PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 113, Number 2, October 1991

MARKOV PARTITIONS FOR THE TWO-DIMENSIONAL TORUS

MARK R. SNAVELY

(Communicated by Kenneth R. Meyer)

ABSTRACT. We examine Markov partitions for hyperbolic automorphisms of $\mathbb{T}^2$ in the spirit of Adler, Weiss, and others and give necessary conditions on the transition matrix of a Markov partition for a given automorphism. We give necessary and sufficient conditions for partitions with two connected rectangles.

1. BACKGROUND DEFINITIONS

We begin by briefly reviewing the notions of a hyperbolic automorphism of $\mathbb{T}^2$ ($\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ where $\mathbb{R}^2 \xrightarrow{\pi} \mathbb{R}^2/\mathbb{Z}^2$) and that of a Markov partition. For a more detailed introduction see [AW, D, Sn]. Let $\mathscr{A}$ be a 2×2 matrix with integer entries. Suppose also that $\det(\mathscr{A}) = \pm 1$ so that $\mathscr{A}^{-1}$ is an integer matrix. In other words, $\mathscr{A} \in GL(2, \mathbb{Z})$. We further suppose that none of the eigenvalues of $\mathscr{A}$ have modulus 1. Therefore by the Perron–Froebenius Theorem we have real eigenvalues λ_u and λ_s satisfying $|\lambda_u| > 1 > |\lambda_s| > 0$. We call λ_u the *unstable eigenvalue* of $\mathscr{A}$ and λ_s the *stable eigenvalue* of $\mathscr{A}$. The corresponding eigenvectors, $\vec{v}_u$ and $\vec{v}_s$, are called the *unstable* and *stable eigenvectors* respectively. We note at this time that $\mathrm{Tr}(\mathscr{A}) \neq 0$. If $\mathrm{Tr}(\mathscr{A}) > 0$, then $\lambda_u > 0$ and if $\mathrm{Tr}(\mathscr{A}) < 0$, then $\lambda_u < 0$. $\mathscr{A}$ induces an automorphism of $\mathbb{T}^2$, which we will also denote $\mathscr{A}$. We define $W^u(x)$ for $x \in \mathbb{T}^2$ as the projection of a line through $\pi^{-1}x$ parallel to $\vec{v}_u$ with $W^s(x)$ defined similarly.

Definition 1.1. We define $\mathscr{P} = \{R_i\}_{i=1}^n$ to be a *Markov partition* for $\mathscr{A}$ if the following are true:

(i) $\mathbb{T}^2 = \bigcup_{i=1}^n R_i$;
(ii) $R_i = \overline{\text{interior}(R_i)}$;
(iii) $\text{interior}(R_i) \cap \text{interior}(R_j) = \varnothing$, $i \neq j$;
(iv) if $x, y \in R_i$, then $W^u(x, R_i) \cap W^s(y, R_i) = z \in R_i$ (the intersection is a single point z);
(v) $\mathscr{A}(W^u(x, R_i)) \supset W^u(\mathscr{A}x, R_j)$ where $x \in R_i$ and $\mathscr{A}x \in R_j$;

Received by the editors December 5, 1989; the contents of this paper were presented at the Conference/Workshop on Ergodic Theory and Symbolic Dynamics, University of Washington, Seattle, on June 21, 1989. This conference was sponsored in part by the NSF.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 58F15.

0002-9939/91 \$1.00 + \$.25 per page

(vi) $\mathscr{A}(W^s(x, R_i)) \subset W^s(\mathscr{A}x, R_j)$ where $x \in R_i$ and $\mathscr{A}x \in R_j$;
where $W^u(x, R_i) = W^u_\varepsilon(x) \cap R_i$.

The R_i's defined above are called *rectangles* and for automorphisms of $\mathbb{T}^2$ they can be unions of actual rectangles in the geometric sense. This is true in this paper, and we assume that all rectangles are connected. Disconnected rectangles are covered in [Sn]. Connected rectangles have two sides parallel to a line through $\mathbf{0}$ in the direction of $\vec{v}_u$ and two sides parallel to a line through $\mathbf{0}$ in the direction of $\vec{v}_s$. The boundary of $\mathscr{P}$, which is the union of the boundaries of the R_i's, can then be written as $\partial\mathscr{P} = \partial_u\mathscr{P} \cup \partial_s\mathscr{P}$ where $\partial_u\mathscr{P}$ is the union of the unstable sides of the R_i and $\partial_s\mathscr{P}$ defined similarly. From the definition of Markov partition, we see that (i) $\mathscr{A}^{-1}(\partial_u\mathscr{P}) \subset \partial_u\mathscr{P}$ and (ii) $\mathscr{A}(\partial_s\mathscr{P}) \subset \partial_s\mathscr{P}$. This allows us to define the following:

Definition 1.2. The *unstable core of* $\mathscr{P}$, denoted $\mathscr{U}$, is defined as $\mathscr{U} = \bigcap_{i=1}^\infty \mathscr{A}^{-i}(\partial_u\mathscr{P})$.
Similarly, define the *stable core of* $\mathscr{P}$, denoted $\mathscr{S}$, as $\mathscr{S} = \bigcap_{i=1}^\infty \mathscr{A}^{i}(\partial_s\mathscr{P})$.
We then define the *core of* $\mathscr{P}$, denoted $\mathscr{C}$, as $\mathscr{C} = \mathscr{U} \cup \mathscr{S}$.

From (v) and (vi) in the definition of Markov partition, we know that if $\mathscr{A}(R_i)$ intersects R_j, then $\mathscr{A}(R_i)$ crosses R_j from one end to the other in the unstable direction. Hence, we can define a transition matrix as follows:

Definition 1.3. We define the *Markov matrix* of a Markov partition $\mathscr{P}$ with n rectangles for $\mathscr{A}$ by $m_{i,j} = \{$ the number of times $\mathscr{A}(R_i)$ crosses the interior of $R_j\}$ for $1 \le i, j \le n$.

Example 1.4. Let $\mathscr{A} = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right]$. The eigenvalues of $\mathscr{A}$ are $\lambda_u = (3+\sqrt{5})/2$ and $\lambda_s = (3-\sqrt{5})/2$ with corresponding eigenvectors $\left[\begin{smallmatrix} 1 \\ (\sqrt{5}-1)/2 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ (-\sqrt{5}-1)/2 \end{smallmatrix}\right]$ respectively.

Figure 1.5 is a Markov partition $\mathscr{P}$ for $\mathscr{A}$ and $\mathscr{A}(\mathscr{P})$. We view $\mathbb{T}^2$ as $[0, 1] \times [0, 1]$ with opposite sides identified.

We see that $\mathscr{A}(R_1)$ crosses R_1 twice and R_2 once and $\mathscr{A}(R_2)$ crosses R_1 and R_2 each once giving us $M = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right]$. Notice that if we reverse the labeling of R_1 and R_2 then $M = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right]$.

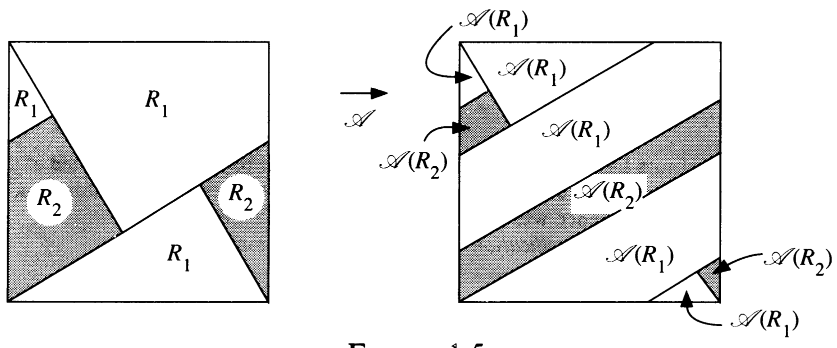


FIGURE 1.5

2. The eigenvalues of the Markov matrix

The goal of this paper is to say what we can about the Markov matrix M. The following theorem tells us about the eigenvalues of M. First, we need the following lemma due to Matt Stafford.

Lemma 2.1 [St]. *Let $h: B \to B$, B finite. Then every eigenvalue of*

$$h_*: \widetilde{H}_0(B) \to \widetilde{H}_0(B)$$

is either 0 or a root of unity. The multiplicity of the 0 eigenvalue is equal to the number of preperiodic points in B. The multiplicity of the eigenvalue 1 is equal to (# of periodic orbits in B) $- 1$.

Proof. If B is a finite union of periodic orbits, $h|_B$ can be represented by a permutation matrix. (Think of each point in B as a basis element.) Clearly, this permutation is cyclic iff B consists of a single periodic orbit.

If B contains preperiodic points as well, then $h|_B$ can be represented by a matrix of the form

$$Q = \begin{bmatrix} P & 0 & 0 & 0 \\ * & 0 & 0 & 0 \\ * & * & \ddots & 0 \\ * & * & * & 0 \end{bmatrix}$$

where P is a permutation matrix representing the action of h on the periodic points in B. The rows below P represent the action of the preperiodic points in B. (Here, one must be careful about the ordering of the basis elements: if x and y are preperiodic and $h(x) = y$, y should be listed before x.) It is clear from the block form of Q that $\chi_Q(t) = \chi_P(t)(-t)^n$, where n is the number of preperiodic points in B. Note that the multiplicity of the root 1 in $\chi_P(t)$ and thus in $\chi_Q(t)$ is equal to the number of periodic orbits in B.

Let $b_0, b_1, \dots, b_{p-1}$ be the elements of B. h induces endomorphisms of $H_0(B) \cong \mathbb{Z}^p$ and $\widetilde{H}_0(B) \cong \mathbb{Z}^{p-1}$. With respect to the basis $b_0, b_1, \dots, b_{p-1}$, the former is represented by the matrix Q described above. Let Q' represent the same transformation in the basis $\sum_{i=0}^{p-1} b_i, b_1 - b_0, b_2 - b_0, \dots, b_{p-1} - b_0$. Q' is of the form

$$Q' = \begin{bmatrix} 1 & * \\ 0 & F \end{bmatrix}.$$

To prove this, it suffices to show that $V = \operatorname{span}\{b_1 - b_0, b_2 - b_0, \dots, b_{p-1} - b_0\}$ is an invariant subspace under h_*. $h_*(b_i - b_0) = h(b_i) - h(b_0)$. If $h(b_i) = h(b_0)$, this difference is $0 \in V$. Otherwise, $h_*(b_i - b_0)$ is in the set $(b_i - b_j | i \neq j)$, which is clearly contained in V. Thus h_* maps the basis elements of V into V; it follows that $h_*(V) \subset V$.

The reason for this change of basis can now be made clear: $b_1 - b_0, b_2 - b_0, \dots, b_{p-1} - b_0$ form a basis of $\widetilde{H}_0(B)$ and $h_*: \widetilde{H}_0(B) \to \widetilde{H}_0(B)$ is represented by the submatrix F (of Q' above) in this basis. Further, $\chi_F(t) = \chi_{Q'}(t)/(1-t) = \chi_Q(t)/(1-t)$. Thus every eigenvalue of F must be 0 or a root of unity. The multiplicity of the eigenvalue 0 for the $p \times p$ matrix Q is not decreased when passing to the $(p-1) \times (p-1)$ matrix F. But the multiplicity of the eigenvalue 1 is one less for F than for Q. This verifies the claim about the multiplicities of the eigenvalues 0 and 1 and completes the proof of the lemma. □

We use both this lemma and its proof in the following theorem:

Theorem 2.2. *Let $\mathscr{P} = \{R_i\}_{i=1}^n$ be a Markov partition for $\mathscr{A}$ where $\mathscr{A}$ is a hyperbolic automorphism of $\mathbb{T}^2$. If $\mathrm{Tr}(\mathscr{A}) > 0$, then the eigenvalues of M are λ_u, λ_s, together with 0's and roots of unity. If $\mathrm{Tr}(\mathscr{A}) < 0$, then the eigenvalues of M are $-\lambda_u$, $-\lambda_s$, together with 0's and roots of unity.*

Proof. A detailed proof of this theorem can be found in [Sn]. We give the proof under the further assumption that both $\partial_u\mathscr{P}$ and $\partial_s\mathscr{P}$ are connected.

Suppose $\mathscr{P}$ has n rectangles, and that $\mathrm{Tr}(\mathscr{A}) > 0$. Let $X = \mathbb{T}^2 - \partial_u^\circ\mathscr{P}$, and $A = X \cap \partial_s^\circ\mathscr{P}$. By $\partial_u^\circ\mathscr{P}$ we mean consider $\partial_u\mathscr{P}$ as a line segment and $\partial_u^\circ\mathscr{P}$ is then the interior of that line segment. In our case, this is just $\partial_u\mathscr{P}$ without the endpoints. A has $n-1$ components and therefore, is homologous to $n-1$ points. The best way to see that A has $n-1$ components is to visualize yourself walking the length of $\partial_s\mathscr{P}$ looking say to the left, counting rectangles once you have crossed them. You pass each rectangle once. Each rectangle is counted at a point where $\partial_u\mathscr{P}$ crosses $\partial_s\mathscr{P}$ completely except two; one is counted at a point where $\partial_u\mathscr{P}$ has an endpoint in $\partial_s\mathscr{P}$ and the last one sees an endpoint of $\partial_s\mathscr{P}$ in $\partial_u\mathscr{P}$. Now $n = \{\text{number of crossings}\} + 2$, hence $\{\text{number of crossings}\} = n-2$. Also, the number of components of A is $\{\text{number of crossings}\} + 1$. Therefore, the number of components in A is $n-1$. X is homologous to $\mathbb{T}^2 - \{\text{one point}\}$. Then we know that $\widetilde{H}_1(A) = 0$, $\widetilde{H}_1(X) = \mathbb{Z}^2$, $\widetilde{H}_0(A) = \mathbb{Z}^{n-2}$, and $\widetilde{H}_0(X) = 0$. The exact relative homology sequence for a pair (X, A) gives us the following commutative diagram.

$$\begin{array}{ccccccccc}
\widetilde{H}_1(A) & \longrightarrow & \widetilde{H}_1(X) & \xrightarrow{\phi} & \widetilde{H}_1(X, A) & \longrightarrow & \widetilde{H}_0(A) & \longrightarrow & \widetilde{H}_0(X) \\
0 & \longrightarrow & \mathbb{Z}^2 & \longrightarrow & \mathbb{Z}^n & \longrightarrow & \mathbb{Z}^{n-2} & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \mathscr{A}_* = \mathscr{A}} & & \downarrow{\scriptstyle M} & & \downarrow{\scriptstyle F} & & \\
0 & \longrightarrow & \mathbb{Z}^2 & \xrightarrow{\phi} & \mathbb{Z}^n & \longrightarrow & \mathbb{Z}^{n-2} & \longrightarrow & 0
\end{array}$$

Because $\mathbb{Z}^m$ is free abelian for $m \geq 0$, the top line of the above sequence is split exact and hence $\widetilde{H}_1(X, A) = \mathbb{Z}^n$. $\mathscr{A}_*$ is the map induced on $\widetilde{H}_1(X)$ by $\mathscr{A}$, which is in fact $\mathscr{A}$ itself. One set of generators for $\widetilde{H}_1(X, A)$ is a line segment across each rectangle in the unstable direction so the induced map on homology is just M. F is from Lemma 2.1. Because the top line of the above sequence is split exact, with possibly a different choice of basis, M can be written in the form:

$$\begin{bmatrix} \mathscr{A} & 0 \\ * & F \end{bmatrix}.$$

Thus the eigenvalues of M are exactly the eigenvalues of $\mathscr{A}$ (λ_u and λ_s) together with the eigenvalues of F (roots of unity and zeros) when $\operatorname{Tr}(\mathscr{A}) > 0$. When $\operatorname{Tr}(\mathscr{A}) < 0$, the induced map on $\widetilde{H}_1(X, A)$ is $-M$ (because $\lambda_u < 0$, $\mathscr{A}$ reverses the orientation of the generators in $\widetilde{H}_1(X, A)$) so in the diagram replace M with $-M$ and the theorem follows similarly, thus concluding the proof for our case; namely, $\partial_u\mathscr{P}$ and $\partial_s\mathscr{P}$ being connected. □

The proof of this theorem gives us the following two results:

Corollary 2.3. *Let $\mathscr{P}$ be a Markov partition for $\mathscr{A}$ with $n > 2$ rectangles. Then the Markov matrix M is similar over $\mathbb{Z}$ to an $n \times n$ matrix of the following form:*

$$\begin{bmatrix} \mathscr{A} & 0 & 0 & 0 \\ * & A_1 & 0 & 0 \\ * & * & \ddots & 0 \\ * & * & * & A_k \end{bmatrix}$$

where the eigenvalues of the nonzero A_i are either {all the pth roots of unity for some p} or {all the pth roots of unity except 1 for some p}.

Proof. By noting that F in the above proof has as eigenvalues only roots of unity and zeros, the corollary follows directly from the proof of Theorem 2.2 and linear algebra [N]. □

Roy Adler, in an as yet unpublished work, proved that every toral automorphism $\mathscr{A}$ has a Markov partition with two rectangles. He in fact proved the existence of a Markov partition for which the Markov matrix is $\mathscr{A}$ itself. We use the first result to give us the following corollary, which is Corollary 2.3 in the 2×2 case:

Corollary 2.4. *Let $\mathscr{P}$ be a Markov partition for $\mathscr{A}$ with two rectangles and Markov matrix M. Then if $\operatorname{Tr}(\mathscr{A}) > 0$, there exists $\phi \in \mathrm{GL}(2, \mathbb{Z})$ such that $\phi^{-1}\mathscr{A}\phi = M$. Similarly, if $\operatorname{Tr}(\mathscr{A}) < 0$, there exists $\phi \in \mathrm{GL}(2, \mathbb{Z})$ such that $\phi^{-1}\mathscr{A}\phi = -M$.*

Proof. Letting X and A be as in Theorem 2.2, with $n = 2$ and $\operatorname{Tr}(\mathscr{A}) > 0$, we have the following diagram.

$$\begin{array}{ccccccc}
\widetilde{H}_1(A) & \longrightarrow & \widetilde{H}_1(X) & \xrightarrow{\phi} & \widetilde{H}_1(X, A) & \longrightarrow & \widetilde{H}_0(A) \\
0 & \longrightarrow & \mathbb{Z}^2 & \longrightarrow & \mathbb{Z}^2 & \longrightarrow & 0 \\
 & & \big\downarrow{\scriptstyle \mathscr{A}_* = \mathscr{A}} & & \big\downarrow{\scriptstyle M} & & \\
0 & \longrightarrow & \mathbb{Z}^2 & \xrightarrow{\phi} & \mathbb{Z}^2 & \longrightarrow & 0
\end{array}$$

ϕ is an isomorphism hence in $\mathrm{GL}(2, \mathbb{Z})$ and in the corollary is proven for $\mathrm{Tr}(\mathscr{A}) > 0$ with the proof for $\mathrm{Tr}(\mathscr{A}) < 0$ being entirely similar. □

At this time we note that $\partial_u \mathscr{P}$ and $\partial_s \mathscr{P}$ being connected implies that $\mathscr{C}(\mathscr{P})$ is at least one fixed point and at most two fixed points. Also, the maps $\mathscr{A}|_{\mathscr{U}}$ and $\mathscr{A}|_{\mathscr{S}}$ together with the signs of λ_u and λ_s determine the nonzero eigenvalues of M. This relationship is explored in [Sn].

3. Some algebraic results

The natural question to ask at this point is whether or not the converse of either of these corollaries is true. Let us examine the 2×2 case. In order to do so we need the following algebraic results.

3.1. *Let $\mathscr{A} \geq 0$ be an element of* $\mathrm{GL}(2, \mathbb{Z})$ *and suppose $\mathscr{A}$ is hyperbolic. Then $\mathscr{A}$ has a dominant row and a dominant column; i.e. the first column of $\mathscr{A} = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is dominant if $a \geq b$ and $c \geq d$ and similarly for a row being dominant.*

Proof. Suppose $\mathscr{A}$ does not have a dominant column. Then either (i) $a > b$ and $c < d$ or (ii) $b > a$ and $d < c$.

Case (i). $a > b$ and $c < d \Rightarrow a \geq b + 1 \Rightarrow ad > bc + c \Rightarrow ad - bc > c$. Therefore, since $\det(\mathscr{A}) = \pm 1$ and $c \geq 0$, we have that $c = 0 \Rightarrow a = d = 1 \Rightarrow b = 0$ (since $a > b$) $\Rightarrow \mathscr{A} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, which is a contradiction to the hyperbolicity of $\mathscr{A}$.

Case (ii). $b > a$ and $d < c \Rightarrow b \geq a + 1 \Rightarrow bc > ad + d \Rightarrow bc - ad > d$. Therefore, since $\det(\mathscr{A}) = \pm 1$ and $d \geq 0$, we have that $d = 0 \Rightarrow b = c = 1 \Rightarrow a = 0$ (since $b > a$) $\Rightarrow \mathscr{A} = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, which is a contradiction. □

3.2. *Suppose $\mathscr{A} \in \mathrm{GL}(2, \mathbb{Z})$, $\mathscr{A} \geq 0$, $\mathscr{A} \neq \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. Then $\mathscr{A} = \prod_{i=1}^{n} \left[\begin{smallmatrix} x_i & 1 \\ 1 & 0 \end{smallmatrix}\right]$ where $x_i = 0$ or 1 for $1 \leq i \leq n$.*

Sketch of proof. We can see the above by observing the following:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} c & d \\ y & z \end{bmatrix}$$

where if $a \geq c$ and $b \geq d$ then let $x = 1$, $y = a - c$, and $z = b - d$ thus reducing the entries in the dominant row. Otherwise let $x = 0$, $y = a$, and $z = b$. We can then view this factorization as a sequence of 1's and 0's $\{x_i\}_{i=1}^{n}$ and this factorization is unique up to two consecutive 0's $(\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]^2 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right])$. We call $\{x_i\}_{i=1}^{n}$ the *defining sequence for* $\mathscr{A}$. From here on we denote $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ as x, $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ as y, and $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right]$ as y'.

3.3. *Under the same hypotheses as in* 3.2, *we see that $\mathscr{A}$ also factors as follows: We define a shear matrix to be a matrix with ones on the diagonal and a single one off the diagonal while all other off-diagonal entries are* 0. *Note that if a shear matrix has a one as its i, jth entry then its inverse has a minus one as its*

i, jth entry. If $\mathscr{A} \neq x$, then $\mathscr{A} = \prod_{i=1}^{m} M_i$ where M_i is a 2×2 shear matrix for $1 \leq i \leq m-1$ and M_m is either a 2×2 shear matrix or y or y'.

Proof. We can see that $xy = y'x$, $yx = \left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]$, a shear, and $y'x = \left[\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right]$, a shear. Given the defining sequence for $\mathscr{A}$, $\{x_i\}_{i=1}^{n}$, we replace 1's with y's and 0's with the x's. Now using the first relation above, put the sequence in the form $s_1 x s_2 x \cdots s_k x s_{k+1} s_{k+2} \cdots s_m$ where s_i can be either y or y' for $1 \leq i \leq m$. (If there are no x's in the sequence, we let $k = 0$.) If $m = k$ or $m = k+1$ then we are done. If $m > k+1$ then insert xx between s_{k+2j-1} and s_{k+2j} for $1 \leq j \leq \lfloor m - k/2 \rfloor$ (where $\lfloor r \rfloor = \{$ the largest integer smaller than $r\}$), replace xs_{k+2j} with sx where s is either y or y' and we are done. □

3.4. *Suppose $\mathscr{A}$ and $\mathscr{B}$ are nonnegative, hyperbolic matrices in* $\mathrm{GL}(2, \mathbb{Z})$. *The following are equivalent:*

(i) *$\mathscr{A}$ and $\mathscr{B}$ are similar over $\mathbb{Z}$.*
(ii) *There is a matrix $P \geq 0$, $P \in \mathrm{GL}(2, \mathbb{Z})$, such that $P^{-1}\mathscr{A}P = \mathscr{B}$.*
(iii) *The defining sequence for $\mathscr{A}$ is a cyclic permutation of the defining sequence for $\mathscr{B}$ up to two consecutive 0's.*

Sketch of proof. (ii) $\Leftrightarrow$ (i) is obvious as (iii) $\Leftrightarrow$ (ii). We then must show that (i) $\Rightarrow$ (iii). One way to do this is to see that (i) $\Rightarrow$ shift equivalent over $\mathbb{Z} \Rightarrow [KR]$ shift equivalent over $\mathbb{Z}^+ \Rightarrow$ (iii). It can also be proven directly by letting $w = \left[\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right]$ and showing that the group generated by x, y, and w has a solvable word problem and is in fact $\mathrm{GL}(2, \mathbb{Z})$. Then using this machinery and a defining sequence for any element of $\mathrm{GL}(2, \mathbb{Z})$ involving x, y, and w, show (i) $\Leftrightarrow$ new defining sequence is cyclic permutation $\Leftrightarrow$ (iii).

4. Partitions with two rectangles

We are now ready to begin examining partitions with two rectangles with the following lemma.

Lemma 4.1. *Let $\mathscr{P}$ be a Markov partition with two rectangles for $\mathscr{A}$ with Markov matrix M and suppose there is a shear matrix S and a matrix $M' \geq 0$ such that $S^{-1}MS = M'$. Then there is a Markov partition $\mathscr{P}'$ for $\mathscr{A}$ with Markov matrix M'.*

Proof. We give the proof for $S = \left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]$ and $\mathrm{Tr}(\mathscr{A}) > 0$ with the proofs for $\left[\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right]$ and $\mathrm{Tr}(\mathscr{A}) < 0$ being similar. Suppose first that $\mathscr{C}(\mathscr{P}) =$ origin. Consider

$$S^{-1}MS = \begin{bmatrix}1 & -1\\ 0 & 1\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} = \begin{bmatrix}a-c & a-c+b-d\\ c & c+d\end{bmatrix} = M'.$$

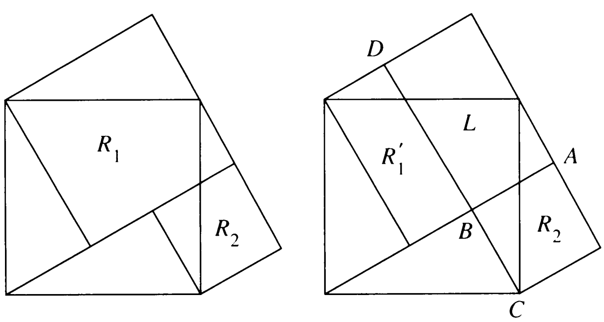
In order for this matrix to be nonnegative, row 1 must be the dominant row in M. Row 1 being dominant means that $\mathscr{A}(R_1)$ crosses each rectangle at least as many times as $\mathscr{A}(R_2)$, which says that R_1 is strictly longer than R_2 in the unstable direction. If R_1 is strictly longer than R_2, we can divide R_1 into two pieces, R_1' and L as shown in Figure 4.2b. To divide R_1 in this manner, draw R_1 and R_2 with their unstable sides furthest from the origin side by side (Figure 4.2a). We then extend segment $\overline{CB}$ to $\overline{CD}$ and now have a Markov partition $\mathscr{T}$ with three rectangles, R_1', L, and R_2. The Markov matrix for $\mathscr{T}$ is

$$M(\mathscr{T}) = \begin{bmatrix} a-c & a-c & b-d \\ c & c & d \\ c & c & d \end{bmatrix}.$$

For row 3, $\mathscr{A}(R_2)$ crosses R_2 the same number of times as in $\mathscr{P}$, and L and R_1' the number of times it crossed R_1 in $\mathscr{P}$. For row 2, $\mathscr{A}(L)$ crosses the same rectangles as $\mathscr{A}(R_2)$; because of the expansive nature of the unstable boundary, the boundary between R_2 and L (segment $\overline{AB}$) gets mapped to the interior of any rectangle it crosses (see Example 1.4). $\mathscr{A}(R_1')$ then crosses everything that R_1 used to cross that is not crossed by $\mathscr{A}(L)$. Then we have seen that $S^{-1}MS$ being nonnegative (in particular, multiplication on the left by S^{-1}) means that we can split R_1 into two rectangles, one of which has the same "personality" as R_2. Now we can remove segment $\overline{AB}$ from $\mathscr{T}$ and get a new partition $\mathscr{P}'$ (Figure 4.2c). This is possible because $\mathscr{A}(\overline{AB}) \cap \partial_u \mathscr{P} = \emptyset$. We claim that the Markov matrix for $\mathscr{P}'$ is M'. The rectangles for $\mathscr{P}'$ are $R_2' = R_2 \cup L$ and R_1'. R_2' is crossed by everything that crossed either R_2 or L, hence, column 2 of M' is {{column 2 of $M(\mathscr{T})$}+ {column 3 of $M(\mathscr{T})$} with either the second or third entry deleted}. Column 1 of M' is {column 1 of $M(\mathscr{T})$ with either the second or third entry deleted}. Hence,

$$M' = \begin{bmatrix} a-c & a-c+b-d \\ c & d+c \end{bmatrix}.$$

This concludes the proof for $\mathscr{C}(\mathscr{P}) = \{\text{origin}\}$. If $\mathscr{C} = \{\text{any fixed point not the origin}\}$, then by translation the above proof works. Then we must deal with the case when $\mathscr{C}(\mathscr{P}) = \{\text{two fixed points}\}$. Notice that if there are more than two fixed points in the core or any periodic point of higher period, then there are more than two rectangles [Sn]. We assume that one of these fixed points is the origin. The only part of the above proof that we must modify is how to



FIGURES 4.2a AND 4.2b

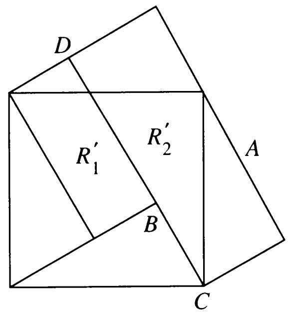


FIGURE 4.2c

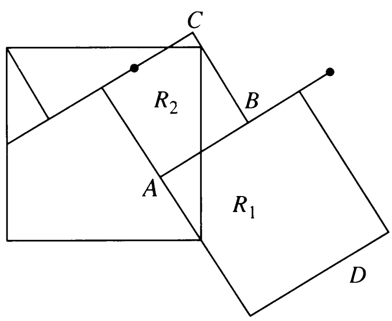


FIGURE 4.3

divide R_1. First, we must have $\mathscr{U} = \{\text{one fixed point}\}$ and $\mathscr{S} = \{\text{the other fixed point}\}$. We assume WLOG that $\mathscr{S} = 0$ and $\mathscr{U} = \{\text{the other fixed point}\}$, call it q. We want to draw $\mathscr{P}$ so that there is no fixed point on $\overline{AB}$ as in Figure 4.3. It is then possible to extend $\overline{CB}$ to $\overline{CD}$ and obtain a partition $\mathscr{T}$. then remove $\overline{AB}$ as above. □

We use the lemma to prove the following proposition that in turn gives us the theorem.

Proposition 4.4. *Let $\mathscr{P}$ be a Markov partition with two rectangles for $\mathscr{A}$ with Markov matrix M and suppose there is a matrix $M' \geq 0$ and $\phi \in \mathrm{GL}(2, \mathbb{Z})$ such that $\phi^{-1}M\phi = M'$. Then there exists a Markov partition $\mathscr{P}'$ for $\mathscr{A}$ with Markov matrix M'.*

Proof. Suppose such a ϕ exists. Then by 3.4 there exists a matrix $P \in \mathrm{GL}(2, \mathbb{Z})$ with $P \geq 0$ such that $P^{-1}MP = M'$. We know by 3.3 that $P = M_1 M_2 \cdots M_k$ where one of the following is true: (i) $k = 1$ and $M_1 = x$; (ii) M_i is a shear matrix for $1 \leq i \leq k$; or (iii) M_i is a shear for $1 \leq i \leq k-1$ and M_k is either y or y' where x and y are as in 3.2.

Case (i). Conjugating M by x corresponds to relabeling the rectangles; i.e. R_1 and R_2 and R_2 as R_1.

Case (ii). We know by Lemma 4.1 that conjugating M by a shear matrix is "legal" as far as partitions go. Therefore, since $P^{-1}M_k^{-1}M_{k-1}^{-1}\cdots M_1^{-1}$, we need only show that $M_j^{-1} \cdots M_1^{-1} M M_1 \cdots M_j \geq 0$ for $1 \leq j \leq k$. It is sufficient to show that $M_1^{-1} M M_1 \geq 0$. So, suppose $P^{-1} M P \geq 0$ with P as above. At this point note that multiplication on the right by a shear matrix does not change the dominant row of M and hence, MP and M have the same dominant row. Therefore, since $M_k^{-1} \cdots M_1^{-1} M P \geq 0$, we know that $M_1^{-1} M P \geq 0$. Since M_1^{-1} is the inverse of a shear matrix and MP and M have the same dominant row, $M_1^{-1} M \geq 0$ also. Hence $M_1^{-1} M M_1 \geq 0$ which concludes case (ii).

Case (iii). If $M_1 = y$ or y' then let $P' = Px$. Now P' is a product of shear matrices and reduces to case (ii). We see, however, that $xP'^{-1} M P' x$ is just $P^{-1} M P$ and hence by case (i) we have proven our proposition. □

Now we give necessary and sufficient conditions on the Markov matrix for a hyperbolic automorphism in the two rectangle case.

Theorem 4.5. *Let $\mathscr{A}$ be a hyperbolic automorphism of $\mathbb{T}^2$, and let M be a nonnegative 2×2 integer matrix.*

(i) *If $\mathrm{Tr}(\mathscr{A}) > 0$, then there exists a Markov partition $\mathscr{P}$ for $\mathscr{A}$ with Markov matrix $M \Leftrightarrow$ there exists $\phi \in \mathrm{GL}(2, \mathbb{Z})$ such that $\phi^{-1} \mathscr{A} \phi = M$.*

(ii) *If $\mathrm{Tr}(\mathscr{A}) < 0$, then there exists a Markov partition $\mathscr{P}$ for $\mathscr{A}$ with Markov matrix $M \Leftrightarrow$ there exists $\phi \in \mathrm{GL}(2, \mathbb{Z})$ such that $\phi^{-1} \mathscr{A} \phi = -M$.*

Proof. ($\Rightarrow$) In both cases this is Corollary 2.4.

($\Leftarrow$) Suppose $\mathrm{Tr}(\mathscr{A}) > 0$, and suppose there exists $\phi \in \mathrm{GL}(2, \mathbb{Z})$ such that $\phi^{-1} \mathscr{A} \phi = M$. Let $\mathscr{P}'$ be any Markov partition for $\mathscr{A}$ with two rectangles. $\mathscr{P}'$ gives rise to a Markov matrix, call it M'. By Corollary 2.4, there exists $\psi \in \mathrm{GL}(2, \mathbb{Z})$ such that $\psi^{-1} \mathscr{A} \psi = M'$. Then $\phi^{-1} \psi M' \psi^{-1} \phi = M \Rightarrow (\psi^{-1}\phi)^{-1} M' (\psi^{-1}\phi) = M \Rightarrow$ by Proposition 4.4 there exists a partition $\mathscr{P}$ for $\mathscr{A}$ with Markov matrix M. If $\mathrm{Tr}(\mathscr{A}) < 0$, let $\mathscr{P}'$ be any Markov partition for $\mathscr{A}$ with two rectangles and Markov matrix $-M$, and Proposition 4.4 gives us the theorem as above. □

5. A CONJECTURE

We know that the Markov matrix M for any hyperbolic automorphism of $\mathbb{T}^2$ must be aperiodic (there is some integer $n \geq 1$ such that $M^n > 0$) because the image of every rectangle is dense in $\mathbb{T}^2$. It is unknown whether or not this is enough for the converse of Corollary 2.3 to be true.

Conjecture 5.1. Let $\mathscr{A}$ be a hyperbolic automorphism of $\mathbb{T}^2$ ($\mathrm{Tr}(\mathscr{A}) > 0$) and let M be an $n \times n$ integer matrix, $n > 2$. Then there is a Markov partition $\mathscr{P}$ with n rectangles for $\mathscr{A}$ with Markov matrix $M \Leftrightarrow M$ is aperiodic and

M is similar over $\mathbb{Z}$ to a matrix of the following form:

$$\begin{bmatrix} \mathscr{A} & 0 & 0 & 0 \\ * & A_1 & 0 & 0 \\ * & * & \ddots & 0 \\ * & * & * & A_k \end{bmatrix}$$

where at least one but at most two of the nonzero A_i have eigenvalues {the pth roots of unity excluding 1 for some p}, and the rest of the nonzero A_i have eigenvalues {the pth roots of unity for some p}.

References

[AW] R. L. Adler and B. Weiss, *Similarity of automorphisms of the torus*, Mem. Amer. Math. Soc. **98** (1970), 17–36.

[D] R. L. Devaney, *An introduction to chaotic dynamical systems*, Benjamin/Cummings, Menlo Park, CA, 1986.

[KR] K. H. Kim and F. W. Roush, *Some results on decidability of shift equivalence*, J. Combin. Inform. System Sci. (1979), 123–146.

[N] M. Newman, *Integral matrices*, Academic Press, New York, 1972, pp. 50–52.

[Sn] M. R. Snavely, *Markov partitions for hyperbolic automorphisms of the two-dimensional torus*, Ph.D. thesis, Northwestern Univ., 1990.

[St] M. W. Stafford, *Markov partitions for the doubling map*, Ph.D. thesis, Northwestern Univ., 1989.

Department of Mathematics, Northwestern University, Evanston, Illinois 60208
Current address: Department of Mathematics, Carthage College, Kenosha, Wisconsin 53140
E-mail address: snavely@rufus.math.nwu.edu