Ergod. Th. & Dynam. Sys. (1991), **11**, 535–546
Printed in Great Britain

Can one always lower topological entropy?

M. SHUB[1]† AND B. WEISS[2]

[1] *Department of Mathematical Sciences, IBM Research Division, T. J. Watson Research Center, PO Box 218, Yorktown Heights, NY 10598, USA*
[2] *Department of Mathematics, Hebrew University, Jerusalem, Israel*

(*Received* 30 *June* 1989)

Abstract. We consider the problem of when does a positive entropy topological system have a continuous factor with strictly smaller entropy. In many cases it is shown that such small entropy factors exist. On the other hand, classes of examples are given where differentiable factors must preserve some of the original entropy.

0. *Introduction*

For measure theoretic entropy, it is well known and quite easy to see that a positive entropy transformation always has factors of smaller entropy. Indeed the factor generated by a two-set partition with one of the sets having very small measure will always have small entropy. It is our purpose here to treat the analogous question for topological entropy. A topological system (X, T) will be a continuous mapping T of a metric space X which we shall usually suppose to be compact. A second system (Y, S) is a *factor* of (X, T) if there is a *continuous* surjective map $\pi: X \to Y$ such that $S\pi = \pi T$. We will exclude the trivial factor, where Y reduces to one point, and use *factor* to mean non-trivial factor.

BASIC PROBLEMS. *If* $h_{\text{top}}(T) > 0$ *does* (X, T) *have a factor with strictly smaller entropy*?

While we do not have a complete answer we have developed a technique for constructing lower entropy factors which works in some situations such as:

(a) systems (X, T) with only a countable number of ergodic invariant measures – in particular, uniquely ergodic systems.
(b) smooth transformations of compact manifolds without too many periodic points – in particular Anosov diffeomorphisms.

For smooth mappings it makes sense to ask about the existence of smooth factors with small entropy. We shall show that in some cases such as transitive Anosov diffeomorphisms there is a positive lower bound to the entropy that can be achieved by a differentiable factor. To explain this a little more precisely let us consider how one can construct factors.

Let $\varphi: X \to [0, 1]$ be a continuous function. Then $\{\varphi T^n\}$ defines a factor which is a closed shift invariant subset of $[0, 1]^{\mathbb{N}}$. Indeed define $\Phi: X \to [0, 1]^{\mathbb{N}}$ by

$$(\Phi(x))_n = \varphi(T^n x), \quad n \geq 0.$$

† Partially supported by an NSF grant.

If $\sigma:[0,1]^{\mathbb{N}}\to[0,1]^{\mathbb{N}}$ denotes the shift,

$$(\sigma\omega)_n=\omega_{n+1},\quad n\geq 0$$

then $\sigma\Phi=\Phi T$, and $\Phi(X)$ is clearly σ-invariant. If there is any factor that has lower entropy there must be a factor of this type with lower entropy since a non-trivial factor has non-constant continuous functions which can be pulled back to X. By a *differentiable factor* we mean a factor of this type that is defined by a differentiable function φ.

A generically chosen differentiable φ will have the property that already a finite number of values of $\varphi(T^n x)$ serve to uniquely determine x, so that Φ is invertible and cannot lower entropy. Thus a generically chosen 'measurement' φ made on the system (X, T) will give the full system. Our construction shows that for continuous functions this need not be always the case, whereas for differentiable functions, in some cases, there is a positive lower bound on the entropy that will be observed.

Even in the situations mentioned above, where we do know how to lower the entropy and make it arbitrarily small, we do not know what the range of values of $h_{\text{top}}(S)$ is as S ranges over the factor (X, T). Presumably it always contains the interval $(0, h_{\text{top}}(x))$ but we cannot even prove that it is dense there. Note that for measure-theoretic reasons there may be no factor with zero topological entropy. Note too that there are so-called 'prime' systems that have no factors [2] and so a general affirmative answer would imply in particular that there are no positive entropy prime systems. Discussions with S. Glasner about this possibility gave the original impetus to this investigation.

Finally, let us conclude this introduction with a special case of the basic problem:

Special Case. $X=[0,1]^{\mathbb{Z}}$, with the product topology, $\sigma: X\to X$ the shift, i.e. $(\sigma x)_n = x_{n+1}$, $n\in\mathbb{Z}$. Does (X,σ) have a finite entropy factor?

While we do not have an explicit reduction of the general problem to this case it seems clear that a positive answer in this case will shed a great deal of light on the general question.

After the first version of this paper was completed, Jonathan Ashley showed us how to show that any non-trivial factor of the one-sided shift on $\Omega=[0,1]^{\mathbb{N}}$ has infinite entropy. Further discussions with E. Glasner led to a proof of the fact that if (X, T) is a non-trivial factor of the shift on Ω then that shift is a factor of (X, T^N) for some power N. This proof relies very strongly on the non-invertible character of the one-sided shift and appears to shed little light on the invertible case of this question.

In § 1 we develop the notion of 'small sets' that will play a crucial role in the rest of the study. § 2 contains the basic construction, and § 3 details the cases where it can be made to work. § 4 contains the lower bound on the entropy of a differentiable factor and in § 5 with J. Ashley's permission we give an answer to the non-invertible version of the special case.

1. *Small sets*

If the space X is totally disconnected then there is an easy construction of low entropy factors that exactly mimics what is done in the case of the measure-theoretic entropy. Since $h_{top}(T)>0$, there is some non-periodic point, say x_0, and thus given N, there is some closed and open set U that contains x_0, such that for all $1\le n\le N$, $T^{-n}U\cap U=\phi$. Let $\varphi=1_U$, the indicator function of the set U. Then since $\partial U=\phi$, 1_U is continuous! Now Φ maps X into $\{0,1\}^{\mathbb{N}}$, and $\Phi(X)$ has the property that occurrences of ones are separated by at least N zeros. Thus the topological entropy of the shift restricted to $\Phi(X)$ is at most

$$\frac{1}{N}\log N+\left(1-\frac{1}{N}\right)\log(N/(N-1)) \qquad (*)$$

since the number of such n-blocks of zeros and ones is at most $\binom{n}{n/N}$, and Stirling's formula gives $(*)$.

The problem is that when X is connected then φ must assume an interval of values and then one cannot easily control the topological, or combinatorial entropy. Our solutions depends upon cutting up the space, disconnecting it, by means of small sets. In effect we shall construct φ by means similar to those used when proving Urysohn's lemma and arrange to have the change of φ take place on a set which cannot contribute to the topological entropy. It turns out that the following is the relevant notion:

Definition. A set $E\subset X$ is called T-small (or simply small if T is understood) if uniformly in $x\in X$

$$\lim_{N\to\infty}\frac{1}{N}\sum_1^N 1_E(T^nx)=0,$$

explicitly, given $\varepsilon>0$ for some N, and all $x\in X$

$$\frac{1}{N}\sum_1^N 1_E(T^nx)<\varepsilon. \qquad (**)$$

If X is compact and E is a closed small set then given $\varepsilon>0$ there is an open neighbourhood $V\supset E$, and an N such that

$$\frac{1}{N}\sum_1^N 1_V(T^nx)<\varepsilon \quad \text{all } x\in X.$$

To see this for given $\varepsilon>0$, find N so that $(**)$ holds. Now for each x there is some open set $V_x\supset E$ such that

$$\frac{1}{N}\sum_1^N 1_{\bar V_x}(T^nx)<\varepsilon,$$

where $\bar V_x$ is the closure of V_x. Since T is continuous there is also some neighbourhood U_x of x, such that for all $y\in U_x$

$$\frac{1}{N}\sum_1^N 1_{V_x}(T^ny)<\varepsilon.$$

Now since X is compact, there are finitely many x_i, $1\le i\le I$, such that $X=\bigcup_{i=1}^I U_{x_i}$ and $V=\bigcap_{i=1}^I V_{x_i}$ will satisfy our requirements.

It does not seem to be easy to verify directly that a set is small. Luckily, there is an alternative characterization which gives a useful method to produce small sets.

PROPOSITION. *A closed set $E \subset X$ is small if and only if for every T-invariant measure μ, $\mu(E)=0$.*

Proof. (1) In one direction, the restriction to closed sets is irrelevant. Indeed, integrating (**) with respect to a probability measure μ, gives

$$\frac{1}{N}\sum_1^N \mu(T^{-n}E) = \int \frac{1}{N}\sum_1^N I_E(T^n x)\, d\mu < \varepsilon.$$

If μ is T-invariant this implies $\mu(E) \leq \varepsilon$ and since ε was arbitrary $\mu(E)=0$.

(2) Now suppose that E is closed and *not small.* Then for some ε_0, and all N, there are points x_N such that

$$\frac{1}{N}\sum_1^N 1_E(T^n x_N) \geq \varepsilon_0.$$

Let $\mu_N = N^{-1}\sum_1^N \delta_{T^n x_N}$, and let μ be an accumulation point of the μ_N's in the w^*-topology. Since for any bounded function f

$$\left| \int f\, d\mu_N - \int Tf\, d\mu_N \right| \leq \frac{2}{N}\|f\|_\infty$$

it follows that μ is T-invariant. If now $\mu(E)=0$, there would be an open set $V \supset E$ such that

$$\mu(V) \leq \varepsilon_0/2.$$

Since E is closed there is a continuous function f which equals one on E, vanishes on $X \setminus V$ and is between 0 and 1 everywhere. Then on the one hand, for all N,

$$\int f\, d\mu_N \geq \varepsilon_0$$

while on the other hand

$$\int f\, d\mu \leq \mu(V) < \varepsilon_0/2$$

which is a contradiction. □

The example $E = \{T^n x_0 : n \in \mathbb{N}\}$ for a non-periodic point shows that a non-closed set E can be μ-null for all invariant measures μ without being small.

2. *The construction*

For this section we suppose that $T: X \to X$ is a homeomorphism, $x_0 \in X$ a non-periodic point and d a metric on X such that there are an abundance of T-small sets. Our notation for the ball, annulus and sphere about x_0 is:

$$B(r) = \{x \in X : d(x, x_0) \leq r\}$$
$$A(I) = \{x \in X : d(x, x_0) \in I\}, \quad I \subset [0, +\infty) \text{ an interval}$$
$$S(r) = \{x \in X : d(x, x_0) = r\}.$$

Specifically, we assume that for a dense set of r's $S(r)$ is a *small* set. In the next section we shall discuss several situations where this hypothesis actually holds. Parameters for the construction of a continuous function

$$\varphi : X \to [0, 1]$$

such that $\{\varphi T^n\}$ has small entropy, will be integers $L_0 < L_1 < L_2 < \cdots$ that control just how small the entropy will be.

Since x_0 is not periodic there is some $b_0 > 0$ so that

$$T^i B(b_0) \cap B(b_0) = \phi \quad 1 \le i \le L_0 \tag{1}$$

$$S(b_0) \text{ is a small set.} \tag{2}$$

Since $S(b_0)$ is small there is a neighbourhood of it V_0 and N_1 such that

$$\frac{1}{N_1} \sum_1^{N_1} 1_{V_0}(T^i x) < \frac{1}{L_1} \quad \text{for all } x \in X. \tag{3}$$

Choose $a_0 < b_0$ so that

$$A([a_0, b_0]) \subset V_0 \tag{4}$$

$$S(a_0) \text{ is a small set.} \tag{5}$$

At this point we can commit ourselves to defining φ for $x \notin A([a_0, b_0])$ as follows:

$$\varphi(x) = \begin{cases} 0, & x \in B(a_0) \\ 1, & x \notin B(b_0). \end{cases}$$

So far, along any orbit $\{\varphi(T^i x)\}$ values that are neither 0 nor 1 occur with a frequency at most $1/L_1$, while of the two values $\{0, 1\}$, 1 occurs with a frequency at least $(L_0 - 1)/L_0$.

Let us denote the interval $[a_0, b_0]$ by I. The next step will be to find intervals $I_0 = [a_0, b_{00}]$, $I_1 = [a_{10}, b_0]$ inside I so that:

$$A(I_0) \cup A(I_1) \subset V_1 \quad \text{where } V_1 \text{ is a neighbourhood of } S(a_0) \cup S(b_0) \text{ that satisfies for some } N_2, \tag{6}$$

$$\frac{1}{N_2} \sum_{i=1}^{N_2} 1_{V_1}(T^i x) < 1/L_2 \quad \text{for all } x \in X \tag{7}$$

and

$$S(b_{00}),\ S(a_{10}) \text{ are small sets.} \tag{8}$$

This will enable the next stage in the construction to proceed. Now the definition of φ may be extended to $A([b_{00}, a_{10}])$ by setting φ equal to $\frac{1}{2}$ there. As a result of this, given that $\varphi(x)$ equals neither 0, nor 1, it is overwhelmingly probable, with a uniform frequency that exceeds $(L_2 - 1)/L_2$, that $\varphi(x) = \frac{1}{2}$.

Next, in each of I_0 and I_1 we find a pair of intervals $I_{00}, I_{01} \subset I_0$; $I_{10}, I_{11} \subset I_1$ all of whose endpoints c are such that

$$S(c) \text{ is a small set.} \tag{9}$$

$$A(I_{00}) \cup A(I_{01}) \cup A(I_{10}) \cup A(I_{11}) \subset V_2 \tag{10}$$

$$\frac{1}{N_3} \sum_1^{N_3} 1_{V_2}(T^i x) < 1/L_3 \tag{11}$$

and are located as in the following:

$$\begin{array}{ccccccc} & & & \varphi=\frac{1}{2} & & & \\ & \varphi=\frac{1}{4} & & & & \varphi=\frac{3}{4} & \\ I_{00} & & I_{01} & & I_{10} & & I_{11} \\ & I_0 & & & & I_1 & \\ & & & I & & & \end{array}$$

We set now

$$\varphi(x)=\begin{cases}\frac{1}{4} & x\in A(I_0)\backslash(A(I_{00})\cup A(I_{01}))\\ \frac{3}{4} & x\in A(I_1)\backslash(A(I_{10})\cup A(I_{11})).\end{cases}$$

Thus values that are neither 0, 1, $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$ can occur only with a frequency that is less than $1/L_3$.

It should be clear now how to continue this construction. At the kth stage φ will be undefined on 2^k annuli, let us denote a typical one by $A([c,d])$. By the preceding steps in the construction $S(c)$ and $S(d)$ will be small sets. Also φ will be defined on either 'side' of this annulus to be $j/2^k$, $(j+1)/2^k$ for some $j=0,1,\ldots,2^{k-1}$. Let V_k be a neighbourhood of these 2×2^k spheres such that for some N_{k+1}

$$\text{(i)} \qquad \frac{1}{N_{k+1}}\sum_{i=1}^{N_{k+1}} 1_{V_k}(T^ix)<1/L_{k+1}, \quad \text{all } x\in X.$$

New points e, f are now found $c<e<f<d$ so that

$$\text{(ii)} \qquad A([c,e])\cup A([f,d])\subset V_k$$

(and the same of course for the other annuli)

$$\text{(iii)} \qquad S(e), S(f) \text{ are small sets.}$$

Finally φ is defined to be $(2j+1)/2^{k+1}$ for $x\in A([c,d])\backslash(A([c,e])\cup A([f,d]))$ and similarly for the other intervals.

In fact we are defining φ as a function of $d(x,x_0)$ which is of course continuous. The Cantor-like function of $d(x,x_0)$ that we have defined above is clearly continuous although *not* Lipschitz.

It remains to see what kind of upper bound we have on the entropy of $\{\varphi T^n\}_{n\in\mathbf{Z}}$. Fixing some $\varepsilon>0$ in the computation of the topological entropy of $\{\varphi T^n\}$ is tantamount to fixing some level in the construction of φ, say the kth, and replacing φ on all of $A([c,d])$ (as above) by the single value $(2j+1)/2^{k+1}$. If this is done, the error that we make in the value of φ is at most $1/2^{k+1}$. Now the combinatorics of the number of possible sequences of values that we see is as follows: Let n be very large, then the number of distinct n-blocks that can appear is at most

$$\binom{n}{\varepsilon_1 n}\cdot 2\binom{n}{\varepsilon_2 n}\cdot 2^2\binom{n}{\varepsilon_3 n}\cdots 2^k\binom{n}{\varepsilon_{k+1} n} \qquad (*)$$

where $\varepsilon_i=1/L_i$, and the typical term $2^i\binom{n}{\varepsilon_{i+1}n}$ arises from the values of the $2j+1/2^{i+1}$ which are 2^i in number, and they occur (as a totality) with a uniform frequency dominated by $1/L_{i+1}$. For given $\delta>0$, the L_i's can be chosen so that $(*)$ is bounded by $2^{\delta n}$ which gives the required estimate.

3. *Transversality*

In § 2 we required an abundance of small spheres. To establish assertion (a) of the introduction we merely remark that for any x_0, and any metric, only countably many spheres $S(r)$ can carry positive measure for a fixed measure μ. Since the complement of any countable set is dense the argument of the last section can be carried out in this case.

The spheres need not be exactly at a constant distance from a given point; as is evident from the construction, perturbations will do. Here we prove that if f is a C^1 diffeomorphism of a compact manifold with the property that f has only finitely many periodic points of any given period, then a generic set of C^1 embeddings of spheres of codimension one are small. This proves our assertion (b) of the introduction.

We recall some of the main and simple features of transversality theory from Abraham and Robbin [1]. A map $f: M \to N$ of differentiable manifolds is *transversal* to a submanifold $W \subset N$, if whenever $f(m) \in W$ the image of the derivative $T_m f(T_m M)$ projects surjectively onto $T_{f(m)}N/T_{f(m)}W$ and $T_m f^{-1}(T_{f(m)}W)$ has a closed complement in $T_m M$. (If W has finite codimension the closed complement follows automatically.)

Transversality of $f: M \to N$ to W insures that $f^{-1}(W)$ is a submanifold of M. If $f(M) \cap W = \phi$, transversality is immediate and $f^{-1}(W)$ is the empty submanifold of M.

The next slightly more complicated situation is to consider a map depending on parameters.

LEMMA. *Let*

$$P \times M \xrightarrow{g} N.$$

Suppose g is transversal to the finite codimension submanifold W, then the regular values of the projection $\pi: P \times M \to P$ restricted to $g^{-1}(W)$ are those parameter values $p \in P$ for which $g(p, -): M \to N$ is transversal to W.

Proof. If p is not in the image $\pi(g^{-1}(W))$, $g(p, -)(M) \cap W = \phi$ so there is nothing to prove. If $(p, m) \in g^{-1}(W)$ we need to check surjectivity of the projection of $T_m g(p, -) T_m M$ onto $T_{g(p,m)}N/T_{g(p,m)}W$, assuming that p is a regular value of π. Given $n \in T_{g(p,m)}N/T_{g(p,m)}W$ there is a $(u, v) \in T_{(p,m)}P \times M$ such that the projection of $T_{(p,m)}g(u, v)$ equals n. Since p is a regular value of $\pi | g^{-1}(W)$ there is a $(u', v') \in T_{(p,m)}g^{-1}(W)$ such that $u' = u$. Now $(0, v - v')$ is tangential to M, $T_{(p,m)}g$ applied to $(0, v - v')$ is the same as $T_m g(p, -)$ and the projection into $T_{g(p,m)}N/T_{g(p,m)}W$ is unchanged since $T_{(p,m)}g(u, v') \in T_{g(p,m)}W$, so n is in the image. Similarly transversality of $g(p, -)$ to W implies regularity of p.

The next step is to assure that the projection has many regular values, this is taken care of by Sard's theorem in finite dimensions, with enough smoothness. In infinite dimensions one uses Smale's infinite-dimensional version of this theorem [6] (and see [1]). To apply the theorem to transversality theory we again assume that M is finite dimensional, the main point is that P may be infinite dimensional. □

THEOREM (Abraham–Robbin). *Let P, M, N be C^r manifolds and $W \subset N$ a submanifold where:*

(a) *M has finite dimension m and W has finite codimension q.*

(b) *P and M are second countable.*

(c) $r > \max(0, m-q)$

(d) *$g : P \times M \to N$ is C^r and is transversal to W.*

Then the set of $p \in P$ such that $g(p, -) : M \to N$ is transversal to W is residual (and hence dense) in P. Furthermore, if M is compact this set is open and dense.

We will only apply this theorem when N is also finite dimensional. Note that if P is also finite dimensional then $\dim g^{-1}(w) = \dim P + m - q$ so $\dim g^{-1}(W) - \dim P = m - q$ and the hypotheses allow the application of the usual Sard theorem to the projection taking $g^{-1}(W) \to P$. Thus we may add to the theorem that the residual set is of full measure.

Now we apply this transversality theory to study invariant measures on submanifolds.

PROPOSITION. *Let M be a finite-dimensional manifold and $V \subset M$ a compact submanifold. Suppose that $f : M \to M$ is a C^r diffeomorphism and j a positive integer with $r > \max(j \dim V - (j-1)\dim M, 0)$. Moreover for given distinct positive integers $n_1, \ldots, n_{j-1}$ suppose that the periodic points of f for any fixed period $n \le \max_{i=1,j-1} n_i$ are isolated. Then for an open and dense set of C^r embeddings i of V into M in the C^r topology the map*

$$(i, f^{n_1} \circ i, \ldots, f^{n_{j-1}} \circ i) : V \times \cdot \underset{j}{\cdot} \cdot \times V \to M \times \cdot \underset{j}{\cdot} \cdot \times M$$

is transversal to the small diagonal $\Delta \subset M \times \cdot \underset{j}{\cdot} \cdot \times M$ where $\Delta = \{(m, m, \cdot \underset{j}{\cdot} \cdot, m) \mid m \in M\}$.

COROLLARY 1. *Suppose that $f : M \to M$ is a C^1 diffeomorphism of the finite-dimensional manifold M, and that for each period n the periodic points of f of period n are isolated. Let $V \subset M$ be a compact submanifold of codimension ≥ 1. Then a residual (and hence dense) set of embeddings i of V into M have the following property: let k be a positive integer with*

$$\left(\frac{k}{k-1}\right) < \left(\frac{\dim M}{\dim V}\right).$$

For any distinct positive integers $n_1, \ldots, n_{k-1}$

$$i(V) \cap f^{n_1} i(V) \cap \cdots f^{n_{k-1}} i(V) = \phi.$$

Proof. By the proposition we may suppose that for a residual set of embeddings i that

$$(i, f^{n_1} \circ i, \ldots, f^{n_{k-1}} \circ i) : V \times \cdot \underset{k}{\cdot} \cdot \times V \to M \times \cdot \underset{k}{\cdot} \cdot \times M$$

is transversal to Δ for all k with

$$\frac{k}{k-1} < \frac{\dim M}{\dim V}$$

and all the countable number of sequences of distinct integers $n_1, \ldots, n_{k-1}$. But the codimension of $\Delta = (k-1) \dim M$ and $k \dim V < (k-1) \dim M$ so the image does not intersect Δ. That is to say there is no $m \in M$ such that $m = i(v_0)$ and $m = f^{n_j}(i(v_j))$ for all $i \leq j \leq k-1$.

COROLLARY 2. *Let f, M, V be as in Corollary* 1. *Then for a residual set of embeddings i taking V into M, the first return map of f on i(V) has no invariant probability measure.*

Proof. If there is an invariant probability measure use the Poincaré recurrence theorem to contradict Corollary 1. □

Now we prove the proposition: Let $\mathrm{Emb}^r(V, M)$ be the C^r embeddings of V in M with the C^r topology. The map $ev: \mathrm{Emb}^r(V, M) \times V \to M$, $(i, v) \to i(v)$ is C^r and its derivative at (i, v) on the target vector (h, u) is $T_{(i,v)}ev(h, u) = h(i(v)) + T_v i(u)$ where h is a C^r section of $TM|i(V)$ and u is a tangent vector to V at v, see [**1, 4, 5**]. Now for any $n < \max_{1 \leq i \leq j-1} n_i$ we can easily see that the set $U \subset \mathrm{Emb}^r(V, M)$ of embeddings i such that $i(V)$ contains no periodic point of period $n \leq \max_{1 \leq i \leq j-1} n_i$ is open and dense, for example by noting that $ev: \mathrm{Emb}^r(V, M) \times V \to M$ is transversal to any point $m \in M$ so that embeddings which do not contain a particular point in their image are open and dense. Now we let

$$ev_f: \mathrm{Emb}^r(V, M) \times \underbrace{V \times \cdots \times V}_{j} \to \underbrace{M \times \cdots \times M}_{j}$$

$$(i, v_0, \ldots, v_{j-1}) \to (i(v_0), f^{n_1}(i(v_1), \ldots, f^{n_{j-1}}(i(v_j))))).$$

By the chain rule

$$\begin{aligned} &T_{(i, v_0, \ldots, v_{j-1})} ev_f(h, u_0, \ldots, u_{j-1}) \\ &\quad = (h(i(v_0)) + T_{v_0} i(u_0), \ldots, T_{i(v_r))} f^{n_r}(h(i(v_r)) + Tf^{n_r}_{i(v_r)} T_{v_r} i(u_r), \ldots). \end{aligned}$$

Restrict ev_f to $U \times V \times \cdots \times V$. It is sufficient to prove that this restriction is transversal to Δ.

Assume that

$$i(v_0) = f^{n_1}(i(v_1)) = \quad = f^{n_{j-1}}(i(v_{j-1})).$$

Since $i(V)$ contains no periodic points of V of period less than or equal $\max_{i=1,\ldots,j-1} n_i$, $i(v_0), \ldots, i(v_{j-1})$ is a set of j distinct points. To show that

$$T_{(i, v_0, \ldots, v_{j-1})} ev_f(T_{(i, v_0, \ldots, v_{j-1})} U \times V \times \cdots \times V)$$

projects onto $T_{(i(v_0), \ldots, i(v_{j-1}))} M \times \cdots \times M / T_{(i(v_0), \ldots, i(v_{j-1}))} \Delta$ we show that the element

$$(w, 0, \ldots, 0, \underset{k}{-w}, 0, \ldots, 0)$$

is in the image of $T_{(i, v_0, \ldots, v_{j-1})} ev_f$ for any $w \in T_{i(v)} M$. Since the $i(v_k)$ are all distinct we are free to specify $h(i(v_0)) = w$ and $T_{i(v_k)} f^{n_k}(h(i(v_k))) = -w$, $h(i(v_l)) = 0$ for $l \neq 0$, k and $u_l = 0$, $l = 1, \ldots, j-1$.

4. *Differentiable functions*

In this section we give a lower bound on the entropy of a factor of an Anosov diffeomorphism given by a differentiable function ϕ, and consequently on any factor where the quotient map is C^1.

Definitions. Given a function $\phi: X \to Y$, Y metric, we say a set $E \subset X$ is δ, ϕ-separated if $\forall x_1 \neq x_2 \in E$

$$d(\phi(x_1), \phi(x_2)) > \delta.$$

Given in addition $T: X \to X$ consider $Y^n = Y \times \underset{n}{\cdots} \times Y$ with the sup metric and

$$\phi_n = (\phi, \phi \circ T, \phi \circ T^2, \ldots, \phi \circ T^{n-1}): X \to Y^n.$$

Let $s(n, \delta)$ be the maximum cardinality of a δ, ϕ_n separated set. The δ, ϕ-entropy of T is

$$h_{\delta,\phi}(T) = \limsup \frac{1}{n} \log s(n, \delta)$$

and the ϕ entropy of T is

$$h_\phi(T) = \lim_{\delta \to 0} h_{\delta,\phi}(T).$$

It is simple to see that if $\phi: X \to I$, then the ϕ entropy of T is the same as the entropy of the factor of T given by $\Phi = \{\phi \circ T^n\}$ as in the introduction.

Let M be a compact differentiable manifold without boundary. Let $f: M \to M$ be a C^1 diffeomorphism. Then f is Anosov if there is a Riemannian metric with norm $\| \; \|$, two reals $0 < \lambda < 1 < \mu$ and a splitting of the tangent bundle of M: $TM = E^s \oplus E^u$ such that

(1) $Tf \,|\, E^s = E^s$ and the operator norm $\| Tf \,|\, E^s_x \| < \lambda \quad \forall x \in M$

(2) $Tf^{-1} \,|\, E^u = E^u$ and the operator norm $\| Tf^{-1} \,|\, E^u_x \| < \mu^{-1} \quad \forall x \in M.$

The maximum of all such μ is called the minimal expansion $\mu_{\min}$.

The bundles E^s and E^u have f invariant foliations tangent to them, see [3]. The leaves of the foliation through x are denoted $W^s(x)$ and $W_u(x)$ respectively, and the open ε-discs in the leaf metric around x in these leaves $W^s_\varepsilon(x)$ and $W^u_\varepsilon(x)$. From (2) it follows immediately that

(3) $f^n(W^u_\varepsilon(x)) \supset W^u_{\mu^n\varepsilon}(f^n(x))$ for all n.

(4) The $W^s_\varepsilon(x)$, $W^u_\varepsilon(x)$ are given as images via the exponential map of Lipschitz functions with Lipschitz constant one from the ε disc in $E^s(x)$ to $E^u(x)$ or $E^u(x)$ to $E^s(x)$ respectively.

THEOREM. *Let f be a C^1 transitive Anosov diffeomorphism of the compact manifold M, with minimal expansion $\mu_{\min}$. Let $\phi: M \to R$ be a non-constant C^1 function. Then the ϕ entropy of f is greater than or equal to $\log \mu_{\min}$.*

Proof. Since f is transitive every unstable manifold is dense. Thus since ϕ is non-constant $D\phi \,|\, E^u$ is somewhere non-zero, say at x.

Fix $1 < \mu < \mu_{\min}$ a Riemannian metric $\varepsilon > 0$ small, and take a neighbourhood of x, given as $\bigcup_{y \in W^s_\varepsilon} W^u_\varepsilon(y)$. For ε small enough there is a constant $c > 0$ such that for any $0 < \delta \leq \varepsilon$ and any $y \in W^s_\varepsilon(x)$, $W^u_\varepsilon(y)$ contains at least $c\varepsilon/\delta$ disjoint δ balls around points $y_1, \ldots, y_{c\varepsilon/[\delta]}$ and any selection of points $z_1, \ldots, z_{c\varepsilon/[\delta]}$ in these balls are ϕ, $c\delta$ separated, i.e. $|\phi(z_i) - \phi(z_j)| > c\delta$.

LEMMA 1. *For $\varepsilon > 0$ small, there is an $R > 0$ with the following property. Given x, $z \in M$ then the R ball around z in $W^u(z)$ contains $W^u_\varepsilon(y)$ for some $y \in W^s_\varepsilon(x)$.*

Proof. For any $z \in M$, $W^u(z)$ is dense, hence there is a minimum $R(z)$ such that the $R(z)$ ball around z is $\varepsilon/3$ dense. Since the unstable manifolds are continuous on compact discs in the C^1 topology $R(z)$ is continuous, hence bounded by R_0. Adding 3ε to R_0 gives R.

Proof of Theorem. Beginning with $W^u_\varepsilon(x)$ we find $(c\varepsilon/\delta)$ disjoint δ balls which are $c\delta$, ϕ separated. Let n_0 be minimal such that $\mu^{n_0}\delta \geq R$. Apply f^{n_0}, each of these balls contains a $W^u_\varepsilon(y)$ which can in turn be divided into $(c\varepsilon/\delta)$ disjoint δ balls which are $c\delta$, ϕ separated. Inductively we find $(c\varepsilon/\delta)^j$, $c\delta$, $\phi_{n_0 j}$ separated sets. Thus the $c\delta$, ϕ entropy of f is

$$\geq \frac{1}{n_0 j}\, j \log\left(\frac{c\varepsilon}{\delta}\right) \quad \text{and} \quad n_0 \leq \frac{\log R - \log \delta}{\log \mu} + 1$$

so the $c\delta$, ϕ entropy of f is

$$\geq \frac{(\log c + \log \varepsilon - \log \delta)\log \mu}{\log R - \log \delta + \log \mu}$$

letting $\delta \to 0$ and taking the limit gives $\log \mu$.

5. *The one-sided shift*

Let $\pi : [0,1]^{\mathbb{N}} \to X$ be a continuous surjective map with $\pi\sigma = T\pi$, where σ is the shift on $\Omega = [0,1]^{\mathbb{N}}$ and $T : X \to X$ is a continuous mapping. Fix metrics on Ω and X, and normalize the metric on X so that for some $x_0 \in X$, $\sup_{y \in X} d(y, x_0) = 1.2$. Observe that given any $\delta > 0$, for some N and *all* $\omega \in \Omega$, $\sigma^{-N}(\omega)$ is δ-dense in Ω. Furthermore $\sigma^{-N}(\omega)$ is the N-cube and is therefore connected. Since π is continuous, there is some N so that for all $\omega \in \Omega$, $\pi(\sigma^{-N}(\omega))$ is $\frac{1}{100}$-dense in X.

Define now

$$\varphi(x) = \begin{cases} 0 & \text{if } d(x, x_0) \leq \frac{1}{10} \\ d(x, x_0) - \frac{1}{10} & \text{if } \frac{1}{10} \leq d(x, x_0) \leq \frac{11}{10} \\ 1 & \text{if } d(x, x_0) \geq \frac{11}{10}. \end{cases}$$

Clearly φ is continuous, and maps X to $[0,1]$.

Claim. For any $x \in X$, $\varphi(T^{-N}x) = [0,1]$.

Proof. Since π is surjective there is some $\omega \in \Omega$ with $\pi(\omega) = x$. Notice that $\pi(\sigma^{-N}(\omega)) \subset T^{-N}x$, and that $\pi(\sigma^{-N}(\omega))$ is a *connected* set that is $\frac{1}{100}$-dense in X. It follows that $\varphi(\pi(\sigma^{-N}(\omega))) = [0,1]$ and *a fortiori* the same holds for $T^{-N}x$. $\square$

As before we now define

$$\Phi : X \to [0,1]^{\mathbb{N}}$$

by

$$\Phi(x) = (\varphi(x), \varphi(T^N x), \varphi(T^{2N} x), \ldots).$$

By the claim, Φ is onto and this proves the following theorem.

THEOREM. *If (X, T) is a continuous factor of the one-sided shift on $[0,1]^{\mathbb{N}}$ then for some power N, the one-sided shift is a factor of T^N.*

REFERENCES

[1] R. Abraham & J. Robbin. *Transversal Mappings and Flows.* Benjamin, New York, 1967.

[2] H. Furstenberg, H. Keynes & L. Shapiro. Prime flows in topological dynamics. *Israel J. Math.* **14** (1973), 26–38.

[3] M. W. Hirsch & C. C. Pugh. Stable manifolds and hyperbolic sets. In: *Proc. Symp. Pure Mathematics*, Vol. XIV, (S. S. Chern and S. Smale, eds.) American Mathematics Society, Providence, RI (1970) pp. 133–163.

[4] M. Shub. Endomorphisms of compact differentiable manifolds. *Amer. J. Math.* **XCI** (1969), 175–199.

[5] M. Shub. *Global Stability of Dynamical Systems.* Springer, New York (1986).

[6] S. Smale. An infinite dimensional version of Sard's theorem. *Amer. J. Math.* **LXXXVII** (1965), 861–866.