Ergod. Th. & Dynam. Sys. (1997), **17**, 349–368
Printed in the United Kingdom © *1997 Cambridge University Press*

Entropy and finite generators for locally compact subshifts

DORIS FIEBIG and ULF-RAINER FIEBIG

Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany
(e-mail: fiebig@math.uni-heidelberg.de)

(*Received* 8 *November* 1994)

Abstract. We introduce transition entropy and periodic entropy for locally compact subshifts. Finiteness of both characterizes the existence of a finite generator. Finiteness of the transition entropy characterizes the existence of a generator with bounded degree. We extend Krieger's embedding theorem and characterize the locally compact non-dense subsystems of a (compact) mixing shift of finite type.

1. *Introduction and basic facts*

A *locally compact subshift* is an expansive homeomorphism S on a locally compact zero-dimensional Polish space X. (Expansiveness means that there is some metric for the topology such that the homeomorphism is expansive with respect to this metric.)

The term subshift is justified as follows. Consider $\mathbf{N}^{\mathbf{Z}}$ endowed with the product topology of the discrete topology on $\mathbf{N}$ together with the (left) shift map σ. Then any locally compact subshift S is conjugate to the action of σ on some locally compact shift invariant subset T of $\mathbf{N}^{\mathbf{Z}}$ (that is, there is a homeomorphism from X onto T that commutes with S and σ) [**Se**].

We call any such $(T, \sigma|_T)$ a *presentation* of S. To reduce notation we shall use the same symbol for the space and the map, thus we will talk about a *presentation* $T \subset \mathbf{N}^{\mathbf{Z}}$ of a subshift S.

We will investigate which locally compact subshifts admit 'good presentations' (for example presentations that use only finitely many symbols). There is a close relationship between presentations and generators (there are various definitions of 'generators' to be found in the literature, the following just suits our purposes).

Definition 1.1. A *generator* for a locally compact subshift S is a countable partition P of S into closed-open sets $P_1, P_2, \ldots$ such that for any sequence $(n(i))_{i\in\mathbf{Z}} \in \mathbf{N}^{\mathbf{Z}}$:

(1) the set $\{s \in S | S^i s \in P_{n(i)} \text{ for all } i \in \mathbf{Z}\}$ contains at most one point;

(2) if the set $\{s \in S | S^i s \in P_{n(i)} \text{ for all } i \in \mathbf{Z}\}$ is non-empty, then for some $k \geq 0$ the set $\{s \in S | S^i s \in P_{n(i)} \text{ for all } -k \leq i \leq k\}$ is compact.

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

A generator P is *finite* if it is a finite partition. It is *compact* if the sets $P_1, P_2, \ldots$ are compact. P has *bounded out-degree* if there is some N such that $\text{card}\{j \in \mathbf{N} | P_i \cap S^{-1}P_j \neq \emptyset\} \leq N$ for all $i \in \mathbf{N}$. P has *bounded in-degree* if there is some N such that $\text{card}\{j \in \mathbf{N} | SP_j \cap P_i \neq \emptyset\} \leq N$ for all $i \in \mathbf{N}$. P has *bounded degree* if it has bounded in- and out-degree.

Any compact subshift has a finite generator. However, non-compact subshifts may also have finite generators, as $S = \{1, 2\}^{\mathbf{Z}} - \{1^{\infty}\}$ shows: its zero-partition $P_i = \{s \in S | s_0 = i\}$, $i = 1, 2$ is a finite generator for S.

Obviously, a finite generator has bounded degree. There are locally compact subshifts with a generator with bounded degree which do not have a finite generator. For this case let $S = \{(s_i)_{i \in \mathbf{Z}} \in \mathbf{N}^{\mathbf{Z}} \mid |s_i - s_{i+1}| \leq 1 \text{ for all } i \in \mathbf{Z}\}$. Let $P_i = \{s \in S \mid s_0 = i\}$, $i \in \mathbf{N}$. Since S has infinitely many fixed points it cannot have a finite generator.

Not every locally compact subshift S has a generator with bounded degree, as Example 3.10 shows in combination with Theorem A. If S has a generator with bounded in-degree or with bounded out-degree, then S also has a generator with bounded degree (Theorem B).

Obviously, if S has a compact finite generator then S itself is compact. But, any S has a compact generator (Lemma 2.6) and if S has a generator with bounded degree then it also has a compact generator with bounded degree (Lemma 2.6).

Our main objective of this paper is to introduce two notions of entropy for locally compact subshifts and to use these to characterize those locally compact subshifts which have a finite generator and those which have a generator with bounded degree.

There is the following relation between generators and presentations. Given a presentation $T \subset \mathbf{N}^{\mathbf{Z}}$ of S and a conjugacy $f : S \to T$, then the partition of S into the sets $f^{-1}\{t \in T \mid t_0 = n\}$, $n \in \mathbf{N}$ is a generator for S. On the other hand, given a generator $P = P_1, P_2, \ldots$ let $f : S \to \mathbf{N}^{\mathbf{Z}}$ be the map that assigns to each point $s \in S$ its P-name, i.e. $f(s)_i = n(i)$ iff $S^i s \in P_{n(i)}$, $i \in \mathbf{Z}$. This gives a presentation $f(S) \subset \mathbf{N}^{\mathbf{Z}}$ of S. For notational convenience, we shall also denote the above map f by P; in this sense any generator P gives a presentation $P(S) \subset \mathbf{N}^{\mathbf{Z}}$ of S.

We will give necessary and sufficient conditions (in terms of entropy) for a locally compact subshift to have a finite generator. There are different ways to phrase this property. Clearly, a locally compact subshift S which is conjugate to a subsystem of a compact subshift has a finite generator. But the converse is also true: use the finite generator to encode S into $\mathbf{N}^{\mathbf{Z}}$ as described above. Then the closure of the image will be a compact subshift (it uses only finitely many symbols). The same reasoning shows that a locally compact subshift S has a finite generator iff it is conjugate to a dense subsystem of a compact subshift (i.e. it has a *subshift-compactification* [**FF**]).

There are two obvious obstructions for a locally compact subshift S having a finite generator.

The (Artin–Mazur) zeta function of S has to be analytic in some neighborhood of 0 since this holds for any subsystem of a compact subshift.

The Gurevic entropy of S (which is the supremum over all measure-theoretical entropies of S) has to be finite since the measure-theoretical entropies of a subsystem of a compact subshift T are bounded by the topological entropy of T.

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

For locally compact Markov shifts (for a definition see Definition 3.8) these two necessary conditions will be sufficient for the existence of a finite generator.

In general, another entropy is needed, which we call *transition entropy*. Roughly speaking, it is the smallest growth rate of words with fixed initial and terminal symbol achievable when choosing the right presentation. The transition entropy will be at least the Gurevic entropy, it can actually be larger, and for Markov shifts it coincides with the Gurevic entropy.

We will show that transition entropy by itself controls the existence of a compact generator with bounded degree ('a locally compact subshift has a presentation with compact zero partition and bounded degree iff its transition entropy is finite', Theorem A).

We shall prove that a locally compact subshift has a finite generator iff its zeta function is analytic in some neighborhood of 0 and its transition entropy is finite (Theorem C).

From this we extend Krieger's embedding theorem to a theorem about 'the locally compact non-dense subsystems of a mixing shift of finite type' (Theorem D).

We fix some frequently used notations.

Definition 1.2. Let $P = P_1, P_2, \ldots$ be a partition of a locally compact subshift S. The *P-n-name* of a point $x \in S$ is the finite sequence $i(0), i(1), \ldots, i(n-1)$ with $S^k(x) \in P_{i(k)}$ for $0 \leq k \leq n-1$.

The number $C_P(i, j, n)$ is the number of P-n-names with initial symbol i and terminal symbol j, i.e. the number of sequences $i(0), i(1), \ldots, i(n-1)$ such that $i(0) = i$, $i(n-1) = j$ and $i(0), i(1), \ldots, i(n-1)$ is the P-n-name of some point $x \in S$.

The *P-name* of a point $x \in S$ is the sequence $(i(n))_{n \in \mathbf{Z}} \in \mathbf{N}^{\mathbf{Z}}$ with $S^n(x) \in P_{i(n)} \forall n$. Let $S \subset \mathbf{N}^{\mathbf{Z}}$ be a presentation of a locally compact subshift. For $x = (x_n)_{n \in \mathbf{Z}} \in S$ and $-\infty \leq n \leq m \leq \infty$ let $x[n, m]$ denote the word $x_n, \ldots, x_m$ (as an element of $\mathbf{N}^{m-n+1}$), and let ${}_n[x_n, \ldots, x_m]$ denote the set of $y \in S$ with $y_n, \ldots, y_m = x_n, \ldots, x_m$. A set ${}_n[x_n, \ldots, x_m]$ will also be called an '$[n, m]$-cylinder'.

2. *Entropies for locally compact subshifts*

In this section we present the entropy concepts needed in our main theorems.

An essential invariant of any dynamical system is the growth rate of periodic points. For large classes of systems, as for compact expansive dynamical systems with specification, this growth rate coincides with topological entropy [**DGS**, Theorem 22.7]. For general compact subshifts it is still bounded above by topological entropy. But for locally compact subshifts even this bound (now replaced by the topological entropy of the 1-point compactification) breaks down. Since this growth rate seems not to be controlled by any other entropy, we called it 'periodic entropy'.

Definition 2.1. Let S be a locally compact subshift. The *periodic entropy* of S is $h_{\mathrm{per}}(S) := \limsup_n 1/n \log |P_n(S)|$, where $P_n(S)$ denotes the fixed points of S^n.

Observe that $1/e^{h_{\mathrm{per}}(S)}$ is the radius of convergence of the (Artin–Mazur) zeta function $\zeta(S)(z) = \exp \sum_{n=1}^{\infty} \frac{1}{n} |P_n(S)| z^n$ of S. In particular, $h_{\mathrm{per}}(S)$ is finite iff the zeta function is analytic in some neighborhood of 0.

There is no unique 'topological entropy' for locally compact, non-compact subshifts. Nonetheless, the following entropy (originally introduced by Gurevic for countable state

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

Markov shifts [**G1**]) seems to be an adequate analogue, also because it satisfies a variational principle.

Definition 2.2. [**G1**] The *Gurevic entropy* of S is $h_{\mathrm{G}}(S) := h_{\mathrm{top}}(S')$, where S' is the dynamical one-point-compactification of S (the extended transformation fixes the point at infinity), and $h_{\mathrm{top}}(S')$ denotes the topological entropy of the compact system S'.

Since conjugate locally compact shifts have conjugate 1-point compactifications, the Gurevic entropy is a conjugacy invariant.

For the moment let $M(S)$, respectively $M(S')$, denote the set of invariant Borel probability measures on S, respectively S'. The spaces S and S' differ only by the point at infinity, thus $\sup_{m\in M(S)} h_m(S) = \sup_{m\in M(S')} h_m(S') = h_{\mathrm{top}}(S')$ (the latter by the variational principle for compact spaces). Thus also $h_{\mathrm{G}}(S) = \sup_{m\in M(S)} h_m(S)$.

The crucial entropy for our purposes will be the following 'transition entropy'. Recall that $C_P(i, j, n)$ is the number of P-n-names with initial symbol i and terminal symbol j.

Definition 2.3. Let S be a locally compact subshift of $\mathbf{N}^{\mathbf{Z}}$. Let P be a generator for S. The *transition entropy of S with respect to P* is

$$h_{\mathrm{t}}(S, P) := \sup_{i,j} \limsup_{n} 1/n \log C_P(i, j, n).$$

The *transition entropy* of S is $h_{\mathrm{t}}(S) := \inf_P h_{\mathrm{t}}(S, P)$, where P varies over all generators.

Remark 2.4. If S is a locally compact non-compact subshift, which is 'transitive non-wandering' (i.e. for any pair U, V of non-empty open subsets $S^{-n}U \cap V$ is non-empty for some $n \in \mathbf{N}$), then there is an (even compact) generator P with $h_{\mathrm{t}}(S, P) = \infty$. We omit the construction, we only wanted to indicate why the infimum and not the supremum is used in the definition of $h_{\mathrm{t}}(S)$.

Remark 2.5. For general locally compact subshifts the only relation between the above entropies h_{per}, h_{G}, and h_{t} is $h_{\mathrm{G}} \le h_{\mathrm{t}}$ (see Theorem 3.7). For Markov shifts $h_{\mathrm{per}} \ge h_{\mathrm{G}} = h_{\mathrm{t}}$ (see Lemma 3.9 for $h_{\mathrm{G}} = h_{\mathrm{t}}$).

The following simple lemma will be crucial for the embedding Theorems C and D, since the necessity of the entropy condition can be easily seen from the definition of transition entropy, but for sufficiency we need that the transition entropy can be approximated arbitrarily well by the transition entropies of compact generators.

LEMMA 2.6. *$h_{\mathrm{t}}(S) = \inf_P h_{\mathrm{t}}(S, P)$, where P varies over all compact generators.*

Proof. Using P to recode S we may assume that S is given in a presentation as a subshift of $\mathbf{N}^{\mathbf{Z}}$ with zero-partition P.

For each S-point x, let $k(x) := \min\{k \ge 0 \mid {}_{-k}[x_{-k}, \ldots, x_k]$ is compact$\}$. Let $Q(x) := {}_{-k}[x_{-k}, \ldots, x_k]$. Varying x over all points of S we obtain at most countably many different $Q(x)$-sets, which form a countable compact partition $Q = Q_1, Q_2, \ldots$. Since Q refines the zero-partition it is a compact generator. We claim that $h_{\mathrm{t}}(S, Q) \le h_{\mathrm{t}}(S, P)$.

We have to estimate the numbers $C_Q(i, j, n)$ of Q-n-names which start with some fixed i and end with some fixed j. Let $Q_i = {}_{-k}[x_{-k} \ldots x_0 \ldots x_k]$ and

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

$Q_j = {}_{-m}[y_{-m} \dots y_0 \dots y_m]$. For this consider S-points z and w in Q_i, with $S^{n-1}z$ and $S^{n-1}w$ in Q_j, but with different Q-n-names. Then, since ${}_{-k}[z_{-k}, \dots, z_{m+n-1}]$ and ${}_{-k}[w_{-k}, \dots, w_{m+n-1}]$ are compact cylinders, $z[0, n-1] \neq w[0, n-1]$. Thus $C_Q(i, j, n) \leq C_P(x_0, y_0, n)$ for all n and any pair i, j, thus $h_t(S, Q) \leq h_t(S, P)$ by definition. (Observe that Q has bounded degree if P has). □

3. *More about transition entropy*

We show that transition entropy is the infimum over a certain class of metric entropies and thus bounded below by the Gurevic entropy (with equality in the Markov shift case). An example shows that h_t may be infinite even if h_G is finite.

The following sections are independent of this one, except that the definition of h_{out} (Definition 3.4) will be used in Theorem B and the fact that $h_t(S) \geq h_G(S)$ (Theorem 3.7), will be used in the proof of the embedding theorem (Theorem C).

A well known way to define entropies for non-compact dynamical systems is to consider a metric d which generates the topology of S and to use Bowen's entropy definition [**B**], which we recall now. For a compact subset $K \subset S$ a set $E \subset K$ is (n, ε, K)-*separated* if for every $x, y \in E$ there is an $0 \leq i < n$, so that $d(S^i x, S^i y) > \varepsilon$. Let $r_d(n, \varepsilon, K)$ be the maximal cardinality of an (n, ε, K)-separated set and let

$$h_d(S, K) = \lim_{\varepsilon \to 0} \overline{\lim_{n \to \infty}} \frac{1}{n} \log r_d(n, \varepsilon, K).$$

Then define the *entropy of S with respect to d* by $h_d(S) = \sup_K h_d(S, K)$, where the supremum is over all compact $K \subset S$ (see [**W**]).

Definition 3.1. From now on let d denote the *standard metric* on $\mathbf{N}^{\mathbf{Z}}$, that is

$$\begin{aligned} &d(x, y) = 1 && \text{if } x_0 \neq y_0 \\ &d(x, y) = 1/(k+1) && \text{if } x_0 = y_0, x[-k+1, k-1] = y[-k+1, k-1] \\ & && \text{and } x[-k, k] \neq y[-k, k]. \end{aligned}$$

If we consider $h_d(T)$ for a presentation $T \subset \mathbf{N}^{\mathbf{Z}}$ of a locally compact subshift S, then there might be other presentations T' of S with $h_d(T') \neq h_d(T)$. To get an invariant entropy we take the infimum over all presentations of S.

Definition 3.2. Let S be a locally compact subshift. Let P be a generator for S. Define $h^d(S, P) = h_d(P(S))$, where $P(S) \subset \mathbf{N}^{\mathbf{Z}}$ is the shift conjugate to S obtained by the map which assigns to any S-point its P-name. Let $h^d(S) = \inf_P h^d(S, P)$, where P ranges over all generators of S.

Remark 3.3. (1) The entropy $h^d(S)$ can be viewed as the infimum of the Bowen entropies $h_{P^{-1}d}(S)$ of S, where the metric $P^{-1}d$ is the pull-back of the metric d on $\mathbf{N}^{\mathbf{Z}}$ under the P-name-embedding $S \to P(S) \subset \mathbf{N}^{\mathbf{Z}}$, and P ranges over all generators of S.

(2) Although we shall not go into this here, one could of course consider other metrics on $\mathbf{N}^{\mathbf{Z}}$ and obtain other invariant entropies for locally compact subshifts (for example, there is a metric on $\mathbf{N}^{\mathbf{Z}}$ such that the above definition yields the Gurevic entropy [**G1**]).

We shall see that $h_t(S) = h^d(S)$ for locally compact subshifts. An intermediate concept used to prove this fact is the *out-going entropy* $h_{out}(S)$. It is inspired by the

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

Salama entropy [**Sa**], but the definition is made conjugacy invariant, the step which is missing in the definition of the Salama entropy. (Salama in fact proposed an invariant entropy via open covers, but this entropy is trivial (i.e. ∞) for most (of the transitive non-wandering) locally compact non-compact subshifts.)

Definition 3.4. Let S be a locally compact subshift of $\mathbf{N}^{\mathbf{Z}}$. Let P be a generator for S. The *out-going entropy of S with respect to P* is

$$h_{\text{out}}(S, P) := \sup_i \limsup_n 1/n \log C_P(i, n),$$

where $C_P(i, n)$ is the number of occurring P-names of length n with initial symbol i. The *out-going entropy of S* is $h_{\text{out}}(S) = \inf_P h_{\text{out}}(S, P)$, where P varies over all generators.

LEMMA 3.5. *Let S be a locally compact subshift. Then* $h_{\text{out}}(S) = h^d(S)$.

Remark 3.6. The following proof also shows that $h_{\text{out}}(S) = \inf_P h_{\text{out}}(S, P)$ and $h^d(S) = \inf_P h^d(S, P)$, where P varies only over all compact generators of S.

Proof. First we show $h^d(S, P) \le h_{\text{out}}(S, P)$ for any generator P. Let $\delta > 0$ be given. By definition, $h^d(S, P) = h_d(P(S))$. Thus there is a compact set $K \subset P(S)$ and an $\varepsilon > 0$ and a sequence $E(n)$ of (n, ε, K)-separated sets such that $\limsup_n 1/n \log \operatorname{card} E(n) \ge h^d(S, P) - \delta$. Choose $N = N(\varepsilon)$ such that $x[-N, n+N] \ne y[-N, n+N]$ for any two (n, ε)-separated points x, y in $P(S)$. By compactness, K intersects only finitely many $[-N, 0]$-cylinders, say m. Thus for each n there is a $[-N, 0]$-cylinder, say $Z(n)$, such that $\operatorname{card}\{E(n) \cap Z(n)\} \ge 1/m E(n)$. There is a subsequence n_k such that $\lim_k 1/n_k \log \operatorname{card} E(n_k) = \limsup_n 1/n \log \operatorname{card} E(n)$ and that the zero-coordinate of $Z(n_k)$ is constant, say i. For these n_k we have $C_P(i, n_k + N) \ge 1/m \operatorname{card} E(n_k)$, so $h_{\text{out}}(S, P) \ge \limsup_k 1/n_k \log C_P(i, n_k) \ge \lim_k 1/n_k \log \operatorname{card} E(n_k) \ge h^d(S, P) - \delta$.

We show that $h^d(S, P) = h_{\text{out}}(S, P)$ for any compact generator P. By [**W**, Theorem 7.5] it holds that $h_d(P(S)) = \sup_i h_d(P(S), [i]_0)$. For $\varepsilon < 1$ we have $C_P(i, n) \le r_d(n, \varepsilon, [i]_0)$. This shows $\limsup_n 1/n \log C_P(i, n) \le h_d(P(S), [i]_0)$ for all i, thus $h_{\text{out}}(S, P) \le h^d(S, P)$. Combined with the above we get equality for compact generators.

So far we have $h^d(S) \le h_{\text{out}}(S) \le \inf_{\text{compact } Q} h_{\text{out}}(S, Q) = \inf_{\text{compact } Q} h^d(S, Q)$. Finally, we show $h^d(S) = \inf_{\text{compact } Q} h^d(S, Q)$. Let $\delta > 0$ and let P be a generator with $h^d(S, P) \le h^d(S) + \delta$. We shall define a compact generator Q with $h_d(Q(S)) \le h^d(S) + 2\delta$. Use P to map S to the conjugate system $P(S)$, thus we may assume that S is given as a subshift of $\mathbf{N}^{\mathbf{Z}}$ with zero-partition P. Thus $h^d(S, P) = h_d(S) \le h^d(S) + \delta$ and it suffices to define a compact generator Q for S with $h_d(Q(S)) \le h_d(S) + \delta$. For this fix an integer $M > h_d(S)/\delta$. For each S-point x let $k(x) := \min\{k \ge 0 \mid {}_{-kM}[x_{-kM}, \dots, x_k] \text{ is compact}\}$. Let $Q(x) := {}_{-kM}[x_{-kM}, \dots, x_k]$. Varying x over all points of S we obtain at most countably many different $Q(x)$-sets, which form a countable compact partition $Q = Q_1, Q_2, \dots$. Since Q refines the zero-partition it is a compact generator. Observe that $h_d(Q(S)) = \sup_{i \in \mathbf{N}} h_d(Q(S), [i]_0)$ by [**W**, Theorem 7.5], since the cylinders $[i]_0$ are compact. For given $\varepsilon > 0$ there is an $N = N(\varepsilon)$ such that $q^1[-N, N+n] \ne q^2[-N, N+n]$ for any two (n, ε)-separated points q^1, q^2 in $Q(S)$. Now consider a fixed i. Then Q_i is a $[-k_i M, k_i]$-cylinder for some k_i.

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

For two S-points x, y denote their Q-names (i.e. images in $Q(S)$) by $Q(x), Q(y)$. Let $E(n, \varepsilon, [i]_0)$ be an (n, ε)-separated subset of $[i]_0 \subset Q(S)$ of maximal cardinality. Consider x, y in S s.t. $Q(x), Q(y)$ are distinct elements of $E(n, \varepsilon, [i]_0)$. By the choice of N, $Q(x)[-N, N+n] \neq Q(y)[-N, N+n]$. Since $x, y \in Q_i$ and Q_i is compact there is some integer p with $-N-(k_i+N)M \leq p \leq n+N+1+(n+N+k_iM+M)/M$ and $x_p \neq y_p$. Thus there is some ε' such that x and y are $(n+r(n), \varepsilon')$-separated in S, where $r(n)$ is the least integer larger than n/M. Thus $\operatorname{card} E(n, \varepsilon, [i]_0) \leq$ card of a maximal $(n+r(n), \varepsilon')$-separated subset of Q_i in S. Thus $h_d(Q(S), [i]_0) \leq (1+1/M)h_d(S, Q_i)$. Since $h_d(S, Q_i) \leq h_d(S)$ and $1/M \leq \delta/h_d(S)$, this gives $h_d(Q(S), [i]_0) \leq h_d(S)+\delta$, and thus $h_d(Q(S)) \leq h_d(S)+\delta$. □

THEOREM 3.7. *Let S be a locally compact subshift. Then $h_{\text{out}}(S) = h^d(S) = h_t(S) \geq h_G(S)$.*

Proof. In Theorem B we will show that $h_{\text{out}}(S) = h_t(S)$. Combined with Lemma 3.5 this shows the equality of the first three entropies. It remains to show that $h^d(S) \geq h_G(S) - \varepsilon$ for any $\varepsilon > 0$. Remember that $h_G(S) = h_{\text{top}}(S')$ where S' is the one-point compactification of S. Let '∞' denote the point at infinity. By the variational principle for compact spaces there is an ergodic measure m on S' with $h_m(S') > h_{\text{top}}(S') - \varepsilon$. We may assume that $h_{\text{top}}(S') > 0$ (otherwise we are done) and thus by ergodicity of m, $m(\infty) = 0$. Thus m is an S-invariant Borel probability measure on S and $h_m(S) = h_m(S')$. It holds that $h_m(S) \leq h_\rho(S)$ for any metric ρ on S [**HK**, Proposition 1.4]. Since $h^d(S)$ is the infimum over certain Bowen entropies of S (Remark 3.3(1)), we have $h_m(S') \leq h^d(S)$, thus $h^d(S) \geq h_G(S) - \varepsilon$. □

We show that $h_t(S) = h_G(S)$ in the case of Markov shifts.

Definition 3.8. A generator P is called a *Markov generator* if it has the *Markov property*, i.e. given any sequence $(n(i))_{i \in \mathbf{Z}} \in \mathbf{N}^{\mathbf{Z}}$ such that $S^{-i}(P_{n(i)}) \cap S^{-i-1}(P_{n(i+1)})$ is non-empty for all $i \in \mathbf{Z}$, then the set $\{s \in S \mid S^i s \in P_{n(i)} \text{ for all } i \in \mathbf{Z}\}$ is non-empty. A *Markov shift* is a subshift which has a Markov generator. A Markov shift is *irreducible* if for any pair of non-empty open subsets U, V the set $S^{-n}U \cap V$ is non-empty for some $n \in \mathbf{N}$.

LEMMA 3.9. *If S is an irreducible locally compact Markov shift, then $h_t(S) = h_G(S)$, and $h_t(S, P) = h_t(S)$ for every (necessarily compact) Markov generator P.*

Proof. Let S be given in a presentation such that the zero partition is a Markov generator. (The compactness of P follows from [**FF**]). Then, $h_G(S) = \sup_{i,j} \limsup_n 1/n \log C(i, j, n)$, where $C(i, j, n)$ denotes the blocks of length n starting in i ending in j [**G2**, Theorem 5]. But the latter is $h_t(S, P)$. This proves the lemma, since $h_t(S) \geq h_G(S)$ by Theorem 3.7. □

Though for Markov shifts Gurevic entropy and transition entropy coincide, this is not the case in general, as the following example shows.

Example 3.10. A transitive locally compact subshift S of $\mathbf{N}^{\mathbf{Z}}$ with $h_G(S) \leq \log 3$, $h_{\text{per}}(S) \leq \log 3$, and $h_t(S) = \log k$, where $k \in \mathbf{R}$, is any pre-assigned number > 9 (or ∞).

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

Let c_n be a sequence in $\mathbf{N}$ such that c_n is the largest number smaller than k^{2n} $\forall n$. If $k = \infty$ then let $c_n = n^n$ $\forall n$. Let $C_1, C_2, \ldots$ be a partition of the even numbers $2\mathbf{N}$ with $\mathrm{card}(C_n) = c_n$ $\forall n$. Let (d_n) be a sequence in $\mathbf{N}$ with $2^{2d_n+2n-1} > c_n$ $\forall n$. Let B_n be the family of words $\{1^{d_n}357\ldots(2n+1)x(2n+1)\ldots7531^{d_n} \mid x \in C_n\}, n \in \mathbf{N}$, where 1^{d_n} denotes the string of length d_n that consists only of 1's. Let B_0 be the set that contains the single word '1'. Let S be the closure of B^∞, where $B^\infty = \{x \in \mathbf{N}^\mathbf{Z} \mid x$ can be written as an infinite concatenation of words in $\bigcup_{n\geq 0} B_n\}$.

It is not hard to see that $S - B^\infty$ consists only of the orbits $1^\infty3579\ldots$ and $\ldots97531^\infty$. In particular, $S - B^\infty$ contains no periodic points.

S is locally compact. For any $n \in \mathbf{N}$ the number of $m \in \mathbf{N}$ such that nm or mn occurs in an S-point is finite. Thus, for any n, any sequence of S-points in $[n]_0 = \{x \in S \mid x_0 = n\}$ has a subsequence which converges in $\mathbf{N}^\mathbf{Z}$, thus also in $[n]_0$ (S is closed by definition). Thus all $[n]_0$ are compact.

$h_{\mathrm{per}}(S) \leq \log 3$. Let a, b be two new symbols. By the choice of d_n, for each $n \geq 1$ there is an injection f_n from B_n into the set of words $\{1x_1x_2\ldots x_{2d_n+2n-1}1 \mid x_1x_2\ldots x_{2d_n+2n-1} \in \{a, b\}^{2d_n+2n-1}\}$. Any periodic point is in B^∞. Thus S has for all n less n-periodic points than $\{1, a, b\}^\mathbf{Z}$.

$h_{\mathrm{G}}(S) \leq \log 3$. Define a map $\phi : S \to \{1, a, b\}$ as follows. The points in B^∞ are uniquely decipherable into words from $\bigcup_{n\geq 0} B_n$, thus we map B^∞ using the above defined mappings f_n (let $f_0(1) = 1$). Map $1^\infty357\ldots$ to $1^\infty a^\infty$ and $\ldots7531^\infty$ to $b^\infty1^\infty$, where the symbol '3' is mapped to the first appearance of 'a', respectively last appearance of 'b'. Now it is easy to see that ϕ is a Borel isomorphism from S onto $\phi(S)$. Let $M(S)$ and $M(\phi(S))$ denote the sets of invariant Borel probability measures on S and $\phi(S)$. Then $h_{\mathrm{G}}(S) = \sup_{m\in M(S)} h_m(S) = \sup_{m\in M(\phi(S))} h_m(\phi(S)) \leq \log 3$, since $\phi(S)$ is a subsystem of $\{1, a, b\}^\mathbf{Z}$.

$h_{\mathrm{t}}(S) = \log k$. Let $P = P_1, P_2, \ldots$ be a generator. W.l.o.g. $1^\infty \in P_1$. Thus ${}_{-N}[1, 1, \ldots, 1]_N$ is contained in P_1 for some N. For x in C_n, consider the P-name of a point in the orbit $1^\infty357\ldots(2n+1)x(2n+1)\ldots7531^\infty$. It is of the form $1^\infty w(x)1^\infty$ with a word $w(x)$ of length $2N + 2n + 1$. If $x, x' \in C_n$ with $x \neq x'$, then $w(x) \neq w(x')$ since P is a generator. Thus $C_P(1, 1, 2N + 2n + 3) \geq c_n$. So $\limsup_n 1/n \log C_P(1, 1, n) \geq \limsup_n 1/2n \log c_n$. Thus if $k = \infty$, then $h_{\mathrm{t}}(S) = \infty$. Now let $k < \infty$. We have shown that $h_{\mathrm{t}}(S) \geq \log k$. Now we show $h_{\mathrm{t}}(S, P) \leq \log k$, where P is the zero-partition of the presentation of S constructed above. For any symbol i there are two S-blocks $u_1(i)$ and $u_2(i)$ (respectively $w_1(i), w_2(i)$) which start (respectively end) with 1 such that whenever v is an S-block that starts (respectively ends) with i then $u_1(i)v$ or $u_2(i)v$ (respectively $vw_1(i)$ or $vw_2(i)$) is an S-block. This shows that for any pair i, j we have $C(x, y, n) \leq C(1, 1, n+n_1)+C(1, 1, n+n_2)+C(1, 1, n+n_3)+C(1, 1, n+n_4)$ $\forall n$, where n_1, n_2, n_3, n_4 depend only on (i, j). This shows that $h_{\mathrm{t}}(S, P) = \limsup_n 1/n \log C(1, 1, n)$. Since $C(1, 1, n)$ is monotone increasing, it suffices to consider $\limsup_n 1/2n \log C(1, 1, 2n)$. Let E_n be the S-blocks from 1 to 1 of length $2n$ which are a subblock of a single block in $\bigcup_{n\geq 0} B_n$. Let F_n be the remaining S-blocks from 1 to 1 of length $2n$. For E_n, obviously $\mathrm{card}\, E_n \leq 2n(c_{n-1} + c_{n-2} + \cdots + c_1 + 1) \leq 2n^2c_n$. Any block in F_n is a subblock of a finite concatenation of length $i \leq 4n$ of blocks from $\bigcup_{n\geq 0} B_n$. Using

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

the above maps f_n (see the estimate of $h_{\mathrm{per}}(S)$), one sees that for each length $i \leq 4n$ there are at most 3^i such concatenations, and trivially each has at most $2n+1$ subblocks of length $2n$. This gives $\operatorname{card} F_n \leq (2n+1)(3^{4n} + 3^{4n-1} + \cdots) \leq 12n^2 3^{4n}$. Thus $C(1, 1, 2n) = \operatorname{card} E_n + \operatorname{card} F_n \leq 24n^2(3^{4n} + c_n)$. Since $c_n \approx k^{2n}$ and $k > 9$, this gives $\limsup_n 1/2n \log C(1, 1, 2n) \leq \limsup_n 1/2n \log 16n^2 c_n \leq \log k$. □

4. *The marker lemma*

Let $S \subset \mathbf{N}^{\mathbf{Z}}$ be a presentation of a locally compact subshift. The constructions of special presentations (Theorems A–C) will use the entries of the S-points into some 'marking subset' M of S. In these constructions two consecutive entries into M will be considered as 'close' or 'far apart', depending on a 'weight function' L (defined on symbols). This basic tool will be provided by the marker lemma, an extension of Krieger's combinatorial marker lemma for compact subshifts [**KR**, Lemma 2] to the case of locally compact subshifts.

Definition 4.1. Given $M \subset S$ and some function $L : \mathbf{N} \to \mathbf{N}$, we use the following suggestive terms. Let $M(x) := \{i \in \mathbf{Z} \mid S^i x \in M\}$.

If $i \in M(x)$, $j \in M(x)$, $k \notin M(x)$ $i < k < j$, then call $[i, j] \subset \mathbf{Z}$ a *marked interval* of x. A marked interval $[i, j]$ of x will be called *long* if $j - i \geq \max\{L(x_i), L(x_j)\}$, i.e. if its length succeeds the weights of x_i and x_j. Otherwise, we call it a *short* marked interval of x.

If $i \in M(x)$, $k \notin M(x)$ $k < i$ then call $(-\infty, i]$ a *marked ray of x.* Analogously, call $[i, \infty)$ a *marked ray of x* if $i \in M(x), k \notin M(x)$ $i < k$.

MARKER LEMMA 4.2. *Let $S \subset \mathbf{N}^{\mathbf{Z}}$ be a presentation of a locally compact subshift. Assume that the zero-partition of P of S is compact and $h_{\mathrm{t}}(S, P) < \infty$. Given $\varepsilon > 0$ and an arbitrary function $a : \mathbf{N} \to \mathbf{N}$, there is a closed-open set $M \subset S$ and a function $L : \mathbf{N} \to \mathbf{N}$ with $L(n) \leq L(n+1)$ $\forall n$, such that:*

(1) *for any S-point x there is a $-L(x_0) \leq i \leq L(x_0)$ with $i \in M(x)$;*

(2) *if $[i, j]$ is a short marked interval of an S-point x, then $x[i - 6L(x_i), j + 6L(x_i)]$ is periodic with least period $j - i$ (in particular $x_i = x_j$), and there is a $(j-i)$-periodic S-point z with $z[i - 6L(x_i), j + 6L(x_i)] = x[i - 6L(x_i), j + 6L(x_i)]$;*

(3) *there is a non-decreasing function $f : \mathbf{Z}^+ \to \mathbf{Z}^+$ with $\lim_n f(n) = \infty$, such that for any S-point x and any long marked interval $[i, j]$ of x it holds that*

$$C_P(x_i, x_j, j - i + 1) \leq \exp((h_{\mathrm{t}}(S, P) + \varepsilon) \cdot (j - i - 2a(k) - 2f(n)),$$

where $k = \max\{x_i, x_j\}$ and $n = j - i + 1 - L(k)$.

A few remarks might help to digest the above statements.

By (1), $M(x)$ is non-empty for any S-point x. Thus the whole index set $\mathbf{Z}$ of x is a concatenation of marked rays and/or intervals of x.

By (2), close entries into M force periodicity along a longer time interval ($L(x_i) > j - i$).

By (3), far apart entries into M leave enough space (which gets even greater if one increases their distance of the fixed symbols x_i, x_j) to encode the block $x[i, j]$ by an

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

alphabet with about $\exp(h_t(S, P)+\varepsilon)$ symbols. We incorporated the function a to be able to claim some extra space, depending on the special needs of the particular application.

COROLLARY 4.3. *Let M, L be as in the marker lemma. If $[i, j]$ is a marked interval of a point x and ($j - i \geq L(x_i)$ or $j - i \geq L(x_j)$), then $j - i \geq \max\{L(x_i), L(x_j)\}$, i.e. $[i, j]$ is a long marked interval.*

Proof. Let $j - i \geq L(x_i)$. If $[i, j]$ is a short marked interval, then by definition $j - i < \max\{L(x_i), L(x_j)\}$. Since $x_i = x_j$ by (2) of the marker lemma, this implies $j - i < L(x_i)$, a contradiction. The other case is analogous. □

Proof of the Marker Lemma. First we prove (3) which does not actually depend on the set M, but is merely a statement about the existence of functions L and f as claimed.

Let $\lambda = \exp(h_t(S, P) + \varepsilon)$ and $\mu = \exp(h_t(S, P) + \varepsilon/2)$. For any pair i, j we have $\limsup_n 1/n \log C_P(i, j, n) \leq h_t(S, P) < \log \mu$, thus there is a $N(i, j)$ such that $C_P(i, j, n) < N(i, j)\mu^n$ for all n. Let $N(k) = \max(N(k, m), N(m, k) \mid m \leq k\}$, $k \in \mathbf{N}$.

Choose L with $L(n) \leq L(n + 1)$ $\forall n$ and $\lambda^{L(k)-2a(k)-1} \geq N(k)\mu^{L(k)}$, $k \in \mathbf{N}$. Choose f with $f(n) \leq f(n + 1)$ $\forall n$, $\lim_n f(n) = \infty$, and $\lambda^{n-2f(n)} \geq \mu^n$, $n \in \mathbf{N}$.

Then, with $k = \max\{x_i, x_j\}$, $C_P(x_i, x_j, L(k) + n) \leq N(x_i, x_j)\mu^{L(k)+n} \leq N(k)\mu^{L(k)}\mu^n \leq \lambda^{L(k)-2a(k)-1}\lambda^{n-2f(n)} = \exp((h_t(S, P) + \varepsilon) \cdot (L(k) + n - 1 - 2a(k) - 2f(n)))$. This implies (3), since there $L(k) + n = j - i + 1$.

Given L, we will define the set M. Compactness of the zero-cylinders implies that there is some function $k : \mathbf{N} \to \mathbf{N}$ with $k(n) > 6L(n)$ such that any non-empty $[-k(n), k(n)]$-cylinder with zero-coordinate n which is p-periodic with $p \leq L(n)$ contains a (then unique) p-periodic point of S. Fix such a function k.

We define M inductively. For each n split $[n]_0$ into the set of non-empty $[-k(n), k(n)]$-cylinders. Choose some enumeration $Z(n, 1), Z(n, 2), \ldots, Z(n, r(n))$ of these cylinders. Order the set $I = \{(n, m) \mid n \in \mathbf{N} \text{ and } 1 \leq m \leq r(n)\}$ by $(n, m) \leq (n', m')$ if $n < n'$, or ($n = n'$ and $m < m'$). Let $(n, m)-1$ denote the predecessor of (n, m), i.e. $(n, m) - 1 = (n, m - 1)$ if $m > 1$, and $(n, m) - 1 = (n - 1, r(n - 1))$ if $m = 1$.

We define a family of sets $M_{n,m}$, $(n, m) \in I$. Let $M_{1,1} = Z(1, 1)$. Using the above order, the sequence $M_{n,m}$ is defined inductively by

$$M_{n,m} := M_{(n,m)-1} \cup (Z(n, m) - \bigcup_{-L(n)\leq j\leq L(n)} S^j(M_{(n,m)-1})).$$

By definition the $M_{n,m}$ are an increasing sequence of compact-open sets.

Let M be the union of the $M_{n,m}$. Disjointness of the cylinders $Z(n, m)$ implies

$$M \cap Z(n, m) = Z(n, m) - \bigcup_{-L(n)\leq j\leq L(n)} S^j(M_{(n,m)-1}). \qquad (*)$$

Thus $M \cap [n]_0 = \bigcup_{1\leq m\leq r(n)} M \cap Z(n, m)$ is a finite union of compact-open sets, thus compact-open. This implies that $M = \bigcup_n M \cap [n]_0$ is closed-open.

We show (1). Let x be a S-point, let $n = x_0$. Then $x \in Z(n, m)$ for some m. Thus by $(*)$ x is either in $\bigcup_{-L(n)\leq j\leq L(n)} S^j(M_{(n,m)-1})$ or in $M_{n,m}$.

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

We show (2). By the definition of a 'short' marked interval, $j-i < \max\{L(x_i), L(x_j)\}$. Assume $x_i > x_j$. The observation

$$\begin{aligned} &\text{if } S^i(x) \text{ is in } M, \text{ then } S^i(x) \in Z(x_i, m) \text{ for some } m, \text{ and then, by}(*),\\ &S^i(x) \text{ is not in } \bigcup_{-L(x_i)\le k\le L(x_i)} S^k(M_{n',m'}) \text{ for any } (n', m') < (x_i, m), \end{aligned} \tag{**}$$

shows that $S^j(x)$, which is in M, has to be in some $M_{x_j,m'}$ with $(x_j, m') \ge (x_i, m)$. Thus $x_j \ge x_i$, a contradiction. Analogously one shows that $x_i < x_j$ cannot hold, thus $x_i = x_j = n$ for some n. Again, $S^j(x)$ is in M, thus in some $M_{n,m'}$. Since $j-i \le L(n)$, by $(**)$ we must have $m' = m$, i.e. $S^i(x) \in Z(n, m)$ and $S^j(x) \in Z(n, m)$. Since $k(n) > 6L(n)$ this implies that $x[i-k(n), j+k(n)]$ is a p-periodic block with p dividing $j-i$. Assume $p < j-i$. Then also $x[i+p-k(n), i+p+k(n)] = x[i-k(n), i+k(n)]$, i.e. $S^{i+p}(x) \in Z(n, m)$. We have that $S^i(x)$ and $S^j(x)$ are not in $X = \bigcup_{-L(n)\le k\le L(n)} S^k(M_{(n,m)-1})$. Since $j-i \le L(n)$, this implies that $S^{i+p}(x)$ is not in X, thus $S^{i+p}(x)$ is in $Z(n, m) - X$, thus in M, contradicting the condition $S^k(x) \notin M$, $i < k < j$. Thus, $x[i-k(n), j+k(n)]$ is periodic with least period $j-i$. Since $k(n) > 6L(n)$, the same holds for the block $x[i-6L(n), j+6L(n)]$. The existence of the periodic point z follows from the choice of $k(n)$. □

5. *Finite transition entropy implies bounded degree.* $h_t(S) = h_{out}(S)$

An over abundance of periodic points, such as infinitely many fixed points, may prevent a locally compact subshift from having a presentation with finitely many symbols. Still, it might have presentations with bounded degree. The following theorem characterizes this situation.

THEOREM A. *A locally compact subshift S has a presentation as a subshift of $\mathbf{N}^{\mathbf{Z}}$ with a compact zero-partition with bounded degree if and only if $h_t(S) < \infty$.*

Proof. A slightly stronger version of the 'only if' part holds: if S has some presentation with a zero-partition with just bounded out-degree, then this zero-partition obviously has finite transition entropy, thus $h_t(S) < \infty$.

Now assume $h_t(S) < \infty$, let $h_t(S) = \log\lambda$. We will show that there is a presentation with compact zero-partition, and out-degree and in-degree both bounded by 400λ.

Let $\varepsilon > 0$ such that $\lambda + \varepsilon/2$ is less than the integral part of $\lambda + 1$. By Lemma 2.6 there is a compact generator P such that $h_t(S, P) < \log(\lambda + \varepsilon/2)$. Replacing S by the presentation $P(S)$ we may assume that P is the zero-partition of S. Choose two disjoint alphabets A and B, both of cardinality at most $\lambda + 1$ and disjoint from $\{0, 1\}$. Choose compact subshifts $S_A \subset A^{\mathbf{Z}}$ and $S_B \subset B^{\mathbf{Z}}$ with entropies between $\log(\lambda + \varepsilon/2)$ and $\log(\lambda + \varepsilon)$ (for example, see [**Gr**]).

After some preparations, we will define an embedding $\phi = (g, h, e_1, e_2, e_3)$ of S into the product $\{A \cup B \cup \{0, 1\}\}^{\mathbf{Z}} \times \{0, 1\}^{\mathbf{Z}} \times \mathbf{N}^{\mathbf{Z}} \times \mathbf{N}^{\mathbf{Z}} \times \mathbf{N}^{\mathbf{Z}}$.

An auxiliary function $c : \mathbf{N} \to \mathbf{N}$ helps us to handle periodic stretches. Let $PB_{p,n}(S) = \{x[0, p-1] \mid x \in S \text{ with } x_0 = n, S^p x = x, S^i x \ne x \text{ for } i = 1, \ldots, p-1\}$. Recall that $B_p(S_B)$ denotes the blocks of length p in S_B. Then $\limsup_p 1/p \log \operatorname{card} PB_{p,n}(S) \le h_t(S, P) < \limsup_p 1/p \log \operatorname{card} B_p(S_B)$. Thus, for

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

each n there is a $c(n) \in \mathbf{N}$ such that card $PB_{p,n}(S) \leq c(n) \cdot$ card $B_p(S_B)$ $\forall p$. Partition $PB_{p,n}(S)$ into sets $PB_{p,n,i}(S)$, $1 \leq i \leq c(n)$, with card $PB_{p,n,i}(S) \leq$ card $B_p(S_B)$ for all i.

For each p, n, i fix an injective map $g_{p,n,i} : PB_{p,n,i}(S) \to B_p(S_B)$.

Obtain the functions L, f and the marker set M from the marker lemma applied with the above ε and the function $a(k) = 2\max\{c(n) \mid n \leq k\} + 2N + 2k$ (where $N \geq 0$ is some fixed number. Its value will become important in Theorem B, here one might assume that $N = 0$). By the marker lemma and with $k = \max\{n, m\}$ there are injective maps

$$b_{n,m,i} : C(n, m, L(k) + i) \to B_{L(k)+i-1-2a(k)-2f(i)}(S_A) \quad \text{for all } n, m \in \mathbf{N}, i \geq 0.$$

For direct coding of symbols, blocks and rays define a function q on $\mathbf{N}$ by $q(n) = 01^{a(n)/2}$ for $n \geq 1$. Extend q to a function on finite blocks by concatenation. For example, if $a(n) = 2n$, then $q(142) = 0101111011$. Extend q to a function on left- and right-infinite rays.

Finally, for notational convenience, let $B_+(m) = 123\ldots(m-2)(m-1)m$ and $B_-(m) = m(m-1)(m-2)\ldots21$ for $m \in \mathbf{N}$. Let $B_+(0)$ and $B_-(0)$ denote the empty block.

We indicate the ideas behind the conjugacy $\phi = (g, h, e_1, e_2, e_3)$.

- h just encodes the entries into the marking set M.
- Let $[i, j]$ be a short marked interval of x. Then $gx[i, j] = g_{j-i,n,m}(x[i, j])$ for some n, m. Then e_2 encodes the value of n and e_3 the value of m. Thus (e_2, e_3) encode which map $g_{j-i,n,m}$ has been used.
- Let $[i, j]$ be a long marked interval of x. Then $x[i, j]$ is mapped by g to a block of the form cad where c, d are blocks with symbols from $\{0, 1\}$ and a is a block with symbols from A. From c, d one recovers x_i and x_j. Then the length $j - i + 1$ of $x[i, j]$ together with the block a is enough to recover $x[i, j]$. The blocks c, d will also be used to make g continuous.
- e_1 makes the zero-partition compact.

Definition of $\phi = (g, h, e_1, e_2, e_3)$. Let $x \in S$. Define $(hx)_i = 1$ if $S^i x \in M$, $(hx)_i = 0$ if $S^i x \notin M$.

The other coordinates are determined as follows.

Case 1. If $[i, j]$ is a short marked interval of x, then by (2) of the marker lemma $x_i, \ldots, x_{j-1} \in PB_{j-i,x_i,m}(S)$ for some m. Let

$$g(x)[i, j-1] = g_{j-i,x_i,m}(x_i, \ldots, x_{j-1}),$$
$$e_1(x)[i, j-1] = 0^{j-i}, \quad e_2(x)[i, j-1] = x_i^{j-i}, \quad e_3(x)[i, j-1] = m^{j-i}.$$

Case 2. If $[i, j]$ is a long marked interval of x, let $k = \max\{x_i, x_j\}$. Then $s := 1 + j - i - L(k) > 0$. Let $n = a(k) + f(s)$. Let

$$g(x)[i, j-1] = \text{the initial block of length } n \text{ of } q(x[i, j]) \text{ followed by } b_{x_i,x_j,s}(x[i, j]) \text{ followed by the terminal block of length } n \text{ of } q(x[i, j]),$$

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

$$e_1(x)[i, j-1] = B_+(n)n^{j-i-2n}B_-(n),$$
$$e_2(x)[i, j-1] = B_-(x_i)0^{j-i-x_i-(x_j-1)}B_+(x_j-1),$$
$$e_3(x)[i, j-1] = B_-(m_1)0^{i-j-m_1-m_2}B_+(m_2),$$

with $m_1, m_2 \in \mathbf{Z}^+$ defined as follows. If there is some $r \in M(x)$ with $i - L(x_i) < r < i$, then choose the largest r and let m_1 be determined by $x_r, \ldots, x_{i-1} \in PB_{i-r,x_r,m_1}(S)$. Otherwise, $m_1 = 0$.

If there is some $r \in M(x)$ with $j < r < j + L(x_j)$, then choose the smallest r and let m_2 be determined by $x_j, \ldots, x_{r-1} \in PB_{r-j,x_j,m_2}(S)$. Otherwise, $m_2 = 0$.

Case 3. If $x[i, \infty)$ is a marked ray, then let

$$g(x)[i, \infty) = q(x[i, \infty)),$$
$$e_1(x)[i, \infty) = 1234\ldots,$$
$$e_2(x)[i, \infty) = B_-(x_i)000\ldots,$$
$$e_3(x)[i, \infty) = B_-(m_1)0^\infty,$$

with m_1 as in case 2.

Case 4. If $x(-\infty, i]$ is a marked ray, then let

$$g(x)(\infty, i-1] = q(x(\infty, i-1]),$$
$$e_1(x)(\infty, i-1] = \ldots 4321,$$
$$e_2(x)(\infty, i-1] = \ldots 000B_+(x_j-1),$$
$$e_3(x)(\infty, i-1] = 0^\infty B_+(m_2),$$

with m_2 as in case 2.

This gives a map $\phi = (g, h, e_1, e_2, e_3)$ which is shift invariant by definition. It is injective since $x \neq y$, and $h(x) = h(y)$ implies $(g, e_2, e_3)(x) \neq (g, e_2, e_3)(y)$ by inspection. Finally, we show that:

(1) $\phi = (g, h, e_1, e_2, e_3)$ is continuous;
(2) zero-cylinders in $\phi(S)$ are compact and ϕ is a homeomorphism;
(3) $\phi(S)$ has in- and out-degree bounded by 400λ.

(1) ϕ is continuous. First note that h is continuous since M is closed-open. Thus it suffices to show that for any x there is an N such that $\phi(x)_0$ is determined by $h(x)[-3N, 3N]$ and $x[-3N, 3N]$.

At least one of the sets $M(x) \cap \{-i \mid i \geq 0\}$, $M(x) \cap \{i > 0\}$ is non-empty. Assume the first one. Let $-i, i \geq 0$, denote its maximal element. (The other case is treated similarly.) Choose an $N \geq i + L(i) + L(x_{-i})$ such that $f(N + i - L(i) - L(x_{-i})) > i$. We claim that $h(x)[-3N, 3N]$ and $x[-3N, 3N]$ determine $\phi(x)[-i, 0]$, thus in particular $\phi(x)_0$.

If $h(x)_j = 1$ for some $0 < j \leq N$, then choose j minimal. If $0 < j < L(x_{-i}) - i$ for this minimal j, then $j + i < L(x_{-i})$, i.e., $[-i, j]$ is a short marked interval and $\phi(x)[-i, j]$ is determined by a case 1 coding, i.e. by $h(x)[-i, j]$ and $x[-i, j]$. If $L(x_{-i}) - i \leq j \leq N$ for this minimal j, then $L(x_{-i}) \leq i + j$, thus $[-i, j]$ is a

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

long marked interval by Corollary 4.3. So $\phi(x)[-i, j]$ is determined by a case 2 coding, i.e. by $h(x)[-i - L(x_{-i}), j + L(x_j)]$ and $x[-i - L(x_{-i}), j + L(x_j)]$. Now by definition $-3N \leq -i - L(x_{-i})$, and since $L(x_j) \leq i + j$ (long interval), also $j + L(x_j) \leq 2j + i \leq 3N$. So assume from now on that $h(x)_j = 0$ for all $-i < j \leq N$.

We will show that the (g, e_1, e_2, e_3)-coordinates of $\phi(x)[-i, 0]$ are as if we applied a case 3 coding to $x[-i, \infty)$, i.e.

$$
\begin{aligned}
&g(x)[-i, 0] = \text{the initial block of length } i + 1 \text{ of } q(x[-i, 0]), \qquad (*)\\
&e_1(x)[-i, 0] = 1, 2, \ldots, i + 1,\\
&e_2(x)[-i, 0] = \text{the initial block of length } i + 1 \text{ of } B_-(x_{-i})0^\infty,\\
&e_3(x)[-i, 0] = \text{the initial block of length } i + 1 \text{ of } B_-(m_1)0^\infty,
\end{aligned}
$$

where m_1 is determined as in the case 2 coding (thus m_1 is determined by $h[-i - L(x_{-i}), -i]$ and $x[-i - L(x_{-i}), -i]$).

If $j \notin M(x)$ for all $-i < j$, then $\phi(x)[-i, \infty)$ is in fact determined by a case 3 coding. The other case is $j \in M(x)$ for some $N < j$. Then $j - (-i) > N > L(x_{-i})$ and $[-i, j]$ is a long marked interval of x by the observation following the marker lemma. Thus $\phi(x)[-i, j - 1]$ is determined by a case 2 coding. Then $(*)$ holds, if we can show that in this coding we have $n > i$, where $n = a(k) + f(s)$, $s = 1 + j + i - L(k)$, $k = \max\{x_{-i}, x_j\}$. If $k > i$, then $n \geq a(i) > i$. If $k \leq i$, then $n \geq f(s) \geq f(j + i - L(i)) \geq f(N + i - L(i)) > i$ by the choice of N. This finishes the proof of continuity.

(2) Let $s = (a, b, c, d, e)$ by a symbol of $\phi(S)$. We claim that $Z = \phi^{-1}([s]_0)$, and thus $[s]_0$, is compact. It suffices to show that Z is contained in a finite union of S-cylinders (S-cylinders are compact and Z is closed, since $[s]_0$ is closed in $\phi(S)$).

If a is a symbol from S_B, then for any $x \in Z$ there is a $0 \leq k \leq L(d)$ with $x_{-k} = d$. If a is a symbol from S_A, then for any $x \in Z$ there are $0 \leq k, i, j, n$ with $k < L(\max(i, j)) + n$, $x_{-k} = i$, $x_{-k+L(\max(i,j))+n} = j$ and $c = a(\max(i, j)) + f(n)$. Since $\lim_n f(n) = \infty$, $\lim_n a(n) = \infty$, this condition is met only by a finite number of tuples k, i, j, n. If $a \in \{0, 1\}$, then for any $x \in Z$ we have $x_{-c+1} \leq d + c - 1$ or $x_c \leq d + c$.

To show the continuity of ϕ^{-1}, let A be a closed subset of S. We show that $\phi(A)$ is closed in $\phi(S)$. Consider a sequence of points y_k in $\phi(A)$ converging to some point $y \in \phi(S)$. Discarding finitely many y_k we may assume that all y_k have the same zero-coordinate. For each k let x_k be the pre-image of y_k. Being contained in a compact set (see above), the x_k have a subsequence x_{k_i} converging to some point x, which has to belong to A. By continuity the y_{k_i} converge to $\phi(x)$. Thus $y = \phi(x)$, i.e. $y \in \phi(A)$.

(3) Let $s = (a, b, c, d, e)$ be a symbol of $\phi(S)$. The numbers of symbols $s' = (a', b', c', d', e')$ such that ss' or $s's$ is a $\phi(S)$-block is bounded by $(2\lambda + 4) \cdot 2 \cdot 27 \leq 400\lambda$, since a' has at most $\operatorname{card}(A \cup B \cup \{0, 1\}) \leq 2 + 2(\lambda + 1)$ choices, b' is in $\{0, 1\}$, and $|c - c'| \leq 1$, $|d - d'| \leq 1$, $|e - e'| \leq 1$ by construction, where $|e - e'| \leq 1$ can be seen as follows. Use the definition of e_3 and the observation that $x \in S$, $h(x)_0 = 1$, $e_3(x)_0 = 0$ implies that there is no $r \in M(x)$ with $-L(x_0) < r < L(x_0)$. Thus $e_3(x)_{-1} = 0$ and $e_3(x)_1 = 0$ by Corollary 4.3. □

THEOREM B. *Let S be a locally compact subshift. Then $h_{\mathrm{t}}(S) = h_{\mathrm{out}}(S)$.*

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

Proof. By definition $h_{\text{out}}(S) \geq h_{\text{t}}(S)$. We show $h_{\text{out}}(S) \leq h_{\text{t}}(S)$. Let $\lambda = \exp(h_{\text{t}}(S))$. Let $\varepsilon > 0$ be given such that $\lambda+\varepsilon/2$ is less than the integral part of $\lambda+1$. Choose shifts S_A, S_B as defined in the proof of Theorem A. Choose $Q \geq 1$ such that card $B_n(S_A) \leq Q(\lambda+\varepsilon)^n$, card $B_n(S_B) \leq Q(\lambda+\varepsilon)^n$ $\forall n$. Choose $N \in \mathbf{N}$ so large that

$$\frac{1}{N-1}\log(2^{15}N^{17}(400\lambda)^2Q^6(\lambda+\varepsilon)^N) \leq \log\lambda + 2\varepsilon.$$

Let $\phi(S)$ be the presentation of S given by a code $\phi = (g, h, e_1, e_2, e_3)$ as defined in the proof of Theorem A (the N comes into play in the definition of $a(k)$ in this proof).

For the zero-partition P of $\phi(S)$ we will show $h_{\text{out}}(\phi(S), P) \leq \log\lambda + 2\varepsilon$, which implies $h_{\text{out}}(S) \leq h_{\text{out}}(S, \phi^{-1}P) \leq h_{\text{t}}(S) + 2\varepsilon$, and proves the theorem.

For any symbol $s = (a_0, b_0, c_0, d_0, e_0)$ of $\phi(S)$, let $E(s, n)$ denote the set of n-blocks $(a_1, b_1, c_1, d_1, e_1), \ldots, (a_n, b_n, c_n, d_n, e_n)$ in $\phi(S)$ which start with $(a_1, b_1, c_1, d_1, e_1) = s$. Let $E'(s, n)$ denote the set of n-blocks in $\phi(S)$ which start with s and for which the first coordinate belongs the whole time to the same alphabet, A, B or $\{0, 1\}$, as a_1.

CLAIM. *For $n \leq N$ it holds that $E'(s, n) \leq 2^5n^5Q^2(\lambda+\varepsilon)^n$.*

Assuming that this claim holds, we finish the proof as follows. Any block in $E(s, N)$ can be written as a composition of three blocks $E'(s, n_1)E'(s_1, n_2)E'(s_2, n_3)$ with $n_1 + n_2 + n_3 = N$. This follows easily since by construction stretches of 0's and 1's have length at least N and a direct change between the alphabets A and B is not possible.

There are at most N^2 choices for n_1, n_2, n_3. For any particular one there are $2^5n_1^5Q^2(\lambda+\varepsilon)^{n_1}$ choices for $E'(s, n_1)$, and given the last symbol of $E'(s, n_1)$ there are at most 400λ choices for s_1 etc. Thus in total card $E(s, N) \leq N^2 2^{15}n_1^5n_2^5n_3^5(400\lambda)^2Q^6(\lambda+\varepsilon)^N \leq 2^{15}N^{17}(400\lambda)^2Q^6(\lambda+\varepsilon)^N$ for all s.

Let $D(n) := \sup_s \text{card}\, E(s, n)$. The last inequality and the choice of N show $1/(N-1)\log D(N) \leq \log\lambda + 2\varepsilon$.

By induction on k one sees that $D(n) \leq D(N)^k$ for $n = k(N-1)+1$. For each n there is a $k_n \leq n/(N-1)+1$ s.t. $n \leq k_n(N-1)+1$. Thus, since the $D(n)$ are increasing, we have $D(n) \leq D(N)^{k_n}$ for all n. Thus $\limsup_n 1/n \log D(n) \leq \limsup_n k_n/n \log D(N) \leq 1/(N-1)\log D(N) \leq \log\lambda + 2\varepsilon$. Thus, in particular, $h_{\text{out}}(\phi(S), P) \leq \log\lambda + 2\varepsilon$.

Proof of the claim. Investigating the definition of ϕ shows the following.

(a) If $a_j \in A$ for all j, then $(a_1, \ldots, a_n)$ is in $B_n(S_A)$, thus there are at most $Q(\lambda+\varepsilon)^n$ possibilities for $(a_1, \ldots, a_n)$. The other coordinates are constant by construction.

(b) Assume $a_j \in B$ for all j. Then $c_j = c_1$, $d_j = d_1$, $e_j = e_1$ for all j. If all $b_j = 0$, then $(a_1, \ldots, a_n)$ is a $B_n(S_B)$-block, thus there are at most $Q(\lambda+\varepsilon)^n$ possibilities.

If exactly one $b_j = 1$, then $(a_1, \ldots, a_n)$ is the composition of at most 2 S_B blocks whose lengths add up to n, thus there are at most $nQ^2(\lambda+\varepsilon)^n$ possibilities.

If $b_1 + \cdots + b_n = k > 2$, then the block $(b_1, \ldots, b_n)$ is determined by (k, x, y), where $1 \leq x < y \leq n$ denote the first two occurrences of '1' (the other 1's have to occur at the same distance from each other). Given such a b-block, there are at most $Q(\lambda+\varepsilon)^{y-x-1} \leq Q(\lambda+\varepsilon)^n$ possibilities for the a-block $(a_1, \ldots, a_n)$, since it is determined by the S_B-block $a_x, \ldots, a_{y-1}$. In total, there are at most $n^3Q(\lambda+\varepsilon)^n$ possibilities.

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

Adding up all the estimates shows that the claimed upper bound holds.

(c) Assume $a_j \in \{0, 1\}$ for all j. The a-block can contain at most one '0' since consecutive 0's are separated by at least N 1's. Thus there are at most $n + 1$ a-blocks. The b-block may contain at most one '1', thus there are at most $n + 1$ b-blocks. By inspection, given the first symbol (a, b, c, d, e) the other 3 coordinates each have at most $2n$ possibilities, which gives in total at most $(n + 1)^2(2n)^3 \leq (2n)^5$ possibilities. □

Remark 5.1. If one could prove that $h_t(T^n) = nh_t(T)$, for $n \geq 0$, or that $h_t(S \times T) = h_t(S) + h_t(T)$, then there would be a more direct derivation of $h_t(S) = h_{out}(S)$ from the fact that the out-degree of the presentation constructed in Theorem A is bounded by 400λ.

6. *Finite generators and embeddings into compact subshifts*

If a locally compact subshift S has a finite generator, then, for some N, it has a presentation as a shift invariant subset of the full shift S_N on $\{1, \ldots, N\}^{\mathbf{Z}}$. Thus $\zeta(S)$ is analytic at 0, since $\zeta(S_N)$ is, and the zero-partition of this presentation has transition entropy at most $\log N$. Thus $h_{per}(S) < \infty$ and $h_t(S) < \infty$. The following theorem shows that the converse, and more, is true.

THEOREM C. *Let S be a locally compact subshift with $h_t(S) < \infty$ and $h_{per}(S) < \infty$. Then, for any $\varepsilon > 0$, there is a compact subshift T with $h_{top}(T) < \max\{h_t(S), h_{per}(S)\} + \varepsilon$ and $\zeta(T) = \zeta(S)$, and a dense embedding ϕ of S into T.*

We will need some suitable compact subshifts which are supplied by the following lemma. Recall that $B_n(T) = \{x[0, n-1] \mid x \in T\}$ for subshifts T.

LEMMA 6.1. *Given $h \geq 0$, $\varepsilon > 0$ and $c \in \mathbf{R}$, there is a compact subshift T with no periodic points, $h < h_{top}(T) < h + \varepsilon$, and* card $B_n(T) \geq ce^{nh}$ $\forall n$.

Proof. There is a minimal subshift T' with $h < h_{top}(T') < h + \varepsilon$ [**Gr**]. Thus for some N, card $B_n(T') \geq ce^{nh}$ for all $n \geq N$. Let k be the least integer larger than ce^{Nh}. Let T be the disjoint union of k copies of T'. □

Proof of Theorem C. By Lemma 2.6 we may assume that S is given in a presentation with compact zero-partition P for which $h_t(S, P) < h_t(S) + \varepsilon/4$.

We choose some compact subshifts S_A, S_B, S_C with disjoint alphabets A, B, C. By Lemma 6.1 there is a compact subshift S_A with $h_t(S) + \varepsilon/2 < h_{top}(S_A) < h_t(S) + \varepsilon$, no periodic points, and card $B_n(S_A) \geq \exp(n \cdot (h_t(S) + \varepsilon/2))$ $\forall n$.

Since $\zeta(S)$ is analytic at 0, there is some $c \in \mathbf{R}$ such that the number of points of period n of S is less than $c(\exp(h_{per}(S) + \varepsilon/2))^n$ $\forall n$. By Lemma 6.1 there is a compact subshift S_B with finite alphabet B with $h_{per}(S) + \varepsilon/2 < h_{top}(S_B) < h_{per}(S) + \varepsilon$, no periodic points, and card $B_n(S_B) \geq$ the number of points of period n of S, $\forall n$.

Let S_C be a minimal compact subshift with finite alphabet C and zero entropy.

After some more preparations, we will define an embedding $\phi = (g, h)$ of S into the product $\{A \cup B \cup C \cup \{0\}\}^{\mathbf{Z}} \times \{0, 1\}^{\mathbf{Z}}$.

Apply the marker lemma with $\varepsilon/2$ and $a(k) = 2k$, $k \in \mathbf{N}$, to obtain the functions L, f, and the marker-set M. By (3) of the marker lemma for all $n, m \in \mathbf{N}$, $s \geq 0$, we can

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

choose an injective function $b_{n,m,s}$ from the set of blocks from n to m of length $L(k)+s$ into $B_{L(k)+s-1-4k-2f(s)}(S_A)$, where $k = \max\{n, m\}$.

For all $n \in \mathbf{N}$ let $PB_n(S) = \{x[0, n-1] \mid x \in S \text{ with } S^n x = x,\ S^i x \neq x \text{ for } i = 1, \ldots, n-1\}$. Let g_1 be an injective map $\cup PB_n(S) \to \cup B_n(S_B)$ with $g_1(PB_n(S)) \subset B_n(S_B)$ for $\forall n$.

For direct coding of symbols, blocks, and rays, fix a point c of S_C. Let 0 be a symbol disjoint from A, B, C. Define a function q on $\mathbf{N}$ by $q(n) = 0c_1 \ldots c_n$ for $n \geq 1$. Extend q to a function on finite blocks by concatenation. Thus, for example, $q(142) = 0c_1 0c_1c_2c_3c_4 0c_1c_2$. Analogously extend q to a function on left- or right-infinite rays.

We now define $\phi = (g, h)$. Let $x \in S$. Define $(hx)_i = 1$ if $S^i x \in M$, $(hx)_i = 0$ if $S^i x \notin M$.

Now we describe g.

Case 1. $[i, j]$ is a short marked interval of x. Then $gx[i, j-1] = g_1(x_i, \ldots, x_{j-1})$.

Case 2. $[i, j]$ is a long marked interval of x. Then with $k = \max\{x_i, x_j\}$ we have $j - i = L(k) + s$ with $s \geq 0$. Let $gx[i, j-1]$ be the concatenation of the following blocks: the initial block of length $2k + f(s)$ of $q(x[i, j])$, the block $b_{x_i, x_j, s}(x[i, j])$, and the terminal block of length $2k + f(s)$ of $q(x[i, j])$.

Case 3. $[i, \infty)$ is a marked ray of x. Then $gx[i, \infty) = q(x[i, \infty))$.

Case 4. $(-\infty, i]$ is a marked ray of x. Then $gx(-\infty, i-1] = q(x(-\infty, i])$.

We show that the code $\phi = (g, h)$ is a conjugacy. ϕ is shift invariant by definition and injective since $x \neq y$ and $h(x) = h(y)$ implies $g(x) \neq g(y)$.

We show that ϕ is continuous. Since M is closed-open, h is continuous. Thus, g will be continuous if for any x there is an N such that $g(x)_0$ is determined by $h(x)[-N, N]$ and $x[-N, N]$. This is obvious when 0 is contained in a marked interval.

So assume now that $[-i, \infty)$ is a marked ray for x, for some $-i \leq 0$ (the case that there is a marked ray $(-\infty, i]$ is treated similarly). Choose $N \geq \max\{L(i), L(x_{-i})\}$ such that $f(N + i - L(i)) > i$. We show that $g(x)_0$ is determined by $h(x)[-N, N]$ and $x[-N, N]$. For this let y be another S-point with $h(y)[-N, N] = h(x)[-N, N]$ and $y[-N, N] = x[-N, N]$. If $[-i, \infty)$ is also a marked ray of y, then by definition $g(y)_0 = g(x)_0$ since $q(x[-i, 0])$ is longer than i, and thus determines the zero-coordinate. If $S^j y \in M$ for some $-i < j$, then choose the first one. Then $N < j$, since $h(y)[-N, N] = h(x)[-N, N]$. Because $N \geq L(x_{-i})$, $[-i, j]$ is a long marked interval for y. Thus $gy[-i, j-1]$ starts with an initial block of length $2k + f(s)$ of $q(y[-i, j])$, where $k = \max\{y_i, y_j\}$ and $s = j + i - L(k)$. We have to show that $2k + f(s) > i$. Assume $k \leq i$. Then $f(s) \geq f(j + i - L(i)) \geq f(N + i - L(i)) > i$.

We show that ϕ^{-1} is continuous. By shift commutativity it suffices to show that for any point (y, z) in $(g, h)(S)$ the zero-coordinate of the pre-image is determined by $y[i, j]$, $z[i, j]$ for some $i \leq 0 \leq j$. Let x be the unique pre-image with $(y, z) = (gx, hx)$. Choose k such that there is a symbol '1' in $z[-k, k]$, say at i. Thus, $S^i x$ is in M. Assume $i \leq 0$ ($i \geq 0$ will be symmetric). Then choose the largest such $i \leq 0$.

If y_i is a symbol in B, then there is some $j > i$ s.t. $[i, j]$ is a short marked interval

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

of x and we can recode $x[i, j-1]$ as $g_1^{-1}(y[i, j-1])$. By the maximality of i we have $0 < j$, thus, in particular, x_0 is determined by $(y, z)[i, j]$.

If y_i is not in B, then by construction y_i is 0. Then either we see an initial block of $y[i, \infty)$ as $0c_1 \dots c_{n(1)}0c_1 \dots c_{n(2)}0 \dots 0c_1 \dots c_{n(k)}0$ with $k = |i| + 1$, then we know that $x[i, 0] = n(1)n(2) \dots n(k)$ is determined by $(y, z)[i, i + n(1) + n(2) + \cdots + n(k) + k]$ or there will be some $i < j$ s.t. $[i, j]$ is a long marked interval, and again by the maximality of i, $0 < j$. Thus $x[i, 0]$ is determined by $(y, z)[i, j]$.

So far we have shown that $\phi = (g, h)$ is an embedding of S into the compact shift $\{A \cup B \cup C \cup \{0\}\}^{\mathbf{Z}} \times \{0, 1\}^{\mathbf{Z}}$. Let T be the closure of $\phi(S)$ in $\{A \cup B \cup C \cup \{0\}\}^{\mathbf{Z}} \times \{0, 1\}^{\mathbf{Z}}$. Then $S^* = T - \phi(S)$ is the 'boundary' of this subshift-compactification [**FF**].

We will show that there are no periodic points in S^* and that $h_{\text{top}}(T) \leq \max\{h_{\text{t}}(S), h_{\text{per}}(S)\} + \varepsilon$. For the latter we will show that $h_{\text{top}}(S^*) \leq \max\{h_{\text{t}}(S), h_{\text{per}}(S)\} + \varepsilon$ (Claim 5), which, combined with $\max_{\mu \text{ ergodic on } S} h_\mu(S) = h_{\text{G}}(S)$, the variational principle for compact spaces, and the fact that $h_{\text{G}}(S) \leq h_{\text{t}}(S)$ (Theorem 3.7), gives

$$\begin{aligned} h_{\text{top}}(T) &= \max_{\mu \text{ ergodic}} h_\mu(T) = \max\left\{ \max_{\mu \text{ ergodic on } S} h_\mu(S), \max_{\mu \text{ ergodic on } S^*} h_\mu(S^*) \right\} \\ &= \max(h_{\text{G}}(S), h_{\text{top}}(S^*)) \leq \max(h_{\text{t}}(S), h_{\text{top}}(S^*)) \leq \max\{h_{\text{t}}(S), h_{\text{per}}(S)\} + \varepsilon. \end{aligned}$$

CLAIM 1. *For each n there is a $Q(n)$, such that whenever $x \in S$ has $x_j = n$ and this x_j gets coded as a block $0c_1c_2 \dots c_n0$, then this block is contained in $gx[-Q(n) + j, j + Q(n)]$.*

Proof. W.l.o.g. assume $j = 0$. We know that for some $-L(n) \leq i \leq L(n)$ we have $S^i x \in M$. Choose an i with smallest absolute value. Assume $i \leq 0$ (the other case is symmetric). Since n is encoded as $0c_1c_2 \dots c_n0$ we have that $0c_1c_2 \dots c_n0$ is the end block of $q(x[i, 0])$. Let $N(n) = \max\{k \mid$ there is a S-block of length $\leq L(n)$ with first symbol k and last symbol n, or with first symbol n and last symbol $k\}$. Then all symbols in $x[i, 0]$ are less than $N(n)$. Thus the length of $q(x[i, 0])$ is at most $(N(n) + 1)(|i| + 1) \leq (N(n) + 1)(L(n) + 1)$. Let $Q(n) = (N(n) + 1)(L(n) + 1)$.

The following claim contains essential information on the structure of S^*.

CLAIM 2. *Let $(y, z) \in S^* = T - \phi(S)$.*
(a) *Then $z_i = 1$ for at most one index i.*
(b) *There are no $i < j$ such that $y_i = 0$, $y_k \in C$ for $i < k \leq j$, and $y_{j+1} = 0$.*

Proof. (a) Assume $i < j$ and $z_i = z_j = 1$, $z_k = 0$ for $i < k < j$. Let $x \in S$ such that $\phi(x)[i, j] = (y, z)[i, j]$. Then $y[i, j]$ has been obtained by a case 1 or case 2 coding, in any case the coordinate x_i is determined and equal to n, say. Let x^m be a sequence of S points such that $\phi(x^m)$ converges to (y, z). Then by the above, eventually all x^m are contained in $[n]_i$. Since $[n]_i$ is compact, the x^m have a subsequence convergent to some x in S and $\phi(x) = (y, z)$, thus $(y, z) \in \phi(S)$, a contradiction.

(b) Assume $y_i = 0$, $y_k \in C$ for $i < k \leq j$, and $y_{j+1} = 0$. Let $x \in S$ such that $\phi(x)[i, j + 1] = (y, z)[i, j + 1]$. Thus the block $y[i, j + 1]$ codes an occurrence of the symbol $n = j - i$ in x. Thus, by Claim 1, there is a k with $-Q(n) \leq k \leq Q(n)$ and $x_k = n$. Thus x is contained in the compact subset $\bigcup_{-Q(n) \leq k \leq Q(n)} S^k([n]_0)$ of S. As above, this implies that (y, z) is in $\phi(S)$, a contradiction.

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

We are now going to describe the structure of the subblocks of points in S^*.

Let E denote the family of blocks which are:

- a subblock of a point in S_B; or
- a subblock of a block $C_1 0 C_2 A C_3 0 C_4$, where A is a subblock of a point in S_A and each C_i is a subblock of a point in S_C, including the empty block.

CLAIM 3. *Let $(y, z) \in S^*$. Let $i \le j$. If $z_k = 0$ for $i + 1 \le k \le j$ then $y[i, j] \in E$. In any case, $y[i, j]$ is a subblock of a concatenation of two E-blocks.*

Proof. We may assume $i < j$ (otherwise we are done). Assume $z_k = 0$ for $i+1 \le k \le j$. Choose $x \in S$ with $\phi(x)[i, j] = (y, z)[i, j]$. If $[i, j]$ is contained in a short marked interval of x, then $y[i, j]$ is a subblock of a point in S_B. If $[i, j]$ is contained in a long marked interval of x, then $y[i, j]$ is a subblock of $C_1 0 C_2 A C_3 0 C_4$ with C_i and A as in the definition of E, by Claim 2(b). If $z_k = 1$ for some $i + 1 \le k \le j$, then this k is unique by Claim 2(a). Now the above applies to the intervals $[i, k - 1]$ and $[k, j]$.

CLAIM 4. *There are no periodic points in the boundary S^*.*

Proof. Let $(y, z) \in S^*$. If $z_i = 1$ for some i, then it is for this i only by Claim 2(a), thus (y, z) is not periodic. Now assume $z_i = 0$ for all i and that y is periodic. Then Claim 3 implies that $y[-n, n]$ is a subblock of a point in S_A or S_B or S_C for each n. Thus y is contained in the union $S_A \cup S_B \cup S_C$. But by definition none of these shifts contains a periodic point, by contradiction.

CLAIM 5. $h_{\text{top}}(S^*) \le \max\{h_{\text{t}}(S), h_{\text{per}}(S)\} + \varepsilon$.

Proof. Since $z_i \ne 0$ for at most one index by Claim 2(a), the entropy of S^* equals the entropy of $pr_1(S^*) = \{y \mid (y, z) \in S^* \text{ for some } z\}$. By Claim 3 any subblock of a point $y \in pr_1(S^*)$ is a subblock of a concatenation of two E-blocks. Thus the entropy of $pr_1(S^*)$ is bounded by the growth rate of E-blocks. Let $E(n)$ denote the set of E-blocks of length n. The definition of E shows

$$\operatorname{card} E(n) \le \operatorname{card} B_n(S_B) + \left(1 + \sum_{m \le n} \operatorname{card} B_m(S_C)\right)^4 \left(1 + \sum_{m \le n} \operatorname{card} B_m(S_A)\right).$$

Since S_C has entropy 0, $h_{\text{top}}(S_B) \le h_{\text{per}}(S) + \varepsilon$, and $h_{\text{top}}(S_A) \le h_{\text{t}}(S) + \varepsilon$, we then have $\limsup_n 1/n \log \operatorname{card} E(n) \le \max\{h_{\text{t}}(S), h_{\text{per}}(S)\} + \varepsilon$. □

A *shift of finite type* is a compact Markov shift (see also [**DGS**, p. 118]). A shift of finite type is *mixing* if for any pair of non-empty open subsets U, V there is an $N \in \mathbf{N}$ such that $S^{-n}U \cap V$ is non-empty for all $n \ge N$.

THEOREM D. *Let S be a locally compact subshift and T a mixing shift of finite type. The following are equivalent.*

(1) *S is conjugate to a non-dense subsystem of T;*

(2) *S satisfies the periodic point and the entropy condition, i.e. for each n the number of periodic points of least period n in S is at most the number of periodic points of least period n in T, and* $\max\{h_{\text{t}}(S), h_{\text{per}}(S)\} < h_{\text{top}}(T)$.

This is an extension of Krieger's embedding theorem [**Kr**] since for compact subshifts $h_{\text{per}}(S) \le h_{\text{t}}(S)$, and any compact proper subsystem of T is non-dense.

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press

Proof. (1) implies (2). Let S be conjugate to a non-dense subsystem R of T. The periodic point condition is obvious. Let $\bar{R}$ denote the closure of R in T. Then $h_t(S) = h_t(R) \leq h_t(\bar{R}, P)$, where P is the zero-partition of $\bar{R}$. Since $h_t(\bar{R}, P) = h_{top}(\bar{R})$ for compact subshifts, we obtain $h_t(S) \leq h_{top}(\bar{R}) < h_{top}(T)$, the latter since $\bar{R}$ is a compact proper subsystem of T. Analogously, $h_{per}(S) \leq h_{per}(\bar{R}) \leq h_{top}(\bar{R}) < h_{top}(T)$.

(2) implies (1). By Theorem C there is an embedding of S into a compact subshift X with the same zeta function and $h_{top}(X) < h_{top}(T)$. Then apply Krieger's embedding theorem. (One could construct the embedding directly using the method of the proof of Theorem C). Then the image of S is contained in the image of X which is non-dense in T. □

Remark 6.2. If S is a locally compact subshift which is conjugate to a non-dense subsystem of a mixing shift of finite type T, then any embedding of S into T has a non-dense image, which can be seen as follows.

Let R be the image of some embedding. If R would be dense in T, then $T = \bar{R}$ would be a metric compactification of R. Thus R would be open in T (since R is locally compact [**D**, Theorem 8.3]). Now let μ be an ergodic measure on T with maximal entropy, i.e. $h_\mu(T) = h_{top}(T)$. Then μ has full support [**DGS**, pp. 130, 157]. Since R is open, $\mu(R) > 0$. Thus $\mu(R) = 1$ by ergodicity. Thus $h_\mu(R) = h_{top}(T)$, which implies $h_G(R) \geq h_{top}(T)$ and thus $h_t(R) \geq h_{top}(T)$ by Theorem 3.7, a contradiction to the entropy condition in (2).

REFERENCES

[B] R. Bowen. Entropy for group endomorphisms and homogenous spaces. *Trans. Amer. Math. Soc.* **153** (1971), 401–414.

[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.

[FF] D. Fiebig and U. Fiebig. Topological boundaries for countable state Markov shifts. *Proc. LMS*. III, **70** (1995), 625–643.

[G1] B. M. Gurevic. Topological entropy of enumerable Markov chains. *Dokl. Akad. Nauk SSSR* **187** (1969); English Transl. in *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.

[Gr] C. Grillenberger. Constructions of strictly ergodic systems I. Given entropy. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **25** (1973), 323–334.

[HK] M. Handel and B. Kitchens. Metrics and entropy for non-compact spaces. *Israel J. Math.* **91** (1995), 253–271.

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

[Sa] I. Salama. Topological entropy and recurrence of countable chains. *Pac. J. of Math.* **134** (1988), 325–341; errata in *Pac. J. of Math.* **140** (1989), 397.

[Se] M. Sears. Expansiveness on locally compact spaces. *Math. Sys. Th.* **7** (1973), 377–382.

[W] P. Walters. *An Introduction to Ergodic Theory*. GTM, Springer, 1981.

https://doi.org/10.1017/S0143385797069873 Published online by Cambridge University Press