# LECTURE NOTES ON BIRKHOFF BILLIARDS: DYNAMICS, INTEGRABILITY AND SPECTRAL RIGIDITY

CORENTIN FIEROBE, VADIM KALOSHIN, AND ALFONSO SORRENTINO

## Contents

1.	Lecture I: Billiard dynamics	1
2.	Lecture II: Variational principle and periodic orbits.	8
3.	Lecture III: Caustics, invariant curves and integrability	17
4.	Lecture IV: Aubry-Mather theory and billiard dynamics	33
References		38

#### 1. LECTURE I: BILLIARD DYNAMICS

In these lecture notes we would like to introduce and investigate an interesting class of dynamical systems, known as *mathematical billiards*. Billiard is a generic term to refer to a very wide range of dynamical models with impacts; we refer the interested reader to [69, 95, 101, 102] for a more exhaustive presentation. Billiard-like models have been capturing the attention of researchers in various areas of mathematics for many years: not only their law of motion is very physical and intuitive, but billiard-type dynamics is ubiquitous. Mathematically, according to the shape of the billiard, they offer models in every subclass of dynamical systems (integrable, regular, chaotic, etc.). Moreover, thanks to their manifold nature, they provide a fruitful laboratory where different ideas and expertises (from dynamical systems, analysis, geometry, etc...), have the possibility to interface and beneficially integrate.

In these notes, we will be mostly interested in the study of the so-called *Birkhoff billiards* (see Definition 1.2). Such billiards are designed to describe the trajectory of a ray of light evolving inside a homogeneous convex cavity bounded by perfectly reflecting mirrors. From a mechanical point of view, if we consider a convex bounded domain  $\Omega \subset \mathbb{R}^d$ , where  $d \geq 2$ , with a  $C^1$ -smooth boundary  $\partial\Omega$ , we can think of a massless ball moving with unit velocity and without friction following a geodesic path: when the ball hits the boundary, it reflects elastically according to the standard *reflection law*, namely the angle of reflection is equal to the angle of incidence<sup>1</sup>; see Figure 1.

The dynamics of the ball (or of the ray of light) evolving in  $\Omega$  can be described by two approaches, a time-continuous one (the *billiard flow*), and a time-discrete one (the *billiard map*).

<sup>&</sup>lt;sup>1</sup>One could also assume that the boundary  $\partial \Omega$  has some point of non-differentiability; in this case, when the ball hits the boundary at one of these points (a sort of "holes"), then the motion stops.



FIGURE 1. Two consecutive reflections of a billiard trajectory in a strictly convex planar domain  $\Omega$  with smooth boundary  $\partial \Omega$ .

- The billiard flow: a billiard trajectory consists of the path followed by a ray of light evolving inside  $\Omega$ , and bouncing on its boundary  $\partial\Omega$  according to the law of reflection of optics, that is angle of incidence = angle of reflection. This first continuous approach can be translated in the existence of a quantity  $B^t(p, v)$ , where p is a point in the domain  $\Omega$ , v is a non-zero vector and  $t \in \mathbb{R}$  represents a time variable, giving the position at time t of the ray of light emitted from p with speed v. The map  $(p, v, t) \mapsto B^t(p, v)$  is called the billiard flow <sup>2</sup>.
- The billiard map: complementary to the billiard flow, which provides a continuous model for the billiard, there is a discrete way to describe the billiard dynamics, namely using the so-called billiard map. Roughly speaking, we keep track only of bouncing points on the boundary and their bouncing direction. More specifically, a ray of light can be modeled by an oriented line intersecting  $\Omega$ . The set of all such oriented lines defines the phase space  $\mathcal{L}$  of the billiard map. The reflection of an oriented line is given by a transformation  $T : \mathcal{L} \to \mathcal{L}$  which associates to an oriented line  $\ell \in \mathcal{L}$ , the oriented line  $\ell' = T(\ell)$  which corresponds to the trajectory of the light after being reflected from  $\ell$ . The map T is called the billiard map associated to  $\Omega$ .

**Remark 1.1.** In the above models, one can easily replace  $\mathbb{R}^d$  by any manifold with a complete Riemannian metric g and oriented lines by oriented geodesics of g. This allows one to define a notion of billiard in other geometrical contexts. The reader can consult [102] for more details.

1.1. **Birkhoff billiards.** In the following, we will be interested in the time-discrete point of view. Moreover, we assume that the billiard domain satisfies some (strict) convexity assumption. Although in these lecture notes we will focus on planar billiards, we present some of the first properties for billiards in any dimension.

**Definition 1.2.** A Birkhoff billiard is a triple  $(\Omega, \mathcal{L}, T)$ , where  $\Omega$  is a strictly convex domain of  $\mathbb{R}^d$ , and  $\mathcal{L}$  and T are the associated phase-space and billiard map. A billiard orbit is a sequence  $(T^n(\ell))_{n \in \mathbb{Z}}$  for a given  $\ell \in \mathcal{L}$ .

The unit-bundle model. A way to define the phase space of a Birkhoff billiard  $\Omega$  is to consider the set  $\Phi$  of pairs (p, v) where  $p \in \partial \Omega$  lies on the boundary of  $\Omega$  and v is a unit vector

<sup>&</sup>lt;sup>2</sup>The billiard flow might not be defined for any time t > 0, as it was shown by Halpern [50]: there exist Birkhoff billiards of  $C^2$ -smooth boundary such that the billiard flow is not defined for all times. However, when the boundary is  $C^3$ -smooth this phenomenon cannot occur.

pointing inside  $\Omega$  when we choose p as its origin. Indeed, any oriented line intersecting  $\Omega$  is uniquely defined by such a pair.

The billiard map can be then defined as the map  $F : \Phi \to \Phi$  associating to a pair  $(p, v) \in \Phi$  a new pair  $F(p, v) = (p', v') \in \Phi$  such that the ray of light emitted from (p, v) is reflected by the boundary as a ray emitted from (p', v').

**Proposition 1.3.** Suppose that the boundary of  $\Omega$  is  $C^r$ -smooth,  $r \geq 2$ . Then  $\Phi$  can be endowed with a structure of  $C^{r-1}$ -smooth manifold which is 2(d-1)-dimensional, and  $F : \Phi \to \Phi$  is a  $C^{r-1}$ -smooth diffeomorphism.

*Proof.* (See also [66, Theorem 4.2 in Part V]) Fix  $p \in \partial\Omega$ . Since the boundary is  $C^r$ -smooth, one can find a  $C^r$ -smooth local isometry  $j: U \to V \subset \partial\Omega$  between an open set  $U \subset \mathbb{R}^{d-1}$  and an open set  $V \subset \partial\Omega$  containing p. For any point  $q \in V$ , one can define

- a unit normal vector N(q) pointing inside  $\Omega$  when we fix its origin at q;

- the tangent space of  $\partial \Omega$  at q = j(x) by  $T_q \Omega = dj(x) \left( \mathbb{R}^{d-1} \right)$ .

Consider the set

$$B(0,1) = \{ r \in \mathbb{R}^{d-1} \mid ||r|| < 1 \}$$

of vectors of  $\mathbb{R}^{d-1}$  with Euclidean norm less than 1. Given  $v \in B(0,1)$  and  $q = j(x) \in V$ , one can define the unit vector of  $\mathbb{R}^d$ 

$$V(x,r) := dj(x) \cdot r + \sqrt{1 - \|dj(x) \cdot r\|^2} N(x).$$

The reader can check that V(x,r) is the unit vector pointing inside  $\partial\Omega$  when we choose its origin at q, and projecting onto  $T_q\partial\Omega$  with the component  $dj(x) \cdot r$ . Note that dj(x) is a linear isometry between B(0,1) and the unit open ball of  $T_q\partial\Omega$ . Hence, the  $C^{r-1}$ -smooth, injective map given by

$$(x,r) \in U \times B(0,1) \mapsto (j(x),V(x,r))$$

defines a chart on  $\Phi$ .

The fact that the billiard map is  $C^{r-1}$ -smooth comes from the implicit function theorem: consider two distinct points  $p, p' \in \partial \Omega$  and a vector  $v \in T_p\Omega$  of norm < 1. The map Gdefined by

$$G(p, v, p') = \frac{p' - p}{\|p' - p\|} - \left\langle \frac{p' - p}{\|p' - p\|} \,|\, N(p) \right\rangle N(p) - v$$

is  $C^{r-1}$ -smooth and vanishes if and only if p' is the point of impact of a ray emitted from p with unit vector given by  $V_{v,\nu(p)}$ . Indeed, the two first terms in G define the projection of the unit vector  $\frac{p'-p}{\|p'-p\|}$  on the tangent space  $T_p\partial\Omega$ . Hence it is enough to show that locally, there is a  $C^{r-1}$  smooth map g in the variables (p, v) such that

$$G(p, v, p') = 0 \quad \Leftrightarrow \quad p' = g(p, v).$$

By the implicit function theorem, it is enough to check that  $\tilde{G} : q \in \partial\Omega \mapsto G(p, v, q) \in T_p \partial\Omega$ is a local diffeomorphism at q = p'. We let the reader check this match by computing the differential of  $\tilde{G}$  at p'.

Hence we have shown that if the billiard map writes as  $(p, v) \mapsto (p', v')$ , then p' = p'(p, v) is  $C^{r-1}$  smooth. Now v' = v'(p, v) is also  $C^{r-1}$ -smooth since it is obtained by projecting  $\frac{p'-p}{\|p'-p\|}$  orthogonally to  $T_p \partial \Omega$ .

**Exercise 1.4.** Let  $i : \Phi \to \mathcal{L}$  be the map associating to a pair (p, v) the oriented line  $\ell$  containing p and directed by v. Show that i is a bijection satisfying  $i \circ F = T \circ i$ .

The cylinder model in dimension 2. From now we shall consider planar Bikrhoff billiards, namely we assume that d = 2, and  $\Omega \subset \mathbb{R}^2$  is a strictly convex bounded domain with with a  $C^r$ -smooth boundary,  $r \geq 3$ . The boundary  $\partial\Omega$  is a closed curve which can be parametrized by an arc-length coordinate s, viewed modulo the perimeter of the boundary  $L = |\partial\Omega|$ . In this case any pair  $(p, v) \in \Phi$  in the unit bundle model can be encoded by a unique pair  $(s, \varphi) \in \mathbb{R}/L\mathbb{Z} \times (0, \pi)$  where s is the arc-length coordinate of p and  $\varphi$  is the angle made by v with the tangent line of  $\partial\Omega$  at p oriented according to the parametrization of the boundary. The phase-space in this case is the annulus (or cylinder)

$$\mathbb{A}_L := \mathbb{R}/L\mathbb{Z} \times (0,\pi).$$

The billiard map is the one which naturally arises from this construction, acting therefore on  $\mathbb{A}_L$ ; see Figure 2.



FIGURE 2. The phase space  $\mathbb{A}_L$  in dimension 2. On the left, its representation as a cylinder; on the right, its representation as an affine chart on  $\mathbb{R}/L\mathbb{Z} \times (0, \pi)$ .

Let us start with an example.

**Example: billiard in the unit disk.** We consider the billiard in the unit disk  $\mathbb{D}^1 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ , whose boundary has perimeter  $L = 2\pi$  and we assume to be parametrized by arc-length. We denote by  $\theta$  the length of the arc joining a point  $p_{\theta}$  from a fixed origin  $p_0 \in \partial \mathbb{D}^1 = \mathbb{S}^1$  in a given orientation (we count  $\theta$  modulo  $2\pi$ ). Given an oriented line  $\ell$  intersecting  $\partial \mathbb{D}^1$  at a point p, we denote by  $\varphi$  the oriented angle between the tangent line  $T_p \partial \mathbb{D}^1$  and  $\ell$ . See Figure 3.

**Proposition 1.5.** In  $(\theta, \varphi)$ -coordinates, the billiard map in  $\mathbb{D}^1$  is the map  $F : \mathbb{A}_{2\pi} \to \mathbb{A}_{2\pi}$  given by

(1) 
$$F(\theta,\varphi) = (\theta + 2\varphi,\varphi).$$

*Proof.* Let us write  $(\theta', \varphi') := F(\theta, \varphi)$ . We consider the following geometric configuration. Denote by O the center of  $\mathbb{D}^1$  and consider an oriented line intersecting  $\partial \mathbb{D}^1$  at a point  $p = p_{\theta}$ and such that the oriented angle between the tangent line  $T_p \partial \mathbb{D}^1$  and  $\ell$  is  $\varphi \in (0, \pi)$ . Let  $p' = p_{\theta'}$  be the second point of intersection of  $\ell$  with  $\partial \mathbb{D}^1$ . See Figure 3.

By definition of  $\mathbb{D}^1$ , the triangle Opp' is isosceles in O and the lines Op and Op' intersect  $\mathbb{D}^1$  orthogonally. This allows to deduce the following statements:

- The oriented angle between the tangent line  $T_{p'}\partial \mathbb{D}^1$  and  $\ell$  is  $-\varphi$ .
- The oriented angle between Op and Op' is  $2\varphi$ .

Hence  $\theta' = \theta + 2\varphi$  and  $\ell$  is reflected at p' into an oriented line making an angle  $\varphi$  with the tangent line  $T_{p'}\partial \mathbb{D}^1$ , *i.e.*,  $\varphi' = \varphi$ .



FIGURE 3. The billiard reflection in a disk: two successive impact points p and p', such that the line pp' makes an angle  $\varphi$  with the boundary at p.

Periodic and dense orbits. In particular, Propositin 1.5 implies that  $\varphi$  stays constant along the orbit (*i.e.*, it represents what is called an *integral of motion* for the map); hence, the property of the orbits are determined by the corresponding angle  $\varphi = \pi \omega$ , with  $\omega \in (0, 1)$ ;  $\omega$  is called the *rotation number* of the orbit. Then:

• If  $\omega = \frac{m}{n} \in (0,1) \cap \mathbb{Q}$ , for two coprime integers m, n > 0, then for any  $\theta \in \mathbb{R}$ , by induction F satisfies

$$F^n(\theta,\varphi) = (\theta + 2n\varphi,\varphi) = (\theta + 2\pi m,\varphi) = (\theta,\varphi) \quad \text{mod. } 2\pi\mathbb{Z}.$$

Hence, the orbit is periodic with minimal period n. We say that  $(\theta, \varphi)$  is a periodic point of rotation number m/n. In fact it corresponds to a polygonal periodic trajectory (if m = 1) or a star-shaped one (if m > 1), with n denoting its period, and m the number of times that the trajectory winds around the boundary  $\partial \mathbb{D}^1$  before closing (see Figure 4).



FIGURE 4. Two periodic trajectories of period 5 in a disk, with rotation numbers 1/5 (on the left) and 2/5 (on the right).

• If  $\omega \in (0,1) \setminus \mathbb{Q}$ , then the orbit is not periodic and it hits the boundary  $\partial \mathbb{D}^1$  on a dense set of points (by Kroenecker's theorem).

Invariant curves. By Equation 1, we observe that given a fixed  $\varphi_0 \in (0, \pi)$ , the curve  $\mathcal{C}_{\varphi_0} := \{(\theta, \varphi) \in \mathbb{A}_{2\pi} \mid \varphi = \varphi_0\}$  is invariant by F, that is  $F(\mathcal{C}_{\varphi_0}) = \mathcal{C}_{\varphi_0}$ . Moreover, the set of all possible  $\mathcal{C}_{\varphi_0}$ , for  $\varphi_0 \in (0, \pi)$ , foliates the annulus  $\mathbb{A}_{2\pi}$  (see Figure 5).



FIGURE 5. The phase space  $\mathbb{A}_{2\pi}$  of the billiard map in a disk with its foliation by horizontal invariant curves  $\mathcal{C}_{\varphi}$ , where  $\varphi \in (0, \pi)$ .

**Remark 1.6.** Equation (2) implies that all points in  $C_{\frac{m}{n}\pi}$  are periodic of rotation number m/n (and in fact, all such points lie on  $C_{\frac{m}{n}\pi}$ ). We say that  $\mathbb{D}^1$  has 1-parameter families (i.e., curves in the phase space) of periodic points. In fact, it was proven (see [84]) that this phenomenon is degenerate in the sense that for a generic domain (i.e., an intersection of open and dense set of domains), the set of periodic orbits in this domain of any given period  $n \geq 2$  is finite (see also [4] for some related results for general maps).

1.2. Generating function. Consider  $\gamma : \mathbb{R} \to \mathbb{R}^2$  be a  $C^r$ -smooth *L*-periodic parametrization of the boundary  $\partial \Omega$ . Define the map  $H : \mathbb{R}^2 \to \mathbb{R}^+$  given by

$$H(s, s') = -\operatorname{dist}(\gamma(s), \gamma(s')) \qquad \forall s, s' \in \mathbb{R}.$$

**Definition 1.7.** The map H is called the *generating function* of the billiard in  $\Omega$ .

**Remark 1.8.** The negative sign in the definition of H is just a convention, that will make sense a bit later when we will consider minimizers of H, which otherwise would be maximizers of the distance.

**Exercise 1.9.** Compute H when  $\Omega$  is a unit disk.



FIGURE 6. The generating function encodes the dynamics of the billiard

The function H is called generating because it contains the dynamics of the billiard (see Figure 6), as stated in the following proposition.

**Proposition 1.10.** The map H is  $C^r$ -smooth outside the set  $\Delta = \{(s,s) \mid s \in \mathbb{R}\}$ . Moreover, if  $(s, \varphi), (s', \varphi') \in \mathbb{A}_L$  are such that  $s \neq s'$ , then the following statements are equivalent: (i)  $F(s, \varphi) = (s', \varphi');$ (ii)  $\partial_1 H(s, s') = -\cos \varphi$  and  $\partial_2 H(s, s') = \cos \varphi'$ , where  $\partial_k$  denotes the derivative with respect to the k-th variable (k = 1, 2).

*Proof.* The smoothness condition is obvious since  $\partial \Omega$  is  $C^r$ -smooth. Given  $s \neq \tilde{s}$ , let us compute  $\partial_1 H(s, \tilde{s})$ . Differating the Euclidean norm, we obtain

$$\partial_1 H(s,\tilde{s}) = -\left\langle \frac{\gamma(\tilde{s}) - \gamma(s)}{\|\gamma(\tilde{s}) - \gamma(s)\|} \,|\, \gamma'(s) \right\rangle.$$

Here  $u := \frac{\gamma(\tilde{s}) - \gamma(s)}{\|\gamma(\tilde{s}) - \gamma(s)\|}$  is a unit vector from  $p = \gamma(s)$  to  $\tilde{p} = \gamma(\tilde{s})$  and  $\gamma'(s)$  has unit norm (since  $\gamma$  parametrizes  $\partial\Omega$  with arc-length). It follows that  $\partial_1 H(s, \tilde{s})$  is  $-\cos\varphi$  where  $\varphi$  is the angle between the line  $p\tilde{p}$  and the tangent line to  $\partial\Omega$  at p. Similarly,  $\partial_2 H(s, \tilde{s}) = \cos\tilde{\varphi}$  where  $\tilde{\varphi}$  is the angle between the line  $p\tilde{p}$  and the tangent line to  $\partial\Omega$  at p. From this we obtain  $F(s,\varphi) = (\tilde{s},\tilde{\varphi})$  and hence  $F(s,\varphi) = (s',\varphi')$  if and only if  $\tilde{s} = s'$  and  $\tilde{\varphi} = \varphi'$ , which is tantamount to item (ii).

**Exercise 1.11.** Let  $\Omega$  be a Birkhoff billiard with  $C^2$ -smooth boundary. Let  $\kappa(s)$  be the radius of curvature of  $\partial\Omega$  at a point of arc-length coordinate s.

(i) Show that (see [66, Theorem 4.2 in Part V]) :

$$\begin{cases} \partial_s s' = \frac{\kappa(s)d(s,s') - \sin\varphi}{\sin\varphi'} & \partial_\varphi s' = \frac{d(s,s')}{\sin\varphi'} \\ \partial_s \varphi' = \frac{\kappa(s)\kappa(s')d(s,s') - \kappa(s)\sin\varphi' - \kappa(s')\sin\varphi}{\sin\varphi'} & \partial_\varphi \varphi' = \frac{\kappa(s')d(s,s') - \sin\varphi'}{\sin\varphi'} \end{cases}$$

where d(s, s') denotes the distance in the plane between the points on the boundary of the billiards corresponding to s and s'.

(ii) Assume that  $\Omega$  is a Birkhoff billiard with  $C^4$ -smooth boundary. Show that when  $(s'-s) \rightarrow 0$ , the generating function H associated to  $\Omega$  can be written as

$$H(s,s') = -|s'-s| \left( 1 - \frac{\kappa(s)^2}{24} (s'-s)^2 - \frac{\kappa(s)\kappa'(s)}{24} (s'-s)^3 + \mathcal{O}\left( (s'-s)^4 \right) \right).$$

Moreover, for small  $\varphi$ , the differential of the billiard map has the following form

$$Df(s,\varphi) = L(s) + \varphi A(s) + O(\varphi^2)$$

where

$$L(s) := \left(\begin{array}{cc} 1 & \frac{2}{\kappa(s)} \\ 0 & 1 \end{array}\right)$$

and

$$A(s) := \begin{pmatrix} -2\frac{\kappa'(s)}{\kappa^2(s)} & -\frac{8}{3}\frac{\kappa'(s)}{\kappa^3(s)} \\ 0 & \frac{4}{3}\frac{\kappa'(s)}{\kappa^2(s)} \end{pmatrix}.$$

**Area-preservation.** Consider the 1-form  $\lambda$  on  $\mathbb{A}_L$  defined by  $\lambda = \cos \varphi \, ds$ . The exterior differential of  $\lambda$  is the area form

$$\omega = d\lambda = \sin\varphi \, ds \wedge d\varphi.$$

Proposition 1.10 has the following consequence. We recall that given a k-form  $\alpha$  on a manifold M, and a smooth map  $F: M \to M$ , we define  $\alpha$ 's pullback by F as the k-form  $\beta := F^* \alpha$  such that

$$\beta_p(v_1,\ldots,v_k) = \alpha_{F(p)}(dF_p(v_1),\ldots,dF_p(v_k)) \qquad \forall p \in M \quad \forall v_1,\ldots,v_k.$$

**Exercise 1.12** (Exact symplecticity). Prove that

$$F^*\lambda - \lambda = dH$$

(F is a so-called exact-symplectic map). Deduce that F preserves  $\omega$ , i.e.,  $F^*\omega = \omega$ .

This suggests the following change of coordinates:

$$\begin{cases} x := s \\ y := -\cos\varphi. \end{cases}$$

Notice that

$$\omega = ds \wedge d(-\cos\varphi) = dx \wedge dy.$$

The billiard map F in these new coordinates is given by

 $f: \mathbb{R}/L\mathbb{Z} \times (-1,1) \to \mathbb{R}/L\mathbb{Z} \times (-1,1),$ 

defined by the condition  $f(s, -\cos \varphi) = (s', -\cos(\varphi'))$  where  $(s', \varphi') = F(s, \varphi)$  for all  $(s, \varphi) \in \mathbb{A}_L$ ; by by construction f is conjugated to F. From Exercise 1.12 we can deduce the following proposition.

**Proposition 1.13.** The billiard map f is area-preserving, i.e. it preserves the area form  $dx \wedge dy$ .

Billiard maps are examples of so-called *exact-symplectic twist maps*; we refer to [43, 95] for more details about these maps. See also section 4 for their relation with the so-called Aubry-Mather theory.

# 2. Lecture II: Variational principle and periodic orbits.

Let  $\Omega \subset \mathbb{R}^2$  be a strictly convex bounded domain with  $C^3$ -smooth boundary and assume that its boundary  $\partial\Omega$  is parametrized by an arc-length coordinate s, viewed modulo the perimeter of the boundary that, without loss of generality, we assume to be equal to 1. We denote by  $\gamma : \mathbb{R} \to \mathbb{R}^2$  this parametrization.

Let us denote as before  $\mathbb{A} := \mathbb{R}/\mathbb{Z} \times (0, \pi)$  and consider the billiard map  $F : \mathbb{A} \longrightarrow \mathbb{A}$  and a lift of F to the universal cover  $\mathbb{R} \times (0, \pi)$  of  $\mathbb{A}$ :

$$\widetilde{F} : \mathbb{R} \times (0, \pi) \longrightarrow \mathbb{R} \times (0, \pi)$$
$$(s, \varphi) \mapsto (s', \varphi').$$

Let us also denote by  $H : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$  its generating function

$$H(s, s') = -\operatorname{dist}(\gamma(s), \gamma(s')) \qquad \forall s, s' \in \mathbb{R}.$$

**Definition 2.1.** An *orbit* of the map  $\widetilde{F}$  is a sequence  $\underline{p} = (s_k, \varphi_k)_{k \in \mathbb{Z}}$  such that for any  $k \in \mathbb{Z}$ ,  $\widetilde{F}(s_k, \varphi_k) = (s_{k+1}, \varphi_{k+1})$ . The *rotation number* of the orbit, if it exists, is defined as the quantity

$$\lim_{k \to +\infty} \frac{s_k - s_0}{k}$$

An orbit is said to be *periodic* if there exist an integer n > 0 and  $m \in \mathbb{Z}$ , such that for any  $k \in \mathbb{Z}$   $s_{k+n} = s_k + m$ . The period of the orbit is the minimal n satisfying this condition.

**Exercise 2.2.** Check that in the case of a periodic orbit, rotation number exists and it is given by the ratio m/n, where n is the period and m the corresponding integer (see definition above).

**Remark 2.3.** The rotation number m/n has a simple geometric interpretation in the case of billiard maps: n is the number of time an orbit bounces on the billiard boundary before repeating itself - or its period, and m is the number of time it winds around the boundary.

**Definition 2.4.** A stationary configuration for  $\widetilde{F}$  is a sequence  $\underline{s} = (s_k)_{k \in \mathbb{Z}} \subset \mathbb{R}^{\mathbb{Z}}$  such that for any  $k \in \mathbb{Z}$ 

$$\partial_2 H(s_{k-1}, s_k) + \partial_1 H(s_k, s_{k+1}) = 0.$$

From Proposition 1.10, we can easily deduce a correspondance between stationary configurations and orbits. More specifically:

**Proposition 2.5.** The map associating to an orbit  $(s_k, \varphi_k)_{k \in \mathbb{Z}}$  the stationary configuration  $(s_k)_{k \in \mathbb{Z}}$  is a bijection between orbits and stationary configurations.

Given a Birkhoff billiard, we can ask about if it has periodic orbits. The answer is in fact yes, and it is given by the famous theorem of Birkhoff (see [16] and also the presentations in [43, 95]).

**Theorem 2.6** (Birkhoff). Let  $\Omega$  be a Birkhoff billiard with  $C^2$ -smooth boundary. Then for any  $m/n \in (0, 1) \cap \mathbb{Q}$  there exist at least two geometrically distinct periodic orbits of rotation number m/n.

*Proof.* The proof of Theorem 2.6 relies on the variational principle given by Proposition 2.5. Fix coprime integers m, n > 0 such that  $m/n \in (0, 1) \cap \mathbb{Q}$ .

The first periodic orbit of rotation number m/n is given by minimizing the functionnal  $h_{m,n}$ :  $K_{m,n} \to \mathbb{R}$  defined on the compact set

$$K_{m,n} = \{(s_0, \dots, s_n) \in \mathbb{R}^{n+1} \mid s_n = s_0 + m, s_0 \in [0, 1] \text{ and } s_k \le s_{k+1} \forall k\}$$

by

$$h_{m,n}(s_0,\ldots,s_n) := \sum_{k=0}^{n-1} H(s_k,s_{k+1}).$$

The function  $h_{m,n}$  is continuous hence reaches its minimal value at a certain  $\underline{s} = (s_0, \ldots, s_n)$ . In fact, one can show that  $\underline{s}$  can be taken in the interior of  $K_{m,n}$ . Indeed, if for some k we have  $s_{k-1} = s_k$ , but  $s_k < s_{k+1}$  then for a fixed  $t \in (s_k, s_{k+1})$  if we denote by

$$\underline{\tilde{s}} = (s_0, \dots, s_{k-1}, t, s_{k+1}, \dots, s_n)$$

we have by construction  $h_{m,n}(\underline{\tilde{s}}) < h_{m,n}(\underline{s})$  because the sum of distances of the chords from  $s_{k-1}$  to t and from t to  $s_{k+1}$  is strictly greater than the distance of the chord from  $s_{k-1} = s_k$  to  $s_{k+1}$  hence

$$H(s_{k-1}, t) + H(t, s_{k+1}) < H(s_{k-1}, s_k) + H(s_k, s_{k+1}).$$

This contradicts the minimality of  $\underline{s}$ . Moreover, adding or substracting 1 to the components of  $\underline{s}$  does not change the minimality, so we can assume  $s_0 \in [0, 1)$ . Now, since  $\underline{s}$  is in the interior of  $K_{m,n}$ , it is a critical point of  $h_{m,n}$ : in particular for any  $k \in \{1, \ldots, n-1\}$ ,

$$0 = \partial_k h_{m,n}(\underline{s}) = \partial_2 H(s_{k-1}, s_k) + \partial_1 H(s_k, s_{k+1})$$

and by Proposition 2.5, this corresponds to a periodic orbit of the billiard map of rotation number m/n.

The other one is obtain by minmax techniques (see figure below for an idea of the so-called mountain pass method) and one needs to show that also this critical configuration must lie in the interior of the set, hence correspond to an orbit of the map.  $\Box$ 



FIGURE 7. Mountainpass critical points

**Remark 2.7.** The lower bound on the number of periodic orbits with a given rational rotation number is optimal. For example, if one takes an ellipse with eccentricity 0 < e < 1, then there are exactly two periodic orbits corresponding, respectively, to the minor and major axis.

**Exercise 2.8.** Let  $\varepsilon > 0$  and consider the domain  $\Omega_{\varepsilon} \subset \mathbb{R}^2$  whose boundary is given in polar coordinates  $(\varphi, r)$  by

 $r = 1 + \varepsilon \cos 3\varphi.$ 

1) Show that there exists  $\varepsilon_0 > 0$  such that for any  $0 \le \varepsilon \le \varepsilon_0$  the domain  $\Omega_{\varepsilon}$  is strictly convex. 2) Consider the points  $p_1, p_2, p_3$  on  $\partial \Omega_{\varepsilon}$  corresponding to  $\varphi = 0, \frac{2\pi}{3}, \frac{4\pi}{3}$ . Show that  $(p_1, p_2, p_3)$  is a 3-periodic orbit. Does this orbit maximize the length functional?

3) Same question as in 2) with the points  $q_1, q_2, q_3$  on  $\partial \Omega_{\varepsilon}$  corresponding to  $\varphi = \frac{\pi}{3}, \pi, \frac{5\pi}{3}$ .

2.1. On the quantity of periodic orbits and Ivrii's conjecture. Describing the set of periodic orbits of a dynamical system like a billiard allows to understand it better. The question of their existence is given by Birkhoff's theorem (Theorem 2.6) which states the existence of periodic orbits of any rotation number. A further question one can ask is about their *quantity*. The following statement was conjectured by Ivrii in [60].

**Ivrii's conjecture.** Given a strictly convex billiard in  $\mathbb{R}^d$ , the set of its periodic orbits has zero measure.

The conjecture can be understood as follows: the billiard map inside a strictly convex domain  $\Omega \subset \mathbb{R}^d$  acts on the pairs (p, v) where p lies on the boundary of  $\Omega$  and v is a unit vector pointing inside  $\Omega$  from p. It can hence be viewed as a diffeomorphism  $F : B(T\partial\Omega) \to B(T\partial\Omega)$  of pairs  $(p, v) \in T\partial\Omega$  such that v is tangent to  $\Omega$  at p and has norm  $\leq 1$ . In this setting, Ivrii's conjecture states that the set of periodic points of F, namely

$$\{(p,v) \in B(T\partial\Omega) \mid \exists n \in \mathbb{Z}^{>0} F^n(p,v) = (p,v)\}$$

has zero measure.

This conjecture is related to the so-called Weyl law related to the Dirichlet eigenvalues of the Laplace-Beltrami operator  $\Delta$  in a bounded convex domain  $\Omega \subset \mathbb{R}^d$ . For  $\lambda \geq 0$ , define  $N(\lambda)$  the number of eigenvalues (with multiplicities) of  $\Delta$  for which the corresponding eigenvector vanish on the boundary  $\partial\Omega$ . Weyl [108] proved the following asymptotics

$$N(\lambda) \sim a_d \operatorname{vol}(\Omega) \lambda^{d/2}$$

where  $a_d > 0$  is a constant depending only on d and conjectured the second order term, so that  $N(\lambda)$  should follow the following asymptotics:

(3) 
$$N(\lambda) = a_d \operatorname{vol}(\Omega) \lambda^{d/2} + a_{d-1} \operatorname{area}(\partial \Omega) \lambda^{\frac{d-1}{2}} + o\left(\lambda^{\frac{d-1}{2}}\right)$$

where  $a_{d-1} \in \mathbb{R}^*$  only depends on d. Ivrii showed that if the set of periodic orbits inside a domain  $\Omega$  has measure zero, then the asymptotics (3) holds.

Ivrii's conjecture remains open until now; however partial answers have been given. A generic positive answer has been given in [84], where it is proven using a transversality theorem that for a generic domain the set of periodic orbits of any given period is finite. Note that in fact

LECTURE NOTES ON BIRKHOFF BILLIARDS: DYNAMICS, INTEGRABILITY AND SPECTRAL RIGIDITY

Ivrii's conjecture holds if and only if it is true for the set of periodic orbits of any given period. Following this idea, [91] and later [99] proved the conjecture for 3-periodic orbits. There result was extended to any dimension in [107]. See also [20, 109] for different proofs of this result which also apply for other geometries. These results were also extended to 4-periodic orbits, see [41, 42], but also [37, 38, 39]. The latter result also applies to billiards in other geometries or with different reflection laws, see for example [20, 34, 35].

This conjecture was proven by [104] under the assumption that the boundary of the domain is analytic. Let us sketch the proof of this result.

**Theorem 2.9** ([104]). Let  $\Omega \subset \mathbb{R}^d$  be a strictly convex domain with analytic boundary. Then the set of periodic orbit for the billiard inside  $\Omega$  has zero measure.

Proof. Let us sketch the proof of this result for d = 2. In this setting, the billiard map F inside  $\Omega$  is an analytic diffeomorphism of the cylinder  $\mathbb{A}$ . If the set of periodic points of F has strictly positive measure, then there should exists an  $n \geq 2$  such that the set of periodic points of F of period n has strictly positive measure. Hence the identity  $F^n = \text{id}$  is satisfied on a set of non zero measure, and by analytic continuation it is satisfied everywhere. But this contradicts the existence of periodic orbits of rotation number 1/q where q is bigger than n.

**Remark 2.10.** The proof of Theorem 2.9 shows that in the analytic setting the existence of a set of non-zero measure of periodic orbits implies the existence of an open set of periodic orbits (here the whole phase space).

**Remark 2.11.** Consider a billiard F in a strictly convex domains. Let  $x_0 = F^q(x_0)$  be a periodic of some period q > 1. Call  $x_0$  absolutely periodic (resp. absolutely periodic of order n) if the differential of  $F^q$  is the identity and all second order and higher partial derivatives of  $F^q$  (resp. of order up to n) at  $x_0$  are 0. In a recent preprint [24], K. Callis showed that for any natural number n, the set of domains containing absolutely periodic orbits of order n are dense in the set of bounded strictly convex domains with  $C^{\infty}$  smooth boundary. This result is a step toward disproving the following conjecture by Safarov-Vassiliev in [94], namely, that no absolutely periodic billiard orbits of infinite order exist in Euclidean billiards. This is also an indication that Ivrii's Conjecture about the measure set of periodic orbits might not be true.

Let us conclude this section by recalling that the set of perimeters of periodic orbits in a strictly convex billiard has zero measure. This is a simple consequence of Sard's lemma and the fact that periodic orbits are critical points of the length functional. This however does not allow to answer Ivrii's conjecture.

**Theorem 2.12** (Birkhoff [16]). Let  $\mathcal{L}(\Omega) \subset \mathbb{R}$  be the set of perimeters of periodic orbits inside a Birkhoff's billiard. Then  $\mathcal{L}$  has zero Lebesgue measure.

*Proof.* It is enough to show this result for the set  $\mathcal{L}_{m/n}(\Omega)$  of periodic orbits of a fixed rotation number  $m/n \in (0, 1)$ . As in the proof of Theorem 2.6, define the map  $h_{m,n} : V_{m,n} \to \mathbb{R}$  on the open set

$$V_{m,n} = \{ (s_0, \dots, s_n) \in \mathbb{R}^{n+1} \mid s_n = s_0 + m, \, \forall k \quad s_k < s_{k+1} \}$$

by

$$h_{m,n}(s_0,\ldots,s_n) = \sum_{k=0}^{n-1} H(s_k,s_{k+1}).$$

As we saw, the set  $\operatorname{Crit}(h_{m,n})$  corresponds to the stationary configurations, and hence to periodic orbits of rotation number m/n of the billiard map. Moreover, the perimeter of the orbit obtained from  $\underline{s} \in V_{m,n}$  corresponds to the quantity  $h_{m,n}(\underline{s})$ . Hence we just showed that

$$\mathcal{L}_{m/n}(\Omega) = h_{m,n}(\operatorname{Crit}(h_{m,n}))$$

is the so-called set of critical values of  $h_{m,n}$ . By Sard's Lemma, this set has zero measure.  $\Box$ 

**Exercise 2.13.** Construct a  $C^{\infty}$  domain with the Length spectrum of positive Hausdorff dimension.

2.2. Laplace spectrum and Length spectrum. We define the *length spectrum* of  $\Omega$  as the set

 $\mathcal{L}_{\Omega} := \mathbb{N}^+ \cdot \{ \text{lengths of periodic orbits in } \Omega \} \cup \mathbb{N}^+ \cdot |\partial \Omega|,$ 

where  $|\partial \Omega|$  denotes the length of the boundary, namely the set of multiples of the lengths of all periodic orbits and multiples of the perimeter of  $\Omega$ .

As we have already pointed out before (see Ivrii's conjecture), a remarkable relation exists between the length spectrum of a billiard in a convex domain  $\Omega$  and the spectrum of the Laplace operator in  $\Omega$  with Dirichlet boundary condition:

$$\begin{cases} \Delta f + \lambda^2 f = 0 & \text{in } \Omega \\ f|_{\partial\Omega} = 0. \end{cases}$$

From the physical point of view, the eigenvalues are the eigenfrequencies of the membrane  $\Omega$  with a fixed boundary. Denote by  $\operatorname{Spec}_{\Delta}(\Omega) = \{0 < \lambda_1 \leq \lambda_2 \leq \ldots\}$  the Laplace spectrum of eigenvalues solving this problem.

The famous question of M. Kac in its original version asks *if one can recover the domain from the Laplace spectrum*. For general manifolds there are counterexamples (see [44]).

K. Anderson and R. Melrose [2] proved the following relation between the Laplace spectrum and the length spectrum (see also [47, 85, 94]):

## Theorem (Anderson-Melrose). The wave trace

$$w(t) := \operatorname{Re}\left(\sum_{\lambda_n \in \operatorname{Spec}_{\Delta}(\Omega)} e^{i\lambda_n t}\right)$$

is well-defined as a distribution and smooth away from the length spectrum:

(4)  $sing. supp.(w(t)) \subseteq \pm \mathcal{L}(\Omega) \cup \{0\}.$ 

Namely, if  $\xi > 0$  belongs to the singular support of this distribution, then there exists either a closed billiard trajectory of length  $\xi$ , or a closed geodesic of length  $\xi$  in the boundary of the billiard table. Generically, equality holds in (4).

**Exercise 2.14.** An example to convince about this relation is the following. Consider  $\Omega = (0, \pi) \times (0, \pi)$ ; then, its Laplace spectrum is given by

 $\operatorname{Spec}(\Omega) = \{ n^2 + m^2 : (n, m) \in \mathbb{N} \times \mathbb{N} \setminus \{ (0, 0) \} \}.$ 

Discuss its relation with the lengths of periodic orbits in  $\overline{\Omega}$ .

**Remark 2.15.** (i) The above inclusion holds for non-convex  $C^{\infty}$  domains in arbitrary dimension (see [85, Theorem 5.4.6]).

(ii) Observe that there are not known examples of domains in which the singular support of the wave trace is strictly included inside the Length spectrum: the equivalence between these sets is strictly related to the problem whether Laplace spectral rigidity implies Length spectral rigidity. Observe that it may be possible that the singular support of the wave trace is strictly included inside the Length spectrum. If they are the same it is closely related to the problem whether Laplace spectral rigidity implies Length spectral rigidity. In an unpublished manuscript Hezari and Zelditch constructed an example of an analytic domain where the equality fails. In a recent preprint Kaloshin-Koval-Vig [62], it was show that for a dense set of eccentricities  $e \in (0, 1)$ , there is a small perturbation  $\Omega$  of an ellipse  $\mathcal{E}$  of eccentricity e such that the equality of the singular support of  $w_{\Omega}(t)$  and the length spectrum (4) fails.

A very interesting result in this direction has been recently provided in [53], where the authors prove that ellipses of sufficiently small eccentricities are Laplace spectrally unique (up to isometry) among all smooth domains (without any assumption on symmetry, convexity, or closeness to other ellipses). A key result in their proof consists in showing that for nearly circular domains, the lengths of periodic orbits of rotation number 1/q is contained in the singular support of the wave trace [53, Theorem 1.4]. This observation was used in a recent work of Koval [68] about local Laplace spectrum rigidity of ellipses.

2.3. Laplace Spectral Rigidity. Given a class  $\mathcal{M}$  of domains and a domain  $\Omega \in \mathcal{M}$ , we say that  $\Omega$  is *spectrally determined in*  $\mathcal{M}$  if it is the unique element (modulo isometries) of  $\mathcal{M}$  with its Laplace Spectrum: if  $\Omega, \Omega' \in \mathcal{M}$  are *isospectral*, *i.e.*,  $\operatorname{Spec}_{\Delta}(\Omega') = \operatorname{Spec}_{\Delta}(\Omega)$ , then  $\Omega'$  is the image of  $\Omega$  by an isometry (*i.e.*, a composition of translations and rotations).

The question of Kac can be thus formulated as follows, assuming we have fixed a class of domains  $\mathcal{M}$ : Is every  $\Omega \in \mathcal{M}$  spectrally determined?

If  $\mathcal{M}$  is the space of all planar domains, the answer is well known to be negative (see *e.g.*, [44], which generalizes some results previously obtained for compact manifolds without boundary (see [100, 106])).<sup>3</sup> However, all known examples of domains that are not spectrally determined are not convex, moreover, they are bounded by curves that are only piecewise analytic (e.g. plane domains with corners). On the other hand, Zelditch proved in [111] that the inverse spectral problem has a positive answer when  $\mathcal{M}$  is a generic class of analytic  $\mathbb{Z}_2$ -symmetric convex domains (*i.e.*, symmetric with respect to reflection about a given axis). More recently, as we have already mentioned, Hezari and Zelditch [53] proved that ellipses of sufficiently small eccentricities are Laplace spectrally unique (up to isometry) among all smooth

The problem for non-analytic domains is substantially more challenging. In the  $C^{\infty}$  category, Osgood-Phillips-Sarnak [80, 81, 82] showed that isospectral sets are necessarily compact in the  $C^{\infty}$  topology. Sarnak (see [93]) also conjectured that an isospectral set consists of isolated domains. In other words,  $C^{\infty}$ -close to a  $C^{\infty}$  domain there should be no isospectral domains, except those that can be obtained by an isometry.

domains (withouth any assumption on symmetry, convexity, or closeness to other ellipses).

A weaker version of this conjecture can be stated as follows: a domain  $\Omega$  is said to be *spectrally* rigid in  $\mathcal{M}$  if any  $C^1$ -smooth one-parameter isospectral family  $(\Omega_{\tau})_{|\tau|\leq 1} \subset \mathcal{M}$  with  $\Omega_0 = \Omega$  is necessarily an isometric family. We can then ask: "Are all  $C^{\infty}$  domains spectrally rigid?"

The problem of spectral rigidity is in principle much simpler than the inverse spectral problem; yet it turns out to be extremely challenging. Hezari–Zelditch (see [52]) provided a result in the affirmative direction: let  $\Omega_0$  be bounded by an ellipse  $\mathcal{E}$ , then any one-parameter isospectral  $C^{\infty}$ -deformation  $(\Omega_{\tau})_{|\tau|<1}$  which additionally preserves the  $\mathbb{Z}_2 \times \mathbb{Z}_2$  symmetry group of the

 $<sup>^{3}</sup>$ Remarkably, Sunada (see [100]) exhibits isospectral sets (*i.e.*, sets of isospectral manifolds) of arbitrarily large cardinality.

ellipse is necessarily flat (i.e., all derivatives have to vanish for  $\tau = 0$ ).<sup>4</sup> Popov–Topalov [87] recently extended these results (see also [88]).

Further historical remarks on the inverse spectral problem can also be found in [52] and in the surveys [110] and [112].

In the case of Riemannian manifolds, we mention that Guillemin–Kazhdan in [46] showed that any negatively curved surface is spectrally rigid among negatively curved surfaces. This result has been later extended to compact manifolds of negative curvature in [27].

2.4. Length spectral rigidity. The relation between the Laplace Spectrum and the Length Spectrum, immediately raises the following question:

Does the knowledge of the lengths of periodic orbits determine the shape of the billiard domain?

All counterexamples to the inverse spectral problem mentioned earlier also constitute a negative answer to this question. Likewise, at present, there is no known counterexample realized by either convex domains or domains with a  $C^{\infty}$  smooth boundary. Moreover, the above mentioned result by Zelditch (in [111]) also holds in the dynamical context. In the case of sufficiently smooth convex domain, the problem is open and presents the same challenges as the inverse spectral problem.

In [29], the following dynamical problem corresponding to spectral rigidity has been investigated: we say that a domain  $\Omega_0 \in \mathcal{M}$  is dynamically spectrally rigid in  $\mathcal{M}$  if any  $C^1$ -smooth one-parameter dynamically isospectral family  $(\Omega_{\tau})_{|\tau|\leq 1} \subset \mathcal{M}$  is necessarily an isometric family. More specifically, the authors proved the following theorem:

**Theorem 2.16** (De Simoi, Kaloshin, Wei [29]). Let  $\mathcal{M}$  be the set of strictly convex domains with sufficiently (finitely) smooth boundary and axial symmetry and that are sufficiently close to a circle. Then,  $\Omega \in \mathcal{M}$  is dynamically spectrally rigid in  $\mathcal{M}$ .

**Remark 2.17.** This work leaves several natural open problems:

- Remove axial symmetry. Similar symmetry assumption appears in a work of Zelditch [111] and double symmetry assumption appears in Colin de Verdière [32].
- Nonlocal dynamically spectrally rigid in  $\mathcal{M}$  is another exciting open problem.
- A closely related setting is dynamically spectrally rigid for standard maps of a cylinder  $(x, y) \mapsto (x + y + V'(x), y + V'(x))$ , where the x-component is taken (mod. 1) and V(x) is a smooth period function, i.e.  $V(x + 1) \equiv V(x)$ .
- Questions of dynamically spectrally rigid for geodesic flows on 2-torus seems a closely related open problem. Even when geodesic flow is close to integrable.

Let us point out that the above-mentioned results are concerned with spectral rigidity for smooth domains. Some results in the analytic category, yet for non-Birkhoff billiards, are contained in:

- [30], where under suitable symmetry and genericity assumptions, it is proved that the Marked Length Spectrum determines the geometry of billiard tables obtained by removing from the plane finitely many strictly convex analytic obstacles satisfying the so-called non-eclipse condition;

<sup>&</sup>lt;sup>4</sup>Results of this kind are usually referred to as *infinitesimal spectral rigidity*.

- Bunimovich stadia and squash-type stadia are beautiful examples of chaotic billiards. A Bunimovich stadium is a convex domain whose boundary is formed by four segments: two parallel segments forming a rectangle and two strictly convex segments connecting the end points of segments. For example, one can take semicircles. Bunimovich showed that such billiards are chaotic. One can consider segments not to be parallel and connect them by two strictly convex segments forming a squash. In [25] it is established the dynamical spectral rigidity for piecewise analytic Bunimovich stadia and squash-type stadia satisfying an additional "symmetry" assumption.

2.5. Some ideas on the proof of deformational spectral rigidity (Theorem 2.16). Here we introduce the key elements of the proof of Theorem 2.16. Let  $(\Omega_{\tau})_{|\tau|\leq 1} \subset \mathcal{M}$  be a isospectral family of domains.

The first step is to establish existence of a countable family of maximal periodic orbits given by q-gons for all  $q \ge 2$ .

**Lemma 2.18.** (see Lemma 4.3 [29]) Let  $\Omega \in \mathcal{M}$ ; for any  $q \geq 2$ , there exists a periodic orbit of rotation number 1/q passing through the marked point of  $\partial\Omega$  and having maximal length among other periodic orbits passing through the marked point. We call such an orbit marked symmetric maximal periodic orbit and denote it by  $S^q(\Omega)$ .

Let  $S^q = (s_q^k, \varphi_q^k)_{k=0}^{q-1}$  be the maximal symmetric periodic orbit. Associate to  $S^q$  and a continuous function  $\nu : \mathbb{T} \to \mathbb{R}$  a linear functional

$$\ell_{\Omega,q}(\nu) = \sum_{k=0}^{q-1} \nu(s_q^k) \sin \varphi_q^k.$$

Given a parameterization  $\gamma$  of a family  $(\Omega_{\tau})_{|\tau|\leq 1}$  in  $\mathcal{M}$ , we define the infinitesimal deformation function:

$$n_{\gamma}(\tau,\xi) = \langle \partial_{\tau}\gamma(\tau,\xi), N_{\tau}(\tau,\xi) \rangle,$$

where  $\langle \cdot, \cdot \rangle$  is the usual scalar product in  $\mathbb{R}^2$  and  $N_{\tau}(\tau, \xi)$  is the outgoing unit normal vector to  $\partial \Omega_{\tau}$  at the point  $\gamma(\tau, \xi)$ . Observe that  $n_{\gamma}$  is continuous in  $\tau$  and  $n_{\gamma}(\tau, \cdot)$  is smooth for any  $|\tau| \leq 1$ . By the normalization condition of  $(\Omega_{\tau})_{|\tau|\leq 1}$ , we conclude that  $n_{\gamma}(\tau, \cdot)$  is an even function, *i.e.*,  $n_{\gamma}(\tau, \xi) = n_{\gamma}(\tau, -\xi)$ , and  $n_{\gamma}(\tau, 0) = 0$  for any  $|\tau| \leq 1$ . Naturally the space of perturbations can be identified with the space of smooth even functions on the circle denoted  $C_{\text{sym}}$ .

**Proposition 2.19.** (See [29, Proposition 4.6]) Let  $(\Omega_{\tau})_{|\tau|\leq 1}$  be an isospectral family, then for any  $|\tau| \leq 1$ ,  $q \geq 2$  and having fixed arbitrarily  $\bar{S}^q_{\tau}$  a maximal marked symmetric periodic orbit for  $\Omega_{\tau}$ , we have

$$\ell_{\Omega_{\tau},q}(n(\tau,\cdot)) = 0.$$

For any domain  $\Omega$  (parameterized by the length s) with the radius of curvature  $\rho$ , we define the linear functional

$$\ell_{\Omega,0}(\nu) := \int_0^1 \frac{\nu(s)}{\rho(s)} \, ds$$

As it is shown in (4.3), [29] if  $(\Omega_{\tau})_{|\tau|\leq 1}$  be an isospectral family, then for any  $|\tau| \leq 1$  we have  $\ell_{\Omega,0}(n(\tau, \cdot)) = 0$ .

Define the following key notion. Call the linearized isospectral operator  $\mathcal{T}_{\Omega} : C_{\text{sym}} \to \mathbb{R}^{\mathbb{N}}$ :

$$\mathcal{T}_{\Omega}\nu = (\ell_{\Omega,0}(n(\tau,\cdot)), \ell_{\Omega,1}(n(\tau,\cdot)), \dots \ell_{\Omega,q}(n(\tau,\cdot)), \dots).$$

In fact,  $\mathcal{T}_{\Omega}$  has range in  $\ell^{\infty}$ , by definition of the functionals  $\ell_{\Omega,q}$ , since by [8, Lemma 8] there exists some C > 0 so that for any  $q \ge 2$  we have  $\sin \varphi_q^k \le \frac{C}{q}$ .

The linearized isospectral operator bears a strong analogy with the X-transform (see [45, Section 2.2]).

**Theorem 2.20.** ([29, Theorem 4.9]) In the space of sufficiently smooth axis symmetric domains there is a neighborhood of the circular domain such that the operator  $\mathcal{T}_{\Omega}: C_{sym} \to \ell^{\infty}$  is injective.

This Theorem implies the rigidity Theorem above. In the case of the domain  $\Omega_0$  being the circle the linearized isospectral operator  $\mathcal{T}_{\Omega_0}$  is easy to compute. For  $j \ge 1$  and  $q \ge 2$ 

$$\ell_q(e_j) = \delta_{q|j}$$

where  $\delta_{q|j} = 1$  is j is divisible by q and zero otherwise. For the circle  $\mathcal{T}_{\Omega_0}$  is clearly indective. In [29, Lemma B.1] we compute a perturbative expression for  $\ell_{\Omega,q}(e_j)$  when a domain  $\Omega$  is close to the circle. In a proper sense perturbation of  $\mathcal{T}_{\Omega_0}$  is also injective.

*Related prior results.* The problem of isospectral deformations of manifolds without boundary were considered in some early works on variations of the spectral functions and wave invariants.

Let (M, g) be a compact boundaryless Riemannian manifold. A family  $(g_{\tau})_{|\tau|\leq 1}$  of Riemannian metrics on M depending smoothly on the parameter  $|\tau| \leq 1$  is called a *deformation of the metric* g if  $g_0 = g$ . A deformation is called *trivial* if there exists a one-parameter family of diffeomorphisms  $\varphi_{\tau} : M \to M$  such that  $\varphi_0 = \text{Id}$ , and  $g_{\tau} = (\varphi_{\tau})^* g_0$ . For each homotopy class of closed curves in M, consider the infimum of g-lengths of curves belonging to the given homotopy class. The *Length Spectrum*  $\mathcal{L}(M, g)$  is defined as the union of these lengths over all homotopy classes. The *inverse spectral problem* in this setting is to show that two metrics with the same Length Spectrum are isometric.

Likewise, a deformation  $(g_{\tau})_{|\tau|\leq 1}$  is said to be *isospectral* if  $\mathcal{L}(M, g_{\tau}) = \mathcal{L}(M, g)$ . We say that a Riemannian manifold (M, g) is *length spectrally rigid* if it does not admit non-trivial isospectral deformations.

It is worth mentioning that for there is a partial solution of the inverse spectral problem due independently to Croke [26] and Otal [83] which can be stated as follows: any negatively curved manifold is uniquely determined by its *Marked Length Spectrum* (see subsection 2.6 for the corresponding billiard problem).<sup>5</sup>

Recently, Guillamou and Lefeuvre [45] proved that in all dimensions, the marked length spectrum of a Riemannian manifold (M, g) with Anosov geodesic flow and non-positive curvature, locally determines the metric in the sense that two close enough metrics with the same marked length spectrum are isometric.

Another example of deformational spectral rigidity appears in De la Llave, Marco and Moriyón [28]. Recall that one can associate to a symplectic map a generating function. Then, for each periodic orbit, one can define the corresponding *action* by summing the generating function along the orbit. This value of the action is invariant under symplectic coordinate changes. The union of the values all these actions over all periodic orbits is called the action spectrum of the symplectic map. In [28, Theorem 1.3], it is proved that there are no non-trivial deformations of exact symplectic mappings  $B_{\tau}$ ,  $\tau \in [-1, 1]$ , leaving the action spectrum fixed, when  $B_{\tau}$  are Anosov's mappings on a symplectic manifold. One of the reasons for symplectic rigidity in [28]

<sup>&</sup>lt;sup>5</sup>The *Marked Length Spectrum* in the case of negatively curved surfaces without boundary consists of the set of pairs of homotopy classes and length of the shortest geodesic in that homotopy class.

LECTURE NOTES ON BIRKHOFF BILLIARDS: DYNAMICS, INTEGRABILITY AND SPECTRAL RIGIDITY7

is that all periodic points of  $B_{\tau}$  are hyperbolic and form a dense set.

2.6. Marked length Spectral Rigidity. One of the difficulties in working with the length spectrum is that all of this information on the periodic orbits come in a non-formatted way. For example, we lose track of the rotation numbers corresponding to each length. A way to overcome this difficulty is to "organize" this set of information in a more systematic way, for instance by labelling each length with corresponding rotation number. This new set is called the *Marked Length Spectrum* of  $\Omega$  and denoted by  $\mathcal{ML}_{\Omega}$ :

 $\mathcal{ML}_{\Omega} := \{ (\operatorname{length}(\gamma), \operatorname{rot}(\gamma)) : \gamma \text{ periodic orbit of the billiard in } \Omega \},\$ 

where  $rot(\gamma)$  denotes the rotation number of  $\gamma$ .

One could also reduce this set of information by considering not the lengths of all orbits, but selecting some of them. More precisely, for each rotation number p/q in lowest terms, one could consider the maximal length among those having rotation number p/q. We call this map the Maximal Marked Length Spectrum of  $\Omega$ , namely  $\mathcal{ML}_{\Omega}^{\max} : \mathbb{Q} \cap [0, 1/2] \to \mathbb{R}$  given by:

$$\mathcal{ML}_{\Omega}^{\max}(p/q) = \max\left\{ \text{lengths of periodic orbits with rot. number } p/q \right\}.$$

**Marked Spectral Rigidity Question.** Let  $\Omega_1$  and  $\Omega_2$  be two strictly convex planar domains with smooth boundaries and assume that they are isospectral, i.e.,  $\mathcal{ML}_{\Omega_1} \equiv \mathcal{ML}_{\Omega_2}$ . Is it true that  $\Omega_1$  and  $\Omega_2$  are isometric?

Similarly, one could ask whether this same question has an affirmative answer by asking only that  $\mathcal{ML}_{\Omega_1}^{\max} \equiv \mathcal{ML}_{\Omega_2}^{\max}$ .

**Remark 2.21.** (i) The above question could be reformulated – and it remains still meaningful and interesting – by asking that they two domains are only isospectral near the boundary, i.e.,  $\mathcal{ML}_{\Omega_1}(p/q) = \mathcal{ML}_{\Omega_2}(p/q)$  for all  $p/q \in \mathbb{Q} \cap [0, \varepsilon)$ , for some  $0 < \varepsilon \leq 1/2$ .

See section 4 for a reformulation of this question in terms of the so-called Mather's minimal average action (or  $\beta$ -function) and for some partial answers to the Marked Length rigidity question related to the proof of the perturbative Birkhoff conjecture (see section 4.3).

## 3. LECTURE III: CAUSTICS, INVARIANT CURVES AND INTEGRABILITY

3.1. Caustics and invariant curves. In this section, we would like to recall the concept of *caustic* of a billiard and discuss its relations with invariant curves for the billiard map. Let us first start to introduce the concepts of caustic and integrability, starting with some motivating means of two examples; the definition of caustic will be given in subsection 3.2.

**Example 1: Circular billiards.** As we saw, the billiard in the unit disk defines a map  $f : \mathbb{A}_{2\pi} \to \mathbb{A}_{2\pi}$  for which the sets

$$\mathcal{C}_{\varphi} := \{ (\theta, \varphi') \in \mathbb{A}_{2\pi} \mid \varphi' = \varphi \}$$

are invariants by f, that is if  $(\theta, \varphi) \in C_{\varphi}$  then  $f(\theta, \varphi) = (\theta + 2\varphi, \varphi) \in C_{\varphi}$ . In particular, as we already notices,  $\varphi$  stays constant along the orbit and it represents an *integral of motion* for the map, and the property of the orbits are determined by the corresponding angle  $\varphi = \pi \omega$ , with  $\omega \in (0, 1)$  (see Section 3.2).

Moreover, this billiard enjoys the peculiar property that all orbits with  $\varphi = \pi \omega$  are tangent to the same concentric circle of radius  $R \cos \pi \omega$  (see Figure 8); this concentric circle is an example of *caustics* (see Definition 3.8) and it is related to the existence of a homotopically non-trivial invariant curve for the corresponding billiard map, namely the  $C_{\omega} = \mathbb{R}/2\pi R\mathbb{Z} \times {\pi \omega}$  (this relation between caustics and invariant curves is more subtle, see Remark ??). Observe that the whole phase space of the circular billiard map – which is topologically a cylinder – is completely by these  $C_{\omega}$  and, looking at the billiard table, this is completely foliated by caustics (this foliation is a singular foliation, due to the special role of the center of the disc): in this regard, circular billiards are example of *integrable billiards*; see Figure 9.



FIGURE 8. A billiard trajectory in the disk remains tangent to the same concentric circle after successive reflections.



FIGURE 9. Circular billiard and its phase space

**Example 2: Elliptic billiard.** Consider the billiard in an ellipse  $\mathcal{E}$  given, for 0 < b < a by

$$\mathcal{E} = \left\{ (x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\}.$$

Since  $\mathcal{E}$  is not a circle, it has two distinct foci  $F_1$  and  $F_2$  lying on the *x*-axis. The billiard in the ellipse can be described geometrically as follows (see Figure 11):

LECTURE NOTES ON BIRKHOFF BILLIARDS: DYNAMICS, INTEGRABILITY AND SPECTRAL RIGIDITY9

**Proposition 3.1.** Let  $\ell$  be an oriented line intersecting the ellipse  $\mathcal{E}$  transversally, and consider the billiard flow induced by  $\ell$  in  $\mathcal{E}$ . Then one of the following situations is satisfied:

1.  $\ell$  contains one of the foci; in this case, the successive reflections of  $\ell$  contain a focus and the latter differs from the focus of previous reflection. Moreover, the succesive reflected lines will converges to the x-axis.

2.  $\ell$  crosses the x-axis between the foci; in this case, the successive reflections of  $\ell$  always cross the x-axis between the foci. Moreover, they are supported by lines which are tangent to one and the same confocal hyperbola.

3.  $\ell$  crosses the x-axis outside the foci; in this case, the successive reflections of  $\ell$  always cross the x-axis outside the foci. Moreover, they remain tangent to one and the same confocal ellipse.



FIGURE 10. The three different possibilities for a billiard trajectory in a ellipse. Left: the trajectory always alternatively contains one of the two foci. Center: the trajectory always crosses the segmente between the foci, and remains tangent to a confocal hyperbola. Right: the trajectory never crosses the foci line and remain tangent to a confocal ellipse.

This proposition tells us that the trajectory of a billiard in an ellipse are tangent to curves, which are confocal conics to  $\mathcal{E}$ , including hyperbolae.

**Exercise 3.2.** Using the description of an ellipse as the data of two distinct points  $F_1, F_2$  and a constant r > 0 such that

$$\mathcal{E} = \{ M \in \mathbb{R}^2 \, | \, F_1 M + F_2 M = r \},\$$

show item 1 of Proposition 3.1.

The proof of Proposition 3.1 can be shown using the so-called *Joachimsthal invariant*:

**Proposition 3.3.** Let  $p = (x, y) \in \mathcal{E}$  and  $v = (v_x, v_y)$  be a unit vector starting at p and pointing inside of  $\mathcal{E}$ . Then the quantity J(p, v) defined by

$$J(p,v) = \frac{xv_x}{a} + \frac{yv_y}{b}$$

is invariant by the billiard map inside  $\mathcal{E}$ . It is called the Joachimsthal invariant of the ellipse.

A proof of Proposition 3.3 can be found in [33, 102].

**Exercise 3.4.** Let 0 < b < a and consider the pencil of confocal conics  $C_{\lambda}$  given for any  $0 < \lambda < a$  and  $\lambda \neq b$  by the equations

$$\mathcal{C}_{\lambda}: \quad rac{x^2}{a^2-\lambda}+rac{y^2}{b^2-\lambda}=1.$$

Show that given a trajectory which does not contain the foci, it remains tangent to the conics  $C_{\lambda_0}$  where

$$\lambda_0 = (abJ(p,v))^2$$

and J(p, v) is Joachimsthal invariant associated to the trajectory.

Description of the phase space of the billiard in an ellipse. Optical properties of conics (an alternative way to consider the billiard ball motion inside a conic) were already well known to ancient Greeks. We refer to [102] for a more detailed discussion (see also [95]). Proposition 3.1 leads to a nice description of the phase space of the billiard in  $\mathcal{E}$ . Consider an arc-length coordinate s on  $\mathcal{E}$  defined modulo the perimeter L of the ellipse, and  $\varphi$  the usual angle of reflection. The phase space in  $\mathcal{E}$  we consider is

$$\mathbb{A}_L = \{ (s, \varphi) \mid s \in \mathbb{R}/L\mathbb{Z}, \, \varphi \in (0, \pi) \}.$$

 $\mathbb{A}_L$  is the disjoint union of the following objects (see Figure 11):

1. Hyperbolic 2-periodic points. The two points  $O_1 = (0, \pi/2)$  and  $O_2 = (L/2, \pi/2)$ , corresponding to the major semi-axis, are called hyperbolic and corresponds to the billiard flow induced by the x-axis. The billiard map permutes  $O_1$  and  $O_2$ .

1'. Elliptic 2-periodic points. The two points  $O'_1 = (L/4, \pi/2)$  and  $O_2 = (3L/4, \pi/2)$ , corresponding to the minor semi-axis, are called *elliptic* and corresponds to the billiard flow induced by the y-axis. The billiard map permutes  $O'_1$  and  $O'_2$ .

2. Stable and unstable manifolds. Two graphs over s,  $\Gamma_1$ ,  $\Gamma_2$ , intersecting at  $O_1$  and  $O_2$  and forming two eyes in the phase space. The curve  $\Gamma_i$  consists of the point  $(s, \varphi)$  such that the line intersecting  $\mathcal{E}$  at s with an angle  $\varphi$  contains the focus  $F_i$ . They correspond to the stable and unstable manifolds of  $O_1$  and  $O_2$  and satisfies the following properties: the billiard map f permutes them, they are invariant under  $f^2$  and the points in  $\Gamma_i$  converge to  $O_{1-i}$  under iteration of  $f^2$ . The graph  $\Gamma_2$  is called stable manifolds of  $O_1$  since the points on it converges to  $O_1$  under positive iteration of  $f^2$ ;  $\Gamma_1$  is called unstable manifold of  $O_1$  since the points on it converges to  $O_1$  under negative iteration of  $f^2$ . The same remarks hold for  $O_2$  by permuting the roles of  $\Gamma_1$  and  $\Gamma_2$ .

3. Homotopically trivial invariant curves. Given a confocal hyperbola H, consider the set of pairs  $(s, \varphi)$  such that the line intersecting  $\mathcal{E}$  at s with angle  $\varphi$  is tangent to H. It consists of two closed invariant curves (depending on the orientation of the tangency points of  $\ell$  with  $\gamma$  - one is given by the positive tangencies, the other one by negative tangencies) located in the eyes described at point 2. These curves are contractible in the cylinder  $\mathbb{A}_L$ .

4. Homotopically non-trivial (or essential) invariant curves. Given a confocal ellipse  $\gamma$  nested in  $\mathcal{E}$ , consider the set of pairs  $(s, \varphi)$  such that the line intersecting  $\mathcal{E}$  at s with angle  $\varphi$  is tangent to  $\gamma$ . This consists of two closed invariants graphs over s (depending on the tangency orientation, as in 3.) located outside the eyes described at point 2. These curves are homotopically non-trivial in the cylinder  $\mathbb{A}_L$ .



FIGURE 11. Elliptic billiard and its phase space

LECTURE NOTES ON BIRKHOFF BILLIARDS: DYNAMICS, INTEGRABILITY AND SPECTRAL RIGIDITY

**Remark 3.5.** Confocal ellipses are therefore examples of caustics and they foliate everything but the closed segment between the two foci (see Figure 11). Hence, this could be also considered as an example of integrable billiard; see Figure 11. Observe that also hyperbolae can be considered examples of caustics, although, differently from concentric circles or confocal ellipses, they are not connected, closed or convex; see subsection 3.2 for a more precise discussion.

3.2. Caustics. Let us introduce the concept of  $convex \ caustic^6$  and its relation with invariant curves for the billiard map. We refer to [49] for a more detailed (and extended) presentation of these topics. We also discuss some results and questions about their existence.

Let us start by recalling the definition of *invariant circle* (or *homotopically non-trivial invariant curve*, or *essential invariant curve*) for a billiard map.

**Definition 3.6.** We say that a curve  $\gamma \subset \mathbb{A}_L$  is an *invariant circle* for the billiard map  $f : \mathbb{A}_L \to \mathbb{A}_L$  if  $\gamma$  is isotopic to a boundary component of  $\mathbb{A}_L$  and  $f(\gamma) = \gamma$ .

**Remark 3.7.** (i) Observe that both boundary components of  $\mathbb{A}_L$  are trivial invariant circles. It follows from Birkhoff's theorem that invariant circles must be Lipschitz graphs (see [17] and also [95, Theorem 1.3.3]).

(ii) Clearly, a billiard map may possess invariant curves that are not invariant circles: see for example the billiard map in an ellipse (see discussion above) and its homotopically trivial (disconnected) invariant curves, corresponding to orbits intersecting the segment between the foci).

In the spirit of what we have seen in the examples of circular and elliptic billiards, let us give the following definition.

**Definition 3.8.** A  $C^1$  simple closed curve  $\Gamma$  in the interior of  $\Omega$  is called a *convex caustic* for the billiard map f, if  $\gamma$  bounds a convex set  $D_{\Gamma}$  and any supporting line to  $D_{\Gamma}$  remains a supporting line to  $D_{\Gamma}$  after the billiard reflection in  $\Omega$ . In other words, every time a trajectory is tangent to  $\Gamma$ , then it remains tangent after every each reflection.



FIGURE 12. Caustic and Lazutkin invariant (figure credits [95, Fig. 3.6]).

In our discussion, we will focus on convex caustics, however one could consider a more general notion of caustic that does not require the properties of bounding a convex region, of being closed (see for example, confocal hyperbola for elliptic billiards, see discussion above), nor to be necessarily  $C^1$ . Since this will not be the object of our investigation, we refer to the discussion

 $<sup>^{6}</sup>Caustic$  comes from the greek word χαυστιχός (kaustikós), meaning "burning"; this terminology is related to optics and refers to the envelope of reflected or refracted rays of light, namely concentration of lights that can potentially lead to burns.

in [5, 49, 67]. See Figure 13 for some examples.

**Remark 3.9.** An interesting example of billiard maps with invariant circles are billiards whose boundary is a curve of constant width, i.e., namely a curve that bounds a convex planar region whose width (defined as the perpendicular distance between two distinct parallel lines each having at least one point in common with the region's boundary but none with its interior) is the same regardless of the orientation of the curve (to construct such curves, see, for example, [67, Section 4] and [102, Exercise 3.13]. The corresponding billiard map has an invariant circle consisting of 2-periodic orbits. These curves corresponds to caustics that, in general, may have cusps; see [67, Section 4 and Fig. 6].

Billiard tables (other than ellipses) with a 1-parameter family of 3-periodic trajectories have been constructed by Innami in [58].



FIGURE 13. Examples of non-convex caustics in billiards of constant width (figure credits [67, Fig. 6]).

As in the case of billiard in a disk, convex caustics and invariant curves are related (see Figure 14). One can prove the following.

**Proposition 3.10.** Let  $\Gamma$  be a caustic of a Birkhoff billiard  $\Omega$  with a chosen orientation. Consider the set  $\gamma$  of pairs  $(p, \varphi)$  where p is any point on  $\partial\Omega$  and v is a unit inward pointing vector supporting the line containing p and tangent to  $\Gamma$  with the same orientation at the tangency point. Then  $\gamma$  is a graph in s which is invariant by the billiard map associated to  $\Omega$ .

Observe that to every convex caustic has a well-defined rotation number. In fact, the dynamics tangent to it, induces a circle homeomorphism from the boundary to itself; the rotation number of the caustic corresponds to the Poincaré rotation number of this circle homeomorphism.

**Exercise 3.11.** In the case of the billiard in the unit disk, show that the invariant curve  $C_{\varphi} := \{(\theta, \varphi') \in \mathbb{A}_{2\pi} \mid \varphi' = \varphi\}$  has rotation number  $2\varphi$ .

**Remark 3.12.** (i) The notion of caustics is often connected to the so-called whispering gallery, a phenomenon that can be detected under some particular domes, in which whispers can be clearly transmitted and received from distant parts of the gallery, as long as the talker/listener are close the wall.

(ii) If  $\Gamma_{\omega}$  is a convex caustic with rotation number  $\omega \in (0, 1/2]$ , then one can associate to it an invariant, the so-called Lazutkin invariant  $Q(\Gamma_{\omega})$ . More precisely

(5) 
$$Q(\Gamma_{\omega}) = |A - P| + |B - P| - |AB|$$

where  $|\cdot|$  denotes the Euclidean length and |AB| the length of the arc on the caustic joining A to B (see figure 12). This quantity is connected to the value of Mather's  $\alpha$ -function, as it will be discussed in section 4.



FIGURE 14. Left: A Birkhoff billiard  $\Omega$  having a caustic  $\Gamma$ ; once we fix a positive orientation, for each point p of arc-length parameter s there is only one possible oriented line emitted from p and tangent to  $\Gamma$ , we denote by  $\varphi(s)$  its angle at pwith  $\partial \Omega$ . Right: The invariant graph  $s \mapsto (s, \varphi(s))$  associated to the caustics  $\Gamma$ and drawn in the phase space  $\mathbb{A}_L$  where L is the perimeter of  $\Omega$ .

One could wonder about the relation between caustics for the billiard in  $\Omega$  and invariant circles for the corresponding billiard map f. While one can show that to a convex caustic in  $\Omega$  (not necessarily  $C^1$ ) corresponds an invariant circle for the billiard map (see Proposition 3.10), however, the converse is however is not always true: given an invariant curve  $\gamma$  of pairs (p, v) where p is a point on the bounary and v is a unit inward pointing vector, one can consider the envelope of the lines passing through p and directed by v; this curve might be not convex nor smooth. We refer to [5, 49, 67, 102] for more details.

3.3. Existence of (convex) caustics. A natural question that one could wonder is whether the existence of (convex) caustics is a common or a rare phenomenon. As we have seen before, circular and elliptic billiards possess many convex caustics.

**Questions:** Are there other Birkhoff billiards with (convex) caustics? And in case of an affirmative answer: How many of them is reasonable to expect?

**Note:** In the following we will often write caustic in place of convex caustic (unless differently specified). However, most of these questions can be addressed for more general notions of caustics.

String construction. Constructing a Birkhoff billiard with **at least one** caustic is easy: it is enough to perform the so-called string construction, similarly to the well-known one to draw a circle as the set of points equidistant from a fixed center, or to construct an ellipse as the locus of points whose distances from two fixed points have a constant sum. Pictoriallky, (see for example [102, Chapter 5] for a more precise construction), given  $\gamma \subset \mathbb{R}^2$  a smooth closed convex curve of length  $L_0$ , let  $L > L_0$  and  $\Gamma$  be the closed curve obtained as follows: consider a closed inextensible string of length L wrapped around  $\gamma$ , pull it tight at a point and move this point around  $\gamma$ : the curve that one obtains, corresponds to a billiard domain that has  $\gamma$  as a caustic.



FIGURE 15. The string construction around a fixed curve  $\gamma$ : an inextensible string is wrapped around  $\gamma$  and stretched so that the quantity  $MA + MB + \underline{AB}$  remains the same. This is the same as saying that Lazutkin's invariant, namely  $Q(M) = |MA| + |MB| - |\widehat{AB}|$ , is constant.

More precisely,  $\Gamma$  can be defined as the set of point M outside  $\gamma$  such that, if A and B are the two points of tangencies of the lines tangent to  $\gamma$  and containing M, then the length

 $|AM| + |BM| + |\underline{AB}|$ 

is constant equal to L, where |AB| is the length of the arc AB of the curve  $\gamma$  between A and B and located the furthest from M. Note that the level sets of this quantity are the same as the one of

$$Q(M) := AM + BM - |AB|$$

where |AB| is the arc of the curve  $\gamma$  between A and B the closest from M (compare this quantity with *Lazutkin invariant* of  $\gamma$ , see Remark 3.12).

Lazutkin's KAM caustics. Are there other billiards with **infinitely many** caustics? Quite surprisingly, the answer is affirmative: all (sufficiently smooth) Birkhoff billiards have infinitely many smooth convex caustics that accumulate to the boundary of the billiard domain. In [70], in fact, V. Lazutkin introduced a very special change of coordinates that reduces the billiard map f to a very simple form (as usual,  $L = |\partial \Omega|$ ). Let  $L_{\Omega} : \mathbb{R}/L\mathbb{Z} \times [0, \pi] \to \mathbb{R}/\mathbb{Z} \times [0, \delta]$  with small  $\delta > 0$  be given by

$$L_{\Omega}(s,\varphi) := \left(x = C_{\Omega}^{-1} \int_{0}^{s} \rho^{-2/3}(s) ds, \qquad y = 4 C_{\Omega}^{-1} \rho^{1/3}(s) \sin \varphi/2\right),$$

where  $\rho$  denotes the radius of curvature of  $\partial\Omega$ , and  $C_{\Omega} := \int_{0}^{\ell} \rho^{-2/3}(s) ds$  (sometimes called the *Lazutkin perimeter*). In these new coordinates the billiard map has a more simple expression:

$$B(x,y) = \left(x + y + O(y^3), y + O(y^4)\right).$$

In particular, near the boundary  $\{y = 0\}$ , this map can be seen as a small perturbation of the integrable map  $(x, y) \mapsto (x + y, y)$ , and hence, under suitable regularity assumptions, KAM

theorem can be applied (it is sufficient, for example, that  $\partial\Omega$  is  $C^6$ , so that the map is at least  $C^5$ ). Hence, there exists a positive measure Cantor set of smooth invariant circles for the map which accumulates on  $\{y = 0\}$  and on which the motion is smoothly conjugate to a rigid rotation with Diophantine rotation number (see [70] and also [86] for a refined version); this translates into the existence of a positive measure set of caustics, accumulating to the boundary of the billiard table.

Non-existence of caustics. Observe that in this context it is extremely important that  $\Omega$  is strictly convex. In [77], in fact, Mather proved the non-existence of caustics if the curvature of the boundary vanishes at one point. An alternative proof of this result has been provided by Gutkin and Katok in [49], where the authors also investigate how the shape of the domain determines the location of caustics, establishing the existence of open regions which are free of caustics and estimating (from below) the size of these regions. More specifically, given a caustic  $\Gamma$  with Lazutkin invariant  $L = L(\Gamma)$ , if we denote by  $\delta_{\max}(\Gamma, \partial\Omega)$  the maximum distance of  $\Gamma$ from the boundary  $\partial\Omega$ , they proved the following estimates (see [49, Propositions 1.2-3]):

$$\frac{\delta_{\max}^2(\Gamma, \partial \Omega)}{d} \le L \le \min\{2d^3\kappa^2, 2/K\},\$$

where  $d = d(\Omega)$  denotes the diameter of  $\Omega$ , while  $\kappa$  and K are respectively the minimum and the maximum of the curvature of  $\partial \Omega$ .

It follows from this that if  $\kappa = 0$  at some point, then caustics cannot exist.

3.4. Integrability and Birkhoff conjecture. Next step then consists in asking in which cases these caustics foliate the whole billiard table or an open dense subset of it, as it happens in the circular and elliptic cases. In other words: *are there other examples of integrable billiards*?

This appearantly naïve question turns out to be much more difficult to extricate, and it has given rise to one of the most famous (and somehow impenetrable) open problem in dynamical systems: the so-called *Birkhoff conjecture*.

**Conjecture (Birkhoff)** Circular and elliptic billiards are the only examples of integrable Birkhoff billiards.

**Remark 3.13.** Although some vague indications of this question can be found in [16], to the best of our knowledge, its first appearance as a conjecture was in a paper by Poritsky [89], where the author attributes it to Birkhoff himself<sup>7</sup>. Thereafter, references to this conjecture (either as Birkhoff conjecture or Birkhoff-Poritsky conjecture) repeatedly appeared in the literature: see, for example, Gutkin [48, Section 1], Moser [79, Appendix A], Tabachnikov [101, Section 2.4], etc.

This conjecture assumes very different connotations and levels of complexity, according to the notion of integrability that one takes into account. Despite its long history and the amount of attention that it has captured over the last decades, many interesting formulations of this conjecture still remain unanswered.

<sup>&</sup>lt;sup>7</sup>Poritsky was a National Research Fellow in Mathematics at Harvard University, presumably under the supervision of Birkhoff and refers to Birkhoff and stated that he wrote that paper while in Harvard. However, [89] was published several years after Birkhoff's death.

We shall see in the section 4 how also this conjecture/question can be rephrased as a regularity question for Mather's minimal average action (or  $\beta$ -function).

3.4.1. *Global Integrability*. In [9], Bialy proved the following result under the assumption of full global integrability.

**Theorem (Bialy).** If the phase space of the billiard ball map is fully foliated by continuous invariant circles, then it is a circular billiard.

**Remark 3.14.** An integral-geometric approach to prove Bialy's result was proposed by Wojtkowski in [109], by means of the so-called mirror formula. This approach was later exploited by Bialy [10] for billiards on the sphere and the hyperbolic plane, as well as for magnetic billiards.

Observe that Bialy and Wojtkowski's result is not in contrast with what we have discussed in the case of elliptic billiards. In fact, in that case the family of convex caustics represented by confocal ellipses do not foliate the whole domain (the segment between the two foci is left out) neither the set of homotopically non-trivial invariant curves (invariant circles) have full  $\omega$ -measure in the phase space: the homotopically trivial invariant curves corresponding to orbits tangent to confocal hyperbolae, foliate a positive  $\omega$ -measure set (in the phase portrait – see Figure 11 – this set corresponds to the area below the separatrix, *i.e.*, the stable/unstable manifold of the hyperbolic 2-periodic orbit corresponding to the major semi-axis of the ellipse).

What about other notions of integrability? In the study of integrable systems, in fact, in most of the cases integrals of motion are non-degenerate not everywhere, but either on an open-dense subset of the phase space (we shall refer to this as *global integrability*) or just a proper (non-trivial) open subset (we shall refer to this as *local integrability*).

**Remark 3.15.** (i) An interesting result by Innami [59] shows that the existence of convex caustics with rotation numbers accumulating to 1/2 implies that the billiard must be an ellipse. This regime of integrability is somehow opposite to the one we are interested in, which is concerned with caustics near the boundary of the billiard table, i.e., with small rotation numbers. Innami's proof is based on Aubry-Mather theory; a simpler and more geometric proof of Innami's result has been recently given in [5]. Observe that in this result it is decisive that the caustics are convex.

(ii) In this regard, Treschev in [105] gave numerical indication that there might exist analytic billiards, different from ellipses, for which the dynamics in a neighborhood of the elliptic period-2 orbit is conjugate to a rigid rotation. These billiards could be seen as an instance of local integrability; however, as we have already remarked above, this regime is somehow complementary to the one usually considered for Birkhoff conjecture, since it is concerned with integrability a neighborhood of an elliptic periodic orbit of period 2. Very interestingly, this fact – if verified – would provide an intriguing indication that these regimes of integrability are significantly different.

3.4.2. *Perturbative Birkhoff conjecture*. Instead of considering all possible Birkhoff billiards, one could restrict the analysis to what happens for domains that are sufficiently close to ellipses and try to study the Birkhoff conjecture in this class of domains, which can be considered as *perturbations* of ellipses. More specifically, we can state the following perturbative version

LECTURE NOTES ON BIRKHOFF BILLIARDS: DYNAMICS, INTEGRABILITY AND SPECTRAL RIGIDIT27

of Birkhoff conjecture.

**Birkhoff Conjecture (Perturbative version).** A smooth strictly convex domain that is sufficiently close (w.r.t. some topology) to an ellipse and whose corresponding billiard map is integrable, is necessarily an ellipse.

First results in this direction were obtained:

- Levallois [72] and Levallois Tabanov [73]: Non-integrability of certain algebraic perturbations of elliptic billiards.
- Delshams and Ramírez-Ros [31]: Non-integrability of entire symmetric perturbations of ellipses (these perturbations break integrability near the homoclinic solutions.

More recently, Avila, De Simoi and Kaloshin proved in [8] that the claim of the perturbative version of Birkhoff conjecture is true, for domains that are sufficiently close to a circular billiard. The complete proof for domains sufficiently close to an ellipse of any eccentricity, has been provided in [64].

Let us describe this result more precisely, starting with the following definition.

# **Definition 3.16.** Let $\Omega$ be a strictly convex domain.

(i) We say  $\Gamma$  is an integrable rational caustic for the billiard map in  $\Omega$ , if the corresponding invariant circle  $\Gamma$  consists of periodic points; in particular, the corresponding rotation number is rational.

(ii) Let  $q_0 \geq 2$  be a positive integer. If the billiard map inside  $\Omega$  admits integrable rational caustics for all rotation numbers  $0 < \frac{p}{q} < \frac{1}{q_0}$ , we say that  $\Omega$  is  $q_0$ -rationally integrable.

The main result proved in [64] is the following.

**Theorem 3.17** (Kaloshin–Sorrentino [64]). For any eccentricity  $0 \le e_0 < 1$  outside of locally finite set in [0,1)the following holds. Let  $\mathcal{E}_0$  be an ellipse of eccentricity  $e_0$  and semifocal distance c; let  $k \ge 39$ . For every K > 0, there exists  $\varepsilon = \varepsilon(e_0, c, K)$  such that the following holds: if  $\Omega$  is a 2-rationally integrable  $C^k$ -smooth domain, whose boundary  $\partial\Omega$  is

- K-close to  $\mathcal{E}_0$ , with respect to the  $C^k$ -norm,
- $\varepsilon$ -close to  $\mathcal{E}_0$ , with respect to the  $C^1$ -norm,

then  $\Omega$  is an ellipse.

**Remark 3.18.** Actually, it is sufficient to ask only the existence of integrable rational caustics of rotation number  $\frac{1}{q}$ , for all  $q \ge 3$ .

3.4.3. Local integrability and Birkhoff conjecture. What can be said for locally integrable Birkhoff billiards? As we have noticed in Remark 3.15, the correct regime that one should consider seems to be integrability in a neighborhood of the boundary of the billiard table, *i.e.*, for small rotation numbers.

Let us denote with  $\mathcal{E}_{e,c} \subset \mathbb{R}^2$  an ellipse of eccentricity e and semifocal distance c. We state the following local version of Birkhoff conjecture.

**Local Birkhoff Conjecture.** For any integer  $q_0 \ge 3$ , there exist  $e_0 = e_0(q_0) \in (0,1)$ ,  $m_0 = m_0(q_0)$ ,  $n_0 = n_0(q_0) \in \mathbb{N}$  such that the following holds. For each  $0 < e \le e_0$  and  $c \ge 0$ , there exists  $\varepsilon = \varepsilon(e, c, q_0) > 0$  such that the following holds.

If  $\mathcal{E}_{e,c}$  is an ellipse of eccentricity e and semi-focal distance c, and  $\Omega$  is a  $q_0$ -rationally integrable  $C^{m_0}$ -smooth domain, whose boundary  $\partial \Omega$  is  $\varepsilon$ -close to  $\mathcal{E}_0$ , with respect to the  $C^{m_0}$ -norm, then  $\Omega$  must be an ellipse.

This conjecture has been first studied in [55]. More precisely, the following results have been proved (see also Section 3.6 for more recent advances by Bialy-Mironov [14], Kaloshin-Koudjinan-Zhang [63], Koval [68]).

**Theorem 3.19** (Huang, Kaloshin, Sorrentino [55]). (i) The Local Birkhoff Conjecture holds true for  $q_0 = 2, 3, 4, 5$ , with  $m_0 = 40q_0$  and  $n_0 = 3q_0$ . (ii) The Local Birkhoff Conjecture holds true for  $q_0 > 5$  with  $m_0 = 40q_0$  and  $n_0 = 3q_0$ , subject to checking that  $q_0 - 2$  matrices (which are explicitly described) are invertible.

**Remark 3.20.** (i) Case  $q_0 = 2$  was proven in [8] (see also [59, 64]).

(ii) Smoothness exponents are probably not optimal.

(iii) Notice that in the proof we actually need only the existence of rationally integrable caustics of rotation numbers, less than  $1/q_0$ , of the form j/q for j = 1, 2, 3.

(iv) The invertibility condition on finitely many matrices, to which the claim of part (ii) of Theorem 3.19 is subject, is explicit and computable. In [56] it is described how to implement an algorithm to verify it by means of symbolic computations. The coefficients of these matrices are completely determined by the e-expansion of the action-angle parametrisation of the ellipse, which, in turn, is explicitly given by elliptic integrals; it turns out that the entries of these matrices are either 0, 1 or of the form  $\xi \cos^{-2j}(w\pi)e^{2j}$ , where  $\xi \in \mathbb{Q}$ ,  $j \in \mathbb{N}$ ,  $w \in \{\frac{1}{2k+1}, \frac{2}{2k+1}, \frac{1}{2k}, \frac{3}{2k} : k > j\}$ .

Recently a complete proof of this conjecture for ellipses of almost all eccentricities has been given in [68].

# 3.5. Some ideas on the proofs of Perturbative Birkhoff conjecture and its local version (Theorems 3.17 and 3.19).

3.5.1. *Perturbative Birkhoff conjecture (Theorem 3.17).* Let us provide a description of the strategy that we adopted in [64] to prove Theorem 3.17.

For small eccentricities, Theorem 3.17 was proven in [8]. Let us start by describing the simplified setting of integrable infinitesimal deformations of a circle. This provides an insight into the strategy of the proof in the general case.

Let  $\Omega_0$  be a circle centered at the origin and radius  $\rho_0 > 0$ . Let  $\Omega_{\varepsilon}$  be a one-parameter family of smooth deformations given in the polar coordinates  $(\rho, \varphi)$  by

$$\partial\Omega_{\varepsilon} = \{(\rho, \varphi) = (\rho_0 + \varepsilon \rho(\varphi) + O(\varepsilon^2), \varphi)\},\$$

Consider the Fourier expansion of  $\rho$ :

$$\rho(\varphi) = \rho'_0 + \sum_{k>0} \rho_k \sin(k\varphi) + \rho_{-k} \cos(k\varphi).$$

**Theorem 3.21** (Ramírez-Ros [90]). If  $\Omega_{\varepsilon}$  has an integrable rational caustic  $\Gamma_{1/q}$  of rotation number 1/q, for any  $\varepsilon$  sufficiently small, then we have  $\rho_{kq} = \rho_{-kq} = 0$  for any integer k.

Let us now assume that the domains  $\Omega_{\varepsilon}$  are 2-rationally integrable for all sufficiently small  $\varepsilon$  and ignore for a moment the dependence on the parametrisation: then the above theorem implies that  $\rho'_k = \rho''_k = 0$  for k > 2, *i.e.*,

$$\rho(\varphi) = \rho'_0 + \rho'_1 \cos \varphi + \rho''_1 \sin \varphi + \rho'_2 \cos 2\varphi + \rho''_2 \sin 2\varphi$$
$$= \rho'_0 + \rho_1^* \cos(\varphi - \varphi_1) + \rho_2^* \cos 2(\varphi - \varphi_2)$$

where  $\varphi_1$  and  $\varphi_2$  are appropriately chosen phases.

Remark 3.22. Observe that

- $\rho_0$  corresponds to an homothety;
- $\rho_1^*$  corresponds to a translation in the direction forming an angle  $\varphi_1$  with the polar axis  $\{\varphi = 0\};$
- $\rho_2^*$  corresponds to a deformation of the circle into an ellipse of small eccentricity, whose major axis forms an angle  $\varphi_2$  with the polar axis.

This implies that, infinitesimally (as  $\varepsilon \to 0$ ), rationally integrable deformations of a circle are tangent to the 5-parameter family of ellipses.

In order to extend these ideas to the case of an integrable perturbation (not necessarily a deformation) of an ellipse, a more elaborate strategy is needed, involving more quantitative estimates and approximation procedure (we refer to [8, 64] for more technical details). In particular, Fourier modes are replaced by new functions determined by the dynamics inside the approximating ellipse, that we call dynamical modes  $\{c_q, s_q\}_{q\geq 3}$ , which are given by:

$$c_q(\varphi) := \frac{\cos\left(\frac{2\pi q}{4K(k_q)}F(\varphi;k_q)\right)}{\sqrt{1-k_q^2\sin^2\varphi}}$$
$$s_q(\varphi) := \frac{\sin\left(\frac{2\pi q}{4K(k_q)}F(\varphi;k_q)\right)}{\sqrt{1-k_q^2\sin^2\varphi}}$$

where  $k_q$  denotes the eccentricity of the confocal ellipse corresponding to the caustic of rotation number 1/q, while

$$F(\varphi;k) := \int_0^{\varphi} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \tau}} \quad \text{and} \quad K(k) := F\left(\frac{\pi}{2};k\right).$$

are the elliptic integrals of first kind (see, for example, [1] for more details on these functions and their properties).

The core of the proof consists in showing that these dynamical modes together with the infinitesimal generators of homotheties, translations, rotations and hyperbolic rotations (*i.e.*, those transformations preserving the set of ellipses), form a basis of  $L^2(\mathbb{R}/2\pi\mathbb{Z})$ . This is one of the main difficulties (maybe the hardest one) involved in the extension of the perturbative result in [8] to the case of perturbations of any ellipse, as studied in [64]. While in the former case, one can take advantage of the fact that these functions can be considered small perturbations of the Fourier modes, in the latter new strategies need to be exploited.

In [64], we consider analytic extensions of the action-angle coordinates of the elliptic billiard, more specifically, of the boundary parametrizations induced by each integrable caustic (these functions can be explicitly expressed in terms of elliptic integrals and Jacobi elliptic functions. ). A detailed study of their complex singularities and the size of their maximal strips of analiticity, allowed us to deduce their linear independence (both for finite and infinite combinations) and, by a suitable codimension argument, to show that they form a complete set of generators, thus completing the proof that they are a basis of  $L^2(\mathbb{R}/2\pi\mathbb{Z})$ .

3.5.2. Local Birkhoff conjecture for nearly circular domains (Theorem 3.19). The main difficulty in this case – in comparison with the one discussed in Theorem 3.17 and Section 3.5.1 – is that we cannot use the preservation of integrable rational caustics for all rotation number 1/q, with  $q \ge 3$ ; hence, we need to recover the missing conditions on the corresponding Fourier coefficients of the perturbation.

Our key idea is the following: for ellipses of small eccentricity e > 0, we study the Taylor expansion, with respect to e, of the corresponding action-angle coordinates. Using this expansion, we derive the necessary condition for the preservation of integrable rational caustics, in terms of the Fourier coefficients of the perturbation, up to the precision of order  $e^{2N}$ , for some positive integer  $N = N(q_0)$ .

Let us outline our strategy, starting from some special cases.

- CASE  $q_0 = 3$ : We lose a pair of conditions corresponding to Fourier coefficients of order 3. We exploit the conditions obtained from the existence of integrable rational caustics of rotation numbers 1/5, 1/7, 2/7: we use the corresponding expansions, with respect to e, up to the precision  $O(e^6)$ , to derive a system of linear equations for the  $3^{rd}, 5^{th}, 7^{th}$  Fourier coefficients. Solving this linear system will provide us with the needed estimates for Fourier coefficients of order 3.
- CASE  $q_0 = 4$ : In this case we lose two pairs of conditions corresponding to Fourier coefficients of order q = 3, 4. These will be recovered in two steps:
  - To recover the one corresponding to Fourier coefficients of order 3, we study the necessary conditions for the existence of integrable rational caustics of rotation numbers 1/5, 1/7, 1/9, 2/9, written in terms of the Fourier coefficients of the perturbation, and consider their expansions, with respect to e, up to order  $O(e^8)$ . We then derive a linear system for the  $3^{rd}$ ,  $5^{th}$ ,  $7^{th}$ ,  $9^{th}$  Fourier coefficients, whose solution will provide us with the needed estimates for the Fourier coefficients of order 3.
  - To recover the one corresponding to Fourier coefficients of order 4, we study the necessary conditions for the existence of integrable rational caustics of rotation numbers 1/6, 1/8, 1/10, 1/12, 1/14, 3/14, which give rise to a system of linear equation for the 4<sup>th</sup>, 6<sup>th</sup>, 8<sup>th</sup>, 10<sup>th</sup>, 12<sup>th</sup>, 14<sup>th</sup> Fourier coefficients; as before, the solution of this linear system will give us the needed estimates for the Fourier coefficients of order 4.
- THE GENERAL CASE: Along the same lines described in the previous two items, we outlined in [56] a general (conditional) procedure to deal with this problem for any  $q_0 \geq 3$ ; the implementation of this scheme is based on the assumption that certain explicit non-degeneracy conditions for the corresponding linear systems hold. We remark however that all of these conditions are very explicit and the algorithm is explicitly described, so to be implemented on a computer.

3.6. More recent advances on Birkhoff conjecture. We recall here some more recent breakthroughs on Birkhoff conjecture that appeared after the CIME summer school.

(1) In [14] Bialy and Mironov proved the Birkhoff conjecture for *centrally-symmetric*  $C^2$ smooth convex planar billiards. More specifically, they assume that the domain between
the invariant curve the invariant curve foliated by 4-periodic orbits and the boundary
of the phase cylinder is foliated by  $C^0$ -invariant curves and prove that the billiard table
must be elliptic. In [63] the authors proved that this condition is equivalent to having

integrable rational caustics of rotation number  $\frac{1}{q}$ , for every  $q \ge 4$ . The main ingredients of the proof are the use of a non-standard generating function for convex billiards and the observation that invariant curve consisting of 4 -periodic orbits enjoys some special properties; combining these ingredients with the integral-geometry approach for rigidity results that was was introduced by Bialy in [9], they establish a Hopf-type rigidity for billiards in ellipses. See also [15] for an effective version of this result.

- (2) Combining the method of Bialy-Mironov [14] and Kaloshin-Sorrentino [64], recently Kaloshin, Koudjinan and Zhang [63] proved a perturbative version of Birkhoff conjecture nearby centrally symmetric strictly convex domain, under the assumption that the billiard admits integrable rational caustics for rotation numbers  $\frac{1}{3}$  and for all  $0 < \frac{p}{a} \leq \frac{1}{4}$ .
- (3) Recently, Koval [68] proved a Local Birkhoff conjecture for nearly elliptic domains. More precisely, for any positive integer  $q_0$  and any eccentricity e outside of locally finite set in [0, 1), a small  $q_0$ -integrable perturbation of an ellipse of eccentricity e is an ellipse.

3.7. Local period-two Birkhoff Conjecture. Notice that when we discuss integrability in the context of Birkhoff conjecture, usually we refer to integrability near the boundary, namely, integrability for nearly glancing orbits (*i.e.*, caustics with small rotation numbers).

There is an alternative notion of local integrability defined as follows. Fix an ellipse  $\mathcal{E}$  of positive eccentricity and let AB denote its minor axis. This corresponds to an elliptic period-two orbit of the billiard map in  $\mathcal{E}$ . We say that it is locally integrable in the sense that there is a neighborhood foliated by local invariant curves for the square of the billiard map. It turns out that for the square of the billiard map there is a twist. Passing to local polar coordinates  $(r, \varphi) \in (\mathbb{R}_+, \mathbb{T})$  with r = 0 corresponding to say A. Then, one can define the polar rotation number of each of such invariant curves. The billiard trajectories for these invariant curves are tangent to confocal hyperbolae, for this reason we call them hyperbolic caustics. Notice that a hyperbolic caustic consists of two connected components, both on the billiard table  $(s, \theta) \in A$  and on the plane  $\mathcal{E} \subset \mathbb{R}^2$ . Call the projection of a hyperbolic caustic onto the boundary (via the map  $\pi(s, \theta) = s$ ) its support.

A natural extension of the notion of rational integrability is the following one.

Let  $\Omega$  be convex domain of non-constant width has a period two orbit, whose length is strictly less than the diameter of  $\Omega$ . We say that  $\Omega$  period-two locally integrable if there exists  $\delta > 0$ such that in a neighborhood of period-two periodic orbit there exists invariant curves of polar rotation number p/2q for each  $0 < p/2q < \delta$ . It is natural to ask:

**Question 1.** Let  $(\Omega_{\tau})_{\tau}$  be an analytic deformation of an ellipse  $\Omega_0 = \mathcal{E}$  that is period-two locally integrable. Is it true that  $\Omega_{\tau}$  is an ellipse for every  $\tau$ ?

**Question 2.** Let  $\Omega$  be an analytic convex period-two locally integrable domain. Is it true that  $\Omega$  is an ellipse?

In the case of smooth deformations one needs to impose an additional requirement. Let  $\Omega_0$  be an ellipse, and  $\omega_0^+ = \omega^+(\Omega_0)$  be the rotation number of the period two orbit. One can check there is a twist. Then in the polar coordinates all rotation numbers in  $(0, \omega_0^+)$  can be realized near the period-two orbit. For a perturbation  $\Omega$  of  $\Omega_0$ , there exists an interval  $(\omega^-(\Omega), \omega^+(\Omega))$ of rotation numbers that are admissible.

Fix an interval of polar rotation numbers  $[\rho_1, \rho_2]$ ,  $\rho_1 < \rho_2 \subset [0, \omega_0^+]$ . We say that the billiard  $\Omega$  is rationally integrable on  $[\rho_1, \rho_2]$  if for all rational rotation numbers in  $[\rho_1, \rho_2] \cap (\omega^-(\Omega), \omega^+(\Omega))$ , there is a hyperbolic caustic with this rotation number.

Associate to this interval the union of the supports of the associated hyperbolic caustics. Similarly, we can associate to a smooth deformation of an ellipse  $(\Omega_{\tau})_{\tau}$ ,  $\Omega_0 = \mathcal{E}$  its support, namely, the part of the boundary, where  $\Omega_{\tau} \setminus \mathcal{E} \neq \emptyset$ .

**Question 3.** Let  $(\Omega_{\tau})_{\tau}$  be a smooth deformation of an ellipse  $\Omega_0 = \mathcal{E}$  which is  $[\rho_1, \rho_2]$ -integrable for an interval of admissible polar rotation numbers  $\subset [0, \omega_{\tau}]$ . Is it true that  $\Omega_{\tau}$  restricted to the support of the hyperbolic caustics coincide with an ellipse?

3.8. Integrable Riemannian geodesic flows on the torus. We want to conclude this section by drawing some connections between Birkhoff conjecture and a problem in Riemannian geometry. The Birkhoff conjecture can be also thought as an analogue, in the case of billiards, of the following task: classifying integrable (Riemannian) geodesic flows on  $\mathbb{T}^2$ . The complexity of this question, of course, depends on the notion of integrability that one considers. If one assumes that the whole phase space is foliated by invariant Lagrangian graphs (*i.e.*, the system is  $C^0$ -integrable, see [3, Définition 4.19], in particular, the integral of motion is only assumed to be continuous), then it follows from Hopf's result [54] (see also [23] for the proof in dimension greater than 2) that the associated metric must be flat. Bialy and Wojtkowski's results in the billiard setting, can be considered as the analogs of this result.

However, the question becomes more challenging – and it is still open – if one considers integrability only on an open and dense set (global integrability), or assumes the existence of an open set foliated by invariant Lagrangian graphs (local integrability). Example of globally integrable (non-flat) geodesic flows on  $\mathbb{T}^2$  are those associated to *Liouville-type metrics*, namely metrics of the form

$$ds^{2} = (f_{1}(x_{1}) + f_{2}(x_{2}))(dx_{1}^{2} + dx_{2}^{2}).$$

A folklore conjecture states that these metrics are the only globally (resp. locally) integrable metrics on  $\mathbb{T}^2$ , which, in some sense, can be interpreted as the analogue of Birkhoff conjecture, in the realm of integrable geodesic flows on  $\mathbb{T}^2$ .

A partial answer to this conjecture (global case) is provided in [22], where the authors prove it under the assumption that the system admits an integral of motion which is quadratic in the momenta. Observe that while the case of quadratic integral of motion reduces to a system of linear PDEs, the case of higher degree integrals of motions is very challenging and it turns out to be equivalent to delicate questions on non-linear PDEs of hydrodynamic type (see, for example, [12, 13]).

Recently, some advances about deformational rigidity of some Liouville metrics on the torus have been provided by Henheik in [51].

This notion of integrability is related to the so-called *algebraic integrability*, namely the existence of integrals of motion that are polynomial in the velocity. The relation between this notion of integrability and the Birkhoff conjecture (*algebraic Birkhoff conjecture*) has been studied and has lead to interesting results [11, 21]. Recently, using previous results of [11], Glutsyuk [40] proved the algebraic Birkhoff conjecture.

Finally, we point out that the topological structure of the torus plays a fundamentel role in the above-mentioned conjectures and results. For example, on the two dimensional sphere there are plenty of non-trivial integrable metrics: the so-called *Zoll surfaces*. A Zoll surface is a surface homeomorphic to the 2-sphere, equipped with a Riemannian metric all of whose geodesics are closed and of equal length (the first non-trivial example was discovered by Zoll in [113]). While the usual unit-sphere metric on  $\mathbb{S}^2$  obviously has this property, there also exists an infinite-dimensional family of geometrically distinct deformations that are still Zoll surfaces. In particular, most Zoll surfaces do not have constant curvature. See [71] for more details. LECTURE NOTES ON BIRKHOFF BILLIARDS: DYNAMICS, INTEGRABILITY AND SPECTRAL RIGIDIT33

# 4. Lecture IV: Aubry-Mather theory and billiard dynamics

In this section we would like to discuss how the study of *action-minimizing properties* of billiards can be used to shed some light on their dynamical properties. In particular, we shall see how many of the questions discussed in the previous sections can be rephrased in these terms. Let us start by briefly recalling the main ideas at the heart of this approach.

4.1. Aubry-Mather theory for twist maps of the annulus. At the beginning of the eighties Serge Aubry and John Mather developed, independently, what nowadays is commonly called *Aubry–Mather theory*. This novel approach to the study of the dynamics of twist diffeomorphisms of the annulus, pointed out the existence of many *action-minimizing orbits* for any given rotation number. For a more detailed introduction, see for example [36, 95, 97]).

More precisely, let  $a, b \in \mathbb{R} \cup \{\pm \infty\}$ , with a < b, and let

$$f: \mathbb{R}/\mathbb{Z} \times (a, b) \longrightarrow \mathbb{R}/\mathbb{Z} \times (a, b)$$

be a monotone twist map, *i.e.*, a  $C^1$  diffeomorphism such that its lift to the universal cover  $\tilde{f}$  satisfies the following properties (we denote  $(x_1, y_1) = \tilde{f}(x_0, y_0)$ ):

- (i)  $\tilde{f}(x_0 + 1, y_0) = \tilde{f}(x_0, y_0) + (1, 0)$  and  $x_0 \le x_1 < x_0 + 1$ ;
- (ii)  $\tilde{f}$  is orientation preserving and it preserves the boundaries of  $\mathbb{R} \times (a, b)$ :

 $y_1(x_0, y_0) \to a$  as  $y_0 \to a$  and  $y_1(x_0, y_0) \to b$  as  $y_0 \to b$ ;

(iii) If  $a > -\infty$ , then  $\tilde{f}$  extends continuously to  $\mathbb{R} \times \{a\}$  by a rotation:

$$f(x,a) = (x + \omega_{-}, a)$$

similarly, If  $b < +\infty$ , then  $\tilde{f}$  extends continuously to  $\mathbb{R} \times \{b\}$  by a rotation:

$$\tilde{f}(x,b) = (x + \omega_+, b);$$

- (iv)  $\frac{\partial x_1}{\partial y_0} \ge c > 0$  (monotone twist condition),
- (v)  $\tilde{f}$  admits a (periodic) generating function h (*i.e.*, it is an exact symplectic map):

$$y_1 \, dx_1 - y_0 \, dx_0 = dh(x_0, x_1).$$

We call the interval  $(\omega_{-}, \omega_{+}) \subset \mathbb{R}$  the twist interval of f (we remark that if  $a = -\infty$ , then  $\omega_{-} = -\infty$  and if  $b = +\infty$ , then  $\omega_{+} = +\infty$ .

In particular, it follows from (v) that:

(6) 
$$\begin{cases} y_1 = \frac{\partial h}{\partial x_1}(x_0, x_1) \\ y_0 = -\frac{\partial h}{\partial x_0}(x_0, x_1) . \end{cases}$$

**Remark 4.1.** The billiard map is an example of monotone twist map (to fit with the above definition, one can normalize the boundary length to be equal to 1). In particular, as we have already pointed out, its generating function is given by  $h(x_0, x_1) = -\ell(x_0, x_1)$ , where  $\ell(x_0, x_1)$  denotes the euclidean distance between the two points on the boundary of the billiard domain corresponding to  $\gamma(x_0)$  and  $\gamma(x_1)$ .

**Exercise 4.2.** As it follows from Proposition 2.5, a billiard map f associated to a Birkhoff billiard of perimeter 1 and given in  $(s, -\cos \varphi)$ -coordinates is an exact-symplectic twist map: we already saw the existence of the generating function. Prove that it also satisfies the other conditions, in particular the twist condition.

**Exercise 4.3** (Completely Integrable map & standard map). (i) The map  $F : \mathbb{R}^2 \to \mathbb{R}^2$  defined for all (x, y) by

$$F(x,y) = (x+y,y)$$

is the lift to  $\mathbb{R} \times \mathbb{R}$  of an exact-symplectic twist map whose generatin function is the map  $H : (\mathbb{R}/\mathbb{Z})^2 \to \mathbb{R}$  given by

$$H(x, x') := \frac{1}{2}(x' - x)^2.$$

Describe its dynamics and characterize it orbits in terms of their rotation number (compare with the billiard in a disk).

(ii) This example can be generalized as follows. Consider  $v : \mathbb{R} \to \mathbb{R}$ , a 1-periodic smooth map with zero average. Then  $F : \mathbb{R}^2 \to \mathbb{R}^2$  defined for all (x, y) by

$$F(x, y) = (x + y + v(x), y + v(x))$$

is the lift of an exact-symplectic twist map; show that its generating function is given by

$$H(x, x') = \frac{1}{2}(x' - x)^2 + V(x)$$

where  $V : \mathbb{R}/\mathbb{Z} \to \mathbb{Z}$  satisfies V' = v.

As it follows from (6), orbits  $(x_i)_{i \in \mathbb{Z}}$  of the monotone twist diffeomorphism f correspond to 'critical points' of the *action functional* 

$$\{x_i\}_{i\in\mathbb{Z}}\longmapsto\sum_{i\in\mathbb{Z}}h(x_i,x_{i+1})$$

Birkhoff's theorem stated in Theorem 2.6 still holds in this more general setting:

**Theorem 4.4.** Let  $f : \mathbb{A} \to \mathbb{A}$  be an exact-symplectic twist map. Then for any rational  $m/n \in (0,1)$  there exist at least two distinct periodic orbits of rotation number m/n.

We recall this interesting result about invariant curves of twist maps (hence, it applies to billiard maps): the so-called *graph property*. This is a famous result due to Birkhoff [18, 19].

**Theorem 4.5** (Birkhoff invariant curve Theorem). Let  $f: \mathbb{R}/\mathbb{Z} \times (a, b) \longrightarrow \mathbb{R}/\mathbb{Z} \times (a, b)$  be an exact-symplectic twist map. Assume that f admits an embedded, homotopically nontrivial, invariant curve  $\gamma \subset \mathbb{R}/\mathbb{Z} \times (a, b)$ . Then, it the graph of a Lipschitz function.

Numerous proofs of this result have been given, see for instance [36, 57, 65].

These results are the starting point of an important and deep theory called Aubry-Mather theory, developed by S. Aubry and J. Mather in 1980s (see [6, 7, 76]). Aubry-Mather theory is concerned with the study of orbits that minimize this action-functional amongst all configurations with a prescribed rotation number; recall that the rotation number of an orbit  $\{x_i\}_{i\in\mathbb{Z}}$ is given by  $\omega = \lim_{i\to\pm\infty} \frac{x_i}{i}$ , if this limit exists (in the billiard case, this definition leads to the same notion of rotation number introduced in subsection 1.2). In this context, *minimizing* is meant in the statistical mechanical sense, *i.e.*, every finite segment of the orbit minimizes the action functional with fixed end-points.

**Theorem (Aubry** [6, 7], **Mather** [76, 36]) A monotone twist map possesses minimal orbits for every rotation number in its twist interval  $(\omega_{-}, \omega_{+})$ . For rational numbers there are always at least two periodic minimal orbits. Moreover, every minimal orbit lies on a Lipschitz graph over the x-axis. LECTURE NOTES ON BIRKHOFF BILLIARDS: DYNAMICS, INTEGRABILITY AND SPECTRAL RIGIDIT35

We refer to [36, 95, 97] for self-contained presentations on Aubry-Mather theory for twist maps and Hamiltonian flows.

Let us now introduce the minimal average action (or Mather's  $\beta$ -function).

**Definition 4.6.** Let  $x^{\omega} = \{x_i\}_{i \in \mathbb{Z}}$  be any minimal orbit with rotation number  $\omega$ . Then, the value of the *minimal average action* at  $\omega$  is given by (this value is well-defined, since it does not depend on the chosen orbit):

(7) 
$$\beta(\omega) := \lim_{N \to +\infty} \frac{1}{2N} \sum_{i=-N}^{N-1} h(x_i, x_{i+1}).$$

This function  $\beta : \mathbb{R} \longrightarrow \mathbb{R}$  enjoys many properties and encodes interesting information on the dynamics. In particular:

- i)  $\beta$  is strictly convex and, hence, continuous (see [36]);
- ii)  $\beta$  is differentiable at all irrationals (see [78]);
- iii)  $\beta$  is differentiable at a rational p/q if and only if there exists an invariant circle consisting of periodic minimal orbits of rotation number p/q (see [78]).

In particular,  $\beta$  being a convex function, one can consider its convex conjugate:

$$\alpha(c) = \sup_{\omega \in \mathbb{R}} \left[ \omega \, c - \beta(\omega) \right].$$

This function – which is generally called *Mather's*  $\alpha$ -function – also plays an important rôle in the study of minimal orbits and in Mather's theory (particularly in higher dimension, see for example [75, 98]). We refer interested readers to surveys [36, 95, 97].

Observe that for each  $\omega$  and c one has:

$$\alpha(c) + \beta(\omega) \ge \omega c,$$

where equality is achieved if and only if  $c \in \partial \beta(\omega)$  or, equivalently, if and only if  $\omega \in \partial \alpha(c)$ (the symbol  $\partial$  denotes in this case the set of 'subderivatives' of the function, which is always non-empty and is a singleton if and only if the function is differentiable).

4.2. Action-minimizing properties of billiards. In the billiard case, since the generating function of the billiard map is the euclidean distance  $-\ell$ , the action of the orbit coincides – up to a sign – to the length of the trajectory that the ball traces on the table  $\Omega$ . In particular, these two functions encode many dynamical properties of the billiard (see [95] for more details):

- For each  $0 < p/q \le 1/2$ , one has:  $\beta(p/q) = -\frac{1}{q}\mathcal{M}L_{\Omega}^{\max}(p/q)$ .
- $\beta$  is differentiable at p/q if and only if there exists an invariant circle of rotation number p/q foliated by periodic orbits.
- If  $\Gamma_{\omega}$  is a convex caustic with rotation number  $\omega \in (0, 1/2]$ , then  $\beta$  is differentiable at  $\omega$ and  $\beta'(\omega) = -\text{length}(\Gamma_{\omega}) =: -|\Gamma_{\omega}|$  (see [95, Theorem 3.2.10]). In particular,  $\beta$  is always differentiable at 0 and  $\beta'(0) = -|\partial\Omega|$ , where  $|\partial\Omega|$  denotes the length of the boundary of  $\Omega$ .
- If  $\Gamma_{\omega}$  is a convex caustic with rotation number  $\omega \in (0, 1/2]$ , then its Lazutkin invariant  $Q(\Gamma_{\omega})$  (see subsection 3.2) can be related to the value of the  $\alpha$ -function. In fact, one can show that (see [95, Theorem 3.2.10]):

$$Q(\Gamma_{\omega}) = \alpha(\beta'(\omega)) = \alpha(-|\Gamma_{\omega}|).$$



In [95, 96] properties of Mather's  $\beta$  and  $\alpha$  functions have been studied more in depth. In particular, explicit expressions for their (formal) Taylor expansions at, respectively,  $\omega = 0$  and  $c = -|\partial \Omega|$  have been obtained. The coefficients in these expressions will be obtained in terms of the curvature of the boundary and its derivatives.

**Theorem 4.7.** Let  $\Omega$  be a strictly convex planar domain with smooth boundary. Denote by k(s) > 0 the curvature of  $\partial \Omega$  with arc-length parametrization s. Let  $\ell_0 := |\partial \Omega|$  be the length of the boundary and denote:

$$\begin{split} \mathcal{I}_{1} &:= \int_{0}^{\ell_{0}} ds = \ell_{0} \\ \mathcal{I}_{3} &:= \int_{0}^{\ell_{0}} k^{2/3} ds \\ \mathcal{I}_{5} &:= \int_{0}^{\ell_{0}} \left(9 \ k^{4/3} + \frac{8 \ \dot{k}^{2}}{k^{8/3}}\right) ds \\ \mathcal{I}_{7} &:= \int_{0}^{\ell_{0}} \left(9 \ k^{2} + \frac{24 \ \dot{k}^{2}}{k^{2}} + \frac{24 \ \ddot{k}^{2}}{k^{4}} - \frac{144 \ \dot{k}^{2} \ddot{k}}{k^{5}} + \frac{176 \ \dot{k}^{4}}{k^{6}}\right) ds \end{split}$$
$$:= \int_{0}^{\ell_{0}} \left[\frac{281}{44800} k^{8/3} + \frac{281 \ \dot{k}^{2}}{8400 \ k^{4/3}} + \frac{167 \ \ddot{k}^{2}}{4200 \ k^{10/3}} - \frac{167 \ \dot{k}^{2} \ \ddot{k}}{700 \ k^{13/3}} + \frac{\ddot{k}^{2}}{42 \ k^{16/3}} + \frac{559 \ \dot{k}^{4}}{2100 \ k^{16/3}} \\ &- \frac{473 \ \ddot{k}^{3}}{4725 \ k^{19/3}} - \frac{10 \ \ddot{k} \ \dot{k} \ \ddot{k}}{21 \ k^{19/3}} + \frac{5 \ \ddot{k} \ \dot{k}^{3}}{7 \ k^{22/3}} + \frac{13142 \ \dot{k}^{2} \ \ddot{k}^{2}}{4725 \ k^{22/3}} - \frac{10777 \ \dot{k}^{4} \ \ddot{k}}{127575 \ k^{28/3}} \right] ds. \end{split}$$

Then:

 $\mathcal{I}_9$ 

• the formal Taylor expansion of  $\beta$  at  $\omega = 0$ ,  $\beta(\omega) \sim \sum_{k=0}^{\infty} \beta_k \frac{\omega^k}{k!}$ , has coefficients:

$$\begin{split} \beta_{2k} &= 0 \quad for \ all \ k \\ \beta_1 &= -\mathcal{I}_1 \\ \beta_3 &= \frac{1}{4} \ \mathcal{I}_3^3 \\ \beta_5 &= -\frac{1}{144} \ \mathcal{I}_3^4 \ \mathcal{I}_5 \\ \beta_7 &= \frac{1}{320} \ \mathcal{I}_3^5 \left( \frac{14}{81} \mathcal{I}_5^2 - \mathcal{I}_3 \mathcal{I}_7 \right) \ = \ \frac{\mathcal{I}_3^5 \left( 14 \ \mathcal{I}_5^2 - 81 \ \mathcal{I}_3 \mathcal{I}_7 \right)}{25920} \\ \beta_9 &= -7 \ \mathcal{I}_3^6 \left( \mathcal{I}_3^2 \ \mathcal{I}_9 - \frac{1}{5600} \mathcal{I}_3 \ \mathcal{I}_5 \ \mathcal{I}_7 + \frac{7}{583200} \mathcal{I}_5^3 \right); \end{split}$$

• the (formal) Taylor expansion of  $(c + \ell_0)^{-3/2} \alpha(c)$  at  $c = -\ell_0$  (note that  $\alpha$  has in fact a square-root type singularity at the boundary),  $(c + \ell_0)^{-3/2} \alpha(c) \sim \sum_{k=0}^{\infty} \alpha_k \frac{(c+\ell_0)^k}{k!}$ , has

coefficients:

$$\begin{aligned} \alpha_0 &= \frac{4\sqrt{2}}{3} \mathcal{I}_3^{-3/2} \\ \alpha_1 &= \frac{\sqrt{2}}{135} \mathcal{I}_3^{-7/2} \mathcal{I}_5 \\ \alpha_2 &= \frac{1}{56700\sqrt{2}} \left( \frac{72 \mathcal{I}_3 \mathcal{I}_7 + 7 \mathcal{I}_5^2}{\mathcal{I}_3^{-11/2}} \right) \\ \alpha_3 &= \frac{1}{826686000\sqrt{2}} \left( \frac{261273600 \mathcal{I}_3^2 \mathcal{I}_9 + 21384 \mathcal{I}_3 \mathcal{I}_5 \mathcal{I}_7 + 1001 \mathcal{I}_5^3}{\mathcal{I}_3^{-15/2}} \right). \end{aligned}$$

**Remark 4.8.** (i) The techniques used in the proof of the Theorem 4.7, allow one to obtain explicit expressions up to any arbitrary high order (we restrict to order 11 just for the sake of this presentation).

(ii) The coefficients  $\beta_k$  are algebraically related to the set of spectral invariants introduced by Marvizi and Melrose [74] for strictly convex planar regions in order to investigate and give some partial answers to Kac's question on the isospectrality of planar domains. These computations provide explicit expressions for those invariants as well (see the expressions for  $\mathcal{I}_k$ 's).

An easy consequence of these formulae is the following corollary, which is a direct consequence of the isoperimetric inequality (see [96, Corollary 1] and [95]).

**Corollary 4.9.** Let  $\Omega$  be a strictly convex planar domain with smooth boundary. Then:

$$\beta_3 + \pi^2 \beta_1 \le 0$$

and equality holds if and only if  $\Omega$  is a disc.

*Proof.* The proof easily follows from the expressions of  $\beta_1$  and  $\beta_3$ , found in Theorem 4.7. In fact, observe that:

$$\beta_3 + \pi^2 \beta_1 \le 0 \qquad \Longleftrightarrow \qquad \mathcal{I}_3^3 - 4\pi^2 \mathcal{I}_1 \le 0.$$

Now, using Hölder inequality (with  $p = \frac{3}{2}$  and q = 3):

$$\mathcal{I}_3 = \int_0^{\ell_0} k^{2/3} ds \le \left(\int_0^{\ell_0} (k^{2/3})^{3/2} ds\right)^{2/3} \left(\int_0^{\ell_0} 1^3 ds\right)^{1/3} \\ = (2\pi)^{2/3} \ell_0^{-1/3} = (4\pi^2 \mathcal{I}_1)^{1/3}.$$

Moreover, equality holds if and only if it holds in Hölder inequality. This means that k must be constant (and strictly positive) and therefore, the curve must be a circle.

**Remark 4.10.** In particular, the above corollary says that if the first two coefficients  $\beta_1$  and  $\beta_3$  coincide to those of the  $\beta$ -function of a disc, then the domain must be a disc. Therefore, the  $\beta$ -function univocally determines discs amongst all possible Birkhoff billiards. It would be interesting to find a similar characterization for elliptic billiards. We can prove the following result: the  $\beta$ -function determines univocally a given ellipse in the family of all ellipses.

**Proposition 4.11.** If  $\mathcal{E}_1$  and  $\mathcal{E}_2$  are two ellipses such that  $\beta_{\mathcal{E}_1} \equiv \beta_{\mathcal{E}_2}$ , then  $\mathcal{E}_1$  and  $\mathcal{E}_2$  are the same ellipse. More generally: if the Taylor coefficients  $\beta_{\mathcal{E}_1,1} = \beta_{\mathcal{E}_2,1}$  and  $\beta_{\mathcal{E}_1,3} = \beta_{\mathcal{E}_2,3}$ , then the same conclusion remains true.

The proof easily follows from expressing these coefficients by means of elliptic integrals (see [96, Proposition 1])

4.3. Birkhoff conjecture and spectral rigidity questions (revisited). We can now rephrase the Spectral Rigidity Question for the maximal Length spectrum (see subsection 2.2) and Birkhoff Conjecture (see subsection 3.4) in terms of these new objects.

**Spectral Rigidity Question (revisited).** Let  $\Omega_1$  and  $\Omega_2$  be two strictly convex planar domains with smooth boundaries and assume that  $\beta_{\Omega_1} \equiv \beta_{\Omega_2}$ . Is it true that  $\Omega_1$  and  $\Omega_2$  are isometric?

More generally: if  $\beta_{\Omega_1}(\omega) = \beta_{\Omega_2}(\omega)$  for all  $\omega \in (0, \varepsilon)$  for some small  $\varepsilon > 0$ , is it true that  $\Omega_1$  and  $\Omega_2$  are isometric?

Similarly, keeping into account the relation between the differentiability properties of Mather's  $\beta$ -function at rational rotation numbers and the existence of invariant circles foliated by periodic points (see subsection 4.2), we can also rephrase Birkhoff conjecture in this context.

**Birkhoff Conjecture (revisited).** Let  $\Omega$  be a strictly convex planar domain with smooth boundary and assume that  $\beta_{\Omega}$  is differentiable in [0, 1/2). Is it true that  $\Omega$  is an ellipse? More generally: if  $\beta_{\Omega}$  is differentiable in  $[0, \varepsilon)$  for some small  $0 < \varepsilon < 1/2$ , is it true that  $\Omega$  is an ellipse?

In fact, if  $\beta_{\Omega}$  is differentiable in an open interval, then the billiard map is locally integrable in an open set. In fact,  $\beta_{\Omega}$  will be differentiable at all rationals in that interval and therefore there will be caustics corresponding to these rotation numbers. By semi-continuity arguments, one obtains caustics corresponding to irrational rotation numbers and hence a family of caustics that foliate an open set. Observe that if  $\beta$  is differentiable in the whole domain of definition (0, 1/2], then it must be a circle by the aforementioned result by Bialy.

The relation between the integrability of the billiard map and the differentiability of the corresponding Mather's  $\beta$  function, implies that a solution to Birkhoff conjecture would lead to a solution to the question whether ellipses are uniquely spectrally determined among convex domain.

**Exercise 4.12.** Rephrase the results in [8, 64, 68] in terms of Mather's  $\beta$ -function and spectral rigidity of ellipses.

**Remark 4.13.** Compare this result with the previously mentioned result by Hezari and Zelditch [53] (see subsection 2.3), where it is proved that ellipses of sufficiently small eccentricities are Laplace spectrally unique (up to isometry) among all smooth domains (withouth any assumption on symmetry, convexity, or closeness to other ellipses).

#### References

- Naum I. Akhiezer. Elements of the theory of elliptic functions. Translations of Mathematical Monographs, 79. American Mathematical Society, Providence, RI, viii+237, 1990.
- [2] K. G. Andersson and R. B. Melrose. The propagation of singularities along gliding rays. Invent. Math., 41(3): 197–232, 1977.
- [3] Marie-Claude Arnaud. Fibrés de Green et régularité des graphes C<sup>0</sup>-Lagrangiens invariants par un flot de Tonelli. Ann. Henri Poincaré, 9 (5): 881–926, 2008.
- [4] Marie-Claude Arnaud, Jessica Elisa Massetti, Alfonso Sorrentino. On the fragility of periodic tori for families of symplectic twist maps. Preprint, 2023.
- [5] Maxim Arnold and Misha Bialy. Nonsmooth convex caustics for Birkhoff billiards. Pacific J. Math. 295 (2): 257–269, 2018.

- [6] Serge Aubry. The twist map, the extended Frenkel-Kontorova model and the devil's staircase. Phys. D, 7 (1-3): 240–258, 1983.
- [7] Serge Aubry. TThe discrete Frenkel-Kontorova model and its extensions. I. Exact results for the groundstates. *Phys. D*, 8(3): 381–422, 1983.
- [8] Artur Avila, Jacopo De Simoi and Vadim Kaloshin. An integrable deformation of an ellipse of small eccentricity is an ellipse. Ann. of Math. 184: 527–558, 2016.
- [9] Misha Bialy Convex billiards and a theorem by E. Hopf. Math. Z. 124 (1): 147–154, 1993.
- [10] Misha Bialy Hopf rigidity for convex billiards on the hemisphere and hyperbolic plane. Discrete Contin. Dyn. Syst. 33 (9): 3903–3913, 2013.
- [11] Misha Bialy and Andrey Mironov. Angular Billiard and Algebraic Birkhoff conjecture. Adv. in Math. 313: 102–126, 2017.
- [12] Misha Bialy and Andrey Mironov. Cubic and quartic integrals for geodesic flow on 2-torus via a system of the hydrodynamic type. *Nonlinearity* 24 (12) 3541–3554, 2011.
- [13] Misha Bialy and Andrey Mironov. Rich quasi-linear system for integrable geodesic flows on 2-torus. Discrete Contin. Dyn. Syst. 29 (1): 81–90, 2011.
- [14] Misha Bialy and Andrey Mironov. The Birkhoff-Poritsky conjecture for centrally-symmetric billiard tables Annals of Mathematics, 196 (1): 389–413, 2022
- [15] Misha Bialy and Daniel Tsodikovich. Locally maximising orbits for the non-standard generating function of convex billiards and applications. *Nonlinearity* 36 (3), 2023.
- [16] George D. Birkhoff. On the periodic motions of dynamical systems. Acta Math. 50 (1): 359–379, 1927.
- [17] George D. Birkhoff. Collected Mathematical Papers, Vol. II. Providence, RI: Am. Math. Soc., 1950.
- [18] George D. Birkhoff. Surface transformations and their dynamical applications. Acta Math. 43: 1–119, 1922.
- [19] George D. Birkhoff. Sur quelques combes fermées remarquables. Bull. Soc. Math. de France 60: 1–26, 1932.
- [20] V. Blumen, K. Kim, J. Nance, V. Zharnitsky, Three-Period Orbits in Billiards on the Surfaces of Constant Curvature, International Mathematics Research Notices (2012), 10.1093/imrn/rnr228.
- [21] Sergey Bolotin. Integrable Birkhoff billiards. Mosc. Univ. Mech. Bull 45 (2): 10–13, 1990.
- [22] Alexey V. Bolsinov, Anatoly T. Fomenko and Vladimir S. Matveev. Two-dimensional Riemannian metrics with an integrable geodesic flow. Local and global geometries. *Mat. Sb.* 189 (10): 5–32, 1998; translation in *Sb. Math.* 189 (9-10): 1441–1466, 1998.
- [23] Dmitri Burago and Sergei Ivanov. Riemannian tori without conjugate points are flat. Geom. Funct. Anal. 4 (3): 259–269, 1994.
- [24] Keagan G. Callis. Absolutely Periodic Billiard Orbits of Arbitrarily High Order. Preprint, ArXiv: 2209.11721, 2022.
- [25] Jianyu Chen, Vadim Kaloshin and Hong-Kun Zhang. Length Spectrum Rigidity for piecewise analytic Bunimovich Billiards. Preprint 2019.
- [26] C. B. Croke. Rigidity for surfaces of nonpositive curvature. Comment. Math. Helv., 65(1): 150–169, 1990.
- [27] C. B. Croke and V. A. Sharafutdinov. Spectral rigidity of a compact negatively curved manifold. *Topology*, 37 (6): 1265–1273, 1998.
- [28] R. de la Llave, J. M. Marco, and R. Moriyón. Canonical perturbation theory of Anosov systems and regularity results for the Livšic cohomology equation. Ann. of Math. (2), 123(3): 537–611, 1986.
- [29] Jacopo De Simoi, Vadim Kaloshin and Qiaoling Wei, (Appendix B coauthored with H. Hezari) Deformational spectral rigidity among Z<sub>2</sub>-symmetric domains close to the circle. Ann. of Math. 186: 277–314, 2017.
- [30] Jacopo De Simoi, Vadim Kaloshin and Martin Leguil. Marked Length Spectral determination of analytic chaotic billiards with axial symmetries. Preprint 2019.
- [31] Amadeu Delshams and Rafael Ramírez-Ros. Poincaré-Melnikov-Arnold method for analytic planar maps. Nonlinearity 9 (1): 1–26, 1996.
- [32] Colin de Verdière. Sur les longueurs des trajectoires périodiques d'un billard, in Geometrie Symplectique et de Contact: Autour du Theorème de Poincaré-Birkhoff. (P. DAZORD and N. DESOLNEUX-MOULIS, eds.), Hermann, Paris, pp. 122–139, 1984.
- [33] Corentin Fierobe. Complex Caustics of the Elliptic Billiard. Arnold Math. J. 7: 1–30, 2021.
- [34] Corentin Fierobe. On projective billiards with open subsets of triangular orbits. Israel Journal of Mathematics.
- [35] Corentin Fierobe. Examples of reflective projective billiards https://arxiv.org/pdf/2002.09845.pdf
- [36] Giovanni Forni and John N. Mather Action minimizing orbits in Hamiltonian systems. Transition to chaos in classical and quantum mechanics (Montecatini Terme, 1991), Lecture Notes in Math., Vol. 1589: 92–186, 1994.
- [37] Alexey Glutsyuk. On 4-reflective complex analytic billiards. Journal of Geometric Analysis 27: 83–238, 2017.
- [38] Alexey Glutsyuk. On quadrilateral orbits in complex algebraic planar billiards. Moscow Mathematical Journal 14: 239–289, 2014.

- [39] Alexey Glutsyuk. On Odd-periodic Orbits in Complex Planar Billiards. Journal of Dynamical and Control Systems, 20 (3): 293–306, 2014.
- [40] Alexey Glutsyuk. On polynomially integrable Birkhoff billiards on surfaces of constant curvature. *Journal of the European Math. Soc.*, in press, 2020.
- [41] A. A. Glutsyuk, Yu.G. Kudryashov. On quadrilateral orbits in planar billiards. Doklady Mathematics, 83 (3): 371–373, 2011.
- [42] A. A. Glutsyuk, Yu.G. Kudryashov. No planar billiard possesses an open set of quadrilateral trajectories. J. Modern Dynamics, 6 (3): 287–326, 2012.
- [43] Christopher Gole. Symplectic Twist Maps, Global Variational Techniques. Advanced Series in Nonlinear Dynamics Vol. 18, 2001.
- [44] C. Gordon, D. L. Webb, and S. Wolpert. One cannot hear the shape of a drum. Bull. Amer. Math. Soc. (N.S.), 27(1): 134–138, 1992.
- [45] Colin Guillarmou and Thibault Lefeuvre. The marked length spectrum of Anosov manifolds. Ann. of Math. (2) 190 (1): 321–344, 2019.
- [46] V. Guillemin and D. Kazhdan. Some inverse spectral results for negatively curved 2-manifolds. *Topology*, 19 (3): 301–312, 1980.
- [47] V. Guillemin and R. Melrose. The Poisson summation formula for manifolds with boundary. Adv. in Math. 32 (3): 204–232, 1979.
- [48] Eugene Gutkin. Billiard dynamics: a survey with the emphasis on open problems. Regul. Chaotic Dyn. 8 (1): 1–13, 2003.
- [49] Eugene Gutkin and Anatole Katok. Caustics for Inner and Outer Billiards. Commun. Math. Phys., 173: 101–133, 1995.
- [50] Benjamin Halpern. Strange billiard tables. Trans. Amer. Math. Soc. 232: 297–305, 1977.
- [51] Joscha Henheik. Deformational rigidity of integrable metrics on the torus. Preprint, ArXiv: 2210.02961, 2022.
- [52] H. Hezari and S. Zelditch.  $C^{\infty}$  spectral rigidity of the ellipse. Anal. PDE, 5 (5): 1105–1132, 2012.
- [53] Hamid Hezari and Steve Zelditch. One can hear the shape of ellipses of small eccentricity. Annals of Mathematics, 196 (3): 1083–1134, 2022.
- [54] Eberhard Hopf. Closed surfaces without conjugate points. Proc. Nat. Acad. Sci. U. S. A. 34: 47–51, 1948.
- [55] Guan Huang, Vadim Kaloshin and Alfonso Sorrentino Nearly circular domains which are integrable close to the boundary are ellipses. Geom. & Funct. Analysis (GAFA), 28 (2): 334–392, 2018.
- [56] Guan Huang, Vadim Kaloshin and Alfonso Sorrentino On Marked Length Spetrums of Generic Strictly Convex Billiard Tables. Duke Math. J. 167 (1): 175 – 209, 2018.
- [57] A. Katok and B. Hasselblatt, Introduction to the Modern Theory of Dynamical Systems (Encyclopedia of Mathematics and its Applications), Cambridge: Cambridge University Press, 1995.
- [58] Nobohuiro Innami. Convex curves whose points are vertices of billiard triangles. Kodai Math. J., 11: 17–24, 1988.
- [59] Nobohuiro Innami. Geometry of geodesics for convex billiards and circular billiards. Nihonkai Math. J. 13: 73–120, 2002.
- [60] V. Ya. Ivrii. The second term of the spectral asymptotics for a Laplace-Beltrami operator on manifolds with boundary. *Funktsional. Anal. i Prilozhen*, 14 (2): 25–34, 1980.
- [61] M. Kac. Can one hear the shape of a drum? Amer. Math. Monthly, 73 (4, part II): 1–23, 1966.
- [62] V. Kaloshin, I. Koval and A. Vig Wave trace cancellations for hyperbolic orbits in convex nearly elliptic billiard tables. Preprint, 2023.
- [63] V. Kaloshin, C. E. Koudjinan and K. Zhang. Birkhoff Conjecture for nearly centrally symmetric domains. Preprint, 2023.
- [64] Vadim Kaloshin and Alfonso Sorrentino On the local Birkhoff conjecture for convex billiards. Ann. of Math., 188 (1): 315-380, 2018.
- [65] Y. Katznelson and D. S. Ornstein. Twist maps and Aubry-Mather sets. Lipa's legacy, Contemp. Math., 211, Amer. Math. Soc., New York, 1995.
- [66] Anatole Katok, Jean-Marie Strelcyn, François Ledrappier and Feliks Przytycki. Invariant manifolds, entropy and billiards; smooth maps with singularities. *Lecture Notes in Mathematics* 1222, Springer-Verlag, Berlin, viii+283 pp, 1986.
- [67] Oliver Knill. On Nonconvex Caustics of Convex Billiards. Elem. Math. 53: 89–106, 1998.
- [68] I. Koval. Local strong Birkhoff conjecture and local spectral rigidity of almost every ellipse. Preprint, arXiv:2111.12171.
- [69] V. V. Kozlov, D. V. Treshchëv. Billiards, A Genetic Introduction to the Dynamics of Systems with Impacts. Translations of Mathematical Monographs vol. 89, American Mathematical Society, 1991.
- [70] Vladimir F. Lazutkin. Existence of caustics for the billiard problem in a convex domain. (Russian) Izv. Akad. Nauk SSSR Ser. Mat. 37: 186–216, 1973.

- [71] Claude LeBrun and Lionel J. Mason. Zoll manifolds and complex surfaces Journal of Differential Geometry 61 (3): 453–535, 2002.
- [72] Philippe Levallois. Non-intégrabilité des billiards définis par certaines perturbations algébriques d'une ellipse et du flot géodésique de certaines perturbations algébriques d'un ellipsoide. Ph.D. Thesis, Univ. Paris VII, December 1993.
- [73] Philippe Levallois, Mikhail Tabanov. Sparation des séparatrices du billard elliptique pour une perturbation algbrique et symétrique de l'ellipse. C. R. Acad. Sci. Paris Sér. I Math. 316 (6): 589–592, 1993.
- [74] S. Marvizi and R. Melrose. Spectral invariants of convex planar regions. J. Differential Geom. 17: 475–502, 1982.
- [75] D. Massart and A. Sorrentino. Differentiability of Mather's average action and integrability on closed surfaces. *Nonlinearity* 24: 777–1793, 2011.
- [76] John N. Mather. Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. *Topology*, 21 (4): 457–467, 1982.
- [77] John N. Mather. Glancing billiards. Ergodic Theory Dynam. Systems 2 (3-4): 397-403, 1982.
- [78] John N. Mather. Differentiability of the minimal average action as a function of the rotation number. Bol. Soc. Brasil. Mat. (N.S.) 21 (1): 59–70, 1990.
- [79] Jürgen Moser. Selected Chapters of the Calculus of Variations. Lectures in Mathematics, ETH, Zurich, 2003.
- [80] B. Osgood, R. Phillips, and P. Sarnak. Compact isospectral sets of surfaces. J. Funct. Anal., 80(1) :212– 234, 1988.
- [81] B. Osgood, R. Phillips, and P. Sarnak. Extremals of determinants of Laplacians. J. Funct. Anal., 80 (1): 148–211, 1988.
- [82] B. Osgood, R. Phillips, and P. Sarnak. Moduli space, heights and isospectral sets of plane domains. Ann. of Math. (2), 129 (2): 293–362, 1989.
- [83] J.-P. Otal. Le spectre marqué des longueurs des surfaces à courbure négative. Ann. of Math. (2), 131 (1): 151–162, 1990.
- [84] V. Petkov, L. Stojanov. On the number of periodic reflecting rays in generic domains. Ergodic Theory and Dynamical Systems, 8: 81–91, 1988.
- [85] V. M. Petkov and L. N. Stoyanov. Geometry of reflecting rays and inverse spectral problems. Pure and Applied Mathematics (New York). John Wiley & Sons, Ltd., Chichester, 1992.
- [86] Georgi Popov. Invariants of the Length Spectrum and Spectral Invariants of Planar Convex Domains. Commun. Math. Phys., 161: 335–364, 1994.
- [87] G. Popov and P. Topalov. Invariants of Isospectral Deformations and Spectral Rigidity. Communications in Partial Differential Equations, 37 (3): 369–446, 2012.
- [88] G. Popov and P. Topalov. From K.A.M. Tori to Isospectral Invariants and Spectral Rigidity of Billiard Tables. ArXiv e-prints, Feb. 2016.
- [89] Hillel Poritsky, The billiard ball problem on a table with a convex boundary an illustrative dynamical problem, Ann. of Math. 51: 446–470,1950.
- [90] Rafael Ramírez-Ros. Break-up of resonant invariant curves in billiards and dual billiards associated to perturbed circular tables. *Phys. D* 214: 78–87, 2006.
- [91] M.R. Rychlik. Periodic points of the billiard ball map in a convex domain. J. Diff. Geom. 30: 191–205, 1989.
- [92] Yu. Safarov and D. Vassilev. The asymptotic distribution of eigenvalues of partial differential operators. *Translations of Mathematical Monographs* Vol. 155, 1996.
- [93] P. Sarnak. Determinants of Laplacians; heights and finiteness. Analysis, et cetera, pages 601–622. Academic Press, Boston, MA, 1990.
- [94] Yu. Safarov, D. Vassiliev. The Asymptotic Distribution of Eigenvalues of Partial Differential Operators. (American Mathematical Society) 98–106, 1996.
- [95] Karl F. Siburg. The principle of least action in geometry and dynamics. Lecture Notes in Mathematics Vol.1844, xiii+ 128 pp, Springer-Verlag, 2004.
- [96] Alfonso Sorrentino. Computing Mather's beta-function for Birkhoff billiards. Discrete and Continuous Dyn. Systems Series A 35 (10): 5055–5082, 2015.
- [97] Alfonso Sorrentino. Action-Minimizing Methods in Hamiltonian Dynamics. An Introduction to Aubry-Mather Theory. *Mathematical Notes Series* Vol. 50, Princeton University Press, 2015.
- [98] A. Sorrentino and C. Viterbo. Action minimizing properties and distances on the group of Hamiltonian diffeomorphisms. *Geom. Topol.*, 14: 2383–2403, 2010.
- [99] L. Stojanov. Note on the periodic points of the billiard. J. Differential Geom. 34: 835–837, 1991.
- [100] T. Sunada. Riemannian coverings and isospectral manifolds. Ann. of Math. (2), 121 (1): 169–186, 1985.
- [101] Serge Tabachnikov. Billiards. Panor. Synth. No. 1, vi+ 142 pp, 1995.
- [102] Serge Tabachnikov. Geometry and billiards. Student Mathematical Library Vol.30, xii+ 176 pp, American Mathematical Society, 2005.

- [103] Mikhail B. Tabanov. New ellipsoidal confocal coordinates and geodesics on an ellipsoid. J. Math. Sci. 82 (6): 3851–3858, 1996.
- [104] D. Vasiliev, Two-term asymptotics of the spectrum of a boundary value problem in interior reflection of general form, *Funct Anal Appl.* 18 (1984), 267–277.
- [105] Dimitry Treschev, Billiard map and rigid rotation, Phys. D 255: 31–34, 2013.
- [106] M.-F. Vignéras. Variétés riemanniennes isospectrales et non isométriques. Ann. of Math. (2), 112 (1): 21–32, 1980.
- [107] Ya. B. Vorobets. On the measure of the set of periodic points of a billiard. Math. Notes 55: 455-460, 1994.
- [108] H. Weyl. Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung). Math. Ann. 71: 441–479, 1912.
- [109] Maciej P. Wojtkowski, Two applications of Jacobi fields to the billiard ball problem, J. Differential Geom, 40: 155–164, 1994.
- [110] S. Zelditch. Survey of the inverse spectral problem. ArXiv Mathematics e-prints, Feb. 2004.
- [111] S. Zelditch. Inverse spectral problem for analytic domains. II. Z<sup>2</sup>-symmetric domains. Ann. of Math. (2), 170 (1): 205–269, 2009.
- [112] S. Zelditch. Survey on the inverse spectral problem. ICCM Not., 2 (2): 1–20, 2014.
- [113] Otto Zoll. Über Flächen mit Scharen geschlossener geodätischer Linien. (German), Math. Ann., 57: 108– 133, 1903.

INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA, AM CAMPUS 1, 3400 KLOSTERNEUBURG, AUSTRIA

#### $Email \ address: \ \texttt{corentin.fierobekoz} @\texttt{gmail.com} \\$

UNIVERSITY OF MARYLAND, COLLEGE PARK, MD, USA & INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA, AM CAMPUS 1, 3400 KLOSTERNEUBURG, AUSTRIA

#### Email address: vadim.kaloshin@gmail.com

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DEGLI STUDI DI ROMA "TOR VERGATA", ROME, ITALY.

Email address: sorrentino@mat.uniroma2.it