

Using Integrability to Produce Chaos: Billiards with Positive Entropy

Victor J. Donnay*

Department of Mathematics, Bryn Mawr College, Bryn Mawr, PA 19010, USA

Received February 1, 1990; in revised form April 15, 1991

Abstract. A new open class of convex 2 dimensional planar billiards with positive Lyapunov exponent almost everywhere is constructed. We introduce the notion of a focusing arc and show that such arcs can be used to build billiard systems with positive Lyapunov exponents. We prove that under small C^6 perturbations, focusing arcs remain focusing and thereby show that perturbations of the Bunimovich stadium billiard have positive Lyapunov exponents.

0. Introduction

We study the ergodic properties of billiards inside a planar domain Q for which the boundary ∂Q consists of piecewise smooth arcs that are either flat or convex. A billiard in Q is the dynamical system arising from the uniform motion of a point mass inside Q with elastic reflections at the boundary: angle of reflection equals angle of incidence. We introduce a very general class of convex arcs for which the resulting billiard will have positive Lyapunov exponents almost everywhere and hence by Pesin [P] will have positive measure-theoretic entropy (chaos). This class is open in the C^6 topology on curves. Using these results, we prove that C^6 small perturbations of the Bunimovich stadium billiard have positive Lyapunov exponents. Our examples generalize work of Bunimovich [B3], Wojtkowski [W3], and Markarian [Mr1, Mr2] who had described various non-open classes of arcs.

We say a non-closed curve is convex if when we connect the ends of the curve by a straight line, the resulting closed curve has no self-intersections and bounds a convex set. Henceforth, the term convex curve will signify a C^∞ smooth curve that is non-closed and whose curvature is never zero, i.e. the curve is strictly convex.

Let γ be a convex curve. We examine a ray that collides with γ a finite number of times.

* Partially supported by NSF grant DMS 8806067

Definition 1. An incoming ray is focused by γ if it collides a finite number of times with γ and if the infinitesimal family of rays that is parallel to the incoming ray

- (1) focuses (i.e. passes through a focal point) between each pair of collisions with γ ,
- (2) focuses after hitting γ for the last time.

For such a ray, the time the parallel family takes to focus after hitting γ for the last time is called its focusing time.

(See Fig. 1).

Our principal result is

Theorem 1. For any convex curve γ there exists an angle $\Theta(\gamma) > 0$ such that any ray that hits γ with angle less than Θ will be focused, and the focusing times for these rays are bounded.

We want boundary curves which do not trap rays: i.e. every (or almost every) ray starting on the boundary will have a finite number of collisions with the curve and then leave it. We make the following restriction:

Definition 2. A convex arc γ is a convex curve whose curvature $k_\gamma(s)$, $s \in [0, l]$ satisfies

$$\int_0^l k_\gamma(s) ds \leq \pi, \tag{1}$$

where the arc-length of γ is l .

Definition 3. A convex arc γ is focusing if all rays collide with γ only a finite number of times and all rays are focused.

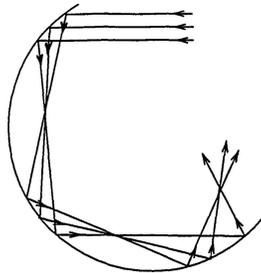


Fig. 1. Parallel variation focusing

Note that any sub-arc of a focusing arc is itself a focusing arc. In the special case that the straight line connecting the endpoints $\{\gamma(0), \gamma(l)\}$ of the arc is a periodic billiard trajectory, we modify this definition slightly (see Definition 1.2).

Using that the focusing times for trajectories that hit a focusing arc with small angles are bounded, we prove that

Theorem 2. Any convex region Q whose boundary consists of focusing arcs connected by straight lines that are sufficiently long will produce a billiard that has positive Lyapunov exponents almost everywhere.

Marlies Gerber [G] pointed out to us the following corollary of Theorem 1.

Theorem 3. Let γ be any convex curve and q any point on γ . Then there exists a neighborhood $U(q) \subset \gamma$ of q such that the convex arc $U(q)$ is focusing.

The set of focusing arcs form an open set in the space of smooth, convex curves satisfying (1).

Theorem 4. *Let γ be a focusing arc. If $\tilde{\gamma}$ is a convex arc of the same length that is sufficiently close to γ in the C^0 topology, then $\tilde{\gamma}$ is also a focusing arc.*

If γ is chosen to be a half circle, then Theorem 4 implies that perturbations of the stadium billiard have positive Lyapunov exponents, a result we announced in [D4]. Fixing the length of the flat sides determines how close to a half circle $\tilde{\gamma}$ must be. Note that the resulting boundary will be, at best, only C^1 smooth. This result on perturbations of the stadium was first claimed by Bunimovich [B2]. His argument was of a heuristic nature, and it was not clear what type of perturbations he allowed.

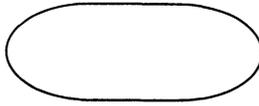


Fig. 2. Elliptical stadium

By Theorem 4, one can replace the half-circles of the stadium by half-ellipses cut along their semi-minor axis and with eccentricity close to 1. We show that a half-ellipse with eccentricity significantly different from one is also a focusing arc (Sect. 7). This allows us to construct what we call an elliptical stadium (two half-ellipses joined by straight lines, see Fig. 2) for which the billiard will have positive Lyapunov exponent almost everywhere. We do not have an explicit estimate on how long the flat sides of the stadium must be.

Theorem 5. *The half-ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, $x \geq 0$, is a focusing arc if and only if $a/b < \sqrt{2}$.*

From the work of Hopf [H], one knew that the divergence of nearby trajectories, such as happens in the geodesic flow on surface of negative curvature, produces ergodic behaviour. First Krylov [Kr] and then, in a rigorous way, Sinai [S1] adapted the ideas of Hopf to show that a billiard with concave boundary (the Sinai scatterer) displays ergodic behaviour. The underlying reason is that a divergent family of rays will be made to diverge even more strongly by a collision with a concave boundary.

Thus one was very surprised by the ergodicity of Bunimovich's stadium example (two half-circles joined by straight lines) [B1, B4] since a collision with a strictly convex boundary component causes a diverging family of rays not to diverge but rather to converge and eventually to focus. Furthermore, for billiards inside a smooth, strictly convex domain, Lazutkin [L] had shown that there always exist caustics, and hence such a system could not be ergodic. A caustic is a closed curve to which a trajectory is repeatedly tangent. The ergodicity of the stadium does not contradict the Lazutkin result since the stadium is neither smooth nor strictly convex.

A heuristic explanation of the chaotic implications of focusing is that after being focused the rays will then diverge. If the time during which the rays diverges is longer than the time during which they converge, there will be a net increase in the divergence. This reasoning is not strictly correct as the cardioid example of Wojtkowski [W3] and our Theorem 3 show, but it serves as a useful heuristic.

Guided by this train of thought, the author [D1, D2] was able to construct a smooth metric on the two-sphere with ergodic geodesic flow.

The idea of separating convex arcs by “sufficiently long” straight segments was advanced by Wojtkowski [W3] and helped clarify the mechanism by which focusing leads to chaotic dynamics. Bunimovich [B3] formulated versions of Definition 1 and Theorem 2 but, because he did not have Theorem 1, his versions required stronger assumptions (see Remark 4.6).

The billiard systems for which we and other authors have proven positive Lyapunov exponents almost everywhere have boundaries that are at most C^1 smooth. For boundaries that are smoother than C^1 , no such results are known. The problem is that if the curvature of the boundary goes smoothly to zero, then the time it takes a family of rays to focus becomes unbounded; present techniques cannot handle this situation. Recently the author [D3] has found a smooth strictly convex billiard that exhibits a weaker form of chaotic behaviour: positive topological entropy.

In this paper, we do not address the question of ergodicity. One should be able to adapt the proofs of ergodicity for systems with singularities [B4, KSS] to our examples. The main technical issue is the behaviour of trajectories that hit the boundary with small angles which can be analyzed with the methods of Sect. 5.

To prove our results, we use Wojtkowski’s version [W2] of the cone-field method Sect. 1. The return map for a billiard inside a convex domain is a twist map (Sect. 1). For integrable and near integrable twist maps, one can use the existence of invariant curves to canonically define a cone-field. We illustrate these ideas by presenting a proof that the stadium billiard has positive Lyapunov exponents (Sect. 2). By viewing the stadium as a linked twist map involving a (non-smooth) perturbation of a circle and using the integrable nature of billiards inside a circle, we are able to give a “one line” proof that does not require any calculations. We use more quantitative information about focusing gained from analyzing the Riccati equation (Sect. 3) to prove Theorem 2 (Sect. 4).

In proving the existence of caustics, which in the phase space corresponds to invariant curves, Lazutkin introduced an especially useful set of coordinates. In these coordinates, one sees that for trajectories that hit the boundary at small angles, the billiard map is nearly integrable. Our intuitive ideas about the relationship between cone-fields and invariant curves for twist maps led us to examine the Lazutkin coordinates which we use (Sect. 5) to prove Theorems 1 and 3. In (Sect. 6) we combine our small angle results with fairly standard perturbation arguments to prove Theorem 4. We use that for a focusing arc, a composition of the billiard map remains a twist map (Proposition 3.6). The results on elliptical arcs are proven in (Sect. 7) and can be read independently of the other sections.

As we were writing up our results, we learned that Markarian [Mr2] had recently proven similar results using quadratic forms. He shows that any C^4 perturbation of a half-circle is still a focusing arc and he proves Theorem 3. In Remark 5.8, we comment on the C^4 versus C^6 smoothness requirement.

1. Background Material

We outline the basic definitions of billiards [W3, C-F-S].

Let Q be a connected domain in the plane with piecewise smooth boundary. By a billiard in Q we mean the dynamical system arising from the uniform motion of a

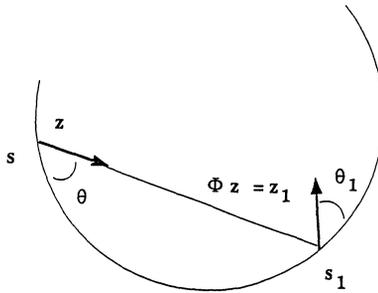


Fig. 3. Billiard map

point mass inside Q with elastic reflections at the boundary: the angle of reflection equals the angle of incidence. This motion produces a flow Ψ^t in the space T_1Q of unit tangent vectors of Q with the obvious identifications. Let $\pi_1 : T_1Q \rightarrow Q$ denote the natural projection.

Let $\Phi : M \rightarrow M$ be the standard section map, where $M \subset T_1Q$ is the two dimensional manifold consisting of unit vectors attached at the boundary ∂Q and pointing inside Q . For $z \in M$, the point Φz is gotten by following the point z under the billiard flow Ψ^t until its next collision with the boundary.

On M , we introduce coordinates (s, θ) , where s is the arc length parameter along ∂Q and $\theta \in [0, \pi]$ is the angle which the unit vector makes with ∂Q (Fig. 3).

We denote by $\{X_s, X_\theta\}$ the associated tangent vectors. Φ preserves the measure $\mu = c \sin \theta ds d\theta$, where c is a normalizing constant.

The points in ∂Q at which the piecewise smooth boundary components meet will lead to singularities for Φ . The singularity set of Φ has μ measure zero and the differential $D\Phi$ satisfies [K-S]

$$\int_M \log^+ \|D\Phi(z)\| d\mu(z) < \infty, \quad \text{where } \log^+ x = \max\{\log x, 0\}. \quad (1.1)$$

We examine the case that Q is a convex domain whose boundary components consist of C^∞ smooth, strictly convex curves and flat pieces (i.e. straight lines). The cases where one has concave components or a combination of concave, convex and flat components are discussed in [W3].

To decide whether a given curve γ is a focusing arc and hence can be used to build up a boundary ∂Q whose billiard will have positive Lyapunov exponents, we analyze the billiard restricted to the curve γ . Given a convex curve γ , we denote the closed convex domain gotten by connecting the endpoints of γ with a line segment \mathcal{L} by Q_γ , the induced map on Q_γ by Φ_γ , and the associated phase space by M_γ . We denote the subset of M_γ that consists of points with basepoint in γ by

$$S_\gamma = \{z = (s, \theta) \in M_\gamma : \pi_1(z) \in \gamma\}. \quad (1.2)$$

We chose the arc-length coordinate so that if $z \in S_\gamma$ and if γ has length l_γ , then $s \in [0, l_\gamma]$.

We say a point $z \in S_\gamma$ is entering (leaving) γ if $\pi_1(\Phi^{-1}z) \in \mathcal{L}$ ($\pi_1(\Phi z) \in \mathcal{L}$). We denote by S_γ^- (S_γ^+) the set of points entering (leaving) γ :

$$\begin{aligned} S_\gamma^- &= \{z \in S_\gamma : \pi_1(\Phi^{-1}z) \in \mathcal{L}\}, \\ S_\gamma^+ &= \{z \in S_\gamma : \pi_1(\Phi z) \in \mathcal{L}\}. \end{aligned} \quad (1.3)$$

We denote by $r_\gamma(s)$ and $k_\gamma(s) = 1/r_\gamma(s)$ the radius of curvature and the curvature of γ . When it is clear to which curve we are referring, we omit the subscript γ in our notation.

In certain cases, such as a half-circle, the points

$$p_0 = (s=0, \theta = \pi/2), \quad p_l = (s=l, \theta = \pi/2) \tag{1.4}$$

form a periodic orbit and hence never leave γ .

Lemma 1.1. *Let γ be a convex arc with*

$$\int_0^l k_\gamma(s) ds \leq \pi. \tag{1.5}$$

Then all points in S_γ , except $\{p_0, p_l\}$ if they are periodic, will leave γ .

Proof. We orient γ in the XY plane so that $\gamma(0)$ is at the origin and $\gamma'(0)$ points along the negative Y axis.

We examine how the Y component of the direction vector of a trajectory changes due to a collision with the boundary. If the trajectory collides at a point where the normal vector to γ (oriented so as to point into Q_γ) has a positive Y component, then the Y component of the direction vector will increase after the collision. Assumption (1.5) implies that the normal vector to γ has a positive Y component everywhere except at $s=0$ and possibly at $s=l$. After successive collisions, the Y component of the direction vector increases monotonically. Since there are no periodic orbits inside γ , trajectories cannot accumulate and hence they must all leave γ . \square

For $z \in S_\gamma$, we denote by $n_\gamma(z)$ the number of hits z has with the boundary before it leaves γ , i.e. $\Phi^{n(z)}z \in S_\gamma^+$. If $z \in S_\gamma^+$ then $n(z)=0$.

We denote by $s_k(s, \theta)$ and $\theta_k(s, \theta)$ the coordinates of $\Phi^k(s, \theta)$. We use $\tilde{}$ to denote quantities associated to a curve $\tilde{\gamma}$.

For the special case that $\{p_0, p_l\}$ form a periodic orbit, we define focusing as follows:

Definition 1.2. Let $\{p_0, p_l\}$ form a periodic orbit for a convex arc γ . The ray which first hits γ at p_α , $\alpha=0, l$ is focused by γ if

- (1) the incoming parallel variation focuses between p_α and Φp_α and between Φp_α and $\Phi^2 p_\alpha$
- (2) this variation is converging when it leaves $\Phi^2 p_\alpha$.

Hence we consider that $n(p_\alpha)=2$. We give a quantitative definition of converging in Definition 3.1. Condition (2) implies that given enough time, the parallel variation would focus after leaving $\Phi^2 p_\alpha$, although the time required might be more than the time to go between $\Phi^2 p_\alpha$ and $\Phi^3 p_\alpha$.

This definition of focusing was chosen to insure that Theorem 4 held (see Appendix A2, case (III)). Theorem 2 holds under a weaker condition ($n(p_\alpha)=1$):

- (1') the incoming parallel variation focuses between p_α and Φp_α .
- (2') this variation is covering when it leaves Φp_α .

The reason that in Theorem 4 we only allow perturbations that satisfy condition (1.5) is to insure that almost all trajectories will leave the perturbed arc. For example, if we started with the half-circle, an arbitrarily small perturbation could result in $\tilde{\gamma}$ being a segment of an ellipse. The billiard motion in an ellipse is integrable [S2] and the semi-minor axis gives a stable periodic orbit. If we had $\int_0^l k_{\tilde{\gamma}}(s) ds > \pi$, then a neighborhood of the stable periodic orbit could be included in

$S_{\tilde{\gamma}}$. Points in this neighborhood would never leave $\tilde{\gamma}$ and they would all have zero Lyapunov exponent. Thus without condition (1.5), Theorem 4 would be false.

Given a convex curve γ , we can scale it so that it will have length l . Scalings do not effect the focusing behaviour of a curve. We define the C^n topology on the space of all smooth curves γ of length l as follows. We let $\gamma: [0, l] \rightarrow R^2$ define the curve as a function of arc-length with the normalization that $\gamma(0) = (0, 0)$ and $\gamma'(0) = (0, -1)$. Then

$$\|\gamma - \tilde{\gamma}\|_{C^n} = \max_{0 \leq k \leq n, s \in [0, l]} |\gamma^{(k)}(s) - \tilde{\gamma}^{(k)}(s)|. \tag{1.7}$$

Lemma 1.3. *The map $\Phi_\gamma(D\Phi_\gamma)$ depends continuously on γ in the $C^2(C^3)$ topology.*

When examining the billiard map on a boundary ∂Q , we will use an associated first return map \mathcal{T} . If ∂Q contains smooth strictly convex arcs $\gamma_i, i = 1, 2, \dots, N$, then we set

$$S = \bigcup_{i=1}^N S_{\gamma_i}, \tag{1.8}$$

and define $\mathcal{T}: S \rightarrow S$ to be the first return map to the convex part of ∂Q .

Lyapunov Exponents and Cone-Fields. For our dynamical system (Φ, M, μ) , the Lyapunov exponents λ_+, λ_- are given by

$$\begin{aligned} \lambda_+(z) &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln \|D\Phi^n(z)\|, \\ \lambda_-(z) &= -\lambda_+(z). \end{aligned} \tag{1.9}$$

Since condition (1.1) holds, we can apply Oseledec's Multiplicative Ergodic Theorem [O] and conclude that the limit sup is actually a limit for μ almost every $z \in M$.

Pesin theory [P], generalized to systems with singularities [K-S], states that the measure theoretic entropy h_μ is given by $h_\mu = \int_M \lambda_+(z) d\mu(z)$.

Wojtkowski's criterion for showing that λ_+ is positive involves cone-fields. A cone \mathcal{C} in a two dimensional vector space V is a subset $\mathcal{C} = \{aX_1 + bX_2 : ab \geq 0\}$, where X_1 and X_2 are linearly independent vectors and $a, b \in R$. We call Interior $(\mathcal{C}) = \{aX_1 + bX_2 : ab > 0 \text{ or } a = b = 0\}$ the interior of \mathcal{C} . A measurable cone-field for our space M is a family of cones $\{\mathcal{C}(z)\} \subset T_z M$ defined for μ almost every $z \in M$ such that the vectors $X_1(z), X_2(z)$ vary measurably with z .

Theorem 1.4 [W2]. *Let $\mathcal{C}(z)$ be a measurable cone-field such that for almost every z*

$$D\Phi(\mathcal{C}(z)) \subset \mathcal{C}(\Phi(z)), \tag{1.10}$$

and for almost every z there exists $k(z)$ for which

$$D\Phi^{k(z)}(\mathcal{C}(z)) \subset \text{Interior } \mathcal{C}(\Phi^{k(z)}z). \tag{1.11}$$

Then the Lyapunov exponents $\lambda_+(z)$ are positive for almost every $z \in M$.

Geometry of Tangent Vectors. We find it useful to think of a tangent vector $\xi = s'X_s + \theta'X_\theta \in T_z M$ geometrically. We associate to ξ the infinitesimal one-parameter family of points that generates ξ : i.e. the variation $\eta_\xi(\sigma) = (s(\sigma), \theta(\sigma)) = (s + \sigma s', \theta + \sigma \theta'), \sigma \in (-\epsilon, \epsilon)$ for which $\eta_\xi(0) = z, \eta'_\xi(0) = \xi$.

Definition 1.5. We say the vector ξ focuses if the associated family of rays $\eta_\xi(\sigma)$ focuses in linear approximation.

We denote by $X_p^-(z) \in T_z M$ the infinitesimal family that under the flow Ψ^t is parallel when it reaches z .

Wojtkowski's results [W3] are based on a simple geometric argument, involving the focusing distance of families of rays, that allows one to compare the slopes of tangent vectors in the $\{X_s, X_\theta\}$ plane.

Let $z_0, z_1 = \Phi z_0 \in M$ have basepoints on convex boundary components. For $\xi_0 \in T_{z_0} M, \bar{\xi} = s' X_s + \bar{\theta}' X_\theta \in T_{z_1} M$, we wish to compare the slope of the vector $\xi = D\Phi \xi_0 = s' X_s + \theta' X_\theta$ with the slope of $\bar{\xi}$.

Denote by L the length of the line segment connecting $\pi_1(z_0)$ and $\pi_1(z_1)$. Suppose that under the flow Ψ^t the vector ξ_0 focuses at a point on this line segment a distance $d_0(\xi_0)$ from $\pi_1(z_0)$. Also suppose that under the backwards flow Ψ^{-t} the vector $\bar{\xi}$ focuses a distance $d_1(\bar{\xi}) > 0$ from $\pi_1(z_1)$.

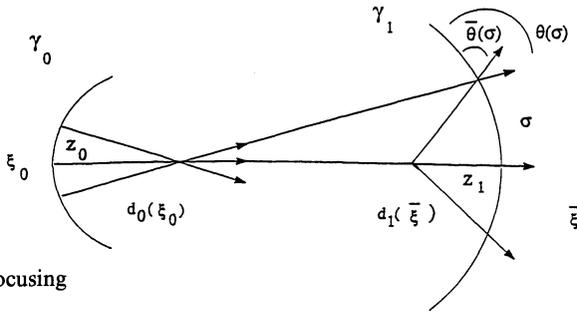


Fig. 4. Order of focusing

Focusing Lemma 1.6. *If*

$$L - d_0(\xi_0) - d_1(\bar{\xi}) > 0, \tag{1.12}$$

then

$$\frac{\theta'}{s'} > \frac{\bar{\theta}'}{\bar{s}'}. \tag{1.13}$$

Proof. Let

$$\eta(\sigma) = \left(s_1 + \sigma, \theta(\sigma) = \theta_1 + \frac{\theta'}{s'} \sigma \right),$$

$$\bar{\eta}(\sigma) = \left(s_1 + \sigma, \bar{\theta}(\sigma) = \theta_1 + \frac{\bar{\theta}'}{\bar{s}'} \sigma \right)$$

be the families corresponding to ξ and $\bar{\xi}$.

We interpret (1.12) to mean that under Ψ^{-t} , the family $\eta(\sigma)$ focuses further from $\pi_1(z_1)$ than does the family $\bar{\eta}(\sigma)$. Therefore, for a fixed value of σ , one has that

$$\theta(\sigma) > \bar{\theta}(\sigma),$$

which proves (1.13) (Fig. 4). \square

We denote the slope of a vector ξ by

$$m(\xi) = \frac{\theta'}{s'}.$$

Twist Maps. The billiard map Φ for a convex domain is a twist map. If we take the vertical family $\eta(\sigma) = (s, \theta + \sigma)$, then the image $\Phi\eta(\sigma) = (s_1(\sigma), \theta_1(\sigma))$ satisfies $\frac{ds_1}{d\sigma} > 0$. This is the defining characteristic of a twist map [Mt1].

2. The Stadium

We give a geometric proof that the stadium billiard has positive Lyapunov exponents almost everywhere by considering the stadium as a perturbation of the circle. This proof is essentially contained in [W3], the difference being that we describe the cones in phase space while Wojtkowski described them in configuration space.

For billiards inside a circle of radius r , the billiard map $\Phi(s, \theta) = (s_1, \theta_1)$ is given by

$$s_1 = s + 2r\theta, \quad \theta_1 = \theta. \tag{2.1}$$

This system is integrable: the angle θ stays constant along an orbit. The phase space decomposes into a union of invariant circles given by $\theta = \text{const}$.

We define a cone-field $\{\mathcal{C}(z)\}$, $z \in M$, by

$$\mathcal{C}(z) = \{aX_s + bX_\theta : ab \geq 0\}. \tag{2.2}$$

Under the map (2.1), this cone-field is invariant but not strictly invariant [i.e. it satisfies (1.10) but not (1.11)]:

$$D\Phi_z(X_\theta) \subset \text{Interior}(\mathcal{C}(\Phi z)), \tag{2.3a}$$

$$D\Phi_z(X_s) = X_s \in \partial(\mathcal{C}(\Phi z)). \tag{2.3b}$$

For any convex billiard, the map Φ is a twist map and hence for such a billiard (2.3a) will always be true. We interpret (2.3b) as saying that the horizontal vector X_s stays horizontal. In terms of the invariant circles, the tangent to the invariant circle gets sent to the tangent to the invariant circle.

Reverting to our geometrical interpretation, let $\eta(\sigma) = (s + \sigma, \theta)$, $\sigma \in (-\varepsilon, \varepsilon)$ be the family that generates X_s . Under Φ this family gets sent to

$$\Phi\eta(\sigma) = (s + 2r\theta + \sigma, \theta), \tag{2.4}$$

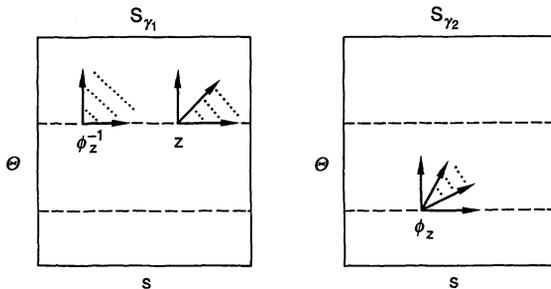


Fig. 5. Strict invariance of cones

which corresponds to the horizontal vector at (s_1, θ_1) . The trajectory connecting (s, θ) and (s_1, θ_1) has length

$$L = 2r \sin \theta. \tag{2.5}$$

By symmetry, we deduce that the family $\eta(\sigma)$ must focus at the midpoint of this trajectory; hence at a distance

$$d = r \sin \theta \tag{2.6}$$

from (s, θ) .

To produce positive Lyapunov exponents, we must find a way to push up the horizontal edge of the cone. We achieve this by perturbing the circle in a non-smooth way to produce the stadium: we cut the circle in half, pull the two halves apart and connect them by straight lines. When a trajectory goes from one half-circle to the other, the horizontal edge of the cone will get pushed up, so that the cone-field becomes strictly invariant (Fig. 5). Here we see the linked twist map phenomena [B-E, W1].

We now study the induced map $\mathcal{T} : S \rightarrow S$ defined in (1.8), where the strictly convex curves $\gamma_i, i = 1, 2$, are the half-circles of the stadium.

For $z \in S$, we define the cone-field by (2.2).

Proposition 2.1. *Under the map $\mathcal{T} : S \rightarrow S$, the cone-field defined by (2.2) is almost everywhere eventually strictly invariant (i.e. it satisfies (1.10), (1.11)).*

Thus the discussion of (Sect. 1) immediately implies

Corollary 2.2. *The Lyapunov exponents for Φ are non-zero almost everywhere.*

Proof of Proposition. If both x and $\mathcal{T}x$ are in the same component of S , then \mathcal{T} is given by (2.1) and the cone-field is still invariant but not strictly invariant.

If $x \in S_i$ and $\mathcal{T}x \in S_j, j \neq i$, then we must show

$$D\mathcal{T}(X_s) \subset \text{Interior}(\mathcal{C}(\mathcal{T}x)). \tag{2.7}$$

We first examine the case that $z = (s, \theta) \in S_i$ and $z_1 = \Phi z \in S_j, j \neq i$. By (2.6), the vector $X_s \in T_z S$ ($X_s \in T_{z_1} S$) will focus at a forward (backward) distance $d = r \sin \theta$ ($d_1 = r \sin \theta_1$). Let L be the length of the trajectory connecting z and z_1 .

For the stadium, we claim that

$$L > d + d_1. \tag{2.8}$$

Note that $2d(2d_1)$ is the length of that part of the trajectory that lies inside the osculating circle at $\pi_1(z)$ ($\pi_1(z_1)$). By pulling the two half-circles apart, we insure that both osculating circles lie completely inside the billiard table. Hence $d + d_1 \leq \max\{2d_0, 2d_1\} < L$.

Given (2.8), (2.7) now follows from the Focusing Lemma 1.6 and the twist property of Φ .

For those $z \in S$ for which $\Phi z \notin S$, i.e. the ray collides with the flat boundary piece, we again use the Focusing Lemma. Notice that after a family of rays collides with a flat boundary piece, its behaviour is the same as if it went straight through the flat piece with no collision. Thus the behaviour of such a family of rays can be understood by looking at their motion in an expanded table, gotten by unfolding the original table across the flat sides. In this setting one sees that (2.8) clearly holds, and hence (2.7). \square

The heuristic lesson we have learned from the stadium example is

Remark 2.3. For points that lie on an invariant curve of a twist map, there exists a canonical cone field that is invariant but not strictly invariant. One edge of the

cone is given by the vertical vector, the other edge is given by the tangent to the invariant curve.

In [W3], Wojtkowski produced a large class of convex billiards that had non-zero Lyapunov exponent almost everywhere. For each point $z \in M$ with base point $\pi_1(z)$ in a convex boundary component, the cone defined by Wojtkowski, when translated into our language, is given by (2.2).

The convex curve γ need no longer be an arc of a circle. If at a point $z = (s, \theta)$ the radius of curvature of γ is $r(s)$, then the horizontal family at z will focus at a distance

$$d = r(s) \sin \theta. \tag{2.9}$$

This result holds because to first order, the curve $\gamma(s)$ is approximated in a neighborhood of the point $\gamma(s)$ by the osculating circle.

For points (s, θ) and (s_1, θ_1) both sitting in convex boundary components, the Focusing Lemma shows that (2.7) holds providing that

$$L - d - d_1 > 0, \tag{2.10}$$

where now d and d_1 are given by (2.9). When equality holds in (2.10), then the horizontal vector stays horizontal. This relation is the meaning of the statement that two convex components must be sufficiently far apart.

When both points are in the same convex component γ_b , Wojtkowski showed that the condition $\frac{d^2 r(s)}{ds^2} < 0$, for all $s \in \gamma_b$, implies (2.10) (the ‘‘convex scattering’’ condition).

3. Jacobi Fields and Riccati Solutions

We define a local coordinate system $\{v, v^\perp, \phi\}$ on T_1Q which will give us more quantitative information on the evolution of variations. The following results are well known but are usually phrased in the language of wave fronts [S1, B1, W3].

Let $q \in Q$ be a point inside the domain and \mathbf{v} a unit vector at q . We define v to be the distance from q in the \mathbf{v} direction, v^\perp the distance measured in the direction perpendicular to \mathbf{v} , and ϕ the angle measured from \mathbf{v} .

We associate to a vector $\xi = J(0)X_{v^\perp} + J'(0)X_\phi$ in the perpendicular subspace of $T_{(q, \mathbf{v})}(T_1Q)$, i.e. in the span $\{X_{v^\perp}, X_\phi\}$, the variation $\eta^\perp(\sigma) = (v(\sigma) \equiv 0, v^\perp(\sigma), \phi(\sigma))$ which satisfies $(v^\perp)'(0) = J(0), \phi'(0) = J'(0)$. Such a variation is an example of a Jacobi field [K1], hence the choice of notation. Under the flow the vector ξ evolves to

$$D\Psi^t \xi = \frac{d}{d\sigma} \Big|_{\sigma=0} \Psi^t \eta^\perp(\sigma) = J(t)X_{v^\perp} + J'(t)X_\phi, \text{ where}$$

$$J(t) = J(0) + J'(0)t, \quad J'(t) = J'(0). \tag{3.1}$$

Definition 3.1. A vector ξ **focuses** when its associated Jacobi field satisfies $J(t) = 0$. We say the vector is **converging** when $J'(t)/J(t) \leq 0$ and is **diverging** when $J'(t)/J(t) \geq 0$.

Suppose that $\Psi^t(q, v)$ hits the boundary at a point $z = (s, \theta)$ at which the curvature is k . The family of rays $\eta^\perp(\sigma)$ will intersect the boundary, producing a family of rays $\eta(\sigma) \subset M$ with $\frac{d}{d\sigma} \Big|_{\sigma=0} \eta(\sigma) = s'X_s + \theta'X_\theta$. Let $(J_-, J'_-), (J_+, J'_+)$ denote

the values of the Jacobi field immediately before and immediately after collision with the boundary. Then straightforward calculations give

Lemma 3.2.

- (1) $J_- = s' \sin \theta, \quad J'_- = -\theta' + s'k,$
- (2) $J_+ = -s' \sin \theta, \quad J'_+ = \theta' + s'k,$
- (3) $J_+ = -J_-, \quad J'_+ = \frac{2J_-k}{\sin \theta} - J'_-.$

If we define $u(t) = \frac{J'(t)}{J(t)}$ then u satisfies the equation

$$u'(t) = -u^2(t). \tag{3.2}$$

This is a special case of the Riccati equation of differential geometry: $u'(t) = -K(t) - u^2(t)$ where $K(t)$ is the Gaussian curvature along a geodesic. We identify $\pm \infty$ and set $u(t) = \infty$ when $J(t) = 0$. Solutions of (3.2) then live on a circle, and they evolve in a simple way [Mn].

If $u(0) < 0$, then u decreases to $-\infty$ in a time $\frac{1}{|u(0)|}$, at which point the corresponding variation has focused. In an additional time τ , the value of u decreases from $+\infty$ to $\frac{1}{\tau}$.

Using Lemma 3.2, we find that the values of u before and after a collision satisfy:

$$\begin{aligned} u^- &= \frac{1}{\sin \theta} \left(k - \frac{\theta'}{s'} \right), \\ u^+ &= \frac{-1}{\sin \theta} \left(k + \frac{\theta'}{s'} \right), \\ u^+ &= u^- - \frac{2k}{\sin \theta}. \end{aligned} \tag{3.3}$$

Combining (3.2) and (3.3), we can define $u(t)$ for all time. We call such a function a Riccati solution based at z and say it focuses when $u(t) = -\infty$. We denote by $u(t; \xi)$ the Riccati solution that at $t=0$ corresponds to the vector $\xi \in T_z(T_1Q)$.

Definition 3.3. Focusing Times. We denote by $\tau^+(z; \xi), \tau^-(z; \xi)$ the times at which the Riccati solution $u(t; \xi)$ focuses in forward and backward time,

$$u(\tau^+(z; \xi); \xi) = u(-\tau^-(z; \xi); \xi) = -\infty,$$

and by

$$\tau_k^+(z; \xi),$$

the k^{th} time at which the Riccati solutions focuses.

The property of uniqueness of solutions to differential equations implies

Ordering Property 3.4. Let u_1, u_2, u_3 be Riccati solutions based on the same trajectory. Suppose $u_2(0) \in (u_1(0), u_3(0))$, where we take the ordering coming from the circle that we get by identifying $\pm \infty$. Then $u_2(t) \in (u_1(t), u_3(t))$ for all t .

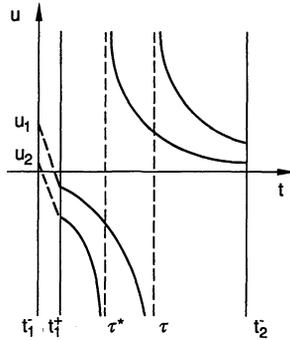


Fig. 6. Ordering of Riccati solutions

Corollary 3.5. Let $u_i, i=1, 2$ be Riccati solutions based at $z \in M$. Let z collide with the boundary at times $\{t_1, t_2\}$. Suppose that $u_1^-(t_1) > u_2^-(t_1) \geq 0$ and that there exists a $\tau \in (t_1, t_2]$ for which $u_1(\tau) = -\infty$. Then there exists $\tau^* \in (t_1, \tau)$ for which $u_2(\tau^*) = -\infty$.

(Fig. 6).

The map $\Phi^k z$ is a composition of twist maps. Such a map is called a tilt map and is typically not itself a twist map. If γ is a focusing arc however, then $\Phi_\gamma^k z$ will be a twist map.

Proposition 3.6. Let γ be a focusing arc. For $k=0, 1, \dots, n(z)$, suppose that $\Phi_\gamma^k z \in \text{Interior}(S_\gamma)$, and let $U(z)$ a neighborhood of z such that if $w \in U(z)$, then $\Phi_\gamma^k w \in \text{Interior}(S_\gamma)$. Then $\Phi_\gamma^k|_{U(z)}$ is a twist map for $k=0, 1, \dots, n(z)$: i.e., for $w=(s, \theta)$ and $s_k(s, \theta)$ the arc-length coordinate of $\Phi_\gamma^k(s, \theta)$, we have $\frac{ds_k}{d\theta} > 0$.

Proof. If z is focused by γ then we claim that

$$D\Phi_\gamma^k(X_\theta) = (s'_k, \theta'_k), \quad k=1, 2, \dots, n(z) \tag{3.4}$$

satisfies

$$s'_k > 0, \tag{3.5}$$

which implies the result.

Let u_p be the Riccati solution corresponding to the parallel variation X_p^- so that $u_p^-(0) = 0$ and $u_p^+(0) < 0$. Between the points $\Phi^k z = \Psi^{t_k} z$ and $\Phi^{k+1} z = \Phi^{t_{k+1}} z$, $k=0, 1, \dots, n(z)$, the variation must focus; i.e., there exists a $\tau_{k+1} \in (t_k, t_{k+1})$ such that $u_p(\tau_{k+1}) = -\infty$.

Let u_θ be the Riccati solution corresponding to X_θ , so that $u_\theta^+(0) = \pm \infty$, and hence $0 < u_\theta^-(t_1) < u_p^-(t_1)$. By Corollary 3.5, u_θ must also focus between (t_{k-1}, t_k) , $k=2, 3, \dots, n(z)$. If we look at the basepoints of the curve $\Phi_\eta(\sigma)$, where $\eta(\sigma)$ is the curve that generates X_θ , then we see immediately that $s'_1 > 0$. If $s'_k > 0$ and the variation u_θ focuses between t_k and t_{k+1} , then by looking at $\pi_1(\Phi^{k+1} \eta(\sigma))$, we see that $s'_{k+1} > 0$. \square

4. Cone of Focusing Variations

In [L], Lazutkin used K.A.M. theory to show that any smooth, strictly convex billiard has caustics and hence that the billiard flow cannot be ergodic. In phase

space, these caustics correspond to invariant circles. These invariant circles fill up a set of positive measure and they accumulate at $\theta=0$ and $\theta=\pi$.

We wish to use these invariant circles to control the cones (Remark 2.3) in each convex boundary component. Then, when a trajectory goes from one convex component to another, the cones which have been invariant will become strictly invariant. Actually, for a non-closed convex arc, it does not make sense to talk about an invariant circle for the billiard map: the trajectories leave the arc and hence are not recurrent. However, as with the half-circle in the stadium, one can image a “ghost” of the invariant circle existing: this is a subset of the invariant circle that would have existed if the arc were part of a closed convex curve. For our purposes, we do not really need an invariant curve; just the possibility that such a curve could exist.

Birkhoff [Bi] showed that an invariant curve for a twist map is given as the graph of a function. Using this fact, Mather [Mt1, Mt2] developed criteria for the non-existence of invariant curves which were used in [P-M, Hu]. We apply these results to our situation.

Let $X_\theta \in T_zM$ be the vertical vector at the point z and $D\Phi^k X_\theta = (s'_k, \theta'_k)$, $k = 1, 2, \dots$ its image. If $s'_k < 0$, then any curve through z will, under iteration by Φ , stop being the graph of a function, and hence could not be an invariant curve. In applying this criteria to the standard map, Percival and MacKay [M-P] defined a cone-field $\mathcal{K}(z) = \{(s', \theta') \in T_zM : s' \geq 0\}$, and showed that if after n iterations, one had

$$s'_k > 0, \quad k = 1, \dots, n, \tag{4.1}$$

then a tangent vector to an invariant curve through z would have to lie inside

$$\bigcap_{k=0}^n D\Phi^{-k}(\mathcal{K}(\Phi^k z)). \tag{4.2}$$

If this intersection becomes empty, then there could not exist an invariant curve through z . (Stark [St] developed a converse to this result. He showed that if $s'_k > 0$ for all $k > 0$, and some auxiliary conditions were also met, then there would exist an invariant curve through z .)

For a focusing arc γ , (4.1) holds as long as a z continues to collide with γ . Thus through every point $z \in S_\gamma$, there could exist an invariant curve: the intersection of cones in (4.2) is not empty. This non-empty intersection will allow us to construct the cones $\mathcal{C}(z)$ that we need to prove positive Lyapunov exponents.

We define a cone $\bar{\mathcal{C}}(z)$ that consists of all variations ξ that

$$\text{are diverging } (u^- \geq 0) \text{ when they reach } \Phi^k z, \quad k=0, 1, \dots, n(z). \tag{4.3a}$$

$$\text{are converging } (u^+ \leq 0) \text{ when they leave } \Phi^k z, \quad k=0, 1, \dots, n(z). \tag{4.3b}$$

Note that a variation satisfying (4.3) will have to focus between each pair of collisions with the boundary γ .

We set

$$\mathcal{K}^+(z) = \{(s', \theta') \in T_zS : -k(z) \leq \theta'/s' \leq +k(z)\}. \tag{4.4}$$

Then, by (3.3), the set of variations satisfying (4.3) is given by

$$\bar{\mathcal{C}}(z) = \bigcap_{k=0}^{n(z)} D\Phi^{-k}(\mathcal{K}^+(\Phi^k z)). \tag{4.5}$$

Proposition 4.1. *If z is focused by γ then the cone $\bar{\mathcal{C}}(z)$ defined by (4.5) has non-empty interior and is given by*

$$\bar{\mathcal{C}}(z) = \left\{ (s', \theta') \in T_z S : m(X_p^+) \leq \frac{s'}{\theta'} \leq +k(z) \right\}. \tag{4.6}$$

Proof. If z is focused (Definition 1) then the parallel variation $X_p^- \in T_z S$ satisfies $D\Phi_z^k(X_p^-) \in \text{Interior}(\mathcal{K}^+(\Phi^k z))$, $k = 1, 2, \dots, n(z)$. Hence there is an open neighborhood around X_p^- with this property. Since X_p^- forms one edge of the cone $\mathcal{K}^+(z)$, $\bar{\mathcal{C}}(z)$ has non-empty interior.

Let $X_p^+ \in T_z \mathcal{S}$ be the variation that is parallel when it leaves $\Phi^{n(z)} z$: $u^+(t_{n(z)}; X_p^+) = 0$. We claim that $X_p^+ \in \bar{\mathcal{C}}(z)$. This follows by noting that the incoming parallel variation at the point $z^* = (s(\Phi^{n(z)} z), \pi - \theta(\Phi^{n(z)} z))$ is gotten by time reversing the variation $D\Phi^{n(z)} X_p^+$, and that the point z^* is focused by γ .

The variation X_p^+ is the limiting case of variations that are converging after leaving $\Phi^{n(z)} z$ and $\tau_{n(z)}^+(z; X_p^+) = +\infty$. If the incoming parallel variation $X_p^- \in T_z \mathcal{S}$ is focused by γ , then $\tau_{n(z)}^+(z; X_p^-)$ is finite. The Ordering Property 3.4 and (3.3) imply that the slopes of the vectors satisfy $m(X_p^+) < m(X_p^-) = k(z)$. \square

For any $m_l > m(X_p^+)$ we choose a cone $C(z, m_l) \subset \bar{\mathcal{C}}(z)$ such that

$$\mathcal{C}(z, m_l) = \left\{ (s', \theta') \in T_z \mathcal{S} : m_l \leq \frac{\theta'}{s'} \leq k(z) \right\}. \tag{4.7}$$

We denote by ξ_l the variation satisfying $m(\xi_l) = m_l$.

Given such a cone $C(z, m_l)$, we define the forward focusing time of z with respect to C using Definition 3.3

$$\tau_C^+(z) = \sup_{\xi \in C(z, m_l)} \tau_{n(z)}^+(z, \xi),$$

and the backward focusing time of z with respect to C by

$$\tau_C^-(z) = \inf_{\xi \in C(z, m_l)} \tau^-(z, \xi). \tag{4.8}$$

The ordering property implies

Lemma 4.2. *The forward and backward focusing times of z with respect to the cone C are finite and are given by*

$$\begin{aligned} \tau_C^+(z) &= \tau_{n(z)}^+(z, \xi_l), \\ \tau_C^-(z) &= \tau^-(z, \xi_l). \end{aligned}$$

Once we have $\tau_C^\pm(z)$ finite, we can make the cones go strictly inside one another. Let γ_1 and γ_2 be two convex arcs. Let $z_1 \in \mathcal{S}_{\gamma_1}^-$ and $z_2 \in \mathcal{S}_{\gamma_2}^-$ be points focused by their respective arcs. We make a choice of cones $C_i(z_i, m_i(z_i)) \subset \bar{C}_i(z)$, $i = 1, 2$.

Proposition 4.3. *Let γ_1, γ_2 be pieces of a convex billiard table ∂Q . Suppose that after leaving γ_1, z_1 next returns to the strictly convex part of Q after a time L at the point z_2 , i.e.,*

$$\Psi^{t_{n(z_1)} + L} z_1 = z_2.$$

Then if

$$L > \tau_{C_1}^+(z_1) + \tau_{C_2}^-(z_2), \tag{4.9}$$

we have that

$$D\mathcal{F}^{n(z_1)+1}(C_1(z_1)) \subset \text{Interior}(C_2(z_2)). \tag{4.10}$$

Proof. Using (1.8) we show that for every $\xi \in C_1(z)$, $m(D\mathcal{F}^{n(z_1)+1}\xi) \in (m(\xi_i(z_2)), +k(z_2))$. Applying the Focusing Lemma 1.6 to (4.9) implies that $m(D\mathcal{F}^{n(z_1)+1}\xi) > m(\xi_i(z_2))$. On the other hand, since $X_p^- \in C_2(z_2)$ never focuses in backward time, (i.e., $\tau^-(z_2; X_p^-) = \infty$), then we have that $m(D\mathcal{F}^{n(z_1)+1}\xi) < k(z_2)$. \square

As we let the value of m_i decrease, we get a bigger cone $\mathcal{C}(z)$ and hence it is easier for variations to enter the cone ($\tau_{\mathcal{C}}^-(z)$ decreases). Making the cone bigger has a trade-off: it is now more difficult for all the vectors in $\mathcal{C}(z)$ to enter the next cone ($\tau_{\mathcal{C}}^+(z)$ increases). For a general focusing arc, it is not clear what the optimal choice of $\{m_i(z)\}$ would be.

To define a cone-field for all $w \in S_\gamma$, we push forward the cones defined for $z \in S_\gamma^-$:

$$\mathcal{C}(\Phi^k z) = D\Phi_z^k(\mathcal{C}(z)), \quad k = 1, 2, \dots, n(z). \tag{4.11}$$

Theorem 4.4. *Given a focusing arc γ , there exists a choice of cone-field $\{C(z)\}$, $z \in \mathcal{S}_\gamma^-$, defined as in (4.7, 4.11), such that the forward and backward focusing time of the arc*

$$\tau_{\gamma, C}^+ = \sup_{z \in \mathcal{S}_\gamma^-} \tau_C^+(z), \quad \tau_{\gamma, C}^- = \sup_{z \in \mathcal{S}_\gamma^-} \tau_C^-(z)$$

are both finite. These focusing times vary continuously with the curve γ .

We define the focusing time of γ with respect to a cone-field $\{C(z)\}$, by

$$\tau_{\gamma, C} = \max\{\tau_{\gamma, C}^+, \tau_{\gamma, C}^-\}.$$

Combining Proposition 4.3 and Theorem 4.4 gives

Corollary 4.5. *Let ∂Q be a convex table with strictly convex arcs $\{\gamma_i\}_{i=1}^N$ all of which are focusing joined by straight lines. The straight lines do not intersect one another. If there exists a choice of cone fields*

$$\{C_i(z, m_i(z)), z \in \mathcal{S}_{\gamma_i}^-, \quad i = 1, \dots, N$$

such that

$$\text{distance}(\gamma_i, \gamma_j) > \tau_{\gamma_i, C_i} + \tau_{\gamma_j, C_j}, \quad i \neq j, \tag{4.12}$$

then the billiard map has positive Lyapunov exponents almost everywhere.

Theorem 2 follows by making a choice of cones and then making the distance (γ_i, γ_j) large enough that (4.12) holds.

Proof of Corollary. Almost every point $z \in M$ will enter S infinitely often. If the Lyapunov exponents of \mathcal{F} are positive almost everywhere, the same is true for the exponents of Φ . Applying Theorem 1.4 to the map \mathcal{F} proves the result. \square

Proof of Theorem 4.4. Given $\varepsilon > 0$, Lemma 5.9 shows that there exists $\Theta > 0$ such that for any $z \in \mathcal{S}_\gamma^-$ with $\theta(z) \in \{(0, \Theta] \cup [\pi - \Theta, \pi)\}$, there exists a choice of cone-field $\{C(z, m_i(z))\}$ for which

$$\begin{aligned} \sup_z \tau^-(z, \xi_i(z)) &\leq \varepsilon, \\ \sup_z \tau_{n(z)}^+(z, \xi_i(z)) &\leq \varepsilon. \end{aligned}$$

Suppose $\{p_0, p_i\}$ are not periodic. Let z satisfy $\theta(z) \in [\Theta, \pi - \Theta]$ and make a choice of $m_l(z)$. For all neighboring points w , we will use the same bottom edge of the cone: i.e., $m_l(w) = m_l(z)$.

Since $\tau^-(z, \xi_l(z))$ and $\tau_k^+(z, \xi_l(z))$, $k = 1, 2, \dots, n(z)$ are finite, these functions vary continuously with the variable z . Thus for $\varepsilon > 0$ sufficiently small, there exists a $\delta(z) > 0$ such that for all $w \in \bar{B}(z, \delta(z)) = \{w \in S_\gamma : |w - z| < \delta(z)\}$,

$$\begin{aligned} \tau^-(w, \xi_l(z)) &\leq \tau^-(z, \xi_l(z)) + \varepsilon, \\ \tau_k^+(w, \xi_l(z)) &\leq \tau_k^+(z, \xi_l(z)) + \varepsilon, \quad k = 0, 1, 2, \dots, n(w), \end{aligned}$$

where $n(w) \leq n(z)$. Covering the compact set $\{z \in \mathcal{S}_\gamma^- : \theta(z) \in [\Theta, \pi - \Theta]\}$ by a finite union $\bigcup_{i=1}^N \bar{B}(z_i, \delta(z_i))$ gives the result.

If $\{p_0, p_i\}$ are periodic then p_α , $\alpha = 0, l$, do not leave γ . However, the twist condition $\frac{ds_2}{d\theta} > 0$ (Proposition 3.6) implies the existence of a neighborhood $\bar{B}(p_\alpha, \delta(p_\alpha))$ such that all $w \in \bar{B} \setminus p_\alpha$ will have $n(w) \leq 1$. For such w and $k = 0$, $n(w)$, we have that

$$\tau_k^+(w, \xi_l(p_\alpha)) \leq \tau_k^+(z, \xi_l(p_\alpha)) + \varepsilon,$$

which by Definition 1.2' is finite. \square

Remark 4.6. Bunimovich [B3, Eq. (4)] gave a condition which implies that an arc is focusing and that its focusing times $\tau_{\gamma, \varphi}^+$, $\tau_{\gamma, \varphi}^-$ are finite. Apriori, his condition is stronger than our focusing condition. For an arc could be focusing, and yet as $\theta \rightarrow 0$, the focusing times $\tau_{n(z)}^+(z, X_p^-)$ could become unbounded. Lemma 5.9 shows however that this is not the case: as $\theta \rightarrow 0$, the focusing times remain bounded and in fact go to zero.

5. Lazutkin Coordinates

We use the coordinates introduced by Lazutkin [L] to analyze the motion along a convex curve γ for θ near zero. For θ near π , the analysis is analogous. In (s, θ) coordinates, Lazutkin showed that the transformation

$$\Phi : (s, \theta) \mapsto (s_1, \theta_1)$$

can be written as

$$\begin{aligned} s_1 &= s + \alpha_1(s)\theta + \alpha_2(s)\theta^2 + \alpha_3(s)\theta^3 + F(s, \theta)\theta^4, \\ \theta_1 &= \theta + \beta_2(s)\theta^2 + \beta_3(s)\theta^3 + G(s, \theta)\theta^4. \end{aligned} \tag{5.1}$$

The functions α_i and β_i are given by

$$\begin{aligned} \alpha_1(s) &= 2r(s), & \alpha_2(s) &= \frac{4}{3}r(s)r'(s), \\ \alpha_3(s) &= \frac{2}{3}r^2(s)r''(s) + \frac{4}{9}r(s)r'^2(s), \\ \beta_2(s) &= -\frac{2}{3}r'(s), & \beta_3(s) &= -\frac{2}{3}r(s)r''(s) + \frac{4}{9}r'^2(s). \end{aligned} \tag{5.2}$$

Lazutkin then made the following change of coordinates:

$$\begin{aligned} x(s, \theta) &= C_1 \int_0^s r^{-2/3}(s) ds, \\ y(s, \theta) &= C_2 r^{1/3}(s) \sin(\theta/2), \end{aligned}$$

where

$$C_1 = \left(\int_0^l r^{-2/3}(s) ds \right)^{-1}, \quad l = \text{length of } \gamma, \tag{5.3}$$

$$C_2 = 4C_1.$$

In these new coordinates, the transformation Φ is given by

$$\begin{aligned} x_1 &= x + y + y^3 f(x, y), \\ y_1 &= y + y^4 g(x, y). \end{aligned} \tag{5.4}$$

The choice of C_1 normalizes the x variables so that $x \in [0, 1]$. C_2 is chosen to make the coefficient for y in the first line of (5.4) equal to one.

We develop some lemmas describing how Φ and $D\Phi$ behave under iteration (see Appendix A1 for the proofs).

Define

$$c_1 = \max f, \quad c_2 = \max g. \tag{5.5}$$

Lemma 5.1. *For y sufficiently small, we have that*

$$|y_k - y| < \frac{1}{2}y \tag{5.6}$$

for $k=1, 2, \dots, N$, where N grows like constant/ y^3 .

We estimate how many successive collisions a trajectory can have with a fixed convex arc γ .

Lemma 5.2. *There exists a constant c_3 such that for y sufficiently small, all points (x, y) with $x \in [0, 1]$ will have $x_n > 1$ for some $n \leq c_3/y$.*

Corollary 5.3. *Let γ be a convex curve. There exists $Y > 0$ and a constant c_3 such that all trajectories (x, y) that hit γ with $y \leq Y$ will move monotonically along γ and will have at most c_3/y collisions before leaving.*

Proof. For y sufficiently small, the condition $x_n > 1$ implies that $\Phi^n(x, y) \notin S_\gamma$. If y is large, it is possible that the trajectory would bounce to the other side of γ and hence not leave. \square

For Φ given by (5.4), the differential $D\Phi$ at the point (x_k, y_k) is given by the matrix

$$D\Phi = \begin{pmatrix} 1 + y_k^3 f_x & 1 + 3y_k^2 f + y_k^3 f_y \\ y_k^4 g_x & 1 + 4y_k^3 g + y_k^4 g_y \end{pmatrix}.$$

If we assume that $|y_k - y| < y/2 < 1/2$ then we can bound the entries of the matrix by

$$\begin{pmatrix} 1 \pm c_4 y^3 & 1 \pm c_4 y^2 \\ \pm c_4 y^4 & 1 \pm c_4 y^3 \end{pmatrix}, \tag{5.7}$$

where the constant c_4 satisfies

$$c_4 = \sup \left\{ \left(\frac{3}{2}\right)^3 \|f_x\|, \left(\frac{3}{2}\right)^4 \|g_x\|, 3\left(\frac{3}{2}\right)^2 c_1 + \left(\frac{3}{2}\right)^3 y_M \|f_y\|, 4\left(\frac{3}{2}\right)^3 c_2 + \left(\frac{3}{2}\right)^4 y_M \|g_y\| \right\} \tag{5.8}$$

and

$$y_M = \max \{y\}.$$

Remark 5.4. The functions f_x, f_y, g_x, g_y are smooth functions of (x, y) and have a limit as $y \rightarrow 0$. Hence c_4 is finite. The value of c_4 depends in a C^6 way upon the curve γ . The value of $g_x(x, 0) = \lim_{y \rightarrow 0} g_x(x, y)$ for example is a function of the radius of curvature and its derivatives up to the fourth order:

$$r^{(j)}(s(x)), \quad j=0, 1, \dots, 4.$$

If $n(x, y)$ is the number of iterations of Φ between entering and leaving γ , then we are interested in

$$D\Phi^k \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (x'_k, y'_k), \quad k=1, 2, \dots, n(x, y).$$

Lemma 5.5. For all (x, y) with y sufficiently small and for $k=1, 2, \dots, n(x, y)$,

$$\begin{aligned} |x'_k - 1| &\leq y^3 k c_4 + y^4 k^2 c_4 \\ |y'_k| &\leq y^4 k c_4 + k y^5, \end{aligned}$$

and therefore there exists a constant $c_5 > 0$ so that

$$\left| \frac{y'_k}{x'_k} \right| \leq c_5 y^3.$$

Theorem 5.6. 1. There exists a $Y(\gamma) > 0$ such that all $z=(x, y)$ with $y \leq Y$ will be focused by γ .

2. The value $Y(\gamma)$ depends C^6 on γ .

Setting $\Theta = 2 \sin^{-1} \left(\frac{Y}{C_2 r_m^{1/3}} \right)$, where $r_m = \min_{s \in [0, 1]} r(s)$, proves the first part of Theorem 1. The second part follows from Lemma 5.9.

Proof of Theorem 5.6. 1. We show that the variation X_x satisfies (4.3). Hence since $u^-(X_x) > u^-(X_p^-)$, the ordering property implies that X_p^- also satisfies (4.3), and hence z is focused by γ . In our definition of $C(z; m_i(z))$, we could then choose $m_i(z) = m(X_x(z))$.

We wish to show that $u^-(t_k, X_x) > 0$ and $u^+(t_k, X_x) < 0, k=0, 1, \dots, n(z)$ which in terms of (s', θ') variables becomes

$$-k(z_k) < m(D\Phi^k X_x) < k(z_k), \quad k=0, 1, \dots, n(z). \tag{5.9}$$

From Lemma 5.5 we have that

$$\frac{y'_k}{x'_k} \leq c_5 y^3, \quad k=0, 1, \dots, n(z).$$

To change from (x, y) to (s, θ) coordinates, we differentiate (5.3) to get

$$\begin{aligned} x' &= c_1 r^{-2/3}(s) s', \\ y' &= \frac{c_2}{3} r^{-2/3} r'(s) \sin(\theta/2) s' + \frac{c_2}{2} r^{1/3} \cos(\theta/2) \theta'. \end{aligned} \tag{5.10}$$

Inverting, we get

$$\frac{\theta'}{s'} = \frac{1}{2r(s) \cos(\theta/2)} \frac{y'}{x'} - \frac{2r'(s)}{3r(s)} \tan(\theta/2). \tag{5.11}$$

We can get uniform estimates for the maximum and minimum of $r(s)$, $r'(s)$, and hence there exist constants c_6, c_7 such that

$$\left| \frac{\theta'}{s'} \right| \leq \frac{c_6}{\cos(\theta/2)} \frac{y'}{x'} + c_7 \tan(\theta/2).$$

For y sufficiently small, we have $y_k \in [\frac{1}{2}y, \frac{3}{2}y]$, and since $\theta_k = 2 \sin^{-1} \left(\frac{y_k}{C_2 r^{1/3}(s_k)} \right)$, there exists a constant c_8 such that $\frac{1}{c_8} y \leq \frac{\theta_k}{2} \leq c_8 y$. Hence

$$\left| \frac{\theta'_k}{s'_k} \right| \leq \frac{c_6}{\cos\left(\frac{y}{c_8}\right)} \frac{y'_k}{x'_k} + c_7 \tan(c_8 y).$$

Using Lemma 5.5, we conclude that given $\varepsilon > 0$, there exists a $Y > 0$ such that if $y \leq Y$ then

$$\left| \frac{\theta'_k}{s'_k} \right| \leq \varepsilon \quad \text{for } k=0, 1, 2, \dots, n(z).$$

Choosing $\varepsilon < \min_{s \in [0, 1]} k(s)$ will prove (5.9).

2. The constants $\{c_i\}_{i=1}^8$ that were used in determining $Y(y)$ depended on at most the sixth derivative of γ . \square

Theorem 5.7 (Theorem 3). *Let γ be any convex curve and q any point on γ . Then there exists a neighborhood $U(q) \subset \gamma$ of q such that the convex arc $U(q)$ is focusing.*

Proof. We define $S_U, S_U^-,$ and S_U^+ as in (1.3). If $z \in S_U^- \cap S_U^+$, then z hits U only once and hence is trivially focused by U .

To handle the trajectories that hit U more than once, we define

$$A_U = \{w \in U : w = \Phi^k z \text{ for some } k \in [0, n(z)], \text{ where } z \in S_U^- \text{ with } y(z) \leq Y(y)\}, \tag{5.12}$$

where $Y(y)$ is chosen by Theorem 5.6. Such a z is focused by γ , i.e. $X_p^- \in T_z M$ is focused, and hence by the Ordering Property 3.4 we have that $X_p^- \in T_w M$ is also focused. Hence w is focused by U .

Let l be the length of the arc U_l . The main observation is that as l decreases, the set $S_{U_l}^- \cap S_{U_l}^+$ increases. More precisely, if $l_1 < l_2$ and $s \in U_{l_1} \cap U_{l_2}$, then

$$\{\theta : (s, \theta) \in S_{U_{l_2}}^- \cap S_{U_{l_2}}^+\} \subset \{\theta : (s, \theta) \in S_{U_{l_1}}^- \cap S_{U_{l_1}}^+\}.$$

As l decreases, the θ component of A_{U_l} stays fixed: for $s \in U_{l_1} \cap U_{l_2}$,

$$\{\theta : (s, \theta) \in A_{U_{l_1}}\} = \{\theta : (s, \theta) \in A_{U_{l_2}}\}.$$

Thus for l sufficiently small,

$$S_{U_l} = A_{U_l} \cup \{S_{U_l}^- \cap S_{U_l}^+\},$$

at which point all points in S_{U_l} are focused by U_l . \square

For U_l a focusing curve, one choice of cone-field is the following. For $z \in S_U^- \cap S_U^+$, we set $\mathcal{C}(z) = \{aX_s + bX_\theta : ab \geq 0\}$, and then $\tau_{\mathcal{C}}^+(z) = \tau_{\mathcal{C}}^-(z) = r(s) \sin \theta$. For

$w \in A_U$, we set $\mathcal{C}(w) = \{aX_h(w) + bX_\theta : ab \geq 0\}$, where the bottom edge of the cone X_h is gotten by pushing forward the vector $X_x \in TS_\gamma^+$,

$$X_h(w) = D\Phi^k(X_x).$$

Remark 5.8. Markarian uses the language of quadratic forms. Translating into our language, he shows [Mr1] that if an arc satisfies

$$\frac{d^2(r^{1/3})}{ds^2} > 0,$$

then any sufficiently short piece of the arc will be focusing. Our Theorem 5.7 generalizes this result. In [Mr2], he finds a change of coordinates whose tangent vectors give the edges of the cones. He constructs this change of variables in an ad-hoc fashion, and the remainder terms he gets depend only on values of $r^{(j)}$ for $j=0, 1, 2$. Hence he can prove his results for C^4 perturbations. Effectively, the Lazutkin coordinates (5.3) are too much of a good thing: by eliminating so many lower order terms, they force the remainder to depend on $r^{(4)}$. The $O(y^3)$ estimate of Lemma 5.5 is indicative of this overkill.

Lemma 5.9. *Let γ be a convex curve. Given $\varepsilon > 0$, there exists a $\Theta > 0$ such that for all $z \in S_\gamma^-$ with $\theta(z) \leq \Theta$ when we define $C(z, m_t(z))$ by $m_t(z) = m(X_x(z))$, then*

$$\max\{\tau_C^+(z), \tau_C^-(z)\} < \varepsilon.$$

The value of Θ varies continuously with the curve.

Proof. We need to examine $\tau_{n(z)}^+(z, X_x) = \tau^+(\Phi^{n(z)}z, D\Phi^{n(z)}X_x)$ and $\tau^-(z, X_x)$. By (3.3),

$$\begin{aligned} \tau_1^+(\omega, \xi) &= \frac{\sin \theta}{k + m(\xi)}, \quad \text{if } m(\xi) > -k, \\ \tau^-(\omega, \xi) &= \left| \frac{\sin \theta}{k - m(\xi)} \right|, \quad \text{if } m(\xi) < k. \end{aligned}$$

The values of $|m(D\Phi^k X_x)|$ and $\theta(\Phi^k z)$, $k=0, 1, \dots, n(z)$ can be made uniformly close to zero by suitable choice of Θ , which implies the result. \square

6. Large Angle Behaviour

Trajectories that hit the boundary with an angle larger than Θ will have only a bounded number of hits before leaving γ . Under a small perturbation of the boundary, these trajectories will still have a bounded number of hits with the new arc $\tilde{\gamma}$. Up to a bounded time N , we can make the difference $\|D\Phi^N(X_p^-) - D\tilde{\Phi}^N(X_p^-)\|$ uniformly small by making $\|\gamma - \tilde{\gamma}\|$ small. This implies that if trajectories focus for γ , they will focus for the perturbed arc $\tilde{\gamma}$.

The one technical issue is that we might not be able to compare $D\Phi^{n(z)}z$ and $D\tilde{\Phi}^{n(z)}z$ directly because z might have more hits with $\tilde{\gamma}$ than with γ : $\tilde{n}(z) > n(z)$. We use the Twist Property (Proposition 3.6) to circumvent this problem.

Let γ be a convex focusing arc with $\int_0^l k_\gamma(s) ds \leq \pi$. Chose $\Theta > 0$ small enough that all the results of Sect. 5 will hold. Let

$$A_\Theta = \{(s, \theta) : s \in [0, l], \theta \in [\Theta, \pi - \Theta]\}. \tag{6.1}$$

- Proposition 6.1.** 1. *There exists $N(\Theta) > 0$ such that any $z \in A_\Theta \setminus \{p_0, p_1\}$ will leave γ in less than N hits: $n(z) < N$.*
 2. *There exists a $\delta > 0$ such that for all smooth convex arcs $\tilde{\gamma}$ of length l with $\int_0^l k_{\tilde{\gamma}}(s) ds \leq \pi$, if $\|\gamma - \tilde{\gamma}\|_{C^2} < \delta$ and if $z \in A_\Theta \setminus \{p_0, p_1\}$, then $\tilde{n}(z) < n(z) + 1$.*
 3. *Given $\varepsilon > 0$ there exists $\delta > 0$ such that if $\|\gamma - \tilde{\gamma}\|_{C^2} < \delta$, then to any $z \in A_\Theta \setminus \{p_0, p_1\}$ we can associate a $z^* = (s^*, \theta^*)$ so that*

$$\tilde{n}(z) \leq n(z^*), \quad \text{and} \quad |z - z^*| < \varepsilon. \tag{6.2}$$

See Appendix A2 for the proof. Our main result is

Theorem 6.2 (Theorem 4). *Let γ be a smooth convex focusing arc with $\int_0^l k_\gamma(s) ds \leq \pi$. Then all smooth convex arcs $\tilde{\gamma}$ of length l satisfying $\int_0^l k_{\tilde{\gamma}}(s) ds \leq \pi$ for which $\|\gamma - \tilde{\gamma}\|_{C^6}$ is sufficiently small will also be focusing arcs.*

Proof. Choose $\Theta > 0$ so small that the results of Sect. 5 will hold for $\theta \leq 2\Theta$. Use Proposition 6.1 (1) to choose $N(\Theta)$.

By Theorem 5.6, all $z \in S \setminus A_\Theta$ will be focused by γ . The value at which the small angle analysis becomes valid depends on γ in a C^6 way. Thus there exists a $\delta_1 > 0$ such that if $\|\gamma - \tilde{\gamma}\|_{C^6} < \delta_1$ then all $z \in S \setminus A_\Theta$ will be focused by $\tilde{\gamma}$.

We now show that given Θ as above there exists a δ_2 such that if $\|\gamma - \tilde{\gamma}\|_{C^3} < \delta_2$, then any $z \in A_\Theta \setminus \{p_0, p_1\}$ will be focused by $\tilde{\gamma}$. The theorem then follows by taking $\|\gamma - \tilde{\gamma}\|_{C^6} < \min\{\delta_1, \delta_2\}$.

By (3.3), z will be focused by $\tilde{\gamma}$ if

$$-\tilde{k}(\tilde{s}_k(z)) < m(D\tilde{\Phi}_z^k(X_p^-)) < \tilde{k}(s_k(z)), \quad k=0, 1, 2, \dots, \tilde{n}(z) < N, \tag{6.3}$$

where \tilde{k} is the curvature of $\tilde{\gamma}$. Since γ is a focusing curve, we know that z^* , given by Proposition 6.1(3), is focused:

$$-k(s_k(z^*)) < m(D\Phi_{z^*}^k(X_p^-)) < k(s_k(z^*)), \quad k=0, 1, 2, \dots, n(z^*) < N. \tag{6.4}$$

Given $\varepsilon > 0$ there exists $\delta_3 > 0$ such that if $\|\gamma - \tilde{\gamma}\|_{C^2} < \delta_3$, then

$$|k(s_k(z^*)) - \tilde{k}(\tilde{s}_k(z))| < \varepsilon, \quad z \in A_\Theta, \quad k=0, 1, 2, \dots, \tilde{n}(z) \leq n(z^*) < N. \tag{6.5}$$

The norm $\|D\Phi_z^k(X_p^-)\|$ is uniformly bounded away from zero for all $z \in A_\Theta$ and $k=0, 1, 2, \dots, n(z) < N$. Therefore if we show that $\|D\Phi_{z^*}^k(X_p^-) - D\tilde{\Phi}_z^k(X_p^-)\|$ can be made uniformly small for all $z \in A_\Theta$, $k=0, 1, 2, \dots, \tilde{n}(z) < N$, then $|m(D\Phi_{z^*}^k(X_p^-)) - m(D\tilde{\Phi}_z^k(X_p^-))|$ can be made uniformly small and hence (6.3) will hold.

The triangle inequality gives that

$$\begin{aligned} & \|D\Phi_{z^*}^k(X_p^-) - D\tilde{\Phi}_z^k(X_p^-)\| \\ & \leq \|D\Phi_{z^*}^k(X_p^-) - D\Phi_z^k(X_p^-)\| + \|D\Phi_z^k(X_p^-) - D\tilde{\Phi}_z^k(X_p^-)\| \\ & \leq \|D^2\Phi^N\| |z - z^*| + \|D\Phi_z - D\tilde{\Phi}_z\|^N, \end{aligned}$$

which by Proposition 6.1(3), Lemma 1.3 can be made uniformly small if $\|\gamma - \tilde{\gamma}\|_{C^3}$ is sufficiently small. \square

7. Ellipses

We describe new types of focusing arcs that are pieces of ellipses. Since scaling does not effect the focusing property, we parameterize the ellipse $E(t) = \{x(t) = a \cos t, y(t) = \sin t\}$, $a \geq 1$, $t \in [-\pi, \pi]$, and let $E_{[t_a, t_b]}$ denote the arc of the ellipse for which $t \in [t_a, t_b]$. The focal points of the ellipse occur at $x = \pm c = \pm \sqrt{a^2 - 1}$, $y = 0$.

Theorem 7.1. *The half-ellipse $E_{[-\pi/2, \pi/2]}$ is a focusing arc if and only if $a < \sqrt{2}$.*

Note that as mentioned in Sect. 1, the arc $E_{[-\pi/2 - \varepsilon, \pi/2 + \varepsilon]}$ for $\varepsilon > 0$, cannot be included in a billiard that has positive Lyapunov exponent almost everywhere.

Billiard motion in an ellipse is integrable [S2]. The phase space decomposes into invariant circles of two types corresponding to trajectories that cross the x axis with $-c < x < c$ (\mathcal{H} trajectories) and to trajectories that cross the x axis with $|x| > c$ (\mathcal{E} trajectories). \mathcal{H} trajectories are tangent to caustics that are confocal hyperbolas; \mathcal{E} trajectories have caustics that are confocal ellipses. In [D3] we use the integrable behaviour of billiards in an ellipse to construct a smooth convex billiard with positive topological entropy.

Proposition 7.2. *The elliptical arc $E_{[t_a, t_b]}$, $t_a \leq 0 < t_b \leq \pi$, is focusing if the line connecting $E(t_a)$ and $E(t_b)$ crosses the x axis at $x \geq c$ or if $t_a = 0$ and $t_b = \pi$.*

Proof. Our assumption implies that only \mathcal{E} trajectories can hit $E_{[t_a, t_b]}$ more than once. For such a trajectory, let ξ_T be the variation consisting of trajectories that are tangent to the same confocal ellipse. In phase space, ξ_T is the tangent vector to the invariant circle. This variation focuses in linear approximation at the caustic. The ordering property implies that the parallel variation X_p^- will focus before the variation ξ_T . Since ξ_T is focused between each pair of collisions, so is X_p^- . \square

The previous results on elliptical arcs implied that $E_{[0, \pi]}$ was a focusing arc [W3] and that any sufficiently short sub-arc of $E_{[-\pi/4, \pi/4]}$ was a focusing arc [Mr1].

The proof of Proposition 7.2 implies that any \mathcal{E} trajectory will be focused by $E_{[\pi/2, -\pi/2]}$. Thus to prove Theorem 7.1 we need only show that all \mathcal{H} trajectories are focused.

Lemma 7.3. (1) *If $1 < a < \sqrt{2}$, then any \mathcal{H} trajectory other than the periodic orbit $\{p_0, p_1\}$ has at most 2 successive collisions with the arc $E_{[\pi/2, -\pi/2]}$.*
 (2) *If $a > \sqrt{2}$, there exist \mathcal{H} trajectories that have at least 3 collisions with $E_{[\pi/2, -\pi/2]}$.*

Proof. (1) The normal line through a point $(x(t), y(t))$, $t \in [0, \pi/2]$ intersects the y axis at $y = (1 - a^2) \sin t$. For $a < \sqrt{2}$, an \mathcal{H} trajectory z that starts in $E_{[\pi/2, -\pi/2]}$ and for which Φz is also in $E_{[\pi/2, -\pi/2]}$ will have $\theta(\Phi z) < \pi/2$. Since a \mathcal{H} trajectory crosses the x axis at some $x \in (-c, c)$, Φz will leave $E_{[\pi/2, -\pi/2]}$.

(2) If $a > \sqrt{2}$, there exists a t' such that the normal line through $(x(t), y(t))$, $t \in (t', \pi/2)$, intersects the y axis at $y < -1$. If $z = (s=0, \theta) = \Phi^{-1}(x(t), y(t))$ is the pre-image of such a point, then $\theta(\Phi z) > \pi/2$ and $\Phi^2 z$ is still in $E_{[\pi/2, -\pi/2]}$. \square

In what follows we will use an angle coordinate ϕ (see Fig. 3) defined at the point $E(t)$,

$$\phi = \begin{cases} \theta, & \text{if } t \in [-\pi/2, 0) \\ \pi - \theta, & \text{if } t \in [0, \pi/2]. \end{cases}$$

Let $z_1 = (t_1, \phi_1) \in E_{[\pi/2, -\pi/2]}$ be an \mathcal{H} trajectory and $z_2 = (t_2, \phi_2)$ be its image. We denote by $\phi_0(t)$ the angle at $E(t)$ whose ray connects the points $E(t)$ and $E(-t)$.

Lemma 7.4. *Let $1 < a < \sqrt{2}$ and $t_2 \in (0, \pi/2]$. If $\phi_2 \in (0, \phi_0(t_2))$, then*

$$r_2 \sin \phi_2 < L < 2r_2 \sin \phi_2, \tag{7.1}$$

where r_2 is the radius of curvature at $E(t_2)$ and L is the distance between $E(t_2)$ and $E(t_1)$.

Proof. We denote by D_1 and D_2 the circles tangent to the ellipse at $E(t_2)$ with radius of curvature r_2 and $r_2/2$ respectively. These are the osculating and half-osculating circles (Fig. 7). Equation (7.1) is equivalent to the following relation: for $x \geq x(t_2)$,

$$\text{Interior } D_2 \subset \text{Interior } E \subset \text{Interior } D_1. \tag{7.2}$$

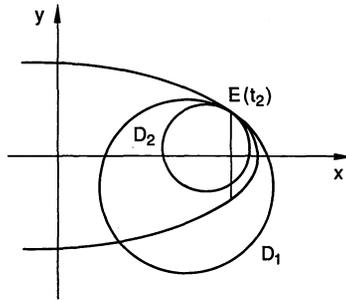


Fig. 7. Half ellipse and osculating circles

Denoting by (X_1, Y_1) and (X_2, Y_2) the centers of the circles D_1 and D_2 , we have that

$$Y_1 = y(t_2) - r_2 \sin \phi_0(t_2),$$

$$Y_2 = y(t_2) - \frac{r_2}{2} \sin \phi_0(t_2).$$

Since $r(t) = \frac{1}{a} (a^2 \sin^2 t + \cos^2 t)^{3/2}$ and $\sin \phi_0(t) = a \sin |t| / \sqrt{a^2 \sin^2 t + \cos^2 t}$, we get for $t_2 \in (0, \pi/2]$ and $a \in (1, \sqrt{2})$ that

$$Y_1 < 0 < Y_2. \tag{7.3}$$

The relation $r(t) \leq r(t_2)$ for $t \in [0, t_2]$ implies that in the region $x \geq x(t_2)$ and $y \geq 0$ we have

$$\text{Interior } E \subset \text{Interior } D_1. \tag{7.4}$$

Now the symmetry of the ellipse and the relation $Y_1 \leq 0$ implies that (7.4) also holds for $y \leq 0$.

In the case that $r_2(t_2)/2 \leq r(t)$ for all $t \in [0, t_2]$, the previous argument implies that

$$\text{Interior } D_2 \subset \text{Interior } E. \tag{7.5}$$

Now suppose that for some $t^* \in (0, t_2)$, we have that $r(t_2)/2 = r(t^*)$. Let $X_{2, \max}$ denote the maximum x value obtained by D_2 . Then for $t_2 \in (0, \pi/2]$ and $a < 2$, we have that

$$X_{2, \max} = x(t_2) + \frac{r(t_2)}{2} (1 - \cos \phi_0(t_2)) < a. \tag{7.6}$$

It is conceivable that as t decreases from t^* to 0, $E(t)$, which was originally outside D_2 , would cross D_2 . Past this crossing point, E would remain inside D_2 since $r(t) < r(t_2)/2$. But this would imply that $a < X_{2, \max}$, contradicting (7.6).

Hence (7.5) holds for $x \geq x(t_2)$ and $y \geq 0$. Symmetry and the relation $Y_2 \geq 0$ then implies the result for $y \leq 0$. \square

Proof of Theorem 7.1. Let $a < \sqrt{2}$. We examine \mathcal{H} trajectories starting at $z_1 = (t_1, \phi_1)$, $t_1 \in [-\pi/2, 0]$ and for which the image point z_2 is still in E . We must show that the variation X_p^- , which is parallel when reaching z_1 , focuses between z_1 and z_2 and is converging after leaving z_2 . By (3.3), this variation focuses at a distance $\frac{1}{2}r_1 \sin \phi_1$ from z_1 . The variation that is parallel when leaving z_2 focuses at a distance $\frac{1}{2}r_2 \sin \phi_2$ before z_2 .

We define

$$\Delta = L - \frac{1}{2}r_1 \sin \phi_1 - \frac{1}{2}r_2 \sin \phi_2, \tag{7.7}$$

where L is the distance between $E(t_1)$ and $E(t_2)$ and $r_i = r(t_i)$ is the radius of curvature at $E(t_i)$. If $\Delta > 0$, then Lemma 1.6 and Property 3.4 imply that the desired focusing behaviour occurs.

We fix $t_1 \in [-\pi/2, 0)$ and examine $\Delta(t_1, \phi_1)$ for $\phi_1 \in [\phi_0(t_1), \phi_f(t_1)]$. The angle $\phi_f(t_1)$ gives the ray connecting $E(t_1)$ and $E(\pi/2)$. We claim that

$$\Delta(t_1, \phi_1) > 0, \quad \forall \phi_1 \in [\phi_0(t_1), \phi_f(t_1)]. \tag{7.8}$$

For $\phi_1 = \phi_0(t_1)$ we have by Lemma 7.4 that

$$\Delta(t_1, \phi_0(t_1)) = L - r_2 \sin \phi_2 > 0. \tag{7.9}$$

We consider L , r_2 and ϕ_2 all as functions of ϕ_1 . Differentiating (7.7) gives

$$\frac{\partial}{\partial \phi_1} \Delta = \frac{1}{2} \frac{\partial L}{\partial \phi_1} - \frac{1}{2} r_1 \cos \phi_1 + \frac{1}{2} \frac{\partial L}{\partial \phi_1} - \frac{1}{2} \frac{\partial r_2}{\partial \phi_1} \sin \phi_2 - \frac{1}{2} r_2 \cos \phi_2 \frac{\partial \phi_2}{\partial \phi_1}.$$

Integrating $\frac{\partial r_2}{\partial \phi_1} \sin \phi_2$ by parts, we get

$$\begin{aligned} \int_{\phi_0}^{\phi_1} \frac{\partial \Delta}{\partial \phi_1} d\phi_1 &= \frac{1}{2} (L(\phi_1) - r_2(\phi_1) \sin \phi_2(\phi_1)) - \frac{1}{2} (L(\phi_0) - r_1 \sin \phi_0) \\ &\quad + \frac{1}{2} \int_{\phi_0}^{\phi_1} \left(\frac{\partial L}{\partial \phi_1} - r_1 \cos \phi_1 \right) d\phi_1. \end{aligned}$$

Calculations give that

$$\frac{\partial L}{\partial \phi_1} = L \cot \phi_2 \quad \text{and} \quad \frac{\partial \phi_2}{\partial \phi_1} = -1 + \frac{L}{r_2 \sin \phi_2}. \tag{7.10}$$

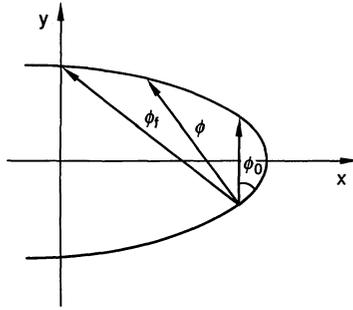


Fig. 8. One parameter family of rays

Since $L(\phi) \leq 2r_2 \sin \phi_2$ by Lemma (7.4), we have that $\frac{\partial \phi_2(\phi_1)}{\partial \phi_1} \leq 1$ for all $\phi_1 \in [\phi_0(t_1), \phi_f(t_1)]$. Thus

$$\phi_2(\phi_1) - \phi_1 = \int_{\phi_0}^{\phi_1} \left(\frac{\partial \phi_2}{\partial \phi_1} - 1 \right) d\phi_1 < 0,$$

and hence $\cos \phi_2(\phi_1) \geq \cos \phi_1$.

Therefore

$$\frac{\partial L}{\partial \phi_1} - r_1 \cos \phi_1 = L \frac{\cos \phi_2}{\sin \phi_2} - r_1 \cos \phi_1 \geq \frac{\cos \phi_1}{\sin \phi_2} (L - r_2 \sin \phi_2) \geq 0. \quad (7.11)$$

Using that $\Delta(\phi_0) = L(\phi_0) - r_1 \sin \phi_0$, we get that

$$\begin{aligned} \Delta(\phi_1) &= \Delta(\phi_0) + \int_{\phi_0}^{\phi_1} \frac{\partial \Delta}{\partial \phi_1} d\phi_1 \\ &= \frac{1}{2} (L(\phi_1) - r_2(\phi_1) \sin \phi_2(\phi_1)) + \frac{1}{2} \Delta(\phi_0) \\ &\quad + \frac{1}{2} \int_{\phi_0}^{\phi_1} \left(\frac{\partial L}{\partial \phi_1} - r_1 \cos \phi_1 \right) d\phi_1 > 0, \end{aligned} \quad (7.12)$$

by Lemma 7.4 and Eqs. (7.9, 7.10, 7.11). Thus we have proved (7.8).

If $a \geq \sqrt{2}$, then $\Delta \leq 0$ for the periodic orbit connecting p_0 and p_b , since $r(s_0) = a^2 \geq 2$. Hence these trajectories do not focus. For $a > \sqrt{2}$, there are orbits starting at $(s=0, \phi)$, $\phi < \pi/2$, for which $\Delta > 0$. \square

For the Bunimovich stadium, the length of the flat sides can be arbitrarily short. In the elliptical stadium, the length of the flat sides must grow as a increases to $\sqrt{2}$. One could try to calculate bounds on these lengths.

In the $a = \sqrt{2}$ case, all trajectories except the periodic orbit $\{p_0, p_b\}$ are focused by $E_{[-\pi/2, \pi/2]}$. However the time it takes for the parallel variation that starts at $(t_1 = -\pi/2, \phi_1)$ to focus after leaving the half-ellipse goes to infinity as ϕ_1 approaches $\pi/2$.

The methods of this section could be used to determine conditions under which the elliptical arc $E_{[-t, t]}$, $t \in (0, \pi/2)$ is a focusing arc.

Appendix A1

We prove the Lemmas of Sect. 5.

Proof of Lemma 5.1. By induction. For $k=1$,

$$\begin{aligned} |y_1 - y| &= |y^4 g(x_0, y)| \leq y^4 c_2, \\ &\leq \frac{1}{2} y, \end{aligned}$$

for y sufficiently small.

If we assume (5.6) is true for $k-1$, then

$$\begin{aligned} |y_k - y| &\leq |y_k - y_{k-1}| + |y_{k-1} - y| \\ &\leq y_{k-1}^4 c_2 + |y_{k-1} - y|. \end{aligned}$$

By induction, $|y_{k-1} - y_0| \leq \frac{1}{2} y$ so that $y_{k-1} \leq \frac{3}{2} y$. Hence,

$$|y_k - y| \leq \left(\frac{3}{2} y\right)^4 c_2 + |y_{k-1} - y|.$$

Iterating this estimate gives

$$|y_k - y| \leq k \left(\frac{3}{2} y\right)^4 c_2 \leq \frac{1}{2} y,$$

providing that

$$k \leq \frac{2^4}{2c_2 3^4 y^3}. \quad \square \tag{A1.1}$$

Proof of Lemma 5.2. By (5.4),

$$x_{k+1} - x_k = y_k + y_k^3 f(x_k, y_k).$$

If we assume that

$$\frac{3}{2} y \geq y_k \geq \frac{1}{2} y, \tag{A1.2}$$

then

$$|x_{k+1} - x_k| \geq \frac{1}{2} y - \left(\frac{3}{2} y\right)^3 c_1 > 0,$$

for y sufficiently small. If (A1.2) holds for $k=0, 1, 2, \dots, n-1$, then

$$\begin{aligned} |x_n - x| &\geq \sum_{k=0}^{n-1} |x_{k+1} - x_k| \geq n \left(\frac{1}{2} y - \left(\frac{3}{2} y \right)^3 c_1 \right) \\ &= \frac{ny}{2} \left(1 - \frac{3^3 y^2 c_1}{2^2} \right) \\ &\geq \frac{ny}{4}, \end{aligned}$$

again for y sufficiently small.

Hence providing that (A1.2) holds for $k=1, 2, \dots, n$, then $|x_n - x| > 1$, when $n > \frac{4}{y}$.

But for y sufficiently small, (A1.2) holds for $n = \text{const}/y^3$ by Lemma 5.1. \square

Proof of Lemma 5.5. For $k=1$, using the bounds on the entries of $D\Phi$ given by (5.7), one has

$$\begin{pmatrix} x'_1 \\ y'_1 \end{pmatrix} = D\Phi \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \pm c_4 y^3 \\ \pm c_4 y^4 \end{pmatrix}.$$

Assume the result holds for $k-1$. We prove the upper estimate for k .

$$\begin{aligned} x'_k &\leq x'_{k-1} + y'_{k-1} + c_4 y^2 y'_{k-1} + c_4 y^3 x'_{k-1} \\ &\leq 1 + y^3 k c_4 + y^4 c_4 \{(k-1)^2 + (k-1)\} + y^5 (k-1) \\ &\quad + y^6 2(k-1)c_4^2 + y^7 \{c_4(k-1) + c_4^2(k-1)^2\}. \end{aligned} \tag{A1.3}$$

Since from Corollary 5.3 we have that

$$k-1 \leq \frac{c_3}{y}, \tag{A1.4}$$

we can estimate the terms of order 5 and above. For y sufficiently small,

$$y^4(k-1)c_4 \{y/c_4 + 2y^2 c_4 + y^3 + y^3 c_4(k-1)\} \leq y^4(k-1)c_4. \tag{A1.5}$$

Hence

$$x'_k \leq 1 + y^3 k c_4 + y^4 c_4 k^2. \tag{A1.6}$$

For y'_k we get,

$$\begin{aligned} y'_k &\leq y'_{k-1} + c_4 y^3 y'_{k-1} + c_4 y^4 x'_{k-1} \\ &\leq y^4 k c_4 + y^5 (k-1) + y^7 2c_4^2(k-1) + y^8 \{c_4(k-1) + c_4^2(k-1)^2\}. \end{aligned}$$

For y sufficiently small the terms of order seven and higher can be estimated by

$$y^5 \{y^2 2c_4^2(k-1) + y^3(c_4(k-1) + c_4^2(k-1)^2)\} \leq y^5. \tag{A1.7}$$

This gives that

$$y'_k \leq y^4 k c_4 + y^5 k.$$

The estimates for the lower bounds are obtained similarly. \square

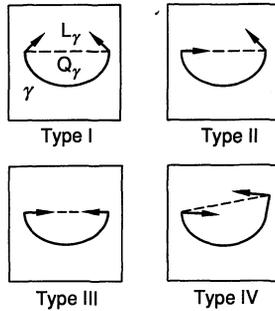


Fig. 9. Types of arcs

Appendix A2

We split our proof of Proposition 6.1 into cases depending on how the normals at $\gamma(0)$ and at $\gamma(l)$ are oriented. By reversing orientation, we can suppose that any convex arc is of one of the following types (Fig. 9):

- I. Both normals lie outside the domain Q_γ .
- II. The normal at $\gamma(0)$ coincides with the line \mathcal{L}_γ , the normal at $\gamma(l)$ lies outside Q_γ .
- III. Both normals coincide with \mathcal{L}_γ .
- IV. The normal at $\gamma(0)$ lies inside Q_γ , the normal at $\gamma(l)$ lies outside Q_γ .

First we discuss case (I). When there are similar arguments to deal with the $\theta \leq \pi/2$ and the $\theta \geq \pi/2$ cases, we discuss only the former.

Phase Diagram. We construct a phase (space) diagram that describes the behaviour of the function $n(z)$ (Fig. 10).

For $s \in (0, 1)$, the monotone twist property (Proposition 3.6) implies the existence of a monotone sequence of angles

$$0 < \dots < \theta_{2R}(s) < \theta_{1R}(s) < \theta_{1L}(s) < \theta_{2L}(s) < \dots < \pi, \tag{A2.1}$$

with the property that

$$s_k(s, \theta_{kR}(s)) = 1, \quad s_k(s, \theta_{kL}(s)) = 0. \tag{A2.2}$$

The indices R and L indicate whether the iterates of a point lie to the right or left of the original point.

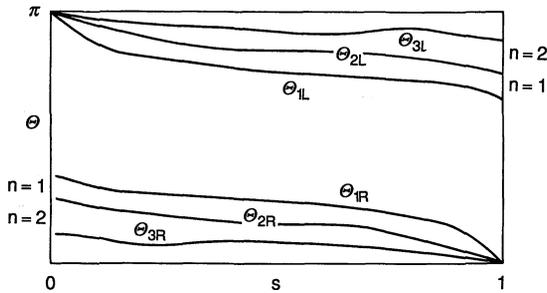


Fig. 10. Phase diagram case (I)

If $s=0(s=1)$, the situation degenerates in an obvious way: $\theta_{kL}(0) \equiv \pi(\theta_{kR}(1) \equiv 0)$, $k=1, 2, \dots$

We can construct the curve $\theta = \theta_{kR}(s)$, $s \in [0, 1]$ using time reversal of the billiard. Examine the curve $\Phi^k(s=1, \theta)$, $\theta \in [\theta_{kL}(1), \pi)$. By the twist property, this curve is the graph of a function $\theta = \theta_{kL}^t(s)$, $s \in [0, 1]$. We have the relation

$$\theta_{kR}(s) = \pi - \theta_{kL}^t(s). \tag{A2.3}$$

If (s, θ) satisfies $\theta \in (\theta_{(k+1)R}(s), \theta_{kR}(s)] \cup (\theta_{kL}(s), \theta_{(k+1)L}(s)]$, then we claim that

$$\Phi^j(s, \theta) \in S_\gamma \quad j=0, 1, 2, \dots, k, \tag{A2.4a}$$

so that

$$\Phi^{k+1}(s, \theta) \notin S_\gamma \tag{A2.4b}$$

$$n(s, \theta) = k. \tag{A2.5}$$

Our assumption on the normals imply that

$$\theta_k(s, \theta_{kR}(s)) \leq \pi/2, \quad \theta_k(s, \theta_{kL}(s)) > \pi/2, \tag{A2.6}$$

and therefore that

$$\Phi^{k+1}(s, \theta_{k\alpha}(s)) \notin S_\gamma \quad \alpha = R, L. \tag{A2.7}$$

The monotone property implies that for $\theta \in (\theta_{(k+1)R}(s), \theta_{kR}(s)]$,

$$s_j(s, \theta_{(k+1)R}) < s_j(s, \theta) \leq s_j(s, \theta_{kR}), \quad j=1, 2, \dots, k, \tag{A2.8}$$

and hence (A2.4a) holds.

A slight generalization implies (A2.4b). Imagine that γ is contained in a closed, convex curve ∂Q (possibly non-smooth) such that for all $z = (s, \theta) \in (\theta_{(k+1)R}, \theta_{kR}]$ the vertical variation $X_\theta \in T_z S_\gamma$ has time to focus between $\Phi^k z$ and $\Phi^{k+1} z$. Since the times it takes these variations to focus are bounded (Lemma 4.2), such a curve ∂Q can always be constructed: just use some long, straight pieces. Then on this extended billiard, the monotonicity still holds:

$$l = s_{k+1}(s, \theta_{(k+1)R}) < s_{k+1}(s, \theta) \leq s_{k+1}(s, \theta_{kR}). \tag{A2.9}$$

Since $\Phi^{k+1}(s, \theta_{kR})$ is outside of γ , the same must be true for $\Phi^{k+1}(s, \theta)$.

Using out phase diagram, we give the

Proof of Proposition 6.1. 1. Choose N so large that

$$\theta_{NR}(s) < \Theta, \quad \pi - \theta_{NL}(s) < \Theta, \quad s \in [0, l]. \tag{A2.10}$$

Since $\theta_{NR}(s)$ is given by (A2.3), the analysis of Sect. 5 implies that once Θ is sufficiently small, such a choice of N is always possible. Then for $\theta \in (\theta_{NR}(s), \theta_{NL}(s)) \supset [\Theta, \pi - \Theta]$, we have that $n(s, \theta) \leq N - 1 < N$.

2. If the perturbed curve $\tilde{\gamma}$ is close enough to γ , then it is also of type (I), and we can apply the previous analysis to construct the curves $\{\tilde{\theta}_{kR}(s), \tilde{\theta}_{kL}(s)\}, s \in [0, l], k = 1, 2, \dots, N$. For a fixed N , we can make $|\Phi^k(s, \theta) - \tilde{\Phi}^k(s, \theta)|$ uniformly small for all $k = 0, 1, \dots, N$ by making $|\gamma - \tilde{\gamma}|_{C^2}$ sufficiently small. Thus by (A2.3), given $\varepsilon > 0$ there exists $\delta > 0$ such that if $|\gamma - \tilde{\gamma}|_{C^2} < \delta$ then

$$\begin{aligned} |\theta_{kR}(s) - \tilde{\theta}_{kR}(s)| &< \varepsilon, \\ |\theta_{kL}(s) - \tilde{\theta}_{kL}(s)| &< \varepsilon, \quad s \in [0, l], \quad k = 1, 2, \dots, N. \end{aligned} \tag{A2.11}$$

Choosing $\varepsilon < \min_{s \in [0, l]} \{\Theta - \theta_{NR}(s), \theta_{NL}(s) - (\pi - \Theta)\}$ gives the result.

3. If $\tilde{n}(z) \leq n(z)$, then we set $z^* = z$. If $\tilde{n}(z) > n(z)$ and $\theta \in (\theta_{(n(z)+1)R}(s), \theta_{n(z)R}(s))$, then we set $s^* = s$ and $\theta^* = \theta_{\tilde{n}(z)R}(s)$. Thus $\tilde{n}(z) = n(z^*)$ and by (A2.11), $|z - z^*| < \varepsilon$. \square

We now outline how the argument goes for the other cases.

Case (II):

The argument proceeds as in case (I) except that the point $(s = l, \theta = \theta_{1L}(s) = \pi/2)$ has $n = 2$; first $s_1(l, \theta_{1L}(l)) = 0$, then $s_2(l, \theta_{1L}(l)) = l$, and then it leaves γ .

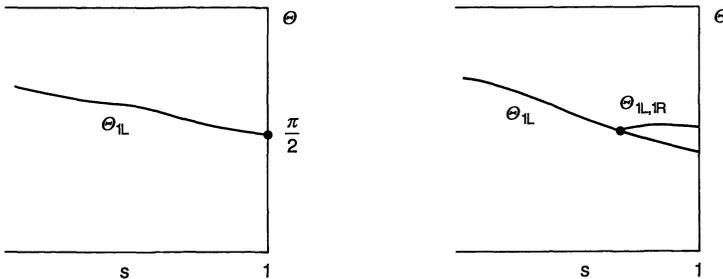


Fig. 11. Case (II) bifurcation

If the perturbed curve $\tilde{\gamma}$ is type (I) or (II), then the phase diagram remains the same from which Proposition 6.1 follows. If $\tilde{\gamma}$ is type (IV), then there is a bifurcation between the phase diagrams of S_γ and $S_{\tilde{\gamma}}$ (Fig. 11). The point

$(l, \theta_{1L}(l)) \in S_\gamma$, turns into a whole curve $\theta = \tilde{\theta}_{1L, 1R}(s) \subset S_\gamma$, $s \in [\tilde{\sigma}_1, l]$, where $\tilde{\sigma}_1 = \tilde{s}_1(s=0, \theta = \pi/2)$. The points $(s, \tilde{\theta}_{1L, 1R}(s))$ have the property that

$$\tilde{s}_2(s, \tilde{\theta}_{1L, 1R}(s)) = l.$$

We let $\theta = \tilde{\theta}_{1R, 1L}(s)$, $s \in [\tilde{\sigma}_1, l]$ be the function whose graph gives the curve $\tilde{\Phi}^2(s=l, \theta)$, $\theta \in [\tilde{\theta}_{1L}(l), \tilde{\theta}_{1L, 2R}(l)]$. Then

$$\tilde{\theta}_{1L, 1R}(s) = \pi - \tilde{\theta}_{1R, 1L}(s). \tag{A2.12}$$

We denote by

$$S_\gamma^{1L, 1R} = \{(s, \theta) : s \in [\sigma_1, l], \theta \in (\tilde{\theta}_{1L}(s), \tilde{\theta}_{1L, 1R}(s))\}. \tag{A2.13}$$

Any $z \in S_\gamma^{1L, 1R}$ will first go left, then right and then leave $\tilde{\gamma}$ after $n(s, \theta) = 2$ hits.

We associate to $z \in S_\gamma^{1L, 1R}$ the point $z^* = (s^* = l, \theta^* = \theta_{1L}(l))$. By choosing δ small enough, we can make $|(s, \tilde{\theta}_{1L, 1R}(s)) - (s = l, \theta_{1L}(l))| < \varepsilon$ proving (3) of the Proposition.

Case (III):

We get the same phase diagram as in case (II) except that the points $\{p_0, p_l\}$ never leave S . For our purposes, we consider that these points have $n = 2$. Then if the perturbed arc $\tilde{\gamma}$ is type (II) or (IV), we compare points $z \in S_\gamma^{1L, 1R}$ with $z^* = (s = l, \theta = \pi/2)$. Here is where Definition 1.2 of focusing gets used.

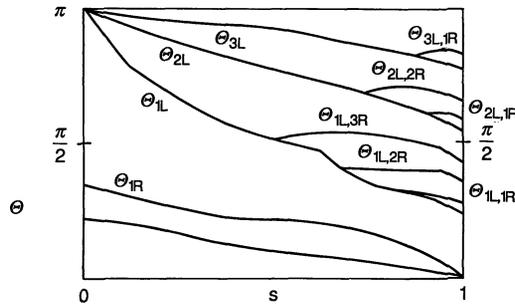


Fig. 12. Phase diagram case (IV)

Case (IV):

Here the phase diagram gets more complicated (see Fig. 12 for an example). Branching off of the curves $\theta_{kL}(s)$ are curves $\theta_{kL, jR}(s)$ that satisfy

$$\begin{aligned} s_i(s, \theta_{kL, jR}(s)) &< s_{i-1}(s, \theta_{kL, jR}(s)), & i = 1, 2, \dots, k, \\ s_i(s, \theta_{kL, jR}(s)) &> s_{i-1}(s, \theta_{kL, jR}(s)), & i = k + 1, \dots, k + j, \\ s_{k+j}(s, \theta_{kL, jR}(s)) &= l. \end{aligned} \tag{A2.14}$$

The twist property implies the following.

The branchings will occur for $k = 1, 2, \dots, K_1$, where K_1 is determined as follows. The point $(s=l, \theta_{1L}(l))$ will first hit at $s=0$ and then monotonically move to the right until it leaves γ . We set

$$K_1 = n(l, \theta_{1L}(l)) - 1. \tag{A2.15}$$

For a fixed $k \leq K_1$ the branchings $\theta_{kL, jR}$ are defined for $j = 1, 2, \dots, J_k$. To determine J_k , take the point $(l, \theta_{kL}(l))$ and set

$$J_k = n(l, \theta_{kL}(l)) - k. \tag{A2.16}$$

Typically we have that $s_{n(l, \theta_{kL}(l))}(l, \theta_{kL}(l)) < l$, so that $\theta_{kL, jR}(s) > \theta_{kL}(s)$. In the degenerate case that $s_{n(l, \theta_{kL}(l))}(l, \theta_{kL}(l)) = l$, the curve $\theta = \theta_{kL, jR}(s)$ reduces to a point $(l, \theta_{kL, jR}(l)) = (l, \theta_{kL}(l))$.

The curve $\theta_{kL, jR}(s)$ is defined for $s \in [\sigma_{kj}, \pi]$. The value $s = \sigma_{kj}$ for which $\theta_{kL, jR}(s)$ is first defined satisfies

$$s_k(\sigma_{kj}, \theta_{kL, jR}(\sigma_{kj})) = 0,$$

and hence $\theta_{kL, jR}(\sigma_{kj}) = \theta_{kL}(\sigma_{kj})$.

The curves $\{\theta_{kL}(s), \theta_{kL, jR}(s)\}$ bound domains $S_\gamma^{kL, jR}$ with the property that if $z \in S_\gamma^{kL, jR}$ then z goes to the left k times, goes to the right j times and then leaves γ . Hence $n(z) = k + j$.

For points in $S_\gamma^{kL, jR}$, $j \neq 0$, the total number of collisions with γ is bounded. The results of Sect. 5 imply that for θ sufficiently small, a trajectory (s, θ) will move monotonically in one direction and then leave γ . Hence given Θ sufficiently small, there exists an N_1 such that

$$\theta_{N_1 L}(s) > \pi - \Theta, \quad \forall s \in [0, L], \quad \text{and} \quad K_1 < N_1. \quad (\text{A2.17})$$

For $z \in S_\gamma^{kL, jR}$, we have

$$\begin{aligned} n(z) &\leq K_1 + J_k = K_1 + n(l, \theta_{kL}(l)) \\ &\leq K_1 + n(l, \theta_{K_1 L}(l)) < 2N_1. \end{aligned}$$

Note that the boundaries $\{\partial S_\gamma^{kL, jR}\}$ of the domains are all determined by iterates of the curves $s = 0$ or $s = l$: $\Phi^k(s = 0, \theta)$, $\Phi^k(s = l, \theta)$ for $k < 2N_1$ and θ in a suitable domain. Thus for $|\gamma - \tilde{\gamma}|_{C^2}$ sufficiently small, we can make $\text{dist}(\partial S_\gamma^{kL, jR}, \partial S_{\tilde{\gamma}}^{kL, jR}) < \varepsilon$. This proves Proposition 6.1 for case (IV).

Acknowledgements. I thank M. Wojtkowski for telling me about the problem of perturbing the stadium and for stimulating discussions. I also benefited from discussions with V. Bangert, P. Doyle, M. Gerber, A. Hubacher, C. Liverani, R. Markarian, J. Mather, and A. Uribe.

References

- [Bi] Birkhoff, G.D.: Surface transformations and their dynamical applications. Acta Math. **43**, 1–119 (1922). Reprinted in Collected Mathematical Papers, Vol. II. New York: Am. Math. Soc., 111–229 (1950)
- [B1] Bunimovich, L.A.: On the ergodic properties of nowhere dispersing billiards. Commun. Math. Phys. **65**, 295–312 (1979)
- [B2] Bunimovich, L.A.: The stochastic dynamics of rays in resonators. Izv. Vyssh. Uchebn. Zaved. Radiofizika **28** (12), 1601–1602 (1985)
- [B3] Bunimovich, L.A.: Many dimensional nowhere dispersing billiards with chaotic behavior. Physica D **33**, 58–64 (1988)
- [B4] Bunimovich, L.A.: A theorem on ergodicity of two-dimensional hyperbolic billiards. Commun. Math. Phys. **130**, 599–621 (1990)
- [B5] Bunimovich, L.A.: On absolutely focusing mirrors. Preprint
- [B-E] Burton, R., Easton, R.W.: Ergodicity of linked twist maps. Lecture Notes in Math., Vol. 819, pp. 35–49. Berlin, Heidelberg, New York: Springer 1980
- [C-F-S] Cornfeld, I.P., Fomin, S.V., Sinai, Ya.G.: Ergodic theory. Berlin, Heidelberg, New York: Springer 1982
- [D1] Donnay, V.J.: Geodesic flow on the two-sphere. Part I. Positive measure entropy. Ergod. Th. Dynam. Sys. **8**, 531–553 (1988)
- [D2] Donnay, V.J.: Geodesic flow on the two-sphere. Part II. Ergodicity. Lecture Notes in Math., Vol. 1342, pp. 112–153. Berlin, Heidelberg, New York: Springer 1988

- [D3] Donnay, V.J.: A smooth, strictly convex billiard with positive topological entropy (in preparation)
- [D4] Donnay, V.J.: Billiards with positive entropy. London Mathematical Society Symposium on Dynamical Systems (1988), University of Durham, Abstracts of Lectures
- [G] Gerber, M.: Personal communication
- [H] Hopf, E.: Statistik der geodatischen Linien in Mannigfaltigkeiten negativer Krümmung. Leipziger Berichte **91**, 261–304 (1939)
- [Hu] Hubacher, A.: Instability of the boundary in the billiard ball problem. Commun. Math. Phys. **108**, 483–488 (1987)
- [K-S] Katok, A. and Strelcyn, J.-M., with the collaboration of Ledrappier, F. and Przytycki, F.: Invariant manifolds, entropy and billiards; smooth maps with singularities. Lecture Notes in Math., Vol. 1222. Berlin, Heidelberg, New York: Springer 1986
- [K1] Klingenberg, W.: Riemannian geometry. de Gruyter Studies in Mathematics, 1982
- [Kr] Krylov, N.S.: Works on the foundations of statistical physics. Princeton, NJ: Princeton University Press 1979
- [KSS] Kramli, A., Simanyi, N., Szasz, D.: A transversal theorem for semi-dispersing billiards. Commun. Math. Phys. **129**, 535–560 (1990)
- [L] Lazutkin, V.F.: On the existence of caustics for the billiard ball problem in a convex domain. Math. USSR Izv. **7**, 185–215 (1973)
- [Mn] Manning, A.: Curvature bounds for the entropy of the geodesic flow on a surface. J. London Math. Soc. **24**, 351–357 (1981)
- [Mr1] Markarian, R.: Billiards with Pesin region of measure one. Commun. Math. Phys. **118**, 87–97 (1988)
- [Mr2] Markarian, R.: Non-uniform hyperbolicity, quadratic forms and billiards. Preprint
- [Mt1] Mather, J.: Glancing billiards. Ergod. Th. Dynam. Sys. **2**, 397–403 (1982)
- [Mt2] Mather, J.: Non-existence of invariant circles. Ergod. Th. Dynam. Sys. **4**, 301–309 (1984)
- [M-P] MacKay, R.S., Percival, I.C.: Converse KAM: theory and practice. Commun. Math. Phys. **98**, 469–512 (1985)
- [O] Oseledec, V.I.: The multiplicative ergodic theorem. The Lyapunov characteristic numbers of dynamical systems. Trans. Mosc. Math. Soc. **19**, 197–231 (1968)
- [P] Pesin, Ya.B.: Lyapunov characteristic exponents and smooth ergodic theory. Russ. Math. Surv. **32**, 55–114 (1977)
- [S1] Sinai, Ya.G.: Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. Russ. Math. Surv. **25**, 137–189 (1970)
- [S2] Sinai, Ya.G.: Introduction to ergodic theory. Princeton, NJ: Princeton University Press 1976
- [St] Stark, J.: An exhaustive criterion for the non-existence of invariant circles for area preserving twist maps. Commun. Math. Phys. **117**, 177–189 (1988)
- [W1] Wojtkowski, M.: Linked twist mappings have the K -property. Annals of the New York Academy of Sciences, Vol. 357, pp. 65–76. Nonlinear Dynamics 1980
- [W2] Wojtkowski, M.: Invariant families of cones and Lyapunov exponents. Ergod. Th. Dynam. Sys. **5**, 145–161 (1985)
- [W3] Wojtkowski, M.: Principles for the design of billiards with nonvanishing Lyapunov exponent. Commun. Math. Phys. **105**, 319–414 (1986)

Communicated by J. N. Mather

Note added in proof. By letting $x = s$, $y = r^{1/3}(s)\theta$ in (5.3), Bunimovich [B5] shows that Theorem 4 holds under C^4 perturbations. He also has an independent, simpler proof of Theorem 5.

