RESIDUAL ENTROPY, CONDITIONAL ENTROPY AND SUBSHIFT COVERS

MIKE BOYLE, DORIS FIEBIG, AND ULF FIEBIG

ABSTRACT. We study the existence of subshift covers for topological dynamical systems, the infimum of the entropy jumps to such covers, and various aspects of conditional entropy and covering maps including a variational principle for covering maps. In particular we show every asymptotically h-expansive system (and therefore by Buzzi every C^∞ homeomorphism of a compact Riemannian manifold) has a subshift cover of equal entropy. Our arguments in dimension zero are extended to higher dimension with theorems of Kulesza and Thomsen.

CONTENTS

1. INTRODUCTION

Good subshift covers have been a useful tool for studying hyperbolic smooth dynamical systems (e.g. [Bow2]), and there is a nice theory of the abstract symbolic dynamics in this setting [F]. One would like to have some general understanding of which topological dynamical systems admit good symbolic dynamics.

The very first question, when is a system T the quotient of any subshift at all, turns out to be very difficult. An affirmative answer has long been known

Date: January 23, 2001.

2000 *Mathematics Subject Classification.* Primary: 37B10; Secondary: 37B40, 37C40, 37C45, 37C99, 37D35.

Key words and phrases. residual entropy, conditional entropy, entropy, variational principle, subshift, symbolic dynamics, cover, defect.

The research of the first author was supported by NSF Grant DMS9706852.

for expansive systems [Re] and for some or all group translations (e.g., Sturmian subshifts cover circle rotations), but we are aware of no general results on this question before the current paper and the work of Downarowicz [Do2]. (On the other hand, in Appendices A and B we see that the work of Kulesza and Thomsen gives excellent information about the existence of good quotient maps from zero dimensional systems onto given higher dimensional systems.) A necessary condition for T to admit a subshift cover is that T must have finite entropy, but this turns out to be not sufficient (Example 3.1 and [Do2]). We define the *residual entropy* $\rho(T)$ of T as the infimum of the entropy gaps $h(S) - h(T)$ over the set of subshifts S covering T. (If this set is empty, we set $\rho(T) = \infty$.)

The residual entropy can be viewed as a descendant of the conditional topological entropy of a system, introduced by Misiurewicz [Mi2]. In Section 4 we review some essentials of the Misiurewicz development. In Section 5 we define the conditional topological entropy of a quotient map and work out some natural results. In Section 6 we prove a variational principle for the conditional entropy of a quotient map, describe its generalization by Downarowicz and Serafin, and give a counterexample to a natural simplifying conjecture. In Section 7, we characterize (Theorem 7.1) the existence of a quotient map from a mixing SFT S to a finite entropy product T of mixing shifts of finite type (SFTs). With this construction and the results on good zero dimensional extensions from the appendices, we go on to prove that any asymptotically h-expansive system T is a quotient of a subshift by a quotient map of conditional entropy zero (and in particular $\rho(T) = 0$). Buzzi [Bu], developing work of Yomdin [Y], has shown that any C^∞ diffeomorphism of a compact Riemannian manifold is asymptotically h-expansive, and it follows that every such system has residual entropy zero.

The characterization of residual entropy turns out to be remarkably complicated. Here the best result, by far, is the Downarowicz characterization in the zero dimensional case [Do2], which we state in Section 8. The Downarowicz characterization is a mixed topological-measurable condition. In Section 8 we also characterize residual entropy in the zero dimensional case in terms of certain functions of words, without reference to measures. By analogy with the usual topological entropy, or even the conditional topological entropy of Misiurewicz, one expects that there should be reasonable definitions of residual entropy in terms of open sets or n, ϵ orbits; but we have been unable to achieve any such definition.

Among the open questions raised we single out two. First, if T has finite residual entropy, must there exist a subshift cover S such that $\rho(T) = h(S) - h(T)$? Second, to what extent is nonzero residual entropy compatible with smoothness? We know that a C^∞ system has residual entropy zero, and in Appendix C we exhibit a finite entropy homeomorphism of a surface with infinite residual entropy. But for $1 \leq k < \infty$, we have no example of a C^k map with nonzero residual entropy, and we know of no obstruction to any value of residual entropy.

Some results of this paper (Example 3.1, Theorem 7.1, the infimum claim of Theorem 8.2, Proposition D.5, parts of Theorem 7.4 and A.1) were announced long ago [B2]. Downarowicz [Do1] came to the problem of residual entropy later but quite independently. In addition to giving the zero dimensional case characterization of residual entropy mentioned above, he gave examples of all allowable combinations of $h(T), h^*(T), \rho(T)$ [Do2]. The paper [Do2] also includes the characterization (done jointly with Boyle) of asymptotically h-expansive zero dimensional systems as subsystems of products of subshifts, and Downarowicz pointed out to us the utility of this characterization in simplifying our own proof that $h^*(T) = 0$ implies $\rho(T) = 0$. The results common to this paper and [Do2] are proved with quite different methods.

We thank Downarowicz for his kind tolerance of our unpublished claims, for the proof simplification mentioned above, and for the stimulus to (finally) finish this work. The first named author also gratefully acknowledges support of the University of Washington in Seattle and the University of Heidelberg at different stages of this work.

2. Background and notation

Throughout the paper, by a *system* we will mean a selfhomeomorphism of a compact metrizable space, e.g. $T\colon X \to X$. A subsystem of T is the restriction of T to a closed invariant subset of X. By the dimension of T we will mean the covering dimension of the domain X. For systems (X,T) and (Y,S), by a homomorphism $\varphi\colon S \to T$ we will mean a continuous map $\varphi\colon X \to Y$ such that $\varphi T = S\varphi$. An embedding $\varphi\colon S \hookrightarrow T$ is an injective homomorphism; a quotient map $\varphi\colon S \twoheadrightarrow T$ is a surjective homomorphism; an isomorphism or topological conjugacy is a bijective homomorphism.

The fixed point set of T will be denoted $\mathrm{Fix}(T)$ and the set of points of least period k will be denoted $\mathcal{P}_k^o(T)$. We let $S \xrightarrow{\text{per}} T$ mean that for all positive integers n,

$$|\mathrm{Fix}(S^n)| > 0 \implies |\mathrm{Fix}(T^n)| > 0 \ .$$

(The condition $S \xrightarrow{\text{per}} T$ is a necessary condition for $S \to T$.) Similarly we let $S \xrightarrow{\text{iper}} T$ mean that for all positive integers k,

$$|\mathcal{P}_k^o(S)| \le |\mathcal{P}_k^o(T)| \ .$$

(The condition $S \xrightarrow{\text{iper}} T$ is a necessary condition for $S \hookrightarrow T$.)

Given a positive integer n, let $\mathcal{A}$ be a set of n elements (usually $\{0,\dots,n-1\}$) with the discrete topology and let $X = \mathcal{A}^{\mathbb{Z}}$ have the product topology. We view a point in X as a doubly infinite sequence $x = \dots x_{-1}x_0x_1\dots$ with each $x_i \in \mathcal{A}$. If T is the shift map on X, defined by requiring $(Tx)_i = x_{i+1}$, then the system T is the *full shift* on n symbols. A *subshift* is a subsystem of some full shift. Any subshift may be described as the set of all points in some full shift in which a countable set of words does not occur. The subshift is a shift of finite type (SFT) if this set of excluded words can be chosen to be finite. For a thorough introduction to subshifts, see [LM, Ki, DGS].

A system T is *expansive* if there exists a metric d compatible with the topology such that there exists $\epsilon > 0$ such that for each pair of points x, y with $x \ne y$ there exists some n in $\mathbb{Z}$ such that $d(T^n x, T^n y) \ge \epsilon$. If this condition holds for one compatible metric, then it holds for every compatible metric. Any zero dimensional expansive system is isomorphic to a subshift.

Now suppose that we have a sequence of systems $T_n\colon X_n \to X_n$ with *bonding maps* $\pi_n\colon T_{n+1} \to T_n$. Let $X = \{x = (x_1, x_2, \dots) \in \prod_n X_n \colon \pi_n(x_{n+1}) = x_n,\ \forall n \in \mathbb{N}\}$. The inverse limit system T is the restriction to X of the infinite product $T_1 \times T_2 \times \dots$. It is elementary to check that every zero dimensional system is isomorphic to an inverse limit of subshifts. (This is an old fact, though for explicit proofs we only know the references [BH, T1], which also give additional structure).

Remark 2.1. The existence of an analogous inverse limit presentation is problematic in higher dimensions. First, a system need not be the inverse limit of expansive systems. For example, one easily checks that an expansive quotient of an isometry is finite, so no nontrivial (more than one point) isometry of a connected compact metric space is an inverse limit of expansive systems. Second, from the work of Elon Lindenstrauss [Li] we know that some infinite entropy systems are not inverse limits of finite entropy systems, and therefore are not inverse limits of subshifts or even quotients of subshifts.

In the sequel, to simplify notation, we will usually use the same symbol (e.g., T) to denote a selfhomeomorphism and its domain.

3. Infinite residual entropy: an example

The purpose of this section is to produce the following

Example 3.1. *There is a selfhomeomorphism T of a compact metric space X such that $h(T) < \infty$ but $\rho(T) = \infty$.*

(X, T) will be the inverse limit system formed from a sequence of mixing SFTs T_n and bonding maps $\pi_n \colon T_{n+1} \twoheadrightarrow T_n$. We will let p_n denote the projection $X \to X_n$; so, $\pi_n p_{n+1} = p_n$, since for $x = (x_1, x_2, \dots)$ we have $\pi_n p_{n+1} x = \pi_n x_{n+1} = x_n = p_n x$. We define the composition bonding maps $\pi_{k,n} \colon T_n \twoheadrightarrow T_k$ by $\pi_{k,n} = \pi_k \pi_{k-1} \cdots \pi_{n-1}$.

Choose the mixing SFTs and bonding maps to have the following properties

1. $h(T_n) < h(T_{n+1}), \qquad n \geq 1.$
2. $\sup_n h(T_n) < \infty$.
3. There exists $\alpha > 1$ such that for all k and for every finite orbit $\mathcal{O}$ of T_k, there exists n (depending on $\mathcal{O}$) such that $h(\pi_{k,n}^{-1}\mathcal{O}) > \log\alpha$.

It remains to prove two claims.

CLAIM 1. A system with properties 1-3 exists.

CLAIM 2. (X, T) has finite entropy and infinite residual entropy.

Proof of CLAIM 2. Clearly $h(T) = \sup_n h(T_n) < \infty$. To show $\rho(T) = \infty$, we argue by contradiction. Suppose S is a subshift and $\varphi \colon S \twoheadrightarrow T$. The map φ yields a commuting diagram of quotient maps in the following way. Define $\varphi_n \colon S \twoheadrightarrow T_n$ as $\varphi_n = p_n \varphi$, then $\varphi_n = \pi_n \varphi_{n+1}$ (since $\varphi_n = p_n \varphi = \pi_n p_{n+1} \varphi = \pi_n \varphi_{n+1}$) and more generally, $\varphi_k = \pi_{k,n} \varphi_n$ whenever $k < n$. Now we have the key observation that for any $n > k > 0$ and finite T_k-orbit $\mathcal{O}$,

$$h(\varphi_k^{-1}\mathcal{O}) = h(\varphi_n^{-1}\pi_{k,n}^{-1}\mathcal{O}) \geq h(\pi_{k,n}^{-1}\mathcal{O}) \quad . \tag{3.2}$$

Define

$$\beta^* = \sup\{\beta \geq 1 : \ h(\varphi_k^{-1}\mathcal{O}) \geq \log\beta \ , \ \forall k, \ \forall \text{ finite } T_k - \text{orbits } \mathcal{O}\} \ .$$

It follows immediately from (3.2) and property 3 that $\beta^* \geq \alpha > 1$. Also $\beta^* < \infty$ since $\log\beta^* \leq h(S) < \infty$.

Now fix any finite T_k-orbit $\mathcal{O}$ and any β such that $0 < \beta < \beta^*$. We will show that $h(\varphi_k^{-1}\mathcal{O}) \geq \log(\alpha\beta)$. (This implies $\beta^* \geq \alpha\beta^*$, which gives the desired contradiction.) Using property 3, choose n such that $h(\pi_{k,n}^{-1}\mathcal{O}) > \log\alpha$. The subshift $E = \pi_{k,n}^{-1}\mathcal{O}$ is SFT (because $\mathcal{O}$ and T_n are SFT), so $h(E)$ is given by the growth rate of the periodic points of E, so for arbitrarily large N there are more than α^N orbits $\mathcal{O}'$ of cardinality N. Fix such an N. For each such orbit $\mathcal{O}'$ in E, $h(\varphi_n^{-1}\mathcal{O}') \geq \log\beta$.

The map φ_n is a block code determined by some rule $w_{i-M} \dots w_{i+M} \mapsto (\varphi_n w)_i$, where M depends on n but not w or i. For each orbit $\mathcal{O}'$ in T_n of cardinality N, there is a set of T_n-words of length N, $\mathcal{W}(\mathcal{O}', T_n) = \{x_0 \dots x_{N-1} : x \in \mathcal{O}'\}$. Let $\mathcal{W}(\mathcal{O}', S) = \{w_{-M} \dots w_{N+M-1} : w \in \varphi_n^{-1}\mathcal{O}'\}$. Because

$$h(\varphi_n^{-1}\mathcal{O}') \leq \frac{\log \# \mathcal{W}(\mathcal{O}', S)}{N + 2M} \quad ,$$

we have $\#\mathcal{W}(\mathcal{O}', S) \geq \beta^{N+2M}$. By the choice of M, for distinct $\mathcal{O}'$ the sets $\mathcal{W}(\mathcal{O}', S)$ are disjoint. Therefore the number of S-words of length $N + 2M$ in $\varphi_n^{-1}E$ is at least

$$\sum_{\mathcal{O}'} \beta^{N+2M} \geq \alpha^N \beta^{N+2M} \quad ,$$

where the sum is over the E-orbits $\mathcal{O}'$ of cardinality N. Because N was arbitrarily large, we conclude

$$h(\varphi_k^{-1}\mathcal{O}) = h(\varphi_n^{-1}E) \geq \lim_N \frac{\log(\alpha^N \beta^{N+2M})}{N + 2M} = \log(\alpha\beta) \ .$$

□

Proof of CLAIM 1. There are many ways to find a sequence (T_n) satisfying (1) and (2), and in addition the condition that every T_n has a fixed point. For example, let T_n be a product $S_1 \times S_2 \times \cdots \times S_n$, where S_n is the mixing SFT whose defining matrix is the companion matrix of the polynomial $x^{k+1} - x^k - 1$, with $k = n^2$.

Now suppose we have such a sequence T_n. Fix α such that $0 < \log\alpha < \lim_n h(T_n)$. Without loss of generality, suppose $\log\alpha < h(T_n)$ for all n. The construction of the π_n is recursive. So suppose $\pi_1, \dots, \pi_{n-1}$ are chosen; then we choose the map π_n as follows. For each $k \leq n$ and each orbit $\mathcal{O}$ of cardinality n or less in T_k, pick a finite orbit $\overline{\mathcal{O}}$ in T_n such that $\pi_{k,n}$ sends $\overline{\mathcal{O}}$ to $\mathcal{O}$. (For $n = k$, set $\overline{\mathcal{O}} = \mathcal{O}$.) Enumerate these orbits $\overline{\mathcal{O}}$ as $\overline{\mathcal{O}}_1, \dots, \overline{\mathcal{O}}_m$.

Let W be an SFT which is the disjoint union of irreducible SFTs $W_1, \dots, W_m$ satisfying the following conditions (in which $\mathcal{P}_k^o(T)$ denotes the set of points in orbits of points of least period k in a subshift T):

- $\log\alpha < h(W_i) < h(T_{n+1})$, $\quad 1 \leq i \leq m$,
- the period of $\overline{\mathcal{O}_i}$ divides the period of W_i, $\quad 1 \leq i \leq m$,
- $\sum_{i=1}^m \#\mathcal{P}_k^o(W_i) \leq \#\mathcal{P}_k^o(T_n)$, $\quad k \in \mathbb{N}$.

There are many ways to produce W. For example, irreducible SFTs W_i satisfying the entropy condition can be chosen, and then each W_i can be replaced by some $W_i \times P_i$, where P_i is some finite orbit of sufficiently large cardinality which is divisible by $\#\overline{\mathcal{O}_i}$.

Now by Krieger's Embedding Theorem [Kr2], we may identify W with a subsystem of T_n. Choose $\psi_i \colon W_i \twoheadrightarrow \overline{\mathcal{O}_i}$, $1 \leq i \leq m$, and let $\psi \colon W \twoheadrightarrow \cup\overline{\mathcal{O}_i}$ be the union of these maps. Because T_n is a mixing SFT with a fixed point, by the Extension Lemma (2.4 of [B1]) we may extend ψ to a quotient map $\pi_n \colon T_{n+1} \twoheadrightarrow T_n$.

The sequence (π_n) has the property 3. □

4. Conditional entropy of a homeomorphism

In this section, T is a selfhomeomorphism of a compact metric space. (Keep in mind our notational convention of using the same letter for a selfhomeomorphism and its domain.) First we recall the definition by open covers of the conditional topological entropy $h^*(T)$ of T, introduced by Misiurewicz [Mi2]. The basic idea of $h^*(T)$ is to give useful uniform estimates for conditional measure theoretic entropies. The definition is done in stages as follows.

$$N(\mathcal{U}|\mathcal{B}) = \max_{V\in\mathcal{B}} \min\{\text{card } \mathcal{U}' : \mathcal{U}' \text{ is a subcover of } \mathcal{U}|V\} \tag{4.1}$$

$$h(T,\mathcal{U}|\mathcal{B}) = \lim_n \frac{1}{n}\log N(\mathcal{U}_0^{n-1}|\mathcal{B}_0^{n-1}) = \inf_n \frac{1}{n}\log N(\mathcal{U}_0^{n-1}|\mathcal{B}_0^{n-1}) \tag{4.2}$$

$$h(T|\mathcal{B}) = \sup_{\mathcal{U}} h(T,\mathcal{U}|\mathcal{B}) = \lim_{\mathcal{U}} h(T,\mathcal{U}|\mathcal{B}) \tag{4.3}$$

$$h^*(T) = \inf_{\mathcal{B}} h(T|\mathcal{B}) = \lim_{\mathcal{B}} h(T|\mathcal{B}) \ . \tag{4.4}$$

Here, $\mathcal{U}$ and $\mathcal{B}$ represent open covers of T, and e.g. $\mathcal{U}_0^{n-1}$ denotes the open cover which is the common refinement of the covers $T^{-i}\mathcal{U}$, $0 \le i \le n-1$. For a number a and a function α of open covers, the notation $a = \lim_{\mathcal{U}} \alpha(\mathcal{U})$ means that for any sequence $\mathcal{U}_n$ of open covers with mesh going to zero, $\lim_n \alpha(\mathcal{U}_n) = a$. It is easy to see that

$$h^*(T) \le h(T) \ , \tag{4.5}$$

$$h^*(T) = \infty \ \text{ if and only if } \ h(T) = \infty \ , \text{ and} \tag{4.6}$$

$$h^*(R) \le h^*(T) \ , \text{ for any subsystem } R \text{ of } T \ . \tag{4.7}$$

Next suppose T is zero dimensional; we will give a description of $h^*(T)$ using words in this case. Without loss of generality, suppose T is an inverse limit of subshifts T_n with surjective bonding maps $\pi_n : T_{n+1} \twoheadrightarrow T_n$. For $n > k$, let $\pi_{k,n}$ denote the composition bonding map $\pi_k \cdots \pi_{n-1} : T_n \twoheadrightarrow T_k$. Then for $k < n$ we define

$$N(T_n, T_k, M) = \max_{x\in T_k}\{y_0 \dots y_{M-1} : y \in T_n, \ \pi_{k,n} y = x\} \tag{4.8}$$

$$h(T_n|T_k) = \lim_M \frac{1}{M}\log \text{card}\mathcal{N}(T_n, T_k, M) \tag{4.9}$$

$$h(T|T_k) = \lim_n h(T_n|T_k) \tag{4.10}$$

and then it is not difficult to verify

$$h^*(T) = \lim_k h(T|T_k) \ . \tag{4.11}$$

For context and meaning, we recall from [Bow1] the metric roots of conditional topological entropy. Recall, in a system T, a set C is an n,δ spanning set for K if for any x in K there exists y in C such that $\text{dist}(T^k x, T^k y) \le \delta$ for $0 \le k < n$. For a compact (but not necessarily invariant) set K, the minimum cardinality of an n,δ spanning set for K is finite and is denoted by $r_n(K,\delta)$. The entropy of K is defined to be

$$h(K) = \lim_{\delta\to 0} \limsup_{n\to\infty} \ (1/n)\log \ r_n(K,\delta) \ . \tag{4.12}$$

For n a nonnegative integer or $n = \infty$, and for $\epsilon > 0$, define

$$B^n_\epsilon(x) = \{y\colon \mathrm{dist}(T^i x, T^i y) \le \epsilon,\ 0 \le i < n\}$$

and let $h^*(x,\epsilon) = h(B^\infty_\epsilon(x))$. Bowen [Bow1] defined

$$h^*(\epsilon) = \sup_{x\in T} h^*(x,\epsilon)\ .$$

In general the inequality $\lim_{\epsilon\to 0}\sup_{x\in T} h^*(x,\epsilon) \ge \sup_{x\in T}\lim_{\epsilon\to 0} h^*(x,\epsilon)$ can be strict, as in the following example.

Example 4.13. A system T in which $\sup_{x\in T} h^*(x,\epsilon) = \log 2$ for all ϵ, and $\lim_{\epsilon\to 0} h^*(x,\epsilon) = 0$ for all x.

Description. Let S be the 2-shift and let $O_1, O_2, \dots$ be an enumeration of the finite orbits of S. Let $T_0 = S$, and for $n \ge 1$ let $T_n = S \cup (S \times \cup_{i=1}^n O_i)$. Thus $T_n \subset T_{n+1}$ for all n. For $n \ge 1$ define $\pi_n : T_{n+1} \to T_n$ by $\pi_n(x) = x$ if $x \in T_n$ and $\pi_n(x) = z$ if $x = (y,z) \in S \times O_{n+1}$. It is not difficult to verify that the inverse limit system T constructed from the bonding maps π_n gives the required example. □

However, Bowen proved an inequality (Prop. 2.2 of [Bow1]) which implies the interchange of operations result

$$h^*(\epsilon) = \lim_{\delta\to 0}\limsup_n (1/n)\max_{x\in T}\log\, r_n(B^n_\epsilon(x),\delta)\ .$$

With this result, it is not very difficult (pp. 163-164 of [DGS] or Theorem 2.1 of [Mi2]) to verify the following claim: if $\mathcal{U}$ and $\mathcal{B}$ are open covers with $\epsilon > 0$ such that every element of $\mathcal{U}$ has diameter less than ϵ and ϵ is a Lebesgue number for $\mathcal{B}$ (i.e. any ϵ-ball is contained in some element of $\mathcal{B}$), then

$$h(T|\mathcal{U}) \le h^*(\epsilon) \le h(T|\mathcal{B})\ .$$

It then follows easily that $h^*(T) = \lim_{\epsilon\to 0} h^*(\epsilon)$.

Bowen defined a system to be *h-expansive* (entropy expansive) if $h^*(\epsilon) = 0$ for some $\epsilon > 0$. Bowen's interest in [Bow1] was that this condition allowed computation of topological entropy from any open cover of sufficiently small mesh, and computation of measure theoretic entropy from any partition of sufficiently small mesh.

Misiurewicz [Mi1] defined a system to be *asymptotically h-expansive* in the case that $\lim_{\epsilon\to 0} h^*(\epsilon) = 0$. For such a system, Misiurewicz pointed out that $\mu \mapsto h_\mu(T)$ defines an uppersemicontinuous function on the compact space of T-invariant Borel probabilities, and in particular T has a measure of maximal entropy. Denker [De] finally characterized the finite entropy systems admitting a measure of maximal entropy by introducing as a further refinement of these ideas the local conditional topological entropy (see Ch. 20 of [DGS]).

We finish this section by recalling Ledrappier's variational characterization of the conditional topological entropy of a selfhomeomorphism T of a compact metric space. (We will not apply this result, but it gives some context for Section 6.) Let T_1 and T_2 be two copies of T. For a $T_1 \times T_2$ invariant Borel probability μ, let $h(\mu|T_1)$ denote the conditional measure theoretic entropy of $T_1 \times T_2$ with respect to the measure μ given the sigma algebra corresponding to projection onto T_1. Define

$$\begin{aligned} h^*(m|T_1) &= \limsup_{\mu\to m} h(\mu|T_1) - h(m|T_1)\ , && \text{if } h(m|T_1) \text{ is finite}\ ,\\ &= \infty\ , && \text{if } h(m|T_1) = \infty\ . \end{aligned}$$

Now we can state Ledrappier's characterization from [Le]:

Theorem 4.14. [Le] **Ledrappier Variational Principle:**

$$h^*(T) = \max_m h^*(m|T_1) \ .$$

5. Conditional entropy of a quotient map

In this section, S and T are selfhomeomorphisms of compact metric spaces and $\varphi\colon S \twoheadrightarrow T$. We assume the definitions and notation of the previous section. We define the conditional entropy of the quotient map φ to be

$$e^*(\varphi) = \inf_{\mathcal{V}} h(S|\varphi^{-1}\mathcal{V}) = \lim_{\mathcal{V}} h(S|\varphi^{-1}\mathcal{V}) \tag{5.1}$$

where $\mathcal{V}$ represents an arbitrary open cover of T. If $\varphi = \mathrm{Id}_S$, then $e^*(\varphi) = h^*(S)$.

In the case that S and T are zero dimensional inverse limit systems of sequences of subshifts S_n and T_n, we can give a description of $e^*(\varphi)$ with words as follows. We let p_n denote the projection $T \twoheadrightarrow T_n$ or $S \twoheadrightarrow S_n$, and similarly let π_n denote a given bonding map $S_{n+1} \twoheadrightarrow S_n$ or $T_{n+1} \twoheadrightarrow T_n$. For $y \in T_k$, let $N(S_n, T_k, M, y)$ be the cardinality of $\{(p_n x)_0 \dots (p_n x)_{M-1} : x \in S \text{ and } p_k \varphi x = y\}$. Set

$$N(S_n, T_k, M) = \max_{y \in T_k} N(S_n, T_k, M, y) \tag{5.2}$$

$$h(S_n|T_k) = \lim_M \frac{1}{M} \log N(S_n, T_k, M) \tag{5.3}$$

$$h(S|T_k) = \lim_n h(S_n|T_k) \tag{5.4}$$

and then it is not difficult to verify

$$e^*(\varphi) = \lim_k h(S|T_k) \ . \tag{5.5}$$

Above, if S is a subshift then we may regard each S_n as S, and the limit over n is unnecessary.

The following properties are evidence for the reasonableness of the definition of the conditional entropy of a quotient map.

Facts 5.6. Suppose S and T are selfhomeomorphisms of compact metric spaces, and $\varphi\colon S \twoheadrightarrow T$. Then the following hold.

1. $e^*(\varphi|R) \le e^*(\varphi)$, for any subsystem R of S.
2. $h(\varphi^{-1}R) \le h(R) + e^*(\varphi)$, for any subsystem R of T.
3. $h_\mu(S) \le h_{\varphi_*\mu}(T) + e^*(\varphi)$, for any S-invariant Borel probability μ.
4. $\max\{e^*(\varphi), e^*(\psi)\} \le e^*(\varphi\psi) \le e^*(\varphi) + e^*(\psi)$, for any quotient map ψ.
5. $h^*(S) \le e^*(\varphi)$.
6. $h^*(T) \le h^*(S) + e^*(\varphi)$.
7. $h^*(T) \le 2e^*(\varphi)$.

Proof. We will verify the sixth property and leave the other verifications to the reader. Let $\mathcal{U}$ be an open cover of S and let $\mathcal{V}, \mathcal{B}$ be open covers of T. Then

$$N(\mathcal{V}_0^{n-1}|\mathcal{B}_0^{n-1}) = N((\varphi^{-1}\mathcal{V})_0^{n-1}|(\varphi^{-1}\mathcal{B})_0^{n-1}) \tag{5.7}$$

$$\le N((\varphi^{-1}\mathcal{V})_0^{n-1}|\mathcal{U}_0^{n-1}) \cdot N(\mathcal{U}_0^{n-1}|(\varphi^{-1}\mathcal{B})_0^{n-1}) \tag{5.8}$$

and so

$$h(T,\mathcal{V}|\mathcal{B}) \le h(S,(\varphi^{-1}\mathcal{V})|\mathcal{U}) + h(S,\mathcal{U}|(\varphi^{-1}\mathcal{B})) \tag{5.9}$$

$$\le h(S|\mathcal{U}) + h(S|(\varphi^{-1}\mathcal{B})) \ . \tag{5.10}$$

Therefore

$$h^*(T) = \inf_{\mathcal{B}} \sup_{\mathcal{V}} h(T,\mathcal{V}|\mathcal{B}) \tag{5.11}$$

$$\le h(S|\mathcal{U}) + e^*(\varphi) \tag{5.12}$$

and because $\mathcal{U}$ was arbitrary we have

$$h^*(T) \le \inf_{\mathcal{U}} h(S|\mathcal{U}) + e^*(\varphi) = h^*(S) + e^*(\varphi) \ . \tag{5.13}$$

□

Remark 5.14. One easily sees that when $e^*(\varphi) = 0$, all the inequalities in (5.6) above become equality. In particular, if $e^*(\varphi) = 0$, then both S and T must be asymptotically h-expansive.

6. A variational principle for conditional entropy of a quotient map

In this section we will establish a variational principle for the conditional entropy of a quotient map; briefly describe the Downarowicz-Serafin and Ledrappier-Walters conditional variational principles; and give Example 6.11, a quotient map with positive entropy jumps on measures but not subsystems.

For a quotient map $\varphi\colon S \twoheadrightarrow T$, we use the notation

$$\mathrm{SM}(\varphi) = \sup_m \big(h_m(S) - h_{\varphi_* m}(T)\big)$$

where the supremum is taken over the S-invariant Borel probabilities. Given m and finite m-measurable partitions P, Q we let $H_m(P|Q) = H_m(P \vee Q) - H_m(Q)$, the measure theoretic conditional entropy of P given Q. We first observe that the well known concavity of the map $m \mapsto H_m(P)$ implies the concavity of $m \mapsto H_m(P|Q)$.

Lemma 6.1. ([DS]; Lemma 3.2 of [LeW]) *Suppose P and Q are finite measurable partitions, μ and ν are probabilities, $0 < \lambda < 1$, and $m = \lambda\mu + (1-\lambda)\nu$. Then*

$$\lambda H_\mu(P|Q) + (1-\lambda)H_\nu(P|Q) \ \le \ H_m(P|Q) \ .$$

Proof. We can assume the sets B in Q have positive μ and ν measure. Define

$$\lambda_B \ = \ \lambda\mu(B) \ / \ [\lambda\mu(B) + (1-\lambda)\nu(B)]$$

and let e.g. μ_B denote the conditional measure, $\mu_B(C) = \mu(B \cap C)/\mu(B)$. Then

$$\lambda H_\mu(P|Q) + (1-\lambda)H_\nu(P|Q) = \sum_{B\in Q} \lambda\mu(B)H_{\mu_B}(P) + (1-\lambda)\nu(B)H_{\nu_B}(P)$$

$$= \sum_B [\lambda\mu(B) + (1-\lambda)\nu(B)][\lambda_B H_{\mu_B}(P) + (1-\lambda_B)H_{\nu_B}(P)]$$

$$\le \sum_B [\lambda\mu(B) + (1-\lambda)\nu(B)][H_{\lambda_B\mu_B + (1-\lambda_B)\nu_B}(P)] \ = \ H_m(P|Q) \ .$$

□

Lemma 6.2. *Suppose S and T are finite entropy selfhomeomorphisms of zero dimensional compact metric spaces, $\varphi\colon S \twoheadrightarrow T$ is a quotient map, and $e^*(\varphi)$ is the topological conditional entropy of the quotient map. Then*

$$\mathrm{SM}(\varphi) \geq e^*(\varphi) - h^*(T) \ .$$

Proof. Let $\epsilon > 0$. We show that there is an S-invariant Borel probability measure m such that

$$h_m(S) - h_{\varphi_* m}(T) > e^*(\varphi) - h^*(T) - 2\epsilon \ .$$

For a clopen partition α of S let α_n denote $\alpha \vee S^{-1}\alpha \cdots \vee S^{-(n-1)}\alpha$. Similarly, for a clopen partition β of T set $\beta_n := \beta \vee T^{-1}\beta \cdots \vee T^{-(n-1)}\beta$. Then

$$h_m(S,\alpha) - h_{\varphi_* m}(T,\beta) = \lim_{n\to\infty} \frac{1}{n} H_m(\alpha_n|\varphi^{-1}\beta_n)$$

if α is finer than $\varphi^{-1}\beta$.

Now choose a clopen partition β of T, a clopen partition α of S which is finer than $\varphi^{-1}\beta$, and an integer N such that for all $n \geq N$ it holds that

$$|e^*(\varphi) - \frac{1}{n}\log N(\alpha_n|\varphi^{-1}\beta_n)| < \epsilon \ \text{ and}$$
$$|h^*(T) - h(T|\beta)| < \epsilon.$$

For each $n \geq N$, fix a set $B_n \in \beta_n$ such that

$$\#\{A \in \alpha_n | A \subset \varphi^{-1}B_n\} \ = \ N(\alpha_n|\varphi^{-1}\beta_n) \ .$$

Let $E_n \subset S$ such that for each $A \in \alpha_n$ with $A \subset \varphi^{-1}B_n$, it holds that $\#(E_n \cap A) = 1$. Let

$$\sigma_n = \frac{1}{\#E_n} \sum_{x\in E_n} \delta_x \ ,$$

where δ_x denotes the point mass at x. Observe that $\sigma_n(\varphi^{-1}B_n) = 1$. Then

$$\begin{aligned}
H_{\sigma_n}(\alpha_n|\varphi^{-1}\beta_n) &= -\sum_{B\in\beta_n} \sigma_n(\varphi^{-1}B)\cdot \sum_{A\in\alpha_n} \sigma_n(A|\varphi^{-1}B)\log\sigma_n(A|\varphi^{-1}B) \\
&= -\sigma_n(\varphi^{-1}B_n)\cdot \sum_{x\in E_n} \frac{1}{\#E_n}\cdot\log\frac{1}{\#E_n} \\
&= \log\#E_n \ .
\end{aligned}$$

Fix an integer q. For $n > q$ large enough, for $0 \leq b < q$ and $a \geq 1$, write $n = aq + b$. Then for $0 \leq j < q$ it holds that

$$\alpha_n = \alpha_j \vee S^{-j}\alpha_q \vee S^{-(q+j)}\alpha_q \vee \cdots \vee S^{-((a-2)q+j)}\alpha_q \vee S^{-((a-1)q+j)}\alpha_{b+q-j}.$$

Thus for each $0 \le j < q$ we have for $Q(j) = \{0 \le k < j\} \cup \{(a-1)q + j \le k < n\}$ and $\sigma_{n,k} := (S^k)_* \sigma_n$ that

$$\begin{aligned}
\log \# E_n &\le H_{\sigma_n}(\alpha_n|\varphi^{-1}\beta_n) \\
&\le \sum_{r=0}^{a-2} H_{\sigma_n}(S^{-(rq+j)}\alpha_q|\varphi^{-1}T^{-(rq+j)}\beta_q) + \sum_{k\in Q(j)} H_{\sigma_n}(S^{-k}\alpha|\varphi^{-1}T^{-k}\beta) \\
&= \sum_{r=0}^{a-2} H_{\sigma_{n,rq+j}}(\alpha_q|\varphi^{-1}\beta_q) + \sum_{k\in Q(j)} H_{\sigma_{n,k}}(\alpha|\varphi^{-1}\beta) \\
&\le \sum_{r=0}^{a-2} H_{\sigma_{n,rq+j}}(\alpha_q|\varphi^{-1}\beta_q) + 3q\cdot\log(\#\alpha)
\end{aligned}$$

Adding these q inequalities, dividing by n and appealing to Lemma 6.1 we get

$$\begin{aligned}
q\cdot\frac{1}{n}\log\#E_n &\le \frac{1}{n}\sum_{p=0}^{n-1} H_{\sigma_{n,p}}(\alpha_q|\varphi^{-1}\beta_q) + 3q^2\cdot\frac{1}{n}\log(\#\alpha) \\
&\le H_{\mu_n}(\alpha_q|\varphi^{-1}\beta_q) + 3q^2\cdot\frac{1}{n}\log(\#\alpha)
\end{aligned}$$

where $\mu_n = 1/n\sum_{k=0}^{n-1}(S^k)_*\sigma_n$. Thus

$$\frac{1}{n}\log\#E_n < \frac{1}{q}H_{\mu_n}(\alpha_q|\varphi^{-1}\beta_q) + 3q\cdot\frac{1}{n}\log(\#\alpha) .$$

For a suitable subsequence n_i and a measure m we have that $\mu_{n_i} \to m$ and, since the sets of the finite partitions α_q and $\varphi^{-1}\beta_q$ are closed open, we get thus that

$$\lim_{n\to\infty}\frac{1}{n}\log\#E_n \le \frac{1}{q}H_m(\alpha_q|\varphi^{-1}\beta_q) .$$

Since this holds for all q, we conclude that

$$\lim_{n\to\infty}\frac{1}{n}\log\#E_n \le h_m(S,\alpha) - h_{\varphi_* m}(T,\beta) .$$

We thus have $e^*(\varphi) - \epsilon \le h_m(S,\alpha) - h_{\varphi_* m}(T,\beta)$, by the choice of the partitions α and β. Therefore

$$e^*(\varphi) - \epsilon \le h_m(S) - h_{\varphi_* m}(T,\beta) . \tag{6.3}$$

Now if β' is a clopen partition of T finer than β, then

$$\begin{aligned}
|h_{\varphi_* m}(T,\beta) - h_{\varphi_* m}(T,\beta')| &= \lim_{n\to\infty}\frac{1}{n}H_{\varphi_* m}(\beta'_n|\beta_n) \\
&\le \lim_{n\to\infty}\frac{1}{n}\log N(\beta'_n|\beta_n) \\
&= h(T,\beta'|\beta) .
\end{aligned}$$

Thus $h_{\varphi_* m}(T,\beta) \ge h_{\varphi_* m}(T,\beta') - h(T,\beta'|\beta)$ for every partition β' finer than β. Therefore

$$\begin{aligned}
h_{\varphi_* m}(T,\beta) &\ge h_{\varphi_* m}(T) - h(T|\beta) \ge h_{\varphi_* m}(T) - (h^*(T)+\epsilon) , \quad \text{so} \\
h_{\varphi_* m}(T,\beta) &\ge h_{\varphi_* m}(T) - h^*(T) - \epsilon .
\end{aligned}$$

Using this last inequality to substitute for $h_{\varphi_* m}(T,\beta)$ in (6.3), we get

$$h_m(S) - h_{\varphi_* m}(T) \ge e^*(\varphi) - h^*(T) - 2\epsilon$$

as required. ◻

Lemma 6.4. *Suppose $\varphi_1\colon S_1 \twoheadrightarrow T$, $\varphi_2\colon S_2 \twoheadrightarrow T$, and F is the fibered product of S_1 and S_2 by φ_1 and φ_2, with projections $p_1\colon F \twoheadrightarrow S_1$ and $p_2\colon F \twoheadrightarrow S_2$. Then* $\mathrm{SM}(\varphi_1) = \mathrm{SM}(p_2)$ *and* $\mathrm{SM}(\varphi_2) = \mathrm{SM}(p_1)$.

Remark 6.5. We only appeal to this lemma in the case $\mathrm{SM}(\varphi_1) = 0$, which follows from the easier inequality $\mathrm{SM}(\varphi_1) \geq \mathrm{SM}(p_2)$.

Proof of Lemma 6.4. Suppose $\mu, \mu_1, \mu_2, \overline{\mu}$ are Borel probabilities on F, S_1, S_2, T with $\varphi_1\mu_1 = \overline{\mu} = \varphi_2\mu_2$, $p_1\mu = \mu_1$, and $p_2\mu = \mu_2$. Let $\mathcal{A}, \mathcal{B}_1, \mathcal{B}_2, \mathcal{C}$ be the Borel σ-algebras of F, S_1, S_2, T. Then

$$\begin{aligned} h_{\mu_1}(S_1,\mathcal{B}_1) - h_{\overline{\mu}}(T,\mathcal{C}) &= h_{\mu_1}(S_1,\mathcal{B}_1|\varphi_1^{-1}\mathcal{C}) \;=\; h_\mu(F,p_1^{-1}\mathcal{B}_1|p_1^{-1}\varphi_1^{-1}\mathcal{C}) \\ &= h_\mu(F,p_1^{-1}\mathcal{B}_1|p_2^{-1}\varphi_2^{-1}\mathcal{C}) \;\geq\; h_\mu(F,p_1^{-1}\mathcal{B}_1|p_2^{-1}\mathcal{B}_2) \\ &= h_\mu(F,p_1^{-1}\mathcal{B}_1 \vee p_2^{-1}\mathcal{B}_2|p_2^{-1}\mathcal{B}_2) \;=\; h_\mu(F) - h_{\mu_2}(S_2)\ . \end{aligned} \tag{6.6}$$

It follows that $\mathrm{SM}(\varphi_1) \geq \mathrm{SM}(p_2)$. For the other direction, given $\varphi_1 : \mu_1 \mapsto \overline{\mu}$, choose μ_2 such that $\varphi_2\mu_2 = \mu$, and choose μ to be the relatively independent joining of μ_1 and μ_2 [Ru]. Then the inequality in (6.6) becomes equality, and it follows that $\mathrm{SM}(\varphi_1) = \mathrm{SM}(p_2)$. Likewise of course, $\mathrm{SM}(\varphi_2) = \mathrm{SM}(p_1)$. ◻

Recall the definition of Ledrappier [Le]: $\varphi\colon S \twoheadrightarrow T$ is a *principal extension* of T if $h_\mu(S) = h_{\varphi\mu}(T)$ for every S-invariant Borel probability μ. A system S has a principal extension to a zero dimensional system if S is finite dimensional (by Theorem B.2) or if S is asymptotically h-expansive (by Corollary A.2 and Facts 5.6). We expect that all finite entropy selfhomeomorphisms of compact metric spaces admit principal extensions to zero dimensional systems.

Theorem 6.7. Variational Principle for Quotient Maps. *Suppose $\varphi\colon S \twoheadrightarrow T$, where T is asymptotically h-expansive, $h(S) < \infty$ and S admits a zero dimensional principal extension. Then*

$$e^*(\varphi) = \mathrm{SM}(\varphi)\ .$$

Proof. We have $e^*(\varphi) \geq \mathrm{SM}(\varphi)$ (without the hypotheses on S and T) by Facts (5.6). It remains to prove the reversed inequality.

By Corollary A.2, there is a zero dimensional extension $\pi_1 : T_1 \twoheadrightarrow T$ such that $h^*(T_1) = e^*(\pi_1) = 0$. The map π_1 is a principal extension by Facts 5.6. Let $\pi_2\colon S_2 \twoheadrightarrow S$ be the assumed zero dimensional principal extension of S. Let F be the fibered product of T_1 and S_2 by the maps π_1 and $\varphi\pi_2$, with projections $p_1 : F \twoheadrightarrow T_1$ and $p_2\colon F \twoheadrightarrow S_2$. By the last lemma, $\mathrm{SM}(p_2) = \mathrm{SM}(\pi_1) = 0$ and $\mathrm{SM}(p_1) = \mathrm{SM}(\varphi\pi_2)$. Because $\mathrm{SM}(\pi_2) = 0$, we have $\mathrm{SM}(\varphi\pi_2) = \mathrm{SM}(\varphi)$, so $\mathrm{SM}(p_1) = \mathrm{SM}(\varphi)$. We also have

$$e^*(p_1) \geq e^*(\varphi) - e^*(\pi_1)$$

since $e^*(\pi_1) + e^*(p_1) \geq e^*(\pi_1 p_1) = e^*(\varphi\pi_2 p_2) \geq e^*(\varphi)$. Finally, because F and T_1 are zero dimensional, by appeal to Lemma 6.2 we get

$$\begin{aligned} \mathrm{SM}(\varphi) \;=\; \mathrm{SM}(p_1) &\geq e^*(p_1) - h^*(T_1) \\ &\geq e^*(\varphi) - e^*(\pi_1) - h^*(T_1) \\ &= e^*(\varphi)\ . \end{aligned}$$

◻

Remark 6.8. We can have $h^*(T) > 0$ with $e^*(\varphi)$ at either end of the interval $[\mathrm{SM}(\varphi), \mathrm{SM}(\varphi) + h^*(T)]$. For example, if φ is the identity map, then $\mathrm{SM}(\varphi) = 0$ and $e^*(\varphi) = h^*(T)$; whereas if φ is projection of $S = R \times T$ onto T, then $\mathrm{SM}(\varphi) = h(R) = e^*(\varphi)$.

Remark 6.9 (Downarowicz-Serafin Relative Variational Principle). Downarowicz and Serafin have proved a more general variational principle, in a manuscript [DS] we received after finishing the writing above of Theorem 6.7, near the completion of this paper. For a quotient map $\varphi\colon S \to T$, they defined the *topological conditional entropy of (X,S) given the factor (Y,T)*, $\mathbf{h}(X|Y)$, which in our notation is given by

$$\mathbf{h}(X|Y) = \sup_{\mathcal{U}} \inf_{\mathcal{B}} h(S,\mathcal{U}|\varphi^{-1}\mathcal{B})$$

where $\mathcal{B}$ ranges over open covers of Y and $\mathcal{U}$ ranges over open covers of X. This contrasts with the definition

$$e^*(\varphi) = \inf_{\mathcal{B}} \sup_{\mathcal{U}} h(S,\mathcal{U}|\varphi^{-1}\mathcal{B}).$$

As explained in [DS], it is not difficult to check that the relation of the two definitions is given by

$$\mathbf{h}(X|Y) \le e^*(\varphi) \le \mathbf{h}(X|Y) + h^*(T) \ ,$$

so $e^*(\varphi) = \mathbf{h}(X|Y)$ in the case that $h^*(T) = 0$, and therefore our Theorem 6.7 can be obtained as a corollary of their result.

Remark 6.10 (Ledrappier-Walters Relative Variational Principle). Ledrappier and Walters [LeW] proved a relative variational principle for pressure for a quotient map $\varphi\colon S \twoheadrightarrow T$, which for entropy takes the form

$$\sup_{\mu} h_\mu(S) = h_\nu(T) + \int h(S, \pi^{-1}(y))d\nu(y)$$

where the supremum is taken over all invariant measures μ such that $\pi_*\mu = \nu$. Subsequently Walters extended these developments and others in [W1].

We finish this section with an example, which in particular shows that one cannot simplify the proof of Lemma 6.2 by using a drop of topological entropy of a suitable restricted map.

Example 6.11. *There are transitive subshifts S and T and a quotient map $f\colon S \to T$ with the following properties:*

(1) *For every subsystem R of T (including $R = T$), $h(f^{-1}R) = h(R)$.*
(2) *$e^*(f) > 0$.*

We first define T as a subshift of $\{1,2,3,4,5\}^{\mathbb{Z}}$, which is obtained from the following skeleton construction. Let x denote a symbol not in $\{1,2,3,4,5\}$. Skeletons will be blocks with symbols in $\{1,x\}$. The skeleton of order 0 is $s_0 := 1$. The skeleton of order 1 is $s_1 := (s_0x)^4s_0 = s_0xs_0xs_0xs_0xs_0$, in which the 0-skeleton occurs $4^1 + 1$ times. Inductively, the $k+1$-skeleton is $s_{k+1} = (s_kx^{k+1})^{4^{k+1}}s_k = s_kx^{k+1}s_kx^{k+1}\dots x^{k+1}s_k$, in which the k-skeleton occurs $4^{k+1}+1$ times. Now we replace the symbols x in the skeletons by some symbol from $\{2,3,4,5\}$ as follows. In the first step, replace each occurrence of the symbol x in the 1-skeleton s_1 by symbols from $\{2,3,4,5\}$ such that every symbol from $\{2,3,4,5\}$ occurs, and call the resulting block s_1^1. This block has symbols in $\{1,2,3,4,5\}$; for example, s_1 could be the block $s_1^1 := 121314151$. In the second step, first replace in s_2 every occurence of s_1 with s_1^1, to obtain $(s_1^1 x^2)^{4^2} s_1^1$, and then replace in this block each occurence of x^2 with an element from $\{2,3,4,5\}^2$ such that each of these 2-blocks is used. Call the resulting block s_2^1; it has symbols in $\{1,2,3,4,5\}$. Inductively, for $k \geq 2$, first replace in s_{k+1} each of the $4^{k+1}+1$ occurences of the block s_k with s_k^1 to obtain the block $(s_k^1 x^{k+1})^{4^{k+1}} s_k^1$, and then replace the blocks x^{k+1} with elements from $\{2,3,4,5\}^{k+1}$ such that each of the elements from $\{2,3,4,5\}^{k+1}$ is used. Call the resulting block s_{k+1}^1. In this way we obtain a family of blocks $(s_k^1)_{k \geq 1}$ with symbols in $\{1,2,3,4,5\}$. Now suppose $t \in \{1,2,3,4,5\}^{\mathbb{Z}}$. Then by definition $t \in T$ if and only if for every $n \geq 0$ there is some $k \geq 1$ such that $t[-n,n]$ is a subblock of s_k^1.

Define a 1-block map $g \colon \{0,1,2,3,4,5\}^{\mathbb{Z}} \to \{1,2,3,4,5\}^{\mathbb{Z}}$ by $g(y)_0 = 1$ if $y_0 \leq 1$ and $g(y)_0 = y_0$ if $y_0 \geq 2$. Let $S := g^{-1}(T)$ and let $f \colon S \to T$ be the restriction of g to S.

Since $\{2,3,4,5\}^{\mathbb{Z}}$ is contained in T, we have $h(T) \geq \log 4$. To get an upper estimate for the entropy of S, consider the subshift T' with symbols $\{1,x\}$ such that a point t is in T' if every subblock $t_i \dots t_j$ is contained in some skeleton s_k. Consider $w = w_1 \dots w_n \in B_n(T')$ which sees at least two 1's and let $m(w) = \max\{p | \exists i, w_i \dots w_{i+p+1} = 1x^p1\}$. Thus w is a subblock of $x^{m(w)+1} s_{m(w)} x^{m(w)+1}$. Therefore the first occurrence of $1x^{m(w)}1$ in w and the first and last occurence of the symbol 1 in w determine the whole block w. Thus there are at most n^3 blocks w in $B_n(T')$ such that $m(w) = m$ for a fixed $m > 0$. There are at most $n+1$ blocks in $B_n(T')$ which do not see at least two 1's. Since $m(w) < n$, it follows that $\#B_n(T') \leq n + 1 + n(n^3) \leq 3n^4$.

For $0 \leq k \leq n$, let $B_{n,k}(T)$ denote the set of T-blocks of length n in which the symbol 1 occurs exactly k times. Then

$$\begin{aligned} \#B_n(S) &= \sum_{k=0}^{n} \#B_{n,k}(T) \cdot 2^k \leq \sum_{k=0}^{n} \#B_n(T') \cdot 4^{n-k} \cdot 2^k \\ &\leq \sum_{k=0}^{n} 3n^4 \cdot 4^{n-k} \cdot 2^k \leq 3n^5 \cdot 4^n. \end{aligned}$$

Thus $\log 4 \geq h(S) \geq h(T) \geq \log 4$.

Now let R be a subshift of T. First consider the case that there is a $t \in R$ with $t_0 = 1$. Let $n_1 > |s_1^1|$. Since $t[-n_1,n_1]$ is a subblock of some s_k^1 and $t_0 = 1$, we get that $t[-n_1,n_1]$ contains s_1^1. Consider $n_2 > n_1 + |s_2^1|$. Then $t[-n_2,n_2]$ is a subblock of some s_k^1, and since $t[-n_1,n_1]$ contains s_1^1, we get that $t[-n_2,n_2]$ contains s_2^1. Inductively we see in this way that every s_k^1 is a subblock of t. Thus t has a dense orbit in T and thus $R = T$ and $f^{-1}R = S$. If $t \in R$ implies that $t_i \neq 1$ for all i, then R is contained in $\{2,3,4,5\}^{\mathbb{Z}}$ and thus every point in R has a unique preimage. In any case, $h(f^{-1}R) = h(R)$.

Simple estimates show that the relative frequency of the symbol 1 in every k-skeleton block is greater than $1/4$, thus we get $e^*(f) \geq (1/4)\log 2 > 0$. This completes the example.

7. Asymptotically h-expansive systems

In this section we will show that an asymptotically h-expansive system has residual entropy zero. The heart of the argument is the coding construction used in the next result.

Theorem 7.1. *Suppose S is a mixing SFT and $T_1, T_2, \dots$ is a sequence of mixing SFT's such that $h(T_n) > 0$ for all n. Let T be the product system $T_1 \times T_2 \times \dots$. Then the following are equivalent.*

(1) *There exists a quotient map $\varphi \colon S \twoheadrightarrow T$.*
(2) *$h(T) < h(S)$ and $S \xrightarrow{per} T$.*

Proof. (1) $\implies$ (2) Clearly $h(T) \le h(S)$ and $S \xrightarrow{\text{per}} T$. To rule out the possibility $h(S) = h(T)$, we will appeal to the following result (Corollary 6.8 in [BT]): for a given mixing SFT S, the set of entropies of its uniform mixing SFT quotients is finite. (Here V is a uniform quotient of S if there exists a quotient map $\psi \colon S \twoheadrightarrow V$ such that $\psi_* \colon \max_S \mapsto \max_V$, where e.g. $\max_S$ denotes the unique measure of maximal entropy of S.)

So, suppose $h(T) = h(S)$ and $\varphi \colon S \twoheadrightarrow T$. Each T_n has a unique measure of maximal entropy μ_n, and because the product system T has finite entropy it follows that $\mu = \prod_n \mu_n$ is the unique measure of maximal entropy of T. There exists an S-invariant Borel probability ν such that $\varphi_* \colon \nu \mapsto \mu$ (Prop. 3.11 of [DGS]), and this measure ν must satisfy $h_\nu(S) \ge h_\mu(T)$. Because $h_\nu(S) \le h(S) = h(T) = h_\mu(T)$, the only possibility is that $\nu = \max_S$. For each n, by postcomposing φ with projection onto T_n we see that T_n is a uniform quotient of S. But the set $\{h(T_n) : n \in \mathbb{N}\}$ is infinite, since the numbers $h(T_n)$ are positive and sum to $h(T) < \infty$, and this is a contradiction.

(2) $\implies$ (1) Let R_0 denote S. For $n \ge 1$, choose mixing SFTs R_n to satisfy the following conditions:

$$(i) \quad \sum_{k=n+1}^{\infty} h(T_k) \ < \ h(R_n) \ < \ h(R_{n-1}) - h(T_n)$$
$$(ii) \quad T_n \times R_n \xrightarrow{\text{iper}} R_{n-1}$$
$$(iii) \quad \lim_n h(R_n) = 0 \ .$$

(The choice of R_n may be carried out recursively as follows. Given $h(R_{n-1}) > \sum_{k=n}^{\infty} h(T_k)$, we have $h(R_{n-1}) - h(T_n) > \sum_{k=n+1}^{\infty} h(T_k)$, so we can choose a mixing SFT satisfying (i), and for (iii) also satisfying $h(R_n) - \sum_{k=n+1}^{\infty} h(T_k) < 1/n$. Now by (i), $h(T_n \times R_n) < h(R_{n-1})$, so $|\mathcal{O}_k(T_n \times R_n)| < |\mathcal{O}_k(R_{n-1})|$ except for perhaps finitely many k. The inequality can be achieved for all k by replacing R_n with a suitable equal entropy mixing SFT cover, by appeal to the Covering Lemma 2.1 of [B1].)

For $n \ge 1$, R_n is a mixing SFT; $h(T_n \times R_n) < h(R_{n-1})$; and $T_n \times R_n \xrightarrow{\text{iper}} R_{n-1}$. Therefore, by Krieger's Embedding Theorem [Kr2], we may choose for each $n \ge 1$ an embedding i_n from $T_n \times R_n$ into R_{n-1}. Then define embeddings $j_n : T_1 \times \dots \times T_n \times R_n \hookrightarrow S$ by composition:

$$\begin{aligned} j_1 \colon (t_1, r_1) &\mapsto i_1(t_1, r_1) \\ j_2 \colon (t_1, t_2, r_2) &\mapsto i_1(t_1, i_2(t_2, r_2)) \\ j_3 \colon (t_1, t_2, t_3, r_3) &\mapsto i_1(t_1, i_2(t_2, i_3(t_3, r_3))) \\ j_4 \colon (t_1, t_2, t_3, t_4, r_4) &\mapsto i_1(t_1, i_2(t_2, i_3(t_3, i_4(t_4, r_4)))) \end{aligned}$$

and so on. Regarding j_n as an isomorphism to a subsystem S_n of S, for $n \ge k$ let $p_{k,n}$ denote the map $S_n \twoheadrightarrow T_k$ defined by following j_n^{-1} with the projection π_k

onto T_k. For every $n \geq k$ and $x \in S_n$, we have $p_{k,n}(x) = \pi_k(i_1 i_2 \cdots i_k)^{-1}(x)$, so for $n > k$ it holds that $p_{k,n}$ equals the restriction of $p_{k,n-1}$ to S_n. Also for every $n \geq k$, the map $p_{1,n} \times \cdots \times p_{k,n} \colon S_n \to T_1 \times \cdots \times T_k$ is surjective.

For each $n \geq 1$, extend the map $p_{n,n} \colon S_n \twoheadrightarrow T_n$ to some quotient map $\varphi_n \colon S \twoheadrightarrow T_n$. This is possible by the Extension Lemma 2.4 of [B1] because T_n is a mixing SFT and $S \xrightarrow{\text{per}} T_n$. If $n \geq k$, then the restriction to S_n of the map $\varphi_1 \times \cdots \times \varphi_k$ agrees with the surjection $p_{1,1} \times \cdots \times p_{k,k} \colon S_n \twoheadrightarrow T_1 \times \cdots \times T_k$. Let $\varphi = \prod_{n=1}^{\infty} \varphi_n$. Then it follows from compactness that the map $\varphi \colon S \to T$ is surjective as required. $\square$

Remark 7.2. Let us note that some obvious candidate conditions are not sufficient to ensure that T is a quotient of a shift of finite type. Suppose T is an inverse limit $T_1 \twoheadleftarrow T_2 \twoheadleftarrow T_3 \ldots$ of mixing SFT's T_n, with each bonding map $T_{n+1} \twoheadrightarrow T_n$ finite to one and noninjective. It is not difficult to verify that $h^*(T) = 0$, and therefore, by Theorem 7.4, T is a quotient of a subshift of equal entropy. However, regardless of whether T has a fixed point, T is not the quotient of any SFT (Theorem 2.10 of [B3]).

The following result is the key ingredient in the proof of Theorem 7.4.

Lemma 7.3. *Suppose $T_1, T_2, \ldots$ are subshifts, $T = T_1 \times T_2 \times \ldots$, and $h(T) < \infty$. Then there is a subshift V and a quotient map $\psi \colon V \twoheadrightarrow T$ such that $e^*(\psi) = 0$. In particular, $h(V) = h(T)$ and $\rho(T) = 0$.*

Proof. Let T_n' be a mixing SFT with a fixed point such that T_n is isomorphic to a subsystem of T_n' and $h(T_n') < h(T_n) + 2^{-n}$. Then T is isomorphic to a subsystem of $T' = \prod T_n'$ and $h(T') < \infty$. It follows from Facts 5.6 that the collection of systems for which the conclusion of the theorem holds is closed under passage to subsystems. So without loss of generality, we may assume each T_n is a mixing SFT, and there is a mixing SFT S such that $h(S) > h(T)$ and $S \xrightarrow{\text{per}} T$.

Now we return to the end of the proof of Theorem 7.1 and continue from there. For $n \geq 1$, φ maps the subshift S_n of S onto T, and the S_n are a nested sequence $S_1 \supset S_2 \supset \cdots$. Let $V = \cap S_n$. It follows from compactness that φ maps the subshift V onto T. Let ψ be the restriction of φ to V. (Remark: already we know $\rho(T) = 0$, because $h(V) = \lim h(S_n) = h(T)$.)

Fix k in $\mathbb{N}$. Let $U = U_k = T_1 \times \cdots \times T_k$; so for $u \in U$ and $i \in \mathbb{Z}$, we have $u_i = (t_i^{(1)}, \ldots, t_i^{(k)})$. By (5.5),

$$e^*(\psi) = \lim_k \lim_M \frac{1}{M} \log N(V, U_k, M) \ .$$

Let $j = j_k$. Now $S_k \supset V$ and we have

$$\begin{array}{ccccc} U \times R_k & \overset{j}{\twoheadrightarrow} & S_k & \twoheadrightarrow & U \\ (u,r) & \mapsto & s & \mapsto & u \end{array}$$

in which the map $S_k \twoheadrightarrow U$ is φ followed by the projection p_k onto U. Because j is a surjective block code, there is a positive integer J such that for any s, s' in S_k with $s_0 \ldots s_{n-1} \neq s_0' \ldots s_{n-1}'$, there exist $(u,r), (u',r')$ in $U \times R_k$ and i in $[-J, n-1+J]$ such that $j(u,r) = s$ and $j(u',r') = s'$ and $(u_i, r_i) \neq (u_i', r_i')$. Therefore, for every

u in U,

$$\begin{aligned}
&\#\{v_0 \dots v_{M-1}\colon\ (p_k\psi v)_i = u_i,\ 0 \le i \le M-1\} \\
\le\ &\#\{(u_{-J}, r_{-J}) \dots (u_{M-1+J}, r_{M-1+J})\colon\ (p_k\varphi j(u,r))_i = u_i,\ 0 \le i \le M-1\} \\
\le\ &\#\{r_{-J} \dots r_{M-1+J} : r \in R_k\} \cdot \#\{(u_{-J} \dots u_{-1})(u_M \dots u_{M-1+J})\colon\ u \in U\}
\end{aligned}$$

where the last inequality follows from $p_k\varphi j\colon (u,r) \mapsto u$. Consequently

$$\lim_M \frac{1}{M} \log N(S, U_k, M) \le h(R_k)$$

and $e^*(\psi) \le \lim_k h(R_k) = 0$. □

The following theorem was proved independently in the zero dimensional case by Downarowicz [Do2]. (The information on $e^*(\varphi)$ is not in his statement but can be derived from his construction.)

Theorem 7.4. *Suppose T is an asymptotically h-expansive selfhomeomorphism of a compact metric space. Then there exists a subshift S and a quotient map $\varphi\colon S \twoheadrightarrow T$ with $e^*(\varphi) = 0$. In particular, $h(S) = h(T)$ and $\rho(T) = 0$.*

Proof. By Corollary A.2, there is an asymptotically h-expansive zero dimensional system T' and a quotient map $\beta\colon T' \twoheadrightarrow T$ such that $e^*(\beta) = 0$. It was shown in [Do2] that any zero dimensional asymptotically h-expansive system embeds in a finite entropy product of subshifts, so without loss of generality we may assume that T' is a subsystem of a system T'' such that $h(T'') < \infty$ and T'' is a product of subshifts. By Lemma 7.3, there is a subshift S'' and a quotient map $\alpha\colon S'' \twoheadrightarrow T''$ such that $e^*(\alpha) = 0$. Define $S = \alpha^{-1}(T')$, define α' as the restriction of α to S, and define $\varphi = \beta\alpha'$. Using Facts 5.6, we have

$$e^*(\varphi) \ = \ e^*(\beta\alpha') \ \le \ e^*(\beta) + e^*(\alpha') \ \le \ e^*(\beta) + e^*(\alpha) \ = \ 0\ ,$$

so $e^*(\varphi) = 0$ and $h(S) = h(T)$. □

Remark 7.5. We thank Downarowicz, who pointed out to us [Do1] that when T is zero dimensional asymptotically h-expansive, $\rho(T) = 0$ follows easily from the product-of-subshifts case, by appeal to the characterization of asymptotically h-expansive zero dimensional systems as subsystems of products of subshifts [Do2]. Our original more complicated proof still appealed to this special case but also used additional marker arguments. What is really needed for those additional arguments is cleanly isolated by the product-of-subshifts characterization in [Do2].

Corollary 7.6. *If $h(T) = 0$, then there is a zero entropy subshift S and a quotient map $\varphi\colon S \twoheadrightarrow T$.*

Remark 7.7. Theorem 7.4 shows that $h^*(T) = 0$ implies $\rho(T) = 0$. But there is no general inequality between $h^*(T)$ and $\rho(T)$. In Example 3.1 we have T such that $0 < h^*(T) < \infty = \rho(T)$. On the other hand it is not difficult to construct from a mixing SFT S an inverse limit T of mixing SFTs such that $S \twoheadrightarrow T$ and $h(S) = h(T)$ (so $\rho(T) = 0$) but $h^*(T) > 0$.

Buzzi [Bu], extending work of Yomdin [Y], showed that if T is a C^r selfmap of a compact Riemannian m-dimensional manifold with boundary, then $h^*(T) \le \frac{m}{r} \log R(T)$, where $R(T)$ is the spectral radius of the map DT on the tangent bundle. (Actually, Buzzi defined a local entropy $h_{loc}(T)$, and this appears in his formula where we use $h^*(T)$. We avoid this $h_{loc}(T)$ because it is always equal to $h^*(T)$, and because there has been some conflicting usage of the term "local entropy": Newhouse [Ne] used a different (possibly equivalent?) definition, while Brin and Katok ([BrKa], [Ma2]) earlier gave the term a different meaning.)

Buzzi's result implies that a C^∞ system is asymptotically h-expansive. So we have the following immediate corollary to Theorem 7.4 and the theorem of Buzzi.

Theorem 7.8. *A C^∞ diffeomorphism of a compact Riemannian manifold has residual entropy zero.*

In Appendix C we give an example of a homeomorphism of a disc which has finite entropy and infinite residual entropy. For $1 \le r < \infty$, we have no results on the compatibility of positive residual entropy with C^r smoothness.

The considerations above show that some reasonable symbolic dynamics exist for a C^∞ system. A much harder problem is to construct them in an explicit and useful way.

8. Characterizing residual entropy in dimension zero

We begin with the Downarowicz characterization. Let T be a zero dimensional system, presented as an inverse limit of subshifts T_n by surjective bonding maps. Let $\mathcal{M}(T)$ denote the compact convex set of T-invariant Borel probabilities. Each μ in $\mathcal{M}(T)$ projects to a measure μ_n in $\mathcal{M}(T_n)$. Define functions h_n on $\mathcal{M}(T)$ by

$$\begin{aligned} h_n(\mu) &= h_{\mu_1}(T_1) && \text{if } n = 1\ , \\ &= h_{\mu_n}(T_n) - h_{\mu_{n-1}}(T_{n-1}) && \text{if } n > 1\ , \end{aligned}$$

so $h_\mu(T) = \sum h_n(\mu)$. Let $\mathcal{F}$ denote the set of sequences of continuous functions $f_n \colon \mathcal{M}(T) \to \mathbb{R}$ such that $f_n \ge h_n$, and let $||f||$ denote the supremum of $|f|$. Now we can state the Downarowicz characterization.

Theorem 8.1. [Do2] *Let the notation be as above for a zero dimensional system T. Then*

$$\rho(T) = \inf_{\mathcal{F}} ||\sum_{n=1}^{\infty} f_n|| - h(T)\ .$$

The rest of this section is devoted to a different (and much more modest) characterization of the residual entropy of T, also just for the case that T is zero dimensional, and also in terms of some given presentation of T as an inverse limit of subshifts T_n with surjective bonding maps $\pi_n \colon T_{n+1} \twoheadrightarrow T_n$. Without loss of generality, assume that the alphabets of the T_k are disjoint and assume that the maps π_n are one-block codes (if x and y are in T_{n+1} and $x_0 = y_0$, then $(\pi_n x)_0 = (\pi_n y)_0$). Let $\mathcal{W}_k(T_n)$ denote the set of words of length k occurring in T_n, let $\mathcal{W}(T_n) = \cup_k \mathcal{W}_k(T_n)$, let $\mathcal{W}(T) = \cup_n \mathcal{W}(T_n)$. By a *word oracle* for T we will mean a function $\alpha \colon \mathcal{W}(T) \to \mathbb{N} = \{1, 2, \dots\}$ satisfying the following two properties (in which α_n denotes the restriction of α to $\mathcal{W}(T_n)$):

- (Submultiplicative Property) If the concatenation W_1W_2 is in $\mathcal{W}(T_n)$, then $\alpha_n(W_1W_2) \le \alpha_n(W_1) \cdot \alpha_n(W_2)$
- (Extension Property) There exist positive constants $c_1, c_2, \dots$ such that for all n and all W in $\mathcal{W}(T_n)$,

  $$\sum \alpha_{n+1}(W') \le c_n \alpha_n(W)$$

  where the sum is over all the words W' such that $\pi_n(W') = W$.

Given a word oracle α, we define

$$h(\alpha_n) = \lim_k \frac{1}{k} \log \sum_{W \in \mathcal{W}_k(T_n)} \alpha_n(W)$$

(where the finite limit exists as a consequence of the submultiplicative property). For any k, n, as a consequence of the extension property we have

$$c_n \sum_{W \in \mathcal{W}_k(T_n)} \alpha_n(W) \geq \sum_{W \in \mathcal{W}_k(T_{n+1})} \alpha_{n+1}(W)$$

and therefore $h(\alpha_n) \geq h(\alpha_{n+1})$ for all n. We define the entropy of the word oracle α to be $h(\alpha) = \lim_n h(\alpha_n)$.

Theorem 8.2. *Let T be a zero dimensional system, with notations as above. Then the set of entropies of subshift covers of T equals the set of entropies of word oracles for T. In particular, if T has finite entropy then $\rho(T) = \inf_\alpha h(\alpha) - h(T)$, where the infimum is over all word oracles for T.*

Proof. First, given a subshift cover $\varphi\colon S \twoheadrightarrow T$, we will define a word oracle α for T such that $h(\alpha) = h(S)$.

For any quotient map $\psi\colon R \twoheadrightarrow V$ of subshifts and $W \in \mathcal{W}_j(V)$, we define

$$\psi^{-1}W = \{x_0 \dots x_{j-1} : x \in R, (\psi x)_0 \dots (\psi x)_{j-1} = W\} \ .$$

Let $\varphi_n = p_n \varphi$ (i.e., φ_n is φ followed by projection onto T_n). Define $\alpha = \cup \alpha_n$ by setting $\alpha_n(W) = |\varphi_n^{-1}W|$. Clearly the Submultiplicative Property holds for α.

For the Extension Property, let $r = r_{n+1}$ be a coding radius for φ_{n+1} (if j is any nonnegative integer, then $x_{-r} \dots x_{j+r-1}$ determines $(\varphi_{n+1}x)_0 \dots (\varphi_{n+1}x)_{j-1}$). Set $c_n = |\mathcal{W}_r(S)|^2$. Fix any j and any word W in $\mathcal{W}_j(T_n)$. For any U in $\varphi_n^{-1}W$, there are at most c_n words W' in $\pi_n^{-1}W$ such that $U \in \varphi_{n+1}^{-1}W'$. Also,

$$\varphi_n^{-1}W = \bigcup \{\varphi_{n+1}^{-1}W' \colon W' \in \pi_n^{-1}W\} \ .$$

Therefore

$$\begin{aligned}\sum_{W' \in \pi_n^{-1}W} \alpha_{n+1}(W') &= \sum_{W' \in \pi_n^{-1}W} |\varphi_{n+1}^{-1}W'| \\ &\leq c_n \left| \bigcup_{W' \in \pi_n^{-1}W} \varphi_{n+1}^{-1}W' \right| = c_n |\varphi_n^{-1}W| = c_n \alpha_n(W) \ .\end{aligned}$$

Therefore the extension condition holds for α. It is not difficult to check that $h(\alpha_n) = h(S)$ for all n, giving $h(\alpha) = h(S)$.

For the remaining, more difficult inclusion, let a word oracle α be given. We will construct for T a subshift cover whose entropy equals $h(\alpha)$. We will use notation of the following sort: $x[i, i+j)$ denotes the word $x_i x_{i+1} \dots x_{i+j-1}$ of length j.

Let S be a mixing SFT with entropy $h(S)$ satisfying $h(\alpha_1) < h(S)$. Also suppose S is a 1-step SFT, i.e., if $x_0 \dots x_i$ and $y_i \dots y_j$ are S-words with $x_i = y_i$, then $x_0 \dots x_i y_{i+1} \dots y_j$ is an S-word. Let Z be a zero entropy subshift containing no periodic points. We will pick certain nested subshifts

$$(Z \times S)_1 \supset (Z \times S)_2 \supset (Z \times S)_3 \supset \cdots$$

of $Z \times S$, and for each n define a quotient map ψ_n from $(Z \times S)_n$ onto some subshift containing T_n, such that $\psi_1 \times \cdots \times \psi_n$ maps $(Z \times S)_n$ onto a supersystem of $(p_1\times\cdots\times p_n)(T)$. Let $(Z\times S)_\infty=\cap_n(Z\times S)_n$. Then it is clear from compactness that $\prod_{n=1}^\infty\psi_n=\psi$ maps $(Z\times S)_\infty$ onto a supersystem of T. We will arrange that $h(Z\times S)_\infty=h(\alpha)$ and that $\psi^{-1}(T)$ is a subsystem of full entropy in $(Z\times S)_\infty$.

For this scheme, we will choose in Z certain nested clopen (marker) sets $F_1\supset F_2\supset\cdots$. For each n, there will be (large) positive integers N_n and P_n with $N_n<P_n$ such that with $(F,N,P)=(F_n,N_n,P_n)$ the following marker conditions are satisfied:

$$\text{the clopen sets } Z^iF \text{ are disjoint for } 0\le i<N, \text{ and} \tag{8.3}$$

$$\cup_{i=0}^{P-1} Z^iF=Z\ . \tag{8.4}$$

(In the last line, on the left Z is the map and on the right Z is the space.) First we define F_1. Choose $\delta>0$ such that $h(\alpha_1)+\log(1+\delta)<h(S)$. Pick some symbol a from $\mathcal{W}_1(S)$. Let $\mathcal{W}_j^a(S)$ be the set of S-words U of length j such that U begins with a and Ua is an S-word. Using the fact that S is a mixing SFT, we pick N_1 in $\mathbb{N}$ such that (using the notation $\lceil x\rceil=\min\{k\in\mathbb{Z}:k\ge x\}$) we have the following:

$$|\mathcal{W}_j^a(S)|\ \ge\ \lceil(1+\delta)^j\rceil\sum_{W\in\mathcal{W}_j(T_1)}\alpha_1(W)\ ,\qquad j\ge N_1\ . \tag{8.5}$$

(The condition (8.5) will be used to guarantee ψ_1 has image containing T_1.) Then set $P_1=2N_1$ and by the standard argument (see [Kr2] or [B1]), we choose a set F_1 in Z satisfying the marker conditions (8.3)-(8.4) with $(F,N,P)=(F_1,N_1,P_1)$.

Next, we give the recursive definition for $(F_{n+1},N_{n+1},P_{n+1})$, supposing that (F_n,N_n,P_n) has been defined. First we choose $N_{n+1}>P_n$ such that

$$(1+\frac{\delta}{n})^j\ >\ c_n\lceil(1+\frac{\delta}{n+1})^j\rceil\ ,\qquad j\ge N_{n+1}\ . \tag{8.6}$$

Then we set $N'=N_{n+1}+P_n$ and choose a set F' in Z such that $(F,N,P)=(F',N',2N')$ satisfies the marker conditions (8.3)-(8.4). Finally (to achieve $F_{n+1}\subset F_n$), for $z\in F'$ we let $h(z)=\min\{i\ge 0\colon Z^iz\in F_n\}$, and define $F_{n+1}=\{Z^{h(z)}z\colon z\in F'\}$. Let $P_{n+1}=2N'+P_n$. Then $F_{n+1}\subset F_n$ and $(F,N,P)=(F_{n+1},N_{n+1},P_{n+1})$ satisfies the marker conditions (8.3)-(8.4). This finishes the definition of the marking sets F_n.

When $Z^ix\in F_n$ and $Z^{i+j}x\in F_n$ and $Z^kx\notin F_n$ for $i<k<i+j$, then we say the integer interval $[i,i+j)$ is an F_n-marker block (of length j) for x. An F_n-marker block $[i,i+j)$ is a normalized F_n-marker block if $i=0$. For each n, there are only finitely many normalized F_n-marker blocks. If $[i,i+j)$ is an F_n-marker block for x, then we say $[0,j)$ is a normalized F_n-marker block for x at i. A point x in Z via F_n produces a tiling of the integers by normalized F_n-marker blocks.

For each n, let R_n denote the full shift on the symbols of T_n, and let $\pi_n\colon R_{n+1}\to R_n$ denote the one-block code given by the one-block coding rule which defines $\pi_n\colon T_{n+1}\to T_n$.

For a normalized F_1-marker block $B=[0,j)$, fix a subset $W_B^{a,1}(S)$ of $W_j^a(S)$ such that $\#W_B^{a,1}(S)=\lceil(1+\delta)^j\rceil\cdot\sum_W\alpha_1(W)$. Define a map $\psi_B^{a,1}\colon W_B^{a,1}(S)\to W_j(T_1)$ such that $\#(\psi_B^{a,1})^{-1}(W)=\lceil(1+\delta)^j\rceil\cdot\alpha_1(W)$ for all $W\in W_j(T_1)$. This is possible by (8.5). Then define $(Z\times S)_1$ to be the set of points (x,y) in $Z\times S$ satisfying the following condition: if $B=[0,j)$ is a normalized F_1-marker block of x at i, then $y[i,i+j)\in W_B^{a,1}(S)$. Finally, set $\psi_1(x,y)[i,i+j)=\psi_B^{a,1}(y[i,i+j))$, when

$B = [0, j)$ and B is a normalized F_1-marker block of x at i. This defines the map $\psi_1 : (Z \times S)_1 \to R_1$. Obviously $T_1 \subset \psi_1(Z \times S)_1$.

The recursive step (the definition of ψ_{n+1} assuming the definitions of $\psi_1, \dots, \psi_n$) is the main step of the proof and for this we must endure some further notation for the marker structure. For $x \in Z$, if $[i, i + j)$ is an F_n-marker block, then x determines a factorization of $[i, i + j)$ as a concatenation of F_{n-1}-marker blocks (if $n > 1$), a concatenation of F_{n-1}-marker blocks into F_{n-2}-marker blocks (if $n > 2$), and so on. We call this whole structure an $F_{[1,\dots,n]}$-marker block (of length j). Formally, if $[i, i + j)$ is an F_n-marker block of x, then the $F_{[1,\dots,n]}$-marker block B of x at i is the n-tuple $B = (B_1, \dots, B_n)$, where $B_n = [i, i + j)$ and for $1 \le k < n$, B_k is the set of F_k-marker blocks of x which are contained in $[i, i + j)$. If B is an $F_{[1,\dots,n]}$-marker block of x at i, then we define its normalization B' to be the $F_{[1,\dots,n]}$-marker block of $Z^i x$ at 0, and we say B' is the normalized $F_{[1,\dots,n]}$-marker block of x at i. For each n, there are only finitely many normalized $F_{[1,\dots,n]}$-marker blocks.

Next we state the inductive hypothesis for the recursive argument. We suppose, for each $1 \le k \le n$ and for each normalized $F_{[1,\dots,k]}$-marker block B of length j, that a subset $W_B^{a,k}(S) \subset W_j^a(S)$ and a map $\psi_B^{a,k} : W_B^{a,k}(S) \to W_j(T_k)$ are given such that the following properties hold:

(1) For each $1 \le k \le n$ and each normalized $F_{[1,\dots,k]}$-marker block B of length j it holds that $\#(\psi_B^{a,k})^{-1}(W') = \lceil (1 + \delta/k)^j \rceil \cdot \alpha_k(W')$ for each $W' \in W_j(T_k)$.

(2) For each $1 \le k < n$ and each normalized $F_{[1,\dots,k+1]}$-marker block B it holds that $W_B^{a,k+1}(S) \subset W_{b(1)}^{a,k}(S) * \cdots * W_{b(l)}^{a,k}(S)$, where $b(1) \cdots b(l)$ is the factorization of B into $F_{[1,\dots,k]}$-marker blocks.

(3) For each k with $1 \le k < n$, and each normalized $F_{[1,\dots,k+1]}$-marker block B with factorization $b(1) \cdots b(l)$ into $F_{[1,\dots,k]}$-marker blocks, it holds for each $U = U_1 \cdots U_l \in W_B^{a,k+1}(S)$ with $U_i \in W_{b(i)}^{a,k}(S)$ that $\pi_k(\psi_B^{a,k+1}(U)) = \psi_{b(1)}^{a,k}(U_1) * \cdots * \psi_{b(l)}^{a,k}(U_l)$.

We shall now define for each normalized $F_{[1,\dots,n+1]}$-marker block B of length j a subset $W_B^{a,n+1}(S) \subset W_j^a(S)$ and a map $\psi_B^{a,n+1} : W_B^{a,n+1}(S) \to W_j(T_{n+1})$ such that the inductive hypothesis holds with $n+1$ in place of n. Let $B = (B_1, \dots, B_{n+1})$ be a normalized $F_{[1,\dots,n+1]}$-marker block, where $B_{n+1} = [0, j)$ and $b(1) \cdots b(l)$ is the factorization of B into $F_{[1,\dots,n]}$-marker blocks of lengths $j_1, \dots, j_l$. For each $W \in W_j(T_n)$ it holds that

$$\begin{aligned}
&\#\{U = U_1 \cdots U_l \in W_{b(1)}^{a,n}(S) * \cdots * W_{b(l)}^{a,n}(S) : \ \psi_{b(1)}^{a,n}(U_1) * \cdots * \psi_{b(l)}^{a,n}(U_l) = W\} \\
&\quad \ge \ \lceil (1 + \frac{\delta}{n})^j \rceil \cdot \alpha_n(\psi_{b(1)}^{a,n}(U_1)) \cdot \cdots \cdot \alpha_n(\psi_{b(l)}^{a,n}(U_l)) \\
&\quad \ge \ \lceil (1 + \frac{\delta}{n})^j \rceil \cdot \alpha_n(W) \\
&\quad \ge \ \lceil (1 + \frac{\delta}{n+1})^j \rceil \cdot \sum_{W'} \alpha_{n+1}(W') \ ,
\end{aligned}$$

where the first inequality holds by the induction hypothesis (1); the second holds by the submultiplicitivity of α_n; and, because $j > N_{n+1}$, the last holds by (8.6) and the extension property of α_n.

Thus we can choose a set $W_B^{a,n+1}(S) \subset W_{b(1)}^{a,n}(S) * \cdots * W_{b(l)}^{a,n}(S)$ such that $\#W_B^{a,n+1}(S) = \lceil (1+\delta/(n+1))^j \rceil \cdot \sum_{W'} \alpha_{n+1}(W')$ and a map $\psi_B^{a,n+1} \colon W_B^{a,n+1}(S) \to W_j(T_{n+1})$ such that $\#(\psi_B^{a,n+1})^{-1}(W') = \lceil (1+\delta/(n+1))^j \rceil \cdot \alpha_{n+1}(W')$ for each $W' \in W_j(T_{n+1})$ and such that $\pi_n(\psi_B^{a,n+1}(U) = \psi_{b(1)}^{a,n}(U_1) * \cdots * \psi_{b(l)}^{a,n}(U_l)$ for all $U = U_1 \dots U_l \in W_B^{a,n+1}(S)$. This new family of subsets of S-blocks and maps satisfies the induction hypotheses with $n+1$ in place of n. This finishes the recursive step.

For $n \geq 1$ we define the subshift $(Z \times S)_{n+1} := \{(x,y) \in Z \times S \colon$ if B is a normalized $F_{[1,\dots,n+1]}$-marker block of length j of x at i, then $y[i,j) \in W_B^{a,n+1}(S)\}$. Define $\psi_{n+1}(x,y)[i,j) = \psi_B^{a,n+1}(y[i,j))$ if B is a normalized $F_{[1,\dots,n+1]}$-marker block of length j of x at i. This defines a map $\psi_{n+1} \colon (Z \times S)_{n+1} \to R_{n+1}$. Obviously $T_{n+1} \subset \psi_{n+1}(Z \times S)_{n+1}$ and $\pi_n \psi_{n+1}(x,y) = \psi_n(x,y)$ for all $(x,y) \in (Z \times S)_{n+1}$.

Next, we check that $(Z \times S)_\infty$ has entropy equal to $h(\alpha)$. Suppose $x' \in Z$ and x' has an $F_{[1,\dots,n]}$-marker block of length j at i and $W \in \mathcal{W}_j(T_n)$. Then

$$\begin{aligned} &|\{(x',y)[i,i+j) \colon (x',y) \in (Z \times S)_n \text{ and } (\psi_n(x',y))[i,i+j) = W\}| \\ &= \lceil (1+\frac{\delta}{n})^j \rceil \alpha_n(W) \ . \end{aligned}$$

Here $N_n \leq j \leq P_n$ and therefore when $n > 1$ we have

$$0 \ \leq \ \frac{1}{j}\log\lceil (1+\frac{\delta}{n})^j \rceil \ \leq \ \frac{1}{j}\log((1+\frac{\delta}{n-1})^j) \ = \ \log(1+\frac{\delta}{n-1}) \ := \ \gamma(n)$$

(where the second inequality holds because the c_n in (8.6) are positive integers). Because $h(Z) = 0$, after considering concatenations of W's we conclude

$$h(\alpha_n) \ \leq \ h((Z \times S)_n) \ \leq \ h(\alpha_n) + \gamma(n) \ .$$

Because $\lim_n \gamma(n) = 0$, we conclude

$$h(\alpha) = \lim h(\alpha_n) = \lim h((Z \times S)_n) = h((Z \times S)_\infty) \ .$$

It only remains to see that the subsystem $\psi^{-1}(T)$ of $(Z \times S)_\infty$ has full entropy. Suppose that (x,y) is a point in $(Z \times S)_\infty$ such that $(\psi_1 \times \psi_2 \times \dots)(x,y) \notin T$. Since $\pi_n \psi_{n+1}(x,y) = \psi_n(x,y)$ for all n, there is thus n_0 such that $\psi_n(x,y) \notin T_n$ for all $n \geq n_0$. Thus there is $k \geq 0$ such that $\psi_{n_0}(x,y)[-k,k] \notin W(T_{n_0})$. Thus, since the π_n are 1-block maps, $\psi_n(x,y)[-k,k] \notin W(T_n)$ for all $n \geq n_0$. Therefore, for all $n \geq n_0$, the interval $[-k,k]$ is not contained in an F_n-marker block of x. Thus, there is some i in $[-k,k]$ such that $Z^i x \in F_n$ for all $n \geq n_0$. This shows that $(Z \times S)_\infty - \psi^{-1}(T)$ is contained in the set $E = \cup_i \cup_{N \geq 1} \cap_{n \geq N} (Z \times S)^{-i}(F_n \times S)$. Because the sets $(Z \times S)^j(F_n \times S)$ are disjoint for $0 \leq j < N_n$, we have that for every invariant measure μ and every n, $\mu(F_n \times S) \leq 1/(N_n)$. Because $\lim_n N_n = \infty$, we conclude that E has measure zero with respect to any invariant probability, and it follows then from the variational principle that $h(\psi^{-1}(T)) = h((Z \times S)_\infty$.

This concludes the proof.

□

Appendix A. Zero dimensional covers

The next result, without the condition on $e^*(\varphi)$, was proved independently by the first author and Klaus Thomsen by essentially the same construction. (Without any entropy condition, this is an old result of Anderson [A].) Thomsen's result was circulated in the preprint [T1] (and he considers more generally systems T which are continuous but not necessarily injective or surjective). In [B2] the first author announced his result, which eventually appeared as the supporting result Prop. 2.5 in [GW]. We revisit the construction below because we want to establish the inequality for $e^*(\varphi)$; the basic construction is unchanged, but additional argument and care are required.

Theorem A.1. *Suppose T is a selfhomeomorphism of a compact metric space Y and $h(T) < \infty$. Then there is a selfhomeomorphism S of a zero dimensional compact metric space X and a quotient map $\varphi\colon S \twoheadrightarrow T$ such that $h(S) = h(T)$ and moreover $h^*(S) \leq e^*(\varphi) \leq h^*(T)$.*

Proof. PART I. In this part we describe an ingredient of the construction.

Let $\mathcal{P}$ be a finite open cover of Y. Given a positive integer N, let $\mathcal{C}$ be a minimum cardinality subcover of the common refinement $\bigvee_{i=0}^{N-1} T^{-i}\mathcal{P}$. Let $\mathcal{C} = \{C_1, \dots, C_m\}$. For each C_i, fix a choice of elements $P(i,j)$ of $\mathcal{P}$ such that $C_i = \cap_{j=0}^{N-1} T^{-j}P(i,j)$. For $1 \leq i \leq m$, let $W(i)$ be the word $C_i 00\dots 0$, where C_i is followed by $N-1$ zeros (and so $W(i)$ has length N.) Let S' be the subshift on all concatenations of the words $W(i)$ (so, S' is conjugate to the tower of height N over the m-shift). Note, for large N the entropy of S' will be close to the entropy of T with respect to the open cover $\mathcal{P}$.

To each x in S', associate a bisequence $\overline{x}$ as follows. If $x_t \cdots x_{t+N-1} = C_i 00\cdots 0$, then set $\overline{x}_{t+j} = P(i,j)$. Define $K(x) = \cap_{n\in\mathbb{Z}} \overline{T^{-n}\overline{x}_n}$, a closed (possibly empty) subset of T. Now define $S(\mathcal{P}, \mathcal{C}, N)$ as the subshift on $\{x \in S'\colon K(x) \neq \emptyset\}$. This subshift has entropy at most $h(S')$; also, T is covered by the sets $K(x)$, $x \in S(\mathcal{P}, \mathcal{C}, N)$.

PART II. For a suitable refining sequence of open covers $\mathcal{P}_n$ of Y, we will construct as above $S_n := S(\mathcal{P}_n, \mathcal{C}_n, N_n)$. For each $n > 0$, we will define bonding maps $S_n \to S_{n-1}$. This will give us an inverse limit system S. For a point x in S, $x = (x^{(1)}, x^{(2)}, \dots)$, the sets $K(x^{(n)})$ will be nested and their intersection will be a point, y. The desired quotient map $\varphi\colon S \twoheadrightarrow T$ will be defined by sending x to y.

The construction is recursive. For $n = 0$, let $\mathcal{P}_0$ be the cover $\{X\}$ of X and let $N_0 = 1$; so, S_0 is the one-point subshift. Notation: given an open cover $\mathcal{U} = \{U_1, \dots, U_m\}$, we let U_i^* be the union of the sets U_j such that $U_i \cap U_j \neq \emptyset$, and we let $\mathcal{U}^*$ denote the open cover $\{U_1^*, \dots, U_m^*\}$.

Fix $\epsilon_1 > \epsilon_2 > \dots$, some arbitrary sequence of positive numbers decreasing to zero, and suppose the construction has been carried out for $0, \dots, n-1$. Choose a finite open cover $\mathcal{P} := \mathcal{P}_n$ such that

- $\mathcal{P}$ refines $\mathcal{C}_{n-1}$ (every element of $\mathcal{P}$ is contained in some element of $\mathcal{C}_{n-1}$),
- the mesh of $\mathcal{P}$ is less than ϵ_n,
- $h(T) - h(T, \mathcal{P}) < \epsilon_n$, and
- $h(T|\mathcal{P}) - h^*(T) < \epsilon_n$.

Then choose $N = N_n$ and $\mathcal{C} = \mathcal{C}_n$ such that

- N is a multiple of N_{n-1}, and
- $\mathcal{C}$ is a minimum cardinality subcover of the join of $\mathcal{P}, \dots, T^{-(N-1)}\mathcal{P}$ such that $|h(T) - (1/N)\log(\#\mathcal{C})| < \epsilon_n$, and
- for $1 \leq k < n$,

$$\frac{1}{N}\log N((\mathcal{P}_n)_0^{N-1} | (\mathcal{P}_k^*)_0^{N-1}) \ < \ h(T|\mathcal{P}_k^*) + \epsilon_n \ .$$

Then define $S_n = S(\mathcal{P}, \mathcal{C}, N)$.

Next we define the bonding map $\pi\colon S_n \to S_{n-1}$ (which typically will not be surjective). To do this, for each of the words $W = W(i)$ of length N used to define $S_n = S(\mathcal{P}, \mathcal{C}, N)$ as in Part I, we will define an S_{n-1} word W' of length N, and set $(\pi x)[j, j+N-1] = W'$ whenever $x[j, j+N-1] = W$. So, consider $W = \mathcal{C}_i 00\dots 0$, and recall $\mathcal{C}_i = \bigcap_{j=0}^{N-1} T^{-j}\mathcal{P}(i,j)$. Let $K = N/(N_{n-1})$. Using the refinement condition, for $0 \le k < K$ pick $C_{I(k)} \in \mathcal{C}_{n-1}$ such that $P(i, kN_{n-1}) \subset C_{I(k)}$. Let V_k denote the word which is the symbol $C_{I(k)}$ followed by $N_{n-1}-1$ zeros. Then set $W' = V_0 V_1 \cdots V_{K-1}$.

This mapping rule on words gives a well defined map $\pi\colon S_n \to S_{n-1}$ because if a concatenation of W's corresponds to a nonempty set B in T, then the corresponding concatenation of W's corresponds to a set which contains B, and is therefore nonempty. For any $x = (x^{(1)}, x^{(2)}, \dots) \in S$, we have $K(x^{(1)}) \supset K(x^{(2)}) \supset \dots$, with the diameters of the $K(x^{(n)})$ going to zero (because the mesh of $\mathcal{P}_n$ goes to zero). So, the rule $x \mapsto \cap_n K(x^{(n)})$ gives a well defined map φ from S to T. The map φ is surjective by a compactness argument because for each n, the union of the sets $K(x^{(n)})$ is all of T. The map φ is obviously equivariant.

Part III. It remains to check the entropy claims. Because $\varphi\colon S \twoheadrightarrow T$ is surjective, $h(T) \le h(S)$. On the other hand, clearly

$$h(S) \le \varlimsup_n h(S_n) \le \lim_n \big(h(T) + \epsilon_n\big) = h(T) \ .$$

So it remains to verify $e^*(\varphi) \le h^*(T)$. We will check that $\lim h(S|\varphi^{-1}\mathcal{P}_k) \le h^*(T)$.

So, fix $\mathcal{P}_k$, fix $n > k$ and let $N = N_n$. Suppose for elements P_i of $\mathcal{P}_k$ that

$$U \ = \ \bigcap_{i=0}^{N-1} (\varphi^{-1}T^{-i}P_i) \ \in \ \bigvee_{i=0}^{N-1} S^{-i}(\varphi^{-1}\mathcal{P}_k) \ .$$

Define the set of words

$$E = \{y_0 y_1 \dots y_{N-1} \colon \exists x \in U, y = x^{(n)}, y_0 \ne 0\} \ .$$

Suppose we have the following CLAIM:

$$\frac{1}{N}\log\#E \ \le \ h(T|\mathcal{P}_k^*) + \epsilon_n \ .$$

Let $\mathcal{Q}_n$ be the open cover/partition of S according to $x_0^{(n)}$. It follows from the claim that

$$\begin{aligned} h(S, \mathcal{Q}_n|\varphi^{-1}\mathcal{P}_k) &= \inf_M \frac{1}{M}\log N\big((\mathcal{Q}_n)_0^{M-1}|(\varphi^{-1}\mathcal{P}_k)_0^{M-1}\big) \\ &\le h(T|\mathcal{P}_k^*) + \epsilon_n \end{aligned}$$

and consequently (because the mesh of $\mathcal{P}_k^*$ goes to zero)

$$\begin{aligned} e^*(\varphi) = \lim_k h(S|\varphi^{-1}\mathcal{P}_k) &= \lim_k \lim_n h(S, \mathcal{Q}_n|\varphi^{-1}\mathcal{P}_k) \\ \le \lim_k h(T|\mathcal{P}_k^*) &= h^*(T) \ . \end{aligned}$$

It remains then to prove the Claim. So suppose $x \in U$, $w = x^{(k)}$ and $y = x^{(n)}$. We have associated sequences $\overline{w}$ and $\overline{y}$ on symbols from $\mathcal{P}_k$ and $\mathcal{P}_n$ respectively. Because $x \in U$, for $0 \le i < N$ the closure of the set $\overline{w}_i$ must intersect P_i, so the open set $\overline{w}_i$ must intersect the open set P_i, and $\overline{w}_i \subset P_i^*$. Then because $\mathcal{P}_n$ refines $\mathcal{P}_k$, the set $\cap_{i=0}^{N-1} T^{-i} P_i^*$ contains the set $\cap_{i=0}^{N-1} T^{-i}\overline{y}_i$. Now the cardinality of E cannot exceed $N((\mathcal{P}_n)_0^{N-1})|(\mathcal{P}_k^*)_0^{N-1})$, because if it did, we could replace in $\mathcal{C}_n$ the subcollection of elements contained in $\cap_{i=0}^{N-1} T^{-i} P_i^*$ with a smaller subcollection covering $\cap_{i=0}^{N-1} T^{-i} P_i^*$, and thus contradict the choice of $\mathcal{C}_n$ as a minimum cardinality cover. Consequently we have

$$\#E \leq N((\mathcal{P}_n)_0^{N-1}|(\mathcal{P}_k^*)_0^{N-1})$$

and now the Claim follows from the construction of $\mathcal{S}_n$. $\square$

Corollary A.2. *Suppose T is asymptotically h-expansive. Then there is an asymptotically h-expansive zero-dimensional system S and a quotient map $\varphi\colon S \twoheadrightarrow T$ such that $h(S) = h(T)$ and $e^*(\varphi) = 0$.*

Proof. By the previous result, we have $\varphi\colon S \twoheadrightarrow T$ with S zero dimensional such that $h(S) = h(T)$ and $h^*(S) \leq h^*(T) = 0$. $\square$

Remark A.3. For $\varphi\colon S \to T$, recall from Facts 5.6 that $e^*(\varphi) \geq \frac{1}{2} h^*(T)$. So in the corollary above, the assumption $h^*(T) = 0$ is necessary for $e^*(\varphi) = 0$.

Remark A.4. It is not possible without further hypotheses to add to the conclusion of Theorem A.1 the requirement $h^*(S) = 0$. This is because an asymptotically h-expansive system can be covered by a subshift, but there are T of finite entropy which cannot be covered by a subshift.

Question A.5. Is every system T covered by an equal entropy zero dimensional system of equal residual entropy?

We see the last question does have an affirmative answer when T is asymptotically h-expansive or (from the next section) finite dimensional.

Appendix B. Zero dimensional covers of finite dimensional systems

If T is a finite dimensional system (that is, its domain has finite covering dimension), then there are very strong results on the existence of covers $\varphi\colon S \twoheadrightarrow T$, with S zero dimensional and φ giving a good approximation of T by S. We will state two theorems, and then explain how they follow from the work of Kulesza [Ku1] and Thomsen[T1, T2, T3].

Theorem B.1. *Suppose T is finite dimensional and the set of periodic points of T is zero dimensional. Then there is a zero dimensional system S and a quotient map $\varphi\colon S \twoheadrightarrow T$ such that the following hold.*

1. *φ is at most $(n+1)^n$ to one.*
2. *φ is residually one-to-one.*
3. *φ has defect zero.*
4. *$\rho(S) = \rho(T)$.*

That φ is residually one-to-one means that there are residual (second category) sets in S and T such that restriction of φ gives a bijection between these sets. The meaning of "defect zero" is explained below.

Theorem B.2. *Suppose T is finite dimensional. Then there is a zero dimensional system S and a quotient map $\varphi\colon S \to T$ such that the following hold.*

1. *For every S-invariant Borel probability μ, $h_\mu(S) = h_{\varphi_*\mu}(T)$.*
2. *For every subsystem R of T, $h(R) = h(\varphi^{-1}R)$.*
3. *$\rho(S) = \rho(T)$.*

The condition 1 in Theorem B.2 is the condition Ledrappier [Le] used to define S as a *principal extension* of T. Let us note that this condition, and even more so the condition of Theorem B.3 below, become particularly subtle to arrange when there are uncountably many T-invariant ergodic Borel probabilities. In this case, when constructing regular closed partitions, itineraries through which will generate symbolic sequences for S, one cannot easily perturb partition boundaries to null sets for all measures of interest.

For $\varphi\colon S \twoheadrightarrow T$ with S zero dimensional, we now recall Thomsen's definition [T2] of the defect $D(\varphi)$ of the factor map φ. (Thomsen considered systems which are not necessarily homeomorphisms; but here as in the rest of this paper we consider only homeomorphisms.) The definition has several layers. Given a finite collection $\mathcal{F} = \{F_i : i \in I\}$ of subsets of T, we set

$$q_k(x,\mathcal{F}) = \#\{(i_1, i_2, \ldots, i_k) : x \in \bigcap_{j=1}^{k} T^{-j+1}(F_{i_j})\}$$

for all $x \in T$, $k \in \mathbb{N}$, and then

$$q_k(T,\mathcal{F}) = \max_{x\in T} q_k(x,\mathcal{F})$$

and

$$Q(T,\mathcal{F}) = \lim_n \frac{1}{n} \log q_n(T,\mathcal{F}) \ .$$

Then we define the defect of φ as

$$D(\varphi) = \sup_{\mathcal{P}} Q(T, \varphi(\mathcal{P}))$$

where the supremum is over all clopen partitions of S (i.e. partitions of S into disjoint nonempty closed open sets). For example, $D(\varphi) = 0$ if φ is bijective. There is an easy but important observation (Lemma 6.6 of [T1]): if $\mathcal{P}_1$ and $\mathcal{P}_2$ are clopen partitions and $\mathcal{P}_2$ refines $\mathcal{P}_1$, then $Q(T,\varphi(\mathcal{P}_1)) \le Q(T,\varphi(\mathcal{P}_2))$. The meaning of the defect is captured by Thomsen's Defect Variational Principle:

Theorem B.3. [T3] *Suppose $\varphi\colon S \twoheadrightarrow T$ and S is zero dimensional. Then*

$$D(\varphi) = \sup_{\mu} \int \log \# \varphi^{-1}(x) \ d\mu(x),$$

where the supremum is over all T-invariant Borel probability measures (or equivalently over all T-invariant ergodic Borel probability measures).

Let us consider the relation of $D(\varphi)$ and $e^*(\varphi)$. Clearly all values of $D(\varphi)$ are compatible with $e^*(\varphi) = 0$. On the other hand, if $\varphi = Id_T$, then $D(\varphi) = 0$ but $e^*(\varphi) = h^*(T)$. In the case that T is asymptotically h-expansive, we have the following result.

Proposition B.4. *Suppose $\varphi\colon S \twoheadrightarrow T$, S is zero dimensional, $h^*(T) = 0$ and $D(\varphi)$ is finite. Then $e^*(\varphi) = 0$.*

Proof. If $D(\varphi)$ is finite, then it follows from the Defect Variational Principle that φ is a principal extension. From the Variational Principle for Quotient Maps (6.7), we then conclude that $e^*(\varphi) = 0$. $\square$

A zero dimensional extension $\varphi\colon S \twoheadrightarrow T$ can be used to study periodic points or invariant measures for T by studying periodic points or invariant measures for S. This approach requires a reasonable correspondence under φ of these objects. If the extension φ has defect zero, then the correspondence is close as possible.

Facts B.5. *Suppose S is zero dimensional and $\varphi : S \twoheadrightarrow T$ has defect zero. Then the following hold.*

1. $h_\mu(S) = h_{\varphi_*\mu}(T)$, *for every S-invariant Borel probability.*
2. *Every T-periodic point has a unique φ preimage.*
3. *For every positive integer n,* $\#\{x : S^n x = x\} \;=\; \#\{x : T^n x = x\}$.
4. *For every subsystem R of T,* $h(\varphi^{-1}R) = h(R)$.

The first three facts are obvious from the Defect Variational Principle, and the fourth follows by also applying the usual variational principle.

Thomsen defined a zero dimensional extension $\varphi\colon S \twoheadrightarrow T$ to be *perfect* when φ is bounded to one with $D(\varphi) = 0$, and constructed perfect extensions for several classes of systems T. He also introducted the *logarithmic covering dimension* of T [T1], and showed this vanishes if and only there is a zero dimensional extension $\varphi\colon S \twoheadrightarrow T$ with defect zero. Positive dimension of the set of periodic points is an obstruction to existence of a defect zero extension [T1].

Next we recall a theorem of Kulesza.

Theorem B.6. [Ku1] *Suppose T is a self homeomorphism of a compact metrizable space of dimension $n < \infty$, and the periodic point set of T is zero-dimensional. Then there is a self homeomorphism S of a zero dimensional compact metrizable space, and a quotient map $\varphi\colon S \twoheadrightarrow T$ such that no point of T has more than $(n+1)^n$ preimages.*

The statement of the theorem is false without the hypothesis on the periodic point set [Ku1]. However, the bound $(n + 1)^n$ can be improved to $(n + 1)$[Ku2], which of course is best possible [HW].

Before proving our two theorems, we isolate a lemma.

Lemma B.7. *Suppose $\varphi\colon S \twoheadrightarrow T$, S is zero dimensional and φ is uniformly finite to one. Then $\rho(S) = \rho(T)$.*

Proof. Clearly $h(S) = h(T)$ and every subshift cover of S is a subshift cover of T, so it suffices to show, given a subshift S' and $\gamma\colon S' \twoheadrightarrow T$, that there is some subshift cover S'' of S such that $h(S'') = h(S')$. For this, let F be the fibered product of S' and S by the maps γ and φ. That is, F is the subset of $S' \times S$ on the points (x, y) such that $\gamma(x) = \varphi(y)$, and the projection map $(x, y) \mapsto y$ maps F onto S. The projection $p\colon F \twoheadrightarrow S'$ (given by $(x, y) \mapsto x$) is uniformly finite to one, so $h(F) = h(S')$ and $SM(p) = 0$. Because F is zero dimensional and the subshift S' has conditional topological entropy zero, it follows from Lemma 6.2 (or the full Variational Principle 6.7) that $e^*(p) = 0$. Then Facts 5.6(5) gives $h^*(F) = 0$, and F is asymptotically h-expansive. It follows from Theorem 7.4 that there exists a quotient map $\beta\colon S'' \twoheadrightarrow F$ such that S'' is a subshift and $h(S'') = h(F)$. Then we have $S'' \twoheadrightarrow F \twoheadrightarrow S$ and $h(S'') = h(F) = h(S')$ as required. □

Now we can prove our two theorems.

Proof of Theorem B.1. Theorem B.6 is the main result of [Ku1] (and from this and the last lemma it follows that $\rho(S) = \rho(T)$). Examining the map φ which Kulesza constructed, we see it is residually one-to-one (this is easy) and has defect zero. We will describe a little of his construction to indicate why $D(\varphi) = 0$.

Kulesza constructed a certain sequence of closed regular covers $\mathcal{D}_i$. The covering zero dimensional system S can be viewed as the inverse limit of subshifts S_i on alphabets $\mathcal{D}_i$. In this construction, for any given clopen partition $\mathcal{P}$ there is some i such that $\mathcal{P}$ is refined by the time-zero partition $\mathcal{P}_i$ for S_i. So, φ will have defect zero if for each i, $Q(T, \varphi(\mathcal{P}_i)) = 0$. Here $\varphi(\mathcal{P}_i)$ is the cover $\mathcal{D}_i$, and the conclusion will follow if for some positive integer M_i, we have for every x in T and $k \in \mathbb{N}$ that $q_k(x, \mathcal{D}_i) \leq M_i$.

The cover $\mathcal{D}_i$ is defined as

$$\mathcal{D}_i \;=\; \bigvee_{-i \leq h \leq i} \left(\bigvee_{j \leq i} T^h(C_j) \right)$$

where the C_j are finite regular closed covers constructed by Kulesza such that for every x,

$$\#\{m \in \mathbb{Z} : T^m x \in \bigcup_{j>0} \mathrm{bd}(C_j)\} \;\leq\; (n+1)^n \;.$$

Here "regular closed"means that each element of the cover is the closure of its interior, and distinct elements have disjoint interiors. It follows that if a point x is contained in more than one element of the cover $T^{-m}\mathcal{D}_i$, then $T^{m+h}x$ must lie in one of the sets $\mathrm{bd}C_j$, for some integer h such that $-i \leq h \leq i$. Thus for all x,

$$q_k(x, \mathcal{D}_i) \;\leq\; [(n+1)^n]^{2i+1}$$

and this shows $D(\varphi) = 0$. $\square$

Proof of Theorem B.2. Let Z be a zero entropy subshift without periodic points. Now $T \times Z$ has no periodic points, so by Kulesza's theorem there is a quotient map $\psi\colon S \twoheadrightarrow T \times Z$ such that S is zero dimensional and ψ is uniformly finite to one. Let π be the projection $T \times Z \twoheadrightarrow T$ and let $\varphi = \pi\psi$. Clearly $\rho(T) = \rho(T \times Z)$. By Lemma B.7, $\rho(T \times Z) = \rho(S)$, so $\rho(T) = \rho(S)$. The straightforward verification of the other claims is left to the reader. $\square$

Remark B.8. There is no general inequality between $\rho(T)$ and the minimum defect of a quotient map from a zero dimensional space onto T. If T is zero dimensional, then the minimum defect is obviously zero, but $\rho(T)$ is arbitary in $[0, +\infty]$. On the other hand, if T is the identity map on a compact metrizable space of dimension $n \in \{0, 1, \ldots, \infty\}$, then the minimum defect is $\log(n+1)$ but $\rho(T) = 0$.

Remark B.9. In the special case that T is expansive, it is a simple consequence of uniform continuity that any zero dimensional extension $S \twoheadrightarrow T$ factors through some subshift extension, $S \twoheadrightarrow S' \twoheadrightarrow T$. Then the defect of $S' \twoheadrightarrow T$ is at most that of $S \twoheadrightarrow T$. For T is expansive, it is well known that $\dim(T)$ is finite [Ma1] and $\mathrm{Per}(T)$ is countable ([DGS], Prop. 16.10). Consequently we have the following corollary of Theorem B.1.

Corollary B.10. *If T is an expansive homeomorphism, then there is a cover $\varphi\colon S \twoheadrightarrow T$ such that S is a subshift and φ has defect zero.*

Appendix C. Infinite residual entropy on a surface

The purpose of this appendix is to construct a selfhomeomorphism T of the unit disc $\mathbb{D} = \{z \in \mathbb{C}\colon |z| \leq 1\}$ such that T has finite entropy and infinite residual entropy. T will be the identity on the boundary of $\mathbb{D}$, so this example can be realized on any surface. Parts of the construction can be done smoothly and parts more generally. We thank Mike Handel, Judy Kennedy, Mike Shub and John Smillie for helpful consultations.

Fix a homeomorphism $T_0\colon \mathbb{D} \to \mathbb{D}$ with the following properties:

- T_0 has finite entropy
- T_0 is C^1 with $\det(DT_0) > 0$
- $T_0 = \mathrm{Id}$ on the boundary of $\mathbb{D}$
- there is a subset E of $\mathrm{int}(\mathbb{D})$ such that $T_0|_E$ is a mixing SFT S of entropy $\log \lambda > 0$.

(We do not have an explicit reference for the existence of such a T_0, but it is not difficult to construct T_0 (with $\lambda = 2$) by suitably extending Smale's horseshoe construction ([S], pp. 772-773) to a disc diffeomorphism.) Fix α such that $1 < \alpha < \lambda$. Below, $D(q, \epsilon)$ represents a closed disc of radius ϵ centered at q, also we use notation for annuli such as the following: $[a < |z - q_n| \leq b]$ represents $\{z : a < |z - q_n| \leq b\}$. Also, $P^0_n(S)$ denotes the set of points in S-orbits of cardinality n.

Lemma C.1. *There is a collection of pairwise disjoint discs $D(q_k, \epsilon_k)$ contained in the interior of $\mathbb{D}$ such that*

- *the points q_k are periodic points of S*
- $\limsup \frac{1}{n}\log|Q_n| \geq \log\alpha$, *where $Q_n = \{q_k\colon k \in \mathbb{N}, q_k \in P^0_n(S)\}$*
- *the set $Q = \cup Q_n$ is invariant*
- *ϵ_k is the same number $\epsilon(n)$ for all $q_k \in |P^0_n(S)|$.*

Proof. Choose N such that $|P^0_n(S)| \geq \alpha^n$ for all $n \geq N$ (here $|P^0_n(S)|$ denotes the set of points in S-orbits of cardinality n). This is possible because S is a mixing SFT and $\log\alpha < h(S)$.

To prove the lemma, it suffices to choose recursively, for $n = N, N+1, \dots$

((i)) a mixing SFT $S^{(n)} \subset S$ such that $h(S^{(n)}) > \log\alpha$,
$$|P^0_k(S^{(n)})| \geq \alpha^k \quad \text{ if } k > n \ ,$$
and $S^{(n)} \subset S^{(n-1)}$ if $n > N$;

((ii)) a set $Q_n \subset P^0_n(S)$ such that $|Q_n| \geq \alpha^n$; and

((iii)) $\epsilon(n) > 0$ such that the family of discs $D(q, \epsilon(k))$, with $q \in Q_k$ and $N \leq k \leq n$, are pairwise disjoint and disjoint from $S^{(n)}$, and are contained in the interior of $\mathbb{D}$.

We begin with $n = N$. Define $Q_N = P^0_N(S)$. Pick a mixing SFT S' such that $\log\alpha < h(S') < h(S)$ and S' has a fixed point (e.g. using [Kr1]). Pick $N_1 > N$ such that $|P^0_k(S)| \geq |P^0_k(S')| \geq \alpha^k$ for $k \geq N_1$. Using the Covering Lemma 2.1 of [B1], produce a mixing SFT S'' such that $h(S'') = h(S')$ and
$$\begin{aligned} |P^0_k(S'')| &= 0 \ , && k \leq N \ , \\ |P^0_k(S'')| &= |P^0_k(S)| \ , && N < k < N_1 \ , \\ |P^0_k(S')| \leq |P^0_k(S'')| &\leq |P^0_k(S)| \ , && k \geq N_1 \ . \end{aligned}$$

By Krieger's Embedding Theorem [Kr2], we may assume $S'' \subset S$ (and necessarily then, Q_N and S'' are disjoint). Now set $S^{(N)} = S''$ and choose $\epsilon(N)$ to satisfy (iii) for $n = N$.

The recursive step is much the same. Suppose $n+1 > N$ and we have carried out the choices above for $N, \dots, n$. Define $Q_{n+1} = P^0_{n+1}(S^{(n)})$ (so, $|Q_{n+1}| \geq \alpha^{n+1}$, and Q_{n+1} as a subset of $S^{(n)}$ is disjoint from the discs previously chosen for $k \leq n$). Pick a mixing SFT S^* such that $\log\alpha < h(S^*) < h(S^{(n)})$ and S^* has a fixed point. Pick $N_2 > n+1$ such that $|P^0_k(S^{(n)})| \geq |P^0_k(S^*)| \geq \alpha^k$ if $k \geq N_2$. Then pick a mixing SFT S^{**} such that $h(S^{**}) = h(S^*)$ and

$$\begin{aligned}
|P^0_k(S^{**})| \;&=\; 0 \;, && k \leq n+1 \;, \\
|P^0_k(S^{**})| \;&=\; |P^0_k(S^{(n)})| \;, && n+1 < k < N_2 \;, \\
|P^0_k(S^*)| \;\leq\; |P^0_k(S^{**})| \;&\leq\; |P^0_k(S^{(n)})| \;, && k \geq N_2 \;.
\end{aligned}$$

As before, by Krieger's Embedding Theorem we may assume $S^{**} \subset S^{(n)}$. Then define $S^{(n+1)} = S^{**}$. Then choose $\epsilon(n+1)$ to satisfy (iii).

This finishes the lemma. □

The map T will be the uniform limit of homeomorphisms $T_n \colon \mathbb{D} \to \mathbb{D}$. First we describe how T_1 is constructed as a modification of T_0. Fix a choice of discs $D_n = D(q_n, \epsilon_n)$ satisfying the statement of Lemma C.1. Define a map $\pi_0 \colon \mathbb{D} \to \mathbb{D}$ by setting

$$\begin{aligned}
\pi_0(z) \;&=\; z \;, && \text{if } z \notin \bigcup_{n=1}^{\infty} D_n \\
\pi_0(q_n + z) \;&=\; q_n \;, && \text{if } |z| \leq \frac{\epsilon_n}{2} \\
\pi_0(q_n + z) \;&=\; q_n + \frac{2}{\epsilon_n}\Big(|z| - \frac{\epsilon_n}{2}\Big) z \;, && \text{if } \frac{\epsilon_n}{2} < |z| \leq \epsilon_n \;.
\end{aligned}$$

The map π_0 maps the half-open annulus $D_n \setminus D(q_n, \epsilon_n/2)$ radially and homeomorphically to the punctured disc $D_n \setminus \{q_n\}$. The restriction of π_0 to $D \setminus \pi_0^{-1} Q$ is a homeomorphism onto its image. For $z \in \mathbb{D} \setminus \pi_0^{-1} Q$, define $T_1(z) = \pi_0^{-1} T_0 \pi_0(z)$. It remains to define T_1 on the discs $D(q_n, \epsilon_n/2)$.

Suppose $T_0(q_n) = q_k$. Because T_0 is differentiable and nonsingular at q_n, the map T_1 defined so far on $[\epsilon_n/2 < |z - q_n| \leq \epsilon_n]$ extends continuously to a map

$$\beta_n \colon \; [|z - q_n| = \frac{\epsilon_n}{2}] \;\mapsto\; [|z - q_k| = \frac{\epsilon_k}{2}] \;.$$

Because $\epsilon_k = \epsilon_n$ and $\det(DT_0) > 0$ at q_n, there is an orientation preserving homeomorphism $h_n \colon (\epsilon_n/2)S^1 \to (\epsilon_n/2)S^1$ such that

$$\beta_n \colon q_n + z \mapsto q_k + h_n(z) \;, \qquad \text{if } |z| = \frac{\epsilon_n}{2} \;.$$

Because an orientation preserving circle homeomorphism is isotopic to the identity, there is a homeomorphism

$$H_n \colon [\frac{\epsilon_n}{4} \leq |z| \leq \frac{\epsilon_n}{2}] \to [\frac{\epsilon_n}{4} \leq |z| \leq \frac{\epsilon_n}{2}]$$

such that $H_n(z) = h_n(z)$ for $|z| = \epsilon_n/2$ and $H_n(z) = z$ for $|z| = \epsilon_n/4$. For $\epsilon_n/4 \leq |z| \leq \epsilon_n/2$, we define $T_1(q_n + z) = q_k + H_n(z)$. Finally, for $|z| \leq \epsilon_n/4$, we define

$$T_1 \colon \; q_n + z \; \mapsto \; q_k + \frac{\epsilon_n}{4} T_0(4z/\epsilon_n) \; .$$

So, T_1 defines a map $D(q_n, \epsilon_n/4) \to D(q_k, \epsilon_k/4)$ which is a miniature copy of T_0. This completes the definition of T_1. The map $T_1 : \mathbb{D} \to \mathbb{D}$ is a homeomorphism. If m is the period of q_n, then the union of the discs $D((T_0)^i q_n, \frac{1}{4}\epsilon_n)$, $0 \le i < m$, is T_1-invariant; and the restriction of T_1^m to $D(q_n, \epsilon_n/4)$ is topologically conjugate to T_0^m. The map $\pi_0 \colon \mathbb{D} \to \mathbb{D}$ is a semiconjugacy from T_1 to T_0, and $\pi_0 = \mathrm{Id}$ on the complement of $\cup D_n$.

To define T_2 as a modification of T_1, we repeat the "blow and sew"process above inside each of those discs $D_k = D(q_k, \frac{1}{4}\epsilon_k)$. Each D_k contains disjoint discs

$$D_{k,n} \;=\; D(q_k + \frac{1}{4}\epsilon_k q_n, \frac{1}{4}\epsilon_k \epsilon_n), \qquad 1 \le n < \infty \; ,$$

centered at T_1-periodic points, and these points are blown up into discs into each of which we sew a miniature copy of T_0. Just as we defined $\pi_0 \colon \mathbb{D} \to \mathbb{D}$ in defining T_1 as a modification of T_0, we define $\pi_{1;k} \colon D_k \to D_k$ in defining T_2 as a modification of T_1, in particular the map $\pi_{1,k}$ is the identity map on the complement of $\cup_n D_{k,n}$. We define $\pi_1 : \mathbb{D} \to \mathbb{D}$ by setting $\pi_1 = \pi_{1;k}$ on D_k and $\pi_1 = \mathrm{Id}$ elsewhere. The map π_1 is a semiconjugacy from T_2 to T_1, and $\pi_1 = \mathrm{Id}$ on the complement of $\cup_{k,n} D_{k,n}$, which is a subset of $\cup_n D_n$. Note, $\pi_1 \pi_0 = \pi_0$.

Recursively, to define T_{n+1} as a modification of T_n, we "blow and sew"in each of a family of disjoint discs $D_{k(1),\dots,k(n)}$, blowing up certain T_n-periodic points to discs and sewing in miniature copies of T_0. Let $Q^{(n)}$ denote the set of T_n-periodic points blown up into discs in the definition of T_{n+1} as a modification of T_n. In the process of constructing T_{n+1} from T_n, we obtain a semiconjugacy π_n from T_{n+1} to T_n, where π_n maps each $D_{k(1),\dots,k(n)}$ onto itself and is the identity elsewhere. We have $\pi_n \pi_{n-1} = \pi_{n-1}$.

For $n \ge 1$, let G_n denote the union of the disks $D_{k(1),\dots,k(n)}$. The sets G_n are nested and $T_n = T_{n-1}$ outside G_n. So for all z and all $m > 0$, $\mathrm{dist}(T_n(z), T_{n+m}(z))$ cannot be more than the maximum diameter of a disc $D_{k(1),\dots,k(n+1)}$, which goes to zero uniformly with n. Therefore the T_n converge uniformly to a continuous map $T \colon \mathbb{D} \to \mathbb{D}$.

Next we will observe that T is topologically conjugate to the homeomorphism T' defined as the inverse limit of the maps T_n, with bonding maps $\pi_n \colon T_{n+1} \twoheadrightarrow T_n$. To show T is conjugate to T', it suffices to produce semiconjugacies $\varphi_n \colon T \twoheadrightarrow T_n$ such that $\varphi_n = \pi_n \varphi_{n+1}$ and such that $\prod_{n=1}^{\infty} \varphi_n$ is injective on $\mathbb{D}$. Simply define $\varphi_n = \pi_n$, then $\varphi_n = \pi_n \varphi_{n+1}$. The images under π_n of G_{n+1} and its complement are disjoint, and the map π_n is one-one on the complement of G_{n+1}. Consequently φ is injective on the complement of $\cap_n G_n$. However, φ is injective on $\cap_n G_n$ as well, because each nested sequence of discs

$$D_{k(1)} \supset D_{k(1),k(2)} \supset D_{k(1),k(2),k(3)} \supset \cdots$$

shrinks to a single point, and for each n, each level-n disc $D_{k(1),\dots,k(n)}$ is π_{n-1}-invariant.

With this inverse limit presentation for T, it is a straightforward matter to see that the argument for infinite residual entropy in Section 3 adapts to show that T has infinite residual entropy.

For each n, under the semiconjugacy $\pi_n \colon T_{n+1} \twoheadrightarrow T_n$ there are some periodic orbits of T_n whose inverse images are subsystems of entropy $h(T_0)$, and the restriction of π_n to the complement of the union of these inverse images is bijective. It follows by induction on n that there is no ergodic T_{n+1}-invariant Borel probability μ with $h_\mu(T_{n+1}) > h(T_0)$; so, by the variational principle and ergodic decomposition, we have $h(T_{n+1}) \leq h(T_0)$. Thus each $h(T_n) = h(T_0)$ and therefore the inverse limit system T satisfies $h(T) = h(T_0) < \infty$. This completes the example.

Remarks C.2. We only needed DT_0 existing and positive at the disc centers q_n. Also, if in the construction we begin with T_0 a C^2 map, then with some modifications to the "blow and sew"operation (and some unpleasantly technical additional arguments), we can arrange each T_n to be C^1, and T to be differentiable on the complement of $\cap_n G_n$. But we see no way to modify the construction to achieve differentiability of T on $\cap_n G_n$: at a point in this Cantor set, the local picture need not approach a linear map as the scale shrinks; for example on a sequence of scales the map could be locally approximated by different linear maps.

The effort to construct the example above naturally raises a technical question:

Question C.3. Suppose T is a homeomorphism of a Cantor set C contained in the interior of a disc D, and T has finite entropy. Does T extend to a finite entropy homeomorphism of D?

It is a well known consequence of the Schoenflies Theorem that any homeomorphism of C above extends to a homeomorphism of D which is the identity on the boundary. If the entropy of the extension can be controlled, then one has a general method for constructing finite entropy, infinite residual entropy homeomorphisms of a surface.

Appendix D. Intermediate residual entropy

We will prove the following result.

Theorem D.1. *Suppose $0 < a < \infty$ and $0 \leq b \leq \infty$. Then there is a zero dimensional system T with $h(T) = a$ and $\rho(T) = b$.*

Remark D.2. These results are also in [Do2]; the constructions are very different. Note, if $h(T) = 0$ then $\rho(T) = 0$ (Cor.7.6 or [Do2]), and if $h(T) = \infty$ then $\rho(T) = \infty$. So Theorem D.1 covers all the possible cases.

The heart of the proof of Theorem D.1 is the explicit construction proving Proposition D.5 below. The rest of the proof rests on the following two lemmas. We use the notation that $T_{(n)}$ denotes the discrete tower of height n built over the transformation T. Explicitly, if X is the domain of T, then $X \times \{1, 2, \ldots, n\}$ is the domain of $T_{(n)}$, which maps by the rule

$$\begin{aligned} T_{(n)} \colon (x, i) &\mapsto (x, i+1) && \text{if } i \neq n \\ &\mapsto (Tx, 1) && \text{if } i = n\ . \end{aligned}$$

It is well known and easy to see that $h(T_{(n)}) = (1/n)h(T)$.

Lemma D.3. *Suppose T is a selfhomeomorphism of a compact metric space. Then*

(1) $\rho(T^n) = n\rho(T)$.
(2) $\rho(T_{(n)}) = (1/n)\rho(T)$.

Proof. (1) If S is a subshift cover of T, then S^n is a subshift cover of T^n, and $h(S^n) - h(T^n) = n[h(S) - h(T)]$. Thus $\rho(T^n) \leq \inf_S n[h(S) - h(T)] = n\rho(T)$.

Conversely, if S is a subshift and $\varphi\colon S \twoheadrightarrow T^n$, then $S_{(n)}$ is a subshift cover of T by the map $(x,i) \mapsto T^{i-1}\varphi x$ if $1 \leq i \leq n$. Therefore $\rho(T) \leq h(S_{(n)}) - h(T) = (1/n)[h(S) - h(T^n)]$ and we obtain $\rho(T) \leq \inf_S (1/n)[h(S) - h(T^n)] = (1/n)\rho(T^n)$.

(2) The system $(T_{(n)})^n$ is the disjoint union of n copies of T, so $\rho((T_{(n)})^n) = \rho(T)$. Then it follows from (1) that $\rho(T_{(n)}) = (1/n)\rho(T)$.

□

Given a sequence T_n of systems, we let $(T_n)_\infty$ denote the one point compactification system (in which the added point is a fixed point). In the lemma below we regard the systems in such a sequence T_n as being disjoint.

Lemma D.4. *Let $T = (T_n)_\infty$. Define*

- $\alpha = \inf\{h(S)\colon S \text{ is a subshift and } S \twoheadrightarrow T\}$,
- $\alpha_n = \inf\{h(S)\colon S \text{ is a subshift and } S \twoheadrightarrow T_n\}$.

Then $\alpha = \sup \alpha_n$.

Proof. Clearly $\alpha \geq \sup \alpha_n$. So suppose $\sup \alpha_n < \infty$ and $\epsilon > 0$. Pick a mixing SFT U such that $0 < h(U) - \sup \alpha_n < \epsilon$. For each n, pick a subshift cover S_n of T_n, with $h(S_n) < h(U)$, and pick a mixing SFT R_n with a fixed point such that $S_n \subset R_n$ and $h(R_n) < h(U)$. Let Z be the identity map on the space which is the convergent sequence $\{0\} \cup \{1/n \colon n \in \mathbb{N}\}$. Using Cor. 7.6, pick some zero entropy subshift W such that $U \times W \twoheadrightarrow U \times Z$. Recall that $U \twoheadrightarrow R_n$ for each n by [B1], so $U \times Z \twoheadrightarrow (R_n)_\infty$. Putting all this together, we see

$$U \times Z \ \twoheadrightarrow\ (R_n)_\infty \ \supset\ (S_n)_\infty \ \twoheadrightarrow\ (T_n)_\infty \ =\ T\ .$$

Taking the inverse image of T inside $U \times Z$, we get a subshift V such that $h(V) \leq h(U \times Z) = h(U) < \sup \alpha_n + \epsilon$. □

Proof of Theorem D.1. Suppose $0 < a, b < \infty$. Given $m \in \mathbb{N}$, pick positive integers k, n such that

$$(a+b) - \frac{1}{m} \ <\ \frac{k}{n}(\log 2) \ <\ a+b\ , \qquad \text{and}$$
$$\frac{1}{n}(\log 2) \ <\ a\ .$$

Using Proposition D.5 below, given $\epsilon > 0$ we may pick T such that

$$\log 2 \ \leq\ h(T) \ \leq\ \log 2 + \epsilon\ , \qquad \text{and}$$
$$(k\log 2) - \epsilon \ \leq\ \rho(T) + \log 2 \ \leq\ k\log 2\ .$$

Then

$$\frac{1}{n}(\log 2) \ \leq\ h(T_{(n)}) \ \leq\ \frac{1}{n}[(\log 2) + \epsilon]\ , \qquad \text{and}$$
$$\frac{k}{n}(\log 2) - \frac{\epsilon}{n} \ \leq\ \rho(T_{(n)}) + h(T_{(n)}) \ \leq\ \frac{k}{n}(\log 2) + \frac{\epsilon}{n}\ ,$$

so for small enough ϵ we have

$$h(T_{(n)}) \ <\ a\ , \quad \text{and}$$
$$(a+b) - \frac{1}{m} \ <\ \rho(T_{(n)}) + h(T_{(n)}) \ <\ (a+b)\ .$$

Choose S_m to be a system $T_{(n)}$ satisfying the last two lines. Let S_0 be a subshift such that $h(S_0) = a$. Regard $S_0, S_1, S_2, \dots$ as pairwise disjoint. Let S be the one point compactification of the systems $S_0, S_1, S_2, \dots$. Then $h(S) = a$, and $\rho(S) = b$ as a consequence of Lemma D.4. This finishes the case $0 < b < \infty$.

If $b = \infty$, then again we may take S to be the one point compactification of systems $S_0, S_1, S_2, \dots$, with $h(S_0) = a$ and $h(S_n) < a$ for $n > 0$, but in this case we require $\rho(S_n) \to \infty$. For the case $b = 0$, let $S = S_0$.

□

The rest of the section is devoted to the following result, which is the heart of the matter.

Proposition D.5. *Suppose $\varepsilon > 0$ and $r \in \mathbb{N}$. Then there is an inverse limit T of mixing sofic shifts T_n and surjective bonding maps $\pi_n\colon T_{n+1} \to T_n$ such that*

- *T is a quotient of the full shift on 2^{r+1} symbols (by construction).*
- *$\log 2 \le h(T) \le \log 2 + \varepsilon$.*
- *$h(R) \ge \log(2^{r+1})$ for every subshift R such that T is a quotient of R.*

Thus we will have $r\log 2 - \varepsilon \le \rho(T) \le r\log 2$.

We prepare for the definition of T. Fix $r \in \mathbb{N}$. Define an ordering $\prec$ on the set $\mathbb{N} \cup \mathbb{N}^2 \cup \dots \cup \mathbb{N}^r$ as follows. Let $<_{lex}$ denote lexicographic ordering on $\mathbb{N} \cup \mathbb{N}^2 \cup \dots \cup \mathbb{N}^r$. For $(n_1, \dots, n_i) \in \mathbb{N} \cup \mathbb{N}^2 \cup \dots \cup \mathbb{N}^r$, let $||(n_1, \dots, n_i)|| := \sum_{1 \le k \le i} n_k$. Let $n, m \in \mathbb{N} \cup \mathbb{N}^2 \cup \dots \cup \mathbb{N}^r$. Then define $n \prec m$ if $(||n|| < ||m||)$ or $(||n|| = ||m||$ and $n <_{lex} m)$. We will use just two properties of this order: when $r = 1$, this is the usual order on $\mathbb{N}$; and for $r \ge 2$ it holds that $(n_1, \dots, n_{k-1}) \prec (n_1, \dots, n_k)$, if $2 \le k \le r$. Let $i\colon \mathbb{N} \to \mathbb{N} \cup \mathbb{N}^2 \cup \dots \cup \mathbb{N}^r$ denote the bijection such that $j < k \Rightarrow i(j) \prec i(k)$.

Let $S = \{0,1\}^{\mathbb{Z}}$ denote the full 2-shift. Fix an enumeration of the finite S-orbits, say $O_1, O_2, \dots$. Fix $\varepsilon > 0$. Fix a sequence $(N_n)_{n \ge 1}$ of natural numbers.

We define $f_0\colon S^{r+1} \to S$ to be the projection onto the first coordinate, $f_0(x^1, \dots, x^{r+1}) := x^1$. For $n \ge 1$ we define $f_n\colon S^{r+1} \to S$ as follows. We have that $i(n) = (k_1, \dots, k_s)$ for some $1 \le s \le r$. Let $f_n(x^1, \dots, x^{r+1})_0 := x_0^{s+1}$ if $O_{k_i} \cap [x^i_{-N_n}, \dots, x^i_{N_n}]_{N_n} \ne \emptyset$ for all $1 \le i \le s$, otherwise $f_n(x^1, \dots, x^{r+1})_0 := 0$. Then each f_n is a shift invariant, continuous onto map. For $n \in \mathbb{N}$, let $T_n := (f_0, \dots, f_{n-1})(S^{r+1})$. Thus $T_n \subset S^n$. Let $\pi_n\colon T_{n+1} \to T_n$ denote the projection onto the first n coordinates. This defines an inverse limit T which is a quotient of S^{r+1}. Note that $T_1 = f_0(S^{r+1}) = S$, thus $h(T) \ge \log 2$.

We shall now show that there is a choice of the sequence N_n such that $h(T) \le \log 2 + \varepsilon$. For each orbit O_n fix a point $p_n \in O_n$. For $n \in \mathbb{N}$ let $i(n) = (k_1, \dots, k_s)$. Define a map $h_n\colon S^{r+1} \to (\prod_{1 \le i \le s}\{1, \dots, |O_{k_i}|\} \cup \{0\})^{\mathbb{Z}}$ as follows. Let $h_n(x^1, \dots, x^{r+1})_0 := (m(1), \dots, m(s))$ if $x^i[-N_n, N_n] = S^{m(i)}p_{k_i}[-N_n, N_n]$ for all $1 \le i \le s$ and let $h_n(x^1, \dots, x^{r+1})_0 := 0$ else. Note, $h_n(x^1, \dots, x^{r+1})_0 \ne 0$ implies $f_n(x^1, \dots, x^{r+1})_0 = x_0^{s+1}$.

Note also, $h_n(x^1, \dots, x^{r+1})_0 = (m(1), \dots, m(s))$ implies $h_n(x^1, \dots, x^{r+1})_1 = 0$ or $h_n(x^1, \dots, x^{r+1})_1 = (m(1) + 1 \mod |O_{k_1}|, \dots, m(s) + 1 \mod |O_{k_s}|)$. Also, if $N_n > 2\max_{1 \le i \le s} |O_{k_i}|$ and $h_n(x^1, \dots, x^{r+1})_0 \ne 0$ and $h_n(x^1, \dots, x^{r+1})_1 = 0$, then $h_n(x^1, \dots, x^{r+1})_m = 0$ for $1 \le m \le N_n/2$. Thus if N_n is large enough we get that $h_n(S^{r+1})$ has entropy $< 2^{-n} \cdot \varepsilon$.

Now fix a sequence N_n such that for each n it holds that

$$h(h_n(S^{r+1})) < 2^{-n} \cdot \varepsilon.$$

We have $T_1 = S$ and thus $h(T_1) = \log 2$. We now estimate the entropy of T_{n+1} for $n \geq 1$. For every $1 \leq m \leq n$ let $1 \leq s(m) \leq r$ denote the integer such that $i(m) \in \mathbb{N}^{s(m)}$. Fix $j \in \mathbb{N}$. Now we consider the map $(h_1, \ldots, h_n)\colon S^{r+1} \to h_1(S^{r+1}) \times \cdots \times h_n(S^{r+1})$. Fix $(a^1, \ldots, a^n) \in B_j((h_1, \ldots, h_n)(S^{r+1}))$. Let $A = A(a^1, \ldots, a^n) := \{(x^1, \ldots, x^{r+1}) | h_m(x^1, \ldots, x^{r+1})[0, j) = a^m \text{ for } 1 \leq m \leq n\}$. Let $0 \leq i < j$ and let $(x^1, \ldots, x^{r+1}) \in A$.

If $a_i^m = 0$ for all $1 \leq m \leq n$ then $f_m(x^1, \ldots, x^{r+1})_i = 0$ for all $1 \leq m \leq n$ and $f_0(x^1, \ldots, x^{r+1})_i \in \{0, 1\}$. Thus $\#\{(f_0, \ldots, f_n)(x^1, \ldots, x^{r+1})_i | (x^1, \ldots, x^{r+1}) \in A(a^1, \ldots, a^n)\} \leq 2$.

Now assume there is $1 \leq m \leq n$ with $a_i^m \neq 0$. Consider all $1 \leq k \leq n$ with $a_i^k \neq 0$. Choose k among those with maximal $s(k)$. Then the definition of h_k implies that $x_i^1, \ldots, x_i^{s(k)}$ are determined by a_i^k. Thus $f_0(x^1, \ldots, x^{r+1})_i = x_i^1$ is determined by a_i^k and for $1 \leq m \leq n$ we get that in case that $(s(m) < s(k)$ or $a_i^m = 0)$ that $f_m(x^1, \ldots, x^{r+1})_i$ is uniquely determined by a_i^m and a_i^k and in case that $(s(m) = s(k)$ and $a_i^m \neq 0)$ we get that $f_m(x^1, \ldots, x^{r+1})_i = x_i^{s(k)+1}$. Again $\#\{(f_0, \ldots, f_n)(x^1, \ldots, x^{r+1})_i | (x^1, \ldots, x^{r+1}) \in A\} \leq 2$.

Thus for each $(a^1, \ldots, a^n) \in B_j((h_1, \ldots, h_n)(S^{r+1}))$ it holds that

$$\#\{(f_0, \ldots, f_n)(x^1, \ldots, x^{r+1})[0, j) | (x^1, \ldots, x^{r+1}) \in A(a^1, \ldots, a^n)\} \leq 2^j.$$

Thus

$$\begin{aligned} B_j(T_{n+1}) &= \#\{(f_0, \ldots, f_n)(x^1, \ldots, x^{r+1})[0, j) | (x^1, \ldots, x^{r+1}) \in S^{r+1}\} \\ &\leq 2^j \cdot \# B_j((h_1, \ldots, h_n)(S^{r+1})). \end{aligned}$$

Thus $h(T_{n+1}) \leq \log 2 + h((h_1, \ldots, h_n)(S^{r+1})) < \log 2 + \varepsilon$.

We now estimate the residual entropy of T. For that we use the following two general lemmata.

Lemma D.6. *Let R be a subshift and $g\colon R \to S$ be a quotient map and $r \in \mathbb{N}$ such that*

$$h(g^{-1}(O)) \geq \log(2^r) \quad \textit{for each finite orbit } O \textit{ of } S.$$

Then $h(R) \geq \log(2^{r+1})$.

Proof of Lemma D.6. Let m be a coding length for g, that is $g(x)_i$ is determined by $x[-m+i, i+m]$ for all $x \in R$ and all $i \in \mathbb{Z}$. Let $b \in B_n(S)$ and let $O(b)$ denote the orbit of b^∞. Let $A(b) := \{x[-m, n+m) | gx[0, n) = y[0, n) \text{ for some } y \in O(b)\}$. Since $h(g^{-1}(O(b))) \geq \log 2^r$ we get $\# A(b) \geq 2^{r(2m+n)}$. Since $O(b) \neq O(b')$ implies that $A(b)$ is disjoint from $A(b')$ and the map $b \to O(b)$ is at most $n{-}to{-}1$ we get $\# B_{2m+n}(R) \geq \sum_{b \in B_n(S)} n^{-1} \cdot \# A(b) \geq n^{-1} \cdot 2^{r(2m+n)} \cdot 2^n = n^{-1} \cdot 2^{n(r+1)+2rm}$. This holds for all n and thus $h(R) \geq \log(2^{r+1})$. □

Lemma D.7. *Let R be a subshift and $g : R \to S$ be a quotient map. Let $s \in \mathbb{N}$. Assume there is a family of quotient maps $g_k\colon R \to S$, $k \in \mathbb{N} \cup \mathbb{N}^2 \cup \cdots \cup \mathbb{N}^s$ such that for each $(j_1, \ldots, j_{s+1}) \in \mathbb{N}^{s+1}$ it holds that*

$$h(g^{-1} O_{j_1} \cap g_{j_1}^{-1} O_{j_2} \cap g_{(j_1, j_2)}^{-1} O_{j_3} \cap \cdots \cap g_{(j_1, \ldots, j_s)}^{-1} O_{j_{s+1}}) \geq \log 2.$$

Then $h(R) \geq \log(2^{s+2})$.

Proof of Lemma D.7. Let

$$R(j_1,\dots,j_s) := g^{-1}O_{j_1} \cap g_{j_1}^{-1}O_{j_2} \cap g_{(j_1,j_2)}^{-1}O_{j_2} \cap \dots \cap g_{(j_1,\dots,j_{s-1})}^{-1}O_{j_s}.$$

Then $R(j_1,\dots,j_s)$ is a subshift and $\alpha_1 := g_{(j_1,\dots,j_s)}|_{R(j_1,\dots,j_s)}$ is a quotient map onto S with $h(\alpha_1^{-1}(O)) \geq \log 2$ for all finite orbits O of S. Thus Lemma D.6 with $r=1$ implies $h(R(j_1,\dots,j_s)) \geq \log 4$. Thus $R(j_1,\dots,j_{s-1})$ is a subshift and $\alpha_2 := g_{(j_1,\dots,j_{s-1})}|_{R(j_1,\dots,j_{s-1})}$ is a quotient map onto S with $h(\alpha_2^{-1}(O)) \geq \log 4$ for all finite orbits O of S. Thus Lemma D.6 with $r=2$ implies $h(R(j_1,\dots,j_{s-1})) \geq \log 8$. Inductively one obtains in this way $h(R(j_1)) \geq \log(2^{s+1})$ for all j_1 and thus a final application of Lemma D.6 with $r=s+1$ shows $h(R) \geq \log(2^{s+2})$. $\square$

Now let R be a subshift and $\varphi\colon R \to T$ be a quotient map. Let $\varphi_n\colon R \to T_n$ denote the map φ followed by the projection from T onto T_n. For $n \geq 1$ let $pr_n\colon T_n \to S$ denote the projection onto the last coordinate. Let $\pi_{i,j}\colon T_{j+1} \to T_i$ denote the composition of the maps $\pi_j,\dots,\pi_i$.

If $r=1$ then by Lemma D.6 it suffices to show that $h(\varphi_1^{-1}O) \geq \log 2$ for every finite orbit O of S. Since $r=1$ we have that the map $i\colon \mathbb{N} \to \mathbb{N}$ is the identity. Let $n \geq 1$. Then $pr_{n+1}((f_0,\dots,f_n)(O_n \times S)) = f_n(O_n \times S) = S$ by definition of f_n. Thus $h(\varphi_1^{-1}O_n) = h((\pi_{1,n}\varphi_{n+1})^{-1}O_n) \geq h((\pi_{1,n})^{-1}O_n) = h((f_0,\dots,f_n)(O_n \times S)) \geq \log 2$.

Now consider the case that $r > 1$. We shall apply Lemma D.7. We define $g := \varphi_1\colon R \to S$. For $k = (j_1,\dots,j_s) \in \mathbb{N} \cup \mathbb{N}^2 \cup \dots \cup \mathbb{N}^{r-1}$ let $n = n(k)$ such that $i(n-1) = (j_1,\dots,j_s)$. Then define $g_k := pr_{n(k)}\varphi_{n(k)}\colon R \to S$. Now let $(j_1,\dots,j_r) \in \mathbb{N}^r$. Let $R(j_1,\dots,j_r) := g^{-1}O_{j_1} \cap (g_{j_1})^{-1}O_{j_2} \cap (g_{(j_1,j_2)})^{-1}O_{j_3} \cap \dots \cap (g_{(j_1,\dots,j_{r-1})})^{-1}O_{j_r}$. Choose n_m such that $i(n_m - 1) = (j_1,\dots,j_m)$ for $1 \leq m \leq r$ and let $n_0 = 0$. (Here $n_m \leq n_r$ by definition of $\prec$ and the bijection i.) Let $P := O_{j_1} \times \dots \times O_{j_r} \times S$. Let $x \in R$ such that $\varphi_{n_r}(x) \in (f_0,\dots,f_{n_r-1})(P) \subset T_{n_r}$. Then $g(x) = \varphi_1(x) = \pi_{1,n_r}\varphi_{n_r}(x) \in f_0(O_{j_1}) = O_{j_1}$. Thus $x \in g^{-1}O_{j_1}$. Now let $1 \leq m < r$. We show $x \in (g_{(j_1,\dots,j_m)})^{-1}O_{j_{m+1}}$. We have $g_{(j_1,\dots,j_m)}(x) = pr_{n_m}\varphi_{n_m}(x)$. Since $\varphi_{n_m}(x) = \pi_{n_m,n_r}\varphi_{n_r}(x) \in (f_0,\dots,f_{n_m-1})(P)$ we get $g_{(j_1,\dots,j_m)}(x) \in f_{n_m-1}(P) \subset O_{j_{m+1}}$. Thus $(f_0,\dots,f_{n_r-1})(P) \subset \varphi_{n_r}(R(j_1,\dots,j_r))$ and since $pr_{n_r}((f_0,\dots,f_{n_r-1})(P)) = f_{n_r-1}(P) = S$ by definition of f_{n_r-1}, it follows that $h(R(j_1,\dots,j_r)) \geq h((f_0,\dots,f_{n_r-1})(P)) \geq \log 2$. Thus the assumptions of Lemma D.7 are satisfied with $s = r-1$ and thus we have that $h(R) \geq \log(2^{r+1})$. This completes the proof of Proposition D.5, and therefore completes the proof of Theorem D.1.

References

[A] R. D. Anderson, *On raising flows and mappings*, Bull. AMS **69** (1963), 259-264.

[Bow1] R. Bowen, *Entropy-expansive maps*, Trans. AMS **164** (1972), 323-331.

[Bow2] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Springer Lec. Notes in Math. 470, Springer-Verlag (1975).

[B1] M. Boyle, *Lower entropy factors of sofic systems*, Erg. Th. & Dyn. Syst. **4** (1984), 541-557.

[B2] M. Boyle, *Quotients of subshifts*, Adler conference lecture (1991) ([W2], p. xi).

[B3] M. Boyle, *Factoring factor maps*, J. London Math. Soc (20) **57** (1998), 491-502.

[BH] M. Boyle and D. Handelman, *Orbit equivalence, flow equivalence and ordered cohomology*, Israel J. Math. **95** (1996), 169-210.

[BT] M. Boyle and S. Tuncel, *Infinite-to-one codes and Markov measures*, Trans. AMS **285**, No. 2 (1984), 657-684.

[BrKa] M. Brin and A. Katok, *On local entropy*, in *Geometric Dynamics*, Springer Lec. Notes in Math. **1007** (1983), Springer-Verlag, 30-38.

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

[DGS] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces*, Springer Lec. Notes in Math. **527**, Springer-Verlag (1976).

[De] M. Denker, *Measures with maximal entropy*, pp. 70-112 in Théorie ergodique (Actes Journées Ergodiques, Rennes, 1973/74), Springer Lec. Notes in Math. **532**, Springer-Verlag (1976).

[Do1] T. Downarowicz, Personal communications, 1998-2000.

[Do2] T. Downarowicz, *Existence of a symbolic extension of a totally disconnected dynamical system*, preliminary manuscript (1999).

[DS] T. Downarowicz and J. Serafin, *Fiber entropy and conditional variational principles in compact non-metrizable spaces*, preprint (2000).

[F] D. Fried, *Finitely presented dynamical systems*, Erg. Th. & Dyn. Syst. **7** (1987), 489-507.

[GW] E. Glasner and B. Weiss, *Quasi-factors of zero-entropy systems*, J. AMS **8** (1995), 665-686.

[HW] W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton University Press (1948).

[Ki] B. Kitchens, *Symbolic dynamics. One-sided, two-sided and countable state Markov shifts*, Springer-Verlag (1998).

[Kr1] W. Krieger, *On the periodic points of topological Markov chains*, Math. Z. **169** (1979), 99–104.

[Kr2] W. Krieger, *On subsystems of topological Markov chains*, Erg. Th. & Dyn. Syst. **2**, 1982, 195-202.

[Ku1] J. Kulesza, *Zero-dimensional covers of finite-dimensional dynamical systems*, Erg. Th. & Dyn. Syst. **15** (1995), 939–950.

[Ku2] J. Kulesza, *Personal communication*, (1994).

[Le] F. Ledrappier, *A variational principle for the topological conditional entropy*, Springer Lec. Notes in Math. **729** (1979), Springer-Verlag, 78-88.

[LeW] F. Ledrappier and P. Walters, *A relativised variational principle for continuous transformations*, J. London Math. Soc. **16** (1977), 568-576.

[Li] E. Lindenstrauss, *Lowering topological entropy*, J. Anal. Math. **67** (1995), 231-267.

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

[Ma1] R. Mañé, *Expansive homeomorphisms and topological dimension*, Trans. Amer. Math. Soc. **252** (1979), 313-319.

[Ma2] R. Mañé, *Ergodic theory and differentiable dynamics*, Springer-Verlag (1987).

[Mi1] M. Misiurewicz, *Diffeomorphism without any measure with maximal entropy*, Bull. Acad. Polon. Sci. , Sér. sci. math. , astr. et phys. **21** (1973), 903-910.

[Mi2] M. Misiurewicz, *Topological conditional entropy* , Studia Math. **55** (1976), 175-200

[Ne] S. Newhouse, *Continuity properties of entropy*, Annals of Math. **129** (1989), 215-235.

[Re] W. L. Reddy, *Lifting expansive homeomorphisms to symbolic flows*, Math. Systems Theory **2**, (1968), 91-92.

[Ru] D. J. Rudolph, *Fundamentals of measurable dynamics. Ergodic theory on Lebesgue spaces*, Oxford University Press (1990).

[S] S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817.

[T1] K. Thomsen, *Covering dimension and topological entropy in dynamical systems*, preprint, Aarhus (1994).

[T2] K. Thomsen, *The defect of factor maps*, Ergod. Th. & Dynam. Sys. **17** (1997), 1233-1256.

[T3] K. Thomsen, *The variational principle for the defect of factor maps*, preprint (1998).

[W1] P. Walters, *Relative pressure, relative equilibrium states, compensation functions and many-to-one codes between subshifts*, Trans. Amer. Math. Soc. **296** (1986), 1-31.

[W2] P. Walters, editor, *Symbolic dynamics and its applications*, Contemporary Mathematics **135**, American Math. Soc. (1992).

[Y] Y. Yomdin, *Volume growth and entropy*, Israel J. Math **57** (1987), 301-318.

Department of Mathematics, University of Maryland, College Park, MD 20742-4015, U.S.A.

E-mail address: `mmb@math.umd.edu`

Institut für Mathematische Stochastik, Universität Göttingen, Lotzestr. 13, 37083 Göttingen, Germany

E-mail address: `fiebig@math.uni-goettingen.de`

Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany

E-mail address: `fiebig@math.uni-heidelberg.de`