Corrigendum: Exponential speed of mixing for skew-products with singularities

2013 Nonlinearity 26 269–87

R Markarian$^1$, M J Pacifico$^2$ and J I Vieitez$^3$

$^1$ Instituto de Matemática y Estadística (IMERL), Facultad de Ingeniería, Brazil

$^2$ Instituto de Matemática, Universidade Federal do Rio de Janeiro, C. P. 68.530, CEP 21.945-970, Rio de Janeiro, RJ, Brazil

$^3$ Regional Norte, Universidad de la República, Rivera 1350, CP 50000, Salto, Uruguay

E-mail: roma@fing.edu.uy, pacifico@im.ufrj.br and jvieitez@unorte.edu.uy

Received 23 June 2013, in final form 29 August 2013
Published 23 September 2013
Online at stacks.iop.org/Non/26/2923

The statement of theorem C was incorrect. The corrected statement is weaker:

Theorem C. For all pair of rectangles $A$ and $B$ with sides parallel to the coordinate axes, there is $\lambda = \lambda(A, B), 0 < \lambda < 1$, such that

$$\left| m(f^{-n}(A) \cap B) - m(A)m(B) \right| < \lambda^n m(A)m(B), \quad \text{for all } n \geq 0.$$ 

It is clear from lemma 5.4 in our original paper that $\lambda$ depends on the length of the sides of the rectangles. The final sentence before section 1.1 should also be replaced by: ‘Finally we prove a result related to the exponential rate of mixing (theorem C).’ We also note that the problem on the decay of correlations considered in the paper is still open.

Acknowledgment

We thank Oliver Butterley for pointing out this error to us.
Exponential speed of mixing for skew-products with singularities

R Markarian\(^1\), M J Pacífico\(^2\) and J L Vieitez\(^3\)

\(^1\) Instituto de Matemática y Estadística (IMERL), Facultad de Ingeniería, Universidad de la República, CC30, CP 11300, Montevideo, Uruguay
\(^2\) Instituto de Matemática, Universidade Federal do Rio de Janeiro, C. P. 68.530, CEP 21.945-970, Rio de Janeiro, RJ, Brazil
\(^3\) Regional Norte, Universidad de la República, Rivera 1350, CP 50000, Salto, Uruguay

E-mail: roma@fing.edu.uy, pacifico@im.ufrj.br and jvieitez@unorte.edu.uy

Received 10 February 2012, in final form 2 November 2012
Published 28 November 2012
Online at stacks.iop.org/Non/26/269

Recommended by L Bunimovich

Abstract
Let \( f : [0, 1] \times [1/2] \rightarrow [0, 1] \times [0, 1] \) be the \( C^\infty \) endomorphism given by \( f(x, y) = (2x - \lfloor 2x \rfloor, y + c \frac{\lfloor x - 1/2 \rfloor}{|x - 1/2|}), \ c \in \mathbb{R}^+ \). We prove that \( f \) is topologically mixing and if \( c > 1/4 \) then \( f \) is mixing with respect to the Lebesgue measure. Furthermore the speed of mixing is exponential. This skew-product can be seen as a toy model related to Lorenz-like attractors.

Mathematics Subject Classification: 37D30, 37C29, 37E30

(Some figures may appear in colour only in the online journal)

1. Introduction

A basic problem in dynamics is understanding the ergodic behaviour of a given dynamical system. Frequently this is translated into the knowledge of mixing properties of the system. Once mixing is established it is natural to ask for the rate or speed of mixing of the system.

For hyperbolic systems and nonuniform hyperbolic ones, with or without singularities, this kind of study is well understood and the techniques to do so have been developed by several authors. See, for instance, [Pe, Ru, BY, Wa] and references therein for the interested reader.

When the system \( T \) under study has singularities, the phase space is not the whole manifold and in this case one requires a zero-Lebesgue measure for the union \( \cup_{n=0}^\infty T^{-n} S \) of the set of singularities \( S \). This is the case of billiards, studied by Sinai [Si], Chernov and Young [CY], Chernov and Markarian [CM], Bunimovich [Bu] and others. In these cases we have the
additional difficulty that the stable and unstable manifolds of points may be arbitrarily short since their length is conditioned by the distance of the points to $S$.

In general, the presence of singularities adds complexity to the problem and makes analysis much more difficult. Nevertheless, in this paper, where we study a certain skew-product with singularities on the fibre, it is the presence of singularities, jointly with the expanding action in the base, that enables us to obtain all the chaotic behaviour of the system.

In this paper we are interested in the mixing properties of the skew-product given by the $C^\infty$-endomorphism $f : ([0, 1] \times [1/2]) \times [0, 1] \rightarrow [0, 1] \times [0, 1]$ defined by

$$f(x, y) = \left( 2x - \lfloor 2x \rfloor, \ y + \frac{c}{|x - 1/2|} - \left\lfloor y + \frac{c}{|x - 1/2|} \right\rfloor, \ c \in \mathbb{R}^+ \right).$$

Here, given a real number $x$, $\lfloor x \rfloor$ stands for the greatest integer less or equal to $x$.

Since the denominator $\frac{1}{x - 1/2}$ vanishes at $x = 1/2$, the line $\{(1/2, y) : y \in [0, 1] \}$ is constituted by singularities of $f$. Besides that, for $c \neq 0$ we have that the vertical projection of $f(x, y)$ varies sharply when $x$ is close to $1/2$.

Identifying $[0, 1] \times [0, 1]$ with the two-dimensional torus $\mathbb{T}^2$, the skew-product may be seen as defined in $\mathbb{T}^2$ where the circle given by $x = 1/2$ is a curve of the singularities of $f$.

The successive iterates by $f$ of a rectangle $R$ with sides parallel to the axes are transformed into a denumerable set of strips accumulating onto the circle $x = 1$ (which is identified with $x = 0$) in the torus. Indeed, when $x \rightarrow 1/2$, $2x$ approaches $1$. This effect together with the fact that the pre-orbit by $x \mapsto 2x \mod (1)$ of the circle $x = 1/2$ is dense in the torus are responsible of the rich chaotic dynamics observed in this system.

Since the length of vertical segments are preserved under $f$, the action of $f$ on the vertical borders of $R$ is just a translation depending continuously on $x \in [0, 1] \setminus \{1/2\}$. Hence, the stretching and accumulation of the iterates of $R$ onto the pre-orbit of the circle $x = 1/2$ in the torus is due to the slipping effect of $f$ in the horizontal borders of $R$.

The skew-product $f$ can be also immersed in a one-parameter family of expanding skew-products with the same line of singularities:

$$f_\lambda(x, y) = \left( 2x, \lambda y + \frac{c}{|x - 1/2|} \right) \mod 1, \ \lambda \geq 1.$$  

Thus, it is interesting to detected the ergodic properties in the limit dynamics given by $\lambda = 1$.

For instance, transitivity, mixing and rate of mixing.

In this paper we prove that the skew-product $f$ defined above is topologically mixing (theorem A), preserves the Lebesgue measure $m$ on the torus and is mixing with respect to $m$ (theorem B). Finally we prove that the rate of mixing is exponential (theorem C).

1.1. Toy model of flows with a singularity: shearing effect

Let $M$ be a three-dimensional manifold and assume that $\Phi : M \rightarrow M$ is a flow containing a transitive attractor $\Lambda \subset M$ with a hyperbolic singularity $p \in \Lambda$. The geometric Lorenz attractor and any Lorenz-like attractor satisfy these conditions, see [GW, Lo, AP].

We consider the case when the singularity has three real eigenvalues $\lambda_i, 1 \leq i \leq 3$, and satisfies $\lambda_2(\sigma) < \lambda_3(\sigma) < 0 < -\lambda_1(\sigma) < \lambda_1(\sigma)$. Via the Hartmann–Großman theorem we

Recall that a skew-product $T$ is an automorphism of the measure space $X \times Y$ where $X$ and $Y$ are measure spaces and the action of $T : X \times Y \rightarrow X \times Y$ has the form

$$T(x, y) = (A(x), B_\lambda(y)); \quad x \in X, \ y \in Y,$$

where $A$ is an automorphism of the space $X$ (the ‘base’) and $B_\lambda(y)$, with $x$ fixed, is an automorphism of $Y$ (the ‘fibre’). The concept of a skew-product extends directly to the case of endomorphisms.
Exponential speed of mixing for skew-products with singularities

Figure 1. A Lorenz-like equilibrium and a cross-section of the flow.

Figure 2. The shearing effect.

Assume that we have linearized coordinates in a neighbourhood $U \supset [-1, 1]^3$ of the singularity $p$ in such a way that $\lambda_1$ corresponds to $x$-axis, $\lambda_2$ to $y$-axis and $\lambda_3$ to $z$-axis.

Let $S = \{(x, y, z) \in U : z = 1\}$ be a transverse section to the flow so that every trajectory eventually crosses $S$ in the direction of the negative $z$-axis. Consider also $\Sigma = \{(x, y, z) : |x| = 1\} = \Sigma^+ \cup \Sigma^-$ with $\Sigma^\pm = \{(x, y, z) : x = \pm 1\}$ (figure 1).

To each $(x_0, y_0, 1) \in S$ the time $\tau$ such that $X^\tau(x_0, y_0, 1) \in \Sigma$ is given by $\tau(x_0) = -\frac{1}{\lambda_1} \log |x_0|$, and it is such that $\tau(x_0) \to \infty$ when $x_0 \to 0$. This fact has the effect that different slices parallel to the $y$-axis of the section $S$ arrive at $\Sigma$ with a delay. Hence, we cannot see the return of each slice to $S$ at the same time, even when the expected delay is bounded.

Assume now that we ignore the effect of the singularity and consider that the return time is the same for points in the same slice. Also ignore the strong stable direction. Note that the strong stable direction does not interfere in the dynamics of the geometric Lorenz attractor [MPP, lemma 2.16].

After these identifications, the dynamics in a neighbourhood of $p$ occur in the $(x, z)$ plane, and may be seen as a shearing in the vertical direction in order to annihilate the delay of time, figure 2. Since the delay goes to infinity as $x \to 0$ the slipping also goes to infinity when $x \to 0$. Thus, the dynamics there is given by $(\phi(x), \psi(x, z))$ with $\psi(x, z) \to \infty$ when $x \to 0$. Moreover, since the ratio $\beta = -\frac{\lambda_2}{\lambda_1}$ is greater than one, the dynamics in the $x$ direction is expanding.

Thus, changing the name of the variable $z$ to $y$, the skew-product

$$f(x, y) = \left(2x \mod (1), y + \frac{c}{|x-1/2|} \mod (1)\right)$$

may be seen as a simplified case of the shearing effect in singular hyperbolic attractors, as is the case of a Lorenz-like attractor, figure 3 displays the features of the map $f$ on a rectangle.
1.2. Statement of results

To announce our results in a precise way let us introduce some definitions and related facts proved elsewhere.

**Definition 1.1.** Let $(X, A, f, \mu)$ be a dynamical system defined on the space $X$, $A$ a $\sigma$-algebra of $X$, and $\mu$ an $f$-invariant probability measure. The map $f$ is mixing if for all pairs of sets $A, B \in A$, we have

$$\lim_{n \to \infty} \mu(f^{-n}(A) \cap B) = \mu(A)\mu(B).$$

A form of mixing that can be defined without appealing to measures is the following

**Definition 1.2.** Let $f : X \to X$ be a continuous map defined in the topological space $X$ ($X$ is not necessarily compact). We say that the dynamical system defined by $f$ is topologically mixing if for every pair of non-empty open subsets $A, B$ of $X$ there is $N > 0$ such that $\forall n \geq N : f^n(A) \cap B \neq \emptyset$.

There is even a commonly used weaker notion: we say that the system defined by $f$ is topologically transitive if for every pair of non-empty open subsets $A, B$ of $X$ there is $n \in \mathbb{Z}$ such that $f^n(A) \cap B \neq \emptyset$.

It is well known that if a dynamical system is defined on topological space $X$, $A$ is the Borel $\sigma$-algebra of $X$ and $\mu$ is a probability invariant measure such that $\mu(A) > 0$ for every open set $A$ of $X$, then if the system is mixing it is topological mixing. This and other general results on ergodic theory may be found in [Wa], for instance.

The main results in this paper are:

**Theorem A.** For all positive $c$ the skew-product $f : \mathbb{T}^2 \to \mathbb{T}^2$ is topologically mixing.

**Theorem B.** The skew-product $f$ preserves the Lebesgue measure $m$ in the torus. For $c > 1/4$, $f$ is mixing with respect to $m$.

**Theorem C.** The rate of mixing is exponential, that is, there is $0 < \lambda < 1$ such that for all pairs of rectangles $A$ and $B$ with sides parallel to the coordinate axes, we have,

$$|m(f^{-n}(A) \cap B) - m(A)m(B)| < \lambda^n m(A)m(B), \quad \text{for all} \quad n \geq 0.$$

2. Preliminaries

In this section we establish some preliminary properties of $f$ that will be used in the proofs. We identify the set $Q = [0, 1] \times [0, 1]$ with the 2-torus $\mathbb{T}^2$. 

![Figure 3. Effect on a rectangle A.](image-url)
First we list two interesting features of the skew-product \( f \)

(*) For all \( p = (x, y) \in \mathbb{T}^2 \), there is no stable manifold \( W^s(p, f) \). Indeed, given \( (x, y) \) and \( (x + \Delta x, y + \Delta y) \), assuming that \( x < 1/2 \), \( \Delta x \neq 0 \), \( x + \Delta x < 1/2 \) and computing

\[
\| f(x + \Delta x, y + \Delta y) - f(x, y) \| = \left\| \left( 2\Delta x, \Delta y + \frac{c}{1/2 - x} \left[ \frac{1}{1 - \frac{\Delta x}{1/2 - x}} - 1 \right] \right) \right\| \mod 1
\]

\[ \geq 2|\Delta x|, \] and similar result holds for \( x > 1/2 \).

Hence, if \( \Delta x \neq 0 \), \( \text{dist}(f^n(x + \Delta x, y + \Delta y), f^n(x, y)) \geq 2^n|\Delta x| \mod 1 \) which does not converges to 0. On the other hand, if \( \Delta x = 0 \) then the distance between \( f^n(x, y + \Delta y) \) and \( f^n(x, y) \) is preserved. Thus, for no \( (\Delta x, \Delta y) \) we have \( \text{dist}(f^n(x + \Delta x, y + \Delta y), f^n(x, y)) \to 0 \) when \( n \to +\infty \).

(**) The unstable manifolds are not unique. Indeed, for any itinerary \( \{(x_n, y_n)\}_{n \in \mathbb{N}} \) such that

\( f(x_n, y_n) = (x_{n-1}, y_{n-1}) \) it is defined an unstable manifold \( W^u((x_0, y_0), f) \) (recall that \( f \) is an endomorphism) and so the unstable manifold of a point is not unique.

Next we analyse another properties of the map \( f \). Note that if \( 0 \leq x < 1/2 \) then \( f(x, y) = (2x, y + \frac{c}{1/2 - x} \cdot \{ y + \frac{c}{1/2 - x} \}) \) while if \( 1/2 < x < 1 \) then \( f(x, y) = (2x - 1, y + \frac{c}{1/2 - x} \cdot \{ y + \frac{c}{1/2 - x} \}) \).

The matrix \( [Df(x, y)] \), in the case \( 0 \leq x < 1/2 \), is given by \( \begin{pmatrix} 2 & 0 \\ \frac{c}{1/2 - x} & 1 \end{pmatrix} \) and, in the case \( 1/2 < x < 1 \), by \( \begin{pmatrix} 2 & 0 \\ \frac{c}{1/2 - x} & 1 \end{pmatrix} \). Therefore it depends only on \( x \). Any vector different from a vertical one is expanded by the action of \( Df \) which presents two eigenvalues: 1 with eigenvector \((0, 1)\), and 2 with eigenvector \((1, \frac{c}{1/2 - x})\) if \( 0 \leq x < 1/2 \) and \((-1, \frac{c}{1/2 - x})\) if \( 1/2 < x < 1 \).

Observe that \( f \) is not an expanding map since for any \( p = (x, y) \) we have \( Df(p)(0, 1) = (0, 1) \).

Points at the left of the line \( x = 1/2 \) have eigenvectors corresponding to the eigenvalue 2 forming an acute angle with the \( OX \) axis such that when \( x \to 1/2 \) the angle the eigenvector associated to 2 tends to be vertical. A similar picture is valid at points at the right of \( x = 1/2 \) taking into account that in that case the eigenvector associated to 2 forms an obtuse angle with the \( OX \) axis. From these facts one can see that no non-trivial splitting is preserved.

Thus, the standard techniques in dynamics using existence of stable and unstable manifolds, for instance, are useless here. To bypass these difficulties, we introduce the specific techniques we shall use to obtain our results.

Given a real number \( a \in (0, 1) \), we write

\[
a = \sum_{i=1}^{\infty} a_i \frac{1}{2^i}, \quad a = (0.a_1 \cdots a_n \cdots)_2 \quad a_i \in \{0, 1\}
\]

for its binary decomposition. Forgetting the leading zero we may identify a point \( x \in [0, 1) \), \( x = (0, b_1 b_2 b_3 \cdots)_2 \), written in base 2, with the string \((b_1, b_2, b_3, \cdots)\). We denote this identification as \( x \sim (b_1, b_2, b_3, \cdots) \). Hence the dynamics in the \( x \)-coordinate given by \( x \mapsto 2x \mod 1 \) acts as the shift

\[
\sigma : [0, 1]^\mathbb{N} \to [0, 1]^\mathbb{N}, \quad \sigma(b_1 b_2 b_3 \cdots) = b_2 b_3 \cdots
\]

Each point \( x \sim (b_1 b_2 \cdots) \) has two pre-images by this map

\[
\sigma^{-1}(b_1 b_2 \cdots) = \begin{cases} x_0 \sim (0 b_1 b_2 \cdots) \sim x/2 \\ x_1 \sim (1 b_1 b_2 \cdots) \sim (1+x)/2. \end{cases}
\]
Since \( f(x, y) = (2x, y + c/|x - 1/2|) \mod (1) \) equation (1) implies that any \( Z = (x, y) \in \mathbb{T}^2 \), with \( x = 0.b_1b_2\cdots \) has two pre-images \( Z_0, Z_1 \) by \( f \) given by:

(a) \( Z_0 = \left( \frac{x}{2}, y - \frac{x}{2} - \frac{\lfloor y - \frac{x}{2} \rfloor}{2} \right) = (x_0, y_0) \),
(b) \( Z_1 = \left( \frac{x}{2}, y - \frac{x}{2} - \frac{\lfloor y - \frac{x}{2} \rfloor}{2} \right) = (x_1, y_1) \).

Inductively, given a sequence \( b = (b_1 b_2 \cdots b_n) \) of length \( |b| = n \), with \( b_j \in \{0, 1\} \) \( \forall j \leq n \), and assuming that \( Z_{b_1 b_2 \cdots b_n} \) (one of the \( 2^{n-1} (n-1) \) pre-images of \( Z \)) is already defined we have that one of the \( 2^n \) \( n \)th pre-images of \( Z \) is \( Z_b = (x_b, y_b) \) with

\[
\begin{align*}
(a) & \quad x_b = \frac{b_1 + x_{b_1 b_2 \cdots b_n}}{2} \\
(b) & \quad y_b = \left( y_{b_1 \cdots b_n} - \frac{2c}{(1-b_1) + (2b_1-1)x_{b_1 \cdots b_n}} \right) \mod (1)
\end{align*}
\]

Remark 2.1. We remark that if \( Z = (x, y) \), \( W = (x', y') \) and \( b = (b_1 b_2 \cdots b_n) \) then \( |x_b - x'_b| = |x - x'|/2^n \). We also remark that for any \( x \in [0, 1) \) the set of pre-images \( S_n \) of \( x \) for all different \( b \) of length \( n \) is almost uniformly distributed in \( [0, 1) \), i.e., for any interval \( I \subset [0, 1) \):

\[
\lim_{n \to \infty} \frac{|S_n \cap I|}{|S_n|} = \ell(I).
\]

Here \( |X| \) means the cardinality of \( X \) \( |S_n| = 2^n \), and \( \ell(I) \) is the length of \( I \).

Convention. We extend this notation to the \( n \)th pre-image of an horizontal segment \( I = [Z, Y] \): \( Z_b(I) \) is the \( n \)th pre-image of \( I \) that has \( Z_b \) as one of its boundaries. In the same way, if \( R \) is a rectangle with borders parallel to the coordinate axes, whose lower bound is \( I \), then \( Z_b(R) \) is the \( n \)th pre-image of \( R \) with \( Z_b(I) \) as one of its ‘sides’.

3. The skew-product is topologically mixing

We start by establishing the following result:

Lemma 3.1. Given a monotone arc \( \gamma(t) = (x(t), y(t)) \) (i.e., an arc such that \( x(t) \) and \( y(t) \) are monotone continuous functions when they are seen as defined in \( [0, 1) \subset \mathbb{R} \)) there exist \( n_0 \in \mathbb{N} \) such that the vertical projection \( \Pi_x(f^n(\gamma)) \) of the image of \( y(t) = (x(t), y(t)) \) covers all \([0, 1) \).

Proof. Given a monotone arc \( \gamma : [0, 1] \to \mathbb{T}^2 \), \( \gamma(t) = (x(t), y(t)) \), the vertical projection of the function \( f(x(t), y(t)) = (2x(t)) \mod (1), y(t) + \lfloor y(t) - 1/2 \rfloor \mod (1) \) varies from \( y(0) + \lfloor y(0) - 1/2 \rfloor \mod (1) \) to \( y(1) + \lfloor y(1) - 1/2 \rfloor \mod (1) \).

If the horizontal projection \( \Pi_x(\gamma) \) contains \( x = 1/2 \) then there is \( 1/8 > \epsilon > 0 \) such that \( \Pi_x(\gamma) \supset (1/2 - \epsilon, 1/2) \) or \( \Pi_x(\gamma) \supset (1/2, 1/2 + \epsilon) \). In both cases the vertical projection of \( f(y) \), with \( x(t) \) restricted to those segments, is monotone and has infinite length. Hence the result follows. Observe that if for no \( t \) it holds that \( x(t) = 1/2 \) then \( \gamma \) is contained either in \([0, 1/2) \) or \((1/2, 1] \). In any case \( f(\gamma) \) is a monotone arc and we can iterate under the hypothesis of the lemma. Taking into account that if \( \ell(\Pi_x(f^{-1}(\gamma))) < 1/2 \) then \( \ell(\Pi_x(f^{-1}(\gamma))) = 2 \cdot \ell(\Pi_x((\gamma))) \) we conclude that there is \( n_0 \) such that the horizontal projection \( \Pi_x(f^n(\gamma)) \) contains \( x = 1/2 \). Thus the proof follows. \( \square \)
Lemma 3.2. With the same hypothesis and \( n_0 \in N \) as in lemma 3.1 there is a sequence of arcs \( \tilde{\gamma}_k \subset f^{n_0}(\gamma) \), images of subarcs of \( \gamma \), such that the vertical projection of \( \tilde{\gamma}_k \) covers \([0, 1]\) and the horizontal projection \( \Pi_x(\tilde{\gamma}_k) \to 0 \) when \( k \to \infty \).

Proof. Let us parametrize \( f^{n_0}(\gamma) \) by \((\tilde{x}(t), \tilde{y}(t))\), \( t \in [0, 1] \). Lemma 3.1 implies that there is a (unique) number \( t_0 \in [0, 1] \) such that \( \tilde{x}(t_0) = 1/2 \). There is no loss in assuming that \( t_0 \in (0, 1) \). Let \( \hat{y}(t_0) = y_0 \). Since both functions are monotone we change coordinates and consider \((x, y(x))\) as the new one, \( a \leq x \leq b \), with \( a < 1/2 < b \). In these coordinates we have that \( y(x) \to y_0 \) when \( x \to 1/2 \). By continuity of the arc \((x, y(x))\) there is \( \delta > 0 \) such that for \( x \in (1/2 - \delta, 1/2) \) it holds that \( \Pi_x(f(x, y(x))) = y(x) + \frac{\epsilon(x)}{1/2-x} = y_0 + \epsilon(x) + \frac{\epsilon(x)}{1/2-x} \) where \( \epsilon(x) \to 0 \) when \( x \to 1/2 \).

Let \( N > 0 \) such that \( \frac{1}{Ne} < \delta \). Hence \( y(x) = y_0 + \epsilon(x) + N \) for \( x = \frac{1}{2} - \frac{1}{2N} \). When \( x \) varies from \( \frac{1}{2} - \frac{1}{2N} \) to \( \frac{1}{2} - \frac{1}{2(N+2)} \) we have that \( y(x) \) varies from \( y_0 + \epsilon(1/2 - 1/2N) + N \) to \( y_0 + \epsilon(1/2 - 1/2(N+2)) + N + 2 \). Since \( \epsilon(x) \to 0 \) when \( x \to 1/2 \), we choose \( \delta \) small enough so that \( |\epsilon(x)| < 1/2 \). Taking values mod \( 1 \) we obtain that \( \Pi_x(f(x, y(x))) \) covers \([0, 1]\) when \( x \) varies in a segment of length \( \frac{1}{2} - \frac{1}{2(N+2)} \). We may repeat the arguments to prove that for all \( k > 0 \) we have that \( \Pi_x(f(x, y(x))) \) covers \([0, 1]\) when \( x \) varies in the segment \( \frac{1}{2} - \frac{1}{2(N+2)} \) whose length tends to zero when \( k \to +\infty \). \(\square\)

Theorem 3.3. If \( c \in \mathbb{R}^+ \) then \( f(x, y) \) is topologically mixing.

Proof. It is enough to prove the statement for open rectangles \( A \) and \( B \) of sides parallel to the coordinate axes since they form a basis for the standard topology of the plane.

Let \( 2\delta_A \) be the length of \( \Pi_x(A) \) and \( 2\delta_B \) the length of \( \Pi_x(B) \). Let also, in base 2, \( x_A = (a_1a_2a_3\ldots)_2 \) and \( x_B = (b_1b_2b_3\ldots)_2 \) and find \( N \) such that for \( \delta = \min(\delta_A, \delta_B) \) we have \( 2^N \delta > 1 \). Now we take

\[
r = 0.a_1a_2\ldots a_N \underbrace{0\ldots 0}_{N}1\ldots 1 0b_1b_2b_3\ldots b_N \quad \text{and} \quad Z = (r, y_A).
\]

Clearly \( r \) is near \( x_A \) and lies in \([x_A - \delta_A, x_A + \delta_A]\) since \( |r - x_A| < 1/2^N \).

After \( N \) iterates by \( f \) we have that \( f^N(Z) \) is at a distance less than \( 1/2^N \) from \( \{x = 1/2\} \) (since \( \Pi_x(f^N(Z)) = 0.011\ldots 10b_1b_2b_3\ldots \) ).

Lemmas 3.1 and 3.2 imply that the vertical projection of \( f^N(A) \) has length greater than 1 and \( \Pi_y(f^{2N+2}(Z)) \in \Pi_x(B) \) and \( \ell(\Pi_y(f^{2N+2}(A))) > 1 \) too. Therefore \( f^{2N+2}(A) \cap B \neq \emptyset \).

Since the length of the horizontal projection doubles under iterations by the action \( x \mapsto 2x \mod 1 \), there is \( N_1 > 0 \) such that for \( n > N_1 \) we have that \( \Pi_x(f^n(A)) \) covers all \([0, 1]\). It follows that \( f^{N_1+2N+2k}(A) \cap B \neq \emptyset \) for all \( k \geq 0 \). Thus \( f \) is topologically mixing, proving theorem A. \(\square\)

4. Lebesgue measure preserved and mixing

In this section we prove theorem B. We start establishing some auxiliary lemmas. The first says that for \( c > 1/4 \) we have that the pre-images by \( f \) expand in length in the vertical direction.

In the next lemma \( \ell \) stands for the length in the torus.

Lemma 4.1. Given \( \epsilon > 0 \) there is \( N(c) = N > 0 \) such that for \( n \geq N \), every horizontal arc \( I, \ell(I) = \Delta x, \) every \( b = b_1b_2 \cdot b_N \) it results \( l(Z_b(I)) > (4 - \epsilon)c \Delta x \).
Proof. Given a segment \([x, x + \Delta x] \subset [0, 1/2]\), the length of its image by any of its branches: \(Z_0, Z_1\) is given by

\[
\int_{x/2}^{x/2+\Delta x/2} \sqrt{1 + \frac{c^2}{(1/2-s)^2}} \, ds \geq \frac{c}{1/2-s} \frac{(x+\Delta x)/2}{x/2} > 2c\Delta x.
\]

Analogously for the four second branches \(Z_{00}, Z_{01}, Z_{10}, Z_{11}\) we have that, in appropriate coordinates \((u, y)\), \(u \in [h, h + \Delta x/4], h < 1/4\), the graph of the pre-images is given by

\[
g(u) = y_0 - 2c - \frac{2}{1 - 2u} - \frac{2c}{1 - 4u}.
\]

Putting \(h' = 4h\), and calculating the length of the graph \(^5\) we get

\[
\int_h^{h + \Delta x/4} \sqrt{1 + (g'(u))^2} \, du \geq \int_h^{h + \Delta x/4} |g'(u)| \, du
\]

\[
= \frac{2c}{1 - h'} \left( \frac{1}{1 - \frac{\Delta x}{1 - h'}} - 1 \right) + \frac{2c}{1 - h'/2} \left( \frac{1}{1 - \frac{\Delta x/2}{1 - h'/2}} - 1 \right)
\]

\[
> 2c\Delta x + c\Delta x = 3c\Delta x.
\]

By induction we obtain in the general case \((n \geq 3)\) that the length of \(Z_b([x_0, x_0 + \Delta x], y)\) is bounded from below by

\[
\ell(Z_b([x_0, x_0 + \Delta x], y)) \geq \left( 2 + \frac{n-2}{2^j} \right) c\Delta x = \left( 3 + \left( 1 - \frac{1}{2^{n-2}} \right) \right) \cdot c \cdot \Delta x.
\]

Choosing \(\frac{1}{2^n} < \epsilon\) the lemma follows if the length of the sequence \(b\) is greater or equal to \(N = N(c)\).

Lemma 4.2. The Lebesgue measure \(m\) is preserved by the map

\[
f(x, y) = \left( 2x - \lfloor 2x \rfloor, y + \frac{c}{|x - 1/2|} - \left\lfloor y + \frac{c}{|x - 1/2|} \right\rfloor \right), \quad c \in \mathbb{R}^+.
\]

Proof. There is no loss of generality in choosing \(A\) a rectangle contained in \(Q\), since the family of rectangles generates the \(\sigma\)-algebra associated to the Lebesgue measure [Wa, theorem 1.17]. Given any small rectangle \(A = (a, b) \times (d, e) \subset (0, 1) \times (0, 1)\) it has two pre-images which are the subsets \(A_0, A_1\) where \(A_0\) is limited by the lines

\[
x = \frac{a}{2}, \quad x = \frac{b}{2},
\]

and the graph of the broken hyperbolas

\[
y = d - \frac{c}{1/2-x} - \left\lfloor d - \frac{c}{1/2-x} \right\rfloor, \quad y = e - \frac{c}{1/2-x} - \left\lfloor e - \frac{c}{1/2-x} \right\rfloor,
\]

\[
a/2 \leq x \leq b/2;
\]

and \(A_1\) is limited by the lines

\[
x = \frac{1+a}{2}, \quad x = \frac{1+b}{2},
\]

\(^5\) Note that \(1 \geq 1 - h' > 0\).
and the graph of the broken hyperbolas
\[ y = d - \frac{c}{x - 1/2}, \quad \left[ d - \frac{c}{x - 1/2} \right]_{(1+a)/2} \leq x \leq (1+b)/2, \]
\[ y = e - \frac{c}{x - 1/2}, \quad \left[ e - \frac{c}{x - 1/2} \right]_{(1+a)/2} \leq x \leq (1+b)/2. \]

Calculating the area of \( A_0 \) by integration we obtain \( (b-a)(e-d)/2 \). Similarly for \( A_1 \). Summing both areas we obtain \( (b-a)(e-d) = \text{Area}(A) \). Since the family of rectangles like \( A \) gives a basis for the \( \sigma \)-algebra associated to Lebesgue measure \( m \) we have proved that \( m \) is \( f \)-invariant. □

Figure 4 shows \( A = [1/4, 1/3] \times [2/3, 3/4] \) and its pre-images, \( A_0 \) and \( A_1 \), where we have chosen \( c = \pi - 3 \approx 0.1416 \). The horizontal sides of \( A \) are mapped into the broken graph of the hyperbolas, the top corresponding to the green line and the bottom to the red one.

In the next lemma we refer to the symbols introduced in section 2 and formula (2).

Lemma 4.3. Given \( b \in [0, 1) \), \( b \sim b_1b_2 \cdots b_{N-1}b_N \),

(a) There are \( N \) vertical asymptotes for (the lift to \( \mathbb{R}^2 \) of) \( y_b(x_b) \) whose horizontal coordinates are
\[ 0.b_1b_2 \cdots b_{N-1} + \frac{1}{2^N}, \quad 0.b_1b_2 \cdots b_{N-2} + \frac{1}{2^{N-1}}, \ldots, \quad 0.b_1 + \frac{1}{2^2}, \quad \frac{1}{2}. \]

(3)

(b) There is at most one point in \((0, 1)\) where the derivative \( y_b'(x_b) \) can vanish. Moreover, if \( y_b'(x_b) \) vanishes, the value of \( x \) at which \( y_b' = 0 \) is in \( \Sigma_{i=0}^1 b_i/2^i \) and \( \Sigma_{i=0}^1 b_i/2^i + 1/2^N \) where \( b_j \) is the last digit in \( b_1b_2 \cdots b_{N-1}b_N \) different from \( b_N \) (i.e., \( b_N = b_{N-1} = \cdots = b_{j+1} \neq b_j \)).

Proof. Let \((x, y)\) be given. For \( N = 1 \) we have that
\[ y_{b_1} = y - \frac{c}{|x_{b_1} - 1/2|} \mod(1) = y - \frac{c \cdot (2b_1 - 1)}{x_{b_1} - 1/2} \mod(1) \]
where \( x_{b_1} = x_{b_1} \in \left(\frac{b_1}{2}, \frac{b_1+1}{2}\right) \). Observe that \( 2b_1 - 1 = -1 \) if \( b_1 = 0 \) and \( 2b_1 - 1 = 1 \) if \( b_1 = 1 \). Thus
\[ y_{b_1}'(x_{b_1}) = \frac{c \cdot (2b_1 - 1)}{(x_{b_1} - 1/2)^2} \]
which does not vanish whenever it exists.
For $N = 2$, on account of
\[
 x_{b_1 b_2} = \frac{x_b + b_1}{2} \in \left( \frac{b_1}{2} + \frac{b_2}{2}, \frac{b_1}{2} + \frac{b_2}{2} + 1 \right)
\]
we have $x_{b_2} = 2x_{b_1 b_2} - b_1$ and
\[
y_{b_1 b_2}(x_{b_1 b_2}) = y_{b_2} = \frac{c}{\left| x_{b_2} - 1/2 \right|} = y_0 - \frac{c}{\left| x_{b_2} - 1/2 \right|} = \frac{c}{(2b_2 - 1)} \mod (1)
\]
from which we conclude that
\[
y_{b_1 b_2}'(x_{b_1 b_2}) = \frac{2c \cdot (2b_2 - 1)}{(2x_{b_2} - b_1 - 1/2)^2} + \frac{c \cdot (2b_1 - 1)}{(x_{b_2} - 1/2)^2},
\]
which does not change sign if $b_1 = b_2$ or changes sign only once in its domain if $b_1 \neq b_2$.

In general the expression of $y_b = y_{b_1 b_2 \cdots b_N}$ as a function of $x_b = x_{b_1 b_2 \cdots b_N}$ is given by
\[
y_b(x_b) = y_0 - \frac{2^{N-1}c \cdot (2b_1 - 1)}{(x_b - 1/2)^2}
\]
\[
= \frac{2^{N-1}c \cdot (2b_1 - 1)}{(x_b - 1/2)^2}.
\]

The last expression can be written as
\[
y_b'(x_b) = \frac{c \cdot (2b_N - 1)}{2^{N-1} \left( x_b - \frac{b_1}{2} + \frac{b_2}{2} + \cdots + \frac{b_{N-2}}{2} + \frac{b_{N-1}}{2} - \frac{1}{2^N} \right)^2} + \frac{c \cdot (2b_1 - 1)}{2^{N-2} \left( x_b - \frac{b_1}{2} + \frac{b_2}{2} + \cdots + \frac{b_{N-2}}{2} + \frac{b_{N-1}}{2} - \frac{1}{2^{N-1}} \right)^2} + \cdots + \frac{c \cdot (2b_1 - 1)}{\left( x_b - \frac{1}{2} \right)^2}.
\]

Hence
\[
y_b'(x_b) = \frac{c \cdot (2b_N - 1)}{2^{N-1} \left( x_b - 0.b_1 b_2 \cdots b_{N-2} b_{N-1} - \frac{1}{2^N} \right)^2} + \frac{c \cdot (2b_1 - 1)}{2^{N-2} \left( x_b - 0.b_1 b_2 \cdots b_{N-2} b_{N-1} - \frac{1}{2^{N-1}} \right)^2} + \cdots + \frac{c \cdot (2b_1 - 1)}{\left( x_b - \frac{1}{2} \right)^2}
\]
where we have written $\frac{b_1}{2} + \frac{b_2}{2} + \cdots + \frac{b_{N-2}}{2} + \frac{b_{N-1}}{2} = 0.b_1 b_2 \cdots b_{N-1}$, and similarly for the other terms.

Observe that each term in the expression or $y_b'(x_b)$ is positive or negative depending on the values of $b_i$, and that there are $N$ vertical asymptotes for (the lift to $\mathbb{R}^2$ of) $y_b(x_b)$ with horizontal coordinates as at (3).
Remark 4.5. Since \( C \) is an asymptote of the positive terms in \( y \), Theorem 4.6. If \( b \) belongs to an interval of length 1 located between the closest asymptotes of different sign.

Proof. The proof is through induction. For \( N = 1 \) there is nothing to prove. For \( N = 2 \), if \( b_2 = 0 \) and \( b_1 = 1 \) then we have an asymptote \( x = \frac{1}{2} \) and the other \( x = \frac{3}{2} \) with derivative

\[
y_{b_2b_1}(x_{b_2}) = y_{10}(x_{10}) = \frac{-c}{2(x_{10} - 3/4)^2} + \frac{c}{(x_{10} - 1/2)^2}.
\]

If \( b_2 = 1 \) and \( b_1 = 0 \) then we have an asymptote \( x = 1/4 \) and the other \( x = 1/2 \) and the derivative is

\[
y_{b_2b_1}(x_{b_2}) = y_{01}(x_{01}) = \frac{c}{2(x_{01} - 3/4)^2} + \frac{-c}{(x_{01} - 1/2)^2}.
\]

Hence the claim is true for \( N = 1, 2 \).

Assume that the claim is true for \( b_2b_3 \ldots b_N \) and let us prove it for \( b = b_1b_2b_3 \ldots b_N \). If \( b_1 = 0 \) then all the values of the asymptotes are divided by 2 and the corresponding asymptotes of the positive terms in \( y_{b}(x_{b}) \) rest to the left of the smallest asymptote corresponding to a negative term (if there is any one). By induction the difference between the smallest asymptote of a negative term and the largest asymptote of a positive term is \( 1/2^{N-1} \) for \( b_2b_3 \ldots b_N \) and when we divide by 2 the difference becomes \( 1/2^N \). Moreover, all terms are less than 1/2 and we add a negative term corresponding to the asymptote \( x = 1/2 \) and we add a positive term corresponding to the asymptote \( x = 1/2 \). Thus the claim is true for \( b_1 = 0 \).

If \( b_1 = 1 \) then all the values of the asymptotes are divided by 2 and to these values we add 1/2. Therefore the corresponding asymptotes of the positive terms in \( y_{b}(x_{b}) \) rest to the left of the smallest asymptote corresponding to a negative term (if there is any). The difference between the smallest asymptote of a negative term and the largest asymptote of a positive term becomes \( 1/2^N \) as above. All terms are greater than 1/2 and we add a positive term corresponding to the asymptote \( x = 1/2 \). Thus the claim is proved.

The proof of the lemma follows readily from claim 4.1, using standard computations similar to those for the case \( N = 2 \).

Remark 4.4. Let \( A = [a, b] \times [d, e] \) be a small rectangle with sides parallel to the coordinate axes, and \( A_0 \) and \( A_1 \) its pre-images. As a consequence of the proof of lemma 4.2 we have that \( m(A_0) = m(A_1) = m(A)/2 \). The measure of each of the \( n \)-pre-images is \( m(A)/2^n \). Moreover, as a consequence of the previous lemma the shape of each pre-image is an ‘almost’ parallelogram bounded by two vertical lines of length \( (e - d) \) and two almost vertical parallel lines, pre-images of \([a, b] \times [d]\) and \([a, b] \times [e]\) respectively.

Remark 4.5. Since \( x \in [0, 1] \), although the number of asymptotes is \( N \), we have that \( x_b \) belongs to an interval of length \( 1/2^N \), determined by two of these asymptotes.

Next, we introduce the following convention that will be used in the following

Convention. Given a positive constant \( C \) and real numbers \( A \) and \( B \), we write \( A \asymp B \) if \( C^{-1} < A/B < C \). Given two sequences of real numbers \( \{A_k\}_{k \in \mathbb{N}} \) and \( \{B_k\}_{k \in \mathbb{N}} \) we write \( A_k \asymp B_k \) if \( A_k \asymp B_k \) for all \( k \in \mathbb{N} \) and the constant \( C_k > 0 \) such that \( C_k^{-1} < A_k/B_k < C_k \) converges to 1 when \( k \to \infty \). When using the symbol \( \asymp \) we usually will not indicate the parameter \( k \).

Theorem 4.6. If \( c > 1/4 \) then the map \( f \) is mixing with respect to the Lebesgue measure.
**Proof.** As mentioned in lemma 4.2, there is no loss of generality in choosing $A$ and $R$ as rectangles with borders parallel to the coordinate axes contained in $Q$. We have to show

$$
\lim_{n \to \infty} |m(f^{-n}(A) \cap R) - m(A) \cdot m(R)| = 0,
$$

where $f^{-n}(A)$ include all the pre-images of $A$: $\cup\{Z_a(A), |a| = n\}$. To do so we proceed as follows. Let $W = (x^R, y^R) \in R$ be the centre point of $R$ and $Z = (x^A, y^A)$ the centre point of $A$. Write $x^R = b_1 b_2 \cdots$ for the binary decomposition of $x^R$ and let $\hat{x} = b_1 b_2 \cdots b_N = b$, that is, $\hat{x}$ is the $N$-periodic point of the map $x \mapsto 2x \text{ mod } 1$.

Taking $N$ sufficiently large, we have that the vertical line $x = \hat{x}$ crosses the rectangle $R$ near its central point $W$. Indeed, $|\hat{x} - x^R| < 2^{-N}$. If $x_b = \Pi_1(Z_b)$, as a consequence of remark 2.1, we obtain $|x_b - \hat{x}| < 2^{-N}$, and also $|x_b - x^R| < 2^{-N}$.

Moreover lemma 4.1 implies that for $n \geq N$, with $N$ large enough, the $n$th pre-image of any horizontal segment of length $\Delta x$ contained in $\pi_1(A)$, is almost vertical and its length is greater than $((4 - \epsilon)c)^{[n/N]} \Delta x$. For $c > 1/4$ we can choose $\epsilon$ such that $d = (4 - \epsilon)c > 1$.

**Claim 4.2.** Let $Z_a(A), a = a_1 a_2 \cdots a_n$, be one of the $2^n$ pre-images of $A$, recall (2). The distance $h_a$ between the pre-images of the horizontal sides (top and bottom) of $A$ is

$$
h_a \simeq \frac{m(A)}{2^n \cdot \Delta x_A \cdot L_a},
$$

where $\Delta x_A$ is the length of the bottom of $A$ and $L_a > d^{[n/N]}$, $d > 1$.

**Proof.** Note that $m(Z_a(A)) = \frac{m(A)}{2^n}$. Moreover, as a consequence of lemma 4.1 the length of $\Pi_1(Z_a(A))$ is given by $\Delta x_A \cdot L_a$ and so the area of $Z_a(A) \simeq (\Delta x_A \cdot L_a) h_a$. Thus the height $h_a$ of the ‘almost’ parallelogram given by $Z_a(A)$ is

$$
h_a \simeq \frac{m(A)}{2^n \cdot \Delta x_A \cdot L_a}
$$

proving the claim. $\square$

Now note that if $n$ is large enough, as a consequence of remark 4.4, $Z_a(A)$ is a long ‘vertical’ strip. And as a consequence of the remarks after equations (2) this happens for each of the $n$th pre-images of $A$ that are almost uniformly distributed in the torus. For one of these strips there are $[\Delta x_A L_a]$ small strips cutting $R$ (the number of connected components of $Z_a(A) \text{ (mod } 1)$ in $[0, 1]^2$). By claim 4.2 the total area of the intersection of a single $Z_a(A)$ with $R$ is

$$
\Delta y_R \cdot h_a \simeq \frac{\Delta y_R \cdot m(A)}{\Delta x_A \cdot 2^n \cdot L_a} = \frac{\Delta y_R \cdot m(A) \cdot [\Delta x_A \cdot L_a]}{2^n \cdot \Delta x_A \cdot L_a} \simeq \frac{\Delta y_R \cdot m(A)}{2^n}.
$$

Since the number of pre-images $Z_a(A)$ satisfying the previous computations is $\simeq 2^n \cdot \Delta x_R$ we obtain that

$$
m(f^{-n}(A) \cap R) \simeq \frac{\Delta y_R \cdot m(A)}{2^n} \cdot 2^n \cdot \Delta x_R \to m(A) \cdot m(R) \quad \text{when } n \to \infty,
$$

completing the proof. $\square$

Lemma 4.2 together with theorem 4.6 prove theorem B.
5. Rate of mixing

In this section we prove that the rate of mixing for $c > 1/4$ is exponential. We start with the following observation

**Remark 5.1.** Given a rectangle $A = [x_A - \Delta x/2, x_A + \Delta x/2] \times [y_A, y_A + \Delta y]$, lemma 4.1 implies that there is a first $N_0 = N_0(\Delta x)$ such that $\ell(Z_0([x_A - \Delta x/2, x_A + \Delta x/2])) \geq 1$ if $|b| = N_0$. Recall $|b| = n$ if $b = (b_1, \ldots, b_n)$.

The following lemma says that far from the asymptotes the growth of lengths of the pre-images is bounded from above.

**Lemma 5.2.** Let $c > 0$. Given $K > 0$ and $\epsilon > 0$ there is $\delta > 0$ such that if $0 < \Delta x \leq \delta$ then $N_0 > K$ for a subset $\Sigma$ of measure greater or equal than $1 - K\epsilon$.

**Proof.** Let us choose a vertical strip $S_\epsilon = [1/2 - \epsilon/2, 1/2 + \epsilon/2] \times [0, 1)$ and assume that $I = [x - \Delta x/2, x + \Delta x/2] \times \{y\}$ does not intersect $S_\epsilon$. Let us bound from above the length of the pre-images of $I$. Recall, see equation (2), that these pre-images are given by

Let us denote by $b(K) = b_1 b_2 b_3 \cdots b_K$ the finite subsequence given by the first $K$ terms of $b = \{b_n\}_{n \in \mathbb{N}} \in \Sigma$ and

$$X_K(x, y) = \{(x_{b(n)}, y_{b(n)}) : f^K(x_{b(n)}, y_{b(n)}) = (x, y)\}.$$

There exist $\epsilon > 0$ such that if $f^j([u - \Delta x/2, u + \Delta x/2], v) \cap [1/2 - \epsilon/2, 1/2 + \epsilon/2] \times [0, 1) = \emptyset$ for all $j = 0, 1, 2, \ldots, K - 1$ then

$$\ell(Z_{b(n)}([u - \Delta x/2, u + \Delta x/2], v)) < \left(\frac{4c + 1}{\epsilon^2 - \delta^2}\right)^K$$

from which the thesis follows the shrinking $\delta$ if necessary.

By remark 5.1 after $N_0$ iterations the length of $Z_{b(n)}$ is at least 1. Thus if $k_0$ denotes the time needed for a monotone arc $\gamma$ to duplicate its length when computing $Z_{(b)}$, (see lemma 4.1) we obtain the following

**Corollary 5.3.** If $N = N_0 + k_0 h$, $h \geq 0$, then for $b$ such that $|b| = N$ we have that

$$\ell(Z_b([x_A - \Delta x/2, x_A + \Delta x/2] \times \{y\})) \geq 2^b.$$
Corollary 5.3 implies that the pre-image $Z_b(A)$ has $2^b$ connected components in $[0, 1] \times [0, 1]$ which are almost vertical strips. The value of $N_0$ is bounded but depends on the length of $\Delta x$ and the position of $x_A$.

The next lemma estimates the width of each of these strips. Before we state it, let us sort out the intersections between $Z_b(A)$ and $[0, 1] \times [0, 1]$ in the following way:

(*) The image in $\mathbb{R}^2$ of the top side $T = [x_A - \Delta x/2, x_A + \Delta x/2] \times \{y + \Delta y\}$ of $A$ is an arc almost parallel to the vertical axis $Oy$ with reverse orientation. We assign the label $n$ to the connected component of this arc whose projection covers the interval $[-n + 1, -n]$ in $Oy$ (see figure 5). Similarly for the bottom segment $B = [x_A - \Delta x/2, x_A + \Delta x/2] \times \{y\}$.

To simplify notations, let us introduce the following convention:

**Convention.** Given $K, L$ depending on $\varepsilon > 0$, we set $K \asymp L$ if given $\delta > 0$ there exists $\varepsilon$ such that $|K - L| < \delta$.

**Lemma 5.4.** Let $A = [x_A - \Delta x/2, x_A + \Delta x/2] \times [y_A, y_A + \Delta y]$ and $N = N_0 + k_b h, h \geq 0$, be as above. Denote by $T$ the top and $B$ the bottom sides of the rectangle $A$. Then, for $b$ such that $|b| = N$ there is a positive constant $C = C(A, b, n) < K(c)$ such that

$$\text{dist}(Z_b(T)_n, Z_b(B)_n) = C \frac{\Delta y}{2^n n^2},$$

where $Z_b(T)_n$ and $Z_b(B)_n$ are the $n$th-connected component of $Z_b(T)$ and $Z_b(B)$ respectively.

**Proof.** For the bottom side $B$ the expression of $y_b = y_{b_1 b_2 \cdots b_N}$ as a function of $x_b = x_{b_1 b_2 \cdots b_N}$ is given by (see the proof of lemma 4.3)

$$y_b(x_b) = y_A - \frac{c \cdot (2b_N - 1)}{(2^{N-1}x_b - (2^{N-2}b_1 + 2^{N-3}b_2 + \cdots + 2b_{N-2} + b_{N-1}) - 1/2)} \cdot \frac{c \cdot (2b^{N-1} - 1)}{(2^{N-2}x_b - (2N - 3b_1 + 2^{N-4}b_2 + \cdots + 2b_{N-3} + b_{N-2}) - 1/2)} \cdots \frac{c \cdot (2b_1 - 1)}{(x_b - 1/2)}.$$

This curve has $N$ asymptotes $x_1, x_2, \ldots, x_N$, see equation (3). As a consequence of the previous formula and (2), close to one of each asymptotes with horizontal coordinate $x_j$, $y_b(x_b)$ can be written as

$$y_b(x_b) = y_A + H(x_b) + \frac{c (2b_j - 1)}{2^{j-1} (x_b - x_j)},$$

where $H(x_b)$ has a finite limit $H_j$ when $x_b \to x_j$. Similarly for the top side $T$ we have

$$y_b(x_b) = y_A + \Delta y + H(x_b) + \frac{c (2b_j - 1)}{2^{j-1} (x_b - x_j)}.$$

The only values that give asymptotes in the domain of $x_b$ correspond to the case $j = N$ that give

$$y_b(x_b) = y_A + H(x_b) + \frac{c (2b_N - 1)}{2^{N-1} (x_b - x_N)},$$

and

$$y_b(x_b) = y_A + \Delta y + H(x_b) + \frac{c (2b_N - 1)}{2^{N-1} (x_b - x_N)}.$$
For a fixed $y_b$, varying from $-n + 1$ to $-n$, there exist $\tilde{x}_b$, $\hat{x}_b$ such that

\[(\star_1) \quad y_b = y_A + H(\tilde{x}_b) + \frac{c(2bN - 1)}{2N-1(\tilde{x}_b - x_N)} \quad \text{and} \quad \frac{c(2bN - 1)}{2N-1(\tilde{x}_b - x_N)}\]

\[(\star_2) \quad y_b = y_A + \Delta y + H(\hat{x}_b) + \frac{c(2bN - 1)}{2N-1(\hat{x}_b - x_N)} \quad \frac{c(2bN - 1)}{2N-1(\hat{x}_b - x_N)}\]

For a given small $\varepsilon > 0$, there is $n_0$ such that for $n > n_0$ it holds that $|H(x_b) - H_N| < \varepsilon$. Thus, from the first equation $(\star_1)$ we have that, for $n > n_0$, it holds

\[y_b \cong y_A + H_N + \frac{c(2bN - 1)}{2N-1(\tilde{x}_b - x_N)} \Rightarrow \tilde{x}_b \cong x_N + \frac{c(2bN - 1)}{(y_b - y_A - H_N)2^{N-1}}\]

From the second one $(\star_2)$ we obtain that

\[y_b \cong y_A + H_N + \Delta y + \frac{c(2bN - 1)}{2N-1(\hat{x}_b - x_N)} \Rightarrow \hat{x}_b \cong x_N + \frac{c(2bN - 1)}{(y_b - y_A - H_N - \Delta y)2^{N-1}}\]

This implies that

\[|\hat{x}_b - \tilde{x}_b| \cong \left| \frac{c(2bN - 1)}{(y_b - y_A - H_N - \Delta y)2^{N-1}} - \frac{c(2bN - 1)}{(y_b - y_A - H_N)2^{N-1}} \right|\]

Taking into account that $-n + 1 > y_b > -n$ and $2bN - 1 = \pm 1$ we conclude that

\[|\hat{x}_b - \tilde{x}_b| \cong \frac{c}{2^{N-1}} \left( \frac{1}{(n - y_A - H_N - \Delta y)} - \frac{1}{(-n - y_A - H_N)} \right)\]

\[= \frac{c}{2^{N-1}} \frac{\Delta y}{(n + y_A + H_N)(n + y_A + H_N + \Delta y)} = E \frac{\Delta y}{2^N n^2}\]

The constant $E$ depends, in particular, on $\varepsilon$. Now we choose the constant $C > 0$ such that the relation of the statement holds for all $n \geq 1$ and not only for $n > n_0$. \qed
**Theorem 5.5.** There is $0 < \theta < 1$ such that after $N = N_0 + k_0 h$ iterates by any branch $Z_b$ of $f^{-1}$, the Lebesgue measure of the set of points that has returned to $A$ is greater or equal to $m(A) \cdot \theta$.

**Proof.** Let $\ell(Z_b([x_A - \Delta x/2, x_A + \Delta x/2] \times \{y\})) \geq 2^h$ with $|b| = N$ (corollary 5.3), and assume that $\Pi_1(Z_b(x_A, y) \in [x_A - \Delta x/2 + \frac{1}{2n}, x_A + \Delta x/2 - \frac{1}{2n}]$. Then $Z_b(A)$ cuts $A$ in at least $2^h$ strips which are almost vertical except perhaps for one where the derivative $y'_b(x_b)$ can vanish. We don’t take into account this strip so that we either have $2^h$ almost vertical strips or $(2^h - 1)$ of them. By lemma 5.4 the (almost) vertical sides of the strips (which are at a distance $\frac{\Delta y}{2n}$ between them) intersected with $A$ are mapped by $f^N$, $N = |b|$, in part of the horizontal sides of length $C \Delta y / n^2$ with $N_0 \leq n \leq 2^h$; recall that $C$ is uniformly bounded although it depends on $A, b$ and $n$.

Thus the area of $A$ covered by the $f^N$-image of one of the strips is equal to a constant $D$ multiplied by the length $\Delta y / n^2$ of the horizontal sub-intervals, by the height $\Delta y$ which gives

$$\text{Area}_{\theta} = D \cdot \frac{\Delta y}{n^2} \cdot \Delta y.$$ 

It follows that the area of the $f^N$-image of the $2^h$ strips is $\simeq D \cdot (\Delta y)^2 \sum_{n=N_0}^{2^h} \frac{1}{n^2}$. Since any point in $A$ has $2^N$ pre-images by the different $Z_b$, $|b| = N$, we have to divide this number by $2^N$ in order not to multiple count. This gives us

$$D \cdot (\Delta y)^2 \sum_{n=N_0}^{2^h} \frac{1}{2n}.$$ 

Since the number of pre-images from $N_0$ to $N$ that cut $A$ is given by the action of $x \mapsto 2x \mod (1)$ in $[0, 1)$ we have that this number is $\simeq 2^{2h} \Delta x$. Hence we have that the covered area of $A$ by the set of points that have returned after $N$ pre-images is

$$\simeq D \cdot (\Delta y)^2 \sum_{n=N_0}^{2^h} \frac{1}{2n} \Delta x \cdot 2^{2h} = \Delta x \Delta y \left( D \cdot \Delta y \cdot \frac{2^{2h} \cdot 2^{2N_0}}{2^{2N_0 + \theta}} \sum_{n=N_0}^{2^h} \frac{1}{n^2} \right)$$

$$= m(A) \left( D \cdot \Delta y \cdot 2^{-N_0} \cdot \sum_{n=N_0}^{2^h} \frac{1}{n^2} \right) = m(A) \cdot \theta,$$

where

$$\theta = D \cdot \Delta y \cdot 2^{-N_0} \cdot \sum_{n=N_0}^{2^h} \frac{1}{n^2} < 1.$$ 

Therefore the measure of the set of points that have not returned yet is at most

$$m(A)(1 - \theta).$$

After taking $2^N$ new pre-images (i.e. by backward iteration $N$ times following all the possible $2^N$ branches $Z_b$, $|b| = N$, from the new starting point) we cover $m(A)(1 - \theta)^2$ points that have not returned to $A$ yet. We conclude by induction that for $|b| = n N$ the measure of points not covered after taking all $2^{nN}$ pre-images is less than $m(A)(1 - \theta)^n \rightarrow 0$ when $n \rightarrow \infty$. \(\square\)

**Corollary 5.6.** We have that after $n N$ iterates, the Lebesgue measure of points that have not yet returned is less than $m(A)(1 - \theta)^n$.

The next corollary gives that the rate of recurrence of $f$ is exponential.
Corollary 5.7. There exists $\tau < 1$ such that $|m(f^{-n}(A) \cap A) - m^2(A)| \leq \tau^n m^2(A)$.

Proof. Note that by theorem 5.5 we have that the measure of points that have returned to $A$ after $N$ iterations is

$$m(A) \left( D \cdot \Delta y \cdot 2^{-N_0} \cdot \sum_{n=N_0}^{2^h} \frac{1}{n^2} \right).$$

We write this expression as

$$m(A) \left( D \cdot \Delta x \cdot \Delta y \cdot \frac{2^{-N_0}}{\Delta x} \cdot \sum_{n=N_0}^{2^h} \frac{1}{n^2} \right) = m^2(A) \left( D \cdot \frac{2^{-N_0}}{\Delta x} \cdot \sum_{n=N_0}^{2^h} \frac{1}{n^2} \right).$$

For $N_0$ sufficiently large we have that $\lambda = (D \cdot \frac{2^{-N_0}}{\Delta x} \cdot \sum_{n=N_0}^{2^h} \frac{1}{n^2}) < 1$ and therefore we obtain that after $n = hN$ iterations

$$|m(f^{-n}(A) \cap A) - m^2(A)| \leq |(m(A))^2(1 - \lambda^{\frac{1}{N}})| = (m(A))^2 \lambda^{\frac{1}{N}}.$$

Putting $\lambda^{\frac{1}{N}} = \tau < 1$ we have

$$|m(f^{-n}(A) \cap A) - m^2(A)| \leq m^2(A) \tau^n,$$

proving the thesis.

The following theorem is similar to theorem 5.5.

Theorem 5.8. Given small rectangles $A = [x_A - \frac{\Delta x}{2}, x_A + \frac{\Delta x}{2}] \times [y_A, y_A + \Delta y_A]$ and $B = [x_B - \frac{\Delta x}{2}, x_B + \frac{\Delta x}{2}] \times [y_B, y_B + \Delta y_B]$ there is $0 < \theta < 1$ such that the set of points of $A$ that has visited $B$ after $N$ iterates is greater or equal than $m(AB) \cdot \theta$.

Proof. By corollary 5.3 we have that $\ell(Z_b([x_A - \frac{\Delta x}{2}, x_A + \frac{\Delta x}{2}] \times \{y\})) = 2^h$ with $|b| = N$. Assume that $\Pi_1(Z_b(x, y)) \in \{x_B - \frac{\Delta x}{2} + \frac{1}{2^N}, x_B + \frac{\Delta x}{2} - \frac{1}{2^N}\}$. Then $Z_b(A)$ cuts $B$ in $2^h$ strips which are almost vertical except perhaps for one of them corresponding to that strip where the derivative $\gamma_b(x_B)$ vanishes. We don’t take into account this strip so that we either have $2^h \Delta x_A$ or $(2^h - 1) \Delta x_A$ almost vertical strips. The area of $Z_b(A)$ is $m(A)/2^N$.

By lemma 5.4 and taking into account the sorting given at (⋆), the intersection of the (almost) vertical sides of the strips with $B$ are mapped by $f^N$, $N = |b|$, in a subsegment of the horizontal sides of $A$ with length $\sim C \Delta y_B/n^2$, $N_0 \leq n \leq 2^h$, recall corollary 5.3.

Thus, the area covered by the $f^{-N}$-image of the $n$th-strip is $(\text{Area}_n \simeq D \cdot \frac{\Delta x}{2^N} \cdot \Delta y_A$, where $D$ is a constant, $\Delta y_B$ is the length of the vertical side of $B$, and $\Delta y_A$ is the length of the vertical side of $A$.

Therefore, the area of the $f^{-N}$-image of all the $(2^h - N_0)$ strips is

$$\sum_{n=N_0}^{2^h} \text{Area}_n \simeq D \cdot (\Delta y_A)(\Delta y_B) \sum_{n=N_0}^{2^h} \frac{1}{n^2}.$$

Since any point in $A$ has $2^N$ pre-images by the different $Z_b$, $|b| = N$, we have to divide this number by $2^N$ in order not to multiple count. This gives us

$$D \cdot (\Delta y_A)(\Delta y_B) \sum_{n=N_0}^{2^h} \frac{1}{n^2}. $$
Since the number of pre-images from $N_0$ to $N = N_0 + 2^k \theta$ that cut $B$ is given by the action of $x \mapsto 2x \mod (1)$ in $[0, 1)$, which is Bernoulli, we have that this number is $\approx 2^{k \theta} \Delta x_B$. Hence we have that the area of the set of points that have cut $B$ after $N$ pre-images is

$$\simeq D \cdot (\Delta y_A)(\Delta y_B) \frac{\sum_{n=N_0}^{2^{k \theta} \Delta x} \frac{1}{2^n}}{\Delta x_B \cdot 2^k \theta} = \Delta x_B \Delta y_B \left( D \cdot \Delta y_A \cdot \frac{2^k \theta}{2^{N_0+k \theta}} \cdot \sum_{n=N_0}^{2^k} \frac{1}{n^2} \right)$$

$$= m(B) \left( D \cdot \Delta y_A \cdot 2^{-N_0} \cdot \sum_{n=N_0}^{2^k} \frac{1}{n^2} \right) = m(B) \cdot \theta,$$

where

$$\theta = D \cdot \Delta y_A \cdot 2^{-N_0} \cdot \sum_{n=N_0}^{2^k} \frac{1}{n^2} < 1.$$

Therefore the measure of the set of points of $A$ that have not yet visited the set $B$ is at most $m(B)(1 - \theta)$. After taking $2^N$ new pre-images (i.e. by backward iteration $N$ times following all the possible $2^N$ branches $Z_b$, $|b| = N$, from the new starting point) we cover $m(B)(1 - \theta) \theta$ which implies that the measure of the points that have not visited $B$ yet is at most $m(B)(1 - \theta)^2$.

By induction we conclude that for $|b| = nN$ the measure of points not covered after taking all $2^n$ pre-images is less than $m(B)(1 - \theta)^n$ when $n \to \infty$. \hfill \Box

The proof of the following corollary is similar to that of corollary 5.7. It finishes the proof of theorem C.

**Corollary 5.9.** The exist $\tau < 1$ such that for any two rectangles $A, B$, it holds that $|m(f^{-n}(A) \cap B) - m(A)m(B)| \leq \tau^n m(A)m(B)$.

**Acknowledgments**

RM and JLV are partially supported by Grupo de Investigación ‘Sistemas Dinámicos’ CSIC (Universidad de la República), SNI-ANII, PEDECIBA, Uruguay. MJP is partially supported by CNPq Brazil, Pronex on Dynamical Systems, FAPERJ, Balzan Research Project of J Palis.

**References**

[AP] Araújo V and Pacifico M J 2010 *Three-Dimensional Flows* (Ergebnisse der Mathematik und ihrer Grenzgebiete vol 53) (Heidelberg: Springer)

[BY] Benedicks M and Young L-S 2000 Markov extensions and decay of correlations for certain Hénon maps *Astérisque* **261** 13–56

[Bu] Bunimovich L A 1979 On the ergodic properties of nowhere dispersing billiards *Commun. Math. Phys.* **65** 295–312

[CM] Chernov N and Markarian R 2006 *Chaotic Billiards* (Mathematical Surveys and Monographs vol 127) (Providence, RI: American Mathematical Society)

[CY] Chernov N and Young L-S 2000 *Decay of Correlations for Lorentz Gases and Hard Balls* (Encyclopaedia of Mathematical Sciences vol 101) (Berlin: Springer) pp 89–120

[GW] Guckenheimer J and Williams R F 1979 Structural stability of Lorenz attractors *Publ. Math. IHES* **50** 59–72
Exponential speed of mixing for skew-products with singularities

[Lo] Lorenz E N 1963 Deterministic nonperiodic flow J. Atmos. Sci. 20 130–41
[MPP] Morales C A, Pacifico M J and Pujals E R 2004 Robust transitive sets for 3-flows are partially hyperbolic attractors or repellers Annal. Math. 160 375–432
[Pe] Pesin Ya B 1977 Characteristic Lyapunov exponents and smooth ergodic theory Russ. Math. Surv. 32 55–112
[Ru] Ruelle D 1979 Ergodic theory of dynamical systems Publ Math IHES 50 27–58
[Si] Sinai Ya G 1970 Dynamical systems with elastic reflections Russ. Math. Surv. 25 137–89
[Wa] Walters P 1982 An Introduction to Ergodic Theory (Graduate Texts in Mathematical vol 79) (Berlin: Springer)