TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 354, Number 7, Pages 2849–2863
S 0002-9947(02)03003-9
Article electronically published on February 26, 2002

A MARKOV PARTITION THAT REFLECTS THE GEOMETRY OF A HYPERBOLIC TORAL AUTOMORPHISM

ANTHONY MANNING

ABSTRACT. We show how to construct a Markov partition that reflects the geometrical action of a hyperbolic automorphism of the n-torus. The transition matrix is the transpose of the matrix induced by the automorphism in u-dimensional homology, provided this is non-negative. (Here u denotes the expanding dimension.) That condition is satisfied, at least for some power of the original automorphism, under a certain non-degeneracy condition on the Galois group of the characteristic polynomial. The $\binom{n}{u}$ rectangles are constructed by an iterated function system, and they resemble the product of the projection of a u-dimensional face of the unit cube onto the unstable subspace and the projection of minus the orthogonal $(n-u)$-dimensional face onto the stable subspace.

1. Introduction

A hyperbolic toral automorphism is the simplest example of a structurally stable diffeomorphism with a dense set of periodic points, see [30, §I.3] or [20]. A Markov partition, [1, 29, 3] or [6, §2.4], is a partition of the state space into subsets called rectangles that enables the orbits of the automorphism to be represented by symbol sequences from some topological Markov chain or subshift of finite type. Counting the periodic points involves also certain auxiliary subshifts of finite type to remedy overcounting of points in the boundaries of the rectangles, see [21, 4]. Equation (2) of [22] (see also [27, section 4]) shows that the main and auxiliary subshifts are related to the automorphism induced in homology. In the present paper we show how to construct a Markov partition with large rectangles that expresses the way the automorphism maps homology classes.

In §2 we explain our main result, Theorem 2.1, and prove the associated Theorem 2.2. Here, the main transition matrix is the transpose M_u^T of the matrix M_u induced by $\overline{M} : \mathbb{T}^n \to \mathbb{T}^n$ in $H_u(\mathbb{T}^n; \mathbb{Z})$, as is suggested by the requirement to have the same topological entropy. The semiconjugacy to $\overline{M}$ is given by a formula

$$\overline{\alpha} : \Sigma(M_u^T) \to \mathbb{T}^n,$$

$$\overline{\alpha}(\omega) := \sum_{r=0}^{\infty} M^{-r}\pi_U z_{\omega_r} - \sum_{r=-\infty}^{-1} M^{-r}\pi_S z_{\omega_r} + \mathbb{Z}^n,\ \omega = (\omega_r)_{r\in\mathbb{Z}},$$

Received by the editors September 4, 2001.
2000 *Mathematics Subject Classification.* Primary 37D20, 37B10; Secondary 28A80, 37B40.
Key words and phrases. Markov partition, hyperbolic toral automorphism, iterated function system.

as in [26, page 3331] or [17, page 360]. See also [11, page 130] and [18, (2.2)] in the one-sided case. (Here $\pi_S z, \pi_U z$ are the stable and unstable components of $z \in \mathbb{Z}^n$.) In §3 we define the term Markov partition and discuss the semi-conjugacy $\overline{\alpha}$. The rectangles formed by $\overline{\alpha}$ are shown in §5 to give a Markov partition, subject to a careful choice of the elements $z \in \mathbb{Z}^n$, which is made in §4.

A different approach via sofic coding may be found in [15, 16, 17, 19].

2. Results

The torus $\mathbb{T}^n$ is given as the quotient group $\mathbb{R}^n/\mathbb{Z}^n$, with its natural projection $\pi_{\mathbb{T}} : \mathbb{R}^n \to \mathbb{T}^n$, $\pi_{\mathbb{T}}(x) := x + \mathbb{Z}^n$, or, equivalently, as the quotient space of the unit cube I^n (where I denotes $[0,1]$) by the usual identifications of opposite faces. In $\mathbb{R}^n$ we use the standard basis $e_j = (\delta_{ij})_{i=1}^n$ for $1 \le j \le n$, which is orthonormal. An $n \times n$ matrix M of integers of determinant ± 1 (so that its inverse is again a matrix of integers) gives an automorphism $\overline{M} : \mathbb{T}^n \to \mathbb{T}^n$, where $\overline{M}(x + \mathbb{Z}^n) = Mx + \mathbb{Z}^n$, for each column vector x in $\mathbb{R}^n$.

The first homology group $H_1(\mathbb{T}^n; \mathbb{R})$ may be identified with $\mathbb{R}^n$ by taking the basis $e_1, \dots, e_n$ to consist of the homology classes of those circles in $\mathbb{T}^n$ that are the images under $\pi_{\mathbb{T}}$ of the one-dimensional edges of the unit cube that join the origin to the vertices $e_1, \dots, e_n$ in $\mathbb{R}^n$. As $\mathbb{T}^n$ is the product of these circles, the Künneth formula [24] gives $H^*(\mathbb{T}^n; \mathbb{R})$ as the exterior algebra generated by $H^1(\mathbb{T}^n; \mathbb{R})$ using the cup product. Dually, the rth homology group $H_r(\mathbb{T}^n; \mathbb{R})$ is a vector space of dimension $\binom{n}{r}$, for $0 \le r \le n$. It has a basis given by the fundamental classes of the r-dimensional tori (whose orientation will be specified in the discussion preceding Theorem 2.2) embedded in $\mathbb{T}^n$ as the image under $\pi_{\mathbb{T}}$ of the r-dimensional vector subspace of $\mathbb{R}^n$ spanned by $\{e_{i_1}, \dots, e_{i_r}\}$ or of the corresponding r-dimensional face of the unit cube in $\mathbb{R}^n$. As an indexing set for this basis we let $C(n,r)$ denote $\{Q \subset \{1, \dots, n\} : \#Q = r\}$. We put an inner product on $H_r(\mathbb{T}^n; \mathbb{R})$ by declaring this basis to be orthonormal.

The automorphism $\overline{M}$ of $\mathbb{T}^n$ induces an automorphism $\overline{M}_{*1}$ of $H_1(\mathbb{T}^n; \mathbb{R})$ which is represented with respect to the basis $\{e_1, \dots, e_n\}$ by the matrix M itself. Dual to the way that the cup products of r generalized eigenvectors in $H^1(\mathbb{T}^n; \mathbb{R})$ are the generalized eigenvectors of $\overline{M}^*$ in $H^r(\mathbb{T}^n; \mathbb{R})$, we find that the eigenvalues of the matrix $\overline{M}_{*r}$ (which we shall write as M_r) are the $\binom{n}{r}$ products of r of the n eigenvalues $\lambda_1, \dots, \lambda_n$ (counted with multiplicity) of $\overline{M}_{*1} = M_1 = M$.

We shall assume throughout that M is hyperbolic, that is, that no eigenvalue of M has modulus 1. Then there is a splitting $\mathbb{R}^n = S \oplus U$ for which $M^k(v) \to 0$ as $k \to \infty$ whenever $v \in S$, and as $k \to -\infty$ whenever $v \in U$. We write π_S, π_U for the projections $\mathbb{R}^n = S \oplus U \to S, U$ onto the factors. Let s and u denote the dimensions of S and U respectively, and let $|\lambda_j| < 1$ for $1 \le j \le s$ while $|\lambda_j| > 1$ for $s < j \le n$. The eigenvalue $\lambda_{s+1} \dots \lambda_n = \mu$, say, of M_u is real, and its modulus is larger than that of any other eigenvalue of $\overline{M}_*$. By [31, Theorem 8.15] the topological entropy $h(\overline{M})$, which is also the entropy of $\overline{M}$ with respect to Lebesgue measure, is $\log(|\mu|)$.

The subshift of finite type σ_A associated with a Markov partition for $\overline{M}$ with transition matrix A has topological entropy $\log(|\mu|)$, the same as $\overline{M}$, see [29, Theorem 5.4] or [20, Theorem 4.9.6]. By [20, Corollary 7.7] or [8, Corollary 3.4], $|\mu|$ must be the dominant eigenvalue of A. Thus the minimum polynomial of $|\mu|$ must divide the characteristic polynomial of A. The matrices M_u and M_u^T (or, if $\mu < 0$, then $-M_u$ and $-M_u^T$), where T denotes transpose, satisfy this property of A. (Under the hypotheses of Theorem 2.2 below, the algebraic number μ has degree $\binom{n}{u}$, so A could not be a smaller matrix.) There is a further constraint on the transition matrix A. Since the subshift of finite type modelling $\overline{M}$ must be mixing, we need A to be non-negative and some power of A to be positive (by which we mean that every entry must be non-negative or positive). Our first theorem asserts that these requirements are sufficient.

Theorem 2.1. *If $\overline{M} : \mathbb{T}^n \to \mathbb{T}^n$ is a hyperbolic toral automorphism with expanding dimension u and the matrix M_u it induces in $H_u(\mathbb{T}^n;\mathbb{R})$ is non-negative and has some power positive, then $\overline{M}$ has a Markov partition with transition matrix M_u^T, the transpose of M_u.*

We shall construct rectangles in §3 depending on a choice of elements of $\mathbb{Z}^n$, and show in §5 that these rectangles do form a Markov partition when the choice is made as in §4.

When *can* we expect that the matrix M_u is non-negative and some power of it is positive? We next derive sufficient conditions and will present them in Theorem 2.2. Notice, first, that the set of eigenvalues of M_u already satisfies one conclusion of the Perron-Frobenius Theorem, namely that some eigenvalue of a positive matrix dominates the modulus of each of the others (if the dominant eigenvalue of M_u is negative we replace M by M^2). However, it need not satisfy the conclusion that the corresponding left and right eigenvectors have each coordinate positive. For example, if N is a 2×2 integer matrix giving a hyperbolic automorphism of $\mathbb{T}^2$ with eigenvalues $\lambda > 1 > \lambda^{-1} > 0$, then $M := \left(\begin{smallmatrix} N & 0 \\ 0 & N \end{smallmatrix}\right)$ determines a hyperbolic automorphism of $\mathbb{T}^4 = \mathbb{T}^2 \times \mathbb{T}^2$. Here $\overline{M}$ has entropy $\log(\lambda^2)$ and λ^2 *is* an eigenvalue of M_2 in $H_2(\mathbb{T}^4;\mathbb{R})$, but the corresponding eigenvectors cannot have all coordinates non-zero because the generator of $H_2(\mathbb{T}^2 \times 0;\mathbb{R}) \subset H_2(\mathbb{T}^4;\mathbb{R})$ is not expanded at the rate of λ^2 but is fixed. Nor is irreducibility of M sufficient to guarantee that the dominant eigenvectors are positive, since $K := \left(\begin{smallmatrix} 0 & L \\ L & 0 \end{smallmatrix}\right)$ is irreducible and $K^2 = M$ if $L^2 = N$, so K_2 has the same eigenvectors for $-\lambda$ as M_2 has for λ^2.

Suppose that the characteristic polynomial $\chi(M)$ of the matrix M is irreducible over the rationals. Then the eigenvalues of M are simple roots. (Also M is irreducible.) Suppose further that the Galois group Γ of the splitting field of $\chi(M)$ over the rationals is isomorphic to the full symmetric group S_n. (This group Γ, which necessarily acts transitively on the eigenvalues [9], thus acts transitively on products of r eigenvalues for each r with $1 \le r < n$.) In [23] we proved in particular that, under these hypotheses, the direct sum of r eigenspaces has trivial intersection with any direct sum of $n - r$ standard basis vectors. It follows that the homology class $[U]$ of U in $H_u(\mathbb{T}^n;\mathbb{R})$ has all its $\binom{n}{u}$ entries nonzero. Let us now choose the sign of each basis vector of $H_u(\mathbb{T}^n;\mathbb{R})$ (or, equivalently, choose an orientation for the corresponding u-dimensional subtori) so that $[U]$ becomes a positive vector.

Again, by [23], S does not meet any u-dimensional coordinate subspace (except at 0). Thus each u-dimensional coordinate subspace approaches U under repeated application of M_u, and eventually has, like $[U]$, all entries positive. Hence some power of M_u has all its columns positive vectors and is a positive matrix. We have now proved the following theorem, which gives sufficient conditions for some power of M to satisfy the hypotheses of Theorem 2.1.

Theorem 2.2. *If the hyperbolic toral automorphism $\overline{M}$ with expanding dimension u has irreducible characteristic polynomial $\chi(M)$ and the Galois group of the splitting field of $\chi(M)$ is the full symmetric group, then there is a choice of sign of the* *standard basis vectors of $H_u(\mathbb{T}^n;\mathbb{R})$ with respect to which some power of the induced map M_u is given by a positive matrix.*

Question. Is there a choice of generators for $H_1(\mathbb{T}^n;\mathbb{Z}) \cong \mathbb{Z}^n$ so that the induced automorphism M_u (rather than some power of M_u) is represented by a non-negative matrix with respect to the corresponding basis of $H_u(\mathbb{T}^n;\mathbb{Z})$ (after an appropriate choice of orientations of those basis vectors)?

3. Markov partitions and the semiconjugacy

In this section we explain our notion of a Markov partition with large rectangles and the corresponding transition matrix, whose entries are non-negative integers. We discuss a semiconjugacy associated to a choice of elements of $\mathbb{Z}^n$, and show that this determines rectangles that arise from fixed points of certain iterated function schemes.

The reader may find it helpful to bear in mind Figure 1, illustrating the rectangles and their images in the case $n = 2$, $M = \left(\begin{smallmatrix} 3 & 1 \\ 2 & 1 \end{smallmatrix}\right)$.

The stable and unstable manifolds of a point $b \in \mathbb{R}^n$ for the hyperbolic automorphism M are the affine subspaces $S + b$ and $U + b$. We use the square brackets notation for the unique point of intersection in $\mathbb{R}^n$ of a stable and an unstable manifold. Thus $[b, c]$ is the unique point in $(S + b) \cap (U + c)$. (Traditionally this notation is used in a *compact* manifold for the intersection of *local* stable and unstable manifolds.) Also $[B, C]$ denotes $\{[b, c] : b \in B, c \in C\}$. Recall that a *rectangle* is a closed set R for which $x, y \in R \Rightarrow [x, y] \in R$. A rectangle is called *proper* if it is equal to the closure of its interior. A subset of $[B, C]$ is called an s-subrectangle if it has the form $[B', C]$ for some $B' \subset B$ and a u-subrectangle if it has the form $[B, C']$ for some $C' \subset C$.

Definition 3.1. We define a *Markov partition* for $\overline{M} : \mathbb{T}^n \to \mathbb{T}^n$ to be a finite set $\mathcal{R}$ of proper rectangles in $\mathbb{R}^n$ so that $\{R + z : R \in \mathcal{R}, z \in \mathbb{Z}^n\}$ tiles $\mathbb{R}^n$ and $\mathcal{R}$ satisfies the following *Markov property.*

(A set of rectangles is said to *tile* $\mathbb{R}^n$ if the union of the rectangles is $\mathbb{R}^n$ while their interiors are disjoint. Traditionally only small rectangles in the compact manifold are used, see [3, §3C] or [14, §18.7], but we use large rectangles, and these are more easily discussed in the universal cover $\mathbb{R}^n$ on which the group $\mathbb{Z}^n$ of covering transformations acts.)

Definition 3.2. The *Markov property* for a Markov partition

$$\mathcal{R} := \{R_1, \dots, R_m\},\ R_j := [B_j, C_j] \text{ for } 1 \le j \le m,$$

is that there exists an $m \times m$ transition matrix A of non-negative integers, and for each i, j there exist closed subsets $B_{i,j,k} \subset B_i, C_{i,j,k} \subset C_j, 1 \le k \le a_{ij}$, whose interiors as subsets of B_i, C_j are non-empty and disjoint, with

$$B_i = \bigcup_{j=1}^{m} \bigcup_{k=1}^{a_{ij}} B_{i,j,k} \text{ and } C_j = \bigcup_{i=1}^{m} \bigcup_{k=1}^{a_{ij}} C_{i,j,k},$$

so that R_i is a union of s-subrectangles $[B_{i,j,k}, C_i]$ and R_j is a union of u-subrectangles $[B_j, C_{i,j,k}]$ that satisfy

$$M[B_{i,j,k}, C_i] + \mathbb{Z}^n = [B_j, C_{i,j,k}] + \mathbb{Z}^n \text{ for } 1 \le k \le a_{ij} \text{ and } 1 \le i, j \le m. \tag{3.1}$$

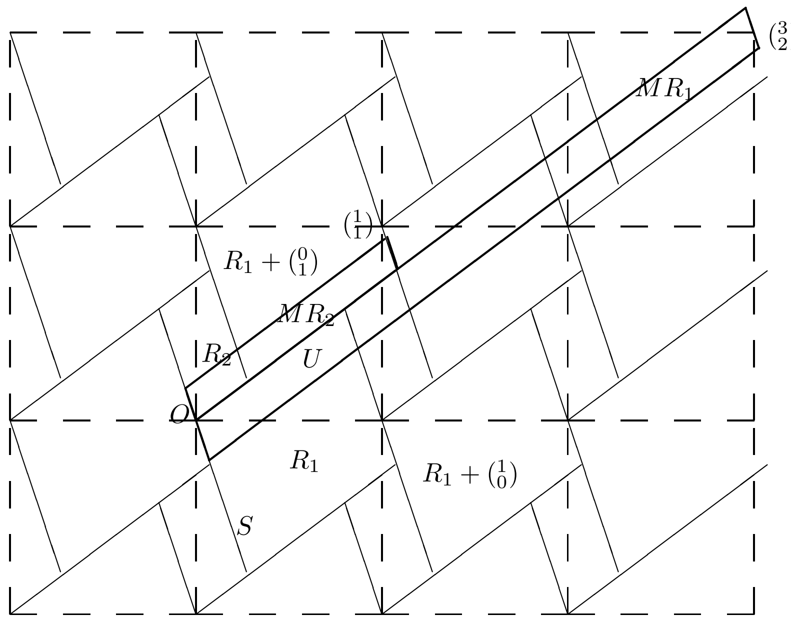

FIGURE 1. The images of rectangles R_1 and R_2 by $M = \left(\begin{smallmatrix}3&1\\2&1\end{smallmatrix}\right)$.

(The traditional formulation with small rectangles allows only $a_{ij} = 0$ or 1 according as $\overline{M}(\mathrm{int}R_i) \cap \mathrm{int}R_j$ is empty or not in $\mathbb{T}^n$. In §4 we shall allocate carefully the elements of $\mathbb{Z}^n$ used in (3.1). A transition matrix with entries larger than 1 was used in [1, Lemma 6.1], [8, Def. 3.1] and [32].) The purpose of a Markov partition for $\overline{M} : \mathbb{T}^n \to \mathbb{T}^n$ is to code $\overline{M}$-orbits by points of a subshift of finite type. The *subshift of finite type* $\sigma_A : \Sigma_A \to \Sigma_A$ is defined by constructing the directed graph Γ_A with m vertices and a_{ij} edges from vertex i to vertex j for $1 \le i, j \le m$. Then a point of Σ_A is a path of *edges* in Γ_A indexed by the integers, and σ_A shifts such a path one place to the left (that is, it reduces the indexing by 1). The *itinerary* of a point $x+\mathbb{Z}^n$ is the path in Γ_A that follows edge k from vertex i at time r to vertex j at time $r+1$, where $M^r(x) \in [B_{i,j,k}, C_i] + \mathbb{Z}^n \subset R_i + \mathbb{Z}^n$, at least for $x + \mathbb{Z}^n$ in the second category subset of $\mathbb{T}^n$ of those points for which each $M^r(x)$ is in a unique such s-subrectangle.

Notation. For $0 \le r \le n$ let

$$C(n, r) := \{Q = (q_1, \dots , q_r) : 1 \le q_1 < q_2 < \dots < q_r \le n\}$$

and order $C(n, r)$ lexicographically. (Put $C(n, 0) = \{\emptyset\}$.) For $P \in C(n, u)$ let $\mathbb{R}^P$ denote $\{x = (x_1, \dots , x_n) \in \mathbb{R}^n : x_j = 0 \text{ for } j \notin P\}$. We use I for the unit interval [0,1] and P^c for the complement of P in $\{1, \dots , n\}$. Thus

$$I^P = \{x = (x_1, \dots , x_n) : j \in P \Rightarrow 0 \le x_j \le 1,\ j \notin P \Rightarrow x_j = 0\}.$$

In proving Theorem 2.1 we shall approach the Markov partition via the semi-conjugacy. Let A denote the $\binom{n}{u} \times \binom{n}{u}$ transition matrix M_u^T. The graph Γ_A has $\binom{n}{u}$ vertices labelled by the elements of $C(n, u)$ and $a_{P,Q}$ edges from vertex P to vertex Q called (P,Q,k), $1\le k\le a_{P,Q}$. A point $\omega\in\Sigma(A)$ is a path consisting of a sequence of *edges* in Γ_A indexed by $\mathbb{Z}$ with $\omega_r=(P_r,Q_r,k_r)$, $r\in\mathbb{Z}$, say.

Suppose now that a label $z_{P,Q,k}\in\mathbb{Z}^n$ has been chosen for each edge in Γ_A. (In §4 we shall make a suitable choice.) We define $\alpha:\Sigma(A)\to\mathbb{R}^n$ by

$$\alpha(\omega):=\sum_{r=0}^{\infty}M^{-r}\pi_U z_{\omega_r}-\sum_{r=-\infty}^{-1}M^{-r}\pi_S z_{\omega_r},\ \omega=(\omega_r)_{r\in\mathbb{Z}},\tag{3.2}$$

and then $\overline{\alpha}:\Sigma(A)\to\mathbb{T}^n$ is given by $\overline{\alpha}=\pi_{\mathbb{T}}\circ\alpha$. By the hyperbolicity of M these series converge to give a point of $\mathbb{R}^n$, and α is continuous.

Then

$$\begin{aligned}M^{-1}\circ\alpha\circ\sigma(\omega) &= M^{-1}\sum_{r=0}^{\infty}M^{-r}\pi_U z_{\omega_{r+1}}-M^{-1}\sum_{r=-\infty}^{-1}M^{-r}\pi_S z_{\omega_{r+1}}\\ &= \sum_{r=1}^{\infty}M^{-r}\pi_U z_{\omega_r}-\sum_{r=-\infty}^{0}M^{-r}\pi_S z_{\omega_r}\\ &= \alpha(\omega)-\pi_U z_{\omega_0}-\pi_S z_{\omega_0}=\alpha(\omega)-z_{\omega_0},\end{aligned}$$

so that $M^{-1}\circ\overline{\alpha}\circ\sigma(\omega)=\overline{\alpha}(\omega)$.

Notice that, for any choice of $\{z_{P,Q,k}\}$, $\overline{\alpha}(\Sigma(A))$ is a closed $\overline{M}$-invariant subset of $\mathbb{T}^n$. Since $\overline{M}$ is ergodic with respect to Lebesgue measure ν_n, either

$$\overline{\alpha}(\Sigma(A))=\mathbb{T}^n\ \text{ or }\ \nu_n\overline{\alpha}(\Sigma(A))=0.\tag{3.3}$$

Consider the image under α of a 0-cylinder:

$$\alpha\{\omega\in\Sigma(A):\ \text{the edge }\omega_0\text{ has initial vertex }P\}=[K_P,L_P]\subset\mathbb{R}^n$$

for $P\in C(n,u)$, where

$$\begin{aligned}K_P &:= \{\pi_U\alpha(\omega):P\text{ is the initial vertex of }\omega_0\},\\ L_P &:= \{\pi_S\alpha(\omega):P\text{ is the initial vertex of }\omega_0\}.\end{aligned}$$

Now, (3.2) shows that

$$\begin{aligned}MK_P &= \bigcup_{Q\in C(n,u)}\bigcup_{k=1}^{a_{P,Q}}(K_Q+M\pi_U z_{P,Q,k}),\\ M^{-1}L_Q &= \bigcup_{P\in C(n,u)}\bigcup_{k=1}^{a_{P,Q}}(L_P-\pi_S z_{P,Q,k}),\end{aligned}$$

so that the families of subsets $(K_P)_{P\in C(n,u)}$, $(L_P)_{P\in C(n,u)}$ of U, S arise as the fixed points of the following iterated function schemes (IFSs) E and F.

The following construction was given in [25, Theorem 1]. Let $\mathcal{K}_0(U)$, $\mathcal{K}_0(S)$ denote the spaces of non-empty compact subsets of U, S with the Hausdorff metric d (see e.g. [7, page 114]). Let

$$\mathcal{K}_u(U):=\prod_{P\in C(n,u)}\mathcal{K}_0(U),\ \ \mathcal{K}_u(S):=\prod_{P\in C(n,u)}\mathcal{K}_0(S)$$

with metric d_u given as the maximum of the distances in the various factors. Then define $E : \mathcal{K}_u(U) \to \mathcal{K}_u(U)$ and $F : \mathcal{K}_u(S) \to \mathcal{K}_u(S)$ by

$$(E(G))_P := \bigcup_{Q \in C(n,u)} \bigcup_{k=1}^{a_{P,Q}} (M^{-1}(G_Q) + \pi_U z_{P,Q,k}), \tag{3.4}$$

$$(F(G))_Q := \bigcup_{P \in C(n,u)} \bigcup_{k=1}^{a_{P,Q}} (M(G_P) - M\pi_S z_{P,Q,k}). \tag{3.5}$$

Because $M^{-1}|U$ and $M|S$ are contractions, E and F are contractions of the complete metric spaces $\mathcal{K}_u(U)$ and $\mathcal{K}_u(S)$. So, as in [12] or [7, page 115],

$$E^r(G) \to (K_P)_{P\in C(n,u)} \quad \text{for any} \quad G = (G_P)_{P\in C(n,u)} \in \mathcal{K}_u(U),$$
$$F^r(G) \to (L_P)_{P\in C(n,u)} \quad \text{for any} \quad G = (G_P)_{P\in C(n,u)} \in \mathcal{K}_u(S).$$

Since some power of A is positive (and σ_A is mixing), we have

$$\text{int} K_P \neq \emptyset \text{ for some } P \in C(n,u) \Rightarrow \text{int} K_Q \neq \emptyset \text{ for all } Q \in C(n,u),$$

and a similar statement for $(L_P)_{P\in C(n,u)}$.

If our choice of labels $z_{P,Q,k}$ yields $\pi_{\mathbb{T}}(\bigcup_P [K_P, L_P]) = \mathbb{T}^n$, then Baire's Theorem gives $\text{int}[K_P, L_P] \neq \emptyset$ for some P, and, arguing as in [18, page 31], we now show that each rectangle $R_P := [K_P, L_P]$ is proper. If for some, and hence all, $P \in C(n,u)$ we have $\text{int} K_P \neq \emptyset$, then we define $K^* \in \mathcal{K}_u(U)$ from K by $K^*_P := \text{cl int} K_P$ for all $P \in C(n,u)$. Then $K^* \subset K$ coordinatewise. Since $M^{-1}|U$ is a homeomorphism, we have $E^r(K^*) \subset K^*$ coordinatewise for $r = 1, 2, \ldots$. But $E^r(K^*) \to K$ as $r \to \infty$, so $K^* = K$ and, for each P, $\text{cl int} K_P = K_P$. Similarly, $\text{cl int} L_P = L_P$ for all $P \in C(n,u)$, provided $\text{int} L_P \neq \emptyset$ for some such P.

4. Allocating transitions in $\mathbb{R}^n$

In this section we make a choice of labels $z_{P,Q,k} \in \mathbb{Z}^n$ by finding a 'step cycle' homologous to the image by M of a subtorus.

The standard basis for $H_u(\mathbb{T}^n; \mathbb{R})$ consists of the fundamental homology classes (with appropriate orientation) of the u-tori that are covered by the u-dimensional coordinate planes; these coordinate planes $\mathbb{R}^Q$ correspond to the elements Q of $C(n,u)$. The dominant right eigenvector of M_u is $[U]$, representing the subspace U (since each is expanded by a factor μ). $[U]$ is given (up to a scalar multiple) by a basis $v_1, \ldots, v_u$ for U, and the Qth coordinate of $[U]$ is the modulus of the determinant of the $n \times n$ matrix with columns $v_1, \ldots, v_u, e_{q'_1}, \ldots, e_{q'_s}$ (where $Q' = (q'_1, \ldots, q'_s) \in C(n,s)$ satisfies $Q \cup Q' = \{1, \ldots, n\}$), which is the volume of the orthogonal projection to $\mathbb{R}^Q$ of the parallelepiped spanned by $v_1, \ldots, v_u$. Since the hypothesis on M_u ensures that each coordinate of $[U]$ is non-zero, we know that

$$\forall Q \in C(n,u) : \ U \cap \mathbb{R}^{Q'} = \{0\}. \tag{4.1}$$

(Provided M is orientation preserving, the functoriality of the intersection form shows that M_s^{-1} is the transpose of M_u (up to the choice of orientation of the basis s-tori and the ordering of this basis), and so $[S]$ is a dominant *left* eigenvector for M_u.)

Our aim of using M_u^T as the transition matrix for our Markov partition whose rectangles correspond to these u-tori accords with the fact that the P-th rectangle has transitions to precisely those rectangles that correspond to the u-tori in the

P-th column of M_u, and so to the P-th row of our transition matrix $A = M_u^T$. The sum of these u-tori gives a cycle in $\mathbb{T}^n$ homologous to $\overline{M}\mathbb{T}^P$; we shall call it a *step* cycle to indicate that it is a sum of singular cubes each parallel to a coordinate plane. Just as we can lift $\overline{M}\mathbb{T}^P$ to a singular cube in $\mathbb{R}^n$ that is a cycle modulo the action of $M\mathbb{Z}^P$, we shall now lift this step cycle to a sum of singular cubes in $\mathbb{R}^n$ that is also a cycle modulo $M\mathbb{Z}^P$ and whose projection to $\mathbb{T}^n$ is homologous to it since they represent the same element of $H_u(\mathbb{T}^n;\mathbb{R})$.

Just as large rectangles are better discussed in $\mathbb{R}^n$ than in $\mathbb{T}^n$, so it will be more convenient to discuss transitions from one rectangle to another in $\mathbb{R}^n$, and this will involve choosing which element of $\mathbb{Z}^n$ specifies the translate of a rectangle that receives a given transition as in (3.1).

For $P = (p_1, \dots, p_r) \in C(n, r)$, we define $\rho_P : I^r \to \mathbb{R}^n$ by $\rho_P(\tau_1, \dots, \tau_r) := (t_1, \dots, t_n)$, where each $t_{p_j} = \tau_j$ and $t_k = 0$ for other k. Here

$$I^r := \{(t_1, \dots, t_r, 0, \dots, 0) \in \mathbb{R}^n : 0 \le t_j \le 1 \text{ for } 1 \le j \le r\},$$

and we let J^P denote the image $\rho_P(I^r)$. In the case $n = 2$, the rectangles

$$[J^P, -J^{P^c}],\ P \in C(2, 1),$$

provide a partition of $\mathbb{T}^2$ as in Figure 1, which motivates us to study

$$[J^P, -J^{P^c}],\ P \in C(n, u),$$

in the general case.

Consider the cubical chain complex $\mathcal{C}_* = \bigoplus_{r=0}^n \mathcal{C}_r$, where $\mathcal{C}_r$ is the free abelian group generated by

$$\{\rho_P + z : P \in C(n, r),\ z \in \mathbb{Z}^n\}.$$

We shall call $\rho_P + z$ a *step* r-cube to indicate that it is parallel to a coordinate r-plane and covers the r-torus $\mathbb{T}^P$.

To illustrate the next definition, Figure 2 shows the points of intersection determining

$$\rho_{\{1\}} + (\rho_{\{2\}} + \tbinom{1}{0}) + (\rho_{\{1\}} + \tbinom{1}{1}) + (\rho_{\{2\}} + \tbinom{2}{1}) + (\rho_{\{1\}} + \tbinom{2}{2})$$

and

$$\rho_{\{2\}} + (\rho_{\{1\}} + \tbinom{0}{1})$$

as step chains in $\mathbb{R}^2$ homologous (when projected to $\mathbb{T}^2$) to the images under $M = \left(\begin{smallmatrix}3 & 1\\ 2 & 1\end{smallmatrix}\right)$ of the two basic 1-chains $\rho_{\{1\}}, \rho_{\{2\}}$. Here $\varepsilon = \left(\begin{smallmatrix}-.2\\ -.16\end{smallmatrix}\right)$.

Definition 4.1. Define a map $Y : \mathcal{C}_* \to \mathcal{C}_*$ as a step version of M. For each r, $0 \le r \le n$, $Y : \mathcal{C}_r \to \mathcal{C}_r$ will be defined from the data $Y(\rho_P)$ by $Y(\rho_P + z) := Y(\rho_P) + Mz$ for $z \in \mathbb{Z}^n$, and by linear extension from its definition on these generators. Thus $M^{-1}Y$ will commute with the action of $\mathbb{Z}^n$ on $\mathcal{C}_*$ by translations.

First fix ε near 0 in $-I^n$ in the full measure set whose complement is the countable union of affine subspaces $M\mathbb{R}^P + \mathbb{R}^{P'} + z$ for $P \in C(n, r), P' \in C(n, r')$, $r + r' < n$ and $z \in \mathbb{Z}^n$. Then define $N : \mathbb{R}^n \to \mathbb{R}^n$ by $Nx := Mx + \varepsilon$. If $N\mathbb{R}^P \cap (\mathbb{R}^{P'} + z) \neq \emptyset$ with P, P', z as above, then, for some $x \in \mathbb{R}^P$, we have $Mx + \varepsilon \in \mathbb{R}^{P'} + z$, and so $\varepsilon \in \mathbb{R}^{P'} + z - Mx \subset M\mathbb{R}^P + (\mathbb{R}^{P'} + z)$, which is impossible, so our choice of ε ensures that the N-image of a subspace does not intersect a lower than complementary dimensional subspace.

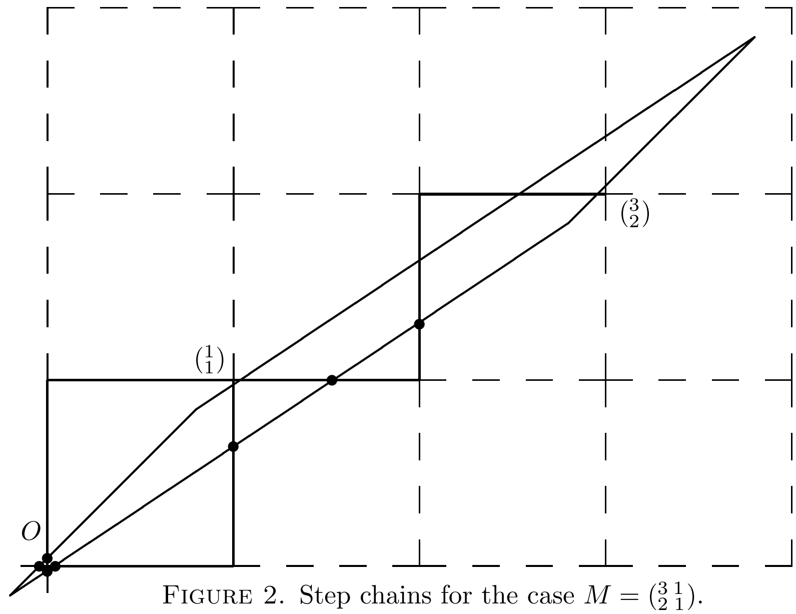

FIGURE 2. Step chains for the case $M = \left(\begin{smallmatrix}3&1\\2&1\end{smallmatrix}\right)$.

Finally, define

$$Y(\rho_P) := \sum_{Q\in C(n,r)} \sum_{i=1}^{|M_r(Q,P)|} \operatorname{sign}(M_r(Q,P))(\rho_Q + Mz_{PQi}), \tag{4.2}$$

where the summand uses those $z_{PQi} \in \mathbb{Z}^n$ for which $NJ^P \cap (-J^{Q^c} + Mz_{PQi}) \neq \emptyset$. Here $M_r(Q,P)$ denotes the entry in row Q and column P in the matrix M_r, and so $\pi_{\mathbb{T}} Y(\rho_P)$ represents the same homology class in $H_r(\mathbb{T};\mathbb{Z})$ as $\overline{M}\pi_{\mathbb{T}}(\rho_P)$, given by column P of M_r.

(The minus sign in $-(\rho_Q + Mz_{PQi})$ means the additive inverse in the abelian group generated by the r-cubes, while $-J^{Q^c}$ means $\{-\rho_{Q^c}(t) : t \in I^{n-r}\}$.) The intuition is that $\{z_{PQi}\}$ is chosen so that $M^{-1}Y\rho_P$ is homologous to ρ_P. The substitutions on stepped surfaces in [13, Fig. 5], which develop [2], provide an illustration of $Y(\rho_P)$.

A point of intersection of NJ^P and $-J^{Q^c} + Mz_{PQi}$ cannot, by the choice of ε, be in the boundary of NJ^P or of $(-J^{Q^c} + Mz_{PQi})$, so there is no ambiguity over whether it should be allocated to $-\rho_{Q^c} + Mz_{PQi}$ or to a neighbouring $(n-r)$-cube. These intersection points are in one-one correspondence via $\pi_{\mathbb{T}}$ with the intersection points of $\overline{M}\mathbb{T}^P$ and $\mathbb{T}^{Q^c}$, and they number $|M_r(Q,P)|$ since $\{[\mathbb{T}^{Q^c}] : Q \in C(n,r)\}$ is a basis for $H_{n-r}(\mathbb{T}^n)$ dual to the basis $\{[\mathbb{T}^Q] : Q \in C(n,r)\}$ for $H_r(\mathbb{T}^n)$ under the intersection form, see [24]. Notice that, in the case $r = 0$, we have $P = \emptyset$, J^P is the origin, $\mathcal{C}_0$ is generated by $\mathbb{Z}^n$, and, since $\varepsilon \in -I^n$, $Y(\rho_\emptyset) = \rho_\emptyset$, so that Y acts as M on the set $\mathbb{Z}^n$ that generates $\mathcal{C}_0$.

Remark 1. We are particularly interested in the case $r = u$, and point out that

$$\begin{aligned} x &\in NJ^P \cap (-J^{Q^c} + Mz_{PQi}) \\ \Rightarrow N^{-1}(x - Mz_{PQi}) &\in N^{-1}(-J^{Q^c}) \cap (-(-J^{P^{cc}}) - N^{-1}Mz_{PQi}). \end{aligned}$$

So our choice of z_{PQi} in (4.2) is suggested not only by the action of N on the u-cube J^P, but also by the action of N^{-1} on the dual s-cube $-J^{Q^c}$. This choice is an analogue for our graph-directed and invertible case of the criterion in [18] of choosing one element from each coset. Indeed, if M is replaced by M^{-1}, then U and E are replaced by S and F.

In §5 we shall show that the above choice of $\{z_{PQi}\}$ makes the rectangles $\{R_P := [K_P, L_P] : P \in C(n,u)\}$ constructed using the IFS in (3.4) and (3.5) have the same volume as $\{[J^P, -J^{P^c}] : P \in C(n,u)\}$. For this we shall need to understand the relationship between Y and the boundary operator.

Definition 4.2. A boundary map $\partial : \mathcal{C}_* \to \mathcal{C}_*$ is defined, see [10, §8.3] or [24, §VII.2], sending each cube to an alternating sum of its faces, by

$$\partial\rho_P := \sum_{k=1}^{r} (-1)^k (\rho_{P\setminus p_k} - (\rho_{P\setminus p_k} + e_{p_k})).$$

Lemma 4.3. *Y commutes with the boundary operator $\partial : \mathcal{C}_* \to \mathcal{C}_*$.*

Proof. It suffices to fix r and $P \in C(n,r)$ and then show that $Y(\partial\rho_P) = \partial(Y\rho_P)$. Clearly both sides are sums over $Q \in C(n, r-1)$ of finite combinations of type $\sum_{z\in\mathbb{Z}^n} k_z(\rho_Q + Mz)$, $k_z \in \mathbb{Z}$. Each contribution of $\rho_Q + Mz$ to $Y(\partial\rho_P)$ comes from the N-image of a face of ρ_P meeting $(-J^{Q^c} + Mz)$, and each such contribution to $\partial(Y\rho_P)$ comes from NJ^P meeting an $(n-r)$-cube that is a face of $(-J^{Q^c} + Mz)$. Thus fix $Q \in C(n, r-1)$, $z \in \mathbb{Z}^n$, and consider $NJ^P \cap (-J^{Q^c} + Mz)$. This intersection is either

1. empty, or
2. a line joining the points of intersection β and β' of NI^P with two $(n-r)$-dimensional faces $-J^{Q^c\setminus k} + Mz - \delta e_k$ and $-J^{Q^c\setminus k'} + Mz - \delta' e_{k'}$ of $-J^{Q^c} + Mz$, where $k, k' \notin Q$ are not necessarily distinct and δ, δ' may be 0 or 1, or
3. a line joining the point of intersection β of NJ^P with one face $-J^{Q^c\setminus k} + Mz - \delta e_k$ of $-J^{Q^c} + Mz$ to the point of intersection β' of $-J^{Q^c} + Mz$ with one face $N(J^{P\setminus k'} + \delta' e_{k'})$ of NJ^P, or else
4. a line joining the points of intersection β and β' of $-J^{Q^c} + Mz$ with two $(r-1)$-dimensional faces $N(J^{P\setminus k} + \delta e_k)$ and $N(J^{P\setminus k'} + \delta' e_{k'})$ of NJ^P.

In case 2, $\partial Y(\rho_P)$ contains a contribution $\rho_Q + Mz$ in $\partial(\rho_{Q\cup k} + Mz - \delta e_k)$ with sign $(-1)^{\#\{j:q_j<k\}+\delta+1}$ and in $\partial(\rho_{Q\cup k'} + Mz - \delta' e_{k'})$ with sign $(-1)^{\#\{j:q_j<k'\}+\delta'+1}$. (Here $Q \cup k$ denotes the r-tuple with k inserted in increasing order, and $P \setminus k'$ denotes the $(r-1)$-tuple with k' deleted.) Now $\operatorname{sign}(M_r(Q\cup k, P))$ is ± 1 according as $\pi_{Q\cup k}M : \mathbb{R}^P \to \mathbb{R}^{Q\cup k}$ respects or reverses orientation. Also

$$(-1)^{\#\{j:q_j<k\}}\operatorname{sign}(M_r(Q\cup k, P)) = \pm 1$$

according as this map respects or reverses orientation when $\mathbb{R}^{Q\cup k}$ is given the orientation $(k, q_1, \ldots, q_{r-1})$; denote this by $\mathbb{R}^{k*Q}$. Thus to show that the two contributions of $\rho_Q + Mz$ in $\partial Y(\rho_P)$ cancel out it suffices to show that $\pi_{k*Q}M : \mathbb{R}^P \to \mathbb{R}^{k*Q}$ and $\pi_{k'*Q}M$ agree or disagree on preserving orientation according as

$\delta-\delta'$ is odd or even. It is equivalent to ask whether π_{k*Q} and $\pi_{k'*Q}$ agree or disagree on preserving orientation as maps from $M\mathbb{R}^P$, or indeed from $\pi_{(k,k')*Q}M\mathbb{R}^P = \operatorname{span}\{\beta'-\beta\}\oplus\mathbb{R}^Q$. And this is equivalent to whether π_k and $\pi_{k'}$ agree or disagree (as maps from $\operatorname{span}\{\beta'-\beta\}$ to $\mathbb{R}^{\{k\}}$ or $\mathbb{R}^{\{k'\}}$). Since $\delta-1<\beta'_k-\beta_k<\delta$ and $-\delta'<\beta'_{k'}-\beta_{k'}<1-\delta'$, these last maps both preserve orientation when $\delta=1,\delta'=0$ and both reverse orientation when $\delta=0,\delta'=1$; in fact they agree precisely when $\delta-\delta'$ is odd, as required.

Now consider case 3. The point β contributes

$$\operatorname{sign}(M_r(Q\cup k,P)).(\rho_{Q\cup k}+Mz-\delta e_k)$$

to $Y(\rho_P)$, while β' contributes

$$\operatorname{sign}(M_{r-1}(Q,P\setminus k')).(\rho_Q+Mz)$$

to $Y(\rho_{P\setminus k'}+\delta' e_{k'})$. Thus, if $k'=p_{j'}$ and $q_{j-1}<k<q_j$, there is a contribution of ρ_Q+Mz from β' to $Y(\partial\rho_P)$ with coefficient

$$(-1)^{j'+\delta'}\operatorname{sign}(M_{r-1}(Q,P\setminus k'))$$

and a contribution of ρ_Q+Mz from β to $\partial(Y\rho_P)$ with coefficient

$$(-1)^{j+\delta}\operatorname{sign}(M_r(Q\cup k,P)).$$

Notice that $\beta'_k-\beta_k$ is positive if $\delta=1$ and negative if $\delta=0$, that $(M^{-1}\beta')_{k'}-(M^{-1}\beta)_{k'}$ is positive if $\delta'=1$ and negative if $\delta'=0$, and that $\pi_Q\beta=\pi_Q\beta'=\pi_Q z$. Thus $\pi_{Q\cup k}M:\mathbb{R}^P\to\mathbb{R}^{Q\cup k}$ respects or reverses orientation according as $(-1)^{\delta'+\delta+j+j'-2}\operatorname{sign}(M_{r-1}(Q,P\setminus k'))$ is positive or negative, where the factor $(-1)^{\delta'+\delta}$ comes from the map $\pi_k M:\operatorname{span}\{M^{-1}(\beta'-\beta)\}\to\mathbb{R}^{\{k\}}$ and the factors $(-1)^{j-1}$ and $(-1)^{j'-1}$ respectively relate $\mathbb{R}^{Q\cup k}$ to $\mathbb{R}^{k*Q}$ and $\mathbb{R}^P$ to $\mathbb{R}^{k'*(P\setminus k')}$. Hence the contributions of ρ_Q+Mz to $Y(\partial\rho_P)$ and $\partial(Y\rho_P)$ agree.

In case 4 we shall show that the two contributions of ρ_Q+Mz to $Y(\rho_{P\setminus k}+\delta e_k)$ and $Y(\rho_{P\setminus k'}+\delta' e_{k'})$ cancel out as contributions to $Y(\partial\rho_P)$. Suppose first that $k<k'$. Now $\operatorname{sign}(M_{r-1}(Q,P\setminus k))$ and $\operatorname{sign}(M_{r-1}(Q,P\setminus k'))$ express respectively whether $\pi_Q M:\mathbb{R}^{P\setminus k}\to\mathbb{R}^Q$ and $\pi_Q M:\mathbb{R}^{P\setminus k'}\to\mathbb{R}^Q$ respect or reverse orientation. These maps are compositions

$$\begin{aligned}\mathbb{R}^{P\setminus k} &\to \mathbb{R}^{k'*(P\setminus k\setminus k')}\to W\oplus\mathbb{R}^{P\setminus k\setminus k'}\to\mathbb{R}^Q,\\ \mathbb{R}^{P\setminus k'} &\to \mathbb{R}^{k*(P\setminus k\setminus k')}\to W\oplus\mathbb{R}^{P\setminus k\setminus k'}\to\mathbb{R}^Q,\end{aligned}$$

where W is the orthogonal complement of $\operatorname{span}\{M^{-1}\beta'-M^{-1}\beta\}$ in $\mathbb{R}^{(k,k')}$. Here the first map is the identity and changes the orientation by $(-1)^{j'-2}$ in the first line and by $(-1)^{j-1}$ in the second. The third map is $\pi_Q M$, the same in each line. The second map is orthogonal projection of the first factor to W and the identity on the second factor. But the projections $\mathbb{R}^{\{k'\}}\to W$ and $\mathbb{R}^{\{k\}}\to W$ agree on whether to respect or reverse orientation if $\delta=\delta'$, and disagree if $\delta\neq\delta'$. Thus

$$\operatorname{sign}(M_{r-1}(Q,P\setminus k))=(-1)^{j-j'+1+\delta-\delta'}\operatorname{sign}(M_{r-1}(Q,P\setminus k')) \tag{4.3}$$

But the two relevant summands in $\partial\rho_P$ are

$$(-1)^{j+\delta}(\rho_{P\setminus k}+\delta e_k)+(-1)^{j'+\delta'}(\rho_{P\setminus k'}+\delta' e_{k'}),$$

so $Y(\partial\rho_P)$ contains contributions of (ρ_Q+Mz) with coefficient

$$(-1)^{j+\delta}\operatorname{sign}(M_{r-1}(Q,P\setminus k))+(-1)^{j'+\delta'}\operatorname{sign}(M_{r-1}(Q,P\setminus k')) \tag{4.4}$$

and, by (4.3), this is 0. The case $k > k'$ is similar. If $k = k'$, then $j = j'$, $\delta \neq \delta'$ and $\operatorname{sign}(M_{r-1}(Q, P \setminus k)) = \operatorname{sign}(M_{r-1}(Q, P \setminus k'))$, so the coefficient given by (4.4) is again 0.

Thus, the contributions of $(\rho_Q + Mz)$ cancel out in $\partial Y(\rho_P)$ in case 2, they cancel out in $Y(\partial\rho_P)$ in case 4, while in case 3 the contributions are the same on both sides, which proves that $\partial Y(\rho_P) = Y(\partial\rho_P)$, as required. □

5. Completing the proof of Theorem 2.1

In this section we investigate the volumes of the rectangles constructed in §3 using the $z_{P,Q,k}$ chosen in §4, and we prove in Theorem 5.4 that they do form a Markov partition. Let ν_u, ν_s, ν_n denote u-, s- and n-dimensional Lebesgue measure.

Lemma 5.1. *For* $\{z_{P,Q,k} : P, Q \in C(n,u), 1 \leq k \leq a_{P,Q}\} \subset \mathbb{Z}^n$ *as in* §4, *we have*

$$\forall P, Q \in C(n,u) \;\; \nu_u(K_P) = \nu_u(\pi_U J^P), \;\; \nu_s(L_Q) = \nu_s(-\pi_S J^{Q^c}) = \nu_s(\pi_S J^{Q^c}).$$

Remark 2. In Figure 1 we see that

$$R_1 = [\pi_U J^{\{1\}}, \pi_S(-J^{\{1\}^c})], \;\; R_2 = [\pi_U J^{\{2\}}, \pi_S(-J^{\{2\}^c})],$$

and, for this example with $u = s = 1$,

$$(\pi_U J^{\{1\}}, \pi_U J^{\{2\}}) \text{ and } (\pi_S(-J^{\{1\}^c}), \pi_S(-J^{\{2\}^c}))$$

are themselves the fixed points, $(K_{\{1\}}, K_{\{2\}}), (L_{\{1\}}, L_{\{2\}})$, of E, F respectively. On the other hand, Bowen pointed out in [5] that, if $u = 2, s = 1$, then ∂K_P cannot be smooth.

Proof. In the Hausdorff metric, K_P is the limit (as $j \to \infty$) of the P-th coordinate in $E^j(\pi_U J)$, where $\pi_U J := (\pi_U J^P)_{P \in C(n,u)}$. Because the definition of E is modelled on that of Y, this is $\pi_U M^{-j} Y^j \rho_P(I^u)$. For any j we have, by Lemma 4.3,

$$\begin{aligned} \partial \pi_U M^{-j} Y^j \rho_P &= \pi_U M^{-j} Y^j \partial\rho_P \\ &= \sum_{k=1}^{u} (-1)^k (\pi_U M^{-j} Y^j \rho_{P \setminus p_k} - (\pi_U M^{-j} Y^j \rho_{P \setminus p_k} + \pi_U e_{p_k})), \end{aligned}$$

so that $\pi_U M^{-j} Y^j \rho_P$ has boundary zero when projected to the torus $U/(\pi_U \mathbb{Z}^P)$, as does $\pi_U \rho_P$. Both these cycles represent the fundamental homology class of this torus. Thus each maps I^u to U, so that its projection to $U/\pi_U \mathbb{Z}^P$ is surjective.

Thus, for each j, the support $K_P^j \subset U$, say, of $\pi_U M^{-j} Y^j \rho_P(I^u)$ (that is, the associated point-set [24, page 317]) has

$$\nu_u(K_P^j) \geq \nu_u(U/(\pi_U \mathbb{Z}^P)) = \nu_u(\pi_U(J^P)). \tag{5.1}$$

By subadditivity of ν_u on the expression for K_P^j as a union in (3.4), we have

$$\nu_u(K_P^j) \leq \mu^{-j} \sum_{Q \in C(n,u)} M_u^{(j)}(Q, P) \nu_u(\pi_U(J^Q)), \tag{5.2}$$

where the index (j) refers to an entry in the j-th power of the matrix M_u.

The Perron-Frobenius Theorem (see e.g. [28, page 4]) now implies that we have equality in (5.1) and (5.2), and that $(\nu_u(\pi_U(J^P)))_{P \in C(n,u)}$ is a left eigenvector for M_u or the zero vector. Thus $\nu_u(K_P^j) = \nu_u(\pi_U(J^P))$ for each j. Any neighbourhood of the Hausdorff limit K_P contains K_P^j for all sufficiently large j, so $\nu_u(K_P) \geq \nu_u(\pi_U(J^P))$.

To prove the opposite inequality we compare ∂K_P^j and ∂K_P^{j+1}. Now the set $M^{j+1}\partial K_P^{j+1}$ is the union of π_U of certain $(u-1)$-dimensional faces of those standard n-cubes ($I^n + z$, $z \in \mathbb{Z}^n$) that meet N of the support of $Y^j(\partial\rho_P)$, and these n-cubes number at most twice the sum of all the entries of M_{u-1}^{j+1} (in fact we only need twice the sum of the entries in those u columns of M_{u-1}^{j+1} that correspond to $(u-1)$-dimensional faces of J^P). Twice the sum of the entries of M_{u-1}^{j+1} is at most $c(\mu/\lambda_{s+1})^{j+1}$ for some constant c, so ∂K_P^{j+1} lies in a neighbourhood of ∂K_P^j of ν_u-volume $c\mu^{-(j+1)}(\mu/\lambda_{s+1})^{j+1}\nu_u(\pi_U(I^n))$.

For any $j' > j$, it follows that $K_P^{j'}$ (and hence K_P) lies in a neighbourhood of K_P^j of ν_u-volume $\nu_u(K_P^j) + \sum_{k=j+1}^{j'} c\lambda_{s+1}^{-k}\nu_u(\pi_U(I^n))$. Since j is arbitrary, we have $\nu_u(K_P) \le \nu_u(\pi_U(J^P))$, and they are equal. The case of ν_s for $(L_Q)_{Q\in C(n,u)} = \lim_{j\to\infty} F^j((\pi_S(-J^{Q^c}))_{Q\in C(n,u)})$ is similar. □

Lemma 5.2.

$$\sum_{P\in C(n,u)} \nu_n[\pi_U J^P, \pi_S J^{P^c}] = 1.$$

Proof. Let the column vectors $e_1, \dots, e_n$ denote the standard basis of $\mathbb{R}^n$. The multilinearity of the determinant function gives

$$\begin{aligned} 1 = \det(e_1, \dots, e_n) &= \det(\pi_S e_1 + \pi_U e_1, \dots, \pi_S e_n + \pi_U e_n) \\ &= \sum_{P\in C(n,u)} \det(\pi_{V(P,1)}e_1, \dots, \pi_{V(P,n)}e_n), \end{aligned}$$

where $V(P, j)$ denotes U if $j \in P$ and S if $j \notin P$. (Other terms in this sum vanish, because more than u vectors in U must be linearly dependent, as must more than s vectors in S.)

Now,

$$\det(\pi_{V(P,1)}e_1, \dots, \pi_{V(P,n)}e_n) = \pm\nu_n[\pi_U J^P, \pi_S J^{P^c}]. \tag{5.3}$$

It remains to prove that the sign in (5.3) is + for each P. This sign remains constant along a continuous path of subspaces $U_t, S_t, 0 \le t \le 1$, that satisfy $\mathbb{R}^n = U_t \oplus S_t\ \forall t \in [0, 1]$, unless, for some t, both sides of (5.3) become zero. And this happens precisely when $\dim(U_t \cap \mathbb{R}^{P^c}) > 0$ or $\dim(S_t \cap \mathbb{R}^P) > 0$.

Now

$$\dim(U \cap \mathbb{R}^{P^c}) = 0 \text{ and } \dim(S \cap \mathbb{R}^P) = 0$$

by (4.1).

So there are linear maps $\beta_U : \mathbb{R}^P \to \mathbb{R}^{P^c}$, $\beta_S : \mathbb{R}^{P^c} \to \mathbb{R}^P$ with graphs U, S. Let U_t, S_t be the graphs of the linear maps $t.\beta_U$, $t.\beta_S$.

Now $U_0 = \mathbb{R}^P$ and $S_0 = \mathbb{R}^{P^c}$, so, for $t = 0$, (5.3) holds with the + sign and both sides being +1. This completes the proof of the lemma. □

Corollary 5.3.

$$\sum_{P\in C(n,u)} \nu_n[K_P, L_P] = \sum_{P\in C(n,u)} \nu_n[\pi_U J^P, \pi_S J^{P^c}] = 1.$$

Proof. The ratio $\nu_n([B, C])/(\nu_u(B)\nu_s(C))$ is constant for $B \subset U, C \subset S$, so the corollary follows from Lemmas 5.1 and 5.2. □

Theorem 5.4. *$\mathcal{R} := \{\pi_{\mathbb{T}} R_P : P \in C(n,u)\}$ is a Markov partition for $\overline{M} : \mathbb{T}^n \to \mathbb{T}^n$ with transition matrix M_u^T.*

Proof. We shall prove that $\mathcal{R}$ has the Markov property and that the rectangles cover $\mathbb{T}^n$ with disjoint interiors. That each rectangle is the closure of its interior was shown at the end of §3.

In $\mathbb{R}^n$, $R_P = [K_P, L_P]$ is a union of s-subrectangles

$$[M^{-1}(K_Q) + \pi_U z_{PQi}, L_P] \text{ for } Q \in C(n,u),\ 1 \le i \le M_u(Q,P) = a_{PQ},$$

while $R_Q = [K_Q, L_Q]$ is a union of u-subrectangles

$$[K_Q, ML_P - \pi_S M z_{PQi}] \text{ for } P \in C(n,u),\ 1 \le i \le M_u(Q,P).$$

Moreover,

$$M[M^{-1}(K_Q) + \pi_U z_{PQi}, L_P] = [K_Q, ML_P - \pi_S M z_{PQi}] + M z_{PQi},$$

so that, in $\mathbb{T}^n$, $\overline{M}$ takes each s-subrectangle to the corresponding u-subrectangle. Thus $\mathcal{R}$ satisfies the Markov property with transition matrix $A = M_u^T$.

Hence the union of the rectangles in $\mathcal{R}$ is a closed $\overline{M}$-invariant subset of $\mathbb{T}^n$. Since $\overline{M}$ is ergodic, this union has n-dimensional Lebesgue measure 1 (and not 0, because some rectangle has positive measure).

Since the union is closed, it is $\mathbb{T}^n$. By Corollary 5.3, each pair of these rectangles in $\mathbb{T}^n$ has intersection a set of measure 0. In particular, their interiors are pairwise disjoint. □

References

[1] R Adler and B Weiss, Similarity of automorphisms of the torus, *Memoirs Amer. Math. Soc.*, **98** (1970). MR **41:**1966

[2] T Bedford, Generating special Markov partitions for hyperbolic toral automorphisms using fractals, *Ergodic Theory Dynam. Systems* **6** (1986), 325–333. MR **88c:**58037

[3] R Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics 470, Springer, New York 1975. MR **56:**1364

[4] R Bowen, *On axiom A diffeomorphisms*, CBMS Regional Conference Series in Mathematics 35, Amer. Math. Soc. 1978. MR **58:**2888

[5] R Bowen, Markov partitions are not smooth, *Proc. Amer. Math. Soc.* **71** (1978), 130–132. MR **57:**14055

[6] R Devaney, *An introduction to chaotic dynamical systems*, Benjamin, Reading, MA, 1986. MR **87e:**58142

[7] K Falconer, *Fractal geometry: mathematical foundations and applications*, Wiley, Chichester 1990. MR **92j:**28008

[8] J Franks, *Homology and dynamical systems*, CBMS Regional Conference Series in Mathematics 49, Amer. Math. Soc. 1982. MR **84f:**58067

[9] D. Garling. *A course in Galois theory*, Cambridge University Press, Cambridge 1986. MR **88d:**12007

[10] P Hilton and S Wylie, *Homology Theory: An introduction to algebraic topology*, Cambridge University Press, Cambridge 1960. MR **22:**5963

[11] M W Hirsch, Expanding maps and transformation groups, *Proc. Symp. Pure Math.* **14** (1970), 125–131. MR **45:**7750

[12] J E Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.* **30** (1981), 713–747. MR **82h:**49026

[13] S Ito and M Ohtsuki, Modified Jacobi-Perron algorithm and generating Markov partitions for special hyperbolic toral automorphisms, *Tokyo J. Math.* **16** (1993), 441–472. MR **95g:**58167

[14] A Katok and B Hasselblatt, *Introduction to the modern theory of dynamical systems*, Cambridge University Press, Cambridge 1995. MR **96c:**58055

[15] R Kenyon, Self-similar tilings, Thesis, Princeton, 1990.

[16] R Kenyon, Self-replicating tilings, *Contemp. Math.* **135** (1992), 239–263. MR **94a:**52043

[17] R Kenyon and A Vershik, Arithmetic construction of sofic partitions of hyperbolic toral automorphisms, *Ergodic Theory Dynam. Systems* **18** (1998), 357–372. MR **99g:**58092
[18] J C Lagarias and Y Wang, Self-affine tiles in R^n, *Adv. Math.* **121** (1996), 21–49. MR **97d:**52034
[19] S Le Borgne, Un codage sofique des automorphismes hyperboliques du tore, *C. R. Acad. Sci. Paris Sér. I Math.* **323** (1996), 1123–1128. MR **97k:**58095
[20] R Mañé, *Ergodic theory and differentiable dynamics*, Springer, Berlin 1987. MR **88c:**58040
[21] A Manning, Axiom A diffeomorphisms have rational zeta functions, *Bull. London Math. Soc.* **3** (1971), 215–220. MR **44:**5982
[22] A Manning, There are no new Anosov diffeomorphisms on tori, *Amer. J. Math.*, **96** (1974), 422–429. MR **50:**11324
[23] A Manning, Irrationality of linear combinations of eigenvectors, *Proc. Indian Acad. Sci. Math. Sci.*, **105** (1995), 269–271. MR **96h:**15009
[24] W Massey, *A Basic Course in Algebraic Topology*, Springer, New York 1991. MR **92c:**55001
[25] R D Mauldin and S C Williams, Hausdorff dimension in graph directed constructions, *Trans. Amer. Math. Soc.*, **309** (1988), 811–829. MR **89i:**28003
[26] B Praggastis, Numeration systems and Markov partitions from self-similar tilings, *Trans. Amer. Math. Soc.* **351** (1999), 3315–3349. MR **99m:**11009
[27] D Ruelle, One-dimensional Gibbs states and axiom A diffeomorphisms, *J. Differential Geom.* **25** (1987), 117–137. MR **88j:**58099
[28] E Seneta, *Nonnegative matrices and Markov chains*, 2nd ed., Springer, New York 1981. MR **85i:**60058
[29] Ya G Sinai, Markov partitions and У-diffeomorphisms, *Func. Anal. Appl.* **2** (1968), 61–82. MR **38:**1361
[30] S. Smale, Differentiable dynamical systems, *Bull. Amer. Math. Soc.* **73** (1967), 747–817. MR **37:**3598
[31] P. Walters, *An introduction to ergodic theory*, Springer, New York-Berlin 1982. MR **84e:**28017
[32] R F Williams, Classification of subshifts of finite type, *Ann. of Math.* **98** (1973), 120–153. Errata, ibid. **99** (1974), 380–381. MR **48:**9769

Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK
E-mail address: `akm@maths.warwick.ac.uk`