar{\omega} \times \bar{\omega} = (-\bar{\omega}) \times \bar{\omega} \neq \pm\omega$. As a result, we must have $\psi^*\bar{\omega} = \bar{\omega}$. □

Motivated by Theorem 6.3, we may define the notion of symplecticity for homeomorphism. Note that if $\psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is a diffeomorphism with $\psi^*\bar{\omega} = -\bar{\omega}$, then $\psi^*\bar{\omega}^n = (-1)^n\bar{\omega}^n$, and if n is odd then ψ can not be orientation preserving. Based on this we have the following definition.

Definition 6.3 Let $\psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a homeomorphism. We say ψ is a symplectic homeomorphism if either n is odd and ψ is an orientation preserving transformation for which $\mathbf{c}(\psi(E)) = \mathbf{c}(E)$ for every ellipsoid E. Or n is even and $\psi \times id : \mathbb{R}^{2n+2} \to \mathbb{R}^{2n+2}$ is a symplectic homeomorphism.

From the proof of Theorem 6.2**(iii)** it is clear that if ψ_k is a sequence of symplectic homeomorphism such that $\psi_k \to \psi$ locally uniformly, then ψ is also a symplectic homeomorphism.

We now turn our attention to the energy displacement. A fundamental theorem of Hofer relates $\mathbf{e}$ to $\mathbf{c}_{HZ}$:

Theorem 6.4 *For every set $A \subset \mathbb{R}^{2n}$,*

$$\mathbf{c}_{HZ}(A) \leq \mathbf{e}(A). \tag{6.13}$$

The proof of Theorem 6.3 will be given in Chapter 8. Following [MS], we use Theorem 6.1**(iii)** to establish a weaker version of Theorem 6.3 but in the more general setting:

Theorem 6.5 *Let (M, ω) be a symplectic manifold such that $\mathbf{c}_{HZ}(M \times B^2(r) \leq \pi r^2$ for every $r > 0$. Then for every open set $U \subset M$,*

$$\underline{\mathbf{c}}(U, \omega) \leq 2\mathbf{e}(U; M, \omega). \tag{6.14}$$

Note that the condition $\mathbf{c}(M \times B^2(r) \leq \pi r^2$ is valid for $M = \mathbb{R}^{2n}$. As a preparation, we start with two straightforward propositions:

Proposition 6.2 *Let (M, ω) be a symplectic manifold, and $H : M \times \mathbb{R} \to \mathbb{R}$ be a time dependent Hamiltonian function. Set*

$$\tilde{M} := M \times \mathbb{R}^2 = \{(x, h, t) : \ x \in M, \ h, t \in \mathbb{R}\}, \qquad \tilde{\omega} := \omega + dt \wedge dh,$$
$$\psi(x, h, t) := \left(\phi_t(x), h + H\left(\phi_t(x), t\right), t\right),$$

with $\phi_t = \phi_t^{H,\omega}$. Then transformation ψ is $\tilde{\omega}$-symplectic.

Proof Observe that if $z = (x, h, t)$ and $\hat{z} = (\hat{x}, \hat{h}, \hat{t}) \in T_x M \times \mathbb{R}^2$, then

$$(d\psi)_z \hat{z} = \Big((d\phi_t)_x \hat{x} + \hat{t} X_H(\phi_t(x)), \hat{h} + \hat{t} H_t\left(\phi_t(x), t\right) + (dH)_{\phi_t(x)}\left(\hat{t} X_H(\phi_t(x)) + (d\phi_t)_x \hat{x}\right), \hat{t}\Big).$$

As a result, $\left(\psi^* \tilde{\omega}\right)_z(\hat{z}, \hat{z}')$ equals to

$$\begin{aligned}
&\left(\phi_t^* \omega\right)_x\left(\hat{x}, \hat{x}'\right) + \hat{t}\ \omega_{\phi_t(x)}\left(X_H(\phi_t(x)), (d\phi_t)_x \hat{x}'\right) - \hat{t}'\ \omega_{\phi_t(x)}\left(X_H(\phi_t(x)), (d\phi_t)_x \hat{x}\right) \\
&\quad + (dt \wedge dk)\Big(\Big(\hat{h} + \hat{t} H_t\left(\phi_t(x), t\right) + (dH)_{\phi_t(x)}\left(\hat{t} X_H(\phi_t(x)) + (d\phi_t)_x \hat{x}\right), \hat{t}\Big), \\
&\qquad\qquad\qquad \Big(\hat{h}' + \hat{t}' H_t\left(\phi_t(x), t\right) + (dH)_{\phi_t(x)}\left(\hat{t}' X_H(\phi_t(x)) + (d\phi_t)_x \hat{x}'\right), \hat{t}'\Big)\Big) \\
&= \omega_x(\hat{x}, \hat{x}') + \left(\phi_t^*\left(i_{X_H}\omega\right)\right)_x\left(\hat{t}\hat{x}' - \hat{t}'\hat{x}\right) + (dt \wedge dk)\big((\hat{h}, \hat{t}), (\hat{h}', \hat{t}')\big) \\
&\quad + (dt \wedge dk)\left(\left(d(H \circ \phi_t)_x \hat{x}, \hat{t}\right), \left(d(H \circ \phi_t)_x \hat{x}', \hat{t}'\right)\right) \\
&= \tilde{\omega}_z(\hat{z}, \hat{z}') - \left(\phi_t^* dH\right)_x\left(\hat{t}\hat{x}' - \hat{t}'\hat{x}\right) + (dt \wedge dk)\left(\left(d(H \circ \phi_t)_x \hat{x}, \hat{t}\right), \left(d(H \circ \phi_t)_x \hat{x}', \hat{t}'\right)\right) \\
&= \tilde{\omega}_z(\hat{z}, \hat{z}'),
\end{aligned}$$

as desired. $\square$

Remark 6.3(i) In the case of $M = \mathbb{R}^{2n}$, simply write $\phi_t(q,p) = (Q,P)$, and $k = h + H\left(\phi_t^H(x), t\right)$, and observe

$$\begin{aligned}\sum_i dP^i \wedge dQ^i + dt \wedge dk &= \sum_i dp_i \wedge dq_i + \sum_i \left[P_t^i \; dQ^i \wedge dt + Q_t^i \; dt \wedge dP^i\right] + dt \wedge dh \\ &\quad + \sum_i \left[H_{q_i}(Q,P,t) \; dQ^i \wedge dt + H_{p_i}(Q,P,t) \; dP^i \wedge dt\right] \\ &= \sum_i dp_i \wedge dq_i + dh \wedge dt.\end{aligned}$$

(ii) Observe that if $\alpha = p \cdot dq - H(x,t)dt$, then $d\alpha = \bar{\omega} + dt \wedge dh$, where $h = H(x,t)$. Regarding h as a new coordinate we get $\tilde{\omega} = \bar{\omega} + dt \wedge dh$. It is worth mentioning that if $\hat{H}(x,h,t) = H(x,t) - h$, and $H^t(x,s) := H(x,t+s)$, then

$$\phi_s^{\hat{H},\tilde{\omega}}(x,h,t) = \left(\phi_s^{H^t,\bar{\omega}}(x), H\left(\phi_s^{H^t,\bar{\omega}}(x), t+s\right) - H(x,t) + h, t+s\right),$$

which resembles ψ. $\square$

Proposition 6.3 *For any pair of sympletic manifolds* (M,ω) *and* (M',ω')*, we have*

$$\mathbf{c}_{HZ}(M \times M', \omega \oplus \omega') \geq \min\left\{\mathbf{c}_{HZ}(M,\omega), \mathbf{c}_{HZ}(M',\omega')\right\}. \tag{6.15}$$

Proof Given $H \in \mathcal{H}(M)$ and $H' \in \mathcal{H}(M')$ with $\max H = \max H' = 1$, set $K(x,y) = H(x)H'(y)$. Note that $K \in \mathcal{H}(M \times M')$ with $\max K = 1$, and $X_K(x,y) = (H'(y)X_H(x), H(x)X_{H'}(y))$. Moreover if $\dot{z} = X_K(z)$, with $z = (x,y)$, then

$$\dot{x} = H'(y)X_H(x), \qquad \dot{y} = H(x)X_{H'}(y).$$

Clearly

$$\frac{d}{dt}H(x) = \frac{d}{dt}H'(y) = 0.$$

So if z is a T-periodic orbit with $\lambda = H(x)$ and $\lambda' = H(y)$, then $\hat{x}(t) = x(t/\lambda)$ is a $T\lambda$ periodic orbit of X_H and $\hat{y}(t) = y(t/\lambda')$ is a $T\lambda'$ periodic orbit of $X_{H'}$. Hence

$$T \geq \lambda T \geq T^{min}(H,\omega), \qquad T \geq \lambda' T \geq T^{min}(H',\omega').$$

If both x and y are non-constant. Similarly, if x or y is constant but not both, then either $T \geq T^{min}(H,\omega)$, or $T \geq T^{min}(H',\omega')$. In summary,

$$T^{min}(K, \omega \oplus \omega') \geq \min\left\{T^{min}(H,\omega), T^{min}(H',\omega')\right\}.$$

By taking the supremum over such pairs of (H, H'), we get (6.15). $\square$

Proof of Theorem 6.4 To ease the notation, we write $\|\cdot\|$ for $\|\cdot\|_{\infty,1}$. First observe that it suffices to establish (6.14) for open sets U that are symplectomorphic to Euclidean balls $B_r = \{x : \ |x| \le r\}$. More precisely, imagine that we can show the the following: If $\varphi : (B_r, \bar{\omega}) \to (U, \omega)$ is a symplectic embedding, and $\phi = \phi_1^{H,\omega}$ displaces $V = \varphi(B_r)$, then

$$\pi r^2 \le 2\|H\|. \tag{6.16}$$

Once this is established, we then take the infimum over such H to deduce

$$\pi r^2 \le 2\mathbf{e}(V; M, \omega) \le 2\mathbf{e}(U; M, \omega).$$

We then take the supremum over such balls to conclude (6.14).

Before embarking on the proof (6.16), let us explain the idea behind the proof. Observe that the condition $\phi_1^H(V) \cap V = \emptyset$ is a property of the flow at time 1, whereas the right-hand side of (6.15) involves the Hamiltonian function H. Because of this, we switch from ϕ_1^H to ψ of Proposition 6.2 that explicitly involves the Hamiltonian function. Let us define

$$h^+(t) = \max_x H(x,t), \quad h^-(t) = \min_x H(x,t), \quad \psi_0(x,t) = \psi(x,0,t), \quad \Gamma = \psi_0(V \times [0,1]).$$

The set Γ is a hypersurface in $\tilde{M}$, and $\Gamma \subseteq M \times A_0$, where

$$E_0 = \{(h,t) : \ t \in [0,1], \ h \in [h^-(t), h^+(t)]\},$$

In other words, ψ_0 embeds $V \times [0,1]$ into a cylinder like set $M \times E_0$, and we wish to show

$$\pi r^2 \le 2 \ area(E_0). \tag{6.17}$$

This indeed has the same flavor as our non-squeezing fact Theorem 6.2**(iii)**, though $V \times [0,1]$ is not a symplectic manifold. To rectify this, we replace $[0,1]$ with a planar set A and use ψ instead of ψ_0. More precisely,

$$A = \big([-a,a] \times [-2a,0]\big) \cup \big([-\varepsilon,\varepsilon] \times [0,1]\big) \cup \big([1,2a] \times [-2a,0]\big) =: A^- \cup A^0 \cup A^+.$$

Note that since V is an open set, we may change H slightly without losing the property $\phi(V) \cap V = \emptyset$. More precisely, if we take a smooth function $\chi = \chi_\delta$ such that $0 \le \chi \le 1$, $\chi(t) = 1$, for $t \in [\delta, 1-\delta]$, and $\chi(t) = 0$ for $t \notin (0,1)$, then for the Hamiltonian function $H'(x,t) = \chi(t)H(x,t)$, we have that $\|H\| - \|H'\| = O(\delta)$, and if $\phi' = \phi_1^{H',\omega}$, then ϕ' is close to $\phi = \phi_1^{H,\omega}$ and we still have $\phi'(V) \cap V = \emptyset$, for sufficiently small δ. So without loss of generality, we may assume that $H(x,t) = 0$ for t close to 0 or 1. This allows us to extend H for all times by setting $H(x,t) = 0$ for $t < 0$ or $t > 1$. We set $U = B_r \times A$ with A a connected open subset of $\mathbb{R}^2$. Note

- $\psi(x,h,t) = (x,h,t)$ for $t \le 0$,
- $\psi(x,h,t) = (\phi(x),h,t)$ for $t \ge 1$.

As a result, if we write $W^{\pm} = V \times A^{\pm}$, $W^0 = V \times A^0$, then

$$\psi(W^-) = W^-, \quad \psi(W^+) = \phi(V) \times A^+, \quad \psi(U^0) \subset M \times E(\varepsilon),$$

where $E(\varepsilon)$ is a small neighborhood of E_0 for small ε. We have the embedding

$$\psi : V \times A \hookrightarrow M \times (A^- \cup E(\varepsilon) \cup A^+).$$

We now want to use $\phi(V) \cap V = \emptyset$ to replace Z with a slimmer cylinder. To achieve this, take a symplectic covering map $\lambda : \mathbb{R}^2 \to \big[\mathbb{R} \times S^1\big] \cup \big[\mathbb{R} \times (0,\infty)\big]$ such that

$$\lambda(\mathbb{R} \times [0,1]) = \mathbb{R} \times S^1, \quad \lambda(h,t) = (h, 1-t), \quad \text{for } t \ge 1.$$

Since $\phi(V) \cap V = \emptyset$, we have that the map $\hat{\psi}(x,h,t) = \psi(x, \lambda(h,t))$ is a symplectic diffeomorphism. This time

$$\hat{\psi} : V \times A \hookrightarrow M \times (A^- \cup E'(\varepsilon)) := M \times E''. \tag{6.18}$$

with $area(E'(\varepsilon)) \to \|H\|$ as $\varepsilon \to 0$. From (6.18) we learn

$$\mathbf{c}_{HZ}(V \times A) \le \mathbf{c}_{HZ}(M \times E''). \tag{6.19}$$

Choose an area preserving diffeomorphism ζ with $\zeta(E'')$ is a disc D with $area(E'') = area(D)$. Using the symplectic map $id \times \zeta$ we learn that $M \times E''$ is symplectomorphic to $M \times D$ and by assumption $\mathbf{c}_{HZ}(M \times D) \le area(D)$. Hence $\mathbf{c}_{HZ}(M \times E'') \le area(E'')$. On the other hand $V \times A$ is symplectomorphic to $B_r \times A$. Hence (6.19) implies

$$\mathbf{c}_{HZ}(B_r \times A) \le area(E''). \tag{6.20}$$

To get a lower bound for the left-hand side, observe that by Theorem 6.1 and Proposition 6.3,

$$\mathbf{c}_{HZ}(B_r \times A) \ge \min\{\pi r^2, area(A)\}.$$

This and (6.20) yield $\min\{\pi r^2, area(A)\} \le area(E'')$. We then send ε to 0 to obtain

$$\min\{\pi r^2, 2(2a)^2\} \le (2a)^2 + \|H\|.$$

Finally choosing r so that $\pi r^2 = 2(2a)^2$, yields $\pi r^2/2 \le \|H\|$. This completes the proof of (6.16), which in turn implies (6.14). $\square$

Exercise 6.1(i) Show that $\mathbf{c}$ given by (6.4) is a strong Euclidean capacity.

(ii) Verify properties **(i)** and **(ii)** for $\underline{\mathbf{c}}$, $\bar{\mathbf{c}}$, $\mathbf{c}_{HZ}$, and Euclidean $\mathbf{e}$. (**Hint**: For $\mathbf{e}$, use $\hat{H}(x) = \lambda^2 H(x/\lambda)$.)

(iii) Verify Proposition 6.1**(ii)**.

(iv) Show that $\mathbf{c}_{HZ}(B^{2n}(1), \bar{\omega}) \geq \pi$.

(v) Show that $\mathbf{c}_{HZ}(K, \bar{\omega}) \geq \mathbf{c}_0(K)$ for every bounded convex subset of $\mathbb{R}^{2n}$ with nonempty interior.

(vi) Show that $\mathbf{e}(U) = area(U)$ for every simply connected open subset $U \subset \mathbb{R}^2$. Use this to show $\mathbf{e}(Z^{2n}(1)) \leq \pi$. **Hint**: Use Exercise 3.1**(vi)** to reduce the problem to the case of $U = [0, a]^2$. $\square$

7 Hofer-Zehnder Capacity

This section is devoted to the proof of Theorem 6.1. In fact we have the following result of Hofer and Zehnder.

Theorem 7.1 *For every convex set K of nonempty interior, $\mathbf{c}_{HZ}(K) = \mathbf{c}_0(K)$.*

On account of Exercise 6.1**(v)**, we only need to show that $\mathbf{c}_{HZ}(K) \le \mathbf{c}_0(K)$. Similarly, for Theorem 6.1, we only need to verify $\mathbf{c}_{HZ}(Z^{2n}(1), \bar{\omega}) \le \pi$. We now state a theorem that would imply this.

Theorem 7.2 *Assume $H_0 \in \hat{\mathcal{H}}(Z^{2n}(1))$ with $\sup H_0 > \pi$. Then the Hamiltonian flow of H_0 has a non-constant periodic orbit of period 1.*

Let $H_0 \in \mathcal{H}(Z^{2n}(1))$. In fact we may assume that H_0 vanishes near the origin. This is because we may replace H with $H \circ \psi$ for a symplectic $\psi : Z(1) \to Z(1)$ that satisfies $\psi(x) = x$ outside a compact subset of $Z(1)$. Indeed we may choose $\psi = \phi_1^h$ where h is a suitable Hamiltonian function that vanishes outside a compact subset of $Z(1)$. For example, first choose $h_0(x) = \bar{J}x \cdot a$ for a fix vector a, so that $\phi_1^{h_0}(x) = x + a$, then set $h = h_0\chi$, where χ is a function of compact support in $Z(1)$ that is identically 1 in a neighborhood of the line segment connecting the origin to a. Using such ψ, we may shift a minimizer of H to the origin. Since the flows of X_H and $X_{H\circ\psi}$ are conjugated, it suffices to find the desired 1-periodic orbit for $H_0 \circ \psi$. From now on we assume that H_0 in Theorem 7.2 vanishes near the origin.

As our next step we extend H_0 to a Hamiltonian function $H : \mathbb{R}^{2n} \to \mathbb{R}$. For this, we take an ellipsoid

$$E^0 = \left\{ x \in \mathbb{R}^{2n} : Q(x) = q_1^2 + p_1^2 + \frac{1}{l^2} \sum_{j=2}^{n} (q_j^2 + p_j^2) < 1 \right\}.$$

Since $H_0 \in \mathcal{H}(Z^{2n}(1))$, we have that $H_0 = \max H_0$ for $x \notin K$ where K is a compact subset of $Z := Z^{2n}(1)$. Choose l sufficiently large so that $K \subseteq E^0$. We now pick $\epsilon > 0$ so that $\max H_0 > \pi + \epsilon$ and pick a smooth function $f : [0, \infty) \to [0, \infty)$ such that $f(r) = \max H$ for $r \in [0, 1]$, $f(r) = (\pi + \epsilon)r$ for large r, $f(r) \ge (\pi + \epsilon)r$ for all r, and $0 < f'(r) \le \pi + \epsilon$ for $r > 1$. We now define

$$H(x) = \begin{cases} H_0(x) & \text{if } x \in E^0, \\ f(Q(x)) & \text{if } x \notin E^0. \end{cases} \tag{7.1}$$

We note that $H(x) = (\pi + \epsilon)Q(x)$ for large x.

Lemma 7.1 *Assume that $x(\cdot)$ is a 1-periodic solution of $\dot{x} = \bar{J}\nabla H(x)$ with*

$$\mathcal{A}(x(\cdot)) = \int_0^1 \left(\frac{1}{2}\bar{J}x \cdot \dot{x} - H(x)\right) dt > 0. \tag{7.2}$$

Then $x(t) \in E^0$ for all t and $x(\cdot)$ is non-constant.

Proof Evidently $X_H = 0$ on ∂E^0. Hence, all points on ∂E^0 are equilibrium points and if $x(t) \equiv a \in \partial E^0$, then $\mathcal{A}(x(\cdot)) = -H(a) \leq 0$ which contradicts (7.2). As a result, either $x(t) \in E^0$ for all t, or $x(t) \notin E^0$ for all t. It remains to rule out the latter possibility.

If $x(t) \notin E^0$ for all t, then

$$\begin{aligned} \dot{x} &= \bar{J}\nabla H(x) = f'(Q(x))\bar{J}\nabla Q(x), \\ \frac{d}{dt}Q(x) &= f'(Q(x))\nabla Q(x) \cdot \bar{J}\nabla Q(x) = 0. \end{aligned}$$

Hence, for such $x(\cdot)$ we have that $Q(x(\cdot)) = Q^0$, and

$$\begin{aligned} \mathcal{A}(x(\cdot)) &= \int_0^1 \left(\frac{1}{2}\bar{J}x \cdot \dot{x} - H(x)\right) dt \\ &= \int_0^1 \left[\frac{1}{2}f'(Q^0)\nabla Q(x) \cdot x - f(Q^0)\right] dt \\ &= f'(Q^0)Q^0 - f(Q^0) \leq (\pi + \epsilon)Q^0 - (\pi + \epsilon)Q^0 = 0, \end{aligned}$$

which contradicts (7.2). Here we used the fact that $2Q(x) = \nabla Q(x) \cdot x$. □

In view of Lemma 7.1, we only need to find a critical point of $\mathcal{A}$ which satisfies (7.2). Let us first observe that $\mathcal{A}$ is not bounded from below or above. Indeed if $y_k(t) = (\cos 2\pi kt)a + (\sin 2\pi kt)\bar{J}a$, for some $a \in \mathbb{R}^{2n}$, then

$$\int_0^1 |y_k(t)|^2 dt = |a|^2, \quad \int_0^1 \bar{J}y_k(t) \cdot \dot{y}_k(t) dt = 2\pi k|a|^2,$$

which in particular implies that $\lim_{k\to\pm\infty} \mathcal{A}(y_k) = \pm\infty$, whenever $a \neq 0$. Because of this, we search for saddle-type critical points of $\mathcal{A}$. A standard way of locating such critical points is by using the celebrated *minimax* principle.

To prepare for this, let us first extend the domain of definition of $\mathcal{A}$ from C^1 to the largest possible Sobolev space which turns out to be the space of function with "half" a derivative. We begin with $\mathcal{H}^0 = L^2$ which consists of measurable functions

$$x(t) = \sum_{k\in\mathbb{Z}} e^{2\pi kt\bar{J}}x_k = \sum_{k\in\mathbb{Z}} [(\cos 2\pi kt)I + (\sin 2\pi kt)\bar{J}]x_k$$

with $x_k \in \mathbb{R}^{2n}$ and $\|x\|_0 = \sum_k |x_k|^2 < \infty$. Here we are using the Fourier expansion of $x(\cdot)$ where instead of $i = \sqrt{-1}$ we use $-\bar{J}$. We write

$$\langle x, y \rangle_0 = \int_0^1 x(t) \cdot y(t) dt,$$

for the standard inner product of $\mathcal{H}^0$. Note that if

$$x(t) = \sum_k e^{2\pi k t \bar{J}} x_k \text{ and } y(t) = \sum_k e^{2\pi k t \bar{J}} y_k,$$

then $\langle x, y \rangle_0 = \sum_k x_k \cdot y_k$ and $\|x\|_0^2 = \langle x, x \rangle_0$. We note that $\int_0^1 H(x(t))dt$ is well defined for every $x \in \mathcal{H}^0$ because $H(x) = Q(x)$ for large x. However, to make sense of $\int_0^1 \bar{J}x \cdot \dot{x} dt$, we need to assume that $x(\cdot)$ possesses half a derivative. To see this first observe that if $x \in C^1$, with $x = \sum_k e^{2\pi k t \bar{J}} x_k$, then

$$\int_0^1 \frac{1}{2} \bar{J}x \cdot \dot{x} dt = \pi \sum_k k |x_k|^2.$$

This suggests defining

$$\mathcal{H}^{1/2} = \left\{ x \in \mathcal{H}^0 : \sum_k |k| |x_k|^2 < \infty \right\}.$$

We turn $\mathcal{H}^{1/2}$ into a Hilbert space by defining

$$\langle x, y \rangle = \langle x, y \rangle_{1/2} = x_0 \cdot y_0 + 2\pi \sum_{k \in \mathbb{Z}} |k| (x_k \cdot y_k).$$

More generally, we define

$$\langle x, y \rangle_s = x_0 \cdot y_0 + (2\pi)^{2s} \sum_{k \in \mathbb{Z}} |k|^{2s} (x_k \cdot y_k),$$

for every $s > 0$ and $\mathcal{H}^s$ consists of function $x \in \mathcal{H}^0$ such that $\|x\|_s^2 = \langle x, x \rangle_s < \infty$. Observe that if $x \in C^1$, then $\int_0^1 |\dot{x}(t)|^2 dt = \sum_k (2\pi k)^2 |x_k|^2$ and that in general $x \in \mathcal{H}^1$ iff x has a weak derivative in L^2.

So far we know that our functional $\mathcal{A}$ is defined on the Hilbert space $\mathcal{H}^{1/2}$. Let us take an arbitrary Hilbert space $\mathcal{E}$ and a function $F : \mathcal{E} \to \mathbb{R}$, and explain the idea of minimax principle for such a function.

Definition 7.1(i) We say that F is continuously differentiable with a derivative ∇F if $\nabla F : \mathcal{E} \to \mathcal{E}$ is a continuous function such that for all x and a,

$$F(x) = F(a) + \langle \nabla F(x), x - a \rangle + o(\|x - a\|).$$

We say that x is a critical point of F if $\nabla F(x) = 0$.

(ii) We say that $F \in C^1(\mathcal{E};\mathbb{R})$ satisfies *Palais–Smale* (PS) condition if the conditions

$$\sup_l |F(x_l)| < \infty, \quad \lim_{l\to\infty} \nabla F(x_l) = 0, \tag{7.3}$$

for a sequence $\{x_l\}$ imply that $\{x_l\}$ has a convergent subsequence. □

Given a family $\mathcal{F}$ of subsets of $\mathcal{E}$, define

$$\alpha(F,\mathcal{F}) = \inf_{A\in\mathcal{F}} \sup_{x\in A} F(x) \in [-\infty,+\infty].$$

Theorem 7.3 *(Minimax Principle). Let $F : \mathcal{E} \to \mathbb{R}$ be a Palais–Smale function and assume that the flow ϕ_t of the gradient* ODE $\dot{x} = -\nabla F(x)$ *is well-defined for all $t \in \mathbb{R}^+$. If $\phi_t(A) \in \mathcal{F}$ for $A \in \mathcal{F}$ and that $\alpha = \alpha(F,\mathcal{F}) \in \mathbb{R}$, then there exists $x^* \in \mathcal{E}$ such that*

$$\nabla F(x^*) = 0 \text{ and } F(x^*) = \alpha(F,\mathcal{F}). \tag{7.4}$$

Proof It suffices to show

$$\inf\left\{\|\nabla F(x)\| : \ \alpha^- < F(x) < \alpha^+\right\} = 0, \tag{7.5}$$

for every pair of constants α^+ and α^- such that $\alpha^- < \alpha < \alpha^+$. Indeed if (7.5) is true, then for every $l \in \mathbb{N}$, we can find x_l such that

$$\alpha - l^{-1} < F(x_l) < \alpha + l^{-1}, \qquad \|\nabla F(x_l)\| \le l^{-1}.$$

We then use the Palais–Smale property of F to assert that $\{x_l\}$ has a convergent subsequence that converges to x^*. Since $F \in C^1$, we deduce that $\nabla F(x^*) = 0$.

To establish (7.5), we argue by contradiction. Suppose to the contrary,

$$\inf\left\{\|\nabla F(x)\| : \ \alpha^- < F(x) < \alpha^+\right\} = \varepsilon > 0. \tag{7.6}$$

By the definition of α, we can find a set $A \in \mathcal{F}$ such that

$$\sup_{x\in A} F(x) < \alpha^+.$$

To get a contradiction, we use the flow of ϕ_t to come up with another set $\hat{A} = \phi_t(A) \in \mathcal{F}$ for which $\sup_{\hat{A}} F \le \alpha^-$ for sufficiently large t. Indeed, if $x \in A$ and $F(\phi_t(x)) > \alpha^-$, then

$$\alpha^- < F\left(\phi_s(x)\right) < \alpha^+,$$

for every $s \in [0,t]$. From this and (7.6) we deduce

$$\alpha^- < F(\phi_t(x)) = F(x) - \int_0^t \|\nabla F(\phi_s(x))\|^2 \, ds < \alpha^+ - t\varepsilon^2,$$

which is impossible if $t > (\alpha^+ - \alpha^-)/\varepsilon^2$. Hence we must have $F(\phi_t(x)) \le \alpha^-$, for such large t. This in turn implies

$$\sup_{\phi_t(A)} F \le \alpha^-,$$

which contradicts our assumption $\alpha^- < \alpha$. Thus (7.5) must be true. $\square$

Example 7.1 Let $F \in C^1(\mathbb{R}^d; \mathbb{R})$ be a function which satisfies the Palais–Smale condition. Assume that F is bounded from below. Then F has a minimizer x^*, i.e., $F(x^*) = \inf F$. This can be shown using Theorem 7.3 by taking $\mathcal{F} = \{\{x\} : x \in \mathbb{R}^d\}$. Similarly, when F is bounded above, take $\mathcal{F} = \{\mathcal{E}\}$ to deduce that the function F has a maximizer. $\square$

We now give two applications of Theorem 7.3.

Proposition 7.1 *Let $F \in C^1(\mathcal{E})$ be a Palais–Smale function.*

(i) *(Mountain Pass Lemma of Ambrosetti and Rabinowitz) Assume that $R \subseteq \mathcal{E}$ is a mountain range relative to F in the following sense:*

- *$\mathcal{E} \setminus R$ is not connected,*
- *$\inf_R F =: \beta > -\infty$,*
- *If A is a connected component of $\mathcal{E} \setminus R$, then $\inf_A F < \beta$.*

Then F has a critical value α satisfying $\alpha \ge \beta$.

(ii) *Let Γ and Σ be two bounded subsets of $\mathcal{E}$ such that $\inf_\Gamma F = \beta > -\infty$, $\phi_t(\Sigma) \cap \Gamma \ne \emptyset$ for all $t \ge 0$, and $\sup_\Sigma F < \infty$. Then F has a critical point x^* such that*

$$F(x^*) = \inf_{t \ge 0} \sup_{x \in \phi_t(\Sigma)} F(x) \ge \beta.$$

(As in Theorem 7.3, ϕ_t denotes the flow of $-\nabla F$.)

Proof **(i)** Let $\mathcal{E}^1$ and $\mathcal{E}^2$ be two connected components of $\mathcal{E} \setminus R$ and set $\hat{\mathcal{E}}^i = \{x \in \mathcal{E}^i : F(x) < \beta\}$, for $i = 1$ and 2. We now define

$$\mathcal{F} = \{\gamma[0,1] \text{ such that } \gamma : [0,1] \to \mathcal{E} \text{ is continuous with } \gamma(0) \in \hat{\mathcal{E}}^1 \text{ and } \gamma(1) \in \hat{\mathcal{E}}^2\}.$$

We set $\alpha = \alpha(F, \mathcal{F})$ and would like to apply Theorem 7.2. Note that if $A = \gamma[0,1] \in \mathcal{F}$, then $A \cap R \neq \emptyset$ and $\sup_A F \geq \beta$. Hence $\alpha \geq \beta$. Evidently $\alpha < \infty$ because $A \in \mathcal{F}$ is compact. On the other hand, if $A = \gamma[0,1] \in \mathcal{F}$ with $\gamma(0) = a_1 \in \hat{\mathcal{E}}^1$ and $\gamma(1) = a_2 \in \hat{\mathcal{E}}^2$, then $\phi_t(a_j) \in \hat{\mathcal{E}}^j$ for $j = 1, 2$ and $t \geq 0$, because

$$F(\phi_t(a_j)) \leq F(a_j) < \beta$$

and $\phi_t(a_j) \notin R$ by $\inf_R F = \beta$. **(ii)** We simply take $\mathcal{F} = \{\phi_t(\Sigma) : t \geq 0\}$. Evidently $\alpha(F, \mathcal{F}) \leq \sup_\Sigma F < \infty$. Moreover, since $\phi_t(\Sigma) \cap \Gamma \neq \emptyset$, we have

$$\sup_{t \in \phi_t(\Sigma)} F(x) \geq \beta$$

for every $t \geq 0$. We can now apply Theorem 7.3 to complete the proof. $\square$

Our goal is proving Theorem 7.2 with the aid of Lemma 7.1 and Proposition 7.1 for $F = \mathcal{A}$ and $\mathcal{E} = \mathcal{H}^{1/2}$. Let us write $\mathcal{A} = \mathcal{A}_0 - \mathcal{C}$ where

$$\mathcal{A}_0(x) = \frac{1}{2} \int_0^1 \bar{J}x \cdot \dot{x} dt, \quad \mathcal{C}(x) = \int_0^1 H(x) dt.$$

To differentiate $\mathcal{A}$, we need the following operators: given $x = \sum_k e^{2\pi k t \bar{J}} x_k$, set

$$P^{\pm} x = \sum_{\pm k > 0} e^{2\pi k t \bar{J}} x_k, \qquad P^0 x = x_0,$$
$$\mathcal{I}(x) = x_0 + \sum_{k \neq 0} \frac{1}{2\pi |k|} e^{2\pi k t \bar{J}} x_k.$$

When we differentiate $x(\cdot)$, the k-th Fourier coefficient is multiplied by $2\pi k \bar{J} = -2\pi k i$, where as in the definition of $\mathcal{I}(x)$, the k-th coefficient is divided by $2\pi |k|$. In some sense $\mathcal{I}$ is an integration operator and it is easy to see

$$\|\mathcal{I}(x)\|_1 = \|x\|_0. \tag{7.7}$$

Proposition 7.2 **(i)** *The function $\mathcal{A}$ is C^1 with*

$$\nabla \mathcal{A}(x) = P^+ x - P^- x - \nabla \mathcal{C}(x) = P^+ x - P^- x - \mathcal{I}\big(\nabla H(x)\big). \tag{7.8}$$

(ii) *The operator $\mathcal{C}$ is a compact operator with*

$$\|\nabla \mathcal{C}(x) - \nabla \mathcal{C}(y)\|_{1/2} \leq c_0 \|x - y\|,$$

where $c_0 = \sup \|D^2 H\|$.

(iii) *The flow for the vector field* $-\nabla\mathcal{A}$ *is well-defined for all times.*

(iv) *If* $\nabla\mathcal{A}(x) = 0$, *then* x *is* C^1 *and* $\dot{x} = \bar{J}\nabla H(x)$.

(v) *If* $\nabla\mathcal{A}(x_l) \to 0$ *in* $\mathcal{H}^{1/2}$, *then the sequence* $\{x_l\}$ *has a convergent subsequence. In particular, the function* $\mathcal{A}$ *satisfies the (PS) condition.*

Proof (i) Note that $\mathcal{A}_0$ is quadratic and therefore smooth. In fact for $x = \sum_k e^{2\pi kt\bar{J}} x_k$,

$$\mathcal{A}_0(x) = \sum_k \pi k |x_k|^2 = \frac{1}{2}\|P^+ x\|^2 - \frac{1}{2}\|P^- x\|^2,$$

which in turn implies

$$\nabla\mathcal{A}_0(x) = (P^+ - P^-)x. \tag{7.9}$$

We now turn to the functional $\mathcal{C}$. Of course $\mathcal{C}$ is differentiable with respect to L^2-inner product and its derivative is given by $\nabla H(x)$. More precisely, since $\sup \|D^2 H\| = c_0 < \infty$, we have

$$|H(x+h) - H(x) - \nabla H(x)\cdot h| \le \frac{c_0}{2}|h|^2.$$

As a result,

$$|\mathcal{C}(x+h) - \mathcal{C}(x) - \langle \nabla H(x), h\rangle_0| \le \frac{c_0}{2}\|h\|_0^2 \le \frac{c_0}{2}\|h\|_{1/2}^2$$

which implies $\nabla\mathcal{C}(x) = \mathcal{I}(\nabla H(x))$, because

$$\langle \nabla H(x), h\rangle_0 = \langle \mathcal{I}(\nabla H(x)), h\rangle_{1/2},$$

completing the proof of (7.7).

(ii) The operator $x(\cdot) \to \nabla H(x(\cdot))$ maps bounded subsets of $\mathcal{H}^0 = L^2$ to bounded subset of $\mathcal{H}^0$ because $|\nabla^2 H(x)| \le c_0$. The operator $\mathcal{I}$ maps bounded subset of $\mathcal{H}^0$ to bounded subsets of $\mathcal{H}^1$ by (7.7). By Exercise 7.1**(iv)** below, bounded subsets of $\mathcal{H}^1$ are precompact in $\mathcal{H}^{1/2}$. Hence $\nabla\mathcal{C}$ is a compact operator.

For Lipschitzness of $\nabla\mathcal{C}$, observe that by (7.7)

$$\begin{aligned}\|\nabla\mathcal{C}(x) - \nabla\mathcal{C}(y)\|_{1/2} &= \|\mathcal{I}(\nabla H(x) - \nabla H(y))\|_{1/2} \le \|\mathcal{I}(\nabla H(x) - \nabla H(y))\|_1 \\ &= \|\nabla H(x) - \nabla H(y)\|_0 \le c_0\|x-y\|_0 \le c_0\|x-y\|_{1/2}.\end{aligned}$$

(iii) Since $\nabla\mathcal{C}$ is Lipschitz and $\nabla\mathcal{A}_0$ is linear, we learn that $\nabla\mathcal{A}$ is Lipschitz. This guarantees that the gradient flow is well-defined.

(iv) Since $\nabla\mathcal{A}(x)=0$, we have that $P^+x-P^-x=\mathcal{I}\big(\nabla H(x)\big)$. If

$$x=\sum_k e^{2\pi kt\bar{J}}x_k,\quad \nabla H(x)=\sum_k e^{2\pi kt\bar{J}}a_k,$$

then we deduce that $a_0=0$ and $\operatorname{sgn}(k)x_k=(2\pi|k|)^{-1}a_k$ for $k\neq 0$, or $2\pi kx_k=a_k$ for all k. Since $\nabla H(x)\in\mathcal{H}^0=L^2$, we learn that $\sum_k|k|^2|x_k|^2<\infty$, or $x\in\mathcal{H}^1$. From $2\pi kx_k=a_k$, we can readily deduce that $\dot{x}=\bar{J}\nabla H(x)$ weakly. Since the right-hand side is continuous by $\mathcal{H}^1\subseteq C(S^1)$, we learn that $x\in C^1(S^1)$.

(v) *Step 1.* Take a sequence $\{x_l\}$ such that

$$\nabla\mathcal{A}(x_l)=P^+x_l-P^-x_l-\nabla\mathcal{C}(x_l)\to 0, \tag{7.10}$$

as $l\to\infty$. We first prove the boundedness of $\{x_l\}$. Assume to the contrary

$$\lim_{l\to\infty}\|x_l\|_{1/2}=\infty.$$

Observe that if $y_l=\frac{x_l}{\|x_l\|_{1/2}}$, then by (7.8),

$$P^+y_l-P^-y_l-\mathcal{I}\left(\frac{1}{\|x_l\|_{1/2}}\nabla H(x_l)\right)\to 0. \tag{7.11}$$

Now we use

$$\left\|\frac{\nabla H(x_l)}{\|x_l\|_{1/2}}\right\|_0\le c_0\frac{\|x_l\|_0}{\|x_l\|_{1/2}}\le c_0,$$

to deduce that the sequence

$$\mathcal{I}\left(\frac{1}{\|x_l\|_{1/2}}\nabla H(x_l)\right),$$

has a convergent subsequence in $\mathcal{H}^{1/2}$. This and (7.11) implies that the sequence $\{P^+y_l-P^-y_l\}$ has a convergent subsequence. Without loss of generality, we may assume that the sequence $\{P^+y_l-P^-y_l\}$ is convergent. Since P^+ and P^- project onto the positive and negative frequencies, we deduce that both sequences $\{P^+y_l\}$ and $\{P^-y_l\}$ are convergent. Since the sequence $\{y_l\}$ is bounded, the sequence $\{P^0y_l\}$ is bounded in $\mathbb{R}^{2n}$. Hence, by switching to a subsequence if necessary, we may assume that $y_l=P^+y_l+P^-y_l+P^0y_l$ is convergent. Let us continue to use $\{y_l\}$ for such a subsequence and write y for its limit. Recall that for $z\in\mathbb{R}^{2n}$ with large $|z|$,

$$H(z)=(\pi+\epsilon)Q(z)=:\hat{Q}(z)$$

where $Q(z)=q_1^2+p_1^2+l^{-2}\sum_{j=2}^n(q_j^2+p_j^2)$. We now argue

$$\lim_{l\to\infty}\left\|\frac{\nabla H(x_l)}{\|x_l\|_{1/2}}-\nabla\hat{Q}(y)\right\|_0=0. \tag{7.12}$$

To see this, observe

$$\left\| \frac{\nabla H(x_l)}{\|x_l\|_{1/2}} - \nabla \hat{Q}(y) \right\|_0 \leq \|\nabla H(x_l) - \nabla \hat{Q}(x_l)\| \|x_l\|_{1/2}^{-1} + \|\nabla \hat{Q}(y_l) - \nabla \hat{Q}(y)\|_0.$$

Now (7.12) follows because $|\nabla H - \nabla \hat{Q}|$ is uniformly bounded and $\|y_l - y\|_{1/2} \to 0$. From (7.12) and (7.7) we deduce

$$\lim_{l\to\infty} \frac{1}{\|x_l\|_{1/2}} \mathcal{I}(\nabla H(x_l)) = \mathcal{I}\big(\nabla \hat{Q}(y)\big),$$

in $\mathcal{H}^{1/2}$. From this, $\lim_{l\to\infty} y_l = y$, and (7.11) we deduce

$$P^+ y - P^- y - \mathcal{I}\big(\nabla \hat{Q}(y)\big) = 0,$$

for y satisfying $\|y\| = 1$. This means that y is a critical point of

$$\mathcal{A}_1(y) = \int_0^1 \left[\frac{1}{2} \bar{J} y \cdot \dot{y} - \hat{Q}(y)\right] \, dt.$$

We then apply part **(iv)** where H is replaced with $\hat{Q}$. As a result y is C^1 and $\dot{y} = \bar{J} \nabla \hat{Q}(y)$. Hence

$$y(t) = \left(e^{-2(\pi+\epsilon)it} a_1, e^{-2\frac{(\pi+\epsilon)}{l^2} it} a_2, \ldots, e^{-2\frac{(\pi+\epsilon)}{l^2} it} a_n\right).$$

This is 1-periodic only if $y \equiv 0$, contradicting $\|y\| = 1$. Hence the sequence $\{x_l\}$ must be bounded.

Step 2. For a sequence $\{x_l\}$ satisfying (7.10), we know that $\{x_l\}$ is bounded. This implies the precompactness of $\{\nabla \mathcal{C}(x_l)\}$ by the compactness of the operator $\nabla \mathcal{C}$. This and (7.10) imply that $x_l^+ - x_l^-$ has a convergent subsequence. As in *Step 1*, we learn that $\{P^+ x_l + P^- x_l\}$ has a convergent subsequence. From this we deduce that $\{x_l\}$ has a convergent subsequence because $\{P^0 x_l\}$ is also bounded. □

We are now ready to establish Theorem 7.2. On the account of Lemma 7.1, we need to find a critical point of $\mathcal{A}$ with $\mathcal{A}(x) > 0$. Note that our Hamiltonian H is supposed to vanish on some neighborhood of the origin.

Proof of Theorem 7.2 *Step 1.* To ease the notation, let us write $x^\pm = P^\pm x$, and $x^0 = P^0 x$. We also write

$$\mathcal{H}^{1/2} = \mathcal{E} = \mathcal{E}^- \oplus \mathcal{E}^0 \oplus \mathcal{E}^+,$$

where $\mathcal{E}^\pm$ and $\mathcal{E}^0$ denote the ranges of the operators $P^\pm$ and P^0 respectively. Let us use the *variation of constants formula* to derive a nice representation for the flow ϕ_t of the vector field $-\nabla \mathcal{A}$:

$$\phi_t(x) = e^t x^- + x^0 + e^{-t} x^t + K(x, t), \tag{7.13}$$

for a function $K : \mathcal{E} \times \mathbb{R} \to \mathcal{E}$ that is continuous and compact. To see this, observe that the flow $\bar{\phi}_t$ of the vector field $-\nabla \mathcal{A}_0 = P^- - P^+$ is simply given by

$$\bar{\phi}_t(x) = e^t x^- + x^0 + e^{-t} x^+.$$

From this, we can readily deduce (7.13) with K given by

$$\begin{aligned} K(x,t) &= \int_0^t (e^{t-s}P^- + P^0 + e^{s-t}P^+)\nabla \mathcal{C}(\phi_s(x))ds \\ &= \mathcal{I} \int_0^t (e^{t-s}P^- + P^0 + e^{s-t}P^+)\nabla H(\phi_s(x))ds. \end{aligned}$$

Here we are using the fact that $P^\pm$ and P^0 commute with $\mathcal{I}$. Now the compactness of K follows from the facts that the flow ϕ_s maps bounded sets to bounded sets (see Exercise 7.1**(iii)**) and that $\mathcal{I}$ is a compact operator.

Step 2. To establish Theorem 7.2 with the aid of Proposition 7.1**(ii)**, we need to come up with suitable candidates for the sets Γ and Σ. Define

$$\begin{aligned} &\Gamma = \Gamma(r) = \{x \in \mathcal{E}^+ : \|x\| = r\}, \\ &\Sigma = \Sigma(\theta) = \{x : x = x^- + x^0 + se^+ : \|x^- + x^0\| \le \theta,\ 0 \le s \le \theta\} \end{aligned}$$

where $e^+(t) = e^{2\pi t \bar{J}} e_1$ with $e_1 = (1, 0, \dots, 0) \in \mathbb{R}^{2n}$. Evidently $\sup_\Sigma \mathcal{A} < \infty$, because

$$\mathcal{A}(x) \le \frac{1}{2}\|P^+x\|_{1/2}^2 \le \frac{\theta^2}{2}\|e^+\|_{1/2}^2.$$

Let us check that indeed $\inf_{\Gamma(r)} \mathcal{A} = \beta > 0$, for $r > 0$ sufficiently small. Recall that $|H(z)| \le c_1|z|^2$ for a constant c_1. This however doesn't do the job and we need to use the fact that H vanishes near the origin. This property implies that for every $p \in (2, \infty)$, we can find a constant $c_1(p)$ such that $|H(z)| \le c_1(p)|z|^p$. On the other hand, by a well-known Sobolev-type inequality,

$$\|x\|_{L^p} = \left(\int_0^1 |x(t)|^p\ dt\right)^{\frac{1}{p}} \le c_0(p)\|x\|_{1/2},$$

for a universal constant $c_0(p)$ (see Appendix B). As a result, if $x \in \Gamma(r)$, and $p > 2$, then

$$\begin{aligned} \mathcal{A}(x) &= \frac{1}{2}\|x^+\|_{1/2}^2 - \int_0^1 H(x(t))\ dt \ge \frac{1}{2}\|x^+\|_{1/2}^2 - c_1(p)\int_0^1 |x(t)|^p\ dt \\ &\ge \frac{1}{2}\|x\|_{1/2}^2 - c_1(p)c_0(p)^p\|x\|^p = \frac{1}{2}r^2 - c_1(p)c_0(p)^p r^p =: \beta. \end{aligned}$$

Evidently $\beta > 0$, if r is sufficiently small.

Step 3. On account of Proposition 7.1**(ii)**, it suffices to show that if θ is sufficiently large, then $\phi_t(\Sigma(\theta)) \cap \Gamma \neq \emptyset$ for all $t \geq 0$. The idea is that in some sense $\partial\Sigma$ and Γ *link* with respect to ϕ_t. That is, $\phi_t(\partial\Sigma)$ can not cross the circle Γ as t increases, so Γ must intersect the "frame" $\phi_t(\Sigma)$. For this to work, we first show

$$\phi_t(\partial\Sigma) \cap \Gamma = \emptyset, \tag{7.14}$$

for every $t \geq 0$. If fact for (7.14) the requirement $H(x) = (\pi+\varepsilon)Q(x)$ with $\pi+\varepsilon > \frac{1}{2}\|e^+\|_{1/2}$ is used in an essential way. This is the only place that the condition of $\max_Z H_0 > \pi$ of Theorem 7.2 is used. Since $\inf_\Gamma \mathcal{A} = \beta > 0$, it suffices to show

$$\sup_{t\geq 0} \sup_{\phi_t(\partial\Sigma)} \mathcal{A} \leq 0.$$

Since $\frac{d}{dt}\mathcal{A}(\phi_t(x)) \leq 0$, it suffices to show

$$\sup_{\partial\Sigma} \mathcal{A} \leq 0. \tag{7.15}$$

We write $\partial\Sigma = \partial_1\Sigma \cup \partial_2\Sigma$ where $\partial_1\Sigma = \{x \in \partial\Sigma : x = x^- + x^0\}$. In the case of $x \in \partial_1\Sigma$, we have

$$\mathcal{A}(x) = -\frac{1}{2}\|x^-\|^2 - \int_0^1 H(x)\ dt \leq 0$$

because $H \geq 0$. It remains to show that $\sup_{\partial_2\Sigma} \mathcal{A} \leq 0$, for sufficiently large θ. Recall that there exists a constant c_1 such that $H(x) \geq (\pi+\epsilon)Q(x) - c_1$. Hence, for $x = x^- + x^0 + se^+$,

$$\begin{aligned}
\mathcal{A}(x) &= \frac{1}{2}s^2\|e^+\|^2 - \frac{1}{2}\|x^-\|^2 - \int_0^1 H(x)\ dt \\
&\leq \pi s^2 - \frac{1}{2}\|x^-\|^2 - (\pi+\epsilon)\int_0^1 Q(x^- + x^0 + se^+)\ dt + c_1 \\
&= \pi s^2 - \frac{1}{2}\|x^-\|^2 - (\pi+\epsilon)\left[\int_0^1 Q(x^-)\ dt + \int_0^1 Q(x^0)\ dt + \int_0^1 Q(se^+)\ dt\right] + c_1 \\
&\leq -\frac{1}{2}\|x^-\|^2 - (\pi+\epsilon)\left[\int_0^1 Q(x^-)\ dt + Q(x^0)\right] - \epsilon s^2 + c_1 \\
&\leq c_1 - c_2(\|x^- + x^0\|^2 + \|se^+\|^2)
\end{aligned}$$

for some constant c_2, where for the first inequality we used $H \geq (\pi+\epsilon)Q - c_1$ and $\|e^+\|^2 = 2\pi$, and for the second equality we used the fact that Q is quadratic and that $\mathcal{E}^+, \mathcal{E}^-, \mathcal{E}^0$ are orthogonal with respect to L^2-inner product. It is now clear that if either $\|x^- + x^0\| = \theta$ or

$s = \theta$ with θ sufficiently large, then $\mathcal{A}(x) \le 0$, proving $\sup_{\partial_2\Sigma} \mathcal{A} \le 0$. This completes the proof of (7.15) which in turn implies (7.14).

Step 3. To show that $\phi_t(\Sigma) \cap \Gamma \neq \emptyset$ for $t \ge 0$, we need to find $x \in \Sigma$ such that $\|\phi_t(x)\| = r$ and $(P^- + P^0)\phi_t(x) = 0$. The latter means

$$e^t x^- + x^0 + (P^- + P^0)K(x,t) = 0,$$

or equivalently,

$$x^0 + P^0 K(x,t) = 0, \qquad x^- + e^{-t} P^- K(x,t) = 0. \tag{7.16}$$

To combine this with the former condition $\|\phi_t(x)\| = r$, define

$$L(x,t) = (e^{-t}P^- + P^0)K(x,t) + P^+\{(\|\phi_t(x)\| - r)e^+ - x\}.$$

We can readily show that $\phi_t(\Sigma) \cap \Gamma \neq \emptyset$ is equivalent to finding $x \in \Sigma$ such that

$$x + L(x,t) = 0. \tag{7.17}$$

We wish to find a solution of (7.17) in the interior of Σ, which is an open bounded subset of

$$\hat{\mathcal{E}} = \mathcal{E}^- \oplus \mathcal{E}^0 \oplus \mathbb{R}e^+.$$

Note that $L : \hat{\mathcal{E}} \times \mathbb{R} \to \hat{\mathcal{E}}$ is a compact operator simply because K is compact and $\mathcal{E}^+$ part of $\hat{\mathcal{E}}$ is one-dimensional.

Step 4. We will use Leray-Schauder degree theory to solve (7.17). (We refer to Appendix C, for a review of Degree Theory.) Note that by (7.14),

$$0 \notin (I + L(\cdot,t))(\partial\Sigma).$$

Because of this, $\deg_0(I + L(\cdot,t))$ is well-defined. Observe that $(I + L(\cdot,s) : s \in [0,t])$ defines a homotopy, which implies

$$\deg_0(I + L(\cdot,t)) = \deg_0(I + L(\cdot,0)). \tag{7.18}$$

From the definition of K given in (7.13), we know that $K(\cdot,0) \equiv 0$. As a result,

$$L(x,0) = P^+\{(\|x\| - r)e^+ - x\}.$$

To calculate the right-hand side of (7.18), let us define

$$L^\alpha(x) = P^+\{(\alpha\|x\| - r)e^+ - \alpha x\},$$

for $\alpha \in [0,1]$. We claim

$$0 \notin (I + L^\alpha)(\partial\Sigma). \tag{7.19}$$

Indeed if $(I + L^\alpha)(x) = 0$ for some $x \in \partial\Sigma$, then $x = se^+$ for some $s \in \{0, \theta\}$ and $s + \alpha s\|e^+\| - r - \alpha s = 0$, or $s((1-\alpha) + \alpha\sqrt{2\pi}) = r$ because $\|e^+\| = \sqrt{2\pi}$. Evidently $s \neq 0$ because $r \neq 0$. To rule out $s = \theta$, observe that we may take θ large enough to have $\theta > r$. But $r = s((1-\alpha) + \sqrt{2\pi}\alpha) > s$ which implies that $\theta > s$. In summary $x + L^\alpha(x) = 0$ has no solution in $\partial\Sigma$, or equivalently (7.19) holds.

By (7.19), $\deg_0(I + L^\alpha)$ is well-defined and by the homotopy invariance of degree,

$$\begin{aligned} \deg_0(I + L(\cdot, 0)) &= \deg_0(I + L^1) = \deg_0(I + L^0) \\ &= \deg_0(I - re^+) = \deg_{re^+}(I) = 1, \end{aligned}$$

provided that $re^+ \in \Sigma$, which is true by our assumption $r < \theta$. From this and (7.18) we deduce that (7.17) has a solution $x \in \Sigma$, for which $\mathcal{A}(x) \geq \beta > 0$. This completes the proof of Theorem 7.2 . □

Remark 7.1 The proof of Theorem 7.1 is similar to the proof of Theorem 7.2. We refer to [HZ] for details and only give a list of adjustments that we need to make to the proof of Theorem 7.2:

(i) Take a bounded compact set K with 0 in its interior. Given $H_0 \in \mathcal{H}(K)$ with $\max H_0 > \mathbf{c}_0(K)$, pick $m > \max H$, and define

$$H(x) = \begin{cases} H_0(x) & \text{if } x \in K, \\ f(Q(x)) & \text{if } x \notin K. \end{cases}$$

where $Q = g_K^2$, and $f : \mathbb{R} \to \mathbb{R}$ is a smooth function such that $f(r) = \max H$ for $r \in [0,1]$, $f(r) = mr$ for large r, $f(r) \geq mr$ for all r, and $0 < f'(r) \leq m$ for $r > 1$. We then have the analog of Lemma 7.1 for H provided that the ellipsoid E is replaced with K.

(ii) As we verify PS condition, let y be as in the proof of Proposition 7.2**(v)**. We need to get a contradiction from the conditions $\|y\|_{1/2} = 1$ and $\dot{y} = m\bar{J}\nabla Q(y)$, where $Q = g_K^2$. Note that y is a 1-periodic solution of X_{mQ}. On the other hand,

$$T^{min}(mQ) = (2m)^{-1}\, T^{min}(H_K) = (2m)^{-1}\, 2\mathbf{c}_0(K) < 1,$$

where we used Remark 5.1**(ii)** for the second equality. This of course does not lead to a contradiction. However, it is possible to find a new compact set K' near K so that K' has no 1-periodic orbit. We then choose $Q = g_{K'}^2$ in part **(i)** in our extension.

(iii) We now explain what the analog of the orbit e^+ in the definition of § is. Set $\bar{m} = \mathbf{c}_0(K)$ and consider the Hamiltonian equation $\cdot x = \bar{m}\bar{J}\nabla Q(x)$. This corresponds to the Hamiltonian function $\bar{m}Q$ with

$$T^{min}(\bar{m}Q) = (2\bar{m})^{-1}\, T^{min}(H_K) = 1,$$

which means that the vector field $\bar{m}X_Q$ has a 1-periodic orbit $\bar{x}$ that lies on ∂K. We then set $e^+ = P^+\bar{x}$. We note that $A(\bar{x}) = \bar{m} > 0$ which implies that $e^+ \neq 0$, because $2A(\bar{x}) = \|P^+\bar{x}\|^2 - \|P^-\bar{x}\|^2$. In fact it was in the proof of

$$\sup_{\partial\Sigma} \mathcal{A} \leq 0, \tag{7.20}$$

in *Step 3* that the form of e^+ played an essential role. We refer to [HZ] for the proof of (7.20). □

Exercise 7.1(i) Assume $F \in C^1(\mathbb{R}^d;\mathbb{R})$ satisfies $\lim_{|x|\to\infty} F(x) = \infty$ and that F possesses two distinct relative minima x_1 and x_2. Show that F has a third critical point x_3 that is different from x_1 and x_2. **Hint:** Use paths connecting x_1 to x_2 for the members of $\mathcal{F}$ in Theorem 7.2.

(ii) Show that $F(x,y) = e^{-x} - y^2$ does not satisfy Palais–Smale condition. Let

$$E^{\pm} = \{(x,y) : F(x,y) \leq 0,\ \pm y \geq 0\}$$

and set

$$\mathcal{F} = \{\gamma[0,1] : \gamma : [0,1] \to \mathbb{R}^2,\ \gamma(0) \in E^-,\ \gamma(1) \in E^+ \text{ and } \gamma \text{ is continuous}\}.$$

Show that $\alpha(F,\mathcal{F}) = 0$ but there is no x^* with $F(x^*) = 0$ and that F has no critical point.

(iii) Let $G : \mathcal{E} \to \mathcal{E}$ be a Lipschitz function and let ϕ_t denote the flow of the ODE $\dot{x} = G(x)$. Show that for every $l > 0$,

$$\sup_{0\leq t\leq l}\ \sup_{\|x\|\leq l} \|\phi_t(x)\| < \infty.$$

(iv) Let $\mathcal{H}^s = \{x \in L^2 : \|x\|_s < \infty\}$. Show that if $s < t$, then a bounded subset of H^t is precompact in H^s.

(v) Show that if $x \in \mathcal{H}^\theta$ and $\theta > 1/2$, then x is Hölder continuous with

$$|x(t) - x(s)| \leq c\|x\||t-s|^\alpha$$

with $\alpha = \min\left(1, \theta - \frac{1}{2}\right)$. **Hint:** Given $x = \sum_k e^{2\pi kt\bar{J}}x_k$, write $x = y + z$ with $z = \sum_{|k|\leq N} e^{2\pi kt\bar{J}}x_k$. Estimate $\sup_{|t-s|<\delta}|y(t)-y(s)|$ and $\sup_t |z(t)|$ in terms of δ and N. This yields a bound for $|x(t)-x(s)|$ that can be minimized with respect to N. □

8 Hofer Geometry

9 Generating Function, Twist Map and Arnold's Conjecture

A Hamiltonian vector field X_H is generated by a scalar-valued function H. It turns out that a similar phenomenon is true for any symplectic diffeomorphism at least locally. To explain this, let us take a symplectic diffeomorphism $\psi : U \to \mathbb{R}^{2n}$ with U a simply connected open subset of $\mathbb{R}^{2n}$, and write

$$\psi(q, p) = (Q, P)$$

with Q and $P \in \mathbb{R}^n$. Since $\psi^* d\bar{\lambda} - d\bar{\lambda} = d(\psi^*\bar{\lambda} - \bar{\lambda}) = 0$ for $\bar{\lambda} = q \cdot dp$, we have that there exists a scalar-valued function S such that $\psi^*\bar{\lambda} - \bar{\lambda} = dS$. In coordinates,

$$P \cdot dQ - p \cdot dq = dS. \tag{9.1}$$

The form of (8.1) suggests that perhaps we should regard S as a function of q and Q so that (8.1) is equivalent to

$$\frac{\partial S}{\partial Q} = P \text{ and } \frac{\partial S}{\partial q} = -p. \tag{9.2}$$

The scalar-valued function S is an example of a *generating function.* Its existence is guaranteed if we make some non-degeneracy assumptions on ψ.

Proposition 9.1 *Let $\psi : U \to \mathbb{R}^{2n}$ be a symplectic transformation and assume that at $(q^0, p^0) \in U$,*

$$\det \frac{\partial Q}{\partial p}(q^0, p^0) \neq 0. \tag{9.3}$$

Then there exist a neighborhood V of q^0 and $Q^0 = Q(q^0, p^0)$, and a C^1 function $S : V \to \mathbb{R}$ such that (9.2) *holds.*

Proof From (9.3) and Implicit Function Theorem, the relation $Q = Q(q, p)$ can be solved for $p = p(q, Q)$ for q and Q near q^0 and Q^0. We then set $P(q, Q) = P(q, p(q, Q))$. To solve (8.2) for S, we need to verify the solvability criterion

$$P_q + p_Q = 0,$$

regarding P and p as functions of q and Q. This is exactly $d(P \cdot dQ - p \cdot dq) = 0$. □

Later we will discuss other types of generating functions. But let us first study some examples. As our first example, consider $\psi = \phi_{t^0 t}$ where $\phi_{t^0 s}$ is the flow of the Hamiltonian ODE

$$\dot{x} = \bar{J}\nabla H(x, s). \tag{9.4}$$

More precisely, $x(s) = \phi_{t^0 s}(a)$ solves (9.4) subject to the condition $\phi(t^0) = a$. Let us write $\alpha(s)$ and $\beta(s)$ for the q and p components of $x(s)$. We assume that for some open set V, the equation (9.4) can be solved if $(q, Q) \in V$ is specified. More precisely, if $x(s) = (\alpha(s), \beta(s))$ with $\alpha, \beta \in \mathbb{R}^n$, then (9.4) has a unique solution subject to the initial and terminal conditions $\alpha(t^0) = q$ and $\alpha(t) = Q$. We then set

$$S(q, Q; t^0, t) = \int_{t^0}^{t} [\beta(s) \cdot \dot{\alpha}(s) - H(x(s), s)]ds, \tag{9.5}$$

with $x(s) = x(s; q, Q)$.

Proposition 9.2 *Under the above conditions, the function S is a generating function for $\psi = \phi_{t^0 t}$. Moreover, S satisfies the Hamilton–Jacobi equation*

$$S_t + H(Q, S_Q, t) = 0. \tag{9.6}$$

Proof Differentiating both sides of (9.5) with respect to q_j yields

$$\begin{aligned} S_{q_j} &= \int_{t^0}^{t} [\beta_{q_j} \cdot \dot{\alpha} + \beta \cdot \dot{\alpha}_{q_j} - \nabla H \cdot x_{q_j}]ds \\ &= \int_{t^0}^{t} [\beta_{q_j} \cdot \dot{\alpha} - \alpha_{q_j} \cdot \dot{\beta} - \nabla H \cdot x_{q_j}]ds + \beta(t) \cdot \alpha_{q_j}(t) - \beta(t^0) \cdot \alpha_{q_j}(t^0) \\ &= \int_{t^0}^{t} [-\bar{J}\dot{x} - \nabla H(x, s)] \cdot x_{q_j} ds + \beta(t) \cdot \alpha_{q_j}(t) - \beta(t^0) \cdot \alpha_{q_j}(t^0). \end{aligned} \tag{9.7}$$

The first term vanishes because of (9.4). On the other hand, since

$$\alpha(t^0; q, Q) = q, \quad \beta(t^0; q, Q) = p \quad \alpha(t; q, Q) = Q,$$

we learn that $\alpha_{q_j}(t) = 0$, and $\alpha_{q_j}(t^0) = e^j$, where e^j denotes the standard unit j-th vector. As a result, $S_{q_j} = -p_j$. The proof of $S_{Q_j} = P_j$ is similar.

As for (9.6), first observe

$$S(q, \alpha(t); t^0, t) = \int_{t^0}^{t} [\beta(s) \cdot \dot{\alpha}(s) - H(\alpha(s), s)]ds.$$

Differentiating both sides with respect to t yields

$$S_Q \cdot \dot{\alpha} + S_t = \beta(t) \cdot \dot{\alpha}(t) - H(\alpha(t), t).$$

This immediately implies (8.6) because $S_Q = P = \beta$ and $\alpha(t) = (Q, P) = (Q, S_Q)$. $\square$

As our second example, let us study generating functions in the simplest case $n = 1$. For this we consider symplectic $\varphi : A \to A$ with

$$A = \{(q,p) : R_-^2 \leq q^2 + p^2 \leq R_+^2\}.$$

Such a transformation was encountered by Poincaré as he used a Poincaré's section to study solutions to Hamiltonian systems. Poincaré was interested in fixed points of φ because they correspond to periodic orbits of the corresponding Hamiltonian system. As we will see such fixed points exist if φ is a twist map.

A function $\varphi : A \to A$ is called a *twist map* if the following conditions are met:

(i) φ is a homeomorphism and the restriction of ψ to A^0 is a diffeomorphism with $\det \varphi' \equiv 1$.

(ii) φ maps the circles $C_\pm = \{q^2 + p^2 = R_\pm^2\}$ to themselves with $\deg \varphi|_{C_\pm} = \pm 1$.

Our main result about twist maps is the following result of Poincaré and Birkhoff.

Theorem 9.1 *Any twist map has at least two fixed points.*

In fact Poincaré established Theorem 9.1 provided that φ has a global generating function. Such a generating function exists if φ is a *monotone twist map*. To prepare for this, let us first observe that any twist map on A yields a twist map on the cylinder $S^1 \times [R^-, R^+]$. Indeed, if $h : S^1 \times [R^-, R^+] \to A$ is given by $h(x,y) = (\sqrt{y} \cos 2\pi x, \sqrt{y} \sin 2\pi x)$, then $\psi = h^{-1} \circ \varphi \circ h$ is again orientation and area preserving because $dq \wedge dp = -\pi dx \wedge dy$. Now $\psi : S^1 \times [R^-, R^+] \to S^1 \times [R^-, R^+]$ has a lift

$$\Psi : \mathbb{R} \times [R^-, R^+] \to \mathbb{R} \times [R^-, R^+]$$

which satisfies the following conditions:

(i) Ψ is a homeomorphism and the restriction of Ψ to $\mathbb{R} \times (R^-, R^+)$ is a diffeomorphism with $\det \Psi' \equiv 1$.

(ii) Ψ maps $\mathbb{R} \times \{R^\pm\}$ onto itself with

$$\Psi(x, R^\pm) = (\pm x + f^\pm(x), R^\pm),$$

where $f^\pm$ is 1-periodic.

(iii) $\Psi(x+1, y) = \Psi(x,y) + (1,0)$.

Such a map Ψ is again called a twist map. We also write $\Psi(x,y) = (X,Y)$ with X and Y functions of (x,y). We note that we may replace (ii) with the condition $Y(x,\pm R) = \pm R$ and $\pm A_{\pm}(x) := \pm X(x,\pm R) > \pm x$.

We now formulate a condition on Ψ that would guarantee the existence of a global generating function $S(x,X)$ for Ψ. A twist map Ψ is called *monotone* if

$$\frac{\partial X}{\partial y}(x,y) > 0, \tag{9.8}$$

for all $(x,y) \in \mathbb{R} \times (R^-, R^+)$.

Proposition 9.3 *Let Ψ be a monotone twist map. Then there exists a C^2 function $S : U \to \mathbb{R}$ with*

$$U = \{(x,X) : A_-(x) < X < A_+(x)\}$$

such that

$$\Psi(x, -S_x(x,X)) = (X, S_X(x,X)).$$

Moreover

$$S(x+1, X+1) = S(x,X), \qquad S_{xX} < 0. \tag{9.9}$$

Proof The image of the line segment $\{x\} \times [R^-, R^+]$ under Ψ is a curve γ with parametrization $\gamma(y) = (X(x,y), Y(x,y))$. By (9.8), the relation $X(x,y) = X$ can be inverted to yield $y = y(x,X)$ which is increasing in X. The set $\gamma[R^-, R^+]$ can be viewed as a graph of the function

$$X \mapsto Y(x, y(x,X))$$

with $X \in [A^-(x), A^+(x)]$. The anti-derivative of this function yields S. This can be geometrically described as the area of the region Δ between the curve $\gamma([R^-, R^+])$, the line $Y = R^-$ and the vertical line $\{x\} \times [R^-, R^+]$. We now apply Ψ^{-1} on this region. The line segment $\{X\} \times [R^-, R^+]$ is mapped to a curve $\hat{\gamma}([R^-, R^+])$ which coincides with a graph of a function $x \mapsto y$. Since Ψ is area preserving the area of $\Psi^{-1}(\Delta)$ is $S(x,X)$. From this we deduce that $S_X = -y$. Here we have used the fact that Ψ^{-1} is a (negative) twist map. (Negative because the degree of Ψ^{-1} restricted to the top boundary is -1 whereas the degree of Ψ^{-1} restricted to the bottom boundary is 1.) This is because if we write $\Psi^{-1}(X,Y) = (x(X,Y), y(X,Y))$, then

$$(\Psi^{-1})' = \begin{bmatrix} x_X & x_Y \\ y_X & y_Y \end{bmatrix} = \begin{bmatrix} X_x & X_y \\ Y_x & Y_y \end{bmatrix}^{-1} = \begin{bmatrix} Y_y & -X_y \\ -Y_x & X_x \end{bmatrix}$$

which implies that $\frac{\partial x}{\partial Y} = -\frac{\partial X}{\partial y} < 0$.

The periodicity (9.9) is an immediate consequence of $\Psi(x+1,y) = \Psi(x,y) + (1,0)$;

$$\Psi(\{x+1\} \times [R^-, R^+]) = \Psi(\{x\} \times [R^-, R^+]) + (1,0).$$

As for the second assertion in (9.9), recall that $y(x, X)$ is increasing in X. Hence

$$S_{xX} = -y_X < 0.$$

$\square$

A partial converse to Proposition 9.3 is true, namely if a function S satisfies (9.9), then it generates a map Ψ which is area preserving. We don't address the behavior of Ψ on the boundary lines and for simplicity assume that S is defined on $\mathbb{R}^2$.

Proposition 9.4 *Let S be a C^2 function satisfying* (9.9)*. Then there exists a C^1-function Ψ such that*

(i) $\Psi(x+1, y) = \Psi(x, y) + (1, 0)$

(ii) $\Psi(x, -S_x(x, X)) = (X, S_X(x, X))$

(iii) $\det \Psi' \equiv 1$.

Proof. Since $S_{xX} < 0$, the function $X \mapsto -S_x(x, X)$ is increasing. As a result, $y = -S_x(x, X)$ can be inverted to yield $X = X(x, y)$. We then set

$$Y(x, y) = S_X(x, X(x, y)) \text{ and } \Psi(x, y) = (X(x, y), Y(x, y)).$$

Evidently (ii) is true and (i) follows from (ii) and (8.9) because $S_x(x+1, X+1) = S_x(x, X)$, and $S_X(x+1, X+1) = S_X(x, X)$. It remains to verify (iii). For this, set $\hat{S}(x, y) = S(x, X(x, y))$. We have

$$\begin{aligned} \hat{S}_x &= S_x + S_X X_x = -y + Y X_x, \\ \hat{S}_y &= S_X X_y = Y X_y. \end{aligned}$$

Differentiating again yields

$$\begin{aligned} \hat{S}_{xy} &= -1 + Y_y X_x + Y X_{xy}, \\ \hat{S}_{yx} &= Y_x X_y + Y X_{yx}. \end{aligned}$$

Since $S \in C^2$, we must have $\hat{S}_{xy} = \hat{S}_{yx}$, which yields $Y_y X_x - Y_x X_y = 1$, as desired. $\square$

We now show how the existence of a generating function can be used to prove the existence of fixed points.

Proof of Theorem 9.1 for a monotone twist map Define $L(x) = S(x, x)$. We first argue that a critical point of L corresponds to a fixed point of Ψ. Indeed, if $L'(x^0) = 0$, then $S_x(x^0,x^0)+S_X(x^0,x^0)=0$. Since $\Psi(x^0,-S_x(x^0,x^0))=(x^0,S_X(x^0,x^0))$, we deduce that $\Psi(x^0,y^0)=(x^0,y^0)$ for $y^0=-S_x(x^0,x^0)=S_X(x^0,x^0)$. On the other hand, by (9.9), we have that $L(x+1)=L(x)$. Either L is identically constant which yields a continuum of fixed points for Ψ, or L is not constant. In the latter case, L has at least two distinct critical points, namely a maximizer and minimizer. These yield two distinct critical points of Ψ. □

Before we discuss the proof of Theorem 9.1 for general twist maps, let us study an example of a map which is not quite a twist map but still possesses a global generating function.

Example 9.1 (Billiard map in a convex domain). Let C be a strictly bounded convex domain in $\mathbb{R}^2$ and denote its boundary by S. Without loss of generality, we assume that the total length of S is 1. First we describe the billiard flow in C. This is the flow associated with the Hamiltonian function $H(q,p)=\frac{1}{2}|p|^2+V(q)$ where

$$V(q)=\begin{cases}0 & \text{if } q\in C\\ \infty & \text{if } q\notin C.\end{cases}$$

Here is the interpretation of the corresponding flow: A ball of velocity p starts from a point $q\in C$ and is bounced off the boundary S by the law of reflection. This induces a transformation for the hitting location and reflection angle. More precisely, if a trajectory $q+tp$, $t>0$ hits the boundary at a point $\gamma(x)$ and a post-reflection angle θ, then we write $\gamma(X)$ and Θ for the location and post-reflection angle of the next reflection. Here x is the length of arc between a reference point $A\in S$ and $\gamma(x)$ on S in positive direction, and θ measures the angle between the tangent at $\gamma(x)$ and the post-reflection velocity vector. We write φ for the map $(x,\theta)\mapsto(X,\Theta)$ with $x,X\in S^1$ and $\theta,\Theta\in[0,\pi]$. It is more convenient to define $y=-\cos\theta$ so that in the (x,y) coordinates, we have a map $\psi:S^1\times[-1,1]\to S^1\times[-1,1]$. As before, we write Ψ for its lift. We claim that Ψ is a monotone twist map except that the twist conditions on the boundary lines $y=\pm1$ are violated. We show this by applying Proposition 8.5. In fact the generating function is simply given by

$$S(x,X)=-|\gamma(x)-\gamma(X)|,$$

because

$$\begin{aligned}-S_x(x,X)&=-\frac{(\gamma(X)-\gamma(x))}{|\gamma(X)-\gamma(x)|}\cdot\dot\gamma(x)=-\cos\theta,\\ S_X(x,X)&=-\frac{(\gamma(X)-\gamma(x))}{|\gamma(X)-\gamma(x)|}\cdot\dot\gamma(X)=\cos\Theta,\\ S_{Xx}(x,X)&=\sin\Theta\frac{\partial\Theta}{\partial x}.\end{aligned}$$

Note that if $\Theta \in (0, \pi)$, then $\sin\Theta > 0$, and Θ is decreasing in x which means that $S_{Xx} < 0$. Here of course we are using the strict convexity. As for the boundary lines, we have $\Psi(x, -1) = (x, -1)$, $\Psi(x, 1) = (x + 1, 1)$. Note that $S(x, X)$ is defined for (x, X) satisfying $X \in [x, x + 1]$. Also note that Ψ has no fixed point inside $\mathbb{R} \times (-1, 1)$. □

The generating function S can be used to study periodic orbits of monotone twist maps. To explain this, let us observe that the twist condition means that if ρ^+ and ρ^- denote the rotation numbers of the top and bottom boundary circles of the cylinder, then $\rho^- < 0 < \rho^+$. By rotation number we mean

$$\rho^{\pm} = \lim_{n\to\infty} \frac{A_{\pm}^n(x)}{n}. \tag{9.10}$$

In fact it is well-known that the limit in (9.10) exists because $A_\pm$ is a lift of a circle homeomorphism, and that $\pm\rho^\pm > 0$ because $\pm A_\pm(x) > x$. Now we can interpret Theorem 9.1 as saying that since $0 \in (\rho^-, \rho^+)$, the map Ψ has an orbit which projects onto an x-sequence of 0 rotation number, namely a fixed point. The following theorem generalizes this property to assert the existence of an orbit which projects onto an x-sequence of rotation number $\rho \in (\rho^-, \rho^+)$ provided that ρ is rational. To this end, let us formulate a definition concerning periodic orbits. We say that the point (x, y) is (r, s)-*periodic point* with $r, s \in \mathbb{N}$ and r, s relatively prime, if $(x_n, y_n) = \Psi^n(x, y)$ satisfies $x_{n+s} = x_n + r$ for every n. This means that on cylinder, the x-projection of the orbit $(\psi^n(x, y) : n \in \mathbb{Z})$ wraps r times around the cylinder in s iterates.

Theorem 9.2 *Let Ψ be a twist map. If $\rho \in (\rho^-, \rho^+)$ with $\rho = \frac{r}{s}$, r and s coprime, then Ψ has at least two (r, s)-periodic orbits.*

We do not present a full proof of Theorem 9.2. We only indicate that its proof is very similar to the proof of Theorem 9.1 and uses a variational principle. If Ψ is a monotone twist map, then the variational principle is the discrete analog of the Lagrange variational principle, as can be seen in the following proposition.

Proposition 9.5 *Let Ψ be a monotone twist map with generating function S. Then the following statements are true.*

(i) *Given x and $X \in \mathbb{R}$, the sequence $x_1, x_2, \ldots, x_{n-1}$ is a critical point of*

$$L(x_1, x_2, \ldots, x_{n-1}) = \sum_{j=0}^{n-1} S(x_j, x_{j+1}),$$

with $x_0 = x$, and $x_n = X$, if and only if there exist $y_0, y_1, \ldots, y_n$ such that $\Psi^j(x_j, y_j) = (x_{j+1}, y_{j+1})$ for $j = 1, 2, \ldots, n-1$.

(ii) *The sequence* $x_0, x_1, x_2, \dots, x_{s-1}$ *is a critical point of*

$$K(x_1, x_2, \dots, x_s) = S(x_{s-1}, x_0 + r) + \sum_{j=0}^{s-2} S(x_j, x_{j+1})$$

if and only if there exist $y_0, y_1, y_2, \dots, y_{s-1}$ *such* $\Psi^j(x_j, y_j) = (x_{j+1}, y_{j+1})$ *for* $j = 0, \dots, s-1$, *with* $x_s = x_0 + r$.

Proof We only prove (ii) because (i) can be proved by a verbatim argument. Let $(x_0, \dots, x_{s-1})$ be a critical point and set $x_s = x_0 + r$. We also set $y_j = -S_x(x_j, x_{j+1})$. The result follows because if $Y_j = S_X(x_j, x_{j+1})$, then

$$K_{x_j} = y_j - Y_{j-1}$$

for $j = 0, 1, 2, \dots, s-1$ and $\Psi(x_j, y_j) = (x_{j+1}, Y_j)$. $\square$

Example 9.2 (Billiard map revisited) Let Ψ be the Billiard map as in Example 8.6. We certainly have $\rho^- = 0$ and $\rho^+ = 1$. According to Theorem 8.7, Ψ has at least two periodic orbits of type (r, s) whenever r and s are relatively prime and $r < s$. Recall that the function K of Proposition 8.8 is defined for $(x_0, x_1, \dots, x_{s-1})$ provided that $x_{j+1} \in [x_j, x_j + 1]$ for $j = 0, 1, \dots, s-1$ with $x_s = x_0 + r$. This however does not reflect the ordering of the orbit. For our purposes we define K on a smaller set Λ which consists of $(x_0, x_1, \dots, x_{s-1})$ such that there exists $z_0 \le z_1 \le \dots \le z_{rs}$, with $z_{i+s} = z_i + 1$ for $i = 0, 1, \dots, (r-1)s$, and $x_j = z_{jr}$ for $j = 0, 1, \dots, s$. Note that once $(x_0, x_1, \dots, x_{s-1}) \in \Lambda$ is known, then all z_j can be determined. Of course $\mathbf{x} \in \Lambda$ imposes various inequalities between $x_0, x_1, \dots, x_{s-1}$. On the other hand, we can regard K as a function of $z_0, z_1, \dots, z_{rs}$. Also there are only s many independent variables among them, say $z_0, z_1, \dots, z_{s-1}$. So, we now have a function $\hat{K}(z_0, z_1, \dots, z_{s-1}) = K(x_0, x_1, \dots, x_{s-1})$. The advantage of $\hat{K}$ to K is that it has a domain which is much easier to describe, namely

$$\hat{\Lambda} = \{(z_0, z_1, \dots, z_{s-1}) : z_0 \le z_1 \le \dots \le z_{s-1} \le z_0 + 1\}.$$

Since $K(x_0 + 1, \dots, x_{s-1} + 1) = K(x_0, \dots, x_{s-1})$, we learn that $\hat{K}(z_0 + 1, \dots, {}_{s-1} + 1) = \hat{K}(z_0, \dots, z_{s-1})$. Introducing $w_j = z_j - z_{j-1}$, we have that $\mathbf{z} = (z_0, \dots, z_{s-1}) \in \hat{\Lambda}$ if and only if $(z_0, w_1, \dots, w_{s-1})$ belongs to the set of points with

$$0 \le w_1, w_2, \dots, w_{s-1}, \; w_1 + w_2 + \dots + w_{s-1} \le 1.$$

Writing $\bar{K}$ for $\hat{K}$ as a function of $z_0, w_1, \dots, w_{s-1}$, then $\bar{K}$ is defined on

$$\bar{\Lambda} = \{(z_0, w_1, \dots, w_{s-1}) : z_0 \in \mathbb{R}, w_1, \dots, w_{s-1} \ge 0, \sum_1^{s-1} w_j \le 1\}.$$

Since $\bar{K}(z_0+1, w_1, \ldots, w_{s-1}) = K(z_0, w_1, \ldots, w_{s-1})$, $\bar{K}$ is a lift of a function $\bar{k}$ which is defined on the set

$$\lambda = S^1 \times \{(w_1, \ldots, w_{s-1}) : w_1, \ldots, w_{s-1} \geq 0, \sum_1^{s-1} w_j \leq 1\} \subseteq S^1 \times [0,1]^{s-1}.$$

Of course $\bar{k}$ has a maximizer and a minimizer. We now argue that a minimizer yields a critical point which is in the interior of λ. To see this, let us assume that to the contrary the minimizer is a point on the boundary. To explain this in its simplest non-trivial case, let us take a boundary point of the form

$$z_{n-1} < z_n = z_{n+1} < z_{n+2}.$$

Recall that $\Psi(x, y(x, X)) = (X, Y(x, X))$ where $y(x, X) = -S_x(x, X)$ is increasing in X and $Y(x, X) = S_X(x, X)$ is decreasing in x. We now examine several cases:

(i) $$y(z_n, z_{n+r}) < Y(z_{n-r}, z_n)$$

In this case $\frac{\partial \hat{K}}{\partial z_n} = S_Y(z_{n-r}, z_n) + S_y(z_n, z_{n-r}) > 0$. Hence by decreasing z_n a little bit, we decrease $\hat{K}$. This contradicts the fact that $\mathbf{z}$ is a minimizer.

(ii) $$y(z_{n+1}, z_{n+r+1}) > Y(z_{n-r+1}, z_{n+1})$$

In this case $\frac{\partial \hat{K}}{\partial z_{n+1}} = S_Y(z_{n-r+1}, z_{n+1}) + S_y(z_n, z_{n-r}) < 0$. Hence by increasing z_{n+1}, the value $\hat{K}$ decreases, contradicting the fact that $\hat{z} = (z_0, \ldots, z_{s-1})$ is a minimizer.

(iii) If (i) and (ii) do not occur, then

$$\begin{aligned} Y(z_{n-r}, z_n) &\leq y(z_n, z_{n+r}) = y(z_{n+1}, z_{n+r}) \\ &\leq y(z_{n+1}, z_{n+r+1}) \leq Y(z_{n-r+1}, z_{n-1}) \\ &= Y(z_{n-r+1}, z_n) \leq Y(z_{n-r}, z_n). \end{aligned}$$

Hence $z_{n-r+1} = z_{n-r}$, $z_{n+r+1} = z_{n+r}$ and

$$\begin{aligned} Y(z_{n-r}, z_n) &= y(z_n, z_{n+r}), \\ Y(z_{n-r+1}, z_{n+1}) &= y(z_{n-1}, z_{n+r+1}) \end{aligned}$$

which means that (z_{n-r}, z_n, z_{n+1}) is the x-coordinate of an orbit. This is what we wanted.

In fact the other critical point is a saddle point in λ. To see this, let us examine the problem when $s = 2$ and $r = 1$ which corresponds to a periodic orbit of period 2. In this case, we simply have

$$K(x_1, x_2) = S(x_1, x_2) + S(x_2, x_1 + 1) = 2S(x_1, x_2)$$

which is defined on the set

$$\Lambda = \{(x_1, x_2) : x_1 \leq x_2 \leq x_1 + 1\}.$$

Note that $K(x_1, x_2) = 0$ if $(x_1, x_2) \in \partial\Lambda$ and we always have $K(x_1, x_2) < 0$ if $(x_1, x_2) \in \Lambda^0$. Writing K in terms of x_1 and $w_2 = x_2 - x_1$ yields $\hat{K}(x_1, w_2) = 2S(x_1, x_1 + w_2)$ which is defined for $(x_1, w_2) \in \mathbb{R} \times [0, 1]$. Since $\hat{K}$ is periodic in x_1, $\hat{K}$ is the lift of $\hat{k} : S^1 \times [0, 1] \to \mathbb{R}$ and evidently its minimum is attained in the interior of $S^1 \times [0, 1]$. Note that $-\min \hat{k}$ is simply the diameter of the convex set C. We now assert that $\hat{k}$ has a saddle critical point which corresponds to the width of C. To see this, for any $x_1 \in S^1$, we can find $\eta(x_1) = x_2 \in S^1$ such that the tangents at x_1 and $x_2 = \eta(x_1)$ are parallel. Now $-\max_{x_1} S(x_1, \eta(x_1)) = -S(x_1^*, \eta(x_1^*))$ yields the *width* of C. We assert that $(x_1^*, x_2^*) = (x_1^*, \eta(x_1^*))$ is the other critical point of K. Indeed, since $\Theta = \pi - \theta$, we have

$$S_x(x_1, \eta(x_1)) = S_X(x_1, \eta(x_1))$$

for all x_1. On the other hand, at a maximizer x_1^* of $S(x_1, \eta(x_1))$, we must have

$$\begin{aligned} 0 &= S_x(x_1^*, \eta(x_1^*)) + S_X(x_1^*, \eta(x_1^*))\eta'(x_1^*) \\ &= S_x(x_1^*, x_2^*)(1 + \eta'(x_1^*)). \end{aligned}$$

It is not hard to show that in fact $\eta'(x_1^*) > 0$. Hence we must have $S_x = S_X = 0$ at (x_1^*, x_2^*). $\square$

So far we have seen that for $\rho \in (\rho^-, \rho^+) \cap \mathbb{Q}$ we can find at least two periodic orbit of rotation number ρ. The variational principle can be used to find orbits corresponding to irrational $\rho \in (\rho^-, \rho^+)$. This is the subject of *Mather Theory.* For any irrational $\rho \in (\rho^-, \rho^+)$, there exists an invariant set on the cylinder which projects onto either a Cantor-like subset of S^1 or the whole S^1. The invariant set lies on a graph of a Lipschitz function defined on S^1. These invariant sets are known as *Aubry–Mather* sets and correspond to the irrational rotations of Exercise 9.1**(i)**.

We now turn to Theorem 9.2. So far we have a proof in the case of a monotone twist map. Following an idea of Chaperon, we try to express a twist map as a composition of monotone twist maps. To this end, let us define some function spaces.

(i) $\mathcal{T}$ denotes a space of homeomorphism ψ from cylinder $C = S^1 \times [0,1]$ onto itself which is an orientation and area preserving diffeomorphism in the interior of C and preserves the boundary circles $S^1 \times \{0\}$ and $S^1 \times \{1\}$. The rotation numbers of ψ restricted to $S^1 \times \{0\}$ and $S^1 \times \{1\}$ are denoted by $\rho_-(\psi)$ and $\rho_+(\psi)$ respectively.

(ii) $\mathcal{T}^*$ denotes the space of $\psi \in \mathcal{T}$ such that $\rho^-(\psi) \neq \rho^+(\psi)$. $\mathcal{T}^+$ denotes the space of $\psi \in \mathcal{T}$ with $\rho^-(\psi) < \rho^+(\psi)$. $\mathcal{T}^-$ denotes the space of $\psi \in \mathcal{T}$ with $\rho^+(\psi) < \rho^+(\psi)$. $\mathcal{M}^+$ denotes the space of monotone twist maps.

We equip the space of $\mathcal{T}$ with the topology of C^1-convergence in the interior of C and uniform-convergence up to the boundary. Evidently $\mathcal{T}$ is a topological group with multiplication given by composition. As an example, note that the shear map ξ with lift $(x,y) \mapsto (x+y,y)$ belongs to $\mathcal{T}^+$ whereas $\xi^{-1} = \lambda$ belongs to $\mathcal{T}^-$. We have the following straightforward lemma.

Lemma 9.1 *Every element ψ in the connected component of identity in $\mathcal{T}$ can be written as*

$$\psi = \lambda \circ \psi_1 \circ \lambda \circ \psi_2 \circ \cdots \circ \lambda \circ \psi_n \tag{9.11}$$

with $\psi_1, \psi_2, \ldots, \psi_n \in \mathcal{M}^+$.

Proof Evidently there exists an open set U in $\mathcal{T}$ such that $\xi \in U \subseteq \mathcal{M}^+$. As a result, $\mathrm{id} \in \xi^{-1}U = \lambda U =: V$ is an open neighborhood of identity and each $\psi \in V$ can be written as $\psi = \lambda \circ \psi_1$ with $\psi_1 \in \mathcal{M}^+$. We now write Ω for the set of ψ in $\mathcal{T}$ for which the decomposition (9.11) exists with $\psi_1, \psi_2, \ldots, \psi_n \in V$. Clearly Ω is open because V is open. If we can show that V is also closed, then we deduce that Ω is the connected component of id in $\mathcal{T}$. To see the closedness of Ω, let $\{\varphi_m\}$ be a convergent sequence in Ω. If $\lim_{m\to\infty} \varphi_m = \varphi$, then $\lim_{m\to\infty} \varphi \circ \varphi_m^{-1} = \mathrm{id}$ and, as a result, $\varphi \circ \varphi_m^{-1} \in V$, for large m. Hence there exists $\bar\psi \in \mathcal{M}^+$, such that $\varphi \circ \varphi_m^{-1} = \lambda \circ \bar\psi$ for a sufficiently large m. That is, $\varphi = \lambda \circ \bar\psi \circ \varphi_m$. Since $\varphi_m \in \Omega$, we deduce that $\varphi \in \Omega$, completing the proof of closedness of Ω. □

Proof of Theorem 9.2 Let Ψ be a twist map. On account of Lemma 9.1, there exist monotone twist maps $\Psi_1, \Psi_2, \ldots, \Psi_n$ such that

$$\Psi = \Lambda \circ \Psi_1 \circ \Lambda \circ \Psi_2 \circ \cdots \circ \Lambda \circ \Psi_n.$$

Each Ψ_j has a generating function $S_j(x,X)$ with $S_j : \Gamma_j \to \mathbb{R}$ with

$$\Gamma_j = \{(x,X) : A_-^j(x) \le X \le A_+^j(x)\}.$$

Note that the generating function for Λ is $T(x,X) = -\frac{1}{2}(X-x)^2$ which is defined on the set

$$\{(x,X) : x-1 \le X \le x\}.$$

We now define

$$L(x_0, x_1, \ldots, x_{2n-1}) = -\sum_{j=0}^{n-1} \frac{1}{2}(x_{ij} - x_{ij+1})^2 + \sum_{j=0}^{n-2} S_j(x_{2j+1}, x_{2j+2}) + S_n(x_{2n-1}, x_0)$$

on the set Γ which consists of points $x_0, x_1, \ldots, x_{2n-1}$ such that for $j = 0, 1, \ldots, n-2$

$$-1 \le x_{2j} - x_{2j+1} \le 0, \;\; A_-^j(x_{2j+1}) \le x_{2j+2} \le A_+^j(x_{2j+1})$$

with $x_{2j} = x_0$. Now as in Proposition 8.8, we can show that if $x_0, x_1, \ldots, x_{2n-1}$ is a critical point of L, then

$$\Lambda(x_{2j}, y_{2j}) = (x_{2j+1}, y_{2j+1}), \;\; \Psi_j(x_{2j+1}, y_{2j+1}) = (x_{2j+2}, y_{2j+2})$$

for $j = 0, 1, \ldots, n-1$, where $y_{2j+1} = T_X(x_{2j}, x_{2j+1}) = x_{2j} - x_{2j+1}$, $y_{2j+2} = S_X(x_{2j+1}, x_{2j=2})$, $x_{2j} = x_0$ and $y_{2j} = y_0$. Of course, in particular (x_0, y_0) is a fixed point of Ψ. $\square$

So far we have used generating functions to study various properties of twist maps on cylinders. We now discuss possible generalizations of such global properties for other manifolds. As a start, let us take a symplectic $\psi : \mathbb{T}^{2n} \to \mathbb{T}^{2n}$ and wonder how many fixed points it can have. Evidently a rotation or translation on $\mathbb{T}^2$ is a symplectic diffeomorphism with no fixed point. The question is what plays the role of the twist condition to guarantee the existence of fixed points.

We first examine the issue of generating function. In section 8, we showed that a generating function $S(q, Q)$ always exists locally provided that $\frac{\partial Q}{\partial p}(q^0, p^0)$ is non-singular. Note that this condition fails for the identity. We now discuss another type of generating function that exists trivially for identity. Again for $\Psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ with $\Psi(q, p) = (Q, P)$ and $P \cdot dQ - p \cdot dq = dS$, we write

$$P \cdot dQ + q \cdot dp = d(p \cdot q) + dS =: d\hat{S},$$

which suggests a generating function $\hat{S}(Q, p)$. The following proposition can be proved as Proposition 9.1.

Proposition 9.6 *Let $\Psi(q, p) = (Q, P)$ be a symplectic diffeomorphism with*

$$\det \frac{\partial Q}{\partial q}(q^0, p^0) \neq 0. \tag{9.12}$$

Then there exist a neighborhood V of $Q^0 = Q(q^0, p^0)$ and p^0, and a C^1-function $\hat{S} : V \to \mathbb{R}$ such that

$$\frac{\partial \hat{S}}{\partial Q} = P, \quad \frac{\partial \hat{S}}{\partial p} = q. \tag{9.13}$$

We note that identity transformation has such a generating function with $\hat{S}(Q,p) = Q\cdot p$. More generally, we may write $\hat{S}(Q,p) = Q \cdot p - V(Q,p)$ with V now satisfying

$$P - p = -V_Q, \quad Q - q = V_p, \tag{9.14}$$

which can be thought of as a discrete analog of a Hamiltonian where $V(Q,p)$ plays the role of the Hamiltonian.

Using the generating function V, it is not hard to come up with a compelling conjecture regarding the fixed points of a symplectic $\psi : \mathbb{T}^{2n} \to \mathbb{T}^{2n}$. Let us write $\Psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ for the lift of ψ. If we assume that Ψ has a globally defined generating function V, then a point (q,p) is a fixed point of Ψ if and only if the corresponding (Q,p) is a critical point of V. This should be compared to our proof of Poincaré–Birkhoff Theorem in the case of a monotone twist map. To have an analogous result, we need to say something about the critical points of a C^1-function of $\mathbb{T}^{2n}$. In this connection we have the following:

Theorem 9.3 *(Ljusternik–Schnirelman) Let M be a compact manifold. Then any C^1-function $V : M \to \mathbb{R}$ has at least $\mathcal{I}(M)$ many critical points where $\mathcal{I}(M)$ denotes the* cup length *of M.*

Here is the definition of cup length: $\mathcal{I}(M)$ is the smallest number l such that there exist open simply connected sets $U_1, \dots, U_l$ such that $M = U_1 \cup \dots \cup U_l$. Alternatively, for closed forms $\alpha_1, \alpha_2, \dots, \alpha_l$, we have that $\alpha_1 \wedge \alpha_2 \wedge \dots \wedge \alpha_l$ is exact. Here each α_j is a k_j-form with $k_j \geq 1$ for $j = 1, \dots, l$.

Example 9.3 $\mathcal{I}(\mathbb{T}^k) = k + 1$.

From Theorem 9.3 and Proposition 9.6 we deduce

Proposition 9.7 *Let $\psi : \mathbb{T}^{2n} \to \mathbb{T}^{2n}$ be a symplectic diffeomorphism which is a small perturbation of identity. Then ψ has at least $2n + 1$ many fixed points.*

We are now in a position to formulate a similar result for more general symplectic diffeomorphism. With our experience from the previous section we search for a condition on ψ that guarantees to a representation $\psi = \psi_1 \circ \psi_2 \circ \dots \circ \psi_N$ where each ψ_j is a small perturbation of identity. To this end, let us assume that $\Psi = \phi_{01}$ where ϕ_{st} is the flow of a Hamiltonian ODE

$$\dot{x} = H(x,t)$$

where H is 1-periodic in x and t, that is

$$H(x + n, t) = H(x, t + 1) = H(x, t)$$

for every $n \in \mathbb{Z}^{2d}$. Such a function Ψ is called a *Hamiltonian symplectomorphism.* Evidently Ψ is a lift of a symplectomorphism on the torus $\mathbb{T}^{2n}$.

Theorem 9.4 *(Conley-Zehnder) Any symplectomorphism $\psi : \mathbb{T}^{2n} \to \mathbb{T}^{2n}$ has at least $2n+1$ many fixed points.*

Note that $\psi = \phi_{0,1} = \phi_{0,\frac{1}{N}} \circ \phi_{\frac{1}{N},\frac{2}{N}} \circ \cdots \circ \phi_{\frac{N-1}{N},1}$ and if N is sufficiently large, then each $\psi_j = \phi_{\frac{j-1}{N},\frac{j}{N}}$ has a generating function V_j.

Exercise 9.1(i) Consider the Billiard map in a circle. Determine the generating function. Find periodic orbits and describe the remaining orbits.

(ii) Consider the Billiard map in an ellipse. Show that the 2-periodic orbits correspond to the reflection along the axes of symmetry. Show that the 2-periodic orbit associated with the shorter axis of symmetry is a saddle point of the generating function. **Hint:** (x_1^*, x_2^*) maximizers of $S(x_1, \eta(x_1))$ but is a local minimum for $S(x_1^*, x_2)$. $\square$

10 Pseudo-Holomorphic Curves and Gromov Width

In Chapters 5-8, we use the invariance (3.2) of the action $A(\gamma)$ to construct various symplectic capacities. For our constructions, we use the periodic orbits of the Hamiltonian systems in an essential role. In this chapter, we focus on the invariance if $\int_\Gamma \omega$ to analyze symplectic maps. Our main tool will be the *pseudo-holomorphic curves* of Gromov. We give two motivations for the relevance of such complex curves (or rather real surfaces):

1. Observe that if $\Gamma = w(D)$ is a 2-dimensional surface with the parametrization $w : D \to \mathbb{R}^{2n}$ for a planar domain D, then

$$\int_\Gamma \bar\omega = \int_D \bar\omega(w_s, w_t)\ dsdt = \int_D \bar J w_s \cdot w_t\ dsdt.$$

This integral is simply $-area(\Gamma)$ when $n = 1$. For $n > 1$, the integral $\int_\Gamma \omega$ is not of a definite sign and does not represent any kind of *size* of the surface Γ. However, if w satisfies

$$w_s = \bar J w_t, \tag{10.1}$$

then $w_s \cdot w_s = 0$, and

$$\int_\Gamma (-\bar\omega) = \int_D |w_s|^2\ dsdt = \int_D \left(|w_s|^2 |w_t|^2 - (w_s \cdot w_s)^2\right)^{1/2}\ dsdt = area(\Gamma). \tag{10.2}$$

This means that for such surfaces, $-\int_\Gamma \bar\omega$ is indeed the area of Γ. If we write $w = (u, v)$ with u and v representing the position and momentum, then (10.1) reads as

$$u_s = v_t, \qquad v_s = -u_t,$$

which are nothing other than Cauchy-Riemann equations. When (10.1) is satisfied, we say that w is a *holomorphic curve.* Holomorphic curves are not preserved under a symplectic change of coordinates. However, if $\varphi \circ \hat w = w$, then

$$\hat w_s = J(\hat w)\hat w_t, \tag{10.3}$$

where

$$J = (\varphi')^{-1} \bar J \varphi',$$

is an example of an *almost complex structure* and $\hat w$ is called a J-holomorphic curve. Observe that if we set $\bar g(a, b) = a \cdot b$, and define a metric $g = \varphi^* \bar g$, then

$$\bar\omega(a, b) = \bar\omega(\varphi'(x)a, \varphi'(x)b) = \bar g(\bar J \varphi'(x)a, \varphi'(x)b) = (\varphi^* \bar g)_x(J(x)a, b). \tag{10.4}$$

We can now repeat our calculation in (10.2) to assert that if $\hat\Gamma = \hat w(D)$, then

$$\int_\Gamma (-\bar\omega) = area_g(\hat\Gamma),$$

where $area_g$ denotes the area with respect to the metric g.

2. Recall that if $H(x,t)$ is a time dependent Hamiltonian function, then the critical points of the functional

$$\mathcal{A}(x) = \int_0^1 \left[p \cdot \dot{q} - H(x,t)\right] \, dt,$$

corresponds to the 1-periodic orbits of $\dot{x} = \bar{J}\nabla H(x,t)$. For example if $H : \mathbb{T}^{2n} \times \mathbb{R} \to \mathbb{R}$ is 1-periodic in t, and if we regard $\mathcal{A} : C^1(S^1, \mathbb{R}^{2n}) \to \mathbb{R}$, then the gradient of $\mathcal{A}$ with respect to the L^2-inner product is given by

$$\partial\mathcal{A}(x(\cdot)) = -\bar{J}\dot{x}(\cdot) - \nabla H(x(\cdot), t).$$

Observe that the gradient with respect to $\mathcal{H}^{1/2}$-inner product, namely $\nabla\mathcal{A}$ is related to $\partial\mathcal{A}$ by the formula $\nabla\mathcal{A} = \mathcal{I}\partial\mathcal{A}$. In fact the critical points of $\mathcal{A}$ corresponds to the fixed points of the map $\psi = \phi_1^H$ and following an idea of Floer, we may study such critical points by developing a Morse-type theory of $\mathcal{A}$. Morse theory may be developed by studying the gradient flow

$$\frac{dw}{ds} = -\partial\mathcal{A}(w). \tag{10.5}$$

Regarding $w : \mathbb{R} \times S^1 \to \mathbb{R}^{2n}$ as a function of two variables s and t, (10.5) reads as

$$w_s = \bar{J}w_t + \nabla H(w,t). \tag{10.6}$$

This is very different from the corresponding

$$\frac{dw}{ds} = -\nabla\mathcal{A}(w) = -\mathcal{I}\partial\mathcal{A}(w); \tag{10.7}$$

the right-hand side of (10.7) is an intergro-differential equation and is well-defined as an ODE, whereas the equation (10.6) is an elliptic PDE and not well-posed as an initial-value problem. Evidently, (10.6) is the same as (10.1) when $H = 0$. The elliptic PDE (10.1) is well-posed for a prescribed $w(\partial D)$. □

Motivated by (10.4), let us give a general definition for almost complex structures. Let M be a C^1 manifold. By an *almost complex structure* on M, we mean a continuous $x \mapsto J_x$ with $J_x : T_xM \to T_xM$ linear function satisfying $J_x^2 = -id$. The pair (M, J) is called an *almost complex manifold.* If (M, ω) is symplectic, then we say (J, g) and ω are compatible if

$$g_x(a,b) = \omega_x(a, J_x b), \ a, b \in T_xM \tag{10.8}$$

is a Riemannian metric on M. We write $\mathcal{I}(M,\omega)$ for the compatible pairs (g, J). We also set

$$\mathcal{G}(M,\omega) = \{g : \ (g,J) \in \mathcal{I}(M,\omega)\}, \quad \mathcal{J}(M,\omega) = \{J : \ (g,J) \in \mathcal{I}(M,\omega)\}.$$

Proposition 10.1 **(i)** *Let (M,ω) be symplectic with a Riemannian metric $\hat{g}$. Then $\mathcal{I}(M,\omega)$ is nonempty.*

(ii) *If $g \in \mathcal{G}(M,\omega)$ and $\dim M = 2n$, then $\mu := (n!)^{-1}\omega^n$ is the Riemannian volume form associated with the Riemannian metric g.*

Proof Fix x. Both $a \mapsto \omega_x(a,\cdot)$ and $a \mapsto \hat{g}_x(a,\cdot)$ are linear isomorphisms between T_xM and $(T_xM)^*$. Hence there exists a linear invertible $A_x : T_xM \to T_xM$ such that $\omega_x(a,b) = \hat{g}_x(A_x a, b)$. Note that A_x is skew symmetric because

$$\hat{g}(A^*a,b) = \hat{g}(a,Ab) = \hat{g}(Ab,a) = \omega(b,a) = -\omega(a,b) = -\hat{g}(Aa,b).$$

In fact a candidate for J is simply the orthogonal matrix that appears in the polar decomposition of A. More precisely, if we use the unique representation $A_x = S_x J_x$, with $S_x = (A_x A_x^t)^{1/2} = (-A_x^2)^{1/2}$, and J_x orthogonal with respect to $\hat{g}_x$, then $J \in \mathcal{J}(M,\omega)$. To verify this, we build this polar decomposition directly. The idea is that since A is real and skew-symmetric, its eigenvalues appear as $\pm i\lambda$ with λ real. Hence the eigenvalues of $-A^{@} = 2$ are positive and of even multiplicities. In fact, the matrix $B = A^2$ is negative definite, and if v is an eigenvector of B associated with the eigenvalue $-\lambda^2$ for some positive λ, then $w = \lambda^{-1}Av$ is another eigenvector for B that is orthogonal to v:

$$\hat{g}(v,w) = \lambda^{-1}\hat{g}(v,Av) = \lambda^{-1}\omega(v,v) = 0, \qquad Bw = \lambda^{-1}ABv = \lambda Av = \lambda^2 w.$$

We now describe S and J on $\{v,w\}$, and hence on the span of $\{v,w\}$:

$$\begin{aligned} Av &= \lambda w, & Sv &= \lambda v, & Jv &= w, \\ Aw &= -\lambda v, & Sw &= \lambda w, & Jw &= -v. \end{aligned}$$

More generally, we can find an orthonormal basis $\{v_1,\dots,v_{2n}\}$ and positive numbers $\lambda_1,\dots,\lambda_{2n}$ such that $Bv_i = -\lambda_i^2 v_i$ for each i. We can readily show that $\{w_i = \lambda_i^{-1}Av_i : i = 1,\dots,2n\}$ is also an orthonormal set and that $Aw_i = -\lambda_i v_i$, $Bw_i = \lambda_i^2 w_i$. But since A is skew symmetric, w_i and v_i are orthogonal eigenvectors of B associated with the same eigenvalue $-\lambda_i^2$. We relabel our eigenvalues and eigenvectors so that $Av_{2i} = \lambda_i v_{2i-1}$ and $Av_{2i-1} = -\lambda_i v_{2i}$ for $i = 1,\dots n$. We then define S and J by $Sv_{2i} = \lambda_i v_{2i}$, $Sv_{2i-1} = \lambda_i v_{2i-1}$, $Jv_{2i} = v_{2i-1}$ and $Jv_{2i-1} = -v_{2i}$. It is straightforward to check that $A = SJ$, $J^2 = -id$ and $J^* = -J$. We also have

$$g(a,b) := \omega(a,Jb) = g(Aa,Jb) = g(Sa,b).$$

We are done because $S > 0$.

(ii) First we claim that there exists a local orthonormal frame for $(M,\hat{g})$ of the form $E_n = \{e_1, Je_1, \dots, e_n, Je_n\}$. This frame is construction inductively; if we already have $2k$ many vector fields $E_k = \{e_1, Je_1, \dots, e_n, Je_k\}$ with $\hat{g}(e_i, e_j) = \delta_{ij}$, $\hat{g}(e_i, Je_j) = 0$ for $i, j = 1, \dots k$, then locally we can find a new vector field e_{k+1} which is orthogonal to vectors in E_k and $\hat{g}(e_{k+1}, e_{k+1}) = 1$. We then use (3.14) and $J^2 = -id$ to deduce that Je_{k+1} is also orthogonal to E_k and e_{k+1} and that $\hat{g}(Je_{k+1}, Je_{k+1}) = 1$.

From $\omega(e_i, e_j) = \omega(Je_i, Je_j) = 0$ and $\omega(e_i, Je_j) = \delta_{ij}$, we can readily deduce

$$\omega^n(e_1, Je_1, \dots, e_n, Je_n) = n!,$$

and this implies that μ is the Riemannian volume. $\square$

In the prominent work [G], Gromov uses the J-Holomorphic curves to establish his non-squeezing result, namely Theorem 6.2**(i)**. The main ingredient for his proof is an existence result:

Theorem 10.1 *(Gromov) For every $J \in \mathcal{J}(\mathbb{R}^{2n}, \bar{\omega})$ and every $x^0 \in Z^{2n}(1)$, there exists a J-holomorphic $\hat{w} : \mathbb{D} \to Z^{2n}(1)$ such that $x^0 \in \hat{w}(\mathbb{D})$, $w(\partial\mathbb{D}) \subset \partial Z^{2n}(1)$, and*

$$\int_{\hat{w}(\mathbb{D})} (-\bar{\omega}) = \pi. \tag{10.9}$$

Let us first see how Theorem 10.1 implies the non-squeezing property of symplectomorphisms.

Proof of Theorem 6.2(i) Let $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a symplectomorphism such that $\varphi(B_R) \subset Z^{2n}(1)$, where $B_R = \{x : \ |x| \le R\}$. Let J be as in (10.3). By Theorem 10.1, we can find $\hat{w} : \mathbb{D} \to Z^{2n}(1)$ such that $\varphi(0) \in \hat{w}(\mathbb{D})$, and (10.9) is true. Let $w = \varphi^{-1} \circ \hat{w}$, so that w solves (10.1), $0 \in w(\mathbb{D})$, and

$$Area(w(\mathbb{D})) = \int_{w(\mathbb{D})} (-\bar{\omega}) = \int_{\hat{w}(\mathbb{D})} (-\bar{\omega}) = \pi,$$

by (10.9). On the other-hand by a classical theorem of Lelong, we must have

$$\pi = Area(w(\mathbb{D})) \ge Area(B_R) = \pi R^2.$$

Thus, $R \le 1$. $\square$

Before embarking on the proof of Theorem 10.1, let us develop some feel for the equation (10.3). To ease the notation, let us write w for $\hat{w}$. Also, to avoid the occurring minus sign, we take $J \in \mathcal{J}(\mathbb{R}^{2n}, -\bar{\omega})$, so that instead of (10.3), we now have

$$w_t = J(w) w_s. \tag{10.10}$$

Note that now J as in (10.3) satisfies

$$J = (\varphi')^{-1} i \varphi', \tag{10.11}$$

where $i = -\bar{J}$ is simply the multiplication by i, when we use the identification $\mathbb{R}^{2n} = \mathbb{C}^n$. We now give two interpretations for the equation (10.10).

1. In our first interpretation, we compare (10.10) with its equivalent formulation that we obtain by multiplying both sides of (10.1) by $J(w)$:

$$w_t = J(w) w_s \qquad - w_s = J(w) w_t. \tag{10.12}$$

Regarding $dw : T\mathbb{D} \to T\mathbb{R}^{2n}$, $i : T\mathbb{D} \to \mathbb{D}$, and $J : T\mathbb{R}^{2n} \to T\mathbb{R}^{2n}$, we may rewrite (10.12) as

$$dw \circ i = J \circ dw; \tag{10.13}$$

if we evaluate both sides of (10.13) at the vector fields $\frac{\partial}{\partial s}$ and $\frac{\partial}{\partial t}$, we obtain the equations of (10.12), because

$$i\left(\frac{\partial}{\partial s}\right) = \frac{\partial}{\partial t}, \qquad i\left(\frac{\partial}{\partial t}\right) = -\frac{\partial}{\partial s}.$$

The reader may compare (10.13) with (10.11) that may be written as

$$i \circ d\varphi = d\varphi \circ J. \tag{10.14}$$

The advantage of the formulation (10.13) is that it has an obvious formulation to arbitrary manifolds.

Definition 10.1 Let M be a manifold and $J : TM \to TM$ be an *almost complex structure* i.e. $J^2 = -id$.

(i) We say that $w : \mathbb{D} \to M$ is a J-holomorphic curve if (10.3) is valid. We also define the operator

$$\bar{\partial}_J w = \frac{1}{2}\big(dw + J \circ dw \circ i\big), \tag{10.15}$$

so that the equation (10.13) may be written as $\bar{\partial}_J w = 0$.

(ii) We say that J is a *complex structure* if for every $x \in M$, we can find $U \subset \mathbb{C}^n$ and a diffeomorphism $\varphi : U \to M$ such that $x \in \varphi(U)$, and (10.14) is valid. □

2. Let us use complex-variable notation to write $z = s + it$ and $\bar{z} = s - it$. We also use the notations

$$\begin{aligned} dz &= ds + idt, & d\bar{z} &= ds - idt, \\ \frac{\partial}{\partial z} &= \frac{1}{2}\left(\frac{\partial}{\partial s} - i\frac{\partial}{\partial t}\right), & \frac{\partial}{\partial \bar{z}} &= \frac{1}{2}\left(\frac{\partial}{\partial s} + i\frac{\partial}{\partial t}\right). \end{aligned}$$

By definition,

$$w_t = i(w_z - w_{\bar{z}}), \qquad w_s = (w_z + w_{\bar{z}}).$$

Hence, if $i + J$ is invertible, the (10.10) may be written as

$$w_{\bar{z}} = K(w)w_z. \tag{10.16}$$

where

$$K(w) := (i + J(w))^{-1}(i - J(w)).$$

Proposition 10.2 *If $J \in \mathcal{J}(\mathbb{R}^{2n}, -\bar{\omega})$, then $i + J$ is invertible and $\|K\| < 1$.*

Proof First observe that since $J \in \mathcal{J}(\mathbb{R}^{2n}, -\bar{\omega})$, we know

$$a \neq 0 \ \Rightarrow (ia) \cdot (Ja) = (-\bar{\omega})(a, Ja) > 0. \tag{10.17}$$

To show that $i + J$ is invertible, note that if to the contrary we can find $a \neq 0$, such that $(i + J)a = 0$, then

$$(ia) \cdot (Ja) = -|ia|^2 < 0,$$

which contradicts (10.17).

To show that $\|K\| < 1$, it suffices to check that $|b| < |a|$ whenever $K(a) = b$ and $a \neq 0$. We have

$$K(w)a = b \ \Leftrightarrow \ (i - J)(a) = (i + J)(b) \ \Leftrightarrow \ b - a = iJ(b + a).$$

Clearly, if $a + b = 0$, then $a = b = 0$. Now if $a, a + b \neq 0$, then by (10.17),

$$|b|^2 - |a|^2 = (b - a) \cdot (b + a) = (iJ)(b + a) \cdot (b + a) = -\big[J(b + a) \cdot i(b + a)\big] < 0,$$

as desired. $\square$

Remark 10.1 The equation (10.13), is equivalent to the classical *Beltrami Equation* when $n = 1$. This equation is related to the theory of *quasi-conformal maps* and a generalization of the *Riemann mapping theorem.* To explain this, first observe that (10.13) and (10.14) are equivalent in the case of $n = 1$ by setting $\varphi = w^{-1}$. As we showed in Example 2.2, the set $\mathcal{I}(\mathbb{R}^2, -\bar{\omega})$ consists of pairs (g, J), with $g(a, b) = Ga \cdot b$, such that

$$G = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}, \qquad J = \begin{bmatrix} -\beta & -\gamma \\ \alpha & \beta \end{bmatrix}, \tag{10.18}$$

with $\alpha, \gamma > 0$ and $\alpha\gamma - \beta^2 = 1$. Now let us write $\varphi(z) = u + iv$ with $z = s + it$, and evaluate both sides of (10.14) at the vector fields $\frac{\partial}{\partial s}$ and $\frac{\partial}{\partial t}$, with J as in (10.18). We obtain

$$i\varphi_s = -\beta\varphi_s + \alpha\varphi_t, \qquad i\varphi_t = -\gamma\varphi_s + \beta\varphi_t.$$

This is equivalent to $\varphi_t = j\varphi_s$ for $j = (\beta + i)/\alpha = \gamma/(\beta - i)$. Equivalently,

$$\varphi_{\bar{z}} = \mu(z)\varphi_z, \qquad \text{with } \mu = \frac{i-j}{i+j}. \tag{10.19}$$

This equation is known as the *Beltrami Equation*, and resembles (10.16) because

$$|\mu| = \left|\frac{\beta + (\alpha - 1)i}{\beta + (\alpha + 1)i}\right| < 1,$$

by our assumption $\alpha > 0$. We refer to Appendix D for more information about the Beltami Equation.

As for the equation (10.16) in dimension 2, observe that when $n = 1$,

$$\begin{aligned} K &= \begin{bmatrix} -\beta & -\gamma - 1 \\ \alpha + 1 & \beta \end{bmatrix}^{-1} \begin{bmatrix} \beta & \gamma - 1 \\ 1 - \alpha & -\beta \end{bmatrix} \\ &= (\alpha + \gamma + 2)^{-1} \begin{bmatrix} \beta & \gamma + 1 \\ -\alpha - 1 & -\beta \end{bmatrix} \begin{bmatrix} \beta & \gamma - 1 \\ 1 - \alpha & -\beta \end{bmatrix} \\ &= (\alpha + \gamma + 2)^{-1} \begin{bmatrix} \gamma - \alpha & -2\beta \\ -2\beta & \alpha - \gamma \end{bmatrix}. \end{aligned}$$

From this, we deduce that in fact the equation (10.16) can be written as

$$w_{\bar{z}} = m(w)\bar{w}_z, \tag{10.20}$$

where $m = (\gamma - \alpha - 2\beta i)/(\alpha + \gamma + 2)$. We note

$$|m|^2 = \frac{(\alpha + \gamma)^2 - 4}{(\alpha + \gamma + 2)^2} = \frac{\alpha + \gamma - 2}{\alpha + \gamma + 2} = 1 - 4(\alpha + \gamma + 2)^{-1} < 1.$$

□

As we discussed in Remark 10.1, the equation (10.16) is a multi-dimensional generalization of the classical Beltrami-equation. Ignoring the boundary condition requirements of Theorem 10.2 we can use the classical transforms of *Cauchy and Beurling* to construct solutions of equation (10.16). We define

$$\mathcal{C}(h)(z) = -\frac{1}{\pi}\int_{\mathbb{D}} \frac{h(\zeta)}{\zeta - z}\, dsdt, \qquad \mathcal{B}(h)(z) = \frac{1}{\pi} PV \int_{\mathbb{D}} \frac{h(\zeta)}{(\zeta - z)^2}\, dsdt, \tag{10.21}$$

where PV stands for the principle value (see Appendix D). The main property of the Cauchy Transform $\mathcal{C}$ is that $\mathcal{C}(h)_{\bar{z}} = h$ and that its z derivative is the Buerling transform; $\mathcal{C}(h)_z = \mathcal{B}(h)$. Moreover, by *Caldron-Zygmund Theory*, the operator $\mathcal{B}$ is bounded on L^p for every $p \in (1, \infty)$. In fact $\mathcal{B}$ is an isometry on L^2 and its norm $C(p)$ on the L^p converges to 1 as $p \to 2$. We refer to Theorem D.1 of Appendix D for more details.

Theorem 10.2 *Assume that* $\|K\|_{L^\infty} = c_0 < 1$*. Let* $q_1, \dots, q_n : \mathbb{D} \to \mathbb{C}$ *be* n *holomorphic functions. Then there exists* $p = p(c_0) > 2$ *and a function* $w = (U_1, \dots, U_n) \in W^{1,p}$*,* $U_i : \mathbb{D} \to \mathbb{C}$*, of the form*

$$U_i = \mathcal{C}(u_i) + q_i, \qquad u_i \in L^p(\mathbb{D}),$$

that solves the equation (10.16) *weakly.*

Proof We use column vectors to write

$$w = \begin{bmatrix} U_1 \\ \vdots \\ U_n \end{bmatrix}, \qquad u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \qquad q = \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix}, \qquad \hat{q}(w) = K(w) \begin{bmatrix} q_1' \\ \vdots \\ q_n' \end{bmatrix}, \qquad \hat{\mathcal{B}}(u) = \begin{bmatrix} \mathcal{B}(u_1) \\ \vdots \\ \mathcal{B}(u_n) \end{bmatrix}.$$

Using the definition of $\mathcal{C}$ and $\mathcal{B}$, we can write (10.17) as

$$w_{\bar{z}} = u = K(w)\hat{\mathcal{B}}(u) + \hat{q}(w). \tag{10.22}$$

We wish to invert the operator $I - K(w)\hat{\mathcal{B}}$ for a given vector w. Since $\mathcal{B}$ is an isometry on L^2, the operator of $\hat{\mathcal{B}}$ is also an isometry. Using Theorem D.1 of the Appendix, we can show that for every $p \in (1, \infty)$, there exists a constant $c(p)$ such that $\lim c(p) = 1$ as $p \to 2$, and

$$\|\hat{\mathcal{B}}(u)\|_{L^p} \leq c(p)\|u\|_{L^p}.$$

From this we learn

$$\|A(w)\hat{\mathcal{B}}(u)\|_{L^p} \leq c(p)c_0\|u\|_{L^p}.$$

Choose $p > 2$ such that $c(p)c_0 < 1$. For such p,

$$\left\| \big(I - A(w)\hat{\mathcal{B}}\big)^{-1} h \right\|_{L^p} \leq \big(1 - c(p)c_0\big)^{-1} \|h\|_{L^p}.$$

This in turn implies that if

$$\mathcal{D}(w) = \big(I - A(w)\hat{\mathcal{B}}\big)^{-1}(\hat{q}(w)),$$

then

$$\|\mathcal{D}(w)\|_{L^p} \leq c_0\big(1 - c(p)c_0\big)^{-1}\|\hat{q}\|_{L^p} := c_1(p).$$

By applying Theorem D.1 of Appendix

$$\|\mathcal{C} \circ \mathcal{D}(w)\|_{W^{1,p}} \leq c_2,$$

for some constant $c_2 = c_2(p)$. Then rewrite (10.16) as

$$w = \big(\mathcal{C} \circ \mathcal{D}\big)(w) + q := \mathcal{E}(w). \tag{10.23}$$

Hence w is a fixed point of the operator $\mathcal{E}$. Set

$$\Gamma_L = \big\{w : \ \|w\|_{L^\infty} \le L\big\}.$$

We wish to show that $\mathcal{E} : \Gamma_L \to \Gamma_L$ and it has a fixed point. As a preparation, we first bound the nonlinear $\mathcal{D}$. Since $p > 2$, we may apply Morrey Inequality to deduce

$$\|\mathcal{C} \circ \mathcal{D}(w)\|_{C^{1-2/p}} \le c_2(p)\|\mathcal{C} \circ \mathcal{D}(w)\|_{W^{1,p}} \le C(p)c_1(p) = c_3(p), \tag{10.24}$$

where C^α denotes the space of α-Hölder continuous functions. In particular

$$\|\mathcal{C} \circ \mathcal{D}(w)\|_{L^\infty} \le c_3(p).$$

Setting $L = c_3(p) + \|q\|_{L^\infty}$, we deduce that $\mathcal{E}(L^p) \subseteq \Gamma_L$. In particular, $\mathcal{E}$ maps Γ_L into itself. On the other hand, the bound (D.12) implies that the image of Γ_L under $\mathcal{E}$ is in fact relatively compact. This allows us to use the *Schauder Fixed Point Theorem* to deduce that $\mathcal{E}$ has a fixed point. □

In this chapter, we will describe two approaches for establishing Theorem 10.1. In our first approach, we will follow a work of Sukhov and Tumanov [ST] that treats (10.16) as a generalization of the Beltrami equation and use Cauchy and Beurling transforms to construct solutions. (We refer to Appendix D for a thorough discussion of these transforms and their use in solving Beltrami-type equations.) The main ingredient of the proof of Theorem 10.16 a la [SK] is a variant of the Cauchy operator that is designed to solve the d-bar problem with a boundary condition. More precisely, given a holomorphic function $Q : \mathbb{D} \to \mathbb{C}$ that is nonzero inside $\mathbb{D}$, we define

$$\mathcal{C}^Q(f)(z) = \frac{Q(z)}{2\pi i} \iint_{\mathbb{D}} \left[\frac{f(\zeta)}{Q(\zeta)(\zeta - z)} + \frac{\overline{f(\zeta)}}{\overline{Q(\zeta)}(z\bar\zeta - 1)}\right] \, d\zeta \wedge d\bar\zeta. \tag{10.25}$$

We note

$$\mathcal{C}^Q(f) = Q\mathcal{C}(f/Q) + h^Q, \tag{10.26}$$

for a function h^Q that is holomorphic inside $\mathbb{D}$. The type of Q we have in mind are

$$Q(z) = a_0 \prod_{j=1}^{l} (z - a_j)^{\alpha_i}, \tag{10.27}$$

with $a_1, \dots, a_l \in \partial\mathbb{D}$ distinct and $\alpha_1, \dots, \alpha_l \in (0, 1]$. We can take a branch of the holomorphic Q that is defined on

$$\Omega_Q = \mathbb{C} \setminus \cup_{\alpha_j<1}\{ra_j : \ r > 0\}.$$

The following theorem of Monakhov guarantees that the operator $\mathcal{C}^Q$ and $\mathcal{B}^Q(f) = \big(\mathcal{C}^Q(f)\big)_z$ satisfy many properties of the Cauchy and Beurling Transforms.

Theorem 10.3 *Let Q be as in* (10.26). *For every $p \in (p_1, p_2)$ with*

$$p_1 = \max_j (1 - \alpha_j/2)^{-1}, \qquad p_2 = 2 \max_j (1 - \alpha_j)^{-1},$$

the operators $\mathcal{C}^Q : L^p(\mathbb{D}) \to W^{1,p}(\mathbb{D})$, and $\mathcal{B}^Q : L^p(\mathbb{D}) \to L^p(\mathbb{D})$, are bounded.

In view of (10.26) and Theorem 10.2, we have

$$\mathcal{C}^Q(h)_{\bar{z}} = h, \qquad \mathcal{C}^Q(h)_z = \mathcal{B}(h),$$

weakly. We note that when when $z \in \mathbb{D}$, then we can write

$$\mathcal{C}^Q(f)(z) = \frac{Q(z)}{\sqrt{z}} \iint_{\mathbb{D}} \left[\frac{\sqrt{z} f(\zeta)}{Q(\zeta)(\zeta - z)} + \frac{\overline{\sqrt{z} f(\zeta)}}{\overline{Q(\zeta)}(z\bar{\zeta} - 1)} \right] \frac{d\zeta \wedge d\bar{\zeta}}{2\pi i}. \tag{10.28}$$

What we learn from this is that $C^Q(f)(z)$ is a real multiple of $Q(z)/\sqrt{z}$ whenever $|z| = 1$. Hence the boundary behavior of $C^Q(f)$ is tied to that of $Q(z)/\sqrt{z}$. It is this property of $\mathcal{C}^Q$ that we use later on in the proof of Theorem 10.1.

Proof of Theorem 10.1 *Step 1.* Without loss of generality, we may assume that $n = 2$. Also, the surface w we wish to construct may embed into a set that is sympletomorphic to $Z^4(R)$ for some R. For our purposes, we choose $T \times \mathbb{C}$ with $T \subset \mathbb{C}$ a triangle with vertices ± 1 and i. We wish to find a transformation $w = (U, V) : \mathbb{D} \to \mathbb{C}^2$ that solves (10.16) and satisfies the following properties:

(i) $U : \mathbb{D} \to T$, with $U(\partial \mathbb{D}) = \partial T$, and $V : \mathbb{D} \to \mathbb{C}$,

(ii) $(u^0, v^0) \in w(\mathbb{D})$,

(iii) $2area(w(\mathbb{D})) = i \int_{\mathbb{D}} dU \wedge d\bar{U} + dV \wedge d\bar{V} = 2$.

Note that we switched from (10.9) to (iii) because $area(T) = 1$. Pick $a \in \mathbb{D}$. We search for solutions of the form

$$U = \mathcal{C}^1(u) + q, \qquad V = \mathcal{C}^2(v) - \mathcal{C}^2(v)(a) + v^0,$$

for $u, v : \mathbb{D} \to \mathbb{C}$ in L^p, and a holomorphic function q. The operators $\mathcal{C}^1$ and $\mathcal{C}^2$ are defined as $\mathcal{C}^i(f) = \mathcal{C}^{Q^i}(f)$, where

$$Q^1(z) = e^{3i\pi/4}(z-1)^{1/4}(z+1)^{1/4}(z-i)^{1/2}, \qquad Q^2(z) = z - 1.$$

$\square$

A Vector Fields and Differential Forms

Lemma A.1 *(De Rham) Let M be a compact connected orientable k-dimensional manifold and let α be a k-form with $\int_M \alpha = 0$. Then α is exact.*

Proof Let $\{U_1, \ldots, U_r\}$ be a finite cover of M with each U_i diffeomorphic to a simply connected subset of $\mathbb{R}^k$. To ease the notation, we assume that $r = 2$. Choose φ_1 and φ_2 with $\varphi_1 + \varphi_2 = 1$, $\varphi_1, \varphi_2 \geq 0$, $\varphi_1, \varphi_2 \in C^1$ and supp $\varphi_i \subseteq U_i$ for $i = 1, 2$. We can readily construct a k-form β such that supp $\beta \subseteq U_1 \cap U_2$ and $\int_M \beta = 1$. Define $\gamma = \varphi_1\alpha - c\beta$ with $c = \int_M \varphi_1\alpha$. Then supp $\gamma \subseteq U_1$ and $\int_{U_1} \gamma = 0$. By Poincaré's lemma, there exists a form $\hat{\gamma}$ such that $d\hat{\gamma} = \gamma$. Similarly, if $\tau = \varphi_2\alpha + c\beta$, then supp $\tau \subseteq U_2$ and $\int_{U_2} \tau = \int_M (1 - \varphi_1)\alpha + c\beta = \int_M \alpha = 0$. Hence, we can apply Poincaré's lemma to find $\hat{\tau}$ with $d\hat{\tau} = \tau$. We now have

$$d(\hat{\gamma} + \hat{\tau}) = \varphi_1\alpha - c\beta + \varphi_2\alpha + c\beta = \alpha.$$

$\square$

Lemma A.2 *Let X be a C^1 vector field of X and write ϕ_t^X for its flow. If $\varphi : N \to M$ is a diffeomorphism, then $\psi_t = \varphi^{-1} \circ \phi_t^X \circ \varphi$, is the flow of the vector field $\varphi^* X := (d\varphi)^{-1} X \circ \varphi$.*

Proof We have

$$\begin{aligned} \frac{d\psi_t}{dt}(x) &= \left(d\varphi^{-1}\right)_{\phi_t^X \circ \varphi(x)} \left(X \circ \phi_t^X \circ \varphi\right)(x) = \left(d\varphi^{-1}\right)_{\varphi \circ \psi_t^X(x)} \left(X \circ \varphi \circ \psi_t^X\right)(x) \\ &= (d\varphi)^{-1}_{\psi_t^X(x)} (X \circ \varphi) \left(\psi_t^X(x)\right), \end{aligned}$$

which means

$$\psi_t = \phi_t^{(d\varphi)^{-1} X \circ \varphi}.$$

$\square$

The vector field $\varphi^* X$ is called the *pull-back* of X. Also, if Y is a vector field on N, we define its *φ-push-forward* by

$$(\varphi_* Y)(y) = (d\varphi)_{\varphi^{-1}(y)} Y\left(\varphi^{-1}(y)\right).$$

We define the *Lie derivative* of a vector field Y with respect to another vector field X by

$$[X, Y] = \mathcal{L}_X Y := \lim_{t \to 0} t^{-1} \left(\left(\phi_t^X\right)^* Y - Y\right). \tag{A.1}$$

Using

$$\left(\phi_t^X\right)^* \left(\phi_s^X\right)^* Y = \left(\phi_{s+t}^X\right)^* Y,$$

we can readily show

$$\text{(A.2)} \qquad \frac{d}{dt}\left(\phi_t^X\right)^* Y = \mathcal{L}_X \left(\phi_t^X\right)^* Y = \left(\phi_t^X\right)^* \mathcal{L}_X Y.$$

In particular $Z(x,t) = \left(\phi_t^X\right)^* Y$ satisfies

$$Z_t = [X, Z] = \mathcal{L}_X Z.$$

Lemma A.3 *For every* 1*-form* α *and vector fields* X *and* Y*, we have*

$$\text{(A.3)} \qquad d\alpha(X,Y) = \mathcal{L}_X\big(\alpha(Y)\big) - \mathcal{L}_Y\big(\alpha(X)\big) + \alpha\big([X,Y]\big).$$

Proof To ease the notation, we write ϕ_t for ϕ_t^X. By definition

$$\left(\phi_t^*\alpha\right)\left(\phi_t^*Y\right) = \alpha(Y) \circ \phi_t.$$

We now differentiate both sides with respect to t and set $t = 0$. We obtain

$$\left(\mathcal{L}_X\alpha\right)(Y) + \alpha\big([X,Y]\big) = \mathcal{L}_X\big(\alpha(Y)\big).$$

By Cartan's formula, the left hand side equals

$$\left(i_X d\alpha\right)(Y) + d\big(\alpha(X)\big)(Y) + \alpha\big([X,Y]\big) = d\alpha(X,Y) + \mathcal{L}_Y\big(\alpha(X)\big) + \alpha\big([X,Y]\big),$$

which implies (A.3). □

Given a sub-bundle ξ of TM of dimension m, we may wonder whether or not there exists a foliation of M that consists of submanifolds N such that for each $x \in N$, we have $T_xN = \xi_x$. If such a foliation exists locally, we say that ξ is *integrable*. According to *Frobenius Theorem* the sub-bundle ξ is integrable iff for every vector fields $X, Y \in \xi$, we have $[X, Y] \in \xi$.

We are particularly interested in the case of $\xi_x = \ker \alpha_x$, for a 1-form α. For example, if $M = \mathbb{R}^k$ and $\alpha = u{\cdot}dx$, then $\xi_x = u^\perp$ consists of vectors that are perpendicular to u. If $k = 3$, $D = B_1(0)$ is the unit disk, and $w : D \to M$ parametrizes a surface with $T_x w(D) = u^\perp$, then by Stokes' theorem,

$$\begin{aligned} 0 &= \int_{w(\gamma)} \alpha = \int_{w(\Gamma)} d\alpha = \int_\Gamma (d\alpha)_{w(s_1,s_2)}\ (w_{s_1}(s_1,s_2), w_{s_2}(s_1,s_2))\ ds_1 ds_2 \\ &= \int_\Gamma \left[(w_{s_1} \times w_{s_2}) \cdot (\nabla \times u)(w)\right]\ ds_1 ds_2. \end{aligned}$$

for every open subset $\Gamma \subset D$ with $\partial\Gamma = \gamma$. By varying Γ and using the assumption that u is parallel to $w_{s_1} \times w_{s_2}$, we learn that $(\nabla \times u) \cdot u \equiv 0$. More generally we have the following consequence of Lemma 10.3.

Lemma A.4 *Let α be a non-degenerate 1-form on M and set $\xi = \ker\alpha$. The following statements are equivalents:*

(i) *The sub-bundle ξ is integrable.*

(ii) *For every vector fields $X, Y \in \xi$, we have $d\alpha(X, Y) = 0$.*

(iii) $\alpha \wedge d\alpha = 0$.

Proof The equivalence of **(i)** and **(ii)** is an immediate consequence of (A.3). We now show that **(ii)** implies **(iii)**. Assume **(ii)**. Take any vector fields X, Y in ξ and any vector field R such that $\alpha(R) = 1$. Define $\pi(v) = v - \alpha(v)R$ so that π_x is the $R(x)$-projection onto ξ_x. We have

$$\begin{aligned}(\alpha \wedge d\alpha)(v_1, v_2, v_3) &= \alpha(v_1)d\alpha(v_2, v_3) - \alpha(v_2)d\alpha(v_1, v_3) + \alpha(v_3)d\alpha(v_1, v_2)\\ &= \alpha(v_1)\big[\alpha(v_2)d\alpha(R, \pi(v_3)) + \alpha(v_2)d\alpha(\pi(v_2), R)\big]\\ &\quad - \alpha(v_2)\big[\alpha(v_1)d\alpha(R, \pi(v_3)) + \alpha(v_3)d\alpha(\pi(v_1), R)\big]\\ &\quad + \alpha(v_3)\big[\alpha(v_1)d\alpha(R, \pi(v_2)) + \alpha(v_2)d\alpha(\pi(v_1), R)\big]\\ &= 0,\end{aligned}$$

which means that $\alpha \wedge d\alpha = 0$.

Conversely, if **(iii)** is true and X, Y are any two vector fields in ξ, then

$$0 = (\alpha \wedge d\alpha)(R, X, Y) = \alpha(R)d\alpha(X, Y) - \alpha(X)d\alpha(R, Y) + \alpha(Y)d\alpha(R, X) = d\alpha(X, Y),$$

as desired. □

B Sobolev Inequality

It is well-known that $\mathcal{H}^{1/2}$ is a subset of the space of functions of *bounded mean oscillation (BMO)*. In particular $\mathcal{H}^{1/2} \subset L^p$ for all $p \geq 2$. We will prove this for $p < 3$.

Lemma B.1 *For every $p \in [2, 3)$, there exists a constant $c_0 = c_0(p)$ such that*

$$\left(\int_0^1 |x(t)|^p \; dt\right)^{1/p} \leq c_0(p)\|x\|.$$

Proof We identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ and write $-i$ for $\bar{J}$. Hence $x(t) = \sum_k e^{2\pi kt\bar{J}}x_k$ can be rewritten as $\sum_k e^{-2\pi kti}x_k$ with $x_k \in \mathbb{C}^n$. Since it suffices to establish the inequality for each component, we may assume without loss of generality that $n = 1$.

We now find an expression of $\|x\|$ that involves the function $x(\cdot)$ directly and does not involve its Fourier coefficients. We claim

$$\int_0^{2\pi}\int_0^{2\pi}\frac{|x(e^{i\theta})-x(e^{i\varphi})|^2}{|e^{i\theta}-e^{i\varphi}|^2}d\theta d\varphi = 4\pi^2\sum_k |k||x_k|^2. \tag{B.1}$$

This follows from a direct calculation; the left-hand side equals

$$\begin{aligned}
&\int_0^{2\pi}\int_0^{2\pi}|e^{i\theta}-e^{i\varphi}|^{-2}\left|\sum_k x_k e^{-ik\theta}-\sum_k x_k e^{-ik\varphi}\right|^2 d\theta d\varphi\\
&\quad=\int_0^{2\pi}\int_0^{2\pi}|1-e^{i\tau}|^{-2}\left|\sum_k x_k(1-e^{-ik\tau})e^{-ik\theta}\right|^2 d\theta d\tau\\
&\quad=2\pi\int_0^{2\pi}|1-e^{i\tau}|^{-2}\sum_k |x_k|^2|1-e^{-ik\tau}|^2 d\tau=(2\pi)^2\sum_k |k||x_k|^2,
\end{aligned}$$

because

$$\int_0^{2\pi}\frac{|1-e^{-ik\tau}|^2}{|1-e^{i\tau}|^2}d\tau=\int_0^{2\pi}(e^{-i(k-1)\tau}+\cdots+1)(e^{i(k-1)\tau}+\cdots+1)d\tau=2\pi.$$

Let us write $\Lambda(x)$ for the left-hand side of (**??**). To simplify the notation write $y(\theta)=x(e^{i\theta})$. We have that for constants c_1 and c_2,

$$\int_0^{2\pi}\int_0^{2\pi}|y(\theta)-y(\varphi)|^2 1\!\!1(|\theta-\varphi|<l^{-1})d\theta d\varphi\le c_1 l^{-2}\Lambda(x)=c_2 l^{-2}\|x\|^2. \tag{B.2}$$

Given $l\ge 1$, define $z^l(t)=l\int_0^{l^{-1}}x(t+\alpha)\,d\alpha$. By (B.2),

$$\begin{aligned}
\int_0^1 |x(t)-z^l(t)|^2 dt &= \int_0^1\left|l\int_0^{l^{-1}}(x(t)-x(t+\alpha))\,d\alpha\right|^2 dt\\
&\le l\int_0^1\int_0^{l^{-1}}|x(t)-x(t+\alpha)|^2\,dtd\alpha\\
&\le c_2 l^{-1}\|x\|^2.
\end{aligned}$$

From this we deduce that for every $l\ge 1$,

$$x=z^l+w^l$$

with

$$\|w^l\|_0^2\le c_2 l^{-1}\|x\|^2,\ \|z^l\|_{L^\infty}\le l\|x\|,$$

because

$$|z^l(t)| \leq l \int_0^1 |x(t)| \, dt \leq l\|x\|_0.$$

Hence we apply Chebyshev's inequality to assert

$$|\{t : |x(t)| > 2l\|x\|\}| \leq |\{t : |w^l(t)| > l\|x\|\}| \leq \frac{\|w^l\|_0^2}{l^2\|x\|^2} \leq c_2 l^{-3},$$

whenever $l \geq 1$. On the other hand

$$\begin{aligned}
\int_0^1 |x(t)|^p \, dt &\leq 1 + \int_0^\infty p l^{p-1} |\{x > l\}| dl \\
&\leq 1 + c_3\|x\|^3 \int_1^\infty p l^{p-1} l^{-3} dl \\
&= 1 + c_4(p)\|x\|^3
\end{aligned}$$

with $c_4(p) < \infty$ whenever $p < 3$. Finally, we replace x with λx, $\lambda > 0$ to deduce

$$\|x\|_{L^p}^p \leq \lambda^{-p} + c_4(p)\lambda^{3-p}\|x\|^3.$$

Minimizing the right-hand side over $\lambda > 0$ yields the desired inequality. □

C Degree Theory

We first review the classical Brouwer degree theory. Consider triplets (f, U, y) with $U \subseteq \mathbb{R}^d$ open and bounded, $f : \bar{U} \to \mathbb{R}^d$ continuous and $y \notin f(\partial U)$. We now would like to assign an integer $\deg(f, U, y)$ to (f, U, y) that, in some sense, counts the solutions to the equation $f(x) = y$, $x \in U$, with a sign. This degree satisfies the following properties:

(i) If $V \subseteq U$ and $f^{-1}(\{y\}) \subseteq V$, then $\deg(f, V, y) = \deg(f, U, y)$.

(ii) For a constant a, $\deg(f + a, U, y + a) = \deg(f, U, y)$.

(iii) If $U \cap V = \emptyset$ and $y \notin f(\partial U) \cup f(\partial V)$, then $\deg(f, U \cup V, y) = \deg(f, U, y) + \deg(f, V, y)$.

(iv) If $f : \bar{U} \times [0, 1] \to \mathbb{R}^d$ is continuous with $y \notin f(\partial U, t)$ for every $t \in [0, 1]$, then

$$\deg(f(\cdot, 1), U, y) = \deg(f(\cdot, 0), U, y).$$

(v) If $\deg(f, U, y) \neq 0$, then $f(x) = y$ has a solution in U.

(vi) $\deg(id, B, 0) = 1$ where $B = \{x : |x| < 1\}$.

It turns out that the above properties determine "deg" uniquely. Indeed one can show that for any triplet as above, we can find (g, U, y) with g smooth and f and g homotopic. As for $g \in C^1$, deg is defined by

$$\deg(g, U, y) = \sum_{x \in f^{-1}\{y\}} \operatorname{sgn}(\det g'(x)).$$

In the same fashion, we can define the degree of a continuous map between manifolds. Given two compact manifolds M and N, and a C^1 map $f : N \to M$, we say $x \in N$ is *regular* if df_x is invertible. We say $x \in M$ is a *regular value* if $f^{-1}\{x\}$ consists of regular points. By inverse mapping theorem, it is not hard to show that if x is a regular value, then $f^{-1}\{x\}$ is finite. For such a value we may define the degree by

$$\deg_x(f) = \sum_{y \in f^{-1}\{x\}} \epsilon_y, \tag{C.1}$$

where $\epsilon_y = \pm 1$ according to whether df_x preserves or reverses orientation. In the Euclidean case $\epsilon_y = \operatorname{sgn} \det D_x f$. The degree of a continuous $f : N \to M$ defined to be the degree of a C^1 function $g : N \to M$ that is sufficiently close to f. As we will see in Lemma C.1, this is well-defined.

Lemma C.1 **(i)** *If $f : N \to M$ is a C^1 function and Ω is a volume form with $\Omega > 0$, $\int_M \Omega = 1$, then for every regular value $x \in M$,*

$$\deg_x f = \int_N f^* \Omega.$$

(ii) *The degree is invariant under homotopies consisting of C^1-maps.*

(iii) *Any two C^1-maps $f, g : N \to M$ that are sufficiently C^0-close are homotopic via C^1-maps.*

(iv) *Let X be an orientable manifold with $\partial X = N$ and let $F : X \to M$ be a continuous map. Then the degree of $f = F|_N$ is zero.*

Proof (i) Let x be a regular value and assume $f^{-1}\{x\} = \{x_1, \dots, x_k\}$. Find an open neighborhood V of x such that $f^{-1}(V) = U_1 \cup \cdots \cup U_k$ with $U_1, \dots, U_k$, open and disjoint, $x_i \in U_i$ for $i = 1, \dots, k$, and $f|_{U_i} : U_i \to V$ a diffeomorphism for every i. We now take an n-form α with support in V such that $\int_V \alpha = 1$. By Lemma 3.10, we may find an $(n-1)$-form β such that $\Omega = \alpha + d\beta$. We now have

$$\begin{aligned} \int_N f^* \Omega &= \int_N f^* \alpha + \int_N df^* \beta = \int_N f^* \alpha \\ &= \sum_{i=1}^k \int_{U_i} f^* \alpha = \sum_{i=1}^k \epsilon_{x_i}. \end{aligned}$$

(ii) Clearly degree is C^1-continuous and locally constant.

(iii) Put a Riemannian metric on M. Given $x \in N$, we may find a geodesic curve connecting $f(x)$ to $g(x)$. We now define $\psi(t,x) = \gamma(t; f(x), g(x))$ where $\gamma(t;a,b)$ is defined to be a point on the geodesic connecting a to b with $\gamma(0;a,b) = a$, $\gamma(1;a,b) = b$. This can be done smoothly for a, b sufficiently close.

(iv) Let Ω be a volume form on M with $\int_M \Omega = 1$, $\Omega > 0$. We then have

$$\deg(f) = \int_N f_*\Omega = \int_{\partial X} f_*\Omega = \int_X d(F_*\Omega) = \int_X F_* d\Omega = 0.$$

□

The *Leray–Schander Theory* allows us to have a similar notion of degree for functions of the form $f = I + L : \bar{U} \to \mathcal{E}$ with $\mathcal{E}$ a Banach space, U a bounded open subset of $\mathcal{E}$, and L a compact operator. Again we wish to define $\deg(f, U, y)$ provided that $y \notin f(\partial U)$. To do so, first we find a sequence $L_m : \bar{U} \to \mathcal{E}$ such that the range of L_m, denoted by $L_m(\bar{U})$, is a subset of finite dimensional space $\mathcal{E}_m$, and

$$\lim_{m\to\infty} \sup_{x\in\bar{U}} \|L_m(x) - L(x)\| = 0.$$

We then set

$$\deg(f, U, y) = \deg(f_m, \mathcal{E}_m \cap U, y), \tag{C.2}$$

for large m, where $f_m = I + L_m : \mathcal{E}_m \cap \bar{U} \to \mathcal{E}_m$. For this to work, we need to check that $y \notin f_m(\partial U)$ for sufficiently large m. By Exercise C(ii), the set $f(\partial U)$ is closed. Since $x \notin f(\partial U)$, we have $\text{dist.}(x, f(\partial U)) = \delta > 0$. Then we find m_0 such that if $m > m_0$, then $\text{dist.}(x, f_m(\partial U)) \geq \delta/2$. Hence the right-hand side of $(A.2)$ is well-defined by Lemma C.2 below.

Lemma C.2 *Let U be an open bounded subset of $\mathbb{R}^n = \mathbb{R}^{n_1} \oplus \mathbb{R}^{n_2}$. Consider $f(x) = x + L(x)$ with $L : \bar{U} \to \mathbb{R}^{n_1}$, $f : \bar{U} \to \mathbb{R}^n$. If $y \notin f(\partial U)$ and $y \in \mathbb{R}^{n_1}$, then*

$$\deg(f, U, y) = \deg(f|_{U_1}, U_1, y)$$

where $U_1 = U \cap \mathbb{R}^{n_1}$.

Proof We may assume $f \in C^1(U)$ and $y = 0$. Let us take two continuous functions φ_1, φ_2 with $\varphi_i : \mathbb{R}^{n_i} \to \mathbb{R}$, $\int \varphi_i dy_i = 1$, for $i = 1, 2$, and both φ_1, φ_2 have support near 0. Set $\varphi(y_1, y_2) = \varphi_1(y_1)\varphi_2(y_2)$ and $\omega = \varphi(y_1, y_2) dy_1 dy_2$ is a volume form of total volume 1. We then have

$$\deg(f, U, y) = \int_U f^*\omega,$$

by Lemma A.1. But $\det(I_n + \nabla L) = \det\left(I_{n_1} + \frac{\partial L}{\partial x_1}\right)$. As a result,

$$\deg(f, U, y) = \int \varphi_1(x_1 + L(x))\varphi_2(x_2) \det\left(I_{n_1} + \frac{\partial L}{\partial x_1}\right) dx_1 dx_2.$$

We may send φ_2 to δ_0 to yield

$$\begin{aligned}\deg(f, U, y) &= \int \varphi_1(x_1 + L(x)) \det\left(I_{n_1} + \frac{\partial L}{\partial x_1}\right) dx_1 \\ &= \deg(f|_{U_1}, U_1, 0).\end{aligned}$$

□

Exercise C Let $\mathcal{E}$ be a Banach space and $K : \Omega \to \mathcal{E}$ be a compact operator with Ω a bounded closed subset of X.

- (i) Show that K is a uniform limit of finite dimensional transformations. **Hint:** Cover the compact set $\overline{K(\Omega)}$ by finitely many open balls, and use a partition of unity.
- (ii) Show that $I + K$ maps closed sets to closed sets.

□

D Cauchy and Beurling Transforms

A classical way of solving the Laplace and Poisson equation in an bounded open subset of $\mathbb{R}^k$ with regular boundary is by first finding its Green's function. That is a function $G : \bar{U} \times \bar{U} \to \mathbb{R}$, such that

$$\begin{cases} \Delta_x G(x, y) = \delta_y(dx), & x \in U, \\ G(x, y) = 0, & x \in \partial U. \end{cases}$$

Once such G is found, we then use Green's identity to derive the following identity:

$$u(x) = \int_U \Delta u(y)\ G(x, y)\ dy + \int_{\partial U} u(y)\ \frac{\partial G}{\partial n}(x, y)\ dy, \tag{D.1}$$

where $\partial G/\partial n$ denotes the normal derivative of G. Once the Green's function G and the *Poisson's function* $\partial G/\partial n$ are known, then we can use (D.1) to solve the PDE

$$\begin{cases} \Delta u(x) = f(x), & x \in U, \\ u(x) = g(x), & x \in \partial U. \end{cases}$$

for the given f and g. In $\mathbb{C}$ an analogous representation formula can be derived that in turn can be used to solve the celebrated *d-bar and Beltrami equations.*

Proposition D.1 *(Cauchy-Pompeiu Formula) Let $U \subset \mathbb{C}$ a bounded domain with a boundary that is positively oriented and parametrized by a curve γ. Then, for every C^1 function $f : \mathbb{C} \to \mathbb{C}$,*

$$
\begin{aligned}
f(z) &= \frac{1}{2\pi i} \oint_\gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta + \frac{1}{2\pi i} \iint_U \frac{f_{\bar\zeta}(\zeta)}{\zeta - z}\, d\zeta \wedge d\bar\zeta, \\
f(z) &= -\frac{1}{2\pi i} \oint_\gamma \frac{f(\zeta)}{\bar\zeta - \bar z}\, d\bar\zeta + \frac{1}{2\pi i} \iint_U \frac{f_{\bar\zeta}(\zeta)}{\bar\zeta - \bar z}\, d\zeta \wedge d\bar\zeta.
\end{aligned} \tag{D.2}
$$

Proof We only derive the first formula in (D.2) because the second identity can be established by a verbatim argument. Write $\zeta = s + it$ and pick any C^1 function $g = u + iv$. Note

$$\frac{i}{2}(d\zeta \wedge d\bar\zeta) = ds \wedge dt.$$

Using the Green's formula,

$$
\begin{aligned}
\oint_\gamma g(\zeta) d\zeta = \oint_\gamma (u\ ds - v\ dt) + i(u\ dt + v\ ds) &= \iint_U \big[-(u_t + v_s) + i(u_s - v_t)\big]\ dsdt \\
&= i \iint_U \big[(u_s - v_t) + i(u_t + v_s)\big]\ dsdt = i \iint_U (g_s + i g_t)\ dsdt \\
&= 2i \iint_U g_{\bar\zeta}\ dsdt.
\end{aligned} \tag{D.3}
$$

A more compact version of the above calculation is

$$\oint_\gamma g(\zeta) d\zeta = \iint_U dg \wedge d\zeta = \iint_U g_{\bar\zeta}\ d\bar\zeta \wedge d\zeta = 2i \iint_U g_{\bar\zeta}\ dsdt.$$

We now take $z \in U$, choose $\varepsilon > 0$ so small that $B_\varepsilon(z) \subset U$, set $g(\zeta) = f(\zeta)/(\zeta - z)$ and replace U with $U/B_\varepsilon(z)$ in (D.3) to deduce

$$\frac{1}{2\pi i} \oint_{|\zeta - z| = \varepsilon} \frac{f(\zeta)}{\zeta - z}\ d\zeta = \frac{1}{2\pi i} \oint_\gamma \frac{f(\zeta)}{\zeta - z}\ d\zeta - \frac{1}{\pi} \iint_U \frac{f_{\bar\zeta}(\zeta)}{\zeta - z}\ dsdt.$$

It remains to show that the right-hand side converges to $f(z)$ as $\varepsilon \to 0$. For this it suffices to check

$$\lim_{\varepsilon \to 0} \frac{1}{2\pi i} \oint_{|\zeta - z| = \varepsilon} \frac{f(\zeta) - f(z)}{\zeta - z}\ d\zeta = 0.$$

This is an immediate consequence of the Lipschitzness of f that implies the boundedness of the integrand. $\square$

Given a C^1 function $h : \bar{U} \to \mathbb{C}$, we wish to find $w : \bar{U} \to \mathbb{C}$ such that $w_{\bar{z}} = h(z)$ in U. Indeed if w is a solution, then by (D.2), $w = \mathcal{C}(h) + \Gamma(h)$, where the Cauchy operators Γ and $\mathcal{C}$ are defined by

$$\Gamma(g) = \frac{1}{2\pi i} \int_\gamma \frac{g(\zeta)}{\zeta - z}\, d\zeta, \qquad \mathcal{C}(h) = -\frac{1}{\pi} \iint_U \frac{h(\zeta)}{\zeta - z}\, dsdt.$$

This means that a solution can be expressed as $w = \mathcal{C}(h) + q$ for a holomorphic function q. Put this differently, what we learn from (D.2) is

$$\frac{\partial}{\partial \bar{z}} \mathcal{C}(h) = h, \qquad \frac{\partial}{\partial z} \bar{\mathcal{C}}(h) = h, \tag{D.4}$$

where

$$\bar{\mathcal{C}}(h) = -\frac{1}{\pi} \iint_U \frac{h(\zeta)}{\bar{\zeta} - \bar{z}}\, dsdt.$$

With the aid of the Cauchy operator/transform $\mathcal{C}$, we can also solve the Beltrami's equation

$$\varphi_{\bar{z}} = \mu \varphi_z. \tag{D.5}$$

This is an elliptic PDE only if $|\mu(z)| < 1$ for all z. We may set $h = \varphi_{\bar{z}}$ and express φ as $\varphi = \mathcal{C}(h) + q$, for a holomorphic function q. Observe that if φ satisfies (D.5), then

$$h = \varphi_{\bar{z}} = \mu \varphi_z = \mu \mathcal{B}(h) + \mu q',$$

where $\mathcal{B}$ is the *Beurling's operator*:

$$\mathcal{B}(h) = \frac{\partial}{\partial z} \mathcal{C}(h).$$

This means that the function h satisfies

$$(I - \mu \mathcal{B}) h = \mu q'. \tag{D.6}$$

In other words, for every holomorphic q, we first solve (D.6) for h and using this $h = h^q$, we have a solution for (D.5) in the form $\varphi = q + \mathcal{C}(h^q)$. Note that $q = 0$ yields the trivial solution $\varphi = 0$. For a nontrivial example, search for a solution of the form $var = \mathcal{C}(h) + 1$, where h solves

$$(I - \mu \mathcal{B}) h = \mu.$$

This or more generally (D.6) can be solve if $\|\mu \mathcal{B}\| < 1$, for a suitable operator norm. As we will see below, if we take the operator norm with respect to the L^2 space, then $\|\mathcal{B}\| = 1$ and if $\sup_z \|\mu(z)\| < 1$, then we can invert $I - \mu \mathcal{B}$ and solve (D.6) in L^2.

The operator $\mathcal{B}$ is the complex-variable analog of the *Hilbert Transform*; from (D.4) we can readily deduce

$$\mathcal{B}(f_{\bar{z}}) = f_z + Q,$$

for some holomorphic Q. It turns out that if we assume $U = \mathbb{C}$ and that f and its first derivative vanish at infinity, then $Q = 0$ (because a bounded entire function is constant), and we simply have $\mathcal{B}(f_{\bar{z}}) = f_z$. If we write

$$\mathcal{F}(h)(\eta) = \hat{h}(\eta) = \int \exp\big(Re(z\bar{\eta})\big)h(z)\ dz,$$

for the Fourier Transform of h, then we have

$$\mathcal{F}(f_z)(\eta) = -\frac{i\bar{\eta}}{2}\mathcal{F}(f)(\eta), \qquad \mathcal{F}(f_{\bar{z}})(\eta) = -\frac{i\eta}{2}\mathcal{F}(f)(\eta).$$

Hence

$$\mathcal{F}(\mathcal{B}(h)) = \frac{\bar{\eta}}{\eta}\mathcal{F}(h)(\eta).$$

From this and Plancherel's equation we deduce that

$$\|\mathcal{B}(h)\|_{L^2} = \|h\|_{L^2}. \tag{D.7}$$

In fact $\mathcal{B}$ is a singular operator that may be defined by

$$\mathcal{B}(h)(z) = \frac{1}{\pi}PV\int \frac{h(\zeta)}{(\zeta - z)^2}\ dsdt = \lim_{\varepsilon\to 0}\frac{1}{\pi}\int_{|\zeta - z|>\varepsilon}\frac{h(\zeta)}{(\zeta - z)^2}\ dsdt,$$

and is also bound on L^p for $p \in (1, \infty)$ by *Calderon-Zygmund Theory.*

Theorem D.1 *Define the operators $\mathcal{C}$ and $\mathcal{B}$ on smooth functions by*

$$\mathcal{C}(h)(z) = -\frac{1}{\pi}\int_{\mathbb{D}}\frac{h(\zeta)}{\zeta - z}\ dsdt, \qquad \mathcal{B}(h)(z) = \frac{1}{\pi}PV\int_{\mathbb{D}}\frac{h(\zeta)}{(\zeta - z)^2}\ dsdt. \tag{D.8}$$

Then for every $p \in (1, \infty)$, there exists constants $C(p)$ and $C'(p)$ such that $\lim C'(p) = 1$ as $p \to 2$, and

$$\|\mathcal{C}(h)\|_{W^{1,p}} \le C(p)\|h\|_{L^p}, \qquad \|\mathcal{B}(h)\|_{L^p} \le C'(p)\|h\|_{L^p}.$$

Writing $\mathcal{C} : L^p(\mathcal{C}) \to W^{1,p}(\mathbb{C})$ and $\mathcal{B} : L^p(\mathcal{C}) \to L^p(\mathbb{C})$ for their extensions, we have

$$\mathcal{C}(h)_{\bar{z}} = h, \qquad \mathcal{C}(h)_z = \mathcal{B}(h),$$

weakly.

Remark D.1(i) Assuming that for some $p_0 > 2$, there exists a constant c_0 such that

$$\|\mathcal{B}(h)\|_{L^{p_0}} \leq c_0 \|h\|_{L^{p_0}},$$

for every h, then we can use (D.7) and *Riesz-Thorin Interpolation Theorem* to assert that if

$$\frac{1}{p_\theta} = \frac{\theta}{2} + \frac{1-\theta}{p_0},$$

then

$$\|\mathcal{B}(h)\|_{L^{p_\theta}} \leq c_0^{1-\theta} \ \|h\|_{L^{p_\theta}}.$$

This allows us to choose $C'(p_\theta) = c_0^{1-\theta}$, for $\theta \in [0,1]$, which enjoys the property

$$\lim_{p \to 2} C'(p) = 1.$$

(ii) Clearly there are many solutions to (D.5). In fact if $\mu = 0$, then (D.5) simply requires that φ to be holomorphic and if we also specify $\varphi(\mathbb{D})$, then we still have many solutions by the Riemann Mapping Theorem. Even if μ is nonzero, we can still require $\varphi(\mathbb{D})$ to be a simply connected domain $U \neq \mathbb{C}$ and solve (D.5) provided that $\|\mu\|_{L^\infty} < 1$. Indeed if φ is any solution to (D.5) and $\varphi(\mathbb{D}) = V$, then by the Riemann Mapping Theorem, we can find a holomorphic function $f : V \to \mathbb{C}$ such that $f(V) = U$. Now $\tilde{\varphi} = f \circ \varphi$ does the job because

$$\tilde{\varphi}_{\bar{z}} = (f' \circ \varphi)\varphi_{\bar{z}}, \qquad \tilde{\varphi}_z = (f' \circ \varphi)\varphi_z.$$

This may be seen from

$$d\tilde{\varphi} = f_\varphi \ d\varphi + f_{\bar{\varphi}} \ d\bar{\varphi} = (f' \circ \varphi) \ d\varphi = (f' \circ \varphi)(\varphi_z \ dz + \varphi_{\bar{z}} \ d\bar{z}).$$

(iii) Observe that if $q, Q : \mathbb{D} \to \mathbb{C}$ are two holomorphic functions with $Q \neq 0$ anywhere inside $\mathbb{D}$, then $f = Q\mathcal{C}(g/Q) + q$ would solve the equation $f_{\bar{z}} = g$.

□

After solving the Beltami Equation, we may wonder whether the solution φ of (D.5) is a homeomorphism. One strategy for verifying the invertibility of φ is to derive an equation for its inverse, verify its solvability, and show that the solution is indeed the inverse of φ. As we have seen in Remark 10.1, the inverse w solves

(D.9) $$w_{\bar{z}} = m(w)\overline{w_z},$$

with $|m| < 1$.

Theorem D.2 *Assume that* $\|m\|_{L^\infty} = c_0 < 1$. *Then there exists* $p = p(c_0) > 2$ *and a function* $w \in W^{1,p}$ *that solves the equation* (D.9) *weakly.*

Proof The method we described above would not work and need to be modified. Indeed if we set $v = w_{\bar{z}}$, we certainly have $w = \mathcal{C}(v) + q$ for a holomorphic q. On the other hand

$$v = m(w)\bar{\mathcal{B}}(v) + m(w)\overline{q'}, \tag{D.10}$$

which cannot be solved as before because m depends on the unknown. To get around this, observe that for given w, the operator $I - m(w)\bar{\mathcal{B}}$ is invertible because $\|m\|_{L^\infty} = c_0 < 1$. Let us define

$$\mathcal{D}(w) = \big(I - m(w)\bar{\mathcal{B}}\big)^{-1}\big(m(w)\overline{q'}\big)$$

Then we use (D.10) to rewrite (D.9) as

$$w = \big(\mathcal{C} \circ \mathcal{D}\big)(w) + q := \mathcal{E}(w). \tag{D.11}$$

Hence w is a fixed point of the operator $\mathcal{E}$. To show that $\mathcal{E}$ has a fixed point, we first decide on its domain of definition. Set

$$\Gamma_L = \big\{w : \ \|w\|_{L^\infty} \le L\big\}.$$

We wish to show that $\mathcal{E} : \Gamma_L \to \Gamma_L$ and it has a fixed point. As a preparation, we first bound the nonlinear $\mathcal{D}$. Let $C'(p)$ be as in Theorem D.1. Choose $p > 2$ such that $c_0 C'(p) < 1$. By Theorem D.1,

$$\Big\|\big(I - m(w)\bar{\mathcal{B}}\big)^{-1} h\Big\|_{L^p} \le \big(1 - c_0 C'(p)\big)^{-1}\|h\|_{L^p}.$$

This in turn implies,

$$\|\mathcal{D}(w)\|_{L^p} \le c_0\big(1 - c_0 C'(p)\big)^{-1}\|q'\|_{L^p} := c_1(p).$$

By applying Theorem D.1 again we learn

$$\|\mathcal{C} \circ \mathcal{D}(w)\|_{W^{1,p}} \le C(p)c_1.$$

Since $p > 2$, we may apply Morrey Inequality to deduce

$$\|\mathcal{C} \circ \mathcal{D}(w)\|_{C^{1-2/p}} \le c_2(p)\|\mathcal{C} \circ \mathcal{D}(w)\|_{W^{1,p}} \le C(p)c_1(p) = c_3(p), \tag{D.12}$$

where C^α denotes the space of α-Hölder continuous functions. In particular

$$\|\mathcal{C} \circ \mathcal{D}(w)\|_{L^\infty} \le c_3(p).$$

Setting $L = c_3(p) + \|q'\|_{L^\infty}$, we deduce that $\mathcal{E}(L^p) \subseteq \Gamma_L$. In particular, $\mathcal{E}$ maps Γ_L into itself. On the other hand, the bound (D.12) implies that the image of Γ_L under $\mathcal{E}$ is in fact compact. This allows us to use the *Schauder Fixed Point Theorem* to deduce that $\mathcal{E}$ has a fixed point. □

References

[AB] A. Abbondandolo, "Morse Theory for Hamiltonian Systems," Chapman and Hall/CRC Research Notes in Mathematics, 425 , 2001.

[AM] A. Ambrosetti and A. Malchiodi, "Nonlinear Analysis and Semilinear Elliptic Problem," Cambridge, 2007.

[C] M. Chaperon, "On generating families", in H. Hofer, C.H. Taubes, A. Weinstein, E. Zehnder (Editors), The Floer Memorial Volume, Birkhäuser (1995).

[CHLS] K. Cieliebak, H. Hofer, J. Latschev and F. Schlenk, *Quantitative symplectic geometry*, preprint (2007).

[E] I. Ekeland, "Convexity Methods in Hamiltonian Mechanics," Springer, 1990.

[G] C. Golé, "Symplectic Twist Maps," World Scientific, 2001.

[HZ] H. Hofer and E. Zehnder, "Symplectic Invariants and Hamiltonian Dynamics," Birkhäuser, 1994.

[H] M. Hutchings, Taubes's proof of the Weinstein conjecture in dimension three. Bull. Amer. Math. Soc. (N.S.) **47** , no. 1, 73125 (2010).

[MS] D. McDuff and D. Salamon, "Introduction to Symplectic Topology," Oxford, 1998.

[Sk] F. Schlenk,"Embedding Problems in Symplectic Geometry," de Gryter, 2005.

[Sz] M. Schwarz, "Morse Homology," Birkhäuser, 1993.

[St] M. Struwe, "Variational Methods," Springer, 1990.

[V] C. Villani, "Topics in Optimal Transportation," Graduate Studies in Mathematics series, vol. 58, AMS (2003).