The restricted three body problem and holomorphic curves

Urs Frauenfelder, Otto van Koert

Contents

CHAPTER 1

Symplectic geometry and Hamiltonian mechanics

1. Symplectic manifolds

The archetypical example of a symplectic manifold is the cotangent bundle of a smooth manifold. Assume that N is a manifold, by physicists also referred to as the *configuration space*. The *phase space* is the cotangent bundle T^*N. The cotangent bundle comes endowed with a canonical one-form $\lambda \in \Omega^1(T^*N)$ called the *Liouville one-form*. It is defined as follows. Abbreviate by $\pi\colon T^*N \to N$ the footpoint projection. If $e \in T^*N$ and $\xi \in T_eT^*N$, the tangent space of T^*N at e, the differential of the footpoint projection at e is a linear map

$$d\pi(e)\colon T_eT^*N \to T_{\pi(e)}N.$$

Interpreting e as a vector in $T^*_{\pi(e)}N$ the dual space of $T_{\pi(e)}N$, we can pair it with the vector $d\pi(e)\xi \in T_{\pi(e)}N$ and define

$$\lambda_e(\xi) := e\big(d\pi(e)\xi\big).$$

In canonical coordinates $(q,p) = \{q_1, \dots, q_n, p_1, \dots, p_n\}$ of T^*N where $n = \mathrm{dim} N$ the Liouville one-form becomes

$$\lambda(q,p) = \sum_{i=1}^n p_i dq_i.$$

The canonical symplectic form on T^*N is the exterior derivative of the Liouville one-form

$$\omega = d\lambda.$$

In canonical coordinates it has the form

$$\omega = \sum_{i=1}^n dp_i \wedge dq_i.$$

The symplectic form ω has the following properties. It is *closed*, i.e., $d\omega = 0$. This is immediate because $d\omega = d^2\lambda = 0$. Further it is *non-degenerate* in the sense that if $e \in T^*N$ and $\xi \neq 0 \in T_eT^*N$, then there exists $\eta \in T_eT^*N$ such that $\omega(\xi, \eta) \neq 0$. These two properties become the defining properties of a symplectic structure on a general manifold M, which does not need to be a cotangent bundle. Namely

Definition 1.1. *A* symplectic manifold *is a tuple (M, ω) where M is a manifold and $\omega \in \Omega^2(M)$ is a two-form satisfying the following two conditions*

(i): *ω is closed.*
(ii): *ω is non-degenerate.*

The two-form ω is called the symplectic structure *on M.*

The assumption that ω is non-degenerate immediately implies that a symplectic manifold is even dimensional. In other words an odd dimensional manifold never admits a symplectic structure.

2. Symplectomorphisms

Symplectic manifolds become a category with morphisms given by symplectomorphisms defined as follows.

Definition 2.1. *Assume that* (M_1, ω_1) *and* (M_2, ω_2) *are two symplectic manifolds. A* symplectomorphism $\phi \colon M_1 \to M_2$ *is a diffeomorphism satisfying* $\phi^*\omega_2 = \omega_1$.

We discuss three examples of symplectomorphisms.

2.1. Physical transformations. Suppose that N_1 and N_2 are manifolds and $\phi \colon N_1 \to N_2$ is a diffeomorphism, for example a change of coordinates of the configuration space. If $x \in N_1$ the differential

$$d\phi(x) \colon T_x N_1 \to T_{\phi(x)} N_2$$

is a vector space isomorphism. Dualizing we get a vector space isomorphism

$$d\phi(x)^* \colon T^*_{\phi(x)} N_2 \to T^*_x N_1.$$

We now define

$$d_*\phi \colon T^*N_1 \to T^*N_2$$

as follows. If $\pi_1 \colon T^*N_1 \to N_1$ is the footpoint projection and $e \in T^*N_1$, then

$$d_*\phi(e) := (d\phi(\pi_1(e))^*)^{-1} e. \tag{1}$$

If λ_1 is the Liouville one-form on T^*N_1 and λ_2 is the Liouville one-form on T^*N_2 one checks that

$$(d_*\phi)^*\lambda_2 = \lambda_1. \tag{2}$$

Because exterior derivative commutes with pullback we obtain

$$(d_*\phi)^*\omega_2 = (d_*\phi)^* d\lambda_2 = d(d_*\phi)^*\lambda_2 = d\lambda_1 = \omega_1$$

which shows that $d_*\phi$ is a symplectomorphism.

Equation (2) might be rephrased in saying that $d_*\phi$ is an *exact symplectomorphism*, i.e., a symplectomorphism which preserves the primitives of the symplectic forms. The notion of exact symplectomorphism in a general symplectic manifold however does not make sense, since usually the symplectic form ω has no primitive. In fact if the symplectic manifold (M, ω) is closed, the non-degeneracy of the symplectic form ω implies that the closed form ω induces a non-vanishing class $[\omega] \in H^2_{dR}(M)$, the second de Rham cohomology group of M. In particular, ω cannot be exact.

2.2. The switch map. The second example of a symplectomorphism is the switch map
$$\sigma\colon T^*\mathbb{R}^n \to T^*\mathbb{R}^n.$$
Namely, if we identify the cotangent bundle $T^*\mathbb{R}^n$ with $\mathbb{R}^{2n}$ with global coordinates $(q,p) = (q_1,\dots,q_n,p_1,\dots,p_n)$ and symplectic form $\omega = \sum_{i=1}^n dp_i \wedge dq_i$ then the switch map is given by
$$\sigma(q,p) = (-p,q)$$
which is actually a linear symplectomorphism on $\mathbb{R}^{2n}$. Note that the switch map is not a physical transformation. In physical terms the q variables, i.e., the variables on the configuration space are referred to as the position variables, while the p variables are referred to as the momenta. Hence the switch map interchanges the roles of the momenta and the positions. We will see, that the switch map plays a major role in Moser's regularization of two body collisions. Note that to define the switch map it is important to have global coordinates on the configuration space. There is no way to define the switch map on the cotangent bundle T^*N of a general manifold N.

2.3. Hamiltonian transformations. The third example of symplectomorphisms we discuss are Hamiltonian transformations. Suppose that (M,ω) is a symplectic manifold and $H \in C^\infty(M,\mathbb{R})$. Smooth functions on a symplectic manifold are referred by physicists as *Hamiltonians*. The interesting point about Hamiltonians is that we can associate to them a vector field $X_H \in \Gamma(TM)$ which is implicitly defined by the condition
$$dH = \omega(\cdot, X_H).$$
Note that the assumption that the symplectic form is non-degenerate guarantees that X_H is well defined. The vector field X_H is called *Hamiltonian vector field*. Let us assume for simplicity in the following that M is closed. Under this assumption the flow of the Hamiltonian vector field exists for all times, i.e., we get a smooth family of diffeomorphisms
$$\phi_H^t\colon M \to M, \quad t \in \mathbb{R}$$
defined by the conditions
$$\phi_H^0 = \mathrm{id}_M, \qquad \frac{d}{dt}\phi_H^t(x) = X_H(\phi_H^t(x)),\ t\in\mathbb{R},\ x\in M.$$
An important property of the Hamiltonian flow is that the Hamiltonian H is preserved under it. If one interprets the Hamiltonian as the energy then this means the the energy is conserved.

Theorem 2.2 (Preservation of energy). *For $x \in M$ it holds that $H(\phi^t(x))$ is constant, i.e., independent of $t \in \mathbb{R}$.*

Proof: Differentiating we obtain
$$\begin{aligned}\frac{d}{dt}H(\phi_H^t(x)) &= dH(\phi_H^t(x))\frac{d}{dt}\phi_H^t(x)\\ &= dH(\phi_H^t(x))X_H(\phi_H^t(x))\\ &= \omega(X_H,X_H)(\phi_H^t(x))\\ &= 0\end{aligned}$$

where the last equality follows from antisymmetry of the two-form ω. $\square$

The next theorem tells us that the diffeomorphisms ϕ_H^t are symplectomorphisms. The intuition from physics might be that a Hamiltonian system has *no friction*.

Theorem 2.3. *For every $t \in \mathbb{R}$ it holds that $(\phi_H^t)^*\omega = \omega$.*

Proof: Note that

$$\frac{d}{dt}(\phi_H^t)^*\omega = (\phi_H^t)^*\mathcal{L}_{X_H}\omega$$

where $\mathcal{L}_{X_H}\omega$ is the Lie derivative of the symplectic form with respect to the Hamiltonian vector field. Using Cartan's formula we obtain by taking advantage of the assumption that ω is closed and the definition of X_H

$$\mathcal{L}_{X_H}\omega = \iota_{X_H}d\omega + d\iota_{X_H}\omega = -d^2H = 0.$$

This proves the theorem. $\square$

3. Examples of Hamiltonians

3.1. The free particle and the geodesic flow. Assume that (N, g) is a Riemannian manifold. Define

$$H_g \colon T^*N \to \mathbb{R}, \quad (q,p) \mapsto \frac{1}{2}|p|_g^2$$

where $|\cdot|_g$ denotes the norm induced from the metric g on the cotangent bundle of N. In terms of physics this is just the *kinetic energy*.

The flow of this Hamiltonian is basically given by the geodesic flow of the metric g on N. To describe this relation we assume for simplicity that N is compact, in order to ensure that the flows exist for all times. If $q \in N$ and $v \in T_qN$ we denote by

$$q_v \colon \mathbb{R} \to N$$

the unique geodesic meeting the initial conditions

$$q_v(0) = q, \quad \partial_t q_v(0) = v.$$

The *geodesic flow*

$$\Psi_g^t \colon TN \to TN$$

is the map

$$(q, v) \mapsto (q_v(t), \partial_t q_v(t)).$$

The metric g gives rise to a bundle isomorphism

$$\Phi_g \colon TN \to T^*N, \quad (q, v) \mapsto (q, g_q(v, \cdot)).$$

This allows us to interpret the geodesic flow as a map from the cotangent bundle to itself

$$\phi_g^t := \Phi_g \Psi_g^t \Phi_g^{-1} \colon T^*N \to T^*N.$$

Theorem 3.1. *The Hamiltonian flow of kinetic energy H_g equals to geodesic flow in the sense that $\phi_{H_g}^t = \phi_g^t$ for every $t \in \mathbb{R}$.*

Proof: Recall that if $q \in C^\infty([0,1],N)$ is a geodesic, then in local coordinates it satisfies the geodesic equation which is the second order ODE

$$\partial_t^2 q^\ell + \Gamma^\ell_{ij}(q)\partial_t q^i \partial_t q^j = 0. \tag{3}$$

Here we use Einstein summation convention. Moreover, the Γ^ℓ_{ij} are the Christoffel symbols which are determined by the Riemannian metric by the formula

$$\Gamma^\ell_{ij} = \frac{1}{2} g^{\ell k}(g_{ki,j} + g_{jk,i} - g_{ij,k}). \tag{4}$$

If we interpret the geodesic equation as a first order ODE on the tangent bundle of N, i.e., if we consider $(q,v) \in C^\infty([0,1],TN)$ with $q(t) \in N$ and $v(t) \in T_{q(t)}N$, then (3) becomes

$$\left\{ \begin{array}{c} \partial_t v^\ell + \Gamma^\ell_{ij}(q) v^i v^j = 0 \\ \partial_t q = v. \end{array} \right. \tag{5}$$

We next rewrite (5) as an equation on the cotangent bundle instead of the tangent bundle. For this purpose we introduce

$$p_i = g_{ij} v^j.$$

Using the p's instead of the v's and the formula (4) for the Christoffel symbols equation (5) translates to

$$\left\{ \begin{array}{c} \partial_t g^{\ell i} p_i + \frac{1}{2} g^{\ell k}(g_{ki,j} + g_{jk,i} - g_{ij,k}) g^{im} p_m g^{jn} p_n = 0 \\ \partial_t q^i = g^{ij} p_j. \end{array} \right. \tag{6}$$

In view of the identity

$$g^{ij} g_{j\ell} = \delta^i_\ell$$

where δ^i_ℓ is the Kronecker Delta we obtain the relation

$$g^{ij}_k g_{j\ell} + g^{ij} g_{j\ell,k} = 0$$

from which we deduce

$$g^{ij}_{,k} = -g^{i\ell} g^{jm} g_{\ell m,k}.$$

Plugging this formula in the first equation in (6) we obtain

$$\begin{aligned} 0 &= g^{\ell i}_k \partial_t q^k p_i + g^{\ell i}\partial_t p_i - \frac{1}{2}\big(g^{\ell m}_{,j} g^{jn} + g^{\ell n}_{,i} g^{im} - g^{mn}_{,k} g^{k\ell}\big) p_m p_n \\ &= g^{\ell i}_{,k} g^{kj} p_j p_i + g^{\ell i}\partial_t p_i - \frac{1}{2} g^{\ell i}_{,k} g^{kj} p_j p_i \\ &= g^{\ell i}\partial_t p_i + \frac{1}{2} g^{\ell i}_{,k} g^{kj} p_j p_i \end{aligned}$$

from which we get

$$\begin{aligned} \partial_t p_r &= g_{r\ell} g^{\ell i} \partial_t p_i \\ &= -\frac{1}{2} g_{r\ell} g^{\ell k} g^{mn}_{,k} p_m p_n \\ &= -\frac{1}{2} \delta^k_r g^{mn}_{,k} p_m p_n \\ &= -\frac{1}{2} g^{mn}_{,r} p_m p_n. \end{aligned}$$

Therefore the geodesic equation (6) on the cotangent bundle becomes

$$\left\{ \begin{array}{c} \partial_t p_i = -\frac{1}{2} g^{mn}_{,i} p_m p_n \\ \partial_t q^i = g^{ij} p_j. \end{array} \right. \tag{7}$$

It remains to check that the right-hand side coincides with the Hamiltonian vector field of H_g. In local canonical coordinates the Hamiltonian H_g is given by

$$H_g(q,p) = \frac{1}{2} g^{ij}(q) p_i p_j.$$

Its differential reads

$$dH_g = \frac{1}{2} g^{ij}_{,k} p_i p_j dq^k + g^{ij} p_j dp_i.$$

Hence the Hamiltonian vector field of H_g with respect to the symplectic form $\omega = dq^i \wedge dp_i$ equals

$$X_{H_g} = -\frac{1}{2} g^{ij}_{,k} p_i p_j \frac{\partial}{\partial p_k} + g^{ij} p_j \frac{\partial}{\partial q^i}.$$

In particular, the flow of X_{H_g} is given by the solution of (7). This finishes the proof of the Theorem. □

3.2. Mechanical Hamiltonians. Assume that (N, g) is a Riemannian manifold and $V \in C^\infty(N, \mathbb{R})$ is a smooth function on the configuration space referred to as the *potential*. We define the Hamiltonian

$$H_{g,V} \colon T^*N \to \mathbb{R}, \quad (q,p) \mapsto \frac{1}{2}|p|_g^2 + V(q),$$

i.e., in the language of physics the sum of kinetic and potential energy.

In the special case where N is an open subset of $\mathbb{R}^n$ endowed with its standard scalar product which we omit in the following from the notation the Hamiltonian vector field of H_V is given with respect to the splitting $T^*\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$

$$X_{H_V}(q,p) = \left(\begin{array}{c} p \\ -\nabla V(q) \end{array} \right).$$

where ∇V is the gradient of V, which in terms of physics can be thought of as a *force*. Given (q,p) in $T^*N \subset T^*\mathbb{R}^n$ if $q_p \colon \mathbb{R} \to N$ is a solution of the second order ODE

$$\partial_t^2 q_p(t) = -\nabla V(q_p(t))$$

meeting the initial conditions

$$q_p(0) = q, \quad \partial_t q_p(0) = p$$

then the Hamiltonian flow of H_V applies to (q,p) is given by

$$\phi^t_{H_V}(q,p) = (q_p(t), \partial_t q_p(t)).$$

We discuss three basic examples of mechanical Hamiltonians. The first example is the *harmonic oscillator*. Its Hamiltonian is given by

$$H \colon T^*\mathbb{R} \to \mathbb{R}, \quad (q,p) \mapsto \frac{1}{2}(p^2 + q^2).$$

Its flow is given by

$$\phi^t_H(q,p) = (q \cos t + p \sin t, -q \sin t + p \cos t).$$

Note that the flow of the harmonic oscillator is periodic of period 1.

Our second example is the case of *two uncoupled harmonic oscillators*. In this example the Hamiltonian is given by

$$H \colon T^*\mathbb{R}^2 \to \mathbb{R}, \quad (q,p) \mapsto \frac{1}{2}(p^2 + q^2) = \frac{1}{2}(p_1^2 + q_1^2) + \frac{1}{2}(p_2^2 + q_2^2).$$

Note that in this example the Hamiltonian flow is just the product flow of two harmonic oscillators with respect to the splitting $T^*\mathbb{R}^2 = T^*\mathbb{R} \times T^*\mathbb{R}$. In particular, the flow is again periodic of period 1. Moreover, note that if $c > 0$, then the level set or energy hypersurface of two uncoupled harmonic oscillators $H^{-1}(c)$ is a three dimensional sphere of radius $\sqrt{2c}$.

Our third example is the *Kepler problem*. In this case the Hamiltonian is given by

$$H\colon T^*(\mathbb{R}^n \setminus \{0\}) \to \mathbb{R}, \quad (q,p) \mapsto \frac{1}{2}p^2 - \frac{1}{|q|}.$$

The from the physical point of view most relevant cases are the case $n = 2$, the *planar Kepler problem* and the case $n = 3$. Using Moser regularization we will see that for negative energy the Hamiltonian flow of the Kepler problem in any dimension n can be embedded up to reparametrization in the geodesic flow of the round n-dimensional sphere. In the planar case the double cover of the geodesic flow on the round two dimensional sphere can be interpreted as the Hamiltonian flow of two uncoupled harmonic oscillators via Levi-Civita regularization.

3.3. Magnetic Hamiltonians. Mechanical Hamiltonians model physical systems where the force just depends on the position. There are however important forces which depend on the velocity as well. Examples are the Lorentz force in the presence of a magnetic field or the Coriolis force. To model such more general systems we twist the kinetic energy with a one form. The set-up is as follows. Assume that (N, g) is a Riemannian manifold, $V \in C^\infty(N, \mathbb{R})$ is a potential, and in addition $A \in \Omega^1(N)$ is a one form on N. We consider the Hamiltonian

$$H_{g,V,A}\colon T^*N \to \mathbb{R}, \quad (q,p) \mapsto \frac{1}{2}|p - A_q|^2 + V(q).$$

Because of its importance in the study of electromagnetism such Hamiltonians are referred to as magnetic Hamiltonians.

3.4. Physical symmetries. The last class of Hamiltonians is less directly associated to Hamiltonian systems arising in physical situations. However, their Hamiltonian flows generate a family of physical transformations which are important to study symmetries of Hamiltonian systems. Assume that N is a manifold and $X \in \Gamma(TN)$ is a vector field on N. We associate to the vector field X a Hamiltonian

$$H_X\colon T^*N \to N$$

as follows. Denote by $\pi\colon T^*N \to N$ the footpoint projection. A point $e \in T^*N$ we can interpret as a vector $e \in T^*_{\pi(e)}N$. We can pair this vector with the vector $X(\pi(e)) \in T_{\pi(e)}N$. Hence we set

$$H_X(e) := e\big(X(\pi(e))\big).$$

Assume that the flow $\phi_X^t\colon N \to N$ exists for every $t \in \mathbb{R}$. Then the Hamiltonian flow of H_X exists as well and is given by

$$\phi^t_{H_X} = d_*\phi^t_X \tag{8}$$

while the symplectomorphism $d_*\phi$ for a diffeomorphism ϕ was defined in (1).

A prominent example of such a Hamiltonian is *angular momentum*. Namely consider on $\mathbb{R}^2$ the vector field

$$X = q_1 \frac{\partial}{\partial q_2} - q_2 \frac{\partial}{\partial q_1} \in \Gamma(T\mathbb{R}^2).$$

Define angular momentum as the Hamiltonian

$$L \colon T^*\mathbb{R}^2 \to \mathbb{R}$$

given for $(q,p) \in T^*\mathbb{R}^2$ by

$$L(q,p) := H_X(q,p) = p_1 q_2 - p_2 q_1.$$

Note that the vector field X generates the clockwise rotation, i.e.,

$$\phi_X^t = R_{-t}$$

where

$$R_t \colon \mathbb{R}^2 \to \mathbb{R}^2$$

is the counterclockwise rotation

$$R_t\big(q_1, q_2\big) = \big((\cos t)q_1 - (\sin t)q_2, (\sin t)q_1 + (\cos t)q_2\big).$$

Therefore in view of (8) we obtain for the Hamiltonian flow of angular momentum

$$\phi_L^t = d_* R_{-t} \colon T^*\mathbb{R}^2 \to T^*\mathbb{R}^2. \tag{9}$$

4. Hamiltonian structures

A Hamiltonian manifold is the odd dimensional analogue of a symplectic manifold.

Definition 4.1. *A* Hamiltonian manifold *is a tuple (Σ, ω), where Σ is an odd dimensional manifold, and $\omega \in \Omega^2(\Sigma)$ is a closed two form with the property that $\ker \omega$ defines a one dimensional distribution in $T\Sigma$. The two form ω is called a* Hamiltonian structure *on Σ.*

Here is how Hamiltonian manifolds arise in nature. Suppose that (M, ω) is a symplectic manifold and $H \in C^\infty(M, \mathbb{R})$ is a Hamiltonian with the property that 0 is a regular value of H. Consider the energy hypersurface

$$\Sigma = H^{-1}(0) \subset M.$$

It follows that the tuple $(\Sigma, \omega|_\Sigma)$ is a Hamiltonian manifold. Moreover, it follows that in this example the one dimensional distribution $\ker \omega$ is given by

$$\ker \omega = \langle X_H|_\Sigma \rangle, \tag{10}$$

i.e., the line bundle spanned by the restriction of the Hamiltonian vector field to Σ. Indeed, note that if $x \in \Sigma$ and $\xi \in T_x\Sigma$ it holds that

$$\omega(X_H, \xi) = -dH(\xi) = 0$$

where the last equality follows since ξ is tangent to a level set of H. In particular, the leaves of the distribution $\ker \omega|_\Sigma$ correspond to the trajectories of the flow ϕ_H^t restricted to Σ. However, by studying the leaves of $\ker \omega|_\Sigma$ instead of the flow of $\phi_H^t|_\Sigma$ we lose the information about their parametrization. One can say that people studying Hamiltonian manifolds instead of Hamiltonian systems are quite relaxed, since they do not care about time. This is often an advantage. Indeed, by regularizing a Hamiltonian system one in general has to reparametrize the flow of the Hamiltonian. Therefore it is convenient to have a notion which remains invariant under these transformations.

Given a Hamiltonian manifold (Σ, ω) a *Hamiltonian vector field* is a non-vanishing section of the line bundle $\ker \omega$. Again if $\Sigma = H^{-1}(0)$ arises as the regular level set of a Hamiltonian function on a symplectic manifold the restriction of X_H to Σ is a Hamiltonian vector field. Note that if X is a Hamiltonian vector field on (Σ, ω) it follows from Cartan's formula that the Lie derivative $\mathcal{L}_X \omega = 0$, so that ω is preserved under the flow of X.

5. Contact forms

Definition 5.1. *Assume that (Σ, ω) is a Hamiltonian manifold of dimension $2n-1$. A* contact form *for (Σ, ω) is a one-form $\lambda \in \Omega^1(\Sigma)$ which meets the following two assumptions*

(i): $d\lambda = \omega$,

(ii): $\lambda \wedge \omega^{n-1}$ *is a volume form on Σ.*

Not every Hamiltonian manifold (Σ, ω) does admit a contact form. An obvious necessary condition for the existence of a contact form is that $[\omega] = 0 \in H^2_{dR}(\Sigma)$. The tuple (Σ, λ) is referred to as a *contact manifold*. The defining property for λ is then the assumption

$$\lambda \wedge (d\lambda)^{n-1} > 0,$$

i.e., $\lambda \wedge (d\lambda)^{n-1}$ is a volume form. Each contact manifold becomes a Hamiltonian manifold by setting $\omega = d\lambda$.

Given a contact manifold (Σ, λ) the Reeb vector field $R \in \Gamma(\Sigma)$ is implicitly defined by the conditions

$$\iota_R d\lambda = 0, \quad \lambda(R) = 1.$$

It follows that the Reeb vector field is a non-vanishing section in the line bundle $\ker d\lambda = \ker \omega \subset T\Sigma$. In particular it is a Hamiltonian vector field of the Hamiltonian manifold (Σ, ω) and we have

$$\ker \omega = \langle R \rangle. \tag{11}$$

If Σ arises as the level set $\Sigma = H^{-1}(0)$ of a Hamiltonian H on a symplectic manifold, it follows from (10) and (11) that the Reeb vector field and the restriction of the Hamiltonian vector field $X_H|_\Sigma$ are parallel. In particular, their flows coincide up to reparametrization.

On the contact manifold (Σ, λ) we can further define the hyperplane field

$$\xi := \ker \lambda \subset T\Sigma.$$

This leads to a splitting

$$T\Sigma = \xi \oplus \langle R \rangle.$$

Note that the restriction of $d\lambda$ to ξ makes ξ a symplectic vector bundle of degree $2n-2$ over Σ.

The hyperplane distribution ξ is referred to as the *contact structure*. While the contact structure ξ is determined by the contact form λ the opposite does not hold. Indeed, if $f > 0$ is any positive smooth function on Σ, we obtain a new contact form

$$\lambda_f := f\lambda \in \Omega^1(\Sigma).$$

That λ_f is indeed a contact form can be checked by computing

$$\lambda_f \wedge (d\lambda_f)^{n-1} = f^n(\lambda \wedge (d\lambda)^{n-1}) > 0.$$

The two contact forms λ and λ_f give rise to the same contact structure

$$\xi = \ker \lambda = \ker \lambda_f.$$

On the other hand the Reeb vector fields of λ and λ_f are in general not parallel to each other. Therefore the Reeb dynamics of the contact manifold (Σ, λ_f) might be quite different from the Reeb dynamics of the contact manifold (Σ, λ). This explains as well that we cannot recover the Hamiltonian structure $\omega = d\lambda$ from the contact structure $\xi = \ker \lambda$.

The study of contact manifolds is nowadays an interesting topic in its own right, see for example the book by Geiges [**40**]. Different than in contact topology our major concern is the Hamiltonian structure $\omega = d\lambda$, since this structure determines the dynamics up to reparametrization. From our point of view the contact form is more an auxiliary structure, which enables us to get information on the dynamics of the Hamiltonian manifold (Σ, ω). Indeed, the contact form turns out to be an indispensable tool in order to apply holomorphic curve techniques. Furthermore, we will see that with the help of contact forms one can rule out blue sky catastrophes.

6. Liouville domains and contact type hypersurfaces

Contact manifolds can sometimes be obtained as certain nice hypersurfaces in a symplectic manifold (M, ω). To be more precise, we consider a hypersurface $S \subset M$, and assume that there is a vector field X defined in a neighborhood of S satisfying

$$\mathcal{L}_X \omega = \omega. \tag{12}$$

In other words, the symplectic form ω is expanding along flow lines of X. So if ϕ_X^t denotes the flow of X, then

$$(\phi_X^t)^* \omega = e^t \omega. \tag{13}$$

We will call a vector field X satisfying (12) a *Liouville vector field*. If X is a Liouville vector field for ω, then we obtain a 1-form λ by $\lambda = \iota_X \omega$. This 1-form is called the Liouville form. We claim that

Proposition 6.1. *Suppose that X is a Liouville vector field defined on a neighborhood of a hypersurface $S \subset M$. Assume that X is transverse to S, so $T_x \partial M \oplus \langle X_x \rangle = T_x M$ for all $x \in M$. Then $(S, (\iota_X \omega)|_S)$ is a contact manifold with contact form $(\iota_X \omega)|_S$.*

Proof: To see this, abbreviate $\lambda = \iota_X \omega$. Given $x \in \Sigma$ choose a basis $\{v_1, \ldots, v_{2n-1}\}$ of $T_x \Sigma$. We compute

$$\begin{aligned} \lambda \wedge (d\lambda)^{n-1}(v_1, \ldots, v_{2n-1}) &= \iota_X \omega \wedge \omega^{n-1}(v_1, \ldots, v_{2n-1}) \\ &= \frac{1}{n} \omega^n (X_x, v_1, \ldots, v_{2n-1}). \end{aligned}$$

Because $X \pitchfork \Sigma$ it follows that $\{X_x, v_1, \ldots, v_{2n-1}\}$ is a basis of $T_x M$. Because ω is non-degenerate it follows that

$$\omega^n (X_x, v_1, \ldots, v_{2n-1}\} \neq 0$$

and we see that $\lambda|_\Sigma$ is indeed a contact form on Σ. $\square$

A hypersurface Σ satisfying the assumptions of the proposition is called a *contact type hypersurface*. We will now define a class of symplectic manifolds that come equipped with a contact type hypersurface. To explain this notion we assume that (M, λ) is an exact symplectic manifold, i.e., $\omega = d\lambda$ is a symplectic structure on M. An exact symplectic manifold cannot be closed, since for a closed symplectic manifold $[\omega] \neq 0 \in H^2_{dR}(M)$. We assume that M is compact, so it must have a non-empty boundary.

Definition 6.2. *A* Liouville domain *is a compact, exact symplectic manifold (M, λ) with the property that the Liouville vector field X, defined by $\iota_X d\lambda = \lambda$, is transverse to the boundary and outward pointing.*

By Proposition 6.1 it follows that the boundary of a Liouville domain (M, λ) is contact with contact form $\lambda|_{\partial M}$, so we have.

Lemma 6.3. *Assume that (M, λ) is a Liouville domain. Then $(\partial M, \lambda|_{\partial M})$ is a contact manifold.*

Given a Liouville domain (M, λ) with Liouville vector field X, we claim that

$$\mathcal{L}_X \lambda = \lambda \tag{14}$$

where $\mathcal{L}_X$ denotes the Lie derivative in direction of X. To see this we first observe that

$$\iota_X \lambda = \omega(X, X) = 0$$

by the antisymmetry of the form ω. Hence we compute using Cartan's formula

$$\mathcal{L}_X \lambda = d\iota_X \lambda + \iota_X d\lambda = \iota_X \omega = \lambda.$$

This proves (14). We give two examples of Liouville domains which play an important role in the following.

Example 6.4. *The cotangent bundle $M = T^*N$ together with the Liouville one-form is an example of an exact symplectic manifold. In local canonical coordinates the Liouville one-form is given by $\lambda = \sum p_i dq_i$ and therefore the associated Liouville vector field reads*

$$X = \sum_{i=1}^{n} p_i \frac{\partial}{\partial p_i}.$$

Now suppose that $\Sigma \subset T^*N$ is fiberwise star-shaped, *i.e., for every $x \in N$ $\Sigma \cap T^*_x N$ bounds a star-shaped domain D_x in the vector space $T^*_x N$. Then $X \pitchfork \Sigma$ and $D = \bigcup_{x \in N} D_x$ is a Liouville domain with $\partial D = \Sigma$.*

Remark 6.5. *If $H \colon T^*N \to \mathbb{R}$ is a mechanical Hamiltonian, i.e., $H(q, p) = \frac{1}{2}|p|^2_g + V(q)$ for a Riemannian metric g on N and a smooth potential $V \colon N \to \mathbb{R}$ and $c > \max V$ Then the energy hypersurface $\Sigma = H^{-1}(c)$ is fiberwise star-shaped. Indeed,*

$$dH(X)(q, p) = |p|^2_g$$

and because $c > \max V$ it follows that p does not vanish on Σ so that we get

$$dH(X)|_\Sigma > 0.$$

Example 6.6. *On the complex vector space $\mathbb{C}^n$ with coordinates $(z_1, \dots, z_n) = (x_1 + iy_1, \dots, x_n + iy_n)$ consider the one-form*

$$\lambda = \frac{1}{2} \sum_{i=1}^{n} (x_i dy_i - y_i dx_i).$$

The pair $(\mathbb{C}^n, \lambda)$ is a (linear) symplectic manifold with symplectic form

$$\omega = \sum_{i=1}^{n} dx_i \wedge dy_i$$

and Liouville vector field

$$X = \frac{1}{2} \sum_{i=1}^{n} \left(x_i \frac{\partial}{\partial x_i} + y_i \frac{\partial}{\partial y_i} \right).$$

If $\Sigma \subset \mathbb{C}^n$ is star-shaped*, i.e., Σ bounds a star-shaped domain D, then $X \pitchfork \Sigma$ and D is a Liouville domain with $\partial D = \Sigma$.*

7. Real Liouville domains and real contact manifolds

A *real Liouville domain* is a triple (M, λ, ρ) where (M, λ) is a Liouville domain and $\rho \in \mathrm{Diff}(M)$ is an exact anti-symplectic involution, i.e.,

$$\rho^2 = \mathrm{id}|_M, \quad \rho^* \lambda = -\lambda.$$

Because the exterior derivative commutes with pullback we immediately obtain that

$$\rho^* \omega = -\omega$$

for the symplectic form $\omega = d\lambda$, i.e., ρ is an anti-symplectic involution. It follows that the Liouville vector field X defined by $\iota_X \omega = \lambda$ is invariant under ρ, meaning

$$\rho^* X = X.$$

The fixed point set $\mathrm{Fix}(\rho)$ of an anti-symplectic involution is a (maybe empty) Lagrangian submanifold. To see that pick $x \in \mathrm{Fix}(\rho)$. The differential

$$d\rho(x) \colon T_x M \to T_x M$$

is then a linear involution. Therefore the vector space decomposes

$$T_x M = \ker(d\rho(x) - \mathrm{id}) \oplus \ker(d\rho(x) + \mathrm{id})$$

into the eigenspaces of $d\rho(x)$ to the eigenvalues ± 1. Note that

$$T_x \mathrm{Fix}(\rho) = \ker(d\rho(x) - \mathrm{id})$$

the eigenspace to the eigenvalue 1. Because ρ is anti-symplectic both eigenspaces are isotropic subspaces of the symplectic vector space $T_x M$, i.e., ω vanishes on both of them. Hence by dimensional reasons they have to be Lagrangian, i.e., isotropic subspaces of the maximal possible dimension, namely half the dimension of M. This proves that $\mathrm{Fix}(\rho)$ is Lagrangian. If ρ is an exact symplectic involution then in addition the restriction of λ to $\mathrm{Fix}(\rho)$ vanishes as well, so that $\mathrm{Fix}(\rho)$ becomes an *exact Lagrangian submanifold*.

Example 7.1. *On $\mathbb{C}^n$ complex conjugation is an involution under which the exact contact form $\lambda = \frac{1}{2}\sum_{i=1}^n (x_i dy_i - y_i dx_i)$ is anti-invariant. Therefore if $\Sigma \subset \mathbb{C}^n$ is a star-shaped hypersurface invariant under complex conjugation then Σ bounds a real Liouville domain. Note that the fixed point set of complex conjugation is the Lagrangian $\mathbb{R}^n \subset \mathbb{C}^n$. This explains the terminology real Liouville domain.*

Example 7.2. *Let N be a closed manifold and $I \in \mathrm{Diff}(N)$ a smooth involution, i.e., $I^2 = \mathrm{id}|_N$. By (2) the physical transformation $d_*I \colon T^*N \to T^*N$ is an exact symplectic involution on T^*N, namely*

$$(d_*I)^*\lambda = \lambda$$

for the Liouville one-form λ. Consider further the involution $rho_1 \colon T^*N \to T^*N$ which on every fiber restricts to minus the identity so that in canonical coordinates we have

$$\rho_1(q,p) = (q,-p).$$

The involution $\mathcal{I}$ is an exact anti-symplectic involution which commutes with the exact symplectic involution

$$\rho_1 \circ d_*I = d_*I \circ \rho_1 =: \rho_I$$

so that ρ_I is an exact anti-symplectic involution on T^*N. Note that

$$\mathrm{Fix}(\rho_I) = N^*\mathrm{Fix}(I)$$

the conormal bundle of the fixed point set of I. If $\Sigma \subset T^*N$ is a fiberwise star-shaped hypersurface invariant under ρ_I, then Σ bounds a real Liouville domain. In particular, if we take I the identity on N, then the anti-symplectic involution is just $\rho_{\mathrm{id}} = \rho_1$ and therefore the requirement is that Σ is fiberwise symmetric star-shaped. Note that a mechanical Hamiltonian is invariant under the involution ρ_1, so that according to Remark 6.5 the energy hypersurface of a mechanical Hamiltonian for energies higher than the maximum of the potential bounds a real Liouville domain.

Now assume that (M,λ,ρ) is a real Liouville domain with boundary $\Sigma = \partial M$. Denote by abuse of notation the restrictions of λ and ρ to Σ by the same letter. Then the triple (Σ,λ,ρ) is a *real contact manifold*, namely a contact manifold (Σ,λ) together with an anti-contact involution $\rho \in \mathrm{Diff}(\Sigma)$, namely an involution satisfying $\rho^*\lambda = -\lambda$. The Reeb vector field $R \in \Gamma(T\Sigma)$ is then anti-invariant under ρ, i.e.,

$$\rho^*R = -R.$$

The fixed point set of an anti-contact involution is a (maybe empty) *Legendrian submanifold* of Σ, namely a submanifold whose tangent space is a Lagrangian subbundle of the symplectic vector bundle $\xi = \ker(\lambda)$. Moreover, the flow $\phi_R^t \colon \Sigma \to \Sigma$ for $t \in \mathbb{R}$ of the Reeb vector field satisfies with ρ the relation

$$\phi_R^t = \rho \circ \phi_R^{-t} \circ \rho. \tag{15}$$

In order to see this, we compute

$$\phi_R^t = \phi_{-R}^{-t} = \phi_{\rho^*R}^{-t} = \rho^{-1} \circ \phi_R^{-t} \circ \rho = \rho \circ \phi_R^{-t} \circ \rho$$

where for the last equality we have used that ρ is an involution.

CHAPTER 2

Symmetries and Noether's theorem

1. Poisson brackets

In order to simplify differential equations, it is important to identify *preserved quantities*, also called integrals. More formally, if X is a vector field on a manifold M, then we call L an *integral* of X if $X(L) = 0$.

The notion of Poisson bracket will be helpful. For a symplectic manifold (M, ω) we define the *Poisson bracket* of smooth functions F and G by

$$\{F, G\} := \omega(X_F, X_G) = -dF(X_G) = -X_G(F) = X_F(G). \tag{16}$$

We see directly from the definition that the Poisson bracket describes the time-evolution of a function. Indeed, suppose that $\gamma(t)$ is a flow line of X_F. Then

$$\frac{dG \circ \gamma(t)}{dt} = X_F(G) = \{F, G\}.$$

From this energy preservation, see Theorem 2.2, follows because $\{H, H\} = 0$ (the Poisson bracket is alternating). Before we turn our attention to conserved quantities, we first need to establish some properties of the Poisson bracket.

Lemma 1.1. *Given smooth functions F, G on a symplectic manifold (M, ω), we have the following relation between the Lie bracket and Poisson bracket,*

$$[X_F, X_G] = X_{\{F,G\}}.$$

Proof: We first rewrite the Lie bracket a bit:

$$[X_F, X_G] = \mathcal{L}_{X_F} X_G = \frac{d}{dt}|_{t=0} {Fl_t^{X_F}}^* X_G = \frac{d}{dt}|_{t=0} X_{G \circ Fl_t^{X_F}}.$$

Now use this identity and the definition:

$$\begin{aligned} i_{[X_F, X_G]}\omega &= \frac{d}{dt}|_{t=0} \omega(X_{G \circ Fl_t^{X_F}}, \cdot) \\ &= \frac{d}{dt}|_{t=0} \left(-d(G \circ Fl_t^{X_F} \right) \\ &= -d \left(\frac{d}{dt}|_{t=0} G \circ Fl_t^{X_F} \right) = -d\left(X_F(G)\right) = -d\{F, G\}. \end{aligned}$$

$\square$

In a Darboux chart $(U, \omega = dp \wedge dq)$ for (M, ω), the Poisson bracket can be written as

$$\{F, G\} = \sum_i \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q^i} - \frac{\partial F}{\partial q^i} \frac{\partial G}{\partial p_i}.$$

In a moment, we shall see that the Poisson bracket endows the space of smooth functions on M with a Lie algebra structure. We briefly recall the definition.

Definition 1.2. *A* **Lie algebra** *consists of a vector space* $\mathfrak{g}$ *together with a binary operation* $[\cdot,\cdot]$ *such that is bilinear, alternating, so* $[v,v]=0$ *for all* $v\in\mathfrak{g}$ *and satisfies the Jacobi identity, i.e. for all* $X,Y,Z\in\mathfrak{g}$*, we have*

$$[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=0.$$

Proposition 1.3. *The pair* $(C^\infty(M),\{\cdot,\cdot\})$ *is a Lie algebra.*

Proof: We check the required properties. First of all, note that for $a\in\mathbb{R}$ and $F,G\in C^\infty(M)$, we have $X_{aF+G}=aX_F+X_G$, since the Hamiltonian vector field is the solution to a linear equation. Hence, $\{aF+G,H\}=\omega(X_{aF+G},X_H)=a\omega(X_F,X_H)+\omega(X_G,X_H)=a\{F,H\}+\{G,H\}$. The same argument works for the second factor, so $\{\cdot,\cdot\}$ is bilinear. Also, $\{F,F\}=\omega(X_F,X_F)=0$, so $\{\cdot,\cdot\}$ is alternating. Alternatively, we can also use Lemma 1.1. Finally, we check that the Jacobi identity by computing the individual terms:

$$\begin{aligned}\{F,\{G,H\}\}&=X_F(\{G,H\})=X_F(X_G(H))\\ \{G,\{H,F\}\}&=-X_G(\{F,H\})=-X_G(X_F(H))\\ \{H,\{F,G\}\}&=-X_{\{F,G\}}(H)=-[X_F,X_G](H)\end{aligned}$$

We have used Lemma 1.1 in the last step. Summing these terms shows that the Jacobi identity holds. □

Lemma 1.4. *The function* G *is an integral of* X_F *if* $\{F,G\}=0$.

Proof: The function G is an integral if and only if $X_F(G)=0$. This holds if and only if $0=-dF(X_G)=\omega(X_F,X_G)=\{F,G\}$. □

Remark 1.5. *By Lemma 1.1* $\{F,G\}=0$ *implies that* $[X_F,X_G]=0$*. On the other hand, for* G *to be an integral of* X_F*, it is not enough to just have* $[X_F,X_G]=0$*. Indeed, consider* $(\mathbb{R}^2,\omega_0=dp\wedge dq)$ *with the Hamiltonians* $F=p$ *and* $G=q$*. Then* $X_F=\partial_q$ *and* $X_G=-\partial_p$*, so* $[X_F,X_G]=0$*. However,* G *is linearly increasing under the flow of* X_F*, so* G *is not an integral of* X_F*.*

1.1. Noether's theorem. Consider a smooth action of a Lie group G on a symplectic manifold (M,ω). Take a vector $\xi\in\mathfrak{g}$. Then we get a path $\exp(t\xi)$ in G and act on M with this path. Take the derivative with respect to t to get a vector field on M, namely

$$X_\xi(x):=\frac{d}{dt}|_{t=0}(exp(t\xi)\cdot x). \tag{17}$$

If we assume that the action preserves the symplectic form ω, then X_ξ is a symplectic vector field.

Remark 1.6. *Note that* X_ξ *does not need to be a Hamiltonian vector field. For instance the vector field* ∂_θ *on* $(T^2,d\theta\wedge d\phi)$ *is symplectic, yet not Hamiltonian.*

We will now assume that X_ξ is Hamiltonian, so there for each X_ξ there is a function H_ξ satisfying $i_{X_\xi}\omega=dH_\xi$. Then we get a map

$$\begin{aligned}\rho:\mathfrak{g}&\longrightarrow C^\infty(M)\\ \xi&\longmapsto H_\xi.\end{aligned}$$

This map has no reason to be nice. We haven't said anything about shifting the H_ξ's by constants. Instead, we make the following definition.

Definition 1.7. *We say the action is* **Hamiltonian** *if the map ρ can be chosen to be a Lie algebra homomorphism.*

There are obstructions for actions to be Hamiltonian, but this will not be relevant here. Instead, we point out a couple of examples

Example 1.8. *We act with an $SO(n)$-matrix A on $T^*\mathbb{R}^n$ by the formula $A \cdot (q,p) = (Aq, Ap)$, or in other words, the standard action on each component. This action preserves the symplectic form.*

The Lie algebra $\mathfrak{so}(n)$ can be identified with skew-symmetric matrices together with the usual commutator as Lie bracket.

Given a skew-symmetric $n \times n$-matrix L, we define the Hamiltonian

$$H_L(q,p) := p^t Lq.$$

The Hamiltonian equations are

$$\dot{p} = -\partial_q H_L = -p^t L = -L^t p = Lp \quad \dot{q} = \partial_p H_L = Lq,$$

which we can solve by exponentiating, so $(q(t), p(t)\,) = (e^{Lt}q, e^{Lt}p)$. This is precisely the $SO(n)$-action on $T^*\mathbb{R}^n$.

Theorem 1.9 (Hamiltonian version of Noether's theorem). *Suppose that G is a Lie group acting Hamiltonianly on a symplectic manifold (M, ω). If $H : M \to \mathbb{R}$ is a Hamiltonian that is invariant under G, then each $\xi \in \mathfrak{g}$ gives an integral H_ξ of X_H, or equivalently $\{H, H_\xi\} = 0$.*

Proof: Take $\xi \in \mathfrak{g}$. We get a vector field X_ξ on M by formula (17), which is the Hamiltonian vector field of the function H_ξ by the assumption of Hamiltonian action. The Hamiltonian H is assumed to be invariant, so by the formula for the Poisson bracket we have

$$0 = X_{H_\xi}(H) = -X_H(H_\xi),$$

so H_ξ is an integral of X_H. □

1.2. Harmonic oscillator. Consider the Hamiltonian

$$H = \frac{1}{2}p^2 + \frac{a^2}{2}q^2$$

on $(T^*\mathbb{R}, d\lambda_c) \cong (\mathbb{R}^2, \omega_0)$. The Hamiltonian flow is hence the linear ODE

$$\begin{aligned} \dot{p} &= -a^2 q \\ \dot{q} &= p \\ p(0) = p_0 q(0) &= q_0, \end{aligned}$$

so we find the explicit solution

$$(p(t), q(t)\,) = (p_0 \cos(at) - a \cdot q_0 \sin(at), \frac{p_0}{a}\sin(at) + q_0 \cos(at).$$

If we don't care about the parametrization, we can just observe that level sets are ellipses, and that solutions lie on them. Later on, we will see several situations where we can say where the solutions are, but not how they are parametrized. The above Hamiltonian is called the *harmonic oscillator*, and it is one of the most basic Hamiltonian systems.

Clearly, all solutions are periodic in this case: the period is given by $2\pi/a$

1.2.1. *Several uncoupled oscillators.* We can generalize the above example to $(\mathbb{R}^{2n}, \omega_0)$. Define the Hamiltonian

$$H = \sum_{j=1}^{n} \frac{1}{2} p_j^2 + \frac{a_j^2}{2} q_j^2$$

This Hamiltonian describes n uncoupled oscillators. We have the ODE

$$\begin{aligned} \dot{p}_j &= -a_j^2 q_j \\ \dot{q}_j &= p_j, \end{aligned}$$

which can clearly also be explicitly solved: just use the above formula n times.

In this case, not all solutions are necessarily periodic. For instance, assume that the frequencies a_j are rationally independent, meaning there are no non-zero rational numbers n_i such that $\sum n_j a_j = 0$. Then we see that the j-th coordinate satisfies

$$(p_j(0)\cos(a_j t) - a_j \cdot q_j(0)\sin(a_j t), \ldots \frac{p_j(0)}{a_j}\sin(a_j t) + q_0\cos(a_j t).$$

Such a solution may not be periodic: the frequencies in the different coordinates differ.

1.3. Central force: conservation of angular momentum. Suppose we are given a Hamiltonian dynamical system H on $T^*\mathbb{R}^3$. Define the **angular momentum** by

$$L := q \times p.$$

Consider the Hamiltonian

$$H = \frac{1}{2}|p|^2 + V(\|q\|) \tag{18}$$

on $\mathbb{R}^{2n} - \{0\} \times \mathbb{R}^{2n}$, where $V : \mathbb{R} \to \mathbb{R}$ is (smooth) function, possibly with some singularities. Such a function V is called the potential for a **central force**, because it only depends on the distance.

We will assume that $n = 3$, although this can be generalized.

Lemma 1.10. *The angular momentum is preserved under the flow of X_H. In other words, the components of the angular momentum $L = (L_1, L_2, L_3)$ satisfy $\{H, L_i\} = 0$.*

Proof: By Example 1.8, the standard $SO(n)$ action acts Hamiltonianly on $T^*\mathbb{R}^n$. The Hamiltonian for a central force is $SO(n)$-invariant, so Noether's theorem, Theorem 1.9, implies the claim. □

Remark 1.11. *The physical interpretation of preservation of angular momentum is that flow lines of the Hamiltonian vector field X_H lie in the plane with normal vector L.*

1.4. The Kepler problem and its integrals. We shall consider the Hamiltonian

$$H = \frac{1}{2}|p|^2 - \frac{1}{|q|} \tag{19}$$

on $\mathbb{R}^{2n} - \{0\} \times \mathbb{R}^{2n}$ with coordinates (q, p) and symplectic form $\omega = dp \wedge dq$. The physically relevant case is $n = 2, 3$, and we shall first consider the case $n = 3$. The equations of motion are

$$\dot{p} = -\frac{q}{|q|^3}$$
$$\dot{q} = p.$$

In other words, the force equals $\ddot{q} = -\frac{q}{|q|^3}$, so its strength drops of with the distance squared.

The strategy is to find as many integrals as possible, and in fact the Kepler problem turns *completely integrable*. We will define this notion later, but roughly we can say that this means that there are so many conserved quantities that we can convert the ODE's into algebraic equations.

Lemma 1.12. *The angular momentum L is an integral of the Kepler problem.*

The Kepler problem has an obvious $SO(3)$-symmetry, or put differently the force is central, so Lemma 1.10 applies.

Remark 1.13. *In higher dimensions this Hamiltonian has a $SO(n)$-symmetry, but we will not consider this more general (and unphysical) situation.*

1.5. The Runge-Lenz vector: another integral of the Kepler problem. The following integral depends on the specific form of Kepler Hamiltonian (19). Define the **Laplace-Runge-Lenz vector** (also called **Runge-Lenz vector**)[1] by

$$A := p \times L - \frac{q}{|q|}.$$

Lemma 1.14. *The Runge-Lenz vector A is preserved under the flow of X_H. In other words, the components of $A = (A_1, A_2, A_3)$ satisfy $\{H, A_i\} = 0$.*

Unlike the preservation of angular momentum, this integral is not obvious from a symmetry of the phase space. We will prove that A is an integral with a short computation.

Proof. We compute the time-derivative of A,

$$\begin{aligned}
\dot{A} &= \dot{p} \times L + p \times \dot{L} - \frac{\dot{q}}{\|q\|} + \frac{q}{\|q\|^2}\frac{q \cdot \dot{q}}{\|q\|} \\
&= -\frac{q}{\|q\|^3} \times (q \times p) - \frac{p}{\|q\|} + \frac{q}{\|q\|^3}(q \cdot p) \\
&= \frac{1}{\|q\|^3}\left(-q \times (q \times p) - (q \cdot q)p + (q \cdot p)q\right) = 0.
\end{aligned}$$

In the second step we have used the Hamilton equations, and in the last step we used the vector product identity

$$(u \times v) \times w = (u \cdot w)v - (v \cdot w)u.$$

□

Lemma 1.15. *The Runge-Lenz vector satisfies the identity*

$$\|A\|^2 = 1 + 2H \cdot \|L\|^2.$$

[1] This integral was discovered by Jakob Hermann (1678-1733). We refer to [**42, 43**] for the history of this vector.

Proof: The following computation makes use of the fact that p and L are orthogonal and the identity $q \cdot p \times L = \det(q, p, L) = q \times p \cdot L$. We find

$$\begin{aligned}\|A\|^2 &= \|p \times L\|^2 - \frac{2}{|q|} q \cdot p \times L + \frac{\|q\|^2}{\|q\|^2} = 1 + \|p\|^2 \|L\|^2 - \frac{2}{|q|}\|L\|^2 \\ &= 1 + 2(\frac{1}{2}\|p\|^2 - \frac{1}{\|q\|})\|L\|^2.\end{aligned}$$

□

1.5.1. *Solving the Kepler problem.* Define the plane $P_L = \{v \in \mathbb{R}^3 \mid \langle L, v\rangle = 0\}$.

Lemma 1.16. *The vector A lies in the plane P_L.*

Recall that $\langle q, L\rangle = 0$ and observe that

$$\langle A, L\rangle = \langle p \times L, L\rangle - \langle \frac{q}{\|q\|}, L\rangle = 0 + 0.$$

To describe the movement of the particle more explicitly, we apply a coordinates change, namely a rotation to move the L-vector to the z-axis. Then $L = (0, 0, \ell)$ for some $\ell > 0$, and hence we can write

$$A = (\|A\| \cos g, \|A\| \sin g, 0).$$

Definition 1.17. *The angle g is called the* **argument of the perigee (perihelion)**[2].

We now determine the radius as function of the angle ϕ. Using the above formula for A and the identity $\langle p \times L, q\rangle = \det(p, L, q)$, we find

$$\|q\| + \langle A, q\rangle = \langle \frac{q}{|q|}, q\rangle + \langle A, q\rangle = \langle p \times L, q\rangle = \det(p, L, q) = \langle q \times p, L\rangle = \|L\|^2.$$

As before, we write q in polar coordinates

$$q = (r \cos \phi, r \sin \phi, 0)$$

and by plugging this into $\|q\| + \langle A, q\rangle = \|L\|^2$, we find

$$r = \frac{\|L\|^2}{1 + \|A\| \cos(\phi - g)}. \tag{20}$$

It is common to call the quantity

$$f := \phi - g$$

the **true anomaly**, and $\|A\|$ is called the *eccentricity.* The geometric picture is indicated in Figure 1.5.1. It is clear that from Equation 20 that the argument of the perigee is the angle of the closest approach in a typical situation (here this means $L \neq 0$).

With the above computations, we can deduce the following classification result for solutions.

[2]Perigee means close to the Earth. Perihelion means close to the Sun. If the heavy mass describes the Earth, one uses perigee, if it is the Sun, one uses the word perihelion.

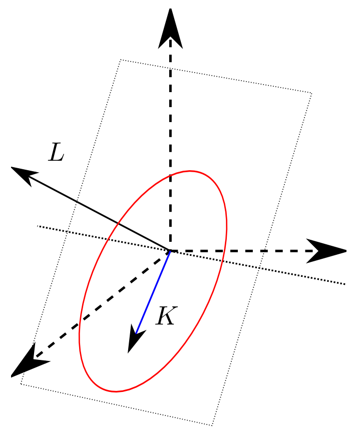

FIGURE 1. A sketch of a Kepler orbit

2. The planar Kepler problem

The above discussion describes the spatial Kepler problem, so the planar problem follows as a special case. For later computations, it will be helpful to develop some explicit formulas for the planar problem.

As before, we can assume that $L = (0, 0, \ell)$. In polar coordinates for q^1q^2-plane, there is a nice expression of ℓ. We write $(q^1, q^2, q^3) = (r\cos\phi, r\sin\phi, z)$. The momentum coordinates (p_x, p_y, p_z) transform with the inverse of the Jacobian, so if we denote the cotangent coordinates dual to (r, ϕ, z) by (p_r, p_ϕ, p_z), then we find

$$(p_x, p_y, p_z) = (\cos\phi \cdot p_r - \frac{\sin\phi}{r} p_\phi, \sin\phi \cdot p_r + \frac{\cos\phi}{r} p_\phi, p_z).$$

The coordinate change for the q-coordinates is

$$q_x = q_r \cos(q_\phi), \quad q_y = q_r \sin(q_\phi).$$

We are looking for a symplectic transformation, so, as mentioned, we just need the inverse and transpose of the Jacobian of this coordinate transformation for the p-part. It is, however, convenient to compute by using the fact that the canonical 1-form, $\lambda = p_x dq_x + p_y dq_y$ is preserved. This gives the equation $p_x dq_x + p_y dq_y = p_r dq_r + p_\phi dq_\phi$, so we find $p_x = p_r \cos q_\phi - \frac{p_\phi}{q_r} \sin q_\phi$ and $p_y = p_r \sin q_\phi + \frac{p_\phi}{q_r} \cos q_\phi$. The Hamiltonian is cylindrical coordinates is hence

$$H = \frac{1}{2}\left(p_r^2 + \frac{p_\phi^2}{q_r^2} \right) - \frac{1}{q_r}.$$

Remark 2.1. *It is important to observe that p_ϕ is the angular momentum. The Hamiltonian equations clearly show that the angular momentum is preserved as we expect for a central force.*

Plug this into $\ell = q^1 p_2 - q^2 p_1$ and use the Hamilton equations $\dot{q}^i = p_i$ to find.

Lemma 2.2 (Kepler's second law). *We have*

$$\frac{1}{2} r^2 \dot{\phi} = \frac{1}{2}\ell = \frac{dArea}{dt}.$$

where Area is the area swept out by an ellipse.

To get the claim about the area, just use that

$$A = \int_{\phi=\phi_1}^{\phi_2} \int_{r=0}^{r=r_1} r d\phi = \int_{\phi=\phi_1}^{\phi_2} \frac{1}{2} r^2 d\phi.$$

We assume that $L \neq 0$. The case $L = 0$ can be worked out separately: it involves collision orbits.

Recall that the Hamiltonian for the planar Kepler problem is given by

$$E\colon T^*(\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}, \quad (q,p) \mapsto \frac{1}{2}p^2 - \frac{1}{|q|}. \tag{21}$$

We prefer to abbreviate the Hamiltonian for the Kepler problem by E, as an abbreviation for energy and not by H as usual. This is because we also have to study the Kepler problem in rotating coordinates, the so called rotating Kepler problem, and we prefer to save the letter H to denote the Hamiltonian in rotating coordinates.

The Kepler problem is rotationally invariant. Because rotation is generated by angular momentum we obtain by Noether's theorem

$$\{E, L\} = 0, \tag{22}$$

a formula, the reader is invited to check as well by direct computation. However, the Kepler problem admits as well some "hidden symmetries". These hidden symmetries do not arise from flows on the configuration space $\mathbb{R}^2 \setminus \{0\}$ but from flows which only live on phase space $T^*(\mathbb{R}^2 \setminus \{0\})$. We introduce the smooth functions

$$A_1, A_2\colon T^*(\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}$$

given by

$$\begin{cases} A_1(q,p) = p_2(p_2 q_1 - p_1 q_2) - \frac{q_1}{|q|} = p_2 L(q,p) - \frac{q_1}{|q|}, \\ A_2(q,p) = -p_1(p_2 q_1 - p_1 q_2) - \frac{q_2}{|q|} = -p_1 L(q,p) - \frac{q_2}{|q|}. \end{cases}$$

Lemma 2.3. *The Poisson bracket of E with A_1 and A_2 vanishes, i.e.,*

$$\{E, A_1\} = \{E, A_2\} = 0.$$

Proof: The Hamiltonian vector field of E is given by

$$X_E = p_1 \frac{\partial}{\partial q_1} + p_2 \frac{\partial}{\partial q_2} - \frac{q_1}{|q|^3} \frac{\partial}{\partial p_1} - \frac{q_2}{|q|^3} \frac{\partial}{\partial p_2}.$$

The differential of A_1 is

$$dA_1 = L dp_2 + p_2 dL - \frac{q_2^2}{|q|^3} dq_1 + \frac{q_1 q_2}{|q|^3} dq_2.$$

Using (22) we compute

$$\{A_1, E\} = dA_1(X_E) = -\frac{q_2 L}{|q|^3} - \frac{q_2^2 p_1}{|q|^3} + \frac{q_1 q_2 p_2}{|q|^3} = 0.$$

This proves that $\{E, A_1\} = -\{A_1, E\} = 0$ and that $\{E, A_2\} = 0$ is shown by a similar computation. □

The two integrals A_1 and A_2 of the Kepler problem give rise to the vector

$$A := (A_1, A_2)\colon T^*(\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}^2. \tag{23}$$

This vector is known as the *Runge-Lenz vector.*

The Runge-Lenz vector can be used to give an elegant explanation why trajectories of the Kepler flow follow conic sections. This is due to the following formula

$$\langle A, q\rangle = L^2 - |q| \tag{24}$$

which follows immediately from the definition of A_1 and A_2. If $A = 0$ this means that q lies on a circle of radius L^2 with center in the origin. If $A \neq 0$ abbreviate

$$r := |q|, \quad \theta = \arccos \frac{\langle A, q\rangle}{r|A|},$$

i.e., θ is the angle between A and q. With this notation (24) becomes

$$r\big(|A|\cos\theta + 1\big) = L^2.$$

In particular, the length of the Runge-Lenz vector is the *eccentricity* of the conic section the Kepler trajectory follows.

The eccentricity can be directly computed in terms of E and L due to the following formula

$$A^2 = 1 + 2EL^2. \tag{25}$$

In order to prove (25) we compute

$$\begin{aligned}
A^2 &= A_1^2 + A_2^2 \\
&= \left(p_2 L - \frac{q_1}{|q|}\right)^2 + \left(p_1 L + \frac{q_2}{|q|}\right)^2 \\
&= p^2L^2 - \frac{2(p_2q_1 - p_1q_2)L}{|q|} + \frac{q^2}{q^2} \\
&= 2L^2\left(\frac{p^2}{2} - \frac{1}{|q|}\right) + 1 \\
&= 2EL^2 + 1.
\end{aligned}$$

Formula (25) gives rise to the following inequality for energy and angular momentum in the Kepler problem

$$2EL^2 + 1 \geq 0. \tag{26}$$

Moreover, equality holds if and only it the trajectory lies on a circle.

CHAPTER 3

Regularization of two body collisions

1. Moser regularization

As we have seen in the discussion about the planar Kepler problem, the Kepler problem admits an obvious rotational symmetry, namely under rotations of the plane, $SO(2)$, but also hidden symmetries which are generated by the Runge-Lenz vector (23). These hidden symmetries played an important role in the early development of quantum mechanics, in particular in Pauli's and Fock's discussion about the spectrum of the hydrogen atom [**35, 92**]. In [**86**] Moser explained how the Kepler flow can be embedded into the geodesic flow of the round sphere. This explains the hidden symmetries because the round metric is invariant under the group $SO(3)$. In the case of the planar Kepler problem we obtain the geodesic flow on the two dimensional sphere and the symmetry group becomes $SO(3)$. In particular, the symmetry group is three dimensional, in accordance with the fact that we have three integrals, namely angular momentum as well as the two components of the Runge-Lenz vector. We refer to the works of Hulthén [**61**] and Bargmann [**11**] for a discussion of these symmetries in terms of quantum mechanics. See also the exposition [**71**] by Kim which explains the relation of the Runge-Lenz vector with the moment map of the Hamiltonian action of $SO(3)$ on the cotangent bundle of the two dimensional sphere.

We first explain Moser's regularization of the Kepler problem at the energy value $-\frac{1}{2}$. Let $E\colon T^*(\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}$ be the Kepler Hamiltonian (21), i.e., the map $(q,p) \mapsto \frac{1}{2}p^2 - \frac{1}{|q|}$. Define

$$K(p,q) = \frac{1}{2}\bigg(|q|\Big(E(-q,p) + \frac{1}{2}\Big) + 1\bigg)^2 = \frac{1}{2}\Big(\frac{1}{2}(p^2+1)|q|\Big)^2.$$

Recall from Section 2.2 the switch map which is symplectic and interchanges the roles of position and momentum. If we do this acrobatics in our mind, the Hamiltonian K is just the kinetic energy of the "momentum" q with respect to the round metric on S^2 in the chart obtained by stereographic projection. Hence by Theorem 3.1 the flow of the Hamiltonian vector field of K is given by the geodesic flow of the round two sphere in the chart obtained by stereographic projection. Note that

$$dK|_{E^{-1}(-\frac{1}{2})}(p,q) = |q| dE|_{E^{-1}(-\frac{1}{2})}(-q,p).$$

In particular, because the switch map $(p,q) \mapsto (-q,p)$ is a symplectomorphism the Hamiltonian vector fields restricted to the energy hypersurface $E^{-1}(-\frac{1}{2}) = K^{-1}(\frac{1}{2})$ are related by

$$X_K|_{E^{-1}(-\frac{1}{2})}(p,q) = |q| X_E|_{E^{-1}(-\frac{1}{2})}(-q,p).$$

That means the flow of X_K is just a reparametrization of the flow of X_E. In particular, after reparametrization the Kepler flow at energy $-\frac{1}{2}$ can be interpreted as the geodesic flow of the round two sphere in the chart obtained by stereographic projection.

The energy hypersurface $E^{-1}(-\frac{1}{2})$ is noncompact. This is due to *collisions*. However, the geodesic flow in the chart on the round two sphere obtained by stereographic projection extends to the geodesic flow on the whole two-sphere. This procedure *regularizes* the Kepler problem. If we think of the two-sphere as $S^2 = \mathbb{R}^2 \cup \{\infty\}$, then the point at infinity corresponds to collisions. Indeed, at a collision point the original momentum p explodes which corresponds to the point at infinity of S^2 after interchanging the roles of momentum and position. It is also useful to note that the geodesic flow on the round two-sphere is completely periodic. The Kepler flow at negative energy is almost periodic. Trajectories are either ellipses which are periodic or collision orbits which are not periodic. Heuristically one can imagine that by regularizing the Kepler problem one builds in some kind of trampoline into the mass at the origin, such that when the body collides it just bounces back. The corresponding bounce orbit then becomes periodic in accordance with the fact that the geodesic flow on the round two-sphere is periodic.

To get a feeling what is going on on a more conceptional level it is also useful to discuss regularization in terms of Hamiltonian manifolds. Recall that by going from the energy hypersurface of a Hamiltonian in a symplectic manifold to the underlying Hamiltonian manifold one loses the information on the parametrization of the trajectories. Because during the regularization procedure one has to reparametrize the trajectories it is conceptually easier to forget the parametrization at all and just discuss the whole procedure via Hamiltonian structures. Given a noncompact Hamiltonian manifold (Σ, ω) the question is if there exists a *closed* Hamiltonian manifold $(\overline{\Sigma}, \overline{\omega})$ and an embedding

$$\iota \colon \Sigma \to \overline{\Sigma}$$

such that

$$\iota^* \overline{\omega} = \omega.$$

Consider now more generally the Kepler problem at a negative energy value $c < 0$. In this case the Kepler flow is still up to reparametrization equivalent to the geodesic flow of the round two sphere in stereographic projection. However the symplectic transformation $(p, q) \mapsto (-q, p)$ now has to be combined with the physical transformation $(q, p) \mapsto (\frac{q}{\sqrt{-2c}}, \sqrt{-2c}p)$. Note

$$K(p,q) = \frac{1}{2}\left\{ -\frac{|q|}{2c}\left[E\Big(-\frac{q}{\sqrt{-2c}}, \sqrt{-2c}p\Big) - c\right] + \frac{1}{\sqrt{-2c}}\right\}^2 = \frac{1}{2}\left(\frac{1}{2}(p^2+1)|q|\right)^2.$$

The two Hamiltonian vector fields on the energy hypersurface

$$\Sigma_c := E^{-1}(c) = K^{-1}(-\tfrac{1}{4c})$$

are related by

$$X_K|_{\Sigma_c}(p,q) = \frac{|q|}{(-2c)^{\frac{3}{2}}} X_E|_{\Sigma_c}\big(-\tfrac{q}{\sqrt{-2c}}, \sqrt{-2c}p\big)$$

respectively

$$X_E|_{\Sigma_c}(q,p) = -\frac{2c}{|q|} X_K|_{\Sigma_c}\big(\tfrac{p}{\sqrt{-2c}}, -\sqrt{-2c}q\big).$$

Suppose now that $\gamma \in C^\infty(S^1, \Sigma_c)$ is a Kepler ellipse, i.e., a solution of the ODE

$$\partial_t \gamma(t) = \tau X_E(\gamma(t)), \quad t \in S^1$$

where $\tau > 0$ is the minimal period of the ellipse. We next express the period τ in terms of the energy value c. In view of the discussion above we can interpret γ as well as a simple closed geodesic for the geodesic flow on the round two sphere. In view of $\Sigma_c = K^{-1}(-\frac{1}{4c})$ the momentum of the geodesic has length $\sqrt{-2c}$. Therefore if $\lambda = -qdp$ is the Liouville one-form on the cotangent bundle of the sphere (note that momentum and position interchanged their role) the action of the simple closed geodesic is

$$\int_{S^1} \gamma^* \lambda = \frac{2\pi}{\sqrt{-2c}}. \tag{27}$$

Consider the one-form

$$\lambda' = \lambda - d(qp) = -2qdp - pdq.$$

Because λ and λ' only differ by an exact one-form we have by Stokes

$$\int_{S^1} \gamma^* \lambda = \int_{S^1} \gamma^* \lambda'. \tag{28}$$

Using that the Hamiltonian vector field of the Kepler Hamiltonian is given by

$$X_E = p \frac{\partial}{\partial q} - \frac{q}{|q|^3} \frac{\partial}{\partial p}$$

we compute

$$\int_{S^1} \gamma^* \lambda' = \int_0^1 \lambda'(\tau X_E(\gamma)) dt = \tau \int_0^1 \Big(-p^2 + \frac{2}{|q|} \Big) dt = -2c\tau. \tag{29}$$

Combining (27), (28). and (29) we obtain

$$\tau = -\frac{\pi}{c\sqrt{-2c}}.$$

We proved the following version of Kepler's third law

Lemma 1.1 (Kepler's third law). *The minimal period τ of a Kepler ellipse only depends on the energy and we have the relation*

$$\tau^2 = \frac{\pi^2}{-2E^3}.$$

2. The Levi-Civita regularization

We embed $\mathbb{C}$ into its cotangent bundle $T^*\mathbb{C}$ as the zero section. We get a smooth map

$$\mathscr{L} \colon \mathbb{C}^2 \setminus (\mathbb{C} \times \{0\}) \to T^*\mathbb{C} \setminus \mathbb{C}, \quad (u, v) \mapsto \Big(\frac{u}{\bar{v}}, 2v^2 \Big).$$

If we think of $\mathbb{C}$ as a chart of S^2 via stereographic projection at the north pole, the map $\mathscr{L}$ extends to a smooth map

$$\mathscr{L} \colon \mathbb{C}^2 \setminus \{0\} \to T^*S^2 \setminus S^2$$

which we denote by abuse of notation by the same letter. The map $\mathscr{L}$ is a covering map of degree 2. If we write (p, q) for coordinates of $T^*\mathbb{C} = \mathbb{C} \times \mathbb{C}$. where a bit unconventional but justified by Moser's regularization we write p for the base coordinate and q for the fiber coordinate, the Liouville one-form on $T^*\mathbb{C}$ is given by

$$\lambda = q_1 dp_1 + q_2 dp_2 = \mathrm{Re}(q d\bar{p}).$$

If we pull back the Liouville one-form by the map $\mathcal{L}$ we obtain

$$\begin{aligned} \lambda_{\mathscr{L}}(u,v) &:= \mathscr{L}^*\lambda(u,v) \\ &= \mathrm{Re}\left(2v^2 d\left(\frac{\bar{u}}{v}\right)\right) \\ &= 2\mathrm{Re}\left(v^2\left(\frac{d\bar{u}}{v} - \frac{\bar{u}dv}{v^2}\right)\right) \\ &= 2\mathrm{Re}\big(v d\bar{u} - \bar{u} dv\big) \\ &= 2\big(v_1 du_1 - u_1 dv_1 + v_2 du_2 - u_2 dv_2\big). \end{aligned}$$

Its exterior derivative is the symplectic form

$$\omega_{\mathscr{L}} = 4\big(dv_1 \wedge du_1 + dv_2 \wedge du_2\big).$$

Note that $\lambda_{\mathscr{L}}$ does not agree with the standard Liouville one-form on $\mathbb{C}^2$ given by

$$\lambda_{\mathbb{C}^2} = \frac{1}{2}\big(u_1 du_2 - u_2 du_1 + v_1 dv_2 - v_2 dv_1\big)$$

and $\omega_{\mathscr{L}}$ differs from the standard symplectic form on $\mathbb{C}^2$

$$\omega_{\mathbb{C}^2} = d\lambda_{\mathbb{C}^2} = du_1 \wedge du_2 + dv_1 \wedge dv_2.$$

Indeed, the two subspaces $\mathbb{C} \times \{0\}$ and $\{0\} \times \mathbb{C}$ of $\mathbb{C}^2$ are Lagrangian with respect to $\omega_{\mathscr{L}}$ but symplectic with respect to $\omega_{\mathbb{C}^2}$. Nevertheless the Liouville vector field of $\lambda_{\mathscr{L}}$ implicitly defined by

$$\iota_{X_{\mathscr{L}}}\omega_{\mathscr{L}} = \lambda_{\mathscr{L}}$$

is given by

$$X_{\mathscr{L}} = \frac{1}{2}\left(u_1\frac{\partial}{\partial u_1} + u_2\frac{\partial}{\partial u_2} + v_1\frac{\partial}{\partial v_1} + v_2\frac{\partial}{\partial v_2}\right)$$

and agrees with the Liouville vector field of $\lambda_{\mathbb{C}^2}$. Recall that the standard Liouville vector field on T^*S^2 defined by $\iota_X d\lambda = \lambda$ for the standard Liouville one-form on T^*S^2 is given by

$$X = q\frac{\partial}{\partial q}$$

where q denotes the fiber variable. Because pull back commutes with exterior derivative we obtain

$$\mathscr{L}^* X = X_{\mathscr{L}}.$$

This implies the following lemma.

Lemma 2.1. *A closed hypersurface $\Sigma \subset T^*S^2$ is fiberwise star-shaped if and only if $\mathscr{L}^{-1}\Sigma \subset \mathbb{C}^2$ is star-shaped.*

Note that a fiberwise star-shaped hypersurface in T^*S^2 is diffeomorphic to the unit cotangent bundle S^*S^2 which itself is diffeomorphic to three dimensional projective space $\mathbb{R}P^3$. On the other hand a star-shaped hypersurface in $\mathbb{C}^2$ is diffeomorphic to the three dimensional sphere S^3 which is a twofold cover of $\mathbb{R}P^3$.

In practice the Levi-Civita regularization of planar two body collisions [**76**] is carried out by the variable substitution $(q,p) \mapsto (2v^2, \frac{u}{v})$. As pointed out by

Chenciner in [**23**] this substitution is already anticipated by Goursat [**44**]. Hence it is much older than Moser's regularization [**86**]. While Moser's regularization works in every dimension the Levi-Civita regularization depends on the existence of complex numbers which only exist in dimension two. Using quaternions instead of complex numbers an analogue of the Levi-Civita regularization can be constructed in the spatial case [**74**].

We illustrate how the Levi-Civita regularization works for the Kepler problem at the energy value $-\frac{1}{2}$. We consider the Hamiltonian

$$H\colon T^*(\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}, \quad (q,p) \mapsto \frac{p^2}{2} - \frac{1}{|q|} + \frac{1}{2}.$$

After variable substitution we obtain

$$H(u,v) = \frac{|u|^2}{2|v|^2} - \frac{1}{2|v|^2} + \frac{1}{2}.$$

We introduce the Hamiltonian

$$K(u,v) := |v|^2 H(u,v) = \frac{1}{2}\big(|u|^2 + |v|^2 - 1\big).$$

The level set

$$\Sigma := H^{-1}(0) = K^{-1}(0)$$

equals the three dimensional sphere. The Hamiltonian flow of K on Σ is just a reparametrization of the Hamiltonian flow of H on Σ. It is periodic. Physically it can be interpreted as the flow of two uncoupled harmonic oscillators. Summarizing we have seen the following. If we apply Moser regularization to the Kepler problem of negative energy we obtain the geodesic flow of S^2 and the regularized energy hypersurface becomes $\mathbb{R}P^3$. It double cover is S^3 which we directly obtain by applying the Levi-Civita regularization to the Kepler problem. The double cover of the geodesic flow on S^2 can be interpreted as the Hamiltonian flow of two uncoupled harmonic oscillators.

CHAPTER 4

The restricted three body problem

1. The restricted three body problem in an inertial frame

The first ingredient in the restricted three body problem are two masses, the *primaries*, which we refer to it as the earth and the moon. We scale the total mass to one so that for some $\mu \in [0,1]$ the mass of the moon equals μ and the mass of the earth equals $1-\mu$. Here we allow the mass of the moon to be bigger than the mass of the earth, although in such a situation one might prefer to change the names of the primaries. The earth and the moon move in 3-dimensional Euclidean space $\mathbb{R}^3$ according to Newton's law of gravitation and we denote their time dependent positions by $e(t) \in \mathbb{R}^3$ respectively $m(t) \in \mathbb{R}^3$ for $t \in \mathbb{R}$.

The second ingredient is a massless object referred to as the satellite. Because the satellite is massless it does not influence the movements of the earth and the moon. On the other hand the earth and the moon attract the satellite according to Newton's law of gravitation. The goal of the problem is to get an understanding of the dynamics of the satellite which can be quite intricate. If q denotes the position of the satellite and p its momentum than the Hamiltonian of the satellite in the inertial system is given according to Newton's law of gravitation by

$$E_t(q,p) = \frac{1}{2}p^2 - \frac{\mu}{|q-m(t)|} - \frac{1-\mu}{|q-e(t)|} \tag{30}$$

namely the sum of kinetic energy and Newton's potential. We abbreviate this Hamiltonian by E and not by H in order to distinguish it from the Hamiltonian of the restricted three body problem in rotating coordinates. Note that because the earth and the moon are moving the Hamiltonian is not autonomous, i.e., it depends on time. Actually, because we have to avoid collisions of the satellite with one of the primaries even the domain of definition of the Hamiltonian is time dependent, namely

$$E_t \colon T^*\big(\mathbb{R}^3 \setminus \{e(t), m(t)\}\big) \to \mathbb{R}.$$

In particular, because the Hamiltonian depends on time it is not preserved under the flow of its time dependent Hamiltonian vector field, i.e., preservation of energy does not hold.

If the satellite moves in the ecliptic, i.e., the plane spanned by the orbits of the earth and the moon, after choosing suitable coordinates such that $e(t), m(t) \in \mathbb{R}^2$ for every $t \in \mathbb{R}$, the domain of definition of the Hamiltonian becomes

$$E_t \colon T^*\big(\mathbb{R}^2 \setminus \{e(t), m(t)\}\big) \to \mathbb{R}.$$

This is referred to as the *planar restricted three body problem*, while the former one is called the *spatial restricted three body problem.* In the following we focus on the planar case. This is due to the fact that the question about global surfaces of section only makes sense in the planar case. A further specialization is obtained by assuming that the earth and moon move on circles about their common center of mass. After choosing suitable coordinates their time dependent position are given by

$$e(t) = -\mu\big(\cos(t), \sin(t)\big), \quad m(t) = (1-\mu)\big(\cos(t), \sin(t)\big). \tag{31}$$

This problem is referred to as the *circular planar restricted three body problem*. Of course there is also a circular spatial restricted three body problem. The amazing thing about the circular case is that after a time dependent transformation which puts the earth and moon at rest, the Hamiltonian of the circular restricted three body problem in rotating coordinates becomes autonomous, i.e., independent of time. In particular, it is preserved along its flow. This surprising observation is due to Jacobi. We first explain time dependent transformations.

2. Time dependent transformations

Suppose that (M, ω) is a symplectic manifold and $E \in C^\infty(M \times \mathbb{R}, \mathbb{R})$ and $L \in C^\infty(M \times \mathbb{R}, \mathbb{R})$ are two time dependent Hamiltonians. For $t \in \mathbb{R}$ abbreviate $E_t = E(\cdot, t) \in C^\infty(M)$ and similarly L_t. This gives rise to two time dependent Hamiltonian vector fields X_{E_t} and X_{L_t}. For simplicity let us assume that the flows of the Hamiltonian vector fields ϕ^t_E and ϕ^t_L exist for all times. One can consider more complicated situations where the domain of definitions of the two Hamiltonians itself depend on time. This actually happens in the restricted three body problem. Nevertheless the treatment of this more general cases does not require basic new ingredients apart from a notational nightmare.

Define the time dependent Hamiltonian function

$$L \diamondsuit E \in C^\infty(M \times \mathbb{R}, \mathbb{R})$$

by

$$(L \diamondsuit E)(x, t) = L(x, t) + E((\phi^t_L)^{-1}x, t), \quad x \in M, \ t \in \mathbb{R}.$$

We claim that

$$\phi^t_{L\diamondsuit E} = \phi^t_L \circ \phi^t_E, \quad t \in \mathbb{R}. \tag{32}$$

To see that pick $x \in M$. Abbreviate $y = \phi^t_L(\phi^t_E(x))$ and pick further $\xi \in T_yM$. We compute using the fact that ϕ^t_E is symplectic from Theorem 2.3

$$\begin{aligned}
\omega\bigg(\frac{d}{dt}(\phi^t_L(\phi^t_E(x)), \xi\bigg) &= \omega\Big(X_{L_t}(y) + d\phi^t_L(\phi^t_E(x))X_{E_t}(\phi^t_E(x)), \xi\Big) \\
&= dL_t(y)\xi + \omega\Big(X_{E_t}((\phi^t_L)^{-1}(y)), (d\phi^t_L)^{-1}(y)\xi\Big) \\
&= dL_t(y)\xi + d(E \circ (\phi^t_L)^{-1})(y)\xi \\
&= d(L\diamondsuit E)_t(y)\xi.
\end{aligned}$$

This establishes (32).

Note that even if E and L are autonomous, i.e., independent of time, the Hamiltonian $L\diamondsuit E$ does not need to be autonomous, unless E is invariant under the flow of L.

3. The circular restricted three body problem in a rotating frame

For simplicity we discuss the planar case. The spatial case works analogously. We apply to the Hamiltonian E_t given by (30) with positions of the earth and moon determined by (31) the time dependent transformation generated by angular momentum[1]

$$L \in C^\infty(T^*\mathbb{R}^2, \mathbb{R}), \quad (q,p) \mapsto p_1 q_2 - p_2 q_1.$$

We abbreviate

$$H := L \diamond E.$$

Note that by (9) angular momentum generates the rotation. If we abbreviate

$$e = (-\mu, 0), \quad m = (1-\mu, 0)$$

the Hamiltonian H becomes

$$H(q,p) = \frac{1}{2}p^2 - \frac{\mu}{|q-m|} - \frac{1-\mu}{|q-e|} + p_1 q_2 - p_2 q_1. \tag{33}$$

Note that this Hamiltonian is *autonomous*. In particular, in the rotating frame the Hamiltonian H is preserved by Theorem 2.2. This surprising observation goes back to Jacobi and therefore H is also referred to as the *Jacobi integral.* More precisely, for some historic reasons the integral $-2H$, which of course is preserved under the Hamiltonian flow of H as well, is traditionally called the Jacobi integral.

We point out that the fact that $H = L \diamond H^i$ is autonomous only holds in the circular case. For example if the primaries move on ellipses with some positive eccentricity, the so called elliptic restricted three body problem, the Hamiltonian H does not become time independent.

Abbreviating by

$$V \colon \mathbb{R}^2 \setminus \{e, m\} \to \mathbb{R}, \quad q \mapsto -\frac{\mu}{|q-m|} - \frac{1-\mu}{|q-e|}$$

the Newtonian potential the Hamiltonian equation of motion become

$$\left\{ \begin{aligned} q_1' &= p_1 + q_2 \\ q_2' &= p_2 - q_1 \\ p_1' &= p_2 - \tfrac{\partial V}{\partial q_1} \\ p_2' &= -p_1 - \tfrac{\partial V}{\partial q_2}. \end{aligned} \right. \tag{34}$$

For the second derivatives of q we compute

$$q_1'' = p_1' + q_2' = p_2 - \frac{\partial V}{\partial q_1} + p_2 - q_1 = 2q_2' + q_1 - \frac{\partial V}{\partial q_1}$$

and

$$q_2'' = p_2' - q_1' = -p_1 - \frac{\partial V}{\partial q_2} - p_1 - q_2 = -2q_1' + q_2 - \frac{\partial V}{\partial q_2}.$$

Therefore the first order ODE (34) is equivalent to the following second order ODE

$$\left\{ \begin{aligned} q_1'' &= 2q_2' + q_1 - \tfrac{\partial V}{\partial q_1} \\ q_2'' &= -2q_1' + q_2 - \tfrac{\partial V}{\partial q_2}. \end{aligned} \right. \tag{35}$$

[1]Actually, one usually defines the angular momentum in physics by $q \times p$, which is minus the quantity we define here.

To give the additional rotational term a physical interpretation we complete the squares and rewrite (33) as

$$H(q,p) = \frac{1}{2}\big((p_1+q_2)^2 + (p_2-q_1)^2\big) - \frac{\mu}{|q-m|} - \frac{1-\mu}{|q-e|} - \frac{1}{2}q^2. \tag{36}$$

The last three terms only depend on q and we introduce the so called *effective potential*

$$U\colon \mathbb{R}^2 \setminus \{e,m\} \to \mathbb{R}, \quad q \mapsto -\frac{\mu}{|q-m|} - \frac{1-\mu}{|q-e|} - \frac{1}{2}q^2 = V(q) - \frac{1}{2}q^2.$$

Using this abbreviation the Hamiltonian H can be written more compactly as

$$H(q,p) = \frac{1}{2}\big((p_1+q_2)^2 + (p_2-q_1)^2\big) + U(q). \tag{37}$$

The effective potential potential consists of the Newtonian potential for the earth and the moon plus the additional term $-\frac{1}{2}q^2$. The additional term gives rise to a new force just experienced in rotating coordinates, namely the *centrifugal force*. The Hamiltonian H in (37) is not a mechanical Hamiltonian anymore, i.e., it doesn't just consist of kinetic plus potential energy. Instead of that the Hamiltonian contains a twist in the kinetic part and is therefore a magnetic Hamiltonian as discussed in Section 3.3. The twist in the kinetic part can be interpreted in terms of physics as an additional force, namely the *Coriolis force*. Different from the gravitational force and the centrifugal force which only depend on the position of the satellite the Coriolis force depends on its velocity, like the Lorentz force for a particle moving in a magnetic field. This explains why the Hamiltonian of the restricted three body problem in rotating coordinates becomes a magnetic Hamiltonian. There are now four forces acting on the satellite in the rotating coordinate system, the gravitational force of the earth, the gravitational force of the moon, the centrifugal force, as well as the Coriolis force. This vividly shows that the dynamics of the restricted three body complex is highly intricate.

4. The five Lagrange points

In this section we discuss the critical points of the Hamiltonian H given by (37). We immediately observe that the projection map $\pi\colon \mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ given by $(q,p) \mapsto q$ induces a bijection

$$\pi|_{\mathrm{crit}(H)}\colon \mathrm{crit}(H) \to \mathrm{crit}(U). \tag{38}$$

Indeed, the inverse map for a critical point $(q_1,q_2) \in \mathrm{crit}(U)$ is given

$$\big(\pi|_{\mathrm{crit}(H)}\big)^{-1}(q_1,q_2) = (q_1,q_2,-q_2,q_1).$$

We explain now that if $\mu \in (0,1)$, the effective potential $U = U_\mu$ has five critical points. The critical points of U are called *Lagrange points*.

Note that U is invariant under reflection at the axis of earth and moon $(q_1,q_2) \mapsto (q_1,-q_2)$. Therefore either a critical point of U lies on the axis of earth and moon, i.e., the fixed point set of the reflection, or it appears in pairs. It turns out that there are three critical points of U on the axis of earth and moon. These *collinear points* were discovered already by Euler and they are saddle points of U. Moreover, there is one pair of non collinear critical points of U discovered by Lagrange. It turns out that these points build equilateral triangles with the earth and the moon and for this reason they are referred to as *equilateral points*. The equilateral points are maxima of U.

We first discuss the two equilateral Lagrange points following the book by Abraham-Marsden [**2**]. Because of the reflection symmetry we can restrict our attention to upper half-space $\mathbb{R} \times (0, \infty)$. Because the distance between earth and moon is one, i.e., $|e - m| = 1$, we have a diffeomorphism

$$\phi\colon \mathbb{R} \times (0, \infty) \to \Theta$$

between the upper half-space and the half-strip

$$\Theta = \{(\rho, \sigma) \in (0, \infty)^2 : \rho + \sigma > 1,\ |\rho - \sigma| < 1\}$$

which is given by

$$\phi(q) = \big(|q - m|, |q - e|\big), \quad q \in \mathbb{R} \times (0, \infty).$$

Consider the smooth function

$$V\colon \Theta \to \mathbb{R}, \quad V := U \circ \phi^{-1}.$$

Critical points of V correspond to critical points of the effective potential U on the upper half-space. To give an explicit description of V in terms of the variables ρ and σ we compute for $q \in \mathbb{R} \times (0, \infty)$

$$\begin{aligned} q^2 &= \mu q^2 + (1-\mu) q^2 \\ &= \mu\big(\rho^2 + 2\langle m, q\rangle - m^2\big) + (1-\mu)\big(\sigma^2 + 2\langle e, q\rangle - e^2\big) \\ &= \mu\rho^2 + 2\mu(1-\mu)\langle 1, q\rangle - \mu(1-\mu)^2 \\ &\quad + (1-\mu)\sigma^2 - 2\mu(1-\mu)\langle 1, q\rangle - (1-\mu)\mu^2 \\ &= \mu\rho^2 + (1-\mu)\sigma^2 - \mu(1-\mu). \end{aligned}$$

Therefore V as function of ρ and σ reads

$$V(\rho, \sigma) = -\frac{\mu}{\rho} - \frac{1-\mu}{\sigma} - \frac{1}{2}\big(\mu\rho^2 + (1-\mu)\sigma^2 - \mu(1-\mu)\big). \tag{39}$$

Its differential is given by

$$dV(\rho, \sigma) = \frac{\mu(1-\rho^3)}{\rho^2} d\rho + \frac{(1-\mu)(1-\sigma^3)}{\sigma^2} d\sigma.$$

Hence V has a unique critical point at $(1, 1) \in \Theta$. The Hessian at the critical point $(1, 1)$ is given by

$$H_V(1, 1) = \begin{pmatrix} -3\mu & 0 \\ 0 & -3(1-\mu) \end{pmatrix}.$$

We conclude that $(1, 1)$ is a maximum. Going back to the original coordinates we define the Lagrange point ℓ_4 as

$$\ell_4 = \phi^{-1}(1, 1) := \big(\tfrac{1}{2} - \mu, \tfrac{\sqrt{3}}{2}\big). \tag{40}$$

We define the Lagrange point ℓ_5 as the point obtained by reflection ℓ_4 at the axis through earth and moon

$$\ell_5 := \big(\tfrac{1}{2} - \mu, -\tfrac{\sqrt{3}}{2}\big). \tag{41}$$

By reflection symmetry of U the Lagrange point ℓ_5 is also a maximum of U and it is the only critical point in the lower half-space $\mathbb{R} \times (-\infty, 0)$. We summarize what we proved so far

Lemma 4.1. *The only critical points of U on $\mathbb{R}^2\setminus(\mathbb{R}\setminus\{0\}$, i.e., the complement of the axis through earth and moon are ℓ_4 and ℓ_5 and they are maximas of U.*

We next discuss the collinear critical points. For this purpose we consider the function

$$u := U|_{\mathbb{R}\setminus\{-\mu,1-\mu\}} \colon \mathbb{R}\setminus\{-\mu, 1-\mu\} \to \mathbb{R}, \quad r \mapsto -\frac{\mu}{|r+\mu-1|} - \frac{1-\mu}{|r+\mu|} - \frac{r^2}{2}.$$

Because U is invariant under reflection at the axis of the earth and moon, it follows that critical points of u are critical points of U as well. The second derivative of u is given by

$$u''(r) = -\frac{2\mu}{|r+\mu-1|^3} - \frac{2(1-\mu)}{|r+\mu|^3} - 1 < 0.$$

Therefore u is strictly concave. Since at the singularities at $-\mu$ and $1-\mu$ as well as at $-\infty$ and ∞ the function u goes to $-\infty$, we conclude that the function u attains precisely three maxima, one at a point $-\mu < \ell_1 < 1-\mu$, one at a point $\ell_2 > 1-\mu$ and one at a point $\ell_3 < -\mu$. The points ℓ_1, ℓ_2, ℓ_3 are referred to as the three collinear Lagrange points. Although one has closed formulas of the position of the Lagrange points ℓ_4 and ℓ_5 a similar closed formula does not exist for ℓ_1, ℓ_2, or ℓ_3. In fact to obtain the exact position of the three collinear Lagrange points one has to solve quintic equations with coefficients depending on μ, see [**2**, Chapter 10].

Lemma 4.2. *The three collinear Lagrange points are saddle points of the effective potential U.*

Proof: Because there is no closed formula for the position of the three collinear Lagrange points we give a topological argument to prove the Lemma. Note that the Euler characteristic of the two fold punctured plane satisfies

$$\chi\big(\mathbb{R}^2 \setminus \{e, m\}\big) = -1.$$

Denote by ν_2 the number of maxima of U, by ν_1 the number of saddle points of U, and by ν_0 the number of minima of U. Because U goes to $-\infty$ at infinity as well as at the singularities e and m it follows from the Poincaré-Hopf index theorem that

$$\nu_2 - \nu_1 + \nu_0 = \chi\big(\mathbb{R}^2 \setminus \{e, m\}\big) = -1. \tag{42}$$

By Lemma 4.1 we know that ℓ_4 and ℓ_5 are maxima, so that

$$\nu_2 \geq 2. \tag{43}$$

Since the three collinear Lagrange points are maxima of u, the restriction of U to the axis through earth and moon, it follows that they are either saddle points or maxima of U. In particular,

$$\nu_0 = 0 \tag{44}$$

and therefore

$$\nu_1 + \nu_2 = 5. \tag{45}$$

Combining (42), (43), (44), and (45) we conclude that

$$\nu_2 = 2, \quad \nu_1 = 3.$$

This finishes the proof of the Lemma. □

The reader might enjoy trying by looking at the Hessian of U to check directly that the three collinear Lagrange points are saddle points. If you get stuck we recommend to have a look at [**2**, Chapter 10].

Because of the reflection symmetry the function U attains the same value at ℓ_4 and ℓ_5. In view of Lemma 4.2 together with the fact that U goes to $-\infty$ at infinity and the singularities e and m, we conclude that it attains it global maximum at the two equilateral Lagrange points. We state this fact in the following Corollary.

Corollary 4.3. *The effective potential attains its global maximum precisely at the two equilateral Lagrange points and it holds that*

$$\max U = U(\ell_4) = U(\ell_5) = -\frac{3}{2} + \frac{\mu(\mu-1)}{2}.$$

Proof: To compute the value of $U(\ell_4)$ we get using (39)

$$U(\ell_4) = V(1,1) = -\mu - (1-\mu) - \frac{1}{2}\big(\mu + 1 - \mu - \mu(1-\mu)\big) = -\frac{3}{2} + \frac{\mu(\mu-1)}{2}.$$

This finishes the proof of the Corollary. $\square$

We next discuss the ordering of the critical values of the saddle points of U.

Lemma 4.4. *If $\mu \in \big(0, \frac{1}{2}\big)$ the critical values of the collinear Lagrange points are ordered as follows*

$$U(\ell_1) < U(\ell_2) < U(\ell_3). \tag{46}$$

If $\mu = \frac{1}{2}$ we have

$$U(\ell_1) < U(\ell_2) = U(\ell_3). \tag{47}$$

Remark 4.5. *If $\mu \in \big(\frac{1}{2}, 1\big)$ one gets from (46) by interchanging the roles of the earth and the moon that*

$$U(\ell_1) < U(\ell_3) < U(\ell_2).$$

Proof of Lemma 4.4: We follow the exposition given by Kim [**70**]. We first show $U(\ell_1) < U(\ell_2)$ for $\mu \in (0,1)$. Suppose that $-\mu < q < 1-\mu$. Abbreviate $\rho := 1-\mu-q > 0$ and set $q' := 1-\mu+\rho$. In the following we identify $\mathbb{R}$ with $\mathbb{R} \times \{0\} \subset \mathbb{R}^2$. We estimate

$$\begin{aligned} U(q') - U(q) &= -\frac{\mu}{\rho} - \frac{1-\mu}{1+\rho} - \frac{1}{2}(1-\mu+\rho)^2 + \frac{\mu}{\rho} + \frac{1-\mu}{1-\rho} + \frac{1}{2}(1-\mu-\rho)^2 \\ &= (1-\mu)\left(\frac{1}{1-\rho} - \frac{1}{1+\rho} - 2\rho\right) \\ &= \frac{2(1-\mu)\rho^3}{1-\rho^2} \\ &> 0. \end{aligned}$$

In particular, by choosing $q = \ell_1$ we get

$$U(\ell_1) < U(\ell_1') \leq U(\ell_2)$$

where for the last inequality we used that ℓ_2 is the maximum of the restriction of U to $(1-\mu, \infty)$.

We next show that for $0<\mu<\frac{1}{2}$ it holds that $U(\ell_2)<U(\ell_3)$. If $q>1-\mu$ we estimate

$$\begin{aligned}
U(-q)-U(q) &= -\frac{\mu}{1-\mu+q}-\frac{1-\mu}{q-\mu}-\frac{q^2}{2}+\frac{\mu}{q-1+\mu}+\frac{1-\mu}{q+\mu}+\frac{q^2}{2}\\
&= \mu\Big(\frac{1}{q-(1-\mu)}-\frac{1}{q+(1-\mu)}\Big)+(1-\mu)\Big(\frac{1}{q+\mu}-\frac{1}{q-\mu}\Big)\\
&= \frac{2\mu(1-\mu)}{q^2-(1-\mu)^2}-\frac{2\mu(1-\mu)}{q^2-\mu^2}\\
&= \frac{2\mu(1-\mu)\big((1-\mu)^2-\mu^2\big)}{(q^2-(1-\mu)^2)(q^2-\mu^2)}\\
&= \frac{2\mu(1-\mu)(1-2\mu)}{(q^2-(1-\mu)^2)(q^2-\mu^2)}\\
&> 0.
\end{aligned}$$

We choose now $q=\ell_2$ to obtain

$$U(\ell_2)<U(-\ell_2)\leq U(\ell_3)$$

because ℓ_3 is the maximum of the restriction of U to $(-\infty,-\mu)$.

We finally note that if $\mu=\frac{1}{2}$ the effective potential U is invariant under reflection at the y-axis $(q_1,q_2)\mapsto(-q_1,q_2)$ as well and ℓ_2 is mapped to ℓ_3 under reflection at the y-axis. This finishes the proof of the Lemma. $\square$

Recall from (38) that projection to position space gives a bijection between critical points of the Hamiltonian H and critical points of the effective potential U. For $i\in\{1,2,3,4,5\}$ abbreviate

$$L_i=\pi|^{-1}_{\mathrm{crit}(H)}(\ell_i)\in\mathrm{crit}(H).$$

If $\ell_i=(q_1^i,q_2^i)$ then $L_i=(q_1^i,q_2^i,-q_2^i,q_1^i)$. Note that

$$H(L_i)=U(\ell_i)$$

and if $\mu(L_i)$ denotes the Morse index of L_i as a critical point of H, i.e., the number of negative eigenvalues of the Hessian of H at L_i, we have

$$\mu(L_i)=\mu(\ell_i).$$

In particular, we proved the following theorem.

Theorem 4.6. *For $\mu\in(0,1)$ the Morse indices of the five critical points of H satisfy*

$$\mu(L_1)=\mu(L_2)=\mu(L_3)=1,\quad \mu(L_4)=\mu(L_5)=2.$$

If $\mu\in\big(0,\frac{1}{2}\big)$ the critical values of H are ordered as

$$H(L_1)<H(L_2)<H(L_3)<H(L_4)=H(L_5).$$

If $\mu=\frac{1}{2}$, then the critical values satisfy

$$H(L_1)<H(L_2)=H(L_3)<H(L_4)=H(L_5).$$

5. Hill's regions

Let H be the Hamiltonian of the planar circular restricted three body problem in rotating coordinates given by (37). Fix $c \in \mathbb{R}$. Because H is autonomous the energy hypersurface or level set

$$\Sigma_c = H^{-1}(c) \subset T^*\big(\mathbb{R}^2 \setminus \{e, m\}\big)$$

is preserved under the flow of the Hamiltonian vector field of H. Consider the footpoint projection

$$\pi \colon T^*\big(\mathbb{R}^2 \setminus \{e, m\}\big) \to \mathbb{R}^2 \setminus \{e, m\}, \quad (q, p) \mapsto q.$$

The Hill's region of Σ_c is the shadow of Σ_c under the footpoint projection

$$\mathfrak{K}_c := \pi(\Sigma_c) \subset \mathbb{R}^2 \setminus \{e, m\}.$$

Because the first two terms in (37) are nonnegative we can obtain the Hill's region $\mathfrak{K}_c$ as well as the sublevel set of the effective potential

$$\mathfrak{K}_c = \big\{q \in \mathbb{R}^2 \setminus \{e, m\} : U(q) \leq c\big\}.$$

If the energy lies below the first critical value, i.e., $c < H(L_1)$, the Hill's region has three connected components

$$\mathfrak{K}_c = \mathfrak{K}_c^e \cup \mathfrak{K}_c^m \cup \mathfrak{K}_c^u$$

where the earth e lies in the closure of $\mathfrak{K}_c^e$ and the moon m lies in the closure of $\mathfrak{K}_c^m$. The connected components $\mathfrak{K}_c^e$ and $\mathfrak{K}_c^m$ are bounded, where the connected component $\mathfrak{K}_c^u$ is unbounded. Trajectories of the restricted three body problem above the unbounded component $\mathfrak{K}_c^u$ are referred to as *comets*. Accordingly the energy hypersurface of the restricted three body problem decomposes into three connected components

$$\Sigma_c = \Sigma_c^e \cup \Sigma_c^m \cup \Sigma_c^u \tag{48}$$

where

$$\Sigma_c^e := \{(q, p) \in \Sigma_c, \quad q \in \mathfrak{K}_c^e$$

and similarly for Σ_c^m and Σ_c^u.

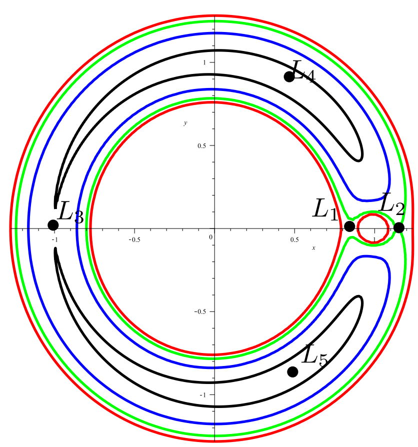

FIGURE 1. The Hill's region for various levels of the Jacobi energy

6. The rotating Kepler problem

The Hamiltonian of the rotating Kepler problem is given by the Hamiltonian of the restricted three body problem (37) for $\mu = 0$. That means that the moon has no zero mass and can be neglected and the satellite is just attracted by the earth like in the Kepler problem. The difference to the usual Kepler problem is that the coordinate system is still rotating. The Hamiltonian of the rotating Kepler problem

$$H\colon T^*(\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}$$

is explicitly given for $(q,p) \in T^*(\mathbb{R}^2 \setminus \{0\})$ by

$$H(q,p) = \frac{1}{2}p^2 - \frac{1}{|q|} + p_1 q_2 - p_2 q_1. \tag{49}$$

The first two terms are just the Hamiltonian E of the planar Kepler problem while the third term is angular momentum L so that we can write

$$H = E + L.$$

Because E and L Poisson commute we get

$$\{H, L\} = \{E, L\} + \{L, L\} = 0$$

meaning that H and L Poisson commute as well. In particular, the rotating Kepler problem is an example of a *completely integrable system.* It is unlikely that for any positive value of μ the restricted three body problem is completely integrable as well and for all but finitely values of μ analytic integrals can be excluded by work of Poincaré and Xia [**90, 108**].

If we complete the squares in (49) we obtain the magnetic Hamiltonian

$$H(q,p) = \frac{1}{2}\big((p_1 + q_2)^2 + (p_2 - q_1)^2\big) - \frac{1}{|q|} - \frac{1}{2}q^2$$

which we write as

$$H(q,p) = \frac{1}{2}\big((p_1 + q_2)^2 + (p_2 - q_1)^2\big) + U(q)$$

for the effective potential

$$U\colon \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$$

given by

$$U(q) = -\frac{1}{|q|} - \frac{1}{2}q^2.$$

Different than for positive μ the effective potential is rotationally invariant. Therefore its critical set is rotationally invariant as well. We write

$$U(q) = f(|q|)$$

for the function

$$f\colon (0,\infty) \to \mathbb{R}, \quad r \mapsto -\frac{1}{r} - \frac{1}{2}r^2.$$

The differential of f is given by

$$f'(r) = \frac{1}{r^2} - r$$

and therefore f has a unique critical point at $r = 1$ with critical value

$$f(1) = -\frac{3}{2}.$$

We have proved

Lemma 6.1. *The effective potential U of the rotating Kepler problem has a unique critical value $-\frac{3}{2}$ and its critical set consists of the circle of radius one around the origin.*

Because critical points of H and U are in bijection via projection as explained in (38) and the value of H at a critical point coincides with the value of U of its projection we obtain the following Corollary.

Corollary 6.2. *The Hamiltonian H of the rotating Kepler problem has a unique critical value $-\frac{3}{2}$.*

7. Moser regularization of the restricted three body problem

In complex notation the Hamiltonian (33) of the restricted three body problem can be written as the map $H\colon T^*(\mathbb{C} \setminus \{-\mu, 1-\mu\}) \to \mathbb{R}$ given by

$$H(q,p) = \frac{1}{2}p^2 - \frac{\mu}{|q-1+\mu|} - \frac{1-\mu}{|q+\mu|} + \langle q, ip \rangle.$$

We shift coordinates and put the origin of our coordinate system to the moon to obtain the Hamiltonian $H_m\colon T^*(\mathbb{C} \setminus \{0,-1\}) \to \mathbb{R}$ defined by

$$\begin{aligned} H_m(q,p) &= H(q+1-\mu, p+i-i\mu) + \frac{(1-\mu)^2}{2} \\ &= \frac{1}{2}p^2 - \frac{\mu}{|q|} - \frac{1-\mu}{|q+1|} + \langle q, ip\rangle + \langle q, \mu - 1\rangle \\ &= \frac{1}{2}p^2 - \frac{\mu}{|q|} + \langle q, ip\rangle - \frac{1-\mu}{|q+1|} - (1-\mu)q_1. \end{aligned} \tag{50}$$

For $c \in \mathbb{R}$ we consider the energy hypersurface $\Sigma_c = H_m^{-1}(c)$. We switch the roles of q and p and think of p as the base coordinate and q as the fiber coordinate. The energy hypersurface is than a subset $\Sigma_c \subset T^*\mathbb{C} \subset T^*S^2$ where for the later inclusion we think of the sphere as $S^2 = \mathbb{C} \cup \{\infty\}$ via stereographic projection.

We examine if the closure $\overline{\Sigma}_c$ is regular in the fiber above ∞. For this purpose we examine how the terms in (50) transform under chart transition. The chart transition from the chart given by stereographic projection at the north pole to the chart given by stereographic projection at the south pole is given by

$$\phi\colon \mathbb{C} \setminus \{0\} \to \mathbb{C} \setminus \{0\}, \quad p \mapsto \frac{1}{p} = \frac{\overline{p}}{|p|^2}$$

where $\overline{p} = p_1 - ip_2$ denotes the conjugate transpose of p. In real notation this corresponds to the map

$$(p_1, p_2) \mapsto \left(\frac{p_1}{p_1^2 + p_2^2}, \frac{-p_2}{p_1^2+p_2^2} \right).$$

The Jacobian of ϕ at $p \in \mathbb{R}^2 \setminus \{0\}$ computes to be

$$d\phi(p) = \frac{1}{(p_1^2+p_2^2)^2} \left(\begin{array}{cc} p_2^2 - p_1^2 & -2p_1p_2 \\ 2p_1p_2 & p_2^2 - p_1^2 \end{array} \right).$$

Its determinant is given by

$$\det\big(d\phi(p)\big) = \frac{1}{(p_1^2+p_2^2)^2}.$$

Therefore the inverse transpose of $d\phi(p)$ reads

$$\big(d\phi(p)^{-1}\big)^T = \begin{pmatrix} p_2^2 - p_1^2 & -2p_1p_2 \\ 2p_1p_2 & p_2^2 - p_1^2 \end{pmatrix}.$$

Consequently the exact symplectomorphism

$$d_*\phi \colon T^*(\mathbb{R}^2 \setminus \{0\}) \to T^*(\mathbb{R}^2 \setminus \{0\})$$

is given by

$$d_*\phi(p,q) = \left(\frac{p_1}{p_1^2+p_2^2}, \frac{-p_2}{p_1^2+p_2^2}, (p_2^2-p_1^2)q_1 - 2p_1p_2q_2, 2p_1p_2q_1 + (p_2^2-p_1^2)q_2\right). \tag{51}$$

Note that because $\phi^{-1} = \phi$ it holds that

$$(d_*\phi)^{-1} = d_*\phi^{-1} = d_*\phi.$$

Therefore the push forward of kinetic energy to the chart centered at the south pole is given by

$$(d_*\phi)_*\left(\frac{p^2}{2}\right) = (d_*\phi)^*\left(\frac{p^2}{2}\right) = \frac{1}{2p^2}.$$

The push forward of Newton's potential is

$$(d_*\phi)_*\left(\frac{1}{|q|}\right) = \frac{1}{p^2|q|}.$$

Angular momentum $L = p_1q_2 - p_2q_1$ transforms as

$$(d\phi_*)_*L = -L.$$

Assume that $\Omega \subset \mathbb{R}^2$ is an open subset containing the origin 0 and $V \colon \Omega \to \mathbb{R}$ is a smooth function. Set

$$F_V \colon \mathbb{R}^2 \times \Omega \to \mathbb{R}, \quad (p,q) \mapsto V(q)$$

and define

$$F^V \colon (d_*\phi)\big((\mathbb{R}^2 \setminus \{0\}) \times \Omega\big) \cup \big(\{0\} \times \mathbb{R}^2\big) \to \mathbb{R}$$

by

$$F^V(p,q) = \begin{cases} F_V\big(d_*\phi(p,q)\big) & p \neq 0 \\ V(0) & p = 0. \end{cases}$$

It follows from the transformation formula (51) that F^V is smooth.

We now choose $\Omega = \mathbb{C} \setminus \{-1\}$ and

$$V \colon \Omega \to \mathbb{R}, \quad q \mapsto -\frac{1-\mu}{|q+1|} - (1-\mu)q_1.$$

It follows that the Hamiltonian H_m in the chart given by stereographic projection at the south pole is given by the map

$$(q,p) \mapsto \frac{1}{2p^2} - \frac{1}{p^2|q|} - L(q,p) - F^V(p,q).$$

Because F^V extends smoothly to $p = 0$ the energy hypersurface Σ_c extends smoothly to $p = 0$ as well and the intersection of its closure with the fiber over 0 is given by

$$\left\{\frac{1}{|q|} = \frac{1}{2} : q \in \mathbb{R}^2\right\} = \big\{|q| = 2 : q \in \mathbb{R}^2\big\},$$

i.e., a circle of radius 2.

The origin in the chart obtained by stereographic projection at the south pole corresponds to the north pole, respectively the point $\infty \in S^2$. The fact that we obtained for the intersection a circle depends on our specific choice of coordinates on the sphere S^2. However, because convexity is preserved under linear transformations we proved the following Lemma.

Lemma 7.1. *For every $c \in \mathbb{R}$ the closure of the energy hypersurface $\Sigma_c = H_m^{-1}(c) \subset T^*S^2$ is regular in the fiber over ∞ and $\overline{\Sigma}_c \cap T^*_\infty S^2$ is convex.*

Recall from (48) that if the energy c is less then the first critical value $H(L_1)$ the energy hypersurface Σ_c has three connected components Σ_c^e, Σ_c^m, Σ_c^u, the first close to the earth, the second close to the moon and the third consisting of comets. The regularization described above compactifies the component Σ_c^m to $\overline{\Sigma}_c^m$. Of course the same can be done with the component around the earth to obtain the regularized component $\overline{\Sigma}_c^e$. Indeed, the roles of the earth and moon can always be interchanged simply by replacing μ by $1-\mu$. In the following we discuss the regularized component $\overline{\Sigma}_c^e$. But by interchanging the roles of the earth and the moon everything said below is true as well for the moon. In [**6**] the following Theorem is proved.

Theorem 7.2 (Albers-Frauenfelder-Paternain-van Koert). *For energy values below the first critical value, i.e., for $c < H(L_1)$, the regularized energy hypersurface $\overline{\Sigma}_c^e \subset T^*S^2$ is fiberwise star-shaped.*

As an immediate corollary of this theorem we have

Corollary 7.3. *Under the assumption of Theorem 7.2 the restriction of the Liouville one-form λ on T^*S^2 gives a contact form on $\overline{\Sigma}_c^e$. In particular, after reparametrization the regularized flow of the restricted three body problem around the earth below the first critical value can be interpreted as a Reeb flow.*

We have seen that in the fiber over ∞ the regularized energy hypersurface bound actually a convex domain. Therefore we ask the following question.

Question 7.4. *Under the assumptions of Theorem 7.2 is $\overline{\Sigma}_c^e$ fiberwise convex in T^*S^2, i.e., after reparametrization can the regularized flow of the restricted three body problem around the earth below the first critical value be interpreted as a Finsler flow?*

In [**75**] Lee proved that below the first critical value Hill's lunar problem is fiberwise convex. Hill's lunar problem is a limit problem of the restricted three body problem where the mass ratio of the earth and moon diverges to ∞ and the satellite moves in a tiny neighborhood of the moon. We discuss this problem in Section 8. It is also known that below the first critical value the regularized rotating Kepler problem is fiberwise convex, see [**25**].

As a further Corollary of Theorem 7.2 we obtain

Corollary 7.5. *Under the assumption of Theorem 7.2 the regularized energy hypersurface $\overline{\Sigma}_c^e$ is diffeomorphic to $\mathbb{R}P^3$.*

We finish this section by explaining how this Corollary follows more elementary from the Fibration theorem of Ehresmann with no reference to Theorem 7.2. We first recall the following theorem of Ehresmann [**30**]. A proof can be found for example in [**20**, Theorem 8.12]

Theorem 7.6 (Fibration theorem of Ehresmann). *Assume that $f\colon Y \to X$ is a proper submersion of differential manifolds, then f is a locally trivial fibration.*

As a Corollary of Ehresmann's fibration theorem we have

Corollary 7.7. *Assume that $F\colon M \times [0,1] \to \mathbb{R}$ is a smooth function such that $F^{-1}(0)$ is compact and for every $r \in [0,1]$ it holds that 0 is a regular value of $F_r := F(\cdot, r)\colon M \to \mathbb{R}$. Then $F_0^{-1}(0)$ and $F_1^{-1}(0)$ are diffeomorphic closed manifolds.*

Proof: That $F_r^{-1}(0)$ is a closed manifold for every $r \in [0,1]$ follows from the assumptions that 0 is a regular value of F_r and $F^{-1}(0)$ is compact. It remains to prove that all these manifolds are diffeomorphic. We show this by applying Theorem 7.6 to the projection map

$$\pi\colon Y := F^{-1}(0) \to [0,1], \quad (x,r) \mapsto r.$$

Because Y is compact by assumption the projection map is proper. It remains to check that is submersive. If $(x,r) \in Y$ the differential of π is the linear map

$$d\pi_{(x,r)}\colon T_{(x,r)}Y \to \mathbb{R}, \quad (\widehat{x}, \widehat{r}) \mapsto \widehat{r}.$$

The tangent space of Y is given by

$$T_{(x,r)}Y = \big\{(\widehat{x},\widehat{r}) \in T_xM \times \mathbb{R} : dF_{(x,r)}(\widehat{x},\widehat{r}) = 0\big\}.$$

Pick $\widehat{r} \in \mathbb{R}$. Because 0 is a regular value of F_r there exists $\widehat{x} \in T_xM$ such that

$$dF_{(x,r)}(\widehat{x}, 0) = dF_r(\widehat{x}) = -dF_{(x,r)}(0, \widehat{r}).$$

This implies that

$$dF_{(x,r)}(\widehat{x},\widehat{r}) = dF_{(x,r)}(\widehat{x},0) + dF_{(x,r)}(0,\widehat{r}) = 0$$

and therefore $(\widehat{x},\widehat{r}) \in T_{(x,r)}Y$. Since $\widehat{r} \in \mathbb{R}$ was arbitrary this shows that $d\pi_{(x,r)}$ is surjective and π is a proper submersion. Now the assertion of the Corollary follows from Theorem 7.6. $\square$

We now use Corollary 7.7 to give a direct proof of Corollary 7.5 with no reference to Theorem 7.2.

Proof of Corollary 7.5: In view of Corollary 7.7 the Corollary 7.5 can now be proved by a deformation argument. Namely we first switch of the moon to end up in the rotating Kepler problem and then we switch of the rotation as well. To make the dependence of the regularized energy hypersurface on the mass of the moon μ visible we write $\overline{\Sigma}^e_{c,\mu}$. For given $\mu_1 \in (0,1)$ and $c_1 < H_{\mu_1}(L_{1,\mu_1})$ we choose a smooth path $c\colon [0,\mu_1] \to \mathbb{R}$ with the property that $c(\mu) < H_\mu(L_{1,\mu})$ and $c(\mu_1) = c_1$. In view of Corollary 7.7 the manifold $\overline{\Sigma}^e_{c_1,\mu_1}$ is diffeomorphic to $\overline{\Sigma}^e_{c(0),0}$ but the last one is just the regularized energy hypersurface of the rotating Kepler problem below the first critical value. A further homotopy which switches of the rotation in the rotating Kepler problem shows that the latter one is diffeomorphic to the regularized energy hypersurface of the (non-rotating) Kepler problem for a negative energy value which by Moser is diffeomorphic to $\mathbb{R}P^3$. This proves the Corollary. $\square$

8. Hill's lunar problem

8.1. Derivation of Hill's lunar problem. While the restricted three body problem considers the case where the two primaries have comparable masses, Hill's lunar problem [**49**] deals with the case where the first primary compared to the second one has a much much bigger mass, the second primary has a much much bigger mass than the satellite and the satellite moves very close to the second primary. We show how to derive the Hamiltonian of Hill's lunar problem from the restricted three body problem by blowing up to coordinates close to the second primary, compare also [**82**]. Recall from (50) that after shifting position and momenta of the Hamiltonian of the restricted three body problem in order to put the origin of our coordinate system to the moon we obtain the Hamiltonian

$$H_m(q,p) = \frac{1}{2}p^2 - \frac{\mu}{|q|} - (1-\mu)\left(\frac{1}{\sqrt{(q_1+1)^2+q_2^2}} + q_1\right) + p_1q_2 - p_2q_1.$$

The diffeomorphism

$$\phi_\mu \colon T^*\mathbb{R}^2 \to T^*\mathbb{R}^2, \quad (q,p) \mapsto (\mu^{\frac{1}{3}}q, \mu^{\frac{1}{3}}p)$$

is conformally symplectic with constant conformal factor $\mu^{\frac{2}{3}}$, i.e.,

$$\phi_\mu^*\omega = \mu^{\frac{2}{3}}\omega$$

for the standard symplectic form on $T^*\mathbb{R}^2$. Define

$$H^\mu \colon T^*(\mathbb{R}^2 \setminus \{(0,0),(-\mu^{-\frac{1}{3}},0)\}) \to \mathbb{R}, \qquad H^\mu := \mu^{-\frac{2}{3}}(H_m \circ \phi_\mu - 1).$$

Because ϕ is symplectically conformal with conformal factor $\mu^{\frac{2}{3}}$ it follows that

$$X_{H^\mu} = \phi_\mu^* X_{H_m}.$$

On each compact subset the Hamiltonian H^μ converges uniformly in the C^∞-topology to the Hamiltonian $H \colon T^*(\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}$ given by

$$H(q,p) = \frac{1}{2}p^2 - \frac{1}{|q|} + p_1q_2 - p_2q_1 - q_1^2 + \frac{1}{2}q_2^2 \tag{52}$$

We refer to the Hamiltonian H as *Hill's lunar Hamiltonian.*

The Hamiltonian equation corresponding to Hill's lunar Hamiltonian is the following first order ODE

$$\left\{\begin{aligned} q_1' &= p_1 + q_2 \\ q_2' &= p_2 - q_1 \\ p_1' &= p_2 + 2q_1 - \tfrac{q_1}{|q|^3} \\ p_2' &= -p_1 - q_2 - \tfrac{q_2}{|q|^3} \end{aligned}\right. \tag{53}$$

which is equivalent to the following second order ODE

$$\left\{\begin{aligned} q_1'' &= 2q_2' + 3q_1 - \tfrac{q_1}{|q|^3} \\ q_2'' &= -2q_1' - \tfrac{q_2}{|q|^3}. \end{aligned}\right. \tag{54}$$

8.2. Hill's lunar Hamiltonian. Apart from its rather simple form which makes Hill's lunar Hamiltonian an important testing ground for numerical investigations, see for example [**103**], a nice feature of it is that it is invariant under two commuting anti-symplectic involutions. Namely $\rho_1, \rho_2 \colon T^*\mathbb{R}^2 \to T^*\mathbb{R}^2$ given for $(q,p) \in T^*\mathbb{R}^2$ by the formula

$$\rho_1(q_1, q_2, p_1, p_2) = (q_1, -q_2, -p_1, p_2), \quad \rho_2(q_1, q_2, p_1, p_2) = (-q_1, q_2, p_1, -p_2)$$

are both anti-symplectic involutions such that

$$H \circ \rho_1 = H, \quad H \circ \rho_2 = H.$$

The two anti-symplectic involutions commute and their product is the symplectic involution

$$\rho_1 \circ \rho_2 = \rho_2 \circ \rho_1 = -\mathrm{id}.$$

In contrast to Hill's lunar Hamiltonian the Hamiltonian of the restricted three body problem is only invariant under ρ_1 but not under ρ_2.

By completing the squares we can write Hill's lunar Hamiltonian (52) in the equivalent form

$$H(q,p) \quad = \quad \frac{1}{2}\big((p_1+q_2)^2 + (p_2-q_1)^2\big) - \frac{1}{|q|} - \frac{3}{2}q_1^2.$$

If one introduces the *effective potential* $U \colon \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$ by

$$U(q) = -\frac{1}{|q|} - \frac{3}{2}q_1^2$$

the Hamiltonian for Hill's lunar problem can be written as

$$H(q,p) = \frac{1}{2}\big((p_1+q_2)^2 + (p_2-q_1)^2\big) + U(q).$$

As in the restricted three body problem there is a one to one correspondence between critical points of H and critical points of U. Namely, the footpoint projection

$$\pi \colon T^*(\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}^2 \setminus \{0\}, \quad (q,p) \mapsto q$$

induces a bijection

$$\pi|_{\mathrm{crit}(H)} \colon \mathrm{crit}(H) \to \mathrm{crit}(U).$$

The partial derivatives of U compute to be

$$\frac{\partial U}{\partial q_1} = \frac{q_1}{|q|^3} - 3q_1 = q_1\left(\frac{1}{|q|^3} - 3\right), \qquad \frac{\partial U}{\partial q_2} = \frac{q_2}{|q|^3}. \tag{55}$$

The latter implies that at a critical point q_2 has to vanish. Since U has a singularity at the origin, we conclude that at a critical point q_1 does not vanish and therefore the critical set of U is given by

$$\mathrm{crit}(U) = \big\{(3^{-\frac{1}{3}}, 0), (-3^{-\frac{1}{3}}, 0)\big\}.$$

The two critical points of U can be thought of as the limits of the first and second Lagrange point ℓ_1 and ℓ_2 of the restricted three body problem under the blow up at the moon. Because the remaining Lagrange points ℓ_3, ℓ_4, and ℓ_5 are too far away from the moon, they are not visible in Hill's lunar problem anymore. In particular, the critical set of H reads

$$\mathrm{crit}(H) = \big\{(3^{-\frac{1}{3}}, 0, 0, 3^{-\frac{1}{3}}), (-3^{-\frac{1}{3}}, 0, 0, -3^{-\frac{1}{3}})\big\}.$$

Note that the two critical points of H are fixed by the anti-symplectic involution ρ_1 but interchanged through the anti-symplectic involution ρ_2. Because H is invariant under ρ_2 it attains the same value at the two critical points. That means that Hill's lunar problem has just one critical value which computes to be

$$\begin{aligned} H(3^{-\frac{1}{3}},0,0,3^{-\frac{1}{3}}) &= U(3^{-\frac{1}{3}},0) \\ &= -3^{\frac{1}{3}} - \frac{3}{2}3^{-\frac{2}{3}} \\ &= -3^{\frac{1}{3}} - \frac{3^{\frac{1}{3}}}{2} \\ &= -\frac{3^{\frac{1}{3}}(2+1)}{2} \\ &= -\frac{3^{\frac{4}{3}}}{2}. \end{aligned}$$

To obtain the Hessian of U at its critical points we compute from (55)

$$\frac{\partial^2 U}{\partial q_1^2}(\pm 3^{-\frac{1}{3}},0) = -\frac{3q_1^2}{|q|^5}\bigg|_{(q_1,q_2)=(\pm 3^{-\frac{1}{3}},0)} = -\frac{3}{|q_1|^3}\bigg|_{q_1=\pm 3^{-\frac{1}{3}}} = -9,$$

$$\frac{\partial^2 U}{\partial q_1 \partial q_2}(\pm 3^{-\frac{1}{3}},0) = \frac{\partial^2 U}{\partial q_2 \partial q_1}(\pm 3^{-\frac{1}{3}},0) = 0,$$

and finally

$$\frac{\partial^2 U}{\partial q_2^2}(\pm 3^{-\frac{1}{3}},0) = \frac{1}{|q|^3}\bigg|_{(q_1,q_2)=(\pm 3^{-\frac{1}{3}},0)} = 3.$$

Therefore the Hessian of U at its critical points is given by

$$H_U(\pm 3^{-\frac{1}{3}},0) = \begin{pmatrix} -9 & 0 \\ 0 & 3 \end{pmatrix}.$$

We conclude that the critical points of U are saddle points. It follows that the two critical points of H have Morse index equal to one. We summarize this fact in the following Lemma.

Lemma 8.1. *Hill's lunar Hamiltonian has a unique critical value at energy* $-\frac{3^{\frac{4}{3}}}{2}$*. At the critical value it has two critical points both of Morse index one.*

If $c \in \mathbb{R}$ we abbreviate by

$$\Sigma_c = H^{-1}(c)$$

the three dimensional energy hypersurface of H in the four dimensional phase space $T^*(\mathbb{R}^2 \setminus \{0\})$. The *Hill's region* to the energy value c is defined as

$$\mathfrak{K}_c = \pi(\Sigma_c) = \big\{q \in \mathbb{R}^2 \setminus \{0\} : U(q) \leq c\big\}. \tag{56}$$

If $c < -\frac{3^{\frac{4}{3}}}{2}$ then the Hills region $\mathfrak{K}_c$ has three connected components one bounded and the other two unbounded. We denote the bounded component of $\mathfrak{K}_c$ by $\mathfrak{K}_c^b$ and abbreviate

$$\Sigma_c^b := \big\{(q,p) \in \Sigma_c : q \in \mathfrak{K}_c^b\big\}.$$

9. Euler's problem of two fixed centers

See [**106**]

CHAPTER 5

Periodic orbits

1. Variational approach

Assume that (M,ω) is a symplectic manifold and $H \in C^\infty(M,\mathbb{R})$ is a Hamiltonian with the property that 0 is a regular value, i.e., $\Sigma = H^{-1}(0) \subset M$ is a regular hypersurface. Abbreviate by $S^1 = \mathbb{R}/\mathbb{Z}$ the circle. A *parametrized periodic orbit* on Σ is a loop $\gamma \in C^\infty(S^1,\Sigma)$ for which there exists $\tau > 0$ such that the tuple (γ,τ) is a solution of the problem

$$\partial_t \gamma(t) = \tau X_H(\gamma(t)), \quad t \in S^1.$$

Because γ is parametrized the positive number $\tau = \tau(\gamma)$ is uniquely determined by γ. We refer to τ as the *period* of γ. Indeed, if we reparametrize γ to $\gamma_\tau \colon \mathbb{R} \to N$ by

$$\gamma_\tau(t) = \gamma\big(\tfrac{t}{\tau}\big)$$

then γ_τ satisfies

$$\partial_t \gamma_\tau(t) = X_H(\gamma_\tau(t)), \quad \gamma_\tau(t+\tau) = \gamma(t),\ t \in \mathbb{R}.$$

By a *period orbit* we mean the trace $\{\gamma(t) : t \in S^1\}$ of a parametrized periodic orbit. If we think of $(\Sigma, \omega|_\Sigma)$ as a Hamiltonian manifold, a periodic orbit corresponds to a closed leaf of the foliation $\ker \omega$ on Σ. However, note that the period of the periodic orbit only makes sense with reference to the Hamiltonian H and cannot be determined directly from the Hamiltonian structure $\omega|_\Sigma$.

We next explain how parametrized periodic orbits can be interpreted variationally as critical points of an action functional. To simplify the discussion we assume that (M,ω) is an exact symplectic manifold, i.e., ω admits a primitive λ such that $d\lambda = \omega$. Abbreviate

$$\mathfrak{L} = C^\infty(S^1, M)$$

the free loop space of M and set $\mathbb{R}_+ = (0,\infty)$ the positive real numbers. Consider

$$\mathcal{A}^H \colon \mathfrak{L} \times \mathbb{R}_+ \to \mathbb{R}.$$

defined for a free loop $\gamma \in \mathfrak{L}$ and $\tau \in \mathbb{R}_+$ by

$$\mathcal{A}^H(\gamma,\tau) = \int_{S^1} \gamma^*\lambda - \tau \int_{S^1} H(\gamma(t))dt. \tag{57}$$

One might think of $\mathcal{A}^H$ as the Lagrange multiplier functional of the area functional of the constraint given by the mean value of H. We refer to $\mathcal{A}^H$ as *Rabinowitz action functional.*

Remark 1.1. *The action functional $\mathcal{A}^H$ has itself an interesting history. Although Moser explicitly wrote in (*[**87**]*) that $\mathcal{A}^H$ is useless in finding periodic orbits, it was used shortly later by Rabinowitz in his celebrated paper* [**93**] *to prove existence of periodic orbits on star-shaped hypersurfaces in $\mathbb{R}^{2n}$ and so opened the way for* *the application of global methods in Hamiltonian mechanics. Moreover, the action functional $\mathcal{A}^H$ can be used to define a semi-infinite dimensional Morse homology in the sense of Floer* [**33**]*, which is referred to as Rabinowitz Floer homology, see* [**5, 24**].

Lemma 1.2. *Critical points of $\mathcal{A}^H$ precisely consist of pairs (γ, τ), where γ is a parametrized periodic orbit of X_H on Σ of period τ.*

Proof: If $\gamma \in \mathfrak{L}$ the tangent space of $\mathfrak{L}$ at γ

$$T_\gamma \mathfrak{L} = \Gamma(\gamma^* TM)$$

consists of vector fields along γ. Suppose that $(\gamma, \tau) \in \mathfrak{L} \times \mathbb{R}_+$ and pick

$$(\widehat{\gamma}, \widehat{\tau}) \in T_{(\gamma,\tau)}(\mathfrak{L} \times \mathbb{R}_+) = T_\gamma \mathfrak{L} \times \mathbb{R}.$$

By applying Cartan's formula to the Lie derivative $\mathcal{L}_{\widehat{\gamma}}\lambda$ we compute for the pairing of the differential of $\mathcal{A}^H$ with $(\widehat{\gamma}, \widehat{\tau})$

$$\begin{aligned}
d\mathcal{A}^H_{(\gamma,\tau)}(\widehat{\gamma}, \widehat{\tau}) &= \int_{S^1} \gamma^* \mathcal{L}_{\widehat{\gamma}}\lambda - \tau \int_{S^1} dH(\gamma)\widehat{\gamma} dt - \widehat{\tau} \int_{S^1} H(\gamma) dt \\
&= \int_{S^1} \gamma^* d\iota_{\widehat{\gamma}}\lambda + \int_{S^1} \gamma^* \iota_{\widehat{\gamma}} d\lambda - \tau \int \omega(\widehat{\gamma}, X_H(\gamma)) dt \\
&\quad -\widehat{\tau} \int_{S^1} H(\gamma) dt \\
&= \int_{S^1} d\gamma^* \iota_{\widehat{\gamma}}\lambda + \int_{S^1} \gamma^* \iota_{\widehat{\gamma}} \omega - \int \omega(\widehat{\gamma}, \tau X_H(\gamma)) dt \\
&\quad -\widehat{\tau} \int_{S^1} H(\gamma) dt \\
&= \int_{S^1} \omega(\widehat{\gamma}, \partial_t \gamma) dt - \int \omega(\widehat{\gamma}, \tau X_H(\gamma)) dt - \widehat{\tau} \int_{S^1} H(\gamma) dt \\
&= \int_{S^1} \omega(\widehat{\gamma}, \partial_t \gamma - \tau X_H(\gamma)) dt - \widehat{\tau} \int_{S^1} H(\gamma) dt
\end{aligned}$$

We conclude that a critical point (γ, τ) of $\mathcal{A}^H$ is a solution of the problem

$$\left\{ \begin{array}{c} \partial_t \gamma - \tau X_H(\gamma) = 0 \\ \int_{S^1} H(\gamma) dt = 0. \end{array} \right. \tag{58}$$

In view of preservation of energy as explained in Theorem 2.2 problem (58) is equivalent to the following problem

$$\left\{ \begin{array}{c} \partial_t \gamma = \tau X_H(\gamma) \\ H(\gamma) = 0, \end{array} \right. \tag{59}$$

i.e., the mean value constraint can be replaced by a pointwise constraint. But solution of problem (59) are precisely parametrized periodic orbits of X_H of period τ on the energy hypersurface $\Sigma = H^{-1}(0)$. This proves the Lemma. □

There are two actions on the free loop space $\mathfrak{L} = C^\infty(S^1, M)$. The first action comes from the group structure of the domain S^1. Given $r \in S^1$ and $\gamma \in \mathfrak{L}$ we reparametrize γ by

$$r_* \gamma(t) = \gamma(r + t), \quad t \in S^1.$$

If we extend this S^1-action of $\mathfrak{L}$ to $\mathfrak{L} \times \mathbb{R}$ by acting trivially on the second factor, the action functional $\mathcal{A}^H$ is S^1-invariant. Its critical set, namely the set of parametrized periodic orbits, is than S^1-invariant as well, which is also obvious from the ODE parametrized periodic orbits meet. We refer to an equivalence class of parametrized periodic orbits under reparametrization by S^1 as an *unparametrized periodic orbit.*

The second action on the free loop space comes from the fact that S^1 is diffeomorphic to its finite covers. Consider the action of the monoid $\mathbb{N}$ on $\mathfrak{L}$ given for $n \in \mathbb{N}$ and $\gamma \in \mathfrak{L}$ by

$$n_*\gamma(t) = \gamma(nt), \quad t \in S^1.$$

The two actions combined give rise to an action on $\mathfrak{L}$ of the semi-direct product $\mathbb{N} \ltimes S^1$ with product defined as

$$(n_1, r_1)(n_2, r_2) = (n_1 n_2, n_1 r_2 + r_1).$$

We extend the action of $\mathbb{N}$ on $\mathfrak{L}$ to an action of $\mathbb{N}$ on $\mathfrak{L} \times \mathbb{R}$ which is given for $\gamma \in \mathfrak{L}$, $\tau \in \mathbb{R}$ and $n \in \mathbb{N}$ by

$$n_*(\gamma, \tau) = (n_*\gamma, n\tau).$$

The action functional $\mathcal{A}^H$ is homogeneous of degree one for the action by $\mathbb{N}$, i.e.,

$$\mathcal{A}^H(n_*(\gamma, \tau)) = n\mathcal{A}^H(\gamma, \tau), \qquad (\gamma, \tau) \in \mathfrak{L} \times \mathbb{R}, \quad n \in \mathbb{N}.$$

Therefore its critical set is invariant under the action of $\mathbb{N}$ as well, again a fact which can immediately understood as well by looking at the ODE. A parametrized periodic orbit γ is called *multiple covered* if there exists a parametrized periodic orbit γ_1 and a positive integer $n \geq 2$ such that $\gamma = n_*\gamma_1$. Note that this notion does not depend on the parametrization of the orbit such that one can talk about multiple covers on the level of unparametrized periodic orbits. A parametrized or unparametrized periodic orbit is called *simple* if it is not multiple covered. Because a periodic orbit is a solution of a first order ODE it follows that for a simple periodic orbit it holds that

$$\gamma(t) = \gamma(t') \quad \Longleftrightarrow \quad t = t' \in S^1.$$

Moreover, for every parametrized periodic orbit γ there exists a unique simple periodic orbit γ_1 and a unique $k \in \mathbb{N}$ such that $\gamma = k_*\gamma_1$. We refer to k as the *covering number* of the periodic orbit γ. Alternatively, the covering number can also be defined as

$$\operatorname{cov}(\gamma) = \max\Big\{k \in \mathbb{N} : \gamma(t + \tfrac{1}{k}) = \gamma(t), \ \forall\, t \in S^1\Big\}.$$

Again the covering number does not depend on the parametrization and can therefore associated as well to unparametrized periodic orbits. With the notion of the covering number at our disposal we can characterize simple periodic orbits as the periodic orbits of covering number one.

We referred to the trace of a parametrized periodic orbit as a periodic orbit. There is a one to one correspondence between periodic orbits and simple unparametrized orbits respectively equivalence classes of parametrized periodic orbits under the action of the monoid $\mathbb{N} \ltimes S^1$.

2. Symmetric periodic orbits and brake orbits

Suppose that (M, ω, ρ) is a real symplectic manifold, namely a symplectic manifold (M, ω) together with an anti-symplectic involution ρ. Moreover, assume that $H \in C^\infty(M, \mathbb{R})$ is a Hamiltonian invariant under ρ, i.e.,

$$H \circ \rho = H$$

and 0 is a regular value of H. Because the Hamiltonian is invariant under ρ while the symplectic form is anti-invariant it follows that the Hamiltonian vector field is anti-invariant as well

$$\rho^* X_H = -X_H.$$

The anti-symplectic involution induces an involution $\mathcal{I}_\rho$ on the free loop space $C^\infty(S^1, \Sigma)$ of the energy hypersurface $\Sigma = H^{-1}(0)$ which is given for $\gamma \in C^\infty(S^1, \Sigma)$ by

$$\mathcal{I}_\rho(\gamma)(t) = \rho(\gamma(1-t)), \quad t \in S^1.$$

Because the Hamiltonian vector field is anti-invariant under ρ it follows that if γ is a periodic orbit of period τ, then $\mathcal{I}_\rho(\gamma)$ is again a periodic orbit of the same period τ. In particular, the involution $\mathcal{I}_\rho$ restricts to an involution of periodic orbits. A periodic orbit fixed under the involution $\mathcal{I}_\rho$ is called a *symmetric periodic orbit*. In particular, a symmetric periodic orbit satisfies

$$\gamma(t) = \rho(\gamma(1-t)).$$

If we plug into this equation $t = \frac{1}{2}$ we obtain

$$\gamma\big(\tfrac{1}{2}\big) = \rho\big(\gamma\big(\tfrac{1}{2}\big)\big)$$

concluding that

$$\gamma\big(\tfrac{1}{2}\big) \in \mathrm{Fix}(\rho) \cap \Sigma.$$

Because γ is periodic, i.e., $\gamma(0) = \gamma(1)$ we further obtain

$$\gamma(0) \in \mathrm{Fix}(\rho) \cap \Sigma$$

as well.

We discussed in Chapter 7 that the fixed point set of an anti-symplectic involution is a (maybe empty) Lagrangian submanifold of M. We further claim that

$$\Sigma \pitchfork \mathrm{Fix}(\rho), \tag{60}$$

meaning that Σ intersects the fixed point set of ρ transversally. To see that let $x \in \Sigma \cap \mathrm{Fix}(\rho) = H^{-1}(0) \cap \mathrm{Fix}(\rho)$. Because 0 is a regular value of H by assumption there exists $v \in T_x M$ such that

$$dH(x)v \neq 0.$$

Because H is invariant under ρ it holds that

$$dH(x) = d(H \circ \rho)(x) = dH(\rho(x))d\rho(x) = dH(x)d\rho(x).$$

Set

$$w := v + d\rho(x)v.$$

We compute

$$dH(x)w = dH(x)v + dH(x)d\rho(x)v = 2dH(x)v \neq 0.$$

Moreover,

$$d\rho(x)w = w$$

so that

$$w \in T_x\mathrm{Fix}(\rho).$$

This proves (60).

If we restrict a symmetric periodic orbit x period τ to $[0, \frac{1}{2}]$, then the path $x|_{[0,\frac{1}{2}]} \in C^\infty\big([0, \frac{1}{2}], \Sigma\big)$ is a solution of the problem

$$\left\{ \begin{array}{c} \partial_t x(t) = \tau X_H(x(t)), \quad t \in \left[0, \frac{1}{2}\right] \\ x(0),\ x\left(\frac{1}{2}\right) \in \mathrm{Fix}(\rho). \end{array} \right.$$

A solution of this problem is referred to as a *brake orbit*. On the other hand given a brake orbit we can obtain a symmetric periodic orbit as follows. Namely we set for $t \in \left(\frac{1}{2}, 1\right]$

$$x(t) := \rho(x(1-t)).$$

In view of the boundary conditions of a brake orbit we obtain a continuous loop $x \in C^0(S^1, \Sigma)$. Because X_H is anti-invariant under ρ and x is an integral curve of the vector field τX_H in the interval $\left[0, \frac{1}{2}\right]$ it follows that x is an integral curve of τX_H on the whole circle. In this case x is automatically smooth and therefore a periodic orbit. Moreover, by construction it is symmetric. This proves that the restriction map

$$x \mapsto x|_{[0,\frac{1}{2}]}$$

induces a one to one correspondence between symmetric periodic orbits and brake orbits.

Brake orbits are an interesting topic of study in its own right. The notion of brake orbits goes back to the work by Seifert [**99**], in which they were studied for mechanical Hamiltonians with anti-symplectic involution mapping p to $-p$ which corresponds to time reversal. We refer to the paper by Long, Zhang, and Zhu [**80**] for a modern study of brake orbits and as guide to the literature. We mention as well the paper by Kang [**64**] in which brake orbits are studied in connection with respect to the restricted three body problem.

3. Blue sky catastrophes

Assume that (M, ω) is a symplectic manifold and $H \in C^\infty(M \times [0,1], \mathbb{R})$. We think of H as a one parameter family of autonomous Hamiltonian functions $H_r = H(\cdot, r) \in C^\infty(M, \mathbb{R})$ and we assume that 0 is a regular value of H_r for every $r \in [0,1]$. the level set $H_r^{-1}(0)$ is connected, and $H^{-1}(0)$ is compact such that $H_r^{-1}(0)$ is a smooth one parameter family of closed, connected submanifolds of M. For $r \in [0,1]$ abbreviate by X_{H_r} the Hamiltonian vector field of H_r implicitly defined by the condition

$$dH_r = \omega(\cdot, X_{H_r}).$$

Suppose that $(\gamma_r, \tau_r) \in C^\infty\big(S^1, H_r^{-1}(0)) \times (0, \infty)\big)$ for $r \in [0,1)$ is a smooth family of loops and positive numbers solving the problem

$$\partial_t \gamma_r(t)) = \tau_r X_{H_r}(\gamma_r(t)), \quad t \in S^1,\ r \in [0,1),$$

i.e., γ_r is a smooth family of periodic orbits of X_{H_r} on the energy hypersurfaces $H_r^{-1}(0)$ and τ_r are its periods. Suppose now that τ_r converges to $\tau_1 \in (0, \infty)$ as r goes to 1. Because $H^{-1}(0)$ is compact it follows from the theorem of Arzela-Ascoli that γ_r converges to a periodic orbit γ_1 of period τ_1. On the other hand if τ_r goes to infinity, then the family of periodic orbits γ_r "disappears in the blue sky" as r goes to 1. Such a scenario is referred to as a *"blue sky catastrophe"*.

We explain how the assumption that the energy hypersurfaces $H_r^{-1}(0)$ are contact prevents blue sky catastrophes.

Theorem 3.1. *Assume that $\omega = d\lambda$ such that $\lambda|_{H_r^{-1}(0)}$ is a contact form for every $r \in [0,1]$. Suppose that (γ_r, τ_r) for $r \in [0,1)$ is a smooth family of periodic orbits γ_r of period τ_r. Then there exists $\tau_1 \in (0,\infty)$ such that τ_r converges to τ_1.*

Proof: Let R_r be the Reeb vector field of $\lambda|_{H_r^{-1}(0)}$. Note that $X_{H_r}|_{H_r^{-1}(0)}$ is parallel to R_r. We first claim that we can assume without loss of generality that

$$R_r = X_{H_r}|_{H_r^{-1}(0)}. \tag{61}$$

To see this note that because the two vector fields are parallel there exist smooth functions $f_r\colon H_r^{-1}(0) \to \mathbb{R}$ such that

$$R_r = f_r X_{H_r}|_{H_r^{-1}(0)}.$$

Moreover, because $H^{-1}(0)$ is compact, there exists $c > 0$ such that

$$\frac{1}{c} \leq |f_r(x)| \leq c, \quad r \in [0,1],\ x \in H_r^{-1}(0).$$

Now choose a smooth extension $\bar{f}\colon M \times [0,1] \to \mathbb{R} \setminus \{0\}$ such that

$$\bar{f}(\cdot, r)|_{H_r^{-1}(0)} = f_r$$

and replace H by $\bar{f} \cdot H$. This guarantees (61). The original family of periodic orbits γ_r gets reparametrized by this procedure, however because of the compactness of $H^{-1}(0)$ the question about convergence of τ_r is unaffected.

We now consider the family of functionals

$$\mathcal{A}_r := \mathcal{A}^{H_r}\colon C^\infty(S^1, M) \times (0,\infty) \to \mathbb{R}.$$

Using (61) we compute now the action of $\mathcal{A}_r$ at the critical point (γ, τ) as follows

$$\mathcal{A}_r(\gamma, \tau) = \int_0^1 \lambda\big(\tau X_{H_r}(\gamma)\big)dt = \tau \int_0^1 \lambda(R_r)dt = \tau \tag{62}$$

i.e., the period of the periodic orbit γ can be interpreted as the action value of $\mathcal{A}_r$. Now let (γ_r, τ_r) for $r \in [0,1)$ be a smooth family of periodic orbits, or according to our new interpretation critical points of $\mathcal{A}_r$. Using (62) we are now in position to compute the derivative of τ_r with respect to the r-parameter as follows

$$\partial_r \tau_r = \frac{d}{dr}\big(\mathcal{A}_r(\gamma_r, \tau_r)\big) = (\partial_r \mathcal{A}_r)(\gamma_r, \tau_r) = -\tau_r \int_{S^1} (\partial_r H_r)(\gamma_r)dt. \tag{63}$$

Here we have used in the second equation that (γ_r, τ_r) is a critical point of $\mathcal{A}_r$. Because $H^{-1}(0)$ is compact there exists $\kappa > 0$ such that

$$\Big|\partial_r H_r|_{H_r^{-1}(0)}\Big| \leq \kappa, \quad \forall\ r \in [0,1]. \tag{64}$$

Combining (63) and (64) we obtain the estimate

$$|\partial_r \tau_r| \leq \kappa \tau_r.$$

In particular, if $0 \leq r_1 < r_2 < 1$, this implies

$$e^{-\kappa(r_2 - r_1)}\tau_{r_1} \leq \tau_{r_2} \leq e^{\kappa(r_2 - r_1)}\tau_{r_1}.$$

This proves that τ_r converges, when r goes to 1 and hence excludes blue sky catastrophes. $\square$

4. Periodic orbits in the rotating Kepler problem

Recall that the Hamiltonian of the rotating Kepler problem is given by

$$H = E + L$$

where E is the Hamiltonian of the (non-rotating) Kepler problem and L is angular momentum. Furthermore, by (25) the relation

$$A^2 = 1 + 2L^2 E$$

holds, where A is the Runge-Lenz vector whose length corresponds to the eccentricity of the corresponding Kepler ellipse. Combining these two facts we obtain the inequality

$$0 \leq 1 + 2E(H-E)^2 = 1 + 2H^2E - 4HE^2 + 2E^3 =: p(H,E).$$

Moreover, equality holds if and only if the corresponding Kepler orbit has vanishing eccentricity, i.e., for circular periodic orbits. Before we investigate these circular orbits in more detail, we first explain some symmetry properties that general orbits enjoy.

4.1. The shape of the orbits if $E < 0$. From Noether's theorem we know that $\{E, L\} = 0$, so $[X_E, X_L] = 0$. It follows that the flows of X_E and X_L commute, so we see that

$$\phi_H^t = \phi_L^t \circ \phi_E^t. \tag{65}$$

We want to investigate how the orbits look like if $E < 0$.

For this we consider the q-components of an orbit in the Kepler problem. Let $\epsilon_T : [0, T] \to \mathbb{R}^2$ denote a Kepler ellipse with period T, i.e. a solution to the Kepler problem with negative energy.

By (65), we also obtain a solution to the rotating Kepler problem, which no longer needs to be periodic. Its q-components are given by

$$\epsilon_T^R(t) = e^{it}\epsilon_T(t),$$

since L just induces a rotation in both the q- and p-plane. There are now two cases

- ϵ_T is a circle. In this case, ϵ_T^R is periodic unless it is a critical point (which can happen if $T = 2\pi$).
- ϵ_T is not a circle, in which case it is either a proper ellipse or a collision orbit (which looks like a line).

We now consider the second, so the orbit ϵ_T is not a circle. We then observe that such an orbit is periodic if the following resonance relation is satisfied for some positive integers k, ℓ

$$2\pi k = T\ell.$$

Hence periodic orbits in the rotating Kepler problem of the second kind have the following symmetry property.

Lemma 4.1. *Periodic orbits in the rotating Kepler problem of the second kind satisfy the following rotational symmetry,*

$$\epsilon_T^R(t+T) = e^{2\pi k/\ell}\epsilon_T^R(t).$$

PROOF. From the periodicity condition we have $T = 2\pi k/\ell$, so we find

$$\epsilon_T^R(t+T) = e^{it+iT}\epsilon_T(t+T) = e^{i2\pi k\ell}e^{it}\epsilon_T(t) = e^{2\pi k/\ell}\epsilon_T^R(t).$$

□

Figure 1 illustrates this lemma.

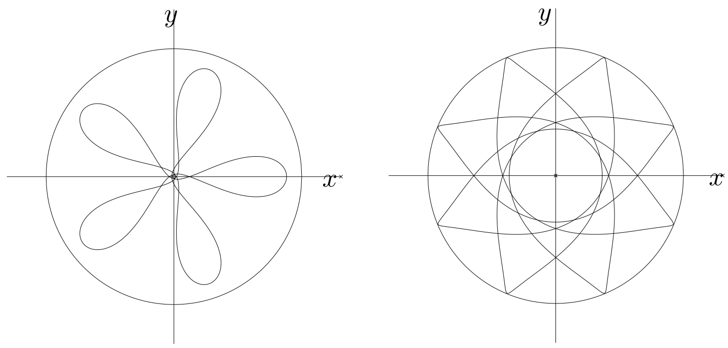

FIGURE 1. Periodic orbits in the rotating Kepler problem for $c = 1.6$: the circle indicates the boundary of the Hill's region.

4.2. More on circular orbits. If we fix H, the function

$$p_H := p(H, \cdot)$$

is a cubic polynomial in E and if we fix E, the function

$$p^E := p(\cdot, H)$$

is a quadratic polynomial in H. By Corollary 6.2 we know that $-\frac{3}{2}$ is the unique critical value of H. At the critical value the cubic polynomial $p_{-\frac{3}{2}}$ splits as follows

$$p_{-\frac{3}{2}}(E) = 2(E+2)(E+\tfrac{1}{2})^2 \tag{66}$$

i.e., $p_{-\frac{3}{2}}$ has a simple zero at -2 and a double zero at $-\frac{1}{2}$. Recall that the discriminant of a cubic polynomial $p = ax^3 + bx^2 + cx + d$ is given by the formula

$$\Delta(p) = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd.$$

In the case that the coefficients are real, the discriminant can be used to determine the number of real roots. Namely if $\Delta(p) > 0$ the polynomial p has three distinct real roots, if $\Delta(p) = 0$ the polynomial has a double root, and if $\Delta(p) < 0$ the polynomial has one real and two complex conjugated roots. For the cubic polynomial p_H the discriminant computes as follows

$$\Delta(p_H) = 64H^6 - 64H^6 + 256H^3 - 108 - 288H^3 = -32H^3 - 108.$$

We see that $\Delta(p_H) > 0$ for $H < -\frac{3}{2}$, vanishes at $H = -\frac{3}{2}$ and satisfies $\Delta(p_H) < 0$ for $H > -\frac{3}{2}$. For $H < -\frac{3}{2}$ we denote by $E^1(H), E^2(H), E^3(H) \in \mathbb{R}$ the zero's of p_H ordered such that

$$E^1(H) < E^2(H) < E^3(H).$$

In view of (66) the three functions extend continuously to $H = -\frac{3}{2}$ such that

$$E^1\big(-\tfrac{3}{2}\big) = -2, \qquad E^2\big(-\tfrac{3}{2}\big) = E^3\big(-\tfrac{3}{2}\big) = -\tfrac{1}{2}.$$

Moreover, E^1 extends to a continuous function on the whole real line such that $E^1(H)$ is the unique real root of p_H if $H > -\frac{3}{2}$. Note that the discriminant of the quadratic polynomial $p^E(H) = 2EH^2 - 4E^2H + 2E^3 + 1$ equals

$$\Delta(p^E) = -8E$$

and therefore for $E < 0$ the polynomial p^E has precisely two real zeros. We conclude that the functions E^1 and E^2 are monotone increasing and the function E^3 is monotone decreasing such that

$$\lim_{H \to -\infty} E^1(H) = \lim_{H \to -\infty} E^2(H) = -\infty, \qquad \lim_{H \to -\infty} E^3(H) = \lim_{H \to \infty} E^1(H) = 0.$$

For later reference we observe that the images of the three functions are

$$\mathrm{im} E^1|_{(-\infty, \frac{3}{2}]} = (-\infty, 2], \quad \mathrm{im} E^2 = (-\infty, \tfrac{1}{2}], \quad \mathrm{im} E^3 = [\tfrac{1}{2}, 0). \tag{67}$$

Note that as an unparametrized, simple orbit a circular orbit in the (non-rotating) planar Kepler problem is uniquely determined by its energy E and its angular momentum L. On the other for given values of E and L a circular periodic orbit only exists if $0 = 1 + 2EL^2$. That means that for a given negative energy value there exist precisely two circular orbits whose angular momenta differ by a sign, i.e., the circle is traversed backwards.

A circular periodic orbit of the Kepler problem also gives rise to a circular periodic orbit in the rotating Kepler problem, since the circle is invariant under rotation. In particular, a circular periodic orbit of the rotating Kepler problem is uniquely determined by the values of H and L or equivalently by the values of H and E. That for given values of H and E a circular periodic orbit exists, it most hold that $p(H, E) = 0$. Hence by the discussion above, for a given value of H less than the critical value $-\frac{3}{2}$ there exist three circular periodic orbits, while for a given energy value H bigger than the critical value $-\frac{3}{2}$ there exists a unique circular periodic orbit.

If the energy value c is less than $-\frac{3}{2}$ the Hill's region $\mathfrak{K}_c$ has two connected components, one bounded and one unbounded. We next discuss which of the three circular periodic orbits lie above the bounded component and which lie above the unbounded one. Because the Runge-Lenz vector for a circular periodic orbit vanishes we obtain from (24) for the radius r of a circular periodic orbit

$$r = L^2 = -\frac{1}{2E}$$

while the second inequality follows from (25) again in view of the fact that the Runge-Lenz vector vanishes. In view of (67) we conclude that the circular periodic orbits corresponding to the energy values $E^1(c)$ and $E^2(c)$ have radius less than one where the circular periodic orbit corresponding to the energy value $E^3(c)$ has radius bigger one. Therefore the first two circular periodic orbits lie above the bounded component of the Hill's region where the third one lies above the unbounded component of the Hill's region.

The circular periodic orbit corresponding to E^1 is referred to as the *retrograde periodic orbit*, while the circular periodic orbit corresponding the E^2 is referred to as the *direct periodic orbit.*

For very small energy the rotating Kepler problem approaches more and more the usual Kepler problem which after Moser regularization is equivalent to the geodesic flow on the two sphere. Due to invariance of the geodesic flow on the round two sphere under rotation the closed geodesics are not isolated. We next explain how the retrograde and direct periodic orbit of the rotation Kepler problem bifurcate from the geodesic flow of the round two sphere.

Periodic orbits can be interpreted variationally as critical points of Rabinowitz action functional. We first explain the bifurcation picture out of a Morse-Bott critical component in the finite dimensional set-up. For this purpose suppose that X is a manifold and $f \in C^\infty(X \times [0,1), \mathbb{R})$. For $r \in [0,1)$ we abbreviate $f_r := f(\cdot, r) \in C^\infty(X, \mathbb{R})$ so that we obtain a one-parameter family of smooth functions on X. Suppose that

$$C \subset \mathrm{crit} f_0$$

is a Morse-Bott component of the critical set of f_0. We mean by this that $C \subset X$ is a closed submanifold which corresponds to a connected component of the critical set of f_0 with the property that for every $x \in C$ it holds that

$$T_x C = \ker H_{f_0}(x)$$

where $H_{f_0}(x)$ is the Hessian of f_0 at x. Suppose that the restriction of the derivative of f_r with respect to the homotopy variable r to the Morse-Bott component $\mathring{f}_0|_C$ is a Morse function. Then it follows from the implicit function theorem that there exists $\epsilon > 0$, an open neighborhood U of C in X and a smooth function

$$x \in \mathrm{crit}(\mathring{f}_0|_C) \times [0, \epsilon) \to U$$

meeting the following conditions.

(i): If $\iota \colon \mathrm{crit}(\mathring{f}_0|_C) \to U$ is the inclusion and $x_0 = x(\cdot, 0) \colon \mathrm{crit}(\mathring{f}_0|_C) \to U$, then it holds that $x_0 = \iota$.

(ii): For every $r \in (0, \epsilon)$ the restriction $f_r|_U$ is Morse and we have $\mathrm{crit}(f_r|_U) = \mathrm{im} x_r$ where $x_r = x(\cdot, r) \colon \mathrm{crit}(\mathring{f}_0|_C) \to U$.

Recall from (49) that the Hamiltonian for the rotating Kepler problem reads

$$H \colon T^*(\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}, \quad (q,p) \mapsto \frac{1}{2}p^2 - \frac{1}{|q|} + p_1 q_2 - p_2 q_1.$$

For an energy value $c < 0$ we regularize the rotating Kepler problem via

$$\begin{aligned} K^c(p,q) &:= \frac{1}{2}\bigg(-\frac{|q|}{2c}\Big(H\big(\tfrac{q}{2c}, \sqrt{-2c}p\big) - c\Big) + 1\bigg)^2 - \frac{1}{2} \\ &= \frac{1}{2}\bigg(\frac{1}{2}(1+p^2) + \frac{(p_1 q_2 - p_2 q_1)}{(-2c)^{\frac{3}{2}}}\bigg)^2 |q|^2 - \frac{1}{2}. \end{aligned}$$

The discussion in Section 7 shows that the Hamiltonian K^c extends to a smooth Hamiltonian on T^*S^2 for every $c < 0$. By abuse of notation we denote the canonical smooth extension of K^c to T^*S^2 by the same letter. There is some small difference in the regularization above compared to the regularization in Section 1. In Section 1 we used to symplectic transformation $(p,q) \mapsto \big(-\frac{q}{\sqrt{-2c}}, \sqrt{-2c}p)$. Here we use the transformation $(p,q) \mapsto \big(\frac{q}{2c}, \sqrt{-2c}p\big)$ which is only conformally symplectic with conformal factor $\frac{1}{\sqrt{-2c}}$. For fixed c we can easily switch between the two transformations because a conformal symplectic factor can always be absorbed in the Hamiltonian in order to get the same Hamiltonian vector field. In particular, a conformal symplectic factor only gives rise to a reparametrization of the Hamiltonian flow. However, this transformation becomes problematic when one wants to study a sequence of symplectically conformal maps where the conformal factor converges to zero. This precisely we intend to do now. Namely we want to study the limit where c goes to $-\infty$. For that purpose we change the variable c in order to get the Hamiltonian

$$K_r(p,q) := K^{-\frac{1}{2r^{\frac{2}{3}}}}(p,q) = \frac{1}{2}\left(\frac{1}{2}(1+p^2) + (p_1q_2 - p_2q_1)r\right)^2 |q|^2 - \frac{1}{2}$$

for $r \in (0,\infty)$. This Hamiltonian smoothly extends to $r=0$, where it becomes

$$K_0(p,q) = \frac{1}{2}\left(\frac{1}{2}(1+p^2)\right)^2 |q|^2 - \frac{1}{2}.$$

This is just the regularized Kepler Hamiltonian which coincides with kinetic energy on the round two-sphere. In particular, its Hamiltonian flow is just the geodesic flow on the round two-sphere.

We abbreviate by $\mathfrak{L} = C^\infty(S^1, T^*S^2)$ the free loop space of T^*S^2 and consider the family

$$\mathcal{A}_r := \mathcal{A}^{K_r} \colon \mathfrak{L} \times \mathbb{R}_+ \to \mathbb{R}$$

of action functionals as defined by (57). Critical points of $\mathcal{A}_0$ correspond to geodesics on the round two-sphere. Unparametrized simple closed geodesics on the round two sphere are in one to one correspondence with unparametrized great circles. Great circles parametrized according to arc length are determined by a point and a unit direction. In particular, the space of parametrized great circles is diffeomorphic to $S^*S^2 \cong \mathbb{R}P^3$ where $S^*S^2 = \{v \in T^*S^2 : ||v|| = 1\}$ is the unit cotangent bundle of S^2. The circle S^1 acts on the space of parametrized great circles by time shift. Hence the space of unparametrized great circles is diffeomorphic to $\mathbb{R}P^3/S^1 \cong S^2$. If $\gamma \colon S^1 \to S^*S^2$ is a (parametrized) periodic orbit of the Hamiltonian vector field of K^0 corresponding to a simple closed geodesic, then because the periodic of a simple closed geodesic on the round two sphere parametrized by arc length is 2π, the tuple $(\gamma, 2\pi)$ is a critical point of $\mathcal{A}_0$. Abbreviate by

$$C \subset \mathrm{crit}\mathcal{A}_0$$

the space of all these tuples. Note that

$$C \cong \mathbb{R}P^3$$

is a Morse-Bott component of $\mathcal{A}_0$. The circle S^1 acts on $\mathfrak{L}$ by time-shift and on $\mathbb{R}_+$ trivially. Because the action functionals $\mathcal{A}_r$ are invariant under this S^1-action they induce action functionals

$$\overline{\mathcal{A}}_r \colon (\mathfrak{L} \times \mathbb{R}_+)/S^1 \to \mathbb{R}.$$

Note that S^1 acts on $\mathfrak{L} \times \mathbb{R}_+$ with finite isotropy so that the quotient $(\mathfrak{L} \times \mathbb{R}_+)/S^1$ is an orbifold. However, at $C \subset \mathfrak{L} \times \mathbb{R}_+$ the S^1-action is free and we denote the quotient by

$$\overline{C} = C/S^1 \subset \mathrm{crit}(\overline{\mathcal{A}}_0).$$

Note that

$$\overline{C} \cong \mathbb{R}P^3/S^1 \cong S^2.$$

We next study the restriction of $\mathring{\overline{\mathcal{A}}}_0$ to $\overline{C}$. Note that

$$\mathring{K}_0(p,q) = \frac{1}{2}(1+p^2)(p_1q_2 - p_2q_1)|q|^2 = \sqrt{2K_0+1}L|q| \tag{68}$$

where $L = p_1 q_2 - p_2 q_1$ is angular momentum. Note that at a point $(\gamma, \tau) \in \mathfrak{L} \times \mathbb{R}_+$ the derivative of $\mathcal{A}_r$ with respect to the r-variable at $r = 0$ is given by

$$\mathring{\mathcal{A}}_0(\gamma, \tau) = -\tau \int_{S^1} \mathring{K}_0(\gamma) dt.$$

If $(\gamma, 2\pi) = (p, q, 2\pi) \in C$ it follows that $K_0(\gamma) = 0$ and therefore

$$\mathring{\mathcal{A}}_0(\gamma, 2\pi) = -2\pi \int_{S^1} L(\gamma)|q| dt.$$

Because angular momentum L is constant along periodic orbits of the Kepler problem, we can write this as

$$\mathring{\mathcal{A}}_0(\gamma, 2\pi) = -2\pi L(\gamma(0)) \int_{S^1} |q| dt.$$

The integral $\int_{S^1} |q| dt$ gives the ratio of the period of a Kepler ellipse of energy $-\frac{1}{2}$ before and after regularization. After regularization a Kepler ellipse becomes a closed geodesic on the round two-sphere and hence has period 2π. Before regularization by the version of the third Kepler law explained in Lemma 1.1 the period for energy $-\frac{1}{2}$ is 2π as well, so that $\int_{S^1} |q| dt = 1$ independent of the orbit. Therefore we can simplify the above formula to

$$\mathring{\mathcal{A}}_0(\gamma, 2\pi) = -2\pi L(\gamma(0)).$$

The induced map $\overline{L}$ of L on the quotient $S^*S^2/S^1 \cong \mathbb{R}P^3/S^1 \cong S^2$ is just the standard height function on the two-sphere. This is not a coincidence but a very special case of a much more general fact. Indeed, the Hamiltonian vector field of $\overline{L}$ on S^2 induces a periodic flow, so that we can think of $\overline{L}$ as a moment map for a circle action on S^2. By a very special case of the convexity theorem of Atiyah-Guillemin-Sternberg [**9, 46**] we know that such a moment map is Morse all whose critical points have even index, in particular, it has no saddle points and a unique maximum and a unique minimum. It is easy to see what the critical points are in our case. By the theory of Lagrange multipliers at a critical point of L on the constraint $K_0^{-1}(0)$ the differential of L and K_0 have to be proportional to each other, so that the Hamiltonian vector fields must be parallel. This happens precisely at the circular periodic orbits of the Kepler problem. For a fixed energy value there are precisely two circular periodic orbits moving in opposite direction.

5. The retrograde and direct periodic orbit

5.1. Low energies. We have seen how the retrograde and direct periodic orbit bifurcate from the geodesic flow in the rotating Kepler problem. The phenomenon described for the rotating Kepler problem is much more general as explained by Conley [**26**] and Kummer [**73**]. It continues to hold if one adds to the rotating Kepler problem some additional velocity independent forces.

Here is the set-up. Let $\Omega \subset \mathbb{R}^2$ be an open subset containing the origin and $V : \Omega \to \mathbb{R}$ be a smooth function with the property that the origin is a critical point of V and $\mu > 0$. We consider the Hamiltonian

$$H := H_{V,\mu} : T^*(\Omega \setminus \{0\}) \to \mathbb{R}, \quad (q, p) \mapsto \frac{1}{2}p^2 - \frac{\mu}{|q|} + p_1 q_2 - p_2 q_1 + V(q).$$

An example of a Hamiltonian of this form is the Hamiltonian H_m in (50) which is obtained from the Hamiltonian of the restricted three body problem by shifting coordinates, or Hill's lunar Hamiltonian. For a given energy value $c < 0$ we regularize H by introducing the Hamiltonian

$$\begin{aligned} K^c(p,q) &= \frac{1}{2}\bigg(-\frac{|q|}{2c}\Big(H\big(\tfrac{q}{2c}, \sqrt{-2c}p\big) - c - V(0)\Big) + \mu\bigg)^2 - \frac{\mu^2}{2} \\ &= \frac{1}{2}\bigg(\frac{1}{2}(1+p^2) + \frac{(p_1q_2 - p_2q_1)}{(-2c)^{\frac{3}{2}}} - \frac{(V\big(\frac{q}{-2c}\big) - V(0))}{2c}\bigg)^2 |q|^2 - \frac{\mu^2}{2} \end{aligned}$$

As in the rotating Kepler problem we change the energy parameter and introduce

$$\begin{aligned} K_r(p,q) &:= K^{\frac{-r^{-\frac{2}{3}}}{2}}(p,q) \\ &= \frac{1}{2}\bigg(\frac{1}{2}(1+p^2) + (p_1q_2 - p_2q_1)r + \big(V(qr^{\frac{2}{3}}) - V(0)\big)r^{\frac{2}{3}}\bigg)^2 |q|^2 - \frac{\mu^2}{2}. \end{aligned}$$

Note that

$$K_0(p,q) = \frac{1}{2}\bigg(\frac{1}{2}(1+p^2)\bigg)^2 |q|^2 - \frac{\mu^2}{2}$$

does not depend on V. In particular, its flow on the energy hypersurface $K_0^{-1}(0)$ coincides with the geodesic flow on the round two sphere up to reparametrization. Abbreviate

$$G_r(p,q) := \frac{1}{2}(1+p^2) + (p_1q_2 - p_2q_1)r + \big(V(qr^{\frac{2}{3}}) - V(0)\big)r^{\frac{2}{3}}$$

so that we can write

$$K_r(p,q) = \frac{1}{2}G_r(p,q)^2|q|^2 - \frac{\mu^2}{2}. \tag{69}$$

For the first derivative of G_r with respect to the homotopy parameter r we get

$$\frac{\partial G_r}{\partial r}(p,q) = (p_1q_2 - p_2q_1) + \frac{2}{3}\big\langle \nabla V(qr^{\frac{2}{3}}), q\big\rangle r^{\frac{1}{3}} + \frac{2\big(V(qr^{\frac{2}{3}}) - V(0)\big)}{3r^{\frac{1}{3}}}$$

In particular,

$$\frac{\partial G_r}{\partial r}\bigg|_{r=0}(p,q) = p_1q_2 - p_2q_1 = L(q,p).$$

For the second derivative of G_r we obtain

$$\begin{aligned} \frac{\partial^2 G_r}{\partial r^2}(p,q) &= \frac{4}{9}\big\langle H_V(qr^{\frac{2}{3}})q, q\big\rangle + \frac{2\big\langle \nabla V(qr^{\frac{2}{3}}), q\big\rangle}{9r^{\frac{2}{3}}} + \frac{4\big\langle \nabla V(qr^{\frac{2}{3}}), q\big\rangle}{9r^{\frac{2}{3}}} \\ &\quad - \frac{2\big(V(qr^{\frac{2}{3}}) - V(0)\big)}{9r^{\frac{4}{3}}} \\ &= \frac{4}{9}\big\langle H_V(qr^{\frac{2}{3}})q, q\big\rangle + \frac{2\big\langle \nabla V(qr^{\frac{2}{3}}), q\big\rangle}{3r^{\frac{2}{3}}} - \frac{2\big(V(qr^{\frac{2}{3}}) - V(0)\big)}{9r^{\frac{4}{3}}} \end{aligned}$$

where by H_V we abbreviate the Hessian of V. Because the origin is a critical point of V we conclude

$$\begin{aligned}\frac{\partial^2 G_r}{\partial r^2}\bigg|_{r=0}(p,q) &= \frac{4}{9}\big\langle H_V(0)q,q\big\rangle + \frac{2}{3}\big\langle H_V(0)q,q\big\rangle - \frac{1}{9}\big\langle H_V(0)q,q\big\rangle \\ &= \big\langle H_V(0)q,q\big\rangle.\end{aligned}$$

In view of (69) we get for the first derivative of K_r

$$\frac{\partial K_r}{\partial r}(p,q) = G_r(p,q)|q|^2\frac{\partial G_r}{\partial r}(p,q) = \sqrt{(2K_r(p,q)+\mu^2)}|q|\frac{\partial G_r}{\partial r}(p,q)$$

so that we have

$$\frac{\partial K_r}{\partial r}\bigg|_{r=0}(p,q) = \sqrt{(2K_0+\mu^2)}|q|L.$$

In particular, this does not depend on V and therefore coincides with the computation for the rotating Kepler problem (68). For the second derivative of K_r it holds that

$$\begin{aligned}\frac{\partial^2 K_r}{\partial r^2} &= \frac{|q|}{\sqrt{2K_r+\mu^2}}\frac{\partial K_r}{\partial r}\frac{\partial G_r}{\partial r} + \sqrt{2K_r+\mu^2}|q|\frac{\partial^2 G_r}{\partial r^2} \\ &= |q|^2\Bigg(\bigg(\frac{\partial G_r}{\partial r}\bigg)^2 + G_r\frac{\partial^2 G_r}{\partial r^2}\Bigg).\end{aligned}$$

In particular, because G_r is two times continuously differentiable in r for every $r \in [0,\infty)$ the same is true for K_r.

As for the rotating Kepler problem we consider for $r \in [0,\infty)$ the family of action functionals

$$\mathcal{A}_r := \mathcal{A}^{K_r} : \mathfrak{L} \times \mathbb{R}_+ \to \mathbb{R}$$

as defined by (57) where $\mathfrak{L} = C^\infty(S^1, T^*S^2)$. Because K_r is two times differentiable and K_0 as well as $\mathring{K}_0$ do not depend on the choice of V we conclude that precisely as in the rotating Kepler problem the geodesic flow bifurcates at $r=0$ into two periodic orbits which we still refer to as the direct and retrograde periodic orbits.

5.2. Birkhoff's shooting method. Recall from (54) that trajectories of Hill's lunar problem satisfy the following second order ODE

$$\left\{\begin{aligned} q_1'' - 2q_2' &= q_1\Big(3 - \tfrac{1}{|q|^3}\Big) \\ q_2'' + 2q_1' &= -\tfrac{q_2}{|q|^3}\end{aligned}\right. \tag{70}$$

The energy constraint for Hill's lunar problem becomes

$$c = \frac{1}{2}\big((q_1')^2 + (q_2')^2\big) - \frac{1}{|q|} - \frac{3}{2}q_1^2. \tag{71}$$

The following theorem is due to Birkhoff [**17**].

Theorem 5.1. *Assume $c < -\frac{3^{\frac{4}{3}}}{2}$. Then there exists $\tau > 0$ and $(q_1,q_2)\colon [0,\tau] \to (-\infty,0] \times [0,\infty)$ solving (70), (71), and*

$$q_2(0) = 0, \quad q_1'(0) = 0, \quad q_1(\tau) = 0, \quad q_2'(\tau) = 0.$$

Proof: Consider the function

$$f_c\colon (0,\infty) \to \mathbb{R}, \quad r \mapsto c + \frac{1}{r} + \frac{3}{2}r^2.$$

Its derivative

$$f_c' = -\frac{1}{r^2} + 3r$$

has a unique zero at $r = \frac{1}{3^{\frac{1}{3}}} = 3^{-\frac{1}{3}}$ and satisfies

$$f_c'|_{\left(0,3^{-\frac{1}{3}}\right)} < 0, \quad f_c'|_{\left(3^{-\frac{1}{3}},\infty\right)} > 0.$$

In particular, f_c has a unique minimum at $r = 3^{-\frac{1}{3}}$ at which it attains the value

$$f_c(3^{-\frac{1}{3}}) = c + \frac{3^{\frac{4}{3}}}{2} < 0.$$

We conclude that there exists a unique $r_c \in (0, 3^{-\frac{1}{3}})$ such that

$$f_c(r_c) = 0.$$

Choose $r \in (0.r_c]$. Let $q^r\colon [0,T_r) \to \mathbb{R}^2$ be the solution of (70) to the initial conditions

$$q_1^r(0) = -r, \quad q_2^r(0) = 0, \quad (q_1^r)'(0) = 0, \quad (q_2^r)'(0) = \sqrt{2f_c(r)}. \tag{72}$$

In view of the initial conditions (71) holds for every t by preservation of energy. In particular, q^r lies in the bounded part of the Hill's region whose only non-compactness comes from collisions at the origin. Hence we choose $T_r \in (0,\infty]$ such that $\lim_{t\to T_r} q^r(t) = 0$ in case that T_r is finite. We introduce the quantity

$$\tau(r) := \inf\left\{t \in (0,T_r) : q_2^r(t) = 0, \text{ or } q_1^r(t) = 0\right\}.$$

Here we understand that if the set is empty, then $\tau(r) = T_r$. If $r < r_c$ in view of the initial conditions (72) we have $(q_2^r)'(0) > 0$ and therefore

$$\tau(r) > 0.$$

We claim that

$$\tau(r) < \infty, \quad r \in (0, r_c). \tag{73}$$

To see that we first integrate the first equation in (70) and use the initial condition (72)

$$\begin{aligned} (q_1^r)'(t) &= (q_1^r)'(0) + 2q_2^r(t) - 2q_2^r(0) + \int_0^t q_1^r\Big(3 - \tfrac{1}{|q^r|^3}\Big)ds \\ &= 2q_2^r(t) + \int_0^t q_1^r\Big(3 - \tfrac{1}{|q^r|^3}\Big)ds. \end{aligned} \tag{74}$$

We further note that by the initial condition and the definition of $\tau(r)$ it holds that

$$q_2^r(t) > 0, \quad q_1^r(t) < 0, \qquad 0 < t < \tau(r). \tag{75}$$

If $\mathfrak{K}_c^b$ is the bounded part of the Hill's region we claim further that

$$\mathfrak{K}_c^b \subset B_{3^{-\frac{1}{3}}}(0), \tag{76}$$

i.e., the bounded part of the Hill's region is contained in the ball of radius $3^{-\frac{1}{3}}$ centered at the origin. To prove (76) suppose that

$$(q_1, q_2) \in \partial B_{3^{-\frac{1}{3}}}(0).$$

In particular,
$$|q| = \frac{1}{3^{\frac{1}{3}}}.$$
We estimate
$$-\frac{1}{|q|} - \frac{3}{2}q_1^2 \geq -\frac{1}{|q|} - \frac{3}{2}|q|^2 = -\frac{3^{\frac{4}{3}}}{2} > c.$$
In view of the characterization (56) of the Hill's region as a sublevel set this implies that
$$\partial B_{3^{-\frac{1}{3}}}(0) \cap \mathfrak{K}_c^b = \emptyset.$$
Because $\mathfrak{K}_c^b$ is connected and contains the origin in its closure (76) follows. Combining (76) with the second inequality in (75) we conclude that
$$q_1^r\Big(3 - \tfrac{1}{|q^r|^3}\Big)(t) > 0, \quad 0 < t < \tau(r). \tag{77}$$
Because $q_1^r(0) = -r$ there exists $t_0 > 0$ such that
$$q_1^r\Big(3 - \tfrac{1}{|q^r|^3}\Big)(t) \geq \mu > 0, \quad 0 \leq t \leq t_0. \tag{78}$$
For $t_0 \leq t < \tau(r)$ we conclude from (74), (77), and (78) in combination with the first inequality in (75) that
$$(q_1^r)'(t) = 2q_2^r(t) + \int_0^t q_1^r\Big(3 - \tfrac{1}{|q^r|^3}\Big)ds \geq \int_0^{t_0} q_1^r\Big(3 - \tfrac{1}{|q^r|^3}\Big)ds \geq \mu t_0. \tag{79}$$
This implies that
$$q_1^r(t) \geq q_1^r(t_0) + \mu t_0(t - t_0), \quad t_0 \leq t < \tau(r).$$
Because the Hill's region $\mathfrak{K}_c^b$ is bounded $\tau(r)$ is finite and (73) is proved.

We define
$$\rho := \inf\big\{r \in (0, r_c) : q_1^r(\tau(r)) = 0,\ q_2^r(\tau(r)) = 0\big\}$$
with the convention that $\rho = r_c$ if the set is empty. We claim that
$$q_1^r(\tau(r)) = 0, \quad r \in (0, \rho). \tag{80}$$
In order to prove (80) we introduce the quantity
$$r_0 := \inf\big\{r \in (0, \rho) : q_2^r(\tau(r)) = 0\big\}$$
with the convention that $r_0 = \rho$ if the set is empty. We need to show that $r_0 = \rho$. For small r the trajectory q_1^r is close to the origin and we conclude from the dynamics of the Kepler problem that $q_1^r(\tau(r)) = 0$. In particular,
$$r_0 > 0.$$
We now argue by contradiction and assume that $r_0 < \rho$. That means that
$$q_2^{r_0}(\tau(r_0)) = 0, \quad (q_2^{r_0})'(\tau(r_0)) = 0, \quad (q_2^{r_0})''(\tau(r_0)) \geq 0.$$
By (79) we have
$$(q_1^{r_0})'(\tau(r_0)) > 0.$$
This contradicts the second equation in (70) and (80) is proved.

Our next claim is
$$\rho < r_c. \tag{81}$$

To prove that we consider the trajectory q^{r_c}. Its initial conditions are

$$q_1^{r_c}(0) = -r_c, \quad q_2^{r_c}(0) = 0, \quad (q_1^{r_c})'(0) = 0, \quad (q_2^{r_c})'(0) = 0.$$

The second equation in (70) implies that

$$(q_2^{r_c})''(0) = -2(q_1^{r_c})'(0) - \frac{q_2^{r_c}(0)}{|q^{r_c}(0)|^3} = 0.$$

Differentiating the second equation in (70) and using $(q_2^{r_c})'(0) = (q_1^{r_c})'(0) = 0$ we conclude

$$\begin{aligned}
(q_2^{r_c})'''(0) &= -2(q_1^{r_c})''(0) \\
&= 2(q_2^{r_c})'(0) - q_1^{r_c}(0)\left(3 - \frac{1}{|q^{r_c}(0)|^3}\right) \\
&= r_c\left(3 - \frac{1}{r_c^3}\right) \\
&< 0.
\end{aligned}$$

Here we have used for the second equality the second equation in (70) and for the last inequality the fact that $r_c \in (0, 3^{-\frac{1}{3}})$. Summarizing we have

$$q_2^{r_c}(0) = 0, \quad (q_2^{r_c})'(0) = 0, \quad (q_2^{r_c})''(0) = 0, \quad (q_2^{r_c})'''(0) < 0.$$

In particular, there exists $\epsilon > 0$ such that

$$q_2^{r_c}(t) < 0, \quad t \in (0, \epsilon). \tag{82}$$

We assume now by contradiction that $\rho = r_c$. If follows from (80) that

$$q_1^r(\tau(r)) = 0, \quad r \in (0, r_c).$$

This implies that there exists $\epsilon > 0$ such that

$$\tau(r) \geq \epsilon, \quad r \in [\tfrac{r_c}{2}, r_c).$$

In particular,

$$q_2^r(t) \geq 0, \quad r \in [\tfrac{r_c}{2}, r_c), \ t \in (0, \epsilon).$$

But this has the consequence that

$$q_2^{r_c}(t) \geq 0, \quad t \in (0, \epsilon)$$

in contradiction to (82). This contradiction proves (81).

In view of (81) the dynamics of the Kepler problem implies that $(q_2^r)'(\tau(r)) < 0$ for r close to ρ. On the other hand the dynamics of the Kepler problem implies as well that $(q_2^r)'(\tau(r)) > 0$ for r close to 0. By the intermediate value theorem we conclude that there exists $r \in (0, \rho)$ such that

$$(q_2^r)'(\tau(r)) = 0.$$

This proves the theorem. □

With more effort such a shooting argument can also be made to work for the restricted three-body problem. The upshot, also due to Birkhoff, is the existence of a retrograde orbit for all mass ratios $\mu < 1$ for energies below the first critical point. Note that such an argument can be implemented on a computer. See Figure 5.2.

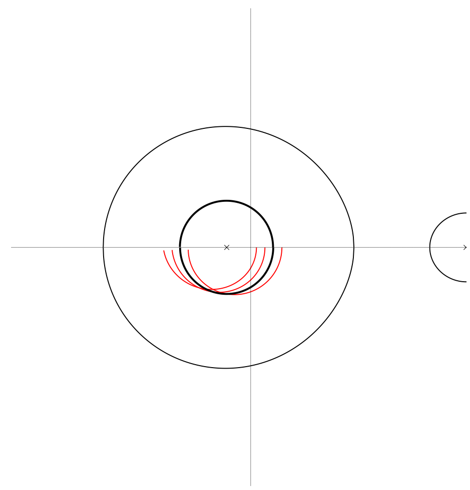

FIGURE 2. Shooting in the restricted three body problem

6. Lyapunov orbits

We first need some result about four times four matrices belonging to the Lie algebra of the symplectic group $\mathrm{Sp}(2)$. Suppose that ω is the standard symplectic form on $\mathbb{R}^4 = \mathbb{C}^2$. Recall that a 4×4-matrix B belongs to the Lie algebra of $\mathrm{Sp}(2)$ if and only if there exists a symmetric 4×4-matrix S such that

$$B = JS$$

where J is the 4×4-matrix obtained by multiplication with i as explained in (107). Note that S can be recovered from B by

$$S = -JB.$$

A symmetric matrix S is called non-degenerate when it is injective. If S is non-degenerate then the *Morse index* $\mu(S) \in \{0, 1, 2, 3, 4\}$ of S is the number of negative eigenvalues of S counted with multiplicity.

Proposition 6.1. *Assume that* $B = JS \in \mathrm{Lie}\,\mathrm{Sp}(2)$ *such that* S *is non-degenerate and* $\mu(S) = 1$*. Then there exists a symplectic basis* $\{\eta_1, \eta_2, \xi_1, \xi_2\}$ *of* $\mathbb{R}^4$ *and* $a \in (0, \infty)$*,* $b \in \mathbb{R} \setminus \{0\}$ *such that*

$$B\eta_1 = a\eta_1, \quad B\eta_2 = -a\eta_2, \quad B\xi_1 = -b\xi_2, \quad B\xi_2 = b\xi_1.$$

If $[B]_{\{\eta_1,\eta_2,\xi_1,\xi_2\}}$ denotes the matrix representation of B in the basis $\{\eta_1, \eta_2, \xi_1, \xi_2\}$ then the Proposition asserts that

$$[B]_{\{\eta_1,\eta_2,\xi_1,\xi_2\}} = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & -a & 0 & 0 \\ 0 & 0 & 0 & -b \\ 0 & 0 & b & 0 \end{pmatrix}.$$

If one complexifies $\mathbb{R}^4$ to $\mathbb{C}^4$, then with respect to the basis $\{\eta_1, \eta_2, \xi_1 + i\xi_2, \xi_1 - i\xi_2\}$ the matrix representation of B is diagonal

$$[B]_{\{\eta_1,\eta_2,\xi_1+i\xi_2,\xi_1-i\xi_2\}} = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & -a & 0 & 0 \\ 0 & 0 & ib & 0 \\ 0 & 0 & 0 & -ib \end{pmatrix}.$$

In particular, the eigenvalues of B are $\{a, -a, ib, -ib\}$.

In order to prove Proposition 6.1 we need the following well known Lemma.

Lemma 6.2. *Assume that $B = JS \in \mathrm{Lie}\,\mathrm{Sp}(2)$ has an eigenvalue $\lambda = a + ib$ with $a, b \in \mathbb{R} \setminus \{0\}$, then the spectrum of B is*

$$\mathfrak{S}(B) = \{\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda}\}.$$

Proof: Using that $J = -J^{-1} = -J^T$ and $S = S^T$ we compute

$$JBJ^{-1} = -SJ^T = -(JS)^T = -B^T$$

which shows that B is conjugated to $-B^T$. Because B and B^T have the same eigenvalues, we conclude that $-\lambda$ is an eigenvalue of B as well. Since B is real $\bar{\lambda}$ and $-\bar{\lambda}$ are eigenvalues of B as well. By assumption neither a nor b are zero so that all the four numbers $\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda}$ are different and hence the assertion about the spectrum follows. □

Proof of Proposition 6.1: We prove the Proposition in four steps. To formulate Step 1 we abbreviate by $\mathrm{Sym}_{\mathrm{inj}}(4)$ the space of injective symmetric 4×4-matrices.

Step 1: *There exists a smooth path $S_r \in \mathrm{Sym}_{\mathrm{inj}}(4)$ for $r \in [0, 1]$ with the property that*

$$S_0 = S, \quad S_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

To see that note that because S is symmetric there exists $R \in SO(4)$ such that $RSR^{-1} = RSR^T$ is diagonal. Moreover, because $\mu(S) = 1$ we can choose R such that the first three diagonal entries of RSR^{-1} are positive and the fourth is negative. Because $SO(4)$ is connected we can connect S and RSR^{-1} by a smooth path in $\mathrm{Sym}_{\mathrm{inj}}(4)$. Combining this path with convex interpolation between RSR^{-1} and S_1 the assertion of Step 1 follows.

Step 2: *The matrix B has two real eigenvalues and two imaginary eigenvalues.*

Let S_r for $r \in [0, 1]$ be the smooth path of injective symmetric matrices obtained in Step 1. This gives rise to a smooth path $B_r = JS_r \in \mathrm{Lie}\,\mathrm{Sp}(2)$. The matrix

$$B_1 = JS_1 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

has the eigenvalues $\{1,-1,i,-i\}$. Because S_r is injective 0 is not an eigenvalue of B_r for every $r \in [0,1]$. Therefore we conclude from Lemma 6.2 that B_r has two real and two imaginary eigenvalues for every $r \in [0,1]$.

The proof of Lemma 6.2 now implies that there exists $a, b \in (0,\infty)$ such that the eigenvalues of $B = B_0$ are $\{a,-a,ib,-ib\}$. In particular, there exists a basis $\{\eta_1,\eta_2,\xi_1,\xi_2\}$ of $\mathbb{R}^4$ such that

$$B\eta_1 = a\eta_1, \quad B\eta_2 = -a\eta_2, \quad B\xi_1 = -b\xi_2, \quad B\xi_2 = b\xi_1. \tag{83}$$

We next examine if we can choose this basis symplectic.

Step 3: *The symplectic orthogonal complement of the span of* $\{\eta_1,\eta_2\}$ *is spanned by* $\{\xi_1,\xi_2\}$*, i.e.,*

$$\langle\eta_1,\eta_2\rangle^\omega = \langle\xi_1,\xi_2\rangle.$$

To prove that we first note that since $B \in \mathrm{Lie}\,\mathrm{Sp}(2)$ for every $\xi,\eta \in \mathbb{R}^4$ the formula

$$\omega(B\xi,\eta) = -\omega(\xi,B\eta)$$

holds. Hence we compute

$$b\omega(\xi_1,\eta_1) = \omega(B\xi_2,\eta_1) = -\omega(\xi_2,B\eta_1) = -a\omega(\xi_2,\eta_1) \tag{84}$$

and

$$b\omega(\xi_2,\eta_1) = -\omega(B\xi_1,\eta_1) = \omega(\xi_1,B\eta_1) = a\omega(\xi_1,\eta_1). \tag{85}$$

Combining (84) and (85) we get

$$\omega(\xi_1,\eta_1) = -\frac{a}{b}\omega(\xi_2,\eta_1) = -\frac{a^2}{b^2}\omega(\xi_1,\eta_1)$$

implying that

$$\left(1+\frac{a^2}{b^2}\right)\omega(\xi_1,\eta_1) = 0$$

and therefore

$$\omega(\xi_1,\eta_1) = 0.$$

From (84) we conclude that

$$\omega(\xi_2,\eta_1) = 0$$

as well. Therefore

$$\eta_1 \in \langle\xi_1,\xi_2\rangle^\omega$$

and the same argument with η_1 replaced by η_2 leads to

$$\eta_2 \in \langle\xi_1,\xi_2\rangle^\omega.$$

Summarizing we showed that

$$\langle\eta_1,\eta_2\rangle \subset \langle\xi_1,\xi_2\rangle^\omega$$

and because ω is non-degenerate the symplectic orthogonal complement of $\langle\xi_1,\xi_2\rangle$ is two dimensional so that we get

$$\langle\eta_1,\eta_2\rangle = \langle\xi_1,\xi_2\rangle^\omega.$$

With

$$\langle\xi_1,\xi_2\rangle = \big(\langle\xi_1,\xi_2\rangle^\omega\big)^\omega = \langle\eta_1,\eta_2\rangle^\omega$$

the assertion of Step 3 follows.

Step 4: *We prove the proposition.*

It follows from Step 3 that $\langle \eta_1, \eta_2 \rangle$ is a two dimensional symplectic subspace of $\mathbb{R}^4$. Hence after scaling η_2 we can assume that

$$\omega(\eta_1, \eta_2) = 1.$$

Note that (83) still remains valid after scaling η_2. Because $\langle \xi_1, \xi_2 \rangle$ by Step 3 is a symplectic subspace as well we have

$$r := \omega(\xi_1, \xi_2) \neq 0.$$

We distinguish two cases. First assume that $r > 0$. In this case we replace ξ_1, ξ_2 by $\frac{1}{\sqrt{r}}\xi_1, \frac{1}{\sqrt{r}}\xi_2$. Then

$$\omega(\xi_1, \xi_2) = 1$$

and (83) still remains valid. It remains to treat the case $r < 0$. In this case we replace b by $-b$ and ξ_1, ξ_2 by $\frac{1}{\sqrt{-r}}\xi_1, -\frac{1}{\sqrt{-r}}\xi_2$. This finishes the proof of the proposition. $\square$

CHAPTER 6

Contacting the moon

1. A contact structure for Hill's lunar problem

The following result was proved in [**6**].

Theorem 1.1. *For any given $\mu \in [0,1)$ assume that $c < H(L_1)$, the first critical value of the restricted three body problem. Then the regularized energy hypersurface $\overline{\Sigma}_c \subset T^*S^2$ of the restricted three body problem is fiberwise star-shaped.*

For Hill's lunar problem a stronger result was proved in [**75**].

Theorem 1.2. *Assume that $c < -\frac{3^{\frac{4}{3}}}{2}$. Then the regularized energy hypersurface $\overline{\Sigma}_c \subset T^*S^2$ of Hill's lunar problem is fiberwise convex.*

Remark 1.3. *It is an open problem if the regularized energy hypersurface of the restricted three body problem for energies below the first critical value are fiberwise convex as well.*

Remark 1.4. *Because $\overline{\Sigma}_c$ is fiberwise star-shaped it follows that it is contact. In particular, a result of Cristofaro-Gardiner and Hutchings* [**28**] *implies that $\overline{\Sigma}_c$ admits two closed characteristics. We already know the existence of one, namely the retrograde periodic orbit from the work of Birkhoff. Lee's result implies that below the first critical value the regularized energy hypersurfaces of Hill's lunar problem are Finsler and in this case the existence of two closed characteristics was already proved by Bangert and Long* [**10**]*. The existence of two closed characteristics in Hill's lunar problem was first proved by Llibre and Roberto in* [**79**]*.*

In this chapter we explain why below the first critical value the energy hypersurface of Hill's lunar problem is fiberwise star-shaped. This is much weaker than the result of Lee [**75**]. However, the advantage of the proof presented here is that the same scheme can also be applied in the restricted three body problem to prove fiberwise starshapedness [**6**], although there the argument gets much more involved.

Note that the vector field
$$X = q\frac{\partial}{\partial q}$$
is a Liouville vector field on $T^*\mathbb{R}^2$. Indeed, for $\omega = dp \wedge dq$ we obtain by Cartan's formula
$$\mathcal{L}_X\omega = d\iota_X\omega + \iota_X d\omega = d(-qdp) = -dq \wedge dp = dp \wedge dq = \omega.$$
Our next theorem tells us that this Liouville vector field is transverse to Σ_c^b and therefore $\iota_X\omega$ defines a contact structure on Σ_c^b.

Proposition 1.5. *Assume that $c < -\frac{3^{\frac{4}{3}}}{2}$. Then $X \pitchfork \Sigma_c^b$.*

Proof: In polar coordinates (r, θ), i.e.,

$$q_1 = r\cos\theta, \quad q_2 = r\sin\theta$$

the Liouville vector field reads

$$X = r\frac{\partial}{\partial r},$$

the effective potential becomes

$$U = -\frac{1}{r} - \frac{3}{2}r^2\cos^2\theta$$

and the Hamiltonian

$$H = \frac{1}{2}\big((p_1 + r\sin\theta)^2 + (p_2 - r\cos\theta)^2\big) + U.$$

To prove the theorem it suffices to show that

$$dH(X)|_{\Sigma_c^b} > 0. \tag{86}$$

We estimate using the Cauchy-Schwarz inequality

$$\begin{aligned} dH(X) &= r\sin\theta(p_1 + r\sin\theta) - r\cos\theta(p_2 - r\cos\theta) + r\frac{\partial U}{\partial r} \\ &\geq r\frac{\partial U}{\partial r} - \sqrt{r^2\sin^2\theta + r^2\cos^2\theta}\sqrt{(p_1 + r\sin\theta)^2 + (p_2 - r\cos\theta)^2} \\ &= r\frac{\partial U}{\partial r} - r\sqrt{2(H-U)}. \end{aligned}$$

This implies that

$$dH(X)|_{\Sigma_c^b} \geq r\bigg(\frac{\partial U}{\partial r} - \sqrt{2(c-U)}\bigg).$$

Note that the right hand side is independent of the variables p_1 and p_2. Therefore to prove (86) it suffices to show

$$\bigg(\frac{\partial U}{\partial r} - \sqrt{2(c-U)}\bigg)\bigg|_{\mathfrak{K}_c^b} > 0. \tag{87}$$

Pick $(r, \theta) \in \mathfrak{K}_c^b$. In particular,

$$U(r, \theta) \leq c.$$

By (76) the bounded part of Hill's region is contained in the ball of radius $3^{-\frac{1}{3}}$ centered at the origin. Observe that

$$U|_{\partial B_{3^{-\frac{1}{3}}}(0)} \geq -\frac{3^{\frac{4}{3}}}{2}.$$

Since $c < -\frac{3^{\frac{4}{3}}}{2}$ it follows that there exists

$$\tau \in \Big[0, \tfrac{1}{3^{\frac{1}{3}}} - r\Big)$$

such that

$$U(r+\tau, \theta) = c. \tag{88}$$

We claim that

$$\frac{\partial U}{\partial r}(q) > 0, \quad q \in B_{3^{-\frac{1}{3}}}(0) \setminus \big\{0, (0, -3^{-\frac{1}{3}}), (0, 3^{-\frac{1}{3}})\big\}. \tag{89}$$

To see that we estimate

$$\frac{\partial U}{\partial r} = \frac{1}{r^2} - 3r\cos^2\theta \geq \frac{1}{r^2} - 3r \geq 0.$$

If $r < 3^{-\frac{1}{3}}$ the inequality is strict and if $r = 3^{-\frac{1}{3}}$ the first inequality is strict because $\cos^2\theta < 1$ because we removed the points $(0, -3^{-\frac{1}{3}})$ and $(0, 3^{-\frac{1}{3}})$ from the ball. This proves (89).

We further claim that

$$\frac{\partial^2 U}{\partial r^2}(q) \leq -1, \quad q \in B_{3^{-\frac{1}{3}}}(0) \setminus \{0\}. \tag{90}$$

In order to prove that we estimate

$$\frac{\partial^2 U}{\partial r^2} = -\frac{2}{r^3} - 3\cos^2\theta \leq -\frac{2}{r^3} \leq -6 \leq -1.$$

In order to prove (87) we estimate using (88), (89), and (90)

$$\begin{aligned}
\left(\frac{\partial U}{\partial r}(r,\theta)\right)^2 &= \left(\frac{\partial U}{\partial r}(r+\tau,\theta)\right)^2 - \int_0^\tau \frac{d}{dt}\left(\frac{\partial U(r+t,\theta)}{\partial r}\right)^2 dt \\
&> -2\int_0^\tau \frac{\partial U(r+t,\theta)}{\partial r}\frac{\partial^2 U(r+t,\theta)}{\partial r^2} dt \\
&\geq 2\int_0^\tau \frac{\partial U(r+t,\theta)}{\partial r} dt \\
&= 2\big(U(r+\tau,\theta) - U(r,\theta)\big) \\
&= 2\big(c - U(r,\theta)\big).
\end{aligned}$$

Using (89) once more this implies

$$\frac{\partial U}{\partial r}(r,\theta) > \sqrt{2\big(c - U(r,\theta)\big)}.$$

Therefore

$$\left(\frac{\partial U}{\partial r} - \sqrt{2(c-U)}\right)(r,\theta) > 0.$$

This proves (87) and the Proposition follows. □

Note that the map $(q,p) \mapsto (-p,q)$ is a symplectomorphism of $T^*\mathbb{R}^2$ to itself which interchanges the roles of position and momentum. Hence the base coordinate q becomes the fiber coordinate where the fiber coordinate p becomes the base coordinate after the transformation. Interchanging the roles of q and p in this way the assertion of Theorem 1.5 can be interpreted as the fact that Σ_c^b is fiberwise star-shaped in $T^*\mathbb{R}^2$, i.e., if $p \in \mathbb{R}^2$ then the fiber

$$\Sigma_{c,p}^b := T_p^*\mathbb{R}^2 \cap \Sigma_c^b \subset T_p^*\mathbb{R}^2 = \mathbb{R}^2$$

bounds a star-shaped domain.

CHAPTER 7

Global surfaces of section

1. Disk-like global surfaces of section

The concept of a global surface of section was introduced by Poincaré shortly before his death in [**91**]. Assume that $X \in \Gamma(TS^3)$ is a non-vanishing vector field on the three dimensional sphere S^3. We denote by ϕ_X^t its flow on S^3.

Definition 1.1. *A* (disk-like) global surface of section *is an embedded disk $D \subset S^3$ satisfying*

- **(i):** *X is tangent to ∂D, the boundary of D,*
- **(ii):** *X is transverse to the interior $\mathring{D}$ of D,*
- **(iii):** *For every $x \in S^3$ there exists $t^+ > 0$ and $t^- < 0$ such that $\phi_X^{t^+}(x) \in \mathring{D}$ and $\phi_X^{t^-}(x) \in \mathring{D}$.*

Remark 1.2. *Requirement (i) implies that the boundary ∂D of the disk is a periodic orbit of X. We refer to ∂D as the* bounding orbit *of the global surface of section.*

Remark 1.3. *Instead of a disk one could consider more generally a Riemann surface with boundary. In particular, an important example is an annulus which has two bounding orbits, see Figure 1. However, in the following we concentrate ourselves on disks and mean by a global surface of section always a disk-like global surface of section unless specified otherwise.*

Let us now assume that $D \subset S^3$ is a global surface of section. We define the *Poincaré return map*

$$\psi\colon \mathring{D} \to \mathring{D}$$

as follows. Given $x \in \mathring{D}$, define

$$\tau(x) := \min\big\{t > 0 : \phi_X^t(x) \in \mathring{D}\big\},$$

i.e., the next return time of x to D. It follows from the conditions of a global surface of section that $\tau(x)$ exists and is finite. Now define

$$\psi(x) := \phi_X^{\tau(x)}(x).$$

If $x \in \mathring{D}$ and $\xi \in T_x\mathrm{D}$ the differential of the Poincaré return map is given by

$$d\psi(x)\xi = d\phi^{\tau(x)}(x)\xi + (d\tau(x)\xi)X. \tag{91}$$

The two dimensional disk D together with the Poincaré return map ψ basically contains all the relevant information on the flow of X on the three dimensional manifold S^3. For example periodic orbits of X different from the bounding orbit ∂D correspond to periodic points of the Poincaré return map. One can say that a global surface of section reduces the complexity of the problem by one dimension.

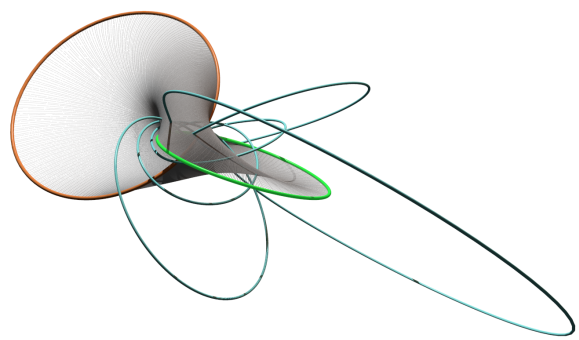

FIGURE 1. The existence of an annular surface of section has been known in the component of the heavy primary for small μ since Poincaré and Birkhoff. This displays such a surface after stereographic projection of the Levi-Civita regularization.

Recall that a Hamiltonian structure on S^3 is a closed two-form $\omega \in \Omega^2(S^3)$ with the property that $\ker \omega$ is a one-dimensional distribution. A non-vanishing section $X \in \Gamma(\ker \omega)$ is referred to as a Hamiltonian vector field. In view of (91) the following Lemma follows.

Lemma 1.4. *Assume that $\omega \in \Omega^2(S^3)$ is a Hamiltonian structure, $X \in \Gamma(\ker \omega)$, and D is a global surface of section for X with Poincaré return map ψ. Then*

$$\psi^* \omega|_{\mathring{D}} = \omega|_{\mathring{D}},$$

i.e., ψ is area preserving with respect to the restriction of ω to the interior of the global surface of section.

2. Obstructions

Given a non-vanishing vector field X on S^3 it is far from obvious that the dynamical system (S^3, X) admits a global surface of section. Moreover, given a periodic orbit γ of the vector field X we would like to know if it bounds a global surface of section. A periodic orbit can be interpreted as a knot in S^3, i.e. an isotopy class of embeddings of $S^1 \to S^3$. The first obstruction is obvious.

Obstruction 1:: If a periodic orbit bounds a global surface of section, then it is unknotted.

Indeed, since the periodic orbit is the boundary of an embedded disk it has Seifert genus zero. This is a characterizing property of the unknot. The second obstruction describes the relation of the binding orbit with all other periodic orbits.

Obstruction 2:: If a periodic orbit bounds a global surface of section, it is linked to every other periodic orbit.

Indeed, a periodic orbit different from the binding orbit is a periodic point of the Poincaré return map, i.e., coincides with a fixed point of an iteration of the Poincaré return map. Because X is transverse to the interior of the global surface of section, each intersection point of the periodic orbit with the global surface of section counts with the same sign and therefore the linking number does not vanish.

The third obstruction holds in the case where the vector field X is the Reeb vector field of a contact form λ on S^3. Abbreviate by $\xi = \ker\lambda$ the hyperplane distribution in T^*S^3. In contact geometry there are two important classes of knots. A *Legendrian knot* is an embedding $\gamma\colon S^1 \to S^3$ with the property that $\partial_t\gamma(t) \in \xi_{\gamma(t)}$ for every $t \in S^1$. The other extreme is a *transverse knots*, which is an embedding $\gamma\colon S^1 \to S^3$ with the property that $\partial_t\gamma(t) \notin \xi_{\gamma(t)}$ for every $t \in S^1$. We refer to the book by Geiges [**40**] about a detailed discussion of Legendrian and transverse knots in contact topology. A periodic Reeb orbit is an example of a transverse knot and therefore we restrict our discussion in the following to transverse knots.

The third obstruction for a periodic orbit of a Reeb flow to bound a global surface of section is that its *selflinking number* has to be equal to minus one. We explain this notion for a transverse unknot $\gamma\colon S^1 \to S^3$. Abbreviating by$D = \{z \in \mathbb{C} : |z| \leq 1\}$ the unit disk in $\mathbb{C}$ we first choose an embedding

$$\overline{\gamma}\colon D \to S^3$$

with the property that

$$\overline{\gamma}(e^{2\pi i t}) = \gamma(t).$$

That such an embedding exists follows from the assumption that γ is the unknot. Consider the vector bundle of rank two $\overline{\gamma}^*\xi \to D$. Because D is contractible we can choose a non-vanishing section $X\colon D \to \overline{\gamma}^*\xi$. Fix a Riemannian metric g on S^3 and define

$$\gamma_X\colon S^1 \to S^3, \quad t \mapsto \exp_{\gamma(t)} X(t).$$

Because γ is a transverse knot we can choose X so small such that

$$\gamma_{rX} \cap \gamma = \emptyset, \quad r \in (0,1].$$

We define the selflinking number of γ to be

$$s\ell(\gamma) := \ell k(\gamma, \gamma_X) \in \mathbb{Z}$$

where ℓk is the linking number. By homotopy invariance of the linking number the selflinking number of γ does not depend on the choice of the section X and the Riemannian metric g. It is as well independent of the choice of the embedded filling disk $\overline{\gamma}$. To see that note that if $\overline{\gamma}\colon D \to S^3$ and $\overline{\gamma}'\colon D \to S^3$ are two embedded filling disks of γ in view of the fact that $\pi_2(S^3) = \{0\}$ the two filling disks are homotopic. Even if the homotopy is not through embedded disks we can use it to construct for a given non-vanishing section $X\colon D \to \overline{\gamma}^*\xi$ a non-vanishing section $X'\colon D \to (\overline{\gamma}')^*\xi$ with the property that the restrictions of the two sections to the boundary $\partial D = S^1$ satisfy

$$X|_{\partial D} = X'|_{\partial D}\colon S^1 \to \gamma^*\xi.$$

This shows that the selflinking number is independent of the choice of the embedded filling disk as well.

Remark 2.1. *The selflinking number can as well be defined for transverse knots* $\gamma\colon S^1 \to S^3$ *which are not necessarily unknots. One uses here the fact that for every knot there exists an oriented surface with boundary* Σ *and an embedding* $\overline{\gamma}\colon \Sigma \to S^3$ *with the property that* $\overline{\gamma}|_{\partial\Sigma} = \gamma$. *Such a surface is referred to as a* Seifert surface, *see for example* [**77**], *and with the help of a Seifert surface one defines the selflinking number of a transverse knot similarly as for unknots. We refer to the book of Geiges* [**40**] *for details. For the unknot the Seifert surface can be chosen as a disk and if one defines the genus of a knot to be*

$$g(\gamma) := \min\big\{g(\Sigma) : \Sigma \textit{ Seifert surface of } \gamma\big\} \in \mathbb{N} \cup \{0\}.$$

then the unknot can be characterized as the knot whose genus vanishes.

We are now ready to formula the third obstruction for a periodic orbit to be the bounding orbit of a global surface of section.

Obstruction 3:: If the vector field on S^3 coincides with the Reeb vector field of a contact form on S^3, and a periodic Reeb orbit γ bounds a global surface of section, then its selflinking number satisfies $s\ell(\gamma) = -1$.

We next explain the reason for Obstruction 3. Let $D \subset S^3$ be a global surface of section. We then have two rank-2 vector bundles over the disk $\xi|_D \to D$ and $TD \to D$. Abbreviate by

$$\pi\colon TS^3 \to \xi$$

the projection along the Reeb vector field R. Since R is transverse to the interior of D we obtain a bundle isomorphism

$$\pi|_{\mathring{D}}\colon T\mathring{D} \to \xi|_{\mathring{D}}.$$

Choose

$$X\colon D \to \xi$$

a non-vanishing section. If (r,θ) are polar coordinates on D we define another section from D to ξ by

$$Y := \pi(r\partial_r)\colon D \to \xi.$$

Because ∂D is a periodic Reeb orbit it holds that $\partial_\theta|_{\partial D}$ is parallel to the Reeb vector field and therefore $Y|_{\partial D}\colon \partial D \to \xi$ is nonvanishing. In particular, we have two nonvanishing section $X|_{\partial D}, Y|_{\partial D}\colon \partial D \to \xi$. The selflinking number of the Reeb orbit γ is then given by

$$s\ell(\gamma) = s\ell(\partial D) = \mathrm{wind}_{\partial D}(Y,X) = -\mathrm{wind}_{\partial D}(X,Y)$$

where $\mathrm{wind}_{\partial D}(Y,X)$ is the winding number of X around Y. For $r \in (0,1]$ abbreviate $D_r = \{z \in D : |z| \leq r\}$ the ball of radius r. The section X is nonvanishing on D while the section Y only vanishes at 0. Therefore the winding number

$$\mathrm{wind}_{\partial D_r}(Y,X) \in \mathbb{Z}$$

is defined for every $r \in (0,1]$. By homotopy invariance of the winding number we conclude that

$$\mathrm{wind}_{\partial D}(Y,X) = \mathrm{wind}_{\partial D_r}(Y,X), \quad \forall\, r \in (0,1].$$

We now look at the situation for $r = \delta$ close to 0. Since $X(0)$ is nonvanishing we see that $Y = \pi(r\partial_r)$ winds once around $X(0)$ and therefore

$$\mathrm{wind}_{\partial D_\delta}(X,Y) = 1.$$

Combining these facts we conclude that

$$s\ell(\gamma) = -1.$$

We point out that to derive Obstruction 3 we only used the local assumption (i) and (ii) of Definition 1.1 and not the global assumption (iii).

3. Existence results from holomorphic curve theory

In the following we assume that $\Sigma \subset \mathbb{C}^2$ is a star-shaped hypersurface. It follows that the restriction of the one-form

$$\lambda = \frac{1}{2}(x_1 dy_1 - y_1 dx_1 + x_2 dy_2 - y_2 dx_2)$$

to Σ defines a contact form, see Example 6.6. The following theorem is due to Hryniewicz [**58, 59**].

Theorem 3.1 (Hryniewicz). *Assume that $\Sigma \subset \mathbb{C}^2$ is a star-shaped hypersurface and $\gamma \in C^\infty(S^1, \Sigma)$ is an unknotted periodic Reeb orbit of period τ whose selflinking number satisfies $s\ell(\gamma) = -1$. Further assume that one of the following conditions holds*

- **(i):** *The period τ of γ is minimal among the periods of all periodic Reeb orbits.*
- **(ii):** *Σ is dynamically convex.*

Then γ bounds a global surface of section. Moreover, each periodic orbit which corresponds to a fixed point of the Poincaré return map of the global surface of section is unknotted and has selflinking number -1.

Remark 3.2. *The fact that a periodic orbit corresponding to a fixed point of the Poincaré return map itself is unknotted and has selflinking number -1 has an interesting consequence. Namely, if Σ is dynamically convex we can apply Theorem 3.1 again to this orbit to see that it bounds as well a global surface of section. This leads to Hryniewicz theory of systems of global surfaces of section* [**59**].

Remark 3.3. *Under the assumption that the star-shaped hypersurface $\Sigma \subset \mathbb{C}^2$ satisfies some non-degeneracy condition, there exists an interesting improvement of the theorem of Hryniewicz, which is due to Hryniewicz and Salomão. Namely if one assumes that γ and all periodic orbits of Σ different from γ of period less than to the period τ of γ are non-degenerate and none of these orbits has Conley-Zehnder index 2 and is unlinked to γ, then γ still bounds a global surface of section.*

The following Theorem is due to Hofer, Wysocki, and Zehnder [**56**].

Theorem 3.4 (Hofer-Wysocki-Zehnder). *Assume that $\Sigma \subset \mathbb{C}^2$ is a star-shaped, dynamically convex hypersurface, then there exists an unknotted periodic Reeb orbit on Σ with selflinking number -1.*

Combining the above two theorems we immediately obtain the following Corollary.

Corollary 3.5 (Hofer-Wysocki-Zehnder). *Assume that $\Sigma \subset \mathbb{C}^2$ is a star-shaped, dynamically convex hypersurface, then Σ admits a global surface of section.*

Remark 3.6. *Historically, this Corollary was first proved by Hofer-Wysocki-Zehnder in the same paper* [**56**] *in which they established the existence of an unknotted periodic Reeb orbit of selflinking number* -1*, before the Theorem of Hryniewicz was available. This comes from the fact that both Hryniewicz as well as Hofer, Wysocki, and Zehnder use holomorphic curves to prove their theorems. Therefore Hofer, Wysocki, and Zehnder could use them to prove directly the existence of a global surface of section in their groundbreaking work* [**56**].

Remark 3.7. *Using the global surface of section obtained in Corollary 3.5 Hryniewicz shows in* [**59**] *that each periodic orbit which corresponds to a fixed point of the Poincaré return map is unknotted and has selflinking number* -1*. In particular, Theorem 3.1 implies that these periodic orbits bound a global surface of section as well.*

4. Contact connected sum - the archenemy of global surfaces of section

We briefly recall the connected sum of two smooth manifolds. Suppose that M_1 and M_2 are two oriented n-dimensional manifolds. We will talk about balls D^n which we give their standard orientation. For M_1, choose an embedded ball $\iota_1 : D^n \to M_1$, where ι_1 is orientation preserving. For M_2, choose an embedded ball $\iota_2 : D^n \to M_2$ which reserves orientation. Intuitively, we take out small balls from M_1 and M_2 and glue collar neighborhoods together.

We do this more precisely. Write D_r for the ball with radius r, so

$$D_r = \{z \in D^n \mid \|z\| < r\}.$$

Fix a number R with $0 < R < 1$. We will use the annulus $A := D^n - \overline{D_R}$, and the orientation reversing map

$$\begin{aligned} r : A &\longrightarrow A \\ x &\longmapsto (1 + R - \|x\|) \cdot \frac{x}{\|x\|}. \end{aligned}$$

In addition, the map reverses the inner and outer sphere of the annulus, meaning the spheres with radii R and 1, respectively. We define the **connected sum** of M_1 and M_2 as

$$M_1 \# M_2 := M_1 \setminus \iota_1(\overline{D_R}) \coprod M_2 \setminus \iota_2(\overline{D_R}) / \sim$$

where $\sim$ is an equivalence relation. Namely, if $\tilde{x} \in M_1$ is given by $\tilde{x} = \iota_1(x)$, and $\tilde{y} \in M_2$ is given by $\tilde{y} = \iota_2(y)$, then we say that $\tilde{x} \sim \tilde{y}$ if and only if $r(x) = y$. Other points are not related. Geometrically, the above just means that we glue the two annuli $\iota_1(A)$ and $\iota_2(A)$ together by reversing inner and outer spheres.

Lemma 4.1. *The connected sum* $M_1 \# M_2$ *defines an oriented smooth manifold.*

Proof: Clearly $M_1 \setminus \iota_1(\overline{D_R})$ and $M_2 \setminus \iota_2(\overline{D_R})$ are smooth manifolds, and the above equivalence relation glues along open sets, so we obtain the structure of a smooth manifold on $M_1 \# M_2$. To see that we get an orientation, we observe that ι_2 and r are orientation reversing, so their composition is orientation preserving. □

4.1. Contact version. The contact version mimics the above construction, but instead of gluing the two annuli directly together, we define a model for the connecting tube. The discussion here follows Weinstein's ideas, and we have drawn a picture describing the construction of the tube in Figure 4.1. This model will appear again when we look at level sets of Hamiltonians.

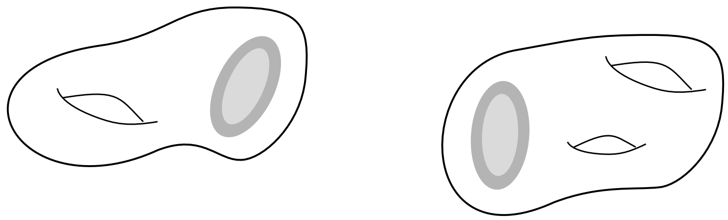

FIGURE 2. Connected sum of smooth, oriented manifolds

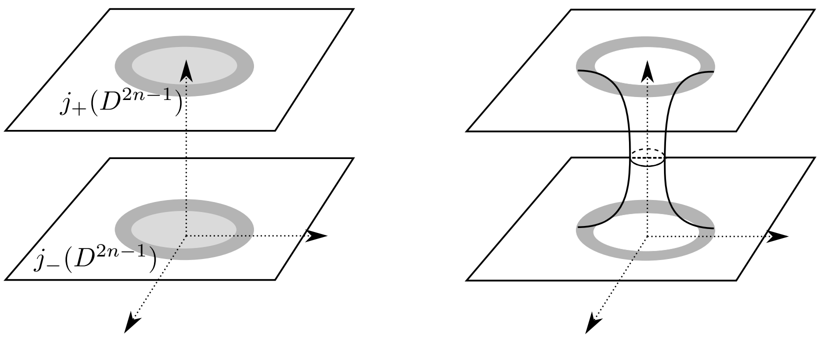

FIGURE 3. The tube of the contact connected sum

Consider contact manifolds (M_1^{2n-1}, α_1) and (M_2^{2n-1}, α_1) together with two points, say $q_1 \in M_1$ and $q_2 \in M_2$. We want to define the contact connected sum. As before, choose a number $0 < R < 1$ for the radius of the ball we are going to cut out.

Choose Darboux balls $\iota_k : D^{2n-1} \to M_k$ containing q_k for $k = 1, 2$. To construct the tube, we also embed these Darboux balls into $\mathbb{R}^{2n}$ by the map

$$\begin{aligned} j_\pm : D^{2n-1} &\longrightarrow \mathbb{R}^{2n} \\ (x, y; z) &\longmapsto (x, y, z, \pm 1) \end{aligned}$$

We observe that the vector field

$$X = \frac{1}{2}\left(x \cdot \partial_x + y \cdot \partial_y + 2z\partial_z - w\partial_w\right)$$

is a Liouville vector field and that it is transverse to $j_\pm(D^{2n-1})$. Furthermore, it induces the Liouville form

$$i_X\omega_0 = \frac{1}{2}\left(xdy - ydx\right) + 2zdw + wdz,$$

which restricts to the contact form $\frac{1}{2}(xdy - ydx) + 2dz$ on $j_+(D^{2n-1})$ and to the contact form $\frac{1}{2}(xdy - ydx) - 2dz$ on $j_-(D^{2n-1})$. These are the standard contact form with the standard orientation, and a variation of the standard contact form with the *opposite* orientation. To construct the tube, choose a smooth, increasing function $f : \mathbb{R} \to \mathbb{R}$ with the following properties

- $f(z) = 1$ if $z > R$
- $f(0) < 0$ and $f(\epsilon) = 0$ and $f'(\epsilon) > 0$.

We define the connecting tube as the set

$$\mathcal{T} := \{(x, y; z, w) \in \mathbb{R}^{2n-2} \times \mathbb{R}^2 \mid w^2 = f(|x|^2 + |y|^2 + z^2)\} \cap \{|x|^2 + |y|^2 + z^2 < 1\}$$

Lemma 4.2. *The connecting tube is a smooth submanifold of $\mathbb{R}^{2n}$. Furthermore, the Liouville vector field $X = \frac{1}{2}(x\partial_x + y\partial_y) + 2z\partial_z - w\partial_w$ is transverse to $\mathcal{T}$, so $\mathcal{T}$ is a contact manifold. For $(x, y; z)$ with $|x|^2 + |y|^2 + z^2 > R$, the contact form coincides with the standard contact form on $j_\pm(D^{2n-1})$.*

Proof: Since $\mathcal{T}$ is a level set of the function $F : (x, y; z, w) \mapsto f(|x|^2 + |y|^2 + z^2) - w^2$, it suffices to check that $X(F) \neq 0$. Indeed, this shows that the Jacobian has everywhere full rank, and of course it shows that X is transverse to $\mathcal{T}$. We now compute

$$X(F) = f'(|x|^2 + |y|^2 + z^2) \cdot (|x|^2 + |y|^2 + 4z^2) + 2w^2.$$

Since f is increasing, all terms are non-negative. To see that their sum is always positive on $\mathcal{T}$, we note the following.

- if $w \neq 0$, then $2w^2$ is positive.
- if $w = 0$, then $f(|x|^2 + |y|^2 + z^2) = 0$ and $|x|^2 + |y|^2 + z^2 = \epsilon$, and by assumption $f'(\epsilon) > 0$.

The last claim follows since $\mathcal{T}$ and $\iota_j(D^{2n-1})$ coincide if $|x|^2 + |y|^2 + z^2 > R$, and their contact forms are induced by the same Liouville vector field. □

We now define the contact connected sum by

$$(M_1, \alpha_1)\#(M_2, \alpha_2) M_1 \setminus \iota_1(\overline{D_R}) \coprod M_2 \setminus \iota_2(\overline{D_R}) \coprod \mathcal{T}/\sim .$$

Here the equivalence relation is defined as follows.

- If $\tilde{x} = \iota_1(x)$ lies in $M_1 \setminus \iota_1(\overline{D_R})$ and $\tilde{y} = j_1(y)$ lies in $\mathcal{T}$, then $\tilde{x} \sim \tilde{y}$ if and only if $x = y$.
- If $\tilde{x} = \iota_2(x)$ lies in $M_2 \setminus \iota_2(\overline{D_R})$ and $\tilde{y} = j_2(y)$ lies in $\mathcal{T}$, then $\tilde{x} \sim \tilde{y}$ if and only if $x = y$.
- Other points are not related.

Theorem 4.3. *The space $(M_1, \alpha_1)\#(M_2, \alpha_2)$ is a contact manifold.*

Proof: The proof that this is a smooth manifold follows the same argument as before: since we glue along open sets, we clearly get a differentiable atlas. To see that this is a contact manifold, we only need to observe that $M_1 \setminus \iota_1(\overline{D_R})$, $M_2 \setminus \iota_2(\overline{D_R})$ and $\mathcal{T}$ are contact manifolds with contact forms that patch together with the gluing relation. □

5. Invariant global surfaces of section

Assume that $\rho \in \mathrm{Diff}(S^3)$ is a smooth involution on the three dimensional sphere, i.e., $\rho^2 = \mathrm{id}$ and $X \in \Gamma(TS^3)$ is a nonvanishing vector field on S^3 which is anti-invariant under the involution ρ in the sense that

$$\rho^* X = -X. \tag{92}$$

A (disk-like) global surface of section $D \subset S^3$ is called *invariant* if $\rho(D) = D$. By abuse of notation we denote the restriction of ρ to D again by the same letter.

Lemma 5.1. *Assume that $D \subset S^3$ is an invariant global surface of section. Then the fixed point set $\mathrm{Fix}(\rho) \subset D$ is a simple arc intersecting the boundary ∂D transversally.*

Proof: We first note that ρ is an orientation reversing involution of D. Indeed, this follows from (92) in view of the fact that the vector field is tangent to the boundary ∂D. Therefore $\mathrm{Fix}(\rho)$ is a one dimensional submanifold of D. That it is transverse to ∂D follows again from (92). It follows from Brouwer's fixed point theorem, see for example [**48**, Theorem 1.9.], that $\mathrm{Fix}(\rho)$ is not empty. In particular, it is a finite union of circles and intervals. We claim that there are no circles. To see that we argue by contradiction and assume that the fixed point set of ρ contains a circle. The complement of this circle consists of two connected components one of them containing the boundary of ∂D. The involution ρ then has to interchange these two connected components. However, the boundary of ∂D is invariant under ρ and this leads to the desired contradiction. Consequently, the fixed points set consists just of a finite union of intervals. It remains to show that there is just one interval. To see that note that the complement of an interval consists again of two connected components which are interchanged by ρ. Therefore there cannot be additional fixed points and the lemma is proved. □

Remark 5.2. *The Lemma above is actually an easy case of a much more general result due to Brouwer* [**22**] *and Kérékjartò* [**69**]*, which says that a topological involution just defined in the interior of the disk is topologically conjugated to a reflection at a line. We refer to* [**27**] *for a modern exposition of this result.*

We observe in particular, that the binding orbit of an invariant global surface of section is necessarily a symmetric periodic orbit. Moreover, if $\psi\colon \mathring{D} \to \mathring{D}$ is the Poincaré return map it follows as in (15) that ψ satisfies with ρ the commutation relation

$$\rho\psi\rho = \psi^{-1}. \tag{93}$$

6. Fixed points and periodic points

In his lifelong quest for periodic orbits Poincaré introduced the concept of a global surface of section in [**90**], because in the presence of a global surface of section the search for periodic orbits is reduced to the search of periodic points of the Poincaré return map. Originally Poincaré thought of a global surface of section as an annulus type global surface of section and it was Birkhoff who showed in [**16**] that an area preserving map of an annulus which moves the two boundary components in different directions has at least two fixed points. In [**37, 38**] Franks proved the following theorem

Theorem 6.1 (Franks). *Assume an area preserving homeomorphism of an open annulus admits a periodic point, then it admits infinitely many periodic points.*

On the other hand Brouwer's translation theorem [**21**] asserts that

Theorem 6.2 (Brouwer). *An area preserving homeomorphism of the open disk admits a fixed point.*

If we restrict an area preserving homeomorphism to the complement of one of its fixed points we obtain an area preserving homeomorphism of the open annulus. This implies the following Corollary.

Corollary 6.3. *An area preserving homeomorphism of an open disk admits either one or infinitely many periodic points.*

In view of Lemma 1.4 together with the fact that if a vector field admits a global surface of section then periodic orbits of the vector field different form the bounding orbit of the global surface of section are in one to one correspondence with periodic points of the Poincaré return map we obtain the further Corollary.

Corollary 6.4. *Assume that $\omega \in \Omega^2(S^3)$ is a Hamiltonian structure on S^3, $X \in \Gamma(\ker \omega)$ a Hamiltonian vector field whose flow admits a global surface of section. Then X has either two or infinitely periodic orbits.*

The fixed point guaranteed by Theorem 6.2 is of special interest in view of Theorem 3.1, because under the assumptions of this theorem the periodic orbit corresponding to the fixed point is itself unknotted and has selflinking number -1 so that it bounds in the dynamically convex case another global surface of section. The amazing thing about Theorem 6.2 is that the homeomorphism of the open disk is not required to extend continuously to the boundary of the disk. If it does than Theorem 6.2 just follows from Brouwer's fixed point theorem, see for example [**48**, Theorem 1.9]. In this case the assumption that the homeomorphism is area preserving is not needed at all. However, if the homeomorphism does not extend continuously to the boundary the condition, that the map is area preserving, is essential. In fact, the open disk is just homeomorphic to the two dimensional plane and a translation of the plane is an example of a homeomorphism without fixed points. To prove Brouwer's translation theorem one first shows that if an orientation preserving homeomorphism of the open disk has a periodic point it has to have a fixed point. We refer to [**32**] for a modern account of this remarkable fact. Hence one is left with the case that the homeomorphism has no periodic point at all and one shows in this case that is has to be a translation which then contradicts the assumption that the homeomorphism is area preserving. A modern treatment of this second step together with the precise definition what a translation is can be found in [**36**], see also [**47**]. It was observed by Kang in [**65**] that a quite different argument for this fact can be given if the global surface of section is symmetric. Moreover, in this case one can find a fixed point of the Poincaré return map which corresponds to a symmetric periodic orbit. We discuss this in the next section.

7. Reversible maps and symmetric fixed points

Let (D, ω) be a closed two-dimensional disk together with an area form $\omega \in \Omega^2(D)$ and suppose that $\rho \in \mathrm{Diff}(D)$ is an anti-symplectic involution, i.e.,

$$\rho^2 = \mathrm{id}, \quad \rho^*\omega = -\omega.$$

Moreover, suppose that $\psi \in \mathrm{Diff}(\mathring{D})$ is an area preserving diffeomorphism of the interior of the disk, i.e., it holds that

$$\psi^*\omega = \omega,$$

which satisfies with ρ the commutation relation

$$\rho\psi\rho = \psi^{-1}. \tag{94}$$

This is the situation one faces by (93) if one considers the Poincaré return map of a symmetric global surface of section for a Hamiltonian vector field of a Hamiltonian structure on S^3. A diffeomorphism satisfying (94) is called *reversible*. The following result was proved by Kang in [**65**].

Lemma 7.1 (Kang). *Under the above assumptions there exists a common fixed point of* ρ *and* ψ *in* $\mathring{D}$*, i.e., a point* $x \in \mathring{D}$ *satisfying*

$$\rho(x) = x, \quad \psi(x) = x.$$

Proof: In view of (94) we obtain

$$(\psi\rho)^2 = \psi\psi^{-1} = \mathrm{id}$$

so that $\psi\rho$ is again an involution. Moreover, because ρ is anti-symplectic and ψ is symplectic the composition is anti-symplectic as well, so that we have

$$(\psi\rho)^*\omega = -\omega.$$

By Lemma 5.1 we know that $\mathrm{Fix}(\rho) \subset D$ is a simple arc intersecting the boundary ∂D transversally. Because ρ is anti-symplectic we conclude that both connected components of the complement of $\mathrm{Fix}(\rho)$ have the same area.

The anti-symplectic involution $\psi\rho$ is only defined in the interior of the disk. However, by the Theorem of Brouwer and Kérékjartò mentioned in Remark 5.2 the involution $\psi\rho$ is topologically conjugated to the reflection at a line and therefore its fixed point set still consists of an arc whose complement has two connected components. Because $\psi\rho$ is anti-symplectic the two connected components of the complement of its fixed point set have the same area as well.

It follows that $\mathrm{Fix}(\rho)$ and $\mathrm{Fix}(\psi\rho)$ intersect. This means that there exists $x \in \mathring{D}$ with the property that

$$\rho(x) = x, \quad \psi\rho(x) = x$$

or equivalently

$$\rho(x) = x, \quad \psi(x) = x.$$

This finishes the proof of the Lemma. □

If $\psi\colon \mathring{D} \to \mathring{D}$ is a reversible map we call following [**65**] a point $x \in \mathring{D}$ a *symmetric periodic point* of ψ, if there exists $k, \ell \in \mathbb{N}$ with the property that

$$\psi^k(x) = x, \quad \psi^\ell(x) = \rho(x).$$

If $k = \ell = 1$ then the symmetric periodic point is called a *symmetric fixed point* whose existence was discussed in Lemma 7.1. The minimal k for which $\psi^k(x) = x$ holds is referred to as the *period* of x and abbreviated by $k(x)$. Then there exists a unique

$$\ell(x) \in \mathbb{Z}/k(x)\mathbb{Z}$$

such that

$$\psi^{\ell(x)}(x) = \rho(x).$$

Note that the period only depends on the orbit of x, in particular it holds that

$$k(\psi(x)) = k(x).$$

This is not true for $\ell(x)$. In particular, from (94) together with the fact that ρ is an involution we obtain the relation

$$\rho = \psi\rho\psi$$

using that we compute

$$\psi^{\ell(x)+1}(x) = \psi\rho(x) = \psi^2\rho\psi(x)$$

implying that

$$\psi^{\ell-2}(\psi(x)) = \rho(\psi(x)).$$

Therefore it holds that

$$\ell(\psi(x)) = \ell(x) - 2 \in \mathbb{Z}/k(x)\mathbb{Z}.$$

If the period is odd then there exists a unique point in the orbit of x which lies on the fixed point set $\mathrm{Fix}(\rho)$. If the period is even, then there are two cases. Either ℓ is even as well for every point in the orbit of x and there are precisely two points in the orbit of x which lie on $\mathrm{Fix}(\rho)$ or ℓ is odd for every point in the orbit of x and there is no point in the orbit of x which lies on the fixed point set of ρ.

In [**65**] Kang proved the following analogue of Franks theorem (Theorem 6.1) for reversible maps.

Theorem 7.2 (Kang). *Assume an area preserving reversible homeomorphism of an open annulus admits a periodic point, then it admits infinitely many symmetric periodic points.*

Remark 7.3. *A surprising feature of Kang's theorem is the fact that the periodic point does not need to be symmetric in order to guarantee infinitely many symmetric periodic points.*

Combining Lemma 7.1 with Theorem 7.2 we obtain the following Corollary.

Corollary 7.4. *A reversible anti-symplectic map of the open disk admits either one or infinitely many symmetric periodic points.*

Because orbits of symmetric periodic of the Poincaré return map of an invariant global surface of section together with the binding orbit correspond to symmetric periodic orbits we obtain further the following Corollary.

Corollary 7.5. *Assume that $\omega \in \Omega^2(S^3)$ is a Hamiltonian structure on S^3 which is anti-invariant under an involution of S^3 and $X \in \Gamma(\ker \omega)$ is a Hamiltonian vector field whose flow admits an invariant global surface of section. Then X has either two or infinitely many symmetric periodic orbits.*

CHAPTER 8

The Maslov Index

Assume that (V,ω) is a finite dimensional symplectic vector space. By choosing a symplectic basis $\{e_1,\cdots,e_n,f_1,\cdots,f_n\}$, i.e., a basis of V satisfying

$$\omega(e_i,e_j)=\omega(f_i,f_j)=0,\ \omega(e_i,f_j)=\delta_{ij},\ 1\le i,j\le n$$

we can identify V with $\mathbb{C}^n$ by mapping a vector $\xi=\xi^1+\xi^2\in V$ with $\xi^1=\sum_{j=1}^n \xi_j^1 e_j$, $\xi^2=\sum_{j=1}^n \xi_j^2 f_j$ to the vector $(\xi_1^1+i\xi_1^2,\ldots,\xi_n^1+i\xi_n^2)\in\mathbb{C}^n$. The *Lagrangian Grassmannian*

$$\Lambda=\Lambda(n)$$

is the manifold consisting of all Lagrangian subspaces $L\subset\mathbb{C}^n$. For example, since any 1-dimensional linear subspace in $\mathbb{C}$ is Lagrangian, we have

$$\Lambda(1)=\mathbb{R}P^1\cong S^1.$$

The Lagrangian Grassmannian has the structure of a homogeneous space. To see that we first observe that the group $U(n)$ acts on the Lagrangian Grassmannian

$$U(n)\times\Lambda(n)\to\Lambda(n),\quad (U,L)\mapsto UL.$$

That this action is transitive can be seen in the following way. On a given Lagrangian $L\subset\mathbb{C}^n$ choose an orthonormal basis $e'_1,\cdots,e'_n$ of L, where orthonormal refers of course to the standard inner product on $\mathbb{C}^n$. Putting $f'_j=ie_j$ for $1\le j\le n$ gives rise to a symplectic, orthonormal basis $\{e'_1,\ldots,e'_n,f'_1,\ldots,f'_n\}$ of $\mathbb{C}^n$. Now define $U\colon\mathbb{C}^n\to\mathbb{C}^n$ as the linear map which maps e_j to e'_j and f_j to f'_j. This proves transitivity. The stabilizer of the $U(n)$ action on $\Lambda(n)$ can be identified with the group $O(n)$, namely the ambiguity in choosing an orthonormal basis on the Lagrangian subspace. Therefore, the Lagrangian Grassmannian becomes the homogeneous space

$$\Lambda(n)=U(n)/O(n).$$

Following [**8**] we next discuss the fundamental group of the Lagrangian Grassmannian. Consider the map

$$\rho\colon U(n)/O(n)\to S^1,\quad [A]\mapsto\det A^2.$$

Note that because the determinant of a matrix in $O(n)$ is plus or minus one, the map ρ is well-defined, independent of the choice of the representative $A\in U(n)$. This gives rise to a fiber bundle

$$\begin{array}{ccc} SU(n)/SO(n) & \longrightarrow & U(n)/O(n) \\ & & \big\downarrow \rho \\ & & S^1 \end{array}$$

Since the fiber $SU(n)/SO(n)$ is simply connected the long exact homotopy sequence tells us that the induced homomorphism

$$\rho_* \colon \pi_1\big(U(n)/O(n)\big) \to \pi_1(S^1)$$

is an isomorphism. We state this fact as a theorem.

Theorem 0.6. *The fundamental group of the Lagrangian Grassmannian satisfies*

$$\pi_1(\Lambda) \cong \mathbb{Z}.$$

Moreover, an explicit isomorphism is given by the map $\rho_* \colon \pi_1(\Lambda) \to \pi_1(S^1)$.

If $\lambda\colon S^1 \to \Lambda$ is a continuous loop of Lagrangian subspaces we obtain a continuous map

$$\rho \circ \lambda \colon S^1 \to S^1$$

and we define the *Maslov index of a loop* as

$$\mu(\lambda) := \deg(\rho\lambda) \in \mathbb{Z}.$$

In view of Theorem 0.6 we can alternatively characterize the Maslov index as

$$\mu(\lambda) = [\lambda] \in \pi_1(\Lambda) = \mathbb{Z}.$$

It is not clear from this definition how to generalize the definition of the Maslov index from a loop of Lagrangian subspaces to a *path* of Lagrangian subspaces $\lambda\colon [0,1] \to \Lambda$. In order to find such a generalization which is needed to define the Conley-Zehnder index, we next discuss Arnold's characterization of the Maslov index as an intersection number with the *Maslov (pseudoco)cycle* [**8**].

In order to define the Maslov pseudo-cocycle we fix a basepoint $L_0 \in \Lambda(n)$ and define for $0 \leq k \leq n$

$$\Lambda^k = \Lambda^k_{L_0}(n) = \{L \in \Lambda(n) : \dim(L \cap L_0) = k\}.$$

For each $0 \leq k \leq n$ the space Λ^k is a submanifold of Λ and the whole Lagrangian Grassmannian is stratified as

$$\Lambda = \bigcup_{k=0}^{n} \Lambda^k. \tag{95}$$

We need the following proposition

Proposition 0.7. *For* $0 \leq k \leq n$ *the codimension of* $\Lambda^k \subset \Lambda$ *is given by*

$$\mathrm{codim}(\Lambda^k, \Lambda) = \frac{k(k+1)}{2}.$$

Proof: We first compute the dimension of the Lagrangian Grassmannian. Because $\Lambda(n) = U(n)/O(n)$ we obtain

$$\begin{aligned} \dim \Lambda(n) &= \dim U(n)/O(n) = \dim U(n) - \dim O(n) \\ &= n^2 - \frac{n(n-1)}{2} = \frac{n(n+1)}{2}. \end{aligned} \tag{96}$$

Recall further that the usual Grassmannian

$$G(n,k) = \{V \subset \mathbb{R}^n : \dim V = k\}$$

of k-planes in $\mathbb{R}^n$ can be interpreted as the homogeneous space

$$G(n,k) = O(n)/O(k) \times O(n-k)$$

and therefore its dimension is given by

$$\begin{aligned}(97)\quad \dim G(n,k) &= \dim O(n) - \dim O(k) - \dim O(n-k)\\ &= \frac{n(n-1)}{2} - \frac{k(k-1)}{2} - \frac{(n-k)(n-k-1)}{2} = k(n-k).\end{aligned}$$

Given a Lagrangian $L \in \Lambda^k_{L_0}(n)$, denote by $(L \cap L_0)^\perp$ the orthogonal complement of $L \cap L_0$ in L_0. We obtain a symplectic splitting

$$\mathbb{C}^n = \Big(L \cap L_0 \oplus i(L \cap L_0)\Big) \oplus \Big((L \cap L_0)^\perp \oplus i(L \cap L_0)^\perp\Big) =: V_0 \oplus V_1.$$

Note that V_0 and V_1 are symplectic subvector spaces of $\mathbb{C}^n$ of dimension $\dim V_0 = 2k$ and $\dim V_1 = 2(n-k)$. Moreover, $(L \cap L_0)^\perp \subset V_1$ as well as $L \cap V_1$ are Lagrangian subspaces of V_1 which satisfy $(L \cap L_0)^\perp \cap (L \cap V_1) = \{0\}$. Therefore we can interpret

$$L \cap V_1 \in \Lambda^0_{(L \cap L_0)^\perp}(n-k).$$

Furthermore, $L \cap V_0 = L \cap L_0$ is a k-dimensional subspace in L and therefore the projection

$$\Lambda^k_{L_0}(n) \to G(n,k), \quad L \mapsto L \cap L_0$$

gives rise to a fiber bundle

$$\begin{array}{ccc} \Lambda^0(n-k) & \longrightarrow & \Lambda^k(n) \\ & & \downarrow \\ & & G(n,k). \end{array}$$

Since $\Lambda^0(n-k)$ is an open subset of $\Lambda(n-k)$ its dimension equals by (96)

$$\dim \Lambda^0(n-k) = \frac{(n-k)(n-k+1)}{2}.$$

Using (97) we obtain

$$\begin{aligned}(98)\quad \dim \Lambda^k(n) &= \dim \Lambda^0(n-k) + \dim G(n,k)\\ &= \frac{(n-k)(n-k+1)}{2} + k(n-k) = \frac{n(n+1)}{2} - \frac{k(k+1)}{2}.\end{aligned}$$

Combining (96) and (98) the Proposition follows. $\square$

Note that if we choose $k = 1$ in the above Proposition we obtain that $\Lambda^1 \subset \Lambda$ is a codimension 1 submanifold of Λ. However, Λ^1 is in general noncompact, its closure is given by

$$\overline{\Lambda^1} = \bigcup_{k=1}^n \Lambda^k.$$

In particular, the boundary of Λ^1 is given by

$$\overline{\Lambda^1} \setminus \Lambda^1 = \bigcup_{k=2}^n \Lambda^k$$

and from Proposition 0.7 we deduce that

$$\dim \Lambda^1 - \dim \Lambda^k \geq 2$$

which means that the boundary of Λ^1 has dimension at least 2 less than Λ^1. This property is crucial for intersection theoretic arguments with Λ^1.

If Λ^1 is in addition orientable it meets the requirements of a pseudo-cycle as in [**81**]. Results of Kahn, Schwarz, and Zinger [**63, 98, 109**], then show how to associate a homology class to a pseudocycle. Unfortunately, Λ^1 is not always orientable. This is related to a Theorem of D. Fuks [**39**] which tells us that the Lagrangian Grassmannian $\Lambda(n)$ is orientable if and only if n is odd. As was noted by Arnold [**8**] Λ^1 is always coorientable. Hence if n is odd $\Lambda^1(n)$ is a pseudocycle in the sense of McDuff and Salamon. Fortunately, the intersection number of Λ^1 with a loop of Lagrangian subspaces can always be defined independently of these theoretic consideration due to the fact that Λ^1 can be cooriented. Our next goal is to show how this coorientation works.

We first describe some natural charts for the Lagrangian Grassmannian. Given $L_1 \in \Lambda$ choose a Lagrangian complement L_2 of L_1, i.e., $L_2 \in \Lambda$ such that $L_2 \cap L_1 = \{0\}$. It follows that $L_1 \in \Lambda^0_{L_2}$. We now explain how to construct a vector space structure on $\Lambda^0_{L_2}$ for which L_1 becomes the origin. This is done by identifying $\Lambda^0_{L_2}$ with $S^2(L_1)$, the quadratic forms on L_1. Namely given $L \in \Lambda^0_{L_2}$ for each $v \in L_1$ there exists a unique $w_v \in L_2$ such that $v + w_v \in L$. We define

$$\Lambda^0_{L_2} \to S^2(L_1), \quad L \mapsto Q_L = Q_L^{L_1, L_2} \tag{99}$$

where

$$Q_L(v) = \omega(v, w_v), \quad v \in L_1.$$

We describe this procedure in coordinates. Namely we choose a basis $\{e_1, \dots, e_n\}$ of L_1. Then there exists a unique basis $\{f_1, \dots, f_n\}$ of the Lagrangian complement L_2 of L_1 such that $\{e_1, \dots, e_n, f_1, \dots, f_n\}$ is a symplectic basis of $\mathbb{C}^n$. Using these bases we identify $L_1 = \mathbb{R}^n \subset \mathbb{C}^n$ and $L_2 = i\mathbb{R}^n \subset \mathbb{C}^n$. Now if $v = v_1 + iv_2, w = w_1 + iw_2 \in \mathbb{C}^n$ with $v_1, v_2, w_1, w_2 \in \mathbb{R}^n$ the symplectic form is given by

$$\omega(v, w) = \langle v_1, w_2 \rangle - \langle v_2, w_1 \rangle$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{R}^n$. Now if $L \subset \mathbb{C}^n$ is a Lagrangian satisfying $L \cap L_2 = \{0\}$ we can write L as

$$L = \{x + iSx : x \in \mathbb{R}^n\} =: \Gamma_S$$

namely as the graph of a linear map $S \colon \mathbb{R}^n \to \mathbb{R}^n$. The fact that L is Lagrangian translates into the fact that for every $x, y \in \mathbb{R}^n$ we have

$$0 = \omega(x + iSx, y + iSy) = \langle x, Sy \rangle - \langle y, Sx \rangle$$

which is equivalent to the assertion that

$$\langle x, Sy \rangle = \langle y, Sx \rangle, \quad x, y \in \mathbb{R}^n$$

meaning that S is symmetric. Now for $x \in \mathbb{R}^n$ we compute

$$Q_L(x) = \omega(x, iSx) = \langle x, Sx \rangle.$$

Remark 0.8. *The natural charts lead to a new proof of the dimension formula for the Lagrangian Grassmannian. Indeed,*

$$\dim S^2(L_1) = \frac{n(n+1)}{2}$$

in accordance with (96).

We next show how the natural charts lead to an identification of the tangent space $T_{L_1}\Lambda$ with the quadratic form $S^2(L_1)$ on L_1. Indeed, given a Lagrangian complement L_2 of L_1, i.e., an element $L_2 \in \Lambda^0_{L_1}$, the natural chart obtained from L_2 leads to a vector space isomorphism

$$\Phi_{L_2} \colon T_{L_1}\Lambda \to S^2(L_1), \quad L \mapsto Q_L^{L_1,L_2}.$$

Our next Lemma shows that this vector space isomorphism does not depend on the chart chosen.

Lemma 0.9. *The vector space isomorphism Φ_{L_2} is independent of L_2, i.e., we have a canonical identification of $T_{L_1}\Lambda$ with $S^2(L_1)$.*

Proof: We can assume without loss of generality that $L_1 = \mathbb{R}^n \subset \mathbb{C}^n$. Then an arbitrary Lagrangian complement of L_1 is given by

$$L_2 = \{By + iy : y \in \mathbb{R}^n\}$$

where B is a real symmetric $n \times n$ matrix. Now consider a smooth path $L \colon (-\epsilon, \epsilon) \to \Lambda$ satisfying $L(0) = L_1$ and such that

$$L(t) = \{x + iA(t)x : x \in \mathbb{R}^n\}$$

where $A(t)$ is a symmetric matrix such that $A(0) = 0$. For $x \in \mathbb{R}^n = L_1$ we next compute $Q_{L(t)}^{L_1,L_2}(x)$. For that purpose let $w_x(t)$ be the unique vector in L_2 such that

$$x + w_x(t) \in L(t).$$

Since $w_x(t) \in L_2$ there exists a unique $y(t) \in \mathbb{R}^n$ such that

$$w_x(t) = By(t) + iy(t).$$

We therefore obtain

$$x + w_x(t) = x + By(t) + iy(t) \in L(t)$$

which implies that

$$y(t) = A(t)(x + By(t)). \tag{100}$$

Moreover,

$$Q_{L(t)}^{L_1,L_2}(x) = \omega(x, w_x(t)) = \omega(x, By(t) + iy(t)) = \langle x, y(t)\rangle$$

such that

$$\frac{d}{dt}\bigg|_{t=0} Q_{L(t)}^{L_1,L_2}(x) = \langle x, y'(0)\rangle. \tag{101}$$

Since $L(0) = L_1$ it follows that $w_x(0) = 0$ and therefore $y(0) = 0$. Using this as well as $A(0) = 0$ we obtain from differentiating (100)

$$y'(0) = A'(0)(x + By(0)) + A(0)(x + By'(0)) = A'(0)x.$$

Plugging this into (101) we get

$$\frac{d}{dt}\bigg|_{t=0} Q_{L(t)}^{L_1,L_2}(x) = \langle x, A'(0)x\rangle. \tag{102}$$

This is independent of B and therefore $\frac{d}{dt}\big|_{t=0} Q_{L(t)}^{L_1,L_2}(x)$ does not depend on the choice of the Lagrangian complement L_2. This proves the lemma. $\square$

In view of the above Lemma we denote for $L \in \Lambda$ and $\widehat{L} \in T_L\Lambda$ the uniquely determined quadratic form on L by

$$Q^{\widehat{L}} \in S^2(L).$$

We can use this form to characterize the tangent space of $\Lambda^k_{L_0}$.

Lemma 0.10. *Assume that $L_0 \in \Lambda$, $k \in \{0, \dots, n\}$, and $L \in \Lambda^k_{L_0}$, then*

$$T_L\Lambda^k_{L_0} = \Big\{\widehat{L} \in T_L\Lambda : Q^{\widehat{L}}\big|_{L_0 \cap L} = 0\Big\}.$$

Proof: We decompose $\mathbb{C}^n$ as

$$\mathbb{C}^n = \mathbb{R}^k \times \mathbb{R}^{n-k} \times \mathbb{R}^k \times \mathbb{R}^{n-k}.$$

We can assume without loss of generality that

$$L = \mathbb{R}^k \times \mathbb{R}^{n-k} \times \{0\} \times \{0\} = \mathbb{R}^n \times \{0\}$$

and

$$L_0 = \mathbb{R}^k \times \{0\} \times \{0\} \times \mathbb{R}^{n-k}.$$

Suppose that $A\colon \mathbb{R}^n \to \mathbb{R}^n$ is a matrix satisfying $A = A^T$. We decompose

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where $A_{11}\colon \mathbb{R}^k \to \mathbb{R}^k$, $A_{12}\colon \mathbb{R}^{n-k} \to \mathbb{R}^k$, $A_{21}\colon \mathbb{R}^k \to \mathbb{R}^{n-k}$, $A_{22}\colon \mathbb{R}^{n-k} \to \mathbb{R}^{n-k}$ satisfy

$$A_{11} = A_{11}^T, \quad A_{22} = A_{22}^T, \quad A_{12} = A_{21}^T.$$

The graph of A then can be written as

$$\Gamma_A = \{(x, y, A_{11}x + A_{12}y, A_{21}x + A_{22}y) : x \in \mathbb{R}^k,\ y \in \mathbb{R}^{n-k}\}.$$

Now suppose that $(x_1, y_1, x_2, y_2) \in \Gamma_A \cap L_0$. Since $(x_1, y_1, x_2, y_2) \in \Gamma_A$ we obtain $x_2 = A_{11}x_1 + A_{12}y_1$ and $y_2 = A_{21}x_1 + A_{22}y_1$. Because $(x_1, y_1, x_2, y_2) \in L_0$ we must have $x_2 = y_1 = 0$ and therefore $A_{11}x_1 = 0$ and $y_2 = A_{21}x_1$. We have proved that

$$\Gamma_A \cap L_0 = \{(x_1, 0, 0, A_{21}x_1) : A_{11}x_1 = 0\}. \tag{103}$$

If $\widehat{L} \in T_L\Lambda$ we can write $\widehat{L}$ as

$$\widehat{L} = \frac{d}{dt}\bigg|_{t=0} \Gamma_{tA} : \quad A\colon \mathbb{R}^n \to \mathbb{R}^n,\ A = A^T.$$

In view of (102) the associated quadratic form is given by

$$Q^{\widehat{L}}\colon L \times L \to \mathbb{R}, \quad z \mapsto \langle z, Az\rangle,\ z = (x, y) \in \mathbb{R}^k \times \mathbb{R}^{n-k} = \mathbb{R}^n.$$

Its restriction to $L \cap L_0$ becomes

$$Q^{\widehat{L}}\big|_{L \cap L_0}\colon L \cap L_0 \times L \cap L_0 \to \mathbb{R}, \quad x \mapsto \langle x, A_{11}x\rangle,\ x \in \mathbb{R}^k. \tag{104}$$

From (103) and (104) we deduce that the dimension of $\Gamma_{tA} \cap L_0 = k$ for every $t \in \mathbb{R}$ if and only if $Q^{\widehat{L}}\big|_{L \cap L_0} = 0$. This implies that

$$T_L\Lambda^k_{L_0} \subset \Big\{\widehat{L} \in T_L\Lambda : Q^{\widehat{L}}\big|_{L_0 \cap L} = 0\Big\}.$$

On the other hand we know from (98) that

$$\dim T_L\Lambda^k_{L_0} = \frac{n(n+1)}{2} - \frac{k(k+1)}{2} = \dim \Big\{\widehat{L} \in T_L\Lambda : Q^{\widehat{L}}\big|_{L_0 \cap L} = 0\Big\}$$

and therefore

$$T_L\Lambda^k_{L_0} = \Big\{\widehat{L} \in T_L\Lambda : Q^{\widehat{L}}\big|_{L_0\cap L} = 0\Big\}.$$

This finishes the proof of the Lemma. □

In view of the previous Lemma we can now define the *coorientation* of $\Lambda^1 = \Lambda^1_{L_0}$ in Λ as follows. If $L \in \Lambda^1$ we define

$$[\widehat{L}] \in T_L\Lambda/T_L\Lambda^1 \text{ positive } \iff Q^{\widehat{L}}\big|_{L_0\cap L} \text{ positive.}$$

Indeed, Lemma 0.10 tells us that this definition is well defined, independent of the choice of $\widehat{L}$ in its equivalence class in $T_L\Lambda/T_L\Lambda^1$.

We now use the coorientation of Λ^1 to give an alternative characterization of the Maslov index via intersection theory. We fix $L_0 \in \Lambda$ and define for a loop $\lambda\colon S^1 \to \Lambda$ the Maslov index as intersection number between λ and $\Lambda^1 = \Lambda^1_{L_0}$. In order to do that we need the following definition.

Definition 0.11. *Assume that $\lambda\colon S^1 \to \Lambda$ is a smooth map. We say that λ intersects Λ^1* transversally *if and only if the following two conditions are satisfied.*

(i): *For every $t \in S^1$ such that $\lambda(t) \in \Lambda^1$ we have*

$$\mathrm{im}(d\lambda(t)) + T_{\lambda(t)}\Lambda^1 = T_{\lambda(t)}\Lambda.$$

(ii): *For every $k \geq 2$ it holds that*

$$\mathrm{im}\lambda \cap \Lambda^k = \emptyset.$$

One writes $\lambda \pitchfork \Lambda^1$ if λ intersects Λ^1 transversally. It follows from Sard's theorem [**83**, Chapter 2] and the fact that $\mathrm{codim}(\Lambda^k, \Lambda) \geq 3$ for $k \geq 2$ that after small perturbation of λ we can assume that $\lambda \pitchfork \Lambda^1$. In fact, for this step $\mathrm{codim}(\Lambda^k, \Lambda) \geq 2$ for $k \geq 2$ would already be sufficient, however, we soon need $\mathrm{codim}(\Lambda^k, \Lambda) \geq 3$ for $k \geq 2$ to show invariance of the intersection number under homotopies.

In the following let us assume that $\lambda \pitchfork \Lambda^1$. It follows that $\lambda^{-1}(\Lambda^1)$ is a finite set. For $t \in \lambda^{-1}(\Lambda^1)$ we use the coorientation of Λ^1 to define

$$\nu(t) := \left\{\begin{array}{cc} 1 & \partial_t\lambda(t) \in T_{\lambda(t)}\Lambda/T_{\lambda(t)}\Lambda^1 \text{ positive} \\ -1 & \text{else.} \end{array}\right.$$

We define now the intersection number of λ with Λ^1 as

$$\widetilde{\mu}(\lambda) := \sum_{t\in\lambda^{-1}(\Lambda^1)} \nu(t).$$

Here we understand that if $t \notin \lambda^{-1}(\Lambda^1)$ then $\nu(t) = 0$ so that only finitely many summand in the above sum are different from zero.

Theorem 0.12. *The intersection number with the Maslov cycle coincides with the Maslov index, i.e., $\widetilde{\mu} = \mu$.*

Proof: We prove the theorem in two steps.

Step 1: $\widetilde{\mu}(\lambda)$ only depends on the homotopy class of λ.

In order to prove Step 1 assume that $\lambda_0, \lambda_1 \pitchfork \Lambda^1$ are two loops of Lagrangian subspaces which are homotopic to each other, i.e., we can choose a smooth map

$$\lambda\colon S^1 \times [0,1] \to \Lambda$$

such that

$$\lambda(\cdot, 0) = \lambda_0, \quad \lambda(\cdot, 1) = \lambda_1.$$

Again taking advantage of Sard's theorem and the fact that by Lemma 0.7 $\mathrm{codim}(\Lambda^k, \Lambda) \geq 3$ for every $k \geq 2$ we can assume maybe after perturbing the homotopy λ that $\lambda \pitchfork \Lambda^1$ in the sense that the following two conditions are met.

(i): For every $(t, r) \in S^1 \times [0, 1]$ such that $\lambda(t, r) \in \Lambda^1$ it holds that

$$\mathrm{im}(d\lambda(t, r)) + T_{\lambda(t,r)}\Lambda^1 = T_{\lambda(t,r)}\Lambda.$$

(ii): For every $k \geq 2$ we have

$$\mathrm{im}\lambda \cap \Lambda^k = \emptyset.$$

It now follows from the implicit function theorem that $\lambda^{-1}(\Lambda^1) \subset S^1 \times [0, 1]$ is a one dimensional manifold with boundary. The boundary is given by

$$\partial\Big(\lambda^{-1}(\Lambda^1)\Big) = \Big(\lambda_0^{-1}(\Lambda^1) \times \{0\}\Big) \cup \Big(\lambda_1^{-1}(\Lambda^1) \times \{1\}\Big).$$

The manifold $\lambda^{-1}(\Lambda^1)$ can be oriented as follows. We orient the cylinder by declaring the basis $\{\partial_r, \partial_t\}$ to be positive at every point $(t, r) \in S^1 \times [0, 1]$. Suppose that $(t, r) \in \lambda^{-1}(\Lambda^1)$ and $v \neq 0 \in T_{(t,r)}\lambda^{-1}(\Lambda^1)$. Choose $w \in T_{(t,r)}(S^1 \times [0, 1])$ such that $\{v, w\}$ is a positive basis of $T_{(t,r)}(S^1 \times [0, 1])$. We now declare v to be a positive basis of the one dimensional vector space $T_{(t,r)}\lambda^{-1}(\Lambda^1)$ if and only if $[d\lambda(w)] \in T\Lambda/T\Lambda^1$ is positive. Let us check that this notion is well defined, independent of the choice of w. However, if w' is another choice it follows that $w' = aw + bv$ where $a > 0$ and $b \in \mathbb{R}$. Since $d\lambda(v) \in T\Lambda^1$ we obtain $[d\lambda(w')] = a[d\lambda(w)]$ which due to the positivity of a is positive if and only if $[d\lambda(w)]$ is positive.

Since $\lambda \pitchfork \Lambda^1$ we have $\lambda^{-1}(\Lambda^1) = \lambda^{-1}(\overline{\Lambda^1})$ and therefore $\lambda^{-1}(\Lambda)$ is compact. It follows, see [**83**, Appendix] that a compact one dimensional manifold with boundary is a finite union of circles and intervals. For a compact submanifold of the cylinder $S^1 \times [0, 1]$ there are three types of intervals. Either both boundary points lie in the first boundary component $S^1 \times \{0\}$, or both boundary points lie in the second component $S^1 \times \{1\}$, or finally one boundary point lies in $S^1 \times \{0\}$ where the other boundary point lies in $S^1 \times \{1\}$. In the first case both boundary points contribute with opposite signs $\widetilde{\mu}(\lambda_0)$, in the second case they contribute with opposite signs to $\widetilde{\mu}(\lambda_1)$ and in the third case they contribute with the same sign to $\widetilde{\mu}(\lambda_0)$ and $\widetilde{\mu}(\lambda_1)$. This proves that $\widetilde{\mu}(\lambda_0) = \widetilde{\mu}(\lambda_1)$ and finishes the proof of the first step.

Step 2: We prove the theorem.

As a consequence of Step 1 the intersection number $\widetilde{\mu}$ induces a homomorphism from $\pi_1(\Lambda)$ to $\mathbb{Z}$. By Theorem 0.6 we know that $\pi_1(\Lambda) = \mathbb{Z}$. Hence it suffices to show that $\widetilde{\mu}$ agrees with μ on a generator of $\pi_1(\Lambda) = \mathbb{Z}$. Such a generator is given by

$$\lambda(t) = (e^{i\pi t}, 1, \ldots, 1)\mathbb{R}^n \subset \mathbb{C}^n, \quad t \in [0, 1].$$

For this path we have

$$\rho \circ \lambda(t) = \det \begin{pmatrix} e^{i\pi t} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}^2 = e^{2\pi i t}$$

so that we obtain for the Maslov index

$$\mu(\lambda) = \deg(t \mapsto e^{2\pi i t}) = 1.$$

It remains to compute $\widetilde{\mu}(\lambda)$. A priori the computation of $\widetilde{\mu}(\lambda)$ depends on the choice of the basepoint $L_0 \in \Lambda$ used to define the Maslov cycle. However, the Lagrangian Grassmannian is connected so that between any two basepoints we can always find a smooth path such that a homotopy argument analogous to the one in Step 1 shows that $\widetilde{\mu}(\lambda)$ is independent of the choice of L_0. Taking advantage of this freedom we choose

$$L_0 = (1, i, \cdots, i)\mathbb{R}^n.$$

It follows that $\lambda \pitchfork \Lambda^1_{L_0}$ and

$$\lambda^{-1}(\Lambda^1) = \{0\}.$$

It remains to compute the sign at $t = 0$. As Lagrangian complement of $\lambda(0) = \mathbb{R}^n$ we choose $i\mathbb{R}^n$. For $t \in (-\frac{1}{2}, \frac{1}{2})$ we can write the path $\lambda(t)$ as graph

$$\lambda(t) = \{x + iA(t)x : x \in \mathbb{R}^n\}$$

where

$$A(t) = \begin{pmatrix} \tan \pi t & 0 & & 0 \\ 0 & 0 & & \\ & & \ddots & \\ 0 & & & 0 \end{pmatrix}.$$

By (102) the quadratic form $Q^{\lambda'(0)} \in S^2(\mathbb{R}^n)$ is given by

$$Q^{\lambda'(0)}(x) = \langle x, A'(0)x \rangle, \quad x \in \mathbb{R}^n.$$

Now the derivative of the matrix A computes to be

$$A'(t) = \begin{pmatrix} \frac{\pi}{\cos^2 \pi t} & 0 & & 0 \\ 0 & 0 & & \\ & & \ddots & \\ 0 & & & 0 \end{pmatrix}$$

and therefore

$$A'(0) = \begin{pmatrix} \pi & 0 & & 0 \\ 0 & 0 & & \\ & & \ddots & \\ 0 & & & 0 \end{pmatrix}.$$

Moreover,

$$L_0 \cap \lambda(0) = \mathrm{span}\{(1, 0, \ldots, 0)\}$$

and therefore the map

$$Q^{\lambda'(0)}\big|_{L_0 \cap \lambda(0)} : x \mapsto \pi x^2$$

is positive. We conclude that

$$\nu(0) = 1$$

and therefore

$$\widetilde{\mu}(\lambda) = 1 = \mu(\lambda).$$

This finishes the proof of the theorem. □

One of the advantages of the intersection theoretic interpretation of the Maslov index is that it generalizes to paths of Lagrangians. This is necessary to define the Conley-Zehnder index. We next explain the definition of the Maslov index for paths due to Robbin and Salamon [**94**].

Suppose that $\lambda\colon [0,1] \to \Lambda$ is a smooth path of Lagrangian subspaces. Fix a basepoint $L \in \Lambda$. For $t \in [0,1]$ the *crossing form*

$$C(\lambda, L, t) := Q^{\lambda'(t)}\big|_{\lambda(t)\cap L} \tag{105}$$

is a quadratic form on the vector space $\lambda(t)\cap L$. An intersection point $t \in \lambda^{-1}(\overline{\Lambda^1_L}) = \lambda^{-1}(\bigcup_{k=1}^n \Lambda^k_L)$ is called a *regular crossing* if and only if the crossing form $C(\lambda, L, t)$ is nonsingular. After a perturbation with fixed endpoints we can assume that the path $\lambda\colon [0,1] \to \Lambda$ has only regular crossings. In fact, we can even assume that for all $t \in (0,1)$ it holds that $\lambda(t) \notin \Lambda^k_L$ for $k \geq 2$. However, $\lambda(0)$ or $\lambda(1)$ might lie in Λ^k_L for some $k \geq 2$.

Suppose now that $\lambda : [0,1] \to \Lambda$ is a path with only regular crossings. Then its Maslov index with respect to the chosen basepoint $L \in \Lambda$ is defined by Robbin and Salamon [**94**] as

$$\mu_L(\lambda) := \frac{1}{2}\mathrm{sign}C(\lambda, L, 0) + \sum_{0<t<1} \mathrm{sign}C(\lambda, L, t) + \frac{1}{2}\mathrm{sign}C(\lambda, L, 1) \in \frac{1}{2}\mathbb{Z}. \tag{106}$$

Here sign refers to the signature of a quadratic form. The Maslov index for paths has the following properties.

Invariance:: If $\lambda_0, \lambda_1\colon [0,1] \to \Lambda$ are homotopic to each other with fixed endpoints, then
$$\mu_L(\lambda_0) = \mu_L(\lambda_1).$$

Concatenation:: Suppose that $\lambda_0, \lambda_1\colon [0,1] \to \Lambda$ satisfy $\lambda_0(1) = \lambda_1(0)$, then
$$\mu_L(\lambda_0 \# \lambda_1) = \mu_L(\lambda_0) + \mu_L(\lambda_1)$$
where $\lambda_0 \# \lambda_1$ refers to the concatenation of the two paths.

Loop: : If $\lambda\colon [0,1] \to \Lambda$ is a loop, i.e., $\lambda(0) = \lambda(1)$, then
$$\mu_L(\lambda) = \mu(\lambda)$$
the Maslov index for loops defined before. In particular, for loops the Maslov index does not depend on the choice of the base point $L \in \Lambda$.

Remark 0.13. *In general the Maslov index for paths depends on the choice of the base point $L \in \Lambda$. However, if $L_0, L_1 \in \Lambda$ and $\lambda\colon [0,1] \to \Lambda$ the three properties of the Maslov index just described imply that the difference $\mu_{L_0}(\lambda) - \mu_{L_1}(\lambda)$ only depends on $L_0, L_1, \lambda(0), \lambda(1)$, i.e., only the endpoints of the Lagrangian path λ matter. Such indices which associate to a collection of four Lagrangian subspaces of a symplectic vector space a number are also known in the literature as Maslov indices and are for example studied in the work by Hörmander* [**50**] *or Kashiwara* [**78**].

Remark 0.14. *There are other ways how to associate to a path of Lagrangian subspaces a Maslov index. We mention here the work of Duistermaat* [**29**]. *The Maslov index of Duistermaat has the property that it is independent of the choice of the base point $L \in \Lambda$ at the expense of the concatenation property. The Maslov index of Duistermaat is related to the Maslov index of Robbin and Salamon by a correction term involving a Hörmander-Kashiwara Maslov index.*

We next explain how to define the *Conley-Zehnder index* as a Maslov index. If (V,ω) is a symplectic vector space the symplectic group $\mathrm{Sp}(V)$ consists of all linear maps $A\colon V \to V$ satisfying $A^*\omega = \omega$. Moreover, if (V,ω) is a symplectic vector space $(V \oplus V, -\omega \oplus \omega)$ is a symplectic vector space as well and has the property that for every $A \in \mathrm{Sp}(V)$ the graph of A

$$\Gamma_A = \{(x, Ax) : x \in V\} \subset V \oplus V$$

is a Lagrangian subspace of $(V \oplus V, -\omega \oplus \omega)$. Indeed, if $(x, Ax), (y, Ay) \in \Gamma_A$ we have

$$(-\omega \oplus \omega)((x, Ax), (y, Ay)) = -\omega(x, y) + \omega(Ax, Ay) = -\omega(x, y) + \omega(x, y) = 0$$

where in the second equality we have used that A is symplectic. In particular, the diagonal

$$\Delta = \Gamma_{\mathrm{id}} = \{(x, x) : x \in V\}$$

is a Lagrangian subspace of $(V \oplus V, -\omega \oplus \omega)$. Suppose now that we have given a smooth path $\Psi\colon [0,1] \to \mathrm{Sp}(V)$, i.e., a smooth path of linear symplectic maps. We associate to such a path a path of Lagrangian subspaces in $V \oplus V$ by

$$\Gamma_\Psi\colon [0,1] \to \Lambda(V \oplus V), \quad t \mapsto \Gamma_{\Psi(t)}.$$

We say that a smooth path of symplectic linear maps $\Psi\colon [0,1] \to \mathrm{Sp}(V)$ starting at the identity $\Psi(0) = \mathrm{id}$ is *non-degenerate* if

$$\det(\Psi(1) - \mathrm{id}) \neq 0,$$

i.e., 1 is not an eigenvector of $\Psi(1)$. This is equivalent to the requirement that

$$\Gamma_{\Psi(1)} \in \Lambda^0_\Delta$$

i.e., $\Gamma_{\Psi(1)}$ does not lie in the closure of the Maslov pseudo-cocycle. We are now in position to define the Conley-Zehnder index

Definition 0.15. *Assume that Ψ $[0,1] \to \mathrm{Sp}(V)$ is a non-degenerate path of symplectic linear maps starting at the identity. Then the* Conley-Zehnder index *of Ψ is defined as*

$$\mu_{CZ}(\Psi) := \mu_\Delta(\Gamma_\Psi) \in \mathbb{Z}.$$

Remark 0.16. *The Maslov index $\mu_\Delta(\Gamma_\Psi)$ is also defined in the case where Ψ is degenerate. However, in the degenerate case we do not define the Conley-Zehnder index via the Maslov index. Instead of that following* [**56**] *we use the spectral flow to associate a Conley-Zehnder index to a degenerate path in Definition 0.29. We point out that in the case of degenerate paths the Conley-Zehnder index might in fact be different from the Maslov index. Indeed, the Conley-Zehnder index as extended to degenerate paths via Definition 0.29 becomes lower semi-continuous where the Maslov index is neither lower nor upper semi-continuous.*

Since the Maslov index is in general only half integer valued it is a priori not clear that the Conley-Zehnder index takes values in the integers. However, this can be seen as follows. Since by assumption the path is non-degenerate it follows that $\mathrm{sign}C(\Gamma_\Psi, \Delta, 1) = 0$. Therefore we obtain the formula

$$\mu_{CZ}(\Psi) = \frac{1}{2}\mathrm{sign}C(\Gamma_\Psi, \Delta, 0) + \sum_{0<t<1} \mathrm{sign}C(\Gamma_\Psi, \Delta, t).$$

Since $\Psi(0) = \mathrm{id}$ we have $\Gamma_\Psi(0) = \Delta$ and therefore $C(\Gamma_\Psi, \Delta, 0)$ is a quadratic form on the vector space Δ. However, Δ is even dimensional and therefore $\mathrm{sign} C(\Gamma_\Psi, \Delta, 0) \in 2\mathbb{Z}$. This proves that the Conley-Zehnder index is an integer.

CHAPTER 9

Spectral flow

In the following we let ω be the standard symplectic form on $\mathbb{C}^n$. We abbreviate by $\mathrm{Sp}(n)$ the linear symplectic group consisting of all real linear transformations $A\colon \mathbb{C}^n \to \mathbb{C}^n$ satisfying $A^*\omega = \omega$.

Suppose that $\Psi\colon [0,1] \to \mathrm{Sp}(n)$ is a smooth path of symplectic matrices which starts at the identity, i.e., $\Psi(0) = \mathrm{id}$. It follows that $\Psi'(t)\Psi^{-1}(t) \in \mathrm{Lie}\,\mathrm{Sp}(n)$, the Lie algebra of the linear symplectic group. The Lie algebra of the linear symplectic group can be described as follows. For the splitting $\mathbb{C}^n = \mathbb{R}^n \times \mathbb{R}^n$ write

$$J = \begin{pmatrix} 0 & -\mathrm{id} \\ \mathrm{id} & 0 \end{pmatrix}. \tag{107}$$

Note that

$$J^2 = -\mathrm{id}, \quad J^T = -J.$$

Using J the condition that a matrix $A \in Sp(n)$ is equivalent to the assertion that A meets

$$A^T J A = J.$$

That means that $B \in \mathrm{Lie}\,\mathrm{Sp}(n)$ if and only if

$$B^T J + JB = 0.$$

Therefore

$$(JB)^T = B^T J^T = -B^T J = JB$$

implying that $JB \in \mathrm{Sym}(2n)$, the vector space of symmetric $2n \times 2n$-matrices. On the other hand one checks immediately that if $JB \in \mathrm{Sym}(2n)$, then $B \in \mathrm{Lie}\,Sp(n)$, which means that the map $B \mapsto JB$ is a vector space isomorphism between $\mathrm{Lie}\,\mathrm{Sp}(n)$ and $\mathrm{Sym}(2n)$.

Therefore to any smooth path $\Psi \in C^\infty([0,1], \mathrm{Sp}(n))$ satisfying $\Psi(0) = \mathrm{id}$ we associate a smooth path $S = S_\Psi \in C^\infty([0,1], \mathrm{Sym}(2n))$ by setting

$$S(t) := -J\Psi'(t)\Psi^{-1}(t).$$

We recover Ψ from S by solving the ODE

$$\Psi'(t) = JS(t)\Psi(t),\ t \in [0,1], \quad \Psi(0) = \mathrm{id}. \tag{108}$$

That means we have a one to one correspondence between path of linear symplectic matrices starting at the identity and paths of symmetric matrices. We associate to these a linear operator

$$A = A_\Psi = A_S\colon W^{1,2}(S^1, \mathbb{C}^n) \to L^2(S^1, \mathbb{C}^n), \quad v \mapsto -J\partial_t v - Sv. \tag{109}$$

Here L^2 refers to the Hilbert space of square integrable functions and $W^{1,2}$ refers to the Hilbert space of square integrable function which admit a weak derivative which is also square integrable. Our goal is to relate the spectral theory of the operators A_Ψ to the Conley-Zehnder index of Ψ. For that purpose we first examine the kernel of A.

Lemma 0.17. *The evaluation map*

$$E\colon \ker A \to \ker(\Psi(1) - \mathrm{id}), \quad v \mapsto v(0)$$

is a vector space isomorphism.

Proof: We prove the lemma in three steps.

Step 1 The evaluation map is well defined.

Pick $v \in \ker A$. We have to show that $v(0) \in \ker(\Psi(1) - \mathrm{id})$. The condition that $Av = 0$ means that v is a solution of the ODE

$$J\partial_t v = -Sv \tag{110}$$

or equivalently

$$\partial_t v = (\partial_t \Psi)\Psi^{-1} v = -\Psi\partial_t(\Psi^{-1})v.$$

This can be rephrased by saying that $\partial_t(\Psi^{-1}v) = 0$ or equivalently that

$$v(t) = \Psi(t)v_0, \quad v_0 \in \mathbb{C}^n.$$

Since v is a loop we obtain

$$v(0) = v(1) = \Psi(1)v(0).$$

This proves that the evaluation map is well defined.

Step 2: The evaluation map is injective.

Suppose that $v \in \ker E$. That means that $v(0) = 0$. However, v is a solution of the ODE (110). This implies that $v(t) = 0$ for every $t \in S^1$.

Step 3: The evaluation map is surjective.

Suppose that $v_0 \in \ker(\Psi(1) - \mathrm{id})$, i.e.,

$$\Psi(1)v_0 = v_0.$$

Define

$$v(t) = \Psi(t)v_0.$$

It follows that

$$v(1) = \Psi(1)v_0 = v_0 = v(0).$$

Hence $v \in W^{1,2}(S^1, \mathbb{C}^n)$ and we have seen in the proof of Step 1 that $v \in \ker A$. This finishes the proof of Step 3 and hence of the Lemma. □

It is worth pointing out that when $\Psi(1)$ has an eigenvector to the eigenvalue 1 this precisely means that its graph $\Gamma_{\Psi(1)}$ lies in the closure of the Maslov pseudocycle Λ^1_Δ. Hence by the Lemma crossing the Maslov pseudocycle is equivalent that an eigenvalue of the operator crosses zero.

The L^2-inner product on $L^2(S^1, \mathbb{C})$ is given for two vectors $v_1, v_2 \in L^2(S^1, \mathbb{C}^n)$ by

$$\langle v_1, v_2 \rangle = \int_0^1 \langle v_1(t), v_2(t) \rangle dt = \int_0^1 \omega(v_1(t), Jv_2(t)\rangle dt.$$

Lemma 0.18. *The operator A is symmetric with respect to the L^2-inner product.*

Proof: Suppose that $v_1, v_2 \in W^{1,2}(S^1, \mathbb{C}^n)$. We compute using integration by parts and the fact that S is symmetric

$$\begin{aligned}\langle Av_1, v_2 \rangle &= -\int_0^1 \langle J\partial_t v_1 + Sv_1, v_2 \rangle dt \\ &= -\int_0^1 \omega(J\partial_t v_1, J\partial_t v_2)dt - \int_0^1 \langle Sv_1, v_2 \rangle dt \\ &= -\int_0^1 \omega(\partial_t v_1, v_2)dt - \int_0^1 \langle v_1, Sv_2 \rangle dt \\ &= -\int_0^1 \omega(v_1, \partial_t v_2)dt - \int_0^1 \langle v_1, Sv_2 \rangle dt \\ &= \langle v_1, Av_2 \rangle.\end{aligned}$$

This finishes the proof of the Lemma. □

A crucial property of the operator $A\colon W^{1,2}(S^1, \mathbb{C}^n) \to L^2(S^1, \mathbb{C}^n)$ is that it is a Fredholm operator of index 0. Before stating this theorem we recall some basic facts about Fredholm operators without proofs. Proofs can be found for example in [**81**, Appendix A.1]. If H_1 and H_2 are Hilbert spaces, then a bounded linear operator $D\colon H_1 \to H_2$ is called *Fredholm operator* if it has a finite dimensional kernel, a closed image and a finite dimensional cokernel. Then the number

$$\mathrm{ind}D := \dim \ker D - \dim \mathrm{coker}D \in \mathbb{Z}$$

is referred to as the *index* of the Fredholm operator D. For example if H_1 and H_2 are finite dimensional, then every linear operator $D\colon H_1 \to H_2$ is Fredholm and its index is given by

$$\mathrm{ind}D = \dim H_1 - \dim H_2.$$

Interestingly in this example the Fredholm index does not depend on D at all, although $\dim \ker D$ and $\dim \mathrm{coker}D$ do. The usefulness of Fredholm operators lies in the fact that similar phenomena happen in infinite dimensions and the Fredholm index is rather stable under perturbations.

Recall that a *compact operator* $K\colon H_1 \to H_2$ is a bounded linear operator with the property that $\overline{K(B_{H_1})} \subset H_2$ is compact, where $B_{H_1} = \{v \in H_1 : ||v|| \leq 1\}$ is the unit ball in H_1 and the closure refers to the topology in H_2. The first useful fact about Fredholm operators is that they are stable under compact perturbations.

Theorem 0.19. *Assume that $D\colon H_1 \to H_2$ is a Fredholm operator and $K\colon H_1 \to H_2$ is a compact operator. Then $D + K\colon H_1 \to H_2$ is a Fredholm operator as well and*

$$\mathrm{ind}(D + K) = \mathrm{ind}D.$$

The second useful fact about Fredholm operators is that they are stable under small perturbations in the operator topology.

Theorem 0.20. *Assume that $D\colon H_1 \to H_2$ is a Fredholm operator. Then there exists $\epsilon > 0$ such that for every $E\colon H_1 \to H_2$ satisfying $||D - E|| < \epsilon$, where the norm refers to the operator norm, it holds that E is still Fredholm and* $\mathrm{ind}E = \mathrm{ind}D$.

In order to prove that an operator is Fredholm the following Lemma is very useful.

Lemma 0.21. *Assume that H_1, H_2, H_3 are Hilbert spaces, $D\colon H_1 \to H_2$ is a bounded linear operator, $K\colon H_1 \to H_3$ is a compact operator, and there exists a constant $c > 0$ such that the following estimate holds for every $x \in H_1$*

$$||x||_{H_1} \leq c\big(||Dx||_{H_2} + ||Kx||_{H_3}\big).$$

Then D has a closed image and a finite dimensional kernel.

An bounded linear operator with a closed image and a finite dimensional kernel is referred to as a *semi Fredholm operator*. Therefore under the conditions of the Lemma D is a semi Fredholm operator. In practice one can often apply the Lemma again to the adjoint of D. Since the kernel of the adjoint of D coincides with the cokernel of D this enables one to deduce the Fredholm property of D. The proofs of the two previous Theorems as well as of the Lemma can be found in [**81**, Appendix A.1.]. We can use these results to prove the following theorem.

Theorem 0.22. *The operator $A\colon W^{1,2}(S^1, \mathbb{C}^n) \to L^2(S^1, \mathbb{C}^n)$ is a Fredholm operator of index 0.*

We present two proofs of this theorem.

Proof 1 of Theorem 0.22: Since the inclusion of $W^{1,2}(S^1, \mathbb{C}^n) \hookrightarrow L^2(S^1, \mathbb{C}^n)$ is compact it follows that the operator $A_S = -J\partial_t + S$ is a compact perturbation of the operator $A_0 = -J\partial_t$. To check that the operator A_0 is Fredholm of index 0 is a straightforward exercise in the fundamental theorem of calculus and left to the reader. Now it follows from Theorem 0.19 that A_S is Fredholm of index 0 as well. $\square$

The second proof does not invoke Theorem 0.19 but uses instead Lemma 0.21. The reason why we present it is that according to a similar scheme many elliptic operators can be proven to be Fredholm operators. These proofs usually contain two ingredients. First one needs to produce an estimate in order to be able to apply Lemma 0.21. Then one has to apply elliptic regularity in order to identify the cokernel of the operator with the kernel of the adjoint. Since the operator A is an operator in just one variable the elliptic regularity part is immediate. The estimate is the content of the following lemma.

Lemma 0.23. *There exists $c > 0$ such that for every $v \in W^{1,2}(S^1, \mathbb{C}^n)$ the following estimate holds*

$$||v||_{W^{1,2}} \leq c\big(||v||_{L^2} + ||Av||_{L^2}\big).$$

Proof: Since $J\partial_t v = -Av - Sv$ we obtain

$$\partial_t v = JAv + JSv.$$

Therefore we estimate

$$\begin{aligned}
||v||^2_{W^{1,2}} &= ||v||^2_{L^2} + ||\partial_t v||^2_{L^2} \\
&= ||v||^2_{L^2} + ||J(A+S)v||^2_{L^2} \\
&= ||v||^2_{L^2} + ||(A+S)v||^2_{L^2} \\
&= ||v||^2_{L^2} + ||Av||^2_{L^2} + 2\langle Av, Sv\rangle + ||Sv||^2_{L^2} \\
&\leq ||v||^2_{L^2} + 2||Av||^2_{L^2} + 2||Sv||^2_{L^2} \\
&\leq c\big(||v||^2_{L^2} + ||Av||^2_{L^2}\big).
\end{aligned}$$

This finishes the proof of the Lemma. □

Proof 2 of Theorem 0.22: Because the inclusion $W^{1,2}(S^1,\mathbb{C}^n) \hookrightarrow L^2(S^1,\mathbb{C}^n)$ is compact we deduce from Lemma 0.21 and Lemma 0.23 that A is a semi Fredholm operator, i.e., $\ker A$ is finite dimensional and $\operatorname{im} A$ is closed. To determine its cokernel choose $w \in \operatorname{im} A^\perp$, the orthogonal complement of the image of A. That means that $w \in L^2(S^1,\mathbb{C}^n)$ satisfies

$$\langle w, Av\rangle = 0, \quad \forall\ v \in W^{1,2}(S^1,\mathbb{C}^n).$$

Hence if $v \in W^{1,2}(S^1,\mathbb{C}^n)$ we have

$$0 = \langle w, Sv\rangle + \langle w, J\partial_t v\rangle = \langle Sw, v\rangle - \langle Jw, \partial_t v\rangle$$

and therefore

$$\langle Jw, \partial_t v\rangle = \langle Sw, v\rangle, \quad \forall\ v \in W^{1,2}(S^1,\mathbb{C}^n).$$

This implies that w which a priori was just an element in $L^2(S^1,\mathbb{C}^n)$ actually lies in $W^{1,2}(S^1,\mathbb{C}^n)$ and satisfies the equation

$$J\partial_t w = -Sw.$$

In particular,

$$Aw = 0.$$

Since by Lemma 0.18 the operator A is symmetric, it holds that every element in the kernel of A is orthogonal to the image of A. We obtain

$$\operatorname{im} A^\perp = \ker A$$

and therefore

$$\dim \operatorname{coker} A = \dim\ker A.$$

In particular, A is Fredholm and its index satisfies

$$\operatorname{ind} A = \dim\ker A - \dim\operatorname{coker} A = 0.$$

This proves Theorem 0.22 again. □

The fact that $A\colon W^{1,2}(S^1,\mathbb{C}^n) \to L^2(S^1,\mathbb{C}^n)$ is a symmetric Fredholm operator of index 0 has important consequences for its spectrum. A reader familiar with the theory of unbounded operators might recognize that the fact that A is Fredholm of index 0 implies that A interpreted as an unbounded operator $A\colon L^2(S^1,\mathbb{C}^n) \to L^2(S^1,\mathbb{C}^n)$ is *self-adjoint* with dense domain $W^{1,2}(S^1,\mathbb{C}^n) \subset L^2(S^1,\mathbb{C}^n)$. Here we do not invoke the theory of self-adjoint unbounded operators but argue directly via the standard properties of Fredholm operators we already recalled. We abbreviate by

$$I\colon W^{1,2}(S^1,\mathbb{C}^n) \to L^2(S^1,\mathbb{C}^n)$$

the inclusion which is a compact operator and define the *spectrum* of A as

$$\mathfrak{S}(A) := \big\{\eta \in \mathbb{C} : A - \eta I \text{ not invertible}\big\}.$$

Its complement is the *resolvent set*

$$\mathfrak{R}(A) := \mathbb{C} \setminus \mathfrak{S}(A) = \big\{\eta \in \mathbb{C} : A - \eta I \text{ invertible}\big\}.$$

An element $\eta \in \mathbb{C}$ is called an *eigenvalue* of A if $\ker(A - \eta I) \neq \{0\}$.

Lemma 0.24. *The spectrum of A consists precisely of the eigenvalues of A, i.e.,*

$$\eta \in \mathfrak{S}(A) \iff \eta \text{ eigenvalue of } A.$$

Proof: The implication "$\Longleftarrow$" is obvious. To prove the implication "$\Longrightarrow$" we observe that in view of the stability of the Fredholm index under compact perturbations as stated in Theorem 0.19 it follows that $A - \eta I$ is still a Fredholm operator of index 0. Now assume that $\eta \in \mathbb{C}$ is not an eigenvalue of A, i.e., $\ker(A - \eta I) = \{0\}$. Since the index of $A - \eta I$ is 0 we conclude that $\operatorname{coker}(A - \eta I) = \{0\}$. This means that $A - \eta I$ is bijective and hence by the open mapping theorem invertible. This finishes the proof of the Lemma. $\square$

Lemma 0.25. $\mathfrak{S}(A) \subset \mathbb{R}$.

Proof: Assume that $\eta \in \mathfrak{S}(A)$. By Lemma 0.24 we know that η is an eigenvalue of A. Now the argument is standard. Indeed, if $v \neq 0$ such that $Av = \eta I v$ we have in view of the symmetry of A establishes in lemma 0.18

$$\eta||v||^2 = \langle Av, v\rangle = \langle v, Av\rangle = \bar{\eta}||v||^2$$

and therefore since $||v||^2 \neq 0$ we conclude $\eta = \bar{\eta}$ or in other words $\eta \in \mathbb{R}$. $\square$

Lemma 0.26. *The spectrum $\mathfrak{S}(A)$ is countable.*

Proof: If $\eta \in \mathfrak{S}(A)$ it follows from Lemma 0.24 that there exists $e_\eta \neq 0$ such that $Ae_\eta = \eta I e_\eta$. Since A is symmetric we have for $\eta \neq \eta' \in \mathfrak{S}(A)$ that $\langle e_\eta, e_{\eta'}\rangle = 0$, i.e., the two eigenvectors are orthogonal to each other. Since the Hilbert space $L^2(S^1, \mathbb{C}^n)$ is separable we conclude that $\mathfrak{S}(A)$ has to be countable. $\square$

As a consequence of Lemma 0.26 we conclude that there exists $\zeta_0 \in \mathbb{R}$ such that $\zeta_0 \notin \mathfrak{S}(A)$. In particular, $A - \zeta_0 I$ is invertible. Its inverse is a bounded linear map

$$(A - \zeta_0 I)^{-1} \colon L^2(S^1, \mathbb{C}^n) \to W^{1,2}(S^1, \mathbb{C}^n).$$

We define the resolvent operator

$$R(\zeta_0) := I \circ (A - \zeta_0 I)^{-1} \colon L^2(S^1, \mathbb{C}^n) \to L^2(S^1, \mathbb{C}^n).$$

Since the inclusion operator $I \colon W^{1,2}(S^1, \mathbb{C}^n) \to L^2(S^1, \mathbb{C}^n)$ is compact we conclude that the resolvent operator $R(\zeta_0)$ is compact. Moreover, since A is symmetric, the resolvent operator $R(\zeta_0)$ is symmetric as well. Furthermore, 0 is not an eigenvalue of $R(\zeta_0)$, because if $R(\zeta_0)v = 0$ we get $v = (A - \zeta_0 I)R(\zeta_0)v = 0$.

Recall that if H is a Hilbert space and $R \colon H \to H$ is a compact symmetric operator then the *spectral theorem for compact symmetric operators* tells us that the spectrum $\mathfrak{S}(R)$ is real, bounded and zero is the only possible accumulation point of $\mathfrak{S}(R)$. Moreover, for all $\eta \in \mathfrak{S}(R)$ there exist pairwise commuting orthogonal projections

$$\Pi_\eta \colon H \to H, \quad \Pi_\eta^2 = \Pi_\eta = \Pi_\eta^*, \ \Pi_\eta \Pi_{\eta'} = \Pi_{\eta'} \Pi_\eta$$

satisfying

$$\sum_{\eta \in \mathfrak{S}(R)} \Pi_\eta = \mathrm{id}$$

such that

$$R = \sum_{\eta \in \mathfrak{S}(R)} \eta \Pi_\eta.$$

Moreover, if $\eta \neq 0 \in \mathfrak{S}(R)$ then $\dim \mathrm{im} \Pi_\eta < \infty$, meaning that the eigenvalue η has finite multiplicity.

From the spectral decomposition of the resolvent $R(\zeta_0)$ we obtain the spectral decomposition of the operator A, namely

$$A = \sum_{\eta \in \mathfrak{S}(R(\zeta_0))} \left(\frac{1}{\eta} + \zeta_0 \right) \Pi_\eta.$$

In particular, we can improve Lemma 0.26 to the following stronger statement.

Proposition 0.27. *The spectral* $\mathfrak{S}(A) \subset \mathbb{R}$ *is discrete.*

If we write for the spectral decomposition of A

$$A = \sum_{\eta \in \mathfrak{S}(A)} \eta \Pi_\eta$$

and $\zeta \in \mathfrak{R}(A)$ is in the resolvent set of A we obtain the spectral decomposition of the resolvent operator $R(\zeta)$ as

$$R(\zeta) = \sum_{\eta \in \mathfrak{S}(A)} \frac{1}{\eta - \zeta} \Pi_\eta.$$

Now choose a smooth loop $\Gamma \colon S^1 \to \mathfrak{R}(A) \subset \mathbb{C}$ such that the only eigenvalue of A encircled by Γ is η. Now we can recover the projection Π_η by the residual theorem as follows

$$\Pi_\eta = \frac{i}{2\pi} \int_\Gamma R(\zeta) d\zeta.$$

This formula plays a central role in Kato's fundamental book [**68**]. According to [**68**] this formula was first used in perturbation theory by Szökevalvi-Nagy [**104**] and Kato [**66, 67**]. More generally if $\Gamma \colon S^1 \to \mathfrak{R}(A)$ is a smooth loop and $\mathfrak{S}_\Gamma(A) \subset \mathfrak{S}(A)$ denotes the set of all eigenvalues of A encircled by Γ we obtain by the residual theorem the following formula

$$\frac{i}{2\pi} \int_\Gamma R(\zeta) d\zeta = \sum_{\eta \in \mathfrak{S}_\Gamma(A)} \Pi_\eta.$$

The reason why this formula is so useful is that the map

$$A \mapsto \frac{i}{2\pi} \int_\Gamma R_A(\zeta) d\zeta$$

is continuous in the operator A with respect the operator topology. In particular, the eigenvalues of A vary continuously under perturbations of A.

Recall that the operator $A = A_S \colon W^{1,2}(S^1, \mathbb{C}^n) \to L^2(S^1, \mathbb{C}^n)$ is given as $v \mapsto -J\partial_t v - S(t)v$. In particular, for $S = 0$ we have the map $A_0 = -J\partial_t$. Its eigenvalues are $2\pi\ell$ for $\ell \in \mathbb{Z}$ and the corresponding eigenvectors are given by $v_0 e^{2\pi i \ell t}$ for $v_0 \in \mathbb{C}^n$. In particular, each eigenvalue has multiplicity $2n$. Abbreviate by

$$\mathcal{P} := C^\infty([0,1], \mathrm{Sym}(2n))$$

the space of paths of symmetric $2n \times 2n$-matrices. We endow the space $\mathcal{P}$ with the metric

$$d(S, S') = \int_0^1 ||S(t) - S'(t)|| dt.$$

The *Spectrum bundle* is defined as

$$\mathfrak{S} := \{S \in \mathcal{P} : \mathfrak{S}(A_S)\} \subset \mathcal{P} \times \mathbb{R}.$$

It comes with a canonical projection $\mathfrak{S} \to \mathcal{P}$. In view of the continuity of eigenvalues of the operators A_S under perturbation and the description of the spectrum of A_0 discussed above there exist continuous sections

$$\eta_k \colon \mathcal{P} \to \mathfrak{S}, \quad k \in \mathbb{Z}$$

which are uniquely determined by the following requirements

- **(i):** The map $S \mapsto \eta_k(S)$ is continuous for every $k \in \mathbb{Z}$.
- **(ii):** $\eta_k(S) \leq \eta_{k+1}(S)$ for every $k \in \mathbb{Z}$, $S \in \mathcal{P}$.
- **(iii):** $\mathfrak{S}(A_S) = \{\eta_k(S) : k \in \mathbb{Z}\}$.
- **(iv):** If $\eta \in \mathfrak{S}(A_S)$, then the number $\#\{k \in \mathbb{Z} : \eta_k(S) = \eta\}$ equals the multiplicity of the eigenvalue η.
- **(v):** The sections η_k are normalized such that $\eta_j(0) = 0$ for $j \in \{1, \ldots, 2n\}$.

Recall that if $S \in \mathcal{P}$ we can associate to S a path of symplectic matrices Ψ_S starting at the identity by formula (108). The following theorem relates the Conley Zehnder index of Ψ_S as defined in Definition 0.15 to the spectrum of A_S, see also [**56, 95**].

Theorem 0.28 (Spectral flow). *Assume that Ψ_S is non-degenerate, meaning that* $\det(\Psi_S(1) - \mathrm{id}) \neq 0$. *Then*

$$\mu_{CZ}(\Psi_S) = \max\{k : \eta_k(S) < 0\} - n. \tag{111}$$

Recall that the Conley-Zehnder index as explained in Definition 0.15 is only associated to non-degenerate paths of symplectic matrices. Inspired by Theorem 0.28 and following [**56**] we extend the Definition of the Conley-Zehnder index to degenerate paths of symplectic matrices, i.e. paths Ψ_S satisfying $\det(\Psi_S(1) - \mathrm{id}) = 0$, in the following way.

Definition 0.29. *Assume that Ψ_S $[0,1] \to \mathrm{Sp}(n)$ is a degenerate path of symplectic linear maps starting at the identity. Then the* Conley-Zehnder index *of Ψ_S is defined as*

$$\mu_{CZ}(\Psi_S) := \max\{k : \eta_k(S) < 0\} - n.$$

In view of Theorem 0.28 formula (111) is now valid for arbitrary paths of symplectic matrices starting at the identity, regardless if they are degenerate or not. The reason why we extend the Conley-Zehnder index to degenerate paths via the spectral flow formula and not the Maslov index is that via the spectral flow formula the Conley-Zehnder index becomes *lower semi-continuous* where the Maslov index is neither lower nor upper semi-continuous.

To prove Theorem 0.28 we first show a Lagrangian version of the spectral flow theorem and then use the Lagrangian version to deduce the periodic version, namely

Theorem 0.28, by looking at its graph. To formulate the Lagrangian version of the spectral flow theorem we fix a smooth path $S\colon [0,1] \to \mathrm{Sym}(2n)$ of symmetric matrices as before. The Hilbert space we consider consists however not of loops anymore by of paths satisfying a Lagrangian boundary condition, namely

$$W^{1,2}_{\mathbb{R}^n}([0,1],\mathbb{C}^n) = \{v \in W^{1,2}([0,1],\mathbb{C}^n) : v(0), v(1) \in \mathbb{R}^n\}.$$

We consider the bounded linear operator

$$L_S\colon W^{1,2}_{\mathbb{R}^n}([0,1],\mathbb{C}^n) \to L^2([0,1],\mathbb{C}^n), \quad v \mapsto -J\partial_t v - Sv.$$

This is the same formula as for the operator A_S but note that the domain of the operator now changed. However, thanks to the Lagrangian boundary condition one easily checks that the operator L_S has the same properties as the operator A_S, namely it is a symmetric Fredholm operator of index 0, or considered as an unbounded operator $L_S\colon L^2([0,1],\mathbb{C}^n) \to L^2([0,1],\mathbb{C}^n)$ a self-adjoint operator with dense domain $W^{1,2}_{\mathbb{R}^n}([0,1],\mathbb{C}^n)$. In particular, L_S has the same spectral properties as the operator A_S. The eigenvalues of the operator L_0 are $\pi\ell$ for every $\ell \in \mathbb{Z}$ with corresponding eigenvectors $v(t) = v_0 e^{\pi i \ell t}$ where $v_0 \in \mathbb{R}^n$. In particular, the multiplicity of each eigenvalue is n. Just as in the periodic case we define the section η_k for $k \in \mathbb{Z}$ to the spectral bundle. Just the normalization condition has to be replaced by

(v'): The section η_k are normalized such that $\eta_j(0) = 0$ for $j \in \{1,\dots,n\}$.

We associate to S the path of Lagrangians

$$\lambda_S\colon [0,1] \to \Lambda(n), \quad \lambda_S(t) = \Psi_S(t)\mathbb{R}^n.$$

We are now in position to state the Lagrangian version of the spectral flow theorem.

Theorem 0.30. *Assume that $\Psi_S(1)\mathbb{R}^n \cap \mathbb{R}^n = \{0\}$. Then*

$$\mu_{\mathbb{R}^n}(\lambda_S) = \max\{k : \eta_k(S) < 0\} - \frac{n}{2}.$$

Proof: Recall from (95) that the Lagrangian Grassmannian is stratified as

$$\Lambda = \bigcup_{k=0}^{n} \Lambda^k_{\mathbb{R}^n}$$

where $\Lambda^1_{\mathbb{R}^n}$ is the Maslov pseudo-cocycle whose closure is given by

$$\overline{\Lambda^1_{\mathbb{R}^n}} = \bigcup_{k=1}^{n} \Lambda^k_{\mathbb{R}^n}.$$

In particular,

$$\Lambda^0_{\mathbb{R}^n} = \big(\overline{\Lambda^1_{\mathbb{R}^n}}\big)^c$$

is the complement of the closure of the Maslov pseudo-cocycle. Recall further that $\Lambda^0_{\mathbb{R}^n}$ is connected. Indeed, it was shown in (99) that

$$\Lambda^0_{\mathbb{R}^n} = S^2(i\mathbb{R}^n)$$

the vector space of quadratic forms on $i\mathbb{R}^n$ such that $\Lambda^0_{\mathbb{R}^n}$ is actually contractible. Note the following equivalences

(112)
$$0 \notin \mathfrak{S}(L_S) \Longleftrightarrow \ker L_S = \{0\} \Longleftrightarrow \Psi_S(1)\mathbb{R}^n \cap \mathbb{R}^n = \{0\} \Longleftrightarrow \Psi_S(1)\mathbb{R}^n \in \Lambda^0_{\mathbb{R}^n}.$$

For a path $S \in C^\infty([0,1], \mathrm{Sym}(2n))$ define

$$\widetilde{\mu}(S) := \max\{k : \eta_k(S) < 0\}.$$

Consider a homotopy in $\mathcal{P}$ namely let $S \in C^\infty([0,1]\times[0,1], \mathrm{Sym}(2n))$ and abbreviate $S_r = S(\cdot, r)$ for the homotopy parameter $r \in [0,1]$. Assume that during the homotopy we never cross the closure of the Maslov pseudo-cocycle, namely

$$\Psi_{S_r}(1)\mathbb{R}^n \in \Lambda^0_{\mathbb{R}^n}, \quad r \in [0,1].$$

We conclude from (112) and the continuity of eigenvalues under perturbation that

$$\widetilde{\mu}(S_0) = \widetilde{\mu}(S_1).$$

By homotopy invariance of the intersection number we have as well

$$\mu_{\mathbb{R}^n}(\lambda_{S_0}) = \mu_{\mathbb{R}^n}(\lambda_{S_1}).$$

It was shown in Theorem 0.6 that the fundamental group of the Lagrangian Grassmannian satisfies $\pi_1(\Lambda) = \mathbb{Z}$ with generator

$$t \mapsto \begin{pmatrix} e^{i\pi t} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \mathbb{R}^n.$$

Because $\Lambda^0_{\mathbb{R}^n}$ is connected it follows that each non-degenerate path $\lambda_S = \Psi_S \mathbb{R}^n$ is homotopic through a path with endpoints in $\Lambda^0_{\mathbb{R}^n}$ to a path of the form

$$\lambda_k \colon [0,1] \to \mathrm{Sp}(n), \quad t \mapsto \begin{pmatrix} e^{i\pi(\frac{1}{2}+k)t} & & & \\ & e^{i\frac{\pi}{2}t} & & \\ & & \ddots & \\ & & & e^{i\frac{\pi}{2}t} \end{pmatrix} \mathbb{R}^n$$

for some $k \in \mathbb{Z}$. If

$$S_k = \frac{\pi}{2} \begin{pmatrix} 1+2k & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

then we obtain

$$\lambda_k = \lambda_{S_k}.$$

In view of the homotopy invariance of $\widetilde{\mu}$ and $\mu_{\mathbb{R}^n}$ it suffices to show that

$$\mu_{\mathbb{R}^n}(\lambda_k) = \widetilde{\mu}(S_k) - \frac{n}{2}.$$

But both sides are equal to $\frac{n}{2} + k$ and hence the theorem is proved. $\square$

We now use Theorem 0.30 to prove Theorem 0.28.

Proof of Theorem 0.28: Recall that if (V, ω) is a symplectic vector space, then the diagonal $\Delta = \{(v,v) : v \in V\} \subset V \oplus V$ is a Lagrangian subspace in the symplectic vector space $(V \oplus V, -\omega \oplus \omega)$. If $\Psi \colon [0,1] \to \mathrm{Sp}(V)$ is a smooth path of linear symplectic transformations starting at the identity, i.e., $\Psi(0) = \mathrm{id}$, and Γ_Ψ is the graph of Ψ, then this graph $\Gamma_\Psi \colon [0,1] \to \Lambda(V \oplus V)$ is a path in the Lagrangian

Grassmannian with the property that $\Gamma_{\Psi(0)} \in \Delta$. The Conley-Zehnder index of Ψ is by definition

$$\mu_{CZ}(\Psi) = \mu_\Delta(\Gamma_\Psi).$$

Choose further a complex structure $J\colon V \to V$, i.e., a linear map satisfying $J^2 = -\mathrm{id}$ which is ω-compatible in the sense that $\omega(\cdot, J\cdot)$ is a scalar product on V. Then $-J \oplus J$ is a complex structure on $V \otimes V$ and

$$-\omega \oplus \omega(\cdot, -J \oplus J\cdot) = \omega(\cdot, J\cdot) \oplus \omega(\cdot, J\cdot)$$

is a scalar product on $V \oplus V$. Define a Hilbert space isomorphism

$$\Gamma\colon W^{1,2}(S^1, V) \to W^{1,2}_\Delta([0,1], V \oplus V)$$

which associates to $v \in W^{1,2}(S^1, V)$ the map

$$\Gamma(v)(t) = \Big(v\big(1 - \tfrac{t}{2}\big), v\big(\tfrac{t}{2}\big)\Big), \quad t \in [0,1].$$

Note that because v was periodic, i.e., $v(0) = v(1)$, it holds that

$$\Gamma(v)(0) = \Big(v(0), v(1)\Big) = \Big(v(0), v(0)\Big) \in \Delta, \quad \Gamma(v)(1) = \Big(v\big(\tfrac{1}{2}\big), v\big(\tfrac{1}{2}\big)\Big) \in \Delta$$

such that $\Gamma(v)$ actually lies in the space $W^{1,2}_\Delta([0,1], V \oplus V)$. Moreover, Γ extends to a map

$$\Gamma\colon L^2(S^1, V) \to L^2([0,1], V \oplus V)$$

by the same formula. With respect to the L^2-inner product Γ is an isometry up to a factor $\sqrt{2}$, indeed

$$\begin{aligned}
||\Gamma(v)||_{L^2} &= \bigg(\int_0^1 ||\Gamma(v)(t)||^2 dt\bigg)^{1/2} \\
&= \bigg(\int_0^1 \Big(\big|\big|v\big(1 - \tfrac{t}{2}\big)\big|\big|^2 + \big|\big|v\big(\tfrac{t}{2}\big)\big|\big|^2\Big) dt\bigg)^{1/2} \\
&= \bigg(\int_0^{1/2} \Big(\big|\big|v\big(\tfrac{1}{2} + t\big)\big|\big|^2 + \big|\big|v\big(t\big)\big|\big|^2\Big) 2dt\bigg)^{1/2} \\
&= \sqrt{2}\bigg(\int_0^1 ||v(t)||^2 dt\bigg)^{1/2} \\
&= \sqrt{2}||v||_{L^2}.
\end{aligned}$$

Abbreviate

$$\mathcal{P}(V) = C^\infty([0,1], \mathrm{Sym}(V))$$

where $\mathrm{Sym}(V)$ denotes the vector space of self-adjoint linear maps with respect to the inner product $\omega(\cdot, J\cdot)$. Define a map

$$\Gamma\colon \mathcal{P}(V) \to \mathcal{P}(V \oplus V)$$

which associates to $S \in \mathcal{P}(V)$ the path of self-adjoint linear maps in $V \oplus V$

$$\Gamma(S)(t) = \frac{1}{2}\Big(S\big(1 - \tfrac{t}{2}\big), S\big(\tfrac{t}{2}\big)\Big).$$

The operators

$$A_S\colon W^{1,2}(S^1, V) \to L^2(S^1, V), \quad L_{\Gamma(S)}\colon W^{1,2}_\Delta([0,1], V \oplus V) \to L^2([0,1], V \oplus V)$$

are related by

$$\Gamma(A_S v) = L_{\Gamma(S)} \Gamma v, \quad v \in W^{1,2}(S^1, V).$$

Moreover, v is an eigenvector of A_S to the eigenvalue η if and only if $\Gamma(v)$ is an eigenvector of $L_{\Gamma(S)}$ to the eigenvalue $\frac{\eta}{2}$, in consistence with the fact that Γ is an isometry up to a factor $\sqrt{2}$ checked above. In particular, we have

$$\mathfrak{S}(L_{\Gamma(S)}) = \frac{1}{2}\mathfrak{S}(A_S).$$

We conclude that

$$\max\{k : \eta_k(S) < 0\} = \max\{k : \eta_k(\Gamma(S)) < 0\}$$

so that we obtain from Theorem 0.30

$$\begin{aligned} \mu_\Delta(\Psi_{\Gamma(S)}\Delta) &= \max\{k : \eta_k(S) < 0\} - \frac{\dim(V \oplus V)}{2} \\ &= \max\{k : \eta_k(S) < 0\} - \dim V. \end{aligned} \tag{113}$$

Note that the path of symplectomorphisms of $V \oplus V$ generated from $\Gamma(S)$ satisfies

$$\Psi_{\Gamma(S)}(t) = \big(\Psi_S\big(1 - \tfrac{t}{2}\big)\Psi_S(1)^{-1}, \Psi\big(\tfrac{t}{2}\big)\big), \quad t \in [0,1].$$

Therefore, if we apply this formula to the diagonal, we obtain

$$\Psi_{\Gamma(S)}(t)\Delta = \Gamma_{\Psi_S(1)\Psi_S(1-\frac{t}{2})^{-1}\Psi_S(\frac{t}{2})}.$$

Note that the path of symplectic matrices

$$t \mapsto \Psi_S(1)\Psi_S\big(1 - \tfrac{t}{2}\big)^{-1}\Psi_S\big(\tfrac{t}{2}\big), \quad t \in [0,1]$$

is homotopic with fixed endpoints to the path

$$t \mapsto \Psi_S(t), \quad t \in [0,1]$$

via the homotopy

$$(t,r) \mapsto \Psi_S(1)\Psi_S\big(1 - \tfrac{t(1-r)}{2}\big)^{-1}\Psi_S\big(\tfrac{t(1+r)}{2}\big), \quad (t,r) \in [0,1] \times [0,1].$$

Consequently, by homotopy invariance of the Maslov index it holds that

$$\mu_\Delta(\Gamma_{\Psi_S}) = \mu_\Delta(\Psi_{\Gamma(S)}\Delta). \tag{114}$$

Combining (113) and (114) we obtain

$$\mu_{CZ}(\Psi_S) = \mu_\Delta(\Gamma_{\Psi_S}) = \max\{k : \eta_k(S) < 0\} - \dim V.$$

This finishes the proof of the Theorem. □

The spectral flow theorem is difficult to apply directly since it requires that one knows the numbering of the eigenvalues. But to obtain this numbering one has to understand the bifurcation of the eigenvalues during a homotopy. A very fruitful idea of Hofer, Wysocki and Zehnder [**53**] is to use the *winding number* to keep track of the numbering of the eigenvalues. This only works if the dimension of the symplectic vector space is two since otherwise the winding number cannot be defined, however in dimension two this idea led to fantastic applications.

We now restrict our attention to the two dimensional symplectic vector space $(\mathbb{C}, \omega)$ and consider a smooth path of symmetric 2×2-matrices $S \in C^\infty([0,1], \mathrm{Sym}(2))$. Recall the operator

$$A_S \colon W^{1,2}(S^1, \mathbb{C}) \to L^2(S^1, \mathbb{C}), \quad v \mapsto -J\partial_t v - S(t)v.$$

Suppose that η is an eigenvalue of A_S and v is a eigenvector of A_S for the eigenvalue η, i.e.

$$A_S v = \eta v.$$

This means that v is a solution of the linear first order ODE
$$-J\partial_t v = (S+\eta)v.$$
Since v as an eigenvector cannot vanish identically it follows from the ODE above that
$$v(t) \neq 0 \in \mathbb{C}, \quad \forall\, t \in S^1.$$
Hence we get a map
$$\gamma_v \colon S^1 \to S^1, \quad t \mapsto \frac{v(t)}{||v(t)||}.$$
We define the *winding number* of the eigenvector v to be
$$w(v) := \deg(\gamma_v) \in \mathbb{Z}$$
where $\deg(\gamma_v)$ denotes the degree of the map γ_v. The following crucial Lemma of Hofer, Wysocki and Zehnder appeared in [**53**].

Lemma 0.31. *Assume that v_1 and v_2 are eigenvectors to the* same *eigenvalue η. Then $w(v_1) = w(v_2)$.*

Proof: If v_2 is a scalar multiple of v_1 the Lemma is obvious. Hence we can assume that v_1 and v_2 are linearly independent. Define
$$v \colon S^1 \to \mathbb{C}, \quad v(t) = v_1(t)\overline{v_2(t)}.$$
It follows that
$$\deg(\gamma_v) = \deg(\gamma_{v_1}) - \deg(\gamma_{v_2}).$$
It therefore remains to show that
$$\deg(\gamma_v) = 0.$$
This follows from the following Claim.

Claim: $v(t) \notin \mathbb{R},\ \forall\, t \in S^1$.

To prove the Claim we argue by contradiction and assume that there exists $t_0 \in S^1$ such that
$$v(t_0) \in \mathbb{R}.$$
Hence there exists $\tau \in \mathbb{R} \setminus \{0\}$ such that
$$v_1(t_0) = \tau v_2(t_0).$$
Define
$$v_3 \colon S^1 \to \mathbb{C}, \quad v_3 = v_1 - \tau v_2.$$
Since $Av_1 = \eta v_1$ and $Av_2 = \eta v_2$ it follows that
$$Av_3 = \eta v_3$$
which implies that v_3 is the solution of a linear first order ODE. On the other hand
$$v_3(t_0) = v_1(t_0) - \tau v_2(t_0) = 0$$
and therefore
$$v_3(t) = 0, \quad \forall\, t \in S^1.$$
By definition of v_3 this implies that v_1 and v_2 are linearly dependent. This contradiction proves the Claim and hence the Lemma. $\square$

As a consequence of Lemma 0.31 we can now associate to every $\eta \in \mathfrak{S}(A_S)$ a winding number, by

$$w(\eta) := w(\eta, S) := w(v) \tag{115}$$

where v is any eigenvector of A_S to the eigenvalue η. Indeed, Lemma 0.31 tells us that this is well-defined, independent of the choice of the eigenvector. We refer to $w(\eta)$ as the *winding number of the eigenvalue* η.

Let us examine the case $S = 0$. In this case $A_0 = -J\partial_t$, the eigenvalues are $2\pi\ell$ for every $\ell \in \mathbb{Z}$ with corresponding eigenvectors $v_0 e^{2\pi i \ell t}$ where $v_0 \in \mathbb{C}$. We conclude that

$$w(2\pi\ell, 0) = \ell. \tag{116}$$

In particular, we see that in the case $S = 0$ the winding is a monotone function in the eigenvalue. This is true in general.

Corollary 0.32 (Monotonicity). *Assume that* $S \in C^\infty([0,1], \mathrm{Sym}(2))$. *Then the map*

$$w\colon \mathfrak{S}(A_S) \to \mathbb{Z}, \quad \eta \mapsto w(\eta)$$

is monotone, i.e.,

$$\eta \leq \eta' \Longrightarrow w(\eta) \leq w(\eta').$$

Proof: Consider the homotopy $r \mapsto rS$ for $r \in [0,1]$. By (116) the assertion is true for $r = 0$. Since the eigenvalues vary continuously under perturbation we conclude that the assertion of the Corollary is true for every $r \in [0,1]$. □

By our convention of numbering the eigenvalues we obtain from (116) that

$$w(\eta_{2\ell}) = w(\eta_{2\ell-1}) = \ell - 1, \quad \ell \in \mathbb{Z}. \tag{117}$$

Define

$$\alpha(S) := \max\big\{w(\eta, S) : \eta \in \mathfrak{S}(A_S) \cap (-\infty, 0)\big\} \in \mathbb{Z} \tag{118}$$

and the *parity*

$$p(S) := \left\{ \begin{array}{ll} 0 & \text{if } \exists\, \eta \in \mathfrak{S}(A_S) \cap [0,\infty),\ \alpha(S) = w(\eta, S) \\ 1 & \qquad\qquad\qquad \text{else.} \end{array} \right. \tag{119}$$

The spectral flow theorem for two dimensional symplectic vector spaces gives rise to the following description of the Conley-Zehnder index.

Theorem 0.33. *Assume that* $S \in C^\infty([0,1], \mathrm{Sym}(2))$. *Then the Conley-Zehnder index satisfies*

$$\mu_{CZ}(\Psi_S) = 2\alpha(S) + p(S).$$

Proof: In view of Theorem 0.28 if Ψ_S is non-degenerate the Conley-Zehnder index is given by

$$\mu_{CZ}(\Psi_S) = \max\{k : \eta_k(S) < 0\} - 1 \tag{120}$$

and if Ψ_S is degenerate we use this formula as definition of the Conley-Zehnder index. By (117) we have

$$\alpha(S) = w(\eta_{2\alpha(S)+1}) = w(\eta_{2\alpha(S)+2}).$$

Therefore

$$\{\eta \in \mathfrak{S}(A_S) : w(\eta) = \alpha(S)\} = \{\eta_{2\alpha(S)+1}, \eta_{2\alpha(S)+2}\}.$$

Since $\eta_k \leq \eta_{k+1}$ the definition of $\alpha(S)$ implies that

$$\eta_{2\alpha(S)+1} < 0, \quad \eta_{2\alpha(S)+3} \geq 0.$$

Therefore, by definition of the parity we get

$$p(S) = \begin{cases} 0 & \text{if } \eta_{2\alpha(S)+2} \geq 0 \\ 1 & \text{if } \eta_{2\alpha(S)+2} < 0. \end{cases}$$

Hence

$$\max\{k : \eta_k(S) < 0\} = 2\alpha(S) + 1 + p(S)$$

and the theorem follows from (120). $\square$

CHAPTER 10

Convexity

Assume that $S \subset \mathbb{R}^{n+1}$ is a closed, connected hypersurface. It follows that we get a decomposition

$$\mathbb{R}^{n+1} \setminus S = M_- \cup M_+$$

into two connected components, where M_- is bounded and M_+ is unbounded. If $p \in S$ we denote by $N(p)$ the unit normal vector of S pointing into the unbounded component M_+. This leads to a smooth map

$$N \colon S \to S^n$$

referred to as the *Gauss map* which defines an orientation on S. Because $T_{N(p)}S^n$ and T_pS are parallel planes we can identify them canonically so that the differential of the Gauss map becomes a linear map

$$dN(p) \colon T_pS \to T_pS.$$

The *Gauss-Kronecker curvature* at $p \in S$ is defined as

$$K(p) := \det dN(p).$$

Note that if $n = 2$, i.e., S is a two dimensional surface, the Gauss-Kronecker curvature coincides with the Gauss curvature of the surface which is intrinsic.

We write S as a level set, i.e., we pick $f \in C^\infty(\mathbb{R}^{n+1}, \mathbb{R})$ such that 0 is a regular value of f and

$$S = f^{-1}(0).$$

We choose f in such a way that

$$M_- = \{x \in \mathbb{R}^{n+1} : f(x) < 0\}, \quad M_+ = \{x \in \mathbb{R}^{n+1} : f(x) > 0\}.$$

It follows that

$$N = \frac{\nabla f}{||\nabla f||}.$$

We abbreviate by

$$H_f(p) \colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$$

the Hessian of f at p. We get for a point $p \in S$ and two tangent vectors $v, w \in T_pS = \nabla f(p)^\perp$ the equality

$$\langle v, dN(p)w \rangle = \frac{1}{||\nabla f(p)||} \langle v, H_f(p)w \rangle.$$

In particular, if

$$\Pi_p \colon \mathbb{R}^{n+1} \to T_pS$$

is the orthogonal projection we can write

$$dN(p) = \frac{1}{||\nabla f(p)||} \Pi_p H_f(p)|_{T_pS} = \frac{1}{||\nabla f(p)||} \Pi_p H_f(p) \Pi_p^*$$

implying that $dN(p)$ is self-adjoint. We make the following definition

Definition 0.34. *S is called* strictly convex*, if $K(p) > 0$ for every $p \in S$.*

If the hypersurface is written as the level set of a function $S = f^{-1}(0)$ then for practical purposes it is useful to note that strict convexity is equivalent to the assertion that

$$\det(\Pi_p H_f(p)|_{\nabla f(p)^\perp}) > 0, \quad \forall\, p \in S.$$

However, observe that the notion of convexity only depends on S and not on the choice of the function f.

Lemma 0.35. *S is strict convex if and only if $dN(p)$ is positive definite for every $p \in S$.*

Proof: The implication " $\Longleftarrow$ " is obvious. It remains to check the implication " $\Longrightarrow$ ". Because $dN(p)$ is self-adjoint it follows that $dN(p)$ has n real eigenvalues counted with multiplicity. If $\mathfrak{S}(dN(p)) \subset \mathbb{R}$ is the spectrum of $dN(p)$ define

$$k \colon S \to \mathbb{R}, \quad p \mapsto \min\{r : r \in \mathfrak{S}(dN(p))\}$$

the smallest eigenvalue of $dN(p)$. Then k is a continuous function on S and we claim

$$k(p) > 0, \quad \forall\, p \in S. \tag{121}$$

We proof (121) in two steps. We first check

Step 1: *There exists $p_0 \in S$ such that $k(p_0) > 0$.*

To prove the assertion of Step 1 we denote $D(r) = \{x \in \mathbb{R}^{n+1} : ||x|| \leq r\}$ for $r \in (0, \infty)$ the closed ball of radius r and set

$$r_S := \min\{r \in (0, \infty) : S \subset D(r)\}.$$

Because S is compact r_S is finite. Moreover, there exists $p_0 \in \partial D_{r_S} = S^n_{r_S} = \{x \in \mathbb{R}^{n+1} : ||x|| = r_S\}$ the sphere of radius r_S such that

$$k^S(p_0) \geq k^{S^n_{r_S}}(p_0) = \frac{1}{r_S} > 0.$$

This finishes the proof of Step 1.

Step 2: *We prove (121).*

Because $K(p) = \det dN(p) > 0$ for every $p \in S$ it follows that $k(p) \neq 0$ for every $p \in S$. Since S is connected and k is continuous we either have $k(p) > 0$ for every $p \in S$ or $k(p) < 0$ for every $p \in S$. By Step 1 we conclude that $k(p) > 0$ for every $p \in S$. This establishes (121).

We are now in position to prove the Lemma. Because $k(p) > 0$ it follows that

$$\mathfrak{S}(dN(p)) \in (0, \infty), \quad \forall\, p \in S.$$

Therefore $dN(p)$ is positive definite for every $p \in S$ and the Lemma is proved. $\square$

For the following Lemma recall that $M_- \subset \mathbb{R}^{n+1}$ denotes the bounded region of $\mathbb{R}^{n+1} \setminus S$.

Lemma 0.36. *Assume that S is strictly convex. Then M_- is convex in the sense that if $x, y \in M_-$ and $\lambda \in [0,1]$ then $\lambda x + (1 - \lambda y) \in M_-$, i.e., the line segment between x and y is contained in M_-.*

Proof: The assertion of the Lemma is true in any dimension. However, the proof we present just works if $n \geq 2$.

For $p \in S$ we introduce the half-space

$$H_p := \{x \in \mathbb{R}^{n+1} : \langle p - x, N(p) \rangle > 0\}.$$

We claim

$$M_- = \bigcap_{p \in S} H_p. \tag{122}$$

We first check

$$\bigcap_{p \in S} H_p \subset M_-. \tag{123}$$

To prove that we note the following equivalences

$$\begin{aligned} \bigcap_{p \in S} H_p \subset M_- \quad &\Longleftrightarrow \quad M_+ \cup S = M_-^c \subset \Big(\bigcap_{p \in S} H_p \Big)^c = \bigcup_{p \in S} H_p^c \\ &\Longleftrightarrow \quad M_+ = \bigcup_{p \in S} \big\{ x \in \mathbb{R}^{n+1} : \langle x - p, N(p) \rangle > 0 \big\} \end{aligned}$$

If $x \in M_+$ choose $p_0 \in S$ and consider the line segment

$$[0,1] \to \mathbb{R}^{n+1}, \quad t \mapsto (1-t)p_0 + tx.$$

Define

$$t_0 := \max \big\{ t \in [0,1] : (1-t)p_0 + tx \in S \big\}.$$

Because $x \in M_+$ it follows that

$$t_0 < 1.$$

We set

$$p := (1 - t_0)p_0 + t_0 x \in S.$$

Since $x \in M_+$ it follows that

$$\langle x - p, N(p) \rangle > 0.$$

This establishes (123). Note that for the proof of (123) we did not use yet the convexity of S. We next check that

$$M_- \subset \bigcap_{p \in S} H_p. \tag{124}$$

We need to show that for every $p \in S$ it holds that

$$M_- \subset H_p.$$

For $u \in S^n = \{x \in \mathbb{R}^{n+1} : ||x|| = 1\}$ consider

$$F_u \colon \mathbb{R}^{n+1} \to \mathbb{R}, \quad x \mapsto \langle x, u \rangle$$

and abbreviate

$$f_u := F_u|_S \colon S \to \mathbb{R}$$

the restriction of F_u to S. We next discuss the critical points of f_u. Recall from the method of Lagrange multipliers that if $F \colon \mathbb{R}^{n+1} \to \mathbb{R}$ is a smooth function and $f = F|_S$, then $p \in \mathrm{crit} f$ if and only if there exists $\lambda_p \in \mathbb{R}$ referred to as the Lagrange multiplier such that

$$\nabla f(p) = \lambda_p N(p).$$

Moreover, if $\Pi_p \colon \mathbb{R}^{n+1} \to T_p S$ denotes the orthogonal projection the Hessian of f at p is given by

$$H_f(p) = \big(\Pi_p H_F(p)|_{T_p S} + \lambda_p dN(p)\big) \colon T_p S \to T_p S.$$

In the case of our interest we have

$$\nabla F_u(p) = u, \quad H_{F_u}(p) = 0$$

and therefore $p \in \mathrm{crit} f_u$ if and only if

$$u = \lambda_p N(p)$$

where

$$\lambda_p = \pm 1$$

because $||u|| = ||N(p)|| = 1$. Moreover, the Hessian is given by

$$H_{f_u}(p) = \lambda_p dN(p).$$

Because S is strictly convex Lemma 0.35 tells us that $dN(p)$ is positive definite for every $p \in S$. Therefore H_{f_u} is either positive definite or negative definite for every $p \in S$. We have shown that $f_u \colon S \to \mathbb{R}$ is a Morse function all whose critical points are either maxima or minima. At this point we need the assumption that $n \geq 2$, i.e., that the dimension of S is at least two. Namely, because S is connected it follows that in this case f_u has precisely one strict maximum and precisely one strict minimum and no other critical points.

For $p \in S$ choose $u = N(p)$. It follows that

$$p \in \mathrm{crit} f_{N(p)}$$

with

$$\lambda_p = 1.$$

Therefore p is the unique strict maximum of $f_{N(p)}$. In particular, for every $q \in S \setminus \{p\}$ we have

$$\langle q, N(p)\rangle = f_{N(p)}(q) < f_{N(p)}(p) = \langle p, N(p)\rangle$$

implying that

$$0 < \langle p, N(p)\rangle - \langle q, N(p)\rangle = \langle p - q, N(p)\rangle, \quad \forall\, q \in S \setminus \{p\}.$$

It follows that

$$S \setminus \{p\} \subset H_p$$

and consequently

$$M_- \subset H_p.$$

This establishes (124) and therefore together with (123) we obtain (122).

In view of (122) the Lemma can now be proved as follows. Note that H_p is convex for every $p \in S$. Because the intersection of convex sets is convex it follows that $\bigcap_{p \in S} H_p$ is convex as well. Hence (122) implies the Lemma. $\square$

Suppose that $N \subset \mathbb{C}^2$ is strictly convex. Because convexity is preserved under affine transformations we can assume without loss of generality that $0 \in M_-$, the bounded part of $\mathbb{C}^2 \setminus N$. Therefore by Lemma 0.36 M_- is star-shaped. It follows that the restriction $\lambda|N$ of the standard Liouville form

$$\lambda = \frac{1}{2}(x_1 dy_1 - y_1 dx_1 + x_2 dy_2 - y_2 dx_2)$$

to N is a contact form on S.

The following Theorem appeared in [**56**]

Theorem 0.37 (Hofer-Wysocki-Zehnder). *Assume that $N \subset \mathbb{C}^2$ is a strictly convex hypersurface such that 0 lies in the bounded part of $\mathbb{C}^2 \setminus N$. Then the contact manifold $(N, \lambda|_N)$ is dynamically convex where λ is the standard Liouville form on $\mathbb{C}$.*

As preparation for the proof suppose that $N \subset \mathbb{C}^2$ meets the assumption of Theorem 0.37. For $z \in \mathbb{C}^2 \setminus \{0\}$ there exists a unique $\lambda_z \in (0, \infty)$ with the property that

$$\lambda_z z \in N.$$

Define

$$F_N \colon \mathbb{C}^2 \setminus \{0\} \to \mathbb{R}, \quad z \mapsto \frac{1}{\lambda_z^2}.$$

It follows that

$$N = F_N^{-1}(1).$$

We denote by X_{F_N} the Hamiltonian vector field of F_N with respect to the standard symplectic structure $\omega = d\lambda$ on $\mathbb{C}^2$ defined implicitly by the condition

$$dF_N = \omega(\cdot, X_{F_N}).$$

Lemma 0.38. *For $z \in N$ it holds that*

$$X_{F_N}(z) = R(z)$$

where R is the Reeb vector field of (N, λ_N).

Proof: For $v \in T_z N$ we compute

$$d\lambda(z)(X_{F_N}(z), v) = \omega_z(X_{F_N}, v) = -dF_N(z)v = 0$$

where for the last equality we have used that $N = F_N^{-1}(1)$. It follows that

$$X_{F_N}|_N \in \ker d\lambda|_N.$$

It remains to show that

$$\lambda(X_{F_N})|_N = 1.$$

Note that F_N is homogeneous of degree 2, i.e.

$$F_N(rz) = r^2 F_N(z), \quad z \in \mathbb{C}^2 \setminus \{0\},\ r > 0.$$

Consequently

$$dF_N(z)z = \frac{d}{dr}\Big|_{r=1} F_N(rz) = 2rF_N(z)\big|_{r=1} = 2F_N(z) \tag{125}$$

which is known as Euler's formula. Note further that if $z, \widehat{z} \in \mathbb{C}$ the equation

$$\lambda(z)\widehat{z} = \frac{1}{2}\omega(z, \widehat{z}) \tag{126}$$

holds true. Using (125) and (126) we compute for $z \in N = F_N^{-1}(1)$

$$\lambda(z)(X_{F_N(z)}) = \frac{1}{2}\omega(z, X_{F_N}(z)) = \frac{1}{2}dF_N(z)z = F_N(z) = 1.$$

This finishes the proof of the Lemma. □

Suppose that $\gamma \in C^\infty(S^1, N)$ is a periodic Reeb orbit of period τ. Abbreviate by

$$\phi_N^t \colon \mathbb{C}^2 \to \mathbb{C}^2, \quad t \in \mathbb{R}$$

the flow of the vector field X_{F_N}. Because γ is a periodic Reeb orbit we have in view of Lemma 0.38 for every $t \in \mathbb{R}$

$$\phi_N^{\tau t}(\gamma(0)) = \gamma(t).$$

For $t \in [0,1]$ consider the smooth path of symplectic matrices

$$\Psi_\gamma^t := d\phi_N^{\tau t}(\gamma(0)) \colon \mathbb{C}^2 \to \mathbb{C}^2. \tag{127}$$

Lemma 0.39. *The path Ψ_γ^t has the following properties*

(i): $\Psi_\gamma^0 = \mathrm{id}$,

(ii): Ψ_γ^t *satisfies the ODE*

$$\partial_t \Psi_\gamma^t = J\tau H_{F_N}(\gamma(t))\Psi_\gamma^t,$$

where $J \colon \mathbb{C}^2 \to \mathbb{C}^2$ is multiplication by i and H_{F_N} is the Hessian of F_N which is positive definite by Lemma 0.35,

(iii): $\Psi_\gamma^1\big(R(\gamma(0))\big) = R(\gamma(0))$ *and* $\Psi_\gamma^1\big(\gamma(0)\big) = \gamma(0)$.

Proof: Properties (i) and (ii) are immediate. To explain why the first equation of property (iii) holds note that because the Hamiltonian vector field X_{F_N} is autonomous (time independent) its flow satisfies

$$\phi_N^{s+t} = \phi_N^t \phi_N^s$$

and therefore

$$X_{F_N} \circ \phi_N^t = \frac{d}{ds}\bigg|_{s=0} \phi_N^{s+t} = \frac{d}{ds}\bigg|_{s=0} \phi_N^t \phi_N^s = d\phi_N^t(X_{F_N}).$$

Because γ is a periodic Reeb orbit of period τ we have

$$\phi_N^\tau(\gamma(0)) = \gamma(0)$$

and therefore

$$R(\gamma(0)) = X_{F_N}(\gamma(0)) = d\phi_N^\tau(\gamma(0))\big(X_{F_N}(\gamma(0))\big) = \Psi_\gamma^1\big(R(\gamma(0))\big).$$

This explains the first equation in property (iii).

It remains to check the second equation in property (iii). We first recall that because F_N is homogeneous of degree 2 Euler's formula (125) holds for every $z \in \mathbb{C}$. Differentiating once more we obtain

$$d^2F_N(z)z + dF_N(z) = 2dF_N(z)$$

which implies that

$$d^2F_N(z)z = dF_N(z).$$

In particular, we have

$$H_{F_N}(z)z = \nabla F_N(z)$$

and therefore

$$JH_{F_N}(z)z = J\nabla F_N(z) = X_{F_N}(z). \tag{128}$$

We claim that

$$d\phi_N^t(z)z = \phi_N^t(z) \in \mathbb{C}^2, \quad \forall\, z \in \mathbb{C}^2. \tag{129}$$

To check this equation we fix $z \in \mathbb{C}^2$ and consider the path

$$w\colon \mathbb{R} \to \mathbb{C}^2, \quad t \mapsto d\phi_N^t(z)z - \phi_N^t(z).$$

Note that

$$w(0) = d\phi_N^0(z)z - \phi_N^0(z) = z - z = 0. \tag{130}$$

Moreover,

$$\begin{aligned} \frac{d}{dt}w(t) &= \frac{d}{dt}d\phi_N^t(z)z - \frac{d}{dt}\phi_N^t(z) \\ &= JH_{F_N}(\phi_N^t(z))d\phi_N^t(z)z - X_{F_N}(\phi_N^t(z)) \\ &= JH_{F_N}(\phi_N^t(z))d\phi_N^t(z)z - JH_{F_N}(\phi_N^t(z))\phi_N^t(z) \\ &= JH_{F_N}(\phi_N^t(z))(d\phi_N^t(z)z - \phi_N^t(z)) \\ &= JH_{F_N}(\phi_N^t(z))w(t). \end{aligned} \tag{131}$$

Here we have used in the third equation (128). Combining (130) and (131) we obtain

$$w(t) = 0, \quad \forall\, t \in \mathbb{R}.$$

This proves (129). Setting $t = \tau$ and $z = \gamma(0)$ we obtain

$$\Psi_\gamma^1\big(\gamma(0)\big) = d\phi_N^\tau(\gamma(0))\big(\gamma(0)\big) = \phi_N^\tau(\gamma(0)) = \gamma(0).$$

This finishes the proof of the Lemma. □

For the next Lemma recall from Lemma 0.9 that if $L \in \Lambda$ the Lagrangian Grassmannian we have a canonical identification

$$T_L\lambda \to S^2(L), \quad \widehat{L} \mapsto Q^{\widehat{L}}$$

where $S^2(L)$ are the quadratic forms on L. Recall further that if $\Psi \in \mathrm{Sp}(n)$ is a symplectic transformation then its graph Γ_Ψ is a Lagrangian in the symplectic vector space $(\mathbb{C}^n \otimes \mathbb{C}^n, -\omega \otimes \omega)$. We can now state the next Lemma.

Lemma 0.40. *Suppose that $\Psi\colon (-\epsilon, \epsilon) \to \mathrm{Sp}(n)$ is a smooth path of symplectic matrices. Then for*

$$\Gamma_{\Psi'(0)} = \frac{d}{dt}\Big|_{t=0} \Gamma_{\Psi(t)} \in T_{\Gamma_{\Psi(0)}}\Lambda(\mathbb{C}^n \otimes \mathbb{C}^n, -\omega \otimes \omega)$$

the corresponding quadratic form

$$Q^{\Gamma_{\Psi'(0)}} \in S^2(\Gamma_{\Psi(0)})$$

is given for $(z, \Psi(0)z) \in \Gamma_{\Psi(0)}$ where $z \in \mathbb{C}^n$ by

$$Q^{\Gamma_{\Psi'(0)}}(z, \Psi(0)z) = \langle \Psi(0)z, S\Psi_0 z\rangle$$

where

$$S = -J\Psi'(0)\Psi(0)^{-1}.$$

Proof: We choose as Lagrangian complement of $\Gamma_{\Psi(0)}$ the Lagrangian

$$\Gamma_{-\Psi(0)} = \big\{(z, -\Psi(0)z) : z \in \mathbb{C}^n\big\}.$$

If $z \in \mathbb{C}^n$ and $t \in (-\epsilon, \epsilon)$ we define $w_z(t) \in \mathbb{C}^n$ by the condition that

$$\big(z, \Psi(0)z\big) + \big(w_z(t), -\Psi(0)w_z(t)\big) \in \Gamma_{\Psi(t)}$$

or equivalently

$$\big(z + w_z(t), \Psi(0)z - \Psi(0)w_z(t)\big) \in \Gamma_{\Psi(t)}.$$

This implies that

$$\Psi(t)z + \Psi(t)w_z(t) = \Psi(t)(z + w_z(t)) = \Psi(0)z - \Psi(0)w_z(t).$$

Differentiating this expression we get

$$\Psi'(t)z + \Psi'(t)w_z(t) + \Psi(t)w_z'(t) = -\Psi(0)w_z'(t).$$

Because $w_z(0) = 0$ we obtain from that

$$\Psi'(0)z + \Psi(0)w_z'(0) = -\Psi(0)w_z'(0)$$

implying

$$w_z'(0) = -\frac{1}{2}\Psi(0)^{-1}\Psi'(0)z.$$

By definition of the quadratic from we compute taking advantage of the fact that $\Psi(0)$ is symplectic

$$\begin{aligned}
Q^{\Gamma_{\Psi'(0)}}\big(z, \Psi(0)z\big) &= -\omega \oplus \omega\Big((z, \Psi(0)z), (w_z'(0), -\Psi(0)w_z'(0)\Big) \\
&= -\omega(z, w_z'(0)) - \omega(\Psi(0)z, \Psi(0)w_z'(0)) \\
&= -\omega(z, w_z'(0)) - \omega(z, w_z'(0)) \\
&= -2\omega(z, w_z'(0)) \\
&= \omega(z, \Psi(0)^{-1}\Psi'(0)z) \\
&= \omega(\Psi(0)z, \Psi'(0)z) \\
&= \omega(\Psi(0)z, JS\Psi(0)z) \\
&= \langle \Psi(0)z, S\Psi(0)z\rangle.
\end{aligned}$$

This finishes the proof of the Lemma. $\square$

Corollary 0.41. *Assume that $\Psi\colon [0,1] \to Sp(n)$ is a smooth path of symplectic matrices satisfying $\Psi'(t) = JS(t)\Psi(t)$ with $S(t)$ positive definite for every $t \in [0,1]$. Then the crossing form $C(\Gamma_\Psi, \Delta, t)$ is positive definite for every $t \in [0,1]$.*

Proof: Be definition of the crossing form we have

$$C(\Gamma_\Psi, \Delta, t) = Q^{\Gamma_{\Psi'(t)}}|_{\Delta \cap \Gamma_{\Psi(t)}}.$$

The Corollary follows now from Lemma 0.40 and the assumption that S is positive definite. $\square$

Corollary 0.42. *Assume that $N \subset \mathbb{C}^2$ is a strictly convex hypersurface and γ is a periodic Reeb orbit of period τ on N. Then*

$$\mu_\Delta(\Gamma_{\Psi_\gamma}) \geq 3 + \frac{1}{2}\dim\ker\big(d^\xi\phi_R^\tau(\gamma(0)) - \mathrm{id}\big) \geq 3$$

Proof: Using the definition of the Maslov index for paths (106),Corollary 0.41 and assertion (i) and (iii) in Lemma 0.39 we estimate

$$\begin{aligned}
\mu_\Delta(\Gamma_{\Psi_\gamma}) &= \frac{1}{2}\mathrm{sign}C(\Gamma_{\Psi_\gamma},\Delta,0)+\sum_{0<t<1}\mathrm{sign}C(\Gamma_{\Psi_\gamma},\Delta,t)+\frac{1}{2}\mathrm{sign}C(\Gamma_{\Psi_\gamma},\Delta,1)\\
&= \frac{1}{2}\dim\big(\Delta\cap\Gamma_{\Psi_\gamma}(0)\big)+\sum_{0<t<1}\dim\big(\Delta\cap\Gamma_{\Psi_\gamma}(t)\big)\\
&\quad+\frac{1}{2}\dim\big(\Delta\cap\Gamma_{\Psi_\gamma}(1)\big)\\
&= \frac{1}{2}\dim\big(\Delta\big)+\sum_{0<t<1}\dim\big(\Delta\cap\Gamma_{\Psi_\gamma}(t)\big)\\
&\quad+\frac{1}{2}\Big(2+\dim\ker\big(d^\xi\phi_R^\tau(\gamma(0))-\mathrm{id}\big)\Big)\\
&= 3+\sum_{0<t<1}\dim\big(\Delta\cap\Gamma_{\Psi_\gamma}(t)\big)+\frac{1}{2}\dim\ker\big(d^\xi\phi_R^\tau(\gamma(0))-\mathrm{id}\big)\\
&\geq 3+\dim\ker\big(d^\xi\phi_R^\tau(\gamma(0))-\mathrm{id}\big).
\end{aligned}$$

This proves the Corollary. □

Proof of Theorem 0.37: It follows Lemma 0.36) that N bounds a star-shaped domain. Therefore N is diffeomorphic to the sphere S^3. In particular, $\pi_2(N)=\{0\}$ and therefore the homomorphism $I_{c_1}\colon\pi_2(N)\to\mathbb{Z}$ vanishes for trivial reasons. Assume that γ is a periodic orbit on N of period τ. Because N is simply connected γ is contractible so that we can choose a filling disk for γ, i.e., a smooth map $\overline{\gamma}\colon D\to N\subset\mathbb{C}^2$ satisfying $\overline{\gamma}(e^{2\pi it})=\gamma(t)$ for every $t\in S^1$. Note that the vector bundle $\overline{\gamma}^*\mathbb{C}^2\to D$ splits as

$$\overline{\gamma}^*\mathbb{C}^2=\overline{\gamma}^*\xi\oplus\overline{\gamma}^*\eta \tag{132}$$

where we abbreviate

$$\eta=\langle X,R\rangle$$

where the vector field X is defined as

$$X(x)=x,\quad x\in\mathbb{C}^2.$$

Note that we have a canonical trivialization

$$\mathfrak{T}_\eta\colon\overline{\gamma}^*\eta\to D\times\mathbb{C},\quad aX+bR\mapsto a+ib,\ a,b\in\mathbb{R}.$$

It follows from Lemma 0.38 and Lemma 0.39 that the map $\Psi_\gamma^t\colon\mathbb{C}^2\to\mathbb{C}^2$ for $t\in[0,1]$ respects the splitting (132), i.e.

$$\Psi_\gamma^t\colon\xi_{\gamma(0)}\to\xi_{\gamma(t)},\quad \Psi_\gamma^t\colon\eta_{\gamma(0)}\to\eta_{\gamma(t)}.$$

Indeed, we have

$$\Psi_\gamma^t|_\xi=d^\xi\phi_R^{t\tau}(\gamma(0))$$

and

$$\mathfrak{T}_{\eta,\gamma(t)}\Psi_\gamma^t\mathfrak{T}_{\eta,\gamma(0)}^{-1}=\mathrm{id}\colon\mathbb{C}\to\mathbb{C}.$$

Choose in addition a symplectic trivialization

$$\mathfrak{T}_\xi\colon\overline{\gamma}^*\xi\to D\times\mathbb{C}.$$

Hence we obtain a symplectic bundle map

$$\mathfrak{T} := \mathfrak{T}_\xi \oplus \mathfrak{T}_\eta \colon D \times \mathbb{C}^2 \to D \times \mathbb{C}^2.$$

Because the disk D is contractible the map $\mathfrak{T}$ is homotopic as a bundle map to the identity map from $D \times \mathbb{C}^2$ to itself. In particular, if we introduce the path of symplectic matrices

$$\Psi_\gamma^{\mathfrak{T}} \colon [0,1] \to \mathrm{Sp}(2)$$

defined for $t \in [0,1]$ as

$$(\Psi_\gamma^{\mathfrak{T}})^t = \mathfrak{T}_{\gamma(t)} \Psi_\gamma^t \mathfrak{T}_{\gamma(0)}^{-1} \colon \mathbb{C}^2 \to \mathbb{C}^2$$

we obtain by homotopy invariance of the Maslov index and Corollary 0.42

$$\mu_\Delta(\Gamma_{\Psi_\gamma^{\mathfrak{T}}}) = \mu_\Delta(\Gamma_{\Psi_\gamma}) \geq 3 + \dim\ker\big(d^\xi \phi_R^\tau(\gamma(0)) - \mathrm{id}\big). \tag{133}$$

If (V_1, ω_1) and (V_2, ω_2) are two symplectic vector spaces, $L_1 \subset V_1$, $L_2 \subset V_2$ are Lagrangian subspaces, $\lambda_1 \colon [0,1] \to \Lambda(V_1, \omega_1)$ is a path of Lagrangians in V_1, and $\lambda_2 \colon [0,1] \to \Lambda(V_2, \omega_2)$ is a path of Lagrangians in V_2 we obtain a Lagrangian

$$L_1 \oplus L_2 \subset V_1 \times V_2$$

and a path of Lagrangians

$$\lambda_1 \oplus \lambda_2 \colon [0,1] \to \Lambda(V_1 \oplus V_2, \omega_1 \oplus \omega_2).$$

The Maslov index satisfies

$$\mu_{L_1 \otimes L_2}(\lambda_1 \oplus \lambda_2) = \mu_{L_1}(\lambda_1) + \mu_{L_2}(\lambda_2).$$

In view of the splitting

$$\Psi_\gamma^{\mathfrak{T}} = \Psi_\gamma^{\mathfrak{T}}|_\xi \oplus \Psi_\gamma^{\mathfrak{T}}|_\eta$$

we obtain

$$\mu_\Delta(\Gamma_{\Psi_\gamma^{\mathfrak{T}}}) = \mu_\Delta(\Gamma_{\Psi_\gamma^{\mathfrak{T}}|_\xi}) + \mu_\Delta(\Gamma_{\Psi_\gamma^{\mathfrak{T}}|_\eta}). \tag{134}$$

Because $\Psi_\gamma^{\mathfrak{T}}|_\eta(t) = \mathrm{id}$ for every $t \in [0,1]$ we conclude that

$$\mu_\Delta(\Gamma_{\Psi_\gamma^{\mathfrak{T}}|_\eta}) = 0. \tag{135}$$

Combining (133), (134), and (135) we obtain the inequality

$$\mu_\Delta(\Gamma_{\Psi_\gamma^{\mathfrak{T}}|_\xi}) \geq 3 + \dim\ker\big(d^\xi \phi_R^\tau(\gamma(0)) - \mathrm{id}\big). \tag{136}$$

We distinguish two cases

Case 1: *The periodic orbit γ is non-degenerate.*

In this case it follows from Definition of the Conley-Zehnder index that

$$\mu_{CZ}(\gamma) = \mu_\Delta(\Gamma_{\Psi_\gamma^{\mathfrak{T}}|_\xi}).$$

It now follows from (136) that

$$\mu_{CZ}(\gamma) \geq 3.$$

Case 2: *The periodic orbit γ is degenerate.*

In this case we consider for $\epsilon > 0$ a smooth path of symplectic matrices $\Psi_\epsilon \colon [0, 1+\epsilon] \to \mathrm{Sp}(1)$ with the property that $\Psi_\epsilon(t) = \Psi_\gamma^{\mathfrak{T}}|_\xi(t)$ for every $t \in [0,1]$ and there are no further crossings of Γ_{Ψ_ϵ} with the closure of the Maslov pseudo-cocycle $\overline{\Lambda^1_\Delta}$ in $(1, 1+\epsilon]$. It follows from the definition of the Maslov index for paths that

$$\begin{aligned}(137)\quad \mu_\Delta(\Gamma_{\Psi_\epsilon}) &= \mu_\Delta(\Gamma_{\Psi^{\mathfrak{T}}_\gamma|_\xi}) + \frac{1}{2}\mathrm{sign} C(\Gamma_{\Psi^{\mathfrak{T}}_\gamma|_\xi}, \Delta, 1)\\ &= \mu_\Delta(\Gamma_{\Psi^{\mathfrak{T}}_\gamma|_\xi}) + \frac{1}{2}\dim\ker\big(d^\xi\phi^\tau_R(\gamma(0)) - \mathrm{id}\big)\\ &\geq 3 + \frac{1}{2}\dim\ker\big(d^\xi\phi^\tau_R(\gamma(0)) - \mathrm{id}\big) + \frac{1}{2}\dim\ker\big(d^\xi\phi^\tau_R(\gamma(0)) - \mathrm{id}\big)\\ &= 3 + \dim\ker\big(d^\xi\phi^\tau_R(\gamma(0)) - \mathrm{id}\big).\end{aligned}$$

Here we have used in the second equation Lemma 0.40 together with the assumption that N is strictly convex and for the inequality we used (136). Because the path Γ_{Ψ_ϵ} is non-degenerate we have by Definition of the Conley-Zehnder index

$$\mu_\Delta(\Gamma_{\Psi_\epsilon}) = \mu_{CZ}(\Psi_\epsilon). \tag{138}$$

In view of the continuity of eigenvalues we have

$$\begin{aligned}(139)\quad \mu_{CZ}(\Psi^{\mathfrak{T}}_\gamma) &\geq \lim_{\epsilon\to 0}\mu_{CZ}(\Psi_\epsilon) - \dim\ker\big(A_{\Psi^{\mathfrak{T}}_\gamma}\big)\\ &= \mu_{CZ}(\Psi_\epsilon) - \dim\ker\big(d^\xi\phi^\tau_R(\gamma(0)) - \mathrm{id}\big)\\ &\geq 3 + \dim\ker\big(d^\xi\phi^\tau_R(\gamma(0)) - \mathrm{id}\big) - \dim\ker\big(d^\xi\phi^\tau_R(\gamma(0)) - \mathrm{id}\big)\\ &= 3\end{aligned}$$

where for the second inequality we have used (137) and (138). Because $\mu_{CZ}(\gamma) = \mu_{CZ}(\Psi^{\mathfrak{T}}_\gamma)$ inequality (139) finishes the proof of the theorem. □

0.1. Connected sum revisited: Hamiltonian flow near a critical point of index 1. In this section we will see how a connected sum can give a dynamical obstruction to convexity.

Consider a symplectic manifold (M^4, ω) with Hamiltonian $H: M \to \mathbb{R}$. We consider a non-degenerate critical point q_0 of index 1, so we can write

$$H(x) = Q(x,x) + R(x),$$

where $R(x) = o(|x|^2)$, so $\lim_{x\to 0}\frac{R(x)}{|x|^2} = 0$.

We first investigate the Hamiltonian $H_Q : x \mapsto Q(x,x)$, and then argue that the results continue to hold qualitatively when R is sufficiently small.

Lemma 0.43. *We consider a quadratic Hamiltonian H_Q with a non-degenerate critical point of index 1. Fix $c > 0$. Then every level set $H_Q = c$ has a unique simple periodic orbit γ_c (up to reparametrization). This periodic orbit is transversely non-degenerate and its Conley-Zehnder index equals 2.*

Proof. By Proposition 6.1, we can find symplectic coordinates such that H_Q has the form

$$(\xi_1, \xi_2; \eta_1, \eta_2) \longmapsto -2a\eta_1\eta_2 + b(\xi_1^2 + \xi_2^2)$$

with $a, b > 0$. The symplectic form is given by $d\eta_1 \wedge d\eta_2 + d\xi_1 \wedge d\xi_2$, so the Hamiltonian vector field is given by

$$X_{H_Q} = 2a\partial_{\eta_1} - 2a\partial_{\eta_2} + -b\xi_2\partial_{\xi_1} + b\xi_1\partial_{\xi_2}.$$

This is linear, so we obtain its time-t flow by exponentiating

$$\phi^t_{H_Q}(\eta_1,\eta_2;\xi_1,\xi_2) = \begin{pmatrix} e^{2at} & 0 & 0 & 0 \\ 0 & e^{-2at} & 0 & 0 \\ 0 & 0 & \cos(bt) & -\sin(bt) \\ 0 & 0 & \sin(bt) & \cos(bt) \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \xi_1 \\ \xi_2 \end{pmatrix}.$$

Clearly, if η_1 or η_2 at time $t=0$ are non-zero, then the solution cannot be periodic: $|\eta_1|$ is strictly increasing when non-zero and $|\eta_2|$ is strictly decreasing when non-zero. This leaves initial conditions with $\eta_1 = \eta_2 = 0$. The solution is clearly periodic in the latter case with period $2\pi/b$. If we fix $s_c = (0,0;\sqrt{\frac{c}{b}},0)$ as initial point, then we find

$$\gamma_c(t) = \sqrt{\frac{c}{b}}(0,0;\cos(bt),\sin(bt).$$

To see that this orbit is transversely non-degenerate, we note that

$$T_{s_c}\Sigma_c = span((1,0,0,0),(0,0,1,0),X_H(s_c)).$$

The linearized time-$t=2\pi/b$ flow sends $(1,0,0,0) \mapsto (e^{4\pi a/b},0,0,0)$ and $(0,1,0,0) \mapsto (0,e^{-4\pi a/b},0,0)$, which clearly does not have any eigenvalue equal to 1.

That leaves the Conley-Zehnder index. Rather than explicitly trivializing the contact structure over a disk, we use the following standard trick. First note that the linearized flow extends to a path of symplectic matrices on $(\mathbb{R}^4,\omega_0)$. Namely, with respect to the standard symplectic basis of $\mathbb{R}^4$ we have

$$\psi_{\mathbb{R}^4} : t \longmapsto \begin{pmatrix} e^{2at} & 0 & 0 & 0 \\ 0 & e^{-2at} & 0 & 0 \\ 0 & 0 & \cos(bt) & -\sin(bt) \\ 0 & 0 & \sin(bt) & \cos(bt) \end{pmatrix}.$$

For later use, let us call the path of abstract (not depending on basis) linear symplectic maps the *extended linearized flow* and denote this path by ψ. Note each symplectic matrix $\psi_{\mathbb{R}^4}(t)$ has the direct sum decomposition

$$\psi_{\mathbb{R}^4}(t) = \begin{pmatrix} e^{2at} & 0 \\ 0 & e^{-2at} \end{pmatrix} \oplus \begin{pmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{pmatrix}.$$

We compute the Maslov index of the former path with the crossing formula. There is only one crossing at $t=0$, and its signature is zero, so we have

$$\mu(\begin{pmatrix} e^{2at} & 0 \\ 0 & e^{-2at} \end{pmatrix}, t=0,\dots,2\pi/b) = \frac{1}{2}\cdot 0.$$

We use the following lemma to compute the Maslov index of the latter path.

Lemma 0.44. *Consider a path of symplectic matrices of the form*

$$\Phi : t \longmapsto \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}$$

with t ranging from 0 to T. Then

$$\mu(\Phi) = \lfloor \frac{T}{2\pi} \rfloor + \lceil \frac{T}{2\pi} \rceil.$$

PROOF. We use the crossing formula to compute the index. First note that there are crossings at every integer multiple of 2π. At each crossing the crossing form is defined on all of $\mathbb{R}^2$ and we have

$$\operatorname{sign}\omega_0(\cdot, \frac{d}{dt}\Phi\cdot) = \operatorname{sign}\omega_0(\cdot, J_0\cdot) = 2,$$

so every interior crossing, meaning a crossing t with $t \in]0,T[$, gives a contribution of 2, and the crossings at $t=0$ and $t=T$ (if there is one) contribute 1. We now count the crossings.

First we assume that T is not divisible by 2π. Then the number of interior crossings equals $\lfloor T/2\pi \rfloor$. There is a crossing at $t=0$, and none at $t=T$, so

$$\mu(\Phi) = 1 + 2\cdot\lfloor T/2\pi\rfloor = \lfloor T/2\pi\rfloor + \lceil T/2\pi\rceil.$$

To see the last step, we just need to observe that $\lfloor T/2\pi\rfloor$ is 1 smaller than $\lceil T/2\pi\rceil$ since T is not divisible by 2π. The case when T is divisible by 2π can be handled similarly by taking care of an additional crossing at the end of the path of symplectic matrices. □

Applying this lemma we find that $\mu(\psi_{\mathbb{R}^4}) = 2$.

On the other hand, we want to compute the Conley-Zehnder index. Take a spanning disk $d_c : D^2 \to \Sigma_c$ capping off γ_c. Choose a symplectic trivialization ϵ_ξ of the contact structure ξ along the disk $d_c(D^2)$. Furthermore, we fix the following symplectic trivialization of the symplectic complement of ξ in $T\mathbb{R}^4$ with respect to ω_0,

$$\begin{aligned} \epsilon_{\xi^{\omega_0}} : D^2 \times \mathbb{R}^2 &\longrightarrow d_c^*\xi^{\omega_0} \\ (z; a_1, a_2) &\longmapsto (d_c(z); (Z\circ d_c(z), R\circ d_c(z)). \end{aligned}$$

Here Z and R are the Liouville and Reeb vector field, respectively.

We write ψ with respect to the trivialization $\epsilon_\xi \oplus \epsilon_{\xi^{\omega_0}}$, and obtain a path of symplectic matrices

$$\psi_\xi \oplus \psi_{\xi^{\omega_0}}.$$

Since $c_1(T\mathbb{R}^4) = 0$, the Maslov index of the extended linearized flow ψ does not depend on the choice of the trivialization of $T\mathbb{R}^4$ along $d_c(D^2)$, so we see that

$$\mu(\psi_\xi \oplus \psi_{\xi^{\omega_0}}) = \mu(\psi_{\mathbb{R}^4}) = 2.$$

On the other hand, we know by the direct sum axiom that

$$\mu(\psi_\xi \oplus \psi_{\xi^{\omega_0}}) = \mu(\psi_\xi) + \mu(\psi_{\xi^{\omega_0}}).$$

We compute the effect of extended linearized flow ψ on vector fields Z and R. We find

$$T\phi^t_{H_Q} Z\circ\gamma_c(0) = Z\circ\gamma_c(t), \quad \text{and} \quad T\phi^t_{H_Q} R\circ\gamma_c(0) = R\circ\gamma_c(t),$$

which means that $\psi_{\xi^{\omega_0}}$ is the constant path. A constant path of symplectic matrices has vanishing Maslov index, so knowing that γ_c is non-degenerate, we conclude that

$$2 = \mu(\psi_\xi \oplus \psi_{\xi^{\omega_0}}) = \mu(\psi_\xi) + \mu(\psi_{\xi^{\omega_0}}) = \mu(\psi_\xi) = \mu_{CZ}(\psi_\xi) = \mu_{CZ}(\gamma_c).$$

□

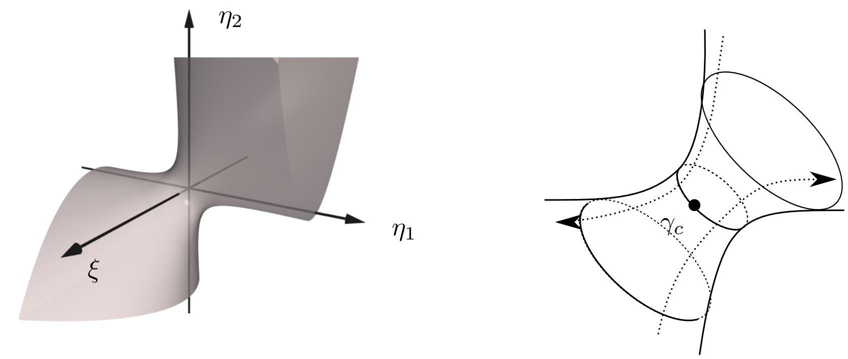

FIGURE 1. Orbits in the tube

CHAPTER 11

Finite energy planes

1. Holomorphic planes

We assume in this section that (N, λ) is a closed, oriented 3-dimensional positive contact manifold, i.e., the contact form $\lambda \in \Omega^1(N)$ satisfies

$$\lambda \wedge d\lambda > 0$$

, or in other words $\lambda \wedge d\lambda$ is a volume form on N inducing the given orientation. The hyperplane distribution

$$\xi = \ker \lambda \subset TN$$

is referred to as the *contact structure*. The *Reeb vector field* $R \in \Gamma(TN)$ is implicitly defined by the conditions

$$\lambda(R) = 1, \quad \iota_R d\lambda = 0.$$

The line bundle $\langle R \rangle$ over N spanned by the Reeb vector field together with ξ leads to a splitting

$$TN = \xi \oplus \langle R \rangle$$

of the tangent bundle of N. By abuse of notation we extend the contact form λ to a one form $\lambda \in \Omega^1(N \times \mathbb{R})$ which at a point $(p, r) \in N \times \mathbb{R}$ is given by

$$\lambda_{p,r} = e^r \lambda_p.$$

Its differential $\omega = d\lambda$ is a symplectic form for $N \times \mathbb{R}$. At a point (p, r) it is given by

$$\omega_{p,r} = e^r d\lambda_p + e^r dr \wedge \lambda_p.$$

The noncompact symplectic manifold $(N \times \mathbb{R}, \omega)$ is called the *symplectization* of the contact manifold (N, λ). The vector field ∂_r is a Liouville vector field on the symplectization, indeed, the Lie derivative of ω with respect to ∂_r is given by Cartan's formula by

$$\mathcal{L}_{\partial_r}\omega = d\iota_{\partial_r}\omega + \iota_{\partial_r} d\omega = d\lambda = \omega$$

where we used $\iota_{\partial_r}\omega = \lambda$ and $d\omega = 0$. By abuse of notation we extend the Reeb vector field to $\Gamma(N \times R)$ by

$$R(p, r) = R(p) \in T_pN \subset T_pN \times \mathbb{R} = T_{(p,r)}(N \times \mathbb{R}), \quad (p, r) \in N \times \mathbb{R}.$$

Similarly we extend the rank-2 bundle $\xi \subset TN$ to a rank-2 subbundle $\xi \subset T(N \times \mathbb{R})$ by

$$\xi_{(p,r)} = \xi_p \subset T_pN \subset T_pN \times \mathbb{R} \subset T_{(p,r)}(N \times \mathbb{R}).$$

We have the splitting

$$T(N \times \mathbb{R}) = \xi \oplus \langle \partial_r, R \rangle. \tag{140}$$

Note that this splitting is symplectic, i.e., both subbundles are symplectic subbundles of TN and they are symplectically orthogonal to each other. Moreover, the symplectic form on ξ is up to the conformal factor e^r just the restriction of $d\lambda$ to ξ. Choose $J \in \mathrm{End}(\xi)$ a $d\lambda$-compatible almost complex structure on ξ invariant under the natural $\mathbb{R}$-action

$$\mathbb{R} \times N \times \mathbb{R} \to N \times \mathbb{R}, \quad (s, p, r) \mapsto (p, s + r).$$

Again by abuse of notation we extend J to an ω-compatible almost complex structure on $T(N \times \mathbb{R})$ given for $v \in \xi$ and $a, b \in \mathbb{R}$ by

$$J(v + a\partial_r + bR) = Jv - b\partial_r + aR.$$

Note that the extension is still $\mathbb{R}$-invariant and respects the symplectic splitting (140). We refer to such an almost complex structure $J \in \mathrm{End}(T(N \times \mathbb{R}))$ as an *SFT-like almost complex structure*. Here SFT stands for Symplectic Field theory [**31**].

We now fix an SFT-like almost complex structure J on $T(N\times\mathbb{R})$. A *(parametrized) holomorphic plane*

$$\widetilde{u} \colon \mathbb{C} \to N \times \mathbb{R}$$

is a solution of the following nonlinear Cauchy-Riemann equation

$$\partial_x \widetilde{u} + J(\widetilde{u}) \partial_y \widetilde{u} = 0 \tag{141}$$

where $z = x + iy$. Recall that the *group of direct similitudes* is the semidirect product

$$\Sigma = \mathbb{C}^* \ltimes \mathbb{C}$$

with multiplication defined as

$$(\rho_1, \tau_1)(\rho_2, \tau_2) = (\rho_1\rho_2, \rho_1\tau_2 + \tau_1).$$

It acts on $\mathbb{C}$ by

$$(\rho, \tau)z = \rho z + \tau, \quad (\rho, \tau) \in \mathbb{C}^* \ltimes \mathbb{C}, \quad z \in \mathbb{C}.$$

Geometrically this amounts to a combination of a rotation, a translation, and a dilation of $\mathbb{C}$. Every biholomorphism of $\mathbb{C}$ to itself is of this form. The group of direct similitudes acts on solutions of (141) by reparametrization

$$(\rho, \tau)_* \widetilde{u}(z) = \widetilde{u}(\rho z + \tau).$$

Note that this is actually a right action. We refer to an equivalence class $[\widetilde{u}]$ of a holomorphic plane $\widetilde{u}$ under the action of the group of direct similitudes as an *unparametrized holomorphic plane*.

While the group of direct similitudes acts on the domain of a holomorphic plane there is an action of the group $\mathbb{R}$ on the target as well. For a holomorphic plane write

$$\widetilde{u} = (u, a), \quad u \colon \mathbb{C} \to N,\ a \colon \mathbb{C} \to \mathbb{R}.$$

Because the SFT-like almost complex structure J is $\mathbb{R}$-invariant it follows that for a solution $\widetilde{u}$ of (141) and $r \in \mathbb{R}$ the map

$$r_*(u, a) = (u, a + r) \colon \mathbb{C} \to N \times \mathbb{R}$$

is still a solution of (141). Note that the actions of the group Σ and $\mathbb{R}$ on solution of (141) commute so that we obtain an action of the group $\Sigma \times \mathbb{R}$ on solutions of (141). In particular, the group $\mathbb{R}$ still acts on unparametrized holomorphic planes.

2. The Hofer energy of a holomorphic plane

In order to describe the *energy* of a holomorphic curve we abbreviate

$$\Gamma := \{\phi \in C^\infty(\mathbb{R}, [0,1]) : \phi' \geq 0\}.$$

For $\phi \in \Gamma$ we define $\lambda^\phi \in \Omega^1(N \times \mathbb{R})$ by

$$\lambda^\phi_{(p,r)} = \phi(r)\lambda_p \quad (p,r) \in N \times \mathbb{R}.$$

The following notion of energy of a holomorphic plane is due to Hofer [**51**]

$$E(\widetilde{u}) := \sup_{\phi \in \Gamma} \int_{\mathbb{C}} \widetilde{u}^* d\lambda^\phi. \tag{142}$$

Note that the energy is invariant under the action of $\Sigma \times \mathbb{R}$ on solutions of (141). For the $\mathbb{R}$-action this follows from the fact that $\mathbb{R}$ acts on Γ as well by

$$r_*\phi(s) = \phi(s-r), \quad s \in \mathbb{R}$$

for $r \in \mathbb{R}$ and $\phi \in \Gamma$. The following lemma tells us that the energy of a holomorphic plane is never negative.

Lemma 2.1. *Assume that $\widetilde{u}$ is a solution of the nonlinear Cauchy-Riemann equation (141). Then its energy satisfies*

$$E(\widetilde{u}) \in [0, \infty].$$

Moreover, $E(\widetilde{u}) = 0$ if and only if $\widetilde{u}$ is constant.

Proof: Pick $\phi \in \Gamma$. At a point $(p,r) \in N \times \mathbb{R}$ the exterior derivative of λ^ϕ is given by

$$d\lambda^\phi_{(p,r)} = \phi(r) d\lambda_p + \phi'(r) dr \wedge \lambda_p.$$

Abbreviate by

$$\pi \colon TN \to \xi$$

the projection along $\langle R \rangle$. By definition of the Reeb vector field we obtain

$$\partial_x u = \pi \partial_x u + \lambda(\partial_x u) R$$

and therefore

$$\partial_x \widetilde{u} = \pi \partial_x u + \lambda(\partial_x u) R + \partial_x a \partial_r.$$

Using (141) we conclude

$$\partial_y \widetilde{u} = J \partial_x \widetilde{u} = J\pi \partial_x u + \partial_x a R - \lambda(\partial_x u)\partial_r.$$

Putting this together we end up with the formula

$$d\lambda^\phi(\partial_x \widetilde{u}, \partial_y \widetilde{u}) = \phi(a) d\lambda(\pi \partial_x u, J\pi\partial_x u) + \phi'(a)\big((\partial_x a)^2 + (\lambda(\partial_x u))^2\big). \tag{143}$$

Since the restriction of J to ξ is $d\lambda$-compatible it follows that $d\lambda(\pi\partial_x u, J\pi\partial_x u) \geq 0$. By definition $\phi(a) \geq 0$ and $\phi'(a) \geq 0$ so that it holds that

$$d\lambda^\phi(\partial_x \widetilde{u}, \partial_y \widetilde{u}) \geq 0.$$

We showed that

$$\int_{\mathbb{C}} \widetilde{u}^* d\lambda^\phi \geq 0, \quad \forall\, \phi \in \Gamma$$

and therefore

$$E(\widetilde{u}) \geq 0.$$

Now assume that $\widetilde{u}$ is not constant. That means that there exists $z \in \mathbb{C}$ such that

$$d\widetilde{u}(z) \neq 0.$$

Since $\widetilde{u}$ satisfies the Cauchy Riemann equation (141) it follows that

$$\partial_x \widetilde{u}(z) \neq 0.$$

It follows from (143) that we can choose $\phi \in \Gamma$ such that

$$d\lambda^\phi(\partial_x \widetilde{u}, \partial_y \widetilde{u})(z) > 0.$$

Hence the energy satisfies

$$E(\widetilde{u}) > 0.$$

This finishes the proof of the Lemma. □

The following Definition is due to Hofer [**51**].

Definition 2.2. *A holomorphic plane* $\widetilde{u} \colon \mathbb{C} \to N \times \mathbb{R}$ *is called a* finite energy plane *if*

$$0 < E(\widetilde{u}) < \infty.$$

The motivation of Hofer to study finite energy planes came from its close relation to the Reeb dynamics on the contact manifold (N, λ). In the following let $S^1 = \mathbb{R}/\mathbb{Z}$ be the circle and $\mathbb{R}_+ = \{r \in \mathbb{R} : r > 0\}$ the positive real numbers. Recall from Section 1 the following definition.

Definition 2.3. *A* (parametrized) periodic orbit of the Reeb vector field R *is a loop* $\gamma \in C^\infty(S^1, N)$ *for which there exists* $\tau \in \mathbb{R}_+$ *such that the tuple* (γ, τ) *is a solution of the problem*

$$\partial_t \gamma(t) = \tau R(\gamma(t)), \quad t \in S^1.$$

Because γ is parametrized the positive number $\tau = \tau(\gamma)$ is uniquely determined by γ and is referred to as the *period* of γ. The following Theorem is due to Hofer. To state it we introduce the following notation. If $u \colon \mathbb{C} \to N$ is a smooth maps and $s \in \mathbb{R}$ we abbreviate

$$u^s \colon S^1 \to N, \quad t \mapsto u(e^{2\pi(s+it)}).$$

Theorem 2.4. *Assume that* $\widetilde{u} = (u, a) \colon \mathbb{C} \to N \times \mathbb{R}$ *is a finite energy plane. Then there exists a periodic Reeb orbit* γ *and a sequence* $s_k \to \infty$ *such that the sequence* u^{s_k} *converges in the* C^∞*-topology to* γ.

For a proof of this fundamental theorem also referred to as the main result of finite energy planes we refer to [**51**, Theorem 31] or [**1**, Chapter 3]. The original interest in this result came from the fact that it enabled Hofer in [**51**] to deduce from it the *Weinstein conjecture* for a broad class of three dimensional contact manifolds. The Weinstein conjecture [**107**] asks if every closed contact manifolds admits a periodic Reeb orbit. The paper by Hofer [**51**] was one of the important breakthroughs concerning this conjecture. Later on Taubes [**105**] proved the Weinstein conjecture in dimension three completely using quite different methods than finite energy planes, namely Seiberg-Witten invariants. In higher dimensions the conjecture is still open in general, we refer to the paper by Albers and Hofer [**7**] and the literature cited therein for partial progress in higher dimensions. The question how far the Weinstein conjecture generalizes to noncompact manifolds is an active topic of research as well. The interested reader might consult the paper by Berg, Pasquotto, and Vandervorst [**15**] or the paper by Suhr and Zehmisch [**102**]. An intriguing point is that without the contact condition the conjecture might fail for a general Hamiltonian system, see the paper by Ginzburg and Gürel [**41**] and the literature cited therein.

3. The Omega-limit set of a finite energy plane

Theorem 2.4 does not claim that the asymptotic periodic Reeb orbit γ is unique. Although we are not aware of an explicit example it is conceivable that asymptotically a finite energy plane starts spiraling around a whole family of periodic Reeb orbits. Let us introduce the Omega-limit set $\Omega(u)$ of the finite energy plane $\widetilde{u}$ as follows. Namely $\Omega(u)$ consists of all periodic Reeb orbits γ for which there exists a sequence s_k going to infinity such that u^{s_k} converges to γ in the C^∞-topology. Note that

$$\Omega(u) \subset C^\infty(S^1, N)$$

and we topologize it as a subset of the free loop space of N. As the notation suggest $\Omega(u)$ only depends on the projection of the finite energy plane $\widetilde{u}$ to N. In particular, the Omega-limit set is invariant under the $\mathbb{R}$-action on finite energy planes. Hofer's theorem tells us that the Omega-limit set is never empty.

Lemma 3.1. *Assume that $\widetilde{u} = (u, a)$ is a finite energy plane. Then its Omega-limit set $\Omega(u)$ is compact and connected.*

Proof: To prove the lemma we use a statement stronger then the one provided by Theorem 2.4. Namely for a given sequence s_k going to infinity there exists a subsequence s_{k_j} and a periodic Reeb orbit γ such that $u^{s_{k_j}}$ converges to γ. However, this improved statement can be shown along the same lines as Theorem 2.4, see [**1**, Theorem 6.4.1]. Armed with this fact we are in position to prove the Lemma.

We first show that $\Omega(u)$ is compact. The free loop space $C^\infty(S^1, N)$ is metrizable. Therefore it suffices to show that $\Omega(u)$ is sequentially compact. Choose a metric d on $C^\infty(S^1, N)$ which induces the given topology of the free loop space. Let γ_ν for $\nu \in \mathbb{N}$ be a sequence in $\Omega(u)$. Since $\gamma_\nu \in \Omega(u)$ for every $\nu \in \mathbb{N}$ there exists a sequence $\{s_k^\nu\}_{k\in\mathbb{N}}$ going to infinity with the property that

$$\lim_{k\to\infty} u^{s_k^\nu} = \gamma_\nu.$$

Set $k_1 := 1$ and define inductively for $\nu \in \mathbb{N}$

$$k_{\nu+1} := \min\left\{k : s_k^{\nu+1} \geq s_{k_\nu}^\nu + 1, d\big(u^{s_k^{\nu+1}}, \gamma_{\nu+1}\big) \leq \frac{1}{\nu+1}\right\}.$$

For $\nu \in \mathbb{N}$ define

$$\sigma_\nu := s_{k_\nu}^\nu.$$

It follows by construction that the sequence σ_ν goes to infinity. By the improved version of Hofer's theorem discussed above there exists a subsequence ν_j and a periodic Reeb orbit γ such that

$$\lim_{\jmath\to\infty} u^{\sigma_{\nu_j}} = \gamma.$$

We claim that

$$\lim_{j\to\infty} \gamma_{\nu_j} = \gamma. \tag{144}$$

To see that pick $\epsilon > 0$. Choose $j_0 = j_0(\epsilon)$ with the property that

$$\nu_{j_0} \geq \frac{2}{\epsilon}, \quad d(u^{\sigma_{\nu_j}}, \gamma) \leq \frac{\epsilon}{2}, \ \forall \ j \geq j_0.$$

We estimate for every $j \geq j_0$

$$d\big(\gamma_{\nu_j}, \gamma\big) \leq d\big(u^{\sigma_{\nu_j}}, \gamma_{\nu_j}\big) + d\big(u^{\sigma_{\nu_j}}, \gamma\big) \leq \frac{1}{\nu_j} + \frac{\epsilon}{2} \leq \frac{1}{\nu_{j_0}} + \frac{\epsilon}{2} \leq \epsilon.$$

This proves (144) and hence $\Omega(u)$ is compact.

It remains to show that $\Omega(u)$ is connected. We assume by contradiction that $\Omega(u)$ is not connected and hence can be written as

$$\Omega(u) = \Omega_1(u) \cup \Omega_2(u)$$

where both $\Omega_1(u)$ and $\Omega_2(u)$ are nonempty, open and closed subsets of $\Omega(u)$ satisfying $\Omega_1(u) \cap \Omega_2(u) = \emptyset$. Since we already know that $\Omega(u)$ is compact the sets $\Omega_1(u)$ and $\Omega_2(u)$ are compact as well and therefore there exist open sets $V_1, V_2 \in C^\infty(S^1, N)$ with the property that

$$V_1 \cap V_2 = \emptyset, \quad \Omega_1(u) \subset V_1, \quad \Omega_2(u) \subset V_2.$$

Since $\Omega_1(u)$ and $\Omega_2(u)$ are nonempty there exist $\gamma_1 \in \Omega_1(u)$ and $\gamma_2 \in \Omega_2(u)$. By definition we can find sequences s_k^1 and s_k^2 going to infinity such that

$$\lim_{k\to\infty} u^{s_k^1} = \gamma_1, \quad \lim_{k\to\infty} u^{s_k^2} = \gamma_2.$$

Set $k_1 = 1$ and define inductively for $\nu \in \mathbb{N}$

$$k_{\nu+1} := \begin{cases} \min\left\{k : s_k^2 > s_{k_\nu}^1\right\} & \nu \text{ odd} \\ \min\left\{k : s_k^1 > s_{k_\nu}^2\right\} & \nu \text{ even.} \end{cases}$$

Note that the sequence k_ν goes to infinity. For any ν consider the path

$$[s_{k_\nu}, s_{k_{\nu+1}}] \to C^\infty(S^1, N), \quad s \mapsto u^s.$$

One of the endpoints of this path lies in $\Omega_1(u)$ while the other one lies in $\Omega_2(u)$. Therefore there exists $\sigma_\nu \in [s_{k_\nu}, s_{k_\nu+1}]$ with the property that

$$u^{\sigma_\nu} \in C^\infty(S^1, N) \setminus (V_1 \cap V_2).$$

Observe that the sequence σ_ν goes to infinity since s_{k_ν} goes to infinity. By the improved version of Hofer's theorem there exists a subsequence ν_j and a periodic Reeb orbit such that

$$\lim_{j\to\infty} u^{\sigma_{\nu_j}} = \gamma.$$

By definition of the Omega-limit set we have

$$\gamma \in \Omega(u).$$

On the other hand V_1 and V_2 were open subsets of the free loop space of N and therefore

$$\gamma \in C^\infty(S^1, N) \setminus (V_1 \cup V_2) \subset C^\infty(S^1, N) \setminus (\Omega_1(u) \cup \Omega_2(u)) = C^\infty(S^1, N) \setminus \Omega(u).$$

This contradiction shows that $\Omega(u)$ is connected and the lemma is proved. $\square$

There is a free action of the group S^1 on the set of parametrized periodic Reeb orbits by time-shift. Indeed, if $\gamma \in C^\infty(S^1, N)$ is a periodic Reeb orbit and $r \in \mathbb{R}/\mathbb{Z}$ the loop $r_*\gamma$ defined as

$$r_*\gamma(t) = \gamma(r + t), \quad t \in S^1$$

is again a periodic Reeb orbit. We refer to an orbit of this action as an unparametrized Reeb orbit, namely

Definition 3.2. *An* unparametrized Reeb orbit $[\gamma] = \{r_*\gamma : r \in S^1\}$ *is an equivalence class of a parametrized Reeb orbit* γ *under the equivalence relation given by time-shift.*

We say that a periodic Reeb orbit γ is *isolated* if $[\gamma]$ is isolated in the space of unparametrized loops $C^\infty(S^1, N)/S^1$. Note that a parametrized Reeb orbit can never be isolated in the free loop space $C^\infty(S^1, N)$ since it always comes in a circle family. If an isolated periodic Reeb orbit γ lies in the Omega-limit set of a finite energy plane $\widetilde{u} = (u, a)$ it follows from Lemma 3.1 that

$$\Omega(u) \subset [\gamma].$$

Therefore we abbreviate for an isolated periodic Reeb orbit γ

$$\widehat{\mathcal{M}}(\gamma) := \widehat{\mathcal{M}}([\gamma]) := \big\{\widetilde{u} = (u, a) \text{ finite energy plane},\ \Omega(u) \subset [\gamma]\big\} \tag{145}$$

the moduli space of finite energy planes asymptotic to the unparametrized periodic orbit $[\gamma]$. Recall that the group of direct similitudes $\Sigma = \mathbb{C}^* \ltimes \mathbb{C}$ acts on finite energy planes by reparametrization. If $(\rho, \tau) \in \Sigma$ with $\rho = |\rho|e^{2\pi i\theta} \in \mathbb{C}^*$ and $\widetilde{u} = (u, a)$ is a finite energy plane it follows that

$$\Omega((\rho, \tau)_* u) = \theta_* \Omega(u).$$

We conclude that the moduli space $\widehat{\mathcal{M}}(\gamma)$ is invariant under the action of Σ and we abbreviate by

$$\mathcal{M}(\gamma) := \widehat{\mathcal{M}}(\gamma)/\Sigma$$

the moduli space of unparametrized finite energy planes asymptotic to $[\gamma]$. Note that since the $\mathbb{R}$-action on finite energy planes given by $r_*(u, a) = (u, a + r)$ commutes with the Σ-action we still have a $\mathbb{R}$-action on the moduli space of unparametrized finite energy planes.

4. Non-degenerate finite energy planes

The situation becomes much nicer if we assume that the periodic Reeb orbit is *non-degenerate*. To explain this notion let us abbreviate by $\phi_R^t \colon N \to N$ for $t \in \mathbb{R}$ the flow of the Reeb vector field on N defined by

$$\phi_R^0 = \mathrm{id}, \quad \frac{d}{dt}\phi_R^t(x) = R(\phi_R^t(x)),\ x \in N,\ t \in \mathbb{R}.$$

Note that the contact form λ is invariant under the Reeb flow. Indeed, the Lie derivative of λ with respect to R computes by Cartan's formula to be

$$\mathcal{L}_R \lambda = \iota_R d\lambda + d\iota_R \lambda = 0$$

by the defining equation for the Reeb vector field. It follows that the differential of the Reeb flow

$$d\phi_R^t(x) \colon T_x N \to T_{\phi_R^t(x)} N$$

keeps the hyperplane distribution $\xi = \ker \lambda$ invariant so that we can define

$$d^\xi \phi_R^t(x) \colon \xi_x \to \xi_{\phi_R^t(x)}, \quad d^\xi \phi_R^t(x) := d\phi_R^t(x)|_{\xi_x}.$$

Again by the fact that λ and therefore $d\lambda$ as well are invariant under the Reeb flow we conclude that the map $d^\xi \phi_R^t(x)$ is a linear symplectic map from the symplectic vector space $(\xi_x, d\lambda)$ to the symplectic vector space $(\xi_{\phi_R^t(x)}, d\lambda)$. In particular, if γ is a periodic Reeb orbit of period τ we obtain a symplectic map

$$d^\xi \phi_R^\tau(\gamma(0)) \colon \xi_{\gamma(0)} \to \xi_{\gamma(0)}.$$

Definition 4.1. *A periodic Reeb orbit γ of period τ is called* non-degenerate *if*

$$\det(d^\xi \phi_R^\tau(\gamma(0)) - \mathrm{id}) \neq 0.$$

Note that if γ is non-degenerate and $[r] \in S^1 = \mathbb{R}/\mathbb{Z}$ the reparametrized periodic orbit $[r]_*\gamma$ is still non-degenerate. Indeed, since ϕ_R^t is a flow we have the relation

$$d^\xi \phi_R^\tau(\gamma(r)) = d^\xi \phi^{\tau r}(\gamma(0)) d^\xi \phi^\tau(\gamma(0)) d^\xi \phi^{\tau r}(\gamma(0))^{-1}.$$

Therefore it makes sense to talk about a non-degenerate unparametrized periodic orbit.

Definition 4.2. *A finite energy plane $\widetilde{u} = (u, a)$ is called* non-degenerate*, if there exists a non-degenerate periodic orbit γ such that $\gamma \in \Omega(u)$.*

A non-degenerate periodic orbit γ is isolated and therefore by Lemma 3.1 it holds that $\Omega(u) \subset [\gamma]$. However, more can be shown [**54**]

Lemma 4.3. *Assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane, then $\Omega(u) = \{\gamma\}$, i.e., the Omega-limit set of a non-degenerate finite energy plane consists of a unique parametrized non-degenerate periodic orbit.*

How a non-degenerate finite energy plane converges to its by the above Lemma unique asymptotic orbit has been described quite precisely by Hofer, Wysocki, and Zehnder in [**54**]. We discuss this in the next section.

5. The asymptotic formula

Assume that $\gamma \in C^\infty(S^1, N)$ is a periodic Reeb orbit and J is an SFT-like almost complex structure. Denote by $\Gamma^{1,2}(\gamma^*\xi)$ the Hilbert space of $W^{1,2}$-sections in ξ and by $\Gamma^{0,2}(\gamma^*\xi)$ the Hilbert space of L^2-sections in ξ. Consider the bounded linear operator

$$A_\gamma := A_{\gamma,J} \colon \Gamma^{1,2}(\gamma^*\xi) \to \Gamma^{0,2}(\gamma^*\xi)$$

which for $w \in \Gamma^{1,2}(\gamma^*\xi)$ is given by

$$A_\gamma(w)(t) = -J(\gamma(t)) d^\xi \phi_R^{t\tau}(\gamma(0)) \partial_t \Big(d^\xi \phi_R^{-t\tau}(\gamma(t)) w(t) \Big), \quad t \in [0,1].$$

Suppose that

$$\mathfrak{T} \colon \gamma^*\xi \to S^1 \times \mathbb{C}$$

is a unitary trivialization, i.e., an orthogonal trivialization, where orthogonality refers to the bundle metric $\omega(\cdot, J\cdot)$ on ξ and the standard inner product on $\mathbb{C}$, which interchanges multiplication by J on $\gamma^*\xi$ with multiplication by i on $\mathbb{C}$. The trivialization $\mathfrak{T}$ gives rise to a Hilbert space isomorphism

$$\Phi_{\mathfrak{T}} \colon \Gamma^{1,2}(\gamma^*\xi) \to W^{1,2}(S^1, \mathbb{C}), \quad w \mapsto \mathfrak{T} w$$

which extends to a Hilbert space isomorphism

$$\Phi_{\mathfrak{T}} \colon \Gamma^{0,2}(\gamma^*\xi) \to L^2(S^1, \mathbb{C})$$

by the same formula. Hence we obtain an operator

$$A_\gamma^{\mathfrak{T}} := \Phi_{\mathfrak{T}} A_\gamma \Phi_{\mathfrak{T}}^{-1} \colon W^{1,2}(S^1, \mathbb{C}) \to L^2(S^1, \mathbb{C}). \tag{146}$$

To describe $A_\gamma^{\mathfrak{T}}$ we write J_0 for the standard complex structure on $\mathbb{C}$ given by multiplication with i. For $t \in [0,1]$ we further abbreviate

$$\Psi(t) := \mathfrak{T}_{\gamma(t)} d^\xi \phi_R^{t\tau}(\gamma(0)) \mathfrak{T}_{\gamma(0)}^{-1} \in \mathrm{Sp}(1)$$

as well as

$$S(t) = -J_0 \partial_t \Psi(t) \Psi(t)^{-1} \in \mathrm{Sym}(2).$$

Assume $v \in W^{1,2}(S^1, \mathbb{C})$ and $t \in [0,1]$. We compute

$$\begin{aligned}
(A_\gamma^{\mathfrak{T}})v(t) &= -\mathfrak{T}_{\gamma(t)} J(\gamma(t)) d^\xi \phi_R^{t\tau}(\gamma(0)) \partial_t \Big(d^\xi \phi_R^{-t\tau}(\gamma(t)) \mathfrak{T}_{\gamma(t)}^{-1} v(t) \Big) \\
&= -J_0 \mathfrak{T}_{\gamma(t)} d^\xi \phi_R^{t\tau}(\gamma(0)) \partial_t \Big(d^\xi \phi_R^{-t\tau}(\gamma(t)) \mathfrak{T}_{\gamma(t)}^{-1} v(t) \Big) \\
&= -J_0 \mathfrak{T}_{\gamma(t)} d^\xi \phi_R^{t\tau}(\gamma(0)) \mathfrak{T}_{\gamma(0)}^{-1} \partial_t \Big(\mathfrak{T}_{\gamma(0)} d^\xi \phi_R^{-t\tau}(\gamma(t)) \mathfrak{T}_{\gamma(t)}^{-1} v(t) \Big) \\
&= -J_0 \Psi(t) \partial_t \Big(\Psi(t)^{-1} v(t) \Big) \\
&= -J_0 \partial_t v(t) + J_0 \partial_t \Psi(t) \Psi(t)^{-1} v(t) \\
&= -J_0 \partial_t v(t) - S(t) v(t).
\end{aligned}$$

That means that

$$A_\gamma^{\mathfrak{T}} = A_S$$

where A_S is the operator defined in (109). Since the operator A_γ is conjugated to the operator $A_\gamma^{\mathfrak{T}}$ it has the same spectral properties as A_S, in particular, its spectrum is discrete and consists of real eigenvalues of finite multiplicity.

The following notion is due to Siefring [**100**].

Definition 5.1. *Assume that $\widetilde{u}$ is a non-degenerate finite energy plane with asymptotic orbit γ of period τ and $U\colon [R,\infty) \times S^1 \to \gamma^*\xi$ is a smooth map such that $U(s,t) \in \xi_{\gamma(t)}$ for all $(s,t) \in [R,\infty) \times S^1$. The map U is called an* asymptotic representative*, if there exists a proper embedding $\phi\colon [R,\infty) \times S^1 \to \mathbb{R} \times S^1$ asymptotic to the identity such that*

$$\widetilde{u}(e^{\phi(s,t)}) = (\exp_{\gamma(t)} U(s,t), \tau s)$$

where exp *is the exponential map of the restriction of the metric $\omega(\cdot, J\cdot)$ to $N = N \times \{0\} \subset N \times \mathbb{R}$.*

The following result is due to Mora [**85**] based on previous work by Hofer, Wysocki, and Zehnder [**54**].

Theorem 5.2. *Assume that $\widetilde{u}$ is a non-degenerate finite energy plane with asymptotic orbit γ. Then $\widetilde{u}$ admits an asymptotic representative. Moreover, there exist a negative eigenvalue η of A_γ and an eigenvector ζ of A_γ to the eigenvalue η such that the asymptotic representative can be written as*

$$U(s,t) = e^{\eta s}\big(\zeta(t) + \kappa(s,t)\big) \tag{147}$$

where κ decays exponentially with all derivatives in the sense that there exist for one and hence every metric constants $M_{i,j}$ for $0 \leq i, j < \infty$ and $d > 0$ such that

$$|\nabla_s^i \nabla_t^j \kappa(s,t)| \leq M_{i,j} e^{-ds}.$$

In view of the requirement that the coordinate change of an asymptotic representative is asymptotic to the identity an asymptotic representative is unique up to restriction of the domain of definition. In particular, the eigenvalue η and the eigenvector ζ are uniquely determined by the finite energy plane $\widetilde{u}$. We denote the eigenvalue by

$$\eta_{\widetilde{u}} \in \mathfrak{S}(A_\gamma) \cap (-\infty, 0)$$

and refer to it as the *asymptotic eigenvalue* and similarly we denote the eigenvector by

$$\zeta_{\widetilde{u}} \in \Gamma(\gamma^*\xi)$$

and refer to it as the *asymptotic eigenvector*.

Recall that if $\widetilde{u} = (u, a)$ is a finite energy plane then the group $\mathbb{R}$ acts on it by $r_*\widetilde{u} = (u, a + r)$. For later reference it will be useful to now how the asymptotic eigenvalue and the asymptotic eigenvector transform under this action. To compute this let U be an asymptotic representative of $\widetilde{u}$ for a proper embedding $\phi\colon [R, \infty) \times S^1 \to \mathbb{R} \times S^1$ asymptotic to the identity. As usual τ stands for the period of the asymptotic Reeb orbit. For $r \in \mathbb{R}$ define

$$U_r\colon [R + \tfrac{r}{\tau}, \infty) \times S^1 \to \gamma^*\xi, \quad (s, t) \mapsto U(s - \tfrac{r}{\tau}, t)$$

and

$$\phi_r\colon [R + \tfrac{r}{\tau}, \infty) \times S^1 \to \mathbb{R} \times S^1, \quad (s, t) \mapsto \phi(s - \tfrac{r}{\tau}, t).$$

We compute

$$\begin{aligned} r_*\widetilde{u}(e^{\phi_r(s,t)}) &= r_*\widetilde{u}(e^{\phi(s-\frac{r}{\tau},t)}) \\ &= (\exp_{\gamma(t)} U(s - \tfrac{r}{\tau}, t), \tau(s - \tfrac{r}{\tau}) + r) \\ &= (\exp_{\gamma(t)} U_r(s, t), \tau s). \end{aligned}$$

Using (147) we compute for U_r

$$\begin{aligned} U_r(s, t) &= U(s - \tfrac{r}{\tau}, t) \\ &= e^{\eta(s-\frac{r}{\tau})}\Big(\zeta(t) + \kappa(s - \tfrac{r}{\tau}, t)\Big) \\ &= e^{\eta s}\Big(e^{-\frac{\eta r}{\tau}}\zeta(t) + e^{-\frac{\eta r}{\tau}}\kappa(s - \tfrac{r}{\tau}, t)\Big). \end{aligned}$$

That means that the asymptotic eigenvalue is unchanged under the $\mathbb{R}$-action while the asymptotic eigenvector gets scaled by a factor $e^{-\frac{\eta r}{\tau}}$. We summarize this computation in the following lemma.

Lemma 5.3. *Assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane with asymptotic orbit γ of period τ. The asymptotic eigenvalue only depends on the projection u, i.e.*

$$\eta_u := \eta_{\widetilde{u}}$$

while the asymptotic eigenvector transforms under the $\mathbb{R}$-action on $\widetilde{u}$ as

$$\zeta_{r_*\widetilde{u}} = e^{-\frac{\eta_u r}{\tau}}\zeta_{\widetilde{u}}.$$

An important Corollary of Theorem 5.2 is the following result.

Corollary 5.4. *If γ is a non-degenerate Reeb orbit, then the action of $\mathbb{R}$ on the moduli space $\widehat{\mathcal{M}}(\gamma)$ of finite energy planes asymptotic to γ is a free action.*

Proof: Pick a finite energy plane $\widetilde{u} = (u, a) \in \widehat{\mathcal{M}}(\gamma)$. In view of the fact that $\widetilde{u}$ admits an asymptotic representative, it follows that the infimum of the function $a\colon \mathbb{C} \to \mathbb{R}$ is attained and we set

$$\underline{a} := \min\{a(z) : z \in \mathbb{C}\} \in \mathbb{R}.$$

Now suppose that $r \in \mathbb{R}$ satisfies $r_* \widetilde{u} = \widetilde{u}$. Since $r_* \widetilde{u} = (u, a + r)$ we obtain

$$\underline{a} + r = \underline{a}$$

implying that $r = 0$. This proves that the $\mathbb{R}$-action on $\widehat{\mathcal{M}}(\gamma)$ is free. $\square$

CHAPTER 12

The index inequality and fast finite energy planes

We first define the Conley-Zehnder index of a non-degenerate finite energy plane $\widetilde{u} = (u, a)$ with asymptotic orbit γ of period τ. Consider the symplectic bundle $u^*\xi \to \mathbb{C}$. Since $\mathbb{C}$ is contractible there exists a symplectic trivialization

$$\mathfrak{T}\colon u^*\xi \to \mathbb{C} \times \mathbb{C}.$$

In view of the asymptotic behavior of $\widetilde{u}$ explained in Theorem 5.2 we can arrange the trivialization such that it extends asymptotically to a symplectic trivialization

$$\mathfrak{T}\colon \gamma^*\xi \to S^1 \times \mathbb{C}.$$

Recall that $d^\xi \phi_R^t(\gamma(0))\colon \xi_{\gamma(0)} \to \xi_{\gamma(t)}$ denotes the restriction of the differential of the Reeb flow to the hyperplane distribution which turns out to be a linear symplectic map. Hence we obtain a smooth path $\Psi\colon [0,1] \to \mathrm{Sp}(1)$ of symplectic maps from $\mathbb{C}$ to itself by

$$\Psi(t) = \mathfrak{T}_{\gamma(t\tau)} d^\xi \phi_R^{t\tau}(\gamma(0)) \mathfrak{T}_{\gamma(0)}^{-1}.$$

Note that $\Psi(0) = \mathrm{id}$. Moreover, due to the assumption that the asymptotic Reeb orbit γ is non-degenerate, the path of symplectic maps Ψ is non-degenerate in the sense that

$$\ker(\Psi(1) - \mathrm{id}) = \{0\}.$$

We define the Conley-Zehnder index of $\widetilde{u}$ as

$$\mu_{CZ}(\widetilde{u}) = \mu_{CZ}(\Psi). \tag{148}$$

Note that since every two symplectic trivializations over the contractible base $\mathbb{C}$ are homotopic it follows that the Conley-Zehnder index does not depend on the choice of the trivialization $\mathfrak{T}$. Moreover, it depends only on the projection u so that we can write as well

$$\mu_{CZ}(u) := \mu_{CZ}(\widetilde{u}).$$

The following index inequality is due to Hofer, Wysocki, and Zehnder [**53**]

Theorem 0.5. *Assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane. Then its Conley-Zehnder index satisfies the inequality*

$$\mu_{CZ}(u) \geq 2.$$

Before starting with the proof of this theorem let us explain how this theorem can be interpreted as an *automatic transversality result.* Denote by γ the asymptotic Reeb orbit of $\widetilde{u}$. Then we can think of $\widetilde{u}$ as an element of the moduli space $\widehat{\mathcal{M}}(\gamma)$ as explained in (145). The unparametrized finite energy plane $[\widetilde{u}]$ is then an element of the moduli space $\mathcal{M}(\gamma) = \widehat{\mathcal{M}}(\gamma)/\Sigma$ where $\Sigma = \mathbb{C}^* \ltimes \mathbb{C}$ is the group of direct similitudes which acts on $\widehat{\mathcal{M}}(\gamma)$ by reparametrizations. We will see later in (210) that the virtual dimension of the moduli space $\mathcal{M}(\gamma)$ at $[\widetilde{u}]$ is given by

$$\mathrm{virdim}_{[\widetilde{u}]}\mathcal{M}(\gamma) = \mu_{CZ}(u) - 1.$$

Here the virtual dimension of the moduli space $\mathcal{M}(\gamma)$ is given by the Fredholm index of a Fredholm operator L which linearizes the holomorphic curve equation in normal direction of $\widetilde{u}$. If the Fredholm operator L is surjective locally around $[\widetilde{u}]$ the moduli space $\mathcal{M}(\gamma)$ is a manifold and the tangent space of $\mathcal{M}(\gamma)$ at $[\widetilde{u}]$ equals

$$T_{[\widetilde{u}]}\mathcal{M}(\gamma) = \mathrm{ker}L.$$

Hence if L is surjective we have

$$\mathrm{virdim}_{[\widetilde{u}]}\mathcal{M}(\gamma) = \mathrm{ind}L = \dim\mathrm{ker}L = \mathrm{dim}T_{[\widetilde{u}]}\mathcal{M}(\gamma).$$

Recall that the group $\mathbb{R}$ acts on finite energy planes by $r_*(u,a) = (u, a+r)$. Since this action commutes with the reparametrization action of the group Σ the group $\mathbb{R}$ still acts on the moduli space $\mathcal{M}(\gamma)$. Moreover, by Corollary 5.4 this action is free. Therefore still assuming that L is surjective we get the inequality

$$\mathrm{dim}T_{[\widetilde{u}]}\mathcal{M}(\gamma) \geq 1.$$

Combining these facts we end up with the inequality $\mu_{CZ}(u) \geq 2$ claimed in Theorem 0.5. However, we point out that this reasoning only works under the assumption that L is surjective. Geometrically the linearization operator L can be thought of as follows. One interprets the moduli space $\mathcal{M}(\gamma)$ as the zero set of a section

$$s\colon \mathcal{B} \to \mathcal{E}, \quad \mathcal{M}(\gamma) = s^{-1}(0).$$

of a space $\mathcal{B}$ into a bundle $\mathcal{E}$ over $\mathcal{B}$. The Fredholm operator L then arises as the vertical differential of the section s at $[\widetilde{u}]$ and the question if L is surjective can be rephrased geometrically as the question if the section s is transverse to the zero section at $[\widetilde{u}]$. That explains why one refers to Theorem 0.5 as an automatic transversality result.

We mention that many transversality results for moduli spaces are so called generic transversality results which hold for a generic choice of datas. In our set-up the data is the SFT-like almost complex structure J. It is therefore highly remarkable that Theorem 0.5 holds for any SFT-like almost complex structure J and not just for a generic choice of it.

We now start with the preparations for the proof of Theorem 0.5. We assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane. We choose a trivialization

$$\mathfrak{T}\colon u^*\xi \to \mathbb{C} \times \mathbb{C}.$$

We further denote by

$$\pi\colon TN = \xi \oplus \langle R \rangle \to \xi$$

the projection along the Reeb vector field R. The major ingredient in the proof of Theorem 0.5 is the map

$$\mathfrak{T}\pi\partial_x u\colon \mathbb{C} \to \mathbb{C}. \tag{149}$$

Here we denote by $z = x + iy$ the coordinates on $\mathbb{C}$. This map is smooth interpreted as a real map from $\mathbb{C} = \mathbb{R}^2$ to itself. Moreover, due to the fact that u is holomorphic the map above is "almost holomorphic" in a sense to be described more precisely below.

We first recall some facts about winding for a general smooth map $f\colon \mathbb{C} \to \mathbb{C}$. We denote the regular set of f by

$$\mathcal{R}_f := \{z \in \mathbb{C} : f(z) \neq 0\}.$$

For a continuous loop $\gamma\colon S^1 \to \mathcal{R}_f$ the map $t \mapsto \frac{f(\gamma(t))}{||f(\gamma(t))||}$ is a continuous map from the circle S^1 to itself. Hence we can consider its degree

$$w_\gamma(f) := \deg\Big(t \mapsto \tfrac{f(\gamma(t))}{||f(\gamma(t))||}\Big) \in \mathbb{Z}. \tag{150}$$

We refer to $w_\gamma(f)$ as the *winding number* of f along the loop γ. It has the following properties.

Homotopy invariance: If $\gamma\colon S^1 \times [0,1] \to \mathcal{R}_f$ is a continuous map, then $\gamma_0 = \gamma(\cdot,0)$ and $\gamma_1 = \gamma(\cdot,1)$ are two homotopic loops in $\mathcal{R}_f$ and its winding numbers are unchanged

$$w_{\gamma_0}(f) = w_{\gamma_1}(f).$$

Concatenation: Suppose that $\gamma_1\colon S^1 \to \mathcal{R}_f$ and $\gamma_2\colon S^1 \to \mathcal{R}_f$ are two continuous maps satisfying $\gamma_1(0) = \gamma_2(0)$. Denote by $\gamma_1 \# \gamma_2$ its concatenation. The winding number is additive under concatenation

$$w_{\gamma_1\#\gamma_2}(f) = w_{\gamma_1}(f) + w_{\gamma_2}(f).$$

These two properties have the following consequences. Assume that the singular set

$$\mathcal{S}_f := \{z \in \mathbb{C} : f(z) = 0\} = \mathbb{C} \setminus \mathcal{R}_f$$

is discrete. Pick $z_0 \in \mathcal{S}_f$. Because $\mathcal{S}_f$ is discrete, there exists $\epsilon > 0$ such that the intersection of $\mathcal{S}_f$ with $D_\epsilon(z_0) = \{z \in \mathbb{C} : ||z - z_0|| \leq \epsilon\}$, the closed ϵ-ball around z_0, satisfies

$$\mathcal{S}_f \cap D_\epsilon(z_0) = \{z_0\}.$$

Consider

$$\gamma^\epsilon_{z_0}\colon S^1 \to \mathcal{R}_f, \quad t \mapsto z_0 + \epsilon e^{2\pi i t}.$$

Define the *local winding number* of f at the singularity z_0 by

$$w_{z_0}(f) := w_{\gamma^\epsilon_{z_0}}(f).$$

Because of homotopy invariance of the winding number the local winding number is well defined, independent of the choice of ϵ.

Suppose that $R > 0$ and that $f(Re^{2\pi i t}) \neq 0$ for every $t \in S^1$, i.e., we get a loop in the regular set of f

$$\gamma^R\colon S^1 \to \mathcal{R}_f, \quad t \mapsto Re^{2\pi i t}.$$

We set

$$w_R(f) := w_{\gamma^R}(f). \tag{151}$$

In view of the homotopy invariance and the concatenation property of the winding number we can express this winding number as the sum of local winding numbers

$$w_R(f) = \sum_{z \in D_R(0) \cap \mathcal{S}_f} w_z(f). \tag{152}$$

Suppose now that $f\colon \mathbb{C} \to \mathbb{C}$ is *holomorphic* and does not vanish identically. We claim that if $f(z_0) = 0$ the local winding number satisfies

$$w_{z_0}(f) \geq 1. \tag{153}$$

To prove this inequality we can assume without loss of generality that $z_0 = 0$. Since f is holomorphic it is given by its Taylor series

$$f(z) = \sum_{n=1}^{\infty} a_n z^n.$$

Abbreviate

$$\ell := \min\{n : a_n \neq 0\}$$

the order of vanishing of f at zero. Since f does not vanish identically ℓ is finite. We can write f now as

$$f(z) = \sum_{n=\ell}^{\infty} a_n z_n.$$

Choose $\epsilon > 0$ such that

$$\sum_{n=\ell+1}^{\infty} |a_n| \epsilon^{n-\ell} < |a_\ell|.$$

Abbreviate

$$g(z) := \sum_{n=\ell+1}^{\infty} a_n z^n$$

so that we obtain

$$f(z) = a_\ell z^\ell + g(z).$$

Consider the map from S^1 to S^1

$$t \mapsto \frac{a_\ell \epsilon^\ell e^{2\pi i t\ell} + g(\epsilon e^{2\pi i t\ell})}{||a_\ell \epsilon^\ell e^{2\pi i t\ell} + g(\epsilon e^{2\pi i t\ell})||}.$$

This map is homotopic to the map from S^1 to S^1 given by

$$t \mapsto \frac{a_\ell \epsilon^\ell e^{2\pi i t\ell}}{||a_\ell \epsilon^\ell e^{2\pi i t\ell}||} = e^{2\pi i t\ell}.$$

This implies that

$$w_0(f) = \ell \geq 1.$$

In other words the local winding of a holomorphic function at a zero is given by the order of vanishing of the function at this point. In combination with (152) we have established the following lemma.

Lemma 0.6. *Assume that $f\colon \mathbb{C} \to \mathbb{C}$ is holomorphic and $R > 0$ such that $f(Re^{2\pi i t}) \neq 0$ for every $t \in S^1$. Then $w_R(f) \geq 0$ and $w_R(f) = 0$ if and only if $f(z) \neq 0$ for every $z \in D_R(0)$.*

We cannot apply Lemma 0.6 to the map $\mathfrak{T}\pi\partial_x u\colon \mathbb{C} \to \mathbb{C}$ from (149) because this map is not actually holomorphic despite the fact that $\widetilde{u}$ was a holomorphic plane. However, $\mathfrak{T}\pi\partial_x u$ is close enough to be holomorphic that the reasoning which established Lemma 0.6 still applies. This is made precise by Carleman's similarity principle. We first examine how far the map $\mathfrak{T}\pi\partial_x u$ deviates from being holomorphic. We first recall *Darboux's Theorem* for contact manifolds which roughly tells us that a contact manifold has no local invariants, see [**40**]

Theorem 0.7 (Darboux). *Assume that (N, λ) is a three dimensional contact manifold and $p \in N$. Then there exists an open neighborhood U of p and a diffeomorphism*

$$\Phi \colon U \to V \subset \mathbb{R}^3$$

such that $\Phi(p) = 0 \in V$ and

$$\Phi^*(dq_1 + q_2 dq_3) = \lambda.$$

In particular, the contact form λ looks locally like $dq_1 + q_2 dq_3$ the standard contact form on $\mathbb{R}^3$. Since the question if the local winding numbers of the map $\mathfrak{T}\pi\partial_x u$ are positive is a local problem we can assume in view of Darboux's theorem without loss of generality that

$$\lambda = dq_1 + q_2 dq_3, \quad d\lambda = dq_2 \wedge dq_3.$$

In these coordinates the Reeb vector field is given by

$$R = \partial_{q_1}.$$

Moreover, a basis of the hyperplane distribution $\xi = \ker\lambda$ is given by the vectors

$$e_1 = \partial_{q_2}, \quad e_2 = -q_2\partial_{q_1} + \partial_{q_3}.$$

Note that

$$d\lambda(e_1, e_2) = 1$$

so that the basis $\{e_1, e_2\}$ is a symplectic basis of ξ. If we write

$$u(x,y) = \big(q_1(x,y), q_2(x,y), q_3(x,y)\big)$$

we have

$$\partial_x u(x,y) = \big(\partial_x q_1(x,y), \partial_x q_2(x,y), \partial_x q_3(x,y)\big)$$

and therefore

$$\begin{aligned} \pi_{u(x,y)}\partial_x u(x,y) &= \big(-q_2(x,y)\partial_x q_3(x,y), \partial_x q_2(x,y), \partial_x q_3(x,y)\big) \\ &= \partial_x q_2(x,y) e_1\big(u(x,y)\big) + \partial_x q_3(x,y) e_2\big(u(x,y)\big). \end{aligned}$$

Because all trivializations on a ball are homotopic we can in view of the homotopy invariance of the winding number choose an arbitrary local trivialization to compute the local winding number. We choose as local trivialization the symplectic trivialization

$$\mathfrak{T}_u \colon \xi_u \to \mathbb{C}, \quad (xe_1 + ye_2) \mapsto x + iy.$$

In this trivialization we obtain

$$\mathfrak{T}\pi\partial_x u = \partial_x q_2 + i\partial_x q_3.$$

Because u is holomorphic the projection to the hyperplane distribution satisfies the equation

$$\pi\partial_x u + J(u)\pi\partial_y u = 0. \tag{154}$$

As we computed above $\pi_u\partial_x u$ we get

$$\pi_u\partial_y u = \partial_y q_2 e_1(u) + \partial_y q_3 e_2(u).$$

Abbreviate by $J(x,y)$ the 2×2-matrix representing $J(u(x,y))$ in the basis $\{e_1(u(x,y)), e_2(u(x,y))\}$. In particular, it holds that

$$J(x,y)^2 = -\mathrm{id}.$$

With this notation (154) translates into

$$\begin{pmatrix} \partial_x q_2 \\ \partial_x q_3 \end{pmatrix} + J(x,y) \begin{pmatrix} \partial_y q_2 \\ \partial_y q_3 \end{pmatrix} = 0.$$

This implies

$$\begin{aligned} \partial_x(\mathfrak{I}\pi\partial_x u) &= \begin{pmatrix} \partial_x^2 q_2 \\ \partial_x^2 q_3 \end{pmatrix} \\ &= -\partial_x J \begin{pmatrix} \partial_y q_2 \\ \partial_y q_3 \end{pmatrix} - J \begin{pmatrix} \partial_x\partial_y q_2 \\ \partial_x\partial_y q_3 \end{pmatrix} \\ &= -(\partial_x J) J \begin{pmatrix} \partial_x q_2 \\ \partial_x q_3 \end{pmatrix} - J \begin{pmatrix} \partial_y\partial_x q_2 \\ \partial_y\partial_x q_3 \end{pmatrix} \end{aligned}$$

Abbreviating

$$A = -(\partial_x J)J$$

we can write this as

$$\partial_x(\mathfrak{I}\pi\partial_x u) = -A(\mathfrak{I}\pi\partial_x u) - J\partial_y(\mathfrak{I}\pi\partial_x u)$$

or equivalently

$$\partial_x(\mathfrak{I}\pi\partial_x u) + J\partial_y(\mathfrak{I}\pi\partial_x u) + A(\mathfrak{I}\pi\partial_x u) = 0. \tag{155}$$

We recall Carleman's similarity principle from [**34**].

Lemma 0.8 (Carleman's similarity principle). *Assume that $f\colon B_\epsilon = \{z \in \mathbb{R}^2 : |z| < \epsilon\} \to \mathbb{R}^2$ is a smooth map and $J, A \in C^\infty(B_\epsilon, M_2(\mathbb{R}))$ are smooth families of 2×2-matrices such that $J(z)^2 = -\mathrm{id}$ for every $z \in B_\epsilon$, i.e., $J(z)$ is a complex structure. Suppose that*

$$\partial_x f + J\partial_y f + Af = 0, \quad f(0) = 0.$$

Then there exists $\delta \in (0,\epsilon)$, $\Phi \in C^0(B_\delta, \mathrm{GL}(\mathbb{R}^2))$, and $\sigma\colon B_\delta \to \mathbb{C}$ holomorphic such that

$$f(z) = \Phi(z)\sigma(z), \quad \sigma(0) = 0, \quad J(z)\Phi(z) = \Phi(z)i.$$

In view of (155) and Carleman's similarity principle all local winding numbers of $\mathfrak{I}\pi\partial_x u$ are positive. Therefore the same reasoning which led to Lemma 0.6 establishes the following proposition.

Proposition 0.9. *Assume that $R > 0$ such that $\pi\partial_x u(Re^{2\pi it}) \neq 0$ for every $t \in S^1$. Then $w_R(\mathfrak{I}\pi\partial_x u) \geq 0$ and $w_R(\mathfrak{I}\pi\partial_x u) = 0$ if and only if $\pi\partial_x u(z) \neq 0$ for every $z \in D_R(0)$.*

In view of the asymptotic behavior of a non-degenerate finite energy plane explained in Theorem 5.2 there exists $R_0 > 0$ such that for all $R \geq R_0$ it holds that $\pi\partial_x u(Re^{2\pi it}) \neq 0$ for every $t \in S^1$. Moreover, again in view of the asymptotic behavior it holds that

$$w_R(\mathfrak{I}\pi\partial_r u) = w(\eta_u) \tag{156}$$

where η_u is the asymptotic eigenvalue of the non-degenerate finite energy plane $\widetilde{u}$. Because η_u is negative this fact has interesting consequences as we will see soon. However, let us first discuss how $w_R(\mathfrak{I}\pi\partial_x u)$ and $w_R(\mathfrak{I}\pi\partial_r u)$ are related. To see that note that

$$\mathfrak{I}\pi\partial_x u = (\mathfrak{I}\pi du)\partial_x, \quad \mathfrak{I}\pi\partial_r u = (\mathfrak{I}\pi du)\partial_r$$

and for each $z \in \mathbb{C}$ the map $\mathfrak{T}\pi du(z)$ is a linear map from $\mathbb{C} = \mathbb{R}^2$ to itself. In general, if $A \in C^\infty(S^1, GL(\mathbb{R}^2))$ and $v_1, v_2 \in C^\infty(S^1, \mathbb{R}^2 \setminus \{0\})$ the formula

$$\deg\left(t \mapsto \frac{A(t)v_2(t)}{|A(t)v_2(t)|}\right) = \deg\left(t \mapsto \frac{A(t)v_1(t)}{|A(t)v_1(t)|}\right) + \mathrm{wind}(v_1, v_2) \tag{157}$$

holds true, where $\mathrm{wind}(v_1, v_2)$ is the winding of v_2 around v_1. Because

$$\mathrm{wind}(\partial_x, \partial_r) = 1$$

we obtain the relation

$$w_R(\mathfrak{T}\pi\partial_r u) = w_R(\mathfrak{T}\pi\partial_x u) + \mathrm{wind}(\partial_x, \partial_r) = w_R(\mathfrak{T}\pi\partial_x u) + 1. \tag{158}$$

Combining Proposition 0.9 with (156) and (158) we obtain

Theorem 0.10. *Assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane. Then the winding number of its asymptotic eigenvalue meets the inequality*

$$w(\eta_u) \geq 1.$$

Moreover, $w(\eta_u) = 1$ if and only if $\pi\partial_x u(z) \neq 0$ for every $z \in \mathbb{C}$.

Theorem 0.5 is a straightforward consequence of Theorem 0.10.

Proof of Theorem 0.5: Assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane. It follows from Theorem 0.10 that the winding number of its asymptotic eigenvalue satisfies $w(\eta_u) \geq 1$. Because η_u is negative it follows from the definition of α from (118) that

$$\alpha \geq 1.$$

Because the parity (119) satisfies $p \in \{0, 1\}$ we obtain from Theorem 0.33 that

$$\mu_{CZ}(u) = 2\alpha + p \geq 2\alpha \geq 2.$$

This finishes the proof of Theorem 0.5. □

Assume that $\widetilde{u} = (u, a)$ is a finite energy plane and $\pi\partial_x u(z) \neq 0$ for every $z \in \mathbb{C}$. Because $\widetilde{u}$ is holomorphic it holds that $\pi\partial_x u + J(u)\pi\partial_y u = 0$, hence because $\pi\partial_x u(z) \neq 0$ we have that $\{\pi\partial_x u(z), \pi\partial_y u(z)\}$ are linearly independent vectors in $\xi_{u(z)}$. This implies that

$$\pi du(z) \colon T_z\mathbb{C} = \mathbb{C} \to \xi_{u(z)}$$

is bijective. In particular, $du(z) \colon T_z\mathbb{C} \to T_{u(z)}N$ is injective and therefore u is an immersion. Moreover, since $TN = \xi \oplus \langle R \rangle$ the Reeb vector field is transverse to the image of u. On the other hand, if u is an immersion transverse to the Reeb vector field we must have $\pi\partial_x u(z) \neq 0$ for every $z \in \mathbb{C}$. Therefore we obtain from Theorem 0.10 the following Corollary.

Corollary 0.11. *Assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane. Then the winding number of its asymptotic eigenvalues satisfies $w(\eta_u) = 1$ if and only if u is an immersion transverse to the Reeb vector field.*

The following definition is due to Hryniewicz [**58**]

Definition 0.12. *A non-degenerate finite energy plane $\widetilde{u} = (u, a)$ is called* fast *if and only if u is an immersion transverse to the Reeb vector field.*

The reason for this terminology is that in view of Theorem 0.5 the winding number of the asymptotic eigenvalue satisfies $w(\eta_u) \geq 1$ and as a consequence of the monotonicity of the winding number from Theorem 0.32 a fast finite energy plane has a fast asymptotic decay. In view of this notion we can rephrase Corollary 0.11 as

Corollary 0.13. *A non-degenerate finite energy plane $\widetilde{u} = (u, a)$ is fast if and only if $w(\eta_u) = 1$.*

If $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane, then in view of the definition of α in (118) and Theorem 0.33 the winding number of $\widetilde{u}$ satisfies the inequality

$$w(\eta_u) \leq \alpha = \left\lfloor \frac{\mu_{CZ}(u)}{2} \right\rfloor$$

where for a real number r we abbreviate by

$$\lfloor r \rfloor := \max\{n \in \mathbb{N} : n \leq r\}$$

the integer part of r. Combining this inequality with Theorem 0.10 we obtain further the following Corollary.

Corollary 0.14. *Assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane. Then the winding number of its asymptotic eigenvalue satisfies*

$$1 \leq w(\eta_u) \leq \left\lfloor \frac{\mu_{CZ}(u)}{2} \right\rfloor.$$

This has the further consequence.

Corollary 0.15. *Assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane such that $\mu_{CZ}(u) \in \{2, 3\}$. Then $\widetilde{u}$ is fast.*

CHAPTER 13

Siefring's intersection theory for fast finite energy planes

1. Positivity of intersection for closed curves

Suppose that (M, J) is a $2n$-dimensional almost complex manifold, i.e., $J \in \mathrm{End}(TM)$ is an almost complex structure which means that $J^2 = -\mathrm{id}$. We define an orientation on M as follows. If $x \in M$ we declare a basis of T_xM the form $\{v_1, Jv_1, v_2, Jv_2, \ldots, v_n, Jv_n\}$ to be positive. Equivalently, that means that the basis $\{v_1, v_2, \ldots, v_n, Jv_1, Jv_2, \ldots Jv_n\}$ is positive if $\frac{n(n+1)}{2}$ is even and negative otherwise. It remains to explain that this notion is well-defined, i.e., independent of the choice of $\{v_1, \ldots, v_n\}$. To see that suppose that we have another basis $\{v_1', Jv_1', v_2', Jv_2', \ldots, v_n', Jv_n'\}$ of this form. Let $B \in GL(\mathbb{C}^n)$ be the basis change matrix from the complex basis $\{v_1, \ldots, v_n\}$ to the complex basis $\{v_1', \ldots, v_n'\}$. Write

$$B = A_1 + iA_2$$

where A_1 and A_2 are real $n \times n$-matrices. It follows that the real $2n \times 2n$-matrix

$$A = \begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix}$$

is the real basis change matrix from the basis $\{v_1, \ldots, v_n, Jv_1, \ldots, Jv_n\}$ to the basis $\{v_1', \ldots, v_n', Jv_1', \ldots, Jv_n'\}$. The determinant of A satisfies

$$\begin{aligned}
\det(A) &= \det \begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix} \\
&= \det\left(\begin{pmatrix} \mathrm{id} & \mathrm{id} \\ i \cdot \mathrm{id} & -i \cdot \mathrm{id} \end{pmatrix}^{-1} \begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix} \begin{pmatrix} \mathrm{id} & \mathrm{id} \\ i \cdot \mathrm{id} & -i \cdot \mathrm{id} \end{pmatrix} \right) \\
&= \det\left(\frac{1}{2} \begin{pmatrix} \mathrm{id} & -i \cdot \mathrm{id} \\ \mathrm{id} & i \cdot \mathrm{id} \end{pmatrix}^{-1} \begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix} \begin{pmatrix} \mathrm{id} & \mathrm{id} \\ i \cdot \mathrm{id} & -i \cdot \mathrm{id} \end{pmatrix} \right) \\
&= \det \begin{pmatrix} A_1 - iA_2 & 0 \\ 0 & A_1 + iA_2 \end{pmatrix} \\
&= \det(\overline{B}) \cdot \det(B) \\
&= |\det(B)|^2 \\
&> 0.
\end{aligned}$$

This proves that the basis $\{v_1, \ldots, v_n, Jv_1, \ldots, Jv_n\}$ has the same sign as the basis $\{v_1', \ldots, v_n', Jv_1', \ldots, Jv_n'\}$. Consequently, the two bases $\{v_1, Jv_1, \ldots, v_n, Jv_n\}$ and $\{v_1', Jv_1', \ldots, v_n', Jv_n'\}$ have the same sign as well. This shows that the orientation of the almost complex manifold (M, J) is well defined. In the following we always endow an almost complex manifold with this orientation.

Now assume that $M = M^4$ is a four dimensional almost complex manifold and (Σ_1, i) and (Σ_2, i) are two closed Riemann surfaces. Suppose that

$$u_1 \colon \Sigma_1 \to M^4, \quad u_2 \colon \Sigma_2 \to M^4$$

are two holomorphic maps, i.e.,

$$du_k \circ i = J du_k, \quad k \in \{1, 2\}.$$

We further assume that

$$u_1 \pitchfork u_2,$$

i.e., the two curves intersect transversally, meaning that if $(z_1, z_2) \in \Sigma_1 \times \Sigma_2$ is such that $u_1(z_1) = u_2(z_2)$ it holds that

$$\mathrm{im} du_1(z_1) + \mathrm{im} du_2(z_2) = T_{u_1(z_1)} M = T_{u_2(z_2)} M.$$

Because M is four dimensional and Σ_1 and Σ_2 are two dimensional this is equivalent to requiring

$$\mathrm{im} du_1(z_1) \oplus \mathrm{im} du_2(z_2) = T_{u_1(z_1)} M.$$

We now compute the intersection index at the intersection point $(z_1, z_2) \in \Sigma_1 \times \Sigma_2$. Note that since Σ_1, Σ_2, and M are all almost complex manifolds they are canonically oriented. Choose a positive basis $\{v_1, iv_1\}$ of $T_{z_1}\Sigma_1$ and a positive basis $\{v_2, iv_2\}$ of $T_{z_2}\Sigma_2$. Using that u_1 and u_2 are holomorphic we get that

$$\begin{aligned} &\{du_1(z_1)v_1, du_1(z_1)(iv_1), du_2(z_2)v_2, du_2(z_2)(iv_2)\} = \\ &\qquad \{du_1(z_1)v_1, J du_1(z_1)v_1, du_2(z_2)v_2, J du_2(z_2)v_2\} \end{aligned}$$

is a positive basis of $T_{u_1(z_1)}M$. Therefore the intersection index equals one for every intersection point. In particular, the intersection number of u_1 and u_2 is nonnegative and it vanishes if and only if there are no intersection points. This phenomenon is referred to as *positivity of intersection*. It is a nontrivial fact due to McDuff and Micallef-White that positivity of intersection continuous to hold for perturbations of non-transverse intersection points. This is explained in [**81**, Appendix E]. That means even without the assumption that u_1 and u_2 intersect transversally their algebraic intersection number is still nonnegative and it vanishes if and only if there are no intersection points.

2. The algebraic intersection number for finite energy planes

In the following we consider a closed 3-dimensional contact manifold (N, λ) and fix an SFT-like almost complex structure J on $N \times \mathbb{R}$. We assume that γ is a non-degenerate Reeb orbit and $\widetilde{u}$ and $\widetilde{v}$ are two finite energy planes whose common asymptotic orbit is γ and $\mathrm{im}(\widetilde{u}) \neq \mathrm{im}(\widetilde{v})$. In this section we explain how to associate to $\widetilde{u}$ and $\widetilde{v}$ an algebraic intersection number

$$\mathrm{int}(\widetilde{u}, \widetilde{v}) \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}.$$

The reason that $\mathrm{int}(\widetilde{u}, \widetilde{v})$ is nonnegative is because of positivity of intersection for the two holomorphic curves. Even if $\widetilde{u}$ and $\widetilde{v}$ intersect transversally it is a priori not obvious that the numbers of intersection points is finite, since the domain $\mathbb{C}$ of $\widetilde{u}$ and $\widetilde{v}$ is not compact. The crucial ingredient which guarantees finiteness of the algebraic intersection number is the following theorem due to Siefring [**100**].

Theorem 2.1. *Suppose that γ is a non-degenerate periodic Reeb orbit which is the common asymptotic of two finite energy planes $\widetilde{u}$ and $\widetilde{v}$. Assume that $U, V \colon [R,\infty) \times S^1 \to \gamma^*\xi$ are two asymptotic representatives of $\widetilde{u}$ and $\widetilde{v}$. If $U \neq V$ then there exists $\eta \in \mathfrak{S}(A_\gamma) \cap (-\infty, 0)$ and ζ an eigenvector of A_γ for the eigenvalue η satisfying*

$$V(s,t) - U(s,t) = e^{\eta s}\big(\zeta(t) + \kappa(s,t)\big)$$

and there exist constants $M_{i,j}, d > 0$ for $0 \leq i, j < \infty$ such that

$$|\nabla_s^i \nabla_t^j \kappa(s,t)| \leq M_{i,j} e^{-ds}.$$

Corollary 2.2. *Assume that $\widetilde{u}$ and $\widetilde{v}$ are two finite energy planes in $N \times \mathbb{R}$ asymptotic to the same non-degenerate periodic Reeb orbit γ. Suppose that $\widetilde{u} \pitchfork \widetilde{v}$. Then the number of intersection points between $\widetilde{u}$ and $\widetilde{v}$ is finite, i.e.,*

$$\#\big\{(z_1, z_2) \in \mathbb{C} \times \mathbb{C} : \widetilde{u}(z_1) = \widetilde{v}(z_2)\big\} < \infty.$$

Proof: We prove the Corollary in 3 steps.

Step 1: *There exists a compact subset $K_0 \subset \mathbb{C}$ with the property that*

$$\big\{(z_1, z_2) \in K_0^c \times K_0^c : \widetilde{u}(z_1) = \widetilde{v}(z_2)\big\} = \emptyset$$

where $K^c = \mathbb{C} \setminus K$ denotes the complement of K in $\mathbb{C}$.

The proof of Step 1 is a bit involved since we have to deal with the possibility that the asymptotic periodic orbit γ is multiply covered. We define the covering number of γ as

$$\mathrm{cov}(\gamma) := \max\Big\{k \in \mathbb{N} : \gamma\big(t + \tfrac{1}{k}\big) = \gamma(t),\ t \in S^1\Big\}.$$

In view of uniqueness of a first order ODE we conclude that

$$\gamma(t) \neq \gamma(t'), \quad t - t' \notin \frac{1}{\mathrm{cov}(\gamma)}\mathbb{Z}. \tag{159}$$

By definition of an asymptotic representative there exists a compact subset $K_0 \subset \mathbb{C}$ such that there exists a bijection between the following two sets

$$\Big\{(z_1, z_2) \in K_0^c \times K_0^c : \widetilde{u}(z_1) = \widetilde{v}(z_2)\Big\} \cong$$
$$\Big\{\big((s_1,t_1),(s_2,t_2)\big) \in \big([R,\infty) \times S^1\big) \times \big([R,\infty) \times S^1\big) :$$
$$\big(\gamma(t_1), U(s_1,t_1), s_1\tau\big) = \big(\gamma(t_2), V(s_2,t_2), s_2\tau\big)\Big\}$$

where $\tau > 0$ is the period of the periodic orbit γ. Assume that $(s_1, t_1) \in [R,\infty) \times S^1$ and $(s_2, t_2) \in [R,\infty) \times S^1$ such that

$$\big(\gamma(t_1), U(s_1,t_1), s_1\tau\big) = \big(\gamma(t_2), V(s_2,t_2), s_2\tau\big).$$

Because $\tau \neq 0$ we conclude that

$$s_1 = s_2$$

and using (159) we infer that

$$t_2 = t_1 + \frac{j}{\mathrm{cov}(\gamma)}, \quad 0 \leq j \leq \mathrm{cov}(\gamma) - 1.$$

Therefore

$$U\Big(s_1, t_1\Big) - V\Big(s_1, t_1 + \tfrac{j}{\mathrm{cov}(\gamma)}\Big) = 0. \tag{160}$$

For $j \in \{0, \dots, \mathrm{cov}(\gamma) - 1\}$ define

$$\widetilde{v}_j \colon \mathbb{C} \to N \times \mathbb{R}, \quad z \mapsto \widetilde{v}(e^{2\pi i \frac{j}{\mathrm{cov}(\gamma)}} z).$$

Note that $\widetilde{v}_j$ is a finite energy plane and in view of the definition of $\mathrm{cov}(\gamma)$ the asymptotic orbit of $\widetilde{v}_j$ is γ as well. An asymptotic representative for $\widetilde{v}_j$ is the map

$$V_j \colon [R, \infty) \times S^1 \to \gamma^*\xi, \quad (s,t) \mapsto V\Big(s, t + \tfrac{j}{\mathrm{cov}(\gamma)}\Big).$$

Equation (160) can be reinterpreted as

$$U(s_1, t_1) - V_j(s_1, t_1) = 0.$$

Because $\widetilde{u} \pitchfork \widetilde{v}$ we conclude that

$$U \neq V_j.$$

Hence by Theorem 2.1 there exists $\eta \in \mathfrak{S}(A_\gamma) \cap (-\infty, 0)$ and ζ an eigenvector of A_γ to the eigenvalue η such that

$$U(s,t) - V_j(s,t) = e^{\eta s}\big(\zeta(t) + \kappa(s,t)\big)$$

where κ decays exponentially with all its derivatives. Because ζ is an eigenvector of A_γ and therefore a solution of a first order ODE it follows that

$$\zeta(t) \neq 0, \quad \forall\, t \in S^1.$$

Because κ decays exponentially we can assume perhaps after enlarging K_0 and R that

$$\sup_{(s,t) \in [R,\infty) \times S^1} |\kappa(s,t)| < \min\big\{|\xi(t)| : t \in S^1\big\}.$$

Hence we can assume that

$$U(s,t) - V_j(s,t) \neq 0, \quad \forall\, (s,t) \in [R, \infty) \times S^1,\ \forall\, 0 \leq j < \mathrm{cov}(\gamma).$$

Therefore

$$\widetilde{u}(z_1) \neq \widetilde{v}(z_2), \quad \forall\, (z_1, z_2) \in K_0^c \times K_0^c.$$

This finishes the proof of Step 1.

Step 2: *There exists a compact subset $K \subset \mathbb{C}$ with the property that*

$$\big\{(z_1, z_2) \in \mathbb{C} \times \mathbb{C} : \widetilde{u}(z_1) = \widetilde{v}(z_2)\big\} = \big\{(z_1, z_2) \in K \times K : \widetilde{u}(z_1) = \widetilde{v}(z_2)\big\}.$$

If $\widetilde{u} = (u, a)$ and $\widetilde{v} = (v, b)$ with $u, v \colon \mathbb{C} \to N$ and $a, b \colon \mathbb{C} \to \mathbb{R}$ and K_0 is as in Step 1 we abbreviate

$$c := \max\big\{a(z), b(z) : z \in K_0\big\}.$$

In view of the asymptotic behavior there exists $K \supset K_0$ compact such that

$$a(z) > c,\ b(z) > c \quad \forall\, z \in K^c.$$

We decompose

$$\begin{aligned}
\big\{(z_1,z_2)\in\mathbb{C}\times\mathbb{C}:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\} \quad &= \quad \big\{(z_1,z_2)\in K^c\times K^c:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\\
&\quad\ \cup\big\{(z_1,z_2)\in K^c\times K:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\\
&\quad\ \cup\big\{(z_1,z_2)\in K\times K^c:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\\
&\quad\ \cup\big\{(z_1,z_2)\in K\times K:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\\
&=: \quad A_{11}\cup A_{12}\cup A_{21}\cup A_{22}.
\end{aligned}$$

Since $K_0\subset K$ we have $K^c\subset K_0^c$ and therefore by Step 1

$$A_{11}\subset\big\{(z_1,z_2)\in K_0^c\times K_0^c:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}=\emptyset.$$

The next two sets in the decomposition we decompose further, namely

$$\begin{aligned}
A_{12} \quad &= \quad \big\{(z_1,z_2)\in K^c\times K_0:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\\
&\quad\ \cup\big\{(z_1,z_2)\in K^c\times K\setminus K_0:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\\
&\subset \quad \big\{(z_1,z_2)\in K^c\times K_0:a(z_1)=b(z_2)\big\}\\
&\quad\ \cup\big\{(z_1,z_2)\in K_0^c\times K_0^c:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\\
&= \quad \emptyset
\end{aligned}$$

and similarly we obtain

$$A_{21}=\emptyset.$$

This finishes the proof of Step 2.

Step 3: *We prove the Corollary.*

Since $\widetilde{u}\pitchfork\widetilde{v}$ the number of intersection points of $\widetilde{u}|_K$ and $\widetilde{v}|_K$ is finite for every compact set K. The Corollary now follows immediately from Step 2. □

In view of Corollary 2.2 if γ is a non-degenerate periodic Reeb orbit and $\widetilde{u}$ and $\widetilde{v}$ are two finite energy plane with common asymptotic γ which intersect transversally we define the algebraic intersection number of $\widetilde{u}$ and $\widetilde{v}$ as

$$\mathrm{int}(\widetilde{u},\widetilde{v}):=\#\big\{(z_1,z_2)\in\mathbb{C}\times\mathbb{C}:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\in\mathbb{N}_0. \tag{161}$$

If $\widetilde{u}$ and $\widetilde{v}$ do not intersect transversally however satisfy

$$\mathrm{im}(\widetilde{u})\neq\mathrm{im}(\widetilde{v})$$

we can still define the algebraic intersection number between $\widetilde{u}$ and $\widetilde{v}$ as follows. Because $\widetilde{u}$ and $\widetilde{v}$ have different image the arguments in the proof of Corollary 2.2 together with the fact that the intersection points of two holomorphic curves with different image are isolated (see [**81**, Appendix E]) imply that there exists a compact subset $K\subset\mathbb{C}$ such that if $\widetilde{u}(z_1)=\widetilde{v}(z_2)$ we necessarily have $(z_1,z_2)\in K\times K$. Now perturb $\widetilde{u}$ to $\widetilde{u}'$ and $\widetilde{v}$ to $\widetilde{v}'$ such that there exists a compact subset $K'\subset\mathbb{C}$ such that

$$\widetilde{u}'|_{(K')^c}=\widetilde{u}|_{(K')^c},\quad \widetilde{v}'|_{(K')^c}=\widetilde{v}|_{(K')^c},\quad \widetilde{u}'\pitchfork\widetilde{v}'.$$

We do not require that positivity of intersection continuous to hold for $\widetilde{u}'$ and $\widetilde{v}'$. We define

$$\nu\colon\big\{(z_1,z_2)\in\mathbb{C}\times\mathbb{C}:\widetilde{u}'(z_1)=\widetilde{v}'(z_2)\big\}\to\{-1,1\}$$

by

$$\nu(z_1, z_2) = \begin{cases} 1 & \{\partial_x \widetilde{u}'(z_1), \partial_y \widetilde{u}'(z_1), \partial_x \widetilde{v}'(z_2), \partial_y \widetilde{v}'(z_2)\} \\ & \text{positive basis of } T_{\widetilde{u}'(z_1)}(N \times \mathbb{R}) = T_{\widetilde{v}'(z_2)}(N \times \mathbb{R}) \\ -1 & \text{else.} \end{cases}$$

The algebraic intersection number of $\widetilde{u}$ and $\widetilde{v}$ is defined as

$$\operatorname{int}(\widetilde{u}, \widetilde{v}) = \sum_{\substack{(z_1, z_2) \in \mathbb{C} \times \mathbb{C} \\ \widetilde{u}'(z_1) = \widetilde{v}'(z_2)}} \nu(z_1, z_2) \in \mathbb{Z}. \tag{162}$$

By homotopy invariance the algebraic intersection number is independent of the choice of the perturbations $\widetilde{u}'$ and $\widetilde{v}'$, see [**83**]. Moreover, if $\widetilde{u}$ and $\widetilde{v}$ intersect transversally, then we do not need to perturb the two maps and since the two maps are holomorphic positivity of intersection implies that (162) coincides with (161) in this case. In particular, the algebraic intersection number is nonnegative if $\widetilde{u} \pitchfork \widetilde{v}$. It is a nontrivial fact that this continues to hold if $\widetilde{u}$ and $\widetilde{v}$ do not intersect transversally but have just different images. In fact we can choose the perturbations $\widetilde{u}'$ and $\widetilde{v}'$ such that $\nu(z_1, z_2) = 1$ for every intersection point $(z_1, z_2) \in \mathbb{C} \times \mathbb{C}$ of $\widetilde{u}'$ and $\widetilde{v}'$. This is due to results of McDuff and Micallef-White, see [**81**, Appendix E]. In particular, we have the following theorem.

Theorem 2.3. *Assume that $\widetilde{u}$ and $\widetilde{v}$ are two finite energy planes with common non-degenerate periodic Reeb orbit γ such that* $\operatorname{im}(\widetilde{u}) \neq \operatorname{im}(\widetilde{v})$. *Then the algebraic intersection number of $\widetilde{u}$ and $\widetilde{v}$ satisfies*

$$\operatorname{int}(\widetilde{u}, \widetilde{v}) \in \mathbb{N}_0.$$

Moreover, $\operatorname{int}(\widetilde{u}, \widetilde{v}) = 0$ *if and only if $\widetilde{u}$ and $\widetilde{v}$ do not intersect.*

3. Siefring's intersection number

Assume that $\widetilde{u} = (u, a)$ is a non-degenerate finite energy plane with asymptotic orbit γ. Choose a trivialization $\mathfrak{T} \colon u^*\xi \to \mathbb{C} \times \mathbb{C}$ which extends to a trivialization $\mathfrak{T} \colon \gamma^*\xi \to S^1 \times \mathbb{C}$. Such a trivialization gives rise to a nonvanishing section $X_{\mathfrak{T}} \colon \mathbb{C} \to u^*\xi$ defined by

$$X_{\mathfrak{T}}(z) = \mathfrak{T}_{u(z)}^{-1} 1 \in \xi_{u(z)}.$$

Since ξ is transverse to the Reeb vector field there exists $\epsilon_0 = \epsilon_0(\gamma, \mathfrak{T}) > 0$ such that for every $0 < \epsilon \leq \epsilon_0$ it holds that

$$\operatorname{im}(\exp_\gamma \epsilon X_{\mathfrak{T}}) \cap \operatorname{im}\gamma = \emptyset$$

where exp is the exponential map with respect to the restriction of the metric $\omega(\cdot, J\cdot)$ to $N \times \{0\} \subset N \times \mathbb{R}$. If $\widetilde{u} = (u, a)$ is a finite energy plane with asymptotic orbit γ we set for $\epsilon \in (0, \epsilon_0]$

$$u_{\mathfrak{T},\epsilon} = \exp_u \epsilon X_{\mathfrak{T}}, \quad \widetilde{u}_{\mathfrak{T},\epsilon} = (u_{\mathfrak{T},\epsilon}, a).$$

Now assume that $\widetilde{u}$ and $\widetilde{v}$ are fast finite energy planes with asymptotic limit the same non-degenerate Reeb orbit γ. We define *Siefring's intersection number* for $\widetilde{u}$ and $\widetilde{v}$ as

$$\operatorname{sief}(\widetilde{u}, \widetilde{v}) = \operatorname{int}(\widetilde{u}_{\mathfrak{T},\epsilon}, \widetilde{v}) + 1 \in \mathbb{Z}.$$

This definition is due to Siefring [**101**] based on previous work by Hutchings [**62**] and [**72**]. Here int denotes the usual algebraic intersection number obtained by the signed count of intersection points of $\widetilde{u}_{\mathfrak{T},\epsilon}$ and $\widetilde{v}$, maybe after a small generic perturbation which makes the two curves intersect transversally. Note that since the curves $\widetilde{u}_{\mathfrak{T},\epsilon}$ and $\widetilde{v}$ have disjoint asymptotics after a small generic perturbation the number of intersection points is necessarily finite. By homotopy invariance of the algebraic intersection number one observes that $\mathrm{sief}(\widetilde{u},\widetilde{v})$ is independent of the trivialization $\mathfrak{T}$ and the choice of ϵ. Since the two curves $\widetilde{u}_{\mathfrak{T},\epsilon}$ and $\widetilde{v}$ have different asymptotics Siefring's intersection number is a homotopy invariant which is not clear for the algebraic intersection number $\mathrm{int}(\widetilde{u},\widetilde{v})$. As for the algebraic one Siefring's intersection number is symmetric, i.e.,

$$\mathrm{sief}(\widetilde{u},\widetilde{v}) = \mathrm{sief}(\widetilde{v},\widetilde{u}).$$

4. Siefring's inequality

The following inequality was discovered by Siefring in [**101**].

Theorem 4.1. *Assume that $\widetilde{u}$ and $\widetilde{v}$ are fast finite energy planes asymptotic to the same non-degenerate periodic Reeb orbit γ such that $\mathrm{im}(\widetilde{u}) \neq \mathrm{im}(\widetilde{v})$. Then*

$$0 \leq \mathrm{int}(\widetilde{u},\widetilde{v}) \leq \mathrm{sief}(\widetilde{u},\widetilde{v}).$$

Before we embark on the proof of Theorem 4.1 let us make the following remarks. Since $\widetilde{u}_{\mathfrak{T},\epsilon}$ is not holomorphic anymore it is far from obvious that Siefring's intersection number turns out to be nonnegative. Later on we are mostly interested in the case where $\mathrm{sief}(\widetilde{u},\widetilde{v})$ is zero. Then it follows from Siefring's inequality that the algebraic intersection number $\mathrm{int}(\widetilde{u},\widetilde{v})$ is zero as well which implies by positivity of intersection that $\widetilde{u}$ and $\widetilde{v}$ do not intersect. Interestingly, Siefring's intersection number is a homotopy invariant. Therefore if it vanishes fast finite energy planes homotopic to $\widetilde{u}$ and $\widetilde{v}$ still do not intersect unless their images coincide.

In [**101**] Siefring defined as well an intersection number for finite energy planes which are not necessarily fast such that the assertion of Theorem 4.1 continues to hold. However, in this case one has to add to the algebraic intersection number of $\widetilde{u}_{\mathfrak{T},\epsilon}$ and $\widetilde{v}$ a number bigger than 1. In the proof it will become clear that the reason why one has to add 1 to fast finite energy planes is that 1 coincides with the winding number of the asymptotic eigenvalue of fast finite energy planes.

We first prove a proposition which does not yet require that $\widetilde{u}$ and $\widetilde{v}$ are fast. Recall that if $\widetilde{u}$ and $\widetilde{v}$ have the same non-degenerate asymptotic Reeb orbit γ but different images, then there exists by Theorem 2.1 an eigenvalue of the asymptotic operator A_γ such that the difference of asymptotic representatives decays exponentially with weight given by this eigenvalue. Because the asymptotic representatives are unique up to restriction of their domain of definition this eigenvalue depends only on $\widetilde{u}$ and $\widetilde{v}$ and we abbreviate it by

$$\eta_{\widetilde{u},\widetilde{v}} = \eta_{\widetilde{v},\widetilde{u}} \in \mathfrak{S}(A_\gamma).$$

We can now formulate the proposition as follows.

Proposition 4.2. *Assume that γ is a non-degenerate periodic Reeb orbit and $\widetilde{u}$ and $\widetilde{v}$ are two finite energy planes with common asymptotic orbit γ such that $\mathrm{im}(\widetilde{u}) \neq \mathrm{im}(\widetilde{v})$. Then*

$$\mathrm{int}(\widetilde{u}_{\mathfrak{T},\epsilon},\widetilde{v}) = \mathrm{int}(\widetilde{u},\widetilde{v}) - w(\eta_{\widetilde{u},\widetilde{v}}).$$

Proof: In Step 2 of Corollary 2.2 we proved that there exists a compact subset $K \subset \mathbb{C}$ with the property that

$$\big\{(z_1,z_2) \in \mathbb{C}\times\mathbb{C} : \widetilde{u}(z_1) = \widetilde{v}(z_2)\big\} = \big\{(z_1,z_2) \in K \times K : \widetilde{u}(z_1) = \widetilde{v}(z_2)\big\}. \tag{163}$$

If $\widetilde{u}=(u,a)$ and $\widetilde{v}=(v,b)$ with $u,v\colon \mathbb{C}\to N$ and $a,b\colon\mathbb{C}\to\mathbb{R}$ we set

$$c:=\max\big\{a(z),b(z):z\in K\big\}.$$

In view of the asymptotic behavior of non-degenerate finite energy planes there exists a compact set $K_0\supset K$ such that

$$a(z),b(z)>c,\quad \forall\ z\in K_0^c$$

where K_0^c denotes as usual the complement of K_0 in $\mathbb{C}$. Now choose a third compact subset $K_1\supset K_0$ with the property that there exists a smooth cutoff function $\beta\in C^\infty(\mathbb{C},[0,1])$ such that

$$\beta|_{K_0}=0,\quad \beta|_{K_1^c}=1.$$

Abbreviate

$$u_{\mathfrak{T},\epsilon,\beta}=\exp_u\beta\epsilon X_{\mathfrak{T}},\quad \widetilde{u}_{\mathfrak{T},\epsilon,\beta}=(u_{\mathfrak{T},\epsilon,\beta},a).$$

Note that $\widetilde{u}_{\mathfrak{T},\epsilon,\beta}$ coincides on the complement of the compact subset K_1 with $\widetilde{u}_{\mathfrak{T},\epsilon}$ and therefore

$$\mathrm{int}(\widetilde{u}_{\mathfrak{T},\epsilon},\widetilde{v})=\mathrm{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon},\widetilde{v}). \tag{164}$$

We write the last intersection number as a sum of four intersection numbers

$$\begin{aligned}\mathrm{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon},\widetilde{v}) \quad = \quad & \mathrm{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_{K^c},\widetilde{v}|_{K^c})+\mathrm{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_{K^c},\widetilde{v}|_{K})\\ &+\mathrm{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_{K},\widetilde{v}|_{K^c})+\mathrm{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_{K},\widetilde{v}|_{K}).\end{aligned} \tag{165}$$

In order to compute these four terms we decompose as in the proof of Corollary 2.2 the set of intersection points of $\widetilde{u}_{\mathfrak{T},\beta,\epsilon}$ and $\widetilde{v}$ into four disjoint subsets

$$\big\{(z_1,z_2)\in\mathbb{C}\times\mathbb{C}:\widetilde{u}_{\mathfrak{T},\beta,\epsilon}(z_1)=\widetilde{v}(z_2)\big\}=A_{11}\cup A_{12}\cup A_{21}\cup A_{22}$$

where

$$\begin{aligned}A_{11}&=\big\{(z_1,z_2)\in K^c\times K^c:\widetilde{u}_{\mathfrak{T},\beta,\epsilon}(z_1)=\widetilde{v}(z_2)\big\}\\ A_{12}&=\big\{(z_1,z_2)\in K^c\times K:\widetilde{u}_{\mathfrak{T},\beta,\epsilon}(z_1)=\widetilde{v}(z_2)\big\}\\ A_{21}&=\big\{(z_1,z_2)\in K\times K^c:\widetilde{u}_{\mathfrak{T},\beta,\epsilon}(z_1)=\widetilde{v}(z_2)\big\}\\ A_{22}&=\big\{(z_1,z_2)\in K\times K:\widetilde{u}_{\mathfrak{T},\beta,\epsilon}(z_1)=\widetilde{v}(z_2)\big\}.\end{aligned}$$

The set A_{12} we decompose further

$$\begin{aligned}A_{12}\quad&=\quad\big\{(z_1,z_2)\in K_0^c\times K:\widetilde{u}_{\mathfrak{T},\beta,\epsilon}(z_1)=\widetilde{v}(z_2)\big\}\\ &\qquad\cup\big\{(z_1,z_2)\in K_0\setminus K\times K:\widetilde{u}_{\mathfrak{T},\beta,\epsilon}(z_1)=\widetilde{v}(z_2)\big\}\\ &=\quad\big\{(z_1,z_2)\in K_0^c\times K:u_{\mathfrak{T},\beta,\epsilon}(z_1)=v(z_2),\ a(z_1)=b(z_2)\big\}\\ &\qquad\cup\big\{(z_1,z_2)\in K_0\setminus K\times K:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\\ &\subset\quad\big\{(z_1,z_2)\in K_0^c\times K:a(z_1)=b(z_2)\big\}\\ &\qquad\cup\big\{(z_1,z_2)\in K^c\times K:\widetilde{u}(z_1)=\widetilde{v}(z_2)\big\}\\ &=\quad\emptyset.\end{aligned}$$

For the last equality we observed that the second set is empty in view of (163) and the first set is empty since on the complement of K_0 the function a takes values bigger than c where one K the function b takes values less than or equal to c by definition of the constant c and the choice of the set K_0. We conclude that

$$\mathrm{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_{K^c},\widetilde{v}|_{K})=0. \tag{166}$$

Since $K \subset K_0$ and β vanishes on K_0 we obtain for the set A_{21}

$$A_{21} = \big\{(z_1, z_2) \in K^c \times K : \widetilde{u}(z_1) = \widetilde{v}(z_2)\big\} = \emptyset$$

where we used again (163). This implies

$$\operatorname{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_K, \widetilde{v}|_{K^c}) = 0. \tag{167}$$

We next examine the set A_{22}. Again taking advantage that β vanishes on K and using (163) we obtain

$$A_{22} = \big\{(z_1, z_2) \in K \times K : \widetilde{u}(z_1) = \widetilde{v}(z_2)\big\} = \big\{(z_1, z_2) \in \mathbb{C} \times \mathbb{C} : \widetilde{u}(z_1) = \widetilde{v}(z_2)\big\}$$

and therefore

$$\operatorname{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_K, \widetilde{v}|_K) = \operatorname{int}(\widetilde{u}, \widetilde{v}). \tag{168}$$

It remains to compute $\operatorname{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_{K^c}, \widetilde{v}|_{K^c})$. Let $U \colon [R, \infty) \times S^1 \to \gamma^*\xi$ be an asymptotic representative for $\widetilde{u}$ and $V \colon [R, \infty) \times S^1 \to \gamma^*\xi$ be an asymptotic representative for $\widetilde{v}$. Because $\widetilde{u}$ and $\widetilde{v}$ have different images their intersection points are isolated, see [**81**, Appendix E] and therefore $U \neq V$. It follows from Theorem 2.1 that there exists $\eta = \eta_{\widetilde{u},\widetilde{v}} \in \mathfrak{S}(A_\gamma) \cap (-\infty, 0)$ and an eigenvector ζ of A_γ to the eigenvalue η such that

$$U(s,t) - V(s,t) = e^{\eta s}\big(\zeta(t) + \kappa(s,t)\big)$$

where κ decays with all its derivative exponentially with uniform weight. Therefore maybe after choosing R larger we can assume that for every $s \geq R$ it holds that $U(s,t) - V(s,t) \neq 0$ for every $t \in S^1$ and

$$\deg\left(t \mapsto \frac{\mathfrak{T}\big(U(s,t) - V(s,t)\big)}{\big|\mathfrak{T}\big(U(s,t) - V(s,t)\big)\big|}\right) = w(\mathfrak{T}\zeta) = w(\eta) \tag{169}$$

where $w(\eta)$ is the winding of the eigenvalue η as explained in (115). Let $\gamma \in C^\infty([R, \infty), [0,1])$ be a smooth cutoff function satisfying

$$\gamma(s) = \begin{cases} 0 & s \in [R, 2R] \\ 1 & s \in [2R+1, \infty) \end{cases}$$

Define

$$F \colon [R, \infty) \times S^1 \to \mathbb{C}, \quad (s,t) \mapsto \mathfrak{T}\big(U(s,t) - V(s,t)\big) + \gamma(s).$$

Abbreviate by

$$Z = [R, \infty) \times S^1$$

the half-infinite cylinder which we identify with the zero section in the bundle $Z \times \mathbb{C}$. Abbreviate further

$$\Gamma_F = \{(z, F(z)) : z \in Z\} \subset Z \times \mathbb{C}$$

the graph of F. By homotopy invariance of the algebraic intersection number as long as boundaries do not intersect we get

$$\operatorname{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_{K^c}, \widetilde{v}|_{K^c}) = \operatorname{int}(Z, \Gamma_F). \tag{170}$$

If $s \in [R, \infty)$ meets the condition that $F(s,t) \neq 0$ for every $t \in S^1$ we introduce the winding number

$$w_s(F) := \deg\left(t \mapsto \frac{F(s,t)}{|F(s,t)|}\right) \in \mathbb{Z}.$$

Note that because $U - V$ decays exponentially in the s-variable and γ is one for s large enough, there exists $S > R$ such that $F(s,t) \neq 0$ for all $s \geq S$ and for all $t \in S^1$. By homotopy invariance we have

$$w_s(F) = w_S(F), \quad s \geq S.$$

By the choice of R and the fact that $\gamma(R) = 0$ we further have $F(R,t) \neq 0$ for every $t \in S^1$ and therefore the winding number $w_R(F)$ is well defined. It follows that

$$\mathrm{int}(Z, \Gamma_F) = w_S(F) - w_R(F). \tag{171}$$

Using again that $\gamma(R) = 0$ we get from (169) that

$$w_R(F) = w(\eta). \tag{172}$$

Moreover, because $U - V$ decays exponentially, the map $t \mapsto \frac{F(S,t)}{|F(S,t)|}$ is homotopic to the constant map $t \mapsto 1$ and therefore

$$w_S(F) = 0. \tag{173}$$

Combining (170) and (171) with (172) and (173) we obtain

$$\mathrm{int}(\widetilde{u}_{\mathfrak{T},\beta,\epsilon}|_{K^c}, \widetilde{v}|_{K^c}) = -w(\eta). \tag{174}$$

The proposition now follows by combining (164) and (165) with (166), (167), (169), and (174). □

The following Lemma enables us to estimate the winding number of $\eta_{\widetilde{u},\widetilde{v}}$ for fast finite energy planes.

Lemma 4.3. *Assume that $\widetilde{u}$ and $\widetilde{v}$ are fast finite energy planes with common asymptotic Reeb orbit γ such that* $\mathrm{im}(\widetilde{u}) \neq \mathrm{im}(\widetilde{v})$.

- **(a):** *Assume that the asymptotic eigenvectors of $\widetilde{u}$ and $\widetilde{v}$ satisfy $\zeta_{\widetilde{u}} \neq \zeta_{\widetilde{v}}$. Then $w(\eta_{\widetilde{u},\widetilde{v}}) = 1$.*
- **(b):** *In general, $w(\eta_{\widetilde{u},\widetilde{v}}) \leq 1$.*

Proof: We first prove assertion (a). In this case we have

$$\eta_{\widetilde{u},\widetilde{v}} = \max\{\eta_u, \eta_v\}.$$

Since $\widetilde{u}$ and $\widetilde{v}$ are fast it follows from Corollary 0.13 that $w(\eta_{\widetilde{u},\widetilde{v}}) = 1$. This proves assertion (a).

To prove assertion (b) it suffices in view of the already proved assertion (a) to consider the case $\zeta_{\widetilde{u}} = \zeta_{\widetilde{v}}$. In this case we have

$$\eta_{\widetilde{u},\widetilde{v}} < \eta_u = \eta_v.$$

Corollary 0.32 tells us that the winding number is monotone and hence using that $\widetilde{u}$ is fast we get

$$w(\eta_{\widetilde{u},\widetilde{v}}) \leq w(\eta_u) = 1.$$

This finishes the proof of the Lemma. □

We are now in position to prove Siefring's inequality

Proof of Theorem 4.1: The fact that $\mathrm{int}(\widetilde{u}, \widetilde{v}) \geq 0$ is a consequence of positivity of intersection for holomorphic curves and was already stated in Theorem 2.3.

To prove the second inequality we combine the definition of Siefring's intersection number with Proposition 4.2 and assertion (b) of Lemma 4.3 to get

$$\mathrm{sief}(\widetilde{u},\widetilde{v}) = \mathrm{int}(\widetilde{u}_{\mathfrak{T},\epsilon},\widetilde{v}) + 1 = \mathrm{int}(\widetilde{u},\widetilde{v}) - w(\eta_{\widetilde{u},\widetilde{v}}) + 1 \geq \mathrm{int}(\widetilde{u},\widetilde{v}). \tag{175}$$

This finishes the proof of the theorem. □

In the proof of Theorem 4.1 we did not use assertion (a) of Lemma 4.3. Plugging assertion (a) into (175) we see that in "most" cases Siefring's inequality is actually an equality, namely

Theorem 4.4. *Assume that $\widetilde{u}$ and $\widetilde{v}$ are fast finite energy planes asymptotic to the same non-degenerate periodic Reeb orbit such that their asymptotic eigenvectors satisfy $\zeta_{\widetilde{u}} \neq \zeta_{\widetilde{v}}$. Then*

$$\mathrm{int}(\widetilde{u},\widetilde{v}) = \mathrm{sief}(\widetilde{u},\widetilde{v}).$$

Recall that $\mathbb{R}$ acts on a fast finite energy plane $\widetilde{u} = (u,a)$ by $r_*(u,a) = (u,a+r)$. Bringing the $\mathbb{R}$-action into play we can give a quantitative statement what me mean that in "most" cases Siefring's inequality is actually an equality.

Corollary 4.5. *Assume that $\widetilde{u} = (u,a)$ and $\widetilde{v} = (v,b)$ are two fast finite energy planes with the same asymptotic Reeb orbit γ. Then there exists a $C \subset \mathbb{R}$ with the property that $\#C \leq 2$ and for every $r \in \mathbb{R} \setminus C$ the algebraic intersection number $\mathrm{int}(r_*\widetilde{u},\widetilde{v})$ is defined and satisfies*

$$\mathrm{int}(r_*\widetilde{u},\widetilde{v}) = \mathrm{sief}(r_*\widetilde{u},\widetilde{v}) = \mathrm{sief}(\widetilde{u},\widetilde{v}).$$

Proof: That $\mathrm{sief}(r_*\widetilde{u},\widetilde{v}) = \mathrm{sief}(\widetilde{u},\widetilde{v})$ is an immediate consequence of the homotopy invariance of Siefring's intersection number and is true for any $r \in \mathbb{R}$. To prove the first equality we distinguish two cases.

Case 1: *We assume that* $\mathrm{im}(u) \neq \mathrm{im}(v)$.

In this case we have

$$\mathrm{im}(r_*\widetilde{u}) \neq \mathrm{im}(\widetilde{v}), \quad \forall\, r \in \mathbb{R}.$$

Therefore the algebraic intersection number $\mathrm{int}(r_*\widetilde{u},\widetilde{v})$ is defined for every $r \in \mathbb{R}$. Recall from Lemma 5.3 that the asymptotic eigenvector transforms under the $\mathbb{R}$-action as $\zeta_{r_*\widetilde{u}} = e^{-\frac{\eta_u r}{\tau}}\zeta_{\widetilde{u}}$ where $\tau > 0$ is the period of the periodic Reeb orbit γ. Therefore if we define

$$C := \{r \in \mathbb{R} : \zeta_{r_*\widetilde{u}} = \zeta_{\widetilde{v}}\}$$

we conclude with the fact that asymptotic eigenvector η_u is different from zero that the set C consists of at most one point. If $r \in \mathbb{R} \setminus C$ Theorem 4.4 implies that $\mathrm{int}(r_*\widetilde{u},\widetilde{v}) = \mathrm{sief}(r_*\widetilde{u},\widetilde{v})$. This finishes the proof of the Corollary for Case 1.

Case 2: *We assume that* $\mathrm{im}(u) = \mathrm{im}(v)$.

We first show that

$$\#\{r \in \mathbb{R} : \mathrm{im}(r_*\widetilde{u}) = \mathrm{im}(\widetilde{v})\} \leq 1. \tag{176}$$

To see that assume that

$$\mathrm{im}\big((r_1)_*\widetilde{u}\big) = \mathrm{im}(\widetilde{v}) = \mathrm{im}\big((r_2)_*\widetilde{u}\big), \quad r_1, r_2 \in \mathbb{R}.$$

We have to show that $r_1 = r_2$. Without loss of generality assume that $r_1 = 0$ and $r_2 = r$. Abbreviate

$$\underline{a} := \inf_{z \in \mathbb{C}} a(z).$$

In view of the asymptotic behavior of the fast finite energy plane $\widetilde{u} = (u, a)$ we conclude that

$$\underline{a} = \min_{z \in \mathbb{C}} a(z) < \infty.$$

Using that the image of $\widetilde{u}$ coincides with the image of $r_* \widetilde{u} = (u, a + r)$ we conclude that

$$\underline{a} = \underline{a + r} = \underline{a} + r$$

and therefore $r = 0$. This establishes the truth of (176).

We now set

$$C_0 := \{ r \in \mathbb{R} : \mathrm{im}(r_* \widetilde{u}) = \mathrm{im}(\widetilde{v}) \}.$$

By (176) we have

$$\# C_0 \leq 1.$$

Case 2 now follows by applying the reasoning of Case 1 to the set $\mathbb{R} \setminus C_0$. $\square$

5. Computations and applications

Lemma 5.1. *Assume that $\widetilde{u} = (u, a)$ is a fast finite energy plane with asymptotic Reeb orbit γ such that $\mathrm{im}(u) \cap \mathrm{im}(\gamma) = \emptyset$. Then Siefring's self-intersection number of $\widetilde{u}$ vanishes, i.e.,*

$$\mathrm{sief}(\widetilde{u}, \widetilde{u}) = 0.$$

Proof: We argue by contradiction and assume that $\mathrm{sief}(\widetilde{u}, \widetilde{u}) \neq 0$. By homotopy invariance of Siefring's intersection number we get

$$\mathrm{sief}(r_* \widetilde{u}, \widetilde{u}) \neq 0, \quad \forall\, r \in \mathbb{R}.$$

It follows from Corollary 4.5 that there exists $C \subset \mathbb{R}$ of cardinality $\# C \leq 2$ such that

$$\mathrm{int}(r_* \widetilde{u}, \widetilde{u}) \neq 0, \quad r \in \mathbb{R} \setminus C.$$

That means for every $r \in \mathbb{R} \setminus C$ there exist $z_1^r, z_2^r \in \mathbb{C}$ with the property that

$$r_* \widetilde{u}(z_1^r) = \widetilde{u}(z_2^r).$$

Equivalently

$$u(z_1^r) = u(z_2^r), \; a(z_1^r) + r = a(z_2^r), \quad r \in \mathbb{R} \setminus C. \tag{177}$$

In view of the asymptotic behavior the function $a \colon \mathbb{C} \to \mathbb{R}$ is bounded from below. It follows that

$$\lim_{r \to \infty} a(z_2^r) = \infty$$

and therefore

$$\lim_{r \to \infty} |z_2^r| = \infty. \tag{178}$$

In view of Step 2 of the proof of Corollary 2.2 there exists a compact subset $K \subset \mathbb{C}$ such that

$$z_1^r \in K, \quad r \in \mathbb{R} \setminus C.$$

Hence we can find a subsequence r_ν for $\nu \in \mathbb{N}$ and $z_* \in K$ such that

$$\lim_{\nu \to \infty} z_1^{r_\nu} = z_* \in K.$$

Taking advantage of (177) and (178) this implies that

$$u(z_*) = \lim_{\nu\to\infty} u(z_1^{r_\nu}) = \lim_{\nu\to\infty} u(z_2^{r_\nu}) \in \mathrm{im}(\gamma).$$

In particular,

$$\mathrm{im}(u) \cap \mathrm{im}(\gamma) \neq \emptyset.$$

This contradicts the assumption of the lemma and hence Siefring's self-intersection number of $\widetilde{u}$ has to vanish. □

The main idea in the proof of Lemma 5.1 was to take advantage of the $\mathbb{R}$-action on finite energy planes. A stronger result can be obtained by letting r go to infinity and interpreting Siefring's self-intersection number of a fast finite energy plane with Siefring's intersection number of the finite energy plane with the orbit cylinder of its asymptotic Reeb orbit. This idea strictly speaking goes beyond the part of Siefring's intersection theory discussed here, since the orbit cylinder is not a finite energy plane anymore but a punctured finite energy plane. The reader is invited to have a look at Siefring's article [**101**] to see how Siefring's intersection theory works as well for punctured finite energy planes. Using this technology one obtains the following theorem which has Lemma 5.1 as an immediate Corollary.

Theorem 5.2. *Assume that $\widetilde{u} = (u, a)$ is a fast finite energy plane with asymptotic Reeb orbit γ. Then Siefring's self-intersection number of $\widetilde{u}$ can be computed as*

$$\mathrm{sief}(\widetilde{u}, \widetilde{u}) = \#\big\{(z,t) \in \mathbb{C} \times S^1 : u(z) = \gamma(t)\big\}.$$

Proof: Abbreviate by $\tau > 0$ the period of the periodic Reeb orbit γ. For $r \in \mathbb{R}$ define

$$\widetilde{u}_r \colon \mathbb{C} \to N \times \mathbb{R}$$

by

$$\widetilde{u}_r(e^{2\pi(s+it)}) = \Big(u(e^{2\pi(s+r+it)}), a(e^{2\pi(s+r+it)}) - \tau r\Big), \quad (s,t) \in \mathbb{R} \times S^1$$

and

$$\widetilde{u}_r(0) = \big(u(0), a(0) - \tau r\big) = (-\tau r)_* \widetilde{u}(0).$$

Then in view of the asymptotic behavior of $\widetilde{u}$ the restriction of $\widetilde{u}_r$ to $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ converges in the C^∞_{loc}-topology to the orbit cylinder

$$\widetilde{\gamma} \colon \mathbb{C}^* \to N \times \mathbb{R}, \quad e^{2\pi(s+it)} \mapsto (\gamma(t), \tau s)$$

as r goes to infinity. In view of the homotopy invariance of Siefring's intersection number we have

$$\mathrm{sief}(\widetilde{u}, \widetilde{u}) = \mathrm{sief}(\widetilde{u}_r, \widetilde{u}), \quad \forall\, r \in \mathbb{R}.$$

Hence by letting r go to infinity we obtain

$$\mathrm{sief}(\widetilde{u}, \widetilde{u}) = \mathrm{sief}(\widetilde{\gamma}, \widetilde{u}). \tag{179}$$

Observe that the map

$$(z,t) \mapsto \Big(z, e^{2\pi\big(\frac{a(z)}{\tau} + it\big)}\Big)$$

gives rise to a bijection between intersection points

$$\big\{(z,t) \in \mathbb{C} \times S^1 : u(z) = \gamma(t)\big\} \cong \big\{(z_1, z_2) \in \mathbb{C} \times \mathbb{C} : \widetilde{u}(z_1) = \widetilde{\gamma}(z_2)\big\}.$$

Because $\widetilde{u}$ is fast u is an immersion transverse to the Reeb vector field R and therefore

$$\widetilde{u} \pitchfork \widetilde{\gamma}.$$

Since both $\widetilde{u}$ and $\widetilde{\gamma}$ are holomorphic by positivity of intersection we obtain for the algebraic intersection number

$$\operatorname{int}(\widetilde{u}, \widetilde{\gamma}) = \#\big\{(z,t) \in \mathbb{C} \times S^1 : u(z) = \gamma(t)\big\}. \tag{180}$$

The asymptotic eigenvector $\zeta_{\widetilde{\gamma}}$ of the orbit cylinder vanishes while the asymptotic eigenvector $\zeta_{\widetilde{u}}$ does not vanish. Therefore it follows from (4.4) that

$$\operatorname{int}(\widetilde{u}, \widetilde{\gamma}) = \operatorname{sief}(\widetilde{u}, \widetilde{\gamma}). \tag{181}$$

Combining (179, (180), and (181) the Theorem follows. □

One reason why the vanishing of Siefring's self-intersection number of a fast finite energy plane $\widetilde{u} = (u, a)$ is so useful is the fact that it implies that u is an embedding.

Theorem 5.3. *Assume that $\widetilde{u} = (u, a)$ is a fast finite energy plane with asymptotic orbit γ such that $\operatorname{im}(u) \cap \operatorname{im}(\gamma) = \emptyset$. Then $u \colon \mathbb{C} \to N$ is an embedding.*

Proof: Since $\widetilde{u}$ is fast the map u is already an immersion and hence it remains to prove that it is injective. Hence assume that $z, z' \in \mathbb{C}$ such that

$$u(z) = u(z'). \tag{182}$$

We first show that

$$a(z) = a(z') \tag{183}$$

In order to prove (183) we argue by contradiction and assume $a(z) \neq a(z')$. Set $c = a(z') - a(z)$. It follows that

$$c_* \widetilde{u}(z) = (u(z), a(z) + c) = (u(z'), a(z')) = \widetilde{u}(z').$$

But since $c \neq 0$ the maps $\widetilde{u}$ and $c_* \widetilde{u}$ have different images. Therefore their algebraic intersection number $\operatorname{int}(\widetilde{u}, c_* \widetilde{u})$ is defined and by positivity of intersection we conclude that

$$\operatorname{int}(\widetilde{u}, c_* \widetilde{u}) \geq 1.$$

Therefore by Theorem 4.1 we obtain

$$\operatorname{sief}(\widetilde{u}, \widetilde{u}_c) \geq 1.$$

Since Siefring's intersection number is a homotopy invariant we get

$$\operatorname{sief}(\widetilde{u}, \widetilde{u}) \geq 1.$$

However, by assumption of the theorem we have $\operatorname{im}(u) \cap \operatorname{im}(\gamma) = \emptyset$ and therefore by Lemma 5.1 it follows that Siefring's self-intersection number of $\widetilde{u}$ vanishes. This contradiction proves (183).
It follows from equations (182) and (183) that

$$\widetilde{u}(z) = \widetilde{u}(z').$$

A point $z \in \mathbb{C}$ is called an *injective point* of $\widetilde{u}$ if and only if $\widetilde{u}^{-1}(\widetilde{u}(z)) = \{z\}$ and $d\widetilde{u}(z) \neq 0$. Abbreviate

$$I := \{z \in \mathbb{C} : z \text{ injective point of } \widetilde{u}\}$$

the set of injective points and let

$$S := \mathbb{C} \setminus I$$

be its complement, i.e., the set of non-injective points. For a finite energy plane the following alternative holds. Either the set S of non-injective points is discrete, or the finite energy plane is multiply covered in the sense that there exists $\widetilde{v} \colon \mathbb{C} \to N \times \mathbb{R}$ a finite energy plane and a polynomial p satisfying $\deg(p) \geq 2$ such that

$$\widetilde{u} = \widetilde{v} \circ p.$$

For a proof of this fact we refer to [**53**, Appendix]. We also recommend to look at the analogue of this fact in the closed case, see [**81**, Proposition 2.5.1,Theorem E.1.2]. We next explain that if a finite energy plane $\widetilde{u} = (u, a)$ is fast it cannot be multiply covered. To see that suppose by contradiction that there exists a finite energy plane $\widetilde{v} = (v, b)$ such that $\widetilde{w} = \widetilde{v} \circ p$ for a polynomial of degree at least 2. It follows that

$$u = v \circ p.$$

Since the polynomial p has degree at least 2 it must have a critical point, i.e., there exists a point $z \in \mathbb{C}$ such that

$$dp(z) = 0.$$

It follows that

$$du(z) = dv(p(z))dp(z) = 0$$

contradicting the fact that u is an immersion, since it is fast. We have shown that $\widetilde{u}$ cannot be multiply covered and therefore its set S of non-injective points is discrete.

In order to finish the proof of the theorem let us now assume that there exists $z \neq z' \in \mathbb{C}$ such that $\widetilde{u}(z) = \widetilde{u}(z')$. For $\epsilon > 0$ let us abbreviate by $D_\epsilon(z) = \{w \in \mathbb{C} : ||w - z|| \leq \epsilon\}$ the closed ϵ-ball around z. Because the set S of non-injective points of $\widetilde{u}$ is discrete we can find $\epsilon > 0$ such that the following condition holds true

$$\mathrm{im}(\widetilde{u}|_{\partial D_\epsilon(z)}) \cap \mathrm{im}(\widetilde{u}|_{\partial D_\epsilon(z')}) = \emptyset.$$

In view of this fact the algebraic intersection number

$$\mathrm{int}(\widetilde{u}|_{D_\epsilon(z)}, \widetilde{u}|_{D_\epsilon(z')}) \in \mathbb{Z}$$

is well defined. Because $\widetilde{u}(z) = \widetilde{u}(z')$ it follows from positivity of intersection that

$$\mathrm{int}(\widetilde{u}|_{D_\epsilon(z)}, \widetilde{u}|_{D_\epsilon(z')}) \geq 1.$$

Choose $c_0 > 0$ such that for every $|c| \leq c_0$ it holds that

$$\mathrm{im}(\widetilde{u}|_{\partial D_\epsilon(z)}) \cap \mathrm{im}(c_* \widetilde{u}|_{\partial D_\epsilon(z')}) = \emptyset.$$

By homotopy invariance of the algebraic intersection number we conclude that

$$\mathrm{int}(\widetilde{u}|_{D_\epsilon(z)}, c_* \widetilde{u}|_{D_\epsilon(z')}) \geq 1.$$

Hence by positivity of intersection if $0 < |c| \leq c_0$ we obtain

$$\mathrm{int}(\widetilde{u}, c_* \widetilde{u}) \geq 1.$$

Again this implies that $\mathrm{sief}(\widetilde{u}, \widetilde{u}) \geq 1$ in contradiction to the assumption of the theorem. This finishes the proof of the theorem. □

We will see later in Lemma 6.3 that the converse of Theorem 5.3 is true as well. With the help of this Lemma we can give now the following characterization of fast finite energy planes with vanishing Siefring self-intersection number.

Theorem 5.4. *Assume that* $\widetilde{u} = (u, a)$ *is a fast finite energy plane with asymptotic orbit* γ*. Then the following assertions are equivalent.*

(i): $\mathrm{sief}(\widetilde{u}, \widetilde{u}) = 0$,
(ii): $\mathrm{im}(u) \cap \mathrm{im}(\gamma) = \emptyset$,
(iii): u *is embedded,*
(iv): $\widetilde{u}$ *is embedded.*

Proof: The equivalence of (i) and (ii) follows from Theorem 5.2. That (ii) implies (iii) is the content of Theorem 5.3. The implication (iii) $\Rightarrow$ (iv) is obvious. Finally the implication (iv) $\Rightarrow$ (i) is proved in Lemma 6.3 below. □

A further important contribution of Siefring's intersection number is that it can be used to make sure that two fast finite energy planes do not intersect.

Corollary 5.5. *Assume that* $\widetilde{u} = (u, a)$ *is a fast finite energy plane with asymptotic orbit* γ *such that* $\mathrm{im}(u) \cap \mathrm{im}(\gamma) = \emptyset$*. Assume that* $\widetilde{v}$ *is a fast finite energy plane which is homotopic to* $\widetilde{u}$ *and such that* $\mathrm{im}(\widetilde{v}) \neq \mathrm{im}(\widetilde{u})$*. Then*

$$\mathrm{im}(\widetilde{u}) \cap \mathrm{im}(\widetilde{v}) = \emptyset.$$

Proof: By Lemma 5.1 we have

$$\mathrm{sief}(\widetilde{u}, \widetilde{u}) = 0.$$

Hence by homotopy invariance of Siefring's intersection number we obtain

$$\mathrm{sief}(\widetilde{u}, \widetilde{v}) = 0.$$

Because $\mathrm{im}(\widetilde{u}) \neq \mathrm{im}(\widetilde{v})$ the algebraic intersection number $\mathrm{int}(\widetilde{u}, \widetilde{v})$ is well-defined and by Theorem 4.1 we have

$$0 \leq \mathrm{int}(\widetilde{u}, \widetilde{v}) \leq \mathrm{sief}(\widetilde{u}, \widetilde{v}) = 0$$

so that we get

$$\mathrm{int}(\widetilde{u}, \widetilde{v}) = 0.$$

By positivity of intersection from Theorem 2.3 we conclude that

$$\mathrm{im}(\widetilde{u}) \cap \mathrm{im}(\widetilde{v}) = \emptyset.$$

This finishes the proof of the Corollary. □

CHAPTER 14

The moduli space of fast finite energy planes

1. Fredholm operators

Before explaining the class of operators we are interested we make some general remarks about Cauchy-Riemann operators. Consider $\mathbb{C}$ with its standard complex structure i but with a measure μ maybe different from the Lebesgue measure, namely

$$\mu = \mu_h := h dx \wedge dy$$

where $h\colon \mathbb{C} \to \mathbb{R}_+$ is a smooth positive function. Abbreviate by

$$\bar{\partial}\colon C^\infty(\mathbb{C}, \mathbb{C}) \to C^\infty(\mathbb{C}, \mathbb{C})$$

the standard Cauchy-Riemann operator on $\mathbb{C}$ which is given for $\zeta \in C^\infty(\mathbb{C}, \mathbb{C})$ by

$$\bar{\partial}\zeta = \partial_x \zeta + i\partial_y \zeta.$$

Then we define the Cauchy-Riemann operator with respect to the measure μ as

$$\bar{\partial}_\mu := \frac{1}{\sqrt{h}}\bar{\partial}\colon C^\infty(\mathbb{C}, \mathbb{C}) \to C^\infty(\mathbb{C}, \mathbb{C}).$$

Of course

$$\ker \bar{\partial}_\mu = \ker \bar{\partial}.$$

The reason why it is natural to consider the operator $\bar{\partial}_\mu$ is, that if we endow the target $\mathbb{C}$ with its standard inner product the norm $||\bar{\partial}_\mu \zeta||$ has an intrinsic description, i.e., a description which only depends on μ, the complex structure i on the domain $\mathbb{C}$, and the inner product on the target $\mathbb{C}$ but not on the coordinates on $\mathbb{C}$ used to define $\bar{\partial}$. Why this is true is explained in the following lemma.

Lemma 1.1. *Assume that $z \in \mathbb{C}$ and $\widehat{z} \neq 0 \in T_z\mathbb{C} = \mathbb{C}$ is an arbitrary nonvanishing tangent vector, then*

$$||\bar{\partial}_\mu \zeta(z)|| = \frac{||d\zeta(z)\widehat{z} + id\zeta(z)i\widehat{z}||}{\sqrt{\mu(\widehat{z}, i\widehat{z})}}.$$

Since $\widehat{z}$ is arbitrary, the right hand side is intrinsic.

Proof: For $\widehat{z} \neq 0 \in T_z\mathbb{C}$ we abbreviate

$$f(\widehat{z}) = \frac{||d\zeta(z)\widehat{z} + id\zeta(z)i\widehat{z}||}{\sqrt{\mu(\widehat{z}, i\widehat{z})}}.$$

By definition we have

$$||\bar{\partial}_\mu \zeta(z)|| = f(1).$$

Therefore it remains to show that the function f is constant. For $\widehat{z} = \widehat{x} + i\widehat{y} = \widehat{r}\cos\widehat{\theta} + i\widehat{r}\sin\widehat{\theta} = \widehat{r}e^{i\widehat{\theta}} \neq 0 \in T_z\mathbb{C} = \mathbb{C}$ we compute

$$\begin{aligned} d\zeta(z)\widehat{z} + id\zeta(z)i\widehat{z} &= d\zeta(z)\widehat{x} + d\zeta(z)i\widehat{y} + id\zeta(z)i\widehat{x} - id\zeta(z)\widehat{y} \\ &= d\zeta(z)\widehat{x} + id\zeta(x)i\widehat{x} - i\Big(d\zeta(z)\widehat{y} - id\zeta(z)i\widehat{y}\Big) \\ &= (\widehat{x} - i\widehat{y})\Big(d\zeta(z)1 + id\zeta(z)i\Big) \\ &= \widehat{r}e^{-i\widehat{\theta}}\bar{\partial}\zeta(z) \end{aligned}$$

and therefore

$$||d\zeta(z)\widehat{z} + id\zeta(z)i\widehat{z}|| = \widehat{r}||\bar{\partial}\zeta(z)||.$$

Moreover,

$$\begin{aligned} \mu(z)(\widehat{z}, i\widehat{z}) &= h(z)(dx \wedge dy)(\widehat{x} + i\widehat{y}, -\widehat{y} + i\widehat{x}) \\ &= h(z)(\widehat{x}^2 + \widehat{y}^2) \\ &= \widehat{r}^2 h(z). \end{aligned}$$

Combining these expressions we conclude that

$$f(\widehat{z}) = \frac{||\bar{\partial}\zeta(z)||}{\sqrt{h(z)}}$$

is independent of $\widehat{z}$. This proves the lemma. □

Choose a function $\gamma \in C^\infty([0,\infty),(0,\infty))$ satisfying

$$\gamma(r) = \begin{cases} 1 & r \leq \frac{1}{2} \\ \frac{1}{4\pi^2 r^2} & r \geq 1. \end{cases}$$

Define the function $h_\gamma \in C^\infty(\mathbb{C}, \mathbb{R}_+)$ by

$$h_\gamma(z) = \gamma(|z|).$$

In the following we endow the complex plane $\mathbb{C}$ with the measure $\mu = \mu_{h_\gamma}$. The measure μ has the following properties. If $D_r = \{z \in \mathbb{C} : ||z|| \leq r\}$ denotes the ball of radius r centered at the origin we have

$$\mu|_{D_{\frac{1}{2}}} = dx \wedge dy.$$

Moreover, if $\phi \colon \mathbb{R} \times S^1 \to \mathbb{C} \setminus \{0\}$ is the biholomorphism

$$\phi(s,t) = e^{2\pi(s+it)}, \quad (s,t) \in \mathbb{R} \times S^1,$$

then we obtain

$$\phi^*\mu|_{\mathbb{C}\setminus D_1} = ds \wedge dt.$$

That means on the cylindrical end $\mathbb{C} \setminus D_1$ the measure μ coincides with the standard measure on the cylinder.

Denote by $M_2(\mathbb{R})$ the vector space of real 2×2-matrices. Let

$$S \in C^\infty(\mathbb{C}, M_2(\mathbb{R}))$$

be a smooth family of 2×2-matrices parametrized by $\mathbb{C}$. Suppose that there exists

$$S_\infty \in C^\infty(S^1, \mathrm{Sym}(2))$$

a smooth loop of symmetric 2×2-matrices such that uniformly in the C^∞-topology it holds that

$$\lim_{s \to \infty} 2\pi e^{2\pi s} S(e^{2\pi(s+it)}) = S_\infty(t).$$

We abbreviate by

$$H_1 := W^{1,2}(\mathbb{C}, \mu; \mathbb{C}) \tag{184}$$

the Hilbert space of $W^{1,2}$-maps from $\mathbb{C}$ to $\mathbb{C}$ where the domain $\mathbb{C}$ is endowed with the measure μ and the target with the standard inner product and similarly for the L^2-space

$$H_0 := L^2(\mathbb{C}, \mu; \mathbb{C}). \tag{185}$$

We consider the bounded linear operator

$$L_S \colon H_1 \to H_0, \quad \zeta \mapsto \frac{1}{\sqrt{h_\gamma}}(\bar{\partial}\zeta + S\zeta) = \bar{\partial}_\mu \zeta + \frac{1}{\sqrt{h_\gamma}} S\zeta. \tag{186}$$

From the assumed behavior of S we also get an asymptotic operator

$$A_{S_\infty} \colon W^{1,2}(S^1, \mathbb{C}) \to L^2(S^1, \mathbb{C}), \quad \zeta \mapsto -J_0 \partial_t \zeta - S_\infty \zeta.$$

Associated to the loop of symmetric matrices S_∞ is the path of symplectic matrices

$$\Psi = \Psi_{S_\infty} \in C^\infty([0,1], \mathrm{Sp}(1))$$

defined by

$$\partial_t \Psi(t) = J_0 S_\infty(t) \Psi, \quad \Psi(0) = \mathrm{id}.$$

The following Theorem is due to Schwarz [**97**]

Theorem 1.2 (Schwarz). *Assume that* $\ker A_{S_\infty} = \{0\}$*, i.e., the path of symplectic matrices* Ψ_{S_∞} *is non-degenerate. Then* L_S *is a Fredholm operator and its index computes to be*

$$\mathrm{ind} L_S = \mu_{CZ}(\Psi_{S_\infty}) + 1.$$

Remark 1.3. *The theorem of Schwarz is very reminiscent of the theorem of Riemann-Roch. Indeed, by noting that the Euler characteristic of the complex plane is one, the index formula of Schwarz can be interpreted as* $\mathrm{ind} L_S = \mu_{CZ}(\Psi_{S_\infty}) + 1 \cdot \chi(\mathbb{C})$*. We refer to the thesis by Bourgeois* [**18**, Section 5.2] *for a derivation of Schwarz theorem out of the Riemann-Roch formula.*

We do not prove Schwarz theorem in detail, however, we next illustrate it for a family of examples.

Illustration of Theorem 1.2: Pick a cutoff function $\beta \in C^\infty(\mathbb{R}, [0,1])$ satisfying

$$\beta(s) = \begin{cases} 1 & s \geq 1 \\ 0 & s \leq 0. \end{cases}$$

Choose further

$$\mu \in \mathbb{R} \setminus 2\pi\mathbb{Z}.$$

Define $S = S^\mu \colon \mathbb{C} \to \mathrm{Sym}(2)$ by

$$S(e^{2\pi(s+it)}) = \frac{\beta(s)\mu}{2\pi e^{2\pi s}} \mathrm{id}.$$

We first examine the kernel of the operator L_S. Suppose that $\zeta \in \ker L_S$ and set

$$\eta = \zeta \circ \phi \colon \mathbb{R} \times S^1 \to \mathbb{C}.$$

It follows that η is a solution of the PDE

$$\partial_s\eta + i\partial_t\eta + \beta\mu\eta = 0. \tag{187}$$

Write η as a Fourier series

$$\eta(s,t) = \sum_{k=-\infty}^{\infty} \eta_k(s)e^{2\pi ikt}.$$

It follows from (187) that each Fourier coefficient is a solution of the ODE

$$\partial_s\eta_k(s) - 2\pi k\eta_k(s) + \beta(s)\mu\eta_k(s) = 0, \quad k \in \mathbb{Z}.$$

Since $\beta(s) = 1$ for $s \geq 1$ we get

$$\eta_k(s) = \eta_k(1)e^{(2\pi k-\mu)(s-1)}, \quad s \geq 1.$$

Because $\eta|_{[0,\infty)\times S^1} \in L^2([0,\infty)\times S^1, \mathbb{C})$ we conclude

$$\eta_k = 0, \quad k > \frac{\mu}{2\pi} \tag{188}$$

Since $\beta(s) = 0$ for $s \leq 0$ we conclude that

$$\eta_k(s) = \eta_k(0)e^{2\pi ks}, \quad s \leq 0.$$

Because $\zeta \in W^{1,2}(\mathbb{C}, \mu; \mathbb{C})$ it is continuous and the limit $\lim_{s\to-\infty}\eta(s,t) = \zeta(0)$ exists. Therefore,

$$\eta_k = 0, \quad k < 0. \tag{189}$$

Denoting by $\lfloor\frac{\mu}{2\pi}\rfloor = \max\{n \in \mathbb{Z} : n \leq \frac{\mu}{2\pi}\}$ the integer part of $\frac{\mu}{2\pi}$ we conclude from (188) and (189) that

$$\dim(\ker L_S) = \begin{cases} 2\big(\lfloor\frac{\mu}{2\pi}\rfloor + 1\big) & \lfloor\frac{\mu}{2\pi}\rfloor \geq 0 \\ 0 & \text{else.} \end{cases} \tag{190}$$

We next compute the dimension of the cokernel of L_S. Suppose that $\eta \in \operatorname{coker} L_S = (\operatorname{im} L_S)^{\perp_{H_0}}$. That means that $\eta \in H_0$ satisfies

$$\langle L_S\zeta, \eta\rangle_{H_0} = 0, \quad \forall\ \zeta \in H_1. \tag{191}$$

We introduce the function

$$f \in C^\infty(\mathbb{R}, \mathbb{R}_+), \quad s \mapsto 2\pi e^{2\pi s}\sqrt{\gamma(e^{2\pi s})}.$$

Note that f satisfies

$$f(s) = \begin{cases} 2\pi e^{2\pi s} & s \leq -\frac{\ln(2)}{2\pi} \\ 1 & s \geq 0. \end{cases}$$

We compute using the coordinate change $z = x + iy = e^{2\pi(s+it)}$

$$\begin{aligned}
\langle L_S\zeta, \eta\rangle_{H_0} &= \int_{\mathbb{C}} \Big\langle \frac{1}{\sqrt{\gamma(|z|)}}\Big(\partial_x\zeta + i\partial_y\zeta + S\zeta\Big), \eta\Big\rangle \gamma(|z|)dxdy \\
&= \int_{\mathbb{R}}\int_0^1 \Big\langle \frac{1}{2\pi e^{2\pi s}\sqrt{\gamma(e^{2\pi s})}}\Big(\partial_s\zeta + i\partial_t\zeta\Big), \eta\Big\rangle 4\pi^2 e^{4\pi s}\gamma(e^{2\pi s})dsdt \\
&\quad + \int_{\mathbb{R}}\int_0^1 \Big\langle \frac{1}{\sqrt{\gamma(e^{2\pi s})}} S\zeta, \eta\Big\rangle 4\pi^2 e^{4\pi s}\gamma(e^{2\pi s})dsdt \\
&= \int_{\mathbb{R}}\int_0^1 \Big\langle \partial_s\zeta + i\partial_t\zeta, \eta\Big\rangle f(s)dsdt + \int_{\mathbb{R}}\int_0^1 \Big\langle \beta\mu\zeta, \eta\Big\rangle f(s)^2dsdt \\
&= \int_{\mathbb{R}}\int_0^1 \Big\langle \partial_s\zeta + i\partial_t\zeta, \eta\Big\rangle f(s)dsdt + \int_{\mathbb{R}}\int_0^1 \Big\langle \beta\mu\zeta, \eta\Big\rangle f(s)dsdt \\
&= \int_{\mathbb{R}}\int_0^1 \Big\langle \partial_s\zeta + i\partial_t\zeta + \beta\mu\zeta, \eta\Big\rangle f(s)dsdt
\end{aligned} \tag{192}$$

From (191) and (192) we obtain

$$\int_{\mathbb{R}}\int_0^1 \Big\langle \partial_s\zeta + i\partial_t\zeta + \beta\mu\zeta, \eta\Big\rangle f(s)dsdt = 0, \quad \forall\ \zeta \in H_1.$$

By elliptic regularity for the Cauchy Riemann equation this implies that the L^2-function η is smooth and therefore using integration by parts η is a solution of the PDE

$$-\partial_s\eta + i\partial_t\eta + \Big(\mu\beta - \frac{\partial_s f}{f}\Big)\eta = 0$$

or equivalently

$$-\partial_s\eta + i\partial_t\eta + \big(\mu\beta - \partial_s(\ln f)\big)\eta = 0.$$

By introducing the function

$$g \in C^\infty(\mathbb{R},\mathbb{R}), \quad s \mapsto \mu\beta - \partial_s(\ln f)$$

this can be written more compactly as

$$-\partial_s\eta + i\partial_t\eta + g\eta = 0. \tag{193}$$

Note that the function g satisfies

$$g(s) = \left\{ \begin{array}{cc} -2\pi & s \leq -1 \\ \mu & s \geq 1. \end{array} \right.$$

We again write η as a Fourier series

$$\eta(s,t) = \sum_{k=-\infty}^{\infty} \eta_k(s)e^{2\pi ikt}.$$

By (193) we obtain for each Fourier coefficient the ODE

$$-\partial_s\eta_k - 2\pi\eta_k + g\eta_k = 0, \quad k \in \mathbb{Z}.$$

Since $g(s) = \mu$ for $s \geq 1$ we obtain

$$\eta_k(s) = \eta_k(1)e^{(\mu-2\pi k)(s-1)}, \quad s \geq 1.$$

Because $\eta|_{[0,\infty)\times S^1} \in L^2([0,\infty)\times S^1, \mathbb{C})$ we conclude from this that

$$\eta_k = 0, \quad k \leq \frac{\mu}{2\pi}. \tag{194}$$

Using that $g(s) = -2\pi$ for $s \leq -1$ we get

$$\eta_k(s) = \eta_k(-1)e^{-2\pi - 2\pi k}s + 1 = \eta_k(-1)e^{-2\pi(k+1)(s+1)}$$

and because η as a smooth function on $\mathbb{C}$ has to converge when s goes to $-\infty$ we must have

$$\eta_k = 0, \quad k \geq 0. \tag{195}$$

Combining (194) and (195) we get for the dimension of the cokernel of L_S the formula

$$\dim(\text{coker}L_S) = \begin{cases} 2\big(-\lfloor\frac{\mu}{2\pi}\rfloor - 1\big) & \lfloor\frac{\mu}{2\pi}\rfloor < -1 \\ 0 & \text{else.} \end{cases} \tag{196}$$

Combining (190) and (196) we obtain for the index of L_S

$$\text{ind}L_S = \dim(\ker L_S) - \dim(\text{coker}L_S) = 2\left(\left\lfloor\frac{\mu}{2\pi}\right\rfloor + 1\right) = \mu_{CZ}(\Psi_{S_\infty}) + 1.$$

This finishes the illustration of Theorem 1.2. $\square$

2. The first Chern class

In this section we explain the first Chern class of complex vector bundles over the two dimensional sphere. Suppose that $E \to S^2$ is a complex vector bundle, satisfying $\text{rk}_{\mathbb{C}}E = n$. We decompose the sphere $S^2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$ into the union of the upper and lower hemisphere

$$S^2 = S^2_+ \cup S^2_- \tag{197}$$

where

$$S^2_\pm = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : \pm x_3 \geq 0\}.$$

Note that both the upper and the lower hemisphere are diffeomorphic to a closed disk. Therefore there exist complex trivializations

$$\mathfrak{T}_\pm \colon E|_{S^2_\pm} \to S^2_\pm \times \mathbb{C}^n.$$

By endowing the vector bundle E with a Hermitian metric we can assume that $\mathfrak{T}_\pm$ are actually unitary trivializations. Let

$$S = S^2_+ \cap S^2_- \subset S^2$$

be the equator which we identify with the circle $S^1 = \mathbb{R}/\mathbb{Z}$ via the map $t \mapsto (\cos 2\pi t, \sin 2\pi t, 0)$. If $t \in S$, we obtain a unitary linear map

$$\mathfrak{T}_{-,t}\mathfrak{T}_{+,t}^{-1} \colon \mathbb{C}^n \to \mathbb{C}^n$$

i.e.,

$$\mathfrak{T}_{-,t}\mathfrak{T}_{+,t}^{-1} \in U(n).$$

We define the first Chern class of E as

$$c_1(E) := \deg\big(t \mapsto \det(\mathfrak{T}_{+,t}\mathfrak{T}_{-,t}^{-1})\big) \in \mathbb{Z}.$$

Note that since $S^2_\pm$ are contractible any two trivializations over the upper or lower hemisphere are homotopic and therefore the first Chern class is independent of the choice of $\mathfrak{T}_\pm$. The first Chern class has the following properties

(i): If $E = S^2 \times \mathbb{C}^n$ is the trivial bundle, it holds that $c_1(E) = 0$.

(ii): If $E_1 \to S^2$ and $E_2 \to S^2$ are two complex vector bundles, the first Chern class of their Whitney sum satisfies

$$c_1(E_1 \oplus E_2) = c_1(E_1) + c_1(E_2) \tag{198}$$

and similarly for their tensor product

$$c_1(E_1 \otimes E_2) = c_1(E_1) + c_1(E_2). \tag{199}$$

(iii): If $L \to S^2$ is a complex line bundle, then $c_1(L) = e(L)$ the Euler number of L. In particular, $c_1(L)$ does not depend on the complex structure of L.

To see why (198) is true choose unitary trivializations

$$\mathfrak{T}^1_\pm \colon E_1|_{S^2_\pm} \to S^2_\pm \times \mathbb{C}^{n_1}, \quad \mathfrak{T}^2_\pm \colon E_2|_{S^2_\pm} \to S^2_\pm \times \mathbb{C}^{n_2}$$

with $n_1 = \mathrm{rk}_\mathbb{C} E_1$ and $n_2 = \mathrm{rk}_\mathbb{C} E_2$. We obtain unitary trivializations

$$\mathfrak{T}^1_\pm \oplus \mathfrak{T}^2_\pm \colon (E_1 \oplus E_2)|_{S^2_\pm} \to S^2_\pm \times \mathbb{C}^{n_1+n_2}.$$

For $t \in S^1$ we get

$$\begin{aligned} \det\big((\mathfrak{T}^1_{+,t} \oplus \mathfrak{T}^2_{+,t})(\mathfrak{T}^1_{-,t} \oplus \mathfrak{T}^2_{-,t})^{-1}\big) &= \det\big(\mathfrak{T}^1_{+,t}(\mathfrak{T}^1_{-,t})^{-1} \oplus \mathfrak{T}^2_{+,t}(\mathfrak{T}^2_{-,t})^{-1}\big) \\ &= \det\big(\mathfrak{T}^1_{+,t}(\mathfrak{T}^1_{-,t})^{-1}\big) \cdot \det\big(\mathfrak{T}^2_{+,t}(\mathfrak{T}^2_{-,t})^{-1}\big). \end{aligned}$$

Therefore the first Chern class of the Whitney sum $E_1 \oplus E_2$ satisfies

$$\begin{aligned} c_1(E_1 \oplus E_2) &= \deg\Big(t \mapsto \det\big(\mathfrak{T}^1_{+,t}(\mathfrak{T}^1_{-,t})^{-1}\big) \cdot \det\big(\mathfrak{T}^2_{+,t}(\mathfrak{T}^2_{-,t})^{-1}\big)\Big) \\ &= \deg\Big(t \mapsto \det\big(\mathfrak{T}^1_{+,t}(\mathfrak{T}^1_{-,t})^{-1}\big)\Big) + \deg\Big(t \mapsto \det\big(\mathfrak{T}^2_{+,t}(\mathfrak{T}^2_{-,t})^{-1}\big)\Big) \\ &= c_1(E_1) + c_1(E_2). \end{aligned}$$

This proves (198) and (199) is proved similarly.

The first Chern class can be associated as well to a *symplectic* vector bundle $(E, \omega) \to S^2$, i.e., each fiber (E_z, ω_z) for $z \in S^2$ is a symplectic vector space. In fact the space $\mathcal{J}(E, \omega)$ consisting of all ω-compatible almost complex structures $J \colon E \to E$ is nonempty and contractible. In particular, $\mathcal{J}(E, \omega)$ is connected. Hence we set

$$c_1(E, \omega) := c_1(E, J), \quad J \in \mathcal{J}(E, \omega)$$

and this is well defined by homotopy invariance of the first Chern class, since $\mathcal{J}(E, \omega)$ is connected.

Assume that $E \to S^2$ is a symplectic vector bundle. As usual we identify the equator of S^2 with the circle $S^1 = \mathbb{R}/\mathbb{Z}$. Suppose that for $t \in [0, 1]$ we have a smooth family of symplectic linear maps

$$\Phi_t \colon E_{[0]} \to E_{[t]}$$

such that

$$\Phi_0 = \mathrm{id} \colon E_{[0]} \to E_{[0]}$$

and for $\Phi_1 \colon E_{[0]} \to E_{[1]} = E_{[0]}$ the non-degeneracy condition

$$\ker(\Phi_1 - \mathrm{id}) = \{0\} \tag{200}$$

holds true. In this set-up we can associate to the path $t \mapsto \Phi_t$ two Conley-Zehnder indices. Namely, if

$$\mathfrak{T}_+ \colon E|_{S^2_+} \to S^2_+ \times \mathbb{C}^n, \quad 2n = \mathrm{rk}_{\mathbb{R}} E$$

is a symplectic trivialization we define the smooth path

$$\Psi^+ \colon [0,1] \to \mathrm{Sp}(n), \quad t \mapsto \mathfrak{T}_{+,[t]} \Phi_t \mathfrak{T}^{-1}_{+,[0]}$$

and set

$$\mu^+_{CZ}(\Phi) := \mu_{CZ}(\Psi^+)$$

which is independent of the trivialization $\mathfrak{T}_+$ by homotopy invariance. Similarly we define $\mu^-_{CZ}(\Phi)$ for a symplectic trivialization over S^2_-.

Lemma 2.1. *The difference of the two Conley-Zehnder indices satisfies*

$$\mu^+_{CZ}(\Phi) - \mu^-_{CZ}(\Phi) = 2c_1(E).$$

Proof: Choose $J \in \mathcal{J}(E, \omega)$ and choose unitary trivializations $\mathfrak{T}_\pm \colon E|_{S^2_\pm} \to S^2_\pm \times \mathbb{C}^n$ with respect to J and $\omega(\cdot, J\cdot)$ such that

$$\mathfrak{T}_{+,0} = \mathfrak{T}_{-,0} =: \mathfrak{T}_0 \colon E_0 \to \mathbb{C}^n.$$

For $t \in [0,1]$ abbreviate

$$U_{[t]} := \mathfrak{T}_{+,[t]} \mathfrak{T}^{-1}_{-,[t]} \in U(n)$$

so that we get a loop

$$U \colon S^1 \to U(n), \quad [t] \mapsto U_{[t]}.$$

We obtain

$$\Psi^+([t]) = \mathfrak{T}_{+,[t]} \Phi_t \mathfrak{T}^{-1}_{+,[0]} = \mathfrak{T}_{+,[t]} \mathfrak{T}^{-1}_{-,[t]} \mathfrak{T}_{-,[t]} \Phi_t \mathfrak{T}_0^{-1} = U([t]) \Psi^-([t]).$$

In particular, the path Ψ^+ is homotopic with fixed endpoints to the concatenation

$$\Psi^+ \cong U\Psi^-(1) \# \Psi^-. \tag{201}$$

In view of the non-degeneracy condition (200) we have by definition of the Conley-Zehnder index

$$\mu^+_{CZ}(\Phi) = \mu_{CZ}(\Psi^+) = \mu_\Delta(\Gamma_{\Psi^+}) \tag{202}$$

where Δ is the diagonal in the symplectic vector space $(\mathbb{C}^n \times \mathbb{C}^n, -\omega \oplus \omega)$ and Γ_{Ψ^+} is the graph of Ψ^+. By homotopy invariance and the concatenation property of the Maslov index we conclude

$$\mu_\Delta(\Gamma_{\Psi_+}) = \mu_\Delta(\Gamma_{\Psi^-}) + \mu_\Delta(\Gamma_{U\Psi^-(1)})$$

Since U is a loop we can write this equation as

$$\mu_\Delta(\Gamma_{\Psi_+}) = \mu_{CZ}(\Psi^-) + \mu(\Gamma_U) = \mu^-_{CZ}(\Phi) + \mu(\Gamma_U). \tag{203}$$

Since

$$\Gamma_U = \begin{pmatrix} U & 0 \\ 0 & \mathrm{id} \end{pmatrix} \Delta$$

the Maslov index for the loop Γ_U computes as

(204)

$$\mu(\Gamma_U) = \deg\left(\det \begin{pmatrix} U & 0 \\ 0 & \mathrm{id} \end{pmatrix}^2\right) = \deg\big(\det(U)^2\big) = 2\deg\big(\det(U)\big) = 2c_1(E).$$

Combining (202), (203), and (204) the Lemma follows. $\square$

3. The normal Conley-Zehnder index

Assume that $\widetilde{u}\colon \mathbb{C} \to N \times \mathbb{R}$ is an *embedded* non-degenerate finite energy plane. Set

$$C := \widetilde{u}(\mathbb{C}) \subset N \times \mathbb{R}.$$

Since $\widetilde{u}$ is an embedding C is a two dimensional submanifold of $N \times \mathbb{R}$. Because $\widetilde{u}$ is holomorphic the tangent bundle TC is invariant under the SFT-like almost complex structure J. The fiber of the normal bundle

$$NC \to C$$

at a point $c \in C$ is given by $N_c C := (T_c C)^{\perp}$ where the orthogonal complement is taken with respect to the metric $\omega(\cdot, J\cdot)$. Because the almost complex structure J is ω-compatible the normal bundle NC is invariant under J as well. Because $\mathbb{C}$ is contractible we can choose a symplectic trivialization

$$\mathfrak{T}_N \colon \widetilde{u}^* NC \to \mathbb{C} \times \mathbb{C}.$$

Moreover, in view of the asymptotic behavior of $\widetilde{u}$ explained in Theorem 5.2 we can arrange that $\mathfrak{T}_N$ extends smoothly to a symplectic trivialization

$$\mathfrak{T}_N \colon \gamma^* \xi \to S^1 \times \mathbb{C}$$

where γ is the asymptotic Reeb orbit of $\widetilde{u}$. We define the *normal Conley-Zehnder index* of $\widetilde{u}$ as

$$\mu_{CZ}^N(\widetilde{u}) := \mu_{CZ}\Big(t \mapsto \mathfrak{T}_{N,t} d^{\xi} \phi_R^{t\tau}(\gamma(0)) \mathfrak{T}_{N,0}^{-1}\Big).$$

Recall from (148) that the usual Conley-Zehnder index for $\widetilde{u}$ is defined as follows. Choose a symplectic trivialization

$$\mathfrak{T}_{\xi} \colon \gamma^* \xi \to S^1 \times \mathbb{C}$$

which extends to a symplectic trivialization

$$\mathfrak{T}_{\xi} \colon \gamma^* \xi \to S^1 \times \mathbb{C}$$

and set

$$\mu_{CZ}(\widetilde{u}) = \mu_{CZ}\Big(t \mapsto \mathfrak{T}_{\xi,t} d^{\xi} \phi_R^{t\tau}(\gamma(0)) \mathfrak{T}_{\xi,0}^{-1}\Big).$$

Note that both the usual Conley-Zehnder index and the normal Conley-Zehnder index are independent of the trivializations chosen, because over the contractible space $\mathbb{C}$ all trivializations are homotopic. On the other hand although $\mathfrak{T}_N$ and $\mathfrak{T}_{\xi}$ trivialize $\gamma^* \xi$ there is no need that the two Conley-Zehnder indices agree, since over the circle two trivializations do not need to be homotopic, in view of the fact that the circle has a nontrivial fundamental group. The following theorem tells us how the two Conley-Zehnder indices are related. It is due to Hofer, Wysocki, and Zehnder, see [**57**].

Theorem 3.1. *Assume $\widetilde{u}$ is an embedded non-degenerate finite energy plane. Then*

$$\mu_{CZ}^N(\widetilde{u}) = \mu_{CZ}(\widetilde{u}) - 2.$$

As preparation for the proof of Theorem 3.1 we associate to an embedded non-degenerate finite energy plane $\widetilde{u}$ two complex vector bundles over the two dimensional sphere S^2

$$\ell_1 \to S^2, \quad \ell_2 \to S^2.$$

Abbreviate by

$$\eta := \langle \partial_r, R \rangle \subset T(N \times \mathbb{R})$$

the subbundle of $T(N \times \mathbb{R})$ spanned by the Reeb vector field R and the Liouville vector field ∂_r. Note that this gives rise to a complex splitting of vector bundles

$$T(\mathbb{R} \times N) = \eta \oplus \xi.$$

We decompose the sphere S^2 into upper and lower hemisphere $S^2 = S^2_+ \cup S^2_-$ as in the discussion about the first Chern class in (197). If

$$\mathring{S}^2_\pm = \{(x_1, x_2, x_3) \in S^2 : 0 < \pm x_3 \leq 1\}$$

denotes the interior of $S^2_\pm$ we get diffeomorphisms

$$\psi_\pm \colon \mathring{S}^2_\pm \to \mathbb{C}, \quad (x_1, x_2, x_3) \mapsto e^{2\pi \arctan \frac{\pi\sqrt{x_1^2+x_2^2}}{2}}(x_1 + ix_2).$$

Note that ψ_+ is orientation preserving, while ψ_- is orientation reversing. In view of the asymptotic behavior of $\widetilde{u}$ explained in Theorem 5.2 we get a complex vector bundle ℓ_1 over the sphere S^2 which is characterized by

$$\ell_1|_{\mathring{S}^2_+} = \psi_+^* \widetilde{u}^* TC = \psi_+^* T\mathbb{C}, \quad \ell_1|_{\mathring{S}^2_-} = \psi_-^* \widetilde{u}^* \eta.$$

Note that if we identify the equator $S = S^2_+ \cap S^2_-$ with the circle S^1 by mapping $t \in S^1$ to $(\cos 2\pi t, \sin 2\pi t, 0)$ we obtain

$$\ell_1|_{S^1} = \gamma^* \eta$$

where γ is the asymptotic Reeb orbit of the finite energy plane $\widetilde{u}$. Similarly, we define a complex line bundle ℓ_2 over the sphere S^2 by

$$\ell_2|_{\mathring{S}^2_+} = \psi_+^* \widetilde{u}^* NC, \quad \ell_2|_{\mathring{S}^2_-} = \psi_-^* \widetilde{u}^* \xi.$$

Over the equator this vector bundle satisfies

$$\ell_2|_{S^1} = \gamma^* \xi.$$

Lemma 3.2. *The first Chern classes of the two complex line bundles ℓ_1 and ℓ_2 over S^2 satisfy*

$$1 = c_1(\ell_1) = -c_1(\ell_2).$$

Proof: We define two more vector bundles ℓ_1' and ℓ_2' over the sphere S^2. The vector bundle ℓ_1' is characterized by the conditions

$$\ell_1'|_{\mathring{S}^2_+} = \psi_+^* \widetilde{u}^* \eta, \quad \ell_1'|_{\mathring{S}^2_-} = \psi_-^* \widetilde{u}^* \eta$$

and therefore satisfies over the equator

$$\ell_1'|_{S^1} = \gamma^* \eta.$$

The vector bundle ℓ_2' is determined by

$$\ell_2'|_{\mathring{S}^2_+} = \psi_+^* \widetilde{u}^* \xi, \quad \ell_1'|_{\mathring{S}^2_-} = \psi_-^* \widetilde{u}^* \xi$$

and therefore meets

$$\ell_2'|_{S^1} = \gamma^* \xi.$$

In view of

$$\widetilde{u}^* TC \oplus \widetilde{u}^* NC = \widetilde{u}^* T(N \times \mathbb{R}) = \widetilde{u}^* \xi \oplus \widetilde{u}^* \eta$$

we obtain

$$\ell_1 \oplus \ell_2 = \ell_1' \oplus \ell_2'.$$

Using the additivity of the first Chern class under Whitney sum (198) we get the formula

$$c_1(\ell_1)+c_1(\ell_2)=c_1(\ell_1\oplus\ell_2)=c_1(\ell_1')+c_1(\ell_2'). \tag{205}$$

We next claim that

$$c_1(\ell_1')=c_1(\ell_2')=0. \tag{206}$$

To see that $c_1(\ell_2')=0$ we first choose a nonvanishing section $\sigma\colon S^1\to\gamma^*\xi$. We extend σ to a transverse section $\sigma\colon\mathbb{C}\to\widetilde{u}^*\xi$. We define a section $s\in\Gamma(\ell_2')$ by

$$s(x)=\left\{\begin{array}{cc}\sigma(\psi_+(x)) & x\in\mathring{S}^2_+\\ \sigma(\psi_-(x)) & x\in\mathring{S}^2_-\\ \sigma(t) & t\in S^1.\end{array}\right.$$

By construction all of the zeros of s lie in $\mathring{S}^2_+\cup\mathring{S}^2_-$ and they appear in pairs. Namely to each zero x of s in $\mathring{S}^2_+$ corresponds the zero $\psi_-^{-1}\psi_+(x)\in\mathring{S}^2_-$ and vice versa. Because ψ_+ is orientation preserving and ψ_- is orientation reversing the signs of the two zeros cancel and therefore the Euler number of ℓ_2' vanishes. Because the first Chern class of a complex line bundle is just its Euler number we have shown that $c_1(\ell_2')=0$. The same argument proves that $c_1(\ell_1')=0$ as well. However, the vanishing of $c_1(\ell_1')$ can be understood even easier by noting that the Reeb and Liouville vector fields R and ∂_r give rise to a trivialization of ℓ_1'. Formula (206) is established.

Combining (205) and (206) we conclude that

$$c_1(\ell_1)=-c_1(\ell_2) \tag{207}$$

and it suffices to show that $c_1(\ell_1)=1$. To see that we construct again a section $s\in\Gamma(\ell_1)$ in order to compute the Euler number of ℓ_1. Since $\ell_1|_{\mathring{S}^2_-}=\psi_-^*\widetilde{u}^*\eta$ we define $s|_{S^2_-}$ as the pullback of the Liouville vector field ∂_r. Note that therefore $s|_{S^2_-}$ has no zeros. Using that $\widetilde{u}^*TC=T\mathbb{C}$ we have to find for the extension to the upper hemisphere a vector field on $\mathbb{C}$ which points outward asymptotically. For example the vector field $x\partial_x+y\partial_y$ meets this requirement. Note that this vector field has one positive zero at the origin. Since ψ_+ is orientation preserving we conclude that there exists a section $s\in\Gamma(\ell_1)$ which has precisely one positive zero. That means that the Euler number of ℓ_1 is one and therefore

$$c_1(\ell_1)=e(\ell_1)=1. \tag{208}$$

Equations (205) and (206) prove the lemma. □

Proof of Theorem 3.1: By construction of ℓ_2 we have by Lemma 2.1

$$\mu^N_{CZ}(\widetilde{u})-\mu_{CZ}(\widetilde{u})=2c_1(\ell_2).$$

The Theorem now follows from Lemma 3.2. □

4. An implicit function theorem

Assume that γ is a non-degenerate periodic Reeb orbit. Recall that by $\widehat{\mathcal{M}}(\gamma)$ we denote the moduli space of (parametrized) finite energy planes with unparametrized asymptotic orbit $[\gamma]$. The group of direct similitudes $\Sigma=\mathbb{C}^*\ltimes\mathbb{C}$ acts on $\widehat{\mathcal{M}}(\gamma)$

by reparametrization and its quotient $\mathcal{M}(\gamma) = \widehat{\mathcal{M}}(\gamma)/\Sigma$ is the moduli space of unparametrized finite energy planes asymptotic to $[\gamma]$. We denote by

$$\Pi\colon \widehat{\mathcal{M}}(\gamma) \to \mathcal{M}(\gamma), \quad \widetilde{u} \to [\widetilde{u}]$$

the projection. Suppose that $\widetilde{u} \in \widehat{\mathcal{M}}(\gamma)$ is embedded and set $C = \widetilde{u}(\mathbb{C})$. Choose a unitary trivialization $\mathfrak{T} = \mathfrak{T}_N \colon NC \to \mathbb{C} \times \mathbb{C}$, i.e., a trivialization which is complex with respect to the SFT-like almost complex structure J and orthogonal with respect to the inner product $\omega(\cdot, J\cdot)$ and such that it extends to a trivialization $\mathfrak{T}\colon \gamma^*\xi \to S^1 \times \mathbb{C}$. Define a smooth path

$$\Psi = \Psi^{\mathfrak{T}} \colon [0,1] \to \mathrm{Sp}(1)$$

by setting for $t \in [0,1]$

$$\Psi(t) = \mathfrak{T}_t d^{\xi}\phi_R^{t\tau}(\gamma(0))\mathfrak{T}_0^{-1}.$$

Note that $\Psi(0) = \mathrm{id}\colon \xi_{\gamma(0)} \to \xi_{\gamma(0)}$ and Ψ is non-degenerate in the sense that $\ker(\Psi(1) - \mathrm{id}) \neq \{0\}$ because γ is non-degenerate. Let

$$S_\infty = S_\infty^{\mathfrak{T}} \in C^\infty([0,1], \mathrm{Sym}(2))$$

generate the path Ψ by

$$\partial_t \Psi = J_0 S_\infty \Psi.$$

In [**57**] Hofer, Wysocki, and Zehnder construct a smooth map

$$S = S_{\widetilde{u}}^{\mathfrak{T}} \colon \mathbb{C} \to M_2(\mathbb{R}) \tag{209}$$

which has the property that $2\pi e^{2\pi s} S(e^{2\pi(s+it)})$ uniformly converges in the C^∞-topology to $S_\infty(t)$, when s goes to infinity. Recall the Hilbert spaces H_1 and H_0 from (184) respectively (185). Then the map S gives rise to a bounded linear operator

$$L_S \colon H_1 \to H_0$$

as defined in (186). One should think of the operator L_S as the linearization of the holomorphic curve equation modulo the action of the group of direct similitudes Σ by reparametrization. That we mod out the reparametrization action is due to the fact that we only consider variations in the normal direction. The importance of the operator L_S lies in the following implicit function theorem proved by Hofer, Wysocki, and Zehnder in [**57**].

Theorem 4.1. *Assume that $\widetilde{u} \in \widehat{\mathcal{M}}(\gamma)$ is embedded and the operator $L_S\colon H_1 \to H_0$ is surjective. Then locally around $[\widetilde{u}]$ the space $\mathcal{M}(\gamma)$ is a manifold and there exists $\mathcal{U} \subset \ker(L_S)$ an open neighborhood of 0 and a map*

$$\widehat{\mathcal{F}} \colon \mathcal{U} \to \widehat{\mathcal{M}}(\gamma)$$

satisfying $\widehat{\mathcal{F}}(0) = \widetilde{u}$, such that the map

$$\mathcal{F} = \Pi\widehat{\mathcal{F}} \colon \mathcal{U} \to \mathcal{M}(\gamma)$$

defines a local chart around $[\widehat{u}]$. Moreover, if $\zeta \in \mathcal{U}$ the intersection points of $\widetilde{u}$ and $\widehat{\mathcal{F}}(\zeta)$ are related by

$$\big\{(z,z') \in \mathbb{C} \times \mathbb{C} : \widetilde{u}(z) = \widehat{\mathcal{F}}(\zeta)(z')\big\} = \big\{(z,z) : z \in \mathbb{C},\ \zeta(z) = 0\big\}.$$

Using Theorem 1.2 and Theorem 3.1 we can under the assumptions of Theorem 4.1 compute the local dimension of the moduli space $\mathcal{M}(\gamma)$ at $[\widetilde{u}]$ by

$$\begin{aligned}\dim_{[\widetilde{u}]}\mathcal{M}(\gamma) &= \dim\ker L_S \\ &= \operatorname{ind} L_S = \mu_{CZ}(\Psi_{S_\infty}) + 1 \\ &= \mu^N_{CZ}(\widetilde{u}) + 1 \\ &= \mu_{CZ}(\widetilde{u}) - 1.\end{aligned} \tag{210}$$

5. Exponential weights

Suppose that γ is a non-degenerate Reeb orbit. Abbreviate by

$$\widehat{\mathcal{M}}_{\text{fast}}(\gamma) \subset \widehat{\mathcal{M}}(\gamma)$$

the subspace of fast finite energy planes asymptotic to γ. Assume that $\widetilde{u} = (u, a) \in \widehat{\mathcal{M}}_{\text{fast}}(\gamma)$ is embedded. Recall that if $U(s,t) \in \xi_{\gamma(t)}$ is an asymptotic representative of $\widetilde{u}$ we can write

$$U(s,t) = e^{\eta s}\big(\zeta(t) + \kappa(s,t)\big)$$

where κ decays with all its derivatives exponentially, $\eta = \eta_u \in \mathfrak{S}(A_\gamma) \cap (-\infty, 0)$ is a negative eigenvalue of the asymptotic operator and $\zeta = \zeta_{\widetilde{u}} \in \Gamma(\gamma^*\xi)$ is an eigenvector of the asymptotic operator A_γ to the eigenvalue η. Pick further a unitary trivialization

$$\mathfrak{T}_\xi \colon u^*\xi \to \mathbb{C} \times \mathbb{C}$$

which extends to a trivialization

$$\mathfrak{T}_\xi \colon \gamma^*\xi \to S^1 \times \mathbb{C}.$$

In particular, $\mathfrak{T}_\xi\zeta \in C^\infty(S^1, \mathbb{C})$ and the winding of the eigenvalue η is defined as

$$w(\eta) = w(\mathfrak{T}_\xi\zeta) = \deg\left(t \mapsto \frac{\mathfrak{T}_\xi\zeta(t)}{|\mathfrak{T}_\xi\zeta(t)|}\right).$$

Because $\widetilde{u}$ is fast we have by Corollary 0.13

$$w(\eta) = 1.$$

Recall further from Corollary 0.32 that the winding is monotone, i.e., if $\eta \leq \eta'$, then $w(\eta) \leq w(\eta')$. If $\mu_{CZ}(u) \geq 3$ choose $\delta \leq 0$ such that

$$\max\big\{\eta \in \mathfrak{S}(A_\gamma) : w(\eta) = 1\big\} < \delta < \min\big\{\eta \in \mathfrak{S}(A_\gamma) : w(\eta) = 2\big\}.$$

That a non-positive δ with this property exists is guaranteed by Theorem 0.33 in view of the assumption that $\mu_{CZ}(u) \geq 3$. If $\mu_{CZ}(u) \leq 2$ it follows from Theorem 0.5 that $\mu_{CZ}(u) = 2$ and we set in this case $\delta = 0$. Because $\widetilde{u}$ is embedded we set $C = \widetilde{u}(\mathbb{C}) \in N \times \mathbb{R}$ and choose in addition a unitary trivialization

$$\mathfrak{T}_N \colon u^*NC \to \mathbb{C} \times \mathbb{C}$$

which extends to a trivialization

$$\mathfrak{T}_N \colon \gamma^*\xi \to S^1 \times \mathbb{C}.$$

This trivialization gives rise to a smooth map

$$S = S^{\mathfrak{T}_N}_{\widetilde{u}} \colon \mathbb{C} \to M_2(\mathbb{R})$$

as mentioned in (209). Recall that $\phi\colon \mathbb{R}\times S^1 \to \mathbb{C}\setminus\{0\}$ denotes the biholomorphism $(s,t)\mapsto e^{2\pi(s+it)}$. For $\delta \leq 0$ we introduce the Hilbert space

$$H_1^\delta = \Big\{\zeta \in H_1 : e^{-\delta s}\zeta\circ\phi|_{[0,\infty)\times S^1} \in W^{1,2}\big([0,\infty)\times S^1,\mathbb{C}\big)\Big\},$$

i.e., the subvector space of functions in H_1 introduced in (184) which decay on the cylindrical end exponentially with weight $|\delta|$. Similarly, we introduce

$$H_0^\delta = \Big\{\zeta \in H_0 : e^{-\delta s}\zeta\circ\phi|_{[0,\infty)\times S^1} \in L^2\big([0,\infty)\times S^1,\mathbb{C}\big)\Big\}.$$

We define the bounded linear operator

$$L_S^\delta\colon H_1^\delta \to H_0^\delta,\quad \zeta\mapsto \bar\partial_\mu\zeta + \frac{1}{\sqrt{h_\gamma}}S\zeta.$$

Note that L_S^δ is given by the same formula as the operator L_S introduced in (186), however, its domain and target differ form the ones of L_S. This fact can be expressed with the following commutative diagram

$$\begin{tikzcd} H_1^\delta \arrow[r, "L_S^\delta"] \arrow[d] & H_0^\delta \arrow[d] \\ H_1 \arrow[r, "L_S"] & H_0 \end{tikzcd}$$

where the vertical arrows stand for the inclusion maps.

The analogue of the implicit function Theorem (Theorem 4.1) in the fast case can now be formulated as follows.

Theorem 5.1. *Assume that $\widetilde{u}\in\widehat{\mathcal{M}}_{\mathrm{fast}}(\gamma)$ is embedded and the operator $L_S^\delta\colon H_1^\delta\to H_0^\delta$ is surjective. Then locally around $[\widetilde{u}]$ the space $\mathcal{M}_{\mathrm{fast}}(\gamma)$ is a manifold and there exists $\mathcal{U}\subset\ker(L_S^\delta)$ an open neighborhood of 0 and a map*

$$\widehat{\mathcal{F}}\colon\mathcal{U}\to\widehat{\mathcal{M}}_{\mathrm{fast}}(\gamma)$$

satisfying $\widehat{\mathcal{F}}(0)=\widetilde{u}$, such that the map

$$\mathcal{F}=\Pi\widehat{\mathcal{F}}\colon\mathcal{U}\to\mathcal{M}_{\mathrm{fast}}(\gamma)$$

defines a local chart around $[\widehat{u}]$. Moreover, if $\zeta\in\mathcal{U}$ the intersection points of $\widetilde{u}$ and $\widehat{\mathcal{F}}(\zeta)$ are related by

$$\big\{(z,z')\in\mathbb{C}\times\mathbb{C}:\widetilde{u}(z)=\widehat{\mathcal{F}}(\zeta)(z')\big\}=\big\{(z,z):z\in\mathbb{C},\ \zeta(z)=0\big\}.$$

Note that if $\widetilde{u}$ is an embedded fast finite energy plane and both $L_S\colon H_1\to H_0$ and $L_S^\delta\colon H_1^\delta\to H_0^\delta$ are surjective, then locally around $[\widetilde{u}]$ both $\mathcal{M}(\gamma)$ and $\mathcal{M}_{\mathrm{fast}}(\gamma)$ are smooth manifolds and we have

$$\ker(L_S^\delta)=T_{[\widetilde{u}]}\mathcal{M}_{\mathrm{fast}}(\gamma)\subset T_{[\widetilde{u}]}\mathcal{M}(\gamma)=\ker(L_S).$$

Our next goal it to compute the Fredholm index of the operator L_S^δ. The following result is due to Hryniewicz [**58**].

Theorem 5.2. *Assume that $\widetilde{u}=(u,a)$ is an embedded fast finite energy plane. Then the Fredholm index of the Fredholm operator L_S^δ satisfies*

$$\mathrm{ind}L_S^\delta=\begin{cases}2 & \mu_{CZ}(u)\geq 3\\ 1 & \mu_{CZ}(u)=2.\end{cases}$$

Proof: If $\mu_{CZ}(u)=2$ we have chosen $\delta=0$ and hence in this case we have $L_S^\delta = L_S$. Hence by Theorem 1.2 and Theorem 3.1 we obtain

$$\mathrm{ind} L_S^\delta = \mu_{CZ}(\Psi_{S_\infty})+1=\mu_{CZ}^N(u)+1=\mu_{CZ}(u)-1=1.$$

Therefore we can assume in the following that

$$\mu_{CZ}(u)\geq 3.$$

In order to compute the Fredholm index of the operator L_S^δ in this case we first choose a smooth function $\rho\colon \mathbb{R}\to\mathbb{R}$ satisfying

$$\rho(s)=\left\{\begin{array}{ll} s & s\geq 1\\ 0 & s\leq 0.\end{array}\right.$$

We define a Hilbert space isomorphism

$$T_\delta\colon H_1^\delta \to H_1$$

which is given for $\zeta\in H_1^\delta$ by

$$T_\delta\zeta(e^{2\pi(s+it)})=e^{-\delta\rho(s)}\zeta(e^{2\pi(s+it)}),\ (s,t)\in\mathbb{R}\times S^1,\quad T_\delta\zeta(0)=\zeta(0).$$

The isomorphism T_δ extends by the same formula to an isomorphism

$$T_\delta\colon H_0^\delta\to H_0.$$

Note that its inverse is given by

$$T_\delta^{-1}=T_{-\delta}\colon H_i\to H_i^\delta,\quad i\in\{0,1\}.$$

We consider

$$T_\delta L_S^\delta T_{-\delta}\colon H_1\to H_0.$$

To describe this map we introduce $S_\delta\in C^\infty(\mathbb{C},M_2(\mathbb{R}))$ by

$$S_\delta(e^{2\pi(s+it)}):=S(e^{2\pi(s+it)})+\frac{\delta\rho'(s)}{2\pi e^{2\pi s}}\cdot\mathrm{id},\quad (s,t)\in\mathbb{R}\times S^1.$$

If $\zeta\in H_1$ we compute

$$\begin{aligned}
& T_\delta L_S^\delta T_{-\delta}\zeta(e^{2\pi(s+it)}) = T_\delta L_S^\delta\big(e^{\delta\rho(s)}\zeta(e^{2\pi(s+it)})\big)\\
= \;& T_\delta\frac{1}{\sqrt{\gamma(e^{2\pi s})}}\big(\partial_x+i\partial_y+S\big)\big(e^{\delta\rho(s)}\zeta(e^{2\pi(s+it)})\big)\\
= \;& T_\delta\frac{1}{\sqrt{\gamma(e^{2\pi s})}}\left(\frac{1}{2\pi e^{2\pi s}}(\partial_s+i\partial_t)+S\right)\big(e^{\delta\rho(s)}\zeta(e^{2\pi(s+it)})\big)\\
= \;& T_\delta\frac{e^{\delta\rho(s)}}{\sqrt{\gamma(e^{2\pi s})}}\left(\frac{1}{2\pi e^{2\pi s}}(\partial_s+i\partial_t)+S+\frac{\delta\rho'(s)}{2\pi e^{2\pi s}}\right)\zeta(e^{2\pi(s+it)})\\
= \;& \frac{1}{\sqrt{\gamma(e^{2\pi s})}}\left(\frac{1}{2\pi e^{2\pi s}}(\partial_s+i\partial_t)+S+\frac{\delta\rho'(s)}{2\pi e^{2\pi s}}\right)\zeta(e^{2\pi(s+it)})\\
= \;& \frac{1}{\sqrt{\gamma(e^{2\pi s})}}\left(\partial_x+i\partial_y+S+\frac{\delta\rho'(s)}{2\pi e^{2\pi s}}\right)\zeta(e^{2\pi(s+it)})\\
= \;& \frac{1}{\sqrt{\gamma(e^{2\pi s})}}\big(\partial_x+i\partial_y+S_\delta\big)\zeta(e^{2\pi(s+it)})\\
= \;& L_{S_\delta}\zeta(e^{2\pi(s+it)}).
\end{aligned}$$

We showed

$$T_\delta L_S^\delta T_{-\delta}=L_{S_\delta}.$$

Note that

$$\begin{aligned}\lim_{s\to\infty} 2\pi e^{2\pi s} S_\delta(e^{2\pi(s+it)}) &= \lim_{s\to\infty} 2\pi e^{2\pi s} S(e^{2\pi(s+it)}) + \delta\cdot \mathrm{id}\\ &= S_\infty(t) + \delta \cdot \mathrm{id}\\ &=: S_\infty^\delta(t).\end{aligned}$$

From Theorem 1.2 we obtain

$$\mathrm{ind}(L_S^\delta) = \mathrm{ind}(L_{S_\delta}) = \mu_{CZ}(\Psi_{S_\infty^\delta}) + 1. \tag{211}$$

It remains to compute the Conley-Zehnder index of the symplectic path $\Psi_{S_\infty^\delta}$. Recall that the linear operator $A_{S_\infty} = -J\partial_t - S_\infty \colon W^{1,2}(S^1,\mathbb{C}) \to L^2(S^1,\mathbb{C})$ equals

$$A_{S_\infty} = A_\gamma^{\mathfrak{T}_N} =: A^N$$

where the operator $A_\gamma^{\mathfrak{T}_N}$ as explained in (146) is conjugated to the asymptotic operator $A_\gamma \colon \Gamma^{1,2}(\gamma^*\xi) \to \Gamma^{0,2}(\gamma^*\xi)$ via the trivialization $\mathfrak{T}_N$. Because the winding numbers of the eigenvalues of A_γ are computed with respect to the trivialization $\mathfrak{T}_\xi$ we also need to consider the operator

$$A^\xi := A_\gamma^{\mathfrak{T}_\xi}.$$

Since both operators A^N and A^ξ are conjugated to A_γ they are conjugated to each other as well. To see how A^N and A^ξ are conjugated we consider the loop of unitary transformations $t \mapsto \mathfrak{T}_{N,t}\mathfrak{T}_{\xi,t}^{-1}$ from $\mathbb{C}$ to itself. Because a unitary transformation of $\mathbb{C}$ is just multiplication by a complex number of norm one we can think of $\mathfrak{T}_{N,t}\mathfrak{T}_{\xi,t}^{-1}$ for each $t \in S^1$ as a unit complex number. Recall from Lemma 3.2 that $c_1(\ell_2) = -1$ which implies that

$$\deg\Big(t \mapsto \mathfrak{T}_{N,t}\mathfrak{T}_{\xi,t}^{-1}\Big) = -1.$$

Therefore by replacing the trivializations $\mathfrak{T}_N$ and $\mathfrak{T}_\xi$ by homotopic ones we can assume without loss of generality that

$$\mathfrak{T}_{N,t}\mathfrak{T}_{\xi,t}^{-1} = e^{-2\pi i t}.$$

Consider the Hilbert space isomorphism

$$\Phi \colon W^{1,2}(S^1,\mathbb{C}) \to W^{1,2}(S^1,\mathbb{C})$$

which is given for $v \in W^{1,2}(S^1,\mathbb{C})$ by

$$\Phi(v)(t) = e^{2\pi i t} v(t), \quad t \in S^1$$

with inverse

$$\Phi^{-1} \colon W^{1,2}(S^1,\mathbb{C}) \to W^{1,2}(S^1,\mathbb{C}), \quad \Phi^{-1}v(t) = e^{-2\pi i t}v(t),\ t \in S^1.$$

Note that both Φ and Φ^{-1} extend to Hilbert space isomorphisms

$$\Phi, \Phi^{-1} \colon L^2(S^1,\mathbb{C}) \to L^2(S^1,\mathbb{C}).$$

Note that if $v \in W^{1,2}(S^1,\mathbb{C})$ we have

$$A^N v = \mathfrak{T}_N A_\gamma \mathfrak{T}_N^{-1} v = \mathfrak{T}_N \mathfrak{T}_\xi^{-1} A^\xi \mathfrak{T}_\xi \mathfrak{T}_N^{-1} v = \mathfrak{T}_N \mathfrak{T}_\xi^{-1} A^\xi (\mathfrak{T}_N \mathfrak{T}_\xi)^{-1} v$$

so that we obtain

$$A^N = \Phi^{-1} A^\xi \Phi.$$

Now if $\eta \in \mathfrak{S}(A^\xi)$ is an eigenvalue of A^ξ and v is an eigenvector of A^ξ to the eigenvalue η it follows that $\Phi^{-1}v$ is an eigenvector of A^N to the eigenvalue η. However, note that the winding number changes under this transformation, namely

$$w(\Phi^{-1}v) = w(v) - 1.$$

Therefore although the spectra of the conjugated operators A^ξ and A^N agree if $\eta \in \mathfrak{S}(A^\xi) = \mathfrak{S}(A^N)$ the corresponding winding numbers differ by

$$w(\eta, A^N) = w(\eta, A^\xi) - 1.$$

Because $S_\infty^\delta = S_\infty + \delta$ we have

$$A_{S_\infty^\delta} = A_{S_\infty} - \delta I = A^N - \delta I =: A_\delta^N$$

where $I \colon W^{1,2}(S^1, \mathbb{C}) \to L^2(S^1, \mathbb{C})$ is the inclusion. We further abbreviate

$$A_\delta^\xi := A^\xi - \delta I.$$

Note that

$$A_\delta^N = \Phi^{-1} A_\delta^\xi \Phi.$$

Therefore

$$\begin{aligned} \alpha(S_\infty^\delta) &= \max\Big\{w(\eta, A_\delta^N) : \eta \in \mathfrak{S}(A_\delta^N) \cap (-\infty, 0)\Big\} \\ &= \max\Big\{w(\eta, A_\delta^\xi) : \eta \in \mathfrak{S}(A_\delta^\xi) \cap (-\infty, 0)\Big\} - 1. \end{aligned} \tag{212}$$

Note that we have a bijection

$$\mathfrak{S}(A^\xi) \cong \mathfrak{S}(A_\delta^\xi), \quad \eta \mapsto \eta - \delta.$$

Indeed, v is an eigenvector of the operator A^ξ to the eigenvalue η if and only if v is an eigenvector of the operator A_δ^ξ to the eigenvalue $\eta - \delta$. In particular,

$$w(\eta, A^\xi) = w(\eta - \delta, A_\delta^\xi).$$

Hence

$$\begin{aligned} &\max\Big\{w(\eta, A_\delta^\xi) : \eta \in \mathfrak{S}(A_\delta^\xi) \cap (-\infty, 0)\Big\} \\ = \ &\max\Big\{w(\eta + \delta, A^\xi) : \eta \in \mathfrak{S}(A_\delta^\xi), \eta < 0\Big\} \\ = \ &\max\Big\{w(\eta, A^\xi) : \eta \in \mathfrak{S}(A_\delta^\xi), \eta < \delta\Big\} \\ = \ &1 \end{aligned} \tag{213}$$

by the choice of δ. Combining (212) with (213) we obtain

$$\alpha(S_\infty^\delta) = 1 - 1 = 0. \tag{214}$$

Recall that the parity is defined as

$$p(S_\infty^\delta) = \begin{cases} 0 & \exists\, \eta \in \mathfrak{S}(A_\delta^N) \cap [0, \infty) \text{ such that } \alpha(S_\infty^\delta) = w(\eta, S_\infty^\delta) \\ 1 & \text{else.} \end{cases}$$

By the choice of δ we have

$$p(S_\infty^\delta) = 1. \tag{215}$$

Combining (214) and (215) with Theorem 0.33 we obtain

$$\mu_{CZ}(\Psi_{S_\infty^\delta}) = 2\alpha(S_\infty^\delta) + p(S_\infty^\delta) = 1. \tag{216}$$

In combination with (211) equation (216) implies

$$\mathrm{ind}(L_S^\delta) = 2.$$

This finishes the proof of the theorem. □

6. Automatic transversality

We first explain the following local version of automatic transversality for fast finite energy planes.

Lemma 6.1. *Assume that $\widetilde{u} = (u, a)$ is an embedded fast finite energy plane with asymptotic orbit γ satisfying $\mu_{CZ}(u) \geq 3$. Then locally around $[\widetilde{u}]$ the moduli space $\mathcal{M}_{\mathrm{fast}}(\gamma)$ is a two dimensional manifold.*

To prove Lemma 6.1 we need the following result, see also [**58**].

Lemma 6.2. *Assume that $\mu_{CZ}(\Psi_{S_\infty}) = 1$, then L_S is surjective.*

Proof: Because $\mu_{CZ}(\Psi_{S_\infty}) = 1$ we obtain from Theorem 1.2 that

$$\mathrm{ind}(L_S) = 2.$$

Hence it suffices to show that

$$\dim \ker L_S \leq 2.$$

Suppose that $v \neq 0 \in \ker L_S$. In view of the asymptotic behavior there exists $R_0 > 0$ such that for every $R \geq R_0$

$$\deg\left(t \mapsto \frac{v(Re^{2\pi i t})}{|v(Re^{2\pi i t})|}\right) = w(\eta)$$

where $w(\eta)$ is the winding of a negative eigenvalue $\eta \in \mathfrak{S}(A_{S_\infty}) \cap (-\infty, 0)$. By Theorem 0.33 we have

$$1 = \mu_{CZ}(\Psi_{S_\infty}) = 2\alpha(S_\infty) + p(S_\infty)$$

where the parity satisfies $p(S_\infty) \in \{0, 1\}$. Therefore

$$0 = \alpha(S_\infty) = \max\{w(\eta) : \eta \in \mathfrak{S}(A_{S_\infty}) \cap (-\infty, 0)\}.$$

Therefore

$$\deg\left(t \mapsto \frac{v(Re^{2\pi i t})}{|v(Re^{2\pi i t})|}\right) \leq 0.$$

On the other hand since $v \in \ker L_S$ it follows form Carleman's similarity principle in Lemma 0.8 that all local winding numbers of the map $v \colon \mathbb{C} \to \mathbb{C}$ are positive so that

$$\deg\left(t \mapsto \frac{v(Re^{2\pi i t})}{|v(Re^{2\pi i t})|}\right) \geq 0$$

and equality holds if and only if $v(z) \neq 0$ for every $z \in \mathbb{C}$. We conclude that if $v \in \ker L_S$ does not vanish identically we necessarily have

$$v(z) \neq 0, \quad \forall\, z \in \mathbb{C}.$$

Now suppose that $v_1, v_2, v_3 \in \ker L_S$. It remains to show that the set $\{v_1, v_2, v_3\}$ is linearly dependent. Pick $z \in \mathbb{C}$. Then $v_1(z), v_2(z), v_3(z) \in \mathbb{C}$ and therefore there exist $a_1, a_2, a_3 \in \mathbb{R}$ such that

$$a_1 v_1(z) + a_2 v_2(z) + a_3 v_3(z) = 0.$$

Because $a_1v_1 + a_2v_2 + a_3v_3 \in \ker L_S$ we conclude that

$$a_1v_1 + a_2v_2 + a_3v_3 = 0.$$

This proves that $\{v_1, v_2, v_3\}$ is linearly dependent and hence $\dim\ker L_S \leq 2$. This finishes the proof of the lemma. $\square$

Proof of Lemma 6.1: By Theorem 5.1 and Theorem 5.2 it suffices to show that the operator $L_S^\delta \colon H_1^\delta \to H_0^\delta$ is surjective. By (216) we have $\mu_{CZ}(\Psi_{S_\infty^\delta}) = 1$ and hence the Lemma follows from Lemma 6.2. $\square$

The proof of Lemma 6.2 actually reveals more. Namely, if $\zeta \neq 0 \in \ker L_S^\delta$ it follows that $\zeta(z) \neq 0$ for every $z \in \mathbb{C}$. Therefore, if $\Phi \colon \mathcal{U} \to \mathcal{M}_{\mathrm{fast}}(\gamma)$ is the local chart from Theorem 5.1 the assertion of the Theorem implies that

$$\{(z, z') \in \mathbb{C} \times \mathbb{C} : \widetilde{u}(z) = \Phi(\zeta)(z')\} = \{(z, z) : z \in \mathbb{C} : \zeta(z) = 0\} = \emptyset.$$

Therefore the algebraic intersection number of $\widetilde{u}$ with $\Phi(\zeta)$ satisfies

$$\mathrm{int}(\widetilde{u}, \Phi(\zeta)) = 0.$$

Hence in view of Siefring's inequality (Theorem 4.1)

$$\mathrm{sief}(\widetilde{u}, \Phi(\zeta)) = 0.$$

Because Siefring's intersection number is a homotopy invariant we proved the following result.

Lemma 6.3. *Assume that $\widetilde{u}$ is an embedded fast finite energy plane. Then its Siefring self-intersection number satisfies*

$$\mathrm{sief}(\widetilde{u}, \widetilde{u}) = 0.$$

We mention that Lemma 6.3 was already used to prove Theorem 5.4. In combination with the local automatic transversality result obtained in Lemma 6.1 we are now in position to prove the following global automatic transversality statement.

Theorem 6.4. *Assume that (N, λ) is a closed three dimensional contact manifold satisfying $\pi_2(N) = \{0\}$ whose symplectization is endowed with an SFT-like almost complex structure. Suppose further that γ is a non-degenerate periodic Reeb orbit with the property that there exists $[\widetilde{u}] \in \mathcal{M}_{\mathrm{fast}}(\gamma)$ such that $\widetilde{u} = (u, a)$ is embedded and $\mu_{CZ}(u) \geq 3$. Then $\mathcal{M}_{\mathrm{fast}}(\gamma)$ is a two dimensional manifold.*

Proof: Suppose that $[\widetilde{v}] \in \mathcal{M}_{\mathrm{fast}}(\gamma)$. Because $\pi_2(N) = \{0\}$ the fast finite energy plane $\widetilde{v} = (v, b)$ is homotopic to $\widetilde{u}$. In particular,

$$\mu_{CZ}(v) = \mu_{CZ}(u) \geq 3$$

and

$$\mathrm{sief}(\widetilde{v}, \widetilde{u}) = \mathrm{sief}(\widetilde{u}, \widetilde{u}).$$

Because $\widetilde{u}$ is embedded it follows from Theorem 5.4 that $\widetilde{v}$ is embedded as well. In particular, we can apply the local automatic transversality result Lemma 6.1 to conclude that locally around $[\widetilde{v}]$ the moduli space $\mathcal{M}_{\mathrm{fast}}(\gamma)$ is a smooth two dimensional manifold. Because $[\widetilde{v}]$ was an arbitrary point in $\mathcal{M}_{\mathrm{fast}}(\gamma)$ the theorem is proved. $\square$

CHAPTER 15

Compactness

1. Negatively punctured finite energy planes

Assume that (N, λ) is a closed three dimensional contact manifold. A *punctured holomorphic plane* is a smooth map

$$\widetilde{u} = (u, a) \colon \mathbb{C} \setminus P \to N \times \mathbb{R}$$

where $P \subset \mathbb{C}$ is a finite subset, such that $\widetilde{u}$ satisfies the nonlinear Cauchy Riemann equation (141) at every point $z \in \mathbb{C} \setminus P$. In particular, if $P = \emptyset$ the map $\widetilde{u}$ is just a holomorphic plane. As in the unpunctured case the energy of a punctured holomorphic plane is defined as

$$E(\widetilde{u}) := \sup_{\phi \in \Gamma} \int_{\mathbb{C}} \widetilde{u}^* d\lambda^{\phi}.$$

A puncture $p \in P$ is called *removable* if there exists an open neighborhood of p in $\mathbb{C}$ such that the restriction of a to this neighborhood is bounded. The reason for this terminology comes from the fact that if the energy of $\widetilde{u}$ is finite, then by the Theorem on removal of singularities [**45, 81**] the map $\widetilde{u}$ can be smoothly extended over a removable puncture. If a puncture $p \in P$ is not removable, then it is called a *non-removable puncture*.

If one thinks of $S^2 = \mathbb{C} \cup \{\infty\}$ then one can interpret a punctured holomorphic plane $\widetilde{u}$ as a map

$$\widetilde{u} \colon S^2 \setminus (P \cup \{\infty\}) \to N \times \mathbb{R}.$$

In particular, we might think of $\widetilde{u}$ as a *punctured holomorphic sphere* with the point $\{\infty\} \in S^2$ an additional puncture.

Assume that $\widetilde{u} = (u, a) \colon \mathbb{C} \setminus P \to N \times \mathbb{R}$ is a punctured holomorphic plane whose energy $E(\widetilde{u})$ is finite and all of whose punctures are non-removable. It is shown in [**54**] that for each puncture $p \in P$ there exists an open neighborhood U of p such that the restriction of a to U is either bounded from below or above. Of course, since p is assumed to be a non-removable puncture in the first case a is unbounded from above and in the second case a is unbounded from below. This fact allows us to write the set of non-removable punctures of a holomorphic plane of finite energy as a disjoint union

$$P = P_+ \cup P_-$$

where P_+ is the subset of punctures on which a remains bounded from below in a small neighborhood and $P_- = P \setminus P_+$ is the subset of punctures on which a remains bounded from above in a small neighborhood. Elements $p \in P_+$ are called *positive punctures*, while elements $p \in P_-$ are called *negative punctures*. Interpreting $\widetilde{u}$ as a punctured holomorphic sphere the same classification applies to the point at infinity, so that $\{\infty\}$ is either a removable, positive or negative puncture.

A special instance of a punctured holomorphic plane of finite energy is an *orbit cylinder*. Namely if γ is a periodic Reeb orbit of period τ define $\widetilde{\gamma}\colon \mathbb{C}\setminus\{0\} \to N\times\mathbb{R}$

$$\widetilde{\gamma}(e^{2\pi(s+it)}) = (\gamma(t), \tau s), \quad (s,t) \in \mathbb{R}\times S^1.$$

We are now in position to define

Definition 1.1. *A* negatively punctured finite energy plane $\widetilde{u}\colon \mathbb{C}\setminus P \to N\times\mathbb{R}$ *is a punctured holomorphic plane satisfying*

- **(i):** $0 < E(\widetilde{u}) < \infty$,
- **(ii):** *All punctures $p \in P$ are non-removable and negative.*
- **(iii):** *If one interprets $\widetilde{u}$ as a punctured holomorphic sphere, the point at infinity becomes an additional positive puncture.*
- **(iv):** *$\widetilde{u}$ is not a reparametrization of an orbit cylinder, i.e., there does not exist $(\rho,\tau) \in \Sigma = \mathbb{C}^* \ltimes \mathbb{C}$ and a periodic Reeb orbit γ such that $(\rho,\tau)_*\widetilde{u} = \widetilde{\gamma}$.*

Remark 1.2. *We mention that condition (iii) in the Definition of a negatively punctured finite energy plane follows from condition (ii) and the maximum principle for a holomorphic curve $\widetilde{u} = (u,a)$. The maximum principle says that the function a does not attain a local maximum. Indeed, using that the almost complex structure J is SFT-like one shows using the nonlinear Cauchy Riemann equation that the Laplacian of a satisfies*

$$\Delta a \geq 0$$

which establishes the maximum principle.

2. Weak SFT-compactness

Assume that $\widetilde{u} = (u,a)\colon \mathbb{C}\setminus P \to N\times\mathbb{R}$ is a negatively punctured finite energy plane. Hofer's theorem (Theorem 2.4) can still be applied in the case where $\widetilde{u}$ is punctured and we conclude that there exists a periodic Reeb orbit γ and a sequence s_k going to infinity such that

$$\lim_{k\to\infty} u(e^{2\pi(s_k+it)}) = \gamma(t)$$

uniformly in the C^∞-topology. If γ is non-degenerate we refer to the negatively punctured finite energy plane as a *non-degenerate negatively punctured finite energy plane*. If $\widetilde{u}$ is a non-degenerate negatively punctured finite energy plane it still admits an asymptotic representative. This asymptotic representative U can either be chosen to decay exponentially like in (147) or to vanish identically. In particular, we can associate to a non-degenerate negatively punctured finite energy plane $\widetilde{u} = (u,a)$ with asymptotic period orbit γ an element

$$\eta_{\widetilde{u}} = \eta_u \in [-\infty, 0)$$

where either η_u is a negative eigenvalue of the asymptotic operator A_γ such that the asymptotic representative decays exponentially with weight η_u or $\eta_u = -\infty$ and the asymptotic representative vanishes identically.

Theorem 2.1 (Weak SFT-compactness). *Assume that γ is a non-degenerate Reeb orbit and $\widetilde{u}_\nu = (u_\nu, a_\nu) \in \widehat{\mathcal{M}}(\gamma)$ for $\nu\in\mathbb{N}$ a sequence of finite energy planes with asymptotic Reeb orbit γ. Then there exists a subsequence ν_j, a sequence of gauge transformations $(r_j, (\rho_j,\tau_j)) \in \mathbb{R}\times\Sigma$ and a negatively punctured finite energy* *plane* $\widetilde{u} = (u, a)\colon \mathbb{C} \setminus P \to N \times \mathbb{R}$ *with positive asymptotic orbit* γ *as well such that* $(r_j, (\rho_j, \tau_j))_* \widetilde{u}_{\nu_j}$ *converges in the* C^∞_{loc}*-topology to* $\widetilde{u}$. *Moreover,*

$$\eta_u \leq \eta_{u_{\nu_j}}, \quad j \in \mathbb{N}. \tag{217}$$

This result is a special case of the SFT-compactness theorem, see [**19, 52, 55**]. We make the following remarks.

Remark 2.2. *The fact that* $\widetilde{u}_\nu \in \widehat{\mathcal{M}}(\gamma)$ *implies that*

$$E(\widetilde{u}_\nu) = \tau$$

where τ *is the period of the periodic orbit* γ. *In particular, the energy of the sequence* $\widetilde{u}_\nu$ *is constant and therefore uniformly bounded.*

Remark 2.3. *That in the limit no positive punctures occur follows from the maximum principle mentioned in Remark 1.2.*

In view of Theorem 2.1 in order to show that the moduli space $\widehat{\mathcal{M}}(\gamma)/\mathbb{R} \times \Sigma = \mathcal{M}(\gamma)/\mathbb{R}$ is compact it suffices to show that the limit has no negative punctures, i.e., is a honest finite energy plane.

3. The systole

Assume that (N, λ) is a closed contact manifold. Denote by

$$\mathcal{R} = \mathcal{R}(N, \lambda) \subset C^\infty(S^1, N)$$

the set of all periodic Reeb orbits of N, i.e., the set of all loops $\gamma \in C^\infty(S^1, N)$ for which there exists a positive number $\tau = \tau_\gamma$, referred to as the period, such that the tuple (γ, τ) is a solution of the ODE $\partial_t \gamma = \tau R(\gamma)$. The *systole* of (N, λ) is defined as

$$\mathrm{sys}(N, \lambda) = \inf\{\tau_\gamma : \gamma \in \mathcal{R}\}.$$

Here we use the convention that infimum of the empty set equals infinity. However, in view of Weinstein's conjecture the systole of every closed contact manifold is expected to be finite. In view of the result by Taubes [**105**] this is definitely true in dimension three. Moreover, because our contact manifold is assumed to be closed it follows from the Theorem of Arzelà-Ascoli that if a periodic Reeb orbit exists the infimum is actually attained so that for a three dimensional contact manifold we can define the systole as well as a minimum

$$\mathrm{sys}(N, \lambda) = \min\{\tau_\gamma : \gamma \in \mathcal{R}\}.$$

We further point out that if the contact manifold (N, λ) satisfies in addition

$$H_1(N; \mathbb{Q}) = \{0\}$$

then the systole only depends on the Hamiltonian structure $(N, d\lambda)$ and not on the choice of the contact form λ. Indeed, if λ and λ' are two contact forms on N satisfying

$$d\lambda = d\lambda'$$

then since the first rational homology group of N vanishes there exists a smooth function $f \in C^\infty(N, \mathbb{R})$ such that

$$\lambda = \lambda' + df.$$

Now in view of Stokes theorem and the definition of the Reeb vector field we compute for $\gamma \in \mathcal{R}$

$$\tau_\gamma = \int_{S^1} \gamma^* \lambda = \int_{S^1} \gamma^* \lambda'.$$

This proves that if the first rational homology group vanishes the systole only depends on the Hamiltonian structure $(N, d\lambda)$.

If $\gamma \in \mathcal{R}$ satisfies $\tau_\gamma = \mathrm{sys}(N, \lambda)$ we say that the periodic Reeb orbit γ *represents the systole* of (N, λ). Note that the systole in general does not have a unique representative. However, each representative of the systole necessarily has minimal period among all periodic Reeb orbits.

Theorem 3.1. *Assume that (N, λ) is a closed, three dimensional contact manifold and γ is a non-degenerate Reeb orbit of (N, λ) which represents the systole. Then the moduli space $\mathcal{M}(\gamma)/\mathbb{R}$ is compact.*

Proof: In view of Theorem 2.1 it suffices to show that each negatively punctured finite energy plane $\widetilde{u} = (u, a)\colon \mathbb{C} \setminus P \to N \times \mathbb{R}$ has no punctures, i.e., $P = \emptyset$. We argue by contradiction and assume that $P \neq \emptyset$. Hence suppose

$$P = \{p_1, \dots, p_\ell\}$$

for $\ell \in \mathbb{N}$. Hofer's theorem (Theorem 2.4) can as well be applied to negative punctures. For a negative puncture at p_j with $j \in \{1, \dots, \ell\}$ it asserts that there exists a sequence s_k^j going to $-\infty$ and a periodic Reeb orbit γ_j such that

$$\lim_{k \to \infty} u(e^{2\pi(s_k^j + it)} + p_j) = \gamma_j(t) \tag{218}$$

uniformly in the C^∞-topology. Because the puncture at infinity is positive there exists moreover a sequence s_k going to infinity such that

$$\lim_{k \to \infty} u(e^{2\pi(s_k + it)}) = \gamma(t) \tag{219}$$

uniformly in the C^∞-topology. Because $\widetilde{u}$ is holomorphic as a special instance of (143) it holds that

$$u^* d\lambda = ||\pi \partial_x u||^2 \geq 0. \tag{220}$$

Abbreviate by $D_R(p) = \{z \in \mathbb{C} : ||z - p|| \leq R\}$ the disk of radius R centered at p. There exists $k_0 \in \mathbb{N}$ such that for every $k \geq k_0$ and every $1 \leq j, j' \leq \ell$ satisfying $j \neq j'$ we have

$$D_{e^{2\pi s_k^j}}(p_j) \cap D_{e^{2\pi s_k^{j'}}}(p_j') = \emptyset, \quad D_{e^{2\pi s_k^j}}(p_j) \subset D_{e^{2\pi s_k}}(0).$$

In view of Stokes theorem we conclude for every $k \geq k_0$

$$0 \leq \int_{D_{e^{2\pi s_k}}(0) \setminus \bigcup_{j=1}^{\ell} D_{e^{2\pi s_k^j}}(p_j)} u^* d\lambda = \int_{\partial D_{e^{2\pi s_k}}(0)} u^* \lambda - \sum_{j=1}^{\ell} \int_{\partial D_{e^{2\pi s_k^j}}(p_j)} u^* \lambda.$$

Since this is true for any $k \geq k_0$ we obtain in view of (201) and (202)

$$0 \leq \int_{S^1} \gamma^* \lambda - \sum_{j=1}^{\ell} \int_{S^1} \gamma_j^* \lambda = \tau_\gamma - \sum_{j=1}^{\ell} \tau_{\gamma_j} \tag{221}$$

where τ_{γ_j} are the periods of the periodic orbits γ_j. Because γ represents the systole the following inequalities hold true

$$\tau_{\gamma_j} \geq \tau_\gamma, \quad 1 \leq j \leq \ell. \tag{222}$$

From (203) and (204) we conclude

$$\ell = 1, \quad \tau_{\gamma_1} = \tau_\gamma.$$

In view of (220) we further get

$$\pi \partial_x u(z) = 0, \quad z \in \mathbb{C}$$

and because $\widetilde{u}$ is holomorphic we get as well

$$\pi \partial_y u(z) = 0 \quad z \in \mathbb{C}.$$

In particular, $\widetilde{u}$ is an orbit cylinder up to reparametrization, in contradiction to assertion (iv) in Definition 1.1. This finishes the proof of the Theorem. □

Corollary 3.2. *Suppose the assumptions of Theorem 3.1. Then the moduli space $\mathcal{M}_{\mathrm{fast}}(\gamma)/\mathbb{R}$ is compact.*

Proof: Recall from Corollary 0.13 that a non-degenerate finite energy plane $\widetilde{u} = (u, a)$ is fast if and only if the winding number of its asymptotic eigenvalue satisfies $w(\eta_u) = 1$. By Theorem 0.5 the inequality $w(\eta_u) \geq 1$ holds. If $\widetilde{u}$ is the limit of fast finite energy planes it follows from (217) that $w(\eta_u) \leq 1$. Therefore, it follows that $w(\eta_u) = 1$ and $\widetilde{u}$ is fast. We have shown that the moduli space $\mathcal{M}_{\mathrm{fast}}(\gamma)/\mathbb{R}$ is closed in $\mathcal{M}(\gamma)/\mathbb{R}$. Now the Corollary follows from Theorem 3.1. □

4. Dynamical convexity

Recall that a periodic Reeb orbit $\gamma \in C^\infty(S^1, N)$ of period τ is called non-degenerate if $\det(d^\xi \phi_R^\tau(\gamma(0)) - \mathrm{id}) \neq 0$.

Definition 4.1. *A contact manifold (N, λ) is called* non-degenerate*, if all periodic Reeb orbits on (N, λ) are non-degenerate.*

After a small perturbation we can always assume that a closed contact manifold is non-degenerate. To make this statement precise, recall that if $f \in C^\infty(N, \mathbb{R}_+)$ is a smooth positive function on N, then the one form $\lambda_f := f\lambda \in \Omega^1(N)$ is still a contact form on N. Note that the contact structure $\xi = \ker \lambda = \ker \lambda_f$ remains unchanged under this procedure, although the Reeb vector field and therefore the dynamics on N might change dramatically. The following result is due to Robinson [**96**].

Theorem 4.2. *Assume that (N, λ) is a closed contact manifold. Then there exists a subset $\mathcal{F} \subset C^\infty(N, \mathbb{R}_+)$ which can be written as a countable intersection of open and dense subsets of $C^\infty(N, \mathbb{R}_+)$ such that for every $f \in \mathcal{F}$ the contact form λ_f is non-degenerate.*

Since $\mathcal{F}$ is a countable intersection of open and dense subsets it follows from Baire's theorem that $\mathcal{F}$ is dense itself. This explains why after small perturbation we can assume that the contact manifold is non-degenerate.

Suppose that γ is a contractible closed Reeb orbit in a contact manifold (N, λ)

of period τ. Since γ is contractible there exists a *filling disk* for γ, i.e., a smooth map $\overline{\gamma}\colon D = \{z \in \mathbb{C} : |z| \leq 1\} \to N$ such that

$$\overline{\gamma}(e^{2\pi it}) = \gamma(t), \quad t \in S^1.$$

Choose a symplectic trivialization

$$\mathfrak{T}\colon \overline{\gamma}^*\xi \to D \times \mathbb{C}.$$

We define the Conley-Zehnder index of the filling disk $\overline{\gamma}$ as

$$\mu_{CZ}(\overline{\gamma}) = \mu_{CZ}\Big(t \mapsto \mathfrak{T}_{e^{2\pi it}} d^\xi \phi_R^{t\tau}(\gamma(0))\mathfrak{T}_1^{-1}\Big).$$

The Conley-Zehnder index is independent of the choice of the symplectic trivialization and depends only on the homotopy class of the filling disk $\overline{\gamma}$. If $\overline{\gamma}'$ is another filling disk for γ, one obtains a sphere $\overline{\gamma}\#(\overline{\gamma}')^-$ by gluing $\overline{\gamma}$ and $(\overline{\gamma}')^-$, the filling disk $\overline{\gamma}'$ with opposite orientation, along γ. In view of Lemma 2.1 the Conley-Zehnder indices with respect to the two filling disks are related by

$$\mu_{CZ}(\overline{\gamma}) - \mu_{CZ}(\overline{\gamma}') = 2c_1\Big((\overline{\gamma}\#(\overline{\gamma}')^-)^*\xi\Big). \tag{223}$$

The first Chern class gives rise to a homomorphism

$$I_{c_1} : \pi_2(N) \to \mathbb{Z}, \quad [v] \mapsto c_1(v^*\xi).$$

Suppose now that the homomorphism I_{c_1} is trivial. This for example happens if $\pi_2(N) = \{0\}$. Then it follows from (223) that the Conley-Zehnder index is independent of the choice of the filling disk $\overline{\gamma}$ and only depends on the periodic Reeb orbit γ. Hence under the assumption that $I_{c_1} = 0$ we can set

$$\mu_{CZ}(\gamma) := \mu_{CZ}(\overline{\gamma})$$

where $\overline{\gamma}$ is any filling disk for the contractible Reeb orbit γ.

Definition 4.3. *A closed three dimensional contact manifold (N, λ) is called* dynamically convex *if $I_{c_1} = 0$ and every closed contractible Reeb orbit γ of N satisfies*

$$\mu_{CZ}(\gamma) \geq 3.$$

If γ is a periodic Reeb orbit recall that the *covering number* of γ is defined as

$$\operatorname{cov}(\gamma) = \max\Big\{k \in \mathbb{N} : \gamma(t + \tfrac{1}{k}) = \gamma(t),\ \forall\ t \in S^1\Big\}.$$

Moreover, a periodic Reeb orbit γ is called *simple* if $\operatorname{cov}(\gamma) = 1$. The following theorem is due to Hryniewicz [**58**].

Theorem 4.4 (Hryniewicz). *Assume that (N, λ) is a non-degenerate, dynamically convex closed three dimensional contact manifold and γ is a simple periodic Reeb orbit of N. Then the moduli space $\mathcal{M}_{\mathrm{fast}}(\gamma)/\mathbb{R}$ is compact.*

Proof: Suppose that $\widetilde{u}_\nu = (u_\nu, a_\nu)$ is a sequence of fast finite energy planes which asymptotic orbit γ which converge to a negatively punctured finite energy plane $\widetilde{u} = (u, a)\colon \mathbb{C} \setminus P \to N \times \mathbb{R}$ with asymptotic orbit γ in the C^∞_{loc}-topology. It remains to show that the set of negative punctures P is empty and that $\widetilde{u}$ is fast. We first rule out the danger that $\widetilde{u}$ is a so called connector, namely a negatively punctured finite energy plane satisfying

$$||\pi\partial_x u||^2 = \frac{1}{2}\Big(||\pi\partial_x u||^2 + ||\pi\partial_y u||^2\Big) = 0$$

where $\pi\colon TN \to \xi$ is the projection along the Reeb vector field. If $\widetilde{u}$ is a connector it follows that

$$\widetilde{u} = \widetilde{\gamma}' \circ p$$

where $\widetilde{\gamma}'$ is the orbit cylinder over a periodic Reeb orbit γ' and $p\colon \mathbb{C} \to \mathbb{C}$ is a holomorphic map, which has to be a polynomial because the energy of $\widetilde{u}$ is finite. Because γ is simple it follows that $\gamma' = \gamma$ and p has degree one, i.e., $\widetilde{u}$ is a reparametrization of an orbit cylinder which is forbidden by condition (iv) in Definition 1.1. Therefore $\widetilde{u}$ is not a connector.

We now suppose by contradiction that the set of punctures is not empty, so that we can write $P = \{p_1, \ldots, p_\ell\}$ for $\ell \in \mathbb{N}$. Because (N, λ) is non-degenerate and $\widetilde{u}$ is not a connector the asymptotic description from Theorem 5.2 can now be applied to the negative punctures as well. Namely for each $1 \leq j \leq \ell$ there exists a periodic Reeb orbit γ_j , a positive eigenvalue η_j of the operator A_{γ_j} and an eigenvector ζ_j of A_{γ_j} to the eigenvalue η_j such that the puncture p_j admits an asymptotic representative of the form

$$U_j(s,t) = e^{\eta_j s}\big(\zeta_j(t) + \kappa_j(s,t)\big)$$

where κ_j decays with all derivatives exponentially with uniform exponential weight. For negative punctures asymptotic representative means that there exists proper embeddings $\phi_j\colon (-\infty, R_j] \times S^1 \to \mathbb{R} \times S^1$ asymptotic to the identity such that

$$\widetilde{u}\Big(e^{\phi_j(s,t)} + p_j\Big) = \Big(\exp_{\gamma_j(t)} U_j(s,t), \tau_j s\Big)$$

where exp is the exponential map for some Riemannian metric on N and τ_j are the periods of the periodic orbits γ_j. Because $\widetilde{u}$ is the limit of the finite energy planes $\widetilde{u}_\nu$ it follows that the periodic orbits γ_j are contractible. Hence we can pick for each periodic orbit γ_j a filling disk $\overline{\gamma_j}$. Pick unitary trivializations $\mathfrak{T}_j\colon \overline{\gamma_j}^*\xi \to D \times \mathbb{C}$, i.e., trivializations which are complex with respect to the complex structure J on ξ and orthogonal with respect to the metric $\omega(\cdot, J\cdot)$ on ξ. The restriction of the trivializations to the periodic Reeb orbits γ_j gives rise to the bounded linear operators

$$A^{\mathfrak{T}_j}_{\gamma_j}\colon W^{1,2}(S^1, \mathbb{C}) \to L^2(S^1, \mathbb{C})$$

as in (146). Because the operator $A^{\mathfrak{T}_j}_{\gamma_j}$ is conjugated to the operator A_{γ_j} the eigenvalue η_j of the operator A_{γ_j} can also be interpreted as an eigenvalue of the operator $A^{\mathfrak{T}_j}_{\gamma_j}$. As an eigenvalue of the operator $A^{\mathfrak{T}_j}_{\gamma_j}$ it has a winding number

$$w(\eta_j, \overline{\gamma}_j) \in \mathbb{Z}.$$

As the notation indicates the winding number is independent of the choice of the trivialization $\mathfrak{T}_j$. A priori it depends at least up to homotopy on the choice of the filling disk $\overline{\gamma_j}$. If $\overline{\gamma_j}'$ is another filling disk then the winding numbers are related by

$$w(\eta_j, \overline{\gamma_j}) - w(\eta_j, \overline{\gamma_j}') = c_1\Big((\overline{\gamma_j}\#(\overline{\gamma_j}')^-)^*\xi\Big).$$

Because the contact manifold (N, λ) is dynamically convex by assumption the homomorphism $I_{c_1}\colon \pi_2(N) \to \mathbb{Z}$ is trivial and therefore the winding number is independent of the choice of the filling disk, so that we can set

$$w(\eta_j) := w(\eta_j, \overline{\gamma_j}).$$

Using again that (N,λ) is dynamically convex it holds that

$$\mu_{CZ}(\gamma_j) \geq 3, \quad 1 \leq j \leq \ell.$$

Because η_j is positive we conclude in view of Theorem 0.33 and the monotonicity of the winding number from Corollary 0.32 that

$$w(\eta_j) \geq 2, \quad 1 \leq j \leq \ell. \tag{224}$$

By gluing the filling disks $\overline{\gamma_j}$ to u along γ_j for $1 \leq j \leq \ell$ we obtain an open disk

$$u \# \bigcup_{j=1}^{\ell} \overline{\gamma_j}$$

whose closure is a filling disk for γ. Choose a trivialization

$$\mathfrak{T} \colon u^*\xi \to (\mathbb{C} \setminus P) \times \mathbb{C}$$

which extends at the positive puncture to a trivialization $\mathfrak{T} \colon \gamma^*\xi \to S^1 \times \mathbb{C}$ and coincides at the negative punctures with $\mathfrak{T}_j \colon \gamma_j^*\xi \to S^1 \times \mathbb{C}$. Inspired by the proof of Theorem 0.5 we consider the smooth map

$$\mathfrak{T}\pi\partial_x u \colon \mathbb{C} \setminus P \to \mathbb{C}.$$

In view of (224) there exists $\epsilon > 0$ such that for the loops

$$\gamma_j^\epsilon \colon S^1 \to \mathbb{C}, \quad t \mapsto p_j + \epsilon e^{2\pi i t}$$

where $1 \leq j \leq \ell$ the winding number as defined in (150) of the map $\mathfrak{T}\pi\partial_r u$ along these loops satisfies

$$w_{\gamma_j^\epsilon}(\mathfrak{T}\pi\partial_r u) \geq 2.$$

In view of (157) we conclude that

$$w_{\gamma_j^\epsilon}(\mathfrak{T}\pi\partial_x u) \geq 1.$$

Because $\widetilde{u}$ is not a connector it follows that η_u is finite and therefore an eigenvalue of the asymptotic operator A_γ. Because $\widetilde{u}$ is the limit of fast finite energy planes it follows from (217) and the monotonicity of winding established in Corollary 0.32 that

$$w(\eta_u) \leq 1.$$

Hence there exists $R > 0$ such that

$$w_R(\mathfrak{T}\pi\partial_r u) \leq 1$$

where we recall from (151) that w_R denotes the winding number of the loop $t \mapsto Re^{2\pi i t}$. Again using (157) we conclude that

$$w_R(\mathfrak{T}\pi\partial_x u) \leq 0.$$

On the other hand since $\widetilde{u}$ is holomorphic we conclude using Carleman's similarity principle as in the proof of Theorem 0.5 that

$$0 \geq w_R(\mathfrak{T}\pi\partial_x u) \geq \sum_{j=1}^{\ell} w_{\gamma_j^\epsilon}(\mathfrak{T}\pi\partial_x u) \geq \ell.$$

This implies that $\ell = 0$. Hence $\widetilde{u}$ has no negative punctures and is therefore a finite energy plane. In view of Theorem 0.10 the winding number of its asymptotic eigenvalue satisfies $w(\eta_u) \geq 1$ and because $\widetilde{u}$ is the limit of fast finite energy planes it follows from (217) that $w(\eta_u) = 1$. This shows that $\widetilde{u}$ is fast and the theorem is proved. $\square$

5. Open book decomposition

The following theorem due to Godement can be found for example in [**3**, Theorem 3.5.25].

Theorem 5.1. *Assume that M is a manifold and $R \subset M \times M$ is an equivalence relation. Denote by $p_1 \colon M \times M \to M$ the projection to the first factor. Suppose that the following conditions hold*

- **(i):** *$R \subset M \times M$ is a closed submanifold.*
- **(ii):** *The restriction of the projection $p_1|_R \colon R \to M$ is a submersion.*

Then the quotient space M/R is a manifold and the quotient projection $\pi \colon M \to M/R$ is a submersion.

Remark 5.2. *All manifolds in the theorem are assumed to satisfy the Hausdorff separation axiom. If one does not require in assertion (i) that $R \subset M \times M$ is closed, then M/R is still a (not necessarily Hausdorff) manifold and the quotient projection a submersion.*

In the special case where G is a Lie group action on a manifold M and the equivalence relation is the orbit relation, then

$$R = \{(x, gx) : x \in M, g \in G\}.$$

If G acts freely on M, then $R \subset M \times M$ is a manifold and assertion (ii) holds. Hence we obtain the following Corollary.

Corollary 5.3. *Assume that a Lie group G acts freely on a manifold M and $\{(x, gx) : x \in M, g \in G\}$ is closed in $M \times M$. Then the quotient M/G is a manifold and the orbit projection a submersion.*

5.1. Global surface of section to open book.

Suppose that $d : D^2 \to S^3$ is a global disk-like surface of section for a Reeb vector field X. Let $\gamma := d|_\partial : S^1 \to S^3$ denote the periodic orbit that bounds the surface of section.

Proposition 5.4. *Assume that γ is transversely non-degenerate and that $\mu_{CZ}(\gamma) \geq 3$. Then there is an open book on S^3 with binding γ whose pages are transverse to X.*

Proof: First of all, note that it is enough to exhibit for every point $y \in int(D^2)$ a return time $t(y)$ depending smoothly on y such that the return map $y \mapsto d^{-1} \circ \phi_X^{t \circ d(y)} d(y)$ extends continuously to the boundary.
Step 1: We claim that there is a neighborhood $\nu(\gamma)$ of γ together with coordinates $\psi : S^1 \times D^2 \to \nu(\gamma)$ such that

$$\psi^{-1}(\nu(\gamma) \cap d(D)\,) = \{(\phi; x, 0) \in S^1 \times D^2) \mid \phi \in S^1, \text{ and } x \geq 0\}.$$

The tubular neighborhood theorem gives us a neighborhood $\nu(\gamma)$ that is diffeomorphic to $S^1 \times D^2$. We need to be a little more explicit here. We choose a trivialization of the contact structure along γ, so a map $S^1 \times \mathbb{R}^2 \to \gamma^*\xi$ with the property that the vector U, the image of $(1, 0)$ is tangent to d along γ and pointing inward. Let V denote the image of the vector $(0,1)$. Extend the vector field U, V to a neighborhood of γ. To find suitable coordinates, define the map

$$\begin{aligned}\psi : S^1 \times D^2 &\longrightarrow \nu(\gamma)\\ (\phi; x, y) &\longmapsto Exp_{yV} \circ \phi_U^x \circ d(e^{i\phi}).\end{aligned}$$

Put $y = 0$ and take a sequence ϵ_n converging to 0. We get a curve $\phi_U^\epsilon \circ d(e^{i\phi})$ to which we apply the map ϕ_X^t. This is the curve

$$\phi_X^t \circ \phi_U^\epsilon \circ d(e^{i\phi})$$

whose equivalence class at $\epsilon = 0$ is by definition

$$T\phi_X^t(U_{d(e^{i\phi})}).$$

This is the time-t linearized flow of X acting on U.

Step 2: Bounding the return time Let θ denote the rotation number of γ with respect to a trivialization of $d^*\xi$. By a standard formula for the Conley-Zehnder index due to Long, $\mu_{CZ}(\gamma)$ equals $2\lfloor\frac{\theta}{2\pi}\rfloor + 1$ if γ is elliptic and $\frac{\theta}{\pi}$ if γ is hyperbolic. As $\mu_{CZ}(\gamma) \geq 3$, the rotation number is more than 2π.

If τ is the period of γ we see the return time of the linearized flow is less than τ. Since the actual flow $\phi_X^t \circ \phi_U^\epsilon \circ d(e^{i\phi})$ converges to the linearized flow as ϵ converges to 0, we find for every $y \in D^2$ with $d(y) \in \nu(\gamma)$ a minimal positive time $0 < \tilde{t}(x) < \tau + \delta$ such that

$$\phi_X^{\tilde{t}(y)} \circ d(y) \in D.$$

Let D_i denote the set of points $y \in D^2$ with $d(y) \notin \nu(\gamma)$. Then D_i is compact, so by smooth dependence on initial conditions we find for all $y \in D_i$ a minimal positive time $t(y)$ such that

$$\phi_X^{\tilde{t}(y)} \circ d(y) \in D.$$

We conclude that $t : D^2 \to \mathbb{R}$ is a continuous function that is smooth in the interior. Define the return map

$$\begin{aligned}rt : D^2 &\longrightarrow D^2\\ x &\longmapsto d^{-1} \circ \phi_X^{t(x)} \circ d(x).\end{aligned}$$

We see that rt is homeomorphism that is smooth on the interior with the property that it is conjugate to a map preserving the $d^*d\alpha$-area. $\square$

Bibliography

[1] C. Abbas, H. Hofer, *Holomorphic Curves and Global Questions in Contact Geometry*, book in preparation.

[2] R. Abraham, J. Marsden, *Foundations of Mechanics*, 2nd ed., Addison-Wesley, Reading (1978).

[3] R. Abraham, J. Marsden, T. Ratiu, *Manifolds, tensor analysis, and applications*, 2nd ed., Applied Mathematical Sciences **75**, Springer (1988).

[4] P. Albers, J. Fish, U. Frauenfelder, H. Hofer, O. van Koert, *Global surfaces of section in the planar restricted 3-body problem*, Arch. Ration. Mech. Anal. **204** (2012), no. 1, 273–284.

[5] P. Albers, U. Frauenfelder, *Rabinowitz Floer homology: A survey*, in Global Differential Geometry, Springer Proc. in Math. (2012), 437–461.

[6] P. Albers, U. Frauenfelder, O. van Koert, G. Paternain, *Contact geometry of the restricted three-body problem*, Comm. Pure Appl. Math. **65** (2012), no. 2, 229–263.

[7] P. Albers, H. Hofer, *On the Weinstein conjecture in higher dimensions*, Comm. Math. Helv. (2009), 429–436.

[8] V. Arnold, *On a characteristic class entering into conditions of quantization*, Func. Anal. **1** (1967), 1–8.

[9] M. Atiyah, *Convexity and commuting Hamiltonians*, Bull. London Math. Soc. **14** (1982), 1–15.

[10] V. Bangert, Y. Long, *The existence of two closed geodesics on every Finsler 2-sphere*, Math. Ann. **346** (2010), 335–366.

[11] V. Bargmann, *Zur Theorie des Wasserstoffatoms: Bemerkungen zur gleich- namigen Arbeit von V. Fock*, Zeitschrift für Physik **99** (1936), 576–582.

[12] E. Belbruno, *Two body motion under the inverse square central force and equivalent geodesic flows*, Celest. Mech. **15** (1977), 467–476.

[13] E. Belbruno, *Regularizations and geodesic flows*, Lecture notes in Pure and Applied Mathematics **80** (1981), 1–11.

[14] E. Belbruno, *Capture Dynamics and Chaotic Motions in Celestial Mechanics*, Princeton University Press (2004).

[15] J. van den Berg, F. Pasquotto, C. Vandervorst, *Closed characteristics on non-compact hypersurfaces in $\mathbb{R}^{2n}$*, Math. Ann. **343** (2009), 247–284.

[16] G. Birkhoff, *Proof of Poincaré's geometric theorem*, Trans. Amer. Math. Soc. **14** (1913), 14–22.

[17] G. Birkhoff, *The restricted problem of three bodies*, Rend. Circ. Matem. Palermo **39** (1915), 265–334.

[18] F. Bourgeois, *A Morse-Bott approach to contact homology*, Ph.D. thesis, Stanford University (2002).

[19] F. Bourgeois, Y. Eliashberg, H. Hofer, K. Wysocki, E. Zehnder, *Compactness results in Symplectic Field theory*, Geom. and Top. **7** (2003), 799–833.

[20] T. Bröcker, K. Jänich, *Introduction to Differential Topology*, Cambridge University Press (1982).

[21] L. Brouwer, *Beweis des ebenen Translationssatzes*, Math. Ann. **72** (1912), 37–54.

[22] L. Brouwer, *Über die periodischen Transformationen der Kugel*, Math. Ann. **80** (1919), 39–41.

[23] A. Chenciner, *Poincaré and the Three-Body Problem*, Poincaré 1912–2012, Séminaire Poincaré XVI (2012), 45–133.

[24] K. Cieliebak, U. Frauenfelder, *A Floer homology for exact contact embeddings* (2009), 251–316.

[25] K. Cieliebak, U. Frauenfelder, O. van Koert, *The Finsler geometry of the rotating Kepler problem*. Publ. Math. Debrecen 84 (2014), no. 3-4, 333–350.

[26] C. Conley, *On Some New Long Periodic Solutions of the Plane Restricted Three Body Problem*, Comm. Pure Appl. Math. **16** (1963), 449–467.

[27] A. Constantin, B. Kolev, *The theorem of Kérékjartò on periodic homeomorphisms of the disc and the sphere*, Enseign. Math. (2) **40** (1994), no. 3–4, 193–204.

[28] D. Cristofaro-Gardiner, M. Hutchings, *From one Reeb orbit to two*, arXiv:1202.4839

[29] J. Duistermaat, *On the Morse index in Variational Calculus*, Adv. Math. **21** (1976), 173–195.

[30] C. Ehresmann, *Les connexions infinitésimales dans un espace fibré différentiable*, Colloque de Topologie, Bruxelles (1950), 29–55.

[31] Y. Eliashberg, A. Givental, H. Hofer, *Introduction to symplectic field theory*, Geom. Funct. Anal. **10** (2000), 560–673.

[32] A. Fathi, *An orbit closing proof of Brouwer's lemma on translation arcs*, L'enseignement Math. **33** (1987), 315–322.

[33] A. Floer, *Morse theory for Lagrangian intersections*, J. Differential Geom. **28** (1988), 513–547.

[34] A. Floer, H. Hofer, D. Salamon, *Transversality in elliptic Morse theory for the symplectic action*, Duke Math. Journal, **80** (1996), 251–292.

[35] V. Fock, *Zur Theorie des Wasserstoffatoms*, Zeitschrift für Physik **98** (1935), 145–154.

[36] J. Franks, *A new proof of the Brouwer plane translation theorem*, Erg. Th. and Dyn. Systs., **12** (1992), 217–226.

[37] J. Franks, *Geodesics on S^2 and periodic points of annulus homeomorphisms*, Invent. Math., **108** (1992), 403–418.

[38] J. Franks, *Area preserving homeomorphisms of open sufaces of genus zero*, New York Jour. Math. **2** (1996), 1–19.

[39] D. Fuks, *The Maslov-Arnold characteristic classes*, Dokl. Akad. Nauk SSSR **178** (1968), 303–306.

[40] H. Geiges, *An introduction to contact topology*, Cambridge Studies in Adv. Math. **109**, Cambridge Univ. Press (2008).

[41] V. Ginzburg, B. Gürel, *A C^2-smooth counterexample to the Hamiltonian Seifert conjecture in $\mathbb{R}^4$*, Ann. of Math. (2) **158** (2003), no.3, 953–976.

[42] H. Goldstein, *Prehistory of the Runge-Lenz vector*, Amer. Jour. Phys. **43** (1975), no. 8, 737–738.

[43] H. Goldstein, *More on the prehistory of the Runge-Lenz vector*, Amer. Jour. Phys. **44** (1976), no. 11, 1123–1124.

[44] E. Goursat, *Sur les transformations isogonales en Mécanique*, Comptes Rendus des Séances de l'Académie des Sciences, Paris, **108**, (1889), 446–448.

[45] M. Gromov, *Pseudo Holomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), 307–347.

[46] V. Guillemin, S. Sternberg, *Convexity Properties of the Moment Mapping*, Inv. Math. **67** (1982), 491–513.

[47] L. Guillou, *Théorème de translation plane de Brouwer et généralisations du théorème de Poincaré-Birkhoff*, Topology **33**(2) (1994), 331–351.

[48] A. Hatcher, *Algebraic Topology*, Cambridge University Press (2001).

[49] G. Hill, *Researches in the lunar theory*, Amer. J. Math.**1** (1878), 5–26, 129–147, 245–260

[50] L. Hörmander, *Fourier integral operators I*, Acta Math. **127** (1971), 79–183.

[51] H. Hofer, *Pseudoholomorphic curves in symplectisations with application to the Weinstein conjecture in dimension three*, Invent. Math. **114** (1993), 515–563.

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

[53] H. Hofer, K. Wysocki, E. Zehnder, *Properties of pseudo-holomorphic curves in symplectizations. II. Embedding controls and algebraic invariants*, Geom. Funct. Anal. **5** (1995), no. 2, 270–328.

[54] H. Hofer, K. Wysocki, E. Zehnder, *Properties of pseudo-holomorphic curves in symplectizations. I. Asymptotics*, Ann. Inst. H. Poincaré Anal. Non Linéaire **13** (1996), no. 3, 337–379.

[55] H. Hofer, K. Wysocki, E. Zehnder, *Correction to: "A characterisation of the tight three sphere"*, Duke Math. J. **89** (1997), no. 3, 603–617.

[56] H. Hofer, K. Wysocki, E. Zehnder, *The dynamics on a strictly convex energy surface in* $\mathbb{R}^4$, Ann. Math., **148** (1998) 197–289.

[57] H. Hofer, K. Wysocki, E. Zehnder, *Properties of pseudo-holomorphic curves in symplectizations. III. Fredholm theory*, Topics in nonlinear analysis, Progr. Nonlinear Differential Equations Appl., **35**, Birkhäuser, Basel (1999), 381–475.

[58] U. Hryniewicz, *Fast finite-energy planes in symplectizations and applications*, Trans. Amer. Math. Soc. **364** (2012), 1859–1931.

[59] U. Hryniewicz, *Systems of global surfaces of section for dynamically convex Reeb flows on the* 3*-sphere*, arXiv:1105.2077v3.

[60] U. Hryniewicz, P. Salomão, *On the existence of disk-like global sections for Reeb flows on the tight 3-sphere*, Duke Math. J. **160**, no. 3, (2011), 415–465.

[61] L. Hulthén, *Über die quantenmechanische Herleitung der Balmerterme*, Zeitschrift für Physik **86** (1933), 21–23.

[62] M. Hutchings, *An index inequality for embedded pseudoholomorphic curves in symplectizations*, J. Eur. Math. Soc. (JEMS) **4** (2002), 313–361.

[63] P. Kahn, *Pseudohomology and Homology*, math.AT/0111223.

[64] J. Kang, *Some remarks on symmetric periodic orbits in the restricted three-body problem*, Disc. Cont.,Dyn. Sys. (A), **34**, no. 12 (2014), 5229–5245.

[65] J. Kang, *On reversible maps and symmetric periodic points*, arXiv:14103997

[66] T. Kato, *On the convergence of the perturbation method I*, Progr.,Theor. Phys. **4** (1949), 514–523.

[67] T. Kato, *On the convergence of the perturbation method II*, Progr.,Theor. Phys. **5** (1950), 207–212.

[68] T. Kato, *Perturbation theory for Linear Operators*, Springer, Grundlehren edition (1976).

[69] B. von Kérékjartò, *Über die periodischen Transformationen der Kreisscheibe und der Kugelfläche*, Math. Ann. **80**, (1919–1920), 36–38.

[70] D. Kim, *Planar Circular Restricted Three Body Problem*, Master thesis, Seoul National University (2011).

[71] S. Kim, *Hamiltonian mechanics and Symmetries*, Master thesis, Seoul National University (2014).

[72] M. Kriener, *An intersection formula for finite energy half cylinders*, PhD thesis, ETH Zurich (1998).

[73] M. Kummer, *On the stability of Hill's Solutions of the Plane Restricted Three Body Problem*, Amer. J. Math. **101**(6), (1979), 1333–1354.

[74] P. Kustaanheimo, E. Stiefel, *Perturbation theory of Kepler motion based on spinor regularization*, J. Reine Angew. Math. **218**,(1965), 204–219.

[75] J. Lee, *Fiberwise Convexity of Hill's lunar problem*, arXiv:1411.7573

[76] T. Levi-Civita, *Sur la régularisation du problème des trois corps*, Acta Math. **42** (1920), 99–144.

[77] R. Lickorish, *An introduction to knot theory*, Graduate texts in mathematics **175**, New York, Springer (1997).

[78] G. Lion, M. Vergne, *The Weil representation, Maslov index and Theta Series*, Progress in Mathematics **6**, Birkhäuser (1980).

[79] J. Llibre, L Roberto, *On the periodic orbits and the integrability of the regularized Hill lunar problem*, J. Math. Phys. **52**(8), 082701, 8 pp. (2011).

[80] Y. Long, D. Zhang, C. Zhu, *Multiple brake orbits in bounded convex symmetric domains*, Adv. Math. **203** (2006), 568–635.

[81] D. McDuff, D. Salamon, *J-holomorphic Curves and Symplectic Topology* 2nd edition, Amer. Math. Soc., Providence, RI (2012).

[82] K. Meyer, D. Schmidt, *Hill's lunar Equations and the Three-Body Problem*, J. Diff. Eq. **44**, (1982), 263–272.

[83] J. Milnor, *Topology from the Differential Viewpoint*, The University Press of Virginia (1965).

[84] J. Milnor, *On the geometry of the Kepler problem*, Amer. Math. Monthly **90**(6), (1983),353–365.

[85] E. Mora, *Pseudoholomorphic cylinders in symplectisations*, Doctoral dissertation, New York University (2003).

[86] J. Moser, *Regularization of Kepler's problem and the averaging method on a manifold*, Comm. Pure Appl. Math. **23** (1970), 609–636.

[87] J. Moser, *Periodic orbits near an equilibrium and a theorem by Alan Wein-stein*, Comm. Pure Appl. Math. **29** (1976), 724–747.

[88] Y. Osipov, *Geometrical interpretation of Kepler's problem*, *Uspehi Mat. Nauk*, **27**(2) (1972), 161 (In Russian).

[89] Y. Osipov, *The Kepler problem and geodesic flows in spaces of constant curvature*, Celest. Mech. **16** (1977), 191–208.

[90] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste I-III*, Gauthiers-Villars, Paris (1899).

[91] H. Poincaré, *Sur un théorème de géométrie*, Rend. Circ. Matem. Palermo **33** (1912), 375–407.

[92] W. Pauli, *Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik*, Zeitschrift für Physik **36** (1926), 336–363.

[93] P. Rabinowitz, *Periodic solutions of Hamiltonian systems*, Comm. Pure Appl. Math. **31** (1978), 157–184.

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

[95] J. Robbin, D. Salamon, *The spectral flow and the Maslov index*, Bull. L.M.S. **27** (1995), 1–33.

[96] C. Robinson, *A global approximation theorem for Hamiltonian systems*, Proc. Symp. Pure Math. **XIV**, Global Analysis, AMS (1970), 233–244.

[97] M. Schwarz, *Cohomology operations from S^1-Cobordism in Floer Homology*, Ph.D thesis, ETH Zürich (1996).

[98] M. Schwarz, *Equivalences for Morse Homology*, Geometry and Topology in Dynamics, Contemp. Math. **246**, AMS, Providence, RI (1999), 197–216.

[99] H. Seifert, *Periodische Bewegungen mechanischer Systeme*, Math. Z. **51** (1948), 197–216.

[100] R. Siefring, *Relative asymptotic behavior of pseudoholomorphic half-cylinders*, Comm.Ṗure Appl. Math. **61** (2008), no. 12, 1631–1684.

[101] R. Siefring, *Intersection theory of punctured pseudoholomorphic curves*, Geom. Topol. **15** (2011), no. 4, 2351–2457.

[102] S. Suhr, K. Zehmisch, *Linking and closed orbits*, arXiv:1305.2799v1.

[103] F. Spirig, J. Waldvogel, *Chaotic motion in Hill's lunar problem*, From Newton to chaos, 217–230, NATO Adv. Sci. Inst. Ser. B Phys. **336**, Plenum, New York (1995).

[104] B. Szökefalvi-Nagy, *Spektraldarstellungen linearer Transformationen des Hilbertschen Raumes*, Ergebnisse der Mathematik und ihrer Grenzgebiete. Berlin: Springer (1942).

[105] C. Taubes, *The Seiberg-Witten equations and the Weinstein conjecture*, Geom. Topol. **11** (2007), 2117–2202.

[106] T. Vozmischeva, *Integrable Problems of Celestial Mechanics in Spaces of Constant Curvature*, Kluwer Academic Publishers, Dordrecht (2003).

[107] A. Weinstein, *On the hypothesis of Rabinowitz periodic orbit theorems*, J. Diff. Eq. **33** (1979), 353–358.

[108] Z. Xia, *Melnikov Method and Transversal Homoclinic Points in the Restricted Three-Body Problem*, Jour. Diff. Eq. **96** (1992), 170–184.

[109] A. Zinger, *Pseudocycles and integral homology*, Trans. Amer. Math. Soc. **360** (2008), no. 5, 2741–2765.