Ergod. Th. & Dynam. Sys. (2005), **25**, 107–131
DOI: 10.1017/S0143385704000689
© *2004 Cambridge University Press*
Printed in the United Kingdom

Embedding theorems for locally compact Markov shifts

DORIS FIEBIG and ULF-RAINER FIEBIG
IMS, Universität Göttingen, Maschmühlenweg 8-10, 37073 Göttingen, Germany
(e-mail: urfiebig@math.uni-goettingen.de)

(*Received* 15 *November* 2001 *and accepted in revised form* 16 *October* 2003)

Abstract. We characterize when a lower entropy locally compact subshift S embeds into a locally compact mixing Markov shift T. As in the compact case (when T is a shift of finite type) the existence of an embedding depends on S only through its periodic orbit counts. However, in contrast to the compact case, the topology of the target system T becomes important. This is demonstrated with the help of the Zeta Function Lemma, which in particular characterizes the periodic orbit counts and entropies of locally compact mixing Markov shifts.

0. *Introduction*

The three ingredients for our main embedding results, Theorems 1.2, 2.7 and 3.2, and the Zeta Function Lemma 2.4, are subshifts, entropy and periodic points. Therefore, we first describe the systems we consider, the notions of entropy used and some growth rates for periodic points and orbits. Then we give an outline of the paper.

A *subshift* is a shift-invariant subset S of $\mathbb{N}^{\mathbb{Z}}$, endowed with the product topology of the discrete topology on $\mathbb{N} = \{1, 2, \dots\}$, together with the left shift map. To reduce notation we use the same symbol for the space and the shift map—which one is meant will always be clear from the context. The topology on S is generated by the cylinder sets $[x_0 \cdots x_n]_{n+i} := \{y \in S \mid y_i = x_0, \dots, y_{i+n} = x_n\}$. For a point $x = (x_i)_{i\in\mathbb{Z}} \in S$ and integers $n \le m$ we use the notation $x[n, m]$ to denote the S-block $x_n x_{n+1} \cdots x_{m-1} x_m$ of x, use $x[n, \infty)$ to denote the right-infinite ray $x_n x_{n+1} \cdots$, and finally use $x(-\infty, m]$ to denote the left-infinite ray $\cdots x_{m-1} x_m$. We call a subshift S *transitive* if for any S-blocks a and b there is $x \in S$ and $n \le m$ such that $x[n, m] = awb$ for some, possibly empty, block w. (This is usually called transitive non-wandering.)

Let S and T be subshifts. An *embedding* from S into T is a shift-commuting continuous map $f : S \to T$ such that $f : S \to f(S)$ is a homeomorphism. We say that S and T are *conjugate* if there is a homeomorphism $f : S \to T$ which commutes with the shift maps. Any subshift T of $\mathbb{N}^{\mathbb{Z}}$ which is conjugate to S is called a *presentation* of S.

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

A *locally compact subshift* is a subshift S where the space S is locally compact. A *Markov shift* S is by definition a subshift which is conjugate to the set S_G of bi-infinite walks on the edges of a countable directed graph G, with the left shift transformation acting on S_G. We consider only non-degenerate graphs here, so that for every edge e there is a point $x \in S_G$ with $x_0 = e$. The subshift S_G is called a graph presentation of S. The Markov shift S is locally compact if and only if G has finite in- and out-degree (at most finitely many incoming and outgoing edges at every vertex), for a discussion of the relation between the possible degrees of G and entropy see [**F3**]. The Markov shift is compact if and only if G is a finite graph if and only if it is a shift of finite type (SFT). Transitivity of S means strong connectedness of G; that is, for every pair α, β of vertices there is a path from α to β. Finally, S is mixing if it is transitive with period 1.

Any locally compact subshift (or locally compact Markov shift) might also be compact, if not stated otherwise.

We will have to deal with different growth rates of sequences of numbers. For conciseness we use the following notation. Suppose $a = (a_n)_{n\in\mathbb{N}}$ is a sequence with $a_n \in \mathbb{Z}^+ \cup \{\infty\}$ and I is an infinite subset of $\mathbb{N}$. Let

$$GR_{n\in I}(a_n) := \limsup_{n\in I} \frac{1}{n} \log_+ a_n \quad \text{and} \quad GR^-_{n\in I}(a_n) := \liminf_{n\in I} \frac{1}{n} \log_+ a_n$$

where $\log_+(0) = 0$, $\log_+(\infty) = \infty$ and $\log_+(n) = \log(n)$ for $n \in \mathbb{N}$. These quantities attain values in $[0, \infty]$. For finite I let $GR_{n\in I}(a_n) = 0$ and $GR^-_{n\in I}(a_n) = \infty$. If $I = \mathbb{N}$ then we use $GR_n(a_n)$ and $GR^-_n(a_n)$.

The crucial entropy for our purposes is the following. Let T be a presentation of a locally compact subshift S. Let $C_T(i, j, n) = \{x[1, n] \mid x \in T,\ x_1 = i,\ x_n = j\}$ be the set of T-blocks of length n which start with the symbol i and end with the symbol j. The *transition entropy of the presentation T of S* is

$$h_t^{\text{pres}}(T) := \sup_{i,j} GR_n(\#C_T(i, j, n)).$$

The *transition entropy* of S is [**FF1**]

$$h_t(S) := \inf\{h_t^{\text{pres}}(T) \mid T \subset \mathbb{N}^{\mathbb{Z}} \text{ is a presentation of } S\}$$

Another natural notion of entropy for locally compact subshifts S is the *Gurevic entropy* $h_G(S)$, [**G1**], defined as $h_G(S) := h_{\text{top}}(S_0)$ where $h_{\text{top}}(S_0)$ denotes the (usual) topological entropy of the dynamical one-point compactification S_0 of S (the extended transformation fixes the point at infinity). For general locally compact subshifts S the only relation is $h_G(S) \le h_t(S)$ [**FF1**]. However, for transitive locally compact Markov shifts this implies $h_G(S) = h_t(S)$ since $h_G(S) = GR_n(\#C_{S_G}(i, j, n))$ for any pair of symbols i, j and any graph presentation S_G of S [**G2**, Theorem 5].

Therefore, a reader only interested in results about embeddings of transitive Markov shifts may replace h_t by h_G throughout the paper (and all proofs work without reference to the fact that $h_G(S) = h_t(S)$). But for general subshifts it is indeed crucial to consider h_t, see the end of this section. We will use the following facts about $h_G(S) = h_t(S)$ for transitive locally compact Markov shifts S.

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

- Let S_G be any graph presentation for S, v any vertex in G, and $C_v(n)$ the set of loops of length n at vertex v. Then $h_G(S) = GR_n(\#C_v(n))$ ('loop-counting').
- $h_G(S) = \sup\{h_G(T) \mid T \subset S,\ T \text{ is an SFT}\}$ [**G1, G2, K**].

For compact subshifts S, $h_G(S) = h_{\text{top}}(S) = h_t(S) = h_t^{\text{pres}}(T)$ for every presentation T of S.

Now suppose S is a selfmap of some set. Let $\text{Per}_n(S) = \{x \mid S^n x = x\}$ denote the points of period n. Let $\text{Per}(S) = \bigcup_n \text{Per}_n(S)$, $\text{per}_n(S) = \#\text{Per}_n(S)$. For $x \in \text{Per}(S)$ let $O(x) = \{S^n x \mid n \in \mathbb{Z}\}$ denote the orbit of x, $O_n(S)$ the set of all orbits of length n, and $o_n(S)$ the number of orbits of length n. We also use $o_n(S, K)$ to denote the number of orbits of length n contained in some subset $K \subset S$. Let

$$h^-_{\text{orb}}(S) = GR^-_n(o_n(S)) \quad \text{and} \quad h_{\text{orb}}(S) = GR_n(o_n(S));$$
$$h^-_{\text{per}}(S) = GR^-_n(\text{per}_n(S)) \quad \text{and} \quad h_{\text{per}}(S) = GR_n(\text{per}_n(S)).$$

By definition $h^-_{\text{orb}} \le \min\{h^-_{\text{per}}, h_{\text{orb}}\} \le \max\{h^-_{\text{per}}, h_{\text{orb}}\} \le h_{\text{per}}$.

We coined the term *periodic entropy* for $h_{\text{per}}(S)$ since it was the essential quantity to describe embeddings of locally compact subshifts into mixing SFTs [**FF1**].

For mixing SFTs all these quantities coincide and are equal to the topological entropy [**LM**]. The non-compact situation is more subtle. In this work, where we consider embeddings into non-compact systems, the growth rates h^-_{orb} and h_{orb} will be essential (Theorem 3.2, Example 5.1). Still, if $o_n(S) < \infty$ for all n (this holds whenever h_{per} is finite), then $h_{\text{orb}} = h_{\text{per}}$ (Lemma 5.2), but h^-_{orb} and h^-_{per} may behave quite differently (Example 5.1, Lemma 5.3).

Remark 0.1. Suppose T is a mixing locally compact Markov shift. Then $h_t(T) \le h^-_{\text{orb}}(T)$ since $h_t(T) = \sup\{h_t(R) \mid R \subset T \text{ mixing SFT}\}$ and $h_t(R) = h^-_{\text{orb}}(R)$ for mixing SFTs.

We now give a condensed outline of the paper—see §1 for some more discussion. Suppose S is a locally compact subshift and T a mixing locally compact Markov shift. When is there an embedding from S into T, or, in other words, when does T have a subsystem S' which is conjugate to S? Trivially necessary for the existence of an embedding are

$$h_t(S) \le h_t(T) \quad \text{and} \quad S \to_{\text{per}} T$$

where $S \to_{\text{per}} T$ means that $o_n(S) \le o_n(T)$ for all $n \in \mathbb{N}$.

To avoid the conjugacy problem we deal only with the case $h_t(S) < h_t(T)$. Krieger's embedding theorem [**Kr**] gives a complete answer in the compact case—if S is a compact subshift and T a mixing SFT with $h_t(S) < h_t(T)$, then S embeds into T if and only if

$$S \to_{\text{per}} T.$$

In [**FF1**, Theorem D] a complete answer was given for the case when S is a locally compact subshift and T is a mixing SFT with $h_t(S) < h_t(T)$. Then S embeds into T if and only if

$$S \to_{\text{per}} T \quad \text{and} \quad h_{\text{per}}(S) < h_t(T).$$

Here we solve the general case, where S is a locally compact subshift and T a mixing locally compact Markov shift with $h_t(S) < h_t(T)$. Our main result, the Embedding

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

Theorem 1.2, says that there is an embedding from S into T if and only if

$$S \to_{\mathrm{per}} D \quad \text{for a non-dense subset } D \subset T,$$

where $S \to_{\mathrm{per}} D$ means that $o_n(S) \leq o_n(T, D)$ for all $n \in \mathbb{N}$.

Thus, as in the SFT case, the possibility of an embedding of a lower-entropy system S is solely determined by its orbit counts $(o_n(S))_{n\in\mathbb{N}}$. However, the topology of the target shift actually becomes important—there are mixing locally compact Markov shifts S, T_1, T_2 where T_1 and T_2 have the same orbit counts, $h_t(S) < h_t(T_1) = h_t(T_2)$, and S embeds into T_1 but not into T_2 (Theorem 2.7 actually proves something stronger). This example is obtained by use of the second major result, the Zeta Function Lemma 2.4, which describes all triples of orbit counts, entropy and orbit counts of a non-dense subset that can be realized by a non-trivial mixing locally compact Markov shift T and a non-dense subset D of T. (In particular, it describes all zeta functions and entropies of mixing locally compact Markov shifts.) It also gives a result about embeddings into some higher-entropy Markov shifts with pre-assigned orbit counts, just in terms of orbit counts and entropy (Theorem 3.4).

We obtain several sufficient conditions for embeddings just in terms of orbit counts and entropy (Theorem 3.2), the simplest one being $h_{\mathrm{orb}}(S) < h^-_{\mathrm{orb}}(T)$. This condition cannot be replaced, for example, by $h_{\mathrm{per}}(S) < h_{\mathrm{per}}(T)$ or $h_{\mathrm{per}}(S) < h^-_{\mathrm{per}}(T)$ (Example 5.1). This shows why it is important to consider growth rates of orbits instead of periodic points.

In §4 we pose a question about orbit counts of non-dense subsets. A positive answer would show that our embedding results in terms of orbit counts and entropy (Theorem 3.2) are best possible.

Finally, §§6 and 7 are devoted to the proofs of the two major results, the Embedding Theorem 1.2 and the Zeta Function Lemma 2.4. These proofs are self-contained and can be read independently of the other parts of this paper.

Why do we consider transition entropy instead of Gurevic entropy? As stated, for transitive locally compact Markov shifts S there is no difference. But for general locally compact subshifts S our embedding results become false if one replaces transition entropy by Gurevic entropy. We give an example. There is a transitive locally compact subshift S with $h_G(S) \leq \log 3$, $S \to_{\mathrm{per}} S_3$ (the full 3-shift) and $h_t(S) = \log 9$, [**FF1**, Example 3.10]. Thus, the Gurevic entropy $h_G(S)$ and the orbit counts of S do not rule out the possibility of an embedding in the full 4-shift S_4. But there is no embedding into S_4, since $h_t(S) > \log 4 = h_t(S_4)$!

1. *The Embedding Theorem*

We discuss the main result of this work, the Embedding Theorem 1.2. Some previously known embedding results are easily obtained as corollaries.

If not stated otherwise, any locally compact subshift or Markov shift might also be compact. In particular, the Embedding Theorem applies if one, or both, systems are compact. Moreover, there are no assumptions on the zeta functions of the systems under consideration. They may have infinitely many orbits of length n for some or all $n \in \mathbb{N}$.

First, we observe that lower-entropy subsystems are non-dense. The converse does not hold—there are mixing locally compact Markov shifts that contain proper closed

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

subsystems of equal entropy (which can be chosen to be mixing Markov shifts); see [**K**], and [**Fi**] for a mixing modification.

LEMMA 1.1. *Let T be a locally compact mixing Markov shift. Let $S \subset T$ be a locally compact subshift with $h_t(S) < h_t(T)$. Then S is non-dense.*

Proof. Let T be given in graph representation and let $S \subset T$ be dense. We show that $h_t(S) = h_t(T)$. Since S is locally compact, there is an open set $O \subset T$ and a closed set $A \subset T$ such that $S = O \cap A$ [**D**, Theorem 6.5]. Since S is dense, A is dense in T, and thus $A = T$. Thus S is open. Now let $S' \subset \mathbb{N}^{\mathbb{Z}}$ be a presentation of S and let $f : S \to S'$ be a conjugacy. Fix $x \in S$ and let $x' = fx$. Since S is open and f is continuous, there is $N \geq 0$ such that $C := [x_{-N}, \ldots, x_N]_N \subset S$ and $f(C) \subset [x'_0]_0$. Thus, mapping $y \in \mathrm{Per}_n(T) \cap C$ to $(fy)[0, n]$ is an injective map $\mathrm{Per}_n(T) \cap C \to C_{S'}(x'_0, x'_0, n+1)$. Thus,

$$h_t^{\mathrm{pres}}(S') \geq GR_n(\#C_{S'}(x'_0, x'_0, n)) \geq GR_n(\#(\mathrm{Per}_n(T) \cap C)) = h_t(T).$$

Thus $h_t(S) = \inf\{h_t^{\mathrm{pres}}(S') \mid S' \subset \mathbb{N}^{\mathbb{Z}} \text{ is a presentation of } S\} \geq h_t(T)$. □

For a locally compact mixing Markov shift T, let

$$ND(T) := \{(o_n(T, D))_{n\in\mathbb{N}} \mid D \subset T \text{ non-dense}\}.$$

EMBEDDING THEOREM 1.2. *Let T be a mixing locally compact Markov shift and S a locally compact subshift with $h_t(S) < h_t(T) \leq \infty$. The following are equivalent:*

(a) *there is an embedding $g : S \to T$;*

(b) *$S \to_{\mathrm{per}} D$ for some non-dense subset $D \subset T$;*

(c) *$(o_n(S))_{n\in\mathbb{N}} \in ND(T)$.*

Proof. Obviously (c) is just a restatement of (b). Lemma 1.1 shows that (a) implies (b). The proof that (b) implies (a) is given in §7. □

Condition (c) shows that the lower-entropy subsystems S of T are still, as in the SFT case, determined by their orbit counts $o_n(S)$, but in a more complicated way, namely by the set $ND(T)$. The new challenge is to get more information on $ND(T)$. By definition, $ND(T) = \{(b_n)_{n\in\mathbb{N}} |$ there is a non-empty compact open subset K of T such that $0 \leq b_n \leq o_n(T, K^c)$ for all $n \in \mathbb{N}\}$, which is not very satisfactory, since it involves countably many sets K.

In the following sections we obtain some general information on the set $ND(T)$. We show that for certain sequences $(b_n)_{n\in\mathbb{N}}$ with $0 \leq b_n \leq o_n(T)$ for all $n \in \mathbb{N}$ it is possible to determine if they belong to $ND(T)$ or not, just from knowing the orbit counts and the entropy of T (Theorem 2.2, Corollary 2.3). This gives some simpler sufficient conditions for embeddings (Theorem 3.2). And we determine for which sequences $(b_n)_{n\in\mathbb{N}}$ there is at least some (possibly different) mixing locally compact Markov shift T' with the same entropy and the same orbit counts as T and such that $(b_n)_{n\in\mathbb{N}}$ belongs to $ND(T')$ (the Zeta Function Lemma 2.4.) On the other hand, Theorem 2.7 shows that generically it is not possible to determine $ND(T)$ just from knowing the orbit counts and the entropy of T—whenever $h_t(T) > 0$, $o_n(T) < \infty$ for infinitely many n, then there is a mixing locally compact Markov shift T' with $o_n(T) = o_n(T')$ for all n, $h_t(T) = h_t(T')$, which contains

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

lower-entropy subsystems (these can be chosen as mixing locally compact Markov shifts) that do not embed into T.

Of course, extra information on the structure of the underlying system T might help us to say more about $ND(T)$. For example, if T is an *SFT* then the set $ND(T)$ depends only on the orbit counts of T (Lemma 1.4). It remains an interesting problem to study $ND(T)$ for other classes of locally compact Markov shifts.

One interesting piece of information on $ND(T)$ is provided by the following lemma. It shows that the set $ND(T)$ remains the same if we replace the term 'non-dense' by 'discrete'. Recall that a subset D of a topological space X is *discrete* if for each point $x \in D$ there is an open neighborhood U of x in X with $U \cap D = \{x\}$.

PERIODIC POINT LEMMA 1.3. *Suppose T is a mixing locally compact Markov shift. Suppose $D \subset T$ is a non-dense subset. Then there is a discrete subset $D' \subset T$ with $o_n(T, D) = o_n(T, D')$ for all n.*

Thus, whenever (b) of the Embedding Theorem 1.2 is satisfied the set D can be chosen to be discrete. This extra information will be used in the proof of the Embedding Theorem to show that the continuous map $g : S \to T$ constructed there is actually an embedding. (Injectivity of g can be seen without reference to the discreteness of D, but it is essential for our proof that the inverse of g is continuous.) The proof of Lemma 1.3 is deferred to §7.

A well-known characterization of $ND(T)$ for SFTs allows us to obtain some known embedding results as corollaries of the Embedding Theorem 1.2.

LEMMA 1.4. *Suppose T is a mixing SFT. Let $b_n \geq 0$ be a sequence of integers with $b_n \leq o_n(T)$ for all n. Then $(b_n)_{n\in\mathbb{N}} \in ND(T)$ if and only if $GR_n(b_n) < h_{\text{top}}(T) = GR_n(o_n(T))$.*

Proof. If there is a non-dense shift-invariant subset $D \subset T$ with $o_n(T, D) = b_n$ for all n then the closure $\text{clos}(D)$ of D is a compact proper subshift in T. Thus $GR_n(b_n) \leq h_{\text{top}}(\text{clos}(D)) < h_{\text{top}}(T)$, since T is mixing [**DGS, LM**]. If $GR_n(b_n) < h_{\text{top}}(T)$ then choose n such that $b_n < o_n(T)$ and choose a point $x \in \text{Per}_n(T)$ of least period n. Let D be the SFT obtained by forbidding $x[-N, N]$. For large enough N this shows $(b_n)_{n\in\mathbb{N}} \in ND(T)$ [**LM**, Lemma 10.1.3]. □

COROLLARY 1.5. (Essentially [**FF1**]) *Suppose T is a mixing SFT and S a locally compact subshift with $h_t(S) < h_t(T)$. Then S embeds into T if and only if $S \to_{\text{per}} T$ and $h_{\text{per}}(S) < h_t(T)$.*

Proof. If there is an embedding $g : S \to T$ then by Lemma 1.1 the closure $C = \text{clos}(g(S))$ is a proper closed subsystem of T, and thus $h_{\text{per}}(S) \leq h_{\text{per}}(C) \leq h_{\text{top}}(C) < h_{\text{top}}(T) = h_t(T)$.

If $S \to_{\text{per}} T$ and $h_{\text{per}}(S) < h_t(T)$ then $h_{\text{orb}}(S) = h_{\text{per}}(S)$ by Lemma 5.2, thus by Lemma 1.4, condition (c) of the Embedding Theorem 1.2 is satisfied. □

COROLLARY 1.6. (Essentially [**FF1**]) *Suppose T is a mixing locally compact Markov shift and S a locally compact subshift. Then S embeds into a compact proper subsystem of T if and only if $S \to_{\text{per}} T$ and $\max\{h_{\text{per}}(S), h_t(S)\} < h_t(T)$.*

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

Proof. If S embeds into a compact proper subsystem, then it embeds into a compact mixing sub-SFT of T. Thus, $\max\{h_{\mathrm{per}}(S), h_t(S)\} < h_t(T)$ and $S \to_{\mathrm{per}} T$.

If $\max\{h_{\mathrm{per}}(S), h_t(S)\} < h_t(T)$ and $S \to_{\mathrm{per}} T$ then there is a mixing SFT $R \subset T$ with $R \neq T$ and $\max\{h_{\mathrm{per}}(S), h_t(S)\} < h_t(R)$ and $S \to_{\mathrm{per}} R$. Thus, by Corollary 1.5 there is an embedding $g : S \to R$. □

COROLLARY 1.7. (Krieger's embedding theorem [**Kr**]) *Suppose T is a mixing locally compact Markov shift and S a compact subshift with $h_t(S) < h_t(T)$. Then there is an embedding of S into T if and only if $S \to_{\mathrm{per}} T$.*

Proof. Since $h_{\mathrm{per}}(S) \leq h_{\mathrm{top}}(S)$ this follows from Corollary 1.5. □

Remark 1.8. The last two corollaries show that:

- the compact lower-entropy subsystems S of T are determined by $S \to_{\mathrm{per}} T$;
- the relatively compact lower-entropy subsystems S of T are determined by $S \to_{\mathrm{per}} T$ and $h_{\mathrm{per}}(S) < h_t(T)$. Note that the condition $h_{\mathrm{per}}(S) < h_t(T)$ is equivalent to $h_{\mathrm{orb}}(S) < h_t(T)$ and $o_n(S) < \infty$ for all n (Lemma 5.2).

We close this section with some remarks on embeddings with closed image. In general the Embedding Theorem 1.2 cannot yield embeddings $g : S \to T$ where the image $g(S)$ is closed in T. This is obvious if S is non-compact and T an SFT. We describe a more general case.

LEMMA 1.9. *Suppose $g : S \to T$ is an embedding of a locally compact subshift S into a locally compact subshift T such that $g(S)$ is closed in T. Suppose that T has a rome, i.e. $T = \bigcup_{n\in\mathbb{Z}} T^n K$ for some compact subset K. Then S has a rome, too.*

Proof. Since $g(S)$ closed in T, $K \cap g(S)$ is a rome in $g(S)$, thus $g^{-1}(K \cap g(S))$ is a rome for S. □

Examples of Markov shifts with rome and without rome can be found in [**F1**, Example 1.4]. The tools used to prove the following result will be provided in the next two sections.

PROPOSITION 1.10. *Suppose T is a mixing locally compact Markov shift with a rome. Then there is a mixing locally compact Markov shift S, $h_t(S) < h_t(T)$, which embeds into T, but no embedding of S into T has a closed image.*

Proof. Since T has a rome $h_t(T) = h^-_{\mathrm{orb}}(T)$ by Lemma 2.1. By the Zeta Function Lemma 2.4 it is easy to obtain a mixing locally compact Markov shift S with $h_t(S) < h_{\mathrm{orb}}(S) < h_t(T)$ and $S \to_{\mathrm{per}} T$. Then S embeds into T by Theorem 3.2, but $h_t(S) < h_{\mathrm{orb}}(S)$ implies by Lemma 2.1 that S cannot have a rome; thus no embedding can have a closed image by Lemma 1.9. □

Remark 1.11. There are further restrictions. Suppose S, T are locally compact non-compact subshifts and $g : S \to T$ an embedding with $g(S)$ closed in T. Then it is not hard to see that g extends to an embedding of the dynamical one-point compactifications S_0 and T_0 (the point at infinity is a fixed point for the extended transformations). There are examples of mixing locally compact Markov shifts S and T both having a rome,

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

$h_t(S) < h_t(T)$, S embeds into T, but never with a closed image since T_0 is expansive and S_0 is not. Examples of Markov shifts with T_0 expansive as well as not expansive can be found in [**F2**, Lemma 4.4].

2. *Orbit counts for non-dense subsets*

Two results contribute to a better understanding of the *ND* condition in the Embedding Theorem 1.2. Theorem 2.2 gives some information on $ND(T)$ just in terms of orbit counts and entropy of T. The Zeta Function Lemma 2.4 characterizes for which triples $((a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}, h)$ there is a mixing locally compact Markov shift T with entropy h, orbit counts a_n and $(b_n)_{n\in\mathbb{N}} \in ND(T)$. This is also a handy tool to construct examples. We use it here to show that the lower-entropy subsystems of a non-trivial mixing locally compact Markov shift T are never determined just by the entropy and the orbit counts of T, as long as $o_n(T) < \infty$ for infinitely many n (Theorem 2.7).

For the first result, Theorem 2.2, we need a characterization of the Gurevic entropy, provided by part (a) of the following lemma.

LEMMA 2.1. *Suppose T is a mixing locally compact Markov shift. Let $K \subset T$ be a non-empty compact open subset. Let $S = \bigcup_{n\in\mathbb{Z}} S^n K$. Then:*

(a) *the numbers $k_n = \#\{O \in O_n(T) \mid O \cap K \neq \emptyset\}$ are finite and $GR_n^-(k_n) = GR_n(k_n) = h_t(T)$;*

(b) *S is a mixing locally compact Markov shift with a rome;*

(c) *$o_n(S) < \infty$ for all n and $h^-_{\mathrm{orb}}(S) = h_{\mathrm{orb}}(S) = h^-_{\mathrm{per}}(S) = h_{\mathrm{per}}(S) = h_t(S) = h_t(T)$.*

Proof. (a) is a consequence of (c) since $k_n = o_n(S)$.

(b) Since S is a shift-invariant open subset of T, it is a locally compact subshift which is mixing and every point sees a synchronizing block. Thus, S is a Markov shift by [**FF2**]. The set K is a rome for S.

(c) Since S has a rome, we can choose a rome consisting of finitely many vertices in a graph presentation of S; thus, using loop-counting shows $h^-_{\mathrm{per}}(S) = h_{\mathrm{per}}(S) = h_t(S)$, since every periodic point visits one of a finite set of vertices. Again, since S has a rome and since the graph presentation has finite out-degree at each vertex by locally compactness, $o_n(S) < \infty$. By Remark 0.1, $h_t(S) \le h^-_{\mathrm{orb}}(S)$; thus, $h_t(S) \le h^-_{\mathrm{orb}}(S) \le h_{\mathrm{per}}(S) = h_t(S)$ by the above. Since S is a dense subsystem, Lemma 1.1 shows $h_t(S) = h_t(T)$. □

Recall the conventions $\infty - \infty = \infty$ and $GR_{n\in I}(b_n) = 0$ if I is a finite set.

THEOREM 2.2. *Suppose T is a mixing locally compact Markov shift. Let $I(T) = \{n \in \mathbb{N} \mid o_n(T) < \infty\}$. Suppose $(b_n)_{n\in\mathbb{N}}$ is given with $0 \le b_n \le o_n(T)$ for all n.*

(a) $(b_n)_{n\in\mathbb{N}} \in ND(T) \Rightarrow GR_n^-(o_n(T) - b_n) \ge h_t(T)$.

(b) $GR_{n\in I(T)}(b_n) < h_t(T) \Rightarrow (b_n)_{n\in\mathbb{N}} \in ND(T)$.

(c) $GR_n^-(o_n(T) - b_n) > h_t(T) \Rightarrow (b_n)_{n\in\mathbb{N}} \in ND(T)$.

Proof. (a) Let $D \subset T$ be non-dense with $b_n = o_n(T, D)$ for all n. By local compactness there is a non-empty $K \subset D^c$ which is compact open. Let $k_n = \#\{O \in O_n(T) \mid O \cap K \neq \emptyset\}$. Then $o_n(T, D) + k_n \le o_n(T)$ for all n. This implies $k_n \le o_n(T) - b_n$,

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

since by convention the right-hand side is ∞ whenever $o_n(T) = \infty$. By Lemma 2.1(a), $h_t(T) = GR_n^-(k_n) \leq GR_n^-(o_n(T) - b_n)$.

(b) Choose a mixing SFT $R \subset T$, $R \neq T$, with $GR_{n\in I(T)}(b_n) < h_t(R)$, [**LM**]. Since $h_t(R) = GR_n(o_n(R))$, there is an N such that $o_n(R) \geq b_n$ for all $n \geq N$, $n \in I(T)$. There are only finitely many orbits of length $n < N$, $n \in I(T)$; let F denote their union. Then $T - (R \cup F)$ is open and non-empty (T has no isolated points) and thus contains a non-empty compact open set K with $o_n(T, K^c) \geq o_n(R) \geq b_n$ for all $n \in I(T)$. Note that $o_n(T, K^c) = \infty$ whenever $n \notin I(T)$ by Lemma 2.1(a). Thus $(b_n)_{n\in\mathbb{N}} \in ND(T)$.

(c) Start with any non-empty compact open subset $L \subset T$. Let $k_n = \#\{O \in O_n(T) \mid O \cap L \neq \emptyset\}$. Now $GR_n^-(o_n(T) - b_n) > h_t(T)$ and Lemma 2.1(a) implies that $o_n(T) - b_n \geq k_n$ for all $n \geq N$, N large enough. This shows that $o_n(T, L^c) \geq b_n$ for all $n \geq N$, $n \in I(T)$. Now choose F and then K as in (b). □

There is a corollary solely in terms of orbit counts.

COROLLARY 2.3. *$GR_{n\in I(T)}(b_n) < h^-_{\text{orb}}(T) \Rightarrow (b_n)_{n\in\mathbb{N}} \in ND(T)$.*

Proof. If $GR_{n\in I(T)}(b_n) < h_t(T)$, then Theorem 2.2(b) applies. Otherwise there are $\mu < \lambda$ such that $h_t(T) \leq GR_{n\in I(T)}(b_n) < \log\mu < \log\lambda < h^-_{\text{orb}}(T)$. Thus, there is an N such that $o_n(T) \geq \lambda^n$ for all $n \geq N$ and $b_n < \mu^n$ for all $n \geq N$, $n \in I(T)$. Since by definition $GR_n^-(o_n(T) - b_n) = GR^-_{n\in I(T)}(o_n(T) - b_n)$, the above shows that $GR_n^-(o_n(T) - b_n) \geq \log\lambda > h_t(T)$ and Theorem 2.2(c) applies. □

Theorem 2.2(a) shows that $GR_n^-(o_n(T) - b_n) \geq h_t(T)$ is necessary for $(b_n)_{n\in\mathbb{N}} \in ND(T)$. It is not sufficient in general, as the sequence (b_n) obtained in the proof of Theorem 2.7 shows. Nonetheless, by the following there is always some mixing locally compact Markov shift T' with the same orbit counts and the same entropy as T such that $(b_n)_{n\in\mathbb{N}} \in ND(T')$.

ZETA FUNCTION LEMMA 2.4. *Let $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ be sequences with a_n, $b_n \in \{0, 1, 2, 3, \ldots\} \cup \{\infty\}$ and $a_n \geq b_n$ for all n, and let $0 < h \leq \infty$. The following are equivalent (with the convention $\infty - \infty = \infty$).*

(1) *There is a mixing locally compact Markov shift T with $h_t(T) = h$, $(o_n(T))_{n\in\mathbb{N}} = (a_n)_{n\in\mathbb{N}}$, and $(b_n)_{n\in\mathbb{N}} \in ND(T)$.*

(2) *$GR_n^-(a_n - b_n) \geq h$.*

Proof. That (1) implies (2) is Theorem 2.2(a). The constructive proof that (2) implies (1) given in §6 produces a non-compact system T and a discrete subset D of T with $o_n(T, D) = b_n$ (the set D is even closed). This gives a direct proof that the Zeta Function Lemma remains true with 'non-dense' replaced by 'discrete' in the definition of $ND(T)$, which, of course, is also implied by the Periodic Point Lemma 1.3. □

Remark 2.5. With $b_n = 0$ for all n, the Zeta Function Lemma 2.4 characterizes all pairs (orbit counts, entropy) which can be realized by a non-trivial mixing locally compact Markov shift.

LEMMA 2.6. *Suppose T is a mixing locally compact Markov shift with $I(T) = \{i \in \mathbb{N} \mid o_i(T) < \infty\}$ finite. Then, a locally compact subshift S with $h_t(S) < h_t(T)$ embeds into T if and only if $S \to_{\text{per}} T$.*

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

Proof. If $S \to_{\text{per}} T$ then $o_n(S) \leq o_n(T)$ for all n, and $I(T)$ finite implies $GR_{n\in I(T)}(o_n(S)) = 0 < h_t(T)$ thus $(o_n(S))_{n\in\mathbb{N}} \in ND(T)$ by Theorem 2.2(b) and the Embedding Theorem 1.2 applies. □

Suppose that T is a non-trivial mixing locally compact Markov shift. What can be said about the lower-entropy subsystems of T just from knowing $h_t(T)$ and $(o_n(T))_{n\in\mathbb{N}}$? By Remark 1.8 this information suffices to determine all those which are compact or relatively compact, and if $|I(T)| < \infty$ then by Lemma 2.6 all lower-entropy subsystems are determined by $h_t(T)$ and $(o_n(T))_{n\in\mathbb{N}}$.

We show that this is special; i.e. for a non-trivial mixing locally compact Markov shift T with $|I(T)| = \infty$ the lower-entropy subsystems (even those which are mixing locally compact Markov shifts) are never determined just by the entropy and the orbit counts of T. (Of course, if T is, for example, known to be compact, then the lower-entropy subsystems are described by $S \to_{\text{per}} T$, $h_{\text{per}}(S) < h_t(T)$ see Corollary 1.5. It is interesting to study notions weaker than compactness that still give good information on these subsystems, but we will not pursue this idea in this paper.)

THEOREM 2.7. *Suppose T is a non-trivial mixing locally compact Markov shift with $I(T) = \{i \in \mathbb{N} \mid o_i(T) < \infty\}$ infinite. Then,*

(a) *there is a mixing locally compact Markov shift T' with $(o_n(T'))_{n\in\mathbb{N}} = (o_n(T))_{n\in\mathbb{N}}$, $h_t(T') = h_t(T)$ with the property that for every $0 < \varepsilon < h_t(T)$ there is a mixing locally compact Markov shift S with $h_t(S) = \varepsilon$ that embeds into T' but not into T;*

(b) *for every $0 < \varepsilon < h_t(T)$ there is a mixing locally compact Markov shift S with $h_t(S) = \varepsilon$ and $S \to_{\text{per}} T$ that does not embed into T.*

Proof. (a) Let $K_1, K_2, \ldots,$ be an enumeration of the countably many non-empty compact open sets of T. By Lemma 2.1(a) there is a sequence $n_1 < n_2 < \cdots$ in $I(T)$ such that $|n_i^{-1} \log(\#\{O \in O_{n_i}(T) \mid O \cap K_i \neq \emptyset\} - 1) - h_t(T)| \leq 1/i$ (*).

Let $b_n = \infty$ if $o_n(T) = \infty$ and otherwise let b_n be the least integer with $b_n \geq o_n(T)/2$ if $n \notin \{n_i \mid i \in \mathbb{N}\}$ and with $b_n \geq \max\{o_n(T)/2, o_n(T) - \#\{O \in O_{n_i}(T) \mid O \cap K_i \neq \emptyset\} + 1\}$ if $n = n_i$ for some i.

This defines a sequence b_n with $o_n(T)/2 \leq b_n \leq o_n(T)$.

By Remark 0.1 $GR_n^-(o_n(T)/2) \geq h_t(T)$. This and (*) imply $GR_n^-(o_n(T) - b_n) \geq h_t(T)$. By the Zeta Function Lemma 2.4 $(b_n) \in ND(T')$ for some mixing locally compact Markov shift T' with $h_t(T') = h_t(T)$ and $o_n(T') = o_n(T)$ for all n.

Now let $0 < \varepsilon < h_t(T)$. Then $GR_n^-(b_n) > \varepsilon$ and again by the Zeta Function Lemma there is a mixing locally compact Markov shift S with $o_n(S) = b_n$ for all n and $h_t(S) = \varepsilon$. Then $(o_n(S)) \in ND(T')$ and S embeds into T' by the Embedding Theorem 1.2.

Assume there is an embedding g of S into T. By Lemma 1.1 $g(S)$ is non-dense in T. By local compactness there is an i such that K_i is disjoint from $g(S)$. In particular, all orbits of length n_i of $g(S)$ belong to K_i^c; thus, $b_{n_i} = o_{n_i}(S) \leq o_{n_i}(T) - \#\{O \in O_{n_i}(T) \mid O \cap K_i \neq \emptyset\}$, which clearly contradicts the definition of b_{n_i}.

(b) This is a consequence of (a) since $S \to_{\text{per}} T'$ implies $S \to_{\text{per}} T$. However, the argument can be simplified. Since $GR_n^-(o_n(T)) \geq h_t(T) > 0$, the Zeta Function Lemma 2.4 produces a mixing locally compact Markov shift S with $h_t(S) = \varepsilon$ and

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

$o_n(S) = o_n(T)$ for all n. Now $|I(T)| = \infty$ implies that $o_n(T) - o_n(S) = 0$ infinitely often; thus, $GR_n^-(o_n(T) - o_n(S)) = 0 < h_t(T)$ and S does not embed into T by Theorem 2.2(a) and Lemma 1.1. □

3. *Embedding results solely in terms of orbit counts and entropy*

Two results from the last section, Theorem 2.2 and Corollary 2.3, immediately translate into embedding results. The first gives a necessary, the second several sufficient numerical conditions in terms of orbit counts and entropy.

THEOREM 3.1. *Suppose T is a mixing locally compact Markov shift and S a locally compact subshift with $h_t(S) < h_t(T) \leq \infty$ and $S \to_{\text{per}} T$. If S embeds into T then $GR_n^-(o_n(T) - o_n(S)) \geq h_t(T)$.*

Proof. Lemma 1.1 and Theorem 2.2(a). □

THEOREM 3.2. *Suppose T is a mixing locally compact Markov shift and S a locally compact subshift with $h_t(S) < h_t(T) \leq \infty$ and $S \to_{\text{per}} T$. Any of the following conditions implies that S embeds into T:*

(1) $GR_{n\in I(T)}(o_n(S)) < h_t(T)$;
(2) $GR_n^-(o_n(T) - o_n(S)) > h_t(T)$;
(3) $GR_{n\in I(T)}(o_n(S)) < h_{\text{orb}}^-(T)$;
(4) $h_{\text{orb}}(S) < h_{\text{orb}}^-(T)$.

Proof. (1), (2) Theorem 2.2 with $b_n = o_n(S)$ and the Embedding Theorem 1.2.

(3) Corollary 2.3 with $b_n = o_n(S)$ and the Embedding Theorem.

(4) This implies (3) since $GR_{n\in I(T)}(o_n(S)) \leq h_{\text{orb}}(S)$ by definition. □

The conditions (3), (4) only use orbit counts. Condition (4) implies (3) which covers more cases, for example the case where $I(T)$ is finite and $o_n(S) = \infty$ whenever $o_n(T) = \infty$. If (3) holds, then at least one of the conditions (1), (2) holds, as the proof of Corollary 2.3 shows.

Examples 3.3. (a) An example where there is an embedding but Theorem 3.2 does not apply (used again in §5). By the Zeta Function Lemma 2.4 there are mixing locally compact Markov shifts S, T with $o_n(S) = 2^n$, $o_n(T) = 2^{n+1}$ for all n, $h_t(S) < h_t(T) = \log 2$ and $(2^n)_{n\in\mathbb{N}} \in ND(T)$. Thus S embeds into T by the Embedding Theorem 1.2. For both S and T all growth rates h_{orb}^-, h_{orb}, h_{per}^-, h_{per} equal $\log 2$ and also $GR_n^-(o_n(T) - o_n(S)) = GR_{n\in I(T)}(o_n(S)) = h_t(T) = \log 2$.

(b) An example where condition (2) of Theorem 3.2 is satisfied but no other condition of Theorem 3.2. Modify the above such that $h_t(T) < \log 2$. Then $GR_n^-(o_n(T) - o_n(S)) = GR_n^-(2^n) = \log 2 > h_t(T)$. Still all growth rates h_{orb}^-, h_{orb}, h_{per}^-, h_{per} equal $\log 2$; thus (1), (3), (4) do not hold. □

Now we slightly change our point of view. We discuss the existence of embeddings into some higher-entropy Markov shift with pre-assigned orbit counts. There is a complete answer in terms of entropy and orbit counts.

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

THEOREM 3.4. *Suppose S is a locally compact subshift. Let $(a_n)_{n\in\mathbb{N}}$ with $a_n \in \{0, 1, 2, 3, \dots\} \cup \{\infty\}$ with $o_n(S) \le a_n$ for all n and let $0 < h \le \infty$ with $h_t(S) < h$. The following are equivalent (again with the convention $\infty - \infty = \infty$).*

(1) *The subshift S embeds into some locally compact mixing Markov shift T with $h_t(T) = h$ and $o_n(T) = a_n$ for all n.*

(2) $GR_n^-(a_n - o_n(S)) \ge h$.

Proof. Given T as in (1) then $(o_n(S))_{n\in\mathbb{N}} \in ND(T)$ by the Embedding Theorem 1.2 and thus (2) holds by Theorem 3.1. (2) implies (1) by the Zeta Function Lemma 2.4 and the Embedding Theorem. □

Given $\varepsilon > 0$ any locally compact subshift S with $h_t(S) < h_{\mathrm{per}}(S) < \infty$ embeds into some mixing SFT T with $h_{\mathrm{per}}(T) = h_t(T) < h_{\mathrm{per}}(S) + \varepsilon$, [**FF1**, Theorem D]. The next result shows that one can obtain equal periodic entropy if one allows T to be non-compact.

COROLLARY 3.5. *Let S be a locally compact subshift with $h_t(S) < h_{\mathrm{per}}(S) \le \infty$. Then for every $0 < \varepsilon \le h_{\mathrm{per}}(S) - h_t(S)$ there is a mixing locally compact Markov shift T with $h_t(T) = h_t(S) + \varepsilon$, $h_{\mathrm{per}}(T) = h_{\mathrm{per}}(S)$ and such that S embeds into T.*

Proof. Let $h = h_t(S) + \varepsilon$. Let $c = h_{\mathrm{per}}(S)$. If $c = \infty$ then Theorem 3.4 applies with $a_n = \infty$ for all n. Now suppose $c < \infty$, which implies $o_n(S) < \infty$ for all n. Let a_n be the least integer with $a_n - o_n(S) \ge e^{cn}$, $n \in \mathbb{N}$. Then $GR_n^-(a_n - o_n(S)) \ge h$ and with T from Theorem 3.4 both systems have finite orbit counts; thus $h_{\mathrm{orb}} = h_{\mathrm{per}}$ by Lemma 5.2; therefore $c = h_{\mathrm{per}}(S) = h_{\mathrm{orb}}(S) = GR_n(o_n(S)) = GR_n(o_n(S) + e^{cn}) = h_{\mathrm{orb}}(T) = h_{\mathrm{per}}(T)$. □

4. *An open problem on orbit counts of non-dense subsets*

Are the results on the set $ND(T)$ given by Theorem 2.2 the best possible if the only information one is allowed to use is the orbit counts and the entropy of the underlying system? We turn this vague question into a precise one, see Question 4.3 and Remark 4.4 below.

Let $(a_n)_{n\in\mathbb{N}}$ be a sequence with $a_n \in \{0, 1, 2, 3, \dots\} \cup \{\infty\}$ and let $0 < h \le \infty$. To increase readability, let us write (a_n) instead of $(a_n)_{n\in\mathbb{N}}$. The pair (a_n), h will be called compatible if $GR_n^-(a_n) \ge h$. By the Zeta Function Lemma 2.4 this ensures that the set $MS((a_n),\ h)$ of mixing locally compact Markov shifts T with $(o_n(T)) = (a_n)$ and $h_t(T) = h$ is non-empty. Recall the convention $\infty - \infty = \infty$.

Definition 4.1. For a compatible pair (a_n), h let us have

$$C((a_n), h) = \{(b_n) \mid \text{there are } T_1, T_2 \in MS((a_n), h) \text{ with } (b_n) \in ND(T_1),\ (b_n) \notin ND(T_2)\},$$

$$B((a_n), h) = \{(b_n) \mid 0 \le b_n \le a_n \text{ for all } n \text{ and } GR_n^-(a_n - b_n) = h \le GR_{\{n \mid a_n < \infty\}}(b_n)\}.$$

LEMMA 4.2. *Suppose (a_n), h is a compatible pair. Then:*

(a) *$C((a_n), h) \subset B((a_n), h)$;*

(b) *if $\{n \mid a_n < \infty\}$ is finite then $C((a_n), h) = B((a_n), h) = \emptyset$;*

(c) *if $\{n \mid a_n < \infty\}$ is infinite then $C((a_n), h)$ contains sequences even with $GR_n^-(b_n) \ge h$.*

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

Proof. (a) Let $(b_n) \in C((a_n), h)$, and T_1, T_2 be as in Definition 4.1. Now $(b_n) \in ND(T_1)$ implies $0 \leq b_n \leq a_n$ for all n and $GR_n^-(a_n - b_n) \geq h$ by Theorem 2.2(a), and $(b_n) \notin ND(T_2)$ implies $GR_{\{n|a_n<\infty\}}(b_n) \geq h$ and $GR_n^-(a_n - b_n) \leq h$ by Theorem 2.2(b) and (c). Thus $(b_n) \in B((a_n), h)$.

(b) If $\{n \mid a_n < \infty\}$ is finite then $GR_{\{n|a_n<\infty\}}(b_n) = 0$ for all (b_n); thus $B((a_n), h) = \emptyset$.

(c) Let $T_1 \in MS((a_n), h)$. Then $|I(T_1)| = \infty$. The proof of Theorem 2.7(a) produces a sequence (b_n) with $GR_n^-(b_n) \geq h$, $(b_n) \notin ND(T_1)$, and a $T_2 \in MS((a_n), h)$ with $(b_n) \in ND(T_2)$. □

Question 4.3. Does $C((a_n), h) = B((a_n), h)$ hold for all compatible pairs (a_n), h? Note that by the Zeta Function Lemma 2.4 $C((a_n), h) = B((a_n), h)$ holds if and only if for every $(b_n) \in B((a_n), h)$ there is a $T \in MS((a_n), h)$ with $(b_n) \notin ND(T)$.

Remarks 4.4.

(a) If (a_n), h are the orbit counts and entropy of a mixing SFT T_1 then $(b_n) \notin ND(T_1)$ for all $(b_n) \in B((a_n), h)$ by Lemma 1.4; thus in this case $C((a_n), h) = B((a_n), h)$.

(b) Replacing $h \leq GR_{\{n|a_n<\infty\}}(b_n)$ by the stronger condition $GR_n^-(a_n) \leq GR_{\{n|a_n<\infty\}}(b_n)$ does not change the set $B((a_n), h)$; see the proof of Corollary 2.3.

(c) If $h < GR_n^-(a_n)$ then $B((a_n), h) = \{(b_n) \mid 0 \leq b_n \leq a_n \text{ for all } n, GR_n^-(a_n - b_n) = h\}$.

5. *Counting periodic points versus counting periodic orbits*

We discuss conditions for embeddings in terms of strict inequalities of growth rates of periodic points or orbits. Suppose S and T are mixing locally compact Markov shifts with $h_t(S) < h_t(T)$ and $S \to_{\text{per}} T$. No strict inequality $h_1(S) < h_2(T)$, $h_1, h_2 \in \{h_{\text{orb}}^-, h_{\text{orb}}, h_{\text{per}}^-, h_{\text{per}}\}$, is necessary for the existence of an embedding of S into T, as Example 3.3(a) shows.

On the other hand, Theorem 3.2 shows that $h_{\text{orb}}(S) < h_{\text{orb}}^-(T)$ is sufficient, and so is $h_{\text{per}}(S) < h_{\text{orb}}^-(T)$ since $h_{\text{orb}}(S) \leq h_{\text{per}}(S)$ by definition. The following example finishes this discussion—no other strict inequality $h_1(S) < h_2(T)$, as for example $h_{\text{per}}(S) < h_{\text{per}}(T)$ or the even stronger $h_{\text{per}}(S) < h_{\text{per}}^-(T)$, is sufficient.

Example 5.1. Two mixing locally compact Markov shifts S and T with finite orbits counts, $h_t(S) < h_t(T)$, $S \to_{\text{per}} T$, $h_1(S) < h_2(T)$ for all pairs $(h_1, h_2) \notin \{(h_{\text{orb}}, h_{\text{orb}}^-), (h_{\text{per}}, h_{\text{orb}}^-)\}$, $h_{\text{orb}}(S) = h_{\text{per}}(S) = h_{\text{orb}}^-(T)$, and there is no embedding of S into T. Let

$$a_n = \begin{cases} 2^n & \text{if } \dfrac{n}{2} \text{ is an odd integer,} \\ 2^{4n} & \text{otherwise.} \end{cases}$$

By the Zeta Function Lemma 2.4 there is a mixing locally compact non-compact Markov shift T with $o_n(T) = a_n$ for all n and $h_t(T) < \log 2$. One calculates $h_{\text{orb}}^-(T) = \log 2$, $h_{\text{orb}}(T) = \log 16$ and $\#\text{Per}_n(T) \geq o_{n/2}(T) = 2^{2n}$ if $n/2$ is an odd integer and $\#\text{Per}_n(T) \geq o_n(T) = 2^{4n}$ otherwise; thus $h_{\text{orb}}^-(T) = \log 2 < \log 4 \leq h_{\text{per}}^-(T) \leq h_{\text{per}}(T) = h_{\text{orb}}(T) = \log 16$.

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

Let $R \subset T$ be a mixing SFT with $h_{\text{top}}(R) > 0$. Let n_i be a sequence such that $\mathbb{N} - \{n_i \mid i \in \mathbb{N}\}$ is infinite and $\lim_i n_i^{-1} \log o_{n_i}(T) = h^-_{\text{orb}}(T)$. Let

$$b_n = \begin{cases} o_n(T) & \text{if } n = n_i \text{ for some } i \in \mathbb{N}, \\ o_n(R) & \text{otherwise.} \end{cases}$$

Then $h_{\text{orb}}(R) = h^-_{\text{orb}}(R) \le h^-_{\text{orb}}(T)$ implies $GR_n(b_n) = h^-_{\text{orb}}(T)$.

Since $GR^-_n(b_n) \ge h_{\text{top}}(R) > 0$, by the Zeta Function Lemma there is a mixing locally compact Markov shift S with $h_t(S) < h_t(T)$ and $o_n(S) = b_n < \infty$ for all n. Thus, by Lemma 5.2, $h_{\text{per}}(S) = h_{\text{orb}}(S) = GR_n(b_n) = h^-_{\text{orb}}(T)$. Assuming that the n_i grow fast enough to ensure $h^-_{\text{per}}(S) \le h_{\text{per}}(R)$, we also obtain $\max\{h^-_{\text{orb}}(S), h^-_{\text{per}}(S)\} \le h_{\text{per}}(R) = h_t(R) \le h_t(T) < \log 2 = h^-_{\text{orb}}(T)$, and all the claimed inequalities follow.

Finally, $S \to_{\text{per}} T$ but $o_n(T) = o_n(S) < \infty$ for infinitely many n implies $GR^-_n(o_n(T) - o_n(S)) = 0$ and thus there is no embedding of S into T by Theorem 3.1. □

If T is a mixing SFT then it is well known that $h_t = h^-_{\text{orb}} = h^-_{\text{per}} = h_{\text{per}} = h_{\text{orb}}$ [**LM**]. The non-compact situation is more subtle. We mostly suppress the letter T in the following.

LEMMA 5.2. *Let T be a selfmap on a set X with $o_n(T) < \infty$ for all n (this holds whenever $h_{\text{per}} < \infty$). Then $h_{\text{orb}} = h_{\text{per}}$. If additionally $h^-_{\text{per}} = h_{\text{per}} < \infty$, then $h^-_{\text{orb}} = h^-_{\text{per}} = h_{\text{per}} = h_{\text{orb}}$.*

Proof. Let $o_n = o_n(T)$. We show $GR_n(o_n) \ge GR\big(n \cdot \sum_{i=1}^n o_i\big)$. Let $\log \lambda = GR_n(o_n)$ and $\varepsilon > 0$. Since $o_n < \infty$ for all n there is some $C > 0$ such that $o_i \le C(\lambda + \varepsilon)^i$ for all i. Thus,

$$n \cdot \sum_{i=1}^n o_i \le n \cdot C \cdot \sum_{i=1}^n (\lambda + \varepsilon)^i \le C \cdot n^2 (\lambda + \varepsilon)^n,$$

which implies $GR\big(n \cdot \sum_{i=1}^n o_i\big) \le \log(\lambda + \varepsilon)$. Now suppose $h^-_{\text{per}} = h_{\text{per}} = \log \lambda < \infty$. Since the case $h_{\text{per}} = 0$ is trivial we may assume $1 < \lambda < \infty$. Choose $\varepsilon > 0$ so small that $\lambda(1 - \varepsilon) > (\lambda + \varepsilon)^{1/2}$ and then C, N with $\text{per}_n(T) \ge (\lambda(1 - \varepsilon))^n$ for all $n \ge N$ and $\text{per}_n(T) \le C(\lambda + \varepsilon)^n$ for all n. For all $n \ge N$ we obtain

$$\begin{aligned} n \cdot o_n(T) &\ge \text{per}_n(T) - \sum_{j=1}^{n/2} \text{per}_j(T) \ge (\lambda(1 - \varepsilon))^n - \sum_{j=1}^{n/2} C(\lambda + \varepsilon)^j \\ &\ge (\lambda(1 - \varepsilon))^n - \frac{Cn}{2}(\lambda + \varepsilon)^{n/2}. \end{aligned}$$

By the choice of ε this implies $GR^-_n(o_n(T)) \ge \log(\lambda(1 - \varepsilon))$ and $h^-_{\text{orb}} = h_{\text{per}}$ follows with $\varepsilon \to 0$. □

LEMMA 5.3. *Suppose T is a mixing locally compact Markov shift. Then*

(1) $h_t \le h^-_{\text{orb}} \le \min\{h^-_{\text{per}}, h_{\text{orb}}\} \le \max\{h^-_{\text{per}}, h_{\text{orb}}\} \le h_{\text{per}}$, *with possibly* $h_{\text{orb}} < h^-_{\text{per}}$.

(2) *If $o_n(T) < \infty$ for all n (this holds whenever $h_{\text{per}} < \infty$), then*

(2a) $h_{\text{orb}} = h_{\text{per}}$, *and thus* $h^-_{\text{orb}} \le h^-_{\text{per}} \le h_{\text{per}} = h_{\text{orb}}$;

(2b) *there are examples with* $h_{\text{per}} < \infty$ *and* $h^-_{\text{orb}} < h^-_{\text{per}}$;

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

(2c) *there are examples with* $h^-_{\mathrm{per}} = h_{\mathrm{per}} = \infty$ *and* $h^-_{\mathrm{orb}} < h^-_{\mathrm{per}}$;

(2d) *if* $h^-_{\mathrm{per}} = h_{\mathrm{per}} < \infty$ *then* $h^-_{\mathrm{orb}} = h^-_{\mathrm{per}} = h_{\mathrm{per}} = h_{\mathrm{orb}}$, *with possibly* $h_t < h^-_{\mathrm{orb}}$.

(3) *If* T *has a rome, see Lemma 1.9, then* $h_t = h^-_{\mathrm{orb}} = h^-_{\mathrm{per}} = h_{\mathrm{per}} = h_{\mathrm{orb}}$, *with possibly* $h_t = \infty$.

Proof. (1) The first inequality follows from Remark 0.1, the others by definition. By the Zeta Function Lemma 2.4 there is a Markov shift T with $o_1(T) = \infty$ and $o_n(T) = 2^n$, $n \geq 2$, thus $h_{\mathrm{orb}} = \log 2 < \infty = h^-_{\mathrm{per}}$.

(2a) Lemma 5.2.

(2b) Example 5.1.

(2c) Take T as in Example 5.1, but with $a_n = 2^n$ if $n/2$ is an odd integer and $a_n = n^{4n}$ otherwise.

(2d) Lemma 5.2 and Example 3.3(b).

(3) This is Lemma 2.1 with $S = T$. It is easy to define a graph for a locally compact Markov shift with a vertex α as a rome and where the number of loops of length n at α grows superexponentially. □

6. *Proof of the Zeta Function Lemma 2.4*

Let $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ be sequences with $a_n, b_n \in \{0, 1, 2, 3, \dots\} \cup \{\infty\}$, $a_n \geq b_n$ for all n, and let $1 < \lambda \leq \infty$ such that $GR^-_n(d_n) \geq \log\lambda$ where $d_n = a_n - b_n$ with the convention $\infty - \infty = \infty$. We shall construct a mixing locally compact Markov shift T and a closed discrete shift-invariant subset $D \subset \mathrm{Per}(T)$ such that $o_n(T) = a_n$, $o_n(T, D) = b_n$ for all n and $h_t(T) = \log\lambda$. Inductively we will choose sequences $(N(i), l(i), k(i))_{i\geq 1} \in (\mathbb{N}^3)^{\mathbb{N}}$ with $N(1) < N(2) < \cdots$ such that the Markov shift T defined by the following graph G has the desired properties. The vertex set of G is $\mathbb{Z} \cup \bigcup_{\{i\geq 1 | l(i)\geq 2\}} \{(h, i) \mid 1 \leq h < l(i)\}$ and the set of directed edges is defined as follows.

- There is an edge from vertex n to vertex $n + 1$, for each $n \in \mathbb{Z}$.
- For each $i \in \mathbb{N}$ with $l(i) = 1$ there is a loop at vertex $N(i)$.
- For each $i \in \mathbb{N}$ with $l(i) > 1$ there is an edge from vertex $N(i)$ to vertex $(1, i)$, an edge from vertex (h, i) to vertex $(h + 1, i)$, $1 \leq h < l(i) - 1$, and an edge from vertex $(l(i) - 1, i)$ to vertex $N(i)$.
- For each $i \in \mathbb{N}$ there are $k(i)$ edges from vertex $N(i)$ to vertex $-N(i)$, for all $i \geq 1$.

We show how to choose the $N(i)$, $l(i)$, $k(i)$ such that the Markov shift T defined by the graph G has the desired properties. For that we fix some notation. Let $i \geq 1$. Let T_i be the SFT given by the finite subgraph G_i of G consisting of the connected component of the vertex 0 after removing the vertex $N(i)+1$. Let E_i denote the edge set of G_i. Let $C(n, T_i)$ denote the set of loops of length n in G_i at vertex 0 and let $C(n)$ denote the set of loops of length n in G at vertex 0.

Fix a strictly increasing sequence δ_i with $1 < \delta_i < \lambda$ for all i and $\lim_i \delta_i = \lambda$. Fix a surjective map $r : \mathbb{N} \to \mathbb{N}$ such that $\#r^{-1}(n) = \infty$ for all $n \in \mathbb{N}$.

We define $N(i), l(i), k(i)$, $i \geq 1$ inductively such that for all $i \geq 2$:

(1) $N(i) > \max\{l(k) \mid 1 \leq k \leq i\}$ and $N(i) > N(i - 1)$;

and for all $i \geq 1$:

(2) $\delta_i^n \leq a_n$ for all $n \geq N(i)$;

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

(3) $\#C(n, T_i) \le \min(\delta_i^n, d_n)$ for all n;
(4) $o_n(T_i) \le a_n$ for all n.

To get started we show that for N large enough the triple $(N(1), l(1), k(1)) = (N, 2N + 1, 1)$ satisfies conditions (2)–(4). Since $GR_n^-(a_n) \ge \log\lambda > \log\delta_1$, condition (2) holds for N large enough. Note that the SFT T_1 is a suspension of the 2-shift with $\#\mathrm{Per}_n(T_1) = 2^{n/(2N+1)}$ for all n divisible by $2N + 1$ and $\#\mathrm{Per}_n(T_1) = 0$ otherwise. Since $o_n(T_1)$ and $\#C(n, T_1)$ are both bounded by $\#\mathrm{Per}_n(T_1)$ and since $GR_n^-(a_n) > 1$, $GR_n^-(d_n) > 1$, $\delta_1 > 1$ it is easy to see that (3), (4) hold for N large enough.

Now assume we have defined $N(j)$, $l(j)$ and $k(j)$ for $1 \le j \le i$ such that (1)–(4) are satisfied. We define $N(i + 1)$, $l(i + 1)$ and $k(i + 1)$. Since $GR_n^-(a_n) \ge \log\lambda > \log\delta_i \ge h_{\mathrm{top}}(T_i) = GR_n(o_n(T_i))$ [**LM**], we can define

$$l(i + 1) := \min\{n \ge r(i + 1) \mid o_n(T_i) < a_n\}.$$

Obviously for all $N(i + 1)$ large enough the conditions (1),(2) are satisfied, where we can use $GR_n^-(a_n) \ge \log\lambda > \log\delta_i$ for (2). We will choose $k(i + 1) \in \mathbb{N}$ as a function of $N(i + 1)$, namely such that

$$k(i + 1) \le \delta_i^{2N(i+1)} < k(i + 1) + 1.$$

We now show that for $N(i + 1)$ large enough $\#C(n, T_{i+1}) \le \min(\delta_{i+1}^n, d_n)$ for all n. First, note that for $n < 2N(i + 1)$ we have $\#C(n, T_{i+1}) = \#C(n, T_i) \le \min(\delta_i^n, d_n)$. Thus, it suffices to show that for $N(i + 1)$ large enough, $\#C(n, T_{i+1}) \le \delta_{i+1}^n$ for all n.

Since $\#C(n, T_i) \le \delta_i^n$ for all $n \ge N(i)$ there is an $M_i > 1$ such that the number of paths in G_i of length n from vertex α to vertex β is at most $M_i\delta_i^n$ for all $n \ge 1$ for all vertices α, β of T_i.

We define a map $f_n : C(n, T_{i+1}) \to \{0, 1\}^n$. For $w = w_1 \cdots w_n \in C(n, T_{i+1})$ let $f_n(w)_j = 1$ if and only if $w_j \in E_i$ and $f_n(w)_j = 0$ if and only if $w_j \in E_{i+1} - E_i$, $1 \le j \le n$.

Let $b = b_1 \cdots b_n \in f_n(C(n, T_{i+1}))$. We estimate $\#f_n^{-1}(b)$. Let $w \in f_n^{-1}(b)$. If $b_j \cdots b_{j+m+1} = 10^m1$ then $m > 2(N(i + 1) - N(i))$ and $w_{j+1} \cdots w_{j+m}$ is one of $k(i + 1)$ paths which start in $N(i)$, move straight towards $N(i + 1)$, then loop $(m - 2(N(i + 1) - N(i)) - 1)/l(i + 1)$ times in the loop of length $l(i + 1)$ at this vertex, use one of the $k(i + 1)$ edges from vertex $N(i + 1)$ to vertex $-N(i + 1)$, and finally move to the vertex $-N(i)$. Thus, $\#\{w_{j+1} \cdots w_{j+m} \mid w \in f_n^{-1}(b)\} = k(i + 1)$ if $b_j \cdots b_{j+m+1} = 10^m1$.

Let $k = \#\{1 \le j \le n \mid b_j = 0\}$. If $k = 0$ then $w \in C(n, T_i)$ and thus $\#f_n^{-1}(b) = \#C(n, T_i) \le \delta_i^n$. If $k > 0$ then consider a maximal subblock $b_{j+1}, \ldots, b_{j+m}$ of b with $b_{j+k} = 1$, $1 \le k \le m$. Then $w_{j+1}, \ldots, w_{j+m}$ is a path in G_i from vertex α to vertex β, where $\alpha, \beta \in \{0, N(i), -N(i)\}$ are determined by b. Thus, there are at most $M_i\delta_i^m$ possible choices for $w_{j+1}, \ldots, w_{j+m}$. Now let $s = \#\{1 \le j \le n \mid b_j \cdots b_{j+m+1} = 10^m1$ for some $m\}$. Then $0 \le s \le n/2N(i + 1)$ and

$$\#f_n^{-1}(b) \le k(i + 1)^s \cdot M_i^s\delta_i^{n-k}.$$

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

Now $k \geq s \cdot 2(N(i+1) - N(i))$ and we have chosen $k(i+1) \leq \delta_i^{2N(i+1)}$; thus

$$\begin{aligned}\#f_n^{-1}(b) &< \delta_i^{s\cdot 2N(i+1)} M_i^s \delta_i^{n-s\cdot 2(N(i+1)-N(i))} \\ &= M_i^s \delta_i^{n+s\cdot 2N(i)} \\ &\leq (M_i^{1/(2N(i+1))} \delta_i^{1+(N(i)/N(i+1))})^n \quad \text{since } s \leq \frac{n}{2N(i+1)}.\end{aligned}$$

Since $M_i > 1$, this estimate holds for all $b \in f_n(C(n, T_{i+1}))$ and thus

$$\#C(n, T_{i+1}) \leq (M_i^{1/(2N(i+1))} \delta_i^{1+(N(i)/N(i+1))})^n \cdot \#\{f_n(w) \mid w \in C(n, T_{i+1})\}.$$

Let $\varepsilon_i > 0$ be so small that $\delta_i(1+\varepsilon_i)^2 < \delta_{i+1}$. Then, since $f_n(w)_1 = f_n(w)_n = 1$ and if $f_n(w)_j \cdots f_n(w)_{j+m+1} = 10^m 1$ implies $m > 2(N(i+1) - N(i))$, we can choose $N(i+1)$ so large that by simple standard estimates

$$\#\{f_n(w) \mid w \in C(n, T_{i+1})\} \leq (1+\varepsilon_i)^n \quad \text{for all } n$$

and that

$$M_i^{1/(2N(i+1))} \delta_i^{(N(i)/N(i+1))} \leq 1 + \varepsilon_i.$$

Thus, we get

$$\#C(n, T_{i+1}) \leq (\delta_i \cdot (1+\varepsilon_i)^2)^n < \delta_{i+1}^n \quad \text{for all } n.$$

Thus, condition (3) holds for $N(i+1)$ large enough. We show that condition (4) holds. Note that $o_n(T_{i+1}) = o_n(T_i)$ for all $n < N(i+1)$ with $n \neq l(i+1)$. Since $a_{l(i+1)} - o_{l(i+1)}(T_i) > 0$ by definition of $l(i+1)$ and since $N(i+1) > l(i+1)$ we get $o_{l(i+1)}(T_{i+1}) = o_{l(i+1)}(T_i) + 1 \leq a_{l(i+1)}$. By (3), for $n \geq N(i+1)$ each orbit in G_{i+1} of length n has to visit the vertex 0 at least once. Thus, $o_n(T_{i+1}) \leq C(n, T_{i+1}) \leq \delta_{i+1}^n \leq d_n \leq a_n$ for all $n \geq N(i+1)$, by (2).

Thus, G_{i+1} satisfies (1)–(4). This finishes the induction.

We show that $h_t(T) = \log \lambda$. By (1), $h_{\text{top}}(T_i) \leq \log \delta_i \leq \log \lambda$ for each i; thus $h_t(T) = h_G(T) = \sup_i h_{\text{top}}(T_i) \leq \log \lambda$ [**G1, FF1**]. On the other hand, by construction $\#C(2N(i)+1) \geq \#C(2N(i)+1, T_i) \geq k(i) > \delta_{i-1}^{2N(i)} - 1$ for all $i \geq 2$, and thus $GR_n(\#C(n)) \geq \log \lambda$ since $\delta_i \to \lambda$. Thus, $h_t(T) \geq \log \lambda$.

We show that $o_n(T) = a_n$ for all n. Since $o_n(T_i) \leq a_n$ for all i and $\lim_{i\to\infty} o_n(T_i) = o_n(T)$ we obtain $o_n(T) \leq a_n$. Assume there is an n with $o_n(T) < a_n$. Then in particular $o_n(T) < \infty$. Then for each $i \geq 2$ with $r(i) = n$ it holds that $o_n(T_{i-1}) \leq o_n(T) < a_n$ and thus, by definition of $l(i)$, at vertex $N(i)$ is a simple loop of length n. Since $\#r^{-1}(n) = \infty$ we get $o_n(T) = \infty$, a contradiction. Thus $o_n(T) = a_n$ for all n.

Let $D' := \{x \in \text{Per}(T) \mid i(x_m) \neq 0 \text{ for all } m\}$, that is D' is the set of periodic points which do not visit the vertex 0. If $a_n = \infty$ then, since $o_n(T_{i-1}) < \infty$ for all i, for each i with $r(i) = n$ there is a loop of length $n = l(i)$ at vertex $N(i)$. Thus $o_n(T, D') = \#r^{-1}(n) = \infty \geq b_n$. If $a_n < \infty$ then there is an i such that $o_n(T_i) = a_n$, since $o_n(T) = a_n$ by the above. Then $o_n(T, D') \geq o_n(T_i, D' \cap T_i) \geq o_n(T_i) - \#C(n, T_i) \geq a_n - d_n = b_n$ by (1). Thus $o_n(T, D') \geq b_n$ for all n. Thus, we can choose a shift-invariant subset $D \subset D'$ such that $o_n(T, D) = b_n$ for all n. Then D is closed and discrete since the loops that define its orbits use disjoint sets of edges. □

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

7. *Proof of the Embedding Theorem 1.2*

Proof of the Periodic Point Lemma 1.3. Fix a sequence $c_i \in \mathbb{N}$, $i \in \mathbb{N}$ such that $\#\{i \mid c_i = n\} = o_n(T, D)$ for all n. Let $\mathrm{per}(p)$ denote the least period of a periodic point p. Assume we have shown the following.

CLAIM. *Given $p \in D$ there is a mixing locally compact Markov shift $S \subset T$ such that S is closed in T with $p \notin S$ and a non-dense subset $C \subset S$ such that $o_n(C) = o_n(D)$ for all $n \neq \mathrm{per}(p)$ and $o_{\mathrm{per}(p)}(C) = o_{\mathrm{per}(p)}(D) - 1$.*

Then, starting with $T_1 = T$, $D_1 = D$ and a point $p_1 \in D_1$ with $\mathrm{per}(p_1) = c_1$ one obtains S and C from the above claim; let $T_2 = S$, $D_2 = C$ and choose $p_2 \in D_2$ with $\mathrm{per}(p_2) = c_2$ and apply the above claim to obtain T_3, D_3, etc. This produces sequences of mixing locally compact Markov shifts $T_1 \supset T_2 \supset \cdots$ which are closed in T, of subsets $D_i \subset T_i$, and of periodic points p_i.

Let D' be the union of the orbits of the p_i, $i \in \mathbb{N}$ obtained in this way. By construction $o_n(D') = o_n(D)$ for all n. Moreover, D' is discrete since every point $p \in D'$ belongs to some $T - T_n$, which is open in T, and $D' \cap (T - T_n)$ is finite by construction.

Proof of the claim. Let T be given in graph presentation, let $n = \mathrm{per}(p)$. Fix a cylinder $Z = [z_0, \ldots, z_{m-1}]_{m-1} \subset T$ with $Z \cap D = \emptyset$; without loss of generality $m > n$. Passing to the higher $2m$-block system of T we obtain a graph presentation of T such that:

(1) p is given by a simple loop $w = w_0 w_1 \cdots w_{n-1}$ at some vertex β, i.e. a loop that does not visit a vertex twice, except for β;

(2) Z is given by some vertex α, i.e. by the set of paths such that α is the initial vertex of their time-zero edge.

Now fix some shortest loop u at β that runs through the vertex α. Let $k = |u|$ denote the length of u. Let $W = w^k$, $U = u^n$. Then $|W| = nk = |U|$.

We define a (generally non-continuous) map $\pi : D - O(p) \to T$. Given $x \in D - O(p)$ call $[i, j]$ a maximal w-interval if $x[i, j] = w^t$ for some $t \geq 1$ and $x[i - n, j] \neq w^{t+1}$, $x[i, j + n] \neq w^{t+1}$. For each maximal w-interval of x of length $l = j - i + 1 \geq 2nk$ write $l = s(2nk) + rn$ with $0 \leq r \leq 2k$ and replace the block $x[i, j] = W^{2s} w^r$ by $(UW)^s w^r$; no other action takes place.

Let $C = \pi(D - O(p))$. Then $C \subset \mathrm{Per}(S)$ since π is shift-invariant. It is not hard to see that π has an inverse which is given by replacing each occurrence of the block U by W. Thus π is an injection and thus $o_n(C) = o_n(D)$ for all $n \neq \mathrm{per}(p)$ and $o_{\mathrm{per}(p)}(C) = o_{\mathrm{per}(p)}(D) - 1$.

Let $S^{(m)}$ be obtained from T by forbidding the block w^m, $m \in \mathbb{N}$. Each $S^{(m)}$ is a locally compact Markov shift and closed in T, since it is given by a subgraph of a higher block presentation of T. For all large enough m the Markov shift $S^{(m)}$ is mixing by standard arguments [**LM**, proof of Lemma 10.1.3], and $C \subset S^{(m-1)}$ for all $m > 3k$. Finally, $S^{(m)} - S^{(m-1)}$ is open in $S^{(m)}$ and non-empty since it contains $(w^m u)^\infty$ thus C is non-dense in $S = S^{(m)}$. □

We now prepare for the proof of the Embedding Theorem 1.2. We will need a modified version of a quite technical, but extremely helpful tool from [**FF1**], the Marker Lemma.

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

Let $S \subset \mathbb{N}^{\mathbb{Z}}$ be a locally compact subshift. Given a set $M \subset S$ and a function $L : \mathbb{N} \to \mathbb{N}$ we use the following suggestive terms. Let $x \in S$. Let

$$M(x) := \{i \in \mathbb{Z} \mid S^i x \in M\}.$$

A *marked interval* of x is a set

$$[i, j] \subset \mathbb{Z} \quad \text{such that } i \in M(x), k \notin M(x) \text{ for } i < k < j, j \in M(x).$$

A marked interval $[i, j]$ of x is *long* if $j - i \geq \max\{L(x_i), L(x_j)\}$ and it is *short* if $j - i < \max\{L(x_i), L(x_j)\}$. If $i \in M(x)$ and $k \notin M(x)$ for $k < i$ then $(-\infty, i]$ is a *marked ray* of x. Analogously, $[i, \infty)$ is a marked ray of x if $i \in M(x)$, $k \notin M(x)$ for $i < k$.

A new feature compared to [**FF1**] is the following. Let $\mathcal{B}_n(S)$ denote the blocks of length n occurring in points of S. Given M, L as above and a function $R : \bigcup_n \mathcal{B}_n(S) \to \mathbb{N}$ we will call $[i, j] \subset \mathbb{Z}$ a *long run of short marked intervals* of x if there are $k, r \geq 1$ such that $j = i + kr$, $[i + k(n-1), i + kn]$ is a short marked interval for x for all $1 \leq n \leq r$, the block $p := x[i + k(n-1), i + kn)$ is the same for all $1 \leq n \leq r$, and $r \geq 3R(p)$. Then each $[i + k(n-1), i + kn]$ is called a *short marked interval contained in a long run*. Note that a maximal run might also be (one-sided or two-sided) infinite.

MARKER LEMMA. *Let $S \subset \mathbb{N}^{\mathbb{Z}}$ be a locally compact subshift with compact zero-partition and $h_t^{\text{pres}}(S) < \infty$. For $n, m, j \in \mathbb{N}$ let $C(n, m, j) = \{x[1, j] \mid x_1 = n, x_j = m\}$. Let $\lambda > 1$ be such that $\log \lambda > h_t^{\text{pres}}(S)$ and let $a : \mathbb{N} \to \mathbb{N}$. Then there are non-decreasing functions $L : \mathbb{N} \to \mathbb{N}$ and $b : \mathbb{Z}^+ \to \mathbb{Z}^+$ with:*

(1) *$3a(k) \leq L(k)$ for all $k \in \mathbb{N}$, $2b(r) \leq r$ for all $r \in \mathbb{Z}^+$, $\lim_r b(r) = \infty$;*

(2) *$\#C(n, m, L(k) + r) \leq \lambda^{L(k)+r-1-2a(k)-2b(r)}$ for all $n, m \in \mathbb{N}$, $r \geq 0$, where $k := \max\{n, m\}$.*

Moreover, given $R : \bigcup_n \mathcal{B}_n(S) \to \mathbb{N}$, there is a closed–open set $M \subset S$ such that:

(3) *for all $x \in S$ it holds that $M(x) \neq \emptyset$;*

(4) *if $[i, j]$ is a short marked interval for $x \in S$, then there is $z \in \text{Per}(S)$ with least period $j - i$ and $z[-6L(x_i) + i, j + 6L(x_i)] = x[-6L(x_i) + i, j + 6L(x_i)]$ and $[i, j]$ is contained in a long run of short marked intervals.*

Remark. If the function a is strictly increasing (as it will be in the upcoming construction), then

(5) for each N_2 there is N_1 such that $L(k) + r \geq N_1$ implies $a(k) + b(r) \geq N_2$.

This can be seen as follows. Fix k_0 such that $a(k_0) \geq N_2$. Choose N_1 such that $b(N_1 - L(k_0)) \geq N_2$. If $L(k) + r \geq N_1$, then either $k \geq k_0$, and we are done, or $k < k_0$ which implies $r \geq N_1 - L(k) \geq N_1 - L(k_0)$ and thus $a(k) + b(r) \geq b(r) \geq N_2$.

Proof. (1), (2) For the construction of L and b we follow the arguments in the proof of the Marker Lemma [**FF1**]. Let $1 < \mu < \lambda$ be such that $\log \mu > h_t^{\text{pres}}(S)$. For any pair i, j it holds that $GR_n(\#C(i, j, n)) \leq h_t^{\text{pres}}(S) < \log \mu$; thus there is an $N(i, j)$ such that $\#C(i, j, n) < N(i, j) \cdot \mu^n$ for all n. Let $N(k) := \max\{N(k, m), N(m, k) \mid m \leq k\}$, $k \in \mathbb{N}$. Choose L with $3a(k) \leq L(k) \leq L(k + 1)$ for all $k \in \mathbb{N}$ and $\lambda^{L(k)-2a(k)-1} \geq N(k)\mu^{L(k)}$.

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

Choose b with $b(r) \leq b(r+1)$ and $2b(r) \leq r$ for all $r \in \mathbb{Z}^+$ and $\lim_r b(r) = \infty$ and $\lambda^{r-2b(r)} \geq \mu^r$ for all r. Then, with $k = \max\{n, m\}$ we obtain

$$\begin{aligned}\#C(n, m, L(k)+r) &\leq N(n,m)\mu^{L(k)+r} \leq N(k)\mu L(k)\mu^r \\ &\leq \lambda^{L(k)-2a(k)-1}\lambda^{r-2b(r)} = \lambda^{L(k)+r-1-2a(k)-2b(r)}.\end{aligned}$$

(3) Given L and b the Marker Lemma from 4.2 [**FF1**] produces a compact open set M' with the properties that $M'(x)$ is non-empty for all $x \in S$ and

(*) if $[i, j]$ is a short M' marked interval for $x \in S$, then there is $z \in \mathrm{Per}(S)$ with least period $j - i$ and $z[-6L(x_i) + i, j + 6L(x_i)] = x[-6L(x_i) + i, j + 6L(x_i)]$.

We note two consequences of (*). If $[i, j]$ is a short M' marked interval of a point x then $x_i = x_j$. This shows that $L(x_i) = L(x_j)$ and thus $j < i + L(x_i)$ (**).

Another feature, implied by (*), is that any two consecutive short M' marked intervals have the same length and see the same block, i.e. whenever $[i, j]$ and $[j, k]$ are short marked intervals of a point x then $k - j = j - i$ and $x[i, j] = x[j, k]$ (***). Now let

$$\begin{aligned}M := \{x \in M' \mid\ &\text{if } [0, r] \text{ is a short } M' \text{ marked interval for } x \text{ for some } r > 0 \\ &\text{then } [0, r] \text{ is contained in a long run of short } M' \text{ marked intervals for } x\}.\end{aligned}$$

We show that M is closed–open. Let $h' : S \to \{0, 1\}^{\mathbb{Z}}$ be defined by $(h'x)_i = 1$ if and only if $i \in M'(x)$. Since M' is closed–open, h' is continuous. Given $x \in M' - M$ then there is $r > 0$ such that $[0, r]$ is a short M' marked interval not contained in a long run of short M' marked intervals. With $N = r \cdot R(x[0, r))$ this is determined by $x[-3N, 3N]$ and $h'(x)[-3N, 3N]$. Therefore $M' - M$ is open; thus M is closed. Now let $x \in M$. Then $x \in M'$. If there is no $r > 0$ such that $[0, r]$ is a short M' marked interval then this is determined by $x[0, L(x_0)]$ and $h'(x)[0, L(x_0)]$, see (**). If there is such an r then $[0, r]$ is contained in a long run of short M' marked intervals. Again, with N as above, this is both determined by $x[-3N, 3N]$ and $h'(x)[-3N, 3N]$. Thus M is open.

We show that $M(x) \neq \emptyset$ for all $x \in S$. If $M(x) = M'(x)$ this follows since $M'(x) \neq \emptyset$. If $M(x) \neq M'(x)$ then there are $i \in M'(x) - M(x)$ and $r > 0$ such that $[i, i + r]$ is a short M' marked interval, which is not contained in a long run. Then there is a $j \geq i + r$ such that $[i, j]$ is a run of short M' marked intervals but $[i, j + r]$ is not; thus $j \in M'(x)$ and by (***) there is no short M' marked interval to the right of j; thus $j \in M(x)$ by definition. This shows $M(x) \neq \emptyset$.

(4) Let $[i, j]$ be a short M marked interval for x. We first show that $[i, j]$ is a short M' marked interval for x, too. Assume there is $i < k < j$ such that $k \in M'(x)$. Choose such k minimal. Since $k \notin M(x)$ and by (***) $[k, j]$ is a run of short M' marked intervals, $x_k = x_j$ by (**). Thus $[i, k]$ is a short M' marked interval; again (***) shows that $[i, j]$ is a run of short M' marked intervals with $i \in M(x)$ and $k \notin M(x)$, a contradiction. Thus $[i, j]$ is a short M' marked interval for x; the existence of the point z claimed follows by (*). Since it is also an M marked interval, it is contained in a long run of short M' marked intervals, which are all M marked by definition. Thus $[i, j]$ is contained in a long run of M marked short intervals. □

Proof of the Embedding Theorem 1.2. It remains to show that (b) implies (a). We show that we may assume that T is non-compact. Suppose T is compact and $D \subset T$ a non-dense

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

set such that $S \to_{\text{per}} D$. Choose a periodic orbit O with $O \cap D = \emptyset$. Let $T' = T - O$. Then T' is a locally compact Markov shift by [**FF2**, Corollary 1.11]. Since T' is open in T, T' is also mixing. Since $T \neq T'$, T' is not compact. Since T' is dense in T, the set D is non-dense in T' and $h_t(T') = h_t(T)$ by Lemma 1.1. Replace T by T'.

1. Preparations for the definition of $g : S \to T$. Fix $1 < \lambda < \infty$ with $h_t(S) < \log\lambda < h_t(T)$. Then by [**FF1**, Lemma 2.6] we may assume that S is given in a presentation as a subshift of $\mathbb{N}^{\mathbb{Z}}$, with compact zero-partition and $h_t^{\text{pres}}(S) < \log\lambda$.

Let T be given by a graph $G = (V, E)$ with vertex set V and edge set E. Thus $T \subset E^{\mathbb{Z}}$. Since T is not compact we may assume that $E = \mathbb{N}$. By (b) and the Periodic Point Lemma 1.3 there is a discrete subset $D \subset T$ such that $S \to_{\text{per}} D$. Fix a shift-commuting injection $f : \text{Per}(S) \to \text{Per}(T)$ with $f(\text{Per}(S)) \subset D$. Since $f(\text{Per}(S))$ is not dense, there is a cylinder set $C \subset T$ such that $C \cap f(\text{Per}(S)) = \emptyset$. Replacing T by a suitable higher block presentation, we may assume that there is a vertex $\alpha \in V$ such that $x \in f(\text{Per}(S))$ implies that x_0 does not begin in vertex α.

Having fixed λ we define suitable functions a, R to apply the Marker Lemma. On our way we also fix some coding structures for later use.

Let W be the SFT given by all infinite concatenations of the first return loops at α of length less than or equal to K, with K as follows. Since T is mixing and since $\log\lambda < h_t(T)$, there is a K such that there are two first return loops at α with coprime lengths less than K, and such that the number of loops of length K at α is strictly larger than λ^K. Let $C_{\alpha,n}$ denote the set of loops in W of length n at vertex α. Concatenating the two first return loops of coprime length with repeated loops of length K shows that there is some N_0 with

$$|C_{\alpha,n}| \geq \lambda^n \quad \text{for all } n \geq N_0.$$

For each $n \geq N_0$ we fix a block $t_n \in C_{\alpha,n}$. These will be used later on to build transitions.

Now fix five distinct first return loops, say m^*, A, B, C, D, at α in T each of a length larger than K. Then none of these loops occurs in a point of W. We shall use the loop m^* as a marker and the loops A, B, C, D for encoding certain blocks and rays of S using the maps q and q' on $\mathbb{N}$ given by

$$q(n) := AB^nA \quad \text{and} \quad q'(n) = CD^nC \quad \text{for } n \in \mathbb{N}.$$

Define $a : \mathbb{N} \to \mathbb{N}$ by $a(k) = |m^*| + |q(k)| + N_0$, $k \in \mathbb{N}$.

For every $p \in \mathcal{B}_n(S)$, $n \in \mathbb{N}$ with the property that there is a periodic point $y \in S$ with least period n and $y[0, n) = p$ we fix $R = R(p) \in \mathbb{N}$ such that:

(1) $R > \max(K, |m^*|, |A|, |B|, |C|, |D|)$;

(2) $R \cdot |p| > |m^*q'(y_0) \cdots q'(y_n)| + \max\{|w(\alpha, \beta)|, |w(\beta, \alpha)|\} + N_0$,

where $\beta = \beta(p)$ is the initial vertex of $(fy)_0$. Extend R arbitrarily to a function $R : \bigcup_n \mathcal{B}_n(S) \to \mathbb{N}$.

Since S has compact zero partition and $h_t^{\text{pres}}(S) < \log\lambda < \infty$ we can apply the Marker Lemma with λ and the functions $a : \mathbb{N} \to \mathbb{N}$, $R : \bigcup_n \mathcal{B}_n(S) \to \mathbb{N}$ to obtain functions L, b and a marker set M satisfying (1)–(5).

2. Defining a map $g : S \to E^{\mathbb{Z}}$. For each $x \in S$ the index set $\mathbb{Z}$ splits into maximal long runs of short marked intervals (which might be infinite), long marked intervals, and marked rays.

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

Not actually needed, but helpful to clarify the picture, is the observation that whenever a long run of short marked intervals terminates to the left or right, it will be followed by a long marked interval or a marked ray, and not by another long run. This is a consequence of (4) of the Marker Lemma.

Definition of g on long runs of short marked intervals. First fix for each pair of vertices $\beta, \gamma \in V$ a shortest path $w(\beta, \gamma)$ from β to γ. Now, if $[i, j]$ is a short marked interval of x, for some $i < j$, then let $p = x[i, j)$ and $R = R(p)$. Let $y \in \mathrm{Per}(S)$ with $y[i, j) = p$ and $S^{|p|}y = y$ and let β be the initial vertex of $(fy)_i$. Choose $0 \le R', R'' \le R$ maximal such that $[-R' \cdot |p| + i, j + R'' \cdot |p|]$ is a run of length $R' + R'' + 1$ for x. Since $[i, j]$ is a short marked interval of x, by (4) of the Marker Lemma $R' = R$ or $R'' = R$.

- If $R' = R$ and $R'' = R$ then let $gx[i, j) = fy[i, j)$.
- If $R' < R$ then by (2) there is a unique $a > N_0$ such that the block $m^* q'(x_i) \cdots q'(x_j) t_a w(\alpha, \beta)$ has length $R \cdot |p|$. Write this block as $w_1 w_2 w_3$ where $|w_1| = R' \cdot |p|$, $|w_2| = |p|$. Let $gx[i, j) = w_2$.
- If $R'' < R$ then by (2) there is a unique $b > N_0$ such that the block $w(\beta, \alpha) t_b q'(x_i) \cdots q'(x_j) m^*$ has length $R \cdot |p|$. Write this block as $w_1 w_2 w_3$ where $|w_3| = R'' \cdot |p|$, $|w_2| = |p|$. Let $gx[i, j) = w_2$.

Note that if $0 \le R_1, R_2 \le \infty$ are maximal such that $[i', j'] := [-R_1 \cdot |p| + i, j + R_2 \cdot |p|]$ is a run of length $R_1 + R_2 + 1$ for x, then by the Marker Lemma $R_1 + R_2 + 1 \ge 3R(p)$. If $R_1 + R_2 + 1 < \infty$ then

$$gx[i', j'] = Q'_{\mathrm{left}} w(\alpha, \beta) fy[i' + R \cdot |p|, j' - R \cdot |p|) w(\beta, \alpha) Q'_{\mathrm{right}}.$$

where $Q'_{\mathrm{left}} = Q'_{\mathrm{left}}(x[i, j)) = m^* q'(x_i) \cdots q'(x_j) t_a$ as defined above and $Q'_{\mathrm{right}} = Q'_{\mathrm{right}}(x[i, j)) = t_b q'(x_i) \cdots q'(x_j) m^*$ are loops at α which visit α in every time interval of length $\max(K, |m^*|, |C|, |D|)$. Note that $w(\alpha, \beta) fy[i' + R \cdot |p|, j' - R \cdot |p|) w(\beta, \alpha)$ is a first return loop at α of length larger than $\max(K, |m^*|, |A|, |B|, |C|, |D|)$.

If $R_1 = R_2 = \infty$, then $x = y$ and $gx = fy \in T$. Note that in this case gx never visits the vertex α. If $R_1 = \infty$, $R_2 < \infty$, then

$$gx(-\infty, j + R_2 \cdot |p|) = fy(-\infty, i) w(\beta, \alpha) Q'_{\mathrm{right}}.$$

Similarly for $R_1 < \infty$, $R_2 = \infty$. In the last two cases the image ray visits α only finitely many times.

Definition of g on long marked intervals. First fix for all $n, m \in \mathbb{N}$, $r \ge 0$, with $k = \max\{n, m\}$, and $N = L(k) + r - 1 - 2a(k) - 2b(r)$ injective functions

$$b_{n,m,r} : C(n, m, L(k) + r) \to C_{\alpha,N}$$

where $C(n, m, j) = \{x[1, j] \mid x \in S \text{ with } x_1 = n,\ x_j = m\}$. These exist since $N \ge a(k) - 1 \ge N_0$; thus $|C_{\alpha,N}| \ge \lambda^N$ and (2) of the Marker Lemma applies.

Now if $[i, j]$ is a long marked interval for x, then let $k = \max(x_i, x_j)$, $r = 1 + j - i - L(k)$ and $n = a(k) + b(r)$. We call n the boundary length for the marked interval $[i, j]$.

Let $0 \le a < j - i$ be maximal with $|m^* q(x_i) \cdots q(x_{i+a})| \le n - N_0$. Then there is a unique $b > 0$ such that $|m^* q(x_i) \cdots q(x_{i+a}) t_b| = n$. Let $Q_{\mathrm{left}} := m^* q(x_i) \cdots q(x_{i+a}) t_b$. Let $0 \le c < j - i$ be maximal with $|q(x_{j-c}) \cdots q(x_j) m^*| \le n - N_0$. Then there is a

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

unique $d > 0$ such that $|t_d q(x_{j-c}) \cdots q(x_j) m^*| = n$. Let $Q_{\text{right}} := t_d q(x_{j-c}) \cdots q(x_j) m^*$. Then let

$$gx[i, j) := Q_{\text{left}} b_{x_i, x_j, r}(x[i, j]) Q_{\text{right}}.$$

Note that $gx[i, j)$ is a loop at vertex α which visits α in each time interval of length $\max(K, |m^*|, |A|, |B|)$.

Definition of g on marked rays. If $[i, \infty)$ is a marked ray of x then let $gx[i, \infty)$ be the right-infinite path in G starting in vertex α defined by

$$gx[i, \infty) = m^* q(x_i) q(x_{i+1}) q(x_{i+2}) \cdots .$$

If $(-\infty, i]$ is a marked ray of x then let $gx(-\infty, i)$ be the left-infinite path in G ending at vertex α defined by

$$gx(-\infty, i) = \cdots q(x_{i-2}) q(x_{i-1}) q(x_i) m^*.$$

This defines a shift-commuting map $g : S \to E^{\mathbb{Z}}$.

For later use, when we show injectivity of g and continuity of g^{-1}, we note that the code is made such that whenever $z = gx$ for some x and

(g1) $z[i, \infty)$ starts with $m^* q(n_1) q(n_2) \ldots q(n_k)$, then $x[i, i + k) = n_1 \cdots n_k$;

(g2) $z[i, \infty)$ starts with $m^* q'(n_1) q'(n_2) \cdots q'(n_k) E$ where E is a first return loop at α different from C, then either there is an $i < j$ such that $[i, j]$ is a finite long run of short marked intervals for x, or $[i, \infty)$ is an infinite run of short marked intervals for x, and for the unique periodic point w with least period $k - 1$ and $w[i, i + k - 1) = n_1 \cdots n_{k-1}$ it holds that $x[i, j] = w[i, j]$ or $x[i, \infty) = w[i, \infty)$, respectively.

3. Verification that the map g is an embedding from S into T. We first show that g maps S into T, then that g is continuous and finally with one argument we show that g is injective and g^{-1} is continuous.

To see that $g(S) \subset T$ recall that $\mathbb{Z}$ splits into a union of marked rays, long marked intervals and maximal runs of short marked intervals. If $[i, j]$ is a finite maximal run of short marked intervals or a long marked interval then $gx[i, j)$ is a loop at α. If $[i, \infty)$ (or $(-\infty, i]$) is a maximal run of short marked intervals or a marked ray then $gx[i, \infty)$ (respectively $gx(-\infty, i]$) is a path in G starting (ending) at α. If $(-\infty, \infty)$ is a run of short intervals then $gx = fx$. In any case, $gx \in T$.

We now show the continuity of g. Let $h : S \to \{0, 1\}^{\mathbb{Z}}$ be defined by $(hx)_i = 1$ if and only if $i \in M(x)$. Since $M(x)$ is closed–open, h is continuous. Thus it suffices to show that $g(x)_0$ is determined by $x[-N, N]$ and $h(x)[-N, N]$ for some $N = N(x) \geq 0$.

If there is $i \leq 0 < j$ such that $[i, j]$ is a long marked interval of x then let $N = \max\{-i, j\}$. If there is $i \leq 0 < j$ such that $[i, j]$ is a short marked interval of x then let $p = x[i, j)$ and $N = (R(p) + 1) \cdot |p|$.

If there is $i \leq 0$ such that $[i, \infty)$ is a marked ray of x then there is $j > 1 - i$ such that $gx[i, i + j) = m^* q(x_i) q(x_{i+1}) \cdots q(x_0)$ and thus $gx[i, 0]$ is the prefix of length $1 - i$ of this block. By (5) of the Marker Lemma there is some $N > L(x_i)$ such that whenever $[i, i+N']$ is a long marked interval for some point y with $N' \geq N$, then the boundary length $a(k) + b(r)$ for the marked interval $[i, i + N']$ of y is at least $|gx[i, i + j)| + N_0$. Let $y \in S$ with $hy[i, i + N] = hx[i, i + N]$ and $y[i, i + N] = x[i, i + N]$. If there is no $N' > 0$ such

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

that $[i, i+N']$ is a marked interval for y then $[i, \infty)$ is a marked ray for y and $gy[i, \infty) = m^*q(y_i)q(x_{i+1})\cdots q(y_0)\cdots$ and thus $(gy)_0 = (gx)_0$. If there is some $N' > 0$ such that $[i, i+N']$ is a marked interval for y then, since $N' \geq N > L(x_i) = L(y_i)$, $[i, i+N']$ is a long marked interval for y with boundary length $a(k) + b(r)$ at least $|gx[i, i+j)| + N_0$. Since $x[i, 0] = y[i, 0]$ we get $gy[i, i+j) = m^*q(y_i)q(y_{i+1})\cdots q(y_0) = gx[i, i+j)$. Thus, in particular, $(gy)_0 = (gx)_0$.

The case that there is $i > 0$ such that $(-\infty, i]$ is a marked ray for x is symmetric. Since $M(x) \neq \emptyset$ this finishes the proof that g is continuous.

We show injectivity of g and continuity of g^{-1}.

Let $x \in S$ and $gx = z \in g(S)$. We show

(*) there is some $N = N(z) \geq 0$ such that $y \in S$ with $gx[-N, N] = gy[-N, N]$ implies $x_0 = y_0$.

We split the proof of (*) into several cases.

Case 1: The block m^*m^* does not occur in z. Then there are $r \geq 1$ and $i \in \mathbb{Z}$ such that $[n \cdot r + i, (n+1) \cdot r + i]$ is a short marked interval for x for every $n \in \mathbb{Z}$. Thus $x \in \mathrm{Per}(S)$ and $z = gx = fx$. Let $p = x[i, i+r)$. Since $f(\mathrm{Per}(S))$ is discrete there is an $N_1 > \max(K, |m^*|, |A|, |B|, |C|, |D|)$ such that

(**) $[z_{-N_1}, \ldots, z_{N_1}]_{N_1} \cap f(\mathrm{Per}(S)) = \{z\}$.

Now consider $y \in S$ with $gy[-N_1 - r, N_1 + r] = z[-N_1 - r, N_1 + r]$. Since this block does not see the vertex α and $N_1 > \max(K, |m^*|, |A|, |B|, |C|, |D|)$ we know by the construction of g that $gy[-N_1 - r, N_1 + r]$ has to be contained in some block of the form $w(\alpha, \beta)uw(\beta, \alpha)$ where $u = (fx[i, i+r))^n$ for some n and β is the initial vertex of $(fx)_i$. Since $w(\alpha, \beta)$ and $w(\beta, \alpha)$ are shortest paths from α to β (respectively from β to α) we get that $gy[-N_1, N_1]$ is contained in u. Thus, by construction, $gy[-N_1, N_1] = fw[-N_1, N_1]$ for some $w \in \mathrm{Per}(S)$ with $y[-N_1, N_1] = w[-N_1, N_1]$. Now $fw \in [z_{-N_1}, \ldots, z_{N_1}]_{N_1} \cap f(\mathrm{Per}(S))$ implies $fw = z = fx$ by (**). Since f is injective this shows that $w = x$ and thus $y_0 = w_0 = x_0$.

Case 2: $z[i - |m^*|, i + |m^*|) = m^*m^*$ for some $i \in \mathbb{Z}$. Fix such i with $|i|$ minimal. Assume that $i \leq 0$ (the argument for $i > 0$ will be symmetric). This implies $i \in M(x)$ by construction.

Case 2.1: $z[i, \infty)$ begins with $m^*q'(x_i)$. Then there is a finite or infinite run of short marked intervals for x starting at i (since $z[i, \infty)$ begins with $m^*q'(x_i)$ and not with $m^*q(x_i)$.) In any case $z[i, \infty)$ begins with some block $m^*q'(x_i)q'(x_{i+1})\cdots q'(x_{i+r})t_n$. Let $N = |m^*q'(x_i)q'(x_{i+1})\cdots q'(x_{i+r})t_n|$ and consider $y \in S$ with $gy[i, i+N) = z[i, i+N)$. Then $gy[i, i+N) = m^*q'(x_i)q'(x_{i+1})\cdots q'(x_{i+r})t_n$ and the property (g2) applies to both x and y and shows $y_0 = x_0$.

Case 2.2: $z[i, \infty)$ begins with $m^*q(x_i)$. There are two possibilities. First, suppose there is an $i < j$ such that $[i, j]$ is marked for x. Then $[i, j]$ is a long marked interval of x (since $z[i, \infty)$ begins with $m^*q(x_i)$ and not with $m^*q'(x_i)$). Let $N = j$ and consider $y \in S$ with $gy[-N, N] = z[-N, N]$. By the minimality of $|i|$, we have $N = j \geq |i|$. Thus $gy[i, j] = z[i, j]$ and $[i, j]$ is a long marked interval for y, too. Thus $z[i, j) = Q_{\mathrm{left}}(x)b_{x_i, x_j, r}(x[i, j])Q_{\mathrm{right}}(x) = Q_{\mathrm{left}}(y)b_{y_i, y_j, r'}(y[i, j])Q_{\mathrm{right}}(y)$. Since the blocks Q_{left} begin with $m^*q(x_i)$ (respectively $m^*q(y_i)$) and $q(n) = AB^nA$

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press

we obtain $x_i = y_i$. Since the blocks Q_{right} end with $q(x_j)m^*$ (respectively $q(y_j)m^*$) we get $x_j = y_j$. Since the length of the blocks Q_{left} and Q_{right} is determined by x_i, x_j and $j - i$ we get $|Q_{\text{left}}(x)| = |Q_{\text{right}}(x)| = |Q_{\text{left}}(y)| = |Q_{\text{right}}(y)|$, which shows that $b_{x_i,x_j,r}(x[i, j]) = b_{y_i,y_j,r'}(y[i, j])$. Since $x_i = y_i$ and $x_j = y_j$, and thus also $r = r'$, the injectivity of the map $b_{x_i,x_j,r}$ shows that $x[i, j] = y[i, j]$.

The only other possibility is that $[i, \infty)$ is a marked ray for x and thus $z[i, \infty) = m^*q(x_i)q(x_{i+1})q(x_{i+2})\cdots$. Let $N = |m^*q(x_i)q(x_{i+1})\cdots q(x_0)|$ and consider $y \in S$ with $gy[-N, N] = z[-N, N]$. Then $gy[i, i + N) = m^*q(x_i)q(x_{i+1})\cdots q(x_0)$, since $N > |i|$, and the property (g1) implies $y[i, 0] = x[i, 0]$, in particular $y_0 = x_0$.

This finishes the proof that (b) implies (a). □

Acknowledgement. Part of this work is from the second authors Habilitationsschrift [**Fi**].

References

[D] J. Dugundji. *Topology*. Allyn and Bacon, Boston, 1968.

[DGS] M. Denker, C. Grillenberger and K. Sigmund. *Ergodic Theory on Compact Spaces (Lecture Notes in Mathematics, 527)*. Springer, New York, 1976.

[F1] D. Fiebig. Factor theorems for locally compact Markov shift. *Forum Math*. **14** (2002), 623–640.

[F2] D. Fiebig. Factor maps, entropy and fibre cardinality for Markov shifts. *Rocky Mountain J. Math.* **31** (2001), 955–986.

[F3] D. Fiebig. Graphs with pre-assigned Salama entropies and optimal degrees for locally compact Markov shifts. *Ergod. Th. & Dynam. Sys.* **23**(4) (2003), 1093–1124.

[FF1] D. Fiebig and U. Fiebig. Entropy and finite generators for locally compact subshifts. *Ergod. Th. & Dynam. Sys.* **17** (1997), 349–368.

[FF2] D. Fiebig and U. Fiebig. Subshift compactifications of locally compact Markov shifts. *Theoret. Comput. Sci*. to appear.

[Fi] U. Fiebig. Symbolic dynamics and locally compact Markov shifts. *Habilitationsschrift*, Universität Heidelberg, 1996.

[G1] B. M. Gurevic. Topological entropy of enumerable Markov chains. *Dokl. Akad. Nauk SSSR* **187** (1969) (English transl. *Soviet Math. Dokl*. **10** (1969), 911–915).

[G2] B. M. Gurevic. Shift entropy and Markov measure in the path space of a denumerable graph. *Soviet Math. Dokl*. **11** (1970), 744–747.

[K] B. Kitchens. *Symbolic Dynamics*. Springer, Berlin, Heidelberg, New York, 1998.

[Kr] W. Krieger. On the subsystems of topological Markov chains. *Ergod. Th. & Dynam. Sys.* **2** (1982), 195–202.

[LM] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge, Cambridge University Press, 1995.

https://doi.org/10.1017/S0143385704000689 Published online by Cambridge University Press