Ergod. Th. & Dynam. Sys. (1998), **18**, 357–372
Printed in the United Kingdom © *1998 Cambridge University Press*

Arithmetic construction of sofic partitions of hyperbolic toral automorphisms

RICHARD KENYON†§ and ANATOLY VERSHIK‡¶

† *CNRS UMR 128, Ecole Normale Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon, France*

‡ *St. Petersburg Branch of Mathematical Institute of the Russian Acad. Sciences, 27 Fontanka, 191011 St. Petersburg, Russia*

(*Received* 11 *December* 1995 *and accepted in revised form* 20 *December* 1996)

Abstract. For each irreducible hyperbolic automorphism A of the n-torus we construct a sofic system (Σ, σ) and a bounded-to-one continuous semiconjugacy from (Σ, σ) to $(\mathbb{T}^n, A)$. This construction is natural in the sense that it depends only on the characteristic polynomial of A and, furthermore, it has an arithmetic interpretation.

1. *Introduction*

Starting from [**1, 4, 12**], many papers have been devoted to Markov partitions for hyperbolic automorphisms of the torus, as well as Markov partitions for arbitrary hyperbolic systems. If we want to give a symbolic realization of such a toral automorphism A, the following question naturally arises: is it possible to interpret other structures which exist on the torus (foliations, homoclinic points, algebraic and arithmetic structures) in terms of the symbolic dynamics? The constructions of [**4, 12**] are very general (they work for arbitrary Anosov systems) and do not respond to this question.

In [**14**] an attempt was made to give a construction, at least for some automorphisms A, which had an arithmetic meaning. In that construction, the symbolic representation was given by a two-sided radix representation (i.e. expansion in terms of a base and digits) of points on the torus. The construction gave symbolically a so-called *sofic system* (a quotient of a shift of finite type) describing the allowed expansions.

In the case where A was a Pisot automorphism (a hyperbolic automorphism with one expanding eigenvalue λ) the corresponding one-sided radix representation coincided with the β-expansion for the value $\beta = \lambda$ (here we mean β-expansion in the sense of Renyi and Parry [**7, 10, 11**]). Two-sided expansions for the Pisot case were also considered

§ Research at MSRI supported in part by NSF grant no. DMS-9022140. Current address: Laboratoire de Topologie, Bât 425, Université Paris-Sud, 91405 Orsay, France.
¶ Partially supported by the Russian Fund for Fundamental Research grant no. 94-01-00921 and the Miller Institute, University of California, Berkeley.

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

in [**3, 8**]. Expansions for general A have not been considered in detail before (although see [**9**]): the ideas of [**14**] were not fully developed in this case.

In this paper we suggest, using an approach similar to [**14**], a new arithmetic construction which works for all automorphisms. It gives a boundedly finite-to-one map from a sofic system to a hyperbolic system on the torus. This construction, motivated by an idea of Thurston (see Theorem 3, below), preserves the properties mentioned earlier: it gives a two-sided expansion of points on the torus, which converge in the ordinary sense in one direction and in the A-adic sense in the other direction. The homoclinic points, foliations, and so on, all have a precise meaning in this expansion.

The digits in our radix representations take values in the algebraic integers. We do not know if it is possible to restrict ourselves to digits in the rational integers; however we remark that this restriction seems less natural in regards to our construction (also, Hirsch [**5**] noted that it is not always possible to use non-negative integer digits). Note that, as in [**14**], even for the two-torus we obtain a different kind of isomorphism than that in [**1, 4, 12**].

The main results are formulated in §2. We consider an example from [**1**] in our point of view, and several other examples, in §3. §4 is devoted to the proofs.

Let us emphasize that this problem is closely related to the theory of self-similar tilings (see [**6, 8, 9, 13**]), adic transformations (see [**14**]), and to the general theory of radix representations.

In a forthcoming paper we extend our construction to commuting toral automorphisms. Specifically, we will show that if (Σ, T) is a shift of finite type (respectively, sofic system) and S is a homeomorphism of Σ commuting with T, then the $\mathbb{Z}^2$-action (Σ, T, S) is a two-dimensional shift of finite type (respectively sofic system). This fact, along with an additional number-theoretic lemma, gives us a simple coding for non-hyperbolic ergodic automorphisms of the torus.

2. *Background and results*

Let $A' \in GL(n, \mathbb{Z})$ be an irreducible hyperbolic automorphism of the n-torus. That is, the characteristic polynomial of A' is irreducible and has no roots of modulus 1. Let $\{\lambda_1, \ldots, \lambda_k, \lambda_{k+1}, \ldots, \lambda_n\}$ be the eigenvalues of A', with

$$|\lambda_i| > 1 \quad \text{for } i \in [1, k] \tag{1}$$

$$|\lambda_i| < 1 \quad \text{for } i \in [k+1, n]. \tag{2}$$

By irreducibility of A', the λ_i are distinct. Let $p(x)$ be the characteristic polynomial of A', let A be the companion matrix of A', and let $K = \mathbb{Q}[x]/(p(x))$. Let ρ_i for $i = 1, \ldots, n$ be the distinct embeddings of K into $\mathbb{C}$. Specifically, for a polynomial $z(x) \in K$ we have $\rho_i(z(x)) = z(\lambda_i)$. When λ_i is real the image of ρ_i is in $\mathbb{R}$; for each complex-conjugate pair of roots, the corresponding embeddings are also complex conjugate embeddings.

Let r be the number of real roots of $p(x)$ and let $2c$ be the number of non-real roots. The product of all the real embeddings ρ_i and half of the complex embeddings (taking one of each pair of complex-conjugate embeddings) is an embedding of K into $\mathbb{R}^r \times \mathbb{C}^c \cong \mathbb{R}^n$. Denote this embedding by ρ.

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

The matrix A acts as a linear map on $\mathbb{R}^n$ with basis $\{\rho(1), \rho(x), \ldots, \rho(x^{n-1})\}$; it satisfies $A\rho(z) = \rho(xz)$ for each $z \in K$. Hence multiplication by A will be denoted using multiplication by x.

In what follows we will coordinatize $\mathbb{R}^n$ using the $r+c$ coordinates given by the above $r+c$ embeddings, r of which are real, c of which are complex. Furthermore, we order these coordinates as in (1) and (2), so that $W_u = \mathbb{R}^k \times \{0\} \subset \mathbb{R}^n$ and $W_s = \{0\} \times \mathbb{R}^{n-k} \subset \mathbb{R}^n$ are the unstable and stable eigenspaces of A, respectively; that is, the eigenvalues of A on W_u have modulus > 1, and those on W_s have modulus < 1.

For simplicity of notation we will usually identify K and its image $\rho(K)$, so that we can always work in the space $\mathbb{R}^n$.

Let Λ be the lattice $\mathbb{Z}[x]/(p(x)) \subset K$. The image $\rho(\Lambda)$ (which we denote again by Λ) is a discrete full-rank lattice in $\mathbb{R}^n$.

Let $|\cdot|$ denote the sup-norm on $\mathbb{R}^r \times \mathbb{C}^c$, that is, $|x| = \max_i\{|x_i|\}$, where the i ranges over the r real coordinates and the c complex coordinates. This norm restricts to a norm on K: for $z \in K$, $|z| = \max_i\{|\rho_i(z)|\}$.

Let

$$B = B_r(0) = \{z \in \mathbb{R}^n : |z| \le r\},$$

where r is the infimum of real numbers t such that any translate of $B_t(0)$ contains a point of Λ. Let $D = \Lambda \cap (B + xB)$, the set of lattice points in the symmetric product neighborhood $B + xB$ of the origin. The set $D \subset \Lambda$ will serve as the set of *digits* for our radix expansions.

LEMMA 1. *D is a finite set. For any point $z \in B$, there exists $d \in D$ such that $xz - d \in B$.*

(See the last section for the proofs of all results.)

For $i \in [1, k]$ define the *base λ_i representation map* S_i as follows. For $d = (d_0, d_1, \ldots) \in D^{\mathbb{N}}$ define

$$S_i(d) = \sum_{j=0}^{\infty} \rho_i(d_j)\lambda_i^{-j}. \tag{3}$$

If $\lambda_i \in \mathbb{R}$ then $S_i : D^{\mathbb{N}} \to \mathbb{R}$; otherwise $S_i : D^{\mathbb{N}} \to \mathbb{C}$. Define $R_u : D^{\mathbb{N}} \to W_u$ by $R_u(d) = (S_1(d), \ldots, S_k(d))$, where as before for each pair of complex conjugate embeddings we only take one of the two.

Similarly, for $i \in [k+1, n]$ define $S_i' : D^{\mathbb{N}} \to (\mathbb{C}$ or $\mathbb{R})$ by

$$S_i'((d_0, d_1, \ldots)) = \sum_{j=0}^{\infty} \rho_i(d_j)\lambda_i^{j+1}.$$

Define $R_s(d) = (S_{k+1}(d), \ldots, S_n(d))$, where again we only take one of each complex conjugate pair. The maps R_u and R_s are continuous maps from $D^{\mathbb{N}}$ (with the product topology) to W_u and W_s, respectively.

LEMMA 2. *The set $R_u(D^{\mathbb{N}})$ contains a neighborhood of the origin in W_u.*

Let $\prec$ be an arbitrary total order on D, and let $\mathbf{a}$ denote the minimal element. Extend $\prec$ to lexicographic order on D^N for each $N \in \mathbb{N}$. If $(y_1, \ldots, y_N) \in D^N$ is a finite

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

sequence, we say it is *non-minimal* if there is another sequence $(y'_1, \ldots, y'_N) \in D^N$ which is lower in lexicographic order and for which

$$\sum_{j=1}^{N} y_j x^{N-j} = \sum_{j=1}^{N} y'_j x^{N-j},$$

the equality holding in K. If a finite sequence is not non-minimal we say it is *minimal*.

Let σ denote the left shift on $D^{\mathbb{N}}$ and let $(\Sigma_u, \sigma) \subset (D^{\mathbb{N}}, \sigma)$ be the closed shift-invariant subset consisting of those sequences whose finite (consecutive) subsequences are all minimal. Let $(\Sigma, \sigma) \subset (D^{\mathbb{Z}}, \sigma)$ be the natural extension of (Σ_u, σ) to bi-infinite sequences; that is, Σ consists of all bi-infinite sequences, all of whose finite consecutive subsequences are minimal.

Note that Σ_u is embedded naturally in Σ by sending $(d_0, d_1, \ldots)$ to the bi-infinite sequence $(\ldots, \mathbf{a}, \mathbf{a}, \mathbf{a}, d_0, d_1, \ldots)$: since if $(d_0, \ldots, d_k)$ is minimal then so is $(\mathbf{a}, d_0 \ldots, d_k)$.

In general, a shift space (X, σ) on an alphabet D is called a *sofic system* if there is a finite graph G with edges labelled by elements of D such that X is the set of labellings of bi-infinite paths in G. This definition is due to Weiss [**16**], who proved the main property that sofic systems comprise of the smallest class of shift spaces which is closed under factors and contains the shifts of finite type (a *factor* is the image under a continuous surjective shift-commuting map).

THEOREM 3. *The shift space (Σ, σ) is a sofic system.*

Remark. Thurston (unpublished) proved this theorem in the special case when A has either only one expanding eigenvalue or only two (complex conjugate) expanding eigenvalues. In the case of a Pisot automorphism A (i.e. one expanding eigenvalue) this was also proved in different terms in [**3**] (based on results of [**11**]) and in [**8, 14**]. Our proof is a generalization of Thurston's method.

There is a natural map $R : \Sigma \to \mathbb{R}^n = W_u \oplus W_s$: defined by

$$R(\ldots, d_{-1}, d_0, d_1, \ldots) = (R_u(d_0, d_1, d_2, \ldots), -R_s(d_{-1}, d_{-2}, \ldots)).$$

As R is continuous and Σ is compact, the image $R(\Sigma)$ is compact. Furthermore, we claim that for $d = (d_j)_{j \in \mathbb{Z}}$, we have $AR(d) = R(\sigma(d)) + \rho(d_0)$.

To see this, consider the ith coordinate. If $i \in [1, k]$, then

$$\begin{aligned}
(AR(d))_i &= \lambda_i(S_i(d_0, d_1, d_2, \ldots)) \\
&= \lambda_i \sum_{j=0}^{\infty} \rho_i(d_j)\lambda_i^{-j} \\
&= \rho_i(d_0) + \sum_{j=0}^{\infty} \rho_i(d_{j+1})\lambda_i^{-j} \\
&= S_i(\sigma(d)) + \rho_i(d_0).
\end{aligned}$$

Similarly, for $i \in [k+1, n]$,

$$(AR(d))_i = -\lambda_i(S_i(d_{-1}, d_{-2}, \ldots))$$

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

$$\begin{aligned} &= -\lambda_i \sum_{j=1}^{\infty} \rho_i(d_{-j})\lambda_i^j \\ &= \rho_i(d_0) - \sum_{j=1}^{\infty} \rho_i(d_{-j+1})\lambda_i^j \\ &= -S_i(\sigma(d)) + \rho_i(d_0). \end{aligned}$$

This completes the claim, since it is true for each coordinate.

Let π be the projection to the torus $\pi : \mathbb{R}^n \to \mathbb{R}^n/\Lambda = \mathbb{T}^n$. Then $\tilde{A} = \pi A \pi^{-1}$ is the associated toral automorphism (which we have, by abuse of notation, been denoting A up until now). Since $d_0 \in D \subset \Lambda$, we have

$$\tilde{A}(\pi R(d)) = \pi A \pi^{-1}(\pi R(d)) = \pi A R(d) = \pi R(\sigma(d)).$$

That is, πR is a semiconjugacy between (Σ, σ) and its image in $(\mathbb{T}^n, \tilde{A})$.

It remains to show that the image $\pi R(\Sigma)$ is the whole torus. By Lemma 2, the image $\pi R(\Sigma_u) \subset \pi R(\Sigma)$ contains an open neighborhood in the unstable leaf $\pi(W_u)$ around the point $\pi R(\ldots, \mathbf{a}, \mathbf{a}, 0, 0, 0, \ldots)) = \pi(0, -R_s((\mathbf{a}, \mathbf{a}, \ldots)))$. By invariance the image $\pi R(\Sigma)$ must contain arbitrarily large balls in the unstable leaf and so must be dense in the torus. Since the image is also closed, it must be the whole torus.

THEOREM 4. *The map πR is a factor map from (Σ, σ) to $(\mathbb{T}^n, \tilde{A})$. It is bounded-to-one everywhere and constant-to-one almost everywhere.*

As noted above, a (continuous) quotient of a sofic system is again sofic, so the theorem proves the existence of a metric isomorphism of $(\mathbb{T}^n, A)$ with a sofic system. In fact since πR is continuous this can be arranged to give an almost homeomorphism (i.e. a metric isomorphism which is in addition continuous and almost everywhere injective) from a sofic system to the torus (also see Remark 4 below.)

What may we now conclude about the original automorphism A'? The matrices A and A' are conjugate over $GL(n, \mathbb{Q})$ but not necessarily over $GL(n, \mathbb{Z})$. Let $M \in GL(n, \mathbb{Q})$ be such that $A' = MAM^{-1}$. We can assume that $M\Lambda \subset \Lambda$ by multiplying by an integer if necessary, so that $M : \mathbb{T}^n \to \mathbb{T}^n$ is a cover of the torus. Hence M is a semiconjugacy from $(\mathbb{T}^n, \tilde{A})$ to $(\mathbb{T}^n, A')$.

The map $M\pi R : \Sigma \to \mathbb{T}^n$ gives a bounded-to-one semiconjugacy between (Σ, σ) and $(\mathbb{T}^n, A')$: we have the diagram

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma} & \Sigma \\ {\scriptstyle \pi R}\downarrow & & \downarrow{\scriptstyle \pi R} \\ \mathbb{T}^n & \xrightarrow{A} & \mathbb{T}^n \\ {\scriptstyle M}\downarrow & & \downarrow{\scriptstyle M} \\ \mathbb{T}^n & \xrightarrow{A'} & \mathbb{T}^n \end{array}$$

which commutes. By the same reasoning as above, we can construct a quotient Σ' which is almost homeomorphic to $(\mathbb{T}^n, A')$.

Remark 1. Our construction depends strongly on our choice of digits D and choice of ordering on D. Many other choices of D will work, and in some cases a good choice

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

of D greatly simplifies the resulting sofic system (see the examples below). It would be nice if we could always take D to consist of rational integers, but we do not know if this is possible for all A. Hirsch [**5**] proved that it is not always possible to take $D \subset \mathbb{N}$.

Remark 2. If 0 is chosen to be the minimal element of D, then $\Sigma_u \subset \Sigma$ since any element of Σ_u can be preceded by an infinite string of zeros. Let Σ^+ be the set of sequences $d \in \Sigma$ which are eventually zero to the left ($d_{-n} = 0$ for all n sufficiently large). From Lemma 2 it follows that $\pi R(\Sigma^+) = \pi W_u$. No similar statement is true for W_s, since Σ may contain no sequences which are eventually zero to the right. However, the proof of Theorem 4 indicates how to write down from the graph G a sofic subshift of Σ consisting of the points mapping to πW_s. So in this sense we have a symbolic description of the homoclinic points on $(\mathbb{T}^n, A)$.

Remark 3. It seems difficult to understand, in general, when πR is almost-everywhere injective. For a given pair D, A, however, as the proof of Theorem 4 indicates, one can tell from the graph G if at least R is injective almost everywhere: for almost every path in G there is no path which 'follows' it.

Remark 4. It is easy to construct geometrically a genuine Markov partition for $(\mathbb{T}^n, A')$ from Σ as follows. A standard argument (see e.g. [**8, 9**]) shows that the boundaries of the cylinder sets U_v have measure zero. (The cylinder set U_v is the image under $M\pi R$ of the set of sequences in Σ passing through a given vertex v of G at the zeroth coordinate.) Take the finest partition of the torus induced by the covering by cylinder sets; that is, the elements of the partition are intersections of cylinder sets which do not contain any other intersections of cylinder sets. The dynamics on this partition is easily defined from the dynamics of Σ.

We do not know a good way to describe this construction arithmetically.

Lemmas 1 and 2 can be reformulated as a statement about the representation of elements of $\mathbb{Z}[x]/(p(x))$ as polynomials, in the same spirit as [**15**]. Specifically, we have the following.

PROPOSITION 5. *Let $p \in \mathbb{Z}[x]$ be irreducible, monic, hyperbolic (no roots on the unit circle), and suppose $|p(0)| = 1$. Let $\Lambda = \mathbb{Z}[x]/(p(x))$. There is a finite set $D' \subset \Lambda$ such that every element of Λ can be represented as a polynomial in x, x^{-1} with coefficients in D'.*

It is now easy to choose a unique such representation for each element of Λ in such a way that the resulting set of polynomials is invariant under the shift (multiplication by x). We conjecture that it is possible to do this in such a way that the resulting sequences are 'sofic', that is, form a regular language.

3. *Examples*

Example 1. Let us start with the following well-known example. Let $p(x) = x^n - a_{n-1}x^{n-1} - \cdots - a_0$, where $a_{n-1} \geq a_{n-2} \geq \cdots \geq a_0 = 1$. As noted above, the sofic system is often simplified with a better choice of digits than that used in Lemma 1. In this case

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

if we choose $D = \{0, 1, 2, \ldots, a_{n-1}\}$ with the reverse natural order ($0 \succ 1 \succ \cdots \succ a_{n-1}$) then our method gives the same coding as in [**8, 14**], namely the sofic system is an n-step Markov system. This sofic system corresponds to the β-expansion in base x, where x is the positive real root of the above polynomial, which is a Pisot number (see [**11**]). To see this connection note that with the above order on D the lowest lexicographic expansions are the 'greedy' expansions: one takes as the kth digit of the expansion of a positive number z the *largest* digit possible which leaves the remainder less than z. This is exactly the definition of the β-expansion.

These examples all behave like the Fibonacci case $p(x) = x^2 - x - 1$, which we now work out in more detail. (This case was also considered in detail in [**15**].) We take $D = \{0, 1\}$ with $1 \prec 0$. In this case we get the graph G (from the proof of Theorem 3) in Figure 1(a). Note that, for example, the sequence $0, 1, 1$ leads from $\emptyset$ to the fail state F, which means that the sequence $0, 1, 1$ is non-minimal. Indeed, the sequence $1, 0, 0$ is lower in lexicographic order and $0 \cdot x^2 + 1 \cdot x + 1 = 1 \cdot x^2 + 0 \cdot x + 0$. Note that we get the same bi-infinite labellings if we identify states $\{-1\}$ and $\{-1, -x\}$, as in Figure 1(b). The classical Fibonacci coding is the one in which the vertex $\emptyset$ of Figure 1(b) does not appear. Our coding is equivalent to this since almost every path in G does not pass through $\emptyset$: the set of sequences which are eventually 1 to the left has image under πR which is a countable union of one-dimensional subsets of $\mathbb{T}^2$, and hence measure 0.

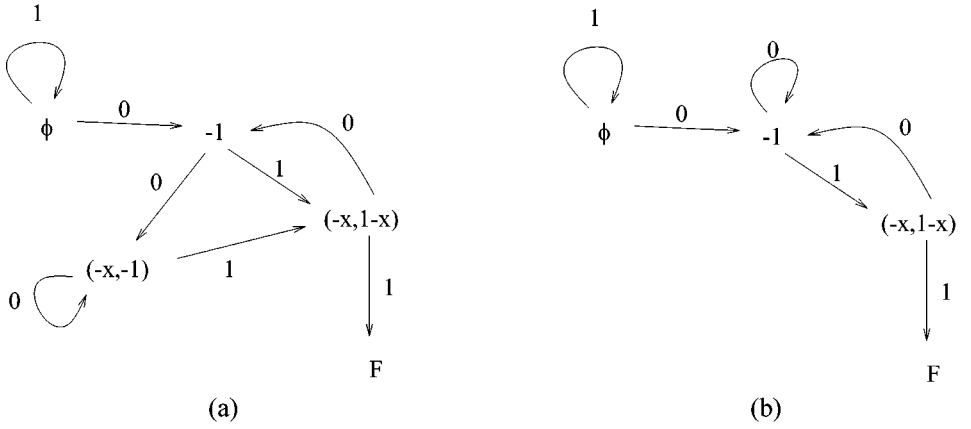


FIGURE 1. The sofic system for $p(x) = x^2 - x - 1$.

Example 2. Let us consider in detail the example of the automorphism $A = \left(\begin{smallmatrix} k & 1 \\ -1 & 0 \end{smallmatrix}\right)$ with $k \geq 3$. If we choose $D = \{0, 1, \ldots, k-1\}$ with the reverse natural order, this example again corresponds to the β-expansion of the Pisot number $x = (k + \sqrt{k^2 - 4})/2$, and was studied in [**3, 11**].

We now choose the *natural* order $0 \prec 1 \prec \cdots \prec k - 1$ on D. For simplicity, let us take the case $k = 6$ (we will need this example below anyway); the other cases are essentially identical.

For this order on D we get the graph G shown in Figure 2.

For the purposes of computing the sofic system (Σ, σ) in this example, we can further

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

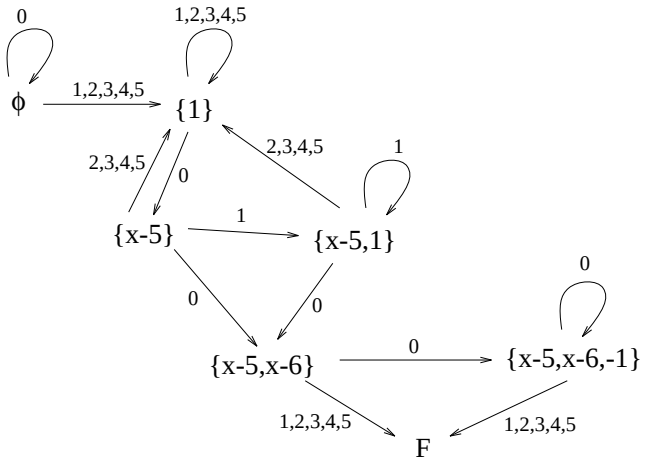


FIGURE 2. The sofic system for the case $k = 6$.

simplify the graph by removing the state $\emptyset$: bi-infinite sequences which pass through this state are eventually zero to the left and so all map under πR into the one-dimensional subset πW_u. Similarly, we can identify the states $\{x-5, x-6, -1\}$ and $\{x-5, x-6\}$ with the fail state, since this would remove only bi-infinite sequences which are eventually 0 to the right. We can further identify states $\{x-5\}$ and $\{x-5, 1\}$ without changing the allowed edge-labelling sequences, leaving the graph of Figure 3.

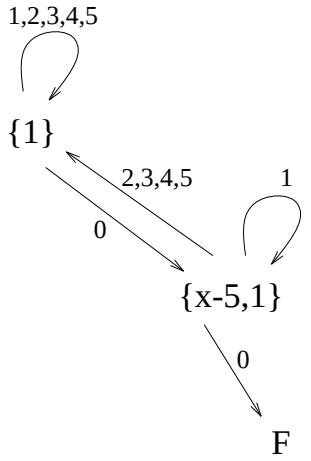


FIGURE 3. The reduced sofic system.

Note that this sofic system is not a shift of finite type: for example, the word $0(1)^m0$ for $m > 0$ is not allowed although all proper subwords of this word are allowed.

The image under R of Σ is shown in Figure 4. In this case R is almost everywhere injective. Furthermore, the image of the set of sequences which are in state $\{1\}$ at index 0

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

is a genuine rectangle. Similarly, we get a genuine rectangle for those sequences in state $\{x-5, 1\}$ at index 0.

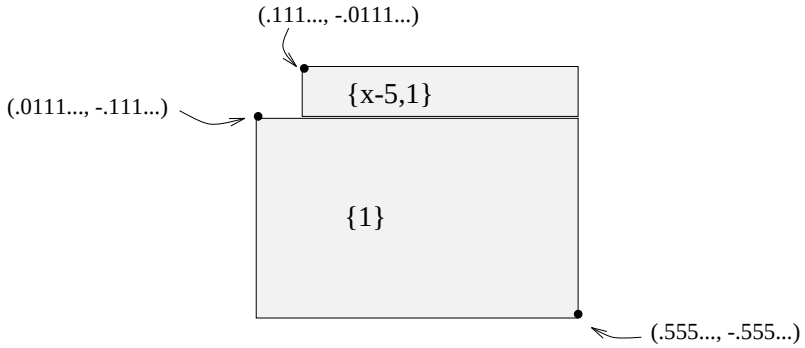


FIGURE 4. The image $R(\Sigma)$ consists of two rectangles. The two rectangles are labelled by the state in G of their zeroth digit d_0.

Example 3. Let us consider the case of two hyperbolic automorphisms which are not continuously conjugate but have the same characteristic polynomial (see [**1**]). Let $A_1 = \left(\begin{smallmatrix} 5 & 4 \\ 1 & 1 \end{smallmatrix}\right)$, and $A_2 = \left(\begin{smallmatrix} 5 & 2 \\ 2 & 1 \end{smallmatrix}\right)$. These two matrices have characteristic polynomial $p(x) = x^2 - 6x + 1$, but are not conjugate over $\mathbb{Z}$: A_2 does not have a single orbit which generates all of $\mathbb{Z}^2$ (although A_1 does). A theorem of Arov [**2**] states that if two ergodic algebraic automorphisms of $\mathbb{T}^n$ are topologically conjugate then they are algebraically conjugate. So $(\mathbb{T}^2, A_1)$ and $(\mathbb{T}^2, A_2)$ are not topologically conjugate, although they are *metrically* conjugate.

Note that A_1 is similar over $\mathbb{Z}$ to the companion matrix $A = \left(\begin{smallmatrix} 6 & 1 \\ -1 & 0 \end{smallmatrix}\right)$; with $M = \left(\begin{smallmatrix} 5 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, we have $A_1 = MAM^{-1}$. We worked out the sofic system for A above using digits $\{0, 1, 2, 3, 4, 5\}$. The fundamental domain for the basis $\{(0, 1), (1, -5)\}$ is shaded in Figure 5(a). (In this figure, the horizontal and vertical directions are the eigendirections.)

The matrix A_2 is not similar over $\mathbb{Z}$ to A. However, we can take $M = \left(\begin{smallmatrix} 5 & 1 \\ 2 & 0 \end{smallmatrix}\right)$, which has determinant 2. The basis $\{(0, 1), (1/2, -5/2)\}$ for A has fundamental domain shaded in Figure 5(b). In this case the map $M\pi R$ is two-to-one almost everywhere.

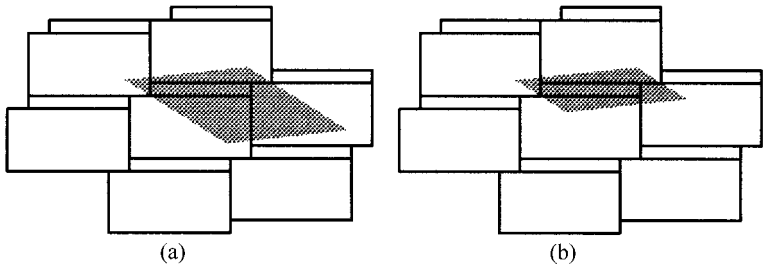


FIGURE 5. The induced partitions for (a) A_1 and (b) A_2.

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

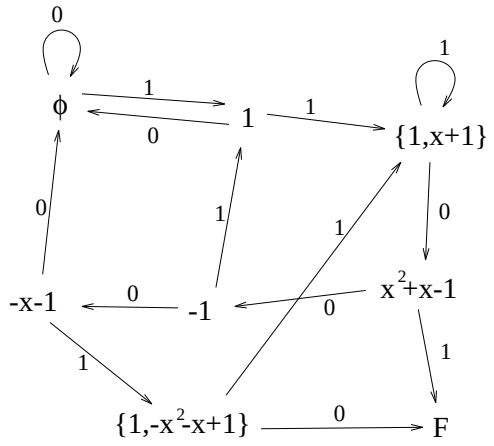


FIGURE 6.

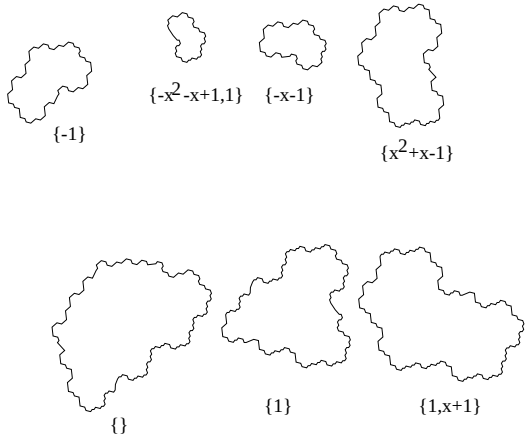


FIGURE 7.

Example 4. Let A be the companion matrix of $p(x) = x^3 + x^2 - x + 1$. In this case $W_u = \mathbb{R}$ and $W_s = \mathbb{C}$. Using $D = \{0, 1\}$ with $0 \prec 1$ we arrive at the graph in Figure 6.

Here, R is again injective. The image of the cylinder sets of Σ are products of intervals in W_u with closed topological disks in W_s. In Figure 7 we show the images of the seven cylinder sets

$$U_\emptyset, U_1, U_{\{1,x+1\}}, U_{x^2+x-1}, U_{-1}, U_{-x-1}, U_{\{-x^2-x+1,1\}}$$

in W_s corresponding to the seven states of G. Figure 8 shows how the image of $A^{-4}(U_0)$ intersects each of the cylinders corresponding to the seven states. In particular, there is one tile in the figure for each backwards path in G of length 4 starting at $\emptyset$.

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

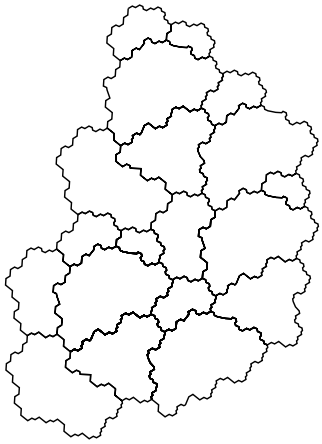


FIGURE 8.

4. *Proofs*

Proof of Lemma 1. Let $z \in B$. Let $d \in \Lambda$ be a point closest to xz (that is, $|xz - d|$ is minimal). Then since $xz + B$ contains a point of Λ by the definition of B, we have $d \in xz + B$. So $xz - d \in B$ since B is symmetric. Since in addition $xz \in xB$, this implies $d \in (B + xB) \cap \Lambda = D$. □

Proof of Lemma 2. We show that the projection (along W_s) of B to W_u is in $R_u(D^{\mathbb{N}})$.

Let $z_1 \in B$. Define $d_1, d_2, \ldots \in D$ such that: $z_2 \stackrel{\text{def}}{=} z_1 x - d_1 \in B$, $z_3 \stackrel{\text{def}}{=} xz_2 - d_2 \in B$, and so on. The d_i exist by Lemma 1. Furthermore, we have

$$z_k = -d_{k-1} - xd_{k-2} - \cdots - x^{k-2}d_1 + x^{k-1}z_1,$$

and so

$$z_1 = \frac{d_1}{x} + \frac{d_2}{x^2} + \cdots + \frac{d_{k-1}}{x^{k-1}} + \frac{z_k}{x^{k-1}}.$$

For each $i \in [1, k]$ we have

$$\rho_i(z_1) = \sum_{j=1}^{k-1} \rho_i(d_j)\lambda_i^{-j} + \rho_i(z_k)\lambda_i^{-(k-1)}$$

(here, by $\rho_i(z_1)$ and $\rho_i(z_k)$ we mean the ith coordinates). Since $\rho_i(z_k)$ is bounded ($z_k \in B$) and $|\lambda_i| > 1$ this series converges as $k \to \infty$. In the limit we get an expansion for the projection of z_1 to W_u. □

Proof of Theorem 3. Let δ_i be the maximum modulus of any element of $\rho_i(D)$. Let $V \subset \Lambda$ be the finite set

$$V = \left\{ z \in \Lambda \mid \forall i \in [1, n], |\rho_i(z)| \leq \left| \frac{2\delta_i}{|\lambda_i| - 1} \right| \right\}.$$

A sketch of the construction is the following. We read a sequence $d \in D^{\mathbb{N}}$ one digit at a time, and remember as we go along all sequences which are lower in lexicographic

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

order than d. If, after j digits have been read, one of our lower sequences $(d'_1, \dots, d'_j)$ has *the same* R_u-value as $(d_1, \dots, d_j)$, then $(d_1, \dots, d_j)$ is non-minimal and so must be rejected. Now the first main idea is that we do not have to remember all the digits of all sequences lower than d: it suffices to remember the *differences*

$$R_u((d_1, \dots, d_j)) - R_u((d'_1, \dots, d'_j)) = \rho_u\left(\sum_{i=1}^{j}(d_i - d'_i)x^{-i}\right) \tag{4}$$

since we are only interested if this difference is ever 0. Since ρ_u is injective on K it suffices to remember $\sum_{i=1}^{j}(d_i - d'_i)x^{-i}$.

The second main idea is that if this difference is ever too big, in the sense that

$$\sum_{i=1}^{j}(d_i - d'_i)x^{-i} \notin x^{-j}V,$$

then there is no way to extend the sequences $(d'_1, \dots, d'_j)$ and $(d_1, \dots, d_j)$ to longer sequences which have the same R_u-value. So in this case we can remove $(d'_1, \dots, d'_j)$ (or rather, the difference (4)) from consideration.

So we need only keep track of all those possible differences

$$x^j \sum_{i=1}^{j}(d_i - d'_i)x^{-i} \tag{5}$$

which are in the set V. Since V is finite, we keep track of only a finite amount of information (a subset of V) at each step. This is exactly what the graph G constructed below does.

We begin by proving that V has the desired property.

LEMMA 6. *If two sequences* $(d_1, \dots, d_j)$ *and* $(d'_1, \dots, d'_j)$ *can be extended to longer sequences* $(d_1, \dots, d_j, \dots, d_{j+\ell})$, $(d'_1, \dots, d'_j, \dots, d'_{j+\ell})$ *having equal values under* R_u, *then the quantity (5) is in* V.

Proof. Suppose we are given $(z_1, z_2, z_3, \dots)$ satisfying $z_1 = 0$ and $z_{j+1} = xz_j + d_j - d'_j$ for all $j \geq 1$. Then z_{j+1} is the difference (5). The ith coordinate $\rho_i(z_j)$ for $i \in [k+1, n]$ is always bounded by $2\delta_i/(1 - |\lambda_i|)$ since $z_j \in V$ implies

$$|\rho_i(z_{j+1})| = |\rho_i(xz_j + d_j - d'_j)| \leq |\lambda_i||\rho_i(z_j)| + 2\delta_i \leq |\lambda_i|\frac{2\delta_i}{1 - |\lambda_i|} + 2\delta_i = \frac{2\delta_i}{1 - |\lambda_i|}.$$

If the ith coordinate $\rho_i(z_j)$ for $i \in [1, k]$ is bounded as $j \to \infty$ then we claim that $z_j \in V$ for all j. This will complete the proof of the lemma, since if $(d_1, \dots, d_j)$ and $(d'_1, \dots, d'_j)$ can be extended to be equal under R_u, then the difference (5) is equal to

$$\sum_{i=j+1}^{\ell}(d'_i - d_i)x^{-i}$$

which has bounded image under R_u.

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

If, say, $z_{j_0} \notin V$, then we must have $|\rho_i(z_{j_0})| > 2\delta_i/(|\lambda_i| - 1)$ for some $i \in [1, k]$. Then

$$\begin{aligned} |\rho_i(z_{j_0+1})| &= |\rho_i(xz_{j_0} + d_{j_0} - d'_{j_0})| > |\lambda_i \rho_i(z_{j_0})| - 2\delta_i \\ &= |\lambda_i| \left(|\rho_i(z_{j_0})| - \frac{2\delta_i}{|\lambda_i| - 1} \right) - 2\delta_i + \frac{|\lambda_i| 2\delta_i}{|\lambda_i| - 1} \end{aligned}$$

and so

$$|\rho_i(z_{j_0+1})| - \frac{2\delta_i}{|\lambda_i| - 1} \geq |\lambda_i| \left(|\rho_i(z_{j_0})| - \frac{2\delta_i}{|\lambda_i| - 1} \right).$$

In other words, since $|\lambda_i| > 1$, the z_{j_0+N} escape to ∞ in the ith coordinate as $N \to \infty$. This completes the proof of the claim. □

Continuation of the proof of Theorem 3. For $d \in D$ let D_d be the set

$$D_d = \{d - d_i \mid d_i \in D, d_i \prec d\} \cap V.$$

The set $D_{\mathbf{a}}$ is empty (recall that $\mathbf{a}$ is the minimal element of D).

Let G' be the graph whose vertices are all subsets of $V - \{0\}$ (including the empty set $\emptyset$), with one additional vertex labelled F called the *fail state*. Let G' have the following edges. From any vertex $U \neq F$, for every $d \in D$ there is an edge labelled d to the vertex

$$V \cap ((xU + d - D) \cup D_d),$$

provided this does not contain 0. If this set contains 0, there is instead an edge labelled d from U to the fail state F. Let G be the connected component of G' containing the vertex $\emptyset$. (For an example of a graph G, see §3.)

The interpretation of this graph G is as follows. If after reading j digits of d we are in state U (i.e. for all $u \in U$ there exists a sequence d' such that the difference (5) is u), then the new differences after reading d_{j+1} are either of the form $xU + d_{j+1} - D$, since we can extend each d' with any digit of D, or of the form $D_{d_{j+1}}$, coming from sequences d' which are identical to d up to the jth coordinate. Of all these new differences, we need only keep those in V, as noted earlier.

Here is the formal proof that G has the desired property. Let $d = (d_1, d_2, \ldots) \in D^{\mathbb{N}}$. We claim that d labels a path in G (that is, there is a path of edges starting from some vertex $U_1 \neq F$ labelled consecutively $d_1, d_2, \ldots,$) if and only if each finite initial sequence $(d_1, \ldots, d_N)$ of d is minimal, that is, there is no lower lexicographic sequence $(d'_1, \ldots, d'_N)$ in D^N with

$$x^{N-1}d_1 + x^{N-2}d_2 + \cdots + d_N = x^{N-1}d'_1 + x^{N-2}d'_2 + \cdots + d'_N. \tag{6}$$

First of all, if there were such a sequence $(d'_1, \ldots, d'_N)$, we will show that the path labelled by d leads from U_1 to F. Assume (by removing an equal initial segment of d and d' if necessary) that $d_1 \neq d'_1$. Define a sequence $(z_1, \ldots, z_{N+1}) \in V^{N+1}$ by $z_1 = 0$ and for $j \in [1, N]$,

$$z_{j+1} = xz_j + d_j - d'_j.$$

Now from (6), we have $z_{N+1} = 0$. By the definition of the graph G, if we start at U_1 and move along the path labelled by d, then the jth vertex U_j will contain the element z_j. Note, in particular, that $z_2 = d_1 - d'_1 \in D_{d_1} \subset U_2$ since $d'_1 \prec d_1$.

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

Thus the path labelled by $(d_1, \dots, d_N)$ will end up at the fail state after the Nth step, since $z_{N+1} = 0$. We conclude that no sequence labelling a finite path in G has a lower lexicographic sequence representing the same point of Λ.

It remains to show that every finite minimal sequence actually does occur in G. It is clear that any sequence of digits labels a path in G starting from the vertex $\emptyset$ except those which eventually end up in the fail state. Let $(d_1, \dots, d_N)$ label a path starting at $\emptyset$ which ends up in the fail state at time N. We will find a lower lexicographic sequence; this will complete the proof.

Assume without loss of generality that the path never passes through $\emptyset$ except where it starts (if it does, just consider the path starting from the last time it goes through $\emptyset$). Let $\{U_i\}$ be the sequence of vertices of G along the path d:

$$\emptyset = U_0 \xrightarrow{d_1} U_1 \xrightarrow{d_2} U_2 \xrightarrow{d_3} \cdots \xrightarrow{d_N} U_N = F.$$

Define a sequence $(z_N, z_{N-1}, \dots, z_0)$ with $z_0 = 0 = z_N$, $z_j \in U_j$ for $j \in [1, N-1]$, and a sequence $(d'_N, d'_{N-1}, \dots, d'_1) \in D^N$ as follows. First, since the edge labelled d_N from U_{N-1} leads to the fail state, there is an element $z_{N-1} \in U_{N-1}$ and a digit d'_N such that

$$z_N = 0 = xz_{N-1} + d_N - d'_N.$$

Similarly, there is an element $z_{N-2} \in U_{N-2}$ and $d'_{N-1} \in D$ such that

$$z_{N-1} = xz_{N-2} + d_{N-1} - d'_{N-1}.$$

Inductively for decreasing j define $z_j \in U_j$ and $d'_{j+1} \in D$ from z_{j+1} by the equation

$$z_{j+1} = xz_j + d_{j+1} - d'_{j+1}.$$

These definitions may not be unique.

At the end we have $z_0 = 0$, since the path began at the state $\emptyset$. This gives us an equation

$$x^{N-1}d_1 + x^{N-2}d_2 + \cdots + d_N = x^{N-1}d'_1 + x^{N-2}d'_2 + \cdots + d'_N$$

and, since $z_1 = d_1 - d'_1 \in U_1 = D_{d_1}$, $d_1 \neq d'_1$ and d'_1 is strictly less than d_1 in the order on D. Thus $(d'_1, \dots, d'_N)$ is lower lexicographically than $(d_1, \dots, d_N)$. This completes the proof. □

Proof of Theorem 4. We have already proved that πR is a factor map onto $\mathbb{T}^n$. We first show that it is bounded-to-one.

Let d, d' be distinct elements of Σ_u. There is an infinite path γ in G starting from $\emptyset$ labelled by d:

$$\left(\emptyset \xrightarrow{d_0} U_1 \xrightarrow{d_1} U_2 \longrightarrow \cdots\right).$$

We say that d' *follows* d *on the path* γ if there is a sequence $\{z_i\}_{i \geq 0}$ with $z_i \in U_i \cup \{0\}$ satisfying $z_0 = 0$ and $z_{i+1} = xz_i + d_i - d'_i$.

From the proof of Theorem 3, we see that there exists such a $\{z_0, z_1, \dots\}$ if $R_u(d_0, d_1, \dots) = R_u(d'_0, d'_1, \dots)$ and $(d_0, d_1, \dots) \succ (d'_0, d'_1, \dots)$: since then all the z_j defined as above are bounded and so $z_j \in U_j$. Conversely, if the $\{z_0, z_1, \dots\}$ defined

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

above are bounded then we must have $R_u(d_0, d_1, \ldots) = R_u(d_0', d_1', \ldots)$. So this is the condition for when R_u is non-injective.

Now take three distinct elements $d, d', d'' \in \Sigma_u$ such that $R_u(d) = R_u(d') = R_u(d'')$. After reordering the three elements d, d', d'' if necessary, we can assume that $d \succ d'$ and $d \succ d''$. Let $\{z_i'\}_{i\geq 0}$ and $\{z_i''\}_{i\geq 0}$ be the corresponding sequences as above for the pairs d, d' and d, d''.

Let i_0 be the first index at which $z_{i_0}' \neq z_{i_0}''$. Then $z_i' \neq z_i''$ for $i \geq i_0$ by minimality of d', d'': if $z_j' = z_j''$ for some $j > i_0$ then

$$\begin{aligned} z_j' &= x^{j-1}(d_0 - d_0') + x^{j-2}(d_1 - d_1') + \cdots + d_{j-1} - d_{j-1}' \\ &= z_j'' = x^{j-1}(d_0 - d_0'') + x^{j-2}(d_1 - d_1'') + \cdots + d_{j-1} - d_{j-1}'', \end{aligned}$$

which implies

$$x^{j-1}d_0' + x^{j-2}d_1' + \cdots + d_{j-1}' = x^{j-1}d_0'' + x^{j-2}d_1'' + \cdots + d_{j-1}''.$$

This is impossible by the minimality of the sequences d', d'' unless $d_\ell' = d_\ell''$ for each $\ell \in [0, j-1]$, contradicting $j > i_0$. By a similar argument using d and d', we have that for i sufficiently large $z_i' \neq 0$.

We conclude that the number of elements $d' \in \Sigma_u$ which satisfy $R_u(d) = R_u(d')$ is bounded by $|V|$ (note $|V|$ is the maximum cardinality of the $U_j \cup \{0\}$): there are at most $|V|$ sequences of $\{z_i\}_{i\geq 0}$, since once two such sequences diverge they must stay apart.

A similar argument holds for the unstable leaf: the condition for $R_s(d) = R_s(d')$ is that d' follows d backwards in G, in the following sense. There exists a sequence $\{z_i\}_{i\leq 0}$ with $z_i \in U_i$, $z_0 = 0$ and with $z_{i+1} = xz_i + d_i - d_i'$. Again, once two sequences $\{z_i'\}$ and $\{z_i''\}$ diverge (as $i \to -\infty$) they must stay apart. Thus the cardinality of $R_s^{-1}(d)$ is also at most $|V|$.

We conclude that the number of sequences $d' \in \Sigma$ with $R(d) = R(d')$ is bounded by $|V|^2$. Since R has bounded preimages and $R(\Sigma)$ is compact, the same is true for πR.

The fact that the map πR is constant-to-one almost everywhere is now a standard argument. Define a function $f : \mathbb{T}^n \to \mathbb{R}$ by $f(x)$ is the cardinality of $(\pi R)^{-1}(x)$. Then f is measurable. Also, f is invariant under the transformation A, by the equation $\pi R(\sigma d) = A\pi R(d)$ and the invertibility of A and σ (that is, $\pi R(d) = x$ iff $\pi R(\sigma(d)) = Ax$). Since A is ergodic for Lebesgue measure, f is constant almost everywhere. □

Proof of Proposition 5. We need to use a larger set of digits than just D. Let $D' = D \cup xV$, where V is the set of lattice points from the proof of Theorem 3.

For each $y \in \Lambda$, take k such that $\pi_u(A^{-k}(y)) \in \pi_u(B)$, where π_u is the projection onto W_u. Since A^{-1} is contracting on W_u and $\pi_u(B)$ contains an open neighborhood of 0 in W_u, such a k exists.

We can construct an expansion for $y' = A^{-k}y$ in $D'[x^{-1}]$ as follows. By Lemma 1, there exists a digit $d_1 \in D$ such that $\pi_u(xy' - d_1) \in \pi_u(B)$. Let $y^{(2)} = xy' - d_1$. There exists a digit $d_2 \in D$ with $\pi_u(xy^{(2)} - d_2) \in \pi_u(B)$, and so on. The coordinates in the W_s-direction of the lattice points $y^{(2)}, y^{(3)}, \ldots$ are contracting exponentially (as in the proof of Lemma 6), until some $y^{(r)}$ falls within the set V. At that point $xy^{(r)} \in xV$.

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press

Define $d_r = xy^{(r)}$. We then have

$$x^r y' - x^{r-1} d_1 - \cdots - x d_{r-1} - d_r = 0,$$

or

$$y' = \sum_{j=1}^{r} d_j x^{-j}.$$

Multiplying this expression by x^k gives an expansion of y in $D'[x, x^{-1}]$. □

REFERENCES

[1] R. Adler and B. Weiss. Entropy, a complete metric invariant for automorphisms of the torus. *Proc. Nat. Acad. Sci.* **57**(6) (1967), 1573–1576.

[2] D. Arov. About topological similarity of automorphisms and shifts on compact abelian groups. *Russian Math. Surveys* ('Uspehi Matematicheskih nauk') **18**(5) (1963), 133–138.

[3] A. Bertrand-Mathis. Developpement en base θ, répartition modulo un de la suite $(x\theta^n)_{n\geq 0}$, langages codés et θ-shift. *Bull. Soc. Math. France* **114** (1986), 271–323.

[4] R. Bowen. Markov partitions for Axiom A diffeomorphisms. *Amer. J. Math.* **92** (1970), 725–747.

[5] E. A. Hirsch. On the construction of a symbolic realization of a hyperbolic automorphism of the torus. *Preprint*, St. Petersburg.

[6] R. Kenyon. Self-similar tilings. *Ph.D. Thesis*, Princeton University, 1990.

[7] W. Parry. On the β-expansions of real numbers. *Acta Math. Acad. Sci. Hungar.* **11** (1960), 401–416.

[8] B. Praggastis. Markov partitions for hyperbolic toral automorphisms. *Ph.D. Thesis*, University of Washington, 1992.

[9] B. Praggastis. Numeration systems and Markov partitions from self-similar tilings. *Preprint*, University of Washington.

[10] A. Renyi. Representations for real numbers and their ergodic properties. *Acta Math. Acad Sci. Hungar.* **8** (1957), 477–493.

[11] K. Schmidt. On periodic expansions of Pisot numbers and Salem numbers. *Bull. London Math. Soc.* **12** (1980), 269–278.

[12] Y. Sinai. Markov partitions and U-diffeomorphisms. *Funkts. Anal. Prilozhen.* **2** (1968), 64–89.

[13] W. P. Thurston. *Groups, Tilings, and Finite-state Automata (Amer. Math. Soc. Colloquium Lectures)* (1990).

[14] A. Vershik. Arithmetic isomorphism of hyperbolic toral automorphisms and sofic shifts. *Funkts. Anal. Prilozhen* **26** (1992), 22–27.

[15] A. Vershik. The fibadic expansions of real numbers and adic transformations. *Preprint*, Institute Mittag-Leffler, 1991–92, No. 4.

[16] B. Weiss. Subshifts of finite type and sofic systems. *Monats. Math.* **77** (1973), 462–474.

https://doi.org/10.1017/S0143385798100445 Published online by Cambridge University Press