On the Birkhoff Conjecture for Convex Billiards

Han Yong Wunrow

Submitted under the supervision of Richard Moeckel to the University Honors Program at the University of Minnesota-Twin Cities in partial fulfillment of the requirements for the degree of Bachelor of Science, *magna cum laude*, in Mathematics.

May 14, 2017

Abstract

On the Birkhoff Conjecture for Convex Billiards

by

Han Yong Wunrow

Bachelor of Science in Mathematics: Genomics Specialization

University of Minnesota, Twin Cities

The Birkhoff conjecture claims that if a convex billiard is integrable, then the boundary is necessarily an ellipse. In this paper, we analyze a local version of this conjecture that a small integrable perturbation of an ellipse must be an ellipse as proven in [3] by Kaloshin and Sorrentino. As an introduction, we discuss the behavior of billiards in a disc and in an ellipse. We then move onto some of the properties of a near-elliptical billiard. The dynamics of this system are modeled by a two-dimensional nonlinear mapping. The phase space of such billiards are of mixed kind with chaotic and regular behavior. We conclude by analyzing the behavior of the positive Lyapunov exponent and integrability of our system.

Contents

List of Figures

Background

Starting with the works of George David Birkhoff in early 1927, billiards have grown to become a popular topic of investigation in the framework of the theory of dynamical systems. Billiard systems are further remarkable for their appearance in a number of problems from optics, accoustics, and classical mechanics.

Similar to billiards in real life, the concept of a billiards system is fairly simple: you have a table and a ball, but in this case the ball has no mass so there is no friction. The ball hits the boundaries of the table and bounces off as it normally would. That is, it moves along in a straight line until it hits the edge of the table where it bounces off such that the angle of incidence equals the angle of reflection. Such systems become increasingly interesting in studying the trajectory of the ball as we change parameters such as the initial position, initial angle, and even the shape of the table.

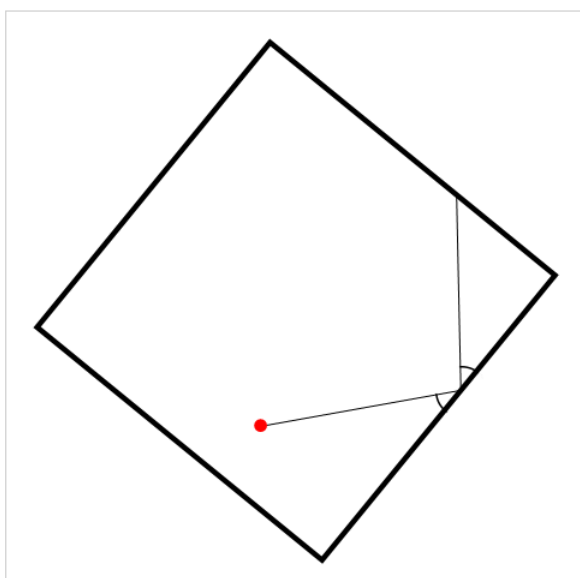


Figure 1.1: Square Billiard where angle of incidence equals the angle of reflection

1.1 Billiard Map

We can express these trajectories in the context of a dynamical system. Let $\theta \in [0, 2\pi)$ be the angular position measured in radians from the x-axis and $\alpha \in (0, \pi)$ be the angle of reflection as shown in Figure 1.2.

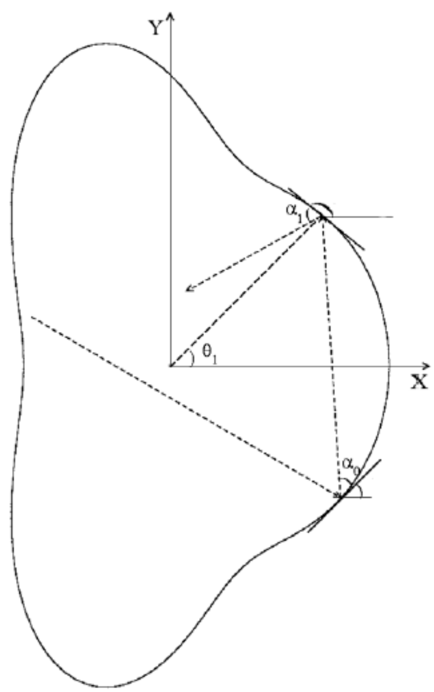


Figure 1.2: Illustration of particle's trajectory

The billiard ball map for this discrete dynamical system is defined as follows

$$\begin{aligned} f : [0, 2\pi) \times (0, \pi) &\to [0, 2\pi) \times (0, \pi) \\ (\theta_n, \alpha_n) &\mapsto (\theta_{n+1}, \alpha_{n+1}) \end{aligned} \tag{1.1}$$

There are two main restrictions on what shape our billiard can be. Let $\Omega \subset \mathbb{R}^2$ be a convex domain. We denote $\partial\Omega$ to be the boundary of the billiard. We require that the boundary of our billiard to be:

- convex, and
- smooth

Note that if the boundary is not convex, then f is not continuous. This is because there may be trajectories that are tangent to the boundary on the inside. If such trajectories are perturbed slightly, the resulting trajectory is drastically changed. We require the boundary to be smooth since the motion of a trajectory that hits a corner is not well defined. Now that we have laid out the basics of the billiard dynamical system, we examine trajectories in a few elementary boundaries.

1.2 Billiard in the Disc

Although the billiard in a disc is one of the simplest billiards to analyze, there are a few interesting properties. Due to its geometry, the billiard map in a disc can be described by a very simple system of equations.

$$\begin{pmatrix} \theta_{n+1} \\ \alpha_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} \theta_n \\ \alpha_n \end{pmatrix} \tag{1.2}$$

The angle of reflection remains constant after each iteration, and the position angle along the boundary is obtained by rotating the previous position by 2α. An interesting remark is that if α_0 is π-rational, then our orbit is periodic. Otherwise if α_0 is π-irrational, then the orbit is dense and hits every position along our boundary. Orbits of both types along with their phase space diagrams can be seen in figures 1.3 and 1.4.

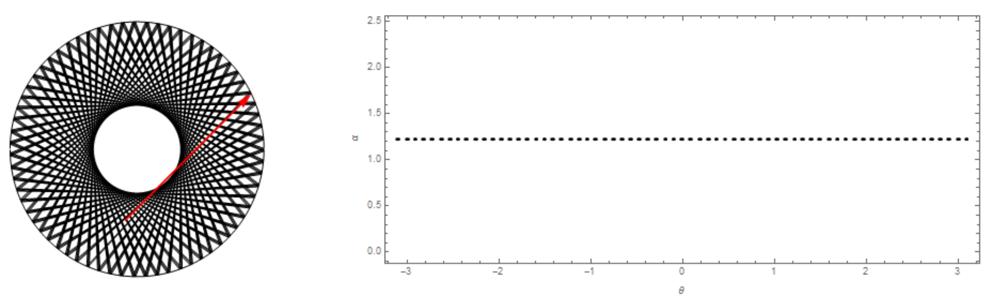


Figure 1.3: Disc Billiard with dense orbit

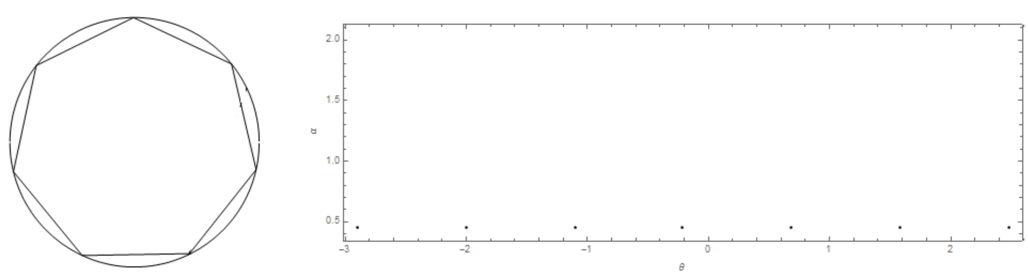


Figure 1.4: Disc Billiard with periodic orbit

One should take note of the concentric circle that is formed by the trajectories that stay tangent to it in Figure 1.3. We refer to such curves as caustics.

Definition 1. *A curve $\Gamma \subset \Omega$ is said to be **caustic** if any billiard orbit having one segment tangent to Γ has all its segments tangent to Γ.*

One should also take note of the formation of invariant curves in the phase space corresponding to particular caustics. These relationships will be examined further in the next section.

1.3 Billiard in the Ellipse

We now consider billiards in slightly more complex shapes - ellipses where the boundary $\partial\Omega$ is defined by

$$\frac{x^2}{a} + \frac{y^2}{b} = 1 \text{ where } a > b \tag{1.3}$$

The trajectory in an elliptical billiard is related to its *confocal* conic sections. We say two conic sections are *confocal* if they share the same foci.

Remark 1. *A billiard trajectory inside Ω stays tangent to a caustic fixed confocal ellipse or hyperbola. More precisely, the billiard trajectories in Ω stay tangent to a caustic*

$$\mathbb{C}_\lambda \in \left\{\frac{x^2}{a-\lambda} + \frac{y^2}{b-\lambda} = 1\right\} \tag{1.4}$$

where $a > b$.

Proof. This result is equivalent to Poncelet's closure theorem and its elementary geometric proof is shown in [6]. □

It is necessarily true that the confocal conics of Ω are either hyperbolas or ellipses. Consider the function

$$f = \frac{x^2}{a-\lambda} + \frac{y^2}{b-\lambda} - 1 = 0 \qquad a > b \tag{1.5}$$

For fixed values of x and y we can graph f as a function of λ as shown in figure 1.5. As seen in the graph the roots are $\lambda_1 \in (-\infty, b)$ and $\lambda_2 \in (b, a)$. Note that $\lambda = 0$ defines the original ellipse $\partial\Omega$ and since we are only interested in confocal conics contained within $\partial\Omega$, $\lambda_1 \in (0, b)$. For $\lambda = \lambda_1$ equation 1.5 defines an ellipse, and for $\lambda = \lambda_2$ it defines a hyperbola. Trajectories of both types are shown below in figures 1.7 and 1.6.

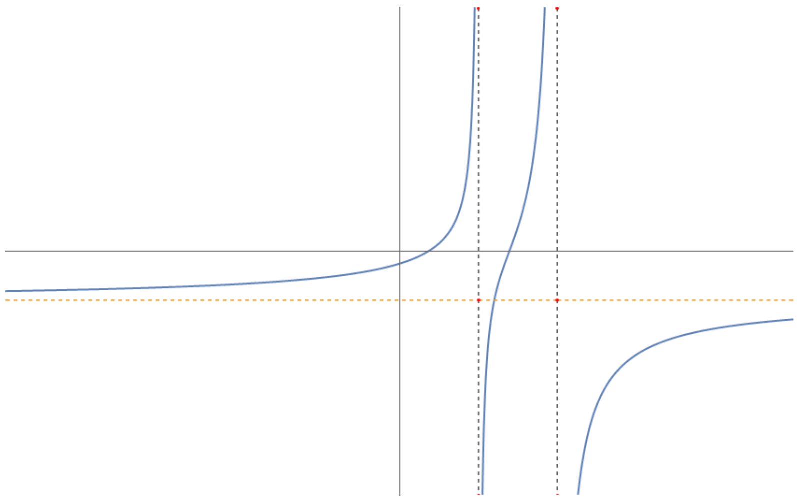


Figure 1.5: Graph of f as a function of λ for fixed x and y

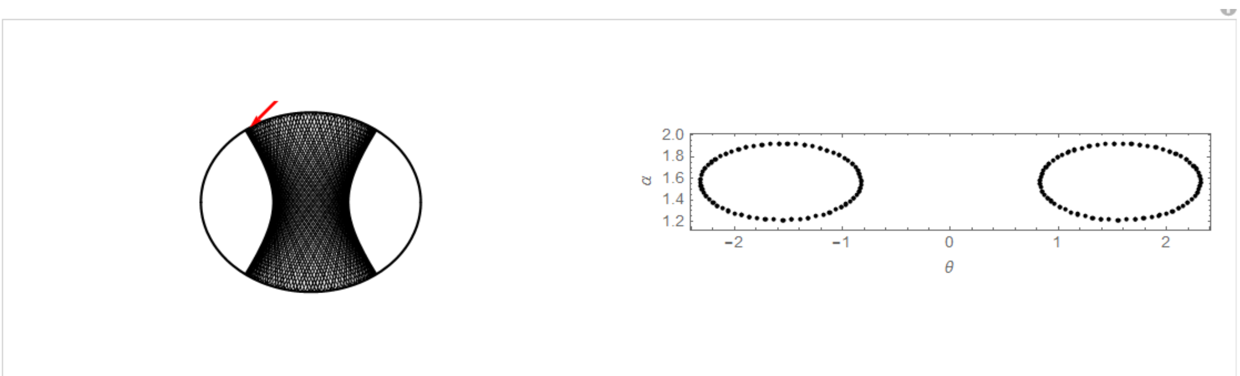


Figure 1.6: Confocal Hyperbola Elliptical Billiard with Phase Space Diagram $a = 1, b = 1.5$

Similarly to a billiard in a disc, one can see a relationship between invariant curves in the phase space that correspond to either a confocal hyperbola or confocal ellipse caustic. If we plot these invariant curves for multiple initial conditions as seen in Figure 1.8, one can see that the phase space is completely foliated by invariant curves of the billiard map f.

This brings us to the topic of *integrability*. A billiard in an ellipse (and also a disc) are examples of *integrable billiards*. There are a few equivalent definitions of what it means for a billiard to be integrable. We provide two, one in the context of caustics and another in the context of invariant curves.

Definition 2. *A billiard Ω is said to be **integrable** if the union of all smooth, convex caustics, has nonempty interior.*

Definition 3. *A billiard is said to be **integrable** if there exists a (smooth) foliation of the whole phase space consisting of invariant curves of the billiard map*

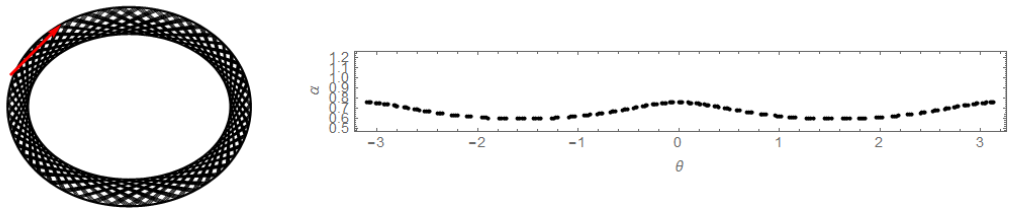


Figure 1.7: Confocal Ellipse Elliptical Billiard with Phase Space Diagram $a = 1, b = 1.5$

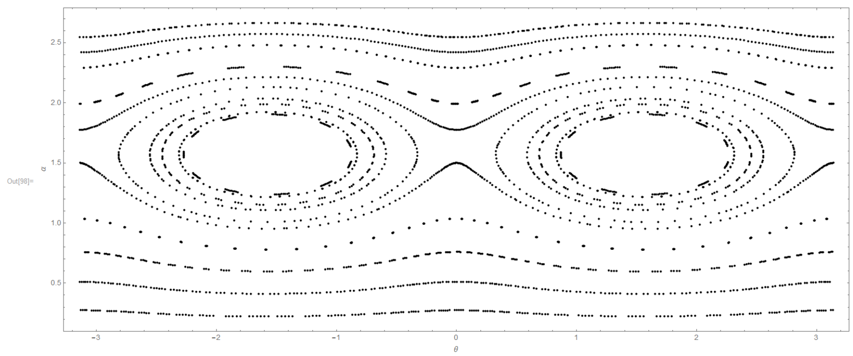


Figure 1.8: Elliptical Billiard Phase Space Diagram $a = 1, b = 1.5$

Remark 2. *A billiard in an ellipse is integrable.*

Proof. It follows from remark 1 that the smooth, convex caustics of an elliptical billiard are confocal ellipses. It is clear that the union of all smooth confocal ellipses has nonempty interior. Hence all elliptical billiards are integrable. □

It is well known that elliptical billiards are integrable. However, a long standing open problem in billiard systems is the converse. That is, proving that the only integrable billiards are ellipses.

1.4 Birkhoff Conjecture

Birkhoff Conjecture *If a billiard in* Ω *is integrable, then the boundary* $\partial\Omega$ *is an ellipse.*

Birkhoff's question of the existence of other integrable billiards other than ellipses has long been studied. We will briefly describe the progress that has been made to towards proving this conjecture. In [4], Mather proved the non-existence of caustics (hence, the non-integrability) if the curvature of the boundary vanishes at one point. This further justifies our restriction to strictly convex billiards. As stated by Mather, We will say a trajectory is ε-glancing if for at least one bounce the angle of reflection (with either the positive or negative tangent of $\partial\Omega$ at the point of reflection) is $< \varepsilon$. If $\varepsilon < \pi/2$, we can distinguish between a positively ε-glancing trajectory and a negatively ε-glancing trajectory according to whether it is the positive or negative tangent to $\partial\Omega$ which the direction of reflection is close to. A trajectory might be positively ε-glancing at one bounce and negatively ε-glancing at another bounce. Thus, one can ask whether for every $\varepsilon > 0$, there exist trajectories which are both positively and negatively ε -glancing. We will prove a theorem which shows in some cases the answer is yes.

Theorem 1. *If the curvature of $\partial\Omega$ vanishes at some point, then for every $\varepsilon > 0$, there exist trajectories which are both positively and negatively ε-glancing. (Mather; 1982)*

In [2], Bialy proved the following result concerning global integrability.

Theorem 2. *If the phase space of the billiard ball map is globally foliated by continuous invariant curves which are not null-homotopic, then it is the circular billiard. (Bialy; 1993)*

Although billiards in a disc are the only examples of global integrable billiards, integrability itself is still an intriguing open problem. More recently, in [1], Avila, De Simoi, and Kaloshin proved a local version of this conjecture for ellipses with small eccentricity.

Theorem 3. *A small integrable perturbation of an ellipse of small eccentricity must be an ellipse. (Avila, Simoi, Kaloshin; 2016)*

In [3], Kaloshin and Sorrentino expand upon this result in which they prove a local version of this conjecture for arbitrary eccentricity.

Theorem 4. *A small integrable perturbation of an ellipse of arbitrary eccentricity must be an ellipse. (Kaloshin, Sorrentino; 2017)*

Note that Theorems 3 and 4 are only local versions of this conjecture. That is, there still may be examples of integrable billiards that are very different from ellipses. In the following section, similarly to in [3], we will analyze the integrability of perturbations of elliptical billiards with arbitrary eccentricity.

Near Elliptical Billiard

2.1 Billiard Mapping

We define the boundary of a near elliptical billiard to be

$$ax^2 + by^2 + \varepsilon x^4 = 1.$$

We begin by discussing the details needed for the construction of the billiard map that describes the dynamics of this problem. The radius of the boundary expressed in polar coordinates is given by

$$1 = R^2(a\cos^2\theta + b\sin^2\theta) + R^4(\varepsilon\cos^4\theta) \tag{2.1}$$

$$0 = -1 + x(a\cos^2\theta + b\sin^2\theta) + x^2(\varepsilon\cos^4\theta) \qquad (x = R^2) \tag{2.2}$$

$$R^2 = x = \frac{-(a\cos^2\theta + b\sin^2\theta) \pm \sqrt{(a\cos^2\theta + b\sin^2\theta)^2 + 4(\varepsilon\cos^4\theta)}}{2(\varepsilon\cos^4\theta)} \tag{2.3}$$

$$R(\theta, a, b, \varepsilon) = \sqrt{\frac{-(a\cos^2\theta + b\sin^2\theta) \pm \sqrt{(a\cos^2\theta + b\sin^2\theta)^2 + 4(\varepsilon\cos^4\theta)}}{2(\varepsilon\cos^4\theta)}}. \tag{2.4}$$

Given an initial position and direction (θ_n, α_n), the angle between the vector tangent to the curve and the vector perpendicular to the reflected trajectory is given as

$$\phi_n = \arctan\left(\frac{Y'(\theta_n)}{X'(\theta_n)}\right) \tag{2.5}$$

where

$$X(\theta_n) = R(\theta_n)\cos(\theta_n) \tag{2.6}$$

$$Y(\theta_n) = R(\theta_n)\sin(\theta_n) \tag{2.7}$$

and their derivatives are

$$X'(\theta_n) = R'(\theta_n)\cos(\theta_n) - R(\theta_n)\sin(\theta_n) \tag{2.8}$$

and

$$Y'(\theta_n) = R'(\theta_n)\sin(\theta_n) + R(\theta_n)\cos(\theta_n). \tag{2.9}$$

To obtain the new position θ_{n+1} we must solve the equation

$$Y(\theta_{n+1}) - Y(\theta_n) = \tan(\alpha_n + \phi_n)(X(\theta_{n+1}) - X(\theta_n)) \tag{2.10}$$

As seen in figure 1.2 the next trajectory angle α_{n+1} is given by

$$\alpha_{n+1} = \phi_{n+1} - (\alpha_n + \phi_n) \tag{2.11}$$

Thus the billiard mapping is defined as

$$\begin{cases} F(\theta_{n+1}) = 0 = Y(\theta_{n+1}) - Y(\theta_n) - \tan(\alpha_n + \phi_n)(X(\theta_{n+1}) - X(\theta_n)) \\ \alpha_{n+1} = \phi_{n+1} - (\alpha_n + \phi_n) \end{cases} \tag{2.12}$$

2.2 Numerical Results

Let us now discuss some of our results. Figures 2.1, 2.2, and 2.3 illustrate the destruction of both the confocal caustic and invariant curve in the phase space as we increase the parameter ε. As shown in Figures 2.4 and 2.5 we can see a complex hierarchy of behaviours in the phase space along with the formation of a large chaotic sea as we increase ε.

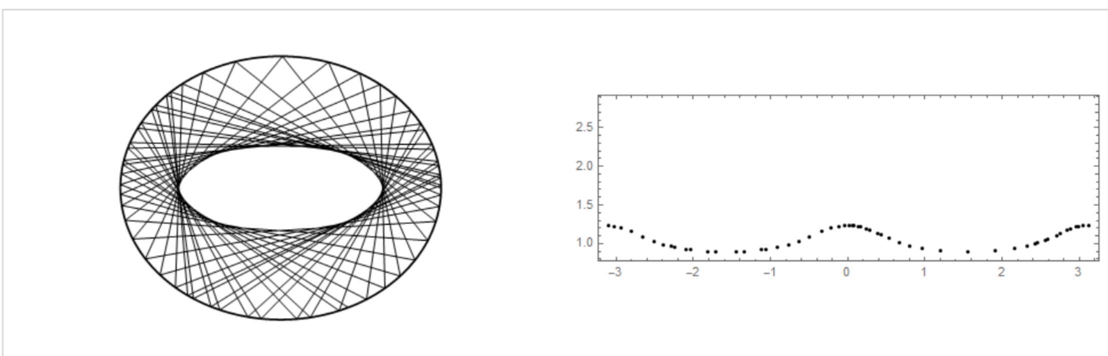


Figure 2.1: Near elliptical Billiard $a = 1, b = 1.5, \varepsilon = .01$ with Phase Space Diagram

Qualitatively, it is simple to see from the phase space diagrams that as we perturb an ellipse by increasing ε, the billiard no longer is integrable. To obtain a more quantitative sense of this chaotic motion we examine the Lyapunov exponent.

Definition 4. *As discussed in [5], the **Lyapunov exponent** is defined as*

$$\lambda_j = \lim_{n\to\infty} n^{-1} \log |\Lambda_j| \qquad j = 1, 2$$

where Λ_j are the eigenvalues of $M = \prod_{i=1}^{n} J_i(\theta_i, \alpha_i)$ and J_i is the Jacobian matrix evaluated over the orbit (θ_i, α_i).

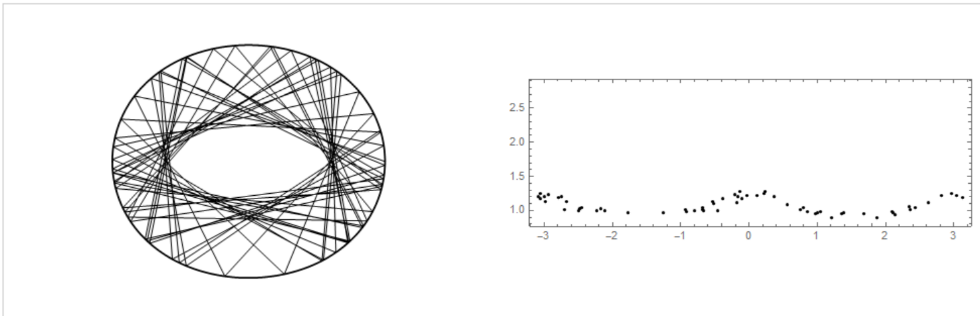


Figure 2.2: Near elliptical Billiard $a = 1, b = 1.5, \varepsilon = .1$ with Phase Space Diagram

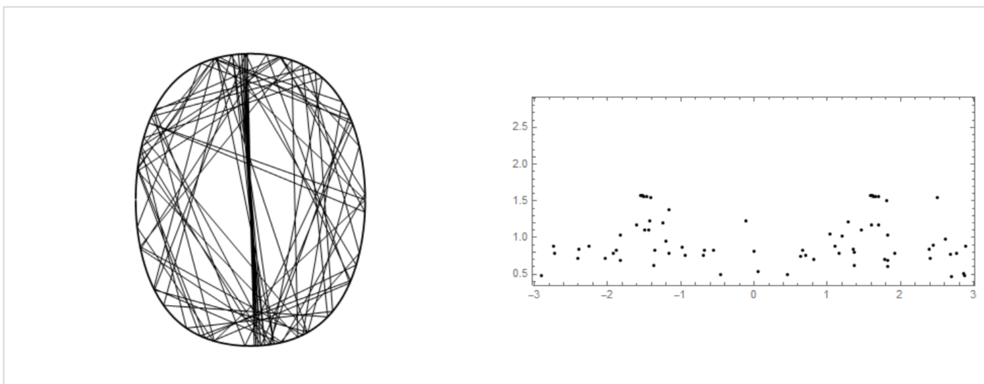


Figure 2.3: Near elliptical Billiard $a = b = \varepsilon = 1$ with Phase Space Diagram

It is well known that the Lyapunov exponent serves as a practical tool for studying the behavior of a dynamical system, and consequently whether the given system is chaotic. If the Lyapunov exponent converges to a positive value after significantly large n we say that our system is chaotic.

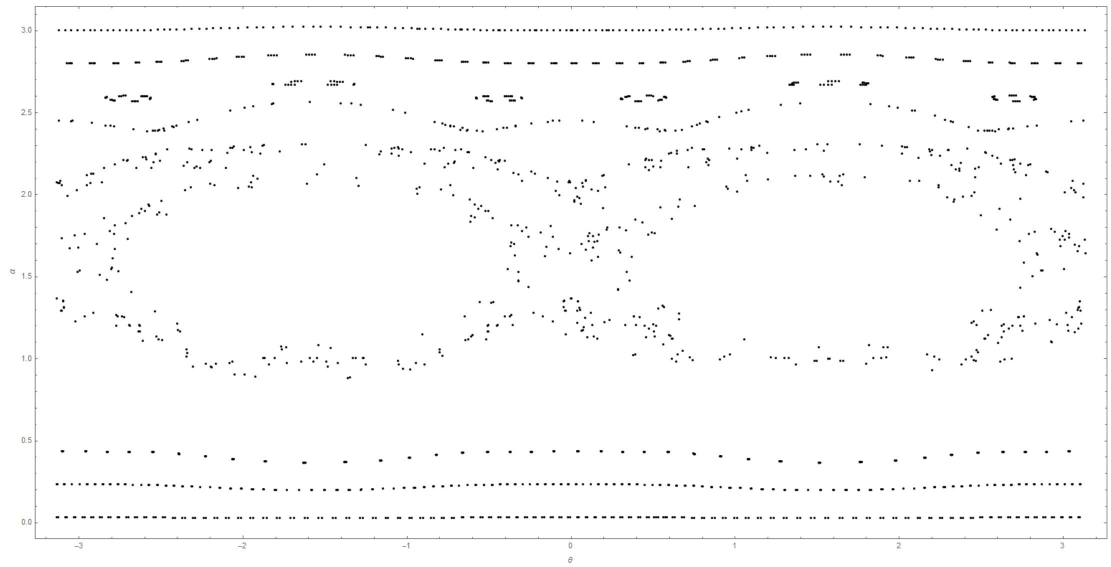


Figure 2.4: Near elliptical Billiard $a = 1, b = 1.5, \varepsilon = .1$ with Phase Space Diagram

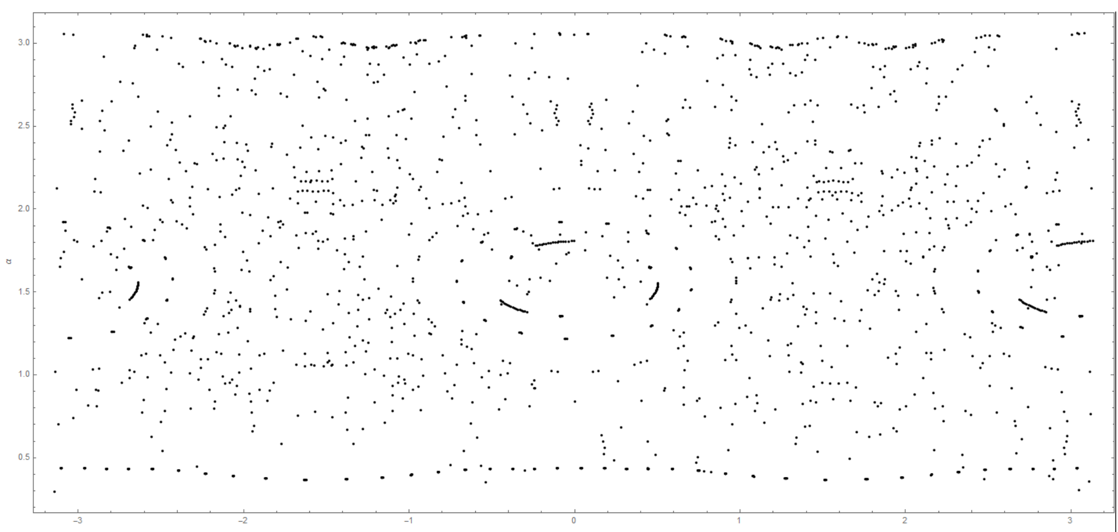


Figure 2.5: Near elliptical Billiard $a = 1, b = 1.5, \varepsilon = 10$ with Phase Space Diagram

Since our billiard mapping 2.12 is not well defined to compute the product of Jacobians, we discuss a method to compute the Lyapunov exponents numerically.

1. Starting with initial conditions θ_{u0} and α_{u0} (call orbit u), choose a nearby point $(\theta_{v0}, \alpha_{v0})$ (call orbit v) that is within distance d_0.

2. Iterate both orbits once and calculate the new distance d_1.

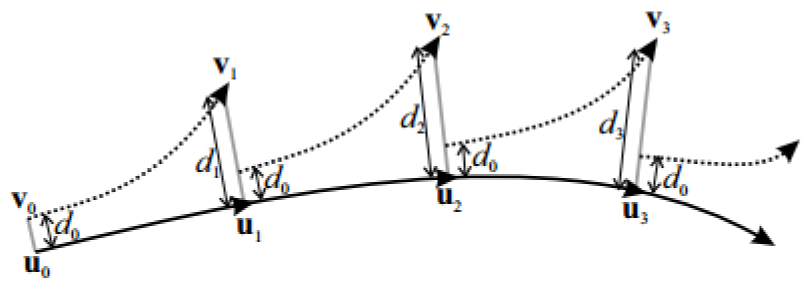


Figure 2.6: Diagram of how to compute the Lyapunov Exponent numerically

3. Evaluate $\ln(d_1/d_0)$.

4. Readjust the point along orbit v so that its distance from the new point on orbit u is d_0 and in the same direction as d_1. i.e.

$$\theta_{v0} = \theta_{u1} + d_0(\theta_{v1} - \theta_{u1})/d_1$$

   and

$$\alpha_{v0} = \alpha_{u1} + d_0(\alpha_{v1} - \alpha_{u1})/d_1$$

5. Repeat steps 2 - 4 n times and calculate a running average of step 3.

This process is illustrated in Figure 2.6.

Due to the precision of floating point numbers in Mathematica, we choose $d_0 = 10^{-8}$. To illustrate that our near elliptical billiard has a chaotic component in the phase space, we show in Figures 2.7 and 2.8 the behavior of the positive Lyapunov exponent as a function of the number of collisions n for various initial conditions chosen in the chaotic sea. Each initial condition was iterated up to 500 times.

In future work, many more iterations are needed to provide a proper convergence plot of the positive Lyapunov exponent. However, it is still clear to see from the plots that it will eventually converge to a positive value. As such we can conclude that the regions of our phase space for near elliptical billiards are chaoitic and that the system is no longer integrable due to the destruction of the invariant spanning curves in the phase space.

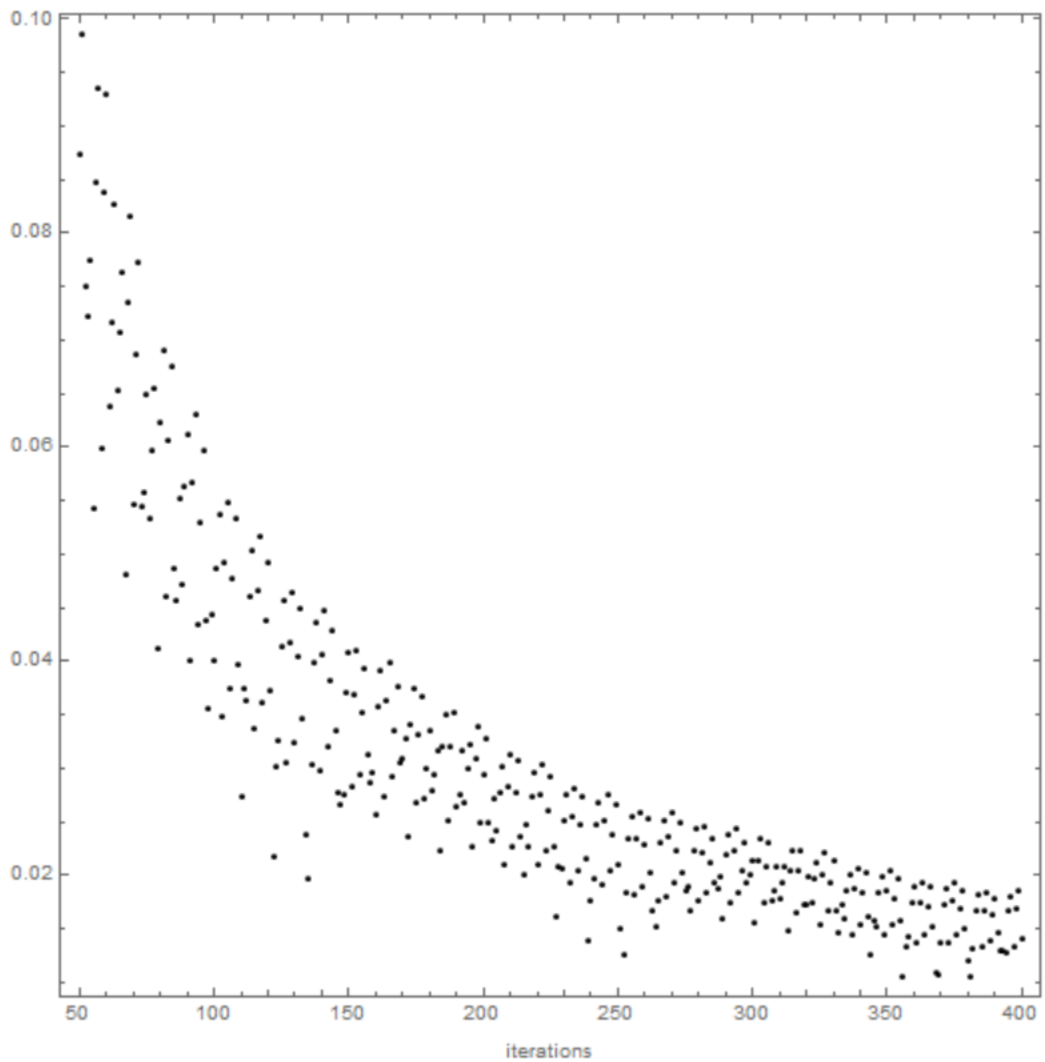


Figure 2.7: Behavior of the Lyapunov Exponent for $a = 1$, $b = 1.4$, and $\varepsilon = 1.5$

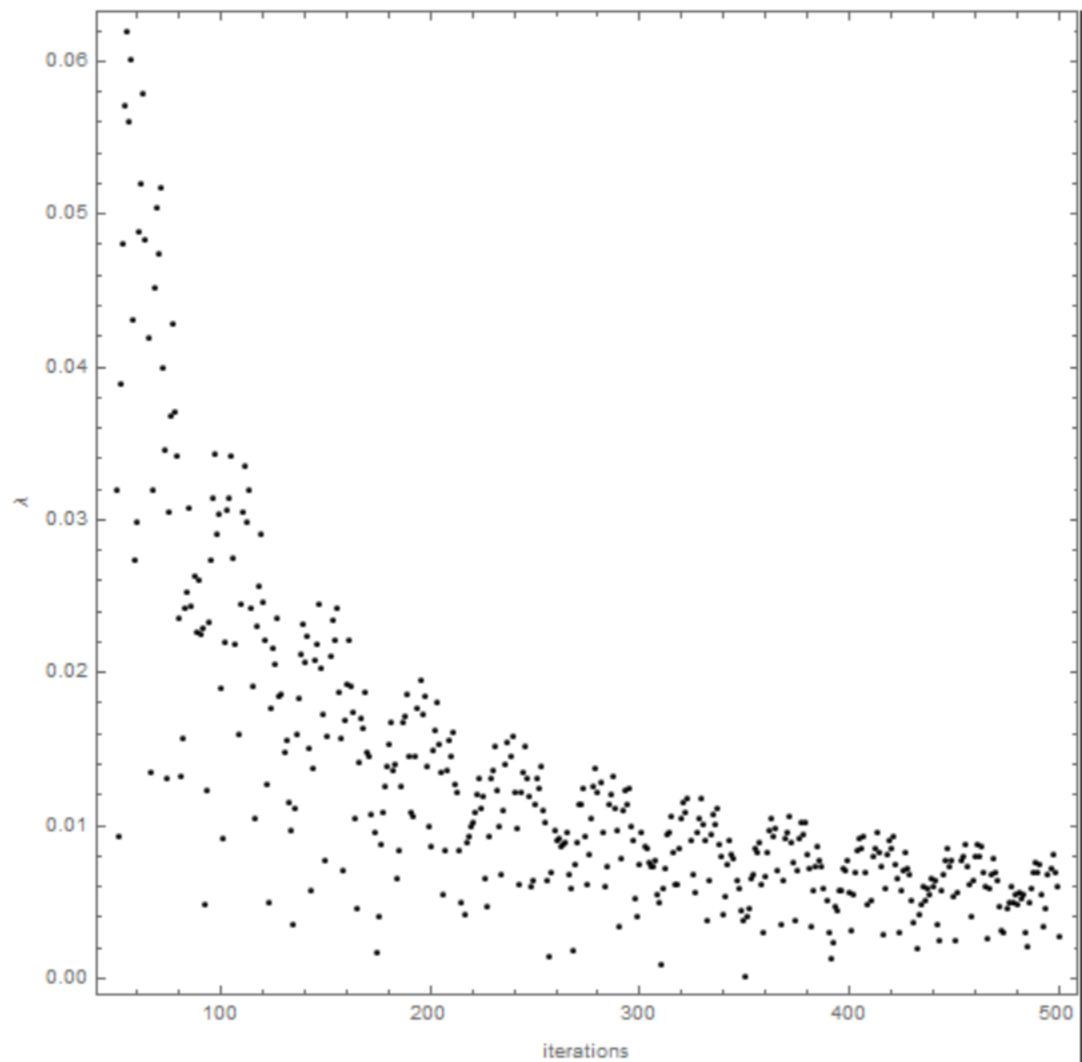


Figure 2.8: Behavior of the Lyapunov Exponent for $a = 1$, $b = 1.4$, and $\varepsilon = 0.1$

Conclusion

As a conclusion, we have provided a brief background of the behavior of billiards in a disc and ellipse. We then obtained a nonlinear mapping that describe the dynamics of a near elliptical billiard. We have shown that the phase space has a mixed form, and lastly, we have shown that the Lyapunov exponent converges to a positive value and thus proves that near elliptical billiards are not integrable due to the destruction of some of the invariant spanning curves in the phase space.

Appendix

We provide the Mathematica used to create the plots of trajectories in the paper along with the computations of the Lyapunov exponent.

```
f[a_, b_, ε_, x_, y_] = a*x^2 + b*y^2 + ε*x^4 - 1;
```

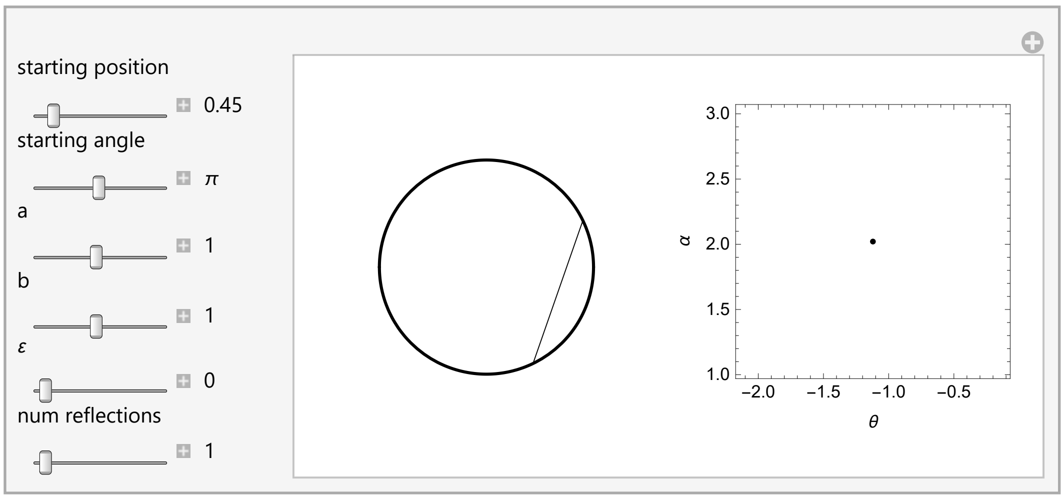
ellplus[a_, b_, ε_, x_] = $\frac{\sqrt{1 - a\,x^2 - x^4\,\varepsilon}}{\sqrt{b}}$;

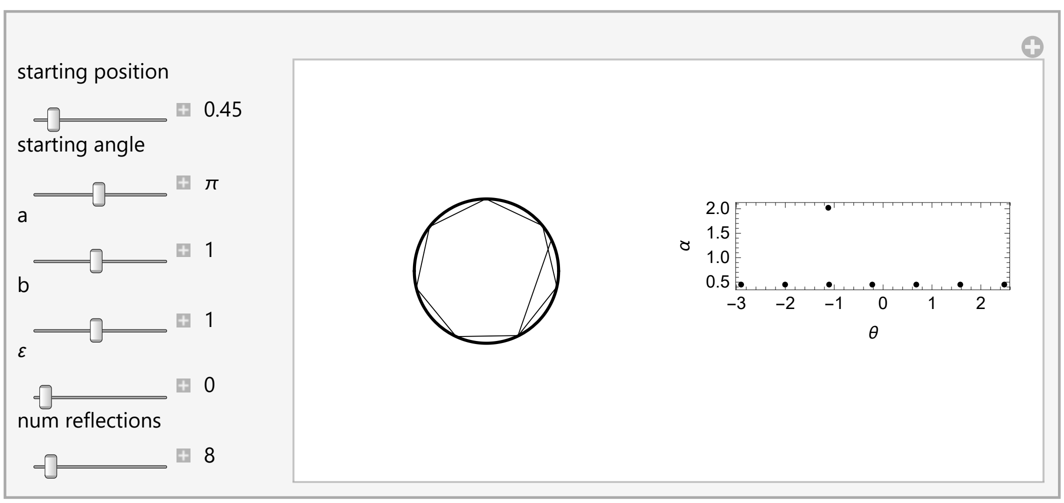
ellminus[a_, b_, ε_, x_] = $-\frac{\sqrt{1 - a\,x^2 - x^4\,\varepsilon}}{\sqrt{b}}$;

```
gradf[a_, b_, ε_, x_, y_] = Grad[f[a, b, ε, x, y], {x, y}];
```

```
nextCoord[a_, b_, startθ_, startα_, ε_, n_] :=
 Module[{start, s, startx, starty, startt, endx,
   endy, angle, list, ss, xx, yy, u, slope, v, ϕ, plot, coord, gr1},
  start = NSolve[{f[a, b, ε, x, y] == 0, x == t * Cos[startθ],
     y == t * Sin[startθ], t ≥ 0}, {x, y, t}, Reals];
  s = {x, y, t} /. start;
  startx = s[[1, 1]];
  starty = s[[1, 2]];
  startt = s[[1, 3]];
  coord = {{startx, starty}};
  angle = startα;
  j = 1;
  list = {}; (* stores position angles and trajectory angles*)
  While[j ≤ n,
   ss = NSolve[{f[a, b, ε, x, y] == 0, x == startx + t * Cos[angle],
      y == starty + t * Sin[angle], t ≥ 0}, {x, y, t}, Reals];
   xx = ss[[All, 1, 2]];
   yy = ss[[All, 2, 2]];
   (*Update for next iteration of while loop*)
   startx =
    If[startx ≤  xx[[1]] + 0.0000001 && startx ≥  xx[[1]] - 0.0000001 , xx[[2]], xx[[1]]];
   starty = If[starty ≤  yy[[1]] + 0.0000001 && starty ≥  yy[[1]] - 0.0000001,
     yy[[2]], yy[[1]]];
   endx = If[startx ≤  xx[[1]] + 0.0000001 && startx ≥  xx[[1]] - 0.0000001 ,
     xx[[2]], xx[[1]]];
   endy = If[starty ≤  yy[[1]] + 0.0000001 && starty ≥  yy[[1]] - 0.0000001,
     yy[[2]], yy[[1]]];
   coord = Append[coord, {startx, starty}];
   u = {endx - startx, endy - starty};
   v = gradf[a, b, ε, startx, starty].{{0, -1}, {1, 0}};
   ϕ = VectorAngle[u, v];
   list = Append[list, {ArcTan[startx, starty], ϕ}];
   angle =  angle + 2 * ϕ;
   j++;
  ];
  plot = Plot[{ellplus[a, b, ε, x], ellminus[a, b, ε, x]}, {x, -1/Sqrt[a] - .5,
     1/Sqrt[a] + .5}, Axes → False, PlotStyle → Black, AspectRatio → Automatic];
  gr1 = Show[plot,
    Graphics[{Line[coord]}]
    (*,Graphics[{PointSize[Large],Red,Thickness[Large],Arrow[{{ s[[1,1]]+1*Cos[startα],
         s[[1,2]]+1*Cos[startα]},{s[[1,1]],s[[1,2]]}}]}]*)];
  gr2 = Graphics[Map[Point, list], Frame → True, FrameLabel → {θ, α},
    AspectRatio → Automatic];
  GraphicsRow[{gr1, gr2}]
 ]
```

```
Manipulate[
 nextCoord[a, b, startθ, startα, ε, n],
 "starting position",
 {{startθ, .45, ""}, 0, 2 * Pi, .01, Appearance → "Labeled", ImageSize → Tiny},
 "starting angle",
 {{startα, Pi, ""}, 0, 2 * Pi, .01, Appearance → "Labeled", ImageSize → Tiny},
 "a",
 {{a, 1, ""}, .1, 2, .01, Appearance → "Labeled", ImageSize → Tiny},
 "b",
 {{b, 1, ""}, .1, 2, .01, Appearance → "Labeled", ImageSize → Tiny},
 "ε",
 {{ε, 0, ""}, 0, 100, .01, Appearance → "Labeled", ImageSize → Tiny},
 "num reflections",
 {{n, 1, ""}, 1, 200, 1, Appearance → "Labeled", ImageSize → Tiny},
 SaveDefinitions → True
]
```






```
gradf[1, 1.5, 0, 1, 0].{{0, -1}, {1, 0}}
```

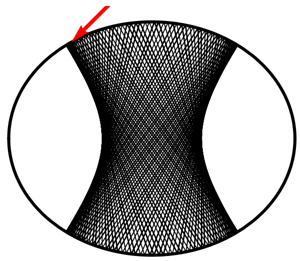
{0., -2.}

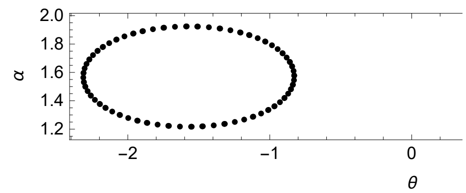
```
nextCoord[1, 1.5, 2.26, 5.28, 0.1 , 20]
```

$Aborted

```
p1 = gr2;
```

```
nextCoord[1, 1.5, 2.26, 5.28 - .2, 0 , 200]
```



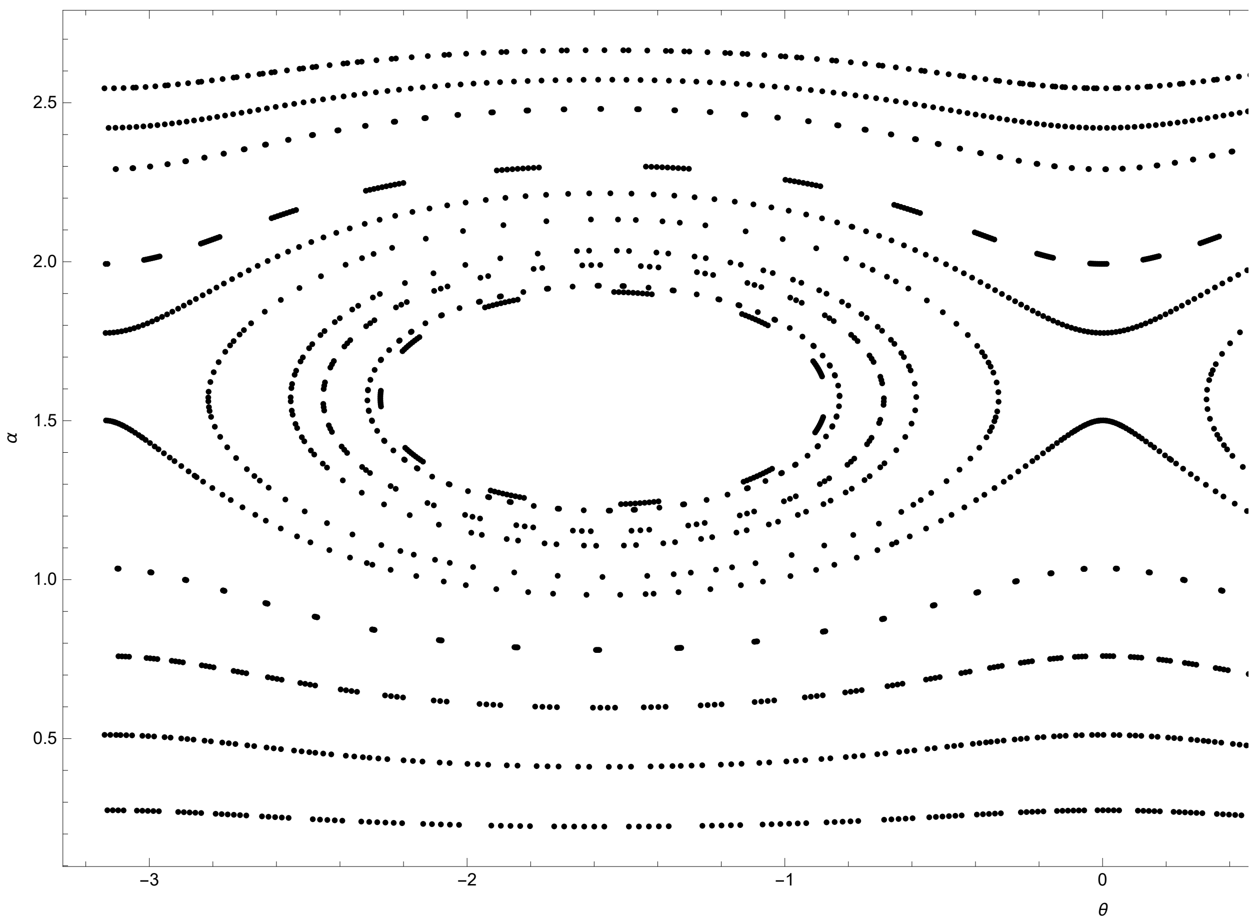



```
Pi
```

π

```
p2 = gr2;
```

```
Show[p1, p2, p3, p4, p5, p6, p7, p8, p9, p10, p11, p12, p13, p14, p15]
```



Bibliography

[1] Artur Avila, Jacopo De Simoi, and Vadim Kaloshin. "An integrable deformation of an ellipse of small eccentricity is an ellipse". In: *Annals of Mathematics* 184.2 (2016), pp. 527–558. ISSN: 0003-486X. DOI: 10.4007/annals.2016.184.2.5. arXiv: 1412.2853. URL: http://annals.math.princeton.edu/2016/184-2/p05.

[2] Misha Bialy. "Convex billiards and a theorem by E. Hops." In: *Mathematische Zeitschrift* 214.1 (1993), pp. 147–154. URL: http://eudml.org/doc/174561.

[3] Vadim Kaloshin and Alfonso Sorrentino. "On Local Birkhoff Conjecture for Convex Billiards". In: figure 1 (2016), pp. 1–47. arXiv: 1612.09194. URL: http://arxiv.org/abs/1612.09194.

[4] J N Mather. "Glancing Billiards". In: *Erg. Theory and Dyn. Sys.* 2.May (1982), pp. 397–403. ISSN: 14694417. DOI: 10.1017/S0143385700001681.

[5] Diego F M Oliveira and Edson D. Leonel. "On the dynamical properties of an elliptical-oval billiard with static boundary". In: *Communications in Nonlinear Science and Numerical Simulation* 15.4 (2010), pp. 1092–1102. ISSN: 10075704. DOI: 10.1016/j.cnsns.2009.05.044. URL: http://dx.doi.org/10.1016/j.cnsns.2009.05.044.

[6] Serge Tabachnikov. "Geometry and Billiards". In: *Göttingen Summer School on Dynamics* (2011).