Journal of Statistical Physics, Vol. 80, Nos. 5/6, 1995

Statistical Properties of Dynamical Systems with Singularities

Roberto Markarian [1]

Received November 1, 1993; final February 24, 1995

We prove statistical properties of two-dimensional hyperbolic dynamical systems with singularities. Bunimovich, Sinai, and Chernov proved a theorem on the subexponential decay of correlations and a central limit theorem for billiard systems. In this paper we use their techniques to prove the same results for "abstract systems."

KEY WORDS: Dynamical systems; hyperbolicity; decay of correlations; billiards.

1. INTRODUCTION

We will study statistical properties of two-dimensional smooth hyperbolic dynamical systems with singularities. We will obtain upper bounds for the rate of the decay of correlations and a central limit theorem.

The (nonuniform) hyperbolicity of our dynamical systems is characterized by nondegenerate increasing quadratic forms. All the conditions will be given in terms of these quadratic forms and the evolution of singularities. Close to the singularities, we also introduce some relations between the measure and the pseudometric defined by the quadratic form. These relations are a little more restrictive than those established for proving ergodic properties in our previous papers.[23,24]

In this paper we suppose that the quadratic form is uniformly expanding on unstable directions (see Section 3.3). This condition simplifies some constructions and evaluations. If this uniformity is not satisfied, a derived

[1] Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia," Facultad de Ingeniería, Universidad de la República, CC No. 30, Montevideo, Uruguay. E-mail: roma@fing.edu.uy.

0022-4715/95/0900-1207$07.50/0

map can be introduced. In this case the presence of new singularities must be studied. This was done, for example, in ref. 6 and 7 to study the same properties in the stadium.

We work in the context of the book by Katok and Strelcyn[14] in order to use Pesin's methods for constructing invariant manifolds. Motivated by billiard problems, they imposed conditions on the derivatives of the maps near the singularities.

We adapt the methods and techniques used by Bunimovich *et al.*[6,7] for some hyperbolic billiards to our "abstract" systems and obtain the same results. So the main results in this paper are generalizations of those obtained by the Russian school, using quadratic forms and conditions on the evolution of singularities instead of specific properties of these mechanical systems.

The ideas behind the proofs of the main results and the previous literature on these topics can be found in the Introduction of ref. 7, especially pp. 50–52, and ref. 9.

The fundamental method of previous investigations consists in the construction of Markov partitions with the subsequent reduction of the system to its symbolic representation as a topological Markov chain with finite alphabet. Roughly speaking, a Markov partition consists in a finite covering by parallelograms with sides parallel to unstable and stable directions of the map f. Almost every point y belongs to an element $U(y)$ of the partition. $U(y)$ and $fU(f^{-1}y)$ intersect regularly. Regularity and nonregularity of intersections are explained (dimension 2) in Fig. 1.

Our dynamical systems are hyperbolic, and the hyperbolicity is close to uniform. But they are not continuous and the singularities eventually destroy the hyperbolicity, although hyperbolicity prevails over fractioning, and some ergodic and statistical properties can be proved.

We studied ergodic properties of dynamical systems with singularities in refs. 23 and 24. In these papers we prove that the K-property is satisfied if some additional conditions are imposed. In these discontinuous systems

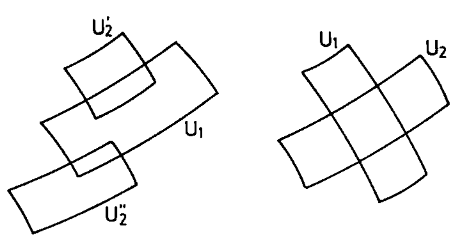

Fig. 1. Nonregular and regular intersections.

the elements of Markov partitions could have rather complicated shape: they can be nowhere dense, form totally disconnected sets of Cantor type, and their number can be uncountable (the symbolic system has infinite alphabet).

This and other circumstances led Bunimovich *et al.*[7] "to work out a new approach to the investigation of statistical properties of billiard systems." They substantially simplified previous approachs "constructing the symbolic system more 'roughly,' but with finite alphabet. For this we introduce a finite family of subsets of M which have the Markov property during a definite finite interval of time. ... This family of subsets does not cover all the phase space, hence we call it a Markov lattice" (Markov sieve).

This construction allows one to prove conditions on strong mixing analogous to those studied by Ibragimov and Linnik[13] and derive the rate of the decay of correlations.

Our paper includes only this construction because the probabilistic proofs of the main theorems are exactly the same of ref. 7: they used the structure and mixing properties of the Markov sieve, but not its mechanical origin.

We remark that one of the most important technical difficulties in the construction comes from the fact that we must use a kind of "local linearization" of the nonlinear dynamical system. A simplified use of the methods in refs. 6 and 7 is done in an interesting and clarifying paper by Chernov.[9] In these papers the most interesting numerical results are reviewed.

This paper is organized as follows. In Section 2 we describe the dynamical system and present some known results describing its hyperbolicity in terms of quadratic forms. In Section 3 we introduce conditions that are sufficient to construct the Markov sieve and state the main results. Section 4 includes the definitions of the "minimal geometrical structures" and how to control their measures. In Section 5 we define the "local linearization" and prove some of its important properties. In Section 6 we study transitivity properties of invariant manifolds. Sections 7–9 are devoted to the construction of the Markov sieve; Section 8 includes the definition and construction of pre-Markov partitions.

These generalizations of the theory created by Sinai, Chernov, and Bunimovich can be applied to the investigation of statistical properties of other mechanical systems. These results are valid in other (nonuniform) hyperbolic billiards, taking a smaller phase space, defining a "new" derived map, and studying the relations between the "new" singularities.

Chernov[12] has studied the same statistical properties for various classes of functions. The bounds he has obtained (for example, on correlations) are intimately related with asymptotic properties of certain types of partitions of the phase space.

Recently, Liverani[19] has studied and obtained results on the same problems using the Perron–Frobenius transfer operator.

2. PHASE SPACE. QUADRATIC FORMS. INVARIANT MANIFOLDS

Our phase space and map are defined in order to apply the generalization done by Katok and Strelcyn of Pesin's theory (see ref. 14, I.1.1, and refs. 23 and 24).

Let M be the union of a finite number of smooth Riemannian compact connected manifolds $M_1, M_2, \ldots, M_s$ (possibly with boundaries and with angles) all of dimension d. Here ρ is the Riemannian metric and $\|\cdot\|$ the Riemannian norm in TM. The boundaries of M_i are contained in S, the union of a finite number of C^1-compact submanifolds of positive codimension in $M_1, \ldots, M_s$. The set $N = M/S$ is an open subset of M, and μ is a probability measure on M such that its restriction to N is absolutely continuous with respect to the volume measure λ induced by ρ ($\mu = h\lambda$, where h is a bounded function). It is obvious that $\mu(N) = 1$.

Let $f: N \to M$ be a C^r-diffeomorphism ($r \geqslant 2$) between N and its image, which preserves the measure μ. If $H = \bigcap_{n=-\infty}^{+\infty} f^n N$, it results that $\mu(H) = 1$. We suppose that f can be extended as a C^r-function to S_0^+ and that f^{-1} can be extended as a C^r-function to S_0^-. The sets $S_0^\pm \subset S$ are C^1 compact submanifolds and $f_{|S_0^+}$ and $f^{-1}_{|S_0^-}$ are C^r-diffeomorphisms. We define

$$S_n = \{f^n x \in N: x \in S_0^+\}, \qquad n > 0$$

$$S_{-n} = \{f^{-n} x \in N: x \in S_0^-\}, \qquad n > 0$$

We will write S_0 instead of $S_0^\pm$ if they are associated to $S_{\pm n}$, $n > 0$, and there is no confusion. The set of singularities of f is $S_0^+ \cup S_0^-$. The discontinuity sets of f and f^{-1} are respectively contained in S_0^- and S_0^+.

Then, the discontinuity sets of f^n (f^{-n}) are contained in S_{-n+1} (S_{n-1}). For $m \leqslant n$, let be $S_{m,n} = S_m \cup S_{m+1} \cup \cdots \cup S_n$.

If $\exp_x$: $T_x M \to N$, $x \in N$, is the exponential map defined on

$$\{u \in T_x M: \|u\| < \min\{d(x), \text{radius of injectivity of } x \text{ in } M\}\}$$

we define $f_{0x} = \exp_{fx}^{-1} \circ f \circ \exp_x$. It is well defined in a neighborhood of $0 \in T_x M$. We will suppose that f satisfies the following condition concerning its growth[14]: there exist $c \geqslant 1$ and $b > 1$ such that $\|d^2 f_{0x}(h)\| < c[d(\exp h)]^{-b}$ for h in a small neighborhood of $0 \in T_x M$. We remark that the second derivative is a map between linear spaces.

We will assume also that $\log^+ \|(f^{\pm 1})'_x\| \in L^1(H, \mu)$ ($\log^+ s = \max\{\log s, 0\}$), a condition that is needed to apply the ergodic multiplicative theorem of Oseledets.

We say that x is a regular (Oseledets) point of f if there exist numbers $\lambda_1(x) > \cdots > \lambda_m(x)$ and a decomposition $T_x M = E_1(x) \oplus \cdots \oplus E_m(x)$ such that

$$\lim_{n \to \pm\infty} \frac{1}{n} \log \|(f^n)'_x w\| = \lambda_i(x)$$

for every $0 \neq w \in E_i(x)$ and every $1 \leqslant i \leqslant m(x)$. $E_i(x)$ is the proper subspace of the Lyapunov exponent $\lambda_i(x)$. $\Sigma = \Sigma(f)$ will denote the Pesin region, that is, the set of regular points that have only nonzero exponents. If $\mu(\Sigma) = 1$, we will say that the map f (or the dynamical system defined by it) is (nonuniformly) hyperbolic or has chaotic behavior.

The function $Q: TM \to \mathbb{R}$ is a quadratic form on $U \subset M$ if $Q_x: T_x M \to \mathbb{R}$ is a quadratic form in the usual sense for each $x \in U$. If $f: U \to U$ is a $\mathbb{C}^1$ map, we denote by $f^{\#}Q$ (pullback of Q by f) the quadratic form $(f^{\#}Q)_x u = Q_{fx}(f'_x u)$, $u \in T_x M$. The quadratic form Q is nondegenerate on a subset $U \subset M$ if Q_x is nondegenerate for every $x \in U$, that is, if the associate matrix of Q_x in any base has nonzero eigenvalues. Q is positive on U if $Q_x u > 0$ for every $x \in U$ and every nonzero $u \in T_x M$.

For any quadratic form Q on the corresponding half orbit of $x \in M$ we define

$$S_x = \{u \in T_x M: Q((f^n)' u) < 0, n \geqslant 0\}$$

$$U_x = \{u \in T_x M: Q((f^n)' u) > 0, n \leqslant 0\}$$

For $x \in \Sigma(f)$ let

$$E^s_x = \bigoplus_{\lambda_i(x) < 0} E_i(x), \qquad E^u_x = \bigoplus_{\lambda_i(x) > 0} E_i(x)$$

We have proved[20,22] the following theorem establishing conditions for the existence of (nonuniform) hyperbolicity in terms of quadratic forms.

2.1. Theorem. Let $Q: TM \to \mathbb{R}$ be a quadratic form such that:

(i) Q_y depends continuously on y, and is nondegenerate, on $\mathbb{N}$.

(ii) $P_y = (f^{\#}Q - Q)_y$ is positive for every $y \in N$.

Then $\mu(\Sigma) = 1$, $S_x = E^s_x$, $U_x = E^u_x$, and they depend continuously on $x \in \Sigma \subset H$.

2.2. Remarks. (a) The main results can be obtained with measurable Q on M of any dimension. In ref. 22 it is proved that the existence of such Q gives a characterization of (nonuniform) hyperbolicity.

(b) A slightly stronger version of Theorem 2.1 can be proved in the same way; instead of condition (ii), we assume that P is eventually positive. $P \geqslant 0$ and for almost every $x \in H$ there exist $k \in \mathbb{N}$ such that

$$Q((f^{k+1})' u) - Q((f^k)' u) > 0 \qquad \text{for every nonzero } u \in T_x M$$

2.3. Definition. For every x where Q is defined, and $b > 0$,

$$C_x^{+b} = \{u \in T_x M \colon Q_x u > b \, \|u\|^2\}$$
$$C_x^{-b} = \{u \in T_x M \colon Q_x u < -b \, \|u\|^2\}, \qquad C_x^{\pm} = C_x^{\pm 0}$$

For every cone C, clos C is the union of C and the vectors in its boundary. These cones satisfy

$$C_x^+ \supset U_x, \qquad C_x^- \supset S_x, \qquad \operatorname{clos}(f' C_x^+) \subset C_{fx}^+$$

2.4. Definitions. If Q_x is defined, let $|u| = |Q_x u|^{1/2}$ for every $u \in T_x M$. This *pseudonorm* satisfies

$$|f'_x u| > |u| \qquad \text{for every} \quad u \in C_x^+$$
$$|f'_x s| < |s| \qquad \text{for every} \quad s \in C_x^-$$

For any C^1-curve $\mathscr{C}$ in M defined by $c\colon [t_1, t_2] \to M$, let $l(\mathscr{C}) = \int_{t_1}^{t_2} |Q_{c(t)}(c'(t))|^{1/2} \, dt$. This length allows us to define a *pseudometric* σ induced by $|\cdot|$ in M.

The Riemannian length is given by

$$\rho(\mathscr{C}) = \int_{t_1}^{t_2} \|c'(t)\| \, dt$$

2.5. Relations Between the Riemannian and l-Lengths. Then, we have that

$$\rho(\mathscr{C}) \simeq (t_2 - t_1) \, \|c'(t_0)\| \quad \text{with } t_0 \in [t_1, t_2]$$
$$l(\mathscr{C}) \simeq (t_2 - t_1) \, |Q_y(c'(\tilde{t}))|^{1/2} \quad \text{with } \tilde{t} \in [t_1, t_2], \quad c(\tilde{t}) = y \in M$$

If $\mathscr{C}$ is a "short" curve $(t_2 - t_1 \simeq 0)$, it results that

$$\rho(\mathscr{C}) \simeq \frac{l(\mathscr{C}) \, \|c'(t_1)\|}{|Q_x(c'(t_1))|^{1/2}}, \qquad x = c(t_1)$$

Finally, if v_x is the unit tangent vector to $\mathscr{C}$ at x, $v_x = c'(t_1)/\|c'(t_1)\|$, we have that

$$\rho(\mathscr{C}) \simeq \frac{l(\mathscr{C})}{|Q_x(v_x)|^{1/2}}$$

2.6. LUMs and LSMs. Pesin's construction of local invariant manifolds can be applied if the conditions of Theorem 2.1 are satisfied at almost every point of $x \in \Sigma$. See refs. 25 and 14 and ref. 24, §3, for a compact version of results.

Now, μ-almost every point $x \in M$ has a local unstable manifold (LUM) $\gamma^u(x)$ and a local stable manifold (LSM) $\gamma^s(x)$, both passing through x such that their tangent spaces are $T_x\gamma^s(x) = S_x$, $T_x\gamma^u(x) = U_x$. Due to the singularities these submanifolds can be small (in the Riemannian sense; if $\dim M = 2$, they are *short*) and there are plenty of arbitrarily small LUMs and LSMs throughout M. In particular, if $\tilde{\partial}\gamma^s(x)$ is a smooth part of the boundary of $\gamma^s(x)$, then

$$f^n(\tilde{\partial}\gamma^s(x)) \subset S_0^- \qquad \text{for some} \quad n \geqslant 0$$

For each $x \in \Sigma$ there exists a neighborhood $U(x)$ such that:

(a) $\gamma^s(y)$, $\gamma^u(y)$ are uniformly transversal: angles between E_y^s and E_y^u are greater than θ_0 for μ-almost $y \in U(x)$).

(b) The family of LSMs $\gamma^s(z)$ [LUMs $\gamma^u(z)$] of the same dimension and $|\lambda_i(z)| > r_0$ for some $r_0 > 0$ is absolutely continuous restricted to $U(x)$. Roughly speaking, this means that for any two submanifolds W_1 and W_2 transversal to the stable foliation, $\nu_{W_1}(A) = 0$ implies that $\nu_{W_2}(p(A)) = 0$. Here ν_W denotes the measure induced on W by the Riemannian metric ρ and p is the so-called holonomy map from W_1 to W_2 along the fibers $\gamma^s(y)$. For details see ref. 14, II.4.

3. STATEMENT OF CONDITIONS AND MAIN RESULTS

We now assume additional properties on the set of singularities, its evolution by f and f^{-1}, and on the hyperbolicity of the system (that is, on the quadratic form Q).

3.1. dim $M = 2$. From here on, we will assume that $d = \dim M = 2$. Then, $\gamma^s(y)$, $\gamma^u(y)$ are C^{r-1} curves and E_y^s, E_y^u have only one direction. All the conditions can be formulated for $d > 2$, but the proofs of the theorems —in their actual version—are valid only for $d = 2$. See Remark 3.14.

In T_xM, provided with the Riemannian inner product, we can take an orthonormal basis. If $(dr, d\varphi)$ are the coordinates of a vector with respect to this basis, then

$$Q_x(dr, d\varphi) = C(x)\, d\varphi^2 + D(x)\, dr\, d\varphi$$

and the unstable direction is close to $dr = 0$.

The sets and distances defined in the following paragraphs are needed to bound the measure of singularity regions.

3.2. Definitions. Suppose that conditions of Theorem 2.1 are satisfied. For any $y \in M$, $\mathscr{G}$ is a *decreasing* (*increasing*) submanifold by y if it is C^1, $y \in \mathscr{G}$, and tangent vectors at any point $w \in \mathscr{G}$ are in C_W^- (C_W^+). In some papers, decreasing (increasing curves were called contracting (expanding) curves.

A C^1-curve γ is m-decreasing (m-increasing) for $m \geqslant 1$ if f^m (f^{-m}) is continuous on γ and $f^m\gamma$ ($f^{-m}\gamma$) is decreasing (increasing). m-decreasing (increasing) curves do not intersect $S_{-m,0}$ ($S_{0,m}$).

3.3. Definitions. $S(y)$ and $U(y)$ are defined for μ-almost every $y \in M$. For such y, let $v_y^s \in S(y)$, $v_y^u \in U(y)$, $\|v_y^{s,u}\| = 1$, and $b = b_y = |Q_y v_y^s|/2$.

For any $A \subset M$, $\mathscr{D}^-(y, A)$ is the union of decreasing curves joining y with some point of A, such that tangent vectors at any of its points w are in C_w^{-b}.

The technical restriction on the decreasing curves (we do not take all the decreasing curves, but those whose tangent vectors are not too close to the boundaries of C_w^-) is required for proving ergodic properties in ref. 24.

Then we define a very well adapted "stable distance" σ^- from y to A on decreasing curves, with the $|\cdot|$ pseudonorm:

$$\text{if} \quad \mathscr{D}^-(y, A) = \varnothing, \qquad \sigma^-(y, A) = 1$$

$$\text{if} \quad \mathscr{D}^-(y, A) \neq \varnothing, \qquad \sigma^-(y, A) = \inf\{l(\mathscr{C}) \cdot \mathscr{C} \in \mathscr{D}^-(y, A)\}$$

$\mathscr{D}^+(y, A)$ and $\sigma^+(y, A)$ are defined using unstable spaces and increasing curves. Finally, we define

$$U_\varepsilon^\pm(A) = \{y \colon \sigma^\pm(y, A) \leqslant \varepsilon\}, \qquad U_\varepsilon(A) = U_\varepsilon^+(A) \cup U_\varepsilon^-(A)$$

In fact, $U_\varepsilon(A)$ is the ε-neighborhood of A with the "distance" defined by Q on increasing and decreasing curves.

3.4. Definitions. For any $i > 0$ and x in an increasing (decreasing) curve γ, we define the *local coefficient of expansion* (contraction) under the action of f^i as the limit, when $l(\gamma_x)$ goes to zero, of $l(f^i\gamma_x)/l(\gamma_x)$, where γ_x is any curve by x contained in γ. In an analogous way we can introduce local coefficients of contraction (expansion) of increasing (decreasing) curves under the action of f^{-i}. Here $\Lambda_i^{\mathrm{u}}(x)$ $[\Lambda_{-i}^{\mathrm{s}}(x)]$ will denote local coefficients of expansion for $\gamma^{\mathrm{u}}(x)$ $[\gamma^{\mathrm{s}}(x)]$ under the action of f^i (f^{-i}).

3.5. Remarks. If γ is an increasing (decreasing) curve, its local coefficient of expansion (contraction) under the action of f^j, $j>0$, is given by $[Q_{f^jx}((f^j)'_x u)/Q_x u]^{1/2}$, where u is a tangent vector to γ at x. Thi is a direct consequence of Definition 2.4 and the continuity of Q. In particular, let $Q_{jx}^{\mathrm{u,s}} = Q_{f^jx}(f^j)'_x v_x^{\mathrm{u,s}}$, where $j \in \mathbb{Z}$ and $v_x^{\mathrm{u,s}}$ is a unit tangent vector to $\gamma_x^{\mathrm{u,s}}$ at u (zero will be omitted if it is the first symbol in jx). Then, we have

$$\Lambda_j^{\mathrm{u}}(x) = (Q_{jx}^{\mathrm{u}}/Q_x^{\mathrm{u}})^{1/2}, \qquad \Lambda_j^{\mathrm{s}}(x) = (Q_{jx}^{\mathrm{s}}/Q_x^{\mathrm{s}})^{1/2}$$

3.6. Definitions. (In all these definitions x is in the set of points in which the symbol where it appears makes sense.) Let $c_1 > 0$, $d \geqslant 1$ be fixed numbers. We define, for every $m, n \in \mathbb{N}$,

$$\mathscr{F}_m^{\mathrm{u,s}} = \{x \colon |Q_x^{\mathrm{u,s}}| = 1/m\}, \qquad \mathscr{F} = \bigcup_n (\mathscr{F}_n^{\mathrm{u}} \cup \mathscr{F}_n^{\mathrm{s}})$$

$$F_m^{\mathrm{u,s}} = \{x \colon 1/(m+1) < |Q_x^{\mathrm{u,s}}| \leqslant 1/m\}$$

3.7. Conditions. We now introduce new conditions on our non-uniformly hyperbolic dynamical systems. These conditions are sufficient to prove the main Theorems A and B. In Remarks 3.8–3.11 we explain the motivations for stating some of these conditions. We also indicate the main points of the proofs where they are used.

C1. The set Δ_n of double singularities [that is, the set of $x \in M$ for which there exist $n_1 \neq n_2$, $|n_i| \leqslant n$, such that $f^{n_i}(x) \in S_0$] is finite, for each $n \in \mathbb{N}$.

C2. Let $\hat{S}_1$ $(\hat{S}_{-1})$ be any C^1 piece of S_1 (S_{-1}). For any $y \in \hat{S}_1$ $(\hat{S}_{-1})$,

$$(+)\ T_y S_1 \subset \operatorname{clos} C_y^+$$

$$(-)\ T_y S_{-1} \subset \operatorname{clos} C_y^-$$

C3. (Sinai–Chernov Ansatz). For ν-almost every $y \in S_{-1}$ (S_1)

$$(+)\ \lim_{n\to+\infty} Q(f^n)'_y u = +\infty \qquad \text{for every} \quad u \in C_y^+$$

$$(-)\ \lim_{n\to-\infty} |Q(f^n)'_y s| = +\infty \qquad \text{for every} \quad s \in C_y^-$$

ν is the Riemannian measure in S_{-1} (S_1) induced by its normal.

C4. There exist $\theta > 0$, $K > 0$ such that $\mu(U_\varepsilon(S_0^+ \cup S_0^- \cup \mathscr{F})) \leqslant K\varepsilon^\theta$ for every $\varepsilon > 0$.

C5. There exists K_0 such that for any $m > 1$ the number of smooth components of $S_{-m,m}$ passing through or ending at some point $x \in M$ does not exceed $K_0 m$.

C6. There exist $c_1 \geqslant 1$, $\beta > 1$, and $n_0 \in \mathbb{N}$ such that:

(a) $\Lambda_1^{\mathrm{u}}(x) > \beta$ for every $x \in M \backslash S_{-1}$ and $\Lambda_{-1}^{\mathrm{s}}(x) > \beta$ for every $x \in M \backslash S_1$.

(b) For $m \geqslant n_0$ the sets F_m^{u} (F_m^{s}) are divided into two types: (b_1) Those for which there exists $d > 1$ such that $\Lambda_1^{\mathrm{u}}(x) > c_1 m^d$ for every $x \in F_m^{\mathrm{u}}$ $(\Lambda_{-1}^{\mathrm{s}} > c_1 m^d$ for every $x \in F_m^{\mathrm{s}})$; and (b_2) Those for which the previous inequalities are true for $d = 1$, and there exists $c_2 > 2$ satisfying $c_1^{-1} \log c_2 < 1$, such that each LUM (LSM) can intersect only finitely many of such F_m^{u} (F_m^{s}) with $M \leqslant m < c_2 M$, for every $M > 0$.

C7. There exists $L \in \mathbb{N}$ satisfying $\beta_1 = L\beta^{-1} < 1$, such that each LUM (LSM) intersects each $F_m^{\mathrm{u,s}}$ in at most L smooth curves.

C8. There exist real numbers $c > 0$, $0 < \delta < 1 < \eta$, such that $\eta - \delta > 1$, and if γ^{u} (γ^{s}) is any LUM (LSM) contained in F_m^{u} or F_m^{s} with boundary points x, y, then:

(a) $|Q_x^{\mathrm{u,s}} - Q_y^{\mathrm{u,s}}| \leqslant c l^\eta(\gamma^{\mathrm{u}})$ and $|Q_x^{\mathrm{u,s}} - Q_y^{\mathrm{u,s}}| \leqslant c l^\eta(\gamma^{\mathrm{s}})$.

(b) $|D(y)/D(x) - 1| < c l^\delta(\gamma^{\mathrm{u,s}})$ (see 3.1).

(c) $|h(x) - h(y)|/h(y) \leqslant c l^\delta(\gamma^{\mathrm{u,s}})$.

3.8. Remarks. (a) Conditions **C1–C3** jointly with $\mu(U_\varepsilon^-(S_0^-) \cup U_\varepsilon^+(S_0^+)) < K\varepsilon^\theta$ (K, ε, and θ as in **C4**) are sufficient to prove the K-property for our system. In ref. 24 we state this result with $\theta = 1$, but the restriction is unnecessary, as can be seen following the proof of the so-called Tail Bound Theorem. See ref. 23, §5, and ref. 16, §6.

(b) The present fourth condition includes a new set whose ε-neighborhood (with the pseudometric) must be ε^θ-small (with the μ-measure).

3.9. Remarks. (a) **C1** can be presented in other versions. For example, we can impose the transversality between S_0 and $S_{\pm n}$. Indeed, this assumption (jointly with **C2** and the continuous dependence of $C_x^\pm$ on x) implies **C1** and could be simpler to prove.

(b) We use S_1 (S_{-1}) instead of $S_0^\pm$ in **C2–C3** only to avoid some unimportant technical problems in the definition of Q_y for $y \in S_0^\pm$.

(c) **C2** can be deduced from the following 2′ condition: $T_y S_0^+ \subset \operatorname{clos} C_y^+$ ($T_y S_0^- \subset \operatorname{clos} C_y^-$) for any y in C^1-smooth pieces of $S_0^+(S_0^-)$. But this condition is in general not satisfied. Condition **C2** is commonly fulfilled and very simple to check.

(d) If we want to prove (nonuniform) hyperbolicity, we need exponential growth of Q μ-almost everywhere. **C3** states a weaker condition on this growth at points of singular submanifolds. It seems to be the weakest requirement that allows us to prove local ergodicity in our systems with "good" quadratic forms. Our condition **C3** is like the weaker Ansatz established by Sinai and Chernov for semidispersing billiards (ref. 28, p. 194). Chernov[10] modified this presentation because his "monotone" metric does not satisfy any continuity condition: he must claim the increasing property all over a neighborhood of $y \in S_{-1}$.

(e) The bound in **C5** for the number of discontinuity curves passing through each point is required for the construction of pre-Markov partitions. This condition was introduced for billiard systems in ref. 6.

3.10. Remarks. (a) **C6**(a) establishes a kind of uniform hyperbolicity. This restriction is not necessary to prove the K-property of our systems, but in the construction of the Markov sieve, we will use it repeatedly. It seems to be difficult to prove good rates of mixing not using some kind of uniform hyperbolicity. That is, we suppose that the nonvanishing of Lyapunov exponents is not enough to obtain good bounds on the decay of correlation.

(b) This restriction on the nonuniform hyperbolicity can be avoided by reducing the phase space and introducing a derived map $\tilde{f}$ related to the original f. This derived map $\tilde{f}$ may have more singularities and both must satisfy conditions of Theorem 2.1 and **C1–C8**. Smaller phase spaces and derived maps were naturally introduced for dispersing (Sinai) billiards with neutral components (segments) of the boundary. The "new" singularities were particularly well studied in the case of Bunimovich's stadium (see, for example, refs. 2 and 3). In a circular billiard table, the Pesin region of the billiard map has zero measure and the coefficients of expansion and contraction of f^n (for any n) are not bounded away from one. But when two semicircles are separated—creating a stadium—a derived map can be introduced considering only the last hit in each semicircumference.

(c) **C6**(b) establishes that the values of the expansion (contraction) rates must be very large at points where the value of the quadratic form is very small in unstable (stable) directions. This condition is very important because, for example, the strips F_m^{u} might be very narrow and then the unstable fibers might be fractioned too much by $\mathscr{F}_m^{\mathrm{u}}, \mathscr{F}_{m+1}^{\mathrm{u}}$. This is not good because we need some kind of uniformity in the lengths of our fractioned manifolds (see 5.1 and 6.4). This problem is avoided by assuming that they expand drastically in $\mathscr{F}^{u}_{m}$ (by more than $c_1 m^d$).

(d) Even more, if the exponent d is equal to one, additional conditions are needed. We have selected the condition satisfied by Bunimovich's stadium (see ref. 7, Lemma 2.3 and Appendix 3). But they may be presented in other forms that are satisfied by specific dynamical systems, and are also useful for proving the last part of Theorem 6.3. In ref. 7, Appendix 3, a second version of the needed condition is used.

(e) **C7** also plays an important role in the proof of Theorem 6.3. Usually $L = 1$, and so β must only be greater than one.

3.11. Remarks. (a) Conditions included in **C8** are related to Hölder-regularity of the quadratic form Q and the density h of the measure. They are very important in the proof of Theorem 5.4. They allow us to apply a kind of mean value theorem for the measure of homogeneous parallelograms (Lemma 4.8).

(b) N. Chernov kindly provided a simple example which shows that regularity conditions on Q are necessary to prove the estimations contained in 5.6 and 5.7. Consider the baker transformation on the unit square M, defined by $f(x, y) = (2x, y/2)$ for $x \leqslant 1/2$ and $f(x, y) = (2x - 1, y/2 + 1/2)$ for $x > 1/2$. Then

$$Df = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

f is hyperbolic and preserves the Lebesgue measure. Let Q be the quadratic form $Q_p(u) = q(p)(u_x^2 - u_y^2)$, $u = (u_x, u_y) \in T_p M$, with q a continuous function with a small range: $0.9 < q(p) < 1.1$. If q oscillates wildly, the estimates in 5.6 and 5.7 may fail.

3.12. The Cardioid. In focusing components of hyperbolic billiards, the rates of expansion and contraction are not uniform: they degenerate on short trajectories.

In these cases, the space M is obtained by eliminating some curvilinear triangles. The lines $\varphi = \pm\pi/2$ are sides of these triangles (φ is the angle between the interior normal to the boundary and the vector of the trajectory, and r is the arc length of the boundary; see ref. 24, §5).

Restricting our analysis to the billiard in the cardioid, we observe that the main computations needed to verify conditions **C1–C4** were done in ref. 24, §6.

The phase space must be reduced in such a way that in the "new" M, the "new" derived map $\tilde{f}$ satisfies **C6**(a), with the quadratic form Q used in the verification of the previous conditions. If f is the billiard map, we obtain $\tilde{f}$ as a finite power of f: $\tilde{f}(x) = f^{k(x)}(x)$ with k depending on the phase point $x \in M$. It is simple to show that **C5** is satisfied by this $\tilde{f}$. It is not difficult to verify conditions **C7** and **C8**, since

$$Q_x(dr, d\varphi) = 2\cos\varphi(x)\, dr\, d\varphi - [\cos\varphi(x)/K(x)]\, d\varphi^2$$

$$h(x) = \cos\varphi(x)$$

We suppose that it may be arduous to verify other conditions of 3.7 for this $\tilde{f}$. The difficulties seem to be similar to, but are not the same as, those that appear in the stadium. In the latter case there is no hyperbolicity between bounces on the semicircumferences; in the cardioid such (non-uniform) hyperbolicity exists.

It is very important to remark that β in **C6**(a) can be as close to 1 as we want. Then, the reduced space M can be close to the original phase space as much as we need.

3.13. Statistical Properties. Since μ is an invariant measure for f, every measurable function F on M defines a stationary stochastic process with discrete time: $X_n = F \circ f^n(x)$, $n \in \mathbb{Z}$. The main results of this paper are on statistical properties of these processes that play important roles in applications of hyperbolic dynamical systems with singularities. The results refer to mixing rates of the process and include, as a consequence, a central limit theorem.

Here $\langle \cdot \rangle$ will denote the averaging with respect to the measure μ: $\langle F \rangle = \int F(x)\, d\mu(x)$. Let $\mathscr{L}$ denote the space of complex-valued *piecewise* Hölder continuous functions on M. If $M = \bigcup_{i=1}^{k} N_i$, where the sets N_i, $i = 1, 2, \ldots, K$, are separated by a finite union of compact smooth curves, $\mathscr{L}$ is defined by

$$\mathscr{L} = \{F\colon M \to \mathbb{C}\colon |F(x) - F(y)| \leqslant C_f[\rho(x, y)]^\alpha \text{ for some } \alpha > 0$$
$$\text{and any } x, y \in N_i,\ i = 1, 2, \ldots, k\}$$

For example, N_i can be the domains where $f^{\pm m}$ are continuous for fixed $m > 0$.

Theorem A (Rate of decay of correlations). If the conditions of Theorem 2.1 and **C1–C8** are satisfied on a two-dimensional M, then, for any two functions $F, G \in \mathscr{L}$ and any $n \geqslant 1$,

$$|\langle F \circ f^n \cdot G \rangle - \langle F \rangle \langle G \rangle| \leqslant C(F, G) \exp(-a\sqrt{n})$$

where $C(F, G) > 0$ depends on F and G, and $a > 0$ is a constant.

Theorem B (Central limit theorem). Under the conditions of Theorem A, for any $F \in \mathscr{L}$ and $n \in \mathbb{N}$, let $S_n(F) = \sum_{i=0}^{n-1} F \circ f^i$. Then $\sum_{i=-n}^{n} [\langle F \circ f^i \cdot F \rangle - \langle F \rangle^2]$ converges to $\sigma^2 = \sigma^2(F) \geqslant 0$ and, if $\sigma \neq 0$,

$$\mu\left(\left\{\frac{S_n(F) - n\langle F \rangle}{\sigma\sqrt{n}} < z\right\}\right) \to \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{z} e^{-u^2/2}\, du \qquad \text{as} \quad n \to +\infty$$

(convergence in distribution to the standard normal law).

3.14. Remarks. (a) The decay of correlations is only subexponential. Everyone would like to have an exponential decay ($<\lambda^n$), but it seems that the influence of singularities slows down the decay of correlations. The influences and relations between hyperbolicity and singularities were discussed in ref. 9, §1, and ref. 7, §1.

(b) The quantity σ^2 of Theorem B is equal to zero if and only if the function F is cohomologous to zero (coboundary); this means that there exists $G \in \mathscr{L}^2(M \cdot \beta, \mu)$ such that $F = G \circ f - G$. See ref. 13, Chapter 18, §2.

3.15. Remarks. Kolya Chernov has sketched a proof of the same statistical properties for a multidimensional Lorentz gas.[11] In this case condition **C5** must be strengthened (number of smooth components of $S_{-m,m}$ meeting at any point $x \in M$ cannot exceed a constant K_0). The new condition is not satisfied by other usual billiard systems. I suppose that his proof works in our "abstract multidimensional systems" with this modified condition.

The main modifications in the proofs concern the theorems on the evolution of locally invariant manifolds and the construction of Markov sieves (this is coarser and simpler than the one done for two-dimensional, nonuniformly hyperbolic systems).

4. MEASURE OF PARALLELOGRAMS

If $x, y \in M$, we will always assume that $\gamma^{\mathrm{u}}(y) \cap \gamma^{\mathrm{s}}(z)$ consists of at most one point. If the LUMS or LSMs are too long, we introduce some more dividing curves in M. If it exists, we put $\gamma^{\mathrm{u}}(y) \cap \gamma^{\mathrm{s}}(z) = [y, z]$. For B, $C \subset M$, let $[B, C] = \{[y, z]: y \in B,\ z \in \mathbb{C}\}$ and $\gamma_A^{\mathrm{u,s}}(x) = A \cap \gamma^{\mathrm{u,s}}(x)$.

4.1. Definition. A *parallelogram* is a subset $A \subset M$ such that $\mu(A) > 0$ and for any two points $y, z \in A$ the point $[y, z]$ exists and again belongs to A.

A parallelogram is obtained intersecting a family $\{\gamma^{\mathrm{u}}\}$ of LUMs with a family $\{\gamma^{\mathrm{s}}\}$ of LSMs so that each γ^{u} intersects each γ^{s}: the parallelogram has the structure of a direct product and A can be represented in the form

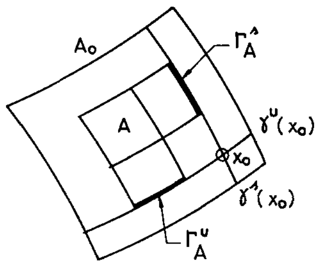

Fig. 2. Parallelogram A in an "ambient" parallelogram A_0, and the coordinate axes $\gamma^u(x_0)$, $\gamma^s(x_0)$.

$A = [\gamma^s_A(x_0), \gamma^u_A(x_0)]$ for an arbitrary point $x_0 \in A$. If x_0 is fixed, $\gamma^u(x_0)$, $\gamma^s(x_0)$ will be called coordinate axes of A_0 (see Fig. 2). As a consequence of the Pesin method of constructing invariant manifolds the limit (in the C^0 topology) of a sequence of LUMs and LSMs can only be a LUM or a LSM. Then, the closure of any parallelogram is again a parallelogram.

4.2. Definition. We say that a subparallelogram $C \subset A$ is *u-inscribed* (*s-inscribed*) in A if

$$\gamma^u_C(x) = \gamma^u_A(x) \quad [\text{resp. } \gamma^s_C(x) = \gamma^s_A(x)] \quad \text{for every } x \in C$$

4.3. Definition. A *quadrilateral* is any domain Q in M, bounded by two LUMs (called the u sides, whose union is $\partial^u Q$) and two LSMs (s sides, whose union is $\partial^s Q$), in alternation.

4.4. Definition. For any parallelogram A the minimal closed quadrilateral containing A is called the *support* (*carrier*) of A; it is denoted by $K(A)$ and the u and s faces of A are taken to be those of its support.

4.5. Definitions. We say that the LUM γ^u (LSM γ^s) is *stretched* on the quadrilateral K (parallelogram A) if its endpoints are precisely on the s faces (u faces) of this quadrilateral (parallelogram).

A parallelogram is said to be *maximal* if it intersects all the LUMs and LSMs stretched on its support $k(A)$.

4.6. Measure of Parallelograms. The μ-measure of sufficiently small parallelograms are approximately equal to $\mu(A) \sim \lambda(A)\, h(x)$ for some point $x \in A$ (see the first paragraphs of Section 2).

Suppose now that $\psi^u(x) - \psi^s(x)$ is the angle between $\gamma^u(x)$ and $\gamma^s(x)$ at the point x, measured with the Riemannian inner product. This means that $\langle v_x^{\mathrm{u}}, v_x^{\mathrm{s}}\rangle = \cos[\psi^{\mathrm{u}}(x) - \psi^{\mathrm{s}}(x)]$ (see 3.3). If the point $x \in A$ is fixed, we may write

$$\lambda(A) \simeq \rho(\gamma_A^{\mathrm{u}}(x))\, \rho(\gamma_A^{\mathrm{s}}(x))\, |\sin[\psi^{\mathrm{u}}(x) - \psi^{\mathrm{s}}(x)]|$$

Using the final relation in 2.5, we obtain

$$\mu(A) \cong \frac{l(\gamma_A^{\mathrm{u}}(x))\, l(\gamma_A^{\mathrm{s}}(x))}{(Q_x v_x^{\mathrm{u}})^{1/2}\, |Q_x v_x^{\mathrm{s}}|^{1/2}} |\sin[\psi^{\mathrm{u}}(x) - \psi^{\mathrm{s}}(x)]|\, h(x)$$

Consider now a parallelogram A included in an "ambient" parallelogram A_0. For a fixed $x_0 \in A_0$ we consider the canonical projections $\Gamma_A^{\mathrm{u}}, \Gamma_A^{\mathrm{s}}$, of A onto the coordinate axes $\gamma^{\mathrm{u}}(x_0)$, $\gamma^{\mathrm{s}}(x_0)$. Once more we have $A = [\Gamma_A^{\mathrm{s}}, \Gamma_A^{\mathrm{u}}]$.

If we partition $\gamma^{\mathrm{u}}(x_0)$ and $\gamma^{\mathrm{s}}(x_0)$ into subsegments Δ_i^{u} and Δ_j^{s}, their "direct product" $[\Delta_i^{\mathrm{s}}, \Delta_j^{\mathrm{u}}]$ gives a partition of A into parallelograms $\Delta_{ij} = [\Delta_i^{\mathrm{s}} \cap \Gamma_A^{\mathrm{s}}, \Delta_j^{\mathrm{u}} \cap \Gamma_A^{\mathrm{u}}]$.

Making up the integral sum $\mu(A) = \sum \mu(\Delta_{ij})$ and passing to the limit, we obtain

$$\mu(A) = \int_A \frac{|\sin[\psi^{\mathrm{u}}(x) - \psi^{\mathrm{s}}(x)]|}{(Q_x^{\mathrm{u}})^{1/2}\, |Q_x^{\mathrm{s}}|^{1/2}} h(x)\, d\mu_A(x)$$

where $d\mu_A(x)$ denotes the derivative of the measure on A which is the direct product of the l-measures on $\gamma_A^{\mathrm{u}}(x)$ and $\gamma_A^{\mathrm{s}}(x)$.

The absolute continuity of stable and unstable families implies that almost every point of the set $\gamma_A^{\mathrm{u}}(x)$ is a density point on $\gamma^{\mathrm{u}}(x)$ and thus the Jacobian of the canonical isomorphism P (Poincaré map) of this set onto its projection $\Gamma_A^{\mathrm{u}} \subset \gamma^{\mathrm{u}}(x_0)$ is defined on it. We refer this Jacobian to the l-length and denote it by $J^{\mathrm{u}}(x)$ [$J^{\mathrm{s}}(x)$ in the stable case]. Then the last expression for $\mu(A)$ takes the form

$$\mu(A) = \int_{\Gamma_A^{\mathrm{u}}} dl(y) \int_{\Gamma_A^{\mathrm{s}}} dl(z)\, B(x)\, J^{\mathrm{u}}(x)\, J^{\mathrm{s}}(x)$$

where $y \in \Gamma_A^{\mathrm{u}}$, $z \in \Gamma_A^{\mathrm{s}}$ are points such that $x = [z, y]$, and

$$B(x) = \frac{|\sin[\psi^{\mathrm{u}}(x) - \psi^{\mathrm{s}}(x)]|}{(Q_x^{\mathrm{u}})^{1/2}\, |Q_x^{\mathrm{s}}|^{1/2}} h(x)$$

We remark that in the same way in which we considered $\psi^{\mathrm{u}}(x) - \psi^{\mathrm{s}}(x)$ as the *angle* between the LUM and LSM by x, we can regard y and z as the *coordinates* of x on the axes $\gamma^{\mathrm{u}}(x_0)$ and $\gamma^{\mathrm{s}}(x_0)$.

The last formula gives the measure of a parallelogram as a product on the metrical sense. However, it will be only useful if we can apply a kind of mean value theorem; that is, if the function that is integrated is almost constant. For this purpose we introduce the notion of weak homogeneity of parallelograms. Let $\alpha_0 < 1$ and C_0 be positive constants.

4.7. Definition. A parallelogram A_0 is called *weakly n-homogeneous* ($n \in \mathbb{N}$) if the following conditions are satisfied:

(i) $|B(x)/B(y) - 1| \leqslant C_0 \alpha_0^n$

(ii) $|J^{u,s}(x) - 1| \leqslant C_0 \alpha_0^n$

for any points $x, y \in A_0$ and any point $x_0 \in A_0$ fixing the coordinate axes $\gamma^u(x_0)$, $\gamma^s(x_0)$.

The next lemmas follow immediately from the formula for $\mu(A)$ and 4.7.

4.8. Lemma. The measure of subparallelograms A of a weakly n-homogeneous parallelogram A_0 can be approximated by $\mu_a(A) = l(\Gamma_A^u)\, l(\Gamma_A^s)\, B(x_0)$.

This approximation satisfies $|\mu_a(A)/\mu(A) - 1| \leqslant C_1 \alpha_0^n$, where $C_1 = C_1(\alpha_0, C_0)$.

4.9. Lemma. Let be A a weakly n-homogeneous parallelogram, A' a subparallelogram u-inscribed in A, and A'' a subparallelogram s-inscribed in A. Then

$$\left| \frac{\mu(A'' \cap A)}{\mu(A)} \cdot \frac{\mu(A')}{\mu(A'' \cap A')} - 1 \right| \leqslant C_2 \alpha_0^n$$

where $C_2(\alpha_0, C_1)$ (conditional measures are close).

5. HOMOGENEOUS LUMS AND LSMS. HOMOGENEOUS PARALLELOGRAMS

We want to describe some definite types of weakly n-homogeneous parallelograms, to which Lemmas 4.8 and 4.9 are applicable. For this purpose it is not enough that the parallelograms have small diameters: it is necessary that the function $B(x)$ does not strongly oscillate [and $J^{u,s}(x)$ do not differ too much from one] inside A and $f^i A$ for small values of $|i|$.

5.1. Definition. A closed segment γ^u (γ^s) of LUM (LSM) is *homogeneous* if each image $f^{-i}\gamma^u$ ($f^i\gamma^s$) is contained in the cosure of some F_m^u (F_m^s), $m = m(i)$, $n = n(i)$, for every $i \geqslant 0$. That is, when moving in the direction of contraction, $f^j\gamma^{u,s}$ "jumps" from one $\bar{N}_m$ to another one.

In other words, the construction of an HLUM (HSLM) can be realized by breaking up an arbitrary LUM (LSM) at points whose dividing image in the past (future) fall in the set of subdividing curves $\mathscr{F}_m^u$ ($\mathscr{F}_m^s$).

5.2. Definition. For $n \in \mathbb{N}$ we will say that a LUM γ^{u} (LSM γ^{u}) is *n-homogeneous* if its image $f^n\gamma^{\mathrm{u}}$ ($f^{-n}\gamma^{\mathrm{s}}$) is an HLUM (HLSM).

An n-homogeneous LUM γ^{u} has each image $f^i\gamma^{\mathrm{u}}$, $i \in \mathbb{Z}$, $i \leqslant n$, in the closure of some F^{u}_m.

5.3. Definition. A parallelogram A is n-homogeneous if for any point x of it the set $\gamma^{\mathrm{u}}_A(x)$ $[\gamma^{\mathrm{s}}_A(x)]$ lies entirely on one n-homogeneous LUM (LSM).

The closure in M of any n-homogeneous parallelogram is also an n-homogeneous parallelogram (see 4.1 and 2.6).

The following theorem establishes the expected relation between n-homogeneity and weak n-homogeneity.

5.4. Theorem. n-homogeneous parallelograms are also weakly n-homogeneous for convenient choices of constants in Definition 4.7.

The proof of this theorem includes evaluations of local coefficients of expansion (contraction) and of the Jacobians $J^{\mathrm{u,s}}(x)$, and will be given in 5.7.

5.5. Proposition. Let $J^{\mathrm{u}}(x, y)$, $y \in \gamma^{\mathrm{s}}(x)$, be the Jacobian of the canonical isomorphism of the LUMs $\gamma^{\mathrm{u}}(x)$ and $\gamma^{\mathrm{u}}(y)$ with respect to the measure l [remember that in 4.6, $J^{\mathrm{u}}(x, P(x))$ was denoted by $J^{\mathrm{u}}(x)$]. Then

$$J^{\mathrm{u}}(x, y) = \lim_{i \to +\infty} [\Lambda^{\mathrm{u}}_i(x)/\Lambda^{\mathrm{u}}_i(y)]$$

where $\Lambda^{\mathrm{u}}_i(x)$ is the coefficient of expansion of $\gamma^{\mathrm{u}}(x)$ under the action of f^i.

Proof. Classical proofs of the absolute continuity of invariant manifolds include the evaluation of the Jacobian of the Poincaré map. See, for example, §5.4 in ref. 1 for a proof of the following formula:

$$J^{\mathrm{u}}(x, y) = \lim_{n \to \infty} \frac{\Lambda^{\mathrm{u}}_1(x)\, \Lambda^{\mathrm{u}}_1(fx) \cdots \Lambda^{\mathrm{u}}_1(f^n x)}{\Lambda^{\mathrm{u}}_1(y)\, \Lambda^{\mathrm{u}}_1(fy) \cdots \Lambda^{\mathrm{u}}_1(f^n y)}$$

The proposition is a consequence of these considerations and Remarks 3.5. ∎

5.6. Lemma. If the LUM γ^{u} is n-homogeneous, then for any $x, y \in \gamma^{\mathrm{u}}$ and $i > 0$, the estimate $|\Lambda^{\mathrm{u}}_{-i}(x)/\Lambda^{\mathrm{u}}_{-i}(y) - 1| \leqslant C_4\alpha_0^n$ holds for local coefficients of contractions under the action of f^{-i}, where C_4 and $\alpha_0 < 1$ are positive real constants. A similar estimate holds for n-homogeneous LSMs.

Proof. As a consequence of 3.5 it is sufficient to prove

$$\left| \frac{(Q^{\mathrm{u}}_{-ix})^{1/2}}{(Q^{\mathrm{u}}_{-iy})^{1/2}} \frac{(Q^{\mathrm{u}}_y)^{1/2}}{(Q^{\mathrm{u}}_x)^{1/2}} - 1 \right| \leqslant C_4\alpha_0^n$$

Since γ^u is homogeneous, $f^{-i}\gamma^u$ is $i+n$-homogeneous, and it is enough to prove that $|[(Q^u_y)^{1/2}-(Q^u_x)^{1/2}]/(Q^u_x)^{1/2}| \leqslant C_5\alpha_0^n$, because, if this is the case, $|(Q^u_{-ix})^{1/2}/(Q^u_{-iy})^{1/2}-1| \leqslant C_5\alpha_0^{n+i}$. But

$$\left|\frac{(Q^u_y)^{1/2}-(Q^u_x)^{1/2}}{(Q^u_x)^{1/2}}\right| \leqslant \frac{|Q^u_y-Q^u_x|}{(Q^u_x)^{1/2}} \leqslant \frac{cl^\eta(\gamma)}{(Q^u_x)^{1/2}}$$

where γ is the part of γ^u bounded by the points x, y (see **C8**). Now $1/(Q^u_x)^{1/2} < (m+1)^{\nu/2}$ and the lemma will be proved if we show this claim: $(m+1)^{\nu/2} \leqslant c_2 l^{\delta-\eta}(\gamma)$ for some $\delta<1$. Indeed, if $|\Lambda^u_{-i}(x)/\Lambda^u_{-i}(y)-1| \leqslant c_4 l^\delta(\gamma)$, condition **C6**(a) implies that $l(\gamma) \leqslant \beta^{-n} l(f^n\gamma)$, and the result is true for $\alpha_0=\beta^{-\delta}$ (we suppose that the l-lengths of LUM are bounded by one). To prove the claim we observe that $l(\gamma) \leqslant c_3(Q^u_z)^{1/2} \leqslant c_3 m^{-\nu/2}$; then we have $l^{\delta-\eta}(\gamma) \geqslant c_3^{\delta-\eta} m^{\nu/2(\eta-\delta)}$ and $l^{\delta-\eta}(\gamma)\, c_3^{\eta-\delta} \geqslant (m+1)^{\nu/2}$ if $\eta-\delta>1$. (The δ in this proof may not be the δ in **C8**; the important point is this last relation.) ▮

5.7. Proof of Theorem 5.4. We must prove that (i) and (ii) in Definition 4.7 are satisfied if A is an n-homogeneous parallelogram. For this purpose it is sufficient to prove:

(a) $|\Lambda^s_i(x)/\Lambda^s_i(y)-1| \leqslant C_0\alpha_0^n$ for $i \geqslant 0$

(b) $|B(x)/B(y)-1| \leqslant C_0\alpha_0^n$

for x, y in a single n-homogeneous LSM γ [$\Lambda^s_i(z)$ is the local coefficient of contraction under the action of f^i].

The proof of (a) is essentially a repetition of 5.6.

The proof of (b) is more subtle. Let $t(z)=|\sin[\psi^u(z)-\psi^s(z)]|$ (see 4.6) and $(\text{“}z\text{”})^{1/2}=(Q^u_z\,|Q^s_z|)^{1/2}$. Then

$$\left|\frac{B(x)}{B(y)}-1\right| \leqslant \frac{|t(x)-t(y)|}{t(y)}\frac{h(x)}{h(y)}\frac{(\text{“}y\text{”})^{1/2}}{(\text{“}x\text{”})^{1/2}}+\frac{|h(x)-h(y)|}{h(y)}\frac{(\text{“}y\text{”})^{1/2}}{(\text{“}x\text{”})^{1/2}} + \frac{|(\text{“}x\text{”})^{1/2}-(\text{“}y\text{”})^{1/2}|}{(\text{“}x\text{”})^{1/2}} \qquad (1)$$

We can suppose (see 3.1) that

$$Q_x(dr, d\varphi)=C(x)\,d\varphi^2+D(x)\,dr\,d\varphi$$

and that the unstable direction is close to $dr=0$. If $v^s_x=(dr_s, d\varphi_s)$, $v^u_x=(dr_u, d\varphi_u)$, are the directions of $S(x)$, $U(x)$, respectively, then

$$\frac{Q^u_x}{d\varphi_u^2}-\frac{Q^s_x}{d\varphi_s^2}=D(x)\left(\frac{dr_u}{d\varphi_u}-\frac{dr_s}{d\varphi_s}\right) \cong D(x)\,\mathrm{tg}[\psi^u(x)-\psi^s(x)]$$

As a consequence of **C6**, the bounds for the first term in the second member of (1) must be studied carefully only when $t(x)$ and $t(y)$ are close to zero.

As $\|v_x^{\mathrm{u,s}}\| = d\varphi_{\mathrm{u,s}}^2 + dr_{\mathrm{u,s}}^2 = 1$ and $\psi^{\mathrm{u}}(x) - \psi^{\mathrm{s}}(x) \cong 0$, we have that

$$|\sin[\psi^{\mathrm{u}}(x) - \psi^{\mathrm{s}}(x)]| \cong \frac{|Q_x^{\mathrm{u}} - Q_x^{\mathrm{s}}|}{D(x)}$$

Then we have

$$\left|\frac{t(x) - t(y)}{t(y)}\right| \cong \left|\frac{D(y)}{D(x)} \frac{Q_x^{\mathrm{u}} - Q_x^{\mathrm{s}}}{Q_y^{\mathrm{u}} - Q_y^{\mathrm{s}}} - 1\right|$$

on F_m^{s} and from **C8**(b) it follows that the needed bounds for $|[t(x) - t(y)]/t(y)|$ are those of $|(Q_x^{\mathrm{u}} - Q_x^{\mathrm{s}})/(Q_y^{\mathrm{u}} - Q_y^{\mathrm{s}}) - 1|$. But

$$\begin{aligned}
\left|\frac{Q_x^{\mathrm{u}} - Q_x^{\mathrm{s}} - Q_y^{\mathrm{u}} - Q_y^{\mathrm{s}}}{Q_y^{\mathrm{u}} - Q_y^{\mathrm{s}}}\right| &\leqslant \left|\frac{Q_x^{\mathrm{u}} - Q_y^{\mathrm{u}}}{Q_y^{\mathrm{u}} - Q_y^{\mathrm{s}}}\right| + \left|\frac{Q_x^{\mathrm{s}} - Q_y^{\mathrm{s}}}{Q_y^{\mathrm{u}} - Q_y^{\mathrm{s}}}\right| \\
&\leqslant \frac{|Q_x^{\mathrm{u}} - Q_y^{\mathrm{u}}|}{Q_y^{\mathrm{u}}} + \frac{|Q_x^{\mathrm{s}} - Q_y^{\mathrm{s}}|}{|Q_y^{\mathrm{s}}|} \\
&\leqslant \frac{|Q_x^{\mathrm{u}} - Q_y^{\mathrm{u}}|}{|Q_y^{\mathrm{u}}|^{1/2}} + \frac{|Q_x^{\mathrm{s}} - Q_y^{\mathrm{s}}|}{|Q_y^{\mathrm{s}}|^{1/2}} \leqslant 2C_4\alpha_0^n
\end{aligned}$$

The second and third terms in (1) are bounded by similar expressions as a consequence of **C8**(c) and 5.6, respectively, jointly with condition **C6**(a). ∎

Theorem 5.4 indicates the way to construct weak n-homogeneous parallelograms. But we do not know if the needed HLUMs (HLUSs) exist: it may be that subdividing points can densely fill each LUM (LSM).

We can not only prove the existence of HLUM and HLSM, but estimate their distribution of lengths. Let $r^{\mathrm{ou}}(x)$ $[r^{\mathrm{os}}(x)]$ be the l-distance from the point $x \in M$ to the nearest endpoint of the maximal smooth segment $\gamma^{\mathrm{ou}}(x)$ $[\gamma^{\mathrm{os}}(x)]$ of HLUM (HLSM) containing x in its interior.

5.8. Theorem. $\mu(\{x \in M: r^{\mathrm{ou}}(x) \leqslant \varepsilon\}) \leqslant C_6 \varepsilon^{\theta}$, where $C_6 = C_6(v, m_0) > 0$ and $\theta = \theta(v, m_0) > 0$. A similar result holds for $r_{\mathrm{os}}(x)$.

5.9. Proposition. For μ-almost any $x \in M$, the subdividing points on the LUM $\gamma^{\mathrm{u}}(x)$ [LSM $\gamma^{\mathrm{s}}(x)$] partitioning this curve into individual HLUMs (HLSMs) can only accumulate at the endpoints of this curve.

Both proofs are almost the same as those of Theorem 3.10 and Proposition 3.11 in ref. 7. Some geometrical properties of scattering billiards that are used in these proofs (ref. 7, Appendix 2) are presented here as condition **C4**.

Formula (A2.1) in ref. 7 is satisfied by our dynamical systems with $\varepsilon_n = p_0 \beta^{-n}$ for some $p_0 > 0$.

6. EVOLUTION OF HLUM IN THE DIRECTION OF EXPANSION

In reading this section, the reader should consider 3.4 and 3.5 in ref. 7. Their first paragraphs describe exactly all the plan and the main ideas behind the statements and proofs: expansion prevails over fractionation caused by discontinuity and subdividing curves (S_0^- and $\mathscr{F}$).

All the results in this section are valid for HLSMs, with n replaced by $-n$; we study the expansion also on HLSMs.

We begin with a result on partitioning by discontinuities (ref. 6 and ref. 7, Appendix 3).

6.1. Theorem. Let γ^{u} be an arbitrary LUM and $D > 0$. For each $N \geqslant 0$ we choose on γ^{u} subsegments γ_i^N such that f^N is smooth on them, $f^n \gamma_i^N$ lies in a smooth component of $f^n \gamma^{\mathrm{u}}$ for any integer n, $0 \leqslant n \leqslant N$, and $l(f^n \gamma_i^N) \leqslant D$. Then there are numbers $D > 0$, $C > 0$, and $0 < \lambda < 1$ such that for all $N \geqslant 1$.

$$l\left(\bigcup_i \gamma_i^N\right) \leqslant C\lambda^N$$

Proof. We first write the obvious estimate $l(\bigcup_i \gamma_i^N) \leqslant D \sum \Lambda_N^{-1}(i)$, where $\Lambda_N(i)$ is a lower bound for the local coefficients of expansion $\Lambda_i^{\mathrm{u}}(x)$ under the action of f^N.

Since our systems satisfy **C6**(a), we may put $\Lambda_N(i) = \beta^N$. It remains to estimate the number of segments γ_i^N for a given N. Let $n \geqslant 1$ be a given integer and D sufficiently small $[D \leqslant D_0(n)]$. Then **C5** implies that any LUM of length less than or equal to D intersects at most $K_0 n$ discontinuity curves in $S_{-n,0}$, and the number of segments γ_i^N does not exceed $(K_0 n)^{[N/n]+1}$. Then

$$l\left(\bigcup_i \gamma_i^N\right) \leqslant D(K_0 n)^{[N/n]+1} \beta^N$$

We choose n such that $\beta(n) = \beta(K_0 n)^{-1/n}$ is larger than 1. For this n let $\lambda^{-1} = \beta(n)$. ∎

We will now obtain a result analogous to 6.1 for HLUMs. We begin with a definition.

6.2. Definition. Let γ^u be an arbitrary HLUM of length $l(\gamma^u) = p$, and D any positive number. For each $n \geqslant 0$, $\gamma^u_{i,n}$, $i = 1, 2, \ldots$, are the subsegments of γ^u that are sent by f^n into homogeneous segments of l-length greater than or equal to D. Then we define the relative fraction of points in γ^u that in the first N steps are sent at least once in an HLUM of l-length greater than or equal to D:

$$\gamma_N = \bigcup_n \left\{ \bigcup_i \gamma^u_{i,n} : n \leqslant N \right\}; \qquad p_{D,\gamma^u}(N) = l(\gamma_N)/l(\gamma^u)$$

6.3. Theorem (Rate of expansion). There are a positive number D and a function $\beta(c)$, $\beta(c) \to 0$ as $c \to +\infty$, independent of γ^u and its length $l(\gamma^u) \leqslant p_0 < 1$, such that for any $c > 0$ we have

$$p_{D,\gamma^u}(N) > 1 - \beta(c) \qquad \text{for} \quad N = -c \log p$$

Proof. By analogy with 6.1, we denote by γ_i^N the subsegments of γ^u such that $f^n \gamma_i^N$ is an HLM and $l(f^n \gamma_i^N) \leqslant D$ for every $0 \leqslant n \leqslant N$. The estimate at the beginning of the proof of 6.1 remains valid. The number of segments γ_i^N can be infinite for subdivision by $\mathscr{F}^u_m$ with $m \geqslant m_1$. This number m_1 will be determined below.

Using **C6** and **C7**, we will obtain an estimate for $\Lambda_N(i)$. We assume that the images of the LUM γ_i^N during the first N iterations fall k_1 times in sets F^u_m of the first type [see **C6**(b)], with indices $m \geqslant m_1 \geqslant n_0$; and k_2 times in sets F^u_m of the second type, also with indices $m \geqslant m_1 \geqslant n_0$. We denote the corresponding indices by $l_1, l_2, \ldots, l_{k_1} \geqslant m_1$ and $j_1, j_2, \ldots, j_{k_2} \geqslant m_1$.

Then, the following estimate holds:

$$\Lambda_N(i) \leqslant \beta^{N - k_1 - k_2} c_1^{k_1 + k_2} \prod_{k=1}^{k_1} l_k^{\alpha} \prod_{k=1}^{k_2} j_k$$

We have not taken into consideration the influence of the discontinuity curves and subdividing set $\mathscr{F}^u_m$ for $m < m_1$. Suppose that an LUM of length D and its future images during n iterations do not intersect any $\mathscr{F}^u_m$ with $m \geqslant m_1$. It follows from **C5** that these are constants $K_1 > 0$ and $D_0 = D_0(n, m_1)$ such that if the length D of the given LUM is less than D_0, then its image intersects at most $K_1 n$ discontinuity curves and subdividing sets $\mathscr{F}^u_m$ with $m < m_1$. Now, K_1 may be larger than K_0 in **C5** as an influence of these subdividing sets.

Then, for $D \leqslant D(n, m)$ and $m_1 \geqslant n_0$ we have

$$l\left(\bigcup \gamma_i^N\right) \leqslant \sum_{k_1, k_2=0}^{N} DL^N c_1^{-k_1-k_2} \beta^{-N+k_1+k_2} \prod_{k=1}^{k_1} l_k^{-d} \prod_{k=1}^{k_2} j_k^{-1} (K_1 n)^{[N/n]+1} \quad (2)$$

In order to simplify the understanding of the estimations, we will discuss separately the cases $k_2 = 0$ and $k_2 > 0$. If $k_2 = 0$,

$$l\left(\bigcup \gamma_i^N\right) \leqslant \sum_{k_1=0}^{N} c_1^{-k_1} DL^N \left(\sum_{m=m_1}^{\infty} m^d\right)^{k_1} \beta^{-N+k_1} (k_1 n)^{[N/n]+1}$$

As $d > 1$, we may choose m_1 such that the series converges to a number smaller than β^{-1}. Since $\beta_1 = L\beta^{-1} < 1$ (**C7**), we have

$$l(\gamma_N) \geqslant p - DC\beta_1^N (K_1 n)^{[N/n]+1}, \qquad \text{with} \quad C = \sum_{k_1=0}^{N} c_1^{-k_1} \leqslant N+1$$

Then, as in the last part of 6.1, we may choose n such that for some $0 < \lambda_1 < 1$, $p_{D,\gamma^u}(N) \geqslant 1 - \lambda_1^N/p$. If $c = -N/\log p$, we have the inequality of the statement for $\beta(c) = p^{-c \log \lambda_1 - 1}$.

If $k_2 > 0$, we will use the fact that condition **C6**(b_2) established that a single LUM intersects only finitely many F_m^u of the second type with $M \leqslant m \leqslant C_2 M$. Such a LUM gives a contribution to (2) smaller than

$$\sum_{n=M}^{C_2 M} \frac{1}{c_1 n} \simeq c_1^{-1} \log c_2$$

Finally, the total contribution of the factor depending on k_2 in (2), after the N iterations, is not essential because $c_1^{-1} \log c_2 < 1$. ∎

This theorem also can be stated in the following way: $f^n\gamma^u$ consists of at most a countable number of HLUMs called components. Let $r_n(f^n x)$ be the l-distance from $f^n x$ to the boundary of its component. Then

$$l(\{x \in \gamma^u \colon \max\{r_n(f^n x) \colon 0 \leqslant n \leqslant N\} \geqslant D\})/p \geqslant 1 - \beta(c)$$

for $N = -c \log p$ (D, β, and c as in the previous statement).

6.4. Transitivity. Consider an HLUM γ^u of length of order one. Then almost all its points expand and the homogeneous segments of its images begin to fill M. General considerations on the mixing property predict that for large n, the homogeneous segments of $f^n\gamma^u$ are uniformly distributed on M. But we need, and can prove, only that the density of filling by these components is asymptotically bounded away from zero as $n \to +\infty$.

6.5. Definition. Let γ^u be an arbitrary HLUM, $p = l(\gamma^u)$, and A_0 an arbitrary 0-homogeneous maximal parallelogram (see 4.5 and 5.4). For each $n \geqslant 1$ we choose inside γ^u subsegments $\tilde{\gamma}^u_{i,n}$ whose images $f^n\tilde{\gamma}^u_{i,n}$ are stretched on the parallelogram A_0. We consider the numbers

$$p_n(\gamma^u, A_0) = l\left(\bigcup_i \tilde{\gamma}^u_{i,n}\right) \Big/ l(\gamma^u)$$

that measure the l-percentage of γ^u whose image by f^n is an HLUM stretched in A_0.

6.6. Theorem (Transitivity). There exist $\delta_0 = \delta_0(p, A_0) > 0$ and $n_0 = n_0(p, A_0)$ (depending on A_0 and the length p of γ^u, but not on its position) such that $p_n(\gamma^u, A_0) \geqslant \delta_0$ for all $n \geqslant n_0$.

The proof of this theorem is the proof of Theorem 3.13 in ref. 7 and we will not repeat it here. We will only state some of the assertions used in the proof.

From 5.1 and our version of the fundamental theorem of Sinai–Chernov,[23,24] we can obtain the following result.

6.7. Lemma. Let γ be an arbitrary increasing curve in M. Then through almost every $x \in \gamma$ there passes an HLSM $\gamma^{os}(x)$.

Using this lemma and the absolute continuity of stable and unstable foliations (see 2.5), we can prove the existence of a parallelogram that is a *dense part* (with *positive parameters* ε *and* $d < 1/10$) of an HLUM γ^u.

6.8. Proposition. In a neighborhood of γ^u there exists a parallelogram A_1 with the following properties:

1. A_1 is 0-homogeneous and maximal.
2. γ^u intersects both s faces of A_1 and the points of intersection have distance larger than $dl(\gamma^u)$ to the endpoints of γ^u.
3. For each HLSM γ^1 stretched on A_1 (see 4.5), the point of intersection $\gamma^u \cap \gamma^s$ has distance larger than $dl(\gamma^s)$ from the endpoints of γ^s.
4. The density of the parallelogram A_1 on every HLSM γ^s stretched on it $[l(A_1 \cap \gamma^s)/l(\gamma^s)]$ is at least $1 - \varepsilon_1$.

6.9. Lemma (On finite collections of dense parts). For any $p > 0$, $\varepsilon_1 > 0$, and $d \in (0, 1/10)$ we can choose in M a finite collection of parallelograms such that for each HLUM of length $\geqslant p$, one of these parallelograms is a dense part with parameters ε_1 and d.

We fix $d \in (0, 1/10)$, a sufficiently small $\varepsilon_1 > 0$, and a dense part A_1 with parameters ε_1 and d of γ^u. Let γ^u and A_0 be as in the condition of 6.6.

Absolute continuity of stable and unstables foliations implies that almost any point $x \in A_0$ is a density point of the measurable set $\gamma^{s}_{A_0}(x)$. Then, for any $\varepsilon_2 > 0$ there are a $p_1 > 0$ and a subset $\tilde{A}_0 \subset A_0$ of nonzero measure such that on any HLSM γ intersecting $\tilde{A}_0$ and l-length less than p_1, the parallelogram has density $p(\gamma \cap A_0)/p(\gamma)$ greater than $1 - \varepsilon_2$. We fix a sufficiently small ε_2 and the corresponding set (not necessarily a parallelogram) $\tilde{A}_0$.

$f^n A_1$ consists of finitely many 0-homogeneous parallelograms for any $n \geqslant 0$. We denote by $A_{1,n},..., A_{k(n),n}$ those parallelograms that intersect $\tilde{A}_0$. Let $\gamma^{u}_{i,n}$ be the projections under the action of the Poincaré map of $A_{i,n}$ on $f^n\gamma^{u}$. Distinct parallelograms $A_{i,n}$ are projected onto distinct homogeneous components $\gamma^{u}_{i,n}$.

6.10. Lemma. Each $\gamma^{u}_{i,n}$, $1 \leqslant i \in k(n)$, intersects both s faces of the parallelogram A_0, provided that n is sufficiently large $[n \geqslant n'(A_0, \varepsilon_z)]$ and $\varepsilon_1, \varepsilon_z$ are sufficiently small.

The proof of 6.6 follows from 6.10, 5.6, 4.6, and 6.9.

Successive application of 6.3 and 6.6 leads to the following general assertion.

6.11. Theorem. Suppose that the HLUM γ^{u} and the parallelogram A_0 satisfy the conditions of Definition 6.5. Then

$$p_n(\gamma^{u}, A_0) > \delta_1 \qquad \text{for all} \quad n \geqslant -\tilde{C} \log p + n_1$$

where the quantities $\delta_1 = \delta_1(A_0)$ and $n_1 = n_1(A_0)$ depend only on A_0 and $\tilde{C} > 0$ is a constant.

7. CONSTRUCTION OF THE MARKOV SIEVE

The purpose of this paper is to construct a finite partition of M that satisfies a kind of strong mixing condition in the sense of Ibragimov (ref. 13, Chapter 17, §2). This result on convergence toward equilibrium is the basis of the proof of Theorem A. The construction will be done in the following three sections.

7.1. Definition. A Markov sieve (*lattice*) with parameters n, N is a finite partition of the space M, $\mathscr{R}_{n,N} = \{V_0, V_1,..., V_I\}$, where $N > n > 0$, $I = I(n, N)$, $\mu(V_i \cap V_j) = 0$ for $i \neq j$, and $\bigcup \bar{V}_i = M$, having the following four properties:

P1 (Sizes). diam $V_i \leqslant e^{-n}$ for all $i \geqslant 1$ (V_0 does not participate in this estimate).

P2 (Measure of marginal set). $\mu(V_0) \leqslant Ne^{-n}$.

P3 (Markovian approximation). For any integers $k > m > 1$ and $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant N$ as well as for indices $j_i, j_2, \ldots, j_k$ taking values from 1 to $I(n, N)$, the following relation holds between conditional probabilities:

$$\begin{aligned} &\mu(f^{i_1} V_{j_1} \cap f^{i_2} V_{j_2} \cap \cdots \cap f^{i_{m-1}} V_{j_{m-1}} / f^{i_m} V_{j_m} \cap \cdots \cap f^{i_k} V_{j_k}) \\ &\quad = \mu(f^{i_1} V_{j_1} \cap f^{i_2} V_{j_2} \cap \cdots \cap f^{i_{m-1}} V_{j_{m-1}} / f^{i_m} V_{j_m})(1 + \Delta) \end{aligned} \tag{7.1}$$

when $|\Delta| \leqslant C_5 \alpha_0^n$ for some $C_5 = C_5(\alpha_0)$ (see 4.7).

Let $\mathscr{Y} = \{1, 2, \ldots, I\}$. For any $k \geqslant 1$ and $i \in \mathscr{Y}$ we choose a subset $R_i(k) \subset \mathscr{Y}$ such that $j \in R_i(k)$ if and only if for some positive constant β_0

$$\mu(f^k V_i \cap V_j) \geqslant \beta_0 \mu(V_i)\, \mu(V_j) \tag{7.2}$$

Then, we will say that $i \in R(k) \subset \mathscr{Y}$ if and only if

$$\sum_{j \in R_i(k)} \mu(V_j) > 1 - e^{-n} \tag{7.3}$$

P4 (Regularity). For each $k \geqslant D_0 n$ we have

$$\sum_{i \in R(k)} \mu(V_i) > 1 - Ne^{-n} \tag{7.4}$$

where D_0 is a positive constant.

7.2. Remarks. (a) N is the length of the time interval on which we will approximate our stationary process $\{X_i\}$ (see 3.7) by a process of Markov type (see Section 1). We will not impose further restrictions on N and n, but it will be useful to keep in mind that we are interested in the case when $n \simeq N^\gamma \to +\infty$ for some $\gamma < 1$.

(b) **P3** establishes, up to a factor $1 + \Delta$, a kind of Markov property for $\{V_1, \ldots, V_I\}$. It is easily proved that the same relation holds for $N \geqslant i_1 > i_2 > \cdots > i_k \geqslant 1$.

(c) **P4** establishes a kind of strong mixing rate in a smaller time interval: (7.2) guarantees a certain degree of mixing after k steps, which holds for the great majority of pairs of indices i, j. More precisely, we consider those V_i such that $f^k V_i$ are "β_0-mixed with a lot of sets V_j," and then (7.4) establishes that for values of k sufficiently n-large, the measure of these V_i is close to 1.

7.3. Construction of the Markov Sieve. This is done in three steps. In the first step (Section 8) we construct a pre-Markov partition that is a covering of M by polygons whose boundaries are segments of LUMs, LSMs, and discontinuity curves. In the second step (Section 9) we turn the pre-Markov partition into a pre-Markov sieve formed by parallelograms. It is a rough approximation to the Markov sieve. Finally, we give two modifications of the pre-Markov sieve and obtain the Markov sieve.

8. PRE-MARKOV PARTITIONS

For the sake of completeness we recall some elements on Markov partitions.

8.1. Definition. For any pair of parallelograms A, B and $n>0$ ($n<0$) we say that the intersection $f^nA \cap B$ is *regular* if it consists of parallelograms u-inscribed (s-inscribed) in B, and its preimage $A \cap f^{-n}B$ consists of parallelograms s-inscribed (u-inscribed) in A. (See Definition 4.2).

Similarly, for any pair of quadrilaterals P, Q and $n>0$, we say that their intersection $f^nP \cap Q$ is *regular* if it consists of quadrilaterals with s faces on ∂Q, and its preimage $P \cap f^{-n}Q$ consists of quadrilaterals with u surfaces on ∂P.

8.2. Definition. A Markov partition for f is a partition (mod 0) of M into parallelograms $\{A_i\}$, $i \in \mathbb{N}$, such that: (i) every parallelogram lies in a connected domain of $\mathbb{N}$ on which f and f^{-i} are continuous; and (ii) $f^nA_i \cap A_j$ intersects regularly for every $i \neq j$ and $n \neq 0$.

8.3. Conjecture. If a dynamical system f satisfies the hypothesis of Theorem A, then for any $\varepsilon>0$ there is a Markov partition for f whose elements have diameters less than ε.

There are good reasons to suppose that this conjecture is true. The first one is that the adaptation that we have done in this paper of the methods of ref. 7 can be done also for ref. 6, where the Markov partition was constructed for dispersing billiards and the stadium.

The second reason is that Markov partitions have been constructed independently for nonuniformly hyperbolic systems with singularities by Krüger and Troubetzkoy.[15] The conditions satisfied by the systems studied in ref. 15 are similar to the hypothesis of our Theorem A.

In fact, as was noted in ref. 7, the proofs of some important statistical properties of these dynamical systems do not require the construction of a Markov partition, but of a kind of approximation to it by a finite number of sets: the Markov sieve.

We begin its construction by introducing a partition into curvilinear polygons.

8.4. Definition. Let $m_1 \geqslant m \geqslant 1$ be integers, and $0 < \varepsilon < \varepsilon_0(m)$ arbitrarily small. Let $\mathcal{N}_0 = \mathcal{N}_0(\varepsilon)$ be a partition of M into curvilinear polygons $P_1, \dots, P_k$ (the boundary of each polygon consists of a finite set of C^1-smooth curves). It is a pre-Markov partition for f^m if:

(i) The boundary $\partial \mathcal{N}_0 = \bigcup \partial P_i$ is the union of S_{-m_1, m_1} and a finite collection of LUMs and LSMs. Respectively we denote these sets by $\partial^0 \mathcal{N}_0$, $\partial^{\mathrm{u}} \mathcal{N}_0$, $\partial^{\mathrm{s}} \mathcal{N}_0$.

(ii) $f^m(\partial^{\mathrm{s}} \mathcal{N}_0) \subset \partial^{\mathrm{s}} \mathcal{N}_0, f^{-m}(\partial^{\mathrm{u}} \mathcal{N}_0) \subset \partial^{\mathrm{u}} \mathcal{N}_0$.

(iii) Any segment of an LUM (LSM) that is part of $\partial^{\mathrm{u}} \mathcal{N}_0$ ($\partial^{\mathrm{s}} \mathcal{N}_0$) ends either in $\partial^0 \mathcal{N}_0$ or *strictly inside* some LSM (LUM) that is a part of $\partial^{\mathrm{s}} \mathcal{N}_0$ ($\partial^{\mathrm{u}} \mathcal{N}_0$).

(iv) The sides of the polygons $P \in \mathcal{N}_0$ lying on LUMs ad LSMs are greater than $k_1 \varepsilon$ and less than $k_2 \varepsilon$ (k_1, k_2 are constants determined by m).

The construction of a pre-Markov partition will be done in three stages, following ref. 6, §3.

Let $N = M \backslash S$, as in the first paragraph of Section 2. The quantities m and $\varepsilon_0(m)$ that appear in the definition of pre-Markov partitions are chosen during the construction process.

8.5. Proposition (Initial partition ξ_0). In N we can choose a finite system of m_2-increasing curves $\Gamma_0^+ = \{\gamma_i^+ : 1 \leqslant i \leqslant I_0^+\}$ and a finite system of m_2-decreasing curves $\Gamma_0^- = \{\gamma_i^- : 1 \leqslant i \leqslant I_0^-\}$ such that, for $m_2 = m+1$, we have the following:

(a) $\lambda\varepsilon \leqslant l(\gamma) \leqslant \lambda^{-1}\varepsilon$ for every $\gamma \in \Gamma_0^+ \cup \Gamma_0^-$.

(b) The curves in Γ_0^+ (Γ_0^-) lie outside $U_{\lambda\varepsilon}(S_{0,m_2})$ [$U_{\lambda\varepsilon}(S_{-m_2,0})$].

(c) (Consistency). The endpoints of each curve in Γ_0^- (Γ_0^+) lie on two curves in Γ_0^+ (Γ_0^-).

(d) (Density). Any 1-decreasing (1-increasing) curve of l-length $\lambda^{-1}\varepsilon$ intersects at least one curve $\gamma \in \Gamma_0^-$ ($\gamma \in \Gamma_0^+$) in such a way that the point of intersection divides γ into segments of l-length at least $\lambda\varepsilon$. Here $\lambda = (200 K_0 m)^{-1}$ depends on the constant K_0 introduced in **C5**.

Proof. It is analogous to the proof of Proposition 2.1 in ref. 6. We must be careful only in the selection of the values of c_1, c_2, and λ. Now c_1 depends on the upper bound of Q and the angles between 1-increasing curves and m_2-decreasing curves at points that are not close to the sets of singularities; these values are controlled by hypothesis (i) in Theorem 2.1. Now c_2 and λ depend on the fact that for any $m > 0$ there is an $\varepsilon_0(m)$ such that any ε_0-Riemannian disk in N intersects at most $K_0 m$ curves in $S_{-m,m}$. ∎

8.6. Lemma. If m is such that $\mu^m > \lambda^{-3}$, then in N we can choose a finite system Γ_∞^+ (Γ_∞^-) of segments of LUMs (LSMs) satisfying the following conditions:

(a) $\forall \gamma \in \Gamma_\infty^\pm$, there exists $\gamma' \in \Gamma_\infty^\pm$ such that $f^{\mp m}\gamma \subset \gamma'$.

(b) Between the curves $\Gamma_{-\infty}^\pm$ and $\Gamma_0^\pm$ there is a natural correspondence under which corresponding curves remain at a distance of at most $\lambda^2\varepsilon/100$ from each other.

(c) (Consistency). The endpoints of each curve $\gamma \in \Gamma_\infty^\pm$ lie on two curves in $\Gamma_\infty^\mp$.

(d) Any 1-increasing (1-decreasing) curve of l-length $\lambda^{-1}\varepsilon$ intersects some segments of an LSM in Γ_∞^- (LUM in Γ_∞^+); moreover, the point of intersection divides this segment into two pieces of l-length not less than $\lambda\varepsilon/2$.

This lemma and the following proposition are proved in §3.2 of ref. 6.

8.7. Proposition. The partition of N determined by the system of curves Γ_∞^+, Γ_∞^-, and $S_{-m,m}$ is finite and pre-Markov for f^m, $m_1 = m$.

This proposition finishes the second stage of the construction. The third stage is resumed in the following proposition; it is proved in §3.3 of ref. 6.

8.8. Proposition. Let $\Gamma^\pm$ be the system of curves consisting of the curves $\gamma \in \Gamma_\infty^\pm$ and the smooth components of their images $\Gamma^{\pm m}\gamma$. The partition ξ determined by $\Gamma^\pm$ and the discontinuity curves $S_{-m,m}$ is finite and pre-Markov for f^m, $m_1 = m$. If an element $A \in \xi$ does not border with $S_{-m,m}$, then it is a curvilinear quadrilateral bounded by two LUMs and two LSMs alternating with each other. If A does border with $S_{-m,m}$, it is bounded by segments of LUM, LSM, and discontinuity curves such that interior angles do not exceed 180°.

8.9. Proposition. The measure of the union of all bordering (with $S_{-m,m}$) elements of ξ does not exceed $\tilde{K}K(k_2\varepsilon)^\delta$, where $\delta = \min\{1, \theta\}$, K and θ were determined in **C4**, and $\tilde{K}$ is a constant that depends on m.

Proof. Recall 4.11(d). If a quadrilateral bordering $S_{-m,m}$ is close to $S_0^\pm$, its measure is controlled by application of **C4**. Otherwise the Riemannian diameter of the quadrilateral is controlled by the bounds of the quadratic form Q in regions were it is continuous (see 2.4–2.5). The result follows from the absolute continuity of μ. ∎

In conclusion, we consider the partition $\xi_1 = \bigvee_{-m}^m f^i\xi$. It is a pre-Markov partition for f, $m_1 = 2m$. It is easily seen that the measure of the union of all bordering sets (with $S_{-2m,2m}$) does not exceed $K_1\varepsilon^\delta$, where $K_1 > 0$ is a constant depending on m, and $\delta > 0$ is defined in 8.9.

9. INITIAL AND PRE-MARKOV SIEVES

The second step of the construction of our Markov sieve consists in the selection of an *initial sieve* (lattice) with good initial Markovian properties and very close (in measures and dimensions) to the pre-Markov partition ξ_1.

9.1. Lemma (4.4 in ref. 7). There exists an (initial) sieve $\mathcal{R}(\varepsilon)$ consisting of a finite number of parallelograms W, one each obtained from an element $Q(W)$ of ξ_1, satisfying the following properties:

(a) (Structure). The parallelograms $W \in \mathcal{R}(\varepsilon)$ are 0-homogeneous and are formed by intersection of all the HLUMs and all the HLSMs stretched on $Q(W)$; the quadrilaterals $Q(W)$ can intersect only along boundaries.

(b) (Measure of the remainder). If $W_0 = M \backslash \bigcup \{W \in \mathcal{R}(\varepsilon)\}$, then $\mu(W_0) < \varepsilon^{b_2}$.

(c) (Markov property in one step). For any two parallelograms $W', W'' \in \mathcal{R}(\varepsilon)$ the intersections $fW' \cap W''$ and $fQ(W') \cap Q(W'')$ either have measure zero or are regular (see Definitions 8.1 and 4.2).

(d) (Density). For each $W \in \mathcal{R}(\varepsilon)$ there are HLUM $\gamma^{\mathrm{u1}}(W)$ and HLSM $\gamma^{\mathrm{s1}}(W)$, stretched on $Q(W)$, on which W has density at least $1 - \delta_1(\varepsilon)$, where $\delta_1(\varepsilon) \to 0$ as $\varepsilon \to 0$.

(e) (Dimensions). For each $W \in \mathcal{R}(\varepsilon)$ we have $\operatorname{diam} W \leqslant \mathrm{const} \cdot \varepsilon^{\delta}$, but $l(\gamma^{\mathrm{u1}}(W)) \geqslant \varepsilon^{b_2}$ and $l(\gamma^{\mathrm{s1}}(W)) \geqslant \varepsilon^{b_2}$.

Proof. For each nonbordering element $U \in \xi_1$ we consider *all* the HLUMs and *all* the HLSMs stretched on U (Definition 4.5).

By intersection we obtain the parallelogram $W(U)$. We will select among them the parallelograms satisfying the following three additional conditions.

(A1) The quadrilateral U and its images f^iU for $|i| \leqslant m$ do not intersect subdividing curves.

(A2) $\mu(W(U))/\mu(U) > 1 - e^{b_1}$.

(A3) $l(\gamma_W^{\mathrm{u,s}}(x)) > \varepsilon^{b_2}$ for every point $x \in W$.

$b_1, b_2, \ldots$ denote positive constants determined by the choice of ν and m_0 in Section 5 and will be determined in the proof of the following statements.

We define the initial sieve $\mathcal{R}(\varepsilon)$ as the set of all parallelograms W constructed and selected above.

The rest of the proof is analogous to the proof of Lemma 4.4 in ref. 7. For proving property (b) we must use **C5** instead of their Theorem 3.10 and, in the evaluation of the measure of the parallelograms not satisfying (A3), we must use the arguments in the proof of our Proposition 8.9. ∎

The third step of the construction of the Markov sieve consists of two modifications of the initial sieve. These modifications are exactly the same as done in §4.3 of ref. 7. We will only describe the construction and give the final result.

9.2. First Modification. It consists in a partition of each $W \in \mathscr{R}(\varepsilon)$ to obtain n-homogeneous parallelograms.

We consider all possible intersections

$$f^{-n}W_{-n} \cap \cdots \cap f^{-1}W_{-1} \cap W_0 \cap fW_1 \cdots \cap f^n W_n$$

having nonzero measure [for all possible $W_i \in \mathscr{R}_\varepsilon(\varepsilon)$]. We denote by $\mathscr{R}(\varepsilon, n)$ the set of these finite intersections, and call it a *pre-Markov sieve*. We put $Z_0 = M \setminus \bigcup \{Z \in \mathscr{R}(\varepsilon, n)\}$.

9.3. Second Modification. It consists of a selection of a refining of the elements in $\mathscr{R}(\varepsilon, n)$ to guarantee the Markov property in $\pm N$ steps.

For each $Z \in \mathscr{R}(\varepsilon, n)$ we define a subparallelogram

$$V = V(Z) = \{x \in Z \colon f^k x \notin Z_0 \text{ for all } |k| \leqslant N\}$$

We select those parallelograms $V \subset Z$ that satisfy the inequality

$$\mu(V) > (1 - \varepsilon^{b_6})\, \mu(z)$$

for some $b_6 > 0$.

For convenience of notation we enumerate all parallelograms $Z_i \in \mathscr{R}(\varepsilon, n)$ satisfying the last inequality, $i = 1, 2, ..., I(\varepsilon, n, N)$. We denote by $\mathscr{R}(\varepsilon, n, N)$ the set of parallelograms V_i and put $V_0 = M \setminus \bigcup \{V_i \colon 1 \leqslant i \leqslant I(\varepsilon, n, N)\}$.

9.4. Lemma (4.6 in ref. 7). The system of sets $\mathscr{R}(\varepsilon, n, N)$ has the following properties:

(a) (Structure). The parallelograms $V \in \mathscr{R}(\varepsilon, n, N)$ are n-homogeneous.

(b) (Measure of the remainder). $\mu(V_0) \leqslant N_n \varepsilon^{b_6}$.

(c) (Markov properties in $\pm N$ steps). For any two parallelograms $V', V'' \in \mathscr{R}(\varepsilon, n, N)$ and any $k \in \mathbb{Z}$, $1 \leqslant |k| \leqslant N$, the intersection $f^k V' \cap V''$ either has measure zero or is regular.

(d) (Dimensions). For each parallelogram $V \in \mathscr{R}(\varepsilon, n, N)$, we have diam $V \leqslant \text{const} \cdot \varepsilon^{\delta}$.

9.5. The Markov Sieve and Its Properties. We define the Markov sieve $\mathscr{R}_{n,N}$ as the system $\mathscr{R}(\varepsilon, n, N)$ for $\varepsilon^{b_7} = e^{-n}$, where $b_7 = \min\{\delta, b_2, b_6\}/4$, and adjoin to it the set V_0.

Properties P1 and P2 in 7.1 follow immediately from 9.5. Property P3 is a consequence of 9.5(c) and Lemma 4.9. Finally, P4 is proved in §4.4 of ref. 7 using our Theorem 6.11 instead of Theorem 3.19 and Lemma 5.6 instead of Lemma 3.7.

ACKNOWLEDGMENTS

This paper was prepared in connection with thoughts on the proof of the statistical properties of the billiard in the cardioid, stimulated by Ya. G. Sinai. I studied the papers by Bunimovich, Sinai, and Chernov, in collaboration with Arthur Lopes (UFRGS, Porto Alegre), in IMPA. During a short visit to Princeton University I discussed the main ideas of this paper with Kolya Chernov, whose sharp observations were crucial for its elaboration.

REFERENCES

1. D. V. Anosov and Ya. G. Sinai, Some ergodic systems, *Russ. Math. Surv.* **22**(5):103–167 (1967).
2. L. A. Bunimovich, On billiards close to dispersing, *Math. USRR Sb.* **23**(1):45–67 (1974).
3. L. A. Bunimovich, On ergodic properties of nowhere dispersing billiards, *Commun. Math. Phys.* **65**:295–312 (1979).
4. L. A. Bunimovich, A theorem on ergodicity of two-dimensional hyperbolic billiards, *Commun. Math. Phys.* **130**:599–621 (1990).
5. L. A. Bunimovich and Ya. G. Sinai, Markov partitions for dispersing billiards, *Commun. Math. Phys.* **78**:247–280 (1980) [Erratum, *Commun. Math. Phys.* **107**:357–358 (1986)].
6. L. A. Bunimovich, Ya. G. Sinai, and N. I. Chernov, Markov partitions for two-dimensional hyperbolic billiards, *Russ. Math. Surv.* **45**(3):105–152 (1990).
7. L. A. Bunimovich, Ya. G. Sinai, and N. I. Chernov, Statistical properties of two-dimensional hyperbolic billiards, *Russ. Math. Surv.* **46**(4):47–106 (1991).
8. I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai, *Ergodic Theory* (Springer, Berlin, 1982).
9. N. I. Chernov, Ergodic and statistical properties of piecewise linear hyperbolic automorphisms of the 2-torus, *J. Stat. Phys.* **69**:111–134 (1992).
10. N. I. Chernov, On local ergodicity in hyperbolic systems with singularities, *Funct. Anal. Appl.* **27**:51–54 (1993).
11. N. I. Chernov, Statistical properties of the periodic Lorentz gas: Multidimensional case, *J. Stat. Phys.* **74**:11–53 (1994).
12. N. I. Chernov, Limit theorems and Markov approximations for chaotic dynamical systems, *Prob. Theory Related Fields*, to appear.
13. I. A. Ibragimov and Y. V. Linnik, *Independent and Stationary Sequences of Random Variables* (Wolters Noordhoff, Gröningen, 1971).

14. A. Katok and J.-M. Strelcyn, in collaboration with F. Ledrappier and F. Przytycki, Smooth maps with singularities, in *Invariant Manifolds. Entropy and Billiards* (Springer, Berlin, 1986).
15. T. Krüger and S. Troubetzkoy, Markov partitions and shadowing for non-uniformly hyperbolic systems with singularities, *Ergodic Theory Dynam. Syst.* **12**:487–508 (1992).
16. A. Krámli, N. Simányi, and D. Szasz, A "transversal" fundamental theorem for semi-dispersing billiards, *Commun. Math. Phys.* **129**:535–560 (1990) [Erratum, *Commun. Math. Phys.* **138**:207–208 (1991)].
17. J. Lewowicz, Lyapunov functions and topological stability, *J. Differ. Equations* **38**:192–209 (1980).
18. J. Lewowicz and E. Lima de Sa, Analytic models of pseudo-Anosov maps, *Ergodic Theory Dynam. Syst.* **6**:385–392 (1986).
19. C. Liverani, Decay of correlations, *Ann. Math.*, to appear.
20. R. Markarian, Billiards with Pesin region of measure one, *Commun. Math. Phys.* **118**:87–97 (1988).
21. R. Markarian, Non uniform hyperbolicity. Billiards, in *Workshop on Dynamical Systems, ICTP, Trieste, 1988*, Z. Coelho and E. Shields, eds. (Longman, 1989), pp. 129–135.
22. R. Markarian, Non-uniformly hyperbolic billiards, *Ann. Fac. Sci. Toulouse* **III**:223–257 (1994).
23. R. Markarian, The fundamental theorem of Sinai–Chernov for dynamical systems with singularities, in *Dynamical Systems, Santiago de Chile 1990*, R. Bamón, R. Labarca, J. Lewowicz, and J. Palis, eds. (Harlow, Longman, 1993), pp. 131–158.
24. R. Markarian, New ergodic billiards: Exact results, *Nonlinearity* **6** (1993).
25. Ya. B. Pesin, Families of invariant manifolds corresponding to non-zero characteristic exponents, *Math. USSR Izv.* **10**:1261–1305 (1976).
26. Ya. B. Pesin, Lyapunov characteristic exponents and smooth ergodic theory, *Russ. Math. Surv.* **32**(4):55–114 (1977).
27. Ya. G. Sinai, Dynamical systems with elastic reflections, *Russ. Math. Surv.* **25**(1):137–189 (1970).
28. Ya. G. Sinai and N. I. Chernov, Ergodic properties of two dimensional discs and three dimensional balls, *Russ. Math. Surv.* **42**(3):181–207 (1987).