Mémoire d'habilitation à diriger des recherches :

PERIODIC ORBITS IN SYMPLECTIC DYNAMICS

MARCO MAZZUCCHELLI

Soutenue le 10 juin 2021, à l'École normale supérieure de Lyon

Composition du jury :

MARIE-CLAUDE ARNAUD (Université de Paris)
PATRICK BERNARD (École normale supérieure)
ALBERT FATHI* (Georgia Institute of Technology)
ALEXANDRU OANCEA* (Sorbonne Université)
PEDRO SALOMÃO* (Universidade de São Paulo)
JEAN-CLAUDE SIKORAV (École normale supérieure de Lyon)

(∗ rapporteurs)

ENS DE LYON

*"Le texte de ce livre est trop formalisé à mon goût,
mais vous savez comme sont devenus les mathématiciens !
De plus, écrire dans la langue du poète W. Hamilton a été
pour moi une dure contrainte. Je crains, cher lecteur,
que vous ne pâtissiez des conséquences."*

Arthur Besse

"Questo libro non sono mai riuscito a terminarlo."

Corto Maltese

Contents

Preface

This memoir, which I am presenting for my French *habilitation à diriger des recherches*, is a promenade along the path that I followed as a researcher since my Ph.D. Its subject is the study of periodic orbits in a few interrelated settings that are part of the broad field of symplectic dynamics: autonomous and non-autonomous Hamiltonian systems, including geodesic flows and more general Reeb flows, and Hamiltonian systems with impacts such as billiards. These dynamical systems have a variational character, meaning that their orbits with suitable boundary conditions are critical points of different versions of the action functional from classical mechanics. Somehow for this reason, these systems are expected to often have many periodic orbits. Nevertheless, finding such periodic orbits is a complicated task, which required, over the course of more than a century, the development of sophisticated techniques of calculus of variations, Morse theory, Lusternik-Schnirelmann theory, and ultimately holomorphic curves in symplectic topology. In this memoir, I tried to put my modest contributions into context, by introducing in some details the different settings and by recalling the relevant state of the art. The study of periodic orbits is an overwhelmingly vast subject, and the choice of arguments in my text is certainly not meant to give a panorama of the field, but only to guide the reader through some of my research in a hopefully accessible way. With few exceptions, all the proofs provided should be intended as sketches, as I often tried to extract and condensate some of the ideas contained in a paper within the few pages of a section.

Chapter 1 is devoted to the study of Tonelli Hamiltonian systems, which are defined by Hamiltonians $H_t : \mathrm{T}^*M \to \mathbb{R}$ over the cotangent bundle of a closed manifold M whose restrictions to any fiber is suitably convex. The importance of this class cannot be overstated: Tonelli Hamiltonians appear in classical mechanics (and in particular in celestial mechanics, at least if one relaxes the compactness of the configuration space M), Aubry-Mather theory, and weak KAM theory. One of the remarkable properties of these Hamiltonians is that the associated dynamics can be defined in terms of dual Lagrangians and of their Lagrangian action functionals, which satisfy most of the common requirements from critical point theory. After recalling the generalities of the Tonelli setting, I will present several results on the existence and multiplicities of periodic orbits. In particular, I will give a very brief sketch of a main result from my Ph.D. thesis (Theorem 1.9): roughly speaking, a time periodic Tonelli Hamiltonian has infinitely many periodic orbits with low average action. Next, I will focus on autonomous Tonelli Hamiltonians, and to the celebrated problem of the existence of periodic orbits on a prescribed energy level. My contributions are mainly in the case of 2-dimensional configuration spaces M, and in particular I will summarize a couple of joint works that are particularly dear to me: the one with Asselle and Benedetti on minimal boundaries (Section 1.7), which provides in particular action minimizing periodic orbits related to the so-called Mather sets from Aubry-Mather theory, and the one with Abbondandolo, Macarini, and Paternain (Section 1.8) on the existence of infinitely many periodic orbits on almost every energy level in a suitable low range. At the end of the chapter (Section 1.9), I will present my joint work with Albers on non-convex billiards, which builds on seminal work of Benci-Giannoni.

Chapter 2 is devoted to the quest of closed geodesics on Riemannian and Finsler manifolds. Strictly speaking, this problem is a special instance of the one of periodic orbits on energy levels of autonomous Tonelli Hamiltonians, when the prescribed energy value is above one of the so-called Mañé critical values. Nevertheless, for historical reasons and due to its connections with Riemannian geometry, this setting is arguably the most important one, with its celebrated closed geodesics conjecture: *every closed Riemannian manifold of dimension at least 2 has infinitely many closed geodesics.* Such a conjecture fails in general if one replaces the Riemannian metric with a non-reversible Finsler one (I will briefly illustrate the celebrated counterexample due to Katok, following Ziller, in Section 2.3). My contributions to the problem of closed geodesics are in three directions. In a joint work with Asselle (Section 2.2), once again building on earlier work of Benci-Giannoni, we extended the celebrated Gromoll-Meyer theorem to a non-compact setting: a complete Riemannian manifold without close conjugate points at infinity and with sufficiently rich loop space homology has infinitely many closed geodesics. In a joint work with Suhr (Section 2.5), we proved a theorem claimed by Lusternik: in particular the result implies that, on a Riemannian 2-sphere, all the simple closed geodesics have the same length if and only if the metric is Zoll, that is, every geodesic is simple closed. Together with De Philippis, Marini, and Suhr, and building on earlier work of Grayson, Angenent, and Oaks, we established the properties of a curve shortening semi-flow for reversible Finsler surfaces (Section 2.4); this allowed us to extend to reversible Finsler 2-spheres the above mentioned result with Suhr, as well as the celebrated theorem of Bangert-Franks-Hingston: we now know that every reversible Finsler 2-sphere has infinitely many closed geodesics. In the last section of the chapter (Section 2.6), I will present my contributions to a variation of the closed geodesics problem: the problem of isometry-invariant geodesics, first studied by Grove.

In **Chapter 3**, the setting is the one of Reeb flows on closed contact manifolds, and the focus is on those Reeb flows all of whose orbits are closed. This is the generalization of the classical subject of Riemannian manifolds all of whose geodesics are closed, which was pioneered by Bott and grew considerably in the course of several decades (I was told as a Ph.D. student that every geometer should have a copy of Besse's "Manifolds all of whose geodesics are closed" on his shelf). I first came into the subject by trying to extend the already mentioned characterization of Zoll Riemannian metrics to higher dimensions. Instead, in a joint work with Cristofaro-Gariner, by employing Hutching's powerful machinery of embedded contact homology, we could provide an ultimate generalization of the characterization of Zoll Riemannian 2-spheres (Section 3.2): the closed Reeb orbits of a closed connected contact 3-manifold have a common period if and only if every Reeb orbit is closed (although not every orbit is required to have the same minimal period).

Remarkably, even for the special case of geodesic flows on Riemannian surfaces, this is a new statement that I would not be able to prove without the arsenal of embedded contact homology. Together with Cristofaro-Gardiner, we asked (or was it a conjecture?) whether such a result hold for higher dimensional contact manifolds as well; unfortunately, for this purpose, the higher dimensional holomorphic curves techniques are not quite as formidable as embedded contact homology, and such a statement seems out of reach. Nevertheless, together with Ginzburg and Gürel, we made a positive step by characterizing those restricted contact-type hypersurfaces (in particular, convex contact spheres, Section 3.3) and those unit tangent bundles (the geodesic setting, Section 3.4) all of whose Reeb orbits are closed in terms of an equality between suitable spectral invariants. The very end of the chapter (Section 3.5) concerns my very recent work on the structure of Besse contact manifold. I will present a joint result with Cristofaro Gardiner which asserts that two contact 3-manifolds all of whose Reeb orbits are closed and with the same prime action spectrum must be strictly contactomorphic. Together with the work on the classification of Seifert fibrations of Geiges and Lange, this implies that a contact 3-sphere all of whose Reeb orbits are closed must be strictly contactomorphic to a rational ellipsoid. Once again, it is a hard open question whether this statement hold in higher dimension. In a joint work with Radeschi, we showed that at least the convex contact spheres (of any odd dimension) "resemble" the rational ellipsoids: for any $\tau > 0$, the set of fixed points of the time-τ map of the Reeb flow is either empty or an integral homology sphere, and the sequence of Ekeland-Hofer spectral invariants coincides with the full sequence of elements in the action spectrum, each one repeated with a suitable multiplicity (as is the case for the ellipsoids).

A significant part of my research did not make it into this monograph. My "childhood" result with Cherubini on the combinatorial theory of inverse semigroups was too far to be integrated into a monograph on symplectic dynamics. Some of my papers, such as one on convex billiards and another one joint with a dream team (Abbondandolo, Asselle, Benedetti, and Taimanov) on non-exact magnetic flows on the 2-spheres were left out for the sake of brevity. The same goes for a result with Suhr on the non-equivariant spectral characterization of Zoll Riemannian metrics, that was actually seminal for my project with Ginzburg and Gürel. Finally, I ultimately decided not to include a full line of research on which I have been active lately together with Guillarmou and Tzou, on geometric inverse problems in Riemannian geometry; at least in spirit, the subject is related to the work that I presented on the characterization of Reeb flows all of whose orbits are closed.

Acknowledgments

First and foremost, I wish to thank Albert Fathi, Alex Oancea, and Pedro Salomão for accepting to be referees of this memoir on such a short notice, and Marie-Claude Arnaud, Patrick Bernard, and Jean-Claude Sikorav for accepting to be in the defense committee together with the referees. It was a great honor for me, and even more so considering their busy schedule.

I also take this occasion to thank some of the people I crossed paths with in the course of my (not anymore so short) career. Let me proceed in rough chronological order. I thank Alessandra Cherubini for encouraging me in pursuing the dream of becoming a mathematician when I was just an engineering undergraduate. I thank Alberto Abbondandolo for tolerating my oddities and my naiveté when I was his Ph.D. student, and for eventually becoming a collaborator and a dear friend; he set an example for me with his sharp thinking, intellectual honesty, and genuine interest in mathematics. I thank Yasha Eliashberg for giving me the opportunity of coming to Stanford for a year during my Ph.D., which turned out to be a crucial experience. In Stanford I met, among many other friends, Colin Guillarmou and Leo Tzou, who became family and eventually partners in crime in the game of geometric inverse problems. I thank Peter Albers for inviting me to the IAS in Princeton during my first year of post-doc, for two exciting weeks in which we ended up playing concave billiards together; we should get back to the game, by the way. I thank Albert Fathi for hiring me as a post-doc at the École normale supérieure de Lyon, the place that eventually became my home, for his encouragement, for his support, and for being a constant source of inspiration; as for many mathematicians of my generation, it is on his unpublished book (which I pre-ordered on Amazon over five years ago) that I learnt Hamiltonian dynamics. I thank Alfonso Sorrentino and Gabriele Benedetti, the best roomates ever, with whom I shared the legendary *Casa Italia* for a semester in Berkeley, so many math conversations, and a few joint papers. I thank my friend and collaborator Leonardo Macarini, with whom I had so much fun seeking periodic orbits and isometry-invariant geodesics in the coffee room at IMPA (quoting Miguel Abreu: "Everybody should have a collaborator in Rio de Janeiro"). I thank Luca Asselle, with whom it often sufficed to have an unplanned short phone call to setup a project for the next few months; I especially thank him for tolerating my email and sms bombing late at night concerning our lemmas. I thank Alexey Glutsyuk for inviting me to Moscow for a few exciting weeks (during which I was lodged right next to the former office of Anosov!). I thank Iskander Taimanov for making me feel home during my visit in Novosibirsk; what a honor it was to work together (with the already mentioned dream team) on magnetic closed geodesics. I have too many reasons to thank Stefan Suhr: for his numerous visits in Lyon, for his numerous invitations to come to Paris, Hamburg, and Bochum, for sharing the pain and the excitement of the quest of closed geodesics in any possible circumstance, including during our bike trips in the Alps. I always considered Nancy Hingston and Viktor Ginzburg as virtual advisors to me (I believe I read pretty much everything they published); I am deeply indebted to them for their teaching, for being sources of inspiration, and for their warm encouragement in the course of my career. Eventually I ended up becoming a collaborator of Viktor Ginzburg and Basak Gürel (one of the most famous duos in Hamiltonian dynamics), a thrilling experience which I hope to repeat soon. I thank Dan Cristofaro Gardiner for allowing me to catch a glimpse of the power of embedded contact homology. I thank my dear friend Gonzalo Contreras, who invited me for a month in his beautiful Guanajuato, and with whom I learnt so much on Tonelli Hamiltonians; this reminds me that I have been delaying our ongoing projects, and that I should get back to them right away. I have also been

delaying for too long the beginning of a project with my friend Richard Siefring. I thank Guido De Philippis for our great time as colleagues in Lyon, for checking the exams that I prepared for my differential geometry course ("It took Guido five minutes to solve all the problems: the exam is definitely too hard!"), and for demolishing any analysis question I came up with in my research. I thank my most recent collaborators, Marco Radeschi, Christian Lange, Luca Baracco, and Olga Bernardi, with whom I hope to write the next chapter of my journey. I thank Klaus Niederkruger, a colleague, a friend, and a former roomate, with whom I somehow never managed to work with. I thank Valentine Roos for her valuable feedback on this memoir. I thank Simon Allais, whom I am proud to call my first Ph.D. student, and who was the reason not to delay any further the writing of this memoir. I thank my colleagues at the UMPA (the best math department in the world); in particular Jean-Claude Sikorav, whose office is right next to mine, for his wisdom and for the many inspiring conversations we had over the last eight years. Outside the world of mathematics, I thank Alba for sharing the ups and downs, and for being a reference in life, a brother despite our respective last names. Finally, I thank Cécile for her suggestions on broken geodesics, and for sharing her life with an odd mathematician, without too many complaints.

Marco Mazzucchelli

Lyon, January 30, 2021.

CHAPTER 1

Tonelli Hamiltonian systems

This chapter is devoted to the study of periodic orbits in the class of Hamiltonian systems arising in classical mechanics [**Arn78**] and weak KAM theory [**CI99, Fat08, Sor15**], which are called Tonelli Hamiltonian systems. Part of the chapter, and specifically Sections 1.1, 1.2, 1.3, and 1.5, is devoted to setup the background, the notation, and some state of the art, in order to put our contributions into perspective.

1.1. Tonelli Hamiltonians and Lagrangians

The phase space of Tonelli Hamiltonian systems is a cotangent bundle T^*M equipped with the Liouville 1-form λ defined by

$$\lambda(w) = p(\mathrm{d}\pi(z)w), \qquad \forall z = (q,p) \in \mathrm{T}^*M,\ w \in \mathrm{T}_z(\mathrm{T}^*M). \tag{1.1}$$

Here, $\pi : \mathrm{T}^*M \to M$, $\pi(q,p) = q$ is the base projection. For us, the base M will always be a closed manifold M of dimension at least 2. The negative exterior differential $-\mathrm{d}\lambda$ is the canonical symplectic form of T^*M. A smooth function $H : \mathrm{T}^*M \to \mathbb{R}$ is called a **Tonelli Hamiltonian** when its restriction to any cotangent fiber $p \mapsto H(q,p)$ is both

- quadratically convex: the Hessian is positive definite at every point,
- superlinear: for every linear function $f : T_q^*M \to \mathbb{R}$, we have $H(q,\cdot) > f$ outside a compact set.

The energy levels $H^{-1}(e)$ of such a Hamiltonian are sometimes called "optical" in the literature, and are always compact. We shall consider the Hamiltonian dynamics defined by H. The Hamiltonian vector field X_H on T^*M is defined by

$$-\mathrm{d}\lambda(X_H, \cdot) = \mathrm{d}H. \tag{1.2}$$

Since each orbit of X_H stays on a compact energy level $H^{-1}(e)$, X_H defines an associated Hamiltonian flow that is complete, i.e.

$$\phi_H^t : H^{-1}(e) \to H^{-1}(e), \quad t, e \in \mathbb{R}.$$

Consider an orbit $z(t) = \phi_H^t(z(0))$ of this flow, which we can write as $z(t) = (q(t), p(t))$ with $q(t) \in M$ and $p(t) \in \mathrm{T}_{q(t)}M$. Equation (1.2) defining the Hamiltonian vector field can be rewritten as the system of first order ODEs

$$\begin{aligned} \dot q(t) &= \partial_p H(q(t), p(t)), \\ \dot p(t) &= -\partial_q H(q(t), p(t)), \end{aligned} \tag{1.3}$$

which are Hamilton's equation from classical mechanics. Notice that, while the second equation only makes sense in local coordinates, the first one for $\dot q$ is intrinsic: $\partial_p H$ is simply the differential of the restriction of H to a fiber of the cotangent bundle.

A remarkable feature of Tonelli Hamiltonians is that their dynamics on T^*M is indeed a second order dynamics on the base manifold M. Indeed, the fiberwise convexity and superlinearity of the Tonelli Hamiltonian H implies that that $\partial_p H$ is a diffeomorphism

$$\partial_p H : \mathrm{T}^*M \xrightarrow{\cong} \mathrm{T}^{**}M = \mathrm{T}M.$$

Therefore, we can rewrite Equation (1.3) as $p = (\partial_p H)^{-1}(q, \dot q)$, and infer that the momentum variable $p(t)$ is completely determined by the curve $q(t)$ in the configuration space M. This second order point of view is the one of the Lagrangian formulation of Hamiltonian dynamics. The dual **Tonelli Lagrangian** to H is the function

$$L : \mathrm{T}M \to \mathbb{R}, \qquad L(q,v) = \max_p \Big(pv - H(q,p) \Big).$$

Classical arguments from convex analysis imply that L is smooth, and indeed given by $L(q,v) = pv - H(q,p)$ with $p = (\partial_p H)^{-1}(q,v) = \partial_v L(q,v)$. Moreover, L has the same properties as H: it is both fiberwise quadratically convex and superlinear, and actually any smooth function on $\mathrm{T}M$ with these properties is the dual of a Tonelli Hamiltonian. Hamilton's equations can be rephrased in terms of the Tonelli Lagrangian as the second order ODE

$$\tfrac{\mathrm{d}}{\mathrm{d}t} \partial_v L(q(t), \dot q(t)) - \partial_q L(q(t), \dot q(t)) = 0, \tag{1.4}$$

which is the Euler-Lagrange equation from classical mechanics.

Example 1.1 (Riemannian geodesic flows). The simplest examples of dual Tonelli Hamiltonian and Lagrangian are the quadratic ones

$$H(q,p) = \tfrac12 \|p\|_g^2, \qquad L(q,v) = \tfrac12 \|v\|_g^2,$$

where g is a Riemannian metric on M, and $\|\cdot\|_g$ denotes the induced norms on tangent vectors and covectors. The Euler-Lagrange equation (1.4) of such a Lagrangian L is the geodesic equation $\nabla_t \dot q = 0$. On any energy level $H^{-1}(e)$ with $e > 0$, the Hamiltonian flow ϕ_H^t is the geodesic flow of (M, g). Its orbits have the form $(q(t), p(t))$, where q is a geodesic and $p = g(\dot q, \cdot)$ its dual velocity.

Example 1.2 (Finsler geodesic flows). For an arbitrary Tonelli Hamiltonian $H : \mathrm{T}^*M \to \mathbb{R}$, it turns out that, on energy levels $H^{-1}(e)$ with e sufficiently large, the Hamiltonian flow ϕ_H^t is always conjugate to a Finsler geodesic flow. Indeed, if e is large enough, any intersection $S_q^*M := \mathrm{T}_q^*M \cap H^{-1}(e)$ is a smooth positively curved sphere of dimension $\dim(M) - 1$ enclosing the origin $0 \in \mathrm{T}_q^*M$. Let $F : \mathrm{T}^*M \to [0, \infty)$ be a function such that $F(q, \lambda p) = \lambda$ for all $\lambda > 0$, $q \in M$, and $p \in S_q^*M$. The Hamiltonian flow ϕ_F^t on $F^{-1}(1)$ is precisely the geodesic flow of the Finsler metric dual to F. Since $F^{-1}(1) = H^{-1}(e)$, the orbits of the Hamiltonian flows ϕ_F^t and ϕ_H^t are the same up to reparametrization.

When the Hamiltonian has the form $H(q,p) = \frac{1}{2}\|p\|_g^2 + U(q)$, the dynamics on high energy levels is actually a Riemannian geodesic flow. Indeed, if $e > \max U$, the Finsler metric constructed before is simply the Riemannian norm $F(q,p) = \frac{1}{2}(e - U(q))^{-1/2}\|p\|_g$. □

In this chapter, we will also consider **non-autonomous Tonelli Hamiltonians**. These are families of Tonelli Hamiltonians H_t smoothly depending on $t \in \mathbb{R}$, whose associated time-dependent Hamiltonian vector field X_{H_t} has integral lines defined for all times $t \in \mathbb{R}$ (in the autonomous case, as we already remarked, this latter condition was automatically guaranteed by the compactness of the energy levels, which are invariant under the autonomous Hamiltonian flow). The vector field X_{H_t} will still define a Hamiltonian flow $\phi_H^t : \mathrm{T}^*M \to \mathrm{T}^*M$ with $\phi_H^0 = \mathrm{id}$. However, unlike in the autonomous case, such a flow may not satisfy $\phi_H^{s+t} = \phi_H^s \circ \phi_H^t$ for all $s, t \in \mathbb{R}$. We will ofter assume H_t to be periodic in time, of minimal period 1 without loss of generality, i.e. $H_t = H_{t+1}$; under this assumption the Hamiltonian flow satisfies

$$\phi_H^{t+1} = \phi_H^t \circ \phi_H^1, \qquad t \in \mathbb{R}.$$

Remark 1.3. According to Gronwall lemma, a sufficient condition for a non-autonomous Tonelli Hamiltonian H_t to define a global Hamiltonian flow is a bound of the form

$$\partial_t H_t \leq c\,(H_t + 1),$$

for some constant $c > 0$. □

A non-autonomous Tonelli Hamiltonian H_t has a dual **non-autonomous Tonelli Lagrangian** $L_t : \mathrm{T}M \to \mathbb{R}$ smoothly depending on $t \in \mathbb{R}$. As in the autonomous case, a curve $z(t) = (q(t), p(t))$ is an orbit of the Hamiltonian flow ϕ_H^t if and only if its base projection $q(t)$ is a solution of the (non-autonomous) Euler-Lagrange equation

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\partial_v L_t(q(t), \dot{q}(t)) - \partial_q L_t(q(t), \dot{q}(t)) = 0.$$

1.2. The Lagrangian action functional

In the next two sections we shall present the easiest among the results on the existence of Hamiltonian periodic orbits: those concerning non-autonomous Tonelli Hamiltonians. The proof of such results is based on a variational principle that we now recall.

Let H_t be a 1-periodic Tonelli Hamiltonian. We make one further assumption beyond the Tonelli one:

$$H_t(q,p) = \|p\|^2, \qquad \forall (q,p) \in \mathrm{T}^*M \setminus K, \tag{1.5}$$

where $K \subset \mathrm{T}^*M$ is a compact subset, and $\|\cdot\|$ is a norm on tangent covectors induced by an auxiliary Riemannian metric on M. Equivalently, the dual Lagrangian satisfies

$$L_t(q,v) = \|v\|^2, \qquad \forall (q,v) \in \mathrm{T}M \setminus K', \tag{1.6}$$

where $K' \subset \mathrm{T}^*M$ is a compact subset, and $\|\cdot\|$ is now the norm on tangent vectors induced by the same Riemannian metric as above. This assumption will allow us to avoid technical details, but is inessential: we will state all the results for general Tonelli Hamiltonians, even though we will sketch the proofs under the assumption (1.5); at the end of this section, we will explain how (1.5) can be completely relaxed.

Let us look for 1-periodic solutions $q : \mathbb{R} \to M$ of the Euler-Lagrange equation (1.4), that is, solutions that satisfy $q(t) = q(t+1)$ for all $t \in \mathbb{R}$; we briefly refer to such solutions as to 1-**periodic orbits**. Their lifts $z(t) = (q(t), p(t)) := (q(t), \partial_v L(q(t), p(t)))$ to the cotangent bundle T^*M are exactly the 1-periodic orbits of the Hamiltonian flow ϕ_H^t, i.e.

$$z(t+1) = z(t) = \phi_H^t(z(0)), \qquad \forall t \in \mathbb{R}.$$

According to the classical principle of stationary action, such orbits are critical points of the **Lagrangian action functional**

$$\mathcal{S} : \Lambda M \to \mathbb{R}, \qquad \mathcal{S}(q) = \int_{S^1} L_t(q(t), \dot{q}(t)) \, \mathrm{d}t,$$

where $S^1 = \mathbb{R}/\mathbb{Z}$ is the 1-periodic circle, and $\Lambda M := W^{1,2}(S^1, M)$ is the space of free loops that are absolutely continuous with square-integrable weak first derivative. A few remarks are in order here. The fact that $\mathcal{S}(q)$ is finite for any $q \in \Lambda M$ is a consequence of assumption (1.6). The fundamental theorem of calculus of variation readily implies that the critical points of $\mathcal{S}$ are weak 1-periodic solutions of the Euler-Lagrange equation (1.4). Finally, a bootstrap argument implies that weak 1-periodic solutions of (1.4) are smooth.

By means of the principle of stationary action, the dynamical problem of finding 1-periodic orbits is translated into the problem of detecting critical points of the Lagrangian action functional. In the course of the last century, since the seminal work of Poincaré, Birkhoff, Morse, Lusternik, and Schnirelmann, several powerful techniques have been developed to detect critical points of "well-behaved" abstract functional. The Lagrangian action functional $\mathcal{S}$ was indeed the functional that motivated the development of the theory, and thus satisfies the common requirements of abstract critical point theory:

• **(Complete domain)** Any auxiliary Riemannian metric on M induces a Riemannian metric on the loop space ΛM, which is the generalization of the inner product of the Sobolev space $W^{1,2}(S^1, \mathbb{R}^n)$. Equipped with such a metric, ΛM is a complete Hilbert manifold (see [**Pal63**]).

• **(Regularity)** The functional $\mathcal{S}$ is $C^{1,1}$, and twice differentiable in the sense of Gateaux (see [**AS09**]). If the restriction of the Lagrangian L to the fibers of $\mathrm{T}M$ is not a polynomial of degree 2, $\mathcal{S}$ is not C^2. Nevertheless, this lack of regularity is not essential: suitable finite dimensional reductions, first developed by Morse in the setting of geodesics [**Mil63**] and further extended to the whole Tonelli class [**Maz12**, Chap. 4], allow to apply to $\mathcal{S}$ all those results from critical point theory that would normally require the C^2, or even the C^∞, regularity.

• **(Compactness of the sublevel sets)** Since the Tonelli Lagrangian L_t is uniformly bounded from below, so is the action functional $\mathcal{S}$. Each sublevel set $\mathcal{S}^{-1}(-\infty, a]$, with $a \in \mathbb{R}$, is compact in a weak sense: any sequence q_n in the sublevel set that satisfies $\|\nabla \mathcal{S}(q_n)\|_{W^{1,2}} \to$

0 admits a converging subsequence. This property is often referred to as the Palais-Smale condition [**PS64**].

• **(Finite Morse indices)** The tangent spaces $\mathrm{T}_q(\Lambda M)$ are the Hilbert spaces of $W^{1,2}$ vector fields w along q such that $w(t) = w(t+1)$ for all $t \in \mathbb{R}$. The Hessian $\nabla^2\mathcal{S}(q)$ of the Lagrangian action functional at a critical point q is the bounded self-adjoint operator on $\mathrm{T}_q(\Lambda M)$ given by

$$\langle \nabla^2\mathcal{S}(q)w, w\rangle_{W^{1,2}} = \int_{S^1} \Big(\partial_{vv}L_t(q,\dot q)[\dot w,\dot w] + 2\partial_{qv}L_t(q,\dot q)[w,\dot w] + \partial_{qq}L_t(q,\dot q)[w,w]\Big)\,\mathrm{d}t.$$

An integration by parts and a bootstrap readily imply that the kernel of $\nabla^2\mathcal{S}(q)$ consists of those w that are solutions of the linearization of the Euler-Lagrange equation (1.4) along q. If $w \in \ker(\nabla^2\mathcal{S}(q))$, the curve

$$y(t) := \big(w(t), \partial_{vv}L(q,\dot q)\dot w + \partial_{qv}L(q,\dot q)w\big)$$

is a 1-periodic solution of the linearized Hamiltonian flow along $z = (q, \partial_v L(q,\dot q))$, i.e.

$$y(t+1) = y(t) = \mathrm{d}\phi_H^t(z(0))y(0), \qquad \forall t \in \mathbb{R}.$$

The **nullity** $\mathrm{nul}(q)$ is defined as

$$\mathrm{nul}(q) := \dim\ker(\nabla^2\mathcal{S}(q)) = \dim\ker(\mathrm{d}\phi_H^1(z(0)) - I). \tag{1.7}$$

The second equality readily implies that $\mathrm{nul}(q) \leq 2\dim(M)$. One can show that the Hessian operator $\nabla^2\mathcal{S}(q)$ is the sum of a positive-definite self-adjoint operator plus a compact one. A standard argument from spectral theory implies that the spectrum of $\nabla^2\mathcal{S}(q)$ consists of eigenvalues of finite geometric multiplicity, only finitely many of which are negative. The **Morse index** $\mathrm{ind}(q)$ is the finite non-negative integer

$$\mathrm{ind}(q) = \sum_{\lambda<0} \dim\ker(\nabla^2\mathcal{S}(q) - \lambda I).$$

Equivalently, $\mathrm{ind}(q)$ is the maximal dimension of a vector subspace of $\mathrm{T}_q(\Lambda M)$ over which $\nabla^2\mathcal{S}(q)$ is negative definite.

If the Tonelli Lagrangian L_t does not satisfy the extra condition (1.6), the associated Lagrangian action will only be a lower semi-continuous function $\mathcal{S} : \Lambda M \to \mathbb{R} \cup \{+\infty\}$, a regularity that is hardly suitable for critical point theory. Nevertheless, the following trick due to Abbondandolo and Figalli [**AF07**] allows to circumvent the issue at once. The crucial remark is that, for every $a \in \mathbb{R}$, there exists a compact subset $K \subset \mathrm{T}M$ with the following property: for every 1-periodic orbit q such that $\mathcal{S}(q) < a$, its lift $(q, \dot q)$ is contained in K. In order to study the periodic orbits in the action sublevel set $\mathcal{S}^{-1}(-\infty, a]$, we can introduce a new Tonelli Lagrangian L_t' that coincides with L_t on a compact subset $K' \subset \mathrm{T}M$ much bigger than K, and satisfies the extra condition (1.6). Let $\mathcal{S}'$ be the Lagrangian action functional of L'. If the compact subset K' was chosen large enough, one can show that all the 1-periodic orbits q for L' with action $\mathcal{S}'(q) \leq a$ have lift contained in the original compact subset K. Therefore, in order to study such orbits, one can employ the Lagrangian L_t' instead of L_t. Thanks to this remark, in the following we will be able to tacitly assume without loss of generality that all the Tonelli Lagrangians satisfy (1.6).

1.3. Existence of 1-periodic orbits

The properties of the Lagrangian action functional described in the previous section immediately imply an elementary existence result for 1-periodic orbits.

Theorem 1.4. *Let $L_t : \mathrm{T}M \to \mathbb{R}$ be a 1-periodic Tonelli Lagrangian. Every connected component of the free loop space $C \subset \Lambda M$ contains at least one 1-periodic orbit: a global minimizer of $\mathcal{S}|_C$.* □

Thus every 1-periodic Tonelli Lagrangian has 1-periodic orbits. It is a simple exercise in topology to see that the connected components of ΛM are in one-to-one correspondence with the conjugacy classes of the fundamental group $\pi_1(M)$. The connected component $C \subset \Lambda M$ corresponding to the trivial conjugacy class $\{1\} \subset \pi_1(M)$ is the one of **contractible** loops: for any $q \in C$ there exists a continuous map $u : B^2 \to M$, where $B^2 \subset \mathbb{C}$ is the closed unit ball, such that $q(t) = u(e^{i2\pi t})$. Theorem 1.4 has a straightforward corollary.

Corollary 1.5. *Let M be a closed manifold whose fundamental group has infinitely many conjugacy classes (e.g. any M with infinite abelian fundamental group). Any 1-periodic Tonelli Lagrangian $L_t : \mathrm{T}M \to \mathbb{R}$ has infinitely many non-contractible 1-periodic orbits.* □

A more difficult task consists in detecting infinitely many periodic orbits in manifolds with finite fundamental group. Clearly, this cannot be done simply by minimizing the Lagrangian action over some connected component of the free loop space. Instead, one needs to apply a recipe that goes back to Poincaré and Birkhoff [**Bir66**], and that can be roughly described as follows: if the sublevel set of a well-behaved function (in the sense of the previous section) has a rich topology, the function must have several critical points therein. In most of the applications, the non-trivial topology is detected by means of singular homology or cohomology. Here, we present a precise statement for the Lagrangian action functional using singular homology, as appeared in the work of Abbondandolo and Figalli [**AF07**], whereas a more general abstract result can be found in Viterbo [**Vit88**]. We denote the sublevel sets of the Lagrangian action functional by

$$\Lambda M^{<a} := \mathcal{S}^{-1}(-\infty, a), \qquad a \in \mathbb{R} \cup \{\infty\}.$$

We will not explicitly indicate the coefficients of singular homology groups, unless a specific choice is necessary.

Theorem 1.6. *Every non-zero homology class $h \in H_d(\Lambda M)$ defines a critical value $c(h)$ of the Lagrangian action functional $\mathcal{S}$ by*

$$c(h) := \inf \big\{ c \in \mathbb{R} \bigm| h \in \mathrm{im}\big(H_*(\Lambda M^{<c}) \xrightarrow{\mathrm{incl}_*} H_*(\Lambda M)\big)\big\},$$

where incl_* *denotes the homomorphism induced by the inclusion. Moreover, there is at least one critical point* $q \in \mathrm{crit}(\mathcal{S}) \cap \mathcal{S}^{-1}(c(h))$ *whose Morse indices satisfy*

$$\mathrm{ind}(q) \leq d \leq \mathrm{ind}(q) + \mathrm{nul}(q).$$

The critical values $c(h)$ provided by this theorem are often called **spectral invariants** in the literature. This terminology refers to the "action spectrum" of the Tonelli Lagrangian at period 1, which is the set of critical values of the Lagrangian action functional.

Proof. The fact that $c = c(h)$ is a critical value of $\mathcal{S}$ can be proved by contradiction: if not, the inclusion $\Lambda M^{<c-\epsilon} \hookrightarrow \Lambda M^{<c+\epsilon}$ would admit a homotopy inverse that can be constructed by "pushing down" $\Lambda M^{<c+\epsilon}$ with the flow of the anti-gradient $-\nabla\mathcal{S}$; however, this would imply that h can be represented by a cycle contained in $\Lambda M^{<c-\epsilon}$, contradicting the definition of $c(h)$.

For the second part of the theorem, let us first make a simplifying assumption: let us require the Lagrangian action functional $\mathcal{S}$ to be a Morse functional, that is, $\mathrm{nul}(q) = 0$ for all $q \in \mathrm{crit}(\mathcal{S})$. Under this assumption, we can invoke the classical Morse lemma [**Pal63**], which implies that we can identify a suitable neighborhood $U \subset \Lambda M$ of a critical point $q \in \mathrm{crit}(\mathcal{S}) \cap \mathcal{S}^{-1}(c)$ with a product $U^- \times U^+$, where $U^\pm$ is an open neighborhood of the origin in a Hilbert space $E^\pm$, with $\dim(E^-) = \mathrm{ind}(q)$ and $\dim(E^+) = \infty$. Under this identification, the critical point $q \in U$ corresponds to the origin $(0, 0) \in U^- \times U^+$, and the action functional takes the form

$$\mathcal{S}(x_-, x_+) = c - \|x_-\|^2 + \|x_+\|^2, \qquad \forall (x_-, x_+) \in U^- \times U^+.$$

A deformation argument implies that

$$H_*(U, U^{<c}) \cong H_*(U^-, U^- \setminus \{0\}).$$

Here, we employed the notation $U^{<c} := \Lambda M^{<c} \cap U$. The homology group $H_d(U^-, U^- \setminus \{0\})$ is non-zero if and only if $d = \mathrm{ind}(q)$.

Now, consider a non-zero homology class $h \in H_d(\Lambda M)$. The definition of $c(h)$ readily implies that $H_d(\Lambda M^{<c+\epsilon}, \Lambda M^{<c})$ is non-trivial for all $\epsilon > 0$. Since we are assuming that $\mathcal{S}$ is a Morse function, $\mathrm{crit}(\mathcal{S}) \cap \mathcal{S}^{-1}(c(h))$ contains only finitely many critical points $q_1, ..., q_k$. We consider a neighborhood U_i of q_i given by the Morse lemma. We choose the U_i's to be sufficiently small, so that they are pairwise disjoint. By taking $\epsilon > 0$ small and pushing down both $\Lambda M^{<c+\epsilon}$ and the U_i's with the anti-gradient flow of $-\nabla\mathcal{S}$, we obtain an isomorphism

$$H_d(\Lambda M^{<c+\epsilon}, \Lambda M^{<c}) \cong H_d(\Lambda M^{<c} \cup U, \Lambda M^{<c}),$$

where $U = U_1 \cup ... \cup U_k$. Finally, by the excision property,

$$H_d(\Lambda M^{<c} \cup U, \Lambda M^{<c}) \cong \bigoplus_{i=1,...,k} H_d(U_i, U_i^{<c}).$$

Therefore, at least one homology group $H_d(U_i, U_i^{<c})$ is non-trivial, and the corresponding critical point q_i has Morse index $\mathrm{ind}(q_i) = d$.

Let us now relax the non-degeneracy assumption that we made on $\mathcal{S}$. The main ingredient is the following genericity statement, which is a variation of the classical bumpy metric theorem from Riemannian geometry [**Ano82**]: for a C^∞ generic 1-periodic Tonelli Lagrangian, the Lagrangian action functional is Morse. This theorem gives us a sequence of Tonelli Lagrangians

$$L_n \xrightarrow{C^\infty} L$$

whose associated action functionals $\mathcal{S}_n$ are Morse. Notice that

$$\mathcal{S}_n^{-1}(-\infty, c-\epsilon_n) \subset \mathcal{S}^{-1}(-\infty, c) \subset \mathcal{S}_n^{-1}(-\infty, c+\epsilon_n), \qquad \forall c \in \mathbb{R}$$

where $\epsilon_n := \|L_n - L\|_{L^\infty}$. If we denote by $c_n(h)$ the spectral invariant defined by the Tonelli Lagrangian L_n, the above inclusions of sublevel sets readily imply $|c_n(h) - c(h)| < \epsilon_n$. Since $\epsilon_n \to 0$, we have $c_n(h) \to c(h)$. We already proved the theorem for the non-degenerate Tonelli Lagrangians: we know that there exist critical points $q_n \in \mathrm{crit}(\mathcal{S}_n) \cap \mathcal{S}_n^{-1}(c_n(h))$ with $\mathrm{ind}(q_n) = d$. A compactness argument implies that, up to a subsequence, q_n converges in C^∞ to a critical point $q \in \mathrm{crit}(\mathcal{S}) \cap \mathcal{S}^{-1}(c(h))$. Finally, the lower semi-continuity of the Morse index and the upper semi-continuity of the Morse index plus nullity imply that, for all n large enough,

$$\mathrm{ind}(q) \leq \mathrm{ind}(q_n) = \mathrm{ind}(q_n) + \mathrm{nul}(q_n) \leq \mathrm{ind}(q) + \mathrm{nul}(q).$$

This provides the index bounds claimed. □

In view of Theorem 1.6, in order to infer the existence of multiple 1-periodic orbits one needs a "rich" loop space homology. For simply connected manifolds, the following result of Vigué Poirrier and Sullivan [**VPS76**] provides the needed information.

Theorem 1.7 (Vigué Poirrier-Sullivan). *If M is a closed simply connected manifold, the rational loop space homology $H^d(\Lambda M;\mathbb{Q})$ is non-trivial in infinitely many degrees d.* □

A corollary of Theorem 1.7, originally due to Benci [**Ben86**] and extended to the full Tonelli class by Abbondandolo and Figalli [**AF07**], provides the multiplicity of 1-periodic orbit in a case not covered by Theorem 1.4.

Corollary 1.8 (Benci, Abbondandolo-Figalli). *Let M be a closed manifold with finite fundamental group. Any 1-periodic Tonelli Lagrangian $L_t : \mathrm{T}M \to \mathbb{R}$ has infinitely many contractible 1-periodic orbits q_n, $n \in \mathbb{N}$. Moreover, both the Morse index and the Lagrangian action diverge along this sequence, i.e.* $\mathrm{ind}(q_n) \to \infty$ *and* $\mathcal{S}(q_n) \to \infty$ *as* $n \to \infty$.

Proof. Without loss of generality, we can assume that M is simply connected. Otherwise, it is enough to prove the theorem for the lifted Tonelli Lagrangian $\widetilde{L}_t : \mathrm{T}\widetilde{M} \to \mathbb{R}$, where $\widetilde{M}$ is the (compact) universal cover of M. Indeed, only finitely many 1-periodic orbits of $\widetilde{L}$ project down to the same 1-periodic orbit of L_t, and such projected orbit must be contractible in M.

By Theorem 1.7, there is a sequence of positive integers $d_n \to \infty$ such that $H^{d_n}(\Lambda M;\mathbb{Q})$ is non-trivial. The min-max Theorem 1.6 thus provides a sequence of 1-periodic orbits q_n such that $\mathrm{ind}(q_n)+\mathrm{nul}(q_n)\to\infty$. Since $\mathrm{nul}(q_n)\leq 2\dim(M)$, we have $\mathrm{ind}(q_n)\to\infty$. This also implies that $\mathcal{S}(q_n)\to\infty$. Indeed, the Palais-Smale condition implies that, for every $c\in\mathbb{R}$, the critical sets $\mathrm{crit}(\mathcal{S})\cap\Lambda M^{<c}$ are compact, and the Morse index is uniformly bounded from above on every compact set. □

Corollaries 1.5 and 1.8 combined provide the existence of infinitely many 1-periodic orbits for a large class of closed manifolds M. For instance, for those closed manifolds with abelian fundamental group, such as the Lie groups. The only closed manifolds not covered by the results are those with fundamental group of infinite order, but with only finitely many conjugacy classes. To the best of the author's knowledge, the only known examples of such groups are not finitely presented, and therefore cannot be fundamental groups of closed manifolds.

1.4. Existence of periodic orbits of arbitrary integer period

The Tonelli Lagrangian $L_t:\mathrm{T}M\to\mathbb{R}$, $L_t(q,v)=\frac{1}{2}\|v\|^2_{g_t}$ given by a 1-periodic family of Riemannian metrics g_t has infinitely many trivial contractible periodic orbits: the constant curves. However, if we add to the Lagrangian a Morse function $U:M\to\mathbb{R}$, the resulting Tonelli Lagrangian $L_t(q,v)=\frac{1}{2}\|v\|^2_{g_t}+U(q)$ has only finitely many constant orbits: the critical points of U. Nevertheless, a remarkable phenomenon occurs: there are always infinitely many contractible periodic orbits of integer period, Morse index bounded from above by $\dim(M)$, and ratio action-period bounded from above by any value larger than the maximum of U.

We now provide the general statement governing this phenomenon. Let $L_t:\mathrm{T}M\to\mathbb{R}$ be a 1-periodic Tonelli Lagrangian. For any positive integer τ, we denote by

$$\Lambda^\tau M = W^{1,2}(\mathbb{R}/\tau\mathbb{Z},M)$$

the space of τ-periodic curves on M. We introduce the **average Lagrangian action functional**

$$\mathcal{S}^\tau:\Lambda^\tau M\to\mathbb{R},\qquad \mathcal{S}^\tau(q)=\frac{1}{\tau}\int_0^\tau L_t(q(t),\dot q(t))\,\mathrm{d}t.$$

Clearly, $\mathcal{S}^\tau$ enjoys the analogous variational principle as the ordinary Lagrangian action: its critical points are the τ-periodic orbits. In the following, we will add a subscript τ to the Morse indices and to the spectral invariants in order to specify that they are referred to the average action $\mathcal{S}^\tau$.

Notice that, with the notation of the previous sections, $\Lambda M=\Lambda^1 M$ and $\mathcal{S}=\mathcal{S}^1$. Moreover, we have an inclusion $\Lambda M\hookrightarrow\Lambda^\tau M$, and $\mathcal{S}=\mathcal{S}^\tau|_{\Lambda M}$. By applying the analysis of the previous section to each action functional $\mathcal{S}^\tau$, one detects periodic orbits of any given integer period τ. The crux of the matter in the problem concerning the multiplicity of periodic orbits of arbitrary period consists in determining whether, by varying τ, one really detects genuine new τ-periodic orbits, rather than periodic orbits already found for lower values of τ. A detected critical point of $\mathcal{S}^\tau$ can be recognized to be a new periodic orbit, for instance, if one can assert that τ is its minimal period. More frequently, the conclusion rather follows by looking at the average action or at the Morse indices.

The announced multiplicity results for periodic orbits of Tonelli Lagrangians is the following. In this form and modulo minor details, it was proved by the author [**Maz11a**] in his Ph.D. thesis, extending previous results of Long [**Lon00**] and Lu [**Lu09**]. The proof is based on techniques first introduced by Bangert and Klingenberg [**Ban80, BK83**] in their seminal work on closed geodesics.

Theorem 1.9. *Let $L_t : \mathrm{T}M \to \mathbb{R}$ be a 1-periodic Tonelli Lagrangian, and c_0 the maximal average action of the constant curves, i.e.*

$$c_0 = \max_{q \in M} \int_0^1 L_t(q, 0)\,\mathrm{d}t. \tag{1.8}$$

If L_t has only finitely many contractible 1-periodic orbits then, for every sufficiently large prime number τ, it has a contractible periodic orbit γ_τ of minimal period τ and average action $\mathcal{S}^\tau(\gamma_\tau) < c_0 + \epsilon_\tau$, where $\epsilon_\tau \to 0$ as $\tau \to \infty$. In particular, for any $\epsilon > 0$, L_t has infinitely many periodic orbits of average action less than $c_0 + \epsilon$.

Remark 1.10. The finiteness assumption on the contractible 1-periodic orbits of L_t cannot be removed. For instance, the only contractible periodic orbits of the autonomous Tonelli Lagrangian $L : \mathrm{T}\mathbb{T}^n \to \mathbb{R}$, $L(q, v) = \frac{1}{2}\|v\|^2$ are the constants. □

The proof of Theorem 1.9 will require an improvement of Theorem 1.6. Such an improvement involves the classical notion of **local homology** of a critical set, which for a critical point $q \in \mathrm{crit}(\mathcal{S}) \cap \mathcal{S}^{-1}(c)$ is defined as

$$C_*(q) := H_d(\Lambda M^{<c} \cup \{q\}, \Lambda M^{<c}).$$

Here, $\Lambda M^{<c} = \mathcal{S}^{-1}(-\infty, c)$. The terminology "local homology" suggests the fact that, by the excision property of singular homology, $C_*(q)$ only depends on the germ of $\mathcal{S}$ at the critical point q. When q is a non-degenerate critical point (i.e. $\mathrm{nul}(q) = 0$), the Morse lemma implies that $\mathcal{S}$ looks like a quadratic form of index $\mathrm{ind}(q)$ near q, and the local homology is

$$C_d(q) \cong \begin{cases} R, & d = \mathrm{ind}(q), \\ 0, & d \neq \mathrm{ind}(q), \end{cases}$$

where R is the coefficient ring employed. The situation is more involved when q is degenerate: the local homology $C_*(q)$ can have arbitrary finite rank, and even spread in several degrees from $\mathrm{ind}(q)$ to $\mathrm{ind}(q) + \mathrm{nul}(q)$. At least when q is isolated in $\mathrm{crit}(\mathcal{S})$, $C_*(q)$ is fully determined by the restriction of $\mathcal{S}$ to a central manifold N for the gradient vector field $\nabla\mathcal{S}$ at q. Such an N is an embedded submanifold $N \subset \Lambda M$ of finite dimension $\dim(N) = \mathrm{nul}(q)$, tangent to $\nabla\mathcal{S}$, containing q in its interior, and such that $\mathrm{T}_q N = \ker(\nabla^2\mathcal{S}(q))$. A generalization of the Morse lemma due to Gromoll and Meyer [**GM69a**] implies that we can identify a tubular neighborhood U of N with a product

$N \times W^- \times W^+$, where $W^\pm$ is an open neighborhood of the origin in a Hilbert space $E^\pm$, with $\dim(E^-) = \mathrm{ind}(q)$ and $\dim(E^+) = \infty$. Under this identification, the critical point $q \in U$ corresponds to $(q, 0, 0) \in N \times W^- \times W^+$, and the action functional takes the form

$$\mathcal{S}(x_0, x_-, x_+) = \mathcal{S}(x_0) - \|x_-\|^2 + \|x_+\|^2, \qquad \forall (x_0, x_-, x_+) \in N \times W^- \times W^+.$$

Namely, $\mathcal{S}$ locally looks like the "stabilization" of $\mathcal{S}|_N$ by a quadratic form of index $\mathrm{ind}(q)$. A standard arguments from non-linear analysis implies that

$$C_*(q) \cong H_{*-\mathrm{ind}(q)}(N^{<c} \cup \{q\}, N^{<c}),$$

where $N^{<c} := \Lambda M^{<c} \cap N$. Since the central manifold N has dimension $\mathrm{nul}(q)$, we readily infer that the local homology $C_d(q)$ can be non-zero only if $\mathrm{ind}(q) \leq d \leq \mathrm{ind}(q) + \mathrm{nul}(q)$.

The local homology groups are in some sense the "building blocks" for the global homology of ΛM. For instance, if c is the only critical value of $\mathcal{S}$ in $[c, c+\epsilon)$ and $K_c := \mathrm{crit}(\mathcal{S}) \cap \mathcal{S}^{-1}(c)$ is finite, then the inclusion

$$\Lambda M^{<c} \cup K_c \hookrightarrow \Lambda M^{<c+\epsilon} \tag{1.9}$$

turns out to be a homotopy equivalence. As usual, the homotopy inverse can be built by "pushing down" the sublevel set $\Lambda M^{<c+\epsilon}$ with the flow of $-\nabla\mathcal{S}$ (some care is needed here, as some points in $\Lambda M^{<c+\epsilon}$ will converge towards critical points in K_c asymptotically along the anti-gradient flow lines). By means of the inclusion (1.9), we obtain an isomorphism

$$\bigoplus_{q \in K_c} C_*(q) \cong H_*(\Lambda M^{<c+\epsilon}, \Lambda M^{<c}).$$

This discussion on the local homology implies the following.

Addendum to Theorem 1.6. *For each non-zero homology class $h \in H_d(\Lambda M)$, if the intersection $K_{c(h)} := \mathrm{crit}(\mathcal{S}) \cap \mathcal{S}^{-1}(c(h))$ is finite and some open interval of the form $(c(h), c(h)+\epsilon)$ does not contain critical values of $\mathcal{S}$, then there exists a critical point $q \in K_{c(h)}$ such that $\mathrm{ind}(q) \leq d \leq \mathrm{ind}(q) + \mathrm{nul}(q)$ and $C_d(q) \neq 0$.* □

Proof of Theorem 1.9. Since we are only interested in contractible periodic orbits, let us redefine $\Lambda^\tau M$ to be the connected component of contractible loops in $W^{1,2}(\mathbb{R}/\tau\mathbb{Z}, M)$. We begin by detecting contractible τ-periodic orbits of average action at most c_0 (the constant defined in (1.8)). In order to to this, observe that we have an inclusion $M \hookrightarrow \Lambda^\tau M$, which amounts to treating any point of M as a τ-periodic curve. Such an inclusion admits a left inverse $\Lambda^\tau M \to M$, $q \mapsto q(0)$. Therefore, the inclusion induces an injective homomorphism in homology

$$H_*(M) \xhookrightarrow{\mathrm{incl}_*} H_*(\Lambda^\tau M).$$

We choose the coefficients of the homology to be in $\mathbb{Z}_2$, so that the top-degree homology $H_d(M)$, $d = \dim(M)$, is non-trivial no matter if the closed manifold M is orientable or not. We thus consider the image of the fundamental class $[M] \in H_d(\Lambda^\tau M)$, whose spectral invariant satisfies $c := c_\tau([M]) \leq c_0$. The addendum to Theorem 1.6 implies that there exists a τ-periodic orbit $q_\tau \in \mathrm{crit}(\mathcal{S}^\tau)$ such that

$$\mathcal{S}^\tau(q_\tau) = c, \qquad \mathrm{ind}_\tau(q_\tau) \leq d \leq \mathrm{ind}_\tau(q_\tau) + \mathrm{nul}_\tau(q_\tau), \qquad C_d^\tau(q) \neq 0.$$

Here, we denoted by $C_*^\tau(q_\tau)$ the local homology of q_τ as a critical point of $\mathcal{S}^\tau$, i.e.

$$C_*^\tau(q) := H_*(\Lambda^\tau M^{<c} \cup \{q_\tau\}, \Lambda^\tau M^{<c}).$$

Let τ varies among the prime numbers, so that every periodic orbit q_τ has minimal period either τ or 1. If all but finitely many of the periodic orbits q_τ have minimal period larger than 1, we are done. Therefore, we are left to consider the case in which there exists a 1-periodic orbit q such that $q = q_\tau$ for infinitely many primes τ. We denote by $\mathbb{K} \subset \mathbb{N}$ the infinite subset of those prime numbers τ with this property. Let us show that the existence of such a periodic orbit q forces the existence of infinitely many other contractible periodic orbits of prime period and average action just above $\mathcal{S}(q)$.

We first focus on the Morse indices of q. The study of the behavior of the function $\tau \mapsto \mathrm{ind}_\tau(q)$ as the period τ varies is the subject of the so-called Bott's iteration theory [**Bot56, Lon02, Maz16**], whose importance in symplectic dynamics cannot be overstated. For the purpose of this theorem, we only need a simple statement that can be proved by hands without any symplectic consideration: for all $\tau \in \mathbb{K}$, since $C_d^\tau(q) \neq 0$, we have $\mathrm{ind}_\tau(q) \leq d$; the fact that $\mathrm{ind}_\tau(q)$ is uniformly bounded from above for infinitely many τ's actually implies that

$$\mathrm{ind}_\tau(q) = 0, \qquad \forall \tau \in \mathbb{N}.$$

The way to prove this assertion is the following. If $\mathrm{ind}(q) > 0$, then for τ large enough one could construct a τ-periodic vector field w such that $\langle \nabla^2 \mathcal{S}^\tau(q) w, w \rangle_{W^{1,2}} < 0$ and $w(0) = w(\tau) = 0$. By means of this vector field, we can show that $\mathrm{ind}_{\tau'}(q) > n$ for all $n \in \mathbb{N}$ and $\tau' \geq n\tau$. Indeed, consider the τ'-periodic vector fields $w_1, ..., w_n$ such that every w_i is supported in $[i\tau, (i+1)\tau]$ and satisfies $w_i|_{[i\tau,(i+1)\tau]} = w|_{[i\tau,(i+1)\tau]}$. Notice that

$$\langle \nabla^2 \mathcal{S}^{\tau'}(q) w_i, w_i \rangle_{W^{1,2}} < 0, \qquad \langle \nabla^2 \mathcal{S}^{\tau'}(q) w_i, w_j \rangle_{W^{1,2}} = 0, \qquad \forall i \neq j.$$

We conclude that the Hessian $\nabla^2 \mathcal{S}^{n\tau}(q)$ is negative definite over the vector subspace $\mathrm{span}\{w_1, ..., w_n\}$.

The study of the behavior of the function $\tau \mapsto \mathrm{nul}_\tau(q)$ is a matter of elementary linear algebra: if $z(t)$ if the orbit of the dual Hamiltonian H_t corresponding to $q(t)$, we have

$$\mathrm{nul}_\tau(q) = \dim\ker(\mathrm{d}\phi_H^\tau(z(0)) - I) = \sum_{\lambda \in \sqrt[\tau]{1}} \dim_{\mathbb{C}} \ker(\mathrm{d}\phi_H^1(z(0)) - \lambda I).$$

If τ is a large enough prime number, say larger than τ_0, the linear symplectic map $\mathrm{d}\phi_H^\tau(z(0))$ has no eigenvalue that is a τ-th root of the unity and is different from 1. Therefore, if we denote by $\mathbb{P}$ the set of all prime numbers larger than τ_0, we have

$$\mathrm{nul}_\tau(q) = \mathrm{nul}(q), \qquad \forall \tau \in \mathbb{P}. \tag{1.10}$$

From now on, τ will be a prime number in $\mathbb{P}$. We claim that the inclusion $\Lambda M \hookrightarrow \Lambda^\tau M$ induces a local homology isomorphism

$$C_*(q) \xrightarrow[\cong]{\text{incl}_*} C_*^\tau(q).$$

Indeed, let $N \subset \Lambda M$ be a central manifold for the anti-gradient $-\nabla\mathcal{S}$ at q. By (1.10) and since $\nabla\mathcal{S}^\tau|_{\Lambda M} = \nabla\mathcal{S}$, N will also be a central manifold for the anti-gradient $-\nabla\mathcal{S}^\tau$. Since $\mathrm{ind}(q) = \mathrm{ind}_\tau(q) = 0$, the inclusion induces the homology isomorphisms i_* and j_* in the commutative diagram

$$\begin{array}{ccc} H_*(N^{<c} \cup \{q\}, N^{<c}) & \xrightarrow[\cong]{i_*} & C_*(q) \\ & \underset{j_*}{\overset{\cong}{\searrow}} & \big\downarrow \text{incl}_* \\ & & C_*^\tau(q) \end{array}$$

The last ingredient that will allow us to complete the proof of the theorem is an instance of the "instability" of the local homology, a phenomenon that appears in the literature in several variations, not all equivalent to one another. For our problem, this property was first discovered by Bangert and Klingenberg [**Ban80, BK83**], but at about the same time an analogous argument appeared in Gromov's study of dilatations of maps [**Gro81**, Prop. 2.26]. The instability of local homology is the fact that, given any $\epsilon > 0$, for all τ large enough the inclusion induces the zero homomorphism

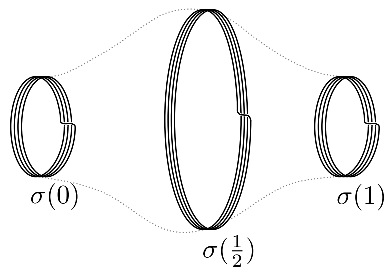
$$C_*(q) \xrightarrow{\ \text{incl}_*=0\ } H_*(\Lambda^\tau M^{<c+\epsilon}, \Lambda^\tau M^{<c}). \tag{1.11}$$

Assuming this for now, we can easily complete the proof of the theorem. Since our special periodic orbit q has isomorphic local homologies $C_*(q) \cong C_*^\tau(q)$, the instability tells us that the homomorphism induced by the inclusion $C_*^\tau(q) \to H_*(\Lambda^\tau M^{<c+\epsilon}, \Lambda^\tau M^{<c})$ is the zero one. This can happen only if $\mathcal{S}^\tau$ has a critical value in the interval $(c, c+\epsilon)$. We thus found a contractible τ-periodic orbit ζ_τ with average action $\mathcal{S}^\tau(\zeta_\tau) \in (c, c+\epsilon)$. If $\epsilon > 0$ is small enough so that, at period 1, the action functional $\mathcal{S}$ does not have critical values in $(c, c+\epsilon)$, we conclude that ζ_τ must be a new periodic orbit of minimal period τ.

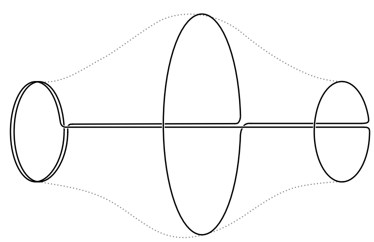
The proof of (1.11) is based on the so-called Bangert trick [**Ban80**]. Despite being technical, the main idea can be illustrated in a special situation: the one of local homology classes in $C_1(q)$ that are represented by maps

$$\sigma : [0,1] \to \Lambda M^{<c} \cup \{q\}$$

with $\mathcal{S}(\sigma(0)) < c$ and $\mathcal{S}(\sigma(1)) < c$. If $[\sigma] \neq 0$ in $C_1(q)$, then $\sigma(s) = q$ for some value of $s \in (0,1)$. For every $s \in [0,1]$, $\sigma(s)$ is an element in the 1-periodic loop space ΛM. Under the inclusion $\Lambda M \subset \Lambda^\tau M$, we treat $\sigma(s)$ as a τ-periodic curve (Figure 1.1(a)). As s varies from 0 to 1, σ "transports" the τ-copies of the loop $\sigma(0)|_{[0,1]}$ to the other end of the path, where they become τ-copies of $\sigma(1)|_{[0,1]}$. Consider instead the path $\tilde\sigma = \tilde\sigma_\tau : [0,1] \to \Lambda^\tau M$ that, roughly speaking, transport one loop $\sigma(0)|_{[0,1]}$ at the time to the other side $\sigma(1)|_{[0,1]}$ (Figure 1.1(b)). As every instant $s \in [0,1]$, $\tilde\sigma(s)$ consists of $\tau - 1$ loops that are either



(a)



(b)

FIGURE 1.1. **(a)** The path $\sigma : [0,1] \to \Lambda M^{<c} \cup \{q\}$ seen inside $\Lambda^\tau M$. The loops $\sigma(0)$ and $\sigma(1)$ have average action less than c. For some value of $s \in (0,1)$ we have $\sigma(s) = q$, and thus $\mathcal{S}^\tau(\sigma(s)) = c$. **(b)** The loop $\tilde\sigma(s)$ for an intermediate value of s in $[0,1]$.

$\sigma(0)|_{[0,1]}$ or $\sigma(1)|_{[0,1]}$, and a remaining portion of path that joins them. The action of $\tilde\sigma(s)$ can be estimated as

$$\mathcal{S}^\tau(\tilde\sigma(s)) \leq \frac{1}{\tau}\Big((\tau-1)\underbrace{\max\{\mathcal{S}(\sigma(0)),\mathcal{S}(\sigma(1))\}}_{<c} + \mathrm{const}\Big),$$

where const > 0 is independent of τ. If τ is large enough, then $\mathcal{S}^\tau \circ \tilde\sigma < c$, and therefore $[\tilde\sigma] = 0$ in $H_1(\Lambda^\tau M^{<c+\epsilon}, \Lambda^\tau M^{<c})$. The paths σ and $\tilde\sigma$ are homotopic relative to their endpoints, and one can find such a homotopy whose image stays inside the sublevel set $\Lambda^\tau M^{<c+\epsilon}$. This implies that $[\sigma] = [\tilde\sigma] = 0$ in $H_1(\Lambda^\tau M^{<c+\epsilon}, \Lambda^\tau M^{<c})$. $\square$

The instability of the local homology of the special periodic orbit q in the proof of Theorem 1.9 is rather mysterious from the dynamical point of view. From the variational point of view, the situation was well described by Hingston in her paper on the Conley conjecture [**Hin09**] (although the instability she deals with in the setting of Hamiltonian diffeomorphisms of tori is different from the current one). For any large prime period τ, the non-vanishing of the local homology $C^\tau_d(q)$ implies that the periodic orbit q is an "essential" critical point of the action functional $\mathcal{S}^\tau$; essential here means that such a critical point cannot be erased with a C^1-small perturbation of the functional $\mathcal{S}^\tau$ (an example of "non-essential" critical point would be the origin for the function $f(x) = x^3$). In the simplest

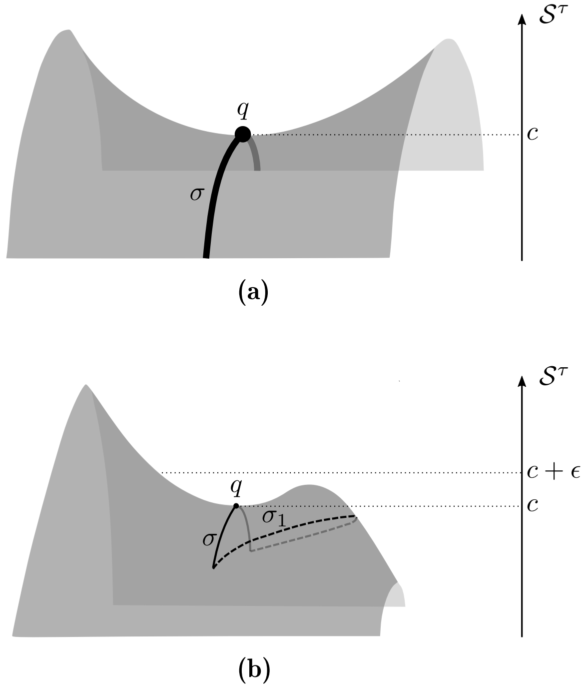


FIGURE 1.2. **(a)** A relative cycle σ generating a non-trivial element of the local homology $C_d(q) \cong C_d^\tau(q)$. **(b)** Instability of the local homology: within the sublevel set $\Lambda^\tau M^{<c+\epsilon}$, σ is homotopic to $\sigma_1 \subset \Lambda^\tau M^{<c}$.

possible situation, q is a degenerate saddle point of $\mathcal{S}^\tau$, and the generator of $C_d^\tau(q)$ is the relative fundamental class of a d-ball σ contained in the 1-periodic loop space $\Lambda M \subset \Lambda^\tau M$ as in Figure 1.2(a). As the period τ grows, q resembles more and more a non-essential critical point: for any $\epsilon > 0$, we can find τ large enough and a continuous deformation $\sigma_t \subset \Lambda^\tau M^{<c+\epsilon}$ from $\sigma_0 = \sigma$ to some $\sigma_1 \subset \Lambda^\tau M^{<c}$ (Figure 1.2(b)).

Figure 1.2(b) suggests that, as τ grows, there should be a critical point of $\mathcal{S}^\tau$ (the local maximum to the right of q in the figure) approaching more and more our original q. However, unlike the picture which is three dimensional, the domain of the functional $\mathcal{S}^\tau$ is infinitely dimensional, and the existence of such sequence of critical points is not at all asserted by the above arguments. Instead, what is asserted is the existence of a sequence of critical values $c_\tau > c$ of $\mathcal{S}^\tau$ such that $c_\tau \to c$ as $\tau \to \infty$.

$\star$ **Open problem:** Let $q \in \mathrm{crit}(\mathcal{S})$ be a 1-periodic orbit whose local homology is unstable, i.e. $C_d(q) \neq 0$ at some positive degree $d > 0$, and $\mathrm{ind}_\tau(q) = 0$ for all $\tau \in \mathbb{N}$. Is q non-isolated in the space of periodic orbits with arbitrary integer periods? Or at least, if $z := (q(0), \partial_v L(q(0), \dot q(0)))$ is the corresponding fixed point of the Hamiltonian diffeomorphism ϕ_H^1 for the dual Tonelli Hamiltonian H_t, does any neighborhood $U \subset \mathrm{T}M$ of z contain infinitely many periodic points of ϕ_H^1?

1.5. The free-period action functional

We now consider an autonomous Tonelli Hamiltonian $H : \mathrm{T}^*M \to \mathbb{R}$, and we address the existence of periodic orbits on a given energy hypersurface $H^{-1}(e)$, $e \in \mathbb{R}$. If we also prescribe the period of the orbits that we look for, the problem becomes overdetermined: in general, there are no periodic orbits with both given energy e and period τ. Therefore, the periodic orbits that we found will have arbitrary period $\tau \in (0, \infty)$. As we will see shortly, this problem is more involved than the existence of periodic orbits with given period or arbitrary integer period which we discussed in the previous sections.

Once again, the problem can be studied by a version of the stationary-action variational principle. Let $L : \mathrm{T}M \to \mathbb{R}$ be the Tonelli Lagrangian dual to H. It will be convenient to transport the Hamiltonian to the tangent bundle $\mathrm{T}M$ by means of the diffeomorphism $\partial_v L$. Namely, we introduce the energy function

$$E : \mathrm{T}M \to \mathbb{R}, \qquad E(q, v) = H(q, \partial_v L(q, v)) = \partial_v L(q, v)v - L(q, v).$$

Notice that, if $z(t) = (q(t), p(t)) \in H^{-1}(e)$ is an orbit of the Hamiltonian flow ϕ_H^t, then $E(q, \dot{q}) \equiv e$. The domain in which we will work consists of periodic curves of any possible period. Such a space can be formally obtained as the product $(0, \infty) \times \Lambda M$, where as before $\Lambda M = W^{1,2}(S^1, M)$ and $S^1 = \mathbb{R}/\mathbb{Z}$. We identify a pair $(\tau, q) \in (0, \infty) \times \Lambda M$ with the τ-periodic curve $\gamma : \mathbb{R}/\tau\mathbb{Z} \to M$, $\gamma(t) = q(t/\tau)$; in the following, we will simply write this identification as $\gamma = (\tau, q)$. For a given energy value $e \in \mathbb{R}$, we introduce the **free-period action functional**

$$\mathcal{S}_e : (0, \infty) \times \Lambda M \to \mathbb{R} \cup \{\infty\}, \qquad \mathcal{S}_e(\tau, q) = \tau \int_0^1 L(q(t), \dot{q}(t)/\tau)\, \mathrm{d}t + \tau e.$$

It is perhaps more informative to express the value $\mathcal{S}_e(\tau, q)$ as the Lagrangian action of the τ-periodic curve $\gamma = (\tau, q)$ with respect to the Lagrangian $L + e$, i.e.

$$\mathcal{S}_e(\tau, q) = \mathcal{S}_e(\gamma) = \int_0^\tau \Big(L(\gamma(t), \dot{\gamma}(t)) + e \Big)\, \mathrm{d}t.$$

The critical points of $\mathcal{S}_e$ are precisely the τ-periodic solutions γ of the Euler-Lagrange equation of L with energy e, i.e.

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\partial_v L(\gamma, \dot{\gamma}) - \partial_q L(\gamma, \dot{\gamma}) = 0, \\ E(\gamma, \dot{\gamma}) = e. \end{cases}$$

The circle S^1 acts on the loop space ΛM by translation:

$$t \cdot q = q(t + \cdot), \qquad \forall t \in S^1,\ q \in \Lambda M.$$

Since the Lagrangian L is autonomous, the functional $\mathcal{S}_e$ is invariant under this action. In particular, every critical point $(\tau, q) \in \mathrm{crit}(\mathcal{S}_e)$ with non-constant q belongs to a critical circle $S^1 \cdot (\tau, q) \subset \mathrm{crit}(\mathcal{S}_e)$.

Most of the properties enjoyed by the fixed-period action functional (see Section 1.2) are still enjoyed by the free-period action functional $\mathcal{S}_e$:

• **(Regularity)** If we focus on the energy level $H^{-1}(e)$, we are at liberty to modify the Hamiltonian H far away from the energy level, and in particular make it quadratic at infinity as in Equation (1.5). The dual Lagrangian L will also become quadratic at infinity as in Equation (1.6). The free-period action functional $\mathcal{S}_e$ of our Tonelli Lagrangian quadratic at infinity is everywhere finite, and indeed $C^{1,1}$ and twice Gateaux differentiable. Finite dimensional techniques developed by Asselle and the author in [**AM19**, Sect. 3] allow to apply to $\mathcal{S}_e$ all those variational methods that normally require the C^2, or even the C^∞, regularity.

• **(Complete domain)** The domain $(0,\infty)\times\Lambda M$ is a Hilbert manifold. We equip it with a product Riemannian metric that is Euclidean on the factor $(0,\infty)$, and is the $W^{1,2}$ metric on ΛM induced by an auxiliary Riemannian metric on M. With this choice, $(0,\infty)\times\Lambda M$ is not complete, as there are Cauchy sequences coverging towards $\{0\}\times\Lambda M$. This is not really an issue in the applications: an argument due to Asselle implies that a sequence $(\tau_n,q_n)\in(0,\infty)\times\Lambda M$ with $\tau_n\to 0$, $\|\nabla\mathcal{S}_e(\tau_n,q_n)\|\to 0$ and $\mathcal{S}_e(\tau_n,q_n)\to c\in\mathbb{R}$ exists only if $c=0$. Therefore the lack of completeness will not manifest itself when working above the level zero of $\mathcal{S}_e$.

• **(Morse indices)** If we freeze the period variable to a specific $\tau\in(0,\infty)$, the restriction $\mathcal{S}_e(\tau,\cdot)$ is essentially the fixed-period action functional plus the constant $e\tau$. Since this latter functional has finite Morse indices, the same will be true for the free-period action functional $\mathcal{S}_e$. If the energy e is a regular value of the Hamiltonian H, a computation analogous to the one in Section 1.2 allows to describe the nullity of a critical point $\gamma=(\tau,q)\in\mathrm{crit}(\mathcal{S}_e)$ in terms of the linearized Hamiltonian flow: if $\Sigma:=\mathrm{T}_zH^{-1}(e)$, and $z:=(\gamma(0),\partial_vL(\gamma(0),\dot\gamma(0)))$ is the fixed point of ϕ_H^τ corresponding to γ, then

$$\mathrm{nul}(\tau,q):=\dim\ker(\nabla^2\mathcal{S}_e(\tau,q))=\dim\ker(\mathrm{d}\phi_H^\tau(z)|_\Sigma-I).$$

It remains one property that does not always hold for $\mathcal{S}_e$: the compactness of the sublevel sets. Actually, $\mathcal{S}_e$ is even unbounded for below for low values of e: it suffices to take $e<-\max L(\cdot,0)$, and any constant curve $q\in\Lambda M$, $q\equiv q_0\in M$ will give us an unbounded line $\mathcal{S}_e(\tau,q)\to-\infty$ as $\tau\to\infty$. On the other hand, if we choose e to be large enough so that $L+e$ is everywhere positive, the functional $\mathcal{S}_e$ will be positive as well. This shows that the behavior of $\mathcal{S}_e$ depends strongly on the energy value e. In the next section, we shall briefly illustrate this dependence. We refer the reader to the article of Contreras [**Con06**] and to the survey of Abbondandolo [**Abb13**] for a comprehensive account.

1.6. Periodic orbits on energy hypersurfaces

Three (possibly coinciding) values of the energy e mark significant changes in the properties of the free-period action functional $\mathcal{S}_e$, which reflect changes on the dynamical properties of the Hamiltonian flow ϕ_H^t on the energy hypersurface $H^{-1}(e)$. The smallest such value is

$$e_0(L) := \max_{q\in M} E(q,0).$$

Since the Tonelli Hamiltonian H is fiberwise convex, and the fiberwise derivative $\partial_p H$ is a diffeomorphism, we must have $E(q,0) = H(q,p)$ for the unique value of $p = p_q$ that minimizes the function $p \mapsto H(q,p)$. This shows that $e_0(L)$ is the largest value with the property that, for all energy values $e < e_0(L)$, the energy hypersurface $H^{-1}(e)$ does not intersects all the fibers of T^*M.

The second and third significant energy values are

$$c_{\mathrm{u}}(L) := \inf\big\{e\in\mathbb{R} \bigm| \mathcal{S}_e(\gamma) > 0,\ \forall \text{ contractible } \gamma\big\},$$
$$c_0(L) := \inf\big\{e\in\mathbb{R} \bigm| \mathcal{S}_e(\gamma) > 0,\ \forall \text{ null-homologous } \gamma\big\}.$$

Here, contractible as usual means that $[\gamma] = 0$ in the fundamental group $\pi_1(M,\gamma(0))$, whereas null-homologous means that $[\gamma] = 0$ in the homology group $H_1(M;\mathbb{Z})$. The energy values $c_{\mathrm{u}}(L)$ and $c_0(L)$ are called the **Mañé critical values** of the universal cover and of the universal abelian cover respectively. The ordinary critical value of the Tonelli Lagrangian L, which was introduced by Mañé in his seminal work [**Mañ97**] on Aubry-Mather theory, is defined by

$$c(L) := \inf\big\{e\in\mathbb{R} \bigm| \mathcal{S}_e^{-1}(-\infty,0) = \varnothing\big\}.$$

Its variations $c_{\mathrm{u}}(L)$ and $c_0(L)$ are thus the ordinary critical values of the lifted Lagrangians $L_{\mathrm{u}} : \mathrm{T}M_{\mathrm{u}} \to \mathbb{R}$ and $L_0 : \mathrm{T}M_0 \to \mathbb{R}$ respectively, where $M_{\mathrm{u}} \to M$ is the universal cover and $M_0 \to M$ is the universal abelian cover. Unlike $c_{\mathrm{u}}(L)$ and $c_0(L)$, the ordinary $c(L)$ does not play a particular role in the study of the multiplicity of periodic orbits on energy hypersurfaces.

The three energy values are ordered as

$$e_0(L) \leq c_{\mathrm{u}}(L) \leq c_0(L).$$

The second inequality is simply due to the fact that contractible curves are nullhomologous. The first inequality follows by the fact that, if we fix a point $q \in M$ such that $E(q,0) = e_0(L)$, and we consider the constant curve $\gamma = (\tau, q) \in (0,\infty)\times\Lambda M$, then

$$\mathcal{S}_{e_0}(\tau,q) = \tau\big(L(q,0) + e_0(L)\big) = \tau\big(e_0(L) - E(q,0)\big).$$

Example 1.11.

(i) For a purely Riemannian Lagrangian $L(q,v) = \frac{1}{2}\|v\|_g^2$, we have $E(q,v) = \frac{1}{2}\|v\|_g^2$ and

$$\min E = e_0(L) = c_{\mathrm{u}}(L) = c_0(L) = 0.$$

(ii) In order to separate e_0 from the minimum of the energy, it is enough to consider a mechanical Lagrangian $L(q,v) = \frac{1}{2}\|v\|^2 - U(q)$ with a non-constant potential U, so that $E(q,v) = \frac{1}{2}\|v\|_g^2 + U(q)$ and

$$\min U = \min E < e_0(L) = c_{\mathrm{u}}(L) = c_0(L) = \max U.$$

(iii) In order to separate e_0 from c_0, the Lagrangian must have a magnetic term. For instance, if $L(q,v) = \frac{1}{2}\|v\|_g^2 + \theta_q(v)$ where θ is a non-exact 1-form on M, then $E(q,v) = \frac{1}{2}\|v\|_g^2$, and a result of Contreras, Iturriaga, G. Paternain, and M. Paternain [**CIPP98**] implies

$$0 = \min E = e_0(L) < c_0(L) = \inf_{u\in C^\infty(M)} \|\theta + \mathrm{d}u\|_{L^\infty}. \qquad \square$$

Remark 1.12. In order for the values $c_{\mathrm{u}}(L)$ and $c_0(L)$ to be different, the notions of being "contractible" or "nullhomologous" for loops in M must be distinct. Namely, $c_{\mathrm{u}}(L) \neq c_0(L)$ only if the fundamental group $\pi_1(M)$ is non-abelian. $\square$

Remark 1.13. The space of Tonelli Lagrangians on $\mathrm{T}M$ has a C^1-dense subspace $\mathcal{U}$ such that $e_0(L) < c_{\mathrm{u}}(L)$ for all $L \in \mathcal{U}$, see [**ABM17**, Section 4]. $\square$

When dealing with a Tonelli Hamiltonian H, we will write $e_0(H)$, $c_{\mathrm{u}}(H)$, and $c_0(H)$ to denote the corresponding energy values of the dual Lagrangian L, i.e.

$$e_0(H) := e_0(L), \qquad c_{\mathrm{u}}(H) := c_{\mathrm{u}}(L), \qquad c_0(H) = c_0(L).$$

We already pointed out in Example 1.2 that the Hamiltonian dynamics on $H^{-1}(e)$ for large values of e is of Finsler type. The following is a more precise statement, due to Contreras, Iturriaga, G. Paternain, and M. Paternain [**CIPP98**].

Theorem 1.14 (Contreras-Iturriaga-Paternain[2]). *For all energy values $e > c_0(H)$, the Hamiltonian flow $\phi_H^t|_{H^{-1}(e)}$ is orbitally equivalent to the geodesic flow of a Finsler metric on M. Namely, there exists a Finsler metric $F : \mathrm{T}^*M \to [0,\infty)$ and a diffeomorphism $\psi : H^{-1}(e) \to F^{-1}(1)$ mapping orbits of $\phi_H^t|_{H^{-1}(e)}$ to orbits of $\phi_F^t|_{F^{-1}(1)}$.*

Proof. The crucial ingredient for the proof is the following characterization of the critical value $c_0(L)$ in terms of subsolutions of the Hamilton-Jacobi equation: for each $e > c_0(L)$ there exists a closed 1-form β on M such that $H(q,\beta_q) < e$. Since β is closed, the diffeomorphism $\psi : \mathrm{T}^*M \to \mathrm{T}^*M$, $\psi(q,p) = (q,\beta_q)$ is symplectic (namely, $\psi^*\mathrm{d}\lambda = \mathrm{d}\lambda$, where λ is the Liouville 1-form (1.1)). Therefore, if we set $K := H \circ \psi$, the Hamiltonian flows $\phi_K^t|_{K^{-1}(e)}$ and $\phi_H^t|_{H^{-1}(e)}$ are orbitally equivalent. Finally, since $K^{-1}(e)$ is a hypersurface that encloses the 0-section, there exists a Finsler metric $F : \mathrm{T}^*M \to [0,\infty)$ such that $F^{-1}(1) = K^{-1}(e)$, which implies that $\phi_F^t|_{F^{-1}(1)}$ and $\phi_K^t|_{K^{-1}(e)}$ have the same orbits up to time reparametrization. $\square$

In view of Theorem 1.14, the study of periodic orbits on energy levels $H^{-1}(e)$ with $e > c_0(H)$ reduces to the study of closed geodesics on closed Finsler manifolds. We will postpone a more detailed discussion concerning closed geodesics to Chapter 2. Here, we just mention that, for a vast class of closed manifold M, there are always infinitely many periodic orbits on every energy levels $H^{-1}(e)$ with $e > c_0(H)$. However, when M is a sphere S^n, a projective space $\mathbb{CP}^n$ or $\mathbb{HP}^n$, or the Cayley projective plane $\mathrm{Ca}\mathbb{P}^2$, a construction due to Ziller [**Zil83**], and based on an earlier result of Katok [**Kat73**] for the 2-sphere, provides a Tonelli Hamiltonian $H : \mathrm{T}^*M \to \mathbb{R}$ such that $H^{-1}(e)$ contains only finitely many periodic orbits for some $e > c_0(H)$. We shall present such construction in the next chapter, in Example 2.10.

More generally, when $e > c_{\mathrm{u}}(L)$, a simple argument shows that the free-period action functional is bounded from below on every connected component of its domain $(0, \infty) \times \Lambda M$. Moreover, the free-period action functional $\mathcal{S}_e$ has sufficiently compact sublevel sets: any sequence $\gamma_n = (\tau_n, q_n)$ in its domain such that $\|\nabla \mathcal{S}_e(\gamma_n)\| \to 0$, $\mathcal{S}_e(\gamma_n) \to c$, and with τ_n bounded from below by a positive constant, admits a converging subsequence. In particular, if M is not simply connected, we recover the simple existence result analogous to Theorem 1.4.

Theorem 1.15. *Let $L : \mathrm{T}M \to \mathbb{R}$ be a Tonelli Lagrangian. For each energy value $e > c_{\mathrm{u}}(L)$ and for every connected component $C \subset (0, \infty) \times \Lambda M$ other than the one of contractible loops, there exists a periodic orbit that is a global minimizer of $\mathcal{S}_e|_C$.* □

On lower energy levels $e < c_{\mathrm{u}}(L)$, the free-period action functional is unbounded from below on every connected component of its domain: for instance, on the connected component of contractible loops, we can find a γ with negative action $\mathcal{S}_e(\gamma) < 0$, and by iterating γ one obtain a sequence of loops with diverging negative action. In view of this unboundedness, one may try to work on suitable strips $\mathcal{S}_e^{-1}[a, b]$. Unfortunately, though, it is not known whether $\mathcal{S}_e$ satisfies the Palais-Smale condition: there might be sequences $\gamma_n = (\tau_n, q_n) \in \mathcal{S}_e^{-1}[a, b]$ such that $\|\nabla \mathcal{S}_e(\gamma_n)\| \to 0$ but $\tau_n \to \infty$. It is currently not known how to control the period variable of such Palais-Smale sequences for any given value of e, but a formidable trick due to Struwe [**Str90**] allows to do it in certain situations provided one is allowed to perturb the energy value e. We will briefly discuss the ideas behind Struwe's argument in the proof of the next statement, which is due to Contreras [**Con06**].

Theorem 1.16 (Contreras). *Let $L : \mathrm{T}M \to \mathbb{R}$ be a Tonelli Lagrangian such that $e_0(L) < c_{\mathrm{u}}(L)$. For almost every $e \in (e_0(L), c_{\mathrm{u}}(L))$, there is a contractible periodic orbit γ of energy e and positive action $\mathcal{S}_e(\gamma) > 0$.*

Proof. The idea of the proof consists in showing that the graph of the free-period action functional $\mathcal{S}_e$ presents a mountain pass geometry around the subspace of constant loops. More precisely, for each $\delta > 0$, consider the open subset

$$\mathcal{U}_\delta := \big\{ (\tau, q) \in (0, 1) \times \Lambda M \ \big| \ \|\dot{q}\|_{L^2} \leq \tau\delta, \ \tau < \delta \big\}.$$

Notice that every $\gamma = (\tau, q) \in \mathcal{U}$ has length less than δ^2. Therefore, up to choosing $\delta > 0$ small enough, $\mathcal{U}_\delta$ is contained in the connected components of contractible loops. From now on, we will implicitly require δ to be small enough to satisfy this assertion.

For every $e > e_0(L)$, a computation shows that

$$\lim_{\delta \to 0^+} \sup_{\mathcal{U}_\delta} \mathcal{S}_e = 0,$$

and, if $\delta > 0$ is small enough,

$$b(e,\delta) := \inf_{\partial \mathcal{U}_\delta} \mathcal{S}_e > 0.$$

Notice that $\mathcal{S}_{e'}(\tau,q) = \mathcal{S}_e(\tau,q) + (e'-e)\tau$. Therefore, if we fix $e' \in (e_0(L), c_{\mathrm{u}}(L))$, we can find $\delta > 0$ and an open neighborhood $I = (e'-\epsilon, e'+\epsilon) \subset (e_0(L), c_{\mathrm{u}}(L))$ such that

$$b(\delta) := \inf_{e\in I} b(e,\delta) > 0.$$

We recall that, if $e < c_u(L)$, there exists $\gamma = (\tau,q) \in (0,\infty) \times \Lambda M$ that is contractible (and thus in the same connected component as $\mathcal{U}_\delta$) and has negative action $\mathcal{S}_e(\gamma) < 0$. For each energy value $e \in I$, we denote by $\mathcal{W}_e$ the (non-empty) family of continuous maps $w : [0,1] \to (0,\infty) \times \Lambda M$ such that $w(0) \in \mathcal{U}_\delta$, $\mathcal{S}_e(w(0)) < b(\delta)$, $w(1) \not\in \mathcal{U}_\delta$, and $\mathcal{S}_e(w(0)) < 0$. We employ this family to define a min-max

$$c(e) := \inf_{w\in\mathcal{W}_e} \max \mathcal{S}_e \circ w \geq b(\delta).$$

If $\mathcal{S}_e$ satisfied the Palais-Smale condition, by the classical min-max theorem from non-linear analysis (which is analogous to Theorem 1.6) we would readily infer that $c(e)$ is a critical value of $\mathcal{S}_e$, and we would have found a periodic orbits with positive action on the energy level e. Since we do not know whether the Palais-Smale condition hold, a priori we may have sequences $w_n \in \mathcal{W}_e$ and $s_n \in [0,1]$ with the following property: if $(\tau_n, q_n) := w_n(s_n)$, then $\mathcal{S}_e(\tau_n,q_n) \to c(e)$, $\|\nabla \mathcal{S}_e(\tau_n,q_n)\| \to 0$, but $\tau_n \to \infty$. The Palais-Smale condition amounts indeed to forbid the sequence of periods τ_n to diverge in this kind of sequences.

Struwe's trick allows to bound the sequence of periods and recover the Palais-Smale condition for almost every $e \in I$. The rough idea goes as follows. Notice first that $\mathcal{S}_e$ is pointwise monotonically increasing in e, and the family $\mathcal{W}_e$ gets bigger as e gets smaller. This readily implies that $e \mapsto c(e)$ is a monotone increasing function, and in particular almost everywhere differentiable according to Lebesgue's theorem. Now, notice that the period τ could be recovered from the action values $\mathcal{S}_e(\tau,q)$ by differentiating with respect to the energy parameter, i.e. $\partial_e \mathcal{S}_e(\tau,q) = \tau$. Suitably elaborated, these observations allow to bound the period variable along suitable sequences $w_n \in \mathcal{W}_e$ such that $\max \mathcal{S}_e \circ w_n \to c(e)$, for those values of $e \in I$ at which the function $e \mapsto c(e)$ is differentiable. □

In the same paper [**CIPP00**], Contreras employed a previous result of Frauenfelder and Schlenk [**Sch06, FS07**] to deal with the lower energy range $(e_0(L), c_{\mathrm{u}}(L))$. An independent proof of the same result, more in the spirit of the one of Theorem 1.16, was given later on by Taimanov [**Tai10a**].

Theorem 1.17 (Contreras). *Let $L : \mathrm{T}M \to \mathbb{R}$ be a Tonelli Lagrangian such that $\min E < e_0(L)$. For almost every $e \in (\min E, e_0(L))$, there is a contractible periodic orbit γ of energy e.*

Proof. We only give a brief outline of the proof. It is enough to work with the dual Tonelli Hamiltonian $H : \mathrm{T}^*M \to \mathbb{R}$. For each $e \in (\min E, e_0(L))$, there is an open subset

$U \subsetneq M$ such that the energy level $H^{-1}(e)$ does not intersect T^*U. Let $K : M \to \mathbb{R}$ be a smooth function all of whose critical points are in U. We treat this function as a Hamiltonian on T^*M independent of the momentum variable p. Its Hamiltonian flow is given by $\phi_K^t(q,p) = (q, p - t\,\mathrm{d}K(q))$. Since $\mathrm{d}K$ is nowhere vanishing on $M \setminus U$, for $t > 0$ large enough we have

$$\phi_K^t(H^{-1}(e)) \cap H^{-1}(e) = \varnothing.$$

This is usually expressed by saying that the energy hypersurface $H^{-1}(e)$ is Hamiltonianly displaceable. Under this condition, a result of Frauenfelder and Schlenk [**Sch06, FS07**] implies that a neighborhood $H^{-1}(e-\delta, e+\delta)$ has finite π_1-sensitive Hofer-Zehnder capacity. This, together with a version of Struwe's trick due to Hofer and Zehnder [**HZ94**], implies that $H^{-1}(e')$ has a periodic orbit for almost every $e' \in (e-\delta, e+\delta)$. □

⋆ **Open problem:** Is there a Riemannian metric on the domain $(0,\infty) \times \Lambda M$ of the free-period action functional that makes such domain complete while at the same time makes $\mathcal{S}_e$ satisfy the Palais-Smale condition for a given energy value $e \in (\min E, c_{\mathrm{u}}(L))$? A theorem of Contreras shows that such a metric exists for an energy level e provided the base projection $\pi : \mathrm{T}^*M \to M$ provides an injective homomorphism $\pi_* : H_1(H^{-1}(e); \mathbb{R}) \to H_1(M; \mathbb{R})$, and the energy level $H^{-1}(e) \subset \mathrm{T}^*M$ is a smooth hypersurface of contact type (i.e. $\ker(\omega|_{H^{-1}(e)})$ is a contact distribution, where ω is the standard symplectic form on T^*M).

1.7. Minimal boundaries

We now focus on Tonelli Hamiltonians and Lagrangians whose configuration space M is an orientable closed surface (with a slight abuse of terminology, we will briefly say "Tonelli Hamiltonians and Lagrangians on surfaces"). The dimension two is special: the fact that embedded loops separate M at least locally (that is, separate a tubular neighborhood of their support) allows to carry over arguments that are not available in general dimension. The results that we are going to present originate from the seminal work of Taimanov [**Tai91, Tai92a, Tai92b**], who showed that when $L : \mathrm{T}M \to \mathbb{R}$ is an electromagnetic Lagrangian as in Example 1.11, every sufficiently small energy level e contains a periodic orbit that is a local minimizer of the free-period action functional $\mathcal{S}_e$. Later, Contreras, Macarini, and Paternain put Taimanov result in the context of Aubry-Mather theory, and in particular pointed out that the energy values for which Taimanov's theorem hold are those in the interval $(0, c_0(L))$. The extension of Taimanov's theorem to arbitrary Tonelli Lagrangians on surfaces was finally proved by the Asselle and the author in [**AM19**], who showed that the existence of local minimizers holds for every energy value in $(e_0(L), c_0(L))$. In this section, we are going to briefly present an enhancement of the above results, that is due to Asselle, Benedetti and the authors [**ABM17**].

We first introduce the main character of these results. Let M be an orientable closed surface. By a multicurve, we mean a finite collection of periodic curves $\gamma_i = (\tau_i, q_i) \in (0,\infty) \times \Lambda M$, $i = 1, ..., n$, which we will write as $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_n)$. In order to avoid technicalities, we will always assume that the components γ_i are piecewise smooth with finitely many singular points. We say that $\boldsymbol{\gamma}$ is simple when all its components are simple (that is, they are piecewise smooth embeddings $\gamma_i : \mathbb{R}/\tau_i\mathbb{Z} \hookrightarrow M$) and pairwise disjoint. We

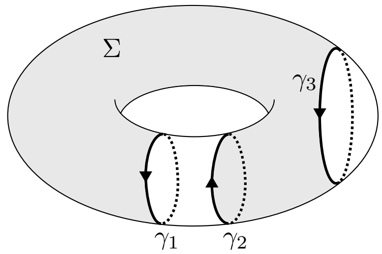


FIGURE 1.3. Example of topological boundary $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3) = \partial\Sigma$.

say that a simple multicurve $\boldsymbol{\gamma}$ is a topological boundary when it is the oriented boundary of a non-empty, possibly disconnected, open subset $\Sigma \subsetneq M$ (see Figure 1.3). We denote by $\mathcal{B}$ the space of topological boundaries on M; we stress that elements $\boldsymbol{\gamma} \in \mathcal{B}$ may have an arbitrary (positive) number of components. The free-period action functional $\mathcal{S}_e$ admits a natural extension to the space of topological boundaries, that is,

$$\mathcal{S}_e : \mathcal{B} \to \mathbb{R}, \qquad \mathcal{S}_e(\boldsymbol{\gamma}) = \mathcal{S}_e(\gamma_1) + ... + \mathcal{S}_e(\gamma_n).$$

We define a **minimal boundary** with energy e to be a topological boundary $\boldsymbol{\gamma} \in \mathcal{B}$ such that

$$\mathcal{S}_e(\boldsymbol{\gamma}) = \inf_{\mathcal{B}} \mathcal{S}_e.$$

Notice that

$$\inf_{\mathcal{B}} \mathcal{S}_e \leq 0. \tag{1.12}$$

Indeed, for every $\epsilon > 0$ there is a short simple contractible loop γ with action $\mathcal{S}_e(\gamma) < \epsilon$, and such a loop is in particular a topological boundary. Any component of a minimal boundary with energy e is a local minimizer of the free-period action functional $\mathcal{S}_e : (0, \infty) \times \Lambda M \to \mathbb{R}$. In particular, the components of minimal boundaries with energy e are simple periodic orbits with energy e.

The reason to study minimal boundaries is that, for energy values $e > e_0(L)$, even though the free-period action functional may be unbounded from below on the loop space, it is bounded from below over the space of topological boundaries. More precisely, we have the following estimate. We define a 1-form θ on M by

$$\theta_q(v) := \partial_v L(q, 0)v,$$

which is the "magnetic part" of the Tonelli Lagrangian L.

Lemma 1.18. *For each topological boundary $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_n) \in \mathcal{B}$, with components of the form $\gamma_i : \mathbb{R}/\tau_i\mathbb{Z} \to M$, and for each energy value $e > e_0(L)$, we have*

$$\mathcal{S}_e(\boldsymbol{\gamma}) \geq (e - e_0(L))(\tau_1 + ... + \tau_n) - \int_M |\mathrm{d}\theta|.$$

Proof. The function $K : \mathrm{T}M \to \mathbb{R}$, $K(q,v) = L(q,v) - \theta_q(v) - L(q,0)$ is a Tonelli Lagrangian that vanishes on the 0-section and is positive outside. We recall that $L(q,0) = -E(q,0) \geq -e_0(L)$. If $\boldsymbol{\gamma}$ as in the statement is the oriented boundary of the open subset $\Sigma \subset M$, by Stokes theorem we estimate

$$\begin{aligned}\mathcal{S}_e(\boldsymbol{\gamma}) &\geq \sum_{i=1}^{n} \left(\int_0^{\tau_i} K(\gamma_i, \dot\gamma_i)\,\mathrm{d}t + \int_{\gamma_i} \theta + (e - e_0(L))\tau_i \right) \\ &\geq \int_\Sigma \mathrm{d}\theta + (e - e_0(L))(\tau_1 + ... + \tau_n) \\ &\geq - \int_M |\mathrm{d}\theta| + (e - e_0(L))(\tau_1 + ... + \tau_n). \end{aligned}$$

□

We now address the question of the existence of minimal boundaries, starting with a negative statement.

Lemma 1.19. *All multicurves $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_n)$ with $[\boldsymbol{\gamma}] = [\gamma_1] + ... + [\gamma_n] = 0$ in $H_1(M;\mathbb{Z})$ have non-negative action $\mathcal{S}_{c_0(L)}(\boldsymbol{\gamma}) \geq 0$. In particular, there are no minimal boundaries with energy $e > c_0(L)$.*

Proof. The definition of $c_0(L)$ implies that $\mathcal{S}_{c_0(L)}(\gamma) \geq 0$ for any nullhomologous periodic curve γ. Assume now that $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_n)$ is a multicurve with $[\boldsymbol{\gamma}] = 0$ in $H_1(M;\mathbb{Z})$. We choose absolutely continuous paths $\zeta_i : [0,1] \to M$ such that $\zeta_i(0) = \gamma_i(0)$ and $\zeta_i(1) = \gamma_{i+1}(0)$. For each positive integer k, we define the loop

$$\xi_k := \gamma_1^k * \zeta_1 * \gamma_2^k * \zeta_2 * ... * \gamma_{n-1}^k * \zeta_{n-1} * \gamma_n^k * \overline{\zeta}_{n-1} * \overline{\zeta}_{n-2} * \dots * \overline{\zeta}_1,$$

where $\overline{\zeta}_i : [0,1] \to M$ denotes the reversed path $\overline{\zeta}_i(t) = \zeta_i(1-t)$ joining $\gamma_{i+1}(0)$ and $\gamma_i(0)$, $*$ denotes concatenation of paths, and the superscript k denotes the k-th iteration of a loop. The loop ξ_k is null-homologous, for

$$[\xi_k] = [\gamma_1^k] + \dots + [\gamma_n^k] + \underbrace{[\zeta_1 * \overline{\zeta}_1]}_{=0} + \dots + \underbrace{[\zeta_{n-1} * \overline{\zeta}_{n-1}]}_{=0} = k[\boldsymbol{\gamma}] = 0,$$

and therefore has non-negative action $\mathcal{S}_{c_0(L)}(\xi_k) \geq 0$. We set $\zeta := (\zeta_1 * \overline{\zeta}_1, ..., \zeta_n * \overline{\zeta}_n)$, so that $\mathcal{S}_{c_0(L)}(\xi_k) = k\,\mathcal{S}_{c_0(L)}(\boldsymbol{\gamma}) + \mathcal{S}_{c_0(L)}(\zeta)$. Therefore $\mathcal{S}_{c_0(L)}(\boldsymbol{\gamma}) \geq -k^{-1}\mathcal{S}_{c_0(L)}(\zeta)$, and by sending $k \to \infty$ we conclude that $\mathcal{S}_{c_0(L)}(\boldsymbol{\gamma}) \geq 0$. This, together with (1.12), implies the first statement of the lemma.

We recall that $\mathcal{S}_e$ is monotone increasing in e. If a minimal boundary $\boldsymbol{\gamma}$ with energy $e > c_0(L)$ existed, we would have $\mathcal{S}_{c_0(L)}(\boldsymbol{\gamma}) < \mathcal{S}_e(\boldsymbol{\gamma}) \leq 0$, contradicting the first statement of the lemma. □

On the energy range $(e_0(L), c_0(L)]$, we have positive existence results. We begin with the interval $(e_0(L), c_0(L))$.

Theorem 1.20. *For each $e \in (e_0(L), c_0(L))$, there exists a minimal boundary with energy e.*

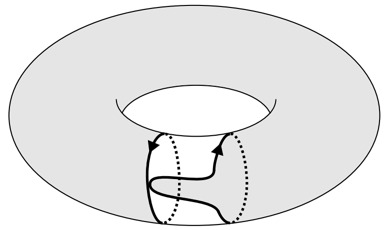


FIGURE 1.4. Topological boundary with a tangency.

Proof. We sketch the the proof in two steps.

• Step 1. *If $\mathcal{S}_e$ attains negative values on the space of topological boundaries $\mathcal{B}$, then there exists a minimal boundary of energy e.*

The proof of this step is unfortunately a technical tour de force (see [**AM19**, Section 2], which is based on Taimanov's [**Tai91**]), but nevertheless the general strategy is rather simple. Notice that we did not endow the space of topological boundaries $\mathcal{B}$ with a topology, but, whatever reasonable topology one considers, the space will not be compact: for instance, a sequence of topological boundaries could converge towards a boundary with a tangency (Figure 1.4), which is not an element of $\mathcal{B}$. By means of a finite dimensional reduction, it is possible to work in the subspace $\mathcal{B}' \subset \mathcal{B}$ of those topological boundaries $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_n)$ whose components γ_i are piecewise smooth solutions of the Euler-Lagrange equation with energy e and have finitely many singular points. For every $\boldsymbol{\gamma} \in \mathcal{B}$, there is a $\boldsymbol{\zeta} \in \mathcal{B}'$ such that $\mathcal{S}_e(\boldsymbol{\zeta}) \leq \mathcal{S}_e(\boldsymbol{\gamma})$; therefore we can look for minimizers of $\mathcal{S}_e$ inside $\mathcal{B}'$ instead of $\mathcal{B}$.

We consider a sequence $\boldsymbol{\gamma}_\alpha \in \mathcal{B}'$ such that $\mathcal{S}_e(\boldsymbol{\gamma}_\alpha) \to \inf \mathcal{S}_e|_{\mathcal{B}'} < 0$ as $\alpha \to \infty$. Without loss of generality, we can assume that every component of every $\boldsymbol{\gamma}_\alpha$ has length larger than some positive constant; indeed, since $e > e_0(L)$, components that are too short are contractible and have positive action, and by removing them we would obtain another element of $\mathcal{B}'$ with lower action (since $\mathcal{S}_e(\boldsymbol{\gamma}_\alpha) < 0$ for α large enough, after removal of the short components we must be left with a non-empty multi-curve). The space $\mathcal{B}'$ can be endowed with a natural topology, and a compactness theorem implies that, up to extracting a subsequence, $\boldsymbol{\gamma}_\alpha$ converges to some $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_n)$ in the closure of $\mathcal{B}'$. The multi-curve $\boldsymbol{\gamma}$ might not be a topological boundary anymore: a priori, $\boldsymbol{\gamma}$ is a topological boundary with tangencies (a source of complication being that the locus of these tangencies may not be a finite set). If a portion of some component $\gamma_i|_{[a,b]}$ does not have self-intersections nor intersections with other components of $\boldsymbol{\gamma}$ or with the remaining portion of γ_i, then $\gamma_i|_{[a,b]}$ is an embedded smooth solution of the Euler-Lagrange equation with energy e; indeed, if this were not the case, we could perturb $\gamma_i|_{[a,b]}$ while keeping its endpoint fixed and lower the action $\mathcal{S}_e(\boldsymbol{\gamma})$, contradicting the fact that the original $\boldsymbol{\gamma}$ was a minimizer.

We are now left to show that the components of $\boldsymbol{\gamma}$ do not have self-intersections nor mutual intersections. Let us show that $\boldsymbol{\gamma}$ cannot have a "simple" tangency, meaning a point $q_0 \in M$ that is an isolated intersection of exactly two components of $\boldsymbol{\gamma}$, as in Figure 1.5(a).

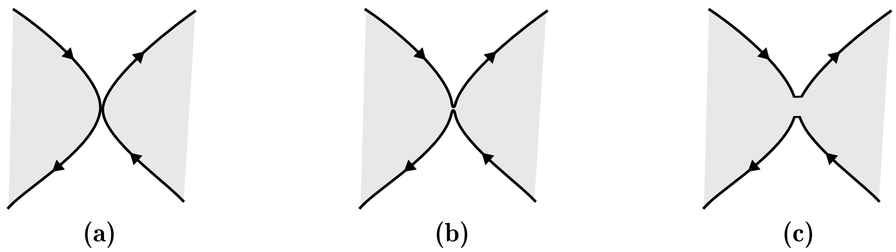


FIGURE 1.5

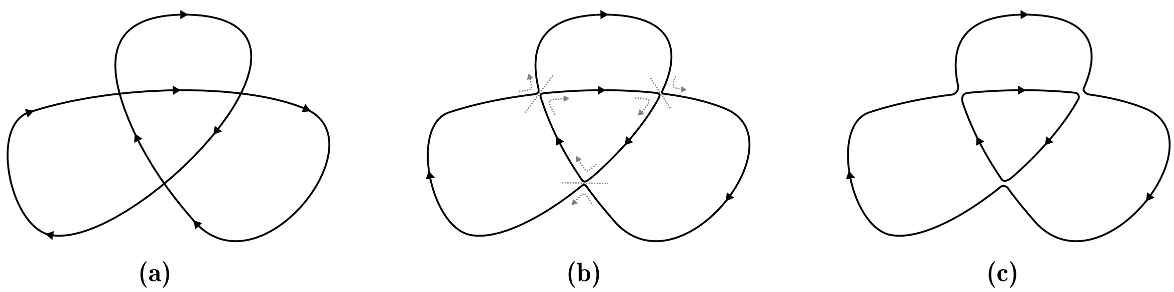


FIGURE 1.6

If this were the case, we could rearrange the components as in Figure 1.5(b), and finally get rid of the tangency point by chamfering the corners as in Figure 1.5(c). This way we would produce a multicurve with lower action, that is still in the closure of $\mathcal{B}$, contradicting the minimality of the original $\boldsymbol{\gamma}$. This gives the idea, but a more sophisticated argument is needed to take care of general tangencies.

• Step 2. *The functional $\mathcal{S}_e$ attains negative values on $\mathcal{B}$.*

Since $e < c_0(L)$, there exists a nullhomologous periodic curve ζ with negative action $\mathcal{S}_e(\zeta) < 0$. Up to a perturbation, we can assume that ζ has finitely many self-intersections, all of whose are transverse double points (Figure 1.6(a)). Out of such a curve ζ, we can produce an embedded multicurve $\boldsymbol{\gamma}$ as follows: at every self-intersection of ζ, we transform a double point into a tangency (Figure 1.6(b)), which we then remove by smoothing the corners (Figure 1.6(c)). Notice that $[\boldsymbol{\gamma}] = [\zeta] = 0$ in $H_1(M;\mathbb{Z})$, and we can do the above operation so that the action $\mathcal{S}_e(\boldsymbol{\gamma})$ is arbitratily close to $\mathcal{S}_e(\zeta)$, and in particular $\mathcal{S}_e(\boldsymbol{\gamma}) < 0$. Finally, since $[\boldsymbol{\gamma}] = 0$, a topological argument implies that $\boldsymbol{\gamma}$ is the disjoint union of topological boundaries $\boldsymbol{\gamma}_1, ..., \boldsymbol{\gamma}_n$. Since $\mathcal{S}_e(\boldsymbol{\gamma}_1) + ... + \mathcal{S}_e(\boldsymbol{\gamma}_n) = \mathcal{S}_e(\boldsymbol{\gamma}) < 0$, at least one of these topological boundaries $\boldsymbol{\gamma}_i$ must have negative action $\mathcal{S}_e(\boldsymbol{\gamma}_i) < 0$. □

Finally, we consider the critical energy level $c_0(L)$.

Theorem 1.21. *There exists a minimal boundary $\boldsymbol{\gamma}$ with energy $c_0(L)$.*

Proof. We consider a sequence of energy values $e_\alpha \in (c_0(L)-\epsilon, c_0(L))$ such that $e_\alpha \to c_0(L)$ as $\alpha \to \infty$, and a sequence of minimal boundaries $\boldsymbol{\gamma}_\alpha = (\gamma_{\alpha,1}, ..., \gamma_{\alpha,n_\alpha})$ of energy e_α. Every component of such boundaries has the form $\gamma_{\alpha,i} : \mathbb{R}/\tau_{\alpha,i}\mathbb{Z} \to M$. By Lemma 1.18, the total period of each minimal boundary $\boldsymbol{\gamma}_\alpha$ is uniformly bounded from above by

$$\tau_{\alpha,1} + ... + \tau_{\alpha,n_\alpha} \leq \frac{1}{e_\alpha - e_0(L)} \left(\mathcal{S}_{e_\alpha}(\boldsymbol{\gamma}_\alpha) + \int_M |\mathrm{d}\theta| \right) \leq \frac{1}{c_0(L) - e_0(L) - \epsilon} \int_M |\mathrm{d}\theta|.$$

Since every component $\gamma_{\alpha,i}$ is a $\tau_{\alpha,i}$-periodic orbit with energy e_α that is approximatively $c_0(L)$, we also have a uniform lower bound

$$\tau_{\alpha,i} \geq \delta > 0,$$

where δ is independent of α and i. Therefore, we have a uniform upper bound for the number of connected components n_α of $\boldsymbol{\gamma}_\alpha$, and up to extracting a subsequence we can assume that $n := n_\alpha$ is independent of α. Up to extracting another subsequence, any connected component $\gamma_{\alpha,i}$ converges in C^∞ to a periodic orbit γ_i with energy $c_0(L)$. A priori, the multicurve $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_n)$ may not be a topological boundary, but is certainly a topological boundary with tangencies. In particular, $[\boldsymbol{\gamma}] = 0$ in $H_1(M;\mathbb{Z})$, since the same was true for the $\boldsymbol{\gamma}_\alpha$'s. Since $0 \geq \mathcal{S}_{e_\alpha}(\boldsymbol{\gamma}_\alpha) \to \mathcal{S}_{c_0(e)}(\boldsymbol{\gamma})$ as $\alpha \to \infty$, Lemma 1.19 implies that $\mathcal{S}_{e_\alpha}(\boldsymbol{\gamma}) = 0$. Finally, an argument analogous to the one in Step 1 of the proof of Theorem 1.20 implies that $\boldsymbol{\gamma}$ is an embedded multicurve, and thus a minimal boundary of energy $c_0(L)$. □

Below the energy level $e_0(L)$, there are no minimal boundaries.

Proposition 1.22. *For all $e < e_0(L)$, we have*

$$\inf_{\mathcal{B}} \mathcal{S}_e = -\infty.$$

Proof. We consider an energy value $e < e_0(L) = \max E(\cdot, 0)$, and we fix $\epsilon \in (0, e_0(L) - e)$, so that the open subset $U := \{q \in M \mid E(q,0) - e > \epsilon\}$ is non-empty. Let $q : S^1 \to U$ be a contractible embedded loop. For each $\tau > 0$, the curve $\gamma_\tau(t) := q(t/\tau)$ is in particular a topological boundary. For τ large enough, we have

$$\mathcal{S}_e(\gamma_\tau) = \tau \int_0^1 \big(L(q, \dot q/\tau) + e\big)\,\mathrm{d}t = \tau \int_0^1 \big(\partial_v L(q, \dot q/\tau)\dot q/\tau + \underbrace{e - E(q, \dot q/\tau)}_{<-\epsilon/2}\big)\,\mathrm{d}t \xrightarrow{\tau\to\infty} -\infty.$$

□

We close this section with one last result relating minimal boundaries to Aubry-Mather theory [**Mat91**]. We denote by $\phi_L^t : \mathrm{T}M \to \mathrm{T}M$ the Euler Lagrange flow, whose orbits are lifts of solutions of the Euler-Lagrange equation, and we consider the associated invariant measures μ on $\mathrm{T}M$. We recall that the rotation vector of μ is the homology class $\rho(\mu) \in H_1(M;\mathbb{R})$ that evaluates against closed 1-forms σ as

$$\langle \sigma, \rho(\mu) \rangle = \int_{\mathrm{T}M} \sigma_q(v)\,\mathrm{d}\mu(q,v).$$

The Lagrangian action of an invariant measure μ is the quantity

$$\mathcal{S}(\mu) := \int_{\mathrm{T}M} L\,\mathrm{d}\mu$$

Aubry-Mather theory implies that

$$\inf_{\mu} \mathcal{S}(\mu) = -c_0(L),$$

where the infimum ranges over the space of invariant measures with zero rotation vector. We denote by $\mathfrak{M}_{\min}$ the space of invariant action-minimizing measures μ with zero rotation vector, i.e. $\rho(\mu) = 0$ and $\mathcal{S}(\mu) = -c_0(L)$. Such measures always exist for any Tonelli Lagrangian L. The union of their supports is a version of the so-called **Mather sets**

$$\mathcal{M} := \bigcup_{\mu\in\mathfrak{M}_{\min}} \operatorname{supp}(\mu),$$

which is invariant under the Euler-Lagrange flow, and is contained in the energy level $E^{-1}(c_0(L))$ according to a theorem of Carneiro [**Car95**]. The celebrated graph theorem of Mather [**Mat91**] implies that the base projection $\pi : \mathrm{T}M \to M$, $\pi(q, v) = q$ restricts to an injective map on the Mather set $\mathcal{M}$ (and indeed has a Lipschitz inverse).

Aubry-Mather theory deals with configuration spaces M of any dimension. When M is a closed orientable surface, minimal boundaries allow to define a Mather set on subcritical energies as well: for each $e \in (e_0(L), c_0(L)]$, we denote by $\mathcal{M}_e$ the set of points of the form $(\gamma(t), \dot\gamma(t))$, where γ is a component of a minimal boundary of energy e. Clearly, each $\mathcal{M}_e$ is an invariant set for the Euler-Lagrange flow, and is contained in the energy level $E^{-1}(e)$. Since a minimal boundary $\boldsymbol{\gamma}$ with energy $c_0(L)$ has action $\mathcal{S}_{c_0(L)}(\boldsymbol{\gamma}) = 0$, it defines a minimal measure on its support; this implies that $\mathcal{M}_{c_0(L)} \subset \mathcal{M}$. Actually, Asselle, Benedetti, and the author [**ABM17**] showed that $\mathcal{M}_{c_0(L)} = \mathcal{M}$, and the subcritical invariant sets $\mathcal{M}_e$ satisfy the graph property too.

Theorem 1.23. *For each $e \in (e_0(L), c_0(L)]$, the base projection $\pi : \mathrm{T}M \to M$ restricts to an injective map on the invariant set $\mathcal{M}_e$.* □

1.8. Waists and multiplicity of periodic orbits on energy levels

We already mentioned that, for general Tonelli Lagrangians $L : \mathrm{T}M \to \mathbb{R}$, nothing is known about the unconditional existence of periodic orbits in the energy range $(e_0(L), c_0(L))$ beyond Contreras' Theorem 1.16: on almost $e \in (e_0(L), c_{\mathrm{u}}(L))$ there exists a contractible periodic orbits α_e with energy e and positive action $\mathcal{S}_e(\alpha_e) > 0$. If we further assume M to be an orientable closed surface, for every $e \in (e_0(L), c_{\mathrm{u}}(L))$ (and not only for almost every) there exists another periodic orbit β_e of energy e that is different from α_e since it has negative action $\mathcal{S}_e(\beta_e) < 0$. Such a β_e is provided by Theorem 1.20 as the component of a minimal boundary with energy e.

The periodic orbits β_e are sometimes referred to as **waists**: they are local minimizer of the free-period action functional $\mathcal{S}_e : (0, \infty) \times \Lambda M \to \mathbb{R}$. The terminology is borrowed from Riemannian geometry, where a waist is a closed geodesic that is a local minimizer of

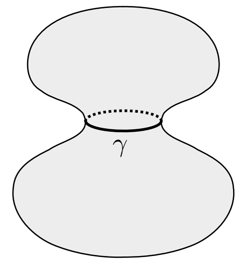


FIGURE 1.7. A waist γ in a Riemannian 2-sphere.

the length functional on the loop space (Figure 1.7). In the quest of multiplicity results for periodic orbits, waists turn out to play an important role: the celebrated waist theorem of Bangert [**Ban80**] implies for instance that the existence of a contractible waist in a closed, orientable, Riemannian surface "forces" the existence of infinitely many other closed geodesics.

Inspired by the work of Bangert, a breakthrough on the multiplicity problem for Tonelli periodic orbits on energy hypersurfaces was proved by the author together with Abbondandolo, Macarini, and Paternain [**AMMP17**], after a partial result in [**AMP15**]. The theorem was originally proved for the class of magnetic Lagrangians of Example 1.11(iii), but was finally extended to the whole class of Tonelli Lagrangians by Asselle and the author [**AM19**]. The final result is the following.

Theorem 1.24. *Let M be a closed surface, and $L : \mathrm{T}M \to \mathbb{R}$ a Tonelli Lagrangian. For almost all $e \in (e_0(L), c_{\mathrm{u}}(L))$ there exist infinitely many periodic orbits with energy e.*

Proof. The proof of this theorem will require several ingredients, some of which are borrowed from the literature while others are novel. We remark that, without loss of generality, we can assume that the surface M is orientable: if this were not the case, we would work on its orientation double cover M' and with the lifted Tonelli Lagrangian $L' : \mathrm{T}M' \to \mathbb{R}$, which has the same relevant energy values as the original Lagrangian, i.e. $e_0(L') = e_0(L')$ and $c_{\mathrm{u}}(L') = c_{\mathrm{u}}(L)$.

Let us consider the waists β_e with negative action $\mathcal{S}_e(\beta_e) < 0$ that we already introduced at the beginning of this section. We recall that each β_e is in particular a critical point of the free-period action functional $\mathcal{S}_e$, and belongs to a critical circle $S^1 \cdot \beta_e$ (see Section 1.5). We fix an energy value $e_1 \in (e_0(L), c_{\mathrm{u}}(L))$ such that

$$\beta := \beta_{e_1}$$

is a strict local minimizer of $\mathcal{S}_{e_1}$, meaning that there exists an open neighborhood $\mathcal{U} \subset (0,\infty) \times \Lambda M$ of the critical circle $S^1 \cdot \beta$ such that $\mathcal{S}_{e_1}(\beta) < \mathcal{S}_{e_1}(\gamma)$ for all $\gamma \in \mathcal{U} \setminus (S^1 \cdot \beta)$. Notice that on those energy levels e for which $S^1 \cdot \beta_e$ is not a strict local minimizer there is nothing to prove: any neighborhood of $S^1 \cdot \beta_e$ contains infinitely many other critical circles of $\mathcal{S}_e$, which give infinitely many periodic orbits at level e. Therefore, in order to prove the

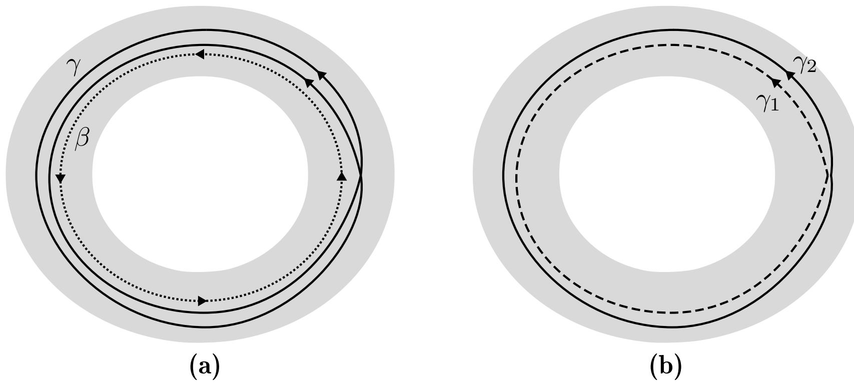


FIGURE 1.8. **(a)** A periodic curve γ close to the second iterate of β. **(b)** As a cycle, γ decomposes as the sum of γ_1 and γ_2, which are both periodic curves close to γ.

theorem, it is enough to show that there are infinitely many periodic orbits with energy e, for almost every e in a neighborhood of e_1.

Let us stress a point that was already touched upon in Section 1.4. A τ-periodic orbit $\gamma = (\tau, q) \in \mathrm{crit}(\mathcal{S}_e)$ is clearly also $m\tau$-periodic for any positive integer m; however, when seen as an $m\tau$-periodic orbit, it corresponds to a different point of the domain of the free period action functional: the point $\gamma^m := (m\tau, q^m) \in (0, \infty) \times \Lambda M$, where $q^m(t) = q(mt)$ is the m-th iterate of the loop q. Therefore, in order to establish multiplicity results for periodic orbits, it is not enough to detect several critical circles of $\mathcal{S}_e$; one further needs to identify critical circles that correspond to iterates of a same "primitive" periodic orbit.

A classical argument due to Hedlund [**Hed32**], reproved in the setting of the free-period action functional by Abbondandolo, Macarini, and Paternain [**AMP15**], implies that the iterates β^m, for all $m \geq 1$, are all strict local minimizers of $\mathcal{S}_{e_1}$. This is due to the fact that we work on an orientable surface M, and therefore a tubular neighborhood of the support of β in M is diffeomorphic to an annulus A. For instance, if a periodic curve γ is close to the second iterate β^2, in particular the support of γ is contained in the annulus A (Figure 1.8(a)); as a cycle, γ decomposes as the sum of two cycles γ_1 and γ_2 (Figure 1.8(b)); each one of these two cycles is a periodic curve in the domain of the free-period action functional, and close to β; since β is a local minimizer of $\mathcal{S}_{e_1}$, we have

$$\mathcal{S}_{e_1}(\gamma) = \mathcal{S}_{e_1}(\gamma_1) + \mathcal{S}_{e_1}(\gamma_2) \geq \mathcal{S}_{e_1}(\beta) + \mathcal{S}_{e_1}(\beta) = \mathcal{S}_{e_1}(\beta^2),$$

and the inequality is strict unless $\gamma \in S^1 \cdot \beta^2$.

We now setup a min-max scheme that will produce the candidate periodic orbits claimed to exist. For each integer $m \geq 1$, since β^m is a strict local minimizer of $\mathcal{S}_{e_1}$, there exists a neighborhood $\mathcal{U}_m$ of the critical circle $S^1 \cdot \beta^m$ such that

$$a_m := \mathcal{S}_{e_1}(\beta^m) = \inf_{\mathcal{U}_m} \mathcal{S}_{e_1} < \inf_{\partial \mathcal{U}_m} \mathcal{S}_{e_1} =: a_m + 2\delta_m.$$

We choose $\mathcal{U}_m$ to be a small neighborhood, so that in particular every $\gamma \in \mathcal{U}_m$ has period close to the period of β^m. We recall that the free-period action functional $\mathcal{S}_e$ depends on the energy parameter e monotonically, and more precisely

$$\mathcal{S}_e(\tau,q)=\mathcal{S}_{e_1}(\tau,q)+(e-e_1)\tau.$$

Therefore, we can find $b_m\leq a_m$ and a sufficiently small neighborhood $I\subset(e_0(L),c_{\mathrm{u}}(L))$ of the energy value e_1 so that

$$b_m<\inf_{\mathcal{U}_m}\mathcal{S}_e<a_m+\delta_m<\inf_{\partial\mathcal{U}_m}\mathcal{S}_e,\qquad\forall e\in I.$$

For each $e\in I$ and $m\geq 1$, we consider the min-max

$$c(e,m):=\inf_u\max_{s\in[0,1]}\mathcal{S}_e(u(s))\geq a_m+\delta_m,$$

where the infimum ranges over the family of continuous paths $u:[0,1]\to(0,\infty)\times\Lambda M$ such that $u(0)\in\mathcal{U}_m$, $\mathcal{S}_e(u(0))<a_m+\delta$, and $\mathcal{S}_e(u(1))<b_m$. Notice that this family of maps is non empty, since the free-period action functional $\mathcal{S}_e$ is unbounded from below on every connected component of its domain. The inequality $c(e,m)\geq a_m+\delta_m$ is due to the fact that every path u in the definition of the min-max must exit the open set $\mathcal{U}_m$.

The fact that $e\mapsto\mathcal{S}_e$ is monotone increasing readily implies the same property for the function $e\mapsto c(e,m)$. In particular, this latter function is differentiable on a full measure subset $J\subset I$ according to Lebesgue's theorem. At every point $e\in J$, the argument of Struwe [**Str90**] that we already employed in the proof of Theorem 1.16 insures that $c(e,m)$ is a critical value of $\mathcal{S}_e$.

We fix, once for all, an energy value $e\in J$. We claim that the sequence of critical values $c_m:=c(e,m)$ correspond to infinitely many periodic orbits with energy e that are geometrically distinct. As a first step, an argument similar to the one sketched in Figure 1.1 implies that $c_m\to-\infty$ as $m\to\infty$; this guarantees that the sequence c_m corresponds to infinitely many distinct critical circles of $\mathcal{S}_e$. Since each critical value c_m is obtained as a min-max over a family of 1-dimensional objects (the paths u), it must be a critical value of mountain pass type: at least one critical circle $S^1\cdot\zeta_m\subset\mathrm{crit}(\mathcal{S}_e)\cap\mathcal{S}_e^{-1}(c_m)$ must join together two different connected components $\mathcal{C}_1,\mathcal{C}_2\subset\mathcal{S}_e^{-1}(-\infty,c_m)$; namely, for any neighborhood $\mathcal{W}$ of $S^1\cdot\zeta_m$, the union $\mathcal{C}_1\cup\mathcal{C}_2$ is contained in a connected component of $\mathcal{S}_e^{-1}(-\infty,c_m)\cup\mathcal{W}$.

The final step, which actually required most of the efforts in the original proof in [**AMMP17**], is the following general principle: if a periodic orbit γ gives critical circles $S^1\cdot\gamma^m$ that are isolated components of $\mathrm{crit}(\mathcal{S}_e)$ for all $m\geq 1$, such critical circles can be of mountain pass type only for finitely many values of m. We do not justify this claim here, but just mention that it requires a study of the Morse indices of the free-period action functional [**AM19**, Section 3], a subtle study of the local properties of the critical circles [**AMMP17**, Section 2], and arguments à la Bangert as in Theorem 1.9. In view of this result, if there were only finitely many periodic orbits with energy e, there would exist a negative value $b<0$ such that none of the critical circles $S^1\cdot\zeta$ of $\mathcal{S}_e$ with critical value less than b are of mountain pass type. But for large enough m, the mountain pass critical value c_m is less than b, which gives a contradiction. □

We already mentioned in the previous section that for energy values $e < e_0(L)$ there are no minimal boundaries. The following example, which the author provided in the joint paper [**AM19**] with Asselle, shows that in general there may be only finitely many periodic orbits with energy e, and no waists with energy e at all.

Example 1.25. We fix real numbers $0 < a_1 < a_2 < 1$ with irrational quotient a_1/a_2, and a smooth monotone increasing function $\chi : [0,\infty) \to [0,1]$ such that $\chi(x) = x$ for all $x \in [0, a_2]$ and $\chi(x) = 1$ for all $x \geq 1$. On the configuration space $\mathbb{T}^2 := [-1,1]^2/\{-1,1\}^2$, we define the Tonelli Hamiltonian

$$H : \mathrm{T}^*\mathbb{T}^2 \to \mathbb{R}, \qquad H(q_1, q_2, p_1, p_2) = \frac{1}{2}\bigg(\frac{\chi(|q_1|^2) + |p_1|^2}{a_1} + \frac{\chi(|q_2|^2) + |p_2|^2}{a_2}\bigg).$$

For such a Hamiltonian, we have

$$0 = \min H < e_0(H) = \frac{1}{2}\bigg(\frac{1}{a_1} + \frac{1}{a_2}\bigg).$$

Every energy level $H^{-1}(e)$ with $e \in (0, 1/a_1)$ is a so-called irrational ellipsoid: in complex notation $z_j := q_j + ip_j$, the Hamiltonian flow is given by

$$\phi_H^t(z_1, z_2) = (\exp(-it/a_1)z_1, \exp(-it/a_2)z_2),$$

and has only two periodic orbits:

$$\gamma_1(t) = (\exp(-it/a_1)2e/a_1, 0), \qquad \gamma_2(t) = (0, \exp(-it/a_2)2e/a_2).$$

One can show that both γ_1 and γ_2 have positive Morse index as critical points of $\mathcal{S}_e$, and in particular they are not local minimizers. □

When M is a closed surface and the energy range $(c_{\mathrm{u}}(L), c_0(L))$ is non-empty, for every $e > c_{\mathrm{u}}(L)$ there are infinitely many periodic orbits of energy e. Indeed, the fundamental group $\pi_1(M)$ must be non-abelian, and thus M must have genus at least 2. The domain of the free-period action functional $(0,\infty) \times \Lambda M$ has infinitely many connected components that do not contain iterated periodic curves, and Theorem 1.15 provides infinitely many periodic orbits with energy e, each one being a global minimizer of $\mathcal{S}_e$ in its connected component.

Finally, when M is simply connected, we have $c_{\mathrm{u}}(L) = c_0(L) = c(L)$, and for all $e > c(L)$ the dual Hamiltonian flow $\phi_H^t : H^{-1}(e) \to H^{-1}(e)$ is orbitally equivalent to the geodesic flow of a Finsler metric F_e (Theorem 1.14). The existence of periodic orbits with energy e reduces to the problem of the existence of closed geodesics for the Finsler metric F_e. As we already mentioned in Section 1.14, there is an example of Finsler metric, for instance on a n-sphere, with only finitely many closed geodesics. Nevertheless, if $e_0(L) < c(L)$, this cannot happen for the Finsler metrics F_e with e just above the critical value $c(L)$. This will be a consequence of the following existence of waists, which was proved by Asselle and the author [**AM20**].

Theorem 1.26. *Let M be a simply connected closed manifold, and $L : \mathrm{T}M \to \mathbb{R}$ a Tonelli Lagrangian such that $e_0(L) < c(L)$. There exists $c_{\mathrm{w}}(L) > c(L)$ and, for every $e \in (c(L), c_{\mathrm{w}}(L))$, a periodic orbit with energy e that is a local minimizer of the free-period action functional $\mathcal{S}_e$.*

Proof. The statement is straightforward when there exists a periodic orbit γ of energy $c(L)$ that is a local minimizer of the free-period action functional $\mathcal{S}_{c(L)}$ (we know that this is always the case if M is an orientable closed surface, thanks to the minimal boundaries provided by Theorem 1.21). Indeed, let $\mathcal{U}$ be a small neighborhood of the critical circle $S^1 \cdot \gamma$ such that $\mathcal{S}_{c(L)}(\gamma) = \inf \mathcal{S}_{c(L)}|_{\mathcal{U}} < \inf \mathcal{S}_{c(L)}|_{\partial\mathcal{U}}$. For values $e > c(L)$ sufficiently closed to $c(L)$, the functional $\mathcal{S}_e$ will still satisfy $\inf \mathcal{S}_e|_{\mathcal{U}} < \inf \mathcal{S}_e|_{\partial\mathcal{U}}$, and therefore it will have a local minimizer within $\mathcal{U}$. If $e_0(L) < c(L)$ but there are no periodic orbits with energy $c(L)$, the proof requires an argument from Aubry-Mather theory. □

Corollary 1.27. *Let $L : \mathrm{T}S^2 \to \mathbb{R}$ be a Tonelli Lagrangian such that $e_0(L) < c(L)$. There exists $c_{\mathrm{w}}(L) > c(L)$ and, for every $e \in (c(L), c_{\mathrm{w}}(L))$, infinitely many periodic orbits with energy e.*

Proof. Theorem 1.26 provides a local minimizer γ_e of $\mathcal{S}_e$ for all $e \in (c(L), c_{\mathrm{w}}(L))$. Since $e > c(L)$, γ_e is not a global minimizer of $\mathcal{S}_e$. Therefore, we can proceed as in the proof of Theorem 1.26 using γ_e instead of β, and detect infinitely many periodic orbits with energy e that are mountain pass critical points of $\mathcal{S}_e$. Notice that the conclusion here is valid for any energy level $e \in (c(L), c_{\mathrm{w}}(L))$, and not only for almost any, since $\mathcal{S}_e$ satisfies the Palais-Smale condition. □

⋆ **Open problem:** For Tonelli Lagrangians $L : \mathrm{T}M \to \mathbb{R}$ on a closed surface M, does the assertion of Theorem 1.24, that is, the existence of infinitely many periodic orbits with energy e, hold for all $e \in (e_0(L), c_{\mathrm{u}}(L))$?

1.9. Billiards

The last topic of this chapter concerns Tonelli Hamiltonian dynamics with obstacles, and in order to keep the presentation simple we will focus on a specific setting, which is nevertheless an important one in the literature: the one of smooth billiards in Euclidean spaces. Any such system is uniquely defined by a connected, bounded, open subset $\Omega \subset \mathbb{R}^n$ with non-empty smooth boundary $\partial\Omega$, which is the billiard table. We denote by ν the unit normal vector field along $\partial\Omega$ pointing outside Ω. The billiard dynamics can be described as follows: a billiard trajectory is a piecewise smooth curve $\gamma : \mathbb{R} \to \overline{\Omega}$ that moves in the interior of Ω along a straight line, say parametrized with constant speed 1, until it hits the boundary at some instant t, in which it "bounces" according to the reflection law

$$\begin{aligned} &\langle \dot\gamma(t^+) + \dot\gamma(t^-), \nu_{\gamma(t)} \rangle = 0, \\ &\dot\gamma(t^+) - \langle \dot\gamma(t^+), \nu_{\gamma(t)} \rangle \nu_{\gamma(t)} = \dot\gamma(t^-) - \langle \dot\gamma(t^-), \nu_{\gamma(t)} \rangle \nu_{\gamma(t)}. \end{aligned} \tag{1.13}$$

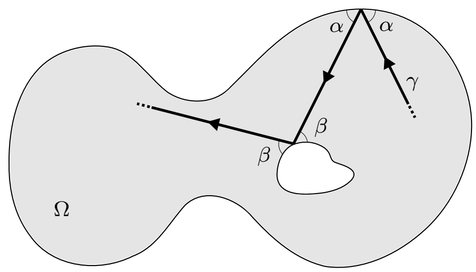


FIGURE 1.9. A billiard trajectory on a 2-dimensional billiard table Ω.

It may happen that a straight line γ coming from within Ω hits the boundary tangentially at time t, i.e. $\dot\gamma(t) \in \mathrm{T}_{\gamma(t)}\partial\Omega$, and $\gamma(t+\epsilon)$ exits $\overline{\Omega}$ for arbitrarily small values of $\epsilon > 0$; we will simply ignore such exiting glancing orbits. If Ω is two-dimensional, the reflection law (1.13) can be expressed as the equality between the angle of incidence and the angle of reflection at the boundary $\partial\Omega$ (Figure 1.9).

We are interested in **periodic billiard trajectories**, that is, those (non-exiting glancing) billiard trajectories γ such that $\gamma(\tau + \cdot) = \gamma$ for some minimal $\tau > 0$. The number of bounces of γ is the cardinality of the set $\{t \in [0,\tau) \mid \gamma(t) \in \partial\Omega\}$. Historically, the existence of periodic billiard trajectories was first studied by Birkhoff [**Bir66**] in dimension 2 under the assumption that the billiard table $\Omega \subset \mathbb{R}^2$ is convex. An application of the celebrated Poincaré-Birkhoff fixed point theorem provides an abundance of periodic billiard trajectories: in particular, for every integer $b \geq 2$ there are at least two periodic billiard trajectories whose number of bounce points divides b. Birkhoff's proof reduces the study of periodic billiard trajectories to the study of periodic points of an area-preserving twist map of the annulus $\mathrm{T}^*\partial\Omega$. For higher dimensional convex billiard tables, the study of periodic billiard trajectories can be carried over by means of a variational principle on the space of polygons inscribed in $\partial\Omega$; results due to Babenko [**Bab90**], Farber and Tabachnikov [**FT02**], and the author [**Maz11b**] still provide plenty of periodic billiard trajectories.

In this section, we consider instead billiard tables $\Omega \subset \mathbb{R}^n$ that are not necessarily convex, and indeed may not even be homeomorphic to a ball (such as the one in Figure 1.9). In this general setting, the mere existence of one periodic billiard trajectory is already a hard statement, which was established by Benci and Giannoni in a seminal paper [**BG89**]. The result actually shows that there exists a periodic billiard trajectory with at most $n+1$ bounces. The beautiful idea in their proof consists in approximating the billiard dynamics with ordinary Tonelli Hamiltonian systems; the involved Tonelli Hamiltonians H_ϵ have a potential term that vanishes on the interior of Ω except in the ϵ-neighborhood of the boundary $\partial\Omega$, and grow to infinity towards the boundary $\partial\Omega$. For every such approximating Hamiltonian system, they found a suitable periodic orbit γ_ϵ by means of variational methods, and with a subtle compactness argument they showed that a subsequence of such orbits tends towards a periodic billiard trajectory as $\epsilon \to 0$.

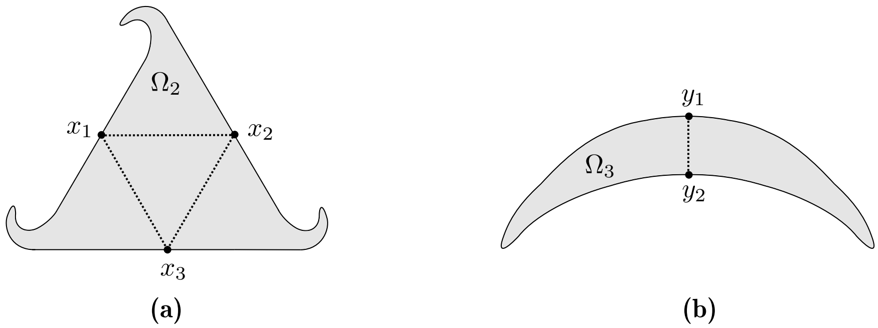


FIGURE 1.10. **(a)** A billiard table $\Omega_2 \subset \mathbb{R}^2$ without any periodic billiard trajectory with 2 bounces, and with a periodic billiard trajectory with 3 bounces at the points $x_1, x_2, x_3 \in \partial\Omega_2$. **(b)** A billiard table $\Omega_3 \subset \mathbb{R}^2$ with a periodic billiard trajectory with 2 bounces at the points $y_1, y_2 \in \partial\Omega_3$, but no periodic billiard trajectories with 3 bounces.

Benci and Giannoni's argument employs the fixed-period action functional of Section 1.2 of the approximating Tonelli Hamiltonian systems. Here we present an improvement of their existence theorem, which is due to Albers and the author [**AM11**], and employs the free-period action functional of Section 1.5 for the approximating Tonelli Hamiltonian systems. Given a subset $U \subset \mathbb{R}^n$, we define its affine displacement by

$$\mathrm{displ}(U) := \inf \big\{ \|v\| \bigm| v \in \mathbb{R}^n \text{ such that } \overline{U} \cap (v + \overline{U}) = \varnothing \big\}.$$

Notice that $\mathrm{displ}(U)$ is bounded from above by the diameter $\mathrm{diam}(\overline{U})$. Given a unit-speed periodic billiard trajectory γ, by length of γ we mean its minimal period τ, or equivalently the length of the closed curve $\gamma|_{[0,\tau]}$.

Theorem 1.28. *There is a universal constant $c > 0$ such that, for every $n \geq 2$, every bounded, open subset $\Omega \subset \mathbb{R}^n$ with non-empty smooth boundary contains a periodic billiard trajectory of length at most $c\,\mathrm{displ}(\Omega)$ and at most $n+1$ bounces.*

As pointed out by Tabachnikov in his monograph [**Tab05**], there exist a billiard table $\Omega_2 \subset \mathbb{R}^2$ without any periodic billiard trajectories with exactly 2 bounces (Figure 1.10(a)), and a billiard table $\Omega_3 \subset \mathbb{R}^2$ without any periodic billiard trajectories with exactly 3 bounces (Figure 1.10(b)). Therefore, the bound of the number of bounces in Theorem 1.28 is sharp, at least for planar billiards.

A result of Irie [**Iri12**], that improves a previous one of Viterbo [**Vit00**], implies that the length of the shortest billiard trajectory in a billiard table $\Omega \subset \mathbb{R}^n$ is bounded from above by $c_n r(\Omega)$, where $c_n > 0$ is a constant depending only on n, and $r(\Omega)$ is the largest radius of a round ball contained in $\overline{\Omega}$. This bound implies the one in Theorem 1.28 for a fixed dimension n, but not in general as the constant c in Theorem 1.28 is independent of n.

The proof of Theorem 1.28 will require the following result from Reeb dynamics originally due to Schlenk [**Sch06**], which we state together with a slight improvement due to Cieliebak, Frauenfelder, and Paternain [**CFP10**]; a partial version of this result was already employed in Theorem 1.17. We recall that a closed hypersurface $\Sigma \subset \mathbb{R}^{2n}$ is of contact type when it admits a contact form α, which is a 1-form on Σ such that $\alpha \wedge (\mathrm{d}\alpha)^{n-1}$ is a volume form. On (Σ, α) there is a preferred vector field R, called the Reeb vector field, uniquely defined by the equations $\alpha(R) \equiv 1$ and $\mathrm{d}\alpha(R, \cdot) \equiv 0$. A closed Reeb orbit is, by definition, a periodic orbit of the flow of R. When the contact form α is the restriction of a 1-form λ on $\mathbb{R}^{2n}$ whose exterior derivative is the standard symplectic form $\mathrm{d}\lambda = \omega$, the hypersurface Σ is said to be of restricted contact type. In this case, any closed Reeb orbit has an associated Maslov index $\mathrm{mas}(\gamma) \in \mathbb{Z}$; we shall not define it here, but simply mention that it coincides with the Morse index of the fixed-period action functional of Section 1.2 when the Reeb flow is the restriction of a Tonelli Hamiltonian flow. The displacement energy of Σ is the quantity

$$e(\Sigma) := \inf_{H_t} \int_0^1 \Big(\max H_t - \min H_t \Big)\, \mathrm{d}t,$$

where the infimum ranges over all the compactly supported time-dependent Hamiltonians $H_t : \mathbb{R}^{2n} \to \mathbb{R}$ whose time-1 map of the Hamiltonian flow displaces Σ from itself, i.e. $\phi_H^1(\Sigma) \cap \Sigma = \varnothing$. If no Hamiltonian realizes such a displacement, we set $e(\Sigma) = \infty$.

Theorem 1.29 (Schlenk, Cieliebak-Frauenfelder-Paternain). *Any restricted contact-type closed hypersurface (Σ, α) of $\mathbb{R}^{2n}$ with finite displacement energy has a closed Reeb orbit γ of period $\tau \leq e(\Sigma)$ and Maslov index* $\mathrm{mas}(\gamma) \in \{n, n+1\}$. □

Proof of Theorem 1.28. We consider the distance function from the boundary $\partial\Omega$

$$b : \Omega \to [0, \infty), \qquad b(q) = \min_{q_0 \in \partial\Omega} \|q - q_0\|.$$

We denote by $U_\rho := b^{-1}(0, \rho)$ the sublevel sets of b, and remark that b is smooth on $U_{2\rho}$ for some $\rho > 0$ small enough. By composing b with a suitable smooth step function, we obtain a smooth function $h : \Omega \to (0, \infty)$ such that $\|\nabla h\|_{L^\infty} < 1$, $h|_{U_{\rho/2}} \equiv b|_{U_{\rho/2}}$, $h|_{\Omega \setminus U_{\rho/2}} > \rho/2$, and $h|_{\Omega \setminus U_\rho} \equiv \mathrm{const}$. For any $\epsilon > 0$, we introduce the Tonelli Lagrangian

$$L_\epsilon : \mathrm{T}\Omega \to \mathbb{R}, \qquad L(q, v) = \tfrac{1}{2}\|v\|^2 - \epsilon h(q)^{-2}.$$

As ϵ gets smaller, the orbits of L_ϵ with energy 1 "resemble" more and more billiard trajectories in Ω (Figure 1.11). This observation is formulated rigorously with the following compactness statement, which is a minor modification of the original argument by Benci and Giannoni. We consider the free-period action functional with energy 1/2 of the Lagrangian L_ϵ, which for simplicity we denote by $\mathcal{S}^\epsilon$, i.e.

$$\mathcal{S}^\epsilon : (0, \infty) \times \Lambda\Omega \to \mathbb{R}, \qquad \mathcal{S}^\epsilon(\tau, q) = \tau \int_0^\tau L_\epsilon(q, \dot q/\tau)\, \mathrm{d}t + \tau/2.$$

We recall that a pair $(\tau, q) \in (0, \infty) \times \Lambda\Omega$ is identifies with a τ-periodic curve $\gamma(t) = q(t/\tau)$, and we express this identification simply by writing $\gamma = (\tau, q)$. Assume that there exist a constant $T_2 > 0$ and, for some sequence $\epsilon_k \to 0^+$, critical points

$$\gamma_k = (\tau_k, q_k) \in \mathrm{crit}(\mathcal{S}^{\epsilon_k})$$

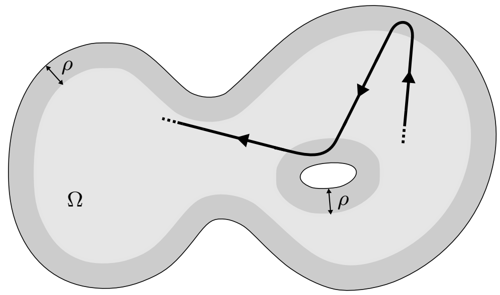


FIGURE 1.11. An orbit of the Lagrangian system L_ϵ, approximating a billiard trajectory.

such that $\tau_k \leq T_2$. Let us further assume that every τ_k is also uniformly bounded from below by a positive constant, i.e. $\tau_k \geq T_1 > 0$. Suitable estimates imply that, up to extracting a subsequence:

- $\epsilon h(\gamma_k)^{-2} \to 0$ almost everywhere,
- $\gamma_k \to \gamma = (\tau, q) \in [T_1, T_2] \times \Lambda\overline{\Omega}$ in the $W^{1,2} \times \mathbb{R}$ topology,
- there exists a finite Borel measure μ on the compact set

$$B := \{t \in \mathbb{R}/\tau\mathbb{Z} \mid \gamma(t) \in \partial\Omega\}$$

such that

$$\int_0^\tau \langle \dot\gamma, \dot\psi \rangle \, \mathrm{d}t = \int_B \langle \nu_\gamma, \psi \rangle \, \mathrm{d}\mu, \qquad \forall \psi \in W^{1,2}(\mathbb{R}/\tau\mathbb{Z}, \mathbb{R}^n). \tag{1.14}$$

By choosing smooth test functions ψ supported away from $\mathrm{supp}(\mu)$, we readily obtain that the restriction of γ to the complement of $\mathrm{supp}(\mu)$ is smooth and has vanishing second derivative. Moreover, since $\frac{1}{2}\|\dot\gamma_k\|^2 - \epsilon h(\gamma_k)^{-2} \equiv \frac{1}{2}$, we have $\|\dot\gamma(t)\| = 1$ for all $t \not\in \mathrm{supp}(\mu)$. Therefore, γ is a unit-speed straight curve on all intervals outside $\mathrm{supp}(\mu)$. Assume now that t is an isolated point of $\mathrm{supp}(\mu)$, so that $[t-\epsilon, t+\epsilon] \cap \mathrm{supp}(\mu) = \{t\}$ for $\epsilon > 0$ small enough. By choosing test functions ψ supported in $[t-\epsilon, t+\epsilon]$ with $v := \psi(t)$, an integration by parts in (1.14) implies

$$\langle \dot\gamma(t^-) - \dot\gamma(t^+), v \rangle = \langle \nu_{\gamma(t)}, v \rangle \mu(\{t\}), \qquad \forall v \in \mathbb{R}^n,$$

which, together with the fact that γ is a unit-speed straight line outside $\mathrm{supp}(\mu)$, implies that the reflection law (1.13) is satisfied at t.

It turns out that a uniform lower bound for the periods $\tau_k \geq T_1 > 0$ is always satisfied for sequences as above. Otherwise, if a subsequence satisfied $\tau_k \to 0$, we could apply the argument of the previous paragraph to suitable iterates of the periodic curves γ_k, but the limit curve γ would need to be both stationary and with unit-speed, a contradiction.

In order for γ to be a periodic billiard trajectory, the set of bouncing instants $\mathrm{supp}(\mu)$ must be finite. By a beautiful argument due to Benci and Giannoni, it turns out that the cardinality of $\mathrm{supp}(\mu)$ can be controlled by means of the Morse indices of the fixed-period action functionals $\mathcal{S}^{\epsilon_k}(\tau_k, \cdot)$ at q_k, which we denote by $\mathrm{ind}(q_k)$. Indeed, consider a point

$t_0 \in \operatorname{supp}(\mu)$; we recall that $q(s) = \gamma(\tau s)$, and we set $s_0 := t_0/\tau \in S^1 = \mathbb{R}/\mathbb{Z}$. Let $\chi : S^1 \to [0,1]$ be a smooth bump function supported in $[s_0 - \delta, s_0 + \delta]$ and identically equal to 1 on $[s_0 - \delta/2, s_0 + \delta/2]$. We set

$$w_k(s) := \chi(s)\nabla h(q_k(s)).$$

If the constant $\delta > 0$ was chosen small enough, then a computation shows that for all k large enough we have

$$\tfrac{\mathrm{d}^2}{\mathrm{d}r^2}\Big|_{r=0}\mathcal{S}^{\epsilon_k}(\tau_k, q_k + r w_k) < 0.$$

Notice that we can build such a w_k around any point $t_0 \in \operatorname{supp}(\mu)$. This provides the desired cardinality bound

$$\#\operatorname{supp}(\mu) \leq \liminf_{k\to\infty} \operatorname{ind}(q_k).$$

In order to prove the theorem, we are left to find the sequence of approximating periodic orbits γ_k. For this, we consider the dual Tonelli Hamiltonians

$$H_\epsilon : \mathrm{T}^*\Omega \to \mathbb{R}, \qquad H_\epsilon(q,p) = \frac{1}{2}\|p\|^2 + \frac{\epsilon}{h(q)^2}.$$

For every $\epsilon > 0$ small enough, the energy level $\Sigma_\epsilon := H_\epsilon^{-1}(\frac{1}{2})$ is a smooth hypersurface in the cotangent bundle. Since H_ϵ is a classical mechanical Hamiltonian of the form kinetic plus potential energy, every Σ_ϵ is of restricted contact type. Indeed, consider the smooth function

$$u : \mathrm{T}^*\Omega \to \mathbb{R}, \qquad u(q,p) = p(\nabla h^{-2}(q)) = -2h(q)^{-3}p(\nabla h(q)).$$

We define the 1-form $\lambda_\epsilon = \lambda - \delta\,\epsilon\,\mathrm{d}u(q,p)$ on $\mathrm{T}^*\Omega$, where λ is the Liouville 1-form on $\mathrm{T}^*\Omega$, and $\delta > 0$ is a universal constant, independent of all the data, that we will soon fix. Notice that

$$\mathrm{d}\lambda_\epsilon = \mathrm{d}\lambda = -\omega. \tag{1.15}$$

A computation shows that, on the energy hypersurface Σ_ϵ,

$$\lambda_\epsilon(X_{H_\epsilon})|_{\Sigma_\epsilon} \geq \|p\|^2\big(\underbrace{1 - \tfrac{3}{2}\delta}_{(*)}\big) + \tfrac{\delta}{2}(\underbrace{1 - \|p\|^2}_{>0})^3\|\nabla h(q)\|^2 - \mathrm{const}\sqrt{\epsilon}.$$

We now fix $\delta \in (0, \frac{2}{3})$, so that $(*)$ is always positive. For all $(q,p) \in \Sigma_\epsilon$, we have

$$\lambda_\epsilon(X_{H_\epsilon}(q,p)) \geq \begin{cases} \frac{1}{4}(1 - \frac{3}{2}\delta) - \mathrm{const}\sqrt{\epsilon}, & \text{if } \|p\|^2 \geq \frac{1}{2}, \\ \frac{\delta}{16} - \mathrm{const}\sqrt{\epsilon}, & \text{if } \|p\|^2 \leq \frac{1}{2}. \end{cases}$$

Overall, for all $\epsilon > 0$ small enough, we have

$$\lambda_\epsilon(X_{H_\epsilon})|_{\Sigma_\epsilon} \geq c, \tag{1.16}$$

where $c > 0$ is a universal constant independent of the billiard table Ω. Equations (1.15) and (1.16) imply that the restriction $\alpha_\epsilon := \lambda_\epsilon|_{\Sigma_\epsilon}$ is a contact form, whose Reeb vector field satisfies $R_\epsilon = \lambda_\epsilon(X_{H_\epsilon})^{-1}X_{H_\epsilon}$.

We claim that the displacement energy $e(\Sigma_\epsilon)$ is bounded in terms of the affine displacement by

$$e(\Sigma_\epsilon) \leq 4\,\mathrm{displ}(\Omega).$$

Indeed, consider a vector $v \in \mathbb{R}^n$ such that $\overline{\Omega} \cap (v + \overline{\Omega}) = \varnothing$, and a smooth function $\chi : [0,\infty) \to [0,1]$ such that $\chi|_{[0,1]} \equiv 1$ and $\mathrm{supp}(\chi) = [0,2]$. We define the Hamiltonian

$$G : \mathbb{R}^n \to \mathbb{R}, \qquad G(q,p) = \chi(\|q\|/r)\chi(\|p\|)p(v).$$

If the constant $r > 0$ is large enough, we have $\phi_G^t(q,p) = (q+t,p)$ for all $(q,p) \in \Sigma_\epsilon$ and $t \in [0,1]$. Therefore $\phi_G^1(\Sigma_\epsilon) \cap \Sigma_\epsilon = \varnothing$, and

$$e(\Sigma_\epsilon) \leq \max G - \min G \leq 4\|v\|$$

By Theorem 1.29, the restricted contact type hypersurface $(\Sigma_\epsilon, \alpha_\epsilon)$ has a closed Reeb orbit ζ_ϵ of period $\sigma_\epsilon \leq e(\Sigma_\epsilon)$ and Maslov index $\mathrm{mas}(\zeta_\epsilon) \in \{n, n+1\}$. Since R_ϵ and X_{H_ϵ} are positively proportional, a reparametrization of ζ_ϵ is a periodic orbit $(\gamma_\epsilon, \dot\gamma_\epsilon)$ of H_ϵ, that is, $\gamma_\epsilon = (q_\epsilon, \tau_\epsilon) \in \mathrm{crit}(\mathcal{S}^\epsilon)$. The fixed-time Morse index of this periodic orbit is $\mathrm{ind}(q_\epsilon) = \mathrm{mas}(\zeta_\epsilon) \in \{n, n+1\}$. By (1.16), the period is bounded by

$$\tau_\epsilon \leq c^{-1}\sigma_\epsilon \leq c^{-1}e(\Sigma_\epsilon) \leq 4c^{-1}\,\mathrm{displ}(\Omega).$$

By the compactness theorem in the first part of the proof, there exists a sequence $\epsilon_k \to 0^+$ such that γ_{ϵ_k} converges towards a unit-speed periodic billiard trajectory of period $\tau \leq 4c^{-1}\,\mathrm{displ}(\Omega)$ and at most $n+1$ bounces. □

CHAPTER 2

Geodesic flows

This chapter is devoted to one of the most important settings for the study of Hamiltonian periodic orbits: the one of closed geodesics in Riemannian and Finsler manifolds. While the subject is rather mature, the first contributions dating back to Hadamard and Poincaré, it is still a source of inspiration in symplectic dynamics, and several celebrated open problems remain to be settled. In the following sections, together with a brief account of the background, we will present the author's (admittedly very modest) contributions.

2.1. Closed geodesics on Riemannian manifolds

A **closed geodesic** in a Riemannian manifold (M,g) is a geodesic $\gamma:\mathbb{R}\looparrowright M$ that is periodic, i.e. $\gamma=\gamma(\tau+\cdot)$ for some $\tau>0$. As we already remarked in Example 1.1, the geodesics are the non-constant solutions of the Euler-Lagrange equation of the Tonelli Lagrangian[1]

$$L:\mathrm{T}M\to[0,\infty),\qquad L(q,v)=\|v\|_g^2.$$

The dual Tonelli Hamiltonian is similarly given by

$$H:\mathrm{T}^*M\to[0,\infty),\qquad H(q,p)=\|p\|_g^2,$$

where $\|\cdot\|_g$ now denotes the Riemannian norm on covectors. A peculiar property of this Hamiltonian system is that the dynamics is the same on every energy level $H^{-1}(e)$ with $e>0$. Indeed, a geodesic $\gamma:\mathbb{R}\looparrowright M$ parametrized with unit-speed $\|\dot\gamma\|_g\equiv1$ provides on every energy level $e>0$ the Hamiltonian orbit

$$\phi_H^t(\gamma(0),e\dot\gamma(0)^\flat)=(\gamma(et),e\dot\gamma(et)^\flat),$$

where $\dot\gamma^\flat=g(\dot\gamma,\cdot)$. In other words, the existence problem for closed geodesics on a Riemannian manifold can be studied as the existence problem for Hamiltonian periodic orbits with arbitrary period in the energy level $H^{-1}(1)$ (as in Section 1.5), or equivalently as the existence problem for Hamiltonian periodic orbits with period 1 in the whole phase space T^*M (as in Section 1.2). We will follow the latter approach, which is more common in the literature (at least for those results based on Morse theory).

In the geodesic setting, the fixed-time action Lagrangian action functional is usually denoted by

$$E:\Lambda M\to[0,\infty),\qquad E(\gamma)=\int_{S^1}\|\dot\gamma(t)\|_g^2\mathrm{d}t,\tag{2.1}$$

[1]In Example 1.1, we rather wrote the Lagrangian $L(q,v)=\frac12\|v\|_g^2$. The factor $\frac12$ is more common in the Hamiltonian literature, whereas in the Riemannian one it is often omitted.

where $S^1 = \mathbb{R}/\mathbb{Z}$ and $\Lambda M = W^{1,2}(S^1, M)$, and is called the **energy functional**.

Remark 2.1. There is unfortunately a clash of notation and terminology, which should nevertheless not cause much confusion in the following: before now, the term "energy" rather referred to the values of the Hamiltonian, and in Section 1.5 the letter E denoted the Hamiltonian seen in the tangent bundle. The first author who referred to the functional (2.1) as to the energy is probably Milnor, who was well aware of this conflict and provided a justification for its terminology in the preface of [**Mil63**]. Since Milnor's celebrated monograph, the terminology sticked in the literature. □

The subspace $E^{-1}(0)$ of global minimizers of E contains precisely the constant curves; in the following we will always identify this space with the manifold M itself, i.e.

$$M \equiv E^{-1}(0) \subset \Lambda M.$$

The closed geodesics $\gamma \in \Lambda M$ are precisely the critical points of E with positive critical values $E(\gamma) > 0$. On a closed manifold with non-trivial fundamental group, the existence of a closed geodesic can be easily established by minimizing the energy on a connected component of non-contractible loops (Theorem 1.4). On simply connected closed manifolds, the existence of closed geodesics requires a min-max procedure similar to the one of Theorem 1.6. The result, which we now sketch, was inspired by the seminal work of Poincaré [**Poi05**] and proved for the 2-sphere by Birkhoff [**Bir66**] and for general closed manifolds by Lusternik and Fet [**LF51**].

Theorem 2.2 (Birhoff, Lusternik-Fet). *Every simply connected closed Riemannian manifold of dimension at least* 2 *has a closed geodesic.*

Proof. Let (M, g) be a simply connected closed Riemannian manifold of dimension at least 2. Algebraic topology guarantees that such a manifold is non-contractible, and indeed that there exists a non-trivial homotopy group $\pi_d(M)$ with degree $2 \leq d \leq \dim(M)$. Since M is simply connected, we can ignore the basepoint in the definition of the homotopy groups: as a set, $\pi_d(M)$ is the family of homotopy classes of continuous maps $u : S^d \to M$. We partition the sphere into smooth loops $\gamma_z : S^1 \to S^n$ parametrized by points z in an $(n-1)$-ball $B^{n-1} \subset S^n$, as in Figure 2.1 (all the γ_z's are embedded loops, except those with $z \in \partial B^{n-1}$ which are constants). Any sufficiently regular map $u : S^d \to M$ can be rewritten as a continuous map $\widetilde{u} : B^{n-1} \to \Lambda M$, $\widetilde{u}(z) = u \circ \gamma_z$, whose restrictions to ∂B^{n-1} take values inside the subspace of constant loops $M \subset \Lambda M$. We fix a non-zero homotopy class $h \in \pi_d(M)$, and define the min-max value

$$c := \inf_{\widetilde{u}} \max_{z} E(\widetilde{u}(z)),$$

where the infimum ranges over the continuous maps $\widetilde{u}$ as above corresponding to a representative u of h. The usual min-max argument as in the first part of the proof of Theorem 1.4 implies that c is a critical value of E.

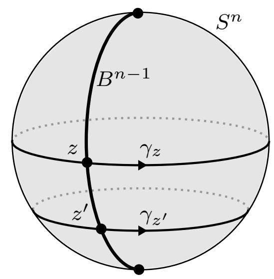


FIGURE 2.1. Partition of S^n into loops γ_z parametrized by $z \in B^{n-1}$.

It remains to show that $c > 0$. Indeed, c is larger than the squared injectivity radius $\mathrm{inj}(M, g)^2$. Otherwise, we could find a representative $\widetilde{u}$ such that $E(\widetilde{u}(z)) < \mathrm{inj}(M, g)^2$ for all $z \in B^{n-1}$. This, together with the Holder inequality, implies that each loop $\widetilde{u}(z)$ has length less than $\mathrm{inj}(M, g)$, and therefore can be shrunk to a constant curve at $\widetilde{u}(z)(0) \in B^{n-1}$ by means of the exponential map. This would produce a new representative v for h such that $\widetilde{v}(z)$ is a constant curve for all $z \in B^{n-1}$, which would imply $h = [u] = [v] = 0$. □

Actually, in his original work [**Poi05**], Poincaré was already addressing a somewhat harder statement: the existence of simple closed geodesics on a convex 2-sphere $M \subset \mathbb{R}^3$, equipped with the Riemannian metric g inherited from the ambient Euclidean one. By **simple closed geodesic** we mean an embedded geodesic $\gamma : S^1 \hookrightarrow M$. A sphere $M \subset \mathbb{R}^3$ is convex when it is a regular level set of a function $F : \mathbb{R}^3 \to \mathbb{R}$ with everywhere positive definite Hessian; equivalently, the induced Riemannian metric g has positive Gaussian curvature. Poincaré's recipe for detecting a simple closed geodesic consisted in looking for the shortest embedded circle in M bisecting the integral Gaussian curvature. This idea was formalized more than half a century later by Berger and Bombieri [**BB81**] under the simplifying assumption that M is C^3-close to the unit sphere in $\mathbb{R}^3$.

The following stronger existence result was proved by Lusternik and Schnirelmann [**LS34**] with a different approach.

Theorem 2.3 (Lusternik-Schnirelmann). *Every Riemannian 2-sphere has three simple closed geodesics.* □

While this theorem was one of the first applications of Lusternik-Schnirelmann cuplength estimates in critical point theory, its proof required an ad hoc variation of the energy gradient-flow deformation in the space of smooth embedded loops, which originally had a gap. For decades, the theorem remained controversial, and several authors proposed their solutions to fix the gap, see, e.g. [**Bal78, Jos89, Jos91, HS94**]. The construction of such a deformation turned out to be tricky business, but nevetheless a solid and universally accepted solution was provided by Grayson [**Gra89**] with its curve shortening flow.

With the existence of at least finitely many closed geodesics settled, the efforts focused on the existence of infinitely many. The idea consists in exploiting more thoroughly the topology of the free loop space ΛM: a rich topology forces the existence of many critical circles $S^1 \cdot \gamma$ of the energy functional E. However, as we already pointed out in Section 1.4, this alone does not guarantee the existence of many closed geodesics. Indeed, any closed geodesic $\gamma \in \mathrm{crit}(E)$ provides an infinite sequence of critical circles $S^1 \cdot \gamma^m \subset \mathrm{crit}(E)$, $m \in \mathbb{N}$, where $\gamma^m(t) = \gamma(t)$ is the m-th iterate of the original geodesic γ. The work of Bott [**Bot56**] and Gromoll-Meyer [**GM69a**] provided a complete understanding of the behavior of the critical point data (the Morse index, the nullity, and the local homology) of the critical circles $S^1 \cdot \gamma^m$ as m varies. This allows in certain cases to recognize whether two critical circles $S^1 \cdot \zeta_1$ and $S^1 \cdot \zeta_2$ obtained via min-max procedures correspond to distinct closed geodesics (that is, are not critical circles belonging to the same sequence $S^1 \cdot \gamma^m$, $m \in \mathbb{N}$).

Still, before the late 1960s, it was not known whether a given simply connected closed manifold M of dimension at least 2 always had infinitely many closed geodesic for any possible choice of a Riemannian metric g. A breakthrough was found by Gromoll and Meyer [**GM69b**], who established the existence of infinitely many closed geodesics for those closed Riemannian manifolds whose free loop space has unbounded Betti numbers. The foundation paper [**VPS76**] of Vigué Poirrier and Sullivan on rational homotopy theory allowed to translate this condition on the loop space homology into a simpler condition of the cohomology of the Riemannian manifold. Overall, the resulting theorem is the following.

Theorem 2.4 (Gromoll-Meyer, Vigué Poirrier-Sullivan). *Any closed simply connected Riemannian manifold (M, g) whose rational cohomology ring $H^*(M;\mathbb{Q})$ is non-monogenic has infinitely many closed geodesics.* □

The cohomology ring $H^*(M;\mathbb{Q})$ is monogenic when it is generated as a ring by a single element $x \in H^d(M;\mathbb{Q})$, and thus $H^*(M;\mathbb{Q}) = \mathbb{Q}[x]/(x^k)$ for some $k > 0$. Among the closed manifolds M with such a cohomology ring are the simply connected compact rank-one symmetric spaces:

- the spheres S^n with $n > 1$ ($d = n$, $k = 2$),
- the complex projective spaces $\mathbb{CP}^n$ ($d = 2$, $k = n + 1$),
- the quaternionic projective spaces $\mathbb{HP}^n$ ($d = 4$, $k = n + 1$),
- the Cayley plane $\mathrm{Ca}\mathbb{P}^2$ ($d = 8$, $k = 3$).

For these spaces, the existence of infinitely many closed geodesics for any choice of the Riemannian metric is still an open problem, with the exception S^2, for which the result followed by a combination of spectacular work by Bangert [**Ban93**], Franks [**Fra92**], and Hingston [**Hin93**] (actually, the result follows by combining the work of Bangert together with either Franks' or Hingston's one).

Theorem 2.5 (Bangert-Franks-Hingston). *Every Riemannian 2-sphere has infinitely many closed geodesics.* □

As for non simply connected closed manifolds, a "rich" fundamental group may directly provide the existence of infinitely many closed geodesics. This is the case, for instance, when the fundamental group is abelian and contains $\mathbb{Z}^2 = \langle e_1, e_2 \rangle$ as subgroup: by minimizing the energy functional over the connected components of loops in the homotopy classes $e_1 + me_2$, with $m \in \mathbb{N}$, we obtain infinitely many distinct closed geodesics. The problem becomes highly non-trivial if the fundamental group is abelian and has rank one, for instant when it is isomorphic to $\mathbb{Z} = \langle e \rangle$: this time, by minimizing the energy functional over the connected components of loops in the homotopy classes me, with $m \in \mathbb{N}$, we may just find the iterates of the same closed geodesic. Nevertheless, the existence of infinitely many closed geodesics is still true, as was established by Bangert and Hingston in their short paper [**BH84**] (a true gem in the literature).

Theorem 2.6 (Bangert-Hingston). *Every closed Riemannian manifold of dimension at least 2 with an infinite abelian fundamental group has infinitely many closed geodesics.* □

In the rest of the chapter, we will sketch the proofs of generalizations of the last four theorems. The subject of closed geodesics being vast, there are many important theorems that we did not mention here: among them, the multiplicity results for closed geodesics on positively curved closed Riemannian manifolds [**BTZ82, BTZ83**], for Riemannian manifolds with a generic choice of the metric [**BTZ81, Hin84, Rad89, Rad94**], and for manifolds of dimension less than 5 [**LD09, DL10**]. We refer the reader to, e.g., [**Kli78, Ber03, Tai10b, Oan15**] for more details and references.

⋆ **Open problem:** Does any closed Riemannian manifold possess infinitely many closed geodesics? At the time of this writing, even the existence of two closed geodesics on any Riemannian sphere of arbitrary dimension would be a breakthrough.

2.2. Complete Riemannian manifolds

Non-compact Riemannian manifolds may have no closed geodesics at all: this is the case, for instance, for the Euclidean spaces. Consequently, any existence and multiplicity result must impose further conditions on the topology of the manifold and on its geometry at infinity. Benci and Giannoni [**BG92**] showed that a complete Riemannian manifold (M, g) always has a closed geodesic provided its sectional curvature is non-positive at infinity and its free loop space has non-trivial homology $H^d(\Lambda M)$ for some coefficient field and some degree $d \geq 2\dim(M)$. From the point of view of critical point theory, the lack of compactness prevents the energy functional to satisfy the Palais-Smale condition. Benci and Giannoni beautiful idea consisted in recovering the compactness by means of a penalization method, taking advantage of the geometry of the manifold at infinity.

In a joint work with Asselle [**AM17**], we improved Benci and Giannoni's theorem in a few directions. The first one consisted in slightly relaxing the geometric condition at infinity as follows. We recall that a **Jacobi vector field** $\xi : [0, \tau] \to \mathrm{T}M$ along a geodesic $\gamma : [0, \tau] \to M$ is a vector field obtained by $\xi(t) = \partial_s|_{s=0}\gamma_s(t)$, where γ_s is a family of geodesics smoothly depending on a parameter $s \in (-\epsilon, \epsilon)$ and such that $\gamma_0 = \gamma$. Two points $\gamma(0)$ and $\gamma(\tau)$ are **conjugate** along a geodesic segment $\gamma : [0, \tau] \to M$ when there is a Jacobi vector field ξ along γ such that $\xi(0) = 0$ and $\xi(\tau) = 0$. We say that a complete Riemannian manifold (M, g) is **without close conjugate points at infinity** when, for every $\ell > 0$, there exists a compact subset $K_\ell \subset M$ such that every geodesic segment $\gamma : [0,1] \to M \setminus K_\ell$ of length at most ℓ does not contain conjugate points along γ. The classical Rauch comparison theorem [**dC92**, Prop. 2.4] implies that a complete Riemannian manifold with non-positive sectional curvature at infinity is indeed without close conjugate points at infinity.

Let (M, g) be a complete Riemannian manifold without close conjugate points at infinity. If $h \in H_d(\Lambda M)$ is a non-zero homology class, unfortunately the associated spectral invariant (defined as in Theorem 1.6) may not be a critical value of E, since the energy functional may not satisfy the Palais-Smale condition when M is non-compact. The Palais-Smale condition can be recovered by means of the following penalization trick due to Benci and Giannoni. We consider a sequence of smooth proper functions $f_\alpha : M \to [0, \infty)$, $\alpha \in \mathbb{N}$, such that

$$f_\alpha \geq f_{\alpha+1}, \qquad \bigcup_{\alpha \geq 1} f_\alpha^{-1}(0) = M$$

Such functions can be easily constructed by means of a partition of unity. The modified energy functional

$$E_\alpha : \Lambda M \to [0, \infty), \qquad E_\alpha(\gamma) = E(\gamma) + f_\alpha(\gamma(0)).$$

now satisfies the Palais-Smale condition; indeed, thanks to the auxiliary proper function f_α, the curves in a sublevel set $E_\alpha^{-1}[0, c]$ are constrained within the compact subset $f_\alpha^{-1}[0, c] \subset M$. The price to pay for this trick is that the variational principle of the modified energy E_α is spoiled: the critical points of E_α are now the geodesics $\gamma : [0,1] \to M$ such that $\gamma(0) = \gamma(1)$ and $\dot\gamma(1) - \dot\gamma(0) = \nabla f_\alpha(\gamma(0))$, where ∇ denotes the gradient with respect to the Riemannian metric g. Nevertheless, it is still possible to detect closed geodesics by means of E_α as follows. We denote by $\mathrm{ind}_\alpha(\gamma)$ and $\mathrm{nul}_\alpha(\gamma)$ the Morse index and the nullity of the energy functional E_α at a critical point $\gamma \in \mathrm{crit}(E_\alpha)$.

Lemma 2.7. *For all $\ell > 0$ there exists $\overline{\alpha}(\ell) \in \mathbb{N}$ such that, for all $\alpha \geq \overline{\alpha}(\ell)$, any $\gamma \in \mathrm{crit}(E_\alpha)$ with $E_\alpha(\gamma) \leq \ell^2$ and $\mathrm{ind}_\alpha(\gamma) + \mathrm{nul}_\alpha(\gamma) > \dim(M)$ is a curve entirely contained in the interior of the compact subset $f_\alpha^{-1}(0) \subset M$, and in particular is a closed geodesic.*

Proof. Since (M, g) is without close conjugate points at infinity, there exists a compact subset $K_\ell \subset M$ such that all geodesic segments of length at most ℓ and not intersecting K_ℓ are without conjugate points. We recall that $f_\alpha \geq f_{\alpha+1}$ pointwise. Therefore, for all α large enough, the support of f_α does not intersect the neighborhood of radius ℓ of the compact set K_ℓ. Notice that any $\gamma \in \mathrm{crit}(E_\alpha)$ with $E_\alpha(\gamma) \leq \ell^2$ is a curve of length at most ℓ. If $\gamma(0)$ is outside the support of f_α, then γ is a closed geodesic. If instead $\gamma(0)$ is in the support of f_α, then γ does not intersect K_ℓ, and in particular is without conjugate points. The classical Morse index theorem from Riemannian geometry [**Mil63**] implies that $\mathrm{d}^2E(\gamma)[\eta, \eta] > 0$ for all non-zero vector fields $\eta : [0,1] \to \mathrm{T}M$ along γ such that $\eta(0) = \eta(1) = 0$. This does not quite imply that the Hessian $\mathrm{d}^2E_\alpha(\gamma)$ is positive definite, as the domain of such Hessian include all periodic vector fields along γ that are possibly non-vanishing at time $t = 0$. Nevertheless, a linear algebra argument implies that $\mathrm{ind}_\alpha(\gamma) + \mathrm{nul}_\alpha(\gamma) \leq \dim(M)$. □

The first existence result that we have, which is a slight improvement of Benci and Giannoni's one, is the following.

Theorem 2.8. *Let (M, g) be a complete Riemannian manifold without close conjugate points at infinity. If the homology $H_d(\Lambda M)$ is non-trivial in some degree $d > \dim(M)$ and for some coefficient ring, then (M, g) has a closed geodesic.*

Proof. We consider a non-zero homology class $h \in H_d(\Lambda M)$ of degree $d > \dim(M)$, and the associated spectral invariants

$$c_\alpha(h) := \inf \big\{ c \geq 0 \; \big| \; h \in \mathrm{im}\big(H_*(E_\alpha^{-1}[0, c]) \to H_*(\Lambda M)\big) \big\},$$

which are critical values of the corresponding E_α. The same argument as in Theorem 1.6 implies that there exists a critical point $\gamma_\alpha \in \mathrm{crit}(E_\alpha) \cap E_\alpha^{-1}(c_\alpha(h))$ with $\mathrm{ind}_\alpha(\gamma_\alpha) + \mathrm{ind}_\alpha(\gamma_\alpha) \geq d > \dim(M)$. Since $E_\alpha > E_{\alpha+1}$ pointwise, we have $c_\alpha(h) > c_{\alpha+1}(h)$. Therefore, by Lemma 2.7, for all α large enough the critical point γ_α is a closed geodesic. □

In order to prove the second result, we first need two ingredients from the general theory of closed geodesics. The first one, which goes back to Bott [**Bot56**], concerns the Morse indices of iterated closed geodesics: if $\gamma \in \mathrm{crit}(E)$ is a closed geodesic, the function $m \mapsto \mathrm{ind}(\gamma^m)$ can be fully understood by means of elementary Fourier analysis. In particular, there is a well define finite limit, called the **average Morse index**

$$\overline{\mathrm{ind}}(\gamma) := \lim_{m \to \infty} \frac{\mathrm{ind}(\gamma^m)}{m} \in [0, \infty).$$

The Morse indices satisfy the inequality

$$m \,\overline{\mathrm{ind}}(\gamma) - \dim(M) \leq \mathrm{ind}(\gamma^m) \leq m \,\overline{\mathrm{ind}}(\gamma) + \dim(M) - \mathrm{nul}(\gamma^m). \tag{2.2}$$

Moreover, $\overline{\mathrm{ind}}(\gamma) = 0$ if and only if $\mathrm{ind}(\gamma^m) = 0$ for all $m \geq 1$.

The second ingredient, which was provided by Gromoll and Meyer [**GM69a, GM69b**], concerns the behavior of the local homology of iterated closed geodesics. If $\gamma \in \mathrm{crit}(E) \cap E^{-1}(c)$ is a closed geodesic, its m-th iterate has energy $E(\gamma^m) = m^2 c$, and the corresponding critical circle has local homology

$$C_*(S^1 \cdot \gamma^m) := H_*(E^{-1}[0, m^2 c] \cup S^1 \cdot \gamma^m, E^{-1}[0, m^2 c]).$$

In Section 1.4 we already discussed some properties of the local homology of periodic orbits that are isolated critical points of the Lagrangian action functional. Analogous conclusions hold in the geodesics settings: if $\mathrm{nul}(\gamma^{m_1}) = \mathrm{nul}(\gamma^{m_2})$ for positive integers m_1, m_2 with $m_2/m_1 \in \mathbb{N}$, then there is an isomorphism of local homologies $C_*(S^1 \cdot \gamma^{m_1}) \cong C_*(S^1 \cdot \gamma^{m_2})$. The Hamiltonian characterization of the nullity (1.7), together with some elementary linear algebra, implies that we have a finite partition of the natural numbers $\mathbb{N} = \mathbb{N}_1 \cup ... \cup \mathbb{N}_r$ such that, if $m_i := \min \mathbb{N}_i$, every $m \in \mathbb{N}_i$ is a multiple of m_i and satisfies $\mathrm{nul}(\gamma^m) = \mathrm{nul}(\gamma^{m_i})$. Overall, this implies that $\mathrm{rank}\, C_*(S^1 \cdot \gamma^m)$ is uniformly bounded from above for all $m \in \mathbb{N}$.

We denote the Betti numbers of the free loop space by $\beta_d(\Lambda M) := \mathrm{rank}\, H_d(\Lambda M)$ with respect to some coefficient ring. The contributions of Vigué Poirrier and Sullivan [**VPS76**] to Theorem 2.4 was to translate the condition on the rational cohomology of M to a condition on the rational Betti numbers of the loop space: if the manifold M is simply connected and its cohomology $H^*(M;\mathbb{Q})$ has finite total rank (as is the case when M is closed), the sequence $\{\beta_d(\Lambda M;\mathbb{Q}) \mid d \geq 0\}$ is unbounded if and only if the cohomology ring $H^*(M;\mathbb{Q})$ is not monogenic. The second result that we present is an extension of Theorem 2.4 to possibly non-compact Riemannian manifolds.

Theorem 2.9. *Let (M, g) be a complete Riemannian manifold without close conjugate points at infinity. If the sequence of Betti numbers $\{\beta_d(\Lambda M) \mid d > \dim(M)\}$ is unbounded with respect to some coefficient ring, then (M, g) has infinitely many closed geodesics.*

Proof. We prove the theorem by contradiction, by assuming that (M, g) has only finitely many closed geodesics $\gamma_1, ..., \gamma_k$. The above discussion on the local homology guarantees that

$$b := \max\big\{\, \mathrm{rank}\, C_*(S^1 \cdot \gamma_i^m) \;\big|\; i = 1, ..., k,\ m \geq 1 \big\} < \infty.$$

The local homology $C_d(S^1 \cdot \gamma_i^m)$ can be non-zero only if $\mathrm{ind}(\gamma^m) \leq d \leq \mathrm{ind}(\gamma^m) + \mathrm{nul}(\gamma^m)$. This, together with the iteration inequality (2.2), implies that there exists $r > 0$ such that, for all $d > \dim(M)$, at most r critical circles of E have non-trivial local homology in degree d.

We fix the degree $d > \dim(M)$ so that $\beta_d(\Lambda M) > rb$, and we choose $h_1, ..., h_{rb+1} \in H_d(\Lambda M)$ that generate a subgroup of rank $rb + 1$. We set

$$a := \max\{c_1(h_1), ..., c_1(h_{rb+1})\} + 1,$$

so that $a > c_\alpha(h_i)$ for all $\alpha \geq 1$ and $i = 1, ..., rb + 1$. This implies

$$\mathrm{rank}\, H_d(E_\alpha^{-1}[0, a)) > rb + 1, \qquad \forall \alpha \geq 1. \tag{2.3}$$

We fix α to be large enough so that, by Lemma 2.7, every critical point $\gamma \in \mathrm{crit}(E) \cap E_\alpha^{-1}[0, a)$ with $\mathrm{ind}_\alpha(\gamma) + \mathrm{nul}_\alpha(\gamma) > \dim(M)$ is contained in the interior of $f_\alpha^{-1}(0)$, and in particular is a closed geodesic; in this case, $\mathrm{ind}_\alpha(\gamma)$ and $\mathrm{nul}_\alpha(\gamma)$ coincide with the corresponding usual indices $\mathrm{ind}(\gamma)$ and $\mathrm{nul}(\gamma)$.

Now consider an interval $[e_1, e_2] \subset [0, a)$ containing a unique critical value c of E_α. Those critical points $\zeta \in \mathrm{crit}(E_\alpha) \cap E_\alpha^{-1}(c)$ with $\mathrm{ind}(\zeta) + \mathrm{nul}(\zeta) \leq \dim(M) < d$ do not contribute to the relative homology $H_d(E_\alpha^{-1}[0, e_2], E_\alpha^{-1}[0, e_1))$. The remaining critical points $\gamma \in \mathrm{crit}(E_\alpha) \cap E_\alpha^{-1}(c)$ with $\mathrm{ind}(\zeta) + \mathrm{nul}(\zeta) \geq d$ are closed geodesics, and contribute to the relative homology group $H_d(E_\alpha^{-1}[0, e_2], E_\alpha^{-1}[0, e_1))$ with their local homology. Namely, the inclusion induces an isomorphism

$$\bigoplus_{S^1 \cdot \gamma} C_d(S^1 \cdot \gamma) \cong H_d(E_\alpha^{-1}[0, e_2], E_\alpha^{-1}[0, e_1)). \tag{2.4}$$

A well known argument with long exact sequences of inclusions implies that the function

$$F(e,f) := \operatorname{rank} H_d(E_\alpha^{-1}[0,f], E_\alpha^{-1}[0,e))$$

is subadditive, meaning that $F(e,h) < F(e,f) + F(f,h)$ for all $e < f < h$. This, together with (2.4), implies

$$\operatorname{rank} H_d(E_\alpha^{-1}[0,a)) \leq \sum_{S^1\cdot\gamma} \operatorname{rank} C_d(S^1\cdot\gamma),$$

where the sum in the right-hand side ranges over all the critical circles of closed geodesics γ with non-trivial local homology in degree d. By the discussion in the first paragraph of the proof, such a sum is at most rb, contradicting (2.3). □

Theorems 2.8 and 2.9 are not the only ones available in the literature on closed geodesics in non-compact Riemannian manifolds. We mention for instance Thorbergsson's result [**Tho78**], which provides a closed geodesic on every complete and non-contractible Riemannian manifold with non-negative sectional curvature outside a compact set. Another beautiful result for surfaces is due to Bangert [**Ban80**], and contains ideas that eventually developed into the existence of infinitely many closed geodesics on Riemannian 2-spheres (Theorem 2.5). Bangert's statement provides infinitely many closed geodesics on every complete surface with finite volume and homeomorphic to either the cylinder $S^1 \times \mathbb{R}$, the Möbius strip, or the plane $\mathbb{R}^2$.

⋆ **Open problem (Bangert):** Does any complete Riemannian manifold with finite volume admit a closed geodesic? Does it admit infinitely many closed geodesics, at least if the sequence of Betti numbers of its free loop space is unbounded?

2.3. Closed geodesics on Finsler manifolds

We now consider the more general class of Finsler geodesics flows, which already appeared in Example 1.2. There are a few slightly different definitions of Finsler metric in the literature. The one that is more adapted to the study of geodesic flows is the following one: on a manifold M, a **Finsler metric** is a continuous function $F : \mathrm{T}M \to \mathbb{R}$, smooth outside the 0-section, positively homogeneous of degree 1 in the fibers of the tangent bundle (i.e. $F(q,\lambda v) = \lambda F(q,v)$ if $\lambda \geq 0$), and such that the fiberwise Hessian $\partial_{vv}F^2$ is everywhere positive definite outside the 0-section. Geometrically, a Finsler metric is completely defined by its unit tangent bundle $SM := F^{-1}(1)$, which is a hypersurface intersecting each fiber in a hypersphere $S_qM := F^{-1}(1) \cap \mathrm{T}_qM$ that encloses the origin and has positive curvature (Figure 2.2). With a general Finsler metric we may have $F(q,v) \neq F(q,-v)$ for some $(q,v) \in \mathrm{T}M$. If instead $F(q,v) = F(q,-v)$ for all $(q,v) \in \mathrm{T}M$, the Finsler metric is said to be **reversible**.

The geodesics of (M,F) are the solutions of the Euler-Lagrange equation of the Lagrangian

$$L : \mathrm{T}M \to \mathbb{R}, \qquad L(q,v) = F(q,v)^2.$$

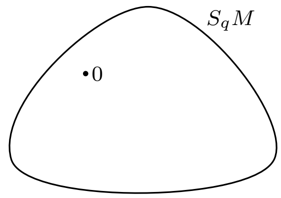


FIGURE 2.2. The unit sphere S_qM associated with a non-reversible Finsler metric on a surface M.

This Lagrangian is almost a Tonelli one, except for the fact that it may not be C^2 at the 0-section. As a consequence, the associated energy functional

$$E : \Lambda M \to [0,\infty), \qquad E(\gamma) = \int_{S^1} F(\gamma(t),\dot\gamma(t))^2 \mathrm{d}t$$

is only $C^{1,1}$, and not C^2 in general. We already encountered a similar lack of regularity in Section 1.2 for the Lagrangian action functional. As we commented there, it can be essentially ignored: with suitable finite dimensional reduction techniques, which in the Finsler setting were first worked out by Rademacher [**Rad92**], the energy can be treated as a C^∞ functional. Those results of Sections 2.1 and 2.2 whose proof does not exploit the reversibility of the Riemannian norm continue to hold if we replace the Riemannian metric with a general Finsler one: this is the case for Theorems 2.2, 2.4, 2.6, 2.8, and 2.9.

A striking example of a Finsler metric on the 2-sphere due to Katok [**Kat73**] shows that Theorems 2.3 and 2.5 fail in the Finsler categories: there are Finsler 2-spheres with only finitely many closed geodesics. Katok's example has been put in the context of Hamiltonian dynamics by Ziller [**Zil83**], who actually generalized it as follows.

Example 2.10 (Ziller). Let (M,g) be a closed Riemannian manifold of dimension at least 2, of so-called **Zoll** type: all its geodesics are closed with the same length, say length 1. We consider the Hamiltonian

$$H : \mathrm{T}^*M \to [0,\infty), \qquad H(q,p) = \|p\|_g.$$

Its Hamiltonian flow outside the zero section has the form

$$\phi_H^t(\gamma(0),\lambda\dot\gamma(0)^g) = (\gamma(t),\lambda\dot\gamma(t)^g),$$

where $\lambda > 0$, γ is a unit-speed geodesic of (M,g), and $\dot\gamma(t)^g = g(\dot\gamma(t),\cdot)$. The Zoll assumption means precisely that $\phi_H^1 = \mathrm{id}$, while ϕ_H^t has no fixed points outside the zero section for all $t \in (0,1)$. We assume that there is a non-trivial Killing vector field V on M whose flow ψ^t is 1-periodic, i.e. $\psi^1 = \mathrm{id}$. We recall that the Killing assumption means that every ψ^t is an isometry of (M,g). We lift this flow to a Hamiltonian flow on the cotangent bundle T^*M by means of the Hamiltonian

$$K : \mathrm{T}^*M \to \mathbb{R}, \qquad K(q,p) = p(V(q)).$$

The Hamiltonian flow has the form $\phi_K^t(q,p) = (\psi^t(q), p \circ \mathrm{d}\psi^t(q)^{-1})$. Since the diffeomorphisms ψ^t are isometries, we readily see that the Hamiltonians H and K commute, i.e.

$$\phi_H^t \circ \phi_K^s = \phi_K^s \circ \phi_H^t, \qquad \forall s,t \in \mathbb{R}. \tag{2.5}$$

For each $\alpha > 0$, we introduce the Hamiltonian

$$H_\alpha : \mathrm{T}^*M \to \mathbb{R}, \qquad H_\alpha(q,p) = H(q,p) + \alpha K(q,p).$$

We require $0 < \alpha < \|V\|_{L^\infty}^{-1}$, so that $\|\alpha V(q)\|_g < 1$ for all $q \in M$. This implies that H_α is positive outside the 0-section, and therefore it is the dual of the non-reversible Finsler metric F_α on M defined by

$$\tfrac{1}{2}F_\alpha(q,v)^2 = \max_{p \in \mathrm{T}_q^*M} \Big(pv - \tfrac{1}{2}H_\alpha(q,p)^2 \Big).$$

The Hamiltonian flow of H_α outside the 0-section has the form

$$\phi_{H_\alpha}^t(\gamma(0), \lambda\dot\gamma(0)^F) = (\gamma(t), \lambda\dot\gamma(t)^F),$$

where $\gamma : \mathbb{R} \to M$ is a geodesic of (M,F) parametrized with unit-speed $F_\alpha(\gamma,\dot\gamma) \equiv 1$, and $\dot\gamma(t)^F = \partial_v F_\alpha(\gamma(t),\dot\gamma(t))$. In particular, there is a one-to-one correspondence between the unit-speed τ-periodic closed geodesic γ of (M,F_α) and the cylinders of fixed points $(\gamma(t), \lambda\dot\gamma(t)^F) \in \mathrm{fix}(\phi_{H_\alpha}^\tau)$, $t \in \mathbb{R}/\tau\mathbb{Z}$, $\lambda > 0$. The commutativity relation (2.5), together with a simple computation, implies that

$$\phi_{H_\alpha}^t = \phi_H^t \circ \phi_K^{\alpha t} = \phi_K^{\alpha t} \circ \phi_H^t. \tag{2.6}$$

We require α to be irrational, and consider a fixed point

$$z = (\gamma(0), \dot\gamma(0)^g) \in \mathrm{fix}(\phi_{H_\alpha}^\tau)$$

for some $\tau > 0$, where γ is a geodesic of (M,g) with unit-speed $\|\dot\gamma\|_g \equiv 1$. By (2.6), the fact that such a z is a fixed point of $\phi_{H_\alpha}^\tau$ is equivalent to

$$\phi_H^s(z) = \phi_H^s \circ \phi_{H_\alpha}^\tau(z) = \phi_{H_\alpha}^\tau \circ \phi_H^s(z) = \phi_K^{\alpha\tau} \circ \phi_H^{\tau+s}(z), \qquad \forall s \in \mathbb{R},$$

and therefore to

$$\psi^{\alpha\tau}(\gamma(s)) = \gamma(s-\tau), \qquad \forall s \in \mathbb{R}, \tag{2.7}$$

We claim that $\alpha\tau$ is irrational. Otherwise, if $\alpha\tau = \frac{n}{m}$ for some integers n,m, we would have

$$\gamma(s) = \psi^n(\gamma(s)) = \psi^{m\alpha\tau}(\gamma(s)) = \gamma(s - m\tau),$$

which would imply that $m\tau$ is an integer, and thus that α is rational. Since $\alpha\tau$ is irrational, for each $t \in S^1 = \mathbb{R}/\mathbb{Z}$ we have a sequence $m_k \to \infty$ and some $s_t \in S^1$ such that $m_k\alpha\tau \to t$ and $m_k\tau \to s_t$ in S^1, and therefore (2.7) implies

$$\psi^t(\gamma(s)) = \gamma(s - s_t), \qquad \forall t,s \in S^1.$$

This show that, for any closed geodesic $\widetilde\gamma$ of (M,F_α) with suitable speed, the curve $\gamma(t) = \psi^{-\alpha t}(\widetilde\gamma(t))$ is a unit-speed geodesic of (M,g) invariant by the whole flow of the Killing vector field V. Conversely, if γ is a unit-speed (closed) geodesic of (M,g) invariant by the whole flow of V, the geodesic $\widetilde{\gamma}(t) = \psi^{\alpha t}(\gamma(t))$ of (M, F_α) must be closed. Overall, we have a one-to-one correspondence between closed geodesics of (M, F_α) and geodesics of (M, g) that are invariant by the flow of V.

This construction can be readily applied to the round sphere $S^2 \subset \mathbb{R}^3$ of radius $(2\pi)^{-1}$, equipped with the restriction of the ambient Euclidean metric of $\mathbb{R}^3$. We choose the Killing vector field $V(q) = -2\pi q_2 \partial_{q_1} + 2\pi q_1 \partial_{q_2}$, which generates the rotation around the q_3 axis. The Finsler metric F_α obtained by the previous construction, with $\alpha \in (0, (2\pi)^{-1})$ irrational, has exactly two unit-speed closed geodesics: the equators

$$\gamma_1(t) = (\cos(2\pi(1+\alpha)t), \sin(2\pi(1+\alpha)t), 0),$$
$$\gamma_2(t) = (\cos(-2\pi(1-\alpha)t), \sin(-2\pi(1-\alpha)t), 0);$$

none of the other geodesics is periodic. Similar constructions can be carried out for:

- higher dimensional spheres S^{2n} and S^{2n-1}, producing a Finsler manifold with $2n$ closed geodesics,
- complex projective spaces $\mathbb{CP}^n$, producing a Finsler manifold with $n(n+1)$ closed geodesics,
- quaternionin projective spaces $\mathbb{HP}^n$, producing a Finsler manifold with $2n(n+1)$ closed geodesics,
- the Cayley plane $\mathrm{Ca}\mathbb{P}^2$, producing a Finsler manifold with 24 closed geodesics. □

2.4. The curve shortening semi-flow

Katok's Finsler metrics F of Example 2.10 are non-reversible, i.e. $F(q, v) \neq F(q, -v)$ for some $(q, v) \in \mathrm{T}M$. It turns out that the lack of reversibility is what makes Theorems 2.3 and 2.5 fail in the Finsler category. In the next section, we shall present our joint results with De Philippis, Marini, and Suhr [**DPMMS20**], which in particular recover these two theorems for reversible Finsler 2-spheres. For what concerns the existence and multiplicity of closed geodesics, one should expect any result valid for general Riemannian metrics to hold for general reversible Finsler metrics as well (crucial differences between Riemannian and reversible Finsler metrics still manifest themselves in other subjects related to the study of geodesic flows, for instance in geometric inverse problems [**Ota90b, Ota90a, BI16**]).

We already mentioned that one of the main ingredients for the proof of Theorem 2.3, which caused trouble for several decades, was a length-decreasing deformation of the space of embedded loops in an orientable closed surface. A formidable solution was provided by Grayson [**Gra89**] with its curve shortening flow (which should more precisely be called a "semi-flow", as the evolution is only defined in forward time). This was later on extended to more general settings, including the one of reversible Finsler metrics, by Oaks [**Oak94**] and Angenent [**Ang08**]. We shall now present the construction in the framework of Morse theory, and in the generality necessary to extend Theorem 2.3.

Let (M, F) be an oriented, reversible Finsler, closed surface. We denote by $\mathrm{Emb}(S^1, M)$ the space of smooth embedded loops, endowed with the C^∞-topology. This is a Fréchet manifold, whose tangent space $\mathrm{T}_\gamma\mathrm{Emb}(S^1, M)$ is the Fréchet vector space of smooth vector fields along γ. Instead of working with the energy functional as in the previous sections, we will work with the length functional

$$L : \mathrm{Emb}(S^1, M) \to (0, \infty), \qquad L(\gamma) = \int_{S^1} F(\gamma(t), \dot{\gamma}(t))\, \mathrm{d}t.$$

In this setting, this functional is smooth. Moreover, it is invariant under the action of the circle diffeomorphisms group

$$\theta \cdot \gamma = \gamma \circ \theta, \qquad \forall \theta \in \mathrm{Diff}(S^1),\ \gamma \in \mathrm{Emb}(S^1, M).$$

Notice that $\mathrm{Diff}(S^1)$ includes the diffeomorphisms that reverse the orientation of the circle, and the invariance of L thus requires the reversibility of the Finsler metric F.

As much as it would be more elegant to proceed canonically, it turns out to be more convenient for the exposition (although not strictly necessary) to introduce an arbitrary auxiliary Riemannian metric g on M. We will employ g only to define an L^2 Riemannian metric G on $\mathrm{Emb}(S^1, M)$ by

$$G(\xi, \eta) = \int_{S^1} g(\xi(t), \eta(t))\, \|\dot{\gamma}(t)\|_g\, \mathrm{d}t, \qquad \forall \xi, \eta \in \mathrm{T}_\gamma \mathrm{Emb}(S^1, M).$$

The group $\mathrm{Diff}(S^1, M)$ acts on $\mathrm{Emb}(S^1, M)$ by G-isometries, i.e.

$$G(\xi \circ \theta, \eta \circ \theta) = G(\xi, \eta), \qquad \forall \theta \in \mathrm{Diff}(S^1, M),\ \xi, \eta \in \mathrm{T}_\gamma \mathrm{Emb}(S^1, M).$$

We denote by ∇L the G-gradient of the length functional, which is defined by

$$\mathrm{d}L(\gamma)\xi = G(\nabla L(\gamma), \xi),$$

and consider the associated anti-gradient flow equation

$$\partial_s \gamma_s = -\nabla L(\gamma_s), \tag{2.8}$$

where $\gamma_s \in \mathrm{Emb}(S^1, M)$ is a family of embedded circles. Since the length $L(\gamma)$ is independent of the parametrization of γ, a simple computation shows that $g(\nabla L(\gamma_s)(t), \dot{\gamma}_s(t)) = 0$ for all $t \in S^1$. Once we suitably mod out the directions parallel to $\dot{\gamma}_s$, Equation (2.8) becomes a parabolic PDE, for which there are well-known local existence and uniqueness results: for every $\gamma_0 \in \mathrm{Emb}(S^1, M)$, there exists $\epsilon > 0$ and a unique smooth solution $\gamma : [0, \epsilon) \times S^1 \to M$, $\gamma(s, t) = \gamma_s(t)$ of (2.8), smoothly depending on the initial condition γ_0. Clearly, if γ_0 is a (simple) closed geodesic, then $\nabla L(\gamma_0) = 0$, and we have a stationary solution $\gamma_s = \gamma_0$ defined for all $s \geq 0$. If γ_0 is not a closed geodesic, the function $s \mapsto L(\gamma_s)$ is strictly decreasing on its domain of definition.

The global existence of solutions of (2.8) is a deep result, originally proved by Grayson [**Gra89**] in the Riemannian case (that is, when $F = \|\cdot\|_g$), and extended by Oaks [**Oak94**] to a more general setting which includes our Finsler one.

Theorem 2.11 (Grayson, Oaks). *For each $\gamma_0 \in \mathrm{Emb}(S^1, M)$, let $[0, s_\infty) \subset [0, \infty)$ be the maximal interval of definition of the corresponding solution of* (2.8)*. For each $s \in (0, s_\infty)$, we have $\gamma_s \in \mathrm{Emb}(S^1, M)$. Let*

$$\ell_\infty := \lim_{s \to s_\infty} L(\gamma_s).$$

Then $s_\infty < \infty$ *if and only if* $\ell_\infty = 0$. □

Since all the closed geodesics of (M, F) have length larger than the injectivity radius, we can modify the PDE (2.8) on the sublevel set $L^{-1}(0, \rho)$, for any $\rho > 0$ smaller than the injectivity radius, in order to have solutions γ_s always defined for all $s \geq 0$. We do this by means of a smooth monotone increasing function $\chi : [0, \infty) \to [0, 1]$ such that $\mathrm{supp}(\chi) = [\rho/2, \infty)$ and $\chi|_{[\rho,\infty)} \equiv 1$. We define the **curve-shortening semi-flow** to be the continuous map

$$\phi : (0, \infty) \times \mathrm{Emb}(S^1, M) \to \mathrm{Emb}(S^1, M), \qquad \phi(s, \gamma_0) = \phi_s(\gamma_0) := \gamma_s,$$

where γ_s is the solution of the PDE

$$\partial_s \gamma_s = -\chi(L(\gamma_s))\nabla L(\gamma_s).$$

The following theorem summarizes its properties. Most of them follow from Theorem 2.11, except property (iv), which is crucial for the applications to Lusternik-Schnirelmann theory (Theorem 2.13) and required rather subtle estimates carried out in [**DPMMS20**] (in the Riemannian case, even though the statement does not explicitly appear, the estimates can be extracted from Grayson's paper [**Gra89**]). Given a subset $\mathcal{U} \subseteq \mathrm{Emb}(S^1, M)$ and $\ell > 0$, we denote

$$\mathcal{U}^{<\ell} := \big\{\gamma \in \mathcal{U} \;\big|\; L(\gamma) < \ell\big\}. \tag{2.9}$$

Theorem 2.12.

(i) $\phi_0 = \mathrm{id}$ *and* $\phi_{s_1} \circ \phi_{s_2} = \phi_{s_1+s_2}$ *for all* $s_1, s_2 \geq 0$*;*
(ii) $\phi_s(\gamma \circ \theta) = \phi_s(\gamma) \circ \theta$ *for all* $\theta \in \mathrm{Diff}(S^1)$ *and* $\gamma \in \mathrm{Emb}(S^1, M)$*;*
(iii) $\frac{\mathrm{d}}{\mathrm{d}s} L(\phi_s(\gamma)) \leq 0$, *with equality if and only if* γ *is a closed geodesic or* $L(\gamma) \leq \rho/2$.
(iv) *For each* $\ell > \rho$ *and* $\epsilon > 0$ *there exists* $\delta > 0$ *such that, for all* $t > 0$ *large enough, we have*

$$\phi_t(\mathrm{Emb}(S^1, M)^{<\ell+\delta}) \subset \mathcal{U}(\ell, \epsilon) \cup \mathrm{Emb}(S^1, M)^{<\ell-\delta},$$

where $\mathcal{U}(\ell, \epsilon)$ *is the* $\mathrm{Diff}(S^1)$*-invariant open subset of the space of simple closed geodesics of length* ℓ *given by*

$$\mathcal{U}(\ell, \epsilon) := \big\{\gamma \in \mathrm{Emb}(S^1, M) \;\big|\; L(\gamma) \in (\ell - \epsilon^2, \ell + \epsilon^2),\ \|\nabla L(\gamma)\|_{L^\infty} < \epsilon\big\}. \tag{2.10}$$

□

Property (ii) implies that the curve shortening semi-flow ϕ_t is well defined on the space of unparametrized embedded loops

$$\Pi = \Pi M := \frac{\mathrm{Emb}(S^1, M)}{\mathrm{Diff}(S^1)},$$

which we endow with the quotient topology. It is in this space that we will now recover the fundamental Lusternik-Schnirelmann theorem for our length functional L. We will apply the same notation (2.9) to subsets $\mathcal{U} \subseteq \Pi$, and with a slight abuse of notation we will still denote by $\mathcal{U}(\ell, \epsilon)$ the projection of the open subsets (2.10) to Π.

For any non-trivial homology class $h \in H_*(\Pi, \Pi^{<\ell})$, we consider the associated spectral invariant

$$c(h) := \inf \big\{ c > 0 \bigm| h \in \operatorname{im}\big(H_*(\Pi^{<c}, \Pi^{<\ell}) \xrightarrow{\text{incl}_*} H_*(\Pi, \Pi^{<\ell})\big)\big\} \geq \ell.$$

Property (iv) of the curve shortening semi-flow implies that $c(h)$ is the length of a simple closed geodesic. Otherwise, by a simple compactness argument, the open subset $\mathcal{U}(c(h), \epsilon)$ would be empty for $\epsilon > 0$ small enough, and property (iv) would imply that

$$h \in \operatorname{im}\big(H_*(\Pi^{<c(h)-\delta}, \Pi^{<\ell}) \xrightarrow{\text{incl}_*} H_*(\Pi, \Pi^{<\ell})\big),$$

contradicting the very definition of $c(h)$.

We now consider the cap product $\frown$ between homology and cohomology classes, which produces $h \frown k \in H_d(X, Y)$ out of $h \in H_{d+j}(X, Y)$ and $k \in H^j(X)$. By means of the curve shortening semi-flow, we can recover the following version of the classical Lusternik-Schnirelmann theorem in the setting Π.

Theorem 2.13. *Let $h \in H_{d+j}(\Pi, \Pi^{<\ell})$ and $k \in H^j(\Pi)$ be such that $h \frown k \neq 0$. Then $c(h) \geq c(h \frown k)$, and if the equality holds then $k|_{\mathcal{U}(c,\epsilon)} \neq 0$ in $H^j(\mathcal{U}(c, \epsilon))$ for all $\epsilon > 0$.*

Proof. The inequality $c(h) \geq c(h \frown k)$ is rather straightforward. Indeed, consider an arbitrary value $c > c(h)$, so that there exists a preimage $h' \in H_{d+k}(\Pi^{<c}, \Pi^{<\ell})$ of h under the inclusion

$$\iota_* := \text{incl}_* : H_{d+k}(\Pi^{<c}, \Pi^{<\ell}) \to H_{d+k}(\Pi, \Pi^{<\ell}).$$

Since $\iota_*(h' \frown \iota^* k) = (\iota_* h') \frown k = h \frown k \neq 0$, we have $c(h \frown k) \leq c$.

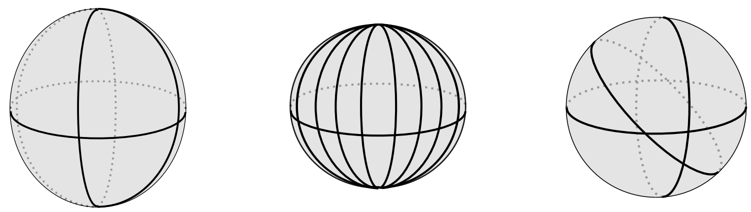
As for the remaining part of the statement, we employ property (iv) of the curve shortening semi-flow: for any given $\epsilon > 0$, if we set $\mathcal{U} := \mathcal{U}(c(h), \epsilon)$ there exists a relative cycle σ that is entirely contained in $\Pi^{<c(h)} \cup \mathcal{U}$ and such that $[\sigma] = h$ in $H_{d+j}(\Pi, \Pi^{<\ell})$. Let us assume that $k|_{\mathcal{U}} = 0$, so that there exists a cocycle κ that vanishes when evaluated on chains contained in $\mathcal{U}$, and represents $[\kappa] = k$ in $H^j(\Pi)$. After applying sufficiently many barycentric subdivisions to σ, we can assume that it splits as a sum $\sigma = \sigma' + \sigma''$, where σ' is contained in $\Pi^{<c(h)}$, while σ'' is contained in $\mathcal{U}$. By computing $h \frown k$ with the representatives that we introduced, we have

$$h \frown k = [(\sigma' + \sigma'') \frown \kappa] = [\sigma' \frown \kappa + \underbrace{\sigma'' \frown \kappa}_{=0}] = [\sigma' \frown \kappa].$$

Since the relative cycle $\sigma' \frown \kappa$ is contained in the sublevel set $\Pi^{<c(h)}$, we conclude that $c(h \frown k) < c(h)$. □

2.5. Reversible Finsler metrics on the 2-sphere

In this section, we prove a stronger version of Theorem 2.3 for general reversible Finsler 2-spheres (S^2, F). We recall that the **length spectrum** $\sigma(S^2, F)$ is the set of all



(a) (b) (c)

FIGURE 2.3. **(a)** A triaxial ellipsoid with only three simple closed geodesics. **(b)** An oblate ellipsoid of revolution, whose only simple closed geodesics are the meridians and the equator. **(c)** The round sphere, all of whose geodesics are simple closed.

the multiples of the lengths of the closed geodesics, that is, the set of all the periods of the unit-speed closed geodesics. A less common object is the **simple length spectrum** $\sigma_{\rm s}(S^2,F)\subset\sigma(S^2,F)$, which is the set of lengths of the simple closed geodesics. The following theorem was claimed for Riemannian 2-spheres by Lusternik [**Lju66**], and a proof was given by the author and Suhr [**MS18**]; by means of the Finsler curve shortening semi-flow, the same proof goes through in the general reversible Finsler case.

Theorem 2.14. *Any reversible Finsler 2-sphere (S^2,F) has at least three simple closed geodesics, and more precisely satisfies one of the following three conditions:*

(i) *$\sigma_{\rm s}(S^2,F)$ contains at least three elements;*

(ii) *$\sigma_{\rm s}(S^2,F)$ contains exactly two elements; for some $\ell\in\sigma_{\rm s}(S^2,F)$ and for all $q\in S^2$ there are infinitely many simple closed geodesics of length ℓ going through q;*

(iii) *$\sigma_{\rm s}(S^2,F)=\{\ell\}$, and every geodesic is a simple closed geodesics of length ℓ.*

Example 2.15 (Morse [Mor96]). Theorem 2.14 is sharp already in the Riemannian case. The three cases can be realized by means of ellipsoids

$$S(r_1,r_2,r_3):=\left\{q=(q_1,q_2,q_3)\in\mathbb{R}^3\ \middle|\ \frac{q_1^2}{r_1^2}+\frac{q_2^2}{r_2^2}+\frac{q_3^2}{r_3^2}=1\right\},\qquad r_1,r_2,r_3>0$$

equipped with the restriction of the Euclidean Riemannian metric of $\mathbb{R}^3$:

(i) The triaxial ellipsoid $S(r_1,r_2,r_3)$ with $r_1<r_2<r_3$ sufficiently close to 1 has exactly three simple closed geodesics $\gamma_1,\gamma_2,\gamma_3$, where $\gamma_i=S(r_1,r_2,r_3)\cap\{q_i=0\}$, see Figure 2.3(a). Any other closed geodesic of $S(r_1,r_2,r_3)$ has self-intersections.

(ii) The oblate ellipsoid of revolution $S(1,1,r)$ with $r<1$ sufficiently close to 1 has only two kind of simple closed geodesics: the equator $\gamma=S(1,1,r)\cap\{q_3=0\}$, which has length 1, and the meridians $\gamma_\theta=S(1,1,r)\cap(\cos\theta,\sin\theta,0)^\perp$, which have length slightly less than 1, see Figure 2.3(b).

(iii) In the round sphere $S(1,1,1)$, every geodesic is simple closed of length 2π; see Figure 2.3(c). □

In order to apply Lusternik-Schnirelmann theory to the 2-sphere, we need to study the topology of the space of unparametrized embedded loops

$$\Pi = \Pi S^2 := \frac{\mathrm{Emb}(S^1, S^2)}{\mathrm{Diff}(S^1)}.$$

In general, the study of the topology of loop spaces can be sophisticated business. However, Π is not the usual free loop space, and its topology (or at least the relevant one for our applications) turns out to be very simple, and completely finite dimensional. It is convenient to enlarge Π slightly, by adding to it the space of constant loops, which we identify with S^2 itself. We denote their union by

$$\overline{\Pi} = \Pi \cup S^2,$$

and we endow it with the topology induced by their inclusion into $C^\infty(S^1, S^2)/\mathrm{Diff}(S^1)$. In this topological discussion, we realize S^2 as the unit sphere in $\mathbb{R}^3$. We consider the real projective space $\mathbb{RP}^2$, whose elements are projective classes $[q] = [-q]$ with $q \in S^2$, and the tautological bundle $\mathrm{pr} : E \to \mathbb{RP}^2$, whose fibers are given by

$$\mathrm{pr}^{-1}([q]) = \big\{([q], \lambda q) \in \mathbb{RP}^2 \times \mathbb{R}^3 \mid \lambda \in [-1, 1]\big\}, \qquad q \in S^2.$$

Throughout this section, all the homology and cohomology groups will be taken with coefficients in $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$. In particular, homology and cohomology groups will be dual of one another. The total space E is homotopy equivalent to $\mathbb{RP}^2$, and therefore its cohomology ring is given by

$$H^*(E) = \frac{\mathbb{Z}_2[w]}{(w^3)},$$

where w is the generator of $H^1(E) \cong \mathbb{Z}_2$. Let $u \in H^1(E, \partial E)$ be the Thom class, which gives the Thom isomorphism

$$H^*(E) \xrightarrow{\cong} H^{*+1}(E, \partial E), \qquad r \longmapsto u \smile r.$$

The relative cohomology $H^*(E, \partial E)$ is generated as a group by $u \smile w^j$, for $j = 0, 1, 2$. Since ∂E is homeomorphic to S^2, the long exact sequence of the inclusion $\partial E \subset E$ gives an isomorphism

$$H^1(E, \partial E) \xrightarrow{\cong} H^1(E)$$

mapping $u \mapsto w$.

The space E embeds into $\overline{\Pi}$ by means of the map

$$\iota : (E, \partial E) \to (\overline{\Pi}, S^2), \tag{2.11}$$

where $\iota([q], v)$ is the (possibly constant) loop given by the intersection of S^2 with the affine plane $P([q], v) := \mathrm{span}\{q\}^\perp + v \subset \mathbb{R}^3$, see Figure 2.4. One would be tempted to show that ι induces an isomorphism in cohomology. Such a statement turns out to be subtle to prove, and instead we will simply show that ι induces a surjective homomorphism in cohomology, which suffices for our applications. We will denote by a two-head arrow $\twoheadrightarrow$ a surjective homomorphism.

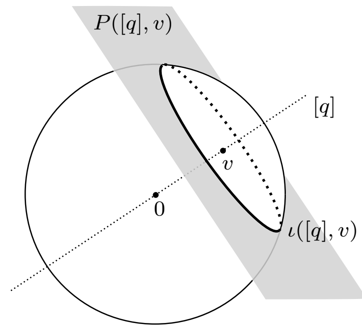


FIGURE 2.4. The map $\iota : (E, \partial E) \to (\overline{\Pi}, S^2)$.

Notice that $\iota|_{\partial E} : \partial E \to S^2$ is a homeomorphism. A simple argument "by hands" shows that ι induces an injective homomorphism of fundamental groups $\iota_* : \pi_1(E) \hookrightarrow \pi_1(\overline{\Pi})$. Since $\pi_1(E) \cong H^1(E) \cong \mathbb{Z}_2$, the map ι also induces surjective and injective homomorphisms

$$\begin{aligned} H^1(\overline{\Pi}) &\twoheadrightarrow H^1(E), \\ H_1(E, \partial E) &\hookrightarrow H_1(\overline{\Pi}, S^2). \end{aligned} \tag{2.12}$$

The relative homology $H_*(E, \partial E)$ is generated as a group by three elements e_1, e_2, e_3, which are related by

$$e_j = e_{j+1} \frown w \in H_j(E, \partial E), \qquad j = 1, 2.$$

We infer that ι induces an injective homomorphism in all degrees

$$\iota_* : H_*(E, \partial E) \hookrightarrow H_*(\overline{\Pi}, S^2), \qquad \overline{h}_j := \iota_*(e_j), \quad j = 1, 2, 3.$$

Indeed, (2.12) implies that $\overline{h}_1 \neq 0$, and for any $\widetilde{w} \in H^1(\overline{\Pi})$ such that $\iota^* \widetilde{w} = w$ we have

$$\overline{h}_{j+1} \frown \widetilde{w} = \iota_*(e_{j+1} \frown \iota^*(\widetilde{w})) = \iota_*(e_{j+1} \frown w) = \iota_*(e_j) = \overline{h}_j.$$

Modulo technical details, this topological description of the space Π was known to Lusternik and Schnirelmann, and in a slightly different setting can be found in Ballmann's memoir [**Bal78**]. The details of the computations were carried out by the author and Suhr in [**MS18**].

We now equip S^2 with any reversible Finsler metric F, and we consider the associated length functional $L : \overline{\Pi} \to [0, \infty)$ and sublevel sets

$$\mathcal{U}^{<b} := \{\gamma \in \mathcal{U} \mid L(\gamma) < b\}, \qquad \mathcal{U} \subset \overline{\Pi}, \ b > 0.$$

We can get rid of the constant loops $S^2 \subset \overline{\Pi}$ as follows. A simple argument, which does not even need the curve shortening semi-flow, implies that the inclusion $S^2 \hookrightarrow (\overline{\Pi})^{<\rho}$ is a homotopy equivalence provided $\rho > 0$ is small enough. This, together with the excision property, implies that the inclusions induce homology isomorphisms

$$\begin{aligned} H_*(\overline{\Pi}, S^2) &\xrightarrow{\cong} H_*(\overline{\Pi}, \overline{\Pi}^{<\rho}) \cong H_*(\Pi, \Pi^{<\rho}), \\ \overline{h}_j &\longmapsto h_j. \end{aligned}$$

Similarly, if $\delta \in (0,1)$ and $F := \big\{([q],v) \in E \ \big|\ \|v\| \leq 1-\delta\big\}$, the inclusions induce homology isomorphisms

$$\begin{aligned} H_*(E,\partial E) &\xrightarrow{\cong} H_*(E,\overline{E\setminus F}) \cong H_*(F,\partial F),\\ e_j &\longmapsto f_j. \end{aligned}$$

Choosing this constant $\delta > 0$ to be small enough, the map ι in (2.11) also has the form

$$\iota : (F,\partial F) \to (\Pi,\Pi^{<\rho}), \tag{2.13}$$

and for all $\widetilde{w} \in H^1(\Pi)$ such that $\iota^*(\widetilde{w}) = w|_F$ we have

$$h_j = \iota_*(f_j) = \iota_*(f_{j+1} \frown w|_F) = h_{j+1} \frown \widetilde{w}.$$

We consider the spectral invariants

$$c_j := c(h_j) := \inf\big\{c > 0 \ \big|\ h_j \in \mathrm{im}\big(H_*(\Pi^{<c},\Pi^{<\rho}) \xrightarrow{\mathrm{incl}_*} H_*(\Pi,\Pi^{<\rho})\big)\big\},\quad j = 1,2,3, \tag{2.14}$$

which are elements of the simple length spectrum $\sigma_{\mathrm{s}}(S^2,F)$, ordered as $c_1 \leq c_2 \leq c_3$ according to Theorem 2.13. Theorem 2.14 will be a direct consequence of the next two statements.

Theorem 2.16. *If $c_j = c_{j+1}$ for some $j \in \{1,2\}$, then for all $q \in S^2$ there are infinitely many simple closed geodesics of length c_j going through q.*

Proof. Assume by contradiction that $c := c_j = c_{j+1}$, but that some point $q \in S^2$ does not lie on a simple closed geodesic of length c. We consider the subset $\mathcal{W} = \{\gamma \in \overline{\Pi} \mid q \notin \gamma\}$. It is easy to see that $\mathcal{W}$ is contractible: if we denote by $B^2 \subset \mathbb{R}^2$ the unit open ball, and we consider a diffeomorphism $\theta : S^2 \setminus \{q\} \to B^2$, the homotopy $r_t : \mathcal{W} \to \mathcal{W}$, $t \in [0,1]$, given by $r_t(\gamma) = \theta^{-1}((1-t)\theta(\gamma))$ defines a contraction of $\mathcal{W}$ onto a constant loop in $S^2 \cap \mathcal{W}$. In particular $H^1(\mathcal{W})$ is trivial. Moreover, for $\epsilon > 0$ small enough, $\mathcal{W}$ contains the open subset $\mathcal{U} := \mathcal{U}(c,\epsilon)$ defined in (2.10).

Choose any $\widetilde{w} \in H^1(\overline{\Pi})$ such that $\iota^*(\widetilde{w}|_\Pi) = w|_F$, so that $h_j = h_{j+1} \frown \widetilde{w}|_\Pi$. Since $c_j = c_{j+1}$, Theorem 2.13 implies that $\widetilde{w}|_\mathcal{U} \neq 0$. However, since $H^1(\mathcal{W}) = 0$, we have $\widetilde{w}|_\mathcal{W} = 0$, and therefore $\widetilde{w}|_\mathcal{U} = (\widetilde{w}|_\mathcal{W})|_\mathcal{U} = 0$, which gives a contradiction. □

Theorem 2.17. *If $c_1 = c_2 = c_3$, then every geodesic of (S^2,F) is a simple closed geodesics of length c_1.*

Proof. We consider a circle bundle $\pi : P \to \Pi$, whose total space is given by

$$P = \{(\gamma,x) \in \Pi \times S^2 \mid x \in \gamma\}$$

and whose projection is $\pi(\gamma,x) = \gamma$. We shall employ the projectivized tangent bundle

$$\mathbb{P}\mathrm{T}S^2 = \big\{V_x \ \big|\ x \in S^2,\ V_x \text{ 1-dimensional vector subspace of } \mathrm{T}_xS^2\big\},$$

and define the continuous evaluation map $\mathrm{ev} : P \to \mathbb{P}\mathrm{T}S^2$, $\mathrm{ev}(\gamma,x) = \mathrm{T}_x\gamma$. Since $\mathbb{P}\mathrm{T}S^2$ is a closed 3-manifold, we have $H^3(\mathbb{P}\mathrm{T}S^2) \cong \mathbb{Z}_2$ (we recall that we are considering homology and cohomology with coefficients in $\mathbb{Z}_2$). We denote by m a generator of $H^3(\mathbb{P}\mathrm{T}S^2)$. We consider the map ι in (2.13), and the zero section $E_0 \subset E$, which is homeomorphic to $\mathbb{RP}^2$. We denote $\iota_0 := \iota|_{E_0} : E_0 \to \Pi$, and form the pull-back bundle

$$P_0 = \iota_0^* P = \big\{(y,p) \in E_0 \times P \mid \iota_0(y) = \pi(p)\big\}.$$

We have a commutative diagram

$$\begin{array}{ccccc} P_0 & \xhookrightarrow{\tilde\iota_0} & P & \xrightarrow{\mathrm{ev}} & \mathbb{P}\mathrm{T}S^2 \\ \big\downarrow{\scriptstyle\pi} & & \big\downarrow{\scriptstyle\pi} & & \\ E_0 & \xhookrightarrow{\iota_0} & \Pi & & \end{array}$$

where $\tilde\iota_0(y,p) = p$ is the projection onto the second factor. Notice that $\mathrm{ev} \circ \tilde\iota_0$ is a homeomorphism. Moreover, since $H^3(E_0)$ and $H^4(E_0)$ are trivial, the Gysin long exact sequence

$$\begin{array}{ccccccccccc} ... & \longrightarrow & H^3(E_0) & \xrightarrow{\pi^*} & H^3(P_0) & \xrightarrow{\pi_*} & H^2(E_0) & \longrightarrow & H^4(E_0) & \longrightarrow & ... \\ & & \| & & & & & & \| & & \\ & & 0 & & & & & & 0 & & \end{array}$$

implies that $\pi_* : H^3(P_0) \xrightarrow{\cong} H^2(E_0)$ is an isomorphism. This implies that $\pi_*\tilde\iota_0^*\mathrm{ev}^*(m) \neq 0$ in $H^2(E_0;\mathbb{Z}_2)$. We set

$$z := \pi_*\mathrm{ev}^*(m) \in H^2(\Pi;\mathbb{Z}_2).$$

Since $\iota_0^*(z) = \pi_*\tilde\iota_0^*\mathrm{ev}^*(m) \neq 0$, we must have $\iota^*(z) = w^2|_F \in H^2(F)$, and therefore $h_1 = h_3 \frown z$.

Now, assume by contradiction that $c := c_1 = c_2 = c_3$, but there exists $(q,v) \in \mathrm{T}S^2$ with $F(q,v) = 1$ and such that the geodesic $\gamma(t) := \exp_q(tv)$ is not a simple closed geodesic of minimal period c (namely, γ is not a closed geodesic, or it is closed but not simple closed, or it is simple closed but its length is not c). We consider the open subset $\mathcal{U} = \mathcal{U}(c,\epsilon) \subset \Pi$ defined in (2.10). We fix $\epsilon > 0$ to be small enough, so that v is not tangent to any curve $\gamma \in \mathcal{U}$ passing through q. Namely, if we set $P' := \pi^{-1}(\mathcal{U})$, the restriction $\mathrm{ev}|_{P'} : P' \to \mathbb{P}\mathrm{T}S^2$ is not surjective. Since $c_1 = c_3$ and $h_1 = h_3 \frown z$, Theorem 2.13 implies that $z|_{\mathcal{U}} \neq 0$ in $H^2(\mathcal{U})$. However, since $z|_{\mathcal{U}} = (\pi|_{P'})_*\mathrm{ev}|_{P'}^* m$, this implies that the homomorphism

$$\mathrm{ev}|_{P'}^* : H^3(\mathbb{P}\mathrm{T}S^2) \to H^3(P')$$

is non-zero, which is impossible since $\mathrm{ev}|_{P'}$ is not surjective. □

The closed geodesics provided by the spectral invariants (2.14), and more precisely the ones provided by the largest one c_3, allows to extend Theorem 2.5 to the class of reversible Finsler metric on the 2-sphere. The arguments remain analogous to the ones for the Riemannian case, and can be found in full details (in particular filling a few expository gaps in the previously available literature) in the joint paper of the author together with De Philippis, Marini, and Suhr [**DPMMS20**].

Theorem 2.18. *Every reversible Finsler 2-sphere has infinitely many closed geodesics.*

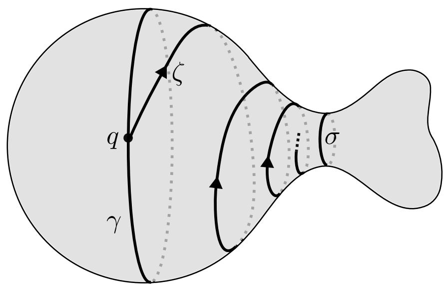


FIGURE 2.5. A geodesic ray ζ issued from $q \in \gamma$ and trapped in the right hemisphere.

Proof. We will not actually provide a proof of this theorem, but only briefly illustrate the required steps. What makes this theorem particularly hard is that there is no sigle phenomenon that guarantees at once the existence of infinitely closed geodesics; all currently available proofs are multiple case ones.

The argument begins with the spectral invariant c_3, defined in (2.14). Since this spectral invariant is associated with the relative homology class h_3, which has degree 3, one would expect to be carried by a closed geodesic with non-trivial local homology in degree 3. There is a technical subtlety here, due to the fact that the spectral invariant c_3 is defined in the setting of the unparametrized loop space $\Pi S^2 = \mathrm{Emb}(S^1, S^2)/\mathrm{Diff}(S^1)$, instead of the more common free loop space $\Lambda S^2 = W^{1,2}(S^1, M)$. Nevertheless, one can show that there exists a simple closed geodesic $\gamma : S^1 \hookrightarrow S^2$ (parametrized with constant speed so as to be a critical point of the energy functional $E : \Lambda M \to [0, \infty)$) having length $L(\gamma) = c_3$ and non-trivial local homology in degree 3, i.e.

$$C_3(\gamma) := H_3((\Lambda S^3)^{<c_3^2} \cup \{\gamma\}, (\Lambda S^3)^{<c_3^2}) \neq 0.$$

For each $t_0 \in S^1 = \mathbb{R}/\mathbb{Z}$, we denote by $t_i > 0$ the i-th positive value such that $\gamma(t_0)$ and $\gamma(t_0+t_i)$ are conjugate along $\gamma|_{[t_0,t_0+t_i]}$. The values $t_i < 0$ for $i < 0$ are defined analogously. The non-vanishing of the local homology $C_3(\gamma)$ implies in a non-trivial way that, for every $t_0 \in S^1$, the corresponding t_2 is at most 1 (i.e. the second conjugate point to $\gamma(t_0)$ occur at latest when γ closes up). We have three possible cases to consider.

In the first case, we have some point $q \in \gamma$ and some tangent vector $v \in \mathrm{T}_q S^2$ such that the geodesic ray $\zeta(t) := \exp_q(tv)$ stays trapped in one connected component $H \subset S^2 \setminus \{\gamma\}$ for all $t > 0$ (Figure 2.5). An argument, originally due to Bangert [**Ban93**] for the Riemannian case, implies that the open hemisphere H contains a waist, that is, a closed geodesic that is a local minimizer of the energy functional E (in Figure 2.5, such a waist σ is the ω-limit of the geodesic ray ζ, but the proof does not guarantee it). The existence of a waist forces the existence of infinitely many other closed geodesics; this was originally proved by Bangert [**Ban80**] in the Riemannian case, and in the general Finsler case follows by an argument similar to the one in the proof of Theorem 1.24.

If we are not in the first case, then no geodesic that intersects γ transversely remains trapped in a connected component of $S^2 \setminus \{\gamma\}$. We denote by H one such connected

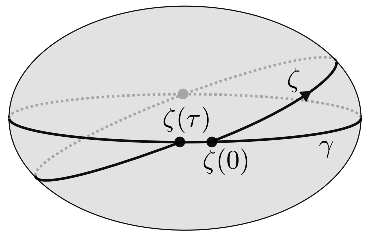


FIGURE 2.6. The first return map $\psi : A \to A$, $\psi(\zeta(0), \dot\zeta(0)) = (\zeta(\tau), \dot\zeta(\tau))$.

components. Following a seminal idea of Poicaré and Birkhoff, the open annulus

$$A := \big\{(q, v) \in \mathrm{T}S^2 \;\big|\; q \in \gamma,\ F(q, v) = 1,\ v \text{ points inside } H\big\}$$

is a surface of section for the geodesic flow. Namely, there is a first return map $\psi : A \to A$ defined by $\psi(q, v) = \zeta(\tau)$, where $\zeta(t) = \exp_q(tv)$ and $\tau > 0$ is the second positive time instant in which ζ intersects γ (Figure 2.6). Notice that there is a one-to-one correspondence between closed geodesics intersecting γ transversely and periodic orbits of ψ. Since, every point $\gamma(t)$ has conjugate points along γ, it turns out that ψ extends continuously to the closure of A by

$$\psi(\gamma(t_0), \pm\dot\gamma(t_0)) = (\gamma(t_{\pm 2}), \pm\dot\gamma(t_{\pm 2}))$$

After a suitable identification $\overline{A} \equiv S^1 \times [-1, 1]$, it turns out that ψ preserves the Lebesgue measure of $\overline{A}$.

Assume that, for every $t_0 \in S^1$, the corresponding t_2 is strictly less than 1 (i.e. the second conjugate point to $\gamma(t_0)$ occurs before γ closes up). Then ψ turns out to satisfy the twist condition of Poincaré-Birkhoff. A simple argument due to Neumann implies that a twist map of the closed annulus has infinitely many periodic orbits in the interior of the annulus. This provides infinitely many closed geodesics intersecting γ.

It remains to consider the case in which, for some t_0, the corresponding t_2 is equal to 1 (i.e. the second conjugate point to $\gamma(t_0)$ occurs exactly the first time that γ closes up). This is the hardest case, and is taken care of by means of a sophisticated argument due to Hingston [**Hin93**] (see also the paper of Asselle and the author [**AM18**] for a detailed argument and an extension of the result to non-compact Riemannian manifolds): under this assumptions, the local homology of the critical circle of γ is unstable (cf. end of Section 1.4), and forces the existence of infinitely many other closed geodesics. □

2.6. Isometry-invariant geodesics

The last topic of this chapter concerns a generalization of the closed geodesics problem for Riemannian manifolds with symmetries, which was first formulated and studied by Grove [**Gro73, Gro74, Gro85b**]. Let (M, g) be a closed Riemannian manifold of dimension at least 2, and $\psi : M \to M$ an isometry (that is, a diffeomorphism satisfying $\psi^* g = g$). Since ψ maps geodesics to geodesics, a natural question is whether there are geodesics preserved by ψ. Grove's notion of invariance is slightly more restrictive: we say that a

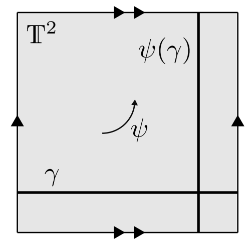


FIGURE 2.7. An isometry ψ of the flat torus $\mathbb{T}^2$ (i.e. the Riemannian square $[0,1]^2$ with the parallel portions of the boundary suitably identified as indicated by the arrows) without invariant geodesics.

geodesic $\gamma : \mathbb{R} \looparrowright M$ is **non-trivially ψ-invariant** when, for some $\tau \neq 0$, we have $\psi \circ \gamma(t) = \gamma(t+\tau)$ for all $t \in \mathbb{R}$; we stress that, with this definition, ψ preserves the orientation of the geodesic and acts on it as a non-trivial shift on its domain. Up to a reparametrization of the geodesic, we can always assume the shift to be $\tau = 1$. Notice that, when $\psi = \mathrm{id}$, the id-invariant geodesics are precisely the closed geodesics.

Theorem 2.2 implies that (M, g) has a closed geodesic. In general, however, it may have isometries not admitting any non-trivially invariant geodesic.

Example 2.19. Consider the 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ equipped with the flat Riemannian metric inherited from the Euclidean space $\mathbb{R}^2$, and the isometry $\psi : \mathbb{T}^2 \to \mathbb{T}^2$ that rotates the fundamental domain $[0,1]^2$ by a counterclockwise angle of $\pi/2$, i.e.

$$\psi([x,y]) = [1-y, x], \qquad \forall (x,y) \in [0,1]^2.$$

The geodesics of $\mathbb{T}^2$ are the straight lines, and none of them is invariant under ψ (Figure 2.7). □

The non-trivially ψ-invariant geodesics admit a variational characterization, which is the generalization of the one of closed geodesics. The ambient is the following space of ψ-invariant curves

$$\Lambda(M,\psi) := \big\{ \gamma \in W^{1,2}_{\mathrm{loc}}(\mathbb{R}, M) \ \big|\ \psi \circ \gamma(t) = \gamma(t+1) \ \ \forall t \in \mathbb{R} \big\},$$

which is the usual free loop space ΛM when $\psi = \mathrm{id}$. The critical points of the energy functional

$$E : \Lambda(M,\psi) \to [0,\infty), \qquad E(\gamma) = \int_0^1 \|\dot\gamma(t)\|_g^2 \,\mathrm{d}t.$$

with positive critical value are precisely the non-trivially ψ-invariant geodesics. All the functional properties of E valid when $\psi = \mathrm{id}$ (Section 1.2) continue to hold for general isometries ψ.

The reason why the isometry ψ of Example 2.19 has no invariant geodesics lies in the fact that ψ is not homotopic to the identity. If instead ψ is homotopic to the identity through continuous map, we can always produce a homotopy $\psi_t : M \to M$ such that

$\psi_0 = \mathrm{id}$, $\psi_1 = \psi$, and every curve $t \mapsto \psi_t(q)$ is piecewise smooth. The space $\Lambda(M, \psi)$ is then homotopy equivalent to the free loop space ΛM via the continuous map

$$\Lambda M \to \Lambda(M, \psi), \qquad \gamma \mapsto \gamma_\psi, \tag{2.15}$$

where $\gamma_\psi(t) = \psi_t(\gamma(t))$ for all $t \in [0, 1]$. With these premises, the proof of Theorem 2.2 goes through for the functional $E : \Lambda(M, \psi) \to [0, \infty)$.

Theorem 2.20 (Grove). *On a closed Riemannian manifold of dimension at least 2, every isometry homotopic to the identity has a non-trivially invariant geodesic.* □

As for the multiplicity of isometry-invariant geodesics, the starting point is the following beautiful observation of Grove, which is a consequence of the Baire category theorem: if $\gamma : \mathbb{R} \looparrowright M$ is a non-trivially ψ-invariant geodesic that is not a periodic curve, then the closure of its image $\overline{\gamma(\mathbb{R})}$ contains uncountably many non-trivially ψ-invariant geodesics. Therefore, as far as multiplicity results are concerned, one can always assume that all the non-trivially ψ-invariant geodesics $\gamma \in \mathrm{crit}(E)$ are also closed geodesics, that is, $\gamma = \gamma(\cdot + p)$ for some period $p \geq 1$.

A non-trivially ψ-invariant closed geodesic $\gamma \in \mathrm{crit}(E)$ produces infinitely many critical circles of the energy functional E. Indeed, if $p \geq 1$ is a period of γ, then for each integer $m > 0$ its reparametrization $\gamma^{mp+1}(t) = \gamma((mp + 1)t)$ is also a critical point of E. The curves γ^{mp+1} play the role of the iterates in the theory of closed geodesics. In the early papers [**GT76, GT78**], Grove and Tanaka studied the behavior of critical point theory indices along sequences $m \mapsto \gamma^{mp+1}$, extending in particular Bott's iteration theory of the Morse indices and Gromoll-Meyer's theory of the local homology. As an outcome, they obtained the following extension of Theorem 2.4.

Theorem 2.21 (Grove-Tanaka). *On every closed simply connected Riemannian manifold (M, g) whose rational cohomology ring $H^*(M; \mathbb{Q})$ is non-monogenic, every isometry homotopic to the identity has infinitely many non-trivially invariant geodesics.* □

We already mentioned in Section 2.2 that the fact that $H^*(M; \mathbb{Q})$ is not a monogenic ring translates into the unboundedness of the rational Betti numbers of the free loop space ΛM, and therefore of the space $\Lambda(M, \psi)$ as well for all $\psi : M \to M$ homotopic to the identity. On the other hand, unlike in the case of closed geodesics, when the homology $H_*(\Lambda(M, \psi))$ is not sufficiently rich there may be isometries having only finitely many non-trivially invariant geodesics.

Example 2.22. Let S^2 be the unit sphere in $\mathbb{R}^3$ equipped with the restriction of the Euclidean Riemannian metric. Any rotation of S^2 of an angle $\theta \notin 2\pi\mathbb{Z}$ is an isometry homotopic to the identity, but has only one invariant geodesic: the closed geodesic orthogonal to the axis of the rotation. □

In a series of papers [**Maz14, Maz15, MM17**], the last one in collaboration with Macarini, we addressed the multiplicity of isometry-invariant geodesics on non simply connected closed Riemannian manifolds. Here, we only mention the following result, which is the outcome of [**Maz15, MM17**], and generalizes Bangert and Hingston's Theorem 2.6.

Theorem 2.23. *On a closed Riemannian manifold of dimension at least 2 with infinite abelian fundamental group, every isometry homotopic to the identity has infinitely many non-trivially invariant geodesics.*

Proof. We only outline the proof for the most difficult case, which was established in [**MM17**]: the one of closed Riemannian manifolds (M, g) with fundamental group $\pi_1(M) \cong \mathbb{Z}$. We denote by $\alpha \in \Lambda M$ a generator of $\pi_1(M)$, and by $D_m \subset \Lambda M$ the connected component containing the m-th iterate α^m. Notice that the evaluation map $\mathrm{ev} : \Lambda M \to M$, $\mathrm{ev}(\gamma) = \gamma(0)$ induces surjective homomorphisms of homotopy groups

$$\mathrm{ev}_* : \pi_*(D_m) \twoheadrightarrow \pi_*(M), \qquad \forall m \in \mathbb{N}. \tag{2.16}$$

Since we are assuming that the closed manifold M has dimension at least 2, it is not homotopy equivalent to a circle; since $\pi_1(M) \cong \pi_1(S^1) \cong \mathbb{Z}$ there must be a non-trivial homotopy group $\pi_d(M)$ in some degree $d \geq 2$. A topological argument due to Bangert and Hingston [**BH84**] implies that there exists $k \in \mathbb{N}$ and, for all $m \in \mathbb{N}$, a non-zero homotopy class $h_m \in \pi_{d-1}(D_{km})$ such that $\mathrm{ev}_*(h_m) = 0$. Notice that, if $d = 2$, the surjectivity of the homomorphisms (2.16) implies that $\pi_1(D_{km}) \neq \mathbb{Z}$.

Let $\psi : M \to M$ be an isometry homotopic to the identity, and $C_{km} \subset \Lambda(M, \psi)$ the connected component corresponding to D_{km} via the homotopy equivalence (2.15). Since we are looking for infinitely many non-trivially ψ-invariant geodesics, we only have to consider the case in which they are all closed geodesics (Theorem 2.20). For each $m \in \mathbb{N}$, we denote by $\zeta_m \in C_{km}$ a global minimizer of the energy functional $E|_{C_{km}}$. We only consider the worst case scenario, in which all the ζ_m are reparametrization of the same closed ψ-invariant geodesic. In order to detect the geodesics claimed by the theorem, we consider a non-zero homotopy class $h_m \in \pi_{d-1}(C_{km}, \zeta_m)$. In case $d = 2$, since we know that $\pi_1(C_{km}, \zeta_m) \ncong \mathbb{Z}$, we choose such a class h_m in such a way that it does not have a representative contained in the critical circle of ζ_m. In any case, the min-max value

$$c_m := \inf \big\{ c \geq 0 \ \big|\ h_m \in \mathrm{im}(\pi_{d-1}(C_{km}^{<c}, \zeta_m) \xrightarrow{\mathrm{incl}_*} \pi_{d-1}(C_{km}, \zeta_m)) \big\},$$

where $C_{km}^{<c} := E^{-1}[0, c) \cap C_{km}$, is a critical value of E strictly larger than $E(\zeta_m)$. The core of the proof consists in showing that, as m grows, the critical points in $\mathrm{crit}(E|_{C_{km}}) \cap E^{-1}(c_m)$ eventually correspond to new non-trivially ψ-invariant geodesics (that is, not the same geodesics already detected for lower values of m). This is carried out by an instability argument for the local homotopy of isometry-invariant geodesics, which is a harder version of the argument employed in the proof of Theorem 1.24. $\square$

We close the section and the chapter by pointing out a contact geometry generalization of the notion of isometry-invariant geodesics. We recall that a contact manifold is a $(2n+1)$-dimensional manifold Y equipped with a 1-form α, called the contact form, such that

$\alpha \wedge (\mathrm{d}\alpha)^n$ is a volume form. A diffeomorphism $\phi : Y \to Y$ is called a contactomorphism when it preserves the contact structure $\ker(\alpha)$, i.e. $\mathrm{d}\phi(z)\ker(\alpha_z) = \ker(\alpha_{\phi(z)})$. If ϕ also preserves the contact form α, i.e. $\phi^*\alpha = \alpha$, it is called a strict contactomorphism. On every contact manifold (Y, α), there is a distinguished vector field R, called the Reeb vector field, defined by the equations $\alpha(R) \equiv 1$ and $\mathrm{d}\alpha(R, \cdot) \equiv 0$. Its flow ϕ_α^t consist of strict contactomorphisms. More generally, any strict contactomorphism ϕ preserves the Reeb vector field, and in particular maps Reeb orbits to Reeb orbits. A Reeb orbit $\gamma : \mathbb{R} \to Y$, $\gamma(t) = \phi_\alpha^t(\gamma(0))$ is called **non-trivially ϕ-invariant** when $\phi \circ \gamma = \gamma(\tau + \cdot)$ for some $\tau \neq 0$.

If (M, g) is a Riemannian manifold, its unit tangent bundle

$$SM = \big\{(q, v) \in \mathrm{T}M \bigm| \|v\|_g = 1\big\}$$

equipped with the Liouville 1-form $\alpha_{(q,v)} = g(v, \mathrm{d}\pi(q, v)\cdot)$, where $\pi : SM \to M$ is the base projection, is a contact manifold. The associated Reeb vector field is precisely the geodesic vector field, and therefore the orbits of its flow are of the form $t \mapsto (\gamma(t), \dot\gamma(t))$, where γ is a unit-speed geodesic of (M, g). Any isometry ψ of (M, g) lifts to a strict contactomorphism

$$\widetilde{\psi} : SM \to SM, \qquad \widetilde{\psi}(q, v) = (\psi(q), \mathrm{d}\psi(q)v),$$

and a non-trivially ψ-invariant unit-speed geodesic γ lifts to a non-trivially $\widetilde{\psi}$-invariant Reeb orbit $(\gamma, \dot\gamma)$.

$\star$ **Open problems:** The celebrated Weinstein conjecture, which is widely open in dimension at least 5, states that every closed contact manifold has a closed Reeb orbit. An even harder problem, which we phrase as an open question, is whether on a closed contact manifold (Y, ψ), any strict contactomorphism ϕ that is isotopic to the identity (say through contactomorphisms) has a non-trivially ϕ-invariant Reeb orbit (the Weinstein conjecture corresponds to the special case $\phi = \mathrm{id}$).

Indeed, we can even ask a harder question by considering the notion of translated point, which was introduced by Sandon [**San10**] as a special case of the notion of leaf-wise intersection from symplectic geometry. If ϕ is a contactomorphism of a closed contact manifold (Y, α), a point $z = \gamma(0)$ on a Reeb orbit $\gamma : \mathbb{R} \to Y$ is called a **ϕ-translated point** when $\phi(\gamma(0)) = \gamma(\tau)$ for some translation $\tau \in \mathbb{R}$, and ϕ preserves the contact form α at z, i.e. $(\phi^*\alpha)_z = \alpha_z$. A conjecture due to Sandon [**San12**], and proved in certain settings [**San13, AFM15, MN18**], states that a generic contactomorphism $\phi : Y \to Y$ that is contact-isotopic to the identity has as many translated points as the total rank of the homology $H_*(Y)$. A slightly different open question, which is harder than the ones in the previous paragraph, is whether any contactomorphism that is contact isotopic to the identity has at least one translated point with non-zero translation.

CHAPTER 3

Besse and Zoll Reeb flows

This final chapter covers the very recent author's results on contact manifolds all of whose Reeb orbits are closed. This line of research originated from the joint result with Suhr [**MS18**] presented in Section 2.5, the initial goal being to extend Theorem 2.14 beyond the class of geodesic flows of Riemannian 2-spheres.

3.1. Basic properties of Besse contact manifolds

The setting of this chapter is the one of closed contact manifolds, which were already introduced at the end of the last chapter. Let (Y, α) be a closed contact manifold of dimension $2n+1$, with Reeb vector field $R = R_\alpha$ and associated Reeb flow $\phi_\alpha^t : Y \to Y$. The **action spectrum** $\sigma(Y, \alpha) \subset (0, \infty)$ is the set of periods of the closed Reeb orbits, i.e.

$$\sigma(Y, \alpha) = \big\{\tau > 0 \;\big|\; \mathrm{fix}(\phi_\alpha^\tau) \neq \varnothing\big\}.$$

Notice that, if some τ belongs to $\sigma(Y, \alpha)$, so do its multiples $k\tau$, with $k \in \mathbb{N}$.

Example 3.1. If (M, F) is a closed Finsler manifold, its unit tangent bundle $SM := \{(q, v) \in \mathrm{T}M \mid F(q, v) = 1\}$ equipped with the Liouville 1-form α is a contact manifold whose Reeb flow is the geodesic flow: its orbits are of the form $\phi_\alpha^t(q, v) = (\gamma(t), \dot\gamma(t))$, where $\gamma : \mathbb{R} \to M$ is the geodesic such that $\gamma(0) = q$ and $\dot\gamma(0) = v$. The action spectrum $\sigma(SM, \alpha)$ is precisely the length spectrum $\sigma(M, F)$, which is the set of multiples of the lengths of the closed geodesics of (M, F). □

A closed contact manifold (Y, α) (or just its contact form, or its Reeb flow) is called **Besse** when every orbit of its Reeb flow is closed, i.e. for all $z \in Y$ there exists a minimal $\tau_z > 0$ such that $\phi_\alpha^{\tau_z}(z) = z$. The terminology comes from the pseudonym "Arthur Besse" of a collective of mathematicians based in France and directed by Berger, who wrote among other things an influential monograph [**Bes78**] on Riemannian manifolds all of whose geodesics are closed (their unit tangent bundle thus being a Besse contact manifold). In the literature, Besse contact manifolds are sometimes called "almost regular" [**Tho76, Gei08**].

The minimal period τ_z, in general, may vary discontinuously with the point z. Nevertheless, when Y is connected, a rather subtle result due to Wadsley [**Wad75**] implies that there is a common multiple $\tau > 0$ for all the τ_z. In other words the whole Reeb flow is periodic:

$$\phi_\alpha^\tau = \mathrm{id}.$$

This is in contrast with a counterexample, due to Sullivan [**Sul76**], of a flow on a closed 5-dimensional manifold all of whose orbits are closed but without a global upper bound for their minimal periods.

What prevents Sullivan's counterexample to apply in the Reeb setting is that Reeb flows are geodesible: every Reeb orbit is a geodesic for any Riemannian metric g_0 on Y such that $g_0(R_\alpha, \cdot) = \alpha$. Assuming that (Y, α) is Besse with minimal common period τ, the same is true with respect to the averaged Riemannian metric

$$g := \frac{1}{\tau} \int_0^\tau (\phi_\alpha^t)^* g_0 \, \mathrm{d}t,$$

which indeed still satisfies $g(R_\lambda, \cdot) = \lambda$, while at the same time is invariant under the Reeb flow, i.e. $(\phi_\lambda^t)^* g = g$. We can employ this Riemannian metric to show that Y admits a stratification in Besse contact submanifolds. For each integer $k > 0$, since $\phi_\alpha^{\tau/k}$ is a g-isometry, its fixed-point set

$$Y_k := \mathrm{fix}(\phi_\alpha^{\tau/k})$$

is a closed submanifold of Y with tangent spaces

$$\mathrm{T}_z Y_k = \ker(\mathrm{d}\phi_\alpha^{\tau/k}(z) - \mathrm{id}), \tag{3.1}$$

see [**Kob95**, Th. 5.1]. For each $z \in Y_k$, the linearized map $\mathrm{d}\phi_\alpha^{\tau/k}(z)$ restricts to a symplectic endomorphism of the symplectic vector space $(\ker(\alpha_z), \mathrm{d}\alpha|_{\ker(\alpha_z)})$. Therefore, the eigenvalue

$$1 \in \sigma(\mathrm{d}\phi_\alpha^{\tau/k}(z)|_{\ker(\alpha_z)})$$

has even algebraic multiplicity. Since $\mathrm{d}\phi_\alpha^{\tau/k}(z)|_{\ker(\alpha_z)}$ is an k-th root of the identity, this algebraic multiplicity is equal to the geometric multiplicity $\dim(\mathrm{T}_z Y_k)$. This, together with the fact that

$$\mathrm{d}\phi_\alpha^{\tau/k}(z) R_\alpha(z) = R_\alpha(z),$$

proves that $\dim(\mathrm{T}_z K_\alpha)$ is odd, and thus (Y_k, α) is a closed contact submanifold of (Y, α).

Example 3.2. The ellipsoid

$$E(a_1, \ldots, a_n) = \left\{ z = (z_1, \ldots, z_n) \in \mathbb{C}^n \;\middle|\; \sum_{i=1}^n \frac{|z_i|^2}{a_i} = \frac{1}{\pi} \right\},$$

where $0 < a_1 \leq a_2 \leq \ldots \leq a_n$, comes equipped with a natural contact form, which we write in complex coordinates as

$$\alpha := \frac{i}{2} \sum_{j=1}^n \Big(z_j \, \mathrm{d}\overline{z}_j - \overline{z}_j \, \mathrm{d}z_j \Big).$$

The associated Reeb flow is the linear one

$$\phi_\alpha^t(z_1, ..., z_n) = (e^{i2\pi t/a_1} z_1, ..., e^{i2\pi t/a_n} z_n).$$

The action spectrum $\sigma(E(a_1,\dots,a_n))$ is the set of multiples of the parameters $a_1,...,a_n$. We readily see that $E(a_1,\dots,a_n)$ is Besse if and only if it is a rational ellipsoid, i.e. $a_h/a_j \in \mathbb{Q}$ for all $h,j=1,...,n$. Its strata are the sub-ellipsoids

$$E(a_1,\dots,a_n)_k = \big\{(z_1,\dots,z_n)\in E(a_1,\dots,a_n)\ \big|\ z_i=0 \text{ if } \tfrac{\mathrm{lcm}(a_1,\dots,a_n)}{ka_i}\not\in\mathbb{N}\big\}. \qquad \square$$

Within the Besse class, a distinguished subclass, which was first introduced and studied by Boothby and Wang [**BW58**], is given by the **Zoll** contact manifolds. They are those closed contact manifolds (Y,α) all of whose Reeb orbits are closed with the same minimal period $\tau>0$. Namely, $\phi_\alpha^\tau=\mathrm{id}$ and $\mathrm{fix}(\phi_\alpha^t)=\varnothing$ for all $t\in(0,\tau)$. The Reeb flow ϕ_α^t thus defines a free $\mathbb{R}/\tau\mathbb{Z}$ action on Y. The simplest examples of Zoll contact manifolds are the round contact spheres $E(1,...,1)$, and the unit tangent bundles of round Riemannian spheres. The terminology comes from a construction due to Zoll [**Zol03**] of a Riemannian 2-sphere of revolution of non-constant curvature and all of whose geodesics are closed with the same length (and thus having a unit tangent bundle that is a Zoll contact manifold).

$\star$ **Open problem:** Let (Y,α) be a closed contact manifold, and $K\subset Y$ be the subset of closed Reeb orbits. For each $z\in K$, we denote by $\tau_z>0$ the minimal period of the corresponding closed Reeb orbit, i.e.

$$z\in\mathrm{fix}(\phi_\alpha^{\tau_z})\setminus\bigcup_{t\in(0,\tau_z)}\mathrm{fix}(\phi_\alpha^t).$$

The **prime action spectrum** $\sigma_{\mathrm{p}}(Y,\alpha)$ is the set of minimal periods of the closed Reeb orbits, i.e.

$$\sigma_{\mathrm{p}}(Y,\alpha)=\big\{\tau_z\ \big|\ z\in K\big\}.$$

The Weinstein conjecture, if true, would guarantee that $\sigma_{\mathrm{p}}(Y,\alpha)$ is always non-empty. Theorem 2.14 prompted Cristofaro-Gardiner and the author to ask whether a closed contact manifold with a singleton prime action spectrum $\sigma_{\mathrm{p}}(Y,\alpha)=\{\tau\}$ is necessarily a Zoll contact manifold. In full generality, of course this question is widely open. Notice that, if the Weinstein conjecture hold, an affirmative answer to our question would imply that every closed contact manifold has at least two closed Reeb orbits, a statement that is currently open even for general geodesic flows (see Section 2.1). Over the next few sections, we shall address this question in several settings.

3.2. Spectral characterization of Besse contact three-manifolds

Several open problems in Reeb dynamics have been actually established for closed contact 3-manifolds. This is the case for the Weinstein conjecture (the existence of a closed Reeb orbit), which was proved by Taubes [**Tau07, Tau09**] a little over a decade ago by means of techniques based on holomorphic curves and Seiberg-Witten theory. As it turned out, the Weinstein conjecture is not sharp in dimension 3: a more recent result of Cristofaro-Gardiner and Hutchings [**CGH16**] asserts that every closed contact 3-manifolds has at least two closed Reeb orbits. By pushing further the arguments in the proof of this result, in a joint work with Cristofaro-Gardiner [**CGM20**] we could answer the open question mentioned at the end of the previous section, and indeed prove the following stronger result.

Theorem 3.3. *For any closed connected contact 3-manifold (Y, α), the following three conditions are equivalent.*

(i) *The contact manifold (Y, α) is Besse.*
(ii) *The closed orbits of the Reeb flow ϕ_α^t have some common period $\tau > 0$, i.e. $\phi_\alpha^\tau = \mathrm{id}$.*
(iii) *The action spectrum $\sigma(Y, \lambda)$ has rank 1, i.e. $\sigma(Y, \alpha) \subset r\mathbb{Z}$ for some $r > 0$.*

Here, the implication (i) $\Rightarrow$ (ii) is provided by the already mentioned Wadsley theorem [**Wad75**], and the implication (ii) $\Rightarrow$ (iii) is a straightforward consequence of the fact that the Reeb vector field is nowhere vanishing. Our novel contribution is the implication (iii) $\Rightarrow$ (i). We shall sketch its proof later on, but first let us discuss the important corollaries of Theorem 3.3. The following one is immediate, and provides the spectral characterization of Zoll contact 3-manifolds.

Corollary 3.4. *A closed contact 3-manifold (Y, α) is Zoll if and only if its prime action spectrum $\sigma_{\mathrm{p}}(Y, \alpha)$ is a singleton.* □

In Section 2.5 we already presented a (slightly stronger) version of this latter corollary for the special class of unit tangent bundles of reversible Finsler 2-spheres (Theorem 2.14). Theorem 3.3, combined with other results available in the Finsler and Riemannian literature, allows us to infer new spectral characterization results for Zoll Finsler surfaces (M, F), reversible or not. We recall that the length spectrum of a Finsler surface is the set $\sigma(M, F) := \sigma(SM, \alpha)$, where α is the standard contact form on the unit tangent bundle SM (Example 3.1)

Corollary 3.5. *Let (M, F) be a closed connected orientable Finsler surface. The length spectrum $\sigma(M, F)$ has rank 1 if and only if $M = S^2$ and F is Besse. Moreover, if F is reversible, the length spectrum $\sigma(M, F)$ has rank 1 if and only if $M = S^2$ and F is Zoll.*

Remark 3.6. The reversibility assumption in the second part of this statement is essential. Indeed, Katok's Finsler metrics F_α of Example 2.10 with a rational parameter α are Besse but not Zoll. □

Proof of Corollary 3.5. Theorem 3.3 guarantees that the length spectrum $\sigma(M, F)$ has rank 1 if and only if the Finsler surface (M, F) is Besse. A theorem due to Frauenfelder-Labrousse-Schlenk [**FLS15**], which extends the classical Bott-Samelson Theorem [**Bot54, Sam63**] from Riemannian geometry, implies that F can be Besse only if the fundamental group of M is finite and the integral cohomology ring of the universal cover of M is monogenic. The only closed orientable surface M with these properties is S^2. Finally, a Besse reversible Finsler metric on S^2 is Zoll according to a theorem of Frauenfelder-Lange-Suhr [**FLS16**], which generalizes the classical Riemannian result of Gromoll-Grove [**GG81**]. □

Corollary 3.7. *Let (M, F) be a closed connected non-orientable Finsler surface. The length spectrum $\sigma(M, F)$ has rank 1 if and only if $M = \mathbb{RP}^2$ and F is Besse. Moreover, if F is Riemannian, the length spectrum $\sigma(M, F)$ has rank 1 if and only if $M = \mathbb{RP}^2$ and F is the norm associated with a Riemannian metric with constant curvature (in particular, F is Zoll).*

Proof. Let M' be the orientation double cover of M, and $F' : \mathrm{T}M' \to [0, \infty)$ the lift of F. By Corollary 3.5, F' is Besse if and only if $\sigma(M', F')$ has rank 1 and $M' = S^2$. Notice that $M' = S^2$ if and only if $M = \mathbb{RP}^2$. The length spectra satisfy $\sigma(M', F') \subseteq \sigma(M, F)$ and $2\sigma(M, F) \subseteq \sigma(M', F')$; in particular, $\sigma(M', F')$ has rank 1 if and only if the same is true for $\sigma(M, F)$. Moreover, F' is Besse if and only if the same if true for F. This proves the first part of the statement. Finally, a Riemannian metric on $\mathbb{RP}^2$ is Besse if and only if it has constant curvature, according to a theorem of Pries [**Pri09**]. □

We now briefly outline the techniques and the arguments needed for the proof of Theorem 3.3; the reader that is only interested in the applications can skip the rest of the section. The closed Reeb orbits of a closed contact manifold (Y, α) can be characterized variationally in several ways. The conceptually simplest way employs the contact action functional

$$\mathcal{A}_\alpha(\gamma) = \int_\gamma \alpha,$$

where $\gamma : \mathbb{R}/\tau\mathbb{Z} \to Y$ is a smooth periodic curve. If we fix the period $\tau > 0$ and define $\mathcal{A}_\alpha$ over the space of immersed curves $\mathrm{Imm}(\mathbb{R}/\tau\mathbb{Z}, Y)$, its critical points are precisely those γ whose derivative $\dot\gamma$ is tangent to the Reeb vector field. In particular, the critical points γ with positive critical value $\mathcal{A}_\alpha(\gamma) > 0$ are the orientation preserving reparametrizations of closed Reeb orbits of period $\mathcal{A}_\alpha(\gamma)$. Unfortunately, $\mathcal{A}_\alpha$ is far from suitable for classical variational methods: beside being invariant under the action of the huge group of orientation-preserving circle diffeomorphisms, its critical points have infinite Morse index, and the sublevel sets are far from being compact in any possible sense (cf. Section 1.2). Nevertheless, when Y is 3-dimensional, there is an available version of Morse theory for the contact action, called **embedded contact homology** and developed by Hutchings. Such a theory is based on holomorphic curves techniques in symplectic geometry, an approach pioneered by Gromov [**Gro85a**] and that paved the way for the further development of Floer homology [**Sal99, AD14**], which is a version of Morse theory for rather general Hamiltonian systems. It is impossible to give an even remotely complete account of such a sophisticated theory within a short section; instead, we shall limit the exposition to a brief description of those properties of embedded contact homology that are needed for our applications. We refer the reader to the survey [**Hut14**] for a more detailed account, and for a guide to the vast literature on the subject.

Let Y be a closed 3-manifold. Its embedded contact homology is a $\mathbb{Z}_2$-vector space $\mathrm{ECH}(Y)$, equipped with a homomorphism $U : \mathrm{ECH}(Y) \to \mathrm{ECH}(Y)$. Both objects are topological invariants of Y. As the name suggests, $\mathrm{ECH}(Y)$ is the homology of a chain complex, and U is the homology homomorphism induced by a chain map. At chain level, the constructions require auxiliary choices:

- We fix a contact form α on Y that is non-degenerate: for all $\tau > 0$ and $z \in \mathrm{fix}(\phi_\alpha^\tau)$, the linearized return map $\mathrm{d}\phi_\alpha^\tau(z)|_{\ker(\alpha_z)}$ does not have the eigenvalue 1; it is well known that a C^∞ generic contact form α is non-degenerate.
- We consider the symplectization of our contact 3-manifold, which is the symplectic manifold $(Y \times \mathbb{R}, \mathrm{d}(e^s\alpha))$, where s denotes the variable on the real line $\mathbb{R}$. We equip $Y \times \mathbb{R}$ with a tame almost complex structure J, which is an endomorphism of $\mathrm{T}(Y \times \mathbb{R})$ such that $J^2 = -\mathrm{id}$, $JR_\alpha = \partial_s$, $J\ker(\alpha) = \ker(\alpha)$, and $\mathrm{d}\alpha(v, Jv) > 0$ for each non-zero $v \in \ker(\alpha)$. Such a J should be chosen in a suitably generic way in order to satisfy certain technical assumptions.
- We fix a point $z \in Y$.

The chain complex $\mathrm{ECC}(Y, \alpha, J)$ that gives the embedded contact homology is the $\mathbb{Z}_2$-vector space freely generated by finite collections

$$\boldsymbol{\gamma} = \{(m_i, \gamma_i) \mid i = 1, ..., k\},$$

where $k \in \mathbb{N}$, the γ_i are distinct simple closed orbits of the Reeb flow ϕ_α^t, and m_i is a positive integer. Here, by "simple" we mean that the closed Reeb orbits γ_i are viewed as maps of the form $\gamma_i : \mathbb{R}/\tau_i\mathbb{Z} \to Y$, where $\tau_i > 0$ is the minimal period of γ_i. Two simple closed Reeb orbits γ_i, γ_j are distinct if they are not of the form $\gamma_i = \gamma_j(\cdot + s)$ for any $s > 0$. One should think of $\boldsymbol{\gamma}$ as to a multicurve whose components are the iterated closed Reeb orbits $\gamma_i^{m_i}$, for $i = 1, ..., k$. Its contact action is given by

$$\mathcal{A}_\alpha(\boldsymbol{\gamma}) = \sum_{i=1}^{k} m_i \mathcal{A}_\alpha(\gamma_i) = \sum_{i=1}^{k} m_i \int_{\gamma_i} \alpha.$$

At chain level, the crucial property of the homomorphism

$$U : \mathrm{ECC}(Y, \alpha, J) \to \mathrm{ECC}(Y, \alpha, J)$$

is related to holomorphic curves in the symplectization. These are smooth maps

$$u : \Sigma \setminus \{p_1^+, ..., p_a^+, p_1^-, ..., p_b^-\} \to Y \times \mathbb{R},$$

where (Σ, j) is a closed Riemann surface equipped with its complex structure, satisfying the Cauchy-Riemann equation $J \circ \mathrm{d}u = \mathrm{d}u \circ j$. We require u to be a proper map, asymptotic towards each puncture $p_i^\pm$ to a closed Reeb orbit $\gamma_i^\pm$ in the end $Y \times \{\pm\infty\}$ of the symplectization (Figure 3.1). We express this by saying that u is positively asymptotic to the multicurve with components $\gamma_1^+, ..., \gamma_a^+$, and negatively asymptotic to the multicurve with components $\gamma_1^-, ..., \gamma_b^-$. The tameness condition of the almost complex structure implies

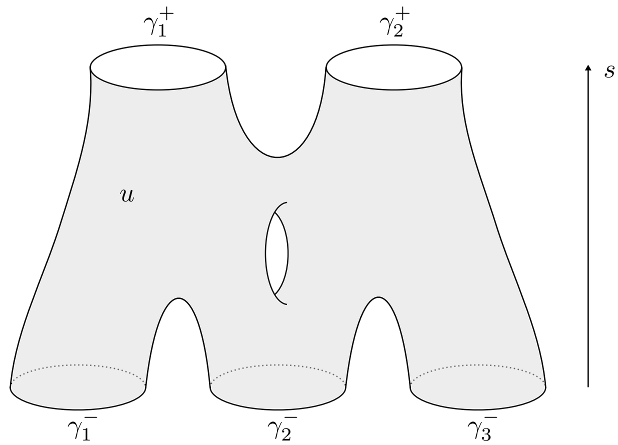


FIGURE 3.1. A J-holomorphic curve $u : \Sigma \setminus \{p_1^+, p_2^+, p_1^-, p_2^-, p_3^-\} \to Y \times \mathbb{R}$ positively asymptotic to γ_1^+, γ_2^+ and negatively asymptotic to $\gamma_1^-, \gamma_2^-, \gamma_3^-$.

that $\mathrm{d}\alpha$ is non-negative on the image of u, and by Stokes theorem

$$0 \le \int_{\Sigma_0} u^*\mathrm{d}\alpha = \mathcal{A}_\alpha(\gamma_1^+) + ... + \mathcal{A}_\alpha(\gamma_a^+) - \mathcal{A}_\alpha(\gamma_1^-) - ... - \mathcal{A}_\alpha(\gamma_b^+),$$

where $\Sigma_0 = \Sigma \setminus \{p_1^+, ..., p_a^+, p_1^-, ..., p_b^-\}$. We can now state the property of the homomorphism U at chain level: if

$$U(\boldsymbol{\gamma}^+) = \boldsymbol{\gamma}_1^- + ... + \boldsymbol{\gamma}_r^-,$$

for some pairwise distinct multicurves $\boldsymbol{\gamma}_i^-$, then for each $i = 1, ..., r$ there exists a holomorphic curve u as above positively and negatively asymptotic to the multicurves $\boldsymbol{\gamma}^+$ and $\boldsymbol{\gamma}_i^-$ respectively, and the image of u contains the point $(z, 0) \in Y \times \{0\}$, where z is the point that we fixed before.

Any non-zero embedded contact homology class $h \in \mathrm{ECH}(Y)$ produces a spectral invariant

$$c_\alpha(h) := \inf\{\mathcal{A}_\alpha(\sigma) \mid h = [\sigma]\}.$$

Here, the contact action of a cycle $\sigma \in \mathrm{ECC}(Y, \alpha, J)$ is defined as follows: if $\sigma = \boldsymbol{\gamma}_1 + ... + \boldsymbol{\gamma}_r$ for some pairwise distinct multicurves $\boldsymbol{\gamma}_i$, then

$$\mathcal{A}_\alpha(\sigma) = \max_{i=1,...,r} \mathcal{A}_\alpha(\boldsymbol{\gamma}_i).$$

Notice that $c_\alpha(h)$ is not an element of the action spectrum $\sigma(Y, \alpha)$, but rather the sum of finitely many elements of $\sigma(Y, \alpha)$. These spectral invariants turn out to be C^0 continuous with respect to the contact form α. This allows to extend their construction to contact forms α on Y that are possibly degenerate (as are the Besse ones): it suffices to consider a sequence of non-degenerate contact forms $\alpha_n = e^{f_n}\alpha$, where $f_n : Y \to \mathbb{R}$ are continuous functions C^0-converging to zero, and set

$$c_\alpha(h) = \lim_{n\to\infty} c_{\alpha_n}(h).$$

It is still true that $c_\alpha(h)$ is the sum of finitely many elements of the action spectrum $\sigma(Y,\alpha)$.

One last ingredient that we need before presenting the proof of Theorem 3.3 is the following statement due to Cristofaro-Gariner and Hutchings [**CGH16**].

Lemma 3.8. *For each contact distribution $\xi \subset \mathrm{T}Y$, there exists an infinite sequence of non-zero embedded contact homology classes $h_k \in \mathrm{ECH}(Y)$ such that $Uh_{k+1} = h_k$ and $c_\alpha(h_k)/k \to 0$ as $k \to \infty$ for each contact form α such that $\ker(\alpha) = \xi$.* □

Proof of Theorem 3.3. We only need to prove the implication (iii) ⇒ (i). Let us assume by contradiction that the closed connected contact 3-manifold (Y,α) is not Besse, but its action spectrum satisfies $\sigma(Y,\alpha) \subset r\mathbb{Z}$ for some $r > 0$. We choose the sequence of embedded contact homology classes $h_k \in \mathrm{ECH}(Y)$ provided by Lemma 3.8. Notice that we cannot have $c_\alpha(h_{k+1}) > c_\alpha(h_k)$ for all k, for the assumption on the action spectrum would imply that $c_\alpha(h_{k+1}) > c_\alpha(h_k) + r$, contradicting the fact that $c_\alpha(h_k)/k \to 0$ as $k \to \infty$. In particular, there exists k such that $c_\alpha(h_{k+1}) = c_\alpha(h_k)$.

Since (Y,α) is not Besse, we can fix a point $z \in Y$ that does not lie on a closed Reeb orbit; we use such a z as the selected point for the homomorphism U at chain level. Let $B \subset Y$ be an embedded compact ball of codimension 1 containing z in its interior and such that $\mathrm{T}_z B = \ker(\alpha_z)$. We fix a constant $\tau > c_\alpha(h_{k+1})$. Up to shrinking B around z, the map

$$\psi : [-\tau/2, \tau/2] \times B \to Y, \qquad \psi(t,w) = \phi_\alpha^t(w)$$

is a diffeomorphism onto its image $K := \psi([-\tau/2,\tau/2] \times B)$. Namely, K is a flow box for the Reeb flow ϕ_α^t containing orbits of length τ.

The contact form α may be degenerate. However, for each integer $n > 0$ we can find a smooth function $b_n : Y \to \mathbb{R}$ such that $\|b_n\|_{L^\infty} < \frac{1}{n}$ and $\alpha_n = e^{b_n}\alpha$ is a non-degenerate contact form. Since the flow box K does not contain closed orbits, we can even find such functions b_n so that $b_n|_K \equiv 0$. This gives us a sequence of non-degenerate contact forms α_n such that $\ker(\alpha_n) = \ker(\alpha)$, $\alpha_n|_K \equiv \alpha$, and $\alpha_n \to \alpha$ as $n \to \infty$. We choose an almost complex structure J tame with respect to α, and a sequence of almost complex structures J_n that are tame with respect to the corresponding α_n, are sufficiently generic to satisfy the technical requirements of embedded contact homology, and satisfy $J_n|_{K\times\mathbb{R}} \equiv J|_{K\times\mathbb{R}}$.

We consider an arbitrary representative $\sigma_n = \boldsymbol{\gamma}_{n,1} + ... + \boldsymbol{\gamma}_{n,a} \in \mathrm{ECC}(Y,\alpha_n)$ of the embedded contact homology class $h_{k+1} = [\sigma_n]$ such that $\mathcal{A}_{\alpha_n}(\sigma_n) \to c_\alpha(h_{k+1})$ as $n \to \infty$. In order to produce a contradiction and thus complete the proof it is enough to show that, for some $\delta > 0$, the following holds: if

$$U\boldsymbol{\gamma}_{n,i} = \boldsymbol{\zeta}_{n,1} + ... + \boldsymbol{\zeta}_{n,b}$$

for some pairwise distinct multicurves $\boldsymbol{\zeta}_{n,h}$, then

$$\mathcal{A}_{\alpha_n}(\boldsymbol{\gamma}_{n,i}) - \mathcal{A}_{\alpha_n}(\boldsymbol{\zeta}_{n,h}) \geq \delta, \qquad \forall h = 1, ..., b.$$

Indeed, this would readily imply $c_\alpha(h_{k+1}) > c_\alpha(h_k)+\delta$, contradicting the equality $c_\alpha(h_{k+1}) = c_\alpha(h_k)$. Let us assume that such a constant $\delta > 0$ does not exists, so that we can find multicurves $\boldsymbol{\gamma}_{n,i_n}$ and $\boldsymbol{\zeta}_{n,h_n}$ as above such that, up to extracting a subsequence,

$$\lim_{n\to\infty}\Big(\mathcal{A}_{\alpha_n}(\boldsymbol{\gamma}_{n,i_n}) - \mathcal{A}_{\alpha_n}(\boldsymbol{\zeta}_{n,h_n})\Big) = 0. \tag{3.2}$$

Since $\boldsymbol{\gamma}_{n,i_n}$ and $\boldsymbol{\zeta}_{n,h_n}$ are related by the chain homomorphism U, there exist J_n-holomorphic curves $u_n : \Sigma_n \to Y \times \mathbb{R}$ positively and negatively asymptotic to the multicurves $\boldsymbol{\gamma}_{n,i_n}$ and $\boldsymbol{\zeta}_{n,h_n}$ respectively, and the image of u contains the point $(z, 0) \in Y \times \{0\}$. We set $C_n := u_n(\Sigma_n)$, and from now on we will not distinguish between the map u_n and its image C_n. Notice that, by Stokes theorem,

$$\int_{C_n} \mathrm{d}\alpha_n = \mathcal{A}_{\alpha_n}(\boldsymbol{\gamma}_{n,i_n}) - \mathcal{A}_{\alpha_n}(\boldsymbol{\zeta}_{n,h_n}), \tag{3.3}$$

and in particular this quantity is uniformly bounded in n. Since $J_n \equiv J$ on $K \times \mathbb{R}$, the intersections $C_n \cap (K \times [-1, 1])$ are J-holomorphic curves. Since $\mathrm{d}\alpha_n = \mathrm{d}\alpha$ is non-negative on $C_n \cap (K \times [-1, 1])$, Equations (3.2) and (3.3) imply that

$$\lim_{n\to\infty} \int_{C_n\cap(K\times[-1,1])} \mathrm{d}\alpha = 0. \tag{3.4}$$

Let $s_0 \in [-2, -1]$ and $s_1 \in [1, 2]$ be such that u_n is transverse to $Y \times \{s_0, s_1\}$. Since both $\mathrm{d}(e^s\alpha_n)$ and $\mathrm{d}\alpha_n$ are non-negative on C_n, we have the uniform bound

$$\begin{aligned}
\int_{C_n\cap(K\times[-1,1])} \mathrm{d}(e^s\alpha) &\leq \int_{C_n\cap(Y\times[s_0,s_1])} \mathrm{d}(e^s\alpha_n)\\
&= e^{s_1}\int_{C_n\cap(Y\times\{s_1\})} \alpha_n - e^{s_0}\int_{C_n\cap(Y\times\{s_0\})} \alpha_n\\
&\leq e^2\bigg(\int_{C_n\cap(Y\times\{s_1\})} \alpha_n + \int_{C_n\cap(Y\times[s_1,\infty))} \mathrm{d}\alpha_n\bigg)\\
&= e^2\mathcal{A}_{\alpha_n}(\boldsymbol{\gamma}_{n,i_n}) \leq e^2 c_\sigma(Y, \alpha) + 1
\end{aligned}$$

for all $n \in \mathbb{N}$ large enough. We now invoke a compactness result due to Taubes [**Tau98**, Prop. 3.3] and infer that, up to extracting a subsequence, $C_n \cap (K \times [-1, 1])$ converges in the sense of currents to a compact J-holomorphic curve $C \subset K \times [-1, 1]$ with boundary in $\partial(K \times [-1, 1])$, and $(0, z) \in C$. Equation (3.4) implies

$$\int_C \mathrm{d}\alpha = 0,$$

and therefore C must have a component of the form $\phi_\alpha^{[-\tau/2,\tau/2]}(z) \times [-1, 1]$. In particular

$$\int_{C\cap(K\times\{s\})} \alpha \geq \tau, \qquad \forall s \in [-1, 1].$$

We fix an arbitrary $\tau' \in (c_\alpha(h_{k+1}), \tau)$. For each $n \in \mathbb{N}$, we choose a point $s_n \in [-1, 1]$ such that u_n is transverse to $Y \times \{s_n\}$. With the suitable convention on the orientation of intersections, the contact form α_n is non-negative along the oriented 1-manifold $C_n \cap (Y \times \{s_n\})$. Therefore, since $C_n \cap (K \times [-1, 1]) \to C$ in the sense of currents, up to removing sufficiently many elements from the sequence $\{C_n \mid n \in \mathbb{N}\}$ we have

$$\int_{C_n \cap (Y \times \{s_n\})} \alpha_n \geq \int_{C_n \cap (K \times \{s_n\})} \alpha_n \geq \tau', \qquad \forall n \in \mathbb{N}.$$

However, if we choose n large enough so that $\mathcal{A}_{\alpha_n}(\gamma_{n,i_n}) < \tau'$, we have

$$\int_{C_n \cap (Y \times \{s_n\})} \alpha_n \leq \int_{C_n \cap (Y \times \{s_n\})} \alpha_n + \int_{C_n \cap (Y \times [s_n, \infty))} \mathrm{d}\alpha_n = \mathcal{A}_{\alpha_n}(\gamma_{n,i_n}) < \tau',$$

which gives a contradiction. □

⋆ **Open problem:** Does Theorem 3.3 hold for higher dimensional contact manifolds, or at least for suitable classes of higher dimensional contact manifolds? This question seems unfortunately out of reach at the moment of this writing. Indeed, there is not even a proof of Theorem 3.3 in the special case of Riemannian or Finsler geodesics flows on the 2-sphere that is purely based on the classical Morse theory of the energy or length functionals. Corollaries 3.5 and 3.7 are rare examples of results on closed geodesics whose only available proof needs holomorphic curves techniques.

3.3. Spectral characterization of Besse convex contact spheres.

While Theorem 3.3 seems currently out of reach in dimension larger than 3, one can hope to provide suitable characterizations of the Besse condition within classes of contact manifolds for which there are available tools to detect closed Reeb orbits. Unlike in Theorems 2.14 and 3.3, the knowledge of the action spectrum alone will not suffice to determine whether the Besse property holds or not. One will often need a "marking" of the action spectrum, provided by suitable spectral invariants.

One of the first class of contact manifolds for which a variational theory of closed Reeb orbits was developed is the one of convex contact sphere, on which we will devote the current section. A **convex contact sphere** is a positively curved hypersurface $Y \subset \mathbb{R}^{2n}$ that bounds a convex set containing the origin. We will equip Y with the contact form $\alpha = \lambda|_Y$, where λ is the 1-form on $\mathbb{R}^{2n}$ given by

$$\lambda = \frac{1}{2} \sum_{i=1}^{n} \Big(x_i \, \mathrm{d}y_i - y_i \, \mathrm{d}x_i \Big). \tag{3.5}$$

Here, as usual, we are denoting by $x_1, y_1, ..., x_n, y_n$ the symplectic coordinates of $\mathbb{R}^{2n}$. The 1-form λ can also be expressed as $\lambda_z = \frac{1}{2}\langle Jz, \cdot \rangle$, where J is the complex structure of $\mathbb{R}^{2n}$ that we can write in matrix form as

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}.$$

The exterior differential of λ is the standard symplectic form $\omega = \mathrm{d}\lambda = \langle J\cdot, \cdot \rangle$ of $\mathbb{R}^{2n}$. Among the contact convex spheres, there are the ellipsoids of Example 3.2.

A convenient variational principle for the study of the closed Reeb orbits of the convex contact sphere Y is the one associated with the Clarke action functional [**Cla79**]. Such a functional appears in the literature under different, although equivalent, formulations; here we present the one in the L^2 setting, following Ekeland-Hofer [**EH87**]. We first embed the Reeb flow $\phi_\alpha^t : Y \to Y$ into a Hamiltonian flow as follows. We consider the Hamiltonian $H : \mathbb{R}^{2n} \to \mathbb{R}$ defined by $H|_Y \equiv 1$ and $H(\lambda z) = \lambda^2$ for all $\lambda > 0$ and $z \in Y$. As usual, we denote the associated Hamiltonian vector field by $X_H := J\nabla H$, and its flow by $\phi_H^t : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$. Such a flow satisfies

$$\phi_H^t(\lambda z) = \lambda \phi_\alpha^t(z), \qquad \forall t \in \mathbb{R},\ \lambda > 0,\ z \in Y.$$

Since H is a convex 2-homogeneous function smooth outside the origin, so is its dual

$$G : \mathbb{R}^{2n} \to [0, \infty), \qquad G(w) = \max_z \Big(\langle w, z \rangle - H(z) \Big).$$

We consider the Hilbert space

$$L_0^2(S^1, \mathbb{R}^{2n}) = \left\{ \zeta \in L^2(S^1, \mathbb{R}^{2n}) \ \middle|\ \int_{S^1} \zeta(t)\, \mathrm{d}t = 0 \right\},$$

where $S^1 = \mathbb{R}/\mathbb{Z}$. Notice that every element in $L_0^2(S^1, \mathbb{R}^{2n})$ is the first derivative $\dot\gamma$ of some $\gamma \in W^{1,2}(S^1, \mathbb{R}^{2n})$. We consider the symplectic action functional

$$\mathcal{A} : L_0^2(S^1, \mathbb{R}^{2n}) \to \mathbb{R}, \qquad \mathcal{A}(\dot\gamma) = \int_{S^1} \gamma^* \lambda = \frac{1}{2} \int_{S^1} \langle J\gamma(t), \dot\gamma(t) \rangle\, \mathrm{d}t.$$

Notice that the expression of $\mathcal{A}$ involves a primitive γ, but the value $\mathcal{A}(\dot\gamma)$ is independent of its choice. We next consider the functional

$$\mathcal{G} : L_0^2(S^1, \mathbb{R}^{2n}) \to [0, \infty), \qquad \mathcal{G}(\dot\gamma) = \int_{S^1} G(-J\dot\gamma(t))\, \mathrm{d}t.$$

Notice that $\mathcal{G}(0) = 0$, $\mathcal{G}$ is positive away from the origin, and $\mathcal{G}(\lambda\dot\gamma) = \lambda^2 \mathcal{G}(\dot\gamma)$ for all $\lambda > 0$ and $\dot\gamma \in L_0^2(S^1, \mathbb{R}^{2n})$. We set

$$\Lambda := \mathcal{G}^{-1}(1) \cap \mathcal{A}^{-1}(0, \infty).$$

The **Clarke action functional** is defined by

$$\Psi : \Lambda \to (0, \infty), \qquad \Psi(\dot\gamma) = \frac{1}{\mathcal{A}(\dot\gamma)}.$$

The critical points $\dot\gamma \in \mathrm{crit}(\Psi) \cap \Psi^{-1}(c)$ are precisely those curves admitting a (unique) smooth primitive $\gamma : S^1 \to H^{-1}(c^{-2})$ such that $\dot\gamma = cX_H(\gamma)$. Namely, the curve $t \mapsto c\,\gamma(t/c)$ is a c-periodic Reeb orbit of Y. In particular, the action spectrum $\sigma(Y) = \sigma(Y, \alpha)$ coincides with the set of critical values of Ψ.

The Clarke action functional Ψ satisfies all the desirable properties required by critical point theory (cf. Section 1.2):

• **(Suitable domain)** The domain Λ is an open subset of $\mathcal{G}^{-1}(1)$, which is a $C^{1,1}$-hypersurface in the Hilbert space $L_0^2(S^1; \mathbb{R}^{2n})$ homeomorphic to its unit sphere via the radial homeomorphism $\dot\gamma \mapsto \dot\gamma / \|\dot\gamma\|_{L^2}$.

• **(Regularity)** The Clarke action functional Ψ is $C^{1,1}$, and suitable finite dimensional reduction techniques allow to treat it as a C^∞ functional.

• **(Compactness of the sublevel sets)** While the domain Λ is not complete as a Riemannian submanifold of $L^2_0(S^1;\mathbb{R}^{2n})$, the functional Ψ is bounded from below, and the closed sublevel sets

$$\Lambda^{\leq b} := \Psi^{-1}(0,b], \qquad b>0,$$

are complete. Moreover, Ψ satisfy the Palais-Smale condition: every sequence $\dot\gamma_k$ in a sublevel set $\Lambda^{\leq b}$ and such that $\|\nabla\Psi(\dot\gamma_k)\|_{L^2}\to 0$ admits a converging subsequence.

• **(Finite Morse indices)** Every critical point $\dot\gamma\in\mathrm{crit}(\Psi)$ has finite Morse indices

$$\mathrm{ind}(\dot\gamma) := \sum_{\lambda<0}\dim\ker(\nabla^2\Psi(\dot\gamma)-\lambda I)<\infty,$$
$$\mathrm{nul}(\dot\gamma) := \dim\ker(\nabla^2\Psi(\dot\gamma))<\infty.$$

The existence of a closed Reeb orbit in the convex contact sphere Y is a straighforward consequence of the properties of the Clarke action functional Ψ: the global minimum

$$c_0(Y) := \min\Psi \tag{3.6}$$

is an element of the action spectrum $\sigma(Y)$, and indeed is the minimum of $\sigma(Y)$. In order to detect other elements of $\sigma(Y)$, one needs to employ Morse theory. It is not hard to show that the domain Λ is contractible, which means that plain Morse theory is not able to detect critical values of Ψ other than the global minimum $c_0(Y)$. Nevertheless, the topology of Λ becomes non-trivial when considered together with the S^1 action

$$t\cdot\dot\gamma=\dot\gamma(t+\cdot), \qquad t\in S^1,\ \dot\gamma\in\Lambda.$$

The Clarke action functional being S^1-invariant, one may detect its critical values by means of equivariant Morse theory.

The tool to handle the topology of the S^1-space Λ is equivariant cohomology, which we briefly recall here (for a detailed account and more references, we refer the reader to Atiyah and Bott [**AB83**]). We consider the universal S^1-bundle $ES^1\to BS^1$, which is any principal S^1-bundle whose total space ES^1 is contractible. Such a bundle can be realized by setting ES^1 to be the unit infinite-dimensional sphere $S^\infty\subset\mathbb{C}^\infty$, equipped with the diagonal action of the unit circle $S^1\subset\mathbb{C}$. With this choice, the base space $BS^1=ES^1/S^1$ is the infinite-dimensional complex projective space $\mathbb{CP}^\infty$. The integral cohomology ring of BS^1 is a polynomial ring $H^*(BS^1)=\mathbb{Z}[e]$ in the variable $e\in H^2(BS^1)$, which is the Euler class of the universal bundle $ES^1\to BS^1$. Now, given a space X equipped with an S^1 action, one obtains a principal S^1-bundle

$$\pi: X\times ES^1\to X\times_{S^1}ES^1, \tag{3.7}$$

whose base space $X\times_{S^1}ES^1$ is the quotient of the total space $X\times ES^1$ under the diagonal S^1-action

$$t\cdot(x,y)=(t\cdot x,t\cdot y), \qquad t\in S^1,\ (x,y)\in X\times ES^1.$$

The Euler class of the principal S^1-bundle (3.7) is $\mathrm{pr}_2^* e$, where $\mathrm{pr}_2 : X \times_{S^1} ES^1 \to BS^1$ is the projection onto the second factor. The equivariant cohomology of X is defined as

$$H^*_{S^1}(X) := H^*(X \times_{S^1} ES^1).$$

Beyond the usual properties, $H^*_{S^1}(X)$ is also a $H^*(BS^1)$-module with the scalar multiplication $f \cdot k := (\mathrm{pr}_2^* f) \smile k$, for $f \in H^*(BS^1)$ and $k \in H^*_{S^1}(X)$. With a common abuse of notation, in the following we shall omit the homomorphism pr_2^*, and simply write e for $e \cdot 1 \in H^2_{S^1}(X)$. The equivariant cohomology is related to the usual cohomology by means of the Gysin long exact sequence

$$... \longrightarrow H^{*+1}(X) \longrightarrow H^*_{S^1}(X) \xrightarrow{\smile e} H^{*+2}_{S^1}(X) \longrightarrow H^{*+2}(X) \longrightarrow ...$$

All the properties mentioned in this paragraph hold for any coefficient ring. With rational coefficients, if the S^1 action on X is without fixed points, the equivariant cohomology is simply the usual cohomology of the quotient $H^*_{S^1}(X;\mathbb{Q}) \cong H^*(X/S^1;\mathbb{Q})$. A "measure" of the cohomological non-triviality of X is given by the **Fadell-Rabinowitz index** [**FR78**]

$$\mathrm{ind}_{\mathrm{FR}}(X) := \begin{cases} -1, & \text{if } X = \varnothing, \\ \sup\big\{ i \ \big|\ e^i \neq 0 \text{ in } H^*_{S^1}(X;\mathbb{Q}) \big\}, & \text{if } X \neq \varnothing, \end{cases}$$

which is usually employed instead of the cup-length in an S^1-equivariant setting.

The domain Λ can be showed to be S^1-equivariantly homotopy equivalent to the unit sphere of the infinite dimensional separable Hilbert space. The contractibility of such sphere, together with the Gysin sequence, implies that $H^*_{S^1}(\Lambda;\mathbb{Q}) \cong \mathbb{Q}[e]$, where $e \in H^2_{S^1}(\Lambda;\mathbb{Q})$ is the Euler class. For each integer $i \geq 0$, the i-th **Ekeland-Hofer spectral invariant** is defined by

$$c_i(Y) := \inf\big\{ c > 0 \ \big|\ \mathrm{ind}_{\mathrm{FR}}(\Lambda^{\leq c}) \geq i \big\}.$$

Equivariant Morse theory guarantees that $c_i(Y)$ is a critical value of Ψ, that is, an element of the action spectrum $\sigma(Y)$. The notation $c_i(Y)$ is consistent with the above one in Equation (3.6): $c_0(Y)$ is the global minimum of Ψ.

Example 3.9. The Ekeland-Hofer spectral invariants can be computed by hands for the ellipsoids $E(a_1, ..., a_n)$ (Example 3.2). The action spectrum $\sigma(E(a_1, ..., a_n))$ is the set of multiples of the parameters a_j. Let $\sigma_1 < \sigma_2 < \sigma_3 < \dots$ be the elements of $\sigma(E(a_1, ..., a_n))$ listed in increasing order. For each $i \geq 0$, we denote by d_i the number of a_j's such that $\sigma_i / a_j \in \mathbb{N}$. These numbers are indeed related to the dimensions of the sub-ellipsoids $\mathrm{fix}(\phi_\alpha^{\sigma_i}) \subset E$, i.e.

$$d_i := (\dim(\mathrm{fix}(\phi_\alpha^{\sigma_i})) + 1)/2.$$

Here, ϕ_α^t denotes as usual the Reeb flow on $E(a_1, ..., a_n)$. The Ekeland-Hofer spectral invariant $c_i(Y)$ is the $(i+1)$-th element in the sequence

$$\underbrace{\sigma_1, \dots, \sigma_1}_{\times d_1}, \underbrace{\sigma_2, \dots, \sigma_2}_{\times d_2}, \underbrace{\sigma_3, \dots, \sigma_3}_{\times d_3}, \dots$$

□

The definition of the Ekeland-Hofer spectral invariants readily implies that $c_i(Y) \leq c_{i+1}(Y)$, and if the equality holds then the critical set $\mathrm{crit}(\Psi) \cap \Psi^{-1}(c_i(Y))$ contains infinitely critical circles, that is, there are infinitely many closed Reeb orbits of period $c_i(Y)$. In a joint work with Ginzburg and Gurel [**GGM19**] we showed that the equality of suitable spectral invariants characterize the Besse and Zoll properties of convex contact spheres.

Theorem 3.10. *Let Y be a convex contact sphere of dimension $2n-1$. Its Ekeland-Hofer spectral invariants satisfy $c_i(Y) = c_{i+n-1}(Y)$ for some $i \geq 0$ if and only if Y is Besse and $c_i(Y)$ is a common period for its closed Reeb orbits.*

Proof. Although none of the two implications of this theorem is straightforward, we begin with a sketch of the harder one. Let us assume that $c := c_i(Y)$ is not a common period for the closed Reeb orbits of Y. Namely, the compact subset $K := \mathrm{fix}(\phi_\alpha^c)$ is not the whole Y. By hands, it is possible to construct arbitrarily small open neighborhoods $\mathcal{U} \subset \Lambda$ of the critical set $\mathcal{K} := \mathrm{crit}(\Psi) \cap \Psi^{-1}(c)$ that are homotopy equivalent to neighborhoods $U \subsetneq Y$ of the fix-point set K (in [**GGM19**], we constructed such neighborhoods by taking advantage of the geodesibility of the Reeb flow, but alternative constructions might be possible as well). In particular,

$$H^d(\mathcal{U}; \mathbb{Q}) \cong H^d(U; \mathbb{Q}) = 0, \qquad \forall d \geq 2n-1. \tag{3.8}$$

Such neighborhoods $\mathcal{U}$ are unfortunately not S^1-invariant, and therefore they do not have a well defined S^1-equivariant cohomology. Instead, we employ them as follows. Since the Ekeland-Hofer spectral invariants $c_j(Y)$ can be shown to diverge to $+\infty$ as $j \to \infty$, we have $\mathrm{ind}_{\mathrm{FR}}(\Lambda^{\leq c+1}) < \infty$. Therefore, we can find a sufficiently small S^1-invariant neighborhood $\mathcal{V} \subset \Lambda^{\leq c+1}$ of the critical set $\mathcal{K}$ such that, for any other S^1-invariant neighborhood $\mathcal{W} \subset \mathcal{V}$ of $\mathcal{K}$, we have $d := \mathrm{ind}_{\mathrm{FR}}(\mathcal{V}) = \mathrm{ind}_{\mathrm{FR}}(\mathcal{W})$. We fix a smaller neighborhood $\mathcal{U} \subset \mathcal{V}$ of $\mathcal{K}$ satisfying (3.8), and another S^1-invariant neighborhood $\mathcal{W} \subset \mathcal{U}$ of $\mathcal{K}$. The Gysin sequence implies that e^d belongs to the image of the homomorphism $\pi_* : H^{2d+1}(\mathcal{V}; \mathbb{Q}) \to H^{2d}_{S^1}(\mathcal{V}; \mathbb{Q})$. Equation (3.8), together with the commutative diagram

$$\begin{array}{ccccc} H^{2d+1}(\mathcal{V};\mathbb{Q}) & \xrightarrow{\mathrm{incl}^*} & H^{2d+1}(\mathcal{U};\mathbb{Q}) & \xrightarrow{\mathrm{incl}^*} & H^{2d+1}(\mathcal{W};\mathbb{Q}) \\ {\scriptstyle \pi_*}\downarrow & & & & \downarrow{\scriptstyle \pi_*} \\ H^{2d}_{S^1}(\mathcal{V};\mathbb{Q}) & & \xrightarrow{\qquad\mathrm{incl}^*\qquad} & & H^{2d}_{S^1}(\mathcal{W};\mathbb{Q}) \end{array}$$

implies $2d+1 = \dim(Y) < 2n-1$, that is,

$$e^{n-1} = 0 \text{ in } H^{2d}_{S^1}(\mathcal{W}; \mathbb{Q}). \tag{3.9}$$

By means of an argument à la Lusternik-Schnirelmann (cf. Theorem 2.13), the equality of the two spectral invariants $c_i(Y) = c_{i+j}(Y)$ implies that $e^j \neq 0$ in $H^{2d}_{S^1}(\mathcal{W}; \mathbb{Q})$. This, together with (3.9), implies the strict inequality $c_i(Y) < c_{i+n-1}(Y)$.

As for the converse implication, we need a property of the Morse indices of the Clarke action functional established by Ekeland [**Eke90**] along the line of the Morse index theorem in Riemannian geometry: the Morse indices of a critical point $\gamma \in \mathrm{crit}(\Psi) \cap \Psi^{-1}(\tau)$ can be computed in terms of conjugate points as

$$\mathrm{ind}(\dot\gamma) := \sum_{t \in (0,\tau)} \dim\ker(\mathrm{d}\phi_H^t(\dot\gamma(0)) - I), \tag{3.10}$$

$$\mathrm{nul}(\dot\gamma) := \dim\ker(\mathrm{d}\phi_H^\tau(\dot\gamma(0)) - I) - 1. \tag{3.11}$$

Assume now that Y is Besse. For every critical value τ of Ψ, every connected component $\mathcal{K} \subset \mathrm{crit}(\Psi) \cap \Psi^{-1}(\tau)$ is homeomorphic to a corresponding connected component $K \subset \mathrm{fix}(\phi_\alpha^\tau)$. Equations (3.11) and (3.1) imply that Ψ is a Morse-Bott functional: $\mathrm{nul}(\mathcal{K}) = \dim(\mathcal{K}) = \dim(K)$, and in particular such nullity is odd, for K is a contact submanifold of Y. It turns out that the Besse assumption further implies that all the Morse indices $\mathrm{ind}(\mathcal{K})$ are even; this can be easily established by reinterpreting $\mathrm{ind}(\dot\gamma)$ as a Maslov index.

Let c be a common period of the closed Reeb orbits, and $\mathcal{K} := \mathrm{crit}(\Psi) \cap \Psi^{-1}(c)$ the corresponding critical manifold. Notice that $\mathcal{K}$ is homeomorphic to Y; in particular, it is connected and has well defined Morse index $2i := \mathrm{ind}(\mathcal{K})$ and nullity $\mathrm{nul}(\mathcal{K}) = \dim(Y) = 2n-1$. The above properties of the Morse indices implies that, every connected component $\mathcal{K}' \subset \mathrm{crit}(\Psi) \cap \Psi^{-1}(0,c)$ satisfy $\mathrm{ind}(\mathcal{K}') + \mathrm{nul}(\mathcal{K}') < 2i-1$, and every connected component $\mathcal{K}'' \subset \mathrm{crit}(\Psi) \cap \Psi^{-1}(c,\infty)$ satisfies $\mathrm{ind}(\mathcal{K}'') > 2i+2n-1$. This, together with Morse theory, implies that

$$H^d_{S^1}(\Lambda^{<c};\mathbb{Q}) = 0, \qquad \forall d \geq 2i-1,$$

$$H^d_{S^1}(\Lambda,\Lambda^{\leq c};\mathbb{Q}) = 0, \qquad \forall d \leq 2i+2n-1,$$

where $\Lambda^{<c} := \Psi^{-1}(0,c)$ and, as before, $\Lambda^{\leq c} := \Psi^{-1}(0,c]$. This readily implies

$$c_i(Y) = c_{i+n-1}(Y) = c. \qquad \square$$

Since $c_0(Y) = \min\sigma(Y)$, Theorem 3.10 has the following immediate consequence.

Corollary 3.11. *Let Y be a convex contact sphere of dimension $2n-1$. Its Ekeland-Hofer spectral invariants satisfy $c_0(Y) = c_{n-1}(Y)$ if and only if Y is Zoll.* $\square$

A long standing open question is whether every convex contact sphere of dimension $2n-1 > 3$ always has at least n closed Reeb orbits (a stronger theorem due to Hofer, Wysozky, and Zehnder [**HWZ98**] implies that any convex contact 3-sphere has either 2 or infinitely many closed Reeb orbits). This conjecture was motivated by a theorem due to Ekeland and Lasry [**EL80**], which guarantees the existence of n closed Reeb orbits on every nearly round convex contact sphere of dimension $2n-1$. More precisely, we say that a convex contact sphere $Y \subset \mathbb{R}^{2n}$ is **δ-pinched** when it is contained in a shell

$$U(r,R) = \left\{ z \in \mathbb{R}^{2n} \;\middle|\; r \leq \|z\| \leq R \right\}$$

with $r > 0$ and $R/r < \delta$. Ekeland and Lasry's theorem holds for $\sqrt{2}$-pinched convex contact spheres. Combining their argument with the above Theorem 3.10, we obtain a characterization of nearly round convex contact spheres purely based on the knowledge of the action spectrum. In the context of geodesic flows (cf. Section 3.4), a theorem in the same spirit was proved by Ballmann, Thorbergsson, and Ziller [**BTZ83**].

Corollary 3.12. *A convex contact δ-pinched sphere Y with $\delta \in (1,\sqrt{2}]$ is Zoll if and only if its action spectrum satisfies $\sigma(Y) \cap (c_0(Y), \delta^2 c_0(Y)) = \varnothing$.*

Remark 3.13. In this corollary, the pinching constant δ cannot be taken larger than $\sqrt{2}$, since $2c_0(Y)$ is always in the action spectrum $\sigma(Y)$. Notice that the pinching condition cannot be relaxed: for instance the ellipsoid $E(1,2)$ has action spectrum $\sigma(E(1,2)) = \mathbb{N} = \{1,2,3,4,...\}$, and in particular $\sigma(Y) \cap (c_0(E(1,2)), 2c_0(E(1,2))) = \varnothing$, but is not Zoll. □

Proof of Corollary 3.12. For each $r > 0$, we consider the round convex contact sphere

$$Y_r := \big\{ z \in \mathbb{R}^{2n} \ \big|\ \|z\| = r \big\}.$$

The associated 2-homogeneous Hamiltonian and dual Hamiltonian are given by

$$H_r(z) = r^{-2}\|z\|^2, \qquad G_r(w) = \tfrac{1}{4} r^2 \|w\|^2.$$

We denote by $\Psi_r : \Lambda_r \to (0,\infty)$ the associated Clarke action functional. Since Y_r is Zoll, its spectral invariants satisfy $c_0(Y_r) = c_{n-1}(Y_r) = \pi r^2$.

Assume now that a convex contact sphere Y is δ-pinched with $\delta < \sqrt{2}$. Therefore, there exists radii $0 < r < R < \delta r$ such that $G_r \leq G \leq G_R$ pointwise, where G is the 2-homogeneous dual Hamiltonian associated to Y. These inequalities can be employed to show the inequalities $c_i(Y_r) \leq c_i(Y) \leq c_i(Y_R)$ among the Ekeland-Hofer spectral invariants. Therefore

$$c_{n-1}(Y) \leq c_{n-1}(Y_R) = \pi R^2 = \delta^2 \pi r^2 = \delta^2 c_0(Y_r) \leq \delta^2 c_0(Y).$$

Since $c_0(Y) = \min \sigma(Y)$, if $\sigma(Y) \cap (\min\sigma(Y), \delta^2 \min\sigma(Y)) = \varnothing$, then the previous inequality implies $c_{n-1}(Y) = c_0(Y)$, and therefore Y is Zoll according to Corollary 3.11. Conversely, if Y is Zoll, then every Reeb orbit has minimal period $\min\sigma(Y)$, and in particular $\sigma(Y) \cap (\min\sigma(Y), 2\min\sigma(Y)) = \varnothing$. □

The results of this section may be extendable to the class of **restricted contact type hypersurfaces** of the symplectic vector space $(\mathbb{R}^{2n}, \omega)$, which are hypersurfaces $Y \subset \mathbb{R}^{2n}$ equipped with a contact form $\theta|_Y$, where θ is a 1-form on $\mathbb{R}^{2n}$ such that $\mathrm{d}\theta = \omega$. Examples of such hypersurfaces are the spheres in $\mathbb{R}^{2n}$ that are star-shaped with respect to the origin (Figure 3.2), for which we can take $\theta = \lambda$ to be the 1-form of Equation (3.5). The Clarke action functional is not available for general restricted contact type hypersurfaces. Instead, Ekeland and Hofer [**EH89, EH90**] successfully employed the Hamiltonian action functional; while such a functional does not satisfy all those desirable properties from critical point theory (most notably its critical points have infinite Morse index), linking methods from non-linear analysis allowed to defined a version of the spectral invariants $c_i(Y)$, which in this context are known as the **Ekeland-Hofer capacities** (this marked the beginning of the theory of symplectic capacities [**HZ94**]). In the same joint work with Ginzburg and Gurel [**GGM19**], we could partially extend Theorem 3.10 as follows: a restricted contact type

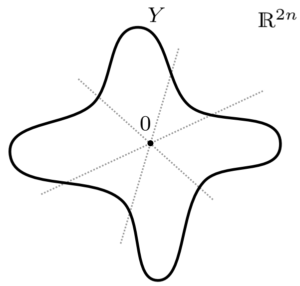


FIGURE 3.2. A hypersurface $Y \subset \mathbb{R}^{2n}$ that is starshaped with respect to the origin: every line in $\mathbb{R}^{2n}$ passing through the origin intersects Y transversely in exactly two points.

hypersurface $(Y, \theta|_Y)$ of dimension $2n-1$ with discrete action spectrum is Besse provided its Ekeland-Hofer capacities satisfy $c_i(Y) = c_{i+n-1}(Y)$ for some $i \geq 0$.

$\star$ **Open problem:** We tend to expect Theorem 3.10, with the Ekeland-Hofer capacities replacing the spectral invariants, to fully hold for restricted contact type hypersurfaces. While the assumption that the action spectrum be discrete may be relaxed with some work, the fact that the Besse assumption implies an equality of Ekeland-Hofer capacities $c_i(Y) = c_{i+n-1}(Y)$ for some $i \geq 0$ is surprisingly more challenging, since the Maslov index does does not necessarily grow monotonically along Reeb orbits as it does in the convex case (Equation (3.10)). It is well known that, for a non-convex restricted contact type hypersurface Y, $c_0(Y)$ is not necessarily the minimum of the action spectrum $\sigma(Y)$; we do not know whether Corollary 3.11 still holds for such a Y.

3.4. Spectral characterization of Zoll geodesic flows

A closed Riemannian manifold (M, g) of dimension at least 2 is called **Besse** when all its geodesics are closed, that is, its unit tangent bundle equipped with its natural contact form (Example 3.1) is a Besse contact manifold. It is called **Zoll** when all the geodesics are closed and have the same length, that is, the unit tangent bundle is a Zoll contact manifold. The theory of Besse and Zoll Riemannian manifolds attracted a lot of attention in the second half of the XX century, and is still an active area of research (although most of the remaining open problems are considered very hard). Despite not being very recent, the monograph of Besse [**Bes78**] is still the main reference for the background.

Examples of Zoll Riemannian manifolds are the compact rank-one symmetric spaces $\mathbb{RP}^n$, S^n, $\mathbb{CP}^{n/2}$, $\mathbb{HP}^{n/4}$, and the Cayley plane $\mathrm{Ca}\mathbb{P}^2$ equipped with their standard Riemannian metrics. On real projective spaces $\mathbb{RP}^n$ of dimension $n \neq 3$, a theorem of Lin and Schmidt [**LS17**] implies that any Besse Riemannian metric has constant sectional curvature (in dimension 2, this was proved earlier by Pries [**Pri09**]). A theorem due to Bott and Samelson [**Bot54, Sam63**] implies that any Besse Riemannian manifold (M, g) has the same integral cohomology ring as a compact rank-one symmetric space of the same dimension, called the model of M.

It is actually not straightforward to build examples of "exotic" Zoll Riemannian metrics: as we already mentioned in Section 3.1, Zoll [**Zol03**] constructed an example of a Zoll Riemannian 2-sphere of revolution of non-constant curvature; with more sophisticated methods based on Nash-Moser implicit function theorem, Guillemin [**Gui76**] constructed an infinite dimensional family of Zoll Riemannian metrics with trivial isometry group on the 2-sphere. These, together with the other compact rank-one symmetric spaces, are the only known examples of Zoll Riemannian manifolds.

The lens space $L(p,q) = S^3/\mathbb{Z}_p$, with $p, q > 1$ relatively prime integers, equipped with the constant curvature Riemannian metric induced by the round metric on the unit sphere $S^3 \subset \mathbb{R}^4$, is an example of Besse Riemannian metric that is not Zoll: most geodesics have length 2π, except those that are projections of $\mathbb{Z}_p$-invariant geodesics of S^3, which have length $2\pi/p$. Notice that such a lens space is not simply connected, its fundamental group being $\pi_1(L(p,q)) \cong \mathbb{Z}_p$. A conjecture due to Berger states that any simply connected Besse Riemannian manifold is actually Zoll. The conjecture was confirmed for S^2 by Gromoll and Grove [**GG81**], and for spheres S^n of dimension $n \geq 4$ by Radeschi and Wilking [**RW17**]; surprisingly, at the moment of this writing, the conjecture is still open for S^3. As we already mentioned in Remark 3.6, Berger conjecture fails in the Finsler setting: Katok's non-reversible Finsler metrics F_α of Example 2.10 with a rational parameter α are Besse but not Zoll.

Let (M, g) be a simply connected closed Riemannian manifold of dimension $n \geq 2$ that has the integral cohomology of some compact rank-one symmetric space M_0 with $\dim(M) = \dim(M_0)$. In this section, we shall employ the S^1-equivariant Morse theory of the energy functional

$$E : \Lambda M \to [0, \infty), \qquad E(\gamma) = \int_{S^1} \|\dot\gamma(t)\|_g^2 \mathrm{d}t$$

in order to provide a spectral characterization of the Zoll condition along the lines of Corollary 3.11. The properties of E that we shall employ are the following.

- **(Morse-Bott)** If g is a Besse Riemannian metric, then E is a Morse-Bott functional: any connected component $K \subset \mathrm{crit}(E)$ has nullity $\mathrm{nul}(K) = \dim(K)$. This readily follows from the property (3.1) of the strata of Besse contact manifolds, together with the characterization (1.7) of the nullity of the energy functional.

- **(Bott formula)** If g is a Besse Riemannian metric whose unit-speed closed geodesics have common period ℓ, and $K := \mathrm{crit}(E) \cap E^{-1}(\ell^2) \cong SM$ is the corresponding critical manifold, the Morse indices of the iterates of K are given by

$$\mathrm{ind}(K^m) = m\,\mathrm{ind}(K) + (m-1)(n-1), \qquad \forall m \geq 1.$$

This is a special instance of Bott's iteration formula for the Morse indices [**Bot56**].

- **(Perfectness)** If g is a Besse Riemannian metric, the energy functional E is perfect for the rational S^1-equivariant relative cohomology relative to the constants: if we denote as usual the energy sublevel sets by $\Lambda M^{<c} := E^{-1}[0, c)$, for each $c > b > a > 0$ the inclusions induce a short exact sequence

$$0 \longrightarrow H^*_{S^1}(\Lambda M^{<c}, \Lambda M^{<b}) \longrightarrow H^*_{S^1}(\Lambda M^{<c}, \Lambda M^{<a}) \longrightarrow H^*_{S^1}(\Lambda M^{<b}, \Lambda M^{<a}) \longrightarrow 0.$$

Namely, in the S^1-equivariant Morse theory, there are no cancellations among the critical manifolds with positive critical values. We should stress, here, that the requirements that $a > 0$ is essential, since the constants $M \equiv E^{-1}(0) \subset \Lambda M$ do interact with the other critical manifolds. The perfectness was established by Hingston in the Zoll case in her seminal paper [**Hin84**]. In the general Besse case, the statement is much harder, and was only established recently by Radeschi and Wilking [**RW17**].

• **(Orientability of the negative bundles)** Assume that M is a spin manifold (which is always the case unless its model compact rank-one symmetric space M_0 is $\mathbb{CP}^{2m}$ of even complex dimension), and g is Besse. For each connected component $K \subset \mathrm{crit}(E)$, we denote by $\pi : N_K \to K$ its negative normal bundle, which is the vector bundle whose fibers $\pi^{-1}(\gamma)$ are given by the negative eigenspaces of the Hessian $\nabla^2 E(\gamma)$, i.e.

$$\pi^{-1}(\gamma) = \bigoplus_{\lambda<0} \ker(\nabla^2 E(\gamma) - \lambda I).$$

As in the tubular neighborhood theorem in differential geometry, we can identify N_K with an open submanifold of dimension $\mathrm{ind}(K) + \mathrm{nul}(K)$ in the free loop space ΛM, containing K in its interior. A deep result due to Radeschi and Wilking [**RW17**] implies that, under the Besse assumption, every negative bundle is orientable. Therefore, the cup product with its Thom class $u_K \in H^{\mathrm{ind}(K)}(N_K, N_K \setminus K)$ gives an isomorphism of $H^*(BS^1)$ modules

$$H^*_{S^1}(K) \xrightarrow[\cong]{\smile u_K} H^{*+\mathrm{ind}(K)}(N_K, N_K \setminus K).$$

• **(Minimal degree of the cohomology)** The S^1-equivariant cohomology $H^*_{S^1}(\Lambda M, M; \mathbb{Q})$ was computed by Hingston in her seminal paper [**Hin84**]. We will not need the full computation for our purposes, but only the following statement:

$$H^{i(M)}(\Lambda M, M; \mathbb{Q}) \cong \mathbb{Q}, \qquad H^d(\Lambda M, M; \mathbb{Q}) = 0,\ \forall d < i(M),$$

where

$$i(M) = i(M_0) = \begin{cases} n-1, & \text{if } M_0 = S^n, \\ 1, & \text{if } M_0 = \mathbb{CP}^{n/2}, \\ 3, & \text{if } M_0 = \mathbb{HP}^{n/4}, \\ 7, & \text{if } M_0 = \mathrm{Ca}\mathbb{P}^2. \end{cases}$$

For each non-zero cohomology class $h \in H^*_{S^1}(\Lambda M, M; \mathbb{Q})$, we consider the associated spectral invariant

$$c_g(h) := \inf \left\{ \sqrt{c} > 0 \mid h \notin \ker \left(H^*_{S^1}(\Lambda M, M; \mathbb{Q}) \xrightarrow{\mathrm{incl}^*} H^*_{S^1}(\Lambda M^{<c}, M; \mathbb{Q}) \right) \right\}.$$

We recall that the S^1-equivariant cohomology groups are $H^*(BS^1; \mathbb{Q})$-modules, and the same properties clearly holds for the relative cohomology group $H^*_{S^1}(\Lambda M, M; \mathbb{Q})$. In particular, as in the previous section we can see the rational Euler class, which is a generator

$e \in H^2(BS^1;\mathbb{Q})$, as an element in $H^*_{S^1}(\Lambda M;\mathbb{Q})$, so that $e \smile h \in H^*_{S^1}(\Lambda M, M;\mathbb{Q})$ for all $h \in H^*_{S^1}(\Lambda M, M;\mathbb{Q})$.

In our joint work with Ginzburg and Gurel [**GGM19**], we characterized the Zoll Riemannian metrics in terms of the above spectral invariants. We will consider simply connected spin closed manifolds M that are integral cohomology compact rank-one symmetric spaces, meaning that there is a ring isomorphism $H^*(M;\mathbb{Z}) \cong H^*(M_0;\mathbb{Z})$, where $\dim(M_0) = \dim(M) = n$ and M_0 is either S^n, $\mathbb{CP}^{n/2}$ (with $n/2$ odd), $\mathbb{HP}^{n/4}$, or $\mathrm{Ca}\mathbb{P}^2$ (with $n = 16$).

Theorem 3.14. *Let M be a simply connected, spin, integral cohomology compact rank-one symmetric space of dimension $n \geq 2$, and k a generator of the cohomology group*

$$H^{i(M)}_{S^1}(\Lambda M, M;\mathbb{Q}) \cong \mathbb{Q}.$$

A Riemannian metric g on M is Zoll if and only if $c_g(k) = c_g(e^{n-1} \smile k)$.

Proof. Throughout this proof, all cohomology groups are assumed to have coefficients in the field of rational numbers $\mathbb{Q}$, and we will remove $\mathbb{Q}$ from the notation.

Let us assume that (M, g) is a Zoll Riemannian manifold (spin or not). If ℓ is the length of the closed geodesics, then

$$\mathrm{crit}(E) \cap E^{-1}(0,\infty) = \bigcup_{m\geq 1} K^m,$$

where $K := \mathrm{crit}(E) \cap E^{-1}(\ell^2) \cong SM$. The circle S^1 acts freely on K, and the quotient K/S^1 is a smooth manifold of dimension $2n-2$ admitting a symplectic form ω such that $\pi^*\omega = \mathrm{d}\alpha$, where $\pi : K \to K/S^1$ is the quotient projection and α is the natural contact form of SM. The quotient projection $\mathrm{pr}_1 : K \times_{S^1} ES^1 \to K/S^1$ is a homotopy equivalence, and we have $\mathrm{pr}_1^*[\omega] = e \in H^2_{S^1}(K)$. This readily implies that $e^{n-1} \neq 0$ in $H^{2n-2}_{S^1}(K)$, since ω^n is a volume form. Since E is a perfect functional and all the negative bundles of its critical manifolds are orientable, we have

$$H^*_{S^1}(\Lambda M, M) \cong \bigoplus_{m\geq 1} H^{*-\mathrm{ind}(K)}_{S^1}(K^m). \tag{3.12}$$

Bott's formula for the Morse indices implies that $m \mapsto \mathrm{ind}(K^m)$ is strictly monotone increasing. This, together with the fact that $i = i(M)$ is the first degree in which the cohomology $H^*_{S^1}(\Lambda M, M)$ is non-trivial, implies that

$$\mathrm{ind}(K) = i(M).$$

Now, consider the homomorphisms of $H^*(BS^1)$-modules

$$H^{*-i}_{S^1}(N_K) \xrightarrow[\cong]{\smile u_K} H^*_{S^1}(N_K, N_K \setminus K) \xleftarrow{\mathrm{incl}^*} H^*_{S^1}(\Lambda M, \Lambda M^{<\ell^2}) \xrightarrow[\cong]{\mathrm{incl}^*} H^*_{S^1}(\Lambda M, M).$$

Since $H^i_{S^1}(\Lambda M, M) \cong \mathbb{Q}$, the second homomorphism is an isomorphism. Therefore, this sequence of homomorphisms sends $1 \in H^0_{S^1}(N_K)$ to a generator $k \in H^i_{S^1}(\Lambda M, M)$, and conversely it sends $e^{n-1} \smile k \in H^i_{S^1}(\Lambda M, M)$ to $e^{n-1} \in H^{2n-2}_{S^1}(N_K)$. We readily conclude that $e^{n-1} \smile k \neq 0$ and $c_g(h) = c_g(e^{n-1} \smile k) = \ell^2$.

Assume now that M is a a simply connected, spin, integral cohomology compact rank-one symmetric space of dimension $n \geq 2$. It was proved by Radeschi and Wilking [**RW17**] that we have a ring isomorphism $H^*_{S^1}(\Lambda M) \cong H^*_{S^1}(\Lambda M_0)$, where M_0 is the model of M. Since M_0 admits a Zoll Riemannian metric, by the result of the previous paragraph we know that $e^{n-1} \smile k \neq 0$ in $H^{i+2n-2}(\Lambda M, M)$, where $i = i(M)$ and k is a generator of $H^i_{S^1}(\Lambda M, M)$. Let g be a Riemannian metric on M such that $\ell := c_g(k) = c_g(e^{n-1} \smile k)$. Since the energy functional E is perfect and all its critical manifolds have orientable negative bundle, i must be the minimum among the Morse indices of the critical manifolds (otherwise $H^d_{S^1}(\Lambda M, M)$ would be non-trivial in some degree $d < i$). By an argument analogous to the one in the proof of Theorem 3.10, g is Besse and the unit-speed closed geodesics have common period ℓ. The critical manifold $K := \mathrm{crit}(E) \cap E^{-1}(\ell^2)$ is homeomorphic to the unit tangent bundle SM. Since the cohomology class k has degree i, we have $\mathrm{ind}(K) \leq i$, and thus $\mathrm{ind}(K) = i$. We are left to show that every geodesic has length ℓ, namely that ℓ^2 is the smallest positive critical value of the energy functional E. Let us assume by contradiction that there exists a smaller critical value $c_0 \in (0, \ell^2)$ of E. We consider a connected component $K_0 \subset \mathrm{crit}(E) \cap E^{-1}(c_0)$. Since $K_0^m \subset K$ for $m = \ell/\sqrt{c_0}$, we have $\mathrm{ind}(K_0) \leq \mathrm{ind}(K) = i$, and thus $\mathrm{ind}(K_0) = i$ (for i is the smallest possible Morse index for the critical manifolds of E). By (3.12), we have

$$\mathrm{rank}\, H^i_{S^1}(\Lambda M, M) \geq \mathrm{rank}\, H^0_{S^1}(K_0) + \mathrm{rank}\, H^0_{S^1}(K) = 2,$$

contradicting the fact that $H^i_{S^1}(\Lambda M, M) \cong \mathbb{Q}$. □

By combining the arguments in the proof of Theorem 3.14 with the validity of the Berger conjecture for spheres of any dimension except possibly 3, we obtain a slightly stronger result for Riemannian spheres.

Theorem 3.15. *For any Riemannian sphere (S^n, g) of dimension $n \neq 3$, the following three conditions are equivalent:*

(i) *(S^n, g) is Zoll.*
(ii) *There exists an integer $m \geq 1$ and a generator $k_m \in H^{(2m-1)(n-1)}_{S^1}(\Lambda S^n, S^n; \mathbb{Q})$ such that $c_g(k_m) = c_g(e^{n-1} \smile k_m)$.*
(ii) *For each integer $m \geq 1$, there exists a generator $k_m \in H^{(2m-1)(n-1)}_{S^1}(\Lambda S^n, S^n; \mathbb{Q})$ such that $c_g(k_m) = c_g(e^{n-1} \smile k_m)$.* □

3.5. On the structure of Besse contact manifolds

In this final section, we investigate the structure of Besse contact manifolds. In the Riemannian setting, this was already addressed in the previous section: Berger's conjecture implies that every simply connected Besse Riemannian manifold is Zoll. While Guillemin [**Gui76**] provided a large family of non-isometric Zoll Riemannian metrics on the 2-sphere, a theorem of Abbondandolo, Bramham, Hryniewicz, and Salomão [**ABHSa17, ABHSa18**] implies that their unit tangent bundles are all strictly contactomorphic (we recall that two contact manifolds (Y_1, α_1) and (Y_2, α_2) are strictly contactomorphic when there exists a diffeomorphism $\psi : Y_1 \to Y_2$ such that $\psi^*\alpha_2 = \alpha_1$). Indeed, extending this latter result, Benedetti and Kang [**BK18**] showed that two diffeomorphic Zoll contact 3-manifolds are strictly contactomorphic.

Beyond Zoll contact manifolds, in a joint work with Cristofaro Gardiner [**CGM20**] we proved the following theorem, which solves an inverse problem for Besse contact 3-manifolds. We adopt the notation of Section 3.1, and denote by $\sigma_{\mathrm{p}}(Y, \alpha)$ the prime action spectrum of the contact manifold (Y, α).

Theorem 3.16. *Let Y be a closed connected 3-manifold, and α_0, α_1 two Besse contact forms on Y. Then $\sigma_{\mathrm{p}}(Y, \alpha_0) = \sigma_{\mathrm{p}}(Y, \alpha_1)$ if and only if there exists a diffeomorphism $\psi : Y \to Y$ such that $\psi^*\alpha_1 = \alpha_0$.*

Proof. Let $\tau_i > 0$ be the minimal common period of the Reeb orbits of (Y, α_i). We denote by Σ_i the quotient of Y by the circle action given by the Reeb flow $\phi^t_{\alpha_i}$. The quotient projection $\pi_i : Y \to \Sigma_i$ is almost a principal circle bundle: it is a Seifert fibration. The total space Y is oriented by the contact volume form $\alpha_i \wedge \mathrm{d}\alpha_i$, and the base space Σ_i, which is a closed surface, is oriented via the symplectic form ω_i pushed forward from $\mathrm{d}\alpha_i$. Such Seifert spaces were classified by Orlik, Vogt, and Zieschang [**OVZ67**] when Y is not a lens space, and the full classification has been extended to lens spaces recently by Geiges and Lange [**GL18**]. Employing this classification, it turns out that the equality of the prime action spectra $\sigma_{\mathrm{p}}(Y, \alpha_0) = \sigma_{\mathrm{p}}(Y, \alpha_1)$ implies that the two Seifert fibrations are isomorphic: there exists a diffeomorphism $\phi : Y \to Y$ such that $\mathrm{d}\phi(z) R_{\alpha_0}(z) = R_{\alpha_1}(z)$ for all $z \in Y$, and the volume forms $\phi^*(\alpha_1 \wedge \mathrm{d}\alpha_1)$ and $\alpha_0 \wedge \mathrm{d}\alpha_0$ define the same orientation on Y. In order to simplify the notation, let us assume without loss of generality that $\phi = \mathrm{id}$, so that α_0 and α_1 define the same Reeb vector field $R = R_{\alpha_0} = R_{\alpha_1}$ and the same orientation on Y via their contact volume forms. For each $t \in [0, 1]$, the convex combination $\alpha_t := t\,\alpha_1 + (1-t)\alpha_0$ is a contact form. Indeed, consider any oriented basis of a tangent space T_zY of the form $R(z), v, w$. We have

$$\alpha_i \wedge \mathrm{d}\alpha_j(R(z), v, w) = \mathrm{d}\alpha_j(v, w) = \alpha_j \wedge \mathrm{d}\alpha_j(R(z), v, w) > 0, \quad \forall i, j \in \{0, 1\}.$$

This readily implies that the 3-form

$$\alpha_t \wedge \mathrm{d}\alpha_t = t^2 \alpha_1 \wedge \mathrm{d}\alpha_1 + (1-t)^2 \alpha_0 \wedge \mathrm{d}\alpha_0 + t(1-t)(\alpha_0 \wedge \mathrm{d}\alpha_1 + \alpha_1 \wedge \mathrm{d}\alpha_0)$$

is a positive volume form on Y, and in particular each α_t is a contact form. We can now complete the proof by applying a Moser trick as follows. We consider the time-dependent vector field X_t on Y defined by $\alpha_t(X_t) \equiv 0$ and $X_t \lrcorner\, \mathrm{d}\alpha_t = \alpha_0 - \alpha_1$. Its flow $\psi_t : Y \to Y$, with $\psi_0 = \mathrm{id}$, satisfies

$$\tfrac{\mathrm{d}}{\mathrm{d}t} \psi_t^* \alpha_t = \psi_t^* \big(\mathrm{d}(\alpha_t(X_t)) + X_t \lrcorner\, \mathrm{d}\alpha_t + \alpha_1 - \alpha_0 \big) = 0,$$

and therefore $\psi_1^* \alpha_1 = \alpha_0$. $\square$

We now focus on convex spheres. In dimension 3, as a byproduct of Theorem 3.16 and of the classification of Seifert fibrations [**GL18**], we can show that there are essentially no Besse contact spheres beside the rational ellipsoids (Example 3.2).

Theorem 3.17. *Every Besse contact 3-sphere (S^3, α) is strictly contactomorphic to an ellipsoid.*

Proof. Let $\tau > 0$ be the minimal common period of the Reeb orbits of (S^3, α). The classification of Seifert fibrations [**GL18**] implies that there are at most two distinct periodic Reeb orbits whose minimal period is strictly less than τ. Therefore, $\sigma_{\mathrm{p}}(S^3, \alpha)$ contains at most three elements, one of which is τ. If $\sigma_{\mathrm{p}}(S^3, \alpha)$ contains at least two elements, we call $a_1 < a_2$ the two smallest elements of $\sigma_{\mathrm{p}}(S^3, \alpha)$; if instead $\sigma_{\mathrm{p}}(S^3, \alpha) = \{\tau\}$, we set $a_1 = a_2 = \tau$. In both cases, we conclude that $\sigma_{\mathrm{p}}(S^3, \alpha)$ is also the simple action spectrum of the Besse ellipsoid $E(a_1, a_2)$. Finally, by Theorem 3.16, two diffeomorphic Besse contact 3-manifolds are strictly contactomorphic if and only if they have the same prime action spectrum. □

In higher dimension, a result such as Theorem 3.17 seems out of reach, and indeed it is not even known whether a Zoll contact $(2n+1)$-sphere is strictly contactomorphic to a Zoll ellipsoid $E(\tau, ..., \tau)$. A celebrated related problem is the uniqueness (up to diffeomorphism) of the symplectic form with unit volume on complex projective spaces, which is known for $\mathbb{CP}^2$ by a result of Taubes [**Tau95**], but open in higher dimension.

Motivated by Theorem 3.17, we close this section and the memoir with a sketch of the proof of our very recent joint result with Radeschi [**MR20**], which shows that Besse convex contact spheres have at least some resemblance with the rational ellipsoids. We recall that the strata of a Besse contact ellipsoid are sub-ellipsoids (Example 3.2), and that the sequence of Ekeland-Hofer spectral invariants of ellipsoids is given by the sequence of elements in the action spectrum, where each value is repeated according to a suitable multiplicity (Example 3.9).

Theorem 3.18. *Let Y be a Besse convex contact sphere with Reeb flow ϕ_α^t, and $\tau > 0$ the minimal common period of its Reeb orbits.*

(i) *For each $k \in \mathbb{N}$, the stratum $Y_k := \mathrm{fix}(\phi_\alpha^{\tau/k})$ is either empty or an integral homology sphere (i.e. $H_*(Y_k; \mathbb{Z}) \cong H_*(S^{2d+1}; \mathbb{Z})$, where $2d + 1 = \dim(Y_k)$).*

(ii) *Let $\sigma_1 < \sigma_2 < \sigma_3 < \dots$ be the elements of the action spectrum $\sigma(Y)$ listed in increasing order. The Ekeland-Hofer spectral invariant $c_i(Y)$ is the $(i+1)$-th element in the sequence*

$$\underbrace{\sigma_1, \dots, \sigma_1}_{\times d_1}, \underbrace{\sigma_2, \dots, \sigma_2}_{\times d_2}, \underbrace{\sigma_3, \dots, \sigma_3}_{\times d_3}, \dots$$

where $d_j = (\dim(\mathrm{fix}(\phi_\alpha^{\sigma_j})) + 1)/2$.

Remark 3.19. To the best of our knowledge, it is unknown whether Theorem 3.18(i) is true for a general locally free S^1 action on a sphere Y. The classical Smith theorem [**Bre72**] implies that, if k is a power of a prime p, then the subspace Y_k, being the set fixed by the cyclic subgroup $\mathbb{Z}_k \subset S^1$, is either empty or a $\mathbb{Z}_p$-homology sphere (i.e. $H_*(Y_k;\mathbb{Z}_p) \cong H_*(S^d;\mathbb{Z}_p)$ for some $d \geq 0$). On the other hand, there are examples of manifolds equipped with a smooth action by a finite cyclic group whose fixed point set is not an integral homology sphere [**Sch82**], although it is not known if these examples can arise as strata of smooth locally free circle actions. □

Proof of Theorem 3.18. We only provide a summary of the proof, which is rather involved. We consider the Clarke action functional $\Psi : \Lambda \to (0,\infty)$ associated with the Besse convex contact sphere Y. We know that Ψ is Morse-Bott: every connected component $K \subset \mathrm{crit}(\Psi)$ is a closed manifold with minimal nullity, that is, $\mathrm{nul}(K) = \dim(K)$. Let $\pi : N_K \to K$ be the negative normal bundle of the critical manifold K, which is the vector bundle whose fibers $\pi^{-1}(\dot\gamma)$ are given by the negative eigenspaces of the Hessian $\nabla^2\Psi(\dot\gamma)$, i.e.

$$\pi^{-1}(\dot\gamma) = \bigoplus_{\lambda<0} \ker(\nabla^2\Psi(\dot\gamma) - \lambda I).$$

The first ingredient of the proof is the fact that $\pi : N_K \to K$ is an orientable vector bundle, a version of a Riemannian result of Radeschi and Wilking [**RW17**] that we already mentioned in Section 3.4. This implies that the critical sets $K := \mathrm{crit}(\Psi) \cap \Psi^{-1}(c)$ are cohomologically visible: if $\pi_0(K)$ denotes as usual the family of connected components of K, we have an isomorphism of $H^*(BS^1)$-modules

$$H^*_{S^1}(\Lambda^{<c+\epsilon}, \Lambda^{<c};\mathbb{Q}) \cong \bigoplus_{K' \in \pi_0(K)} H^{*-\mathrm{ind}(K')}_{S^1}(K';\mathbb{Q}),$$

for $\epsilon > 0$ small enough.

The second ingredient of the proof is the fact that every connected component $K \subset \mathrm{crit}(\Psi)$ has trivial S^1-equivariant cohomology $H^d_{S^1}(K;\mathbb{Q})$ in every odd degree d. This is proved by means of a subtle analysis of the torsion of the cohomology groups $H^d_{S^1}(\Lambda^{<b};\mathbb{Z})$, for $b > 0$. It turns our that, if some rational cohomology group $H^d_{S^1}(K;\mathbb{Q})$ is non-trivial in some odd degree d, it produces non-trivial torsion in $H^p_{S^1}(\Lambda;\mathbb{Z})$ in some odd degree p. This is not possible, for a simple deformation argument shows that

$$H^q_{S^1}(\Lambda;\mathbb{Z}) \cong H^q(BS^1;\mathbb{Z}) \cong \begin{cases} \mathbb{Z}, & \text{if } q \text{ is even,} \\ 0, & \text{if } q \text{ is odd.} \end{cases} \tag{3.13}$$

We already mentioned in the proof of Theorem 3.10 that all the connected components $K \subset \mathrm{crit}(\Psi)$ have even Morse index $\mathrm{ind}(K)$. This, together with the above second ingredient and the lacunary principle from Morse theory, implies that the Clarke action functional is perfect for the S^1-equivariant Morse theory, i.e. for all $b > a > 0$ the inclusions induce a short exact sequence

$$0 \longrightarrow H^*_{S^1}(\Lambda^{<b}, \Lambda^{<a};\mathbb{Q}) \longrightarrow H^*_{S^1}(\Lambda^{<b};\mathbb{Q}) \longrightarrow H^*_{S^1}(\Lambda^{<a};\mathbb{Q}) \longrightarrow 0.$$

Namely, in the S^1-equivariant Morse theory of Ψ, there are no cancellations among the critical manifolds. If we denote by $\sigma_1, \sigma_2, \sigma_3, ...$ the elements of the action spectrum $\sigma(y)$ listed in increasing order, by $K_j := \mathrm{crit}(\Psi) \cap \Psi^{-1}(\sigma_j)$ the corresponding critical set, and by $\pi_0(K_j)$ the family of path-connected components of K_j, the perfectness of Ψ, together with the above first ingredient, provides isomorphisms of $H^*(BS^1; \mathbb{Q})$-modules

$$H^*_{S^1}(\Lambda; \mathbb{Q}) \cong \bigoplus_{j \in \mathbb{N}} H^*_{S^1}(\Lambda^{\leq \sigma_j}, \Lambda^{<\sigma_j}; \mathbb{Q}) \cong \bigoplus_{j \in \mathbb{N}} \bigoplus_{K \in \pi_0(K_j)} H^{*-\mathrm{ind}(K)}_{S^1}(K; \mathbb{Q}) \tag{3.14}$$

Equation (3.13) implies that $H^*_{S^1}(\Lambda; \mathbb{Q}) = \mathbb{Q}[e]$, where e is the Euler class. For each $j \in \mathbb{N}$, let $\iota_0(j)$ and $\iota_0(j)$ be the minimal and maximal degrees d such that $H^d_{S^1}(\Lambda^{\leq \sigma_j}, \Lambda^{<\sigma_j}; \mathbb{Q}) \neq 0$. By the first isomorphism in Equation (3.14), for each non-zero cohomology classes

$$h_i \in H^{\iota_i(j)}_{S^1}(\Lambda^{\leq \sigma_j}, \Lambda^{<\sigma_j}; \mathbb{Q}), \qquad i = 0, 1,$$

there is $\lambda \in \mathbb{Q}$ such that $e^{d_j - 1} \smile h_0 = \lambda h_1$, where $d_j = 1 + (\iota_1(j) - \iota_0(j))/2$. This, together with the second isomorphism in Equation (3.14), implies that K_j is path-connected, $\iota_0(j) = \mathrm{ind}(K_k)$, $2d_j - 1 = \iota_1(j) - \iota_0(j) = \dim(K_j) - 1$, and $\iota_0(j+1) = \iota_1(j) + 2$. This proves point (ii) and

$$H^*_{S^1}(K_j; \mathbb{Q}) \cong \frac{\mathbb{Q}[e]}{(e^{d_j})},$$

which, by means of the Gysin sequence, implies that

$$H_*(K_j; \mathbb{Q}) \cong H^*(K_j; \mathbb{Q}) \cong H^*(S^{2d_j - 1}; \mathbb{Q}).$$

Finally, since $H_d(\Lambda; \mathbb{Z})$ vanishes in positive degrees $d > 0$, the ordinary (i.e. non-equivariant) Morse theory of Ψ implies that $H_*(K_j; \mathbb{Z})$ is torsion free, and therefore $H_*(K_j; \mathbb{Z}) \cong H_*(S^{2d_j - 1}; \mathbb{Z})$. □

Bibliography

[AB83] M. F. Atiyah and R. Bott, *The Yang-Mills equations over Riemann surfaces*, Philos. Trans. Roy. Soc. London Ser. A **308** (1983), no. 1505, 523–615.

[Abb13] A. Abbondandolo, *Lectures on the free period Lagrangian action functional*, J. Fixed Point Theory Appl. **13** (2013), no. 2, 397–430.

[ABHSa17] A. Abbondandolo, B. Bramham, U. L. Hryniewicz, and P. A. S. Salomão, *A systolic inequality for geodesic flows on the two-sphere*, Math. Ann. **367** (2017), no. 1-2, 701–753.

[ABHSa18] A. Abbondandolo, B. Bramham, U. L. Hryniewicz, and P. A. S. Salomão, *Sharp systolic inequalities for Reeb flows on the three-sphere*, Invent. Math. **211** (2018), no. 2, 687–778.

[ABM17] L. Asselle, G. Benedetti, and M. Mazzucchelli, *Minimal boundaries in Tonelli Lagrangian systems*, https://doi.org/10.1093/imrn/rnz246, to appear in Int. Math. Res. Not. IMRN, 2017.

[AD14] M. Audin and M. Damian, *Morse theory and Floer homology*, Universitext, Springer, London; EDP Sciences, Les Ulis, 2014, Translated from the 2010 French original by Reinie Erné.

[AF07] A. Abbondandolo and A. Figalli, *High action orbits for Tonelli Lagrangians and superlinear Hamiltonians on compact configuration spaces*, J. Differential Equations **234** (2007), no. 2, 626–653.

[AFM15] P. Albers, U. Fuchs, and W. J. Merry, *Orderability and the Weinstein conjecture*, Compos. Math. **151** (2015), no. 12, 2251–2272.

[AM11] P. Albers and M. Mazzucchelli, *Periodic bounce orbits of prescribed energy*, Int. Math. Res. Not. (2011), no. 14, 3289–3314.

[AM17] L. Asselle and M. Mazzucchelli, *On the existence of infinitely many closed geodesics on non-compact manifolds*, Proc. Amer. Math. Soc. **145** (2017), no. 6, 2689–2697.

[AM18] ———, *Closed geodesics with local homology in maximal degree on non-compact manifolds*, Differential Geom. Appl. **58** (2018), 17–51.

[AM19] ———, *On Tonelli periodic orbits with low energy on surfaces*, Trans. Amer. Math. Soc. **371** (2019), no. 5, 3001–3048.

[AM20] ———, *Waist theorems for Tonelli systems in higher dimensions*, Manuscripta Math. **163** (2020), no. 1-2, 185–199.

[AMMP17] A. Abbondandolo, L. Macarini, M. Mazzucchelli, and G. P. Paternain, *Infinitely many periodic orbits of exact magnetic flows on surfaces for almost every subcritical energy level*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 2, 551–579.

[AMP15] A. Abbondandolo, L. Macarini, and G. P. Paternain, *On the existence of three closed magnetic geodesics for subcritical energies*, Comment. Math. Helv. **90** (2015), no. 1, 155–193.

[Ang08] S. Angenent, *Self-intersecting geodesics and entropy of the geodesic flow*, Acta Math. Sin. (Engl. Ser.) **24** (2008), no. 12, 1949–1952.

[Ano82] D. V. Anosov, *Generic properties of closed geodesics*, Izv. Akad. Nauk SSSR Ser. Mat. **46** (1982), no. 4, 675–709, 896.

[Arn78] V. I. Arnold, *Mathematical methods of classical mechanics*, Springer-Verlag, New York, 1978.

[AS09] A. Abbondandolo and M. Schwarz, *A smooth pseudo-gradient for the Lagrangian action functional*, Adv. Nonlinear Stud. **9** (2009), no. 4, 597–623.

[Bab90] I. K. Babenko, *Periodic trajectories of three-dimensional Birkhoff billiards*, Mat. Sb. **181** (1990), no. 9, 1155–1169.

[Bal78] W. Ballmann, *Der Satz von Lusternik und Schnirelmann*, Beiträge zur Differentialgeometrie, Heft 1, Bonner Math. Schriften, vol. 102, Univ. Bonn, Bonn, 1978, pp. 1–25.

[Ban80] V. Bangert, *Closed geodesics on complete surfaces*, Math. Ann. **251** (1980), no. 1, 83–96.

[Ban93] ———, *On the existence of closed geodesics on two-spheres*, Internat. J. Math. **4** (1993), no. 1, 1–10.

[BB81] M. S. Berger and E. Bombieri, *On Poincaré's isoperimetric problem for simple closed geodesics*, J. Functional Analysis **42** (1981), no. 3, 274–298.

[Ben86] V. Benci, *Periodic solutions of Lagrangian systems on a compact manifold*, J. Differential Equations **63** (1986), no. 2, 135–161.

[Ber03] M. Berger, *A panoramic view of Riemannian geometry*, Springer-Verlag, Berlin, 2003.

[Bes78] A. L. Besse, *Manifolds all of whose geodesics are closed*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 93, Springer-Verlag, Berlin-New York, 1978.

[BG89] V. Benci and F. Giannoni, *Periodic bounce trajectories with a low number of bounce points*, Ann. Inst. H. Poincaré, Anal. Non Linéaire **6** (1989), no. 1, 73–93.

[BG92] ———, *On the existence of closed geodesics on noncompact Riemannian manifolds*, Duke Math. J. **68** (1992), no. 2, 195–215.

[BH84] V. Bangert and N. Hingston, *Closed geodesics on manifolds with infinite Abelian fundamental group.*, J. Differ. Geom. **19** (1984), 277–282.

[BI16] D. Burago and S. Ivanov, *Boundary distance, lens maps and entropy of geodesic flows of Finsler metrics*, Geom. Topol. **20** (2016), no. 1, 469–490.

[Bir66] G. D. Birkhoff, *Dynamical systems*, With an addendum by Jurgen Moser. American Mathematical Society Colloquium Publications, Vol. IX, American Mathematical Society, Providence, R.I., 1966.

[BK83] V. Bangert and W. Klingenberg, *Homology generated by iterated closed geodesics*, Topology **22** (1983), no. 4, 379–388.

[BK18] G. Benedetti and J. Kang, *A local systolic-diastolic inequality in contact and symplectic geometry*, arXiv:1801.00539, to appear in J. Eur. Math. Soc. (JEMS), 2018.

[Bot54] R. Bott, *On manifolds all of whose geodesics are closed.*, Ann. of Math. **60** (1954), no. 2, 375–382.

[Bot56] ———, *On the iteration of closed geodesics and the Sturm intersection theory*, Comm. Pure Appl. Math. **9** (1956), 171–206.

[Bre72] G. E. Bredon, *Introduction to compact transformation groups*, Academic Press, New York-London, 1972, Pure and Applied Mathematics, Vol. 46.

[BTZ81] W. Ballmann, G. Thorbergsson, and W. Ziller, *Closed geodesics and the fundamental group*, Duke Math. J. **48** (1981), no. 3, 585–588.

[BTZ82] ———, *Closed geodesics on positively curved manifolds*, Ann. of Math. **116** (1982), no. 2, 213–247.

[BTZ83] ———, *Existence of closed geodesics on positively curved manifolds*, J. Differential Geom. **18** (1983), no. 2, 221–252.

[BW58] W. M. Boothby and H. C. Wang, *On contact manifolds*, Ann. of Math. (2) **68** (1958), 721–734.

[Car95] M. J. D. Carneiro, *On minimizing measures of the action of autonomous Lagrangians*, Nonlinearity **8** (1995), no. 6, 1077–1085.

[CFP10] K. Cieliebak, U. Frauenfelder, and G. P. Paternain, *Symplectic topology of Mañé's critical values*, Geom. Topol. **14** (2010), no. 3, 1765–1870.

[CGH16] D. Cristofaro-Gardiner and M. Hutchings, *From one Reeb orbit to two*, J. Differential Geom. **102** (2016), no. 1, 25–36.

[CGM20] D. Cristofaro-Gardiner and M. Mazzucchelli, *The action spectrum characterizes closed contact 3-manifolds all of whose Reeb orbits are closed*, Comment. Math. Helv. **95** (2020), no. 3, 461–481.

[CI99] G. Contreras and R. Iturriaga, *Global minimizers of autonomous Lagrangians*, 22° Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro, 1999.

[CIPP98] G. Contreras, R. Iturriaga, G. P. Paternain, and M. Paternain, *Lagrangian graphs, minimizing measures and Mañé's critical values*, Geom. Funct. Anal. **8** (1998), no. 5, 788–809.

[CIPP00] ———, *The Palais-Smale condition and Mañé's critical values*, Ann. Henri Poincaré **1** (2000), no. 4, 655–684.

[Cla79] F. H. Clarke, *A classical variational principle for periodic Hamiltonian trajectories*, Proc. Amer. Math. Soc. **76** (1979), no. 1, 186–188.

[Con06] G. Contreras, *The Palais-Smale condition on contact type energy levels for convex Lagrangian systems*, Calc. Var. Partial Differ. Equ. **27** (2006), no. 3, 321–395.

[dC92] M.P. do Carmo, *Riemannian geometry*, Birkhäuser Boston, Inc., Boston, MA, 1992.

[DL10] H. Duan and Y. Long, *The index growth and multiplicity of closed geodesics*, J. Funct. Anal. **259** (2010), no. 7, 1850–1913.

[DPMMS20] G. De Philippis, M. Marini, M. Mazzucchelli, and S. Suhr, *Closed geodesics on reversible Finsler 2-spheres*, arXiv:2002.00415, 02 2020.

[EH87] I. Ekeland and H. Hofer, *Convex Hamiltonian energy surfaces and their periodic trajectories*, Comm. Math. Phys. **113** (1987), no. 3, 419–469.

[EH89] ———, *Symplectic topology and Hamiltonian dynamics*, Math. Z. **200** (1989), no. 3, 355–378.

[EH90] ———, *Symplectic topology and Hamiltonian dynamics. II*, Math. Z. **203** (1990), no. 4, 553–567.

[Eke90] I. Ekeland, *Convexity methods in Hamiltonian mechanics*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 19, Springer-Verlag, Berlin, 1990.

[EL80] I. Ekeland and J.-M. Lasry, *On the number of periodic trajectories for a Hamiltonian flow on a convex energy surface*, Ann. of Math. (2) **112** (1980), no. 2, 283–319.

[Fat08] A. Fathi, *Weak KAM theorem in Lagrangian dynamics*, Cambridge Univ. Press, forthcoming, preliminary version number 10, 2008.

[FLS15] U. Frauenfelder, C. Labrousse, and F. Schlenk, *Slow volume growth for Reeb flows on spherizations and contact Bott-Samelson theorems*, J. Topol. Anal. **7** (2015), no. 3, 407–451.

[FLS16] U. Frauenfelder, C. Lange, and S. Suhr, *A Hamiltonian version of a result of Gromoll and Grove*, arXiv:1603.05107, 2016.

[FR78] E. R. Fadell and P. H. Rabinowitz, *Generalized cohomological index theories for Lie group actions with an application to bifurcation questions for Hamiltonian systems*, Invent. Math. **45** (1978), no. 2, 139–174.

[Fra92] J. Franks, *Geodesics on S^2 and periodic points of annulus homeomorphisms*, Invent. Math. **108** (1992), no. 2, 403–418.

[FS07] U. Frauenfelder and F. Schlenk, *Hamiltonian dynamics on convex symplectic manifolds*, Israel J. Math. **159** (2007), 1–56.

[FT02] M. Farber and S. Tabachnikov, *Topology of cyclic configuration spaces and periodic trajectories of multi-dimensional billiards*, Topology **41** (2002), no. 3, 553–589.

[Gei08] H. Geiges, *An introduction to contact topology*, Cambridge Studies in Advanced Mathematics, vol. 109, Cambridge University Press, Cambridge, 2008.

[GG81] D. Gromoll and K. Grove, *On metrics on S^2 all of whose geodesics are closed*, Invent. Math. **65** (1981), 175–177.

[GGM19] V. L. Ginzburg, B. Z. Gürel, and M. Mazzucchelli, *On the spectral characterization of Besse and Zoll Reeb flows*, to appear in Ann. Inst. H. Poincaré Anal. Non Linéaire, https://doi.org/10.1016/j.anihpc.2020.08.004, 2019.

[GL18] H. Geiges and C. Lange, *Seifert fibrations of lens spaces*, Abh. Math. Semin. Univ. Hambg. **88** (2018), no. 1, 1–22.

[GM69a] D. Gromoll and W. Meyer, *On differentiable functions with isolated critical points*, Topology **8** (1969), 361–369.

[GM69b] ———, *Periodic geodesics on compact Riemannian manifolds*, J. Differ. Geom. **3** (1969), 493–510.

[Gra89] M. A. Grayson, *Shortening embedded curves*, Ann. Math. **129** (1989), 71–111.

[Gro73] K. Grove, *Condition (C) for the energy integral on certain path spaces and applications to the theory of geodesics*, J. Differ. Geom. **8** (1973), 207–223.

[Gro74] ———, *Isometry-invariant geodesics*, Topology **13** (1974), 281–292.

[Gro81] M. Gromov, *Structures métriques pour les variétés riemanniennes*, Textes Mathématiques, vol. 1, CEDIC, Paris, 1981.

[Gro85a] ———, *Pseudoholomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), no. 2, 307–347.

[Gro85b] K. Grove, *The isometry-invariant geodesics problem: Closed and open*, Geometry and Topology, Lecture Notes in Mathematics, vol. 1167, Springer Berlin Heidelberg, 1985, pp. 125–140.

[GT76] K. Grove and M. Tanaka, *On the number of invariant closed geodesics*, Bull. Amer. Math. Soc. **82** (1976), no. 3, 497–498.

[GT78] ———, *On the number of invariant closed geodesics*, Acta Math. **140** (1978), no. 1-2, 33–48.

[Gui76] V. Guillemin, *The Radon transform on Zoll surfaces*, Advances in Math. **22** (1976), no. 1, 85–119.

[Hed32] G. A. Hedlund, *Geodesics on a two-dimensional Riemannian manifold with periodic coefficients*, Ann. of Math. (2) **33** (1932), no. 4, 719–739.

[Hin84] N. Hingston, *Equivariant Morse theory and closed geodesics*, J. Differential Geom. **19** (1984), no. 1, 85–116.

[Hin93] ———, *On the growth of the number of closed geodesics on the two-sphere*, Internat. Math. Res. Notices (1993), no. 9, 253–262.

[Hin09] ———, *Subharmonic solutions of Hamiltonian equations on tori*, Ann. Math. **170** (2009), no. 2, 529–560.

[HS94] J. Hass and P. Scott, *Shortening curves on surfaces*, Topology **33** (1994), no. 1, 25–43.

[Hut14] M. Hutchings, *Lecture notes on embedded contact homology*, Contact and symplectic topology, 389–484, Bolyai Soc. Math. Stud., 26, János Bolyai Math. Soc., Budapest, 2014.

[HWZ98] H. Hofer, K. Wysocki, and E. Zehnder, *The dynamics on three-dimensional strictly convex energy surfaces*, Ann. of Math. (2) **148** (1998), no. 1, 197–289.

[HZ94] H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser Advanced Texts: Basler Lehrbücher, Birkhäuser Verlag, Basel, 1994.

[Iri12] K. Irie, *Symplectic capacity and short periodic billiard trajectory*, Math. Z. **272** (2012), no. 3-4, 1291–1320.

[Jos89] J. Jost, *A nonparametric proof of the theorem of Lusternik and Schnirelman*, Arch. Math. (Basel) **53** (1989), no. 5, 497–509.

[Jos91] ———, *Correction to: "A nonparametric proof of the theorem of Lusternik and Schnirelman"*, Arch. Math. (Basel) **56** (1991), no. 6, 624.

[Kat73] A. B. Katok, *Ergodic perturbations of degenerate integrable Hamiltonian systems*, Izv. Akad. Nauk SSSR Ser. Mat. **37** (1973), 539–576.

[Kli78] W. Klingenberg, *Lectures on closed geodesics*, Springer-Verlag, Berlin, 1978, Grundlehren der Mathematischen Wissenschaften, Vol. 230.

[Kob95] S. Kobayashi, *Transformation groups in differential geometry*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1972 edition.

[LD09] Y. Long and H. Duan, *Multiple closed geodesics on 3-spheres*, Adv. Math. **221** (2009), no. 6, 1757–1803.

[LF51] L. A. Lyusternik and A. I. Fet, *Variational problems on closed manifolds*, Doklady Akad. Nauk SSSR (N.S.) **81** (1951), 17–18.

[Lju66] L. A. Ljusternik, *The topology of the calculus of variations in the large*, Translated from the Russian by J. M. Danskin. Translations of Mathematical Monographs, Vol. 16, American Mathematical Society, Providence, R.I., 1966.

[Lon00] Y. Long, *Multiple periodic points of the Poincaré map of Lagrangian systems on tori*, Math. Z. **233** (2000), no. 3, 443–470.

[Lon02] ———, *Index theory for symplectic paths with applications*, Progress in Mathematics, vol. 207, Birkhäuser Verlag, Basel, 2002.

[LS34] L. Lusternik and L. Schnirelmann, *Méthodes topologiques dans les problèmes variationnels*, Hermann, Paris, 1934.

[LS17] S. Lin and B. Schmidt, *Real projective spaces with all geodesics closed*, Geom. Funct. Anal. **27** (2017), no. 3, 631–636.

[Lu09] G. Lu, *The Conley conjecture for Hamiltonian systems on the cotangent bundle and its analogue for Lagrangian systems*, J. Funct. Anal. **256** (2009), no. 9, 2967–3034.

[Mañ97] R. Mañé, *Lagrangian flows: the dynamics of globally minimizing orbits*, Bol. Soc. Brasil. Mat. (N.S.) **28** (1997), no. 2, 141–153.

[Mat91] J. N. Mather, *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Z. **207** (1991), no. 2, 169–207.

[Maz11a] M. Mazzucchelli, *The Lagrangian Conley conjecture*, Comment. Math. Helv. **86** (2011), no. 1, 189–246.

[Maz11b] ———, *On the multiplicity of non-iterated periodic billiard trajectories*, Pacific J. Math. **252** (2011), no. 1, 181–205.

[Maz12] ———, *Critical point theory for Lagrangian systems*, Progress in Mathematics, vol. 293, Birkhäuser/Springer Basel AG, Basel, 2012.

[Maz14] ———, *On the multiplicity of isometry-invariant geodesics on product manifolds*, Algebr. Geom. Topol. **14** (2014), no. 1, 135–156.

[Maz15] ———, *Isometry-invariant geodesics and the fundamental group*, Math. Ann. **362** (2015), no. 1-2, 265–280.

[Maz16] ———, *The Morse index of Chaperon's generating families*, Publ. Mat. Urug. **16** (2016), 81–125.

[Mil63] J. Milnor, *Morse theory*, Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51, Princeton University Press, Princeton, N.J., 1963.

[MM17] L. Macarini and M. Mazzucchelli, *Isometry-invariant geodesics and the fundamental group, II*, Adv. Math. **308** (2017), 671–698.

[MN18] M. Meiwes and K. Naef, *Translated points on hypertight contact manifolds*, J. Topol. Anal. **10** (2018), no. 2, 289–322.

[Mor96] M. Morse, *The calculus of variations in the large*, American Mathematical Society Colloquium Publications, vol. 18, American Mathematical Society, Providence, RI, 1996, Reprint of the 1932 original.

[MR20] M. Mazzucchelli and M. Radeschi, *On the structure of besse convex contact spheres*, arXiv:2012.05389, 12 2020.

[MS18] M. Mazzucchelli and S. Suhr, *A characterization of Zoll Riemannian metrics on the 2-sphere*, Bull. Lond. Math. Soc. **50** (2018), no. 6, 997–1006.

[Oak94] J. A. Oaks, *Singularities and self-intersections of curves evolving on surfaces*, Indiana Univ. Math. J. **43** (1994), no. 3, 959–981.

[Oan15] A. Oancea, *Morse theory, closed geodesics, and the homology of free loop spaces*, Free loop spaces in geometry and topology, IRMA Lect. Math. Theor. Phys., vol. 24, Eur. Math. Soc., Zürich, 2015, With an appendix by Umberto Hryniewicz, pp. 67–109.

[Ota90a] J.-P. Otal, *Le spectre marqué des longueurs des surfaces à courbure négative*, Ann. of Math. (2) **131** (1990), no. 1, 151–162.

[Ota90b] ———, *Sur les longueurs des géodésiques d'une métrique à courbure négative dans le disque*, Comment. Math. Helv. **65** (1990), no. 2, 334–347.

[OVZ67] P. Orlik, E. Vogt, and H. Zieschang, *Zur Topologie gefaserter dreidimensionaler Mannigfaltigkeiten*, Topology **6** (1967), 49–64.

[Pal63] R. S. Palais, *Morse theory on Hilbert manifolds*, Topology **2** (1963), 299–340.

[Poi05] H. Poincaré, *Sur les lignes géodésiques des surfaces convexes*, Trans. Amer. Math. Soc. **6** (1905), 237–274.

[Pri09] C. Pries, *Geodesics closed on the projective plane*, Geom. Funct. Anal. **18** (2009), no. 5, 1774–1785.

[PS64] R. S. Palais and S. Smale, *A generalized Morse theory*, Bull. Amer. Math. Soc. **70** (1964), 165–172.

[Rad89] H.-B. Rademacher, *On the average indices of closed geodesics*, J. Differential Geom. **29** (1989), no. 1, 65–83.

[Rad92] ———, *Morse-Theorie und geschlossene Geodätische*, Bonner Mathematische Schriften, vol. 229, Universität Bonn, Mathematisches Institut, Bonn, 1992, Habilitationsschrift, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, 1991.

[Rad94] ———, *On a generic property of geodesic flows*, Math. Ann. **298** (1994), no. 1, 101–116.

[RW17] M. Radeschi and B. Wilking, *On the Berger conjecture for manifolds all of whose geodesics are closed*, Invent. Math. **210** (2017), 911–962.

[Sal99] D. Salamon, *Lectures on Floer homology*, Symplectic geometry and topology (Park City, UT, 1997), IAS/Park City Math. Ser., vol. 7, Amer. Math. Soc., Providence, RI, 1999, pp. 143–229.

[Sam63] H. Samelson, *On manifolds with many closed geodesics*, Portugal. Math. **22** (1963), 193–196.

[San10] S. Sandon, *An integer-valued bi-invariant metric on the group of contactomorphisms of* $\mathbb{R}^{2n} \times S^1$, J. Topol. Anal. **2** (2010), no. 3, 327–339.

[San12] ———, *On iterated translated points for contactomorphisms of* $\mathbb{R}^{2n+1}$ *and* $\mathbb{R}^{2n} \times S^1$, Internat. J. Math. **23** (2012), no. 2, 1250042, 14 pp.

[San13] ———, *A Morse estimate for translated points of contactomorphisms of spheres and projective spaces*, Geom. Dedicata **165** (2013), 95–110.

[Sch82] R. Schultz, *Differentiability and the P. A. Smith theorems for spheres. I. Actions of prime order groups*, Current trends in algebraic topology, Part 2 (London, Ont., 1981), CMS Conf. Proc., vol. 2, Amer. Math. Soc., Providence, R.I., 1982, pp. 235–273.

[Sch06] F. Schlenk, *Applications of Hofer's geometry to Hamiltonian dynamics*, Comment. Math. Helv. **81** (2006), no. 1, 105–121.

[Sor15] A. Sorrentino, *Action-minimizing methods in Hamiltonian dynamics: an introduction to Aubry-Mather theory*, Mathematical Notes, vol. 50, Princeton University Press, Princeton, NJ, 2015.

[Str90] M. Struwe, *Existence of periodic solutions of Hamiltonian systems on almost every energy surface*, Bol. Soc. Brasil. Mat. (N.S.) **20** (1990), no. 2, 49–58.

[Sul76] D. Sullivan, *A counterexample to the periodic orbit conjecture*, Inst. Hautes Études Sci. Publ. Math. (1976), no. 46, 5–14.

[Tab05] S. Tabachnikov, *Geometry and billiards*, Student Mathematical Library, vol. 30, American Mathematical Society, Providence, RI; Mathematics Advanced Study Semesters, University Park, PA, 2005.

[Tai91] I. A. Taimanov, *Non-self-itersecting closed extremals of multivalued or not everywhere positive functionals*, Izv. Akad. Nauk SSSR Ser. Mat. **55** (1991), no. 2, 367–383.

[Tai92a] ———, *Closed extremals on two-dimensional manifolds*, Uspekhi Mat. Nauk **47** (1992), no. 2(284), 143–185, 223.

[Tai92b] ———, *Closed non-self-intersecting extremals of multivalued functionals*, Sibirsk. Mat. Zh. **33** (1992), no. 4, 155–162, 223.

[Tai10a] ———, *Periodic magnetic geodesics on almost every energy level via variational methods*, Regul. Chaotic Dyn. **15** (2010), no. 4-5, 598–605.

[Tai10b] ———, *The type numbers of closed geodesics*, Regul. Chaotic Dyn. **15** (2010), no. 1, 84–100.

[Tau95] C. H. Taubes, *The Seiberg-Witten and Gromov invariants*, Math. Res. Lett. **2** (1995), no. 2, 221–238.

[Tau98] ———, *The structure of pseudo-holomorphic subvarieties for a degenerate almost complex structure and symplectic form on $S^1 \times B^3$*, Geom. Topol. **2** (1998), 221–332.

[Tau07] ———, *The Seiberg-Witten equations and the Weinstein conjecture*, Geom. Topol. **11** (2007), 2117–2202.

[Tau09] ———, *The Seiberg-Witten equations and the Weinstein conjecture. II. More closed integral curves of the Reeb vector field*, Geom. Topol. **13** (2009), no. 3, 1337–1417.

[Tho76] C. B. Thomas, *Almost regular contact manifolds*, J. Differential Geometry **11** (1976), no. 4, 521–533.

[Tho78] G. Thorbergsson, *Closed geodesics on non-compact Riemannian manifolds*, Math. Z. **159** (1978), no. 3, 249–258.

[Vit88] C. Viterbo, *Indice de Morse des points critiques obtenus par minimax*, Ann. Inst. H. Poincaré Anal. Non Linéaire **5** (1988), no. 3, 221–225.

[Vit00] ———, *Metric and isoperimetric problems in symplectic geometry*, J. Amer. Math. Soc. **13** (2000), no. 2, 411–431.

[VPS76] M. Vigué-Poirrier and D. Sullivan, *The homology theory of the closed geodesic problem*, J. Differential Geometry **11** (1976), no. 4, 633–644.

[Wad75] A. W. Wadsley, *Geodesic foliations by circles*, J. Differ. Geom. **10** (1975), no. 4, 541–549.

[Zil83] W. Ziller, *Geometry of the Katok examples*, Ergodic Theory Dynam. Systems **3** (1983), no. 1, 135–157.

[Zol03] O. Zoll, *Über Flächen mit Scharen geschlossener geodätischer Linien*, Math. Ann. **57** (1903), no. 1, 108–133.

Index