SIAM J. APPL. MATH.
Vol. 56, No. 2, pp. 651–680, April 1996

017

OPEN BILLIARDS: INVARIANT AND CONDITIONALLY INVARIANT PROBABILITIES ON CANTOR SETS*

ARTUR LOPES† AND ROBERTO MARKARIAN‡

Abstract. Billiards are the simplest models for understanding the statistical theory of the dynamics of a gas in a closed compartment. We analyze the dynamics of a class of billiards (the open billiard on the plane) in terms of invariant and conditionally invariant probabilities. The dynamical system has a horseshoe structure. The stable and unstable manifolds are analytically described. The natural probability μ is invariant and has support in a Cantor set. This probability is the conditional limit of a conditional probability μ_F that has a density with respect to the Lebesgue measure. A formula relating entropy, Lyapunov exponent, and Hausdorff dimension of a natural probability μ for the system is presented. The natural probability μ is a Gibbs state of a potential ψ (cohomologous to the potential associated to the positive Lyapunov exponent; see formula (0.1)), and we show that for a dense set of such billiards the potential ψ is not lattice. As the system has a horseshoe structure, one can compute the asymptotic growth rate of $n(r)$, the number of closed trajectories with the largest eigenvalue of the derivative smaller than r. This theorem implies good properties for the poles of the associated Zeta function and this result turns out to be very important for the understanding of scattering quantum billiards.

Key words. open billiards, Cantor sets

AMS subject classifications. 58F13, 58F11

Introduction. The main purpose of this paper is to give a partial answer to a question proposed by G. Pianigiani and J. Yorke [25] about probabilistic properties of trajectories of billiards:

> There is a variety of phenomena in which trajectories appear chaotic for an extended period of time but then settle down. Consider a particularly difficult problem of this type. Picture an energy conserving billiard table with smooth obstacles so that all the trajectories are unstable with respect to the initial data. Now suppose a small hole is cut in the table so that the ball can fall through. We would like to investigate the statistical behavior of such phenomena. In particular, suppose a ball is started on the table in some random way according to some probability distribution. Let $p(t)$ be the probability that the ball stays on the table for at least time t and let $p_E(t)$ be the probability that the ball is in a measurable set E after time t. Does $\frac{p_E(t)}{p(t)}$ tend asymptotically to some constant $\mu(E)$ as t goes to infinity? And if it does, what are the properties of μ? Does it depend on the initial distribution?

We thank S. Martinez, who proposed to one of us to study the existence of quasi-stationary measures and its limit laws for billiard systems. For a Markov process analogous results were obtained first in [18] (see also [10]).

* Received by the editors January 3, 1995; accepted for publication (in revised form) March 3, 1995. Financial support for the authors' travel expenses during the performance of this research was provided by IMPA and Asociación de Universidades "Grupo de Montevideo."

† Instituto de Matemática, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, 91540-000, Brazil.

‡ Facultad de Ingenieria, Instituto de Matemática y de Estadística "Prof. Ing. Rafael Laguardia," Universidad de la Republica, CC30, Montevideo, Uruguay.

We will consider a class of billiards that we call open billiards. In this case we will present mathematical proofs of the results that answer the questions proposed above. For open billiards there is no small hole where the ball can fall through, but the ball can get lost to infinity.

The first contribution in the direction of analyzing this type of problems in billiards was done by Pianigiani and Yorke in their previously mentioned paper, where they consider not billiards but a related problem for one-dimensional C^2 expanding maps on the interval. They show the existence of a density F that plays an important role in the one-dimensional case. The measure $\mu_F = F(x)dx$ associated with this density is not invariant for the one-dimensional expanding map but is conditionally invariant. This result generalizes the Lasota–Yorke theorem to the case where the nonwandering set is a Cantor set. More recently Collet, Martinez, and Schmitt [6] presented another nice result related to the one-dimensional C^2 case. They showed that the measure μ_F obtained by Pianigiani and Yorke conditionally converges to a certain invariant measure ν. We will apply these two results in the context of open billiards. In fact the C^2 case is not enough for our purposes, and we need a $C^{1+\epsilon}$ version, which will be proved in the appendix.

We are able to present a complete picture of the dynamical properties of the billiards we analyze.

The dynamics of these billiards are basically that of a horseshoe (if one considers a certain special metric). Stable and unstable manifolds can be precisely described, and several results about a certain "natural" measure will be presented in the following sections.

The setting and our main results will be briefly presented in the following paragraphs.

The simplest example of the class of billiards we consider is that given by three nonintersecting discs with equal radius and such that the centers of the disks are at the vertices of an equilateral triangle. This is a good example for the reader to bear in mind (even if most of the results that we obtain can be applied to more general open billiards).

This billiard is not what is usually called a Sinai billiard [27], since in our case most trajectories (in the Lebesgue sense) will go to infinity. The set of trajectories that remain on the table in the past and in the future defines a Cantor set. The main obstacles to extend the result presented here to a Sinai billiard (with a hole in the table where the ball could fall through) are the singularities that appear in the system due to the corners and the trajectories that are tangent to the boundaries of such billiards. Therefore, we analyze open billiards where such pathologies do not occur.

What we call the "natural" measure μ (sometimes called the escape measure in the literature) was previously considered by Grebogi, Ott, and Yorke (see, for instance, [22, § 5.6]) and has the following description: suppose that we are considering in the plane a certain expanding map whose nonwandering set is a Cantor set with Lebesgue measure zero. A natural generalization of the Bowen–Ruelle–Sinai measure in this case is obtained in the following way. Given a set B contained in the Cantor set C, we will define the value $\mu(B)$. Consider a grid of squares with side ϵ. Denote by b_ϵ the number of squares that intersect B and by c_ϵ the number of squares that intersect the Cantor set C. Now when ϵ goes to zero, if the limit

$$\lim_{\epsilon \to 0} \frac{b_\epsilon}{c_\epsilon} = \mu(B)$$

exists and is independent of the grid for any Borel set B, then we say that μ is a "natural" measure. This procedure is quite natural from the point of view of an experimental observer. Given what is left after n observations (this will produce a slightly distorted grid with a value ϵ inverse ly proportional to n), then one should consider the proportion of what is left of the set that one wants to measure over the full set that still remains. The role of the grid is to give a computable approximation of the Lebesgue measure. We would like to have a procedure allowing to obtain μ as a limit involving the Lebesgue measure (or a measure equivalent to Lebesgue measure).

We will present a precise definition of the probability μ as a Gibbs state [24, 29] of the potential associated with the positive Lyapunov exponent, but the reader should keep in mind the above procedure. (See [1, 2, 15, 19].)

We will also present a formula relating the entropy h_μ, the positive Lyapunov exponent χ_μ, and the Hausdorff dimension δ of the transverse measure (to be defined later):

$$h_\mu = \delta\chi_\mu.$$

For the general case of axiom-A systems, a proof of this formula appears in [13]. Our result is analogous to the one obtained by Chernov and Markarian [4] for hyperbolic billiards, with a correction term δ due to the fractal structure of the Cantor set.

The Lyapunov exponent of a point x will be expressed in terms of the time between bounces $t(x)$ and $k(x)$ (a continued fraction expression involving the time t between bounces of the trajectory by x, the curvature K of the boundaries of the billiard, and the angles ϕ of the collisions with the boundary of the trajectory by x). More precisely, for x almost everywhere, the Lyapunov exponent χ_μ is as follows:

$$\chi_\mu = \int \log|1 + t(x)k(x)|d\mu(x).$$

The precise definitions will be presented in the next paragraphs.

By definition, a function B is lattice if there exist an integer-valued function G, a real positive constant γ, and a continuous function g such that $B = g \circ T - g + G\gamma$.

The probability μ can be defined as the Gibbs state associated with the potential

$$\psi(x) = -\log|1 + t(x)k(x)|; \tag{0.1}$$

this potential is cohomologous to the potential given by minus log of the positive Lyapunov exponent: $-\log f'_{|E^u|}$ (where f is the billiard map to be defined in the next section). It is therefore natural to ask if the potential ψ is not lattice. We are able to show that for a dense set of billiards, this is so (see § 8). When one considers the statistics of the periodic orbits, it is important to know whether the potential is lattice or not [24].

As the system is axiom A, we are able to estimate the asymptotic growth rate of $n(r)$, the number of periodic trajectories ($f^k(x) = x$) with $\log|(f^{k'})_{E^u(x)}(x)|$ smaller than r. The value $n(r)$ grows like $e^{hr}/(hr)$, when r goes to ∞, for some $h > 0$. In a related result, Morita [20] shows that $t(x)$ is not lattice for a general class of billiards.

The class of billiards that we analyze here is a very important model considered in the theory of quantum scattering (see [7, 8, 21, 22, 26, 28]). The asymptotic growth rate of the number of periodic orbits is of indubitable relevance in this theory. The nonlattice property is related with the distribuition of the poles of the associated Zeta function (see [24]).

In [3], related results about quasi-stationary measures for horseshoe diffeomorphisms were obtained.

In [12] a good result for the wave equation associated with such billiards is presented.

Finally, we would like to point out that all the results (except that in § 8) stated for the billiard with three circles are true for the general case with several convex bodies satisfying condition (M). The proofs are essentially the same that we present below. We do not state the general case in the present paper for the sake of simplification of the notation.

1. The billiard map. Consider a finite number of closed curves δQ_i (where $Q_i, i = 1, 2, \ldots, s,\ s > 2$, are nonintersecting compact convex sets in the plane) that can be either $C^{r+1},\ r > 2$ with nonzero curvature, or real analytic. We will call this system the open billiard.

We will say that the open ball billiard satisfies condition (M) (Morita's and Ikawa's condition) if all curves are simple closed curves and the convex hull of $\delta Q_i \cup \delta Q_j$ does not intersect δQ_k for any triple of three distinct indices i, j, and k. We will assume that all the billiards we consider here satisfy the condition (M).

We will denote by δQ the union of all δQ_i, $i = 1, \ldots, s$, and by $n(q)$ the normal to the curve δQ at the point q. The normal will have norm one and point to the outside of the curve.

Consider the dynamical system describing the free motion of a point mass in the plane with elastic reflections on δQ (angle of incidence with the normal to the curve equal to the angle of reflection). The phase space of such a dynamical system is

$$M = \{(q, v);\quad q \in \delta Q,\quad |v| = 1,\quad \langle v, n(q)\rangle \geq 0\}.$$

A coordinate system is defined on M by the arc length parameter r along δQ (therefore, the state space in these coordinates has more than three connected components because $s > 2$) and the angle ϕ between $n(q)$ and v. Clearly $|\phi| \leq \pi/2$ and $\langle n(q), v\rangle = \cos(\phi)$.

Consider the probability $d\lambda = c\cos(\phi)drd\phi$, where $c = 2|\delta Q|^{-1}$ is just a normalizing factor and $|\delta Q|$ stands for the total length of δQ.

Now we define the transformation map f in the following way:

$$f(x_0) = f(q_0, v_0) = (q_1, v_1),$$

with q_1 the point of δQ (if there exists such a point) where the oriented line through (q_0, v_0) first hits δQ and v_1 the angle with the normal $n(q_1)$ made by that line after reflection on the tangent line through $q_1 \in \delta Q$. Formally, $v_1 = v_0 - 2\langle n(q_1), v_0\rangle n(q_1)$ (see Fig. 1). This transformation map f may not be defined for all $x_0 \in M$.

The measure λ is not globally invariant under f (any invariant measure is singular with respect to Lebesgue measure), but if one considers small neighbourhoods around x and $f(x)$, then the image of the measure λ by f is preserved. f is a C^r diffeomorphism in these small neighbourhoods. The Euclidean length t (or time) between q_0 and q_1 is denoted by t_0. Hence, $q_1 = q_0 + t_0 v_0$ (a trajectory inside the billiard travels with constant velocity equal to one).

The map f is called the billiard map. We are interested in analyzing trajectories with infinite bounces. The trajectories that do not have this property are those that in some finite (positive or negative) time escape to infinity.

We will denote by $x_i = (q_i, v_i) \in M,\ \ i \in \mathbf{N}$, the successive hits of a trajectory beginning at time 0, $x_0 = (q_0, v_0)$, with the boundary δQ, that is, $f(q_i, v_i) = (q_{i+1}, v_{i+1})$.

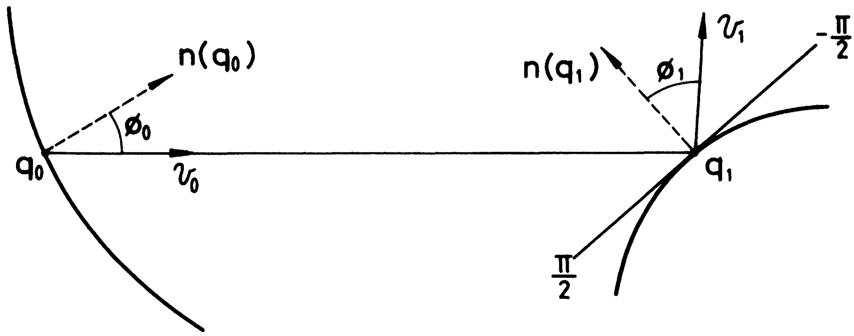

FIG. 1.

We are interested among other things in properties for trajectories with $x_0 = (q_0, v_0)$ in a set of full μ-measure (μ stands for the natural measure).

Given a trajectory beginning at $x_0 = (q_0, v_0) \in \delta Q$, we will denote by $K_i = K(x_i)$, $i \in \mathbf{N}$, the curvature of δQ at q_i. For instance, if one considers the model where all Q_i, $i = 1, \dots, s$, are disks, then the K_i are all constants. The angle between $n(q_i)$ and v_i will be denoted by ϕ_i, and finally, t_i denotes the Euclidean distance between the bounces q_i and q_{i+1}, $i \in \mathbf{N}$ (see Fig. 1). The backward orbit $x_i = (q_i, v_i)$, $i \in \mathbf{Z}$, is analogously defined. In any case, the main property is $f(x_i) = x_{i+1}$, $i \in \mathbf{Z}$.

In the case that we are considering, if f is defined for $x_0 = (q_0, v_0) \in M$, then it is also defined in an open neighbourhood of x_0 unless the trajectory through x_0 hits the image $f(x_0) = f(q_0, v_0) = (q_1, v_1) = x_1$ in a position tangent to δQ, that is, $v_1 = \pi/2$ or $v_1 = -\pi/2$. In this case f is defined in an left or right open neighbourhood. When we speak about neighbourhoods, we are considering any one of the possible cases described above. The set of points $x_0 = (q_0, v_0) \in M$ whose forward or backward trajectory is tangent to δQ for some x_i, $i \in \mathbf{Z}$, has λ-measure zero.

If $\tilde{x}_1 = (\tilde{q}_1, \tilde{v}_1) = f(\tilde{x}_0)$ is defined for $\tilde{x}_0 = (\tilde{q}_0, \tilde{v}_0)$, then for all $x_0 = (q_0, v_0)$ in a small neighbourhood of $\tilde{x}_0$ the derivative matrix is given by (see [27, 16])

$$f'(x_0) = \begin{pmatrix} \frac{t_0 K_0 + \cos\phi_0}{-\cos\phi_1} & \frac{t_0}{\cos\phi_1} \\ K_1 \frac{t_0 K_0 + \cos\phi_0}{\cos\phi_1} + K_0 & -\frac{K_1 t_0}{\cos\phi_1} - 1 \end{pmatrix}. \tag{1.1}$$

Note that when the image of (q_0, v_0) by f is tangent to δQ (that is, $q_1 = \pi/2$ or $q_1 = -\pi/2$), then the entries of the above matrix become infinity.

2. The open billiard with three circumferences. We now consider a particular example where the hypotheses of all results presented in this paper are satisfied. Consider three circular disks of radius one (Fig. 2) whose centers are located in the vertices of an equilateral triangle of side $a > 2$.

The more natural system of coordinates to consider in this problem is to denote by r the angle of the q coordinate in each circle.

In this case the phase space is given by three rectangles M_1, M_2, and M_3, where each one is a copy of a rectangle with base $0 \le r \le 4\pi/3$ and height $-\pi/2 \le \phi \le \pi/2$ (see Figs. 2 and 3).

We will denote by δQ_i the circle corresponding to the set M_i, $i = 1, 2, 3$.

As an example notice that the point $(\pi/2, 0) \in M_1$ is a periodic point with period 2, because $f(\pi/2, 0) = (5\pi/6, 0) \in M_2$ and $f(5\pi/6, 0) = (\pi/2, 0) \in M_1$. There exist several trajectories that are not periodic but have infinitely many bounces. The map

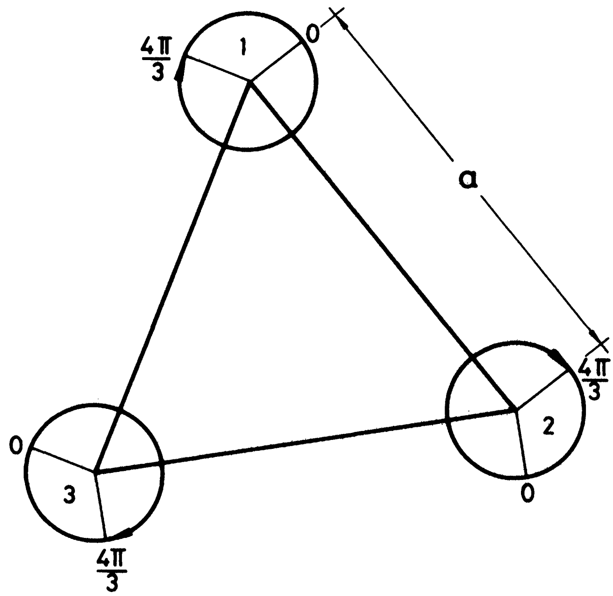

FIG. 2.

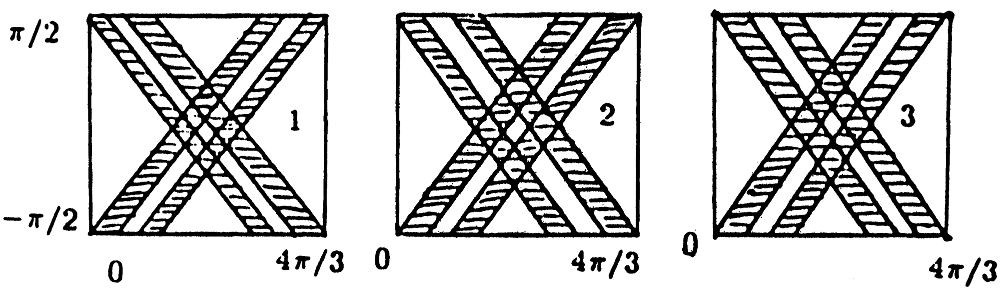

FIG. 3.

f is not defined everywhere (see Fig. 3); for example, it is not defined at the point $(4\pi/3, 0)$.

In fact the map f and its inverse f^{-1} are not defined outside the dashed region in Fig. 3. The horseshoe structure of the map f will be more carefully explained later.

Note that if $a \leq 2$, then the billiard is an example of a classical Sinai billiard, because different components of the boundary intersect with nonzero angle. The statistical properties of this kind of billiards have been studied extensively.

If $2 < a \leq 4/\sqrt{3}$, then it is easy to see that for such open billiards the condition (M) defined above is not satisfied. The case $a = 2$ is extremely interesting but will not be analyzed here.

The three-circle open billiard subject to the condition $a > 4/\sqrt{3}$ satisfies the condition (M), and it is under the assumptions of the theorems that we will prove later. It is the simplest example of such a class of open billiards. Apparently, the results that we present in the following sections can be also extended to the case $2 < a < 4/\sqrt{3}$. We will indicate why we believe that this is true (see the end of § 3).

The dynamics of f in the case $a > 4/\sqrt{3}$ are the same as of a shift of finite type. This can be seen as follows. Denote by π : domain of $f \to \{1, 2, 3\}$ the map that assigns to each $x = (q, v) \in M$ the value i such that $q \in \delta Q_i$. Given a certain sequence $\theta_i \in \{1, 2, 3\}$, $i \in \mathbf{Z}$, such that for any i, $\theta_i \neq \theta_{i+1}$, $i \in \mathbf{Z}$, there exists a unique $x_0 = (q_0, v_0)$ such that

$$\pi(f^n(x)) = \pi(q_n, v_n) = \theta_n.$$

It is also true that $\pi \circ f(x) = \sigma \circ \pi(x)$, where σ is a shift of finite type on three symbols $\{1, 2, 3\}$. In other words, π is a conjugacy of f with the shift σ. Therefore, the dynamics of f are that of a shift of finite type (remember that $\theta_i \neq \theta_{i+1}$, but this is the only restriction). This result was shown by Morita [20]. We will need to analyze metrical questions and therefore will need more delicate properties and estimates about the dynamics; the fact that f is conjugated to the shift σ is not enough. Among other problems, we will need to take special care when the entries of the matrix (1.1) become infinity due to tangencies of the orbit, etc.

Morita [20] also shows that the ceiling function $t(x)$ (the time between bounces) is Hölder continuous and nonlattice. We will consider here another potential ψ (different from t) that is natural in the setting we are working in. We will also show that for a dense set of values $a > 4/\sqrt{3}$, the potential ψ is not lattice. This allows one to estimate the growth number of periodic trajectories subject to weights, as in [24].

3. Trajectories with infinitely many bounces. Our first goal is to analyze geometrical and dynamical properties of the set of points that have infinitely many bounces in the past and in the future. This subset of M will have the structure of the product of two Cantor sets. We will begin considering the trajectories such that there exist infinitely many bounces in the future. We need therefore to analyze the set

$$\bigcap_{0 \leq j} f^{-j}(M_{i_j}), \quad i_j \neq i_{j+1}, \quad \forall j \in \mathbf{N}.$$

We will carefully analyze the case $a > 4/\sqrt{3}$, even if at the end of our reasoning, we will be able to indicate why we believe it is also true for $a > 2$.

From the symmetry of the problem, it follows that we have to analyze the structure of the set M_1 intersected with $\cap_{0 \leq j} f^{-j}(M_{i_j})$, where $i_0 = 1$, because for the other connected components M_2 and M_3 the structure is basically the same (we have of course to assume respectively that $i_0 = 2$ or $i_0 = 3$).

In Fig. 4, we represent some of the backward iterates.

Note that the line $\mathcal{A} = \{(r, -\pi/2), \ \pi/3 \leq r \leq \pi/2\} \subset M_1$ iterated by f goes on the curve $f(\mathcal{A}) \subset M_2$ shown in Fig. 4. The curve $f(\mathcal{A})$ can be also parametrized by r, given by the projection $(\phi, r) \to r$, over $\pi/2 \leq r \leq 4\pi/3$ (see Fig. 4). We draw two strips in M_1 corresponding to the preimages $f^{-1}(M_2)$ and $f^{-1}(M_3)$ in Fig. 4. There are also two other important strips, those corresponding to the images $f(M_2)$ and $f(M_3)$ in M_1 (see the first square in Fig. 3). In Fig. 4 we draw only the set $f(M_2)$ in order to make more clear the other curves and sets that we will describe in the sequel. The intersections of these four strips are four nonlinear rectangles in M_1 that correspond to the cylinders (with coordinates θ in the shift) $\{2, 1, 2\}$, $\{3, 1, 2\}$, $\{2, 1, 3\}$, and $\{3, 1, 3\}$.

Similar pictures can be drawn in M_2 and M_3. From this picture the reader can realize the horseshoe structure of the dynamics of f (see also Fig. 3). It is important to point out that the distortion could be very bad close to the boundaries, and this requires a more delicate analysis. In other words we need extra care with the almost

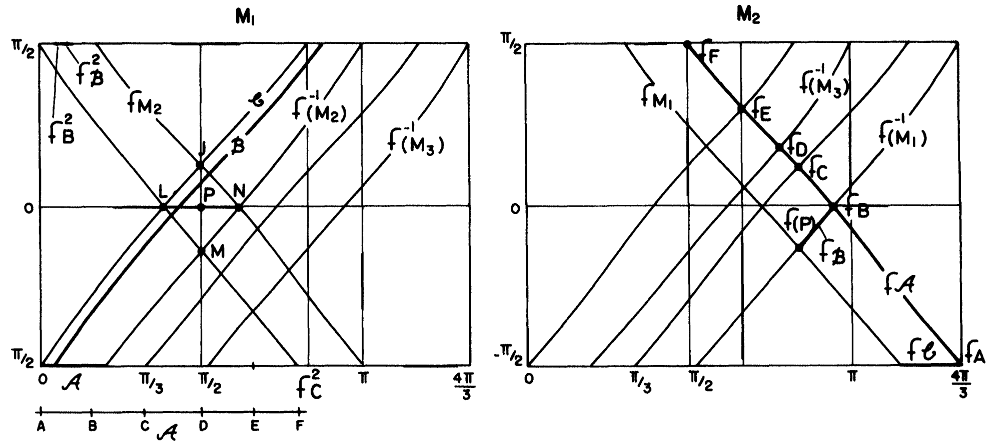

FIG. 4.

tangent trajectories because in this case the expanding properties are not as good. This question will appear in the following sections.

We draw the curve $\mathcal{A}$ in the left square of Fig. 4 and its image $f(\mathcal{A})$ in the right square of Fig. 4. To be more explicit about the dynamics of f we denote by A, B, C, D, E, and F points in the curve $\mathcal{A}$. Note the position of the images of these points in the set $f(\mathcal{A})$ in the right square in Fig. 4. Note also the curve $\mathcal{B}$ and its image $f(\mathcal{B})$ (see Fig. 4). The curve $\mathcal{C}$ represents positions (r_0, ϕ_0) whose image $f(r_0, \phi_0) = (r_1, \phi_1)$ will hit the circle 2 in a tangent position ($\phi_1 = -\pi/2$) (see Fig. 4).

The strip that appears in M_1 between the two strips $\{1, 2\}$ and $\{1, 3\}$ corresponds to the trajectories of M_1 that are lost in the middle of the two circles M_2 and M_3. The two other components in M_1, external to $\{1, 2\}$ and $\{1, 3\}$, correspond to the trajectories that are lost between M_1 and M_2 or between M_1 and M_3. The cylinders $\{1, 2, 1\}$, $\{1, 2, 3\}$, $\{1, 3, 1\}$, and $\{1, 3, 2\}$ correspond in M_1 to four strips contained in the two strips $\{1, 2\}$ and $\{1, 3\}$ (see Fig. 5). These four strips are strictly inside the two previous ones.

Inductively, the cylinders $\{1, i_1, i_2, \dots, i_n\}$, $i_j \neq i_{j+1}$, $j \in \{1, 2, \dots, n-1\}$, correspond to $\cap_{j=1,\dots,n} f^{-j}(M_{i_j})$ and are 2^n thin increasing strips going from the bottom to the top of M_1.

These cylinders form a nested sequence of sets (see Fig. 5). It is easy to see from a geometrical argument that each strip is strictly inside the previous one: note that for a fixed q in M_j, if one considers all possible angles ϕ, then this will determine images $f(q, \phi) = (q_1(\phi), \phi_1(\phi))$ in such a way that $q_1(\phi)$ is monotonical (when defined) and $\phi_1(\phi)$ also is monotonical. For a fixed q_0, as ϕ ranges from $-\pi/2$ to $\pi/2$, half a horizon will be covered by $f(q_0, \phi)$, and the part corresponding to hits in the other circles is strictly inside this half horizon (when observed from the point q_0). Clearly, the boundaries of the cylinder $\{1, i_1, i_2, \dots, i_n\}$ are curves that correspond to trajectories that at the nth bounce are tangent to δQ.

Note the important geometrical property presented in Fig. 4, showing how the set $\mathcal{A}$ goes by f into the curve $f(\mathcal{A})$. The boundary of $M_1 \cap f^{-1}M_2$ goes by f into the upper and lower boundary of M_2. The correct understanding of the geometrical position of all these boundaries and its images by f is essential for the following sections.

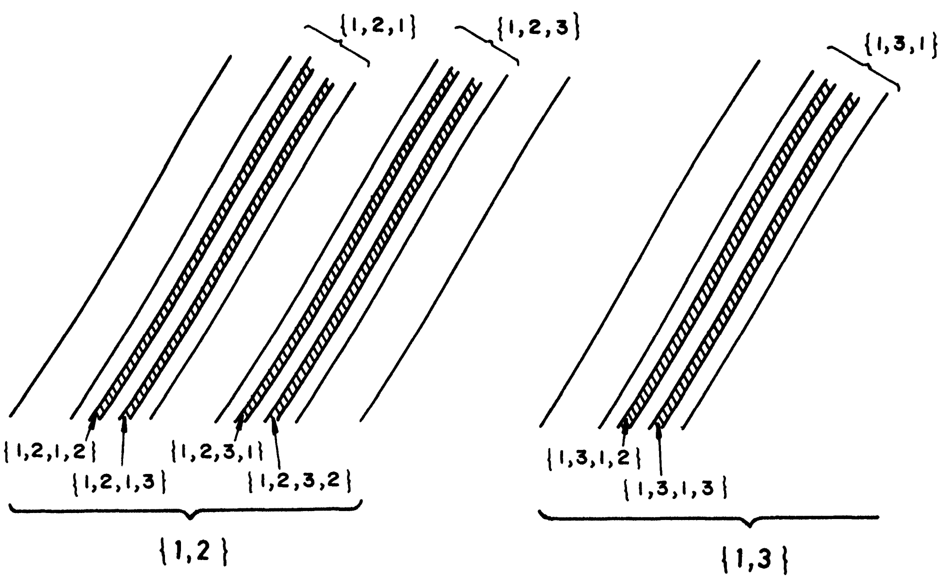

FIG. 5.

The intersection of an infinite sequence of nested sets is given generically by

$$\bigcap_{j=1}^{\infty} f^{-j}(M_{i_j}), \quad i_j \neq i_{j+1}, \quad j \in \mathbf{N},$$

and it is a curve coming from the bottom to the top of M_1. (In order to prove this property, which follows from expansiveness, we need to use an analytical expression that will be shown in §§ 4 and 5.) We will show finally that the union of all such possible nested sequences of sets can be parametrized as the product of a Cantor set by such curves.

The analysis that we have just made is valid for all open billiards satisfying condition (M). But in the case of three circumferences, it seems to be true also if $2 < a < 4/\sqrt{3}$. We will briefly discuss this case in the next paragraph. Note that in this situation there exist trajectories that are tangent to one disk, reflect at another disk, and then escape to infinity.

As we have seen before, Fig. 4 (case $a > 4/\sqrt{3}$) describes the general picture of the dynamics of f. The strip of points between $f^{-1}(M_2)$ and $f^{-1}(M_3)$ corresponds to points that will not hit circle 2 or 3 but will cross between these two circles. More less the same picture will be obtained for the boundary of the band of points x such that $f(x)$ escapes to infinity in the case $2 < a \leq 4/\sqrt{3}$. The difference is that in the present situation the two strips will collapse (see Fig. 6). Proceeding inductively, the trajectories that remain on the table for infinite iterations are in "distorted rectangles" in the same way as it happened in the case $a \geq 4/\sqrt{3}$ (see Fig. 4). In conclusion, the general picture of the case $a \leq 4/\sqrt{3}$ is basically the same as $a > 4/\sqrt{3}$ in topological terms.

4. Analytical expressions. We will now obtain the analytical expression of the differential equations satisfied by the invariant curves that generate the Cantor set, which were mentioned in the last paragraph.

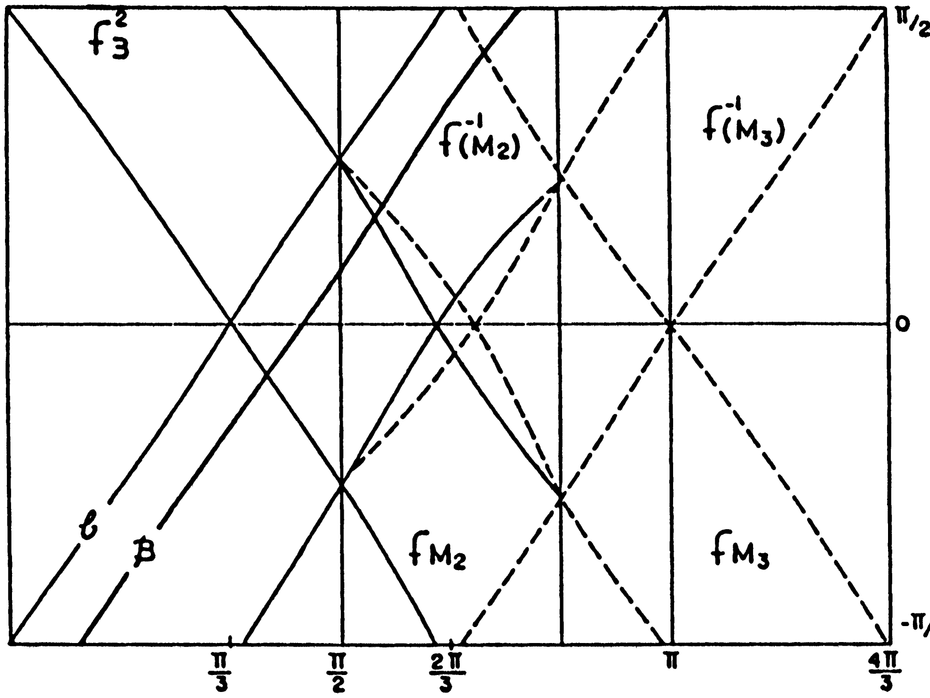

FIG. 6.

To illustrate our reasoning, we will first obtain the equation of the curve $\mathcal{B} \subset M_1$ through x such that $f(\mathcal{B}) \subset M_2$ and $f^2(\mathcal{B}) \subset M_1$, with $\phi_2 = \pi/2$, $\frac{d\phi_2}{dr} = 0$. (We are using the notation $f^i(x) = x_i = (r_i, \phi_i)$.) This curve $\mathcal{B}$ contains the 2-periodic point p_{121}. It follows from [27] (see also [4, 16]) that

$$\frac{d\phi_1}{dr_1}(x) = K_1(x) + \frac{\cos\phi_1(x)}{t_1(x)},$$

$$\frac{d\phi_0}{dr_0}(x) = K_0(x) + \cos\phi_0(x)\frac{1}{t_0(x) + \frac{1}{\frac{2K_1(x)}{\cos\phi_1(x)} + \frac{1}{t_1(x)}}}.$$

The last equation describes the parametrization ϕ_0 of $\mathcal{B}$.

Now, by induction, it follows that the boundary of the strips that successively appear when we remove the trajectories that go to infinity at time n, is given by

$$\frac{d\phi_0}{dr_0} = K_0 + \cos\phi_0 \frac{1}{t_0 + \frac{1}{\frac{2K_1}{\cos\phi_1} + \frac{1}{t_1 + \frac{1}{\frac{2K_2}{\cos\phi_2} + \frac{1}{\cdots + \frac{1}{\frac{2K_n}{\cos\phi_n} + \frac{1}{t_n}}}}}}}.$$

We omitted the reference to the point x in the above formula.

When n goes to infinity, the above equation will converge to the equation of the parametrization of the curve of points $y \in M_1$ with the same future specification of bounces θ_i, $i \in \mathbf{N}$, as x.

The continued fraction that appears multiplying $\cos\phi_0$ is given by

$$k^s(x) = \frac{1}{b_1(x) + \frac{1}{b_2(x)+\frac{1}{b_3(x)+\cdots}}} \tag{4.1}$$

with

$$b_{2k}(x) = \frac{2K(f^k(x))}{\cos\phi(f^k(x))} = \frac{2}{\cos\phi(f^k(x))}, \qquad b_{2k+1}(x) = t(f^k(x)). \tag{4.2}$$

This continued fraction converges if $K(f^k(x)) > 0$ and $\sum_{k=0}^{\infty} t(f^k(x)) = \infty$ (see [5, 20]). For the open billiard that we consider here, this is the case because $K(f^k(x)) = 1$ and $t(f^k(x)) > a - 2$ for all k. Therefore, the curves that in the future have infinitely many bounces are defined by the differential equation

$$\frac{d\phi}{dr}(x) = K(x) + k^s(x)\cos\phi(x).$$

We point out that this is also true for the billiards considered by Morita, when the obstacles are convex and the condition (M) is true.

We will use the notation

$$\frac{d\phi^s}{dr}(x) = K(x) + k^s(x)\cos\phi^s(x) \tag{4.3}$$

to emphasize that this differential equation determines the parametrization $\phi^s(r)$ in the variable r of the stable manifold $(r, \phi^s(r))$ through $x_0 = (r_0, \phi_0)$. Note that the differential equation is nonautonomous because we take derivatives in r but k depends on (r, ϕ).

In an analogous way one can show that the curve through x_0, given by the set of points (r, ϕ) with infinitely many bounces in the past (the unstable manifold passing through x), is parametrized by $(r, \phi^u(r))$, with $\phi^u(r)$ given by

$$\frac{d\phi^u}{dr}(x) = K(x) - k^u(x)\cos\phi^u(x), \tag{4.4}$$

where

$$k^u(x) = a_1(x) + \frac{1}{a_2(x) + \frac{1}{a_3(x)+\frac{1}{a_4(x)+\cdots}}}, \tag{4.5}$$

and

$$a_{2k+1}(x) = \frac{2}{\cos\phi(f^{-k}(x))}, \qquad a_{2k}(x) = t(f^{-k}(x)), \quad k \in \mathbf{N}.$$

5. The hyperbolic structure: Stable and unstable manifolds. Consider in the descending strip of type $\{1,2,1\}$ the unstable manifold of the 2-periodic point $p = p_{121} = (\pi/2, 0) = f^2(\pi/2, 0) \in M_1$. The stable manifold is given by

$$\gamma^s(p) = \{z; \pi(f^{2n}(z)) = 1, \quad \pi(f^{(2n+1)}(z)) = 2, \quad \forall n \in \mathbf{N}\},$$

and the unstable manifold through p is given by

$$\gamma^u(p) = \{z; \pi(f^{-2n}(z)) = 1, \quad \pi(f^{-(2n+1)}(z)) = 2, \quad \forall n \in \mathbf{N}\}.$$

More generally, consider the 2-periodic points p_{iji} in M_i, $i \neq j$, $i, j \in \{1,2,3\}$; there is a total of six such periodic points of period 2.

Unstable manifolds are defined by graphs of decreasing functions and stable manifolds are described by graphs of increasing functions. This follows from the inclination of the parametrizations ϕ given by the analytical expressions (4.3) and (4.4) of the differential equations described in § 4.

Let $\gamma^u(p_{iji})$ be the unstable manifold through p_{iji} intersected with the set M_i and

$$\gamma^u_{ij} = \gamma^u(p_{iji}) \cap f^{-1}(M_j). \tag{5.1}$$

Note that the curve $\gamma^u(p_{iji})$ goes from the bottom to the top of M_i, but for γ^u_{ij} this is not true.

Denote by $M' = \cup_{i,k\neq j} M_{ijk}$ the union of the twelve quadrilaterals, where $M_{ijk} = f(M_i) \cap M_j \cap f^{-1}(M_k)$. The dynamics of the trajectories that do not go to infinity can be studied in M'. These quadrilaterals are far away from $\phi = \pm\pi/2$, and hence, for $x \in M'$, $\cos\phi(x) > c_1 > 0$.

Now we define the p-length of a general curve $\gamma \subset M$ by

$$p(\gamma) = \int_\gamma \cos\phi dr. \tag{5.2}$$

More precisely, if γ is defined by $(r, \phi(r))$, $r_0 \le r \le r_1$, then

$$p(\gamma) = \int_{r_0}^{r_1} \cos\phi(r) dr.$$

If γ is any decreasing curve ($\phi'(r) < 0$), and f is continuous in γ, then

$$p(f(\gamma)) = \int_{r_0}^{r_1} \left(\frac{t(r)(K(r) - \phi'(r))}{\cos\phi} + 1 \right) \cos\phi dr, \tag{5.3}$$

where $t(r) = t(x)$ is the distance to the next bounce beginning at $x = (r, \phi(r))$. Since $p(\gamma)$ is of order $\cos\phi_0 dr$ (for small γ), for small γ passing through $x_0 = (r_0, \phi_0)$,

$$\frac{p(f(\gamma))}{p(\gamma)}$$

is approximately equal to

$$1 + \frac{t(r_0)(K(r_0) - \phi'(r_0))}{\cos\phi(r_0)}$$

with $x_0 \in \gamma$.

This property will lead us to define a kind of partial derivative $\delta f^p_\gamma(x_0)$ using the p-length defined above.

DEFINITION 1. *Given a curve γ through x_0, we define* the p-derivative of γ at x_0 *as the limit*

$$\delta f^p_\gamma(x_0) = \lim_{p(\gamma)\to 0} \frac{p(f(\gamma))}{p(\gamma)}.$$

For decreasing curves parametrized by $(r, \phi(r))$, the p-derivative of γ at x_0 is given by

$$\delta f^p_\gamma(x_0) = 1 + \frac{t(r_0)(K(r_0) - \phi'(r_0))}{\cos\phi(r_0)}. \tag{5.4}$$

Under the hypothesis considered here, the p-derivative of decreasing curves γ given by the last expression is uniformly bounded below by $1+t_{\min}$, where the $t_{\min} = a-2$ is the minimum of the distances between bounces.

Analogously, for the increasing curves γ parametrized by $(r, \phi(r))$, the p-derivative of γ on $x_0 = (r_0, \phi_0)$ is given by

$$\delta f_\gamma^p(x_0) = \frac{1}{1 + \frac{t(r_0)(K(r_1)+\phi'(r_1))}{\cos\phi_1}}. \tag{5.5}$$

In $f^{-1}(M_1)$, any increasing curve γ satisfies $0 < \lambda < \delta f_\gamma^p(x_0) < \frac{1}{1+t_{\min}} < 1$. From (4.4) it follows that for γ^u in M' we have

$$\delta f_{\gamma^u}^p(x_0) = 1 + t(x_0)k^u(x_0) > 1 + t_{\min}a_1(x_0) > 1 + 2t_{\min} = w > 1,$$

and from (4.3) it follows that $\delta f_{\gamma^s}^p < 1/w$. In conclusion, from the above reasoning it follows that there exist $K > w > 1$ and $\lambda < 1/w$ such that such for all x_0 in M'

$$w < \delta f_{\gamma^u}^p(x_0) < K \tag{5.6}$$

and

$$\lambda < \delta f_{\gamma^s}^p(x_0) < \frac{1}{w}. \tag{5.7}$$

These estimates will be important later.

From these last properties ((5.6) and (5.7)) and the way in which the Cantor set structure of the nonwandering set appears (see § 3), we can say that the dynamics of f is that of a horseshoe diffeomorphism. Therefore, all the considerations in chapter 2 of [23] can be applied, and we conclude that there exist $C^{1+\epsilon}$ foliations of stable and unstable manifolds around the nonwandering set

$$\Gamma = \bigcup_{i\neq j,k} \bigcap_{l\in Z} f^l(M_{ijk}).$$

It follows easily (see [23, Chap. 2]) that the projection along stable (and unstable) leaves is $C^{1+\epsilon}$. This property explains why we will need in the future a $C^{1+\epsilon}$ version of the results of class C^2 that were previously obtained by other authors [6, 25].

6. Expanding transformations and invariant measures. We will state in this section the $C^{1+\epsilon}$ results that we will need in § 7. These results will be proved in the appendix.

A piecewise continuous map T is transitive on components if for every two maximal sets B, C where T is continuous there exists $n = n(B,C) \in \mathbf{N}$ such that $T^nB \cap C \neq \emptyset$.

We will say that a probability measure μ, defined on the elements of a σ-algebra $\mathcal{A}$ of A, is conditionally invariant with respect to $T : A \to TA$ if $\mu(T^{-1}C) = \alpha\mu(C)$ for every element $C \in \mathcal{A}$ for some positive constant α.

It results $\alpha = \mu(T^{-1}A)$. Hence μ is conditionally invariant if and only if

$$\frac{\mu(T^{-1}(C) \cap T^{-1}(A))}{\mu(T^{-1}(A))} = \mu(T^{-1}C|T^{-1}A) = \mu(C).$$

This implies that $\alpha^n = \mu(T^{-n}A)$ for every $n \geq 0$.

We will represent by μ_F the probability measure $d\mu_F = Fd\nu$, where ν is another fixed probability measure on A and $\int_A Fd\nu = 1$.

HYPOTHESIS A. *Assume* $T : \bar{A} \to \mathbf{R}$, $B = A \cap T^{-1}A$, *is such that*

a) $A = \bigcup_{i=1}^{k} A_i$, *where* A_i *are disjoint open intervals;*

b) $A \subset T(A)$ *(strictly);*

c) $A \cap T(\partial A) = \emptyset$;

d) $\bar{A}$ *is endowed with some metric* d *such that the derivative* T_d *of* T *with respect to this metric is well defined on* B*; i.e., there exists*

$$T_d(x) = \lim_{y \to x} \frac{d(Ty, Tx)}{d(y, x)} \tag{6.1}$$

for every $x \in B$*;*

e) T_d *is* γ*-Hölder continuous on* B*; i.e., there exist* $k > 0$ *and* $0 < \gamma < 1$ *such that* $|T_d(x) - T_d(y) \le kd^{\gamma}(x, y)$ *for every* x, $y \in B$*;*

f) *there exist* $M > \beta > 1$ *such that* $\inf\{T_d(x) : x \in B\} \ge \beta$ *and* $\sup\{T_d(x) : x \in B\} \le M$*;*

g) $T_{|\bar{A}_i}$ *is an homeomorphism for every* $i = 1, \dots, k$.

Let ν be the probability measure induced by the metric d on Borel sets of A, $\beta_1 = [1/\beta]^{\gamma} < 1$, $\mathcal{P} = \{\bar{A}_i\}_{i=1}^{k}$, and $\mathcal{P}_n = \bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}$.

LEMMA 1. *There exists a constant* $k_1 > 0$ *such that*

a) $|(T^n)_d(x) - (T^n)_d(y)| \le \beta_1 M^{n-1} k_1$,

b)

$$\frac{(T^n)_d(z)}{(T^n)_d(w)} \left[\frac{d(z, w)}{d(T_z^n, T_w^n)} \right]^{\gamma} \le k_1 \beta_1^n$$

for every z, $w \in I \in \mathcal{P}_n$.

Proof. a) T_d satisfies the chain rule for derivatives. Then

$$\begin{aligned} |(T^n)_d(x) - (T^n)_d(y)| &\le M^{n-1} \sum_{l=0}^{n-1} kd^{\gamma}(T_x^l, T_y^l) \\ &\le M^{n-1} k \sum_{l=0}^{n-1} \beta^{n-l} d(T_x^n, T_y^n) \le M^{n-1} \beta_1 k_1. \end{aligned}$$

b) The bounded distortion property for expanding maps establishes that

$$(T^n)_d(z)/(T^n)_d(w) \le k_1$$

(see, for example, [23, § 4.1]).

From b) of Lemma 1, we can choose $N \ge 1$ such that $k_1 \beta_1^N < 1$. Since all our results can be written in terms of T^N instead of T (subdividing A_i), from here we will suppose that T satisfies the last inequality for $N = 1$.

The proof of the next two theorems for the case $C^{1+\epsilon}$ will be made in the appendix.

THEOREM 1 (see Pianigiani and Yorke [25] for the case C^2). *Let* $T : \bar{A} \to \mathbf{R}$ *satisfy Hypothesis* A a)–g). *Then*

i) *there exists a* γ*-Hölder continuous function* $F : A \to \mathbf{R}$,

$$F \in \mathcal{K} = \left\{ G \in C^0(A), \inf_{x \in A} G > 0, \sup_{x \in A} G < \infty, \int G d\nu = 1 \right\}$$

such that μ_F *is absolutely continuous with respect to the measure* ν *and conditionally invariant with respect to* T;

ii) *if T is also transitive on components, then there exists a unique $F \in \mathcal{K}$ such that μ_F is conditionally invariant with respect to T;*

iii) *if, furthermore, T^n is transitive on components for all $n \in \mathbf{N}$, then for every $g \in \mathcal{K}$*

$$\lim \frac{P_1^n g}{\|P_1^n g\|_1} \to F.$$

Here $\| \ \|_1$ means the L^1 norm on $L^1(A,\nu)$ and $P_1 : L^1(A,\nu) \to L^1(TA,\nu)$ is the Perron–Frobenius operator defined by

$$P_1 g(x) = \sum_{y:T_y=x} (T_d(y))^{-1} g(y) = \frac{d(\mu_g \circ T^{-1})}{d\nu}. \tag{6.2}$$

We consider $P_1 : L^1(A,\nu) \to L^1(A,\nu)$ by taking the restriction of $P_1 f$ to A. Then $\int_A f.(g \circ T) d\nu = \int_{TA} (P_1 f) g d\nu$ for $f \in L^1(A,\nu)$, $g \in L^1(TA,\nu)$, and

$$\int_{T^{-n}(A)} f.(g \circ T^n) d\nu = \int_A (P_1^n f) g d\nu$$

for every $f, g \in L^1(A,\nu)$.

Note that $P_1 F = \alpha F$ if and only if μ_F is conditionally invariant with $\alpha = \mu_F(T^{-1}A)$. (This follows from the last expression taking $f = F$ and $g = 1$.)

We now define the operator $Q : L^1(A,\nu) \to L^1(A,\nu)$ by

$$Qg(x) = [\alpha F(x)]^{-1} P_1(gF)(x), \tag{6.3}$$

where $\alpha = \mu_F(T^{-1}A) = \int_{T^{-1}(A)} F d\nu$. Since $P_1 F = \alpha F$, we have $Q1 = 1$.

The reader familiar with thermodynamic formalism (see [24]) will recognize the operator Q as the Ruelle–Perron–Frobenius operator obtained from the potential

$$\log \frac{|T_d|^{-1}(x) F(x)}{\alpha F(T(x))}.$$

This potential is cohomologous to the potential $-\log |T_d(x)|$. The procedure of defining Q by (6.3) above is usual in thermodynamic formalism when one knows the eigenfunction F and the eigenvalue α. This procedure is sometimes called normalization of the operator.

We refer the reader to [24], where the theory of thermodynamic formalism developed initially by Bowen, Ruelle, and Sinai is carefully described.

In terms of the variational problem of the pressure the two cohomologous potentials will determine the same Gibbs state.

The reader should take care with the different domains where the two operators are defined: the Perron–Frobenius operator of Lasota–Pianigiani–Yorke is defined over L_1 functions, and the Ruelle–Perron–Frobenius operator of thermodynamic formalism is defined over Hölder-continuous functions. The most surprising property of the Pianigiani–Yorke result is the existence of the derivative of F in a full neighbourhood of the Cantor set under the C^2 hypothesis. Under the $C^{1+\epsilon}$ hypothesis, we will show in the appendix that F will be Hölder continuous.

Now we will need another result related to Theorem 1.

THEOREM 2 (see Collet, Martinez, and Schmitt [6] for the C^2 case). *Let* $T : \bar{A} \to \mathbf{R}$ *satisfy Hypothesis* A a)–g) *and suppose that* T^n *is transitive on components for all* $n \in \mathbf{N}$. *Then*

i) $Q^n g(x) \to \mu(g)$ *for every* γ*-Hölder-continuous function* g *on* A. $\mu(g)$ *defines a probability measure* μ, *with support on* $K = \bigcap_{n\geq 0} T^{-n}A$;

ii) *the conditional probability measure of staying in* A, *when the evolution occurs with probability* μ_F, *is* $\mu_F(C|T^{-n}A) \to \mu(C)$ *when* $n \to +\infty$ *for every Borel set* $C \subset A$;

iii) μ *is Gibbsian with potential* $-\log T_d(x)$, *i.e.,*

$$c^{-1}[(T^n)_d(z)]^{-1}\alpha^{-n} \leq \mu\left[\bigcap_{l=0}^{n-1} \bar{T}^{-l}(\bar{A}_{i_l})\right] \leq c[(T^n)_d(z)]^{-1}\alpha^{-n}$$

for every $i_0, \dots, i_{n-1} \in \{1, 2, \dots, k\}$, *every* $n \in \mathbf{N}$, *and some* $z \in \bigcap_{l=0}^{n} T_i^{-l}\bar{A}_l$.

So $(A, T, \mathcal{A}, \mu)$ *is a Kolmogorov system and satisfies the property of exponential decay of correlations, and*

$$\log\alpha = h_\mu(T) - \int_A \log T_d(x)d\mu(x) = \sup\left\{h_\eta(T) - \int \log T_d(x)d\eta(x) : \eta \text{ is an invariant probability measure}\right\},$$

where $h_\eta(T)$ *is the entropy of* T *with respect to* η.

Now we will make some comments about different properties claimed by the above theorem.

The precise meaning of the limit in property i) will be explained later in the appendix. The conditions above allows one to apply the Riesz theorem, defining in this way a probability μ such that $\mu(g) = \int g(x)d\mu(x)$. The measure μ is invariant for T, and therefore the support of μ is the nonwandering set of T (having in this case a Cantor set structure on the line). The property ii) is the more important one. It claims that if we calculate $\mu_F(V|T^{-n}(A))$, the part of $V \cap T^{-n}(A)$ in $T^{-n}(A)$ ($T^{-n}(A)$ is the subset of A that still remains in A after n iterations), then when n goes to infinity, the system will determine in the limit a certain measure $\mu(V)$. The analogy of the natural measure that we mention before and the measure μ we just defined (and satisfying property ii)) is transparent.

Property iii) is also very important because a Gibbsian measure has several nice properties: the system is Kolmogorov (therefore ergodic), there exists exponential decay of correlation, and so on (see [24]).

Both theorems can be formulated for $T : \bar{A} \to \mathbf{R}^n, A = \cup A_i$, where $A_i \subset \mathbf{R}^n$ are disjoint connected uniformly arcwise-bounded sets. This means that there exists a number b such that any two points in each A_i can be joined by a polygonal line of lengt at most b (see [24]).

7. Measures for open billiards: Invariant and conditionally invariant. Now we will return to the considerations of § 5 and show how the results of § 6 can be applied to the open billiard.

Note, for example, that for the point $(\pi/2, 0) \in M_1$, $f(\pi/2, 0) = (5\pi/6, 0) \in M_2$ and $f(5\pi/6, 0) = (\pi/2, 0) \in M_1$. Therefore the 2-periodic orbit $(\pi/2, 0) \in M_1$ and $(5\pi/6, 0) \in M_2$ has an unstable manifold with two components. Since there exist three pairs of 2-periodic orbits, we will consider six small pieces of unstable curves around these six periodic points.

Formally, let p_{iji} the 2-periodic point such that $p(f^{2n}(p_{iji})) = i$ and $p(f^{2n+1}(p_{iji})) = j$, where $n \in \mathbf{Z}$. The local unstable manifold of p_{iji} is defined by

$$\gamma^u(p_{iji}) = \{z \in M_i; p(f^{-2n}(z)) = i, \quad p(f^{-(2n+1)}(z)) = j, \quad \forall n \in \mathbf{N}\}.$$

Let us write $\gamma_{ij} = \gamma^u(p_{iji}) \cap f^{-1}(M_j)$; therefore, as we have seen in (5.1),

$$f(\gamma_{ij}) = \gamma^u(p_{jij}).$$

Note that the image of each one of the γ_{ij} six small pieces of unstable manifold is the full unstable manifold (from bottom to the top) through another 2-periodic point.

Denote also by $\Pi^s_{jk} : M_j \cap f^{-1}(M_k) \to \gamma_{jk}$ the projection along stable fibers; as we mentioned before, this projection is $C^{1+\epsilon}$. (This is the reason for the need of $C^{1+\epsilon}$ theorems in the present paper.)

THEOREM 3. *Consider the system described in* § 2. *Let A be the set $\cup_{i\neq j}\gamma_{ij}$. We define T on $\bar{A}$ as a continuous extension of its values on the* 12 *connected pieces of curves $\gamma_{ij} \cap f^{-2}(M_l)$, $l \neq j$. On these curves, T is defined by $T(x) = f(x)$ if $f(x) \in \gamma_{ji}$, and $\Pi^s_{jk} f(x)$ if $f(x) \in f^{-1}(M_k)$, $k \neq i$.*

Then $T : \bar{A} \to T(\bar{A})$ satisfies Hypothesis A *and also the hypotheses of Theorems* 1 *and* 2.

Remark 1. Note that A is now a union of pieces of curves in $\mathbf{R}^2$ and not a union of intervals in $\mathbf{R}$ as in Theorems 1 and 2, but the proof of the analogous result is the same.

Remark 2. To be more precise, we will need to consider f^N, a high iterate of f, as having the hypotheses of Theorems 1 and 2 satisfied, but this is no problem for our purposes, as will be explained later.

In Fig. 7, we represent schematically the graph of T.

Proof of Theorem 3. The verification of conditions a) and b) follows immediately from the definition. Condition c) follows from § 5.

Condition d) can be seen as follows: let $d(x, y) = p(\gamma)$, where γ is the curve contained in γ_{ij} which joins $x, y \in \gamma_{ij}$. Recall that if $T'(x_0)$ is the rate of expansion under the Euclidean norm ($dl = \sqrt{dr^2 + d\phi^2}$) of f at x_0 on unstable directions, then

$$|T'(x_0)| = \delta f^p_{\gamma^u} \frac{\cos\phi(x_0)}{\cos\phi(f(x_0))} \left(\frac{1 + h^2(f(x_0))}{1 + h^2(x_0)}\right)^{-1/2}, \tag{7.1}$$

where $h(y) = \frac{d\phi^u}{dr}(y)$ (see, for instance, § 5 in [4]). Note that $\log|T'|$ and $\log \delta f^p_{\gamma^u}$ are cohomologous. Condition d) is now a consequence of the following considerations: $T'_d(x_0)$ is either $\delta f^p_{\gamma^u}(x_0)$ or $[\delta f^p_{\gamma^u}(x_0)][\delta(\Pi^s_{jk})^p_{\gamma^u}(f(x_0))]$.

Now comes the crucial point: Π^s_{jk} is a $C^{1+\epsilon}$ function in the Euclidean metric, and this metric is equivalent to the p-metric on unstable manifolds because these are not too close to the vertical lines and therefore $\cos\phi$ is bounded away from zero.

T is an expanding map if

$$\min\{\delta(\Pi^s_{jk})^p_{\gamma^u}(y);\ y \in M_j \cap f^{-1}(M_k)\} = m > 1/w.$$

If this condition is not satisfied, we must consider f^N instead of f, with $N \in \mathbf{N}$ such that $w^N m > 1$. Note that m is positive because $\Pi^p_{\gamma^u}$ is a diffeomorphism.

The topological mixing property included in the definition of transitivity is satisfied because of the considerations made at the end of § 3 about the angles varying monotonically and covering half horizons.

Therefore all the conditions listed above are true for our system.

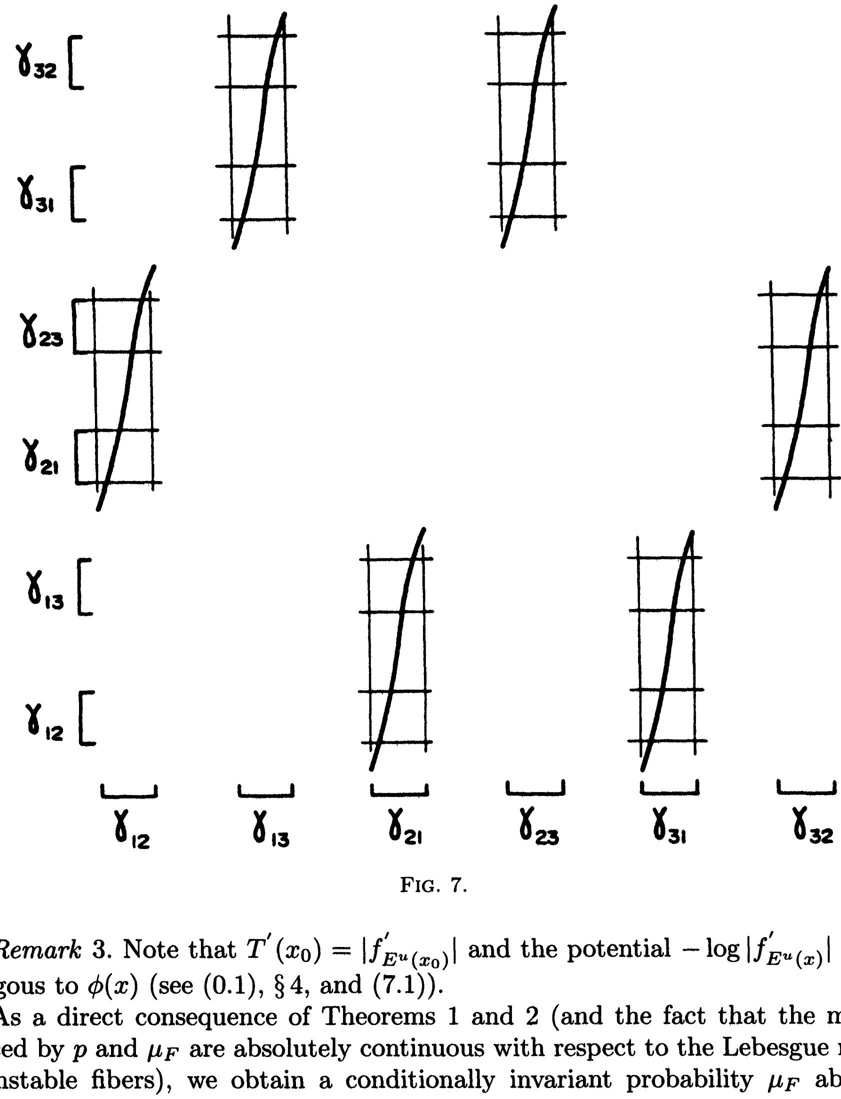

FIG. 7.

Remark 3. Note that $T'(x_0) = |f'_{E^u(x_0)}|$ and the potential $-\log|f'_{E^u(x)}|$ is cohomologous to $\phi(x)$ (see (0.1), § 4, and (7.1)).

As a direct consequence of Theorems 1 and 2 (and the fact that the measures induced by p and μ_F are absolutely continuous with respect to the Lebesgue measure on unstable fibers), we obtain a conditionally invariant probability μ_F absolutely continuous with respect to the Lebesgue measure on A, with density F a positive Hölder-continuous function.

Furthermore, from Theorem 2, there exists a measure μ_1 such that for any Borel set $V \subset A$,

$$\lim_{n\to\infty} \frac{\mu_F(T^{-n}(A) \cap V)}{\mu_F(T^{-n}(A))} = \mu_1(V),$$

where μ_1 is Gibbsian with potential $-\log|T'_d(x)|$. The support of μ_1 is the Cantor set $K_1 = \cap_{n=0}^{\infty} T^{-n}(A)$, the intersection with A of the set of points whose trajectories have infinitely many bounces in the future (do not escape to infinity). μ_1 is an invariant measure for T.

In an analogous way, we can apply to f^{-1} the same reasoning we did before for f.

Consider $C = \cup_{i\neq j}\gamma^s_{ij}$; then applying Theorems 1 and 2 to f^{-1} on C, we are able to find a Hölder-continuous function G defined on C such that μ_G is conditionally invariant. More precisely, let S denote the induced map for f^{-1}, using projection along unstable fibers on the stable manifolds of periodic points of period two; then there exists β such that μ_G satisfies $\mu_G(S^{-1}(D)) = \beta\mu_G(D)$ for every Borel set D. It also follows from Theorem 2 that the measure μ_G conditionally converges to an S-invariant measure μ_2 whose support is in $K_2 = \cap_{n=0}^{\infty} S^{-n}(C)$.

Now we will construct the natural two-dimensional measure for the open billiard problem. Remember that $M' = \cup_{i,k\neq j} M_{ijk}$.

THEOREM 4. *Consider the system $f : M' \to M'$ described in §§2 and 5. Then there exists a conditionally invariant positive measure μ^+ and a measure μ_1^+ such that*

$$\lim_{n\to\infty} \mu^+(V|f^{-n}(M')) = \mu_1^+(V)$$

for every Borel set $V \subset M'$. The measure μ_1^+ is invariant under f, supported on $K_1 \times K_2$, and (M', f, μ_1^+) is a K-system.

Proof. We begin by constructing the measure μ^+ on M' that extends the 1-dimensional measure μ_F. First of all we define a probability measure μ_0 over the σ-algebra $\mathcal{B} = (\Pi_{ij}^s)^{-1}(\mathcal{A})$, where $\mathcal{A}$ is the Borel σ-algebra on $\bar{A}$. For a set $D \in \mathcal{B}$ we define $\mu_0(D) = \mu_F(\Pi_{ij}^s(D))$.

Define the measure μ_n on $f^n(\mathcal{B})$ by $\mu_n(E) = \alpha^{-n}\mu_0(f^{-n}(E))$. It is easy to see that $f^{-1}(\mathcal{B}) \subset \mathcal{B}$ if we restrict the range of f^{-1} to M'. Therefore $f^n(\mathcal{B}) \subset f^{n+1}(\mathcal{B})$ for every $n \in \mathbf{N}$, and we conclude that $\mu_{n+1}(D) = \mu_n(D)$ holds for $D \in f^n(\mathcal{B})$.

This last equality allows one to define a finitely additive measure μ_∞ on the algebra $\uplus_{0\leq n} f^n(\mathcal{B})$ by $\mu_\infty(D) = \mu_n(D)$ if $D \in f^n(\mathcal{B})$. Note that $\cup_{0\leq n} f^n(\mathcal{B})$ is an algebra because if $D \in f^n(\mathcal{B}), E \in f^m(\mathcal{B}), m \leq n$, then $D \cap E \in f^n(\mathcal{B})$ and $\mu_\infty(D\cap E) = \alpha^{-n}\mu_0(f^{-n}(D\cap E))$. The measure μ_∞ satisfies $\mu_\infty(f^{-1}(C)) = \alpha\mu_\infty(C)$ because if $C \in f^n(\mathcal{B})$, then $f^{-1}(C) \in f^{n-1}(\mathcal{B})$ and

$$\begin{aligned}\mu_\infty(f^{-1}(C)) &= \mu_{n-1}(f^{-1}(C)) = \alpha^{-n+1}\mu_0(f^{-n+1}f^{-1}(C)) \\ &= \alpha^{-n+1}\mu_0(f^{-n}(C)) = \alpha^{-n+1}\alpha^n\mu_n(C) = \alpha\mu_\infty(C).\end{aligned}$$

The rest of the construction is exactly the same as that for an Anosov system (see, for example, [14, Chap. III, Thm. 2.3]). Therefore, the measure μ^+ on M' is conditionally invariant: $\mu^+(f^{-1}(D)) = \alpha\mu^+(D)$ for every Borel set $D \subset M'$.

Now we will analyze the limit of the conditioned measure. If $D \in f^k(\mathcal{B})$ for some fixed $k \in \mathbf{N}$, then as n goes to infinity,

$$\begin{aligned}\frac{\mu^+(D \cap f^{-n}(M'))}{\alpha^n} &= \frac{\mu_0(f^{-k}(D \cap f^{-n}(M'))}{\alpha^{n+k}} \\ &= \frac{\mu_F(\Pi^s f^{-k}(D \cap T^{-n}(A))}{\alpha^{n+k}} \to \mu_1(T^{-k}(\Pi^s(D))).\end{aligned}$$

If μ_1^+ is the measure constructed on M' by extending μ_1 (following the same procedure we used to construct μ^+ given μ_F), we have proved that

$$\lim_{n\to\infty} \mu^+(D|f^{-n}M') = \mu_1^+(D)$$

for all Borel sets D in M'. μ_1^+ is an invariant probability measure whose support is contained in $K_1 \times K_2$. Since μ_1 is Gibbsian, (A, T, μ_1) is a Kolmogorov system, and therefore the same is true for (M', f, μ_1^+) (see [14]).

The same procedure applied to the stable conditionally invariant probability μ_G allows one to construct a measure μ^- on M' such that $\mu_-(f(D)) = \alpha\mu_-(D)$. In the case that we analyze here, we cannot compare the two measures μ^+ and μ^-, but for

Anosov systems it is possible to show that the two measures are equivalent to the Lebesgue measure in $\mathbf{R}^2$ (see [14, 13]).

Now considering $S : C \to T(C)$ and f^{-1} instead of, respectively, T and f, we obtain in a similar way the Kolmogorov system (M', f, μ_2^-). The support of μ_2^- is contained in $K_1 \times K_2$.

The set of trajectories in M having infinite bounces in the past and in the future is $K_1 \times K_2 = \Gamma$ (see § 5).

The dynamical system (M, f, μ_1^+) is ergodic, and one can apply the formula for computing the entropy of the f-invariant probability μ_1^+ (see [11, 12]):

$$h = \delta \chi^+. \tag{7.2}$$

In this way, we are able to obtain the measure theoretical entropy h of f with respect to μ_1^+ in terms of δ, the Hausdorff dimension of the transverse measure μ_1^+ and χ^+, the Liapunov exponent of the measure μ_1^+. The Pesin formula is a similar expression (not involving dimension), but for the case when the natural probability is equivalent to Lebesgue measure (see, for instance, [14]).

We point out that from (7.1), χ_μ, the integral of $\log |T'| = \log |f'_{|E_u}|$ with respect to the invariant measure μ, is equal to $-\int \psi(x) d\mu(x)$. (The two potentials are cohomologous.)

Note that the Hausdorff dimension of the measure μ (Gibbsian for $-\log |T'|$) has nothing to do with the Hausdorff measure of the nonwandering set. The Hausdorff measure of the nonwandering set of a one-dimensional expanding system T has a density with respect to the Gibbsian measure of the potential $-s \log |T'|$, where s is the Hausdorff dimension of the nonwandering set (see remarks in [13]).

8. The nonlattice property of the potential $\boldsymbol{\psi}$. This section is the only one that we really need to assume that the billiard is given by three circles with center in the vertices of a equilateral triangle. The results in other sections are true for general billiards satisfying condition (M).

From the introductory section, we recall that by definition a potential B is lattice if there exist an integer-valued function G, γ a real positive constant, and a continuous function g such that $B = g \circ T - g + G\gamma$.

When one wants to prove asymptotic growth-rate properties of the periodic orbits (see [24]), the results (and proofs) are different for the lattice and nonlattice potential. It is possible to obtain such properties by means of Tauberian theorems combined with Fourier series arguments (in the case of lattice potentials) or Fourier transforms arguments (in the case of nonlattice potentials). The lattice potentials appear only in very special situations. One should expect that in general the potentials that occur in mathematical problems are nonlattice. This is the claim of the main theorem of the present section.

In this section we will show that for the billiard given by three circles of radius one centered at the corners of an equilateral triangle with side a, the Lyapunov exponent potential is not lattice for a dense set of possible values a. This claim is equivalent to showing that the potential ψ defined before is not lattice, because these potentials (up to a minus sign) are cohomologous as shown in (7.1).

THEOREM 5. *Consider the system $f : M' \to M'$ described in §§ 2 and 5. For a dense set of parameters $a > 4/\sqrt{3}$, the potential ψ is nonlattice.*

From this result we obtain the asymptotic growth of $(f^{k'})(x)$ on periodic orbits $f^k(x) = x$, which was mentioned in the introductory section. Indeed, the nonlattice property for ψ is equivalent to the weak mixing property of the suspension flow f_ψ (see [24]) defined on the suspension space $M_\psi = \{(x, y) : x \in M', 0 \leq y \leq \psi(x)\}$ with the identification $(x, \psi(x)) = (f(x), 0)$. The function f_ψ is defined by $f_{\psi,t}(x, y) = (x, y+t)$ for small t (see [24, Chap. 6]).

The weak mixing property for f_ψ implies that the Zeta function $\zeta_{-\psi}(s)$ has a nonzero analytic extension to Real$(s) \geq 1$, except for $s = 1$. In $s = 1$, it has a simple pole. We recall the definition of

$$\zeta_{-\psi}(s) = \exp \sum_{n=1}^{\infty} \frac{1}{n} \sum_{f^n(x)=x} e^{-s\psi_n(x)}, \qquad s \in \mathbf{C},$$

where $\psi_n(x)$ is the sum of the potential ψ around the orbit of the n-periodic point x. In the present situation this function is well defined, nonzero, and analytic for Real $s > 1$.

Theorem 6.9 of [24] (see also Chap. 9) implies the following corollary.

COROLLARY. *For the system defined in* §§ 2 *and* 5, *there exists a dense set of parameters* $a > 4/\sqrt{3}$ *such that* $\{$*periodic orbits* $x, f^k(x) = x, \log(f^{k'}_{E_u(x)}(x) \leq r\}$ *grows like* $e^{hr}/(hr)$ *as* $r \to \infty$ *and where* $h > 0$ *is the topological entropy of* f_ψ.

Proof of Theorem 5. The claim is equivalent to showing that there is no continuous function g such that $\psi = g \circ T - g + G\gamma$, where G is an integer-valued function and γ is a real positive constant.

Suppose there exists such a g, G as above; we will arrive at a contradiction as follows. If there exists such a G, the sum of the values of the function ψ along a periodic orbit is always of the form $n\gamma$, with $n \in \mathbf{N}$ depending on the orbit.

We will show that for a dense set of values $a > 4/\sqrt{3}$, the open billiard with this parameter a is such that the period-two and period-three orbits do not have the above-mentioned property. The values a will be rationals of the form

$$a = \frac{2p^2}{p^2 - q^2} = \frac{2}{1 - \frac{1}{(p/q)^2}}, \qquad p, q \in \mathbf{N}.$$

From the continuity of the function

$$r(x) = \frac{2}{1 - \frac{1}{x^2}},$$

it is easy to see that the set of such values a is dense in $a > 4/\sqrt{3}$.

Denote by t_2 and t_3 the length between bounces for, respectively, the period-two and the period-three orbit (see Fig. 2). Denote also by k_2 and k_3, respectively, the expressions $k^u(x_2)$ and $k^u(x_3)$ (see (4.5)), where x_2 and x_3 are, respectively, points on orbits of period two and three.

We want to show that there are no n_2 and n_3 such that

$$2 \log(1 + t_2 k_2) = n_2 \gamma \tag{8.1}$$

and

$$3 \log(1 + t_3 k_3) = n_3 \gamma. \tag{8.2}$$

Equivalently, we will show that there are no n_2 and n_3 such that

$$(1 + t_2 k_2)^{2n_3} = (1 + t_3 k_3)^{3n_2}. \tag{8.3}$$

From simple geometrical arguments (see Fig. 2), it is easy to see that $t_2 = a - 2$ and $t_3 = a - \sqrt{3}$.

Now we will analyze the value k_2. From (4.5) (the continued fraction expression of k^u) it follows that a 2-periodic point x_2 satisfies the property

$$k^u(x_2) = k_2 = \frac{2}{\cos 0} + \frac{1}{t_2 + \frac{1}{k_2}},$$

and therefore k_2 satisfies

$$k_2^2 - 2k_2 - 2/t_2 = 0. \tag{8.4}$$

As $t_2 = a - 2$ and $a = \frac{2p^2}{p^2-q^2}$, it follows from the quadratic formula that $k_2 = 1 + p/q$ is rational. Therefore $(1 + t_2 k_2)^{2n_3}$ is rational.

We will show that up to a finite number of values $a \in \mathbf{Q}$, the value $(1 + t_3 k_3)^{3n_2}$ is not rational, from which Theorem 5 will follow.

Now we will analyze k_3. From the symmetry of the orbit of period 3, it follows from (4.5) that k_3 satisfies

$$k^u(x_3) = k_3 = \frac{2}{\cos \pi/6} + \frac{1}{t_3 + \frac{1}{k_3}},$$

and therefore k_3 satisfies the quadratic equation

$$k_3^2 - \frac{4}{\sqrt{3}}k_3 - \frac{4}{a\sqrt{3} - 3} = 0. \tag{8.5}$$

Theorem 5 follows at once from the next lemma.

LEMMA 2. *Let $a \in \mathbf{Q}, a > 2$, and let ζ be the positive root of*

$$x^2 - (4/\sqrt{3})x - 4/(a\sqrt{3} - 3) = 0. \tag{8.6}$$

There exists a finite set $S \subset \mathbf{Q}$ such that if $2 < a \in \mathbf{Q} - S$, then $(1+(a-\sqrt{3})\zeta)^m = (1 + t_3\zeta)^m$ is not in $\mathbf{Q}$ for any $m \in \mathbf{N}$.

Proof. Remember that a number α is called algebraic if it is a root of an equation

$$x^n + a_1 x^{n-1} + \ldots + a_n = 0, \qquad a_j \in \mathbf{Q},\ n \geq 1.$$

We may assume that this equation is irreducible over $\mathbf{Q}$. We refer the reader to [11] for general properties on algebraic structures that will be used in this section. Denote by α a solution of the above equation. The equation above is uniquely defined in this situation, and all roots are different. The set of solutions of such an equation is called the set of conjugates to α. Therefore, α has n conjugates and α also is conjugate to itself. The degree of the extension $\mathbf{Q}[\alpha]/\mathbf{Q}$ is equal to n. Any automorphism of $\mathbf{C}$ leaves fixed each rational number and transforms α in a conjugate of α. Any conjugate of α is the image of α by some automorphism.

An algebraic number α is called totally real if α is real and all its conjugates are also real.

CLAIM. *Let α be an irrational algebraic totally real number. Suppose that there exists $m \in \mathbf{N}$ such that $\alpha^m \in \mathbf{Q}$. Then $\mathbf{Q}[\alpha]$ is of degree 2 over Q, and the conjugates of α are α and $-\alpha$.*

Proof of the Claim. Let $\alpha_1, \ldots, \alpha_n$ be the conjugates of α. Since $\alpha^m \in \mathbf{Q}$, we have $\alpha_j^m = \alpha^m$, because α_j^m is conjugate to α^m. Therefore, $(\alpha_j/\alpha)^m = 1$. Since $\alpha_j/\alpha \in \mathbf{R}$, we have that α_j/α is equal to 1 or -1. Hence, any conjugate of α is equal to α or $-\alpha$. But α is not in $\mathbf{Q}$, thus $n \geq 2$, and therefore $n = 2$, proving the Claim.

Proof of Lemma 2. The lemma will follow from the four properties listed below.

i) ζ is totally real. This is so because any conjugate of ζ is a root of (8.6) or of $x^2 + (4/\sqrt{3})x + 4/(a\sqrt{3}+3) = 0$, since any automorphism of $\mathbf{C}$ takes $\sqrt{3}$ to $\sqrt{3}$ or $-\sqrt{3}$. Therefore, all conjugates of ζ are real.

ii) $\eta = 1+(a-\sqrt{3})\zeta$ is totally real. Any conjugate of η is of the form $(1+(a-\sqrt{3})\bar{\zeta})$ or $(1+(a+\sqrt{3})\bar{\zeta})$, where $\bar{\zeta}$ is conjugate to ζ. Therefore, any conjugate of η is real.

iii) Suppose (8.6) is irreducible over $\mathbf{Q}[\sqrt{3}]$; then η^m is not in $\mathbf{Q}$ for all $m \in \mathbf{N}$ and for all $a \in \mathbf{Q}, a > 2$.

The proof of iii) is by contradiction. Suppose that there exists $m \in \mathbf{N}$ such that $\eta^m \in \mathbf{Q}$. From the Claim above, $[\mathbf{Q}[\eta] : \mathbf{Q}] \leq 2$. Since $(a - \sqrt{3})\zeta = \eta - 1 \in \mathbf{Q}[\eta]$, $(a - \sqrt{3})\zeta$ is a root of an equation of degree 2 over $\mathbf{Q}$; assume that

$$x^2 + rx + s = 0, \qquad r, s \in \mathbf{Q},$$

is such an equation. Then ζ is a root of the equation

$$x^2 + \frac{r}{a-\sqrt{3}}x + \frac{s}{(a-\sqrt{3})^2} = 0. \tag{8.7}$$

Therefore, we conclude that (8.6) and (8.7) are both equations with coefficients in $\mathbf{Q}[\sqrt{3}]$ with a common root ζ. Since we are assuming that (8.6) is irreducible over this field, equations (8.6) and (8.7) are the same. In particular,

$$-\frac{4}{\sqrt{3}} = \frac{r}{a-\sqrt{3}},$$

and therefore, $-4a + (4 - r)\sqrt{3} = 0$. But $r \in \mathbf{Q}$; hence $a = 0$, which contradicts the fact that $a > 2$.

iv) Now suppose that (8.6) is reducible over $\mathbf{Q}[\sqrt{3}]$; then there exists a finite set $S \subset \mathbf{Q}$ such that if $a \in \mathbf{Q}, a > 2$, and a is not in S, then η^m is not in $\mathbf{Q}$ for all $m \in \mathbf{N}$.

If (8.6) is reducible, then $\zeta \in \mathbf{Q}[\sqrt{3}]$. Hence $\eta \in \mathbf{Q}[\sqrt{3}]$. For the proof, we suppose that there exists $m \in \mathbf{N}$ such that $\eta^m \in \mathbf{Q}$ and prove that a must be contained in some finite set $S \subset \mathbf{Q}$. From the last claim, $\eta \in \mathbf{Q}$ or $-\eta$ is conjugated to η. Write $\zeta = u + v\sqrt{3}, u, v \in \mathbf{Q}$; then

$$\eta = (1 + au - 3v) + (av - u)\sqrt{3},$$

and therefore, either $\eta \in \mathbf{Q}$ and we have

$$av - u = 0 \tag{8.8}$$

or else $-\eta$ is conjugate to η and we have

$$1 + au - 3v = 0. \tag{8.9}$$

Suppose that (8.9) is true. Then $v = \frac{1+au}{3}$ and therefore, $\zeta = u + \frac{1+au}{3}\sqrt{3}$. Equation (8.6) is equivalent to

$$\sqrt{3}x^2 - 4x - \frac{4(a+\sqrt{3})}{a^2-3} = 0,$$

hence

$$\sqrt{3}\left(u + \frac{1+au}{3}\sqrt{3}\right)^2 - 4\left(u + \frac{1+au}{3}\sqrt{3}\right) - \frac{4(a+\sqrt{3})}{a^2-3} = 0.$$

Since $a, u \in \mathbf{Q}$, we should consider the set of two equations:

$$\left(1+\frac{1}{3}a^2\right)u^2-\frac{2}{3}au-1-\frac{4}{a^2-3}=0 \tag{8.10}$$

and

$$2au^2-2u-\frac{4a}{a^2-3}=0. \tag{8.11}$$

If there exist infinitely many values a such that the two equations (8.10) and (8.11) have a common root, the result of these two polynomials would be identically zero. (The coefficients of this polynomials depend on a.) In particular, for $a = 1$ there would be a common root, but this is not true.

Hence, there exists a finite set $S_1 \subset \mathbf{Q}$ such that if a is not in $\mathbf{Q}$, then (8.10) and (8.11) do not have a common root. Therefore, if a is not in S_1, $1 + au - 3v \neq 0$, and this is a contradiction to (8.9).

Now assume alternatively that (8.8) is true; that is, $av - u = 0$. In this case $\zeta = av - v\sqrt{3}$. Proceeding as before, we obtain

$$\sqrt{3}(av-v\sqrt{3})^2-4(av-v\sqrt{3})-\frac{4(a+\sqrt{3})}{a^2-3}=0 \tag{8.12}$$

and therefore the system of equations

$$3v^2+2v+\frac{2}{a^2-3}=0 \tag{8.13}$$

and

$$(a^2+3)v^2+4v-\frac{4}{a^2-3}=0. \tag{8.14}$$

Analogously, as in the case (8.9), the fact that these two equations do not have a common root if $a = 1$ implies that there exist a finite set $S_2 \subset \mathbf{Q}$ such that $av - u \neq 0$ if a is not in S_2.

Finally the set S is obtained as the union of S_1 and S_2, proving iv). This ends the proof of Lemma 2, and therefore Theorem 5 is proved.

Appendix. The $C^{1+\epsilon}$ theorems. In this appendix we will prove Theorems 1 and 2 for the case $C^{1+\epsilon}$, adapting the proofs in [6] and [25].

Proof of Theorem 1. a) Let

$$C^\gamma(A)=\left\{\varphi : \sup\left\{\frac{|\varphi(y)-\varphi(x)|}{d^\gamma(x,y)} : x,y\in A, x\neq y\right\}<\infty\right\}$$

be the set of γ-Hölder-continuous functions defined on A. For every nonnegative function in $C^\gamma(A)$, we define its regularity as

$$R(\varphi)=\sup\left\{\frac{|\varphi(y)-\varphi(x)|}{d^\gamma(x,y)\varphi(x)} : x,y\in A, \varphi(x)>0\right\}.$$

Define $H = \{\varphi \in C^\gamma(A) : \varphi \geq 0, R(\varphi) < \infty, \int \varphi d\nu = 1\}$ and $H_\rho = \{\varphi \in H : R(\varphi) \leq \rho\}$ for every $p > 0$. Let P, the normalized Perron–Frobenius operator, $P : L^1(A,\nu) \to L^1(A,\nu)$, be defined by

$$P(\varphi)=\frac{P_1(\varphi)}{\|P_1(\varphi)\|_1}=\frac{\sum_{y,Tz=x}\frac{\varphi(z)}{T_d(z)}}{\int_{T^{-1}A}\varphi d\nu}.$$

b) We claim that there exists a $\rho > 0$ independent of φ such that

$$\limsup_{n\to+\infty} R(P^n\varphi) \le \rho$$

for all $\varphi \in H$.

We begin by evaluating $R(P\varphi)$. Since $T_{|A_i}$ is an homeomorphism, we can consider the local inverses $S_i : TA \to A_i$ such that $T \circ S_i = Id$, and if $Tw_1 \in A_i$, there exists S_j such that $S_j \circ T(w) = w$. We suppose that w_i and z_i are, respectively, the preimages of y and x in the same "inverse branch," i.e., w_i, $z_i \in A_i$, $Tw_i = y$, $Tz_i = x$.

Then

$$\frac{|\sum_{Tw=y}\varphi(w)(T_d(w))^{-1} - \sum_{Tz=x}\varphi(z)(T_d(z))^{-1}|}{d^\gamma(x,y)\sum_{Tz=x}\varphi(z)(T_d(z))^{-1}}$$

$$\le \frac{\left|\sum_i \frac{\varphi(w_i)-\varphi(z_i)}{T_d(w_i)}\right| + \left|\sum_i \frac{\varphi(z_i)}{T_d(w_i)} - \frac{\varphi(z_i)}{T_d(z_i)}\right|}{d^\gamma(x,y)\sum_i \varphi(z_i)(T_d(z_i))^{-1}}$$

$$\le \max \frac{T_d(z_i)}{T_d(w_i)}\frac{|\varphi(w_i)-\varphi(z_i)|}{d^\gamma(w_i,z_i)\varphi(z_i)}\left[\frac{d(w_i,z_i)}{d(x,y)}\right]^\gamma$$

$$+ \max \frac{|T_d(z_i)-T_d(w_i)|}{d^\gamma(z_i,w_i)}\frac{1}{T_d(w_i)}\left[\frac{d(z_i,w_i)}{d(x,y)}\right]^\gamma.$$

(We have applied Lemma 4.1 of [22]:

$$\left|\frac{\sum a_i}{\sum b_i}\right| \le \max\left|\frac{a_i}{b_i}\right|$$

for any real numbers a_i, b_i, $b_i > 0$, $i = 1, \dots, q$.)

As was remarked immediately after the statement of Lemma 1,

$$\frac{T_d(z_i)}{T_d(w_i)}\left[\frac{d(w_i,z_i)}{d(x,y)}\right]^\gamma \le k_1\left[\frac{1}{\beta}\right]^\gamma = \lambda < 1;$$

then the first term of the last expression is less than $\lambda R(\varphi)$.

The second term is less than

$$k\left[\frac{1}{\beta}\right]^{\gamma+1} = M.$$

So we have that $R(P\varphi) \le M + \lambda R(\varphi)$, and iteration of this inequality yields $R(P^n\varphi) \le M(1+\lambda+\cdots+\lambda^{n-1}) + \lambda^n R(\varphi)$ and finally

$$\limsup_{n\to\infty} R(P^n\varphi) \le \frac{M}{(1-\lambda)} = \rho.$$

c) For the value of ρ that we have just defined, it results that H_ρ is invariant under ρ, since $R(P\varphi) \le M + \lambda\rho = \rho$ if $\varphi \in H_\rho$.

d) H_ρ is convex because if $\varphi \in H_\rho$,

$$R(\alpha\varphi + \beta\psi) \leq \sup \frac{|\alpha(\varphi(x) - \varphi(y)) + \beta(\psi(x) - \psi(y))|}{d^\gamma(x,y)(\alpha\varphi(x) + \beta\psi(x))}$$

$$\leq \sup\max\left\{ \frac{|\varphi(x) - \varphi(y)|}{d^\gamma(x,y)\varphi(x)}, \frac{|\psi(x) - \psi(y)|}{d^\gamma(x,y)\psi(x)} \right\} \leq \rho.$$

(We have once again applied Lemma 4 of [25].)

e) H_ρ is compact. If $\varphi \in H_\rho$, we have

$$\frac{|\varphi(y)|}{\varphi(x)} \leq 1 + \rho d^\gamma(x,y)$$

for every x, $y \in A_i$ and

$$L^{-1}\varphi(x) \leq \varphi(y) \leq L\varphi(x).$$

In particular either φ is zero on A_i or $\inf_{x\in A_i} \varphi(x) > 0$. Furthermore

$$\sup_{y\in A} \varphi(y) \leq L \inf_{x\in A} \varphi(x) \tag{A.1}$$

and $\inf_{x\in A} \varphi(x) < c$ independent of φ because $\int_A \varphi d\nu = 1$. Then H_ρ is equibounded; $\varphi(y) \leq Lc$.

From the equiboundedness and the definition of $R(\varphi)$ it follows that H_ρ is equicontinuous $|\varphi(x) - \varphi(y)| \leq \rho d^\gamma(x,y)\varphi(y) \leq \rho Lcd^\gamma(x,y)$; given ϵ there exists

$$\delta = \left[\frac{\varepsilon}{\rho Lc}\right]^{1/\gamma}$$

such that if $d(x,y) < \delta$, then $|\varphi(y) - \varphi(x)| < \varepsilon$.

It is also closed because there is a uniform Hölder constant equal to ρLc, and all the inequalities hold for the limit functions.

f) Then we can apply the Schauder fixed-point theorem [9] and obtain a function $F \in H_\rho$ such that $PF = F$. The measure μ_F defined by $d\mu_F = Fd\nu$ satisfies the first assertion.

g) The proof of the second assessment is almost the same of that of Theorem 2 in [25].

We remark that if $\psi \in \mathcal{K}$ and $\beta_n(\psi) = \|P_1^n(\psi)\|_1$, then $\sup_n \|P^n(\psi)\|_0 < \infty$ ($\|\cdot\|_0$ is the supremum norm in $C^0(\bar{A})$) as a consequence of the following observation. Since the functions $P^n(\gamma)$ are in H_ρ, from (A.1) it follows that $\sup \|P^n(1)\| < \infty$. We also have that $B_n(1)\inf_{x\in A}\psi(x) \leq \beta_n(\psi) \leq \beta_n(1) \leq B_n(1)\sup_{x\in A}\psi(x)$, and

$$P^n(\psi)(x) = \frac{P_1^n(\psi)(x)}{B_n(\psi)} \leq \frac{\sup\psi}{\inf\psi}\frac{P_1^n(1)(x)}{B_n(1)} = \frac{\sup\psi}{\inf\psi}P^n(1)(x).$$

Then

$$\sup_{x\in A} P^n(\psi)(x) \leq \frac{\sup\psi}{\inf\psi}\sup_{x\in A} P^n(1)(x),$$

and finally $\sup \|P^n(\psi)\|_0 < \infty$.

This remark is used in the proof of Proposition 1 of [25].

h) Our assessment iii) is exactly the same as that of Theorem 3 in [25].

Proof of Theorem 2. The proof of Theorem 2 will be divided into three lemmas, as was done in the proof of the C^2-case [6].

LEMMA 3. *If $I \in \mathcal{P}_n$, denote by $S_I : A \to I$ the inverse branch of T^n : $T^n \circ S_I = Id$, and if $z \in I$, then $S_I \circ T^n z = z$. Since $Q^n\varphi = (\alpha^n F)^{-1} P^n(F\varphi)$, we have that $Q^n\varphi = (\alpha^n F)^{-1} \sum (T_d \circ S_I)^{-1} F \circ (S_I . \varphi) \circ S_I$.*

In $C^\gamma(A)$, consider the norms

$$\|\varphi\|_\gamma = \sup\left\{ \frac{|\varphi(x) - \varphi(y)|}{d(x,y)^\gamma}, x \neq y, x, y \in A \right\}$$

and $\|\varphi\|_B = \|\varphi\|_\gamma + \|\varphi\|_\infty$. Then $\mathbf{B} = \{\varphi \in C^\gamma(A) : \|\varphi\|_B < \infty\}$ in a Banach space.

An operator Q acting on a Banach space is quasi-compact if there exists a compact operator H such that $\|Q^N - H\| < 1$ for some $N \in \mathbf{N}$.

LEMMA 4. *Q is a quasi-compact operator on $\mathbf{B}$. Consider the operator L_n defined by*

$$L_n\varphi = (\alpha^n F)^{-1} \sum_{I \in \mathcal{P}_n} (T_d \circ S_I)^{-1} F \circ S_I \frac{1}{\nu(I)} \int_I \varphi d\nu.$$

If $1_I(z) = 1$ for $z \in I$ and zero is any other point, then $\{Q1_I : I \in \mathcal{P}_n\}$ is a base of the image of L_n. So L_n is (of finite rank and then) compact.

We will prove that for some large enough n, $\|Q^n - L_n\|_B < 1$. We have for $\varphi \in \mathbf{B}$

$$(Q^n - L_n)\varphi = (\alpha^n F)^{-1} \sum_{I \in I} (T^n \circ S_I)^{-1} F \circ S_I \left(\varphi \circ S_I - \frac{1}{\nu(I)} \int_I \varphi d\nu \right)$$

with

$$\left| \phi \circ S_I(x) - \frac{1}{\nu(I)} \int \varphi d\nu \right| = |\varphi(z) - \varphi(w)| \leq \|\varphi\|_\gamma d^\gamma(z,w)$$

$$\leq \|\varphi\|_B (\nu(I))^\gamma \leq \|\varphi\|_B \left[\frac{1}{\beta} \right]^{n\gamma}.$$

Then $|(Q^n - L_n)\varphi| \leq (\alpha^n F)^{-1}(P^n F)\|\varphi\|_B \beta_1^n = \|\varphi\|_B \beta_1^n$, which implies that $\|(Q^n - L_n)\varphi\|_\infty$ goes to zero when $n \to +\infty$. Denote $S_I(x) = z, S_I(y) = y$; then

$$\begin{aligned}
&\frac{((Q^n - L_n)\varphi)(x) - ((Q^n - L_n)\varphi)(y)}{d^\gamma(x,y)} \\
&\times \left(\frac{1}{\alpha^n F(x)} - \frac{1}{\alpha^n F(y)} \right) \alpha^n F(x)((Q^n - L_n)\varphi)(x) \\
&+ \frac{1}{\alpha^n F(y)} \sum_{I \in \mathcal{P}_n} ((F(z) - (F(w))(T_d^n(z))^{-1} \left(\varphi(z) - \frac{1}{\nu(I)} \int_I \varphi d\nu \right) \\
&+ \frac{1}{\alpha^n F(y)} \sum_{I \in \mathcal{P}_n} (F(w))[T_d^n(z)^{-1} - (T_d^n(w))^{-1}] \left(\varphi(z) - \frac{1}{\nu(I)} \int \varphi d\nu \right) \\
&\times \frac{1}{\alpha^n F(y)} \sum_{I \in \mathcal{P}_n} (F(w))[T_d^n(w)]^{-1} [\varphi(z) - \varphi(w)] \frac{1}{d^\gamma(x,y)} \leq a_n + b_n + c_n + d_n
\end{aligned}$$

with

$$\odot\ a_n = \frac{F(y) - F(x)}{F(y) d^\gamma(x,y)}((Q^n - L_n)\varphi)(x)\ ,\ |a_n| \leq \frac{\|F\|_\gamma}{\inf F} \|\varphi\|_B \beta_1 = M_a \|\varphi\|_B \beta_1^n$$

($\inf F > 0$ since $F \in \mathcal{K}$).

$$\odot\ b_n = \frac{1}{\gamma^n F(y)} \sum_{I\in\mathcal{P}_n} \frac{F(z)-F(w)}{d(z,w)^\gamma}\frac{d(z,w)^\gamma}{d(x,y)^\gamma}(T_d^n(z))^{-1}\left(\varphi(z) - \frac{1}{\nu(I)}\int(\varphi d\nu)\right),$$

$$|b_n| \le \frac{1}{\alpha^n \inf F}\|F\|_\gamma \sum_{I\in\mathcal{P}_n} \beta_1^n \frac{1}{\beta^n}\|\varphi\|_B\beta_1^n = \frac{\|F\|_\gamma}{\inf F}\frac{k^n}{\alpha^n}\frac{1}{\beta^n}\beta_1^{2n} \le M_b\beta_1^n\|\varphi\|_B,$$

$$\odot\ c_n = \frac{1}{\alpha^n F(y)} \sum_{I\in\mathcal{P}_n} F(w)(T_d^n(z))^{-1}(T_d^n(w))^{-1}\frac{T_d^n(w)-T_d^n(z)}{d(x,y)^\gamma}$$
$$\times\left(\varphi(z) - \frac{1}{\nu(I)}\int_I \varphi d\nu\right),$$

$$|c_n| \le \frac{(P_1^n F)(y)}{\alpha_1^n F(y)}\sup(T_d^n(z))^{-1}\frac{|T_d^n(w)-T_d^n(z)|}{d(x,y)^\gamma}\left|\varphi(z) - \frac{1}{\nu(I)}\varphi d\nu\right|$$
$$\le \beta^{-n}\beta_1 M^{n-1}K_2\|\varphi\|_B\beta_1^n \le M_e\beta_1^{n/2}\|\varphi\|_B;$$

if n is large enough,

$$\odot\ d_n = \frac{1}{\alpha^n F(y)} \sum_{I\in\mathcal{P}_n} F(w)(T_d^n(w))^{-1}\frac{\varphi(z)-\varphi(w)}{d(x,y)^\gamma},$$

$$|d_n| \le \frac{(P_1^n F)(y)}{\alpha^n F(y)}\sup\frac{|\varphi(z)-\varphi(w)|}{d(z,w)^\gamma}\left(\frac{d(z,w)}{d(x,y)}\right)^\gamma \le \|\varphi\|_\gamma\beta_1^n.$$

We conclude that $\|(Q^n - L_n)\varphi\|_B \le c\beta_1^{n/2}\|\varphi\|_B$ for some constant c and for a large enough n. The lemma is proven.

LEMMA 5. *Restricted to* $\mathbf{B}$, *the operator* Q *has 1 as a simple eigenvalue, and the rest of the spectrum is in a disk of radius* $r < 1$. *The previous lemma allows us to apply VIII.8.6 of [9] and conclude that the spectrum of* Q *can be decomposed into the union of a closed set which lies inside the circle* $|z| < r < 1$ *and a finite number of simple poles* ρ_j, $i = 1, \dots, q$, $|\rho| = 1$.

If $\varphi \in \mathbf{B}$ satisfies $Q\varphi = \rho\varphi$ for $|\rho| = 1$, choose $k > 0$ such that $\varphi + k > 0$. Hence $\varphi + k \in \mathcal{K}$. Since $Q1 = 1$, we have $Q^n(\varphi + k) = \rho^n\varphi + k$. From Theorem iii)

$$\frac{Q^n(\varphi+k)}{\|Q^n(\varphi+k)\|_1} = \frac{F^{-1}P_1^n((\varphi+k)F)}{\|\int P_1^n((\varphi+k)F)\|_1}\frac{\|\int P_1^n((\varphi+k)F)\|_1}{\int F^{-1}P_1^n((\varphi+k)F)d\nu} \to 1.$$

Therefore

$$\frac{\rho^n\varphi+k}{\|\rho^n\varphi+k\|_1} \to 1.$$

But for k large enough, $\rho^n\varphi + k$ is bounded and bounded away from zero for every $|\rho| = 1$, $n \in \mathbf{N}$. Then ρ must be 1. This relation also shows that the eigenfunctions associated with 1 are the constant functions

$$\frac{\varphi+k}{\|\varphi+k\|_1}.$$

Therefore 1 is a simple eigenvalue.

Now let φ_m be the eigenfunctions of the eigenvalues e_m, $|e_m| < r < 1$. Then for any $\psi \in \mathbf{B}$, $\psi = k1 + \sum \alpha_m \varphi_m$, and $Q^n\psi = k1 + \sum \alpha_m e_m^n \varphi_m \to k$ as $n \to +\infty$. So $G : C^0(B) \to \mathbf{R}$ defined by $G(\psi) = k$ is a linear positive functional. From the Riesz theorem there exists a unique measure μ such that $Sgd\mu = G(g)$ for every $g \in C^0(B)$. (The relation is valid for every $g \in \mathbf{B}$.)

This measure is invariant under T since

$$Q(g \circ T) = (\alpha F)^{-1} \sum_i g \circ T \circ S_i \cdot F \circ S_i (T_d \circ S_i)^{-1}$$
$$= (\alpha F)^{-1} g \sum_i F \circ S_i (T_d \cdot S_i)^{-1} = (\alpha F)^{-1} g P_1 F = g.$$

Hence $Q^{n+1}(g \circ T) = Q^n(g)$ and $\int g \circ T d\nu = G(g \circ T) = G(g) = \int g d\nu$ for every $g \in C^0(B)$.

Denote by $K = \bigcap_{n>0} T^{-n}(\bar{A})$ the limit Cantor set. The measure μ is supported by K. In fact, if g vanishes on a neighbourhood of the Cantor set, $Q^n g$ converges to zero. This Cantor set can be coded by the partition of connected components of $\mathcal{P}$. As usual, denote by $[i_0, \dots, i_{n-1}]$ the set

$$\bigcap_{i=0}^{n-1} T^{-l} A_{i_l} \in \mathcal{P}_n,$$

and let $J_{i_0}, \dots, i_{n-1} = T^n_{\mathcal{P}}(y)$ for some $y \in [i_0, \dots, i_{n-1}]$.

LEMMA 6. *There exists a constant $c > 0$ such that for every n*

$$c^{-1} J^{-1}_{i_0, \dots, i_{n-1}} \alpha^{-n} \le \mu([i_0, \dots, i_{n-1}]) \le c J^{-1}_{i_0, \dots, i_{n-1}} \alpha^{-n}.$$

Its proof is exactly the same as that of Lemma 3 in [6].

We have proved parts i) and iii) of Theorem 2. It remains to prove ii). We know that

$$\mu_F(C \cap T^{-n}A) = \int_{T^{-n}A} 1_C \cdot (1_A \circ T^n) \cdot F d\nu = \int_A P^n(1_c F) d\nu$$

(see the remarks between the statements of Theorems 1 and 2).

From the definition of Q we obtain

$$\mu_F(C \cap T^{-n}A) = \int_A \alpha^n F Q^n(1_c) d\nu.$$

But $Q^n(1_c)$ converges to $\mu(C)$ in $L^1(\bar{A}, \nu)$, and F is bounded: then

$$\mu_F(C \cap T^{-n}A)\alpha^{-n} = \int_A F \mu(C) d\nu.$$

It was observed at the beginning of §6 that $\alpha^n = \mu_F(T^{-n}A)$; then iii) is proved.

Acknowledgments. We would like to thank R. Mañé, M. Sebastiani, A. Doering, S. Martinez, R. Ures, and N. Chernov for helpful conversations about several questions related to this paper.

REFERENCES

[1] L. A. Bunimovich, Ya. G. Sinai, and N. I. Chernov, *Markov partitions for the two dimensional billiards*, Russian Math. Surveys, 45 (1990), pp. 105–152.

[2] ———, *Statistical properties of the two dimensional hyperbolic billiards,* Russian Math. Surveys, 46 (1991), pp. 47–106.

[3] N. N. Cencova, *A natural measure on Smale's horseshoe*, Soviet Math. Dokl., 23 (1981), pp. 87–91.

[4] N. I. Chernov and R. Markarian, *Entropy of non-uniformly hyperbolic plane billiards*, Bol. Soc. Brasil. Mat. (N.S.), 23 (1992), pp. 121–135.

[5] N. I. Chernov, *New proof of Sinai's formula for the entropy of hyperbolic billiard systems: Applications to Lorenz gases and Bunimovich stadium*, Functional Anal. Appl., 25 (1991), pp. 204–219.

[6] P. Collet, S. Martinez, and B. Schmitt, *The Yorke-Pianigiani measure and the asymptotic law on the limit to the Cantor set of expanding maps*, Nonlinearity, 7 (1994), pp. 1437–1443.

[7] P. Cvitanovich and B. Eckhardt, *Periodic quantization of chaotic systems*, Phys. Rev. Lett., 63 (1989), pp. 823–826.

[8] M. Ding, C. Grebogi, E. Ott, and J. Yorke, *Transition to chaotic scattering*, Phys. Rev. A, 42 (1990), pp. 7025–7040.

[9] N. Dunford and J. T. Schwartz, *Linear Operators*, Vol. I, Interscience, New York, 1958.

[10] P. Ferrari, H. Kesten, S. Martinez, and P. Picco, *Existence of quasi stationary distribution: A renewal dynamical approach*, Annals of Probab., to appear.

[11] I. N. Herstein, *Topics in Algebra*, Blaisdell, Waltham, 1969.

[12] M. Ikawa, *Decay of solutions of the wave equation in the exterior of several convex bodies*, Ann. Inst. Fourier (Grenoble), 38 (1988), pp. 113–146.

[13] A. Manning, *A relation between Liapunov exponents, Hausdorff dimension and entropy*, Ergodic Theory Dynamic Systems, 1 (1981), pp. 451–459.

[14] R. Mañé, *Ergodic Theory and Differentiable Dynamics*, Springer-Verlag, New York, 1987.

[15] R. Markarian, *Billiards with Pesin region of measure one*, Comm. Math Phys., 118 (1988), pp. 87–97.

[16] ———, *New ergodic billiards: Exact results,* Nonlinearity, 6 (1993), pp. 819–841.

[17] ———, *Non-uniformly hyperbolic billiards,* Ann. Fac. Sci. Toulouse, to appear.

[18] S. Martinez and M. E. Vares, *Markov chain associated to the minimal Q.S.D. of birth-death chains*, J. Appl. Probab., to appear.

[19] L. Mendoza, *The entropy of C^2 surface diffeomorphism in terms of Hausdorff dimension and Liapunov exponents*, Ergodic Theory Dynamic Systems, 5 (1985), pp. 273–283.

[20] T. Morita, *The symbolic representation of billiards without boundary conditions*, Trans. Amer. Math. Soc., 325 (1991), pp. 819–828.

[21] A. Ozorio de Almeida, *Hamiltonian Systems: Chaos and Quantization*, Cambridge University Press, Cambridge, UK, 1988.

[22] E. Ott, *Chaos in Dynamical Systems*, Cambridge University Press, Cambridge, UK, 1993.

[23] J. Palis and F. Takens, *Hyperbolicity and Sensitive Chaotic Dynamics at Homoclinic Bifurcations*, Cambridge University Press, Cambridge, UK, 1993.

[24] W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Asterisque, 187–188 (1990).

[25] G. Pianigiani and J. Yorke, *Expanding maps in sets which are almost invariant: Decay and chaos*, Trans Amer. Math. Soc., 252 (1979), pp. 351–366.

[26] S. Rice, P. Gaspard, and K. Nakamura, *Signature of chaos in quantum dynamics and the controllability of evolution in a quantum system*, Adv. in Classical Trajectory Methods, 1 (1992), pp. 215–313.

[27] Ya. Sinai, *Dynamical systems with elastic reflections*, Russian Math. Surveys, 25 (1970), pp. 137–189.

[28] U. Smilansky, *The classical and quantum theory of chaotic scattering*, in Chaos and Quantum Physics, M. J. Giannoni, A. Voros, and J. Zinn-Justin, eds., Elsevier, New York, 1992, pp. 372–441.

[29] P. Walters, *An Introduction to Ergodic Theory*, Springer, Berlin, 1982.