Ergod. Th. & Dynam. Sys. (2003), **23**, 1785–1806
DOI: 10.1017/S0143385703000221

Entropy and preimage sets

DORIS FIEBIG†, ULF-RAINER FIEBIG† and ZBIGNIEW H. NITECKI‡

† *Institut für Mathematische Stochastik, Universität Göttingen, Lotzestrasse 13, 37083 Göttingen, Germany*
(e-mail: fiebig@math.uni-goettingen.de)
‡ *Department of Mathematics, Tufts University, Medford, MA 02155, USA*
(e-mail: zbigniew.nitecki@tufts.edu)

(*Received* 10 *October* 2002 *and accepted in revised form* 6 *May* 2003)

Abstract. We study the relation between topological entropy and the dispersion of preimages. Symbolic dynamics plays a crucial role in our investigation. For forward expansive maps, we show that the two pointwise preimage entropy invariants defined by Hurley agree with each other and with topological entropy, and are reflected in the growth rate of the number of preimages of a single point, called a preimage growth point for the map. We extend this notion to that of an entropy point for a system, in which the dispersion of preimages of an ε-stable set measures topological entropy. We show that for maps satisfying a weak form of the specification property, every point is an entropy point and that every asymptotically h-expansive homeomorphism (in particular, every smooth diffeomorphism of a compact manifold) has entropy points. Examples are given of maps in which Hurley's invariants differ and of homeomorphisms with no entropy points.

1. *Basic notation and statement of results*

The formulation by Bowen [**Bow71**] and Dinaburg [**Din70**] of topological entropy $h_{\text{top}}(f)$ for a continuous map f of a compact metric space X to itself can be summarized (following [**Bow78**, p. 17]) in an intuitively appealing way. Imagine studying X with measurements of high, but not absolute, resolution—points of X can be distinguished if and only if they are at least some fixed $\varepsilon > 0$ apart. Count the number of distinct points we can detect. Now let the map f act on X and repeat our measurements. Some previously indistinguishable points may be pushed apart by f, so the number of points we can detect grows over time. The rate at which this number grows is an indication of the dynamic complexity of f.

The apparent time-asymmetry of this formulation is illusory for an invertible system, since for any homeomorphism $h_{\text{top}}(f^{-1}) = h_{\text{top}}(f)$. It is less clear how to 'reverse time' in a general system and a number of different entropy-like invariants based on the preimage structure of a map have been formulated and studied in recent years [**LW91, LP92, NP99, Hur95**].

In this paper, we consider a pair of such invariants formulated by Hurley [**Hur95**]. We find that for forward expansive maps, both invariants agree with $h_{\mathrm{top}}(f)$ and in fact are captured by the preimage structure of a single point, which we call a *preimage growth point*. This observation suggests to us the possibility of 'localizing' the entropy of a more general class of maps—including some homeomorphisms—by reinterpreting and extending Hurley's invariants in terms of the preimage structure of stable sets. We are able to establish that when X has finite covering dimension, this invariant agrees with $h_{\mathrm{top}}(f)$, and when f is asymptotically h-expansive there exists an entropy point whose ε-stable sets yield our invariant. In particular, this applies to C^∞ diffeomorphisms of a compact manifold.

Throughout this paper, f is a continuous map of a compact metric space X to itself. We denote the base metric on X by d, but we also consider the Bowen–Dinaburg metrics generated by f,

$$d_n^f(x, y) := \max_{0\le i<n} d(f^i x, f^i y).$$

For $\varepsilon > 0$ and $n \in \mathbb{N} = \{0, 1, 2, \dots\}$, $n \ge 1$, a subset $S \subset X$ is (n, ε)-*separated* (with respect to f) if distinct points are spaced at least ε apart when using the metric d_n^f:

$$x, y \in S, x \ne y \Rightarrow d_n^f(x, y) > \varepsilon.$$

Compactness puts an upper bound on the cardinality of any (n, ε)-separated set in X; for any subset $K \subset X$, let

$$\max\operatorname{sep}(n, \varepsilon, K) := \max\{\#S \mid S \subset K \quad \text{and} \quad S \text{ is } (n, \varepsilon)\text{-separated}\}.$$

It is reasonable to expect the numbers $\max\operatorname{sep}(n, \varepsilon, K)$ to grow roughly exponentially with n; we can take the *exponential growth rate* of any sequence $\{c_n\}_{n\ge1}$ of reals $c_n \ge 0$,

$$GR\{c_n\} := \limsup_{n\to\infty} \frac{1}{n} \log c_n$$

(with $\log 0 := -\infty$). Topological entropy (originally formulated in terms of open covers [**AKM65**]) is characterized by Bowen [**Bow72**] and Dinaburg [**Din70**] as

$$h_{\mathrm{top}}(f) := \lim_{\varepsilon\to0} GR\{\max\operatorname{sep}(n, \varepsilon, X)\}.$$

It is well known that $h_{\mathrm{top}}(f)$ is an invariant of topological conjugacy—in particular, replacing d with any equivalent metric leads to the same value for $h_{\mathrm{top}}(f)$. Furthermore, if f is invertible, then S is (n, ε)-separated with respect to f iff $f^{n-1}S$ is (n, ε)-separated with respect to f^{-1}, yielding the equality $h_{\mathrm{top}}(f) = h_{\mathrm{top}}(f^{-1})$ for homeomorphisms.

When f is not invertible, the 'inverse' is set-valued, yielding the preimage set for $n \ge 0$,

$$f^{-n}x := \{z \in X \mid f^n z = x\}.$$

Hurley's invariants [**Hur95**] try to measure the maximum rate of dispersal of the preimage sets of individual points; they are called *pointwise preimage entropies* in [**NP99**]. The difference between these two invariants is when the maximization takes place:

$$h_m(f) := \lim_{\varepsilon\to0} GR\Big\{\max_{x\in X} \max\operatorname{sep}(n, \varepsilon, f^{-n}x)\Big\}$$
$$h_p(f) := \sup_{x\in X} \lim_{\varepsilon\to0} GR\{\max\operatorname{sep}(n, \varepsilon, f^{-n}x)\}.$$

The inequalities

$$h_p(f) \le h_m(f) \le h_{\text{top}}(f)$$

are clear; Hurley also found a general upper bound for $h_{\text{top}}(f)$ in terms of $h_m(f)$ and a second invariant [**Hur95**]. For f a homeomorphism, the pointwise preimage entropies are automatically both zero (and so, in general, different from $h_{\text{top}}(f)$). It was not known whether these preimage entropies ever differ from each other; in §7 we construct an example for which they do (see Example 7.1).

However, we find that these pointwise preimage entropies agree for a significant family of maps. A map $f : X \to X$ is forward expansive with expansiveness constant $c > 0$ if

$$x, y \in X \quad \text{with } d(f^n x, f^n y) \le c \quad \text{for all } n \in \mathbb{N} \text{ implies } x = y.$$

This condition is independent of the base metric d.

Our first main result is that for forward expansive maps, both preimage entropies agree with topological entropy and there is a point whose preimage sets grow at precisely this rate. The statement can be made clearer if we note that if f is forward expansive with constant c, then for every $x \in X$ and $n \ge 1$ the set $f^{-n}x$ is (n, ε)-separated for any $0 < \varepsilon < c$, so that max sep$(n, \varepsilon, f^{-n}x)$ is just the cardinality $\# f^{-n}x$.

PROPOSITION 4.5. *Let $f : X \to X$ be forward expansive with $h_{\text{top}}(f) = \log\lambda$. Then there is $x \in X$ with $\# f^{-n}x \ge \lambda^n$ for all n and, in particular,*

$$h_p(f) = h_m(f) = h_{\text{top}}(f).$$

We call a point $x \in X$ with $GR\{\# f^{-n}x\} = h_{\text{top}}(f)$ a *preimage growth point* for f.

The proof of $h_m(f) = h_{\text{top}}(f)$ is based on the combinatorial Lemma 2.1, while the proof of $h_p(f) = h_{\text{top}}(f)$ follows from the more general Theorem 4.1. For this, note that preimage sets under a homeomorphism are singletons, so preimage entropy is automatically zero in this case. However, a modification of the definitions of preimage entropy yields a non-trivial extension to a class of maps that also includes some homeomorphisms. Recall that, given $\varepsilon > 0$, the ε-stable set of x under f is the set of points whose forward orbit ε-shadows that of x:

$$S(f, x, \varepsilon) := \{y \in X \mid d(f^n x, f^n y) \le \varepsilon \quad \text{for all } n = 0, 1, 2, \dots\}.$$

The preimages of these sets can be non-trivial and hence can disperse at a non-zero exponential rate. Given $x \in X$ and $\varepsilon > 0$, consider the dispersal rate

$$h_s(f, x, \varepsilon) := \lim_{\delta \to 0} GR\{\max \operatorname{sep}(n, \delta, f^{-n} S(f, x, \varepsilon))\}.$$

We show that forward expansive maps exist only on finite-dimensional spaces (adapting an argument of Mañé). Then we show that in any finite-dimensional space, the topological entropy of a map can be calculated in terms of the dispersion of preimages of ε-stable sets.

THEOREM 4.1. *If $f : X \to X$ is continuous and X is a compact metric space of finite covering dimension, then*

$$\sup_{x \in X} h_s(f, x, \varepsilon) = h_{\text{top}}(f) \quad \textit{for all } \varepsilon > 0.$$

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

(We do not know whether the finite-dimensionality hypothesis is necessary.) This result leads to two extensions of our earlier notion of preimage entropy point. For $\varepsilon > 0$ we call $x \in X$ an *ε-entropy point* for f if $h_s(f, x, \varepsilon) = h_{\text{top}}(f)$; it is simply an *entropy point* if $\lim_{\varepsilon\to 0} h_s(f, x, \varepsilon) = h_{\text{top}}(f)$. (Note, in particular, that this implies that $f^{-n}x \neq \emptyset$ for all n.) Since $h_s(f, x, \varepsilon)$ is decreasing in ε, an entropy point is an ε-entropy point for each $\varepsilon > 0$. Note that while the notion of ε-entropy point depends on the choice of metric, that of entropy point does not. For forward expansive maps, any preimage growth point is an ε-entropy point for all sufficiently small $\varepsilon > 0$. Therefore, by Proposition 4.5, forward expansive maps do have entropy points. Examples in §7 show that, in general, the existence of ε-entropy points for each $\varepsilon > 0$ is not automatic, nor does such existence for all $\varepsilon > 0$ guarantee the existence of an entropy point. However, we are able to establish the existence of entropy points for maps which satisfy a very weak specification property (Definition 3.1), as well as for a large class of maps, defined by Misiurewicz [**Mis76**] (extending ideas of Bowen [**Bow72**]), as follows.

We can apply the calculations involved in defining topological entropy to any subset $K \subset X$. Let

$$h_{\text{top}}(f, K) := \lim_{\delta\to 0} GR\{\max \operatorname{sep}(n, \delta, K)\}.$$

Bowen [**Bow72**] calls a map *h-expansive* if $\sup_{x\in X} h_{\text{top}}(f, S(f, x, \varepsilon)) = 0$ for some $\varepsilon > 0$, while Misiurewicz [**Mis76**] calls f *asymptotically h-expansive* if

$$\lim_{\varepsilon\to 0} \sup_{x\in X} h_{\text{top}}(f, S(f, x, \varepsilon)) = 0.$$

A map is asymptotically h-expansive precisely if its topological conditional entropy vanishes [**Mis76**]. Every C^∞ diffeomorphism of a compact mainfold to itself has this property [**Buz97**]. Then applying a result from [**BFF02**] to the natural extension, that any asymptotically h-expansive homeomorphism is a factor of some two-sided subshift, together with Proposition 2.2, we establish that any asymptotically h-expansive map has entropy points.

THEOREM 6.4. *If f is an asymptotically h-expansive map then there exist entropy points for f (even for the map restricted to its eventual image).*

Proofs of these results involve the interplay between symbolic and topological dynamics.

Recall the definition of the two-sided full shift on N symbols: let $\mathcal{A}$ be an 'alphabet' with N 'letters' (that is, $\mathcal{A} = \{1, 2, \ldots, N\}$) and form the space $\mathcal{A}^{\mathbb{Z}}$ of bisequences $x = (x_i)_{i\in\mathbb{Z}}$, $x_i \in \mathcal{A}$ with the product of the discrete topology on $\mathcal{A}$. This is a compact metric space. The shift map $g : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ defined by $(gx)_i = x_{i+1}$ for all $i \in \mathbb{Z}$ is a homeomorphism. A two-sided subshift is the restriction of g to some closed subset $X \subset \mathcal{A}^{\mathbb{Z}}$ with $gX = X$. If we denote the subwords of $x \in \mathcal{A}^{\mathbb{Z}}$ by $x[n, m] := x_n x_{n+1} \ldots x_m$ for $n \le m$, a two-sided subshift can always be specified by a list $\mathcal{V}$ of 'forbidden words':

$$x \in X \iff x[n, m] \notin \mathcal{V} \quad \text{for all } n \le m.$$

Any subshift is expansive: for some $c > 0$ we have

$$x, y \in X \quad \text{with } d(g^n x, g^n y) \le c \quad \text{for all } n \in \mathbb{Z} \text{ implies } x = y.$$

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

Truncation of negative-index entries gives the one-sided full shift on N symbols $f : \mathcal{A}^{\mathbb{N}} \to \mathcal{A}^{\mathbb{N}}$, which is a factor of g but is N-to-one instead of a homeomorphism. A one-sided subshift is the restriction of f to a closed set $Y \subset \mathcal{A}^{\mathbb{N}}$ with $fY \subset Y$ (a one-sided subshift need not be onto). It is always forward expansive. A one-sided subshift is a factor of some two-sided subshift iff $fY = Y$.

The various entropy invariants reduce to counting words in the case of subshifts. It is well known that the topological entropy of a subshift $f : X \to X$, one-sided or two-sided, equals $GR\{\#W_n\}$, where W_n is the set of admissible words of length n, i.e. the set of words of length n which occur as subwords in some point of X. Similarly, if f is a one-sided subshift, then $f^{-n}x = \{wx \mid w \in W_n \text{ and } wx \in X\}$. This is captured in the nth *predecessor set* of $x \in X$, defined as

$$P_n(x) := \{w \in W_n \mid wx \in X\}.$$

Expansiveness implies that $f^{-n}x$ is (n, ε)-separated for all small $\varepsilon > 0$ and so

$$h_m(f) = GR\Big\{\max_{x\in X} \#P_n(x)\Big\} \quad \text{and} \quad h_p(f) = \sup_{x\in X} GR\{\#P_n(x)\}.$$

We begin with a tree lemma and use it to prove the existence of preimage growth points for one-sided subshifts.

PROPOSITION 2.2. *Let $f : X \to X$ be a one-sided subshift (not necessarily onto) with $h_{\text{top}}(f) = \log\lambda$. Then there is $x \in X$ with*

$$\#P_n(x) \geq \lambda^n \quad \textit{for all } n \geq 1.$$

In particular, x is a preimage growth point for f and $h_p(f) = h_m(f) = h_{\text{top}}(f)$.

This implies that two-sided subshifts do have entropy points. The referee has pointed out to us that, using a conditional Breiman theorem and disintegration of measures, he can show that for subshifts the set of entropy points has measure 1 with respect to each measure of maximal entropy. So from a measure theoretical point of view, the set of entropy points is always large.

There are, however, synchronized subshifts, not of finite type, for which the entropy points form a subshift of finite type. In this case, the set of entropy points is the support of the unique measure of maximal entropy, but is topologically small (first category) and fails to reveal the non-shift of type nature of the synchronized shift. This example will be given in a forthcoming paper.

We prove Proposition 2.2 by a simple combinatorial argument. In §4 we extend this result to forward expansive maps.

2. *Preimage growth points*

In this section we establish the equality $h_p(f) = h_m(f)$ for any forward expansive map and show that in the particular case of one-sided subshifts, these invariants agree with $h_{\text{top}}(f)$. This is done by finding a point for which the preimage sets grow at the required rate. Both results are applications of a result on the growth of branches in a tree.

A *tree rooted at* v_0 is a directed graph $T = T(v_0)$ with a distinguished vertex v_0 (the *root*) such that for every vertex $v \in V(T)$ of T there is a unique path from v_0 to v; the number $|v|$ of edges in this path stratifies T into *levels*:

$$V_k(v_0) := \{v \in V(T) \mid |v| = k\}, \quad k \in \mathbb{N}.$$

For any vertex $v \in V(T)$, we can take v as the root of the subtree $T(v)$ whose vertices can be reached from v; the levels $V_k(v)$ of this subtree are numbered relative to v, so

$$V_k(v) := \{v' \in V(T(v)) \mid |v'| = |v| + k\}.$$

Note that $\#V_1(v)$ is the outdegree of v in T. Our result states that a lower bound on $\#V_N(v_0)$ yields a vertex $v \in V(T)$ for which the sizes $\#V_k(v)$ of all (low) levels in the associated subtree can be *a priori* estimated from below.

LEMMA 2.1. (Tree Lemma) *Suppose $T = T(v_0)$ is a tree all of whose vertices have outdegree at most M, with $\#V_N(v_0) > \lambda^N$ for some $N \in \mathbb{N}$ and $\lambda \geq 1$. Then for every $1 \leq k \leq N$ satisfying*

$$\left(\frac{M}{\lambda}\right)^{k-1} \leq \frac{\#V_N(v_0)}{\lambda^N} \tag{$*$}$$

there exists a vertex $v \in V(T)$ such that

$$\#V_i(v) \geq \lambda^i \quad \textit{for } 1 \leq i \leq k.$$

Proof. Deleting all vertices in levels $n < N$ which do not lead to a vertex in level N, we obtain a tree (again called T) with no dead ends at any level $n < N$, but still with $\#V_N(v_0) > \lambda^N$. Suppose the statement of the lemma is not true—that is, for some $1 \leq k \leq N$ satisfying $(*)$ and every vertex $v \in V(T)$ there is an integer $k(v) \leq k$ such that

$$\#V_{k(v)}(v) < \lambda^{k(v)}. \tag{$**$}$$

Let

$$W := \{v \in V(T) \mid N - k < |v| \leq N\}.$$

Set $Y_0 = \{v_0\}$ and define subsets $Y_j \subset V(T)$ inductively by

$$Y_{j+1} := (Y_j \cap W) \cup \bigcup_{u \in Y_j \setminus W} V_{k(u)}(u).$$

We claim the numbers

$$l_j := \sum_{y \in Y_j} \lambda^{-|y|}$$

are non-increasing in j: to see this, note that

$$\begin{aligned} l_{j+1} &= \sum_{w \in Y_j \cap W} \lambda^{-|w|} + \sum_{u \in Y_j, u \notin W} \sum_{v \in V_{k(u)}(u)} \lambda^{-|v|} \\ &= \sum_{w \in Y_j \cap W} \lambda^{-|w|} + \sum_{u \in Y_j, u \notin W} \#V_{k(u)}(u) \cdot \lambda^{-(k(u)+|u|)} \\ &\leq \sum_{w \in Y_j \cap W} \lambda^{-|w|} + \sum_{u \in Y_j, u \notin W} \lambda^{-|u|} \\ &= l_j. \end{aligned}$$

The inequality in the third line is strict, if $Y_{j+1} \neq Y_j$, by (∗∗). Since $k \leq N$, we have $l_j \leq l_1 < l_0 = 1$ for all $j \geq 1$. Now since $k(u) \leq k \leq N$ the sets Y_j are contained in the finite set $\{v \in V(T) \mid |v| \leq N\}$ and if $\min\{|u| \mid u \in Y_j\} = n \leq N - k$, then $\min\{|u| \mid u \in Y_{j+1}\} > n$; thus for some J we have $Y_J \subset W$ and hence $Y_{J+1} = Y_J$. By construction, for any j and $v \in V_N(v_0)$ the path from v_0 to v hits a unique vertex in Y_j. Since $k(u) \leq k$ we can find for every $v \in V_N(v_0)$ some $u \in Y_J$ so that $v \in V_{N-|u|}(u)$ and $N - |u| < k$. Therefore,

$$\#V_N(v_0) = \sum_{u \in Y_J} \#V_{N-|u|}(u) \leq \sum_{u \in Y_J} M^{N-|u|} = \sum_{u \in Y_J} \lambda^{(N-|u|)} \cdot \left(\frac{M}{\lambda}\right)^{N-|u|}$$
$$\leq \sum_{u \in Y_J} \lambda^{(N-|u|)} \cdot \frac{\#V_N(v_0)}{\lambda^N} = \#V_N(v_0) \cdot l_J < \#V_N(v_0),$$

where the inequality comes from (∗) and the strict inequality from $l_J < 1$; but this contradiction proves the lemma. □

We apply the Tree Lemma 2.1 to establish the existence of preimage growth points (and thus also entropy points) for any one-sided subshift. We use the notation from §1.

PROPOSITION 2.2. *Let $f : X \to X$ be a one-sided subshift (not necessarily onto) with $h_{\mathrm{top}}(f) = \log \lambda$. Then there exists $x \in X$ such that for all $n \geq 1$*

$$\# f^{-n} x = \# P_n(x) \geq \lambda^n.$$

In particular, x is a preimage growth point for f and

$$h_p(f) = h_m(f) = h_{\mathrm{top}}(f).$$

Proof. We will denote $\#W_n$ by b_n, $\#P_n(x)$ by $\varphi_n(x)$ and for $w \in W_m$ let $\varphi_n(w)$ denote $\#\{u \in W_n \mid uw \in W_{n+m}\}$. By compactness there is $x \in X$ with $\varphi_n(x) \geq 1$ for all n, so we can assume $\lambda > 1$. Since $h_{\mathrm{top}}(f) = GR\{\#W_n\}$ and the sequence $\{b_n\}$ is submultiplicative, its growth rate is an infimum; in particular, $b_n \geq \lambda^n$ for all $n \geq 1$. Pick a sequence $\{\lambda_k\}$ strictly increasing to λ with $\lambda_1 > 1$. For each k there is $N(k) \geq k$ such that

$$\left(\frac{b_1}{\lambda_k}\right)^{k-1} \leq \left(\frac{\lambda_{k+1}}{\lambda_k}\right)^{N(k)}. \tag{∗}$$

Fix $M \geq 1$. For each $n \geq 1$, since

$$\lambda^{n+M} \leq b_{n+M} = \sum_{w \in W_M} \varphi_n(w)$$

there exists a word $w_n \in W_M$ with $\varphi_n(w_n) \geq \lambda^{n+M}/b_M \geq \lambda^n/b_M$. Note that this last fraction grows faster than λ_{k+1}^n for any fixed k and so since W_m is finite, there exists $w \in W_M$ and $n(k) \geq N(k)$ satisfying

$$\varphi_{n(k)}(w) \geq \lambda_{k+1}^{n(k)} \quad \text{for all } k. \tag{∗∗}$$

Form a tree $T := T(w)$ with $V_n(T) = P_n(w)$ and an edge from $\alpha \in P_n(w)$ to $\beta \in P_{n+1}(w)$ iff $\beta = a\alpha$ for some $a \in W_1$ and for each $a \in P_1(w)$ there is an edge

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

from w to aw. All vertices have outdegree at most b_1 and $V_i(T) = \varphi_i(w)$, so $(*)$ and $(**)$ give

$$\left(\frac{b_1}{\lambda_k}\right)^{k-1} \leq \frac{\#V_{n(k)}(T)}{\lambda_k^{n(k)}}.$$

Thus the Tree Lemma 2.1 applies to give for each k a vertex $v_k \in V(T)$ with

$$\#V_i(v_k) \geq \lambda_k^i \quad \text{for } 1 \leq i \leq k.$$

Note that $v_k w \in W_{|v_k|+M}$. Passing to a subsequence k_j, we can assume that all $v_{k_j} w$ begin with the same word $v \in W_M$. Fix $i \in \mathbb{N}$. Then for all j with $k_j \geq i$ we have $\varphi_i(v) \geq \varphi_i(v_{k_j} w) = \#V_i(v_{k_j}) \geq \lambda_{k_j}^i$ and thus $\varphi_i(v) \geq \lambda^i$. So far, we have found for each $M \geq 1$ a word $v_M \in W_M$ with $\varphi_i(v_M) \geq \lambda^i$ for all i. Passing to a subsequence, we can assume that these are initial words of a convergent sequence of points in X with limit $x \in X$. Since $\varphi_n(x) = \lim_{m\to\infty} \varphi_n(x[0, m))$ and $\varphi_n(x[0, m)) \geq \varphi_n(v_M)$ if $x[0, m)$ is an initial word of v_M, the limit point x satisfies the condition required by the proposition. □

The equality $h_p(f) = h_m(f)$, which holds for one sided subshifts by the preceding result, can fail for a surjective map on a zero dimensional space (Example 7.1). However, we can formulate general conditions which guarantee that it holds.

A map $f : X \to X$ has *uniform separation of preimages* if for some $\varepsilon > 0$, $d(x, y) \leq \varepsilon$ and $fx = fy$ implies $x = y$. This has two consequences: for any $n \geq 1$, the preimage set $f^{-n}x$ is (n, ε)-separated and its cardinality $\varphi_n(x) := \#f^{-n}x$, as a function of $x \in X$, is upper semicontinuous: $\varphi_n(x) \geq \limsup_i \varphi_n(x_i)$ whenever $x_i \to x$. To see the latter, suppose $\varphi_n(x_i) = k$ for all i. Let $\{z_{i,1}, \ldots, z_{i,k}\} = f^{-n}x_i$ for $i = 1, 2, \ldots$; passing to a subsequence we can assume that for each j the sequence $z_{i,j}$ converges to $z_j \in f^{-n}x$. However, for $j \neq j'$, $d(z_{i,j}, z_{i,j'}) > \varepsilon$ for all $i \Rightarrow d(z_j, z_{j'}) \geq \varepsilon$. Thus $\varphi_n(x) \geq k$.

Note, in particular, that every forward expansive map has these properties. We establish $h_p(f) = h_m(f)$ whenever f has uniform separation of preimages. The core of the argument is that the upper semicontinuity of $\varphi_n(x)$ for every $n \geq 1$ implies the existence of a point with $GR\{\varphi_n(x)\} = h_m(f)$. Recall that an upper-semicontinuous function achieves its maximum on any compact set and if $\varphi_1(x) \leq M$ for all $x \in X$, then $\varphi_n(x) \leq M^n$.

PROPOSITION 2.3. *Suppose $f : X \to X$ is continuous and each of the functions $\varphi_n(x) = \#f^{-n}x$ is upper semicontinuous. If $h_m(f) = \log\lambda$, then there exists $x \in X$ with*

$$\varphi_n(x) \geq \lambda^n \quad \textit{for all } n.$$

Proof. The proposition is trivial if $\lambda = 1$, since any point in the eventual image $\bigcap f^n X$ works; so assume $\lambda > 1$. Set $a_n := \max_{x\in X} \varphi_n(x)$; the sequence $\{a_n\}$ is submultiplicative, so $\log\lambda = \inf_n(\log a_n)/n$ and hence $a_n \geq \lambda^n$ for each n. Fix a sequence $\{\lambda_k\}$ strictly increasing to λ with $\lambda_1 > 1$ and note that for each k we can find $N(k) \geq k$ such that $(a_1/\lambda_k)^{k-1} \leq a_{N(k)}/\lambda_k^{N(k)}$. Pick $x_k \in X$ with $\varphi_{N(k)}(x_k) = a_{N(k)}$ and form a tree T_k rooted at x_k with $V_i(T_k) = f^{-i}x_k$ and an edge from $f(z)$ to z. The Tree Lemma 2.1 gives z_k with $\varphi_i(z_k) \geq \lambda_k^i$ for $1 \leq i \leq k$, so the set $X_k := \{x \in X \mid \varphi_i(x) \geq \lambda_k^i$ for $1 \leq i \leq k\}$ is non-empty and closed (by the upper semicontinuity of φ_i). Since $\lambda_{k+1} > \lambda_k$, $X_{k+1} \subset X_k$ and

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

so by compactness $\bigcap_{k\geq 1} X_k \neq \emptyset$. Since $\lambda_k \to \lambda$, any point $x \in \bigcap_{k\geq 1} X_k$ has $\varphi_n(x) \geq \lambda^n$ for all n. □

Using the fact that $f^{-n}x$ is (n, ε)-separated if f has uniform separation of preimages, we obtain as an immediate corollary the following.

COROLLARY 2.4. *If $f : X \to X$ is continuous with uniform separation of preimages (in particular, if f is forward expansive) and $h_m(f) = \log \lambda$, then there is a point $x \in X$ with*

$$\# f^{-n}x \geq \lambda^n \quad \text{for all } n$$

and, in particular,

$$h_p(f) = h_m(f).$$

In §4 we will see, in fact, that $h_p(f) = h_m(f) = h_{\text{top}}(f)$ for forward expansive maps.

3. *Entropy points*

In this section we study the set $\mathcal{E}(f)$ of entropy points for a map $f : X \to X$. Proposition 2.2 shows that this set is non-empty for any one-sided subshift. A natural question is, how large can $\mathcal{E}(f)$ be in general?

A non-mixing version of weak specification [**DGS76**, Definition 21.1] is that for every $\varepsilon > 0$, there is $N(\varepsilon) \in \mathbb{N}$ such that, given any pair of points $x, y \in X$ and $n, m \geq 1$ there exists a point $z \in X$ and integers $0 \leq k, l < N(\varepsilon)$ such that

$$d_n^f(y, z) \leq \varepsilon, \quad d_m^f(x, f^{n+k}z) \leq \varepsilon \quad \text{and} \quad f^p z = z \quad \text{where } p = n + k + m + l.$$

A weakening of that definition implies $\mathcal{E}(f) = X$.

Definition 3.1. The map $f : X \to X$ has the very weak specification property if for every $\varepsilon > 0$, there is $N(\varepsilon) \in \mathbb{N}$ such that, given any pair of points $x, y \in X$ and $n, m \geq 1$ there exists a point $z \in X$ and an integer $0 \leq k < N(\varepsilon)$ such that

$$d_n^f(y, z) \leq \varepsilon \quad \text{and} \quad d_m^f(x, f^{n+k}z) \leq \varepsilon.$$

PROPOSITION 3.2. *For any continuous map $f : X \to X$ with the very weak specification property, $h_s(f, x, \varepsilon) = h_{\text{top}}(f)$ for every $x \in X$ and every $\varepsilon > 0$. In particular, f is surjective and $\mathcal{E}(f) = X$.*

Proof. By compactness, for every pair of points $x, y \in X$ and $n \geq 1$ and $\varepsilon > 0$ there exists $0 \leq k < N(\varepsilon)$ and a point $z \in f^{-(n+k)}S(f, x, \varepsilon)$ with $d_n^f(y, z) \leq \varepsilon$.

Now fix $x \in X$, $\varepsilon > 0$, and given $0 < \delta < \varepsilon$, let E_n be a maximal $(n, 3\delta)$-separated subset of X. Then for some $k = k(n)$ with $0 \leq k < N(\delta)$ there is a subset $F_n \subset E_n$ with $\#F_n \geq \#E_n/N(\delta)$; for every $y \in F_n$ we can find $z = z_y \in f^{-(n+k)}S(f, x, \varepsilon)$ with $d_n^f(y, z) < \delta$. Then the set $G_n := \{z_y \mid y \in F_n\}$ is (n, δ)-separated, so also $(n + k, \delta)$-separated, and $\#G_n = \#F_n$. Thus for each $n \geq 1$,

$$\max \text{sep}(n + k(n), \delta, f^{-n-k(n)}S(f, x, \varepsilon)) \geq \frac{\max \text{sep}(n, 3\delta, X)}{N(\delta)}.$$

Since $k(n) \leq N(\delta)$ for all n, this proves $GR\{\max \text{sep}(n, \delta, f^{-n}S(f, x, \varepsilon))\} \geq GR\{\max \text{sep}(n, 3\delta, X)\}$. Taking $\delta \to 0$ shows $h_s(f, x, \varepsilon) \geq h_{\text{top}}(f)$. □

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

The induced action on the basic sets of Axiom A diffeomorphisms is forward transitive and has the pseudo-orbit tracing property, which immediately implies the very weak specification property. It is easily seen that transitive shifts of finite type (SFT) have the non-mixing version of the very weak specification property and that this property is preserved under taking factors. Thus sofic shifts also belong to the uncountable class of subshifts with the non-mixing very weak specification property. It is known that for subshifts the very weak specification property is equivalent to the notion of specification. The same argument shows this is also true for the non-mixing versions. Any subshift satisfying the non-mixing version of the very weak specification property is synchronized. This is not true in general, since any irrational rotation of the circle has the very weak specification property too.

We know that $\mathcal{E}(f) \neq \emptyset$ for (one-sided) subshifts; we show that it can be a proper subset of X. In §7 we give two examples that are even more extreme: in general $\mathcal{E}(f)$ can be empty.

Example 3.3. (A transitive subshift with $\emptyset \neq \mathcal{E}(f) \neq X$.) Let $A = \{1, \alpha, \beta\}$ and let $C := \{1\alpha^n\beta^n \mid n \geq 0\}$. Let $X \subset A^{\mathbb{Z}}$ be the closure of the set of points $x \in \mathrm{A}^{\mathbb{Z}}$ which can be written as a bi-infinite concatenation of words from C and let $f : X \to X$ be the left shift map. Since f is expansive, for all small $\varepsilon > 0$, $h_s(f, x, \varepsilon) \leq GR\{\#P_n(x[0, \infty))\}$, so $h_s(f, 1^\infty, \varepsilon) = h_{\mathrm{top}}(f) > 0$ and $h_s(f, x, \varepsilon) = 0$ whenever there is some $n \in \mathbb{Z}$ with $x_i = \beta$ for all $i \geq n$. If $x_i \neq \beta$ for infinitely many $i \geq 0$ then, given $\varepsilon > 0$, there is $y \in S(f, x, \varepsilon)$ and $k < 0$ with $y_k = 1$. Since $P_n(y[k, \infty)) = P_n(1^\infty)$ for any such y and $n \geq 1$, $h_s(f, x, \varepsilon) = h_s(f, 1^\infty, \varepsilon) = h_{\mathrm{top}}(f)$.

In this example, the set $\mathcal{E}(f)$ is a residual set. However, there are also transitive subshifts, with $\mathcal{E}(f)$ a non-empty set of first category. Furthermore, it is interesting how the size of the set $\mathcal{E}(f)$ is related to the existence of a maximal measure. We know examples with a non-empty set $\mathcal{E}(f)$ which do not have a measure of maximal entropy, and there are examples with infinite entropy, so having a measure of maximal entropy, with no entropy points. However, so far we do not know if in the finite entropy case the existence of a measure of maximal entropy implies the existence of entropy points. These issues will be treated in a forthcoming paper.

LEMMA 3.4. *Suppose $f : X \to X$ is continuous and $A \subset X$ is a subset with $f^{-n}A \neq \emptyset$ for all $n \in \mathbb{N}$. Then for all $\varepsilon > 0$*

$$GR\{\max \operatorname{sep}(n-1, \varepsilon, f^{-n}A)\} = GR\{\max \operatorname{sep}(n, \varepsilon, f^{-n}A)\}.$$

Proof. The inequality '$\leq$' is clear. To prove '$\geq$', cover X with open balls $B_1, \dots, B_r$ of radius $\varepsilon/2$ and for $E \subset f^{-n}A$, set $E_k = E \cap f^{-n}B_k$, $1 \leq k \leq r$. If E is (n, ε)-separated then each E_k is $(n-1, \varepsilon)$-separated and at least one k satisfies $\#E_k \geq \#E/r$; the desired inequality follows. □

For any $x \in \mathcal{E}(f)$, $f^{-n}x \neq \emptyset$ for all $n \geq 0$, so $\mathcal{E}(f)$ is always contained in the eventual range $X_\infty := \bigcap_{n\geq 1} f^n X$. However, we can say more.

PROPOSITION 3.5. *For any continuous map $f : X \to X$ (not necessarily surjective), $f(\mathcal{E}(f)) = \mathcal{E}(f)$.*

Proof. The inclusion $f(\mathcal{E}(f)) \subset \mathcal{E}(f)$ holds because $S(f, x, \varepsilon) \subset f^{-1}S(f, fx, \varepsilon)$ for all $x \in X$ and $\varepsilon > 0$. To establish the opposite inclusion (which is easy if f is bijective), suppose $x \in \mathcal{E}(f)$; by definition the set $f^{-1}x$ is non-empty (and of course compact). Now suppose that x is an entropy point. We construct a preimage y which is also an entropy point.

CLAIM. *For all $k > 0$, $f^{-1}x$ contains a $1/k$-entropy point of f.*

Proof. Choose a finite subset E of $K := f^{-1}x$ such that the open balls $B_{1/k}(z) := \{y \in X \mid d(y, z) < 1/k\}$, $z \in E$ cover K, and let U denote their union. Pick $0 < \eta < 1/k$ so that $f^{-1}B_\eta(x) \subset U$, and given $\delta > 0$ for $n \geq 1$, pick F_n a maximal (n, δ)-separated subset of $f^{-(n+1)}S(f, x, \eta)$; then $f^n F_n \subset U$ and we can partition F_n into sets $F_{n,z}$ so that $f^n F_{n,z} \subset B_{1/k}(z)$ for each $z \in E$. Since $fz = x$, $\eta < 1/k$ and $F_{n,z} \subset f^{-(n+1)}S(f, x, \eta)$, we obtain $F_{n,z} \subset f^{-n}S(f, z, 1/k)$, and so for each $z \in E$

$$\#F_{n,z} \geq \max\operatorname{sep}\left(n, \delta, f^{-n}S\left(f, z, \frac{1}{k}\right)\right). \tag{$*$}$$

Now, E is finite so we can pick $z \in E$ such that $\#F_{n,z} \geq \#F_n/\#E$ for infinitely many n, hence $GR\{\#F_{n,z}\} = GR\{F_n\}$. However, since $x \in \mathcal{E}(f)$, this together with $(*)$ and Lemma 3.4 gives for each $\delta > 0$ some element $z \in E$ (depending on δ) such that

$$GR\left\{\max\operatorname{sep}\left(n, \delta, f^{-n}S\left(f, z, \frac{1}{k}\right)\right)\right\} \geq h_{\mathrm{top}}(f)$$

and the finiteness of E lets us pick z independent of δ, ensuring that z is a $1/k$-entropy point for f proving the claim. □

However, taking $z_k \in f^{-1}x$ a $1/k$-entropy point, we can assume that $z_k \to z \in f^{-1}x$; given $\varepsilon > 0$ there is some k such that $S(f, z_k, 1/k) \subset S(f, z, \varepsilon)$ which shows that z is an ε-entropy point of f, hence $z \in f^{-1}x \cap \mathcal{E}(f)$. □

When f is not surjective, one can compare the action of f on X with its restriction $g := f|_{X_\infty}$ to the eventual range: clearly g is surjective and $h_{\mathrm{top}}(f) = h_{\mathrm{top}}(g)$, so $\mathcal{E}(g) \subset \mathcal{E}(f)$. These sets can be different as we show in a forthcoming paper.

4. *Finite-dimensional systems*

In this section we show that if X has finite covering dimension, then for any continuous map $f : X \to X$, $\sup_{x \in X} h_s(f, x, \varepsilon)$ is independent of ε and equals $h_{\mathrm{top}}(f)$. (The supremum might not be attained, as shown in Examples 7.2 and 7.3.) As a consequence, we establish the equality $h_m(f) = h_{\mathrm{top}}(f)$ for forward expansive maps. We note in passing that the analogous quantity obtained by replacing the ε-stable set of x with the ε ball about x can be shown to achieve a maximum equal to $\mathrm{h_{top}}(f)$ by standard arguments.

If $\mathcal{U}$ is a finite cover of X, we define its *mesh*, $\mathrm{mesh}(\mathcal{U})$, with respect to any metric as the maximum diameter of its elements. For any subset $S \subset X$, the *multiplicity* of $\mathcal{U}$ on S is the number of elements of $\mathcal{U}$ intersecting S; the *order* of $\mathcal{U}$ is its maximal multiplicity on singleton sets $S = \{x\}$, $x \in X$, and the *star order* of $\mathcal{U}$ is its maximal multiplicity on its own elements $S \in \mathcal{U}$. The space X has covering dimension at most n if there exist open covers of X with arbitrarily small mesh and order bounded by $n+1$. This notion is independent of the metric used [**Eng78**, Theorem 1.6.12] and replacing 'open' with 'closed' gives an equivalent condition [**Eng78**, Proposition 3.1.3]. Finite covering dimension can be characterized by the condition that X has open (respectively closed) covers of arbitrarily small mesh whose orders (respectively star orders) are bounded.

THEOREM 4.1. *If $f : X \to X$ is continuous and X is a compact metric space of finite covering dimension, then*

$$\sup_{x\in X} h_s(f, x, \varepsilon) = h_{\mathrm{top}}(f) \quad \textit{for all } \varepsilon > 0.$$

Proof. Suppose X has covers of arbitrary small mesh whose star orders are bounded by K.

CLAIM. *Let $\varepsilon > 0$. Then* $\sup_{x\in X} h_s(f, x, \varepsilon) \geq h_{\mathrm{top}}(f) - \log K$.

Proof. Pick $0 < \delta < \varepsilon/2$ and let $GR\{\max \mathrm{sep}(n, \delta, X)\} = \log \lambda$. Suppose $\mathcal{A} = \{A_1, \ldots, A_N\}$ is a closed cover of X with mesh $< \delta$ and star order $\leq K$. A given $x \in X$ has least one itinerary with respect to $\mathcal{A}$. For any $y = (y_i)_{i\geq 0} \in \{1, \ldots, N\}^{\mathbb{N}}$ define $\pi(y) := \bigcap_{i=0}^{\infty} f^{-i} A_{y_i}$; the set

$$Y := \{y = (y_i)_{i\in\mathbb{N}} \in \{1, \ldots, N\}^{\mathbb{N}} \mid \pi(y) \neq \emptyset\}$$

defines a one-sided subshift $g : Y \to Y$ on N symbols. Since $\mathcal{A}$ has mesh less than δ, if $x \in \pi(y)$ and $x' \in \pi(y')$, then $d(f^i x, f^i x') > \delta$ implies $y_i \neq y'_i$. Thus the number of allowed words for g satisfies $\#W_n \geq \max \mathrm{sep}(n, \delta, X)$ and so

$$h_{\mathrm{top}}(g) \geq \log \lambda.$$

Proposition 2.2 gives $y \in Y$ with $\#g^{-n} y \geq \lambda^n$ for all $n \geq 1$; pick $x \in \pi(y)$ and fix $n \geq 1$. Form $E_n \subset X$ by picking a point $z \in \pi(u)$ for each $u \in g^{-n} y$. For any such $z \in E_n$, $f^n z \in \pi(y)$, so, since $\mathcal{A}$ has mesh $< \delta < \varepsilon/2$, we have

$$E_n \subset f^{-n} S(f, x, \varepsilon).$$

Since $\mathcal{A}$ has star order less than or equal to K, given $u \in g^{-n} y$ the number of points $v \in g^{-n} y$ with $A_{v_i} \cap A_{u_i} \neq \emptyset$ for all $0 \leq i < n$ is bounded by K^n. Pick $\eta > 0$ so that $A_i \cap A_j = \emptyset$ implies $d(A_i, A_j) > \eta$; then

$$\max \mathrm{sep}(n, \eta, E_n) \geq \frac{\#g^{-n} y}{K^n} \geq \left(\frac{\lambda}{K}\right)^n.$$

Thus, since $E_n \subset f^{-n} S(f, x, \varepsilon)$ and $\log \lambda = GR\{\max \mathrm{sep}(n, \delta, X)\}$, we obtain

$$\begin{aligned} GR\{\max \mathrm{sep}(n, \eta, f^{-n} S(f, x, \varepsilon))\} &\geq GR\{\max \mathrm{sep}(n, \eta, E_n)\} \\ &\geq GR\{\max \mathrm{sep}(n, \delta, X)\} - \log K \end{aligned}$$

and taking $\delta \to 0$ gives the claim. □

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

Now, for $\varepsilon > 0$ and $m \geq 1$, uniform continuity gives $\varepsilon_m > 0$ such that

$$S(f^m, x, \varepsilon_m) \subset S(f, x, \varepsilon) \quad \text{for every } x \in X.$$

Given $m \geq 1$ and $\delta > 0$, for $n \geq 1$ let $E_n \subset (f^m)^{-n} S(f^m, x, \varepsilon_m)$ be a maximal (n, δ)-separated set for f^m. Note also that $E_n \subset f^{-mn} S(f, x, \varepsilon)$ is also (mn, δ)-separated for f. Thus

$$\begin{aligned} h_s(f, x, \varepsilon) &\geq \lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{nm} \log \# E_n \\ &= \frac{1}{m} \limsup_{n \to \infty} \frac{1}{n} \log \# E_n = \frac{1}{m} h_s(f^m, x, \varepsilon_m). \end{aligned}$$

The claim (with f replaced by f^m and ε by ε_m) then gives

$$\begin{aligned} \sup_{x \in X} h_s(f, x, \varepsilon) &\geq \lim_{m \to \infty} \frac{1}{m} \sup_{x \in X} h_s(f^m, x, \varepsilon_m) \\ &\geq \lim_{m \to \infty} \frac{1}{m} (h_{\text{top}}(f^m) - \log K) = h_{\text{top}}(f). \end{aligned}$$

proving the theorem. □

We adapt to forward expansive maps an argument of Mañé [**Man79**], to show that if $f : X \to X$ is a forward expansive map on a compact metric space X, then X has finite covering dimension. The argument yields a condition on orders, but our application will use the corresponding condition on star orders. For Mañé's argument, it is useful to invoke an alternate formulation of forward expansiveness, which follows from the usual one by uniform continuity arguments.

Remark 4.2. A continuous map $f : X \to X$ on a compact metric space is forward expansive with expansiveness constant $c' > 0$ if and only if, for every $0 < c < c'$ and for every $\varepsilon > 0$, there is $N = N(\varepsilon)$ such that for all $x, y \in X$,

$$d_N^f(x, y) < c \Rightarrow d(x, y) < \varepsilon.$$

The finite dimensionality of a compact metric space carrying a forward expansive map follows from the following variation of Mañé's argument.

LEMMA 4.3. *Given $f : X \to X$ satisfying the condition of Remark 4.2, suppose X can be covered by n open sets $U_1, \ldots, U_n$ of $d^f_{N(c/2)}$-diameter less than $c/2$. Then for all $\varepsilon > 0$ there is a cover of X by open sets of d-mesh less than ε and order at most n^2.*

Proof. Fix an integer $M \geq \max\{N(\varepsilon), N(c/2)\}$. For each pair of indices $i, j \in \{1, \ldots, n\}$, define

$$U_{i,j} := U_i \cap f^{-(M - N(c/2))} U_j.$$

We show that each $U_{i,j}$ is a disjoint union of open sets of d-diameter less than ε, giving the required cover.

On $U_{i,j}$ define $x \sim y \iff d_M^f(x, y) < c$. This relation is symmetric and reflexive; for transitivity, note that $d_M^f(x, y) < c$ implies

$$d(f^r x, f^r y) < \frac{c}{2} \quad \text{for } r = 0, \ldots, M - N\left(\frac{c}{2}\right),$$

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

but since $f^{M-N(c/2)}x$ and $f^{M-N(c/2)}y$ both belong to U_j,

$$d(f^r x, f^r y) < \frac{c}{2} \quad \text{for } r = M - N\left(\frac{c}{2}\right) + 1, \ldots, M-1$$

and hence $d_M^f(x, y) < c/2$. The equivalence classes with respect to '$\sim$' are

$$U_{i,j}(x) := \{y \in U_{i,j} \mid d_M^f(x, y) < c\}.$$

which are open sets of d-diameter less than ε. □

This immediately yields the following.

COROLLARY 4.4. *If $f : X \to X$ is a forward expansive map of a compact metric space X, then X has finite covering dimension.*

The definition of forward expansiveness (with expansiveness constant c) is clearly equivalent to the statement that $S(f, x, \varepsilon) = \{x\}$ for every $x \in X$ and $0 < \varepsilon \leq c$; in particular $h_s(f, x, \varepsilon) = GR\{\# f^{-n}x\}$ and preimage growth points are entropy points. By Corollary 2.4, Corollary 4.4 and Theorem 4.1 we have the following.

PROPOSITION 4.5. *Suppose $f : X \to X$ is a forward expansive continuous map with $h_{\text{top}}(f) = \log \lambda$. Then the set $\mathcal{E}(f)$ of entropy points is the same as the (non-empty) set of preimage growth points for f and there is a point $x \in X$ with*

$$\# f^{-n}x \geq \lambda^n \quad \textit{for all } n.$$

In particular,

$$h_p(f) = h_m(f) = h_{\text{top}}(f).$$

5. *Natural extensions and factor maps*

Recall that the *natural extension* of a surjective map $f : X \to X$ is the homeomorphism $\hat{f} : \hat{X} \to \hat{X}$ where $\hat{X} \subset X^{\mathbb{Z}}$ is the space of 'branches' of f^{-1}

$$\hat{X} := \{\hat{x} = (\hat{x}_i)_{i \in \mathbb{Z}} \in X^{\mathbb{Z}} \mid f\hat{x}_i = \hat{x}_{i+1} \text{ for all } i \in \mathbb{Z}\}$$

endowed with the metric

$$\hat{d}(\hat{x}, \hat{y}) := \sum_{i=0}^{\infty} 2^{-i} d(\hat{x}_{-i}, \hat{y}_{-i})$$

and $\hat{f}$ is the left shift map $(\hat{f}\hat{x})_i = \hat{x}_{i+1}$. The projection $\pi : \hat{X} \to X$ given by $\pi\hat{x} := \hat{x}_0$ is a factor map, that is a continuous surjection satisfying $\pi \circ \hat{f} = f \circ \pi$. Note that, for $\hat{x}$, $\hat{y} \in \hat{X}$ and $n \geq 1$, we have

$$\begin{aligned} \hat{d}(\hat{f}^{n-1}\hat{x}, \hat{f}^{n-1}\hat{y}) &= \sum_{i=0}^{n-1} 2^{-i} d(\hat{x}_{n-1-i}, \hat{y}_{n-1-i}) + \sum_{i=n}^{\infty} 2^{-i} d(\hat{x}_{n-1-i}, \hat{y}_{n-1-i}) \\ &\leq 2 \cdot d_n^f(\pi\hat{x}, \pi\hat{y}) + 2^{-(n-1)} \operatorname{diam} X. \end{aligned}$$

Using the uniform continuity of $\hat{f}$ this yields the following.

LEMMA 5.1. *Given $\varepsilon > 0$, there exists $\delta > 0$ and $m \in \mathbb{N}$ such that, for all $n \geq m$, $d_n^f(\pi \hat{x}, \pi \hat{y}) < \delta$ implies $\hat{d}(\hat{f}^n \hat{x}, \hat{f}^n \hat{y}) < \varepsilon$.*

We show that entropy points for f lift to the natural extension.

LEMMA 5.2. *Suppose $f : X \to X$ is a continuous surjection of a compact metric space. Then $\mathcal{E}(f) \subset \pi(\mathcal{E}(\hat{f}))$.*

Proof. Suppose $x \in \mathcal{E}(f)$. Using Proposition 3.5 we can find an element $\hat{x} \in \pi^{-1}x$ such that $\hat{x}_{-n}$ is an entropy point for every $n \in \mathbb{N}$. Given $\varepsilon > 0$, we will find $i \in \mathbb{N}$ and $\delta > 0$ so that

$$h_s(\hat{f}, \hat{x}, \varepsilon) \geq h_s(f, \hat{x}_{-i}, \delta).$$

Using the fact that $\hat{x}_{-i}$ is an entropy point for f and taking as ε and δ go to zero, we then see that $\hat{x}$ is an entropy point for $\hat{f}$.

Let $\delta > 0$ and $m \in \mathbb{N}$ as in Lemma 5.1. Now fix $i \geq m$. Then $\pi^{-1}S(f, \hat{x}_{-i}, \delta) \subset \hat{f}^{-i}S(\hat{f}, \hat{x}, \varepsilon)$ so that setting $K_n := \hat{f}^{i-n}\pi^{-1}S(f, \hat{x}_{-i}, \delta)$ we have

$$K_n \subset \hat{f}^{-n}S(\hat{f}, \hat{x}, \varepsilon) \tag{$*$}$$

and since π is surjective

$$\pi K_n = f^{i-n}S(f, \hat{x}_{-i}, \delta). \tag{$**$}$$

Given $\gamma > 0$ pick $\eta > 0$ so that $d(\hat{y}, \hat{y'}) < \eta$ guarantees $d(\pi\hat{y}, \pi\hat{y'}) < \gamma$. If E is an (n, γ)-separated subset of πK_n, then its preimage in K_n is (n, η)-separated and has cardinality at least $\#E$. It follows that for each n

$$\max \operatorname{sep}(n, \eta, K_n) \geq \max \operatorname{sep}(n, \gamma, \pi K_n)$$

and taking as $\gamma \to 0$ and $\eta \to 0$ yields

$$\lim_{\eta\to 0} GR\{\max \operatorname{sep}(n, \eta, K_n)\} \geq \lim_{\eta\to 0} GR\{\max \operatorname{sep}(n, \eta, \pi K_n)\}. \tag{$***$}$$

Thus

$$\begin{aligned} h_s(\hat{f}, \hat{x}, \varepsilon) &= \lim_{\eta\to 0} GR\{\max \operatorname{sep}(n, \eta, \hat{f}^{-n}S(\hat{f}, \hat{x}, \varepsilon))\} \\ &\geq \lim_{\eta\to 0} GR\{\max \operatorname{sep}(n, \eta, K_n)\} \quad \text{by (*)} \\ &\geq \lim_{\eta\to 0} GR\{\max \operatorname{sep}(n, \eta, \pi K_n)\} \quad \text{by (***)} \\ &= \lim_{\eta\to 0} GR\{\max \operatorname{sep}(n, \eta, f^{i-n}S(f, \hat{x}_{-i}, \delta))\} \quad \text{by (**)} \\ &= \lim_{\eta\to 0} GR\{\max \operatorname{sep}(n, \eta, f^{n}S(f, \hat{x}_{-i}, \delta))\} \quad \text{by Lemma 3.4.} \end{aligned}$$

The last quantity is $h_s(f, \hat{x}_{-i}, \delta)$. □

COROLLARY 5.3. *If $f : X \to X$ is a forward expansive surjection and $x \in X$ is a preimage growth point for f, then there is an entropy point $\hat{x} \in \hat{X}$ for $\hat{f}$ with $\pi\hat{x} = x$.*

Proof. Since f is forward expansive, preimage growth points for f are entropy points. □

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

A one-sided subshift is forward expansive and so preimage growth points are entropy points. Therefore, by Proposition 2.2, every one-sided subshift has entropy points.

COROLLARY 5.4. *Every two-sided subshift has entropy points.*

Proof. A two-sided subshift is the natural extension of its one-sided projection, which is a one-sided subshift and thus has entropy points. Since it is onto, by Lemma 5.2 this gives an entropy point in the two-sided subshift. □

To prove that $\pi(\mathcal{E}(\hat{f})) = \mathcal{E}(f)$, we introduce a property which ensures that a factor map takes entropy points to entropy points and that the natural projection has this property. Example 7.3 shows that, in general, a factor map does not preserve entropy points.

Definition 5.5. Suppose $g : Y \to Y$, $f : X \to X$ are continuous (not necessarily surjective) maps of compact metric spaces Y and X, respectively, and $\pi : Y \to X$ is a factor map. For $n = 1, 2, \ldots, \delta > 0$ and $x \in X$, set

$$V_{n,\delta}(x) := \{y \in Y \mid d_n^f(\pi(y), x) \leq \delta\}.$$

We say that π is uniformly entropy preserving (u.e.p.) if for every $\varepsilon > 0$ and $\lambda > 1$ there exists $\delta > 0$ and $N \in \mathbb{N}$ such that

$$\max \operatorname{sep}(n, \varepsilon, V_{n,\delta}(x)) < \lambda^n \quad \text{for all } x \in X \text{ and all } n \geq N.$$

Standard arguments show that this definition is independent of the choice of the metrics on X and Y. We show that the u.e.p. condition guarantees that π takes entropy points to entropy points. First we establish a useful technical lemma.

LEMMA 5.6. *Suppose $\pi : Y \to X$ is a u.e.p. factor map from $g : Y \to Y$ to $f : X \to X$. For any sequence K_n of subsets of Y,*

$$\lim_{\varepsilon\to 0} GR\{\max \operatorname{sep}(n, \varepsilon, K_n)\} = \lim_{\varepsilon\to 0} GR\{\max \operatorname{sep}(n, \varepsilon, \pi K_n)\}.$$

Proof. For any $\varepsilon > 0$ and $\lambda > 1$, pick $\delta > 0$ and N as in the u.e.p. condition. If E is a maximal (n, δ)-separated subset of πK_n, then E is also (n, δ) spanning and so $K_n \subset \bigcup_{x\in E} V_{n,\delta}(x)$, which implies

$$\max \operatorname{sep}(n, \varepsilon, K_n) \leq \sum_{x \varepsilon E} \max \operatorname{sep}(n, \varepsilon, V_{n,\delta}(x)) < \lambda^n \max \operatorname{sep}(n, \delta, \pi K_n)$$

and taking the growth rate on both sides, then letting $\lambda \to 1$ and $\varepsilon, \delta \to 0$, we obtain the inequality

$$\lim_{\varepsilon\to 0} GR\{\max \operatorname{sep}(n, \varepsilon, K_n)\} \leq \lim_{\delta\to 0} GR\{\max \operatorname{sep}(n, \delta, \pi K_n)\}.$$

The opposite inequality follows as (***) in the proof of Lemma 5.2. □

Remark 5.7. There are two ways that this lemma can be used. First, taking $K_n = K$ independent of n we can conclude that for any set $K \subset Y$, $h_{\text{top}}(g, K) = h_{\text{top}}(f, \pi K)$ and, in particular, $h_{\text{top}}(g) = h_{\text{top}}(f)$. Second, picking $K \subset Y$ closed and assuming $K_n := g^{-n}K$ is non-empty for all n, the lemma gives the inequality

$$\lim_{\varepsilon\to 0} GR\{\max \operatorname{sep}(n, \varepsilon, g^{-n}K)\} \leq \lim_{\varepsilon\to 0} GR\{\max \operatorname{sep}(n, \varepsilon, f^{-n}\pi K)\}.$$

Note that if K happens to be the full preimage of its projection, then the inclusion, and hence the inequality above, becomes equality.

PROPOSITION 5.8. *If $\pi : Y \to X$ is a u.e.p. factor map between $g : Y \to Y$ and $f : X \to X$, then:*

(1) *for any $\varepsilon_1 > 0$ there exists $\varepsilon_2 > 0$ such that π takes any ε_2-entropy point for g to an ε_1-entropy point for f; and*

(2) $\pi(\mathcal{E}(g)) \subset \mathcal{E}(f)$.

Proof. Given $\varepsilon_1 > 0$, by uniform continuity there exists $\varepsilon_2 > 0$ such that for any $y \in Y$

$$\pi S(g, y, \varepsilon_2) \subset S(f, \pi(y), \varepsilon_1)$$

so that if y is an ε_2-entropy point for g, taking $K = S(g, y, \varepsilon_2)$ in Remark 5.7 gives

$$h_{\text{top}}(g) \geq h_{\text{top}}(f) \geq h_s(f, \pi(y), \varepsilon_1) \geq h_s(g, y, \varepsilon_2) = h_{\text{top}}(g)$$

and so $\pi(y)$ is an ε_1-entropy point for f. In particular, it follows that entropy points for g map to entropy points for f. □

LEMMA 5.9. *For any surjective continuous map $f : X \to X$ of a compact metric space, the natural projection $\pi : \hat{X} \to X$ is u.e.p.*

Proof. Given $\varepsilon > 0$, pick $2\delta > 0$ and $m \in \mathbb{N}$ as in Lemma 5.1. Then for all $n \geq m$,

$$\max \text{sep}(n, \varepsilon, V_{n,\delta}(x)) \leq \max \text{sep}(m, \varepsilon, V_{n,\delta}(x)) \leq \max \text{sep}(m, \varepsilon, \hat{X}),$$

which gives the lemma. □

Other examples of u.e.p. maps are bounded-to-one factor maps and maps with conditional entropy zero between invertible systems. We use this later.

COROLLARY 5.10. *Suppose $f : X \to X$ is a continuous surjection of a compact metric space. Then $\pi(\mathcal{E}(\hat{f})) = \mathcal{E}(f)$. In particular, f has entropy points if and only if $\hat{f}$ has entropy points.*

Proof. Since π is u.e.p. by Lemma 5.9, it maps entropy points for $\hat{f}$ to entropy points for f, by Proposition 5.8. The other inclusion follows from Lemma 5.2. □

6. *Asymptotically h-expansive maps*

In passing, we remark that by a result of Kulesza [**Ku**] any expansive homeomorphism has a bounded-to-one subshift extension (see [**BFF02**, Remark B.9]) and bounded-to-one maps are u.e.p. by an elementary proof. This implies $\mathcal{E}(f) \neq \emptyset$ for expansive systems. Here we show that every asymptotically h-expansive map on a compact metric space has entropy points. We first use a result from [**BFF02**] to obtain the invertible case. To state this result, we recall the definition of conditional entropy of a factor map [**BFF02**]. Suppose $f : X \to X$ and $g : Y \to Y$ are homeomorphisms of compact metric spaces and $\pi : Y \to X$ is a factor map. If $\mathcal{U}$ is an open cover of Y and $V \subset Y$, let $N(\mathcal{U}|V)$ denote the minimum cardinality of a collection $\mathcal{U}$ covering V (this is finite by compactness of Y). For two finite open covers $\mathcal{U}$ and $\mathcal{V}$ of Y, set

$$N(\mathcal{U}|\mathcal{V}) := \max\{N(\mathcal{U}|V) \mid V \in \mathcal{V}\}.$$

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

For $n = 1, 2, \ldots$, set $\mathcal{U}^n$ the common refinement of the preimage covers $g^{-1}\mathcal{U}, \ldots, g^{-n}\mathcal{U}$ and

$$h(g, \mathcal{U}, \mathcal{V}) := GR\{N(\mathcal{U}^n|\mathcal{V}^n)\},$$
$$h(g|\mathcal{V}) := \sup\{h(g, \mathcal{U}, \mathcal{V}) \mid \mathcal{U} \text{ a finite open cover of } Y\}.$$

Finally, we define the *conditional entropy* of π as

$$e^*(\pi) := \inf\{h(g|\pi^{-1}\mathcal{A}) \mid \mathcal{A} \text{ a finite open cover of } X\}.$$

We use the following theorem.

THEOREM 6.1. [**BFF02**, Theorem 7.4] *Suppose $f : X \to X$ is an asymptotically h-expansive homeomorphism of a compact metric space. Then there exists a two-sided subshift $g : Y \to Y$ and a factor map $\pi : Y \to X$ such that $e^*(\pi) = 0$.*

To apply this to our situation, we recall the definition by Downarowicz and Serafin [**DS**] of the topological conditional entropy of (Y, g) given the factor (X, f):

$$h(Y|X) = \sup_{\mathcal{U}} \inf_{\mathcal{A}} h(g, \mathcal{U}|\pi^{-1}\mathcal{A}).$$

Obviously $h(Y|X) \le e^*(\pi)$ by definition. Furthermore, we show the following.

LEMMA 6.2. *If $\pi : Y \to X$ is a factor map from (Y, g) to (X, f) then*

$$h(Y|X) = 0 \iff \pi \text{ is u.e.p.}$$

Proof. '$\Rightarrow$' Pick $\varepsilon > 0$ and $\lambda > 1$. Let $\mathcal{U}$ be a finite open cover of Y with mesh$(\mathcal{U}) < \varepsilon$. Since $h(Y|X) = 0$, we can find a finite open cover $\mathcal{A}$ of X with $h(g, \mathcal{U}|\pi^{-1}\mathcal{A}) < \log \lambda$. We can find N so that for all $n \ge N$,

$$N(\mathcal{U}^n|(\pi^{-1}\mathcal{A})^n) < \lambda^n. \qquad (*)$$

Let $\delta > 0$ be a Lebesgue number for $\mathcal{A}$. Given $x \in X$ and $n \ge N$, there is an element $A \in \mathcal{A}^n$ containing the ball $B = \{x' \in X | d_n^f(x', x) < \delta\}$ and pick $E \subset V_{n,\delta}(x)$ a maximal (n, ε)-separated set for g, noting that

$$V_{n,\delta}(x) = \pi^{-1}B \subset \pi^{-1}A.$$

Since mesh$(\mathcal{U}) < \varepsilon$, distinct points of E belong to different elements of $\mathcal{U}^n$, so

$$N(\mathcal{U}^n|\pi^{-1}\mathcal{A}^n) \ge \#E = \max \operatorname{sep}(n, \varepsilon, V_{n,\delta}(x)).$$

Then $(*)$ shows that π is u.e.p.

'$\Leftarrow$' Given $\mathcal{U}$ and $\lambda > 1$, let 3ε be a Lebesgue number for $\mathcal{U}$. Let N and δ be as in Definition 5.5. Let $\mathcal{A}$ be a finite open cover of Y with mesh$(\mathcal{A}) < \delta$. Fix $n \ge N$. Given an element $A \in \mathcal{A}^n$ pick a point $x \in A$. Then $\pi^{-1}A \subset V_{n,\delta}(x)$. Let $E \subset V_{n,\delta}(x)$ be (n, ε)-separated with cardinality $\max \operatorname{sep}(n, \varepsilon, V_{n,\delta}(x))$. Then E is (n, ε) spanning, i.e. the sets

$$F_y := \{w \in Y \mid d_n^g(w, y) \le \varepsilon\}, \quad y \in E$$

cover $V_{n,\delta}(x)$ and thus $\pi^{-1}\mathcal{A}$. Each set F_y is contained in a single element of $\mathcal{U}^n$, so $\pi^{-1}\mathcal{A}$ can be covered by at most $\#E = \max \operatorname{sep}(n, \varepsilon, V_{n,\delta}(x))$ sets. Therefore, $N(\mathcal{U}^n|\pi^{-1}\mathcal{A}) \le \lambda^n$ for all $n \ge 1$ and $h(Y|X) \le \log \lambda$. The result follows by letting $\lambda \to 1$. □

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

We now can easily prove the following proposition.

PROPOSITION 6.3. *Any asymptotically h-expansive homeomorphism of a compact metric space has entropy points.*

Proof. Let $g : Y \to Y$ be the subshift given by Theorem 6.1. By Corollary 5.4, $\mathcal{E}(g) \neq \emptyset$, and by Lemma 6.2 the projection π is uniformly entropy preserving, hence by Proposition 5.8, $\emptyset \neq \pi(\mathcal{E}(g)) \subset \mathcal{E}(f)$. □

THEOREM 6.4. *For $f : X \to X$ any asymptotically h-expansive map on a compact metric space, $\mathcal{E}(f) \neq \emptyset$.*

Proof. By restricting to the eventual image we may assume that f is onto. Let $\hat{f} : \hat{X} \to \hat{X}$ be the natural extension of f. Given $\varepsilon > 0$, pick $\delta > 0$ such that for any $\hat{x} \in \hat{X}$,

$$\pi S(\hat{f}, \hat{x}, \delta) \subset S(f, \pi(\hat{x}), \varepsilon),$$

so that for all $\hat{x} \in \hat{X}$ and all $\delta' \leq \delta$,

$$\begin{aligned} h_{\text{top}}(\hat{f}, S(\hat{f}, \hat{x}, \delta')) &\leq h_{\text{top}}(\hat{f}, S(\hat{f}, \hat{x}, \delta)) \\ &\leq h_{\text{top}}(f, \pi S(\hat{f}, \hat{x}, \delta)) \leq h_{\text{top}}(f, S(f, \pi(\hat{x}), \varepsilon)). \end{aligned}$$

Taking the supremum over $\hat{x} \in \hat{X}$ and the limit as $\varepsilon \to 0$, it follows that $\hat{f}$ is asymptotically h-expansive and so $\mathcal{E}(\hat{f}) \neq \emptyset$ by Proposition 6.3. However, then Proposition 5.8 gives $\emptyset \neq \pi(\mathcal{E}(\hat{f})) \subset \mathcal{E}(f)$. □

7. *Examples*

In this section we present examples in which the various kinds of 'entropy points' defined earlier fail to exist and we give an example where $h_p(f) \neq h_m(f)$.

Example 7.1. (A surjective map $f : X \to X$ on a zero-dimensional space X with $h_p(f) = 0 < h_m(f)$.) We define $f : X \to X$ as a zero-dimensional factor of a two-sided subshift $g : Y \to Y$. Let $A := \{0, 1, 2, 3, 4\}$ and endow $A^{\mathbb{Z}}$ and $A^{\mathbb{N}\times\mathbb{N}}$ with the product topology of the discrete topology on A. Denote by $g : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ the left shift map and by $f : A^{\mathbb{N}\times\mathbb{N}} \to A^{\mathbb{N}\times\mathbb{N}}$ the left shift map on rows; that is, $(fx)_{m,i} = x_{m,i+1}$ for $m, i \in \mathbb{N}$. Now define a map $\pi : A^{\mathbb{Z}} \to A^{\mathbb{N}\times\mathbb{N}}$ by $\pi y = x$, where $y = (y_i)_{i\in\mathbb{Z}}$ and $x = (x_{m,i})_{m,i\geq 0}$, and

$$x_{m,i} = y_i \quad \text{if } y_i \in \{3, 4\} \quad \text{while} \quad x_{m,i} = y_{i-m} \quad \text{if } y_i \notin \{3, 4\}.$$

The map π is continuous and $\pi \circ g = f \circ \pi$, since $(\pi y)_{m,i}$ is determined by y_{i-m} and y_i. Now for $n \geq 1$ and $w \in \{1, 2\}^n$ let $y^{n,w} \in A^{\mathbb{Z}}$ be the point $w^\infty 0^n (34^n)^\infty$ with $y_{-1}^{n,w} = 0$ and $y_0^{n,w} = 3$. Let Y be the closure of the set

$$\{g^k y^{n,w} \mid k \in \mathbb{Z}, n \geq 1, w \in \{1, 2\}^n\}$$

and $X = \pi(Y)$. From now on by π we mean the map $\pi|_Y : Y \to X$.

We now show that $f : X \to X$ has $h_m(f) \geq (1/2) \log 2$. Let $\delta > 0$ be so small that $x, x' \in X$ with $x_{0,0} \neq x'_{0,0}$ implies $d(x, x') \geq \delta$. For $n \geq 1$ let $y = \lim_{k\to\infty} g^{k(n+1)} y^{n,w}$,

https://doi.org/10.1017/S0143385703000221 Published online by Cambridge University Press

the point $(34^n)^\infty \in Y$ with $y_0 = 3$. For $w \in \{1,2\}^n$ let $x^w := \pi g^{-2n} y^{n,w}$. Then $f^{2n}x^w = \pi y^{n,w} = \pi y$. Note further that for $w \neq w'$ there is some $0 \le i < n$ so that $(x^w)_{0,i} \neq (x^{w'})_{0,i}$ and thus $x = \pi y$ satisfies $\max \operatorname{sep}(2n, \varepsilon, f^{-2n}x) \ge 2^n$ for each $\varepsilon < \delta$. This shows $h_m(f) \ge (1/2)\log 2$.

Next we show that $h_p(f) = 0$. Since $\max \operatorname{sep}(n, \varepsilon, f^{-n}x) \le \# f^{-n}x$ it suffices to show that $c(x) := GR\{\# f^{-n}x\} = 0$ for all $x \in X$. We have $v \in f^{-n}x$ and $u \in \pi^{-1}v \Rightarrow x = f^n v = f^n \pi u = \pi g^n u$ and so $g^n u \in \pi^{-1}x$. Thus we have the following.

(1) $\# f^{-n}x = \#\{\pi u \mid u \in g^{-n}\pi^{-1}x\} \le \#\pi^{-1}x$.

If $x_{0,i} \notin \{3,4\}$ for some $i \ge 0$ then for any $y \in \pi^{-1}x$ we have $y_k = x_{0,k}$ for all $k \ge i$ and $y_k = x_{i-k,i}$ for $k \le i$. Thus $\#\pi^{-1}x = 1$, and from (1) we have the following.

(2) $x_{0,i} \notin \{3,4\}$ for some $i \ge 0$ implies $\# f^{-n}x = 1$ for all n; thus $c(x) = 0$.

Now consider x with $x_{0,i} \in \{3,4\}$ for all $i \ge 0$. Let $y \in Y$ with $\pi y = x$. We distinguish three cases according to the number of threes. If $\#\{i \ge 0 \mid x_{0,i} = 3\} \le 1$ then $\#\{i \in \mathbb{Z} \mid y_i = 3\} \le 1$, since otherwise $y[k,\infty) = (34^n)^\infty$ for some $k \in \mathbb{N}$ and $n \ge 1$. Thus y is in the orbit of one of the points 4^∞, $0^\infty 34^\infty$ or $4^\infty 34^\infty$. If $y_k = 3$ for some $k < -n$ then $\pi y = x = 4^\infty$ and thus by (1) $\# f^{-n}x \le \#\{u \in g^{-n}\pi^{-1}x \mid u_k = 3 \Rightarrow k \ge -n\}$ and, therefore, we have the following.

(3) $\#\{i \ge 0 \mid x_{0,i} = 3\} \le 1 \Rightarrow \# f^{-n}x \le 2n+1$ for all n and thus $c(x) = 0$.

Finally, if $\#\{i \ge 0 \mid x_{0,i} = 3\} \ge 2$ then pick $i \ge 0$ and $n \ge 1$ such that $x_{0,i} = x_{0,i+n+1} = 3$ and $x_{0,j} = 4$ for $i+1 \le j \le i+n$. Then $y[i, i+n+1] = 34^n3$ and thus either y is in the orbit of the point $(34^n)^\infty$ or $y = g^k y^{n,w}$ for some $k \le 0$ and some $w \in \{1,2\}^n$. Thus $x_{0,0} = 4$ implies $\# f^{-1}x = 1$. If $x_{0,0} = 3$ and $v \in f^{-1}x$ with $v_{0,0} \neq 4$ then each $u \in \pi^{-1}v$ equals $g^{-1}y^{n,w}$ for some $w \in \{1,2\}^n$, and if $v_{0,0} = 4$ then $(v_{0,i})_{i \ge 0} = 4(34^n)^\infty$. Thus $\# f^{-1}x = 2^n + 1$ and by (2) we have $\# f^{-m}x \le m \cdot (2^n+1)$, hence $c(x) = 0$.

Example 7.2. (A homeomorphism which has no ε-entropy points for ε sufficiently small.) For $n = 1, 2, \ldots$, let $S_n \subset \{0,1\}^{\mathbb{Z}}$ be the subshift with forbidden words $\mathcal{V}_n := \{1^{n+1}\}$. Then it is straightforward to establish the inequality

$$\frac{n}{n+1}\log 2 \le h_{\mathrm{top}}(S_n) \le \frac{1}{n+1}\log(2^{n+1}-1)$$

so that $h_{\mathrm{top}}(S_n) < \log 2$ and $\lim_{n\to\infty} h_{\mathrm{top}}(S_n) = \log 2$. Let $U_n \subset \{0,1\}^{\mathbb{Z}}$ be the finite orbit of the point $u^n = (10^n)^\infty$ which has $u_0^n = 1$. Finally let $U_0 = \{u \in \{0,1\}^{\mathbb{Z}} \mid \#\{i \in \mathbb{Z} \mid u_i = 1\} \le 1\}$ and put $S_0 := U_0$. Let $U := \bigcup_{n=0}^\infty U_n$. Then U is closed. We denote the nth row of a point $x = (x_{n,i})_{n,i \in \mathbb{N}\times\mathbb{Z}} \in \{0,1\}^{\mathbb{N}\times\mathbb{Z}}$ by x^n, that is

$$x^n = (x_{n,i})_{i\in\mathbb{Z}} \in \{0,1\}^{\mathbb{Z}}.$$

Then for $n \in \mathbb{N}$ let

$$X_n := \{x \in \{0,1\}^{\mathbb{N}\times\mathbb{Z}} \mid x^0 \in U_n, x^n \in S_n \quad \text{and} \quad x^i = 0^\infty \quad \text{for all } i \in \mathbb{N} - \{0,n\}\}.$$

The set $X := \bigcup_{n=0}^\infty X_n$ is closed. Let $f : X \to X$ be the left shift; that is, $(fx)_{n,i} = x_{n,i+1}$ for $x = (x_{n,i})_{n,i\in\mathbb{N}\times\mathbb{Z}} \in X$. Note that $f|_{X_0} : X_0 \to X_0$ is conjugate to the subshift U_0 and for $n \ge 1$ the map $f|_{X_n} : X_n \to X_n$ is conjugate to the subshift $U_n \times S_n$. Since $\max \operatorname{sep}(n, \delta, X) \ge \# W_n(U_m \times S_m)$ for all $\delta > 0$ small enough, $h_{\mathrm{top}}(f) = \lim_m h_{\mathrm{top}}(U_m \times S_m) = \log 2$.

Let d be a metric on X generating the product topology of the discrete topology on $\{0, 1\}$. Fix $\varepsilon_0 > 0$ so that for any $x, y \in X$

$$x_{0,0} \neq y_{0,0} \Rightarrow d(x, y) > \varepsilon_0.$$

Now let $0 < \varepsilon < \varepsilon_0$ and $x \in X$. Then $x \in X_n$ for some $n \geq 0$. If $y \in X$ with $d(f^i x, f^i y) \leq \varepsilon$ for all $i \geq 0$, then $x_{0,i} = y_{0,i}$ for all $i \geq 0$ and thus $y \in X_n$ as well. Thus $S(f, x, \varepsilon) \subset X_n$ and, therefore, since $f(X_n) = X_n$, $h_s(f, x, \varepsilon) = h_s(f|_{X_n}, x^n, \varepsilon) \leq h_{\text{top}}(f|_{X_n})$ and thus $h_s(f, x, \varepsilon) \leq h_{\text{top}}(S_n) < \log 2 = h_{\text{top}}(f)$ which shows that x is not an ε-entropy point.

Example 7.3. (A factor map $\pi : Y \to X$ from $g : Y \to Y$ to $f : X \to X$ with $\mathcal{E}(g) \neq \emptyset$, but $\mathcal{E}(f) = \emptyset$.) We adopt the notation of Example 7.2. First we define for each $n = 1, 2, \ldots$, a factor map from the full 2-shift to the subshift S_n via the $n + 1$ block map that assigns to each word of length $n + 1$ its first letter, provided some letter in the word is 0, and assigns to the word 1^{n+1} the letter 0. This map is a retraction—it equals the identity on S_n.

Now form a set Y of arrays like the set X in Example 7.2, but using the full 2-shift in place of the subshift S_n for $n \geq 1$. The factor maps defined above induce a factor map from the shift $g : Y \to Y$ to the shift $f : X \to X$ and $h_{\text{top}}(g) = h_{\text{top}}(f)$. For $n \geq 1$, every point of Y_n is an entropy point for g, so we have the required example.

Example 7.4. (A homeomorphism which has ε-entropy points for each $\varepsilon > 0$, but no entropy points.) Again X will be a closed subset of $\{0, 1\}^{\mathbb{N}\times\mathbb{Z}}$ and f will be the left shift map on rows. We use the notation, the subshifts S_n and $U(n) := U_n$ and the set X_0 from Example 7.2. Let $P := \{n \in \mathbb{N} \mid n \geq 2 \text{ prime}\}$. For $p, k \in P$ let

$$X_{p,k} := \{x \in \{0, 1\}^{\mathbb{N}\times\mathbb{Z}} \mid x^0 \in U_p, x^p \in U(p^k), x^{p^k} \in S_k \\ \text{and} \quad x^i = 0^\infty \quad \text{for all } i \in \mathbb{N} - \{0, p, p^k\}\}.$$

Then the set

$$X := \bigcup_{p,k \in P} X_{p,k} \cup X_0$$

is closed and the left shift map on rows $f : X \to X$ is a homeomorphism with $h_{\text{top}}(f) = \lim_k h_{\text{top}}(S_k) = \log 2$ with an argument as in Example 7.2. Let $\varepsilon > 0$. Fix $p \in P$ so large that for any $x, y \in X$

$$x_{n,i} = y_{n,i} \quad \text{for } 0 \leq n < p \quad \text{and} \quad -p \leq i \leq p \text{ implies } d(x, y) < \varepsilon.$$

Let $x \in X_{p,p}$ with $x_{0,0} = 1$. We show that x is an ε-entropy point for f. Fix $k \in P$. Consider $y \in X_{p,k}$ with $y_{p,0} = 1$. Then $y^0 = x^0$ and $x^i = y^i = 0^\infty$ for $1 \leq i < p$. Thus $d(f^i x, f^i y) \leq \varepsilon$ for all i and, in particular $y \in S(f, x, \varepsilon)$. It follows that $\{z \in X_{p,k} \mid z_{0,n} = 1\} \subset f^{-n} S(f, x, \varepsilon)$ is a set of entropy $h_{\text{top}}(S_k)$. Therefore, we have $h_s(f, x, \varepsilon) \geq \lim_k h_{\text{top}}(S_k) = \log 2 = h_{\text{top}}(f)$ and x is an ε-entropy point.

To show that f has no entropy points fix $x \in X$. If $x \in X_0$, then pick $\varepsilon_0 > 0$ so that for all $u, v \in X$

$$u_{0,0} \neq v_{0,0} \Rightarrow d(u, v) > \varepsilon_0.$$

Then for $0 < \varepsilon < \varepsilon_0$, $y \in S(f, x, \varepsilon)$ implies $y \in X_0$ and thus $h_s(f, x, \varepsilon) = 0 < h_{top}(f)$. If $x \in X_{p,k}$ for some $p, k \in P$, then choose $\varepsilon_0 > 0$ so that for all $u, v \in X$

$$u_{0,0} \neq v_{0,0} \quad \text{or} \quad u_{p,0} \neq v_{p,0} \Rightarrow d(u, v) > \varepsilon_0.$$

Then for $0 < \varepsilon < \varepsilon_0$, $y \in S(f, x, \varepsilon)$ implies $y^p = x^p$ and so $y \in X_{p,k}$ since p is prime and $x^p \neq 0^\infty$. Thus $h_s(f, x, \varepsilon) \leq h_{top}(S_k) < h_{top}(f)$.

Acknowledgements. We would like to thank Manfred Denker for useful conversations concerning some of this material and the third author thanks the Institut für Mathematische Stochastik, Georg-August Universität Göttingen for its hospitality and, along with the Deutsche Akademische Austauschdienst, for its support during two visits to Göttingen during which much of this project was carried out.

REFERENCES

[AKM65] R. L. Adler, A. G. Konheim and M. H. McAndrew. Topological entropy. *Trans. Amer. Math. Soc.* **114** (1965), 309–319.

[BFF02] M. Boyle, D. Fiebig and U. Fiebig. Residual entropy, conditional entropy and subshift covers. *Forum Math.* **14** (2002), 713–757.

[Bow71] R. Bowen. Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.* **153** (1971), 401–414 (Erratum **181** (1973), 509–510).

[Bow72] R. Bowen. Entropy-expansive maps. *Trans. Amer. Math. Soc.* **164** (1972), 323–331.

[Bow78] R. Bowen. *On Axiom A Diffeomorphisms (Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, 35)*. American Mathematical Society, Providence, RI, 1978.

[Buz97] J. Buzzi. Intrinsic ergodicity of smooth interval maps. *Israel J. Math.* **100** (1997), 125–161.

[DGS76] M. Denker, C. Grillenberger and K. Sigmund. *Ergodic Theory on Compact Spaces (Lecture Notes in Mathematics, 527)*. Springer, New York, 1976.

[Din70] E. I. Dinaburg. The relation between topological entropy and metric entropy. *Soviet Math. Dokl.* **11** (1970), 13–16.

[DS] T. Downarowicz and J. Serafin. Fiber entropy and conditional variational principles in compact, non metrizable spaces. *Fund. Math.* **172** (2002), 217–247.

[Eng78] R. Engelking. *Dimension Theory*. North-Holland, Amsterdam, 1978.

[Hur95] M. Hurley. On topological entropy of maps. *Ergod. Th. & Dynam. Sys.* **15** (1995), 557–568.

[Ku] J. Kulesza. Zero-dimensional covers of finite-dimensional dynamical systems. *Ergod. Th. & Dynam. Sys.* **15** (1995), 939–950.

[LP92] R. Langevin and F. Przytycki. Entropie de l'image inverse d'une application. *Bull. Soc. Math. France* **120** (1992), 237–250.

[LW91] R. Langevin and P. Walczak. Entropie d'une dynamique. *Comptes Rendus, Acad. Sci. Paris* **312** (1991), 141–144.

[Man79] R. Mañé. Expansive homeomorphisms and topological dimension. *Trans Amer. Math. Soc.* **252** (1979), 313–319.

[Mis76] M. Misiurewicz. Topological conditional entropy. *Studia Math.* **55** (1976), 175–200.

[NP99] Z. Nitecki and F. Przytycki. Preimage entropy for mappings. *Int. J. Bifurcation Chaos* **9** (1999), 1815–1843.