MASTER'S DEGREE THESIS

Master of Science in Advanced Mathematics and Mathematical Engineering

Title: On the number of periodic orbits in billiards with flat points

Author: Aida Chaikh Vidal

Advisors: Pau Martín and Rafael Ramírez

Department: Facultat de Matemàtiques i Estadística

Academic year: 2021 - 2022

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya
Facultat de Matemàtiques i Estadística

Master in Advanced Mathematics and Mathematical Engineering

Master's thesis

On the number of periodic orbits in billiards with flat points

Aida Chaikh Vidal

Supervised by Pau Martín, Rafael Ramírez

January, 2022

Gràcies al Rafa i al Pau per haver dedicat tantes hores a ajudar-me, a l'Alex per introduir-me en el món dels billars i a l'Arturo per animar-me a continuar i confiar en mi.

Abstract

A billiard is a map that describes the motion of a ball without mass in a closed region on the plane such that the collisions with the boundary are elastic. The region where the ball moves is the billiard table. In this thesis, we present the convex billiards (the boundary of the billiard table is a convex Jordan curve of class C^2) and some of their properties. In particular, we will study caustics which are curves that often appear in the billiard problem and they are related with rotational invariant curves (RIC). Lazutkin and Douady proved that convex billiards have caustics if all points of the boundary have curvature strictly positive and the boundary has 6 continuous derivatives. Guktin and Katok, under the hypothesis of Lazutkin and Douady, give estimations of the size of the regions free of caustics contained inside the billiard table. Mather proves the non existence of caustics if there is a flat point in the boundary and Hubacher proves the non existence of caustics close to the boundary if the second derivative of the boundary is not continuous. Finally we do a numerical study about symmetric periodic orbits with odd period and we expose a conjecture that relates the number of symmetric periodic orbits with its period.

Keywords

Convex billiards, twist maps, are-preserving maps, generating function, action functional, invariant curves, caustics, periodic orbits, symmetry.

Contents

1 Introduction **3**

2 Convex Billiards **6**

2.1 Properties . . . 7

2.2 Rotational invariant curves . . . 11

3 Caustics **14**

3.1 Definitions and Properties . . . 14

3.2 Existence of caustics . . . 17

3.3 Regions free of convex caustics . . . 19

3.4 Non existence of caustics - Mather . . . 20

3.5 Non existence of caustics - Hubacher . . . 22

4 Periodic Orbits **27**

4.1 Definitions and Poincaré-Birkhoff theorem . . . 27

4.2 Symmetry . . . 31

5 Numerical results **34**

5.1 Description of the program . . . 35

5.2 Analysis of the results . . . 38

1. Introduction

The aim of this thesis is to present the billiard map and some of its properties. The name of the map appears from the well-known billiard game but it models a system with impacts. Suppose that we have a billiard table and a ball inside this table. The mathematical problem arises when we want study the dynamics of this ball in the table. With this approach we can define the billiard map.

Suppose we have a Jordan curve B on the plane that defines the boundary of the billiard table and a ball without mass as a particle. Moreover, suppose that the motion of the ball is a free motion (without friction). This motion draw imaginary lines inside the table that we called trajectories. On the other hand, the collisions with the boundary are elastic and follow the classical reflection rule:

the angle of incidence is equal to the angle of reflection.

In our case, we are going to study the convex billiard, so we will suppose that the boundary of the table B is a C^2 closed convex curve in the plane. Also, we will suppose that B is orientated counterclockwise.

The dynamical system that we have described is a continuous system but we can do a Poincaré section on the boundary to obtain a discrete dynamical system. This new system consists of the following: two consecutive collisions with the boundary are determined by a trajectory, hence this trajectory can be expressed with a contact point of B and a direction vector that determines the trajectory. So, the billiard map, maps a contact point and a direction vector to obtain the next contact point and the next direction vector after the collision with the boundary. We are going to denote the billiard map as f.

Fixing a point p_0 in B, any point of B can be expressed by its arclength parameter $s \in [0, L)$ where L is the length of B. Moreover, for each point s we can define the angle θ between the tangent vector to B at s orientated counterclockwise and the direction vector of the trajectory. Thus, the pair (s, θ) determines the contact point and the direction vector, then $f : S^1 \times (0, \pi) \to S^1 \times (0, \pi)$ since $(s, \theta) \in [0, L) \times (0, \pi) \cong S^1 \times (0, \pi)$. The set of pairs (s, θ) that generate a billiard trajectory is called orbit.

In section 2, we will see some properties of the billiard map as twist condition or area preserving. On the other hand, the concept of generating function H appears. Let $(s, \theta), (s', \theta')$ be two consecutive points of an orbit. The angle θ, as we have explained above, is determined by the line that joins s and s', so if we have s and s' we do not need the angle to determine the orbit. The generating function defines the orbit implicitly using s an s'. The generating function $H(s, s')$ of the billiard map is the function that gives the euclidean length between s and s' but the generating function is not unique.

Also, we will see the definition of action functional L. The action functional is defined from the generating function as $L : \mathbb{R}^n \to \mathbb{R}$ such that $L(s_1, ...s_n) = \sum_{k=i}^{j-1} H(s_k, s_{k+1})$. An important result that relates the action functional and the billiard map is the following: suppose that we have an orbit and a finite segment S of this orbit, then segment S is a critical point of the action functional.

On the other hand, a classical result is the Birkhoff theorem from which we can deduce that the rotational invariant curves (RICs) that appear in the cylinder $S^1 \times (0, \pi)$ under the billiard map are graphs and separate the cylinder in two connected components, i.e., RICs restrict the behaviour of an orbit inside one connected component.

In section 3 we will define the concept of caustic and we will see some results about the existence and non existence of caustics. Let Γ be a closed curve inside the billiard table. Given an orbit, suppose that this orbit has a trajectory between two consecutive points that are tangent to Γ. We say that Γ is a caustic if all trajectories of two consecutive points are tangent to Γ. Caustics are curves that appear in the billiard problem. We will see that there is a two to one correspondence between RICs that appear in the cylinder $S^1 \times (0, \pi)$ and the caustics that are contained inside the billiard table.

The most known convex billiards are the circumference and the ellipse. When the table is a circumference, all concentric circumferences inside the billiard table are caustics. In the case of the ellipse, all cofocal ellipses inside the ellipse, which is the billiard table, are caustics.

Caustics do not always exist. The classical result of Lazutkin [6] says that if the boundary B of the billiard table has all points with curvature strictly positive and B has 558 continuous derivatives then RICs exist. This results implies that there exist caustics since there is a correspondence two to one between RICs and caustics. The proof consists to do an study near to the curves $S^1 \times \{0\}$ and $S^1 \times \{\pi\}$ and, to do this study, Lazutkin defines new coordinates that we will call Lazutkin coordinates. Later, Douady presents a result that reduces the conditions: caustics exists if all point of B has curvature strictly positive and B has 6 continuous derivatives.

Under the existence hypothesis of caustics, Gutkin and Katok in [3] study the regions free of caustics inside the billiard table and give an estimation of the size of these regions. A simple example to see these results consists to fix a one-parameter family such that the minimum value of the curvature of B tends to 0, for example $B_t = \{(x, y); \in \mathbb{R}^2; x^2 + ty^2 + y^4 \leqslant 1\}$ where $t > 0$. Using the size estimations of the regions free of caustics, we can see that size of these regions increase when the minimum value of curvature tends to 0 and, consequently, the size of the regions that contains caustics decrease.

This study of regions leads us to the result of Mather. Mather in [4] supposes that there is a point of B where the curvature vanishes and B has 6 continuous derivatives. Under this hypothesis, he proves the non existence of caustics inside the billiard table. Since RICs separate the cylinder in two connected components, if an orbit states as close as we want to the boundaries $S^1 \times \{0\}$ and $S^1 \times \{\pi\}$ then RICs cannot exist. Mather proves that if there is a flat point then RICs do not exist, then, by Bikfhoff theorem, there exist orbits that state as close as we want to both boundaries. Thus, we can conclude that RICs exist if and only if B does not have a flat point. In other words, caustics exist if and only if B does not have a point with curvature equal to zero.

Other interesting result was written by Hubacher in [5]. This result proves the non existence of caustics near to B and the hypothesis are all points of B have strictly positive curvature but the second derivative of B is not continuous. The proof is based in the following construction: Fix a point s_0 where the second derivative is no continuous but the lateral limits there exist. Consider two osculating circles that approximate s_0. Suppose that there is a neighborhood of s_0 such that contains points of two different orbits. The first orbit has the point s_0 and the second orbit no. Under this construction, we arrive to a contradiction with the Birkhoff theorem, so there cannot exists any caustic near to the boundary.

In section 4, we will study periodic orbits. Suppose that we have an orbit. Let F be the lift of the billiard map defined in $\mathbb{R} \times (0, \pi)$, the orbit is (p, q) periodic if $F^q(s, \theta) = (s + p, \theta)$ for any point of the orbit (s, θ). The rotation number of this orbit is $\frac{p}{q}$. The Poincaré-Birkhoff theorem gives us the existence of, at least, two geometrically different periodic orbits. In other words, the result says that there exists a (p, q) periodic orbit that minimize the action functional L and this minimum remains for translations. Therefore, since the action functional is a continuous function, any path that connects two minimums has a maximum. Now, we can modify this path and we can choose the path which its maximum is the smallest. This point is a saddle, that is a critical point, so the orbit associated to this point is a (p, q) periodic orbit. Thus, we have the existence of two different periodic orbits.

A type of periodic orbits are symmetric periodic orbits. Suppose the curve B that is the boundary of the billiard table has two perpendicular axis of symmetry. Using this type of curve, we can do a classification of the symmetric periodic orbits. We say that a periodic orbit is symmetric (see Figure 17) if the figure that draw the trajectory of one period of the orbit is symmetric respect one axis of symmetry. Suppose even period. Then, we can distinguish two cases (see Figure 18). The first is when the orbit has a point over an axis of symmetry and the second is the opposite. In both cases, the orbit is symmetric respect both axis of symmetry. Now, suppose odd period. In this case the orbit only can be symmetric respect one axis of symmetry, hence we can distinguish two cases according to the axis of symmetry (see Figure 19).

In the last section, we do a numerical study about symmetric periodic orbits when the curve of the boundary B is

$$C = \{(x, y) \in \mathbb{R}^2; x^2 + y^4 = 1\}.$$

In particular, we focus in periodic orbits that has odd period $(2k + 1, k \in \mathbb{Z})$, its rotation number is $\frac{1}{2k+1}$ and contain the point $(0, 1)$, i.e., they just go one time around the curve B and are symmetric respect y-axis. So from now, a symmetric periodic orbit is a periodic orbit of this type.

The main goal is to find an approximation that relates how many times this symmetric periodic orbits appear depending on their period. Let $2k + 1, k \in \mathbb{Z}$ be the period of a symmetric periodic orbit. It is easy to see that this orbit satisfies that the trajectory between the iterates k and $k + 1$ is horizontal. For this reason, consider the function

$$Z_k : (0, \frac{\pi}{2}) \to \mathbb{R} \quad \text{such that} \quad Z_k = P_4 \circ f^k \circ i,$$

where i is the inclusion, f is the billiard map and P_4 is the projection over the fourth component. So, $Z_k(\theta_0)$ define the initial conditions $q_0 = (0, 1), p_0 = (-\cos(\theta_0), -\sin(\theta_0))$, iterates this initial condition k times with the billiard map and, finally, gives the fourth component that is the second component of the direction vector in the iterate k. Therefore, find symmetric periodic orbits of period $(2k + 1)$ corresponds to find the roots of the function Z_k.

In the numerical study we do an study of the function Z_k. We will see when the value of k increase, then the roots of the function Z_k increase and the intervals where roots are located decrease. With the results of the study, we have the following conjecture:

Conjecture 1.1. *Let $C = \{x^2 + y^4 = 1\}$ be the curve that define the billiard map. The orbits that are periodic with period $2k + 1$, symmetric respect the y-axis with initial position $(0, 1)$ and the rotation number is $\frac{1}{2k+1}$ satisfy that there exists a constant α such that*

$$n(k) \approx \alpha k \quad \text{with} \quad \alpha \approx 0.41,$$

where $n(k)$ is the lower bound of the quantity of orbits described above. There exists constants a_*, b_* such that the orbits are contained in the interval $[a(k), b(k)]$ where

$$a(k) \approx \frac{a_*}{k} \quad \text{and} \quad b(k) \approx \frac{b_*}{k} \quad \text{with} \quad a_* \approx 0.91, b_* \approx 1.31.$$

2. Convex Billiards

The aim of this section is to describe the dynamic of billiards in mathematical terms using a physical approach. In our case, we are going study convex billiards. First, we will define the billiard map and we will see some properties that it satisfies, for example the twist condition and the area preservation. Finally, we are going to explain the concept of invariant curves and how the existence of these curves affects the dynamics in our map.

Suppose we have a bounded region in the plane that we called billiard table and the boundary of this region is a regular Jordan curve. The regularity requirements will be fixed later. The billiard map represents the free motion of the ball inside the table (without friction). The trajectory of this ball is given by imaginary straight lines inside the table. When the ball hits the border of the table, the angle of reflection is equal to the incidence angle. In other words, the assumptions are that the collisions with the boundary are plastic and follow the classical reflection rule.

Below we present a mathematical definition of the idea that we have just explained.

Let D be an open convex region in the plane and B its boundary. The main conditions that B has to fulfill is to be a closed regular C^2 curve and we can consider that B is orientated counterclockwise. Let L be the length of B. We denote by C the set of pairs $(p, v) \in B \times S^1$ where p is a point of B and v is a unit vector with footpoint p, directed inward.

Fixing an initial point $p_0 \in B$, for any point $p \in B$ we assign its arclength parameter $s \in [0, L]$ as the counterclockwise distance along the curve from p_0 to p so $s \in [0, L)$. On the other hand, suppose t be a positive vector tangent to B at p, i.e., the vector has the same orientation as the curve B . Then the angle θ between v and t is unique. This implies that any pair $(p, v) \in C$ has a unique representative pair $(s, \theta) \in [0, L) \times (0, \pi)$. Indeed $C \cong S^1 \times (0, \pi)$ since $B \cong S^1$.

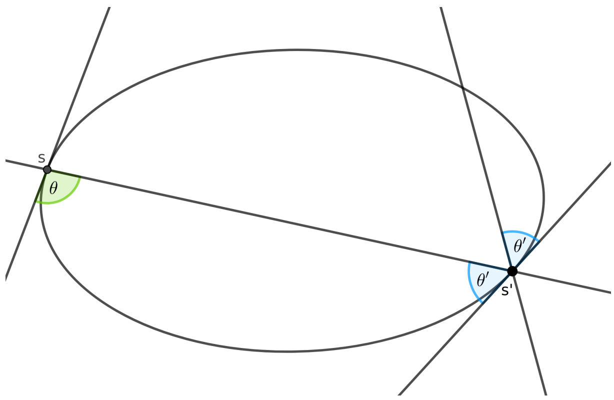


Figure 1: Billiard dynamics

The billiard map is defined as a free motion of a free mass point with the elastic reflections from the boundary, in other words, the rule that it satisfies is that the angle of reflection is equal to the angle of incidence. With this assumption we can determine the successive points through the following process that we can see in Figure 1: fixing a point $s \in S^1$ and an angle $\theta \in (0, \pi)$, we can construct an line l such that $s \in l$ and θ is the angle between l and the positive vector tangent of the boundary in s. This line intersects B in other point s' and we can consider the angle θ' between the line and the tangent vector of B in s'. Then, the billiard map maps (s, θ) to (s', θ') and, to find the next point, we consider (s', θ') and we must repeat the same process.

The map $f : S^1 \times (0, \pi) \to S^1 \times (0, \pi)$ such that $f(s, \theta) = (s', \theta')$ is called billiard map. We can express it as $f(s, \theta) = (S(s, \theta), \Theta(s, \theta))$, where S and Θ are two maps.

Remember that the phase space associated to the map f is the set of all possible values (s, θ). Hence the phase space of the billiard map is the cylinder $S^1 \times (0, \pi)$.

Let $f : S^1 \times (0, \pi) \to S^1 \times (0, \pi)$ be a billiard map. It will be considered the lift of this map, F, so the following diagram commutes:

$$\begin{tikzcd} \mathbb{R} \times (0, \pi) \arrow[r, "F"] \arrow[d, "\pi"] & \mathbb{R} \times (0, \pi) \arrow[d, "\pi"] \\ S^1 \times (0, \pi) \arrow[r, "f"] & S^1 \times (0, \pi) \end{tikzcd}$$

where $\pi : \mathbb{R} \to S^1$ is the projection function over S^1.

In other words, the lift is defined as:

$$\begin{aligned} F : \mathbb{R} \times (0, \pi) &\to \mathbb{R} \times (0, \pi) \\ (s, \theta) &\mapsto (S(s, \theta), \Theta(s, \theta)) \end{aligned} \tag{1}$$

such that $f(s, \theta) = (S(s, \theta) \mod L, \Theta(s, \theta))$.

From now we are going to use the same notation $S(s, \theta)$ to indicate the first term of the functions f and its lift F. Remember that the function S is defined in $\mathbb{R}$ if we use the lift, and it is defined in $\mathbb{R}/L\mathbb{Z}$ if we use f.

2.1 Properties

The billiard map satisfies some very restrictive conditions. Many of their properties are a consequence at the fact that it is a twist map, as we will see. in particular, it will allow us define the generating function. In Section 2.2, we are going to use some properties to apply the Birkhoff's Theorem in the billiard map.

Proposition 2.1. *Let $\alpha : \mathbb{R} \to \mathbb{R}^2$, $\alpha(s) = (\alpha_1(s), \alpha_2(s))$, be an arclength parameterization of B. The map (1) is a differentiable map.*

Proof. Remember that the next point $S(s, \theta)$ under F states over the curve B, then

$$\begin{aligned} &[\alpha_1(S(s, \theta)) - \alpha_1(s)] [\sin(\theta)\dot{\alpha}_1(s) + \cos(\theta)\dot{\alpha}_2(s)] = \\ &= [\alpha_2(S(s, \theta)) - \alpha_2(s)] [\cos(\theta)\dot{\alpha}_1(s) - \sin(\theta)\dot{\alpha}_2(s)] . \end{aligned}$$

Moreover, two consecutive points are in the same period, $S(s, \theta) - s < L$. Hence, for any pair (s, θ) there exists a unique $s' \in [s, s + L]$ such that

$$\begin{aligned} E(s', s, \theta) := &\det(\alpha(s') - \alpha(s) | R_\theta(\dot{\alpha}(s))) \\ &= \left[\alpha_1(s') - \alpha_1(s)\right] [\sin(\theta)\dot{\alpha}_1(s) + \cos(\theta)\dot{\alpha}_2(s)] \\ &- \left[\alpha_2(s') - \alpha_2(s)\right] [\cos(\theta)\dot{\alpha}_1(s) - \sin(\theta)\dot{\alpha}_2(s)] = 0, \end{aligned}$$

where $R_\theta(\dot{\alpha}(s))$ is the tangent vector at s in B with a rotation of angle θ.

Let (s_1, s_0, θ_0) be a solution of the equation $E(s', s, \theta) = 0$. Moreover,

$$\frac{\partial E}{\partial s'}(s', s, \theta) = \sin(\theta') \Rightarrow \frac{\partial E}{\partial s'}(s', s, \theta) \neq 0, \text{ since } \theta' \in (0, \pi).$$

By the implicit function theorem, there exists C_0 neighborhood of (s_0, θ_0) such that $S(s, \theta)$ is a differentiable map. We can do this reasoning for any pair $(s, \theta) \in \mathbb{R} \times (0, \pi)$, so $S(s, \theta)$ is a differentiable map in $\mathbb{R} \times (0, \pi)$.

On the other hand, the new angle $\Theta(s, \theta)$ satisfy the reflection rule, i.e.,

$$\begin{aligned} \cos(\Theta(s, \theta)) = &\left[\cos(\theta)\dot{\alpha}_1(s) - \sin(\theta)\dot{\alpha}_2(s)\right] \dot{\alpha}_2(S(s, \theta)) \\ &+ \left[\sin(\theta)\dot{\alpha}_1(s) + \cos(\theta)\dot{\alpha}_2(s)\right] \dot{\alpha}_1(S(s, \theta)) \end{aligned}$$

so we have

$$\begin{aligned} \Theta(s, \theta) = \arccos[&(\cos(\theta)\dot{\alpha}_1(s) - \sin(\theta)\dot{\alpha}_2(s))\, \dot{\alpha}_2(S(s, \theta)) \\ &+ (\sin(\theta)\dot{\alpha}_1(s) + \cos(\theta)\dot{\alpha}_2(s))\, \dot{\alpha}_1(S(s, \theta))]. \end{aligned}$$

where α, sin, cos and arccos are differentiable maps in our domain, then, the map Θ is a differentiable map.

So, F is a differentiable map since S and Θ are too. $\square$

A direct consequence of this result is that the billiard map f is a differentiable map.

Let $f(s, \theta) = (f_1(s, \theta), f_2(s, \theta))$ be a differentiable map over the cylinder and area preserving. We say that f is twist if

$$\frac{\partial f_1}{\partial \theta}(s, \theta) > 0 \quad (\text{or} < 0),$$

for all pairs (s, θ) in the cylinder, i.e., f_1 is a monotone function of θ. If it is monotonically increasing then f has positive twist and it is negative twist if f is monotonically decreasing. Hence, if we choose a vertical line, after one iterate, the result is a graph that tilts to the right (positive twist) or to the left (negative twist), as we can see in Figure 2.

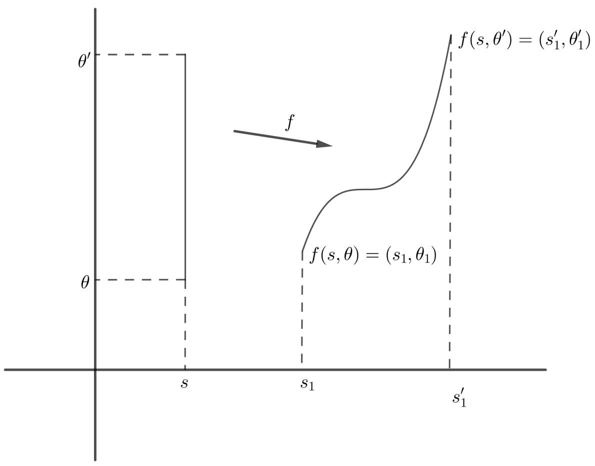


Figure 2: Representation of positive twist condition.

Proposition 2.2. *The billiard map is a positive twist map, i.e., for all pairs* $(s, \theta) \in S^1 \times (0, \pi)$ *we have* $\frac{\partial S}{\partial \theta}(s, \theta) > 0$.

Proof. Let α be an arclength parameterization of B. Fixing $s_0 \in [0, L)$, $\dot{\alpha}(s_0)$ is the tangent vector at s_0 and, for any $\theta \in (0, \pi)$, $v = (\dot{\alpha}_1(s_0)\cos(\theta) - \dot{\alpha}_2(s_0)\sin(\theta), \dot{\alpha}_1(s_0)\sin(\theta) + \dot{\alpha}_2(s_0)\cos(\theta))$ is the director vector of the trajectory.

Suppose $s_1 = S(s_0, \theta)$, then,

$$\begin{aligned} E(s_1, s_0, \theta) &= (\alpha_1(s_1) - \alpha_1(s_0))(\dot{\alpha}_1(s_0)\sin(\theta) + \dot{\alpha}_1(s_0)\cos(\theta)) \\ &\quad - (\alpha_2(s_1) - \alpha_2(s_0))(\dot{\alpha}_1(s_0)\cos(\theta) - \dot{\alpha}_2(s_0)\sin(\theta)) = 0. \end{aligned}$$

and, this implies

$$E(S(s_0, \theta), s_0, \theta) = 0 \Rightarrow \frac{\partial E}{\partial \theta}(S(s_0, \theta), s_0, \theta) + \frac{\partial E}{\partial S}(S(s_0, \theta), s_0, \theta)\frac{\partial S}{\partial \theta}(s_0, \theta) = 0.$$

Computing the last equation we obtain the following result:

$$\frac{\partial E}{\partial \theta}(S(s_0, \theta), s_0, \theta) + \frac{\partial E}{\partial S}(S(s_0, \theta), s_0, \theta)\frac{\partial S}{\partial \theta}(s_0, \theta)) = -\frac{\partial S}{\partial \theta}(s_0, \theta)\sin\theta_1 + h = 0$$

where h is the Euclidean distance between s_0 and $S(s_0, \theta)$ and $\theta_1 = \Theta(s_0, \theta)$. So,

$$\frac{\partial S}{\partial \theta}(s_0, \theta) = \frac{h}{\sin\theta_1} > 0$$

since $\sin\theta_1 > 0$ i $h > 0$ for all $\theta_1 \in (0, \pi)$. □

The following definition introduces the Birkhoff's coordinates. These new coordinates are useful to prove some properties as area preservation.

Definition 2.3. Let $f(s, \theta) = (S(s, \theta), \Theta(s, \theta))$ be the billiard map, we consider $g : S^1 \times (0, \pi) \to S^1 \times (-1, 1)$ the map of change of coordinates such that $g(s, \theta) = (s, r)$ where $r = \cos\theta$.

The map $\bar{f}(s, r) = (S(s, r), R(s, r))$, where S and R are differentiable maps, is called the billiard map in Birkhoff's coordinates.

Definition 2.4. Let $f : S^1 \times (-1, -1) \to S^1 \times (-1, -1)$ be a map. We say $H : S^1 \times S^1 \to (-1, -1)$ is a generating function of f if for all pair $(s, r) \in S^1 \times (-1, -1)$ such that $f(s, r) = (s', r')$, it holds

$$\frac{\partial H}{\partial s}(s, s') = -r \qquad \text{and} \qquad \frac{\partial H}{\partial s'}(s, s') = r'. \tag{2}$$

The generating function H defines implicitly one orbit of f using only two consecutive points of the trajectory on the boundary, so H defines implicitly f.

Proposition 2.5. *A generating function of the billiard map in Birkhoff coordinates* (s, r) *is*

$$H(s, s') = [(\alpha_1(s') - \alpha_2(s))^2 + (\alpha_2(s') - \alpha_1(s))^2]^{\frac{1}{2}},$$

where $\alpha(s)$ *is the arclength parameterization of* B*. The function* H *is the length between two consecutive impact points.*

Proof. Computing the partial derivatives we have

$$\begin{aligned}\frac{\partial H}{\partial s'}(s,s') &= \frac{2[(\alpha_1(s')-\alpha_1(s))^2\dot{\alpha}_1(s')][(\alpha_2(s')-\alpha_2(s))^2\dot{\alpha}_2(s')]}{2[(\alpha_1(s')-\alpha_1(s))^2+(\alpha_2(s')-\alpha_2(s))^2]^{\frac{1}{2}}}\\ &= \frac{\langle(\alpha(s')-\alpha(s)),\dot{\alpha}(s')\rangle}{[(\alpha_1(s')-\alpha_1(s))^2+(\alpha_2(s')-\alpha_2(s))^2]^{\frac{1}{2}}}\\ &= r'\end{aligned}$$

and

$$\begin{aligned}\frac{\partial H}{\partial s}(s,s') &= \frac{2[(\alpha_1(s')-\alpha_1(s))^2(-\dot{\alpha}_1(s))][(\alpha_2(s')-\alpha_2(s))^2(-\dot{\alpha}_2(s))]}{2[(\alpha_1(s')-\alpha_1(s))^2+(\alpha_2(s')-\alpha_2(s))^2]^{\frac{1}{2}}}\\ &= \frac{\langle-(\alpha(s')-\alpha(s)),\dot{\alpha}(s)\rangle}{[(\alpha_1(s')-\alpha_1(s))^2+(\alpha_2(s')-\alpha_2(s))^2]^{\frac{1}{2}}}\\ &= -r.\end{aligned}$$

So, the function H satisfies the equations (2). □

Proposition 2.6. *The billiard map $\bar{f}$ preserves the area form $ds \wedge dr$.*

Proof. To see that $\bar{f}$ preserves $ds \wedge dr$ we have to prove $det(J(\bar{f})) = 1$. Remember $\bar{f}$ is the billiard map in Birkhoff's coordinates.

First, we have

$$det(J(\bar{f})) = \frac{\partial S}{\partial s}(s,r)\frac{\partial R}{\partial r}(s,r) - \frac{\partial S}{\partial r}(s,r)\frac{\partial R}{\partial s}(s,r).$$

On the other hand, using the generating function defined above, we can define a new function as $\tilde{H}(s,r) := H(s,S(s,r))$ and we have the following expressions:

$$\begin{aligned}\frac{\partial \tilde{H}}{\partial s}(s,r) &= \frac{\partial H}{\partial s}(s,S(s,r)) + \frac{\partial H}{\partial S}(s,S(s,r))\frac{\partial S}{\partial s}(s,r) = -r + R(s,r)\frac{\partial S}{\partial s}(s,r),\\ \frac{\partial \tilde{H}}{\partial r}(s,r) &= \frac{\partial H}{\partial S}(s,S(s,r))\frac{\partial S}{\partial r}(s,r) = R(s,r)\frac{\partial S}{\partial r}(s,r)\end{aligned}$$

then,

$$\begin{aligned}\frac{\partial^2 \tilde{H}}{\partial s \partial r}(s,r) &= -1 + \frac{\partial R}{\partial r}(s,r)\frac{\partial S}{\partial s}(s,r) + R(s,r)\frac{\partial^2 S}{\partial s \partial r}(s,r)\\ \frac{\partial^2 \tilde{H}}{\partial r \partial s}(s,r) &= \frac{\partial R}{\partial s}(s,r)\frac{\partial S}{\partial r}(s,r) + R(s,r)\frac{\partial^2 S}{\partial r \partial s}(s,r).\end{aligned}$$

By the Schwarz's lemma,

$$\frac{\partial^2 \tilde{H}}{\partial s \partial r}(s,r) = \frac{\partial^2 \tilde{H}}{\partial r \partial s}(s,r) \Rightarrow det(J(\bar{f})) = \frac{\partial S}{\partial s}(s,r)\frac{\partial R}{\partial r}(s,r) - \frac{\partial S}{\partial r}(s,r)\frac{\partial R}{\partial s}(s,r) = 1.$$

□

Until now, we have seen that the billiard map $f : S^1 \times (0, \pi) \to S^1 \times (0, \pi)$ (in natural coordinates) is a twist map and the billiard map $\bar{f} : S^1 \times (-1, 1) \to S^1 \times (-1, 1)$ (in Birkhoff coordinates) is area preserving. There is a homeomorphism between the intervals $(0, 1)$ and $(-1, 1)$ so we have that both maps are symplectic twist maps.

Definition 2.7. If we fix initial values $(s_0, r_0) \in S^1 \times (-1, 1)$, we say that the set of ordered pairs $\mathcal{O}(s_0, \theta_0) = \{(s_n, \theta_n), n \in \mathbb{Z}\}$ is a orbit of f if $(s_n, \theta_n) = f^n(s_0, \theta_0)$.

Definition 2.8. The function

$$L : \mathbb{R}^n \to \mathbb{R} \text{ such that } L(s_1, ...s_n) = \sum_{k=i}^{j-1} H(s_k, s_{k+1})$$

is called the action functional, where H is a generating function.

The next proposition relates the concept of orbit with the action functional.

Proposition 2.9. *Given $s_n, s_m \in S^1$, the sequence $\{s_n, ..., s_m\} \in S^1$ is a segment of an orbit if and only if $\{s_n, ..., s_m\}$ is a critical point of the action functional $L(s_n, ..., s_m)$ with fixed endpoints s_n and s_m.*

Proof. Fixing $s_n, s_m \in S^1$, the action functional of the segment $\{s_n, ..., s_m\}$ is

$$L(s_n, ..., s_m) = \sum_{i=n}^{m-1} H(s_i, s_{i+1}).$$

Indeed, the segment is a critical point of the action functional if and only if

$$\frac{\partial L}{\partial s_j}(s_n, ..., s_m) = 0 \quad \forall j \in \{n+1, ..., m-1\},$$

and, using the definition of the map L, this implies,

$$\frac{\partial H}{\partial s'}(s_{j-1}, s_j) + \frac{\partial H}{\partial s}(s_j, s_{j+1}) = 0 \quad \forall j \in \{n+1, ..., m-1\}.$$

So, the segment of the orbit is a critical point of L if and only if

$$r_j := \frac{\partial H}{\partial s'}(s_{j-1}, s_j) = -\frac{\partial H}{\partial s}(s_j, s_{j+1}) \quad \forall j \in \{n+1, ..., m-1\},$$

in other words, the equations (2) are satisfied. $\square$

2.2 Rotational invariant curves

Let $F : S^1 \times (-1, 1) \to S^1 \times (-1, 1)$ be an area preserving map on the cylinder.

Definition 2.10. F is end preserving if it satisfies

$$\lim_{r \to -1} f_2(s, r) = -1,$$
$$\lim_{r \to 1} f_2(s, r) = 1,$$

where $F(s, r) = (f_1(s, r), f_2(s, r))$.

Definition 2.11. A curve C is an invariant curve under the map F if $F(C) = C$. Moreover, if this curve is a loop that encircles the cylinder we say that C is a rotational invariant curve (RIC). In other words, a RIC can not be contracted to a point.

Since the map is continuous, end preserving and bijective, this type of curves separate the cylinder in two invariant regions, i.e., they provide an absolute barrier to motion.

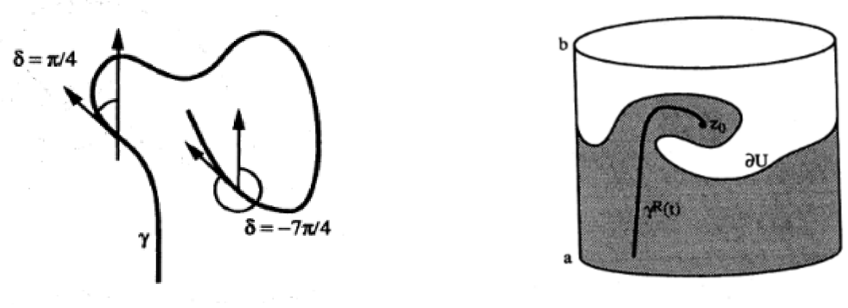
Theorem 2.12. *(Birkhoff's theorem) Let F be a C^1 area preserving, end preserving twist map on the cylinder. Let U be an open invariant set homeomorphic to the cylinder such that there are $a, b \in \mathbb{R}$ with $a < b$ satisfying the following condition*

$$\{(x, y); y < a\} \subset U \subset \{(x, y); y < b\}.$$

Then, the boundary of U is a graph $\{x, Y(x)\}$ of some Lipschitz function Y.

Proof. To do the proof we will follow [1]. First, we present some concepts.
Suppose $\gamma(t) = (x(t), y(t))$ is a curve embedded in U parametrized by $t \in (\alpha, \beta)$ such that $\lim_{t\to\alpha} y(t) = -1$. The deviation of γ from the vertical is defined to be the angle δ between a tangent to γ and the vertical. For those points $\gamma(t)$ such that $y(t) > y(t')$ for all $t' < t$, choose δ in range $[-\frac{\pi}{2, \frac{\pi}{2}}]$; otherwise the branch of δ is chosen to make the deviation a continuous function. (See Figure 3, left image).

Consider two types of curves γ, depending of the range of δ. If $\delta \leqslant 0$ for all points along the curve, we are going to call it γ^R and if $\delta \geqslant 0$ we will denote it γ^L. This means the deviation of γ^R from the vertical is to the right for all points of the curve and the deviation is to the left when $\gamma = \gamma^L$.



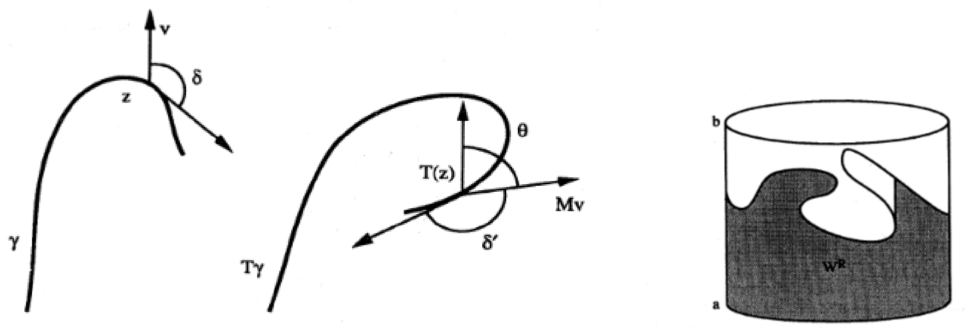
(a) Example of γ that tilts to right. (b) Example of right accessible point z_0.

Figure 3: Images of [1]

Let $z_0 \in U$ be a point. We say that z_0 is right accessible if there exists a $\gamma^R \in U$ such that $\gamma^R(t_0) = z_0$. (See Figure 3, right image)

An interesting property is that this curves under the map F conserve its tilt, i.e., the iterate of the curve that tilts to the right also tilts to the right. To see this, we choose a point z on the curve and the vertical vector v at z. Consider M the linearization of F and the new vector Mv. The twist condition implies that v is mapped to a right-tilt vector and the angle θ between the vertical vector and Mv has a range $[-\pi, 0]$. On the other hand, F preserves orientation, so the angle between Mv and the tangent vector of $F(\gamma^R)$ at $F(z)$ is contained in $[-\pi, 0]$. Therefore, $F(\gamma^R)$ tilts to the right. Similarly, $F(\gamma^L)$ tilts to the left. (See Figure 4, left image).

Suppose W^R and W^L the sets of points that are contained in U and are right and left accessible respectively. The boundary of W^R are the points of the boundary of U that are right accessible together the vertical segments in U that they delimit the parts of U not right accessible (see Figure 4, right image). All points of W^R are on a curve that tilts to right, hence, by the previous property, W^R is mapped to itself ($F(W^R) \subset W^R$). Analogously, if we consider the map F^{-1} we have the twist condition to the left and $F^{-1}(W^L) \subset W^L$.



(a) Tilt property. (b) Region W^R.

Figure 4: The map T in image (a) corresponds to the map F. Images of [1]

Now, we are going to see $W^R = U$. Suppose $W^R \neq U$, then there is a portion of U that is not right accessible and there is a "lobe" bounded by a vertical segment on the right. The points of this vertical segment are mapped into a portion of W^R. On the other hand, we consider a circle $y = y_0$ far below the boundary of U. It is clear that the area of U above y_0 is finite since the points (x, y) of the boundary of U satisfy $y < a$ and $y > b$. Then, using the area preserving property of F, we have that this finite area is mapped to a regions with the same area. Moreover, by the end preserving condition, the area contained under the curve $y = y_0$ is the same under iterations of F and this gives a contradiction. Using the same arguments with the map F^{-1} we have $W^L = U$.

Thus, all points of the boundary of U are right and left accessible so there exists a function $y = Y(x)$ such that defines the boundary of U.

The Lipschitz condition to the graph $y = Y(x)$ is proved in [1]. $\square$

Observation 2.13. The billiard map f satisfies the conditions of the theorem, so if the cylinder where f is defined has RICs, these RICS are graphs and these graphs are Lipschitz.

3. Caustics

Caustics are interesting curves that appears in the billiard problem. They are the "physical counterpart" of the RICs in the billiard map. In general, there is a two to one correspondence between RICs and caustics. In this section we will see some properties about the caustics and the conditions that we need to guarantee the existence of caustics in the convex billiards. The most known examples when the caustics appear are circumferences and ellipses. We will see if the boundary of the billiard table is a circumference, all concentric circumferences inside the table are caustics. On the other hand, is the boundary of the table is an ellipse, all cofocal ellipses inside the table are also caustics. The conditions of the existence of caustics in convex billiards was proved by Lazutkin and Douady. On the other hand, caustics do not exist always. in this seccion we are going to see how Mather proves the non existence of caustics when the boundary of the table has a point where the curvature vanishes and how Hubacher proves the non existence of caustics when the boundary of the table is no sufficiently smooth.

3.1 Definitions and Properties

Let B be the boundary of the billiard table and (s, θ) a point of a billiard's trajectory.

Definition 3.1. Given Γ a curve and (s, θ) a point of an orbit such that the line drawn by the points (s, θ) and its next iterate is tangent to Γ. The curve Γ is called caustic if all of the segments generated by the orbit (s, θ) are tangent to Γ too. Moreover, we say that the caustic is convex if the curve Γ is a smooth closed convex curve contained in the billiard table.

The convex caustics are related with the RICs that we can find in the phase space $[0, L) \times [0, \pi]$. From now, the phase space is defined in $[0, L) \times (0, \pi)$ But we can extend the interval $(0, \pi)$ to $[0, \pi]$ by continuity.

Proposition 3.2. *If Γ is a convex caustic, there exist two RICs in the phase space that corresponds with the orbits that are tangent to the caustic.*

Proof. Given s_0 a point of the curve B we can consider two lines l_1 and l_2 that are tangent to Γ and go through the point s. We define θ_0 the angle between the line l_1 and the tangent line to B at s_0 (see Figure 5).

If we repeat this process for all points s_0 of B, in the phase space appears a rotational curve C in the cylinder $S^1 \times [0, \pi]$. We are going to see that C is invariant: choosing $(s_0, \theta_0) \in C$ and its next iterate (s_1, θ_1) under the billiard map, the line $\overline{s_0 s_1}$ is tangent to Γ and Γ has an other tangent line l_3 that contains s_1. Since Γ is a caustic, the reflection angle θ_1 in s_1 coincides with the angle between the tangent line to B at s_1 and the line l_3, so the point (s_1, θ_1) belongs to C. So C is a RIC.

Analogously, when we fix the point s_0, we can choose θ_0 the angle between the line l_2 and the tangent line to B at s_0. As above, in the phase space appears a curve $\bar{C}$ that is a RIC. □

Observation 3.3. The RIC C and $\bar{C}$ satisfy the relation

$$(s_0, \theta_0) \in C \Leftrightarrow (s_0, \pi - \theta_0) \in \bar{C}.$$

Proposition 3.4. *If there exists a RIC γ sufficiently near to the boundary $S^1 \times \{0\}$ or $S^1 \times \{\pi\}$, there is a caustic Γ such that the trajectories generated by the points of the RIC are tangent to this caustic.*

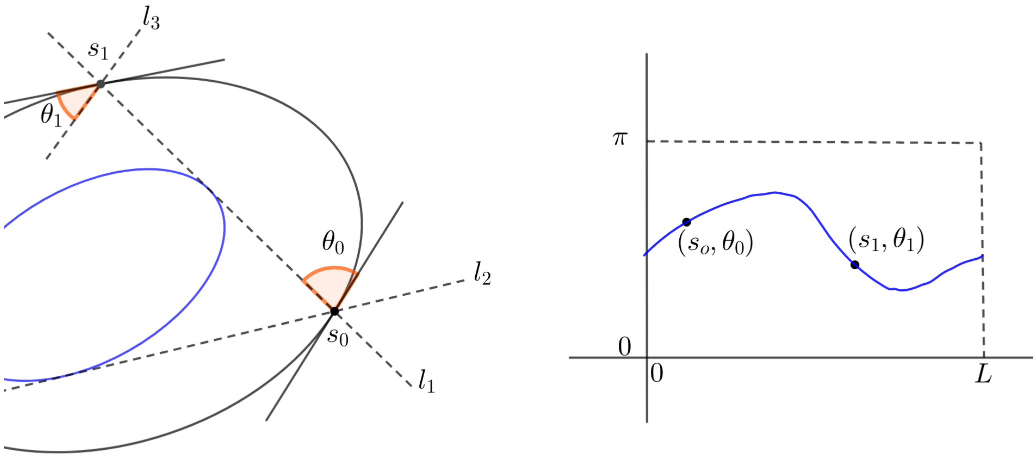


Figure 5: A caustic and its associated RIC

Proof. We recall that an envelope of an uniparametric family of planar curves is a curve that is tangent to each member of the family at some point and these points of tangency together form the whole envelope.

Let T_s be the straight line connecting the points $\alpha(s)$ and $\alpha(g(s))$, where $\alpha : S^1 \to B$ was the parameterization of B. Let $\mathcal{T} = \bigcup_s T_s$ be the familiy of curves formed by these lines. Let Γ be the envelope of the family $\mathcal{T}$. That is, each point in Γ can ve seen as the intersection of two "infinitesimally adjacent" curves:

$$\Gamma = \left\{\beta(s) \in R^2 : s \in S^1\right\}, \qquad \beta(s) = \lim_{s_1 \to s} T_s \cap T_{s_1}.$$

To end the proof, it suffices to check that this curve is well-defined and it is contained in B. That is, we have to check that all limit points $\beta(s)$ exist and are contained in B for any $s \in S^1$. By the Birkhoff's theorem, any RIC γ can be expressed by a graph, hence there exists a function $\delta : S^1 \to [0, \pi]$ such that for any point $(s, \theta) \in \gamma$, $\theta = \delta(s)$. On the other hand, we call g the function that describes the dynamics of the billiard map restricted to the curve. In other words,

$$f(s, \delta(s)) = (g(s), g(\delta(s)), \quad \forall s \in S^1.$$

This function g is monotonously increasing since γ is near to the boundary $S^1 \times \{0\}$ or $S^1 \times \{\pi\}$ and the billiard is area preserving. Hence two close trajectories always intersect (see Figure 6).

□

Caustics are curves very studied because their existence has great implications in the billiard map. In the following sections we are going to see two conditions provided by Lazutkin to prove the existence of caustics near of the boundary of the billiard table and the result of Gutkin-Katok about how regions free of convex caustics change when the curvature of a point of the billiard table tends to zero. On the other hand, we are going to see the result of Mather about the non existence of caustics when a point in the billiard table has curvature zero. Finally, we will see how Hubacher proves the non existence of caustics near the boundary when the table is not sufficiently smooth.

Examples of caustics in convex billiards.

The most known examples of billiard tables are circumferences and ellipses.

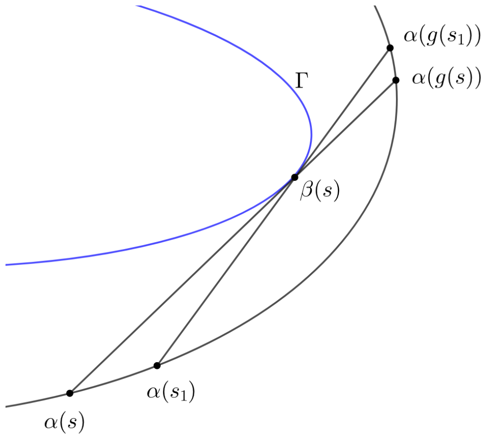


Figure 6: Close trajectories near to the boundary

If we focus in the circumference C, it is easy to see that any concentric circumference $\bar{C}$ contained in C is a caustic, since for all initial condition (s, θ), the iterate $f^n(s, \theta) = (s_n, \theta_n)$ satisfies $\theta_n = \theta \ \forall n \in \mathbb{N}$ (see Figure 7).

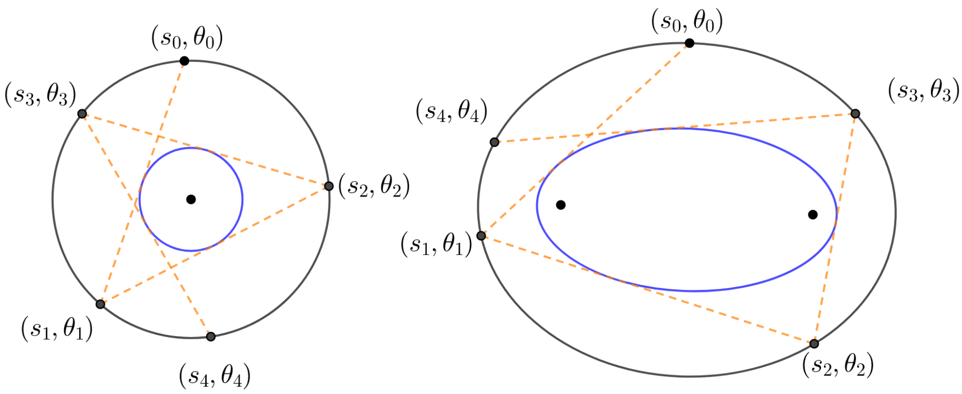


Figure 7: Caustics of a circumference and an ellipse

On the other hand, we can consider an ellipse $\mathcal{E}$ as billiard table. In this case, all cofocal ellipses $\mathcal{E}'$ inside $\mathcal{E}$ are convex caustics as we can see in Figure 7. To prove it we are going to follow [2].

Let A, B, C be three points in the ellipse $\mathcal{E}$ such that satisfy the billiard law. Now we choose two points T_1, T_2 in the tangent line to the ellipse at B (T_1 to the left of B and T_2 to the right of B).

The focus of an ellipse satisfy the following property: given any point P of B, the lines generated by the segments F_1P and PF_2 are part of a billiard map trajectory. This property is esay to prove with a geometric argument and using the followings fact:

given a line l and two points in the same half-plane delimitated by l, the point $\bar{P} \in l$ when the law reflection is satisfied coincides with the point when the sum of distances $d(F_1, \bar{P}) + d(F_2, \bar{P})$ is minimal.

By this property, the angles $\angle F_1BT_1$ and $\angle F_2BT_2$ are equal and this implies that the segment AB is placed inside $\angle F_1BT_1$. Then, the segment BC is placed inside $\angle F_2BT_2$ since $\angle ABT_1$ and $\angle CBT_2$ are equal by the billiard law (orange angles in Figure 8).

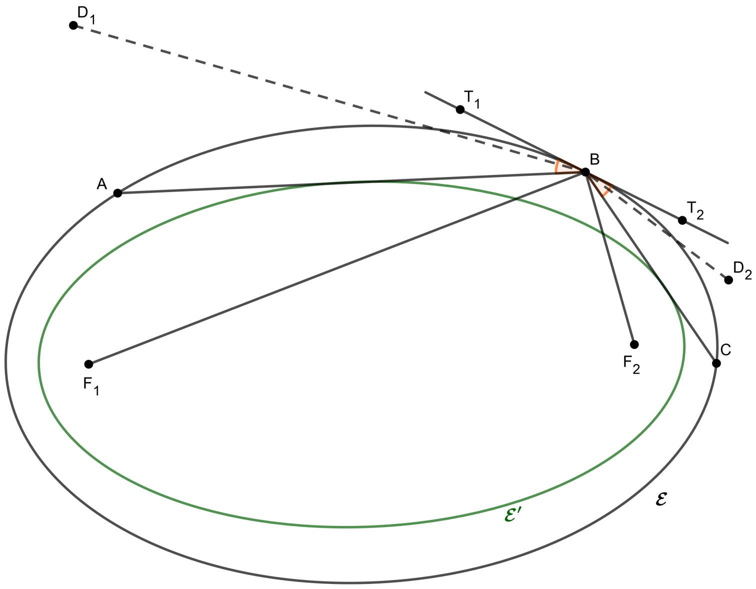


Figure 8

Let D_1, D_2 be points symmetric to F_1, F_2 respect the lines AB and BC respectively. It is easy to prove that the triangles ΔD_1BF_1 and ΔF_1BD_2 are congruent. The distances D_1B, F_1B are the same and F_2B and D_2B have the same length too. Moreover,

$$\angle D_1BF_2 = \angle D_1BF_1 + \angle F_1BF_2,$$
$$\angle F_1BD_2 = \angle F_2BD_2 + \angle F_1BF_2,$$

and, using the symmetry

$$\angle D_1BF_1 = 2\angle F_1BA = 2(\angle F_1BT_1 - \angle ABT_1) = 2(\angle F_2BT_2 - \angle CBT_2) = 2\angle F_2BC = \angle D_2BF_2.$$

Hence, ΔD_1BF_1 and ΔF_1BD_2 are congruent and $D_1F_2 \cong F_1D_2$.
For any point in the line AB, we can consider the sum of distances between this point with the foci F_1, F_2. The minimal value of these sums of distances is equal to the length of the segment D_1F_2. Since ΔD_1BF_1 and ΔF_1BD_2 are congruent, F_1D_2 is equal to the minimal sum of distances from F_1 and F_2 of a point on the line BC. Then, BC is tangent to $\mathcal{E}'$, so $\mathcal{E}'$ is a caustic of $\mathcal{E}$.

3.2 Existence of caustics

The most important known result about the existence of convex caustics was given by Lakutkin. Later, Douady presented the same result with weaker conditions. In this section we are going to see the Lazutkin theorem and an idea about the approach of the proof. Finally, we are going to see the theorem enunciated by Douady.

Theorem 3.5. *(Lazutkin) Let D be the billiard table and let $B = \partial D$ be a closed convex curve such that for all point of B the curvature is positive everywhere and B is $\mathcal{C}^k$ where $k \geqslant 558$. Then, there exist a Cantor set $K \subset [0, \frac{1}{2})\backslash\mathbb{Q}$ such that for all $\alpha \in K$, there exists a RIC with rotation number α.*

Now we are going to remember the notation to see the idea that Lazutkin use to prove the theorem.

We call $f : S^1 \times [0, \pi] \to S^1 \times [0, \pi]$ the billiard map in the natural coordinates and we define $F : \mathbb{R} \times [0, \pi] \to \mathbb{R} \times [0, \pi]$ as its lift. Let $(s, \theta) = (s, 0)$ be the point of the phase space $\mathbb{R} \times [0, \pi]$ such that the angle is zero. By the law reflection, the iterates of this points are same point, i.e. $F(s, 0) = (s, 0)$. Hence this point are stationary and the curve $\mathbb{R} \times \{0\}$ is an stationary curve. Analogously the curve $\mathbb{R} \times \{\pi\}$ is an stationary curve.

Lazutkin in [6] proves the existence of a family of RICs near to the stationary curves $\mathbb{R} \times \{0\}$ and $\mathbb{R} \times \{\pi\}$ with a small rotation number. So, by Proposition 3.4 we have the existence of caustics near to the boundary.

We focus in the stationary curve $\mathbb{R} \times \{0\}$ and the map F near to this curve, i.e., we are going to take small angles θ.

Doing an asymptotic expansion, we can describe $(s_1, \theta_1) = F(s, \theta)$ as

$$\begin{aligned} s_1 &= s + 2\rho(s)\theta + \tfrac{4}{3}\rho(s)\rho'(s)\theta^2 + \tfrac{2}{3}\rho^2(s)\rho''(s)\theta^3 + G_1(s,\theta)\theta^4 \\ \theta_1 &= \theta - \tfrac{2}{3}\rho'(s)\theta^2 + (-\tfrac{2}{3}\rho(s)\rho''(s) + \tfrac{4}{9}\rho'^2(s))\theta^3 + G_2(s,\theta)\theta^4 \end{aligned} \tag{3}$$

where $\rho(s)$ is the radius of curvature of B in s and the functions $G_1(s, \theta)$ and $G_2(s, \theta)$ has three derivatives less than the function $\rho(s)$.

Now, we are going to consider the following change of coordinates:

$$x = C_1 \int_0^{?} \rho^{-2/3}(s)ds$$

$$y = C_2 \rho^{1/3}(s) \sin \frac{\theta}{2}$$

where $C_1 = (\int_B \rho^{-2/3}(s)ds)^{-1}$ and $C_2 = 4C_1$. The new coordinates are called Lazutkin coordinates. This expression appears from some formulas of diffraction theory and from the fact that the invariant measure have the simple form $dm = const|y|dxdy$. The asymptotic expansion (3) of the billiard map F in this new coordinates is the system:

$$\begin{aligned} x_1 &= x + y + y^3 g_1(x, y) \\ y_1 &= y + y^4 g_2(x, y), \end{aligned} \tag{4}$$

where the functions $g_1(x, y)$ and $g_2(x, y)$ has three derivatives less than the curvature function. The factors C_1 and C_2 have been chosen to define the period of x in a unit and the coefficient of y in the first equation of (4) be unitary.

If the functions g_1 and g_2 are zero, the system (4) is

$$\begin{aligned} x_1 &= x + y \\ y_1 &= y. \end{aligned} \tag{5}$$

Then, we have the family of invariant curves $y = const$. On the other hand, is y is small, we can understand the terms $y^3 g_1(x, y)$ and $y^4 g_2(x, y)$ as a small perturbation of (5). At this point, Lazutkin proves the theorem with KAM theory.

Theorem 3.6. *(Douady) If $k > 6$, then there exists a Cantor set K of irrational points contained in $[0, \frac{1}{2})$ such that for all $\alpha \in K$, the there exists a caustic $\gamma \in \mathcal{C}^{1+\varepsilon}, \varepsilon > 0$ with rotation number α.*

Although we have seen the existence of a discontinuous family of convex caustics, in the case of ellipse $\mathcal{E}$ this family is continuous since all cofocal ellipses inside $\mathcal{E}$ are caustics.

3.3 Regions free of convex caustics

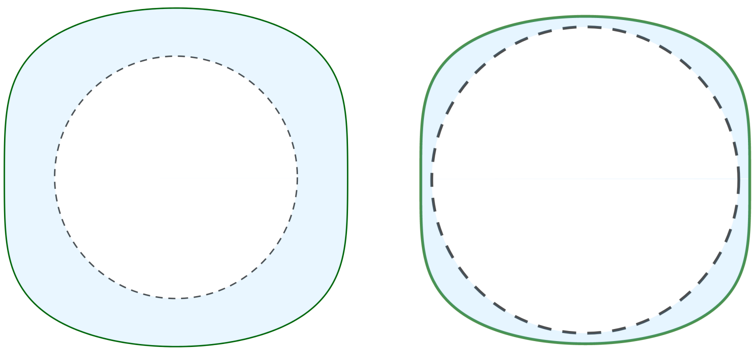
Gutkin and Katok in [3] do an study of the regions free of convex caustics that depends of the billiard table shape. Moreover, they give estimations about the size of these regions.

We have just seen that under some conditions there exists a family of convex caustics near to the boundary. Now, doing some deformations in the billiard table we will see how change the regions where the caustics are contained or, which is the same, how change the region free of convex caustics. There are different factors that appear when we estimate the size of these regions, for example the minimum and maximum value of the curvature.

An simple example to see how can change the size of these regions is the following: suppose that we have a one-parameter family of curves such that the minimum curvature change and tends to zero but the global shape of the table remains essentially the same, for example

$$B_t = \{(x, y); \in \mathbb{R}^2; x^2 + ty^2 + y^4 \leqslant 1\} \text{ where } t > 0.$$

In this case, when we tend the parameter t to zero, the region free of convex caustics increase and, consequently, the caustics are contained in a region near to the boundary which area tends to zero. See Figure 9.



(a) B_t where t=0.05. (b) B_t where t=0.01.

Figure 9: Blue region is the region where the caustics are contained.

Definition 3.7. A region $X \subset D$ is called region free of convex caustics if for any convex caustic γ, $\gamma \cap \mathring{X} = \varnothing$.

The domain of the convex billiard map is a closed convex curve $B = \partial D$ hence the maximum and minimum values if the curvature are finite. Let $\kappa(s)$ be the curvature function of the curve B, we define:

$$\kappa^- = \min_{s \in B} \kappa(s) \quad \text{and} \quad \kappa^+ = \max_{s \in B} \kappa(s).$$

Definition 3.8. Let D be a compact set in the plane. The diameter d is defined as the largest euclidean distance between two points contained in the set D. The width w is the minimum distance between two parallel supporting lines of D. And the inradius r of D is the radius of the biggest circumference inscribed in D.

Now, we are going to state different estimations about regions free of convex caustics. The main tool used to prove this estimations is the mirror equation of the geometric optics.

In the following results we assume that D is the billiard table, B is its boundary and L is the length of B. And we are going to use the notation described above about the curvature.

Proposition 3.9. *If B satisfies*

$$r \geqslant \sqrt{2}\kappa^- d^2$$

then, the billiard table contain a disc free of convex caustics and its radius r' satisfies $r' > r - \sqrt{2}\kappa^- d^2$.

Proposition 3.10. *If B satisfies*

$$\frac{w}{3} \geqslant \sqrt{2}\kappa^- d^2$$

then, the billiard table contain a disc free of convex caustics and its radius r' satisfies $r' > \frac{w}{3} - \sqrt{2}\kappa^- d^2$.

Proposition 3.11. *If B satisfies*

$$1 \geqslant \sqrt{2}d^2\kappa^-\kappa^+$$

then, the billiard table contain a disc free of convex caustics and its radius r' satisfies $r' > \frac{1}{\kappa^+} - \sqrt{2}\kappa^- d^2$.

Theorem 3.12. *If B is $\mathcal{C}^2$ and satisfies*

$$1 \geqslant \sqrt{2}d^2\kappa^-\kappa^+$$

then, the billiard table contain a convex region X free of convex caustics and its area satisfies

$$Area(X) \geqslant Area(D) - \sqrt{2}\kappa^- d^2 L.$$

3.4 Non existence of caustics - Mather

Remember that the conditions of existence of caustics by Lazutkin-Douady are it has the strictly positive curvature and the boundary has 6 continuous derivatives for any point on the boundary of the table. Mather, in [4], changes the first condition and he imposes the existence of a point with zero curvature. Under this conditions, he proves the non existence of caustics. To do it, we need define the ϵ-glancing orbits.

Definition 3.13. An orbit is called ϵ-glancing if, at least, it has a reflection angle less than ϵ.

We differentiate two types of ϵ-glancing orbits: if $\epsilon < \frac{\pi}{2}$, we say that the trajectory is positively ϵ-glancing if the reflection angle is between the trajectory and the positive tangent of B and we say that it is negatively ϵ-glancing if we choose the angle with the negative tangent.

A intuitive way to see these orbits is in the phase space $\mathbb{R} \times [-1, 1]$. Suppose we have some orbit that is positive ϵ-glancing for all $\epsilon > 0$. This orbit has always some iterate such that it states as close as we want to the boundary $B \times \{-1\}$. Now, suppose that we have a RIC. RICs separate the phase space into two connected components then there can be no orbit that is positive and negative ϵ-glancing for all $\epsilon > 0$.This implies if there exists an orbit of this type, there no exist RICs. On the other hand, Mather proves that if there are no RICs then there exists an orbit that is positive and negative ϵ-glancing for all $\epsilon > 0$. To do it he supposes that there is no orbit of this type and, under this hypothesis, he arrives to the conclusion that there exists a RIC. So, there exists RICs if and only if there no exist a positive and negative ϵ-glancing orbits for all $\epsilon > 0$. On the other hand, he proves the existence of this type of orbits when a point of the boundary B has curvature zero. As a consequence, there no exist caustics when a point of the boundary has curvature zero.

Theorem 3.14. *If the curvature of the boundary of B vanishes at some point, then for every $\epsilon > 0$, there are trajectories which are both positively and negatively ϵ-glancing.*

Proof. Suppose that the result of the theorem is false and we are going to arrive at a contradiction. Consider V_- and V_+ the following sets (see Figure 10)

$$V_- = B \times [-1, -1 + \epsilon] \quad \text{and} \quad V_+ = B \times [1 - \epsilon, 1].$$

Remember that the phase space of the billiard map in Birkhoff's coordinates is $S^1 \times [-1, 1]$ that is homeomorphic to $B \times [-1, 1]$.

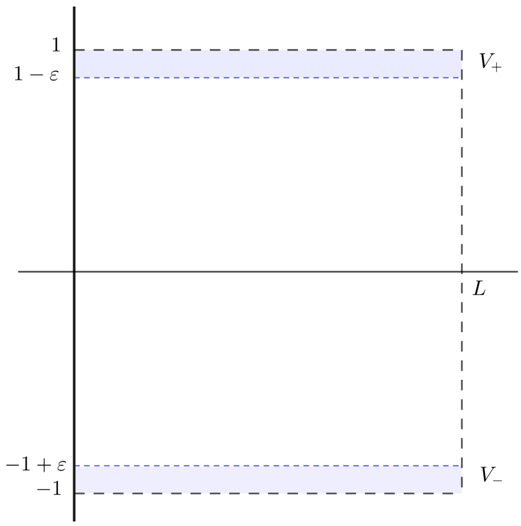


Figure 10: Sets V_- and V_+.

We can construct the new set

$$V = \bigcup_{n=-\infty}^{\infty} \bar{f}^n(V_- \cap \mathring{D}) \cup (B \times \{-1\}),$$

where $D = B \times [-1, 1]$. In other words, V contains all iterates of the orbits that have some point in V_- and the points of the lower boundary.

Using that the conclusion of the theorem is false, there exists an $\epsilon > 0$ sufficiently small such that the intersection of V_+ and V_- is empty ($V_+ \cap V_- = \varnothing$). We choose this ϵ. Consider $C = D \backslash V$ which contains $B \times \{1\}$, that is a connected component, and $U = A \backslash C$. Then, $V_- \subset U$, $U \cap V_+ = \varnothing$, $\bar{f}(U) = U$ and $B \times \{-1\}$ is a deformation retract of U.

By the Birkhoff's theorem, there exists a function γ such that it describes the boundary of U ($\partial U = \{(x, \gamma(x)); x \in B\}$).

Then, $\bar{f}(\partial U) = \partial U$ since $\bar{f}(U) = U$ and $B \times \{1\}$ is a deformation retract of U, so there exists an homeomorphism $g : B \to B$ such that $\bar{f}(x, \gamma(x)) = (g(x), \gamma g(x))$. This implies that $\bar{f}|_{\partial U}$ and g preserve orientation since $\bar{f}$ preserves orientation.

Consider a point $x_0 \in B$ such that its curvature is zero and other point $y \in B$. We are going to denote

$$x_n = g^n(x_0) \text{ and } y_n = g^n(y_0)$$

their respective orbits.

On the other hand, we denote by $h(v, w)$ the euclidean distance between v and w where $v, w \in B$. Remember that this function is a generating function so, by (2),

$$\frac{\partial h}{\partial v}(v, w) = -r$$

where $r = \cos(\theta)$ and θ is the angle associated to the point v in the cylinder under its trajectory (the line that it join the point v and w).

It is clear that $\frac{\partial \theta}{\partial w} < 0$, thus

$$\frac{\partial h^2}{\partial v \partial w} = \frac{\partial}{\partial w}\frac{\partial h}{\partial v} = \frac{\partial}{\partial w}\cos(\theta) \Rightarrow \frac{\partial h^2}{\partial v \partial w} > 0. \tag{6}$$

Moreover, if x_{-1} and x_1 are the previous and posterior points of the trajectory of x_0, using that h is a generating function, we have

$$\frac{\partial}{\partial x_0}(h(x_{-1}, x_0) + h(x_0, x_1)) = 0.$$

On the other hand, at x_0 the curvature of B vanishes, then we can apply the reflection law that is: if we fix two points x and y in the plane and one line under this point, the point z in the line that satisfies the reflection law (the angle of incidence is equal to the angle of reflection) is the one that minimizes the distance $h(x, z) + h(z, y)$. So, in our case we can express this property as

$$\frac{\partial^2}{\partial x_0^2}(h(x_{-1}, x_0) + h(x_0, x_1)) > 0. \tag{7}$$

Choosing $\tilde{y}$ and y' two points on B near to x_{-1} and x_1 respectively and applying the implicit function theorem, there exists a function $\eta = \eta(\tilde{y}, y')$ in a neighborhood of x_0 such that

$$\frac{\partial}{\partial \eta}(h(\tilde{y}, \eta) + h(\eta, y')) = 0$$

and given y_{-1} sufficiently near to x_{-1}, we get

$$y_0 = \eta(y_{-1}, y_1). \tag{8}$$

Using (6) and (7),

$$\left.\frac{\partial \eta}{\partial \tilde{y}}\right|_{\tilde{y}=x_{-1}} < 0 \quad \text{and} \quad \left.\frac{\partial \eta}{\partial y'}\right|_{y'=x_1} < 0. \tag{9}$$

Since y_1 is an increasing function of y_{-1} and using (8) and (9) we have that y_0 is a decreasing function of y_{-1} choosing y_{-1} a point in a small neighborhood of x_{-1}. And this is a contradiction, so the theorem is proved. □

3.5 Non existence of caustics - Hubacher

In [5], Hubacher proves the non existence of convex caustics when the number of continuous derivatives is less than 2. In particular, she supposes the second derivative is not continuous.

Theorem 3.15. *Suppose that the boundary $B = \partial D$ of the billiard table has at least one discontinuity in the curvature (there is a point such that the second derivatives exists from both sides but there is a discontinuity). Then the circles $S^1 \times \{0\}$ and $S^1 \times \{\pi\}$ are isolated invariant curves of the billiard map; hence caustics cannot accumulate at the boundary curve $B = \partial D$.*

First we are going to see an idea about the proof and next we will see the proof of the theorem.

Suppose the boundary B of our table is a convex closed curve that has a point where the curvature function is discontinuous. We can assume that this point is $s = 0$ since otherwise we can do a translation to fix the discontinuity in this point. Assume that this point is the concatenation of two circles with radius r and R respectively, where $r \neq R$. Under this hypotheses, we are going to see the behaviour of two different types of orbits.

The first orbit (orange orbit in Figure 11) we will suppose that it does not impact on the point $s = 0$. Then, this orbit has an impact in $s_0 < 0$ and a second impact $s_1 > 0$. We define the angles as θ_0 and θ_1 respectively, i.e., the points of the orbit are (s_0, θ_0) and (s_1, θ_1). In the phase space, the points $s \in \mathbb{R}$ under the lift of the billiard map evolve as follows: two consecutive points s_i, s_{i+1} of the orbit that impact in the circle with radius r satisfy $s_{i+1} = s_i + 2r\theta_r$. Analogously, $s_{i+1} = s_i + 2R\theta_R$ if the two consecutive points state over the circle with radius R. Remember that the impact angle states the same when the boundary of the table is a circle. On the other hand, the relation between θ_r and θ_R is $\theta_R \leqslant \sqrt{\frac{r}{R}}\theta_r$ (we will see this relation in the proof of the theorem).

The second orbit (green orbit in Figure 11) we will assume it impacts over $s = 0$. In this case, if we fix an impact point s_{-i} with angle θ_{-i} before $s_0 = 0$ and other point s_i with angle θ_i after $s_0 = 0$, the points of the trajectory from s_{-i} to s_i remain the angle, i.e., $\theta_{-i} = \ldots = \theta_0 = \ldots = \theta_i$.

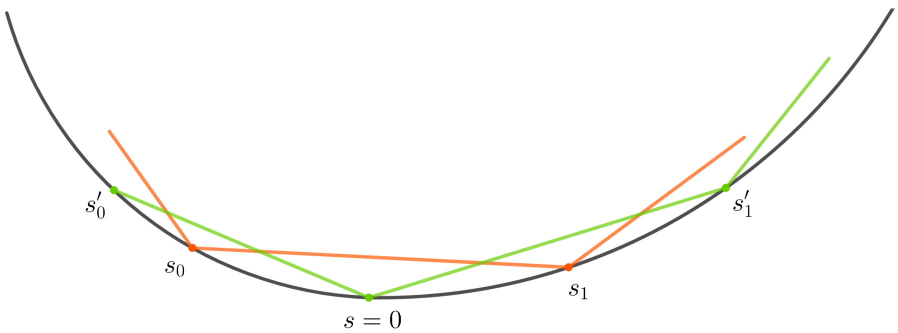


Figure 11

Now, we are focus in the impact point after $s = 0$. The angles of the first orbit (orange) evolve slower than the angles of the second orbit (green). Therefore, the impact points also evolve slower. This implies the orbits will cross.

Suppose that there exists a caustic and we have two orbits as described above. We can assume that these orbits are tangent to the caustic, i.e., the points of their trajectories are in the RIC associated to the caustic. In this case, we have two orbits over the same RIC that they cross. This fact contradicts the Birkhoff theorem, so a caustic does not exist.

Lemma 3.16. *Let (s_n, θ_n) and (s_n', θ_n') the orbits described above. Suppose $s_0 < s_0' < s_1$. For all $\delta > 0, \delta \in \mathbb{R}$, there exist a neighborhood V of s_0' such that if (s_n, θ_n) has at least a point in V, then the orbits do not cross for any $n > 0$ if the following condition holds:*

$$|\theta_0' - \theta_1| < \delta\theta_1. \tag{10}$$

Analogously, they do not cross for any $n < 0$ if if the following condition holds:

$$|\theta'_0 - \theta_0| < \delta\theta_0. \tag{11}$$

Let (s'_n, θ'_n) be an orbit near to the boundary.

Proof. Fix $s = 0$ and consider the set $V_R = V \cap \{s > 0\}$. We can use the Lazutkin coordinates because the orbit is near to the boundary and the boundary is smooth (remember that the Lazutkin coordinates is an asymptotic expansion), then

$$\begin{aligned} s_1 &= s + 2\kappa(s)^{-1}\theta + o(\theta) \\ \theta_1 &= t + o(\theta) \end{aligned}$$

Now, fix δ such that $0 < \delta < 1$ and define

$$m = \inf_{V_R} 2\kappa(s)^{-1}, \quad M = \sup_{V_R} 2\kappa(s)^{-1}.$$

The curvature is bounded then m and M are finite values and we can choose V sufficiently small such that $m > (1-\delta)M$.

Let n be an integer such that $n = [\frac{2m}{(m-(1-\delta)M)}] + 1$ (the smallest integer greater than $\frac{2m}{(m-(1-\delta)M)}$). Therefore all iterates $F^j(s_0, \theta_0)$ and $F^j(s'_0, \theta'_0)$ with $1 \leqslant j \leqslant n$ are contained in V if θ_0 and θ'_0 are sufficiently close.

Suppose (10) is not fulfilled, i.e., $\theta'_0 < (1-\delta)\theta_1$ or $\theta'_0 > (1+\delta)\theta_1$. In the first case, using the Lazutkin coordinates we obtain

$$\begin{aligned} s_n > s_n - s_1 &= \theta_1 \sum_{j=1}^{n-1} 2\kappa(s_j)^{-1} + o(\theta_1) > \theta_1(n-1)m + o(\theta_1) \\ s'_n = s'_n - s'_0 &= \theta_0 \sum_{j=0}^{n-1} 2\kappa(s'_j)^{-1} + o(\theta'_0) < \theta_1(1-\delta)nM + o(\theta_1) \end{aligned}$$

then $s_n > s'_n$ and this implies that the orbits do not cross since the hypothesis was $s_0 < s'_0$. the second case is analogous.

To prove (11) we choose $V_R = V \cap \{s < 0\}$ and we use the same reasoning. □

Proof. We will separate the proof in two steps. The first is a local proof and consist to prove that for all point with a discontinuity in the curvature there exists a neighborhood that it does not have RICs. The second step is a generalize of this result for a neighborhood to the boundary.

Without loss of generality, assume $s = 0$ a point where the curvature function $\kappa(s)$ is discontinuous but the lateral limits of $\kappa(s)$ exist. Since the curvature is bounded, we can define the lateral limits as $r = \kappa(0^-)$ and $R = \kappa(0^+)$ where $r, R \in R$. Moreover, we can assume $\kappa(0^-) < \kappa(0^+)$, otherwise the proof is analogously.

We will use the billiard map f and its lift F to do the proof. Remember that the lift is defined as $F : \mathbb{R}/L\mathbb{Z} \times [0, \pi] \to \mathbb{R}/L\mathbb{Z} \times [0, \pi]$ and it can be extended as a periodic function from $\mathbb{R} \times [0, \pi]$ to $\mathbb{R} \times [0, \pi]$. Then, to do the study at point $s = 0$ is equivalent to do it at $(0, 0)$ in the phase space under the lift F. Consider V a neighborhood of $(0, 0)$ such that we can approximate the boundary of the billiard table B at $s = 0$ by to osculating circles with radius r and R respectively with a first order approximation, i.e., $o(s) \to 0$ when $s \to 0$.

Suppose that there exists a RIC passing through V and we will arrive to a contradiction. Let Γ be a RIC passing through V, since Γ is continuous, there exist a point $q = (s_0, \theta_0)$ such that the segment $q\bar{f}(q)$ is perpendicular to the straight line generates by the normal vector to B at $S = 0$ and is a segment of a trajectory.

We call t_0 the point that is the intersection of the osculating circle with radius r and the segment $q\bar{f}(q)$. We define α_0 the angle between the tangent line of the osculating circle at t_0 and the segment $q\bar{f}(q)$. Analogously, on the osculating circle with radius R we define the point t_1 and the angle α_1. See Figure 12.

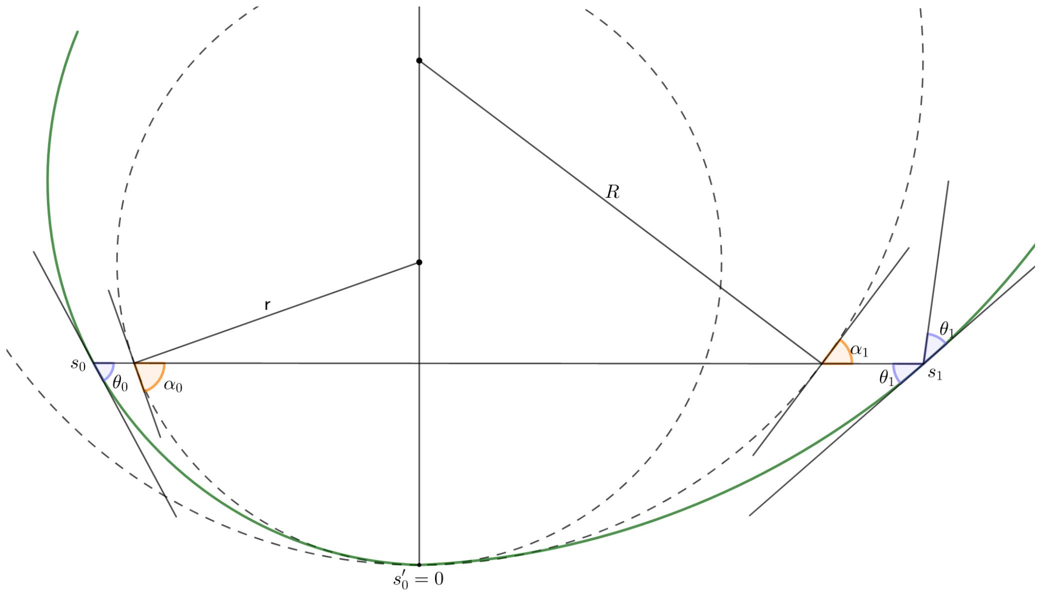


Figure 12

Hence we can conclude that the following relation is satisfied:

$$R\cos(\alpha_1) = R - r + r\cos(\alpha_0).$$

Given this relation we define this function:

$$g(\alpha) = \arccos(1 - \frac{r}{R} + \frac{r}{R}\cos(\alpha)).$$

and doing the derivative of $g(\alpha)$ the have the following bound:

$$g'(\alpha) = \frac{\sqrt{\frac{r}{R}}}{\sqrt{1 + (1 - \frac{r}{R})\tan^2(\frac{\alpha}{2})}} < \sqrt{\frac{r}{R}},$$

then,

$$\alpha_1 = g(\alpha_0) = \alpha_0 \int_0^1 g'(\tau\alpha_0)d\tau < \alpha_0\sqrt{\frac{r}{R}} \Longrightarrow \alpha_1 < \alpha_0\sqrt{\frac{r}{R}}. \tag{12}$$

Using the osculating circles approximation,

$$\alpha_0 = \theta_0 + o(\theta_0) \quad \text{i} \quad \theta_1 = \alpha_1 + o(\alpha_1)$$

when the angles tend to 0 and applying the inequality (12),

$$\theta_1 < \alpha_0\sqrt{\frac{r}{R}} + o(\theta_1) = \theta_0\sqrt{\frac{r}{R}} + o(\theta_0) \longrightarrow \theta_1 < \theta_0,$$

i.e., for a small neighborhood V we can bound $\theta_1 - \theta_0$:

$$\theta_0 - \theta_1 > \frac{1}{2}(1 - \sqrt{\frac{r}{R}})\theta_0 > \frac{1}{2}(1 - \sqrt{\frac{r}{R}})\theta_1. \tag{13}$$

We will refer by (s_n, θ_n) the orbit that we just described.

Now, we consider all orbits that impact in the point $s = 0$ and we denote this type of orbits by (s'_n, θ'_n). Our goal is arrive to a contradiction as follows: fix $s'_0 = 0$ and we will see that for any θ'_0 that we choose, the orbits (s_n, θ_n) and (s'_n, θ'_n) cross. If this happens we arrives a contradiction with the Birkhoff theorem.

By Lemma 3.16, we know that (s_n, θ_n) and (s'_n, θ'_n) do not cross for any positive $n \in \mathbb{Z}$ if for all $\forall\delta > 0$ there exists a neighborhood U such that

$$|\theta'_0 - \theta_1| < \delta\theta_1$$

and, analogously must be satisfied

$$|\theta'_0 - \theta_0| < \delta\theta_0$$

so that two orbits do not cross for some negative $n \in \mathbb{Z}$.

Suppose that the orbits do not cross for any $n > 0$, then, for all $\delta > 0 \exists U$ neighborhood such that $|\theta'_0 - \theta_1| < \delta\theta_1$. Choosing $\delta = \frac{1}{4}(1 - \sqrt{\frac{r}{R}})$ i and using (13) we have:

$$\begin{aligned}|\theta'_0 - \theta_0| &= |\theta'_0 - \theta_1 + \theta_1 - \theta_0| > |\theta_0 - \theta_1| - |\theta'_0 - \theta_1| = \theta_0 - \theta_1 - |\theta'_0 - \theta_1| \\ &> \theta_0 - \theta_1 - \frac{1}{4}(1 - \sqrt{\frac{r}{R}})\theta_1 > \frac{1}{2}(1 - \sqrt{\frac{r}{R}})\theta_1 - \frac{1}{4}(1 - \sqrt{\frac{r}{R}})\theta_1 = \frac{1}{4}(1 - \sqrt{\frac{r}{R}})\theta_1 = \delta\theta_1 \\ &> \delta\theta_0,\end{aligned}$$

so, they cross for some negative n. Analogosuly, is we suppose that they do not cross for any $n < 0$ we conclude that they cross for some positive n. In other words, the orbits (s_n, θ_n) and (s'_n, θ'_n) cross for any θ'_0 that we choose. In particular, we can choose (s'_n, θ'_n) the orbit that impact on $s = 0$ and it is tangent to the caustic. This implies that both orbits state over the RIC but this is a contradiction by the Birkhoff theorem. So, in the neighborhood V does not pass any RIC.

To generalize the result we suppose that there exists an accumulation point of a family of RICs $\{\Gamma_n\}_{m\in\mathbb{Z}}$. The boundary of the table B is piecewise $\mathcal{C}^2$and the curvature is uniform hence the billiard map is twist uniform. On the other hand, by Birkhoff theorem, Γ_n is Lipschitz, then there exists a subsequence of $\{\Gamma_n\}_{m\in\mathbb{Z}}$ that converges to RIC Γ and it intersects with $\mathbb{R}/L\mathbb{Z} \times \{0\}$ at least in one point $(s, 0)$. Since F is periodic, there exists at least other point $(s', 0)$ where Γ intersects with $\mathbb{R}/L\mathbb{Z} \times \{0\}$.

Take two consecutive points where Γ intersects with $\mathbb{R}/L\mathbb{Z} \times \{0\}$ and the region W delimited by Γ and $\mathbb{R}/L\mathbb{Z} \times \{0\}$ contains the point $(0, 0)$ in its boundary. The set W is an invariant set homeomorphic to the plane. Give a vertical line inside W such that this lines divides W in two connected components with positive measure. Under the billiard map this measure remains and, by twist condition, the vertical line tilts to right. These two facts are contradictory, so the RIC Γ cannot exist. $\square$

4. Periodic Orbits

In this section we folllow the references [1] and [7]. We are going to define of the (p, q) periodic orbits and we will prove the existence of at least two (p, q) periodic orbits with different geometry in convex billiards. These periodic orbits are minimizing orbits and minimax orbits. The first are associated with the minimums of the functional action. The existence of these orbits implies the existence of a minimax orbit since the functional action is continuous and between two minimums of the functional action there must be another critical point.

In this section we assume that $F : \mathbb{R} \times (0, \pi) \to \mathbb{R} \times (0, \pi)$ is a lift of the billiard twist map $f : S^1 \times (0, \pi) \to S^1 \times (0, \pi)$ where $S^1 = \mathbb{R}/\mathbb{Z}$. That is, we assume that the length of the billiard map is equal to one, otherwise we can normalize it.

4.1 Definitions and Poincaré-Birkhoff theorem

Definition 4.1. A sequence $\{(s_n, \theta_n)\}_{n\in\mathbb{N}}$ is a (p, q) periodic orbit of a map F if

$$F^q(s_n, \theta_n) = (s_n + p, \theta_n),$$

where $p \in \mathbb{Z}, q \in \mathbb{Z}$.

In terms of the billiard map, the function F is the lift of the billiard map f and s_n is defined in $\mathbb{R}$ for all $n \in \mathbb{Z}$. So, given an orbit $\mathcal{O} = \{(s_n, \theta_n)\}_{n\in\mathbb{N}}$, the components s_n are in S^1 if we consider the billiard map f but, when we use the lift F, s_n is defined in $\mathbb{R}$. In the last case, we say that the space where we are working is the configuration space.

Observation 4.2. If we have a (p, q) periodic orbit, the physical representation of this periodic orbit on the billiard table is a cycle. In other words, suppose the segment $((s_1, \theta_1), ..., (s_q, \theta_q))$ is a segment of a (p, q) periodic orbit, then $s_{q+1} = s_1 + p, s_{q+2} = s_2 + p, ...s_{2q} = s_q + p$ hence

$$(s_{q+1} mod(p), ..., s_{2q} mod(p)) = (s_1, ..., s_q).$$

So the points of the trajectory under the billiard map f generate a cycle with q components.

Definition 4.3. Let $\{(s_n, \theta_n)\}_{n\in\mathbb{N}}$ be an orbit and $F(s, \theta) = (F_1(s, \theta), (F_2(s, \theta))$ a map. The rotation number is the result of the following limit if it exists:

$$\lim_{n\to\infty} \frac{F_1^n(s, \theta) - s}{n},$$

where (s, θ) is an element of the orbit. If the limit exists, it does not depend to the (s, θ) chosen.

Proposition 4.4. *The rotation number of a (p, q) periodic orbit is $\omega = \frac{p}{q}$.*

Proof. Let (s, θ) be an element of a (p, q) periodic orbit, then

$$\lim_{n\to\infty} \frac{F_1^n(s, \theta) - s}{n} = \lim_{n\to\infty} \frac{F_1^{nq}(s, \theta) - s}{nq} = \lim_{n\to\infty} \frac{(s + np) - s}{nq} = \frac{p}{q}$$

□

In section 5 we shall study the symmetric periodic orbits when the billiard table is the curve $x^2 + y^4 = 1$ and the orbits has odd period and one point of the orbit is $(0, 1)$, so the axis of symmetry is y-axis.

Examples of periodic orbits

In the following images we see different types of orbits with same period. In Figure 13 the orbits have period $q = 5$ but $\mathcal{O}_1$ is a $(1, 5)$ periodic orbit and $\mathcal{O}_2$ is a $(2, 5)$ periodic orbit. In Figure 14 the orbits have period $q = 7$ but $\mathcal{O}_3$ is a $(1, 7)$ periodic orbit and $\mathcal{O}_4$ is a $(3, 7)$ periodic orbit.

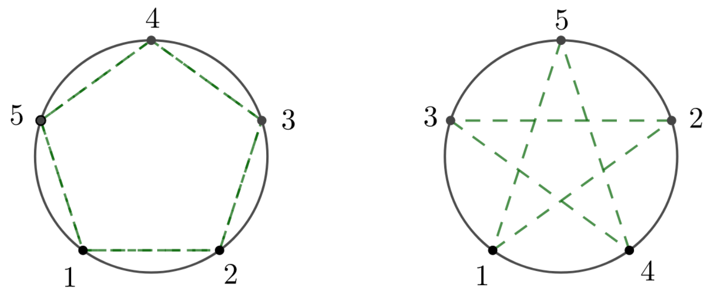


Figure 13: Periodic orbits with period $q = 5$. Left orbit is $\mathcal{O}_1$ and right orbit is $\mathcal{O}_2$.

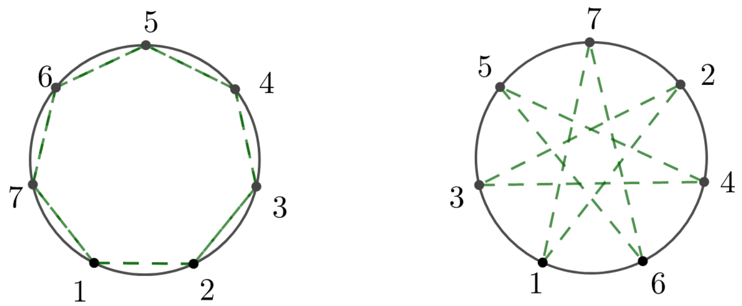


Figure 14: Periodic orbits with period $q = 7$. Left orbit is $\mathcal{O}_3$ and right orbit is $\mathcal{O}_4$.

Figure 15 is a representation of the contact points of one cycle of the orbit in $\mathbb{R}$. This is a good representation to see the meaning of p.

Remember that the generating function of the billiard map is $H(s, s')$ where s and s' are the first component of two consecutive points (s, θ) and (s', θ') of an orbit, that is the length of these two consecutive points in the billiard table. On the other hand, $L(s_1, ..., s_n)$ is the action functional and it is defined in the configuration space. The following Proposition is analogous to Proposition 2.9.

Proposition 4.5. *An orbit $\mathcal{O} = \{(s_n, \theta_n)\}_{n \in \mathbb{N}}$ is a (p, q) periodic orbit if and only if for any q dimensional segment of the orbit $(s_i, ..., s_{i+q-1})$ is a critical point of the functional*

$$L_p(s_i, ..., s_{i+q-1}) := L(s_i, ..., s_{i+q-1}, s_i + p) = \sum_{k=i}^{i+q-1} H(s_k, s_{k+1})|_{s_{i+q} = s_i + p}.$$

Proof. Suppose that $\mathcal{O}$ is a (p, q) periodic orbit and we consider the functional action of a q dimensional

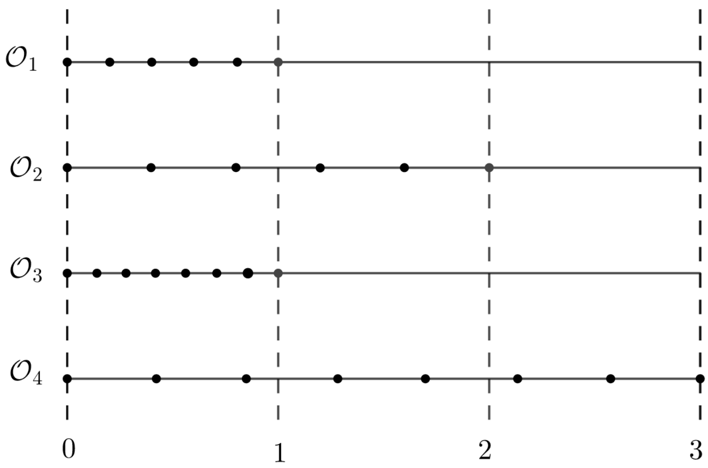


Figure 15

segment of this orbit

$$L_p(s_i, \ldots, s_{i+q-1}) = H(s_i, s_{i+1}) + \ldots + H(s_{i+q-1}, s_i + p)$$

then, the segment $(s_i, \ldots, s_{i+q-1})$ is a critical point if and only if

$$\frac{\partial L_p}{\partial s_j}(s_i, \ldots, s_{i+q-1}) = 0 \quad \forall j \in \{i, \ldots, i+q-1\}.$$

The last equation is equivalent to the following:

$$\frac{\partial H}{\partial s'}(s_{j-1}, s_j) + \frac{\partial H}{\partial s}(s_j, s_{j+1}) = 0 \quad \forall j \in \{i+1, \ldots, i+q-1\}$$
$$\frac{\partial H}{\partial s}(s_i, s_{i+1}) + \frac{\partial H}{\partial s'}(s_{i+q-1}, s_i + p) = 0$$

So, the segment $(s_i, \ldots, s_{i+q-1})$ is a critical point if and only if

$$\begin{aligned} \forall j \in \{i+1, \ldots, i+q-1\} \quad \theta_j &= \frac{\partial H}{\partial s'}(s_{j-1}, s_j) = -\frac{\partial H}{\partial s}(s_j, s_{j+1}) \\ \theta_i &= \frac{\partial H}{\partial s'}(s_{i+q-1}, s_i + p) = -\frac{\partial H}{\partial s}(s_i, s_{i+1}) \\ &= \theta_{i+q}, \end{aligned}$$

thus, the equations (2) are satisfied. □

Given an area preserving end-preserving twist function f, the Poincaré-Birkhoff theorem implies the existence of at least two periodic orbits under the map f. Birkhoff gives us a geometrical proof and his results can be applied to other types of functions. In our case, we can do a simply proof using the variational principle. So we need to introduce the definition of minimizing orbit and the minimax principle.

Consider an orbit $\mathcal{O} = \{(s_n, \theta_n)\}_{n \in \mathbb{N}}$ and a finite segment of this orbit $\{(s_i, \theta_i), \ldots, (s_j, \theta_j)\}$. By Proposition 2.9 this segment is a critical point of the functional action $L(s_i, \ldots, s_j)$ but the second variation can not be zero. Consider the following quadratic form that is the second order expansion of L fixing the end points s_i and s_j:

$$\delta^2 L(s_i, \delta s_{i+1}, ..., \delta s_{j-1}, s_j) = \sum_{k=i+1}^{j-1} \sum_{l=i+1}^{j-1} \delta s_l \frac{\partial^2 L}{\partial s_l \partial s_k} \delta s_k. \tag{14}$$

Definition 4.6. A finite segment $\{(s_i, \theta_i), ..., (s_j, \theta_j)\}$ of an orbit $\mathcal{O} = \{(s_n, \theta_n)\}_{n\in\mathbb{N}}$ is locally minimizing if $\delta^2 L(s_i, \delta s_{i+1}, ..., \delta s_{j-1}, s_j)$ is non negative for all vectors $(\delta s_{i+1}, ..., \delta s_{j-1}) \in \mathbb{R}^{j-i-1}$. And the orbit $\mathcal{O}$ is locally minimizing if every finite segment is locally minimizing.

Definition 4.7. Let $\mathcal{O} = \{(s_n, \theta_n)\}_{n\in\mathbb{N}}$ be an orbit and $(\eta_i, ..., \eta_j) = (s_i, s_{i+1} + \delta s_{i+1}, ..., s_{j-1} + \delta s_{j-1}, s_j)$ be an arbitrary variation of the finite segment $(s_i, ..., s_j,)$ with fixed end points. The segment $(s_i, ..., s_j)$ is minimizing if

$$L(\eta_i, ..., \eta_j) - L(s_i, ..., s_j) \geqslant 0$$

for every variation $(\eta_i, ..., \eta_j)$. And the orbit $\mathcal{O}$ is minimizing if every finite segment of the orbit is minimizing.

The first periodic orbit of the Poincaré-Birkhoff theorem appears as a minimizing periodic orbit. To do the proof we need the growth condition.

Proposition 4.8. *(Growth condition) For any area preserving and end preserving twist map f the generating function H is bounded by*

$$H(s, s') \geqslant A - B|s - s'| + C|s - s'|^2,$$

where $A, B, C \in \mathbb{R}$ and B, C are positive.

Proof. Consider $\xi_\lambda = s + \lambda(s' - s)$ with $\lambda \in [0, 1]$ the line that connects s and s'. Applying Barrow's rule we have the following,

$$H(s, s') - H(s, s) = \int_0^1 \frac{\partial H}{\partial s'}(s, \xi_\lambda) d\lambda (s' - s).$$

Now, we apply again the Barrow's rule and we have

$$H(s, s') - H(s.s) = \int_0^1 \int_0^1 \frac{\partial^2 H}{\partial s' \partial s}(\xi_\mu, \xi_\lambda) d\mu d\lambda (s' - s)^2.$$

Then,

$$\begin{aligned} H(s, s') &= H(s, s) + (H(s, s') - H(s, s)) - (H(s, s') - H(s, s)) \\ &= H(s, s) + \int_0^1 \int_0^1 \frac{\partial^2 H}{\partial s' \partial s}(\xi_\mu, \xi_\lambda) d\mu d\lambda (s' - s)^2 - \int_0^1 \frac{\partial H}{\partial s'}(s, \xi_\lambda) d\lambda (s' - s) \\ &\geqslant A - B \mid s' - s \mid + C \mid s' - s \mid^2 \end{aligned}$$

where $A = \min_{s\in S^1} H(s, s)$, $B = \max_{s,s'\in S^1} |\frac{\partial H}{\partial s'}(s, s')| > 0$ and $C = \frac{1}{2}K > 0$. The constants A and B exist because H is C^∞ and H is defined over the compact $S^1 \times S^1$. The constant K is the constant of the twist condition. □

Theorem 4.9. *(Poincaré-Birkhoff) For an area preserving, end preserving twist mapping F there is a periodic orbit of period (p, q).*

Proof. Let $\mathcal{O} = \{(s_n, \theta_n)\}_{n \in \mathbb{N}}$ be a (p, q) periodic orbit. By Proposition 4.5, any finite segment $(s_i, ..., s_j) \in \mathbb{R}^q$ of this orbit is a critical point of the functional action $L_p(s_i, ..., s_j)$. So, we are going to consider the segment $(s_1, ..., s_q)$.

The first step is to prove that there exists a (p, q) periodic orbit that minimizes the functional L_p. The functional $L_p(s_1, ..., s_q) = H(s_1, s_2) + ... + H(s_q, s_1 + p)$ is defined in $[0, 1] \times \mathbb{R}^{q-2}$ since $s_1 \in [0, 1]$. By Proposition 4.8,

$$L_p(s_1, ..., s_q) = L_p(s_0, ..., s_{q-1}) = H(s_0, s_1) + ... + H(s_{q-1}, s_0 + p) \tag{15}$$

$$\geqslant \sum_{j=0}^{q-1} (A - B|s_{j+1} - s_j| + C|s_{j+1} - s_j|^2). \tag{16}$$

Now, consider $\Omega \subset \mathbb{R}^q$ such that

$$L(s_0, ..., s_{q-1}) \leqslant qA + D \tag{17}$$

where $(s_0, ..., s_{q-1}) \in \Omega$ and D is a constant sufficiently small to satisfy $\Omega \neq \varnothing$. If Ω is a compact set, the function L has a minimum and this implies that there exists a (p, q) periodic orbit.

Using (15) and (17),

$$D \geqslant \sum_{j=0}^{q-1} (-B|s_{j+1} - s_j| + C|s_{j+1} - s_j|^2),$$

hence $|s_{j+1} - s_j|$ is bounded for all j and $|s_t - s_0|$ is bounded for all $0 < t < q$ since we can consider $s_1 \in [0, 1]$ because F if periodic. Then, Ω is a compact set. $\square$

The minimum of L_p is not unique. Suppose that we have a (p, q) periodic orbit and $(s_0, ..., s_{q-1})$ is a minimum, then the translation $(s_j + k, s_{j+1} + k, ..., s_{q-1} + k, s_0 + k, ..., s_{j-1} + k)$ is a minimum for any j where $k \in \mathbb{Z}$ is chosen so that $x_j + k$ is in the unit interval $[0, 1]$.

Definition 4.10. Let $\mathcal{O}, \mathcal{O}'$ be two minimizing orbits such that they do not cross. Consider a path such that connects both minimums of L_p. Since L_p is continuous, it must have a maximum along this path. If we vary this path, we can find the orbit where this maximum is the smallest. This point is a saddle point so it is a critical point of L_p. Then, the saddle point has a (p, q) periodic orbit associated and this orbit is called minimax orbit (see Figure 16).

4.2 Symmetry

In this section we are going to see examples of symmetric periodic orbits and we will do a classification. To do this, we will see some examples where the billiard table is elliptic.

Consider a billiard table such that its boundary C is symmetric respect a line. When C is an ellipse centered in $(0, 0)$, we have that the symmetric axis are x-axis and y-axis. Other example is the circumference whose symmetric axis are all lines that contains the center of the circumference.

We say that a (p, q) periodic orbit is symmetric if the figure created by the trajectories generated by the orbit segment $((s_0, \theta_0), ..., (s_q, \theta_q))$ is symmetric respect one symmetric axis of the curve C. For example, in Figure 17 the $(1, 3)$ periodic orbit is symmetric respect y-axis.

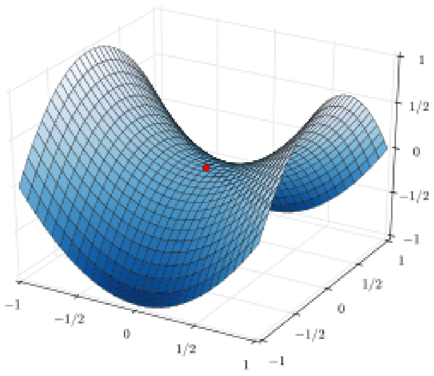


Figure 16

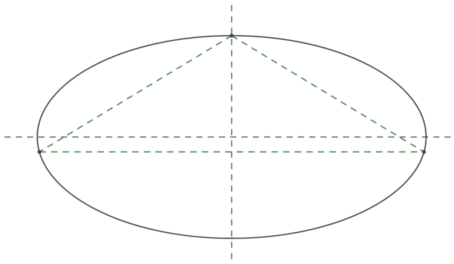


Figure 17: Symmetric periodic orbit respect y-axis

Now, we will focus to the curves C such that has two symmetric axis and they are perpendicular, for example the ellipse or the curve $x^2 + y^4 = 1$. Under these hypothesis we can do a classification that depends on the period and the axis of symmetry.

First, suppose the period is even. We can distinguish two cases, when an axis of symmetry contains a point of the orbit and the opposite as we can see in Figure 18. We can observe that in both cases the period orbit is symmetric respect both axis of symmetry.

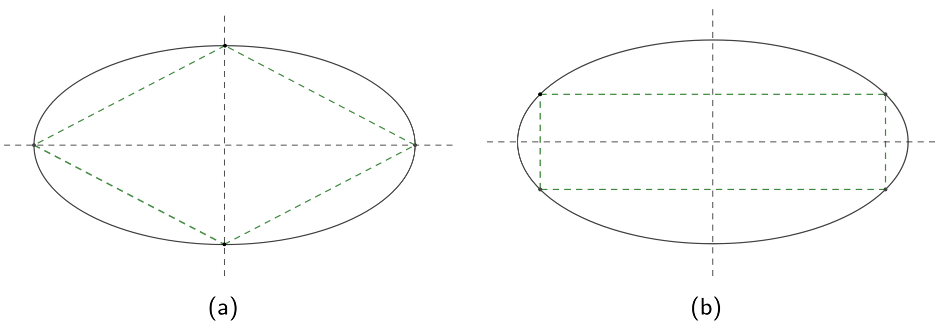


Figure 18: Symmetric periodic orbits with period 4.

Now, suppose that the period is odd. In this case, if the orbit has a point in an axis of symmetry, the other axis of symmetry cannot contain a point of the orbit. This implies that the orbit is only symmetric respect one axis of symmetry as we can see in Figure 19.

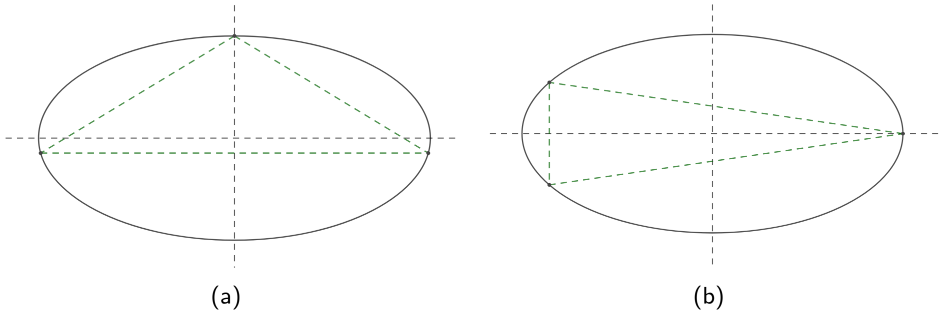


Figure 19: Symmetric periodic orbits with period 3.

In section 5 we shall study of symmetric periodic orbits when the billiard table is the curve $x^2 + y^4 = 1$ and the orbits have odd period and one point of the orbit is $(0, 1)$, so the axis of symmetry is the y-axis. These orbits are to the one shown in Figure 19 (a).

5. Numerical results

In the previous section we have seen a classification of symmetries. In this section we will focus in the orbits that are symmetric respect the y-axis and have an impact point on the y-axis.

Suppose the billiard table is the region of the plane delimited by the curve

$$C = \{(x, y) \in \mathbb{R}^2; x^2 + y^4 = 1\}.$$

Our goal is to study the periodic orbits with odd period, i.e., the period has the form $2k + 1$ where $k \in \mathbb{N}$ and we will impose that the initial point is $(0, 1)$ and the rotation number of this orbits is $1/(2k + 1)$.

Under these conditions, the first question that we raise is how many such orbits are there. Other question that appears is how the number of such orbits grows when the period increases and if there exists some relation between the period and the number of such periodic orbits.

In previous sections, the billiard map in the natural coordinates (s, θ) is defined in the space $S^1 \times [0, \pi]$. Now, we will do a program that reproduce the behaviour of this map. In the program, the points of the boundary will be expressed as a Cartesian coordinates $q_i = (x_i, y_i)$ and the direction vector of the trajectory will be expressed as $p_i = (u_i, v_i)$.

Now, we are going to focus in our orbits. We will study periodic orbits with period $2k + 1$ so from now on, we will only refer to the k that determines the period. On the other hand, since our orbits are symmetric respect the y-axis and have odd period, the trajectory between the iterates k and $k + 1$ should be horizontal, i.e., the direction vector should be $p_k = (u_k, 0)$.

Finding the orbits we describe above consists to solve the following problem:

Fix the initial point $q_0 = (0, 1)$ in the boundary of the billiard table. We will consider the three following functions

1. the inclusion $i : (0, \frac{\pi}{2}) \to \mathbb{R}^4$ such that $i(\theta_0) = (q_0, p_0)$ where $q_0 = (0, 1)$ and $p_0 = (-\cos(\theta_0), -\sin(\theta_0))$,
2. the billiard map $f : [-1, 1]^2 \times \mathbb{R}^2 \to [-1, 1]^2 \times \mathbb{R}^2$ that maps (q_j, p_j) to (q_{j+1}, p_{j+1}),
3. the projection $P_4 : \mathbb{R}^4 \to \mathbb{R}$ such that $P_4(x, y, u, v) = v$, i. e., P_4 gives us the fourth component.

Finally, we consider the function

$$Z_k : (0, \frac{\pi}{2}) \to \mathbb{R} \quad \text{such that} \quad Z_k = P_4 \circ f^k \circ i.$$

So, given an angle θ_0, we have

$$Z_k(\theta_0) = P_4(f^k(i(\theta_0))) = P_4(f^k(q_0, p_0)) = P_4(q_k, p_k) = v_k$$

i.e., the function Z_k gives us the second component of the direction in the iterate k (see Figure 20).

Therefore, the problem consists to find the zeros of the function Z_k since our orbits satisfy $p_k = (u_k, 0)$.

One of the difficulties we encountered in making the program was numerical precision. In Figure 21 we can see that in the interval $[0.0036, 0.0055]$ there are 99 zeros of the function. This implies that the computations should be preformed with high accuracy.

Initially, the program was written in C language and the variable used were type double. But these variables have 16 decimals and the program did not find the solutions with sufficiently precision to apply

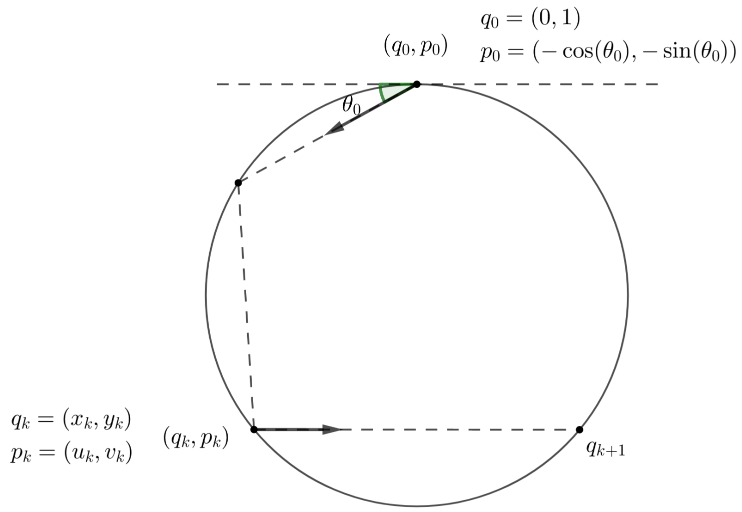


Figure 20: Dynamics of the billiard map.

Newton's method correctly and the program failed when $k = 60$. For this reason we used a library called quadmath.h where we could use a new format that has 28 decimals. However, this change was not sufficient since the program failed when $k = 100$ and finally the program was written in PARI since we can choose the precision of the variables. We have performed the more extreme computations with more than 100 decimal digits.

5.1 Description of the program

First, we have a function that we call *iterateFunction*. Let q_i be a point in C and p_i be a direction. The function returns the next point and the next direction under the billiard map. So this function computes the next iterate of an orbit.

The next step is reproduce the billiard behaviour and to do it we have to iterate the function *iterateFunction* several times. Suppose that we will find three consecutive iterates of an orbit. To do it we fix a initial point and a direction vector (pointing inside the table) of the orbit. Then, we execute the function one time with the initial conditions and we repeat this process two times using the returned values of the previous iterate as a new initial condition.

In our case, the goal is the periodic orbit with odd period, rotation number $1/(2k+1)$ and y-symmetric. So the will fix the initial condition $q_0 = (0, 1)$. As we have explained above, identify this orbits is equivalent to see if the trajectory between the iterate k and $k + 1$ is horizontal, hence we will fix the initial condition and iterate the function k times and we observe if the direction $p_k = (u_k, v_k)$ has $v_k = 0$. This condition is not sufficient to assure that the rotation number of the orbit is $1/(2k + 1)$, as we can see in Figure 22. So, other condition that the orbit need is the first k iterates satisfies that the point $q_i = (x_i, y_i)$ has $x_i < 0$, i.e., the orbits remains in the left half-plane defined by the y axis.

Under this two conditions and fixing the period, to calculate the times that this type of orbit appears we need fix $q_0 = (0, 1)$ and vary the direction vector as following: let θ_0 be the angle between the tangent line of C ant q_0 and the line defined by the direction, we have $p_0 = (u_0, v_0) = (-\cos(\theta_0), -\sin(\theta_0))$ so moving θ_0 between 0 and π provides us about all orbits. In our case, we can fins symmetric orbits, so we

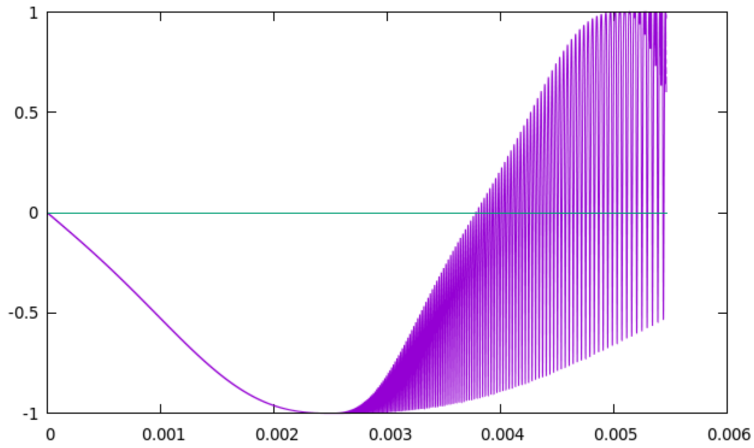


Figure 21: Relation between θ_0 and v_k, for $k = 240$.

only need move θ_0 between 0 an $\frac{\pi}{2}$.

Description of iterateFunction

Let C be the curve defined by $x^2 + y^4 = 1$. Suppose that we have a point $q_i = (x_i, y_i)$ and a unitary direction $p_i = (u_i, v_i)$ (pointing inside the table). We want to find the next iterate under the billiard map defined in the curve C, i.e., the point $q_{i+1} = (x_{i+1}, y_{i+1})$ and the vector $p_{i+1} = (u_{i+1}, v_{i+1})$. We call l the line defined by q_i and p_i. The new point q_{i+1} remains in this line and, simultaneously, remains in the curve C so this new point satisfies

$$(x_{i+1}, y_{i+1}) = (x_i, y_i) + \tau(u_i, v_i),$$

where $\tau > 0$ (see Figure 23).

The function computes τ such that $(x_i, y_i) + \tau(u_i, v_i) \in C \cap l$. To find τ we impose that $x = x_i + \tau u$ and $y = y_i + \tau v$ and $x^2 + y^4 = 1$. Consider the function

$$g(\tau) = (x_i + \tau u_i)^2 + (y_i + \tau v_i)^4 - 1.$$

Applying Newton method over this function we obtain the point (x, y) we the precision that we impose in the method. This point is the new iterate $q_{i+1} = (x_{i+1}, y_{i+1})$.

When we have the point q_{i+1} we need to find the direction vector p_{i+1}. Let n_{i+1} be the normal vector at q_{i+1} pointing inside the table. In Figure 23 we can see the following relation:

$$p_i = \alpha t_{i+1} + \beta n_{i+1},$$

where t_{i+1} is the tangent vector, $\alpha = \frac{\langle t_{i+1}, p_i \rangle}{\langle t_{i+1}, t_{i+1} \rangle}$ and $\beta = \frac{\langle n_{i+1}, p_i \rangle}{\langle n_{i+1}, n_{i+1} \rangle}$, and by the reflection law we have

$$p_{i+1} = \alpha t_{i+1} - \beta n_{i+1} = \alpha t_{i+1} + \beta n_{i+1} - 2\beta n_{i+1} = p_i - \nu n_{i+1},$$

where $\nu = -2\frac{\langle p_i, n_{i+1} \rangle}{\langle n_{i+1}, n_{i+1} \rangle}$.

So, the direction p_{i+1} is determined and the new iterate is (q_{i+1}, p_{i+1}).

Description of the main function

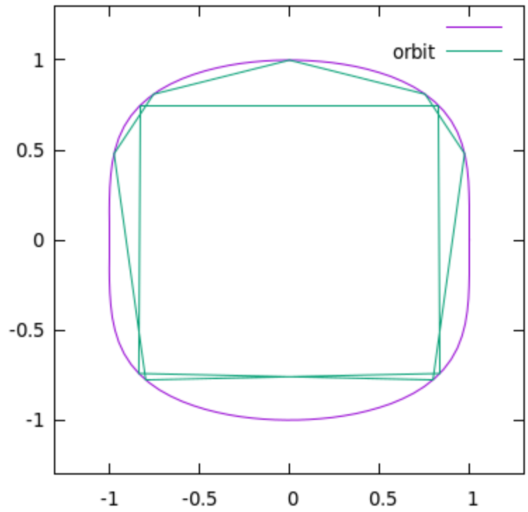


Figure 22: Orbit with period 11.

Now, our goal is find the periodic orbits described above.

Remember that is we call θ_0 the angle of the direction of the trajectory at initial point q_0, we can express the direction vector as $p_0 = (x_0, y_0) = (\cos(\theta_0), \sin(\theta_0))$.

We fix a step h. The program is defined by two loops. The first starts with $\theta_0 = h$ and inside this loop there is other loop that iterates the function *iterateFunction* k times to save the component v_k of the direction vector in the k iterate. Now, we fix again the initial conditions increasing the value of θ_0 ($\theta_0 = \theta_0 + h$) and we repeat the same process to save the component v_k of this new orbit. The main loop stops when some stop condition is satisfied, i.e., if the angle $\theta_0 = \frac{\pi}{2}$ or there is an orbit such that $x_k > 0$ and $y_k > x_k$.

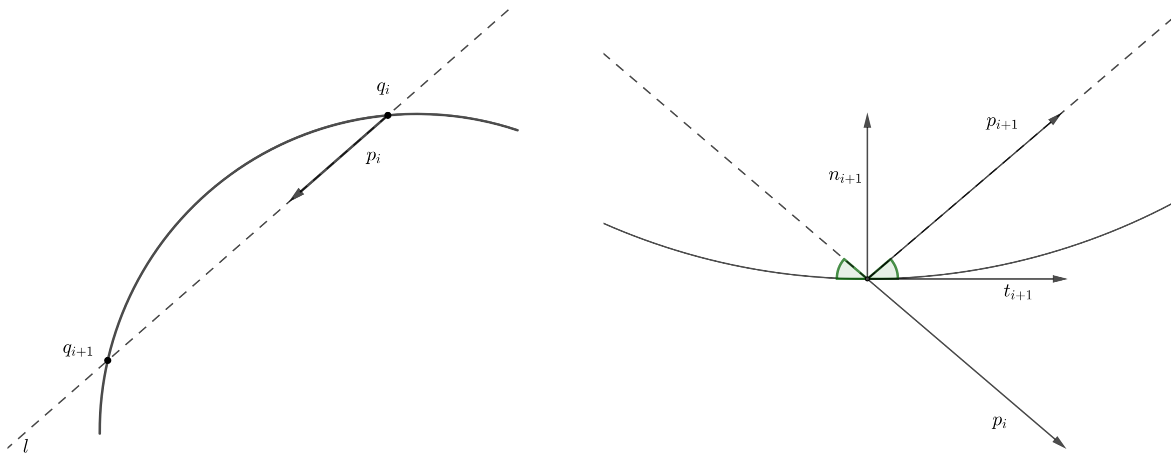
In this main function we obtain the following results:

1. Components v_k: For any k associated with period $2k+1$, we prove different initial conditions for θ_0. The program save the component v_k for each orbit. So if we draw this values we obtain a graphic that relates the angle θ_0 and the values of v_k. The zeros of this graphic correspond with the orbits that we find.

2. Intervals: Our program saves the components v_k. Then, we can save the interval when the component v_k change the sing that is the interval where the zero of the graphic is located. This implies that, fixed k, we can save all interval of the initial condition θ_0 where each orbit are contained therefore we have an estimation of the initial condition of these orbits.

3. For any k, we save the number of orbits (periodic and symmetric with rotation number is 1) appear, the value of k and the interval of θ_0 where appear all these orbits.

The step h depends of the period or, equivalently, the value of k. To decide this step we have executed the program for small k ($k = 10, k = 20$) and we have realized that when we multiply the value of k by 2, the number of periodic orbits doubles and the length of the interval where they appear halves. Then, for example to choose the step for $k = 40$ (h_{40}) we can calculate $h_{40} = h_{20}/4$. In general, if we choose a good step for $k = 20$ (h_{20}) the other steps are $h_k = h_{20}(\frac{20}{k})^2$.



(a) New point q_{i+1}. (b) New direction p_{i+1}.

Figure 23

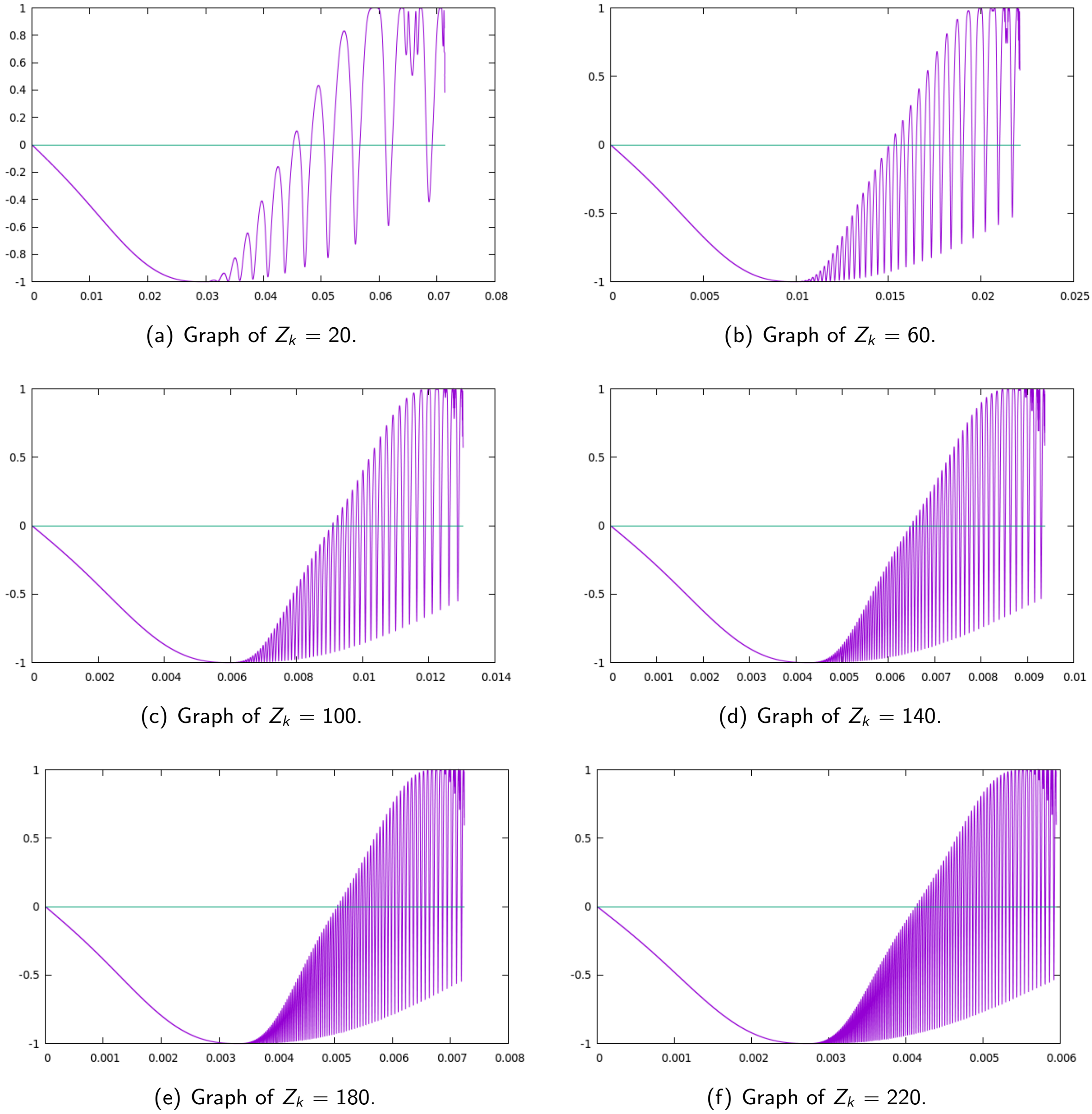
Since our goal is obtain the number of symmetric periodic orbits which rotation number is $1/(2k+1)$, when the program find the first periodic orbit with rotation number greater than Images of $1/(2k+1)$, it stops. But we can not guarantee that after there exists an orbit with rotation number equal to Images of $1/(2k+1)$. Therefore, the program only gives us a lower bound of the number of symmetric periodic orbits with rotation number Images of $1/(2k+1)$.

5.2 Analysis of the results

In this section we are going to see different results for different values of k and they are presented in the following table. The first column indicates the value of k that corresponds with the results of the periodic orbits with period $2k+1$. In the second column there is the lower bound for the quantity of symmetric periodic orbits with rotation number $1/(2k+1)$ and initial position $(0,1)$. In the third and fourth columns there are the minimum and maximum values of the interval $[a,b]$ where the orbits are located.

k	n(k)	a(k)	b(k)
20	11	0.04515819152	0.07135487634
40	17	0.02287729655	0.03295807921
60	25	0.01531634677	0.02210007196
80	35	0.01131739084	0.01660011565
100	41	0.009091481615	0.01303980509
120	49	0.007598678339	0.01090283008
140	57	0.006527308478	0.009376407510
160	67	0.005670984411	0.008221585129
180	73	0.005051525380	0.007238510268
200	81	0.004554418378	0.006532498081
220	89	0.004146496545	0.005949672563
240	99	0.003783215851	0.005463043983

Remember that, for any k we obtain a file where the first column is the initial condition of the angle θ_0 and the second column is the component v_k, i.e., the second component to the direction vector when we iterate the initial condition k times. The graphics of Figure 24 correspond to this file with different values of k.



(a) Graph of $Z_k = 20$.

(b) Graph of $Z_k = 60$.

(c) Graph of $Z_k = 100$.

(d) Graph of $Z_k = 140$.

(e) Graph of $Z_k = 180$.

(f) Graph of $Z_k = 220$.

Figure 24

The first observation is that the number $n(k)$ of roots of Z_k increases when the value of k increase. So, we can do the graphic that relates the number of orbits with the value of k and we can see that there is a lineal relation (see Figure 25). So we can suppose that

$$n(k) \approx \alpha k + \beta$$

where $\alpha, \beta \in \mathbb{R}$.

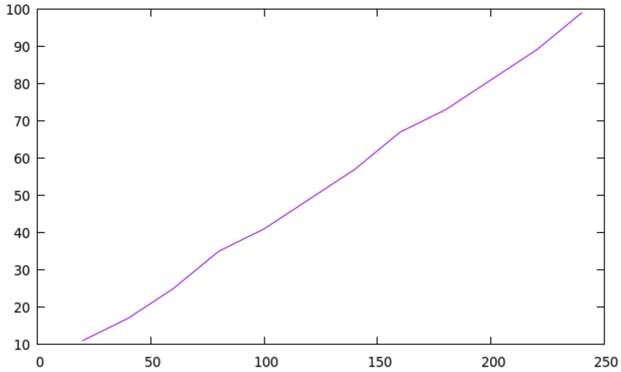


Figure 25: Relation between $n(k)$ and k.

To do an approximate value for β we have executed the program for small values of k and the results are the following:

k	nº Orbits
10	6
5	4
2	2
1	1

The case $k = 0$ corresponds with the orbits with period 1 but this only success if the angle $\theta_0 = 0$ or $\theta_0 = \pi$ because with this conditions the points are stationary as we had seen in a previous section. So we can do the approximation $\beta = 0$ and this implies:

$$n(k) \approx \alpha k$$

To do an approximation of α we are going to see the graphic that relates k and $\frac{n(k)}{k}$ (see Figure 26).

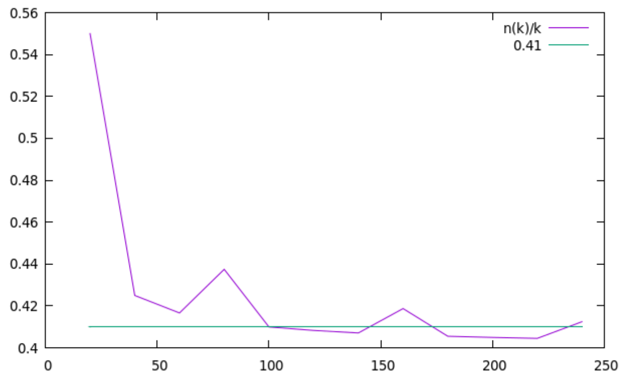
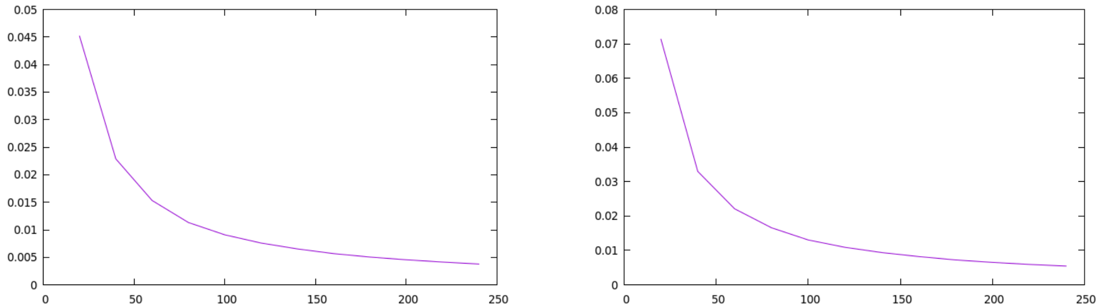


Figure 26: Relation between k and $n(k)/k$.

Therefore, an approximation is $\alpha = 0.41$. So, we can do the following approximation

$$n(k) \approx 0.41k.$$

The second observation about the graphics is when the value of k increase, the length of the interval where the zeros are located decrease. Then, we can see the graphics that relate the minimum and maximum value of the intervals with the value of k (see Figure 27).



(a) Relation of k and the minimum value of the interval where are contained our orbits

(b) Relation of k and the maximum value of the interval where are contained our orbits

Figure 27: Relation between k and the minimum and maximum values of the interval $[a, b]$

So, we can suppose the following approximations:

$$a(k) \approx \frac{a_*}{k}, b(k) \approx \frac{b_*}{k}$$

where $a_*, b_* \in \mathbb{R}$.

Now, to do an approximation of the values a_* and b_* we can observe Figure 28 that relates k with $a(k)k$ and $b(k)k$.

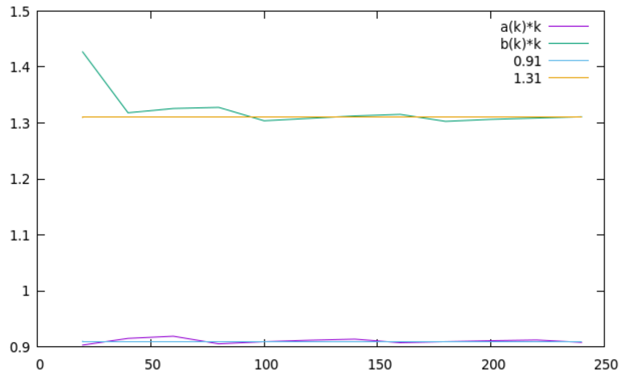


Figure 28: Relation of k with $a(k)k$ and $b(k)k$

and an approximation is $a_* = 0.91$ and $b_* = 1.31$.

Finally, with this results we can do the following conjecture:

Conjecture 5.1. *Let* $C = \{x^2 + y^4 = 1\}$ *be the curve that define the billiard map. The orbits that are periodic with period* $2k + 1$*, symmetric respect the y-axis with initial position* $(0, 1)$ *and the rotation number is* $\frac{1}{2k+1}$ *satisfy that there exists a constant* α *such that*

$$n(k) \approx \alpha k \quad \textit{with} \quad \alpha \approx 0.41,$$

where $n(k)$ is the lower bound of the quantity of orbits described above. There exists constants a_*, b_* such that the orbits are contained in the interval $[a(k), b(k)]$ where

$$a(k) \approx \frac{a_*}{k} \quad \textit{and} \quad b(k) \approx \frac{b_*}{k} \quad \textit{with} \quad a_* \approx 0.91, b_* \approx 1.31.$$

References

[1] J.D. Meiss. *Symplectic maps, variational principles and transport*, Reviews of Modern Physics, Vol.64, No. 3, July 1992.

[2] Vladimir Dragović, Milena Radnović.*Poncelet Porisms and Beyond*, Springer 2010.

[3] Eugene Gutkin, Anatole Katok.*Caustics for Inner and Outer Billiards*, Commun. Math Phys. 173,101-1033 (1995).

[4] John N. Mather.*Glancing billiards*, Print 1982.

[5] Andrea Hubacher.*Inestability of the Boundary in the Billiard Ball Problem*, Commun. Math Phys. 108, 483-488 (1987).

[6] V.F. Lazutkin. *The existence of caustics for a billiard problem in a convex domain*, 1973 Math. USSR.

[7] Markus Himmelstrand, Victor Wilén. *A survey of Dynamical Billiards*, Department of Mathematics, Analysis. Royal Institute of Technology (KTH).

[8] Wei-Feng Chen. *Birkhoff Priodic Orbits, Aubry-Mather sets, Minimal Geodesics and Lyapunov Exponents*, California Institute of technology. Pasadena, California, 1992.

[9] Serge Tabachnikov. *Geometry and Billiards*, Department of Mathematics, Penn State, University Park, PA 16802.

[10] Pablo S. Casas, Rafael Ramírez-Ros. *Classification of symmetric periodic trajectories in ellipsoidal billiards*, Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, December 2011.