

## ON THE PREVALENCE OF HORSESHOES

BY

LAI-SANG YOUNG<sup>1</sup>

*To the memory of my adviser, Rufus Bowen*

**ABSTRACT.** In this paper the symbolic dynamics of several differentiable systems are investigated. It is shown that many well-known dynamical systems, including Axiom A systems, piecewise monotonic maps of the interval, the Lorenz attractor and Abraham-Smale examples, have inside them subsystems conjugate to subshifts of finite type. These subsystems have hyperbolic structures and hence are stable. They can also be chosen to have entropy arbitrarily close to that of the ambient system.

Let  $f: M \rightarrow M$  be a diffeomorphism of a manifold into itself and let  $\Omega(f)$  be its nonwandering set. While it is usually difficult to describe completely the orbit structure of  $f$ , one could look for large invariant subsets of  $\Omega(f)$  on which the dynamics of  $f$  are simple to characterize. This idea is due to Rufus Bowen, who proposed that one asks the following: Given  $\varepsilon > 0$ , does there exist a hyperbolic  $f$ -invariant set  $\Lambda_\varepsilon \subset \Omega(f)$  with the property that

- (1)  $f|_{\Lambda_\varepsilon}$  is conjugate to a subshift of finite type, and
- (2)  $h(f|_{\Lambda_\varepsilon}) > h(f) - \varepsilon$  where  $h(f)$  is the topological entropy of  $f$ ?

We say that  $f$  is, in the sense of entropy, a *limit of hyperbolic subshifts of finite type* if  $\Lambda_\varepsilon$  above exists for every  $\varepsilon > 0$ . The main result of this paper is that many dynamical systems we know are limits of this kind. They include Axiom A diffeomorphisms and flows, piecewise monotonic maps of the interval, the Poincaré map of the Lorenz attractor [15] and certain Abraham-Smale examples [1]. Perhaps this new class of systems accounts for some of the other non-Axiom A examples as well (e.g. [11], [17], [19], [23], [24], [28]), but that remains to be decided.

Recall the definition of a subshift of finite type. Let  $\{1, \dots, n\}$  be given the discrete topology and  $\Sigma = \prod_{-\infty}^{\infty} \{1, \dots, n\}$  the product topology. Let  $\sigma: \Sigma \rightarrow \Sigma$  be defined by  $(\sigma x)_i = (x)_{i+1}$  where  $(x)_i$  denotes the  $i$ th coordinate of  $x$ . Let  $A = [A_{ij}]$  be an  $n \times n$  matrix of 0's and 1's. Let  $\Sigma_A = \{x \in \Sigma: A_{x_i, x_{i+1}} = 1 \text{ for all } i \in \mathbb{Z}\}$ . Then  $\sigma|_{\Sigma_A}$  is called a *subshift of finite type*. If  $\Sigma' = \prod_0^{\infty} \{1, \dots, n\}$  and  $\Sigma'_A = \{x \in \Sigma': A_{x_i, x_{i+1}} = 1 \text{ for all } i \geq 0\}$ , then  $\sigma|_{\Sigma'_A}$  is called a *one-sided* subshift of finite type. From now on "subshifts of finite type" will be abbreviated as "ssft".

---

Received by the editors January 26, 1979.

AMS (MOS) subject classifications (1970). Primary 58F15; Secondary 28A65.

Key words and phrases. Subshift of finite type, topological entropy, hyperbolicity.

<sup>1</sup>This is part of the author's Ph.D. thesis at the University of California at Berkeley.

When the mapping in question is noninvertible, by a ssft we always mean a one-sided ssft. In the case of a flow,  $\Lambda_\varepsilon$  is the suspension of a ssft.

We make a couple of remarks here. First, if  $f$  is a limit of hyperbolic sets then it is stable in the sense that as “large” a subset of  $\Omega(f)$  as one wishes persists under perturbations. Second, since zero-dimensional Axiom A basic sets and topologically transitive<sup>2</sup> ssft are equivalent, Bowen’s idea can be viewed as an attempt to generalize Smale’s Axiom A systems [9], [29].

**1. Axiom A systems.** A diffeomorphism  $f$  of a manifold into itself is said to satisfy *Axiom A* if (a)  $\Omega(f)$  is hyperbolic and (b) periodic points are dense in  $\Omega(f)$ . Axiom A flows are defined similarly. For an exposition on the subject see [29].

**THEOREM 1.1.** *Let  $f \in \text{Diff}(M)$  be an Axiom A diffeomorphism. Let  $\Omega_s \subset M$  be a basic set. Then there is a nested sequence of compact  $f$ -invariant sets  $X_1 \subset X_2 \subset \cdots$  with  $\bigcup X_i = \Omega_s$  satisfying*

- (1)  $f|X_i$  is conjugate to a ssft, and
- (2)  $h(f|X_i) \uparrow h(f|\Omega_s)$ .

*The analogous statement for flows also holds.*

Ssft and Axiom A basic sets have long been known to be intimately related. Via Markov partitions [8] every basic set can be realized as the quotient of a ssft. When constructing our  $X_i$ ’s in the theorem, it is this ssft that is exploited.

The notion of topological pressure defined by Ruelle [27] and studied by Walters [30] is the key to proving the theorem for flows. We state the definitions. If  $f: X \rightarrow X$  is a continuous map of the compact metric space  $X$  into itself then  $E \subset X$  is  $(n, \varepsilon)$ -separated if for each pair of distinct points  $x, y \in E$ ,  $d(f^k x, f^k y) > \varepsilon$  for some  $k$ ,  $0 \leq k \leq n - 1$ . Let  $\phi: X \rightarrow \mathbf{R}$  be a continuous function. Let

$$Z_n(f, \phi, \varepsilon) = \sup \left\{ \sum_{x \in E} \exp \sum_{k=0}^{n-1} \phi f^k x : E \text{ is } (n, \varepsilon)\text{-separated} \right\}$$

and

$$P(f, \phi, \varepsilon) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log Z_n(f, \phi, \varepsilon).$$

Then  $P(f, \phi) = \lim_{\varepsilon \rightarrow 0} P(f, \phi, \varepsilon)$  is called the *topological pressure* of  $f$  for the function  $\phi$ . When  $\phi \equiv 0$  the number  $P(f, \phi)$  is just the topological entropy  $h(f)$  of  $f$ ; the theory of topological pressure generalizes that of topological entropy (as defined in [3]). If  $Y \subset X$  is a compact  $f$ -invariant subset, then  $P_Y(f, \phi)$  denotes the pressure when everything is restricted to  $Y$ .

A homeomorphism  $f: X \rightarrow X$  is said to be *topologically mixing* if for every pair of nonempty open sets  $U, V \subset X$ , there exists  $M \subset \mathbf{Z}^+$  such that  $U \cap f^n V \neq \emptyset$  whenever  $n \geq M$ .

---

<sup>2</sup>The usual definition of ssft requires topological transitivity. If  $A$  is a transition matrix, there is always an irreducible submatrix  $B$  such that  $h(\sigma|_{\Sigma_B}) = h(\sigma|_{\Sigma_A})$ . Thus our results hold using either definition.

LEMMA 1.2. Let  $(\Sigma, \sigma)$  be a topologically mixing ssft. Let  $A, B \subset \Sigma$  be closed subsets with the property that  $A, B \subsetneq \Sigma$ ,  $\sigma A \subset A$  and  $\sigma^{-1}B \subset B$ . Let  $\phi: \Sigma \rightarrow \mathbb{R}$  be a step function depending only on the 0th coordinate. Then given  $\varepsilon > 0$ , there is a ssft  $X \subset \Sigma$  satisfying

- (1)  $X \cap (A \cup B) = \emptyset$ , and
- (2)  $P_X(\sigma, \phi) > P(\sigma, \phi) - \varepsilon$ .

PROOF. Since periodic points of mixing ssft are dense, let us fix one in  $\Sigma \setminus (A \cup B)$ . Call it  $\mathbf{a} = (\cdots a_{-1}a_0a_1\cdots)$ .  $A \cup B$  being closed implies that there is a positive integer  $N_0$  such that for all  $\mathbf{x} \in \Sigma$ , if  $x_i = a_i$  for  $i = -N_0, \dots, N_0$ , then  $\mathbf{x} \notin (A \cup B)$ . Choose  $N \geq N_0$  such that

$$\frac{1}{2N} \log \sum_{\substack{\mathbf{x} \in \Sigma \\ x_0 = a_0 \\ \sigma^{2N}\mathbf{x} = \mathbf{x}}} \exp \sum_{i=0}^{2N-1} \phi \sigma^i \mathbf{x} > P(\phi) - \varepsilon$$

(see [30]). Assume also that the period of  $\mathbf{a}$  divides  $N$ . Let  $\mathcal{S} = \{\mathbf{b}_1, \dots, \mathbf{b}_q\}$  be the set of periodic  $2N$ -blocks starting with  $a_0$  that occur in the elements of  $\Sigma$ . Let  $\mathbf{b}_1 = (a_{-N}, \dots, a_{N-1})$ .  $N$  is fixed for the rest of the proof.

We now define a sequence of ssft. For each  $m = 1, 2, \dots$  let  $S_m = \{\mathbf{x} \in \Sigma: \text{no } m\text{-block of } \mathbf{x} \text{ appears in any element } \mathbf{y} \in A \text{ to the right of } y_{-N} \text{ or in any element } \mathbf{z} \in B \text{ to the left of } z_N\}$ . Notice that if a symbol sequence in  $\Sigma$  has the property that the  $2N$ -block  $\mathbf{b}_1$  occurs at least once in every one of its  $2mN$ -blocks then it is automatically in  $S_{2mN}$ . For if say  $\mathbf{b}_1$  appeared in some  $\mathbf{y} \in A$  to the right of  $y_{-N}$ , then  $\sigma^i \mathbf{y} \notin A$  for some  $i \geq 0$  contradicting  $\sigma A \subset A$ . Clearly  $S_m \cap (A \cup B) = \emptyset$ . We shall see that  $\sup_m P_{S_m}(\sigma, \phi) = P(\sigma, \phi)$ .

Consider  $2Nmn$ -blocks of the form

$$\mathbf{b}_{i_1} \mathbf{b}_{i_2} \cdots \mathbf{b}_{i_{m-2}} \mathbf{b}_1 \mathbf{b}_1 \mathbf{b}_{i_{m-1}} \cdots \mathbf{b}_{i_{2(m-2)}} \mathbf{b}_1 \mathbf{b}_1 \cdots \mathbf{b}_{i_{n(m-2)}} \mathbf{b}_1 \mathbf{b}_1$$

where  $\mathbf{b}_{i_j}$  can be any element of  $\mathcal{S}$ . These blocks appear in elements of  $S_{2mN}$ . Let  $w_i = \sum_{j=0}^{2N-1} \phi \sigma^j \mathbf{b}_i$ . Let  $E_n(S_m)$  be a maximal subset of  $S_m$  whose distinct elements  $\mathbf{x}$ ,  $\mathbf{y}$  have  $x_i \neq y_i$  for some  $i$ ,  $0 \leq i < n-1$ . Then

$$\begin{aligned} Z_{2mnN}(\phi)|_{S_{2mN}} &= \sum_{\mathbf{x} \in E_{2mnN}(S_{2mN})} \exp \sum_{i=0}^{2nmN-1} \phi \sigma^i \mathbf{x} \\ &> \sum_{\substack{(i_1, \dots, i_{n(m-2)}) \\ 1 \leq i_j < q}} \exp \left[ \sum_{j=1}^{n(m-2)} w_{i_j} + 2nw_1 \right] \\ &= e^{2nw_1} \sum_{(i_1, \dots, i_{n(m-2)})} \prod_{j=1}^{n(m-2)} \exp w_{i_j} \\ &= e^{2nw_1} \left[ \sum_{i=1}^q \exp w_i \right]^{n(m-2)}, \end{aligned}$$

so that

$$\begin{aligned}
 P_{S_{2mN}}(\phi, \sigma) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log Z_n(\phi) | S_{2mN} \\
 &\geq \lim_{n \rightarrow \infty} \frac{1}{2nmN} \log e^{2nw_1} \left( \sum_{i=1}^q \exp w_i \right)^{n(m-2)} \\
 &= \frac{m-2}{m} \left( \frac{1}{2N} \log \sum_{i=1}^q \exp w_i \right) + \frac{w_1}{mN}.
 \end{aligned}$$

Since  $(1/2N) \log \sum_{i=1}^q \exp w_i > P(\phi) - \varepsilon$ ,  $X = S_{2mN}$  for large enough  $m$  is our desired invariant subset of  $\Sigma$ .  $\square$

**PROOF OF THEOREM 1.1 FOR DIFFEOMORPHISM CASE** ( $\phi \equiv 0$ ). By the spectral decomposition theorem [8], we have  $\Omega_s = \Lambda_1 \cup \dots \cup \Lambda_r$ , the  $\Lambda_i$ 's being closed disjoint sets with  $f\Lambda_i = \Lambda_{i+1}$ ,  $f\Lambda_r = \Lambda_1$  and where  $f^r|_{\Lambda_i}$  is topologically mixing. Since it suffices to prove our assertion for  $f^r|_{\Lambda_1}$ , we may as well assume that  $f$  itself is topologically mixing.

Let  $\mathcal{R} = \{R_0, \dots, R_k\}$  be a Markov partition on  $\Omega_s$  and  $\Sigma \subset \prod_{-\infty}^{\infty} \{0, \dots, k\}$  its associated ssft.

$$\begin{array}{ccc}
 \Sigma & \xrightarrow{\sigma} & \Sigma \\
 \pi \downarrow & & \downarrow \pi \\
 \Omega_s & \xrightarrow{f} & \Omega_s
 \end{array}$$

We know that  $(\Sigma, \sigma)$  is also topologically mixing and that  $h(\sigma) = h(f)$ . Let  $A = \pi^{-1}\partial^s \mathcal{R}$ ,  $B = \pi^{-1}\partial^u \mathcal{R}$ ,  $\phi \equiv 0$ . All the hypotheses in 1.2 are satisfied.  $\pi|_X$  ( $X$  as in Lemma 1.2) is 1-1 and is therefore an embedding.

To get a nested sequence of  $f$ -invariant subsets in  $\Omega_s$ , suppose we have  $Y_1, Y_2, \dots \subset \Sigma(\text{ssft})$  obtained as above with  $h(\sigma|_{Y_i}) \uparrow h(\sigma)$ . Since  $Y_1, Y_2 \cap (A \cup B) = \emptyset$ , there is an  $M$  such that for all  $\mathbf{x} \in \Sigma$ ,  $(x_{-M}, \dots, x_M)$  appearing in elements of  $Y_1$  or  $Y_2$  implies  $\mathbf{x} \notin (A \cup B)$ . Let  $\tilde{Y}_2 = \{\mathbf{x} \in \Sigma: \text{any } (2M+1)\text{-block is admissible if and only if it appears in } Y_1 \text{ or } Y_2\}$ . One lets  $X_1 = \pi Y_1$ ,  $X_2 = \pi \tilde{Y}_2$  and so on. The proof is complete if we observe that  $\bigcup X_i$  must be all of  $\Omega_s$ , for  $h(\sigma|\pi^{-1}\bigcup X_i) = h(\sigma)$  and no invariant closed proper subsets of a mixing ssft can have full entropy.

**PROOF FOR FLOWS.** Let  $\Psi: \Sigma \rightarrow \mathbf{R}^+$  be a continuous function and let  $\{S_t\}$  be the special flow built on  $\sigma: \Sigma \curvearrowright$  under  $\Psi$ . Let  $\Lambda = \{(\mathbf{x}, t): t \in [0, \psi \mathbf{x}], \mathbf{x} \in \Sigma\} \subset \Sigma \times \mathbf{R}$ . Every Axiom A flow on a basic set is a quotient of one of these flows (where  $\Sigma$  is a topologically mixing ssft) in a way very similar to the diffeomorphism case [7], [10]. We produce a  $\sigma$ -invariant subset  $X \subset \Sigma$  on which the flow has correct entropy and is identification-free when pushed down to the basic set.

Let  $\mathfrak{N}_{S_t}(\Lambda)$  denote the set of  $S_t$ -invariant Borel probability measures on  $\Lambda$  and similarly for  $\mathfrak{N}_\sigma(\Sigma)$ . A theorem of Abramov [2] states that for  $\mu \in \mathfrak{N}_{S_t}(\Lambda)$ ,

$$h_\mu(S_1) = h_\nu(\sigma) / \int \psi d\nu$$

where  $\nu \in \mathfrak{N}_\sigma(\Sigma)$  is the measure induced. Since  $h(S_1) = \sup_{\mu \in \mathfrak{N}_{S_t}(\Lambda)} h_\mu(S_1)$  [12], we have, by the Variational Principle

$$P(\sigma, -h(S_1)\psi) = \sup_{\nu \in \mathfrak{M}_\sigma(\Lambda)} \left( h_\nu(\sigma) - \int h(S_1)\psi \, d\nu \right) = 0.$$

For further details see [10] and [30].

Now fix some arbitrary  $\delta > 0$ . From the formula above,  $P(\sigma, -(h(S_1) - \delta)\psi) = c > 0$ . Choose a step function  $\phi: \Sigma \rightarrow \mathbf{R}$  with  $\|\phi + (h(S_1) - \delta)\psi\| \leq c/3$ . In Lemma 1.2 set  $\varepsilon = c/3$ . Then the set  $X \subset \Sigma$  obtained has the property that

$$\begin{aligned} P(-(h(S_1) - \delta)\psi) - P_X(-(h(S_1) - \delta)\psi) \\ &\leq |P(-(h(S_1) - \delta)\psi) - P(\phi)| + |P(\phi) - P_X(\phi)| \\ &\quad + |P_X(\phi) - P_X(-(h(S_1) - \delta)\psi)| \\ &\leq c/3 + c/3 + c/3. \end{aligned}$$

Thus  $P_X(-(h(S_1) - \delta)\psi) > 0$ . But this means, if we go through the argument again, that

$$h(S_1 \text{ over } X) = \sup_{\mu \in \mathfrak{M}_\sigma(X)} \frac{h_\mu(\sigma)}{\int \psi \, d\mu} > h(S_1) - \delta.$$

This completes the proof.  $\square$

**2. Maps of the interval.** Throughout this section let  $f: [0, 1] \rightarrow [0, 1]$  be a piecewise monotone mapping. That is, there is a partition  $0 = c_0 < \dots < c_q = 1$  such that on each  $(c_{i-1}, c_i)$   $f$  is strictly monotonic. Let  $A_i = (c_{i-1}, c_i)$ ,  $\bar{A}_i = [c_{i-1}, c_i]$  and  $\mathfrak{A} = \{A_1, \dots, A_q\}$ . Let  $f(c_i) = \lim_{x \rightarrow c_i^+} f(x)$  or  $\lim_{x \rightarrow c_i^-} f(x)$ . In the case  $f$  is not continuous at  $c_i$ , it may be convenient to think of  $f$  as taking on both values, though we do not particularly care which.

First define  $\Sigma(f) = \{x \in \Pi_0^\infty\{1, \dots, q\} : \cap_{i=0}^n f^{-i}A_{x_i} \neq \emptyset \text{ for all } n \geq 0\}$ .  $\Sigma(f)$  is compact; the shift operator takes  $\Sigma(f)$  into itself.  $(\Sigma(f), \sigma)$  is called the *symbolic dynamics* of  $f$ .

The following proposition was arrived at independently by several people: J. Rothschild, Misiurewicz and Szlenk [22] and myself.

**PROPOSITION 2.1.** *Let  $f$  be continuous and piecewise monotonic. Then  $h(\sigma|\Sigma(f)) = h(f)$ . It follows that  $\lim_{n \rightarrow \infty} (1/n) \log(\# \text{ turning points of } f^n)$  exists and equals  $h(f)$ .*

**PROOF.** Let  $\Sigma_I(f) = \{(x, x) \in \Sigma(f) \times [0, 1] : f^i x \in \bar{A}_{x_i} \text{ for all } i \geq 0\}$ . Let  $\sigma_I: \Sigma_I(f) \rightarrow \Sigma(f)$  be defined by  $\sigma_I(x, x) = (\sigma x, fx)$ . We have

$$\begin{array}{ccccc} \Sigma_I(f) & \xrightarrow{\sigma_I} & \Sigma_I(f) & \xrightarrow{\sigma_I} & \Sigma_I(f) \\ \pi_1 \downarrow & & \downarrow \pi_1 & & \downarrow \pi_2 \\ \Sigma(f) & \xrightarrow{\sigma} & \Sigma(f) & \xrightarrow{f} & [0, 1] \end{array}$$

where  $\pi_i$  is a projection into the  $i$ th factor. Since [5]

$$h(\sigma) \leq h(\sigma_I) \leq h(\sigma) + \sup_{x \in \Sigma(f)} h(\sigma_I|_{\pi_1^{-1}x})$$

and  $\sigma_I$  restricted to fibers maps intervals monotonically into intervals therefore having entropy zero, one has  $h(\sigma|\Sigma(f)) = h(\sigma_I)$ . Similarly,  $h(\sigma_I) = h(f)$  because  $\pi_2$

is at most a 2-to-1 map. Finally the number of monotonic parts of  $f^n$  is simply the number of distinct  $n$ -strings that appear in elements of  $\Sigma(f)$ . Thus the turning point formula.  $\square$

We state a consequence of Proposition 2.1, even though it is irrelevant in our ensuing discussion. (See also [21], [25], [26].)

**PROPOSITION 2.2.** *Let  $f$  be continuous, piecewise linear with slope  $= \pm \lambda$  for some  $\lambda \geq 1$ . Then  $h(f) = \log \lambda$ .*

**PROOF.** That  $h(f) \leq \log \lambda$  follows from Kushnirenko's formula or from simple spanning set type arguments. For the other inequality consider the length of  $A \in \bigvee_{i=0}^{n-1} f^{-i}\mathfrak{A}$ . Since  $f^n$  is monotone in  $A$ , one has  $l(A) \leq 1/\lambda^n$  so that  $f^n$  has at least  $\lambda^n$  monotonic parts. Proposition 2.1 now gives the result.  $\square$

In general, topological entropy is defined only for continuous maps. In the case of a piecewise monotonic map, discontinuous perhaps at finitely many points, one could define  $h(f)$  using open coverings, spanning sets or separated sets. It is not clear to me that all of these definitions agree, but for the Poincaré map of the Lorenz attractor, our main interest in these maps in this paper, it is natural to define  $h(f)$  to be  $h(\sigma|\Sigma(f))$ . We first check to see that this makes sense.

**PROPOSITION 2.3.** *Let  $\tilde{\mathfrak{A}}$  be a partition of  $[0, 1]$  into arbitrarily short intervals. Assume  $\tilde{\mathfrak{A}}$  is a refinement of  $\mathfrak{A}$ . Then*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \text{card} \bigvee_{i=0}^{n-1} f^{-i}\tilde{\mathfrak{A}} = h(\sigma|\Sigma(f)).$$

**PROOF.** Clearly  $\text{card} \bigvee_{i=0}^n f^{-i}\tilde{\mathfrak{A}} \geq \text{card} \bigvee_{i=0}^n f^{-i}\mathfrak{A}$ . For each  $n$ , let  $a_n$  be the maximum number of elements in  $\bigvee_{i=0}^{n-1} f^{-i}\tilde{\mathfrak{A}}$  contained in any element of  $\bigvee_{i=0}^{n-1} f^{-i}\mathfrak{A}$ . We claim that  $a_n \leq na_1$ . Fix  $A \in \bigvee_{i=0}^{n-1} f^{-i}\mathfrak{A}$ .  $fA \subset B$  for some  $B \in \bigvee_{i=0}^{n-1} f^{-i}\mathfrak{A}$ .  $B$  contains  $\leq a_n$  elements of  $\bigvee_{i=0}^{n-1} f^{-i}\tilde{\mathfrak{A}}$ .  $A$  meets  $\leq a_1$  elements of  $\tilde{\mathfrak{A}}$ . Since  $f|A$  is monotonic,  $\text{card}(\bigvee_{i=0}^n f^{-i}\tilde{\mathfrak{A}}|A) \leq a_1 + a_n$ . This shows  $\text{card} \bigvee_{i=0}^{n-1} f^{-i}\tilde{\mathfrak{A}} \leq na_1 \text{card} \bigvee_{i=0}^{n-1} f^{-i}\mathfrak{A}$ .  $\square$

**THEOREM 2.4.** *Suppose on each open interval  $A_i$ ,  $f$  is  $C^1$  and  $f' \neq 0$ . Assume also that  $h(f) > 0$ . Then given any  $\epsilon > 0$ , there is a compact hyperbolic  $f$ -invariant set  $\Lambda \subset [0, 1]$  such that*

- (1)  $f|_\Lambda \sim \text{ssft}$ , and
- (2)  $h(f|_\Lambda) > h(f) - \epsilon$ .

For any subset  $\mathfrak{B} \subset \mathfrak{A}$  and  $B \in \mathfrak{B}$  let  $\alpha(n, B, \mathfrak{B}) = \text{card}(\bigvee_{i=0}^{n-1} f^{-i}\mathfrak{B}|B)$ . Then by Proposition 2.1,  $h(f) = \lim_{n \rightarrow \infty} (1/n) \log \sum_{A \in \mathfrak{A}} \alpha(n, A, \mathfrak{A})$ . Let  $\mathfrak{E} = \{A \in \mathfrak{A} : \limsup_n (1/n) \log \alpha(n, A, \mathfrak{A}) = h(f)\}$ .  $\mathfrak{E} \neq \emptyset$ .

The next two lemmas are borrowed from [22]. We include their proofs for completeness.

**LEMMA 2.5.** *For all  $A \in \mathfrak{E}$ ,  $\limsup_n (1/n) \log \alpha(n, A, \mathfrak{E}) = h(f)$ .*

PROOF. Fix  $A \in \mathfrak{E}$ . Observe that

$$\alpha(n, A, \mathfrak{A}) \leq \sum_{k=1}^{n-1} \alpha(k, A, \mathfrak{E}) \cdot \left[ \sum_{B \notin \mathfrak{E}} \alpha(n-k, B, \mathfrak{A}) \right].$$

Fix  $B \in \mathfrak{A} \setminus \mathfrak{E}$  with

$$\limsup_n \frac{1}{n} \log \sum_{k=1}^{n-1} \alpha(k, A, \mathfrak{E}) \alpha(n-k, B, \mathfrak{A}) \geq h(f). \quad (*)$$

Suppose our assertion was false. Fix  $u$  with

$$\limsup_n \frac{1}{n} \log \alpha(n, A, \mathfrak{E}), \quad \limsup_n \frac{1}{n} \log \alpha(n, B, \mathfrak{A}) < u < h(f),$$

i.e. there is an  $M > 0$  such that  $\alpha(n, A, \mathfrak{E}), \alpha(n, B, \mathfrak{A}) < Me^{nu}$  for all  $n \in \mathbb{Z}^+$ .

Putting these back in (\*), we have

$$h(f) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log n M^2 e^{nu} = u,$$

a contradiction.  $\square$

For  $A, B \in \mathfrak{A}$  let  $\gamma(A, B, n) = \text{card}\{E \in \bigvee_{i=0}^{n-1} f^{-i}\mathfrak{A} : E \subset A, f^{n+1}E \supset B\}$ .

LEMMA 2.6. *If  $h(f) > 3$ , then there exists  $A_0 \in \mathfrak{A}$  with*

$$\limsup_n (1/n) \log \gamma(A_0, A_0, n) = h(f).$$

PROOF. Fix  $A \in \mathfrak{E}$  and fix  $u$  with  $\log 3 < u < h(f)$ . There are arbitrarily large  $n$ 's with  $(1/n) \log \alpha(n, A, \mathfrak{E}) > u$  and  $\alpha(n+1, A, \mathfrak{E}) > 3u$ . If  $E \in \bigvee_{i=0}^{n-1} f^{-i}\mathfrak{A}$  and  $f^{n+1}E$  meets  $r$  elements of  $\mathfrak{E}$ , then  $f^{n+1}E$  contains at least  $r-2$  of them. Thus

$$\sum_{B \in \mathfrak{E}} \gamma(A, B, n) \geq \alpha(n+1, A, \mathfrak{E}) - 2\alpha(n, A, \mathfrak{E}) \geq \alpha(n, A, \mathfrak{E}).$$

Define  $\phi: \mathfrak{E} \hookrightarrow \mathfrak{E}$  in such a way that  $\phi A = B$  for some  $B$  with

$$\limsup_n (1/n) \log \gamma(A, B, n) \geq u.$$

Being a map of a finite set into itself,  $\phi$  has a periodic point  $A_0$ , say with period  $m$ . Thus for any  $v < u$ , there are arbitrarily large numbers  $n_1, n_2, \dots, n_m$  with  $\gamma(\phi^i A_0, \phi^{i+1} A_0, n_i) \geq e^{n_i v}$  so that

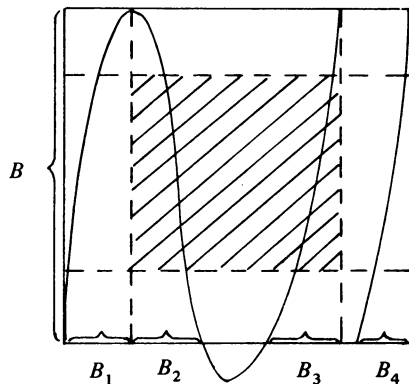
$$\gamma\left(A_0, A_0, \sum_1^m n_i\right) \geq \prod_1^m \gamma(\phi^i A_0, \phi^{i+1} A_0, n_i) \geq e^{(\sum n_i)v}. \quad \square$$

LEMMA 2.7. *Given  $\varepsilon > 0$ , there is a ssft  $\Sigma \subset \Sigma(f)$  such that  $\pi_2|_{\pi_1^{-1}\Sigma}$  is an embedding and  $h(\sigma|\Sigma) > h(\sigma|\Sigma(f)) - \varepsilon$ .*

PROOF. For some  $l > 0$ ,  $h(f^l) > \log 3$ . Let  $B \in \bigvee_{i=0}^{l-1} f^{-i}\mathfrak{A}$  be the  $A_0$  in Lemma 2.6. That is, for some large  $n$ , there are intervals  $B_1, \dots, B_M$  (appearing in that order) with  $B_j \subset B$ ,  $B_j \in \bigvee_{i=0}^{(n+1)l-1} f^{-i}$ ,  $f^{nl} B_j = B$  and  $(1/n) \log M > h(f^l) - \varepsilon/2$ . Fix  $n$ . Let  $C = \bigcup_{j=0}^{nl-1} f^j \{x \in [0, 1] : \text{for all } k \geq 0, f^{knl} x \in \bar{B}_i \text{ for some } i, 1 \leq i \leq M\}$ . Then  $h(f|C) = (1/nl) \log(M-2) > h(f) - \varepsilon$ . Let  $\Sigma$  be the ssft defined by: each  $2nl$ -block  $(x_0, \dots, x_{2nl-1})$  is admissible if and only if it appears in elements

of  $\Sigma(f)$  and no  $c_j \in \bigcap_{x=0}^{2n-1} f^{-i} \bar{A}_x$ . Since the images of  $C$  miss all the  $c_j$ 's,  $C \subset \pi_2 \pi_1^{-1} \Sigma$ .  $\pi_2$  is not 1-1 over  $x \in [0, 1]$  exactly if  $f^i x = c_j$  for some  $i \geq 0$  and some  $j$ . Such an  $x$  is not in  $\pi_2 \pi_1^{-1} \Sigma$ .  $\square$

graph of  $f^{nl}$



LEMMA 2.8. Let  $\Sigma$  be a ssft and let  $\epsilon > 0$  be given. Then for any fixed  $\mathbf{a} \in \Sigma$ , there is a ssft  $\tilde{\Sigma} \subset \Sigma$  with  $\mathbf{a} \notin \tilde{\Sigma}$  and  $h(\sigma|\tilde{\Sigma}) > h(\sigma|\Sigma) - \epsilon$ .

PROOF. We may assume  $\Sigma$  is topologically transitive or even mixing. (See for instance [13].) Let  $\mathcal{O}_+(\mathbf{a})$  denote the forward orbit of  $\mathbf{a}$ .

Case 1.  $\overline{\mathcal{O}_+(\mathbf{a})} \neq \Sigma$ . Then  $\sigma \overline{\mathcal{O}_+(\mathbf{a})} \subset \overline{\mathcal{O}_+(\mathbf{a})}$  and Lemma 1.2 gives the desired result.

Case 2.  $\overline{\mathcal{O}_+(\mathbf{a})} = \Sigma$ . Since  $\Sigma$  is mixing, there are many periodic points. Fix one, say  $\mathbf{b}$ . Choose a ssft  $\tilde{\Sigma}$  missing  $\mathbf{b}$  as in Lemma 1.2. There is a neighborhood  $U \supset \mathbf{b}$  such that  $U \cap \tilde{\Sigma} = \emptyset$ . Since  $\sigma^i \mathbf{a} \in U$  for some  $i \geq 0$ ,  $\mathbf{a} \notin \tilde{\Sigma}$ .  $\square$

LEMMA 2.9. Let  $\Sigma \subset \Sigma(f)$  be the ssft in Lemma 2.7. Let  $\epsilon, \delta > 0$  be given. Then there is a ssft  $\tilde{\Sigma} \subset \Sigma$  and  $N \in \mathbb{Z}^+$  such that

- (1)  $h(\sigma|\tilde{\Sigma}) > h(\sigma|\Sigma) - \epsilon$ , and
- (2) if  $C$  is any  $N$ -cylinder set of  $\tilde{\Sigma}$ , then  $|\pi_2 \pi_1^{-1} C| \leq 4\delta$ .

PROOF. There are only finitely many points  $\mathbf{x} \in \Sigma(f)$  with  $|\pi_2 \pi_1^{-1} \mathbf{x}| > \delta$ . Using Lemma 2.8, choose a ssft  $\tilde{\Sigma} \subset \Sigma$  with  $h(\sigma|\tilde{\Sigma}) > h(\sigma|\Sigma) - \epsilon$  and such that for all  $\mathbf{x} \in \tilde{\Sigma}$ ,  $|\pi_2 \pi_1^{-1} \mathbf{x}| < \delta$ . Now there is an  $N_0$  such that if  $\mathbf{x} \in \Sigma(f)$  has  $|\pi_2 \pi_1^{-1} \mathbf{x}| > \delta$ , then the cylinder set  $(x_0, \dots, x_{N_0}) \cap \tilde{\Sigma} = \emptyset$ .

Let  $0 = z_0 < \dots < z_r = 1$  be a partition of  $[0, 1]$  with  $\delta < z_i - z_{i-1} < 2\delta$ . For each  $i$ , if there is a  $j \geq 0$  such that  $f^j z_{i-1} \in \bar{A}_k, f^j z_i \in \bar{A}_l$  with  $k \neq l$ , then let  $n_i$  be the smallest such  $j$ . Let  $N = 1 + \max(N_0, n_i)$ . We claim that  $\pi_2 \pi_1^{-1}$  (any  $N$ -cylinder set of  $\tilde{\Sigma}$ )  $\leq 4\delta$ .

Let  $\mathbf{x}, \mathbf{y} \in \tilde{\Sigma}$  and  $x \in \pi_2 \pi_1^{-1} \mathbf{x}, y \in \pi_2 \pi_1^{-1} \mathbf{y}$  have  $|x - y| > 4\delta$ . For some  $i$ ,  $x < z_{i-1} < z_i < y$ . If for all  $j$ ,  $f^j z_{i-1}$  and  $f^j z_i$  lie in the closure of the same element of  $\mathcal{A}$ , then  $|\pi_2 \pi_1^{-1} \pi_1 \pi_2^{-1}(z_{i-1}, z_i)| \geq \delta$  and we have  $(x_0, \dots, x_{N_0}) \neq (y_0, \dots, y_{N_0})$ . If on the other hand  $z_{i-1}$  and  $z_i$  admit distinct symbolic representations, then  $(x_0, \dots, x_n) \neq (y_0, \dots, y_n)$ . Thus  $\mathbf{x}, \mathbf{y}$  lie in distinct  $N$ -cylinder sets.  $\square$



LEMMA 2.10. Let  $\tilde{\Sigma}$  be a mixing ssft. Let  $\mu$  be the ergodic measure on  $\tilde{\Sigma}$  with  $h_\mu(\sigma) = h(\sigma|\tilde{\Sigma})$ . Let  $\phi: \tilde{\Sigma} \rightarrow \mathbf{R}$  be a step function depending only on the 0th coordinate. Then given  $\varepsilon > 0$ , there is a ssft  $\Lambda \subset \tilde{\Sigma}$  and  $M \in \mathbf{Z}^+$  such that

- (1)  $h(\sigma|\Lambda) > h(\sigma|\tilde{\Sigma}) - \varepsilon$ , and
- (2)  $|(1/M) \sum_{i=0}^{M-1} \phi \sigma^i \mathbf{x} - \int \phi d\mu| < \varepsilon$  for all  $\mathbf{x} \in \Lambda$ .

PROOF. Let  $\mathcal{P}$  be the 0-time partition of  $\tilde{\Sigma}$ . For some small  $\alpha > 0$ , let  $\mathcal{N}_n = \{P \in \bigvee_{i=0}^{n-1} \sigma^{-i} \mathcal{P} : |(1/n) \sum_{i=0}^{n-1} \phi \sigma^i \mathbf{x} - \int \phi d\mu| < \alpha \text{ for } \mathbf{x} \in P\}$ . We first show that  $\lim_{n \rightarrow \infty} (1/n) \log \text{card } \mathcal{N}_n$  exists and equals  $h = h(\sigma)$ . By the ergodic theorem,  $(1/n) \sum_{i=0}^{n-1} \phi \sigma^i \rightarrow \int \phi d\mu$  a.e. By Egorov's theorem, there is a set  $A \subset \tilde{\Sigma}$  with  $\mu A > 0$  such that on  $A$  convergence is uniform. Since  $\mu$  is the maximal measure of an ergodic ssft, there exists  $c > 0$  such that  $A$  meets  $> c2^{nh}$  elements of  $\bigvee_{i=0}^{n-1} \sigma^{-i} \mathcal{P}$ . (We assume logarithm is to base 2.) Thus for large  $n$ , we have  $\text{card } \mathcal{N}_n \geq c2^{nh}$  so that  $\liminf_{n \rightarrow \infty} (1/n) \log \text{card } \mathcal{N}_n \geq h$ .

Now let  $\mathcal{N}_{n,r}$  be the collection of periodic  $n$ -strings in  $\tilde{\Sigma}$  defined by  $\{(x_0, \dots, x_{n-1}) : x_0 = r \text{ and } |(1/n) \sum_{i=0}^{n-1} \phi x_i - \int \phi d\mu| < \varepsilon/2\}$  (a slight abuse of notation). Since  $\tilde{\Sigma}$  is mixing,  $\lim_{n \rightarrow \infty} (1/n) \log \text{card } \mathcal{N}_{n,r}$  exists and equals  $h$ . Suppose for now that for certain arbitrarily large numbers  $n$ , there is a symbol  $r$  and a periodic  $n$ -string  $\mathbf{a}$  in some element of  $\tilde{\Sigma}$  with  $\mathbf{a} = a_0 \cdots a_{n-1}$ ,  $a_0 = a_n = r$  and  $a_i \neq r$  for  $0 < i < n$ . Let  $A = \{\mathbf{x} \in \tilde{\Sigma} : \mathbf{x} \text{ has the following form } \mathbf{a} \mathbf{b}_{i_1} \mathbf{b}_{i_2} \cdots \mathbf{b}_{i_{m-1}} \mathbf{a} \mathbf{b}_{i_m} \cdots \mathbf{b}_{i_{2(m-1)}} \mathbf{a} \cdots \text{ where } \mathbf{b}_{i_j} \in \mathcal{N}_{n,r}, \mathbf{b}_{i_j} \neq \mathbf{a} \forall i_j\}$ . Let  $\Lambda = A \cup \sigma A \cup \cdots \cup \sigma^{mn} A$ . Since there is absolutely no ambiguity about the position of  $\mathbf{a}$  in each symbol sequence in  $\Lambda$ , the union is disjoint and  $\Lambda$  is precisely the ssft defined by "a  $2mn$ -string is admissible iff it occurs in  $\Lambda$ ". If  $m$  and  $n$  are large enough, then

$$\left| \frac{1}{mn} \sum_{i=0}^{mn-1} \phi \sigma^i \mathbf{x} - \int \phi d\mu \right| < \varepsilon \quad \forall \mathbf{x} \in \Lambda$$

and

$$h(\sigma|\Lambda) \geq \frac{m-1}{m} \frac{1}{n} \log(\text{card } \mathcal{N}_{n,r} - 1) > h - \varepsilon.$$

$\Lambda$  therefore is the ssft we need.

We have yet to justify the existence of  $\mathbf{a}$ . Since  $\tilde{\Sigma}$  is not just a single periodic orbit, there exist symbols  $r$  and  $s$  and a periodic string  $\mathbf{s} = s_0 \cdots s_l$  in  $\tilde{\Sigma}$  with  $s_0 = s_l = s$ ,  $s_i \neq r$  for all  $i$ . Let  $\mathbf{a}_{rs}$  be the shortest path from  $r$  to  $s$  and  $\mathbf{a}_{sr}$  the shortest path from  $s$  to  $r$ . Then  $\mathbf{a} = \mathbf{a}_{rs} \mathbf{s} \cdots \mathbf{s} \mathbf{a}_{sr}$ , with as many  $\mathbf{s}$ 's as necessary is the string we have claimed to exist.  $\square$

LEMMA 2.11. Let  $X \subset [0, 1]$  be a compact  $f$ -invariant set. Suppose  $X \cap \{c_0, \dots, c_q\} = \emptyset$  and  $|f'x| \leq \lambda$  for some  $\lambda \geq 1$  for all  $x \in X$ . Then  $h(f|X) \leq \log \lambda$ .

PROOF. For small  $\varepsilon > 0$ , there exists  $\delta > 0$  such that for  $x \in X$ ,  $y \in [0, 1]$ ,  $|x - y| \leq \delta \Rightarrow |f'x - f'y| < \varepsilon$ . Then for all  $x, y \in X$ ,  $|x - y| \leq \delta \Rightarrow |fx - fy| \leq (\lambda + \varepsilon)|x - y|$ . A spanning set argument gives  $h(f|X) \leq \log(\lambda + \varepsilon)$ .  $\square$

PROOF OF THEOREM 2.4. Assume  $\varepsilon < (1/8)h(f)$ . Let  $\Sigma \subset \Sigma(f)$  be the one chosen in Lemma 2.7. That is,  $h(\sigma|\Sigma) > h(f) - \varepsilon$  and  $\pi_2 \pi_1^{-1} \Sigma \subset [0, 1]$  is essentially a ssft

except that some points are intervals. So first of all let us get rid of at least the longer intervals. We appeal to Lemmas 2.8 and 2.9. Since  $\log|f'|$  is continuous on  $\pi_2\pi_1^{-1}\Sigma$  there exists  $\delta > 0$  such that for  $x, y \in \Sigma$ ,

$$|x - y| < 4\delta \Rightarrow |\log|f'|x| - \log|f'|y|| < \epsilon.$$

Lemma 2.9 gives a ssft  $\tilde{\Sigma} \subset \Sigma$  with  $h(\sigma|\tilde{\Sigma}) > h(f) - 2\epsilon$  and a positive integer  $N$  with  $\sup\{|\log|f'|x| - \log|f'|y| : x, y \in \pi_2\pi_1^{-1}(\text{same } N\text{-cylinder of } \tilde{\Sigma})\} < \epsilon$ . Define  $\phi: \tilde{\Sigma} \rightarrow \mathbf{R}$  by

$$\phi x = \min\{\log|f'|x| : x \in \pi_2\pi_1^{-1}((x_0, \dots, x_{N-1}) \cap \tilde{\Sigma})\}.$$

Renaming symbols so that  $\phi$  is constant on 1-cylinder sets, we now apply Lemma 2.10.

All this tells us is that for some large  $M$ ,  $|(1/M) \sum_{i=0}^{M-1} \phi \sigma^i x - \int \phi d\mu| < \epsilon$  for all  $x \in \Lambda$  or,

$$\left| \frac{1}{M} \log|Df_x^M| - \int \phi d\mu \right| < 2\epsilon \quad \text{for } x \in \pi_2\pi_1^{-1}\Lambda.$$

Since

$$\begin{aligned} h(\sigma^M|\Lambda) &= h(f^M|\pi_2\pi_1^{-1}\Lambda) \\ &\leq \max_{x \in \pi_2\pi_1^{-1}\Lambda} \log|Df_x^M| \quad \text{by Lemma 2.11} \\ &\leq M \left( \int \phi d\mu + 2\epsilon \right), \end{aligned}$$

we have

$$h(f) - 3\epsilon < h(\sigma|\Lambda) \leq \int \phi d\mu + 2\epsilon$$

and finally

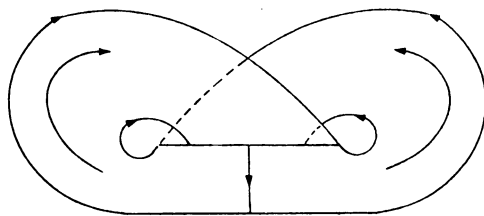
$$\frac{1}{M} \log|Df_x^M| \geq \int \phi d\mu - 2\epsilon > h(f) - 7\epsilon > 0$$

for all  $x \in \pi_2\pi_1^{-1}\Lambda$ . We have thus produced a hyperbolic set in  $[0, 1]$ . Notice that  $\pi_1^{-1}\Lambda$  contains no intervals, for otherwise its corresponding part in  $[0, 1]$  would grow indefinitely in length. This completes the proof.  $\square$

**3. A. THE LORENZ ATTRACTOR.** This interesting non-Axiom A example has its origin in the two-dimensional convection problem. It is the flow in  $\mathbf{R}^3$  generated by the system of O.D.E.

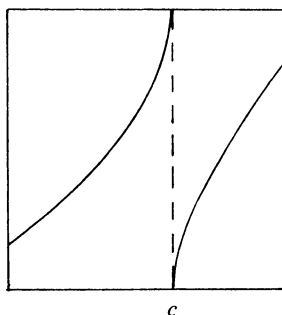
$$\begin{aligned} \dot{x} &= -10x + 10y, \\ \dot{y} &= 28x - y - xz, \\ \dot{z} &= -\frac{8}{3}z + xy. \end{aligned}$$

Following Williams [32], we can picture the flow as a (pinched) inverse limit of a semiflow on a branched 2-manifold as shown below. For details see [15], [16], [20], [32].



The Poincaré map of this semiflow is a map of the interval that looks like the figure below. Theorem 2.4 applies. The inverse limit of the one-sided shift is hyperbolic as long as we stay away from  $c$ . Thus we have

**THEOREM 3.1.** *The Poincaré map of the Lorenz attractor flow is a limit of hyperbolic ssft.*



**B. THE ABRAHAM-SMALE EXAMPLES.** This diffeomorphism is a skew-product  $F: S^2 \times T^2 \hookrightarrow$  given by  $F(x, y) = (gx, f_x y)$  where  $g: S^2 \hookrightarrow$  has a horseshoe  $H$ , a sink and a source. We think of  $H$  as  $\Sigma_2 = \prod_{-\infty}^{\infty} \{0, 1\}$ . For  $x \in \Sigma_2$  with  $x_0 = 0$ , define

$$f_x = f_0 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

the toral Anosov. For  $x \in \Sigma_2$  with  $x_0 = 1$ , let  $f_x = f_1$  be a DA (derived from Anosov) map [31]. For  $x \notin H$ ,  $f_x$  is defined so that we have an isotopy from  $f_0$  to  $f_1$ .

We put a few assumptions on  $f_1$ . Let  $\mathcal{R}$  be the usual Markov partition for  $f_0$  on  $T^2$  and  $\Sigma_A$  be its associated ssft. In eigenvector coordinates let

$$f_0 = \begin{pmatrix} u & 0 \\ 0 & s \end{pmatrix}, \quad u > 1.$$

Write  $f_1 = \alpha \circ f_0$ . We assume

- (1)  $f_1$  is obtained from  $f_0$  by pushing along the stable foliation of  $f_0$ ,
- (2) for each element  $R \in \mathcal{R}$ , as a set  $f_0 R = f_1 R$ , and
- (3)  $f_1$  has a weaker expansion in the  $E^s(f_0)$  direction than  $f_0$  does in the  $E^u(f_0)$  direction, i.e., if

$$Df_1(x, y) = \begin{pmatrix} 1 & 0 \\ a(x, y) & b(x, y) \end{pmatrix} \begin{pmatrix} u & 0 \\ 0 & s \end{pmatrix} = \begin{pmatrix} u & 0 \\ a(x, y)u & b(x, y)s \end{pmatrix},$$

then  $1 < |b(x, y)|s < u$ .

THEOREM 3.2. *With the above assumptions,  $F$  is a limit of hyperbolic ssft.*

Assumption (2) guarantees the following: take any random composition of  $f_0$  and  $f_1$ . Pretend it is done on the same torus. With respect to  $\mathfrak{R}$  exactly the same symbol sequences occur as if one were iterating  $f_0$  alone. Let  $\Sigma_A^+$  be the one-sided ssft corresponding to  $\Sigma_A$ . A model for the nonwandering set of  $F$ ,  $H \times T^2$ , is  $\Sigma_2 \times \Sigma_A^+ \times [0, 1]$ . The first coordinate tells us which toral fiber a point  $x$  is in; this together with the second coordinate determines the location of  $x$  up to the line segment  $W^s(f_0, x) \cap R$  for the appropriate  $R \in \mathfrak{R}$ , and the third coordinate tells where  $x$  is on this line segment.

Defining  $\phi: \Sigma_2 \times \Sigma_A^+ \times [0, 1] \hookrightarrow$  naturally we get a semiconjugacy:

$$\begin{array}{ccc} \Sigma_2 \times \Sigma_A^+ \times [0, 1] & \xrightarrow{\phi} & \Sigma_2 \times \Sigma_A^+ \times [0, 1] \\ \pi \downarrow & & \downarrow \pi \\ H \times T^2 & \xrightarrow{F} & H \times T^2 \end{array}$$

where  $\phi$  is continuous but not one-to-one. This set-up is used in the rest of the discussion.

We compute  $h(F)$ . Consider

$$\begin{array}{ccc} \Sigma_2 \times \Sigma_A^+ \times [0, 1] & \xrightarrow{\phi} & \Sigma_2 \times \Sigma_A^+ \times [0, 1] \\ \rho \downarrow & & \downarrow \rho \\ \Sigma_2 \times \Sigma_A^+ & \xrightarrow{\sigma^+} & \Sigma_2 \times \Sigma_A^+ \end{array}$$

where  $\sigma^+$  is the product of shift operators and  $\rho$  is the projection map. We have  $h(F) = h(\phi)$  because  $\pi$  is a finite-to-one map. Also,  $h(\phi) = h(\sigma^+)$  because  $\rho$ -fibers are intervals and  $\phi$  maps such intervals monotonically into other intervals [5]. Thus  $h(F) = \log 2 + h(\sigma|\Sigma_A)$ .

PROOF OF THEOREM 3.2. Let  $\mathfrak{S}_n = \{n\text{-strings in } \Sigma_2 \text{ starting with 0 and containing at least half 0's}\}$ . We know that  $\lim_{n \rightarrow \infty} (1/n) \log |\mathfrak{S}_n| = \log 2$ . Fix some large  $N$ . Let  $\Lambda = \{x \in \Sigma_2: \text{for all } k \in \mathbb{Z}, (x_{kN}, \dots, x_{(k+1)N-1}) \in \mathfrak{S}_N\}$ . We show that  $F^N|_{\Lambda \times T^2}$  is hyperbolic. This implies that hyperbolicity of  $F$  on  $(\bigcup_{i=0}^{N-1} \sigma^i \Lambda) \times T^2$ . An argument similar to (and actually simpler than) that in Lemma 2.10 gives a ssft  $\tilde{\Lambda} \subset \bigcup_{i=0}^{N-1} \sigma^i \Lambda$  with entropy near that of  $F$ . Finally from Theorem 1.1 we obtain a ssft  $\tilde{\Lambda} \times T^2$  with desired entropy.

It remains to check the hyperbolicity of  $f^N|_{\Lambda \times T^2}$ . We adopt the following notation: if  $A = [A_{ij}]$  and  $B = [B_{ij}]$  are matrices of the same dimension, we say  $A \leq B$  if  $|A_{ij}| \leq |B_{ij}|$  for all  $i, j$ . Suppose  $D\alpha \leq \begin{pmatrix} 1 & 0 \\ a & b \end{pmatrix}$  with  $bs < u$ . For  $i = 1, \dots, N$ , let  $X_i$  be either  $\begin{pmatrix} u & 0 \\ 0 & s \end{pmatrix}$  or  $\begin{pmatrix} u & 0 \\ a & bs \end{pmatrix}$  and let

$$X_n X_{n-1} \cdots X_1 = \begin{pmatrix} u^n & 0 \\ \phi_n & s^n b^{in} \end{pmatrix}.$$

Then  $\phi_{n+1} \leq au^{n+1} + sb\phi_n$  and so

$$\phi_N \leq a[u^N + (bs)u^{N-1} + \cdots + (bs)^{N-1}u] \leq au^N / (1 - bs/u).$$

For  $(x, y) \in \Lambda \times T^2$ ,

$$D(f_{g^{N_x}} \circ \cdots \circ f_{g^x} \circ f_x) \leq \begin{pmatrix} u^N & 0 \\ au^N / \left(1 - \frac{bs}{u}\right) & s^N b^{N/2} \end{pmatrix}.$$

The criteria for hyperbolicity in [18] are satisfied.

#### REFERENCES

1. R. Abraham and S. Smale, *Nongenericity of  $\Omega$ -stability*, Proc. Sympos. Pure Math., vol. 14, Amer. Math. Soc., Providence, R. I., 1970, pp. 5–9.
2. L. M. Abramov, *On the entropy of a flow*, Trans. Amer. Math. Soc. **49** (1966), 167–170.
3. R. L. Adler, A. G. Konheim and M. H. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc. **114** (1965), 309–319.
4. R. Bowen, *Topological entropy and Axiom A*, Proc. Sympos. Pure Math., vol. 14, Amer. Math. Soc., Providence, R. I., 1970, pp. 23–42.
5. ———, *Entropy for group endomorphisms and homogeneous spaces*, Trans. Amer. Math. Soc. **153** (1971), 401–414.
6. ———, *Markov partitions for Axiom A diffeomorphisms*, Amer. J. Math. **92** (1970), 725–747.
7. ———, *Symbolic dynamics for hyperbolic flows*, Amer. J. Math. **95** (1973), 429–460.
8. ———, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Math., vol. 470, Springer-Verlag, Berlin and New York, 1975.
9. ———, *On Axiom A diffeomorphisms*, CBMS Regional Conf. Ser. Math., no. 35, Amer. Math. Soc., Providence, R. I., 1978.
10. R. Bowen and D. Ruelle, *The ergodic theory of Axiom A flows*, Invent. Math. **29** (1975), 181–202.
11. A. Cooks and P. Roberts, *The Rikitake two-disc dynamo system*, Proc. Cambridge Philos. Soc. **68** (1970), 547–569.
12. E. I. Dinaburg, *On the relations among various entropy characteristics of dynamical systems*, Math. USSR Izv. **5** (1971), 337–378 (see p. 340 in particular).
13. F. Gantmacher, *Matrix theory*, Vol. 2, Chelsea, New York, 1959.
14. L. W. Goodwyn, *Topological entropy bounds measure-theoretic entropy*, Proc. Amer. Math. Soc. **23** (1969), 679–688.
15. J. Guckenheimer, *A strange, strange attractor*, Lecture Notes in Appl. Math. Sci., vol. 19, Springer-Verlag, Berlin and New York, 1976, pp. 368–391.
16. J. Guckenheimer and R. F. Williams, *Structural stability of the Lorenz attractor*, Inst. Haute Etude Sci. Publ. Math., no. 50, 1980.
17. M. Hénon, *A 2-dimensional mapping with a strange attractor*, Comm. Math. Phys. **50** (1976), 69–78.
18. M. W. Hirsch and C. C. Pugh, *Stable manifolds and hyperbolic sets*, Proc. Sympos. Pure Math., vol. 14, Amer. Math. Soc., Providence, R. I., 1970, pp. 133–164.
19. M. W. Hirsch, C. C. Pugh and M. Shub, *Invariant manifolds*, Lecture Notes in Math., vol. 583, Springer-Verlag, Berlin and New York, 1977, §8, p. 136.
20. E. N. Lorenz, *Deterministic nonperiodic flows*, J. Atmospheric Sci. **20** (1963), 130–141.
21. J. Milnor and W. Thurston, *On iterated maps of the interval* (to appear).
22. M. Misiurewicz and W. Szlenk, *Entropy of piecewise monotone mappings* (to appear).
23. S. Newhouse, *Nondensity of Axiom A(a) on  $S^2$* , Proc. Sympos. Pure Math., vol. 14, Amer. Math. Soc., Providence, R. I., 1970, pp. 191–203.
24. ———, *Diffeomorphisms with infinitely many sinks*, Topology **13** (1974), 9–18.
25. W. Parry, *Intrinsic Markov chains*, Trans. Amer. Math. Soc. **112** (1964), 55–66.
26. ———, *Symbolic dynamics and transformations of the interval*, Trans. Amer. Math. Soc. **122** (1966), 368–378.

27. D. Ruelle, *Statistical mechanics on a compact set with  $\mathbb{Z}^r$  action satisfying expansiveness and specification*, Trans. Amer. Math. Soc. **185** (1973), 237–251.
28. C. Simon, *A 3-dimensional Abraham-Smale example*, Ph.D. Thesis, Northwestern University, 1970.
29. S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817.
30. P. Walters, *A variational principle for the pressure of continuous transformations*, Amer. J. Math. **97** (1976), 937–971.
31. R. F. Williams, *The “DA” maps of Smale and structural stability*, Proc. Sympos. Pure Math., vol. 14, Amer. Math. Soc., Providence, R. I., 1970, pp. 329–334.
32. ———, *The structure of Lorenz attractors*, Inst. Hautes Études Sci. Publ. Math., no. 50, 1980.

DEPARTMENT OF MATHEMATICS, NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS 60201