TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 154, February 1971

PERIODIC POINTS AND MEASURES FOR AXIOM A DIFFEOMORPHISMS

BY
RUFUS BOWEN

1. **Introduction.** We shall study the distribution of periodic points for a class of diffeomorphisms defined by Smale [16, §I.6].

We recall some of the definitions. Let $f: M \to M$ be a diffeomorphism of a compact manifold. A point $x \in M$ is *wandering* under f if it has a neighbourhood U such that $U \cap \bigcup_{m \neq 0} f^m(U) = \varnothing$; the set of other (i.e. nonwandering points) forms the *nonwandering set* $\Omega(f)$ which is closed and f-invariant. One sees that all periodic points of f are in $\Omega(f)$ and that any finite f-invariant measure on M has its support in $\Omega(f)$. A closed f-invariant subset Λ of M is *hyperbolic* under f if the tangent bundle of M restricted to Λ, $T_\Lambda(M)$, has a continuous splitting $T_\Lambda(M) = E^s + E^u$ which is invariant under Df and such that $Df: E^s \to E^s$ is contracting and $Df: E^u \to E^u$ is expanding (see [16, p. 758] for the meaning of these terms). *f satisfies Axiom* A if

(Aa) $\Omega(f)$ is hyperbolic and

(Ab) the periodic points of f are dense in $\Omega(f)$.

Smale's Spectral Decomposition Theorem [16, p. 777] states that for such an f we can write $\Omega(f) = \Omega_1 \cup \cdots \cup \Omega_r$ where the Ω_i are disjoint closed f-invariant sets and $f|\Omega_i$ is topologically transitive (the Ω_i are called *basic* sets). Our main result is that the periodic points of $f|\Omega_i$ have a definite limiting distribution as the period becomes large; this distribution is given by a measure μ_f on Ω_i. In the algebraic case μ_f turns out to be Haar measure.

We show that μ_f is ergodic, positive on open sets and zero on points (unless Ω_i is finite). In a subsequent paper [7] it is shown that $(f|\Omega_i, \mu_f)$ is a K-automorphism in the C-dense case (in fact that it is isomorphic to a Markov chain) and that μ_f is the unique invariant normalized Borel measure on Ω_i which maximizes entropy.

The Russian school has done much work on the measure theoretic aspects of Anosov diffeomorphisms (i.e. all of M hyperbolic under f); as a sampling we refer the reader to the papers [2], [14] and [15]. We also mention the papers [3], [9] and [11] where various measures are constructed for expanding maps; our methods are easily modified to give results along this direction also.

We now sketch our construction of μ_f. First we decompose $\Omega_i = X_1 \cup \cdots \cup X_m$ into disjoint closed pieces X_j such that $f(X_j) = X_{j+1}$ and $f^m|X_j: X_j \to X_j$ is *C-dense* for all $1 \leq j \leq m$. We do not define C-density here but it implies topological mixing

Received by the editors June 27, 1969.

and the existence of periodic points of all sufficiently large periods; for Markov chains this is the well-known decomposition into transitive pieces.

One then restricts attention to the C-dense case; i.e. assume $f\colon \Omega_i \to \Omega_i$ is C-dense. What we want is a measure μ_f such that (letting $N_n(E)$ be the number of fixed points of f^n lying in E)

$$N_n(E)/N_n(\Omega_i) \to \mu_f(E)$$

as $n \to \infty$ for many subsets E of Ω_i (we save precision for later). A priori we do not know that such a limit exists; using a diagonalization process we can choose sequences of integers $\{n_k\}$ and measures $\mu_{f,\{n_k\}}$ such that

$$N_{n_k}(E)/N_{n_k}(\Omega_i) \to \mu_{f,\{n_k\}}(E)$$

for many $E \subseteq \Omega_i$. We then show that all these measures $\mu_{f,\{n_k\}}$ are ergodic and equivalent; the Radon-Nikodym theorem tells us that they are all equal. When enough subsequences converge to a common limit, the sequence itself converges. Thus we get our desired $N_n(E)/N_n(\Omega_i) \to \mu_f(E)$.

Conversations with W. Parry, S. Smale, P. Walters and R. F. Williams were helpful in preparing this paper. The author wishes to thank the referees for many ideas which improved this paper.

2. **Axiom A* and C-density.** Let $g\colon M \to M$ be a diffeomorphism satisfying Smale's Axiom A. Let $X = \Omega(g) \subseteq M$ and $f = g|X$. Define, for $x \in X = \Omega(g)$ and $\delta > 0$,

$$W^s_\delta(x) = \{y \in X : d(f^n(x), f^n(y)) \leqq \delta \text{ for all } n \geqq 0\}.$$
$$W^u_\delta(x) = \{y \in X : d(f^n(x), f^n(y)) \leqq \delta \text{ for all } n \leqq 0\}.$$
$$W^s(x) = \{y \in X : d(f^n(x), f^n(y)) \to 0 \text{ as } n \to +\infty\}.$$
$$W^u(x) = \{y \in X : d(f^n(x), f^n(y)) \to 0 \text{ as } n \to -\infty\}.$$

Then (Smale [16, pp. 780–782] and Hirsch and Pugh [10]) the following are true:

A1. The periodic points of f are dense in X.

A2. For each $\delta > 0$ there is an $\varepsilon(\delta) > 0$ such that $W^s_\delta(x) \cap W^u_\delta(z) \neq \varnothing$ whenever $d(x, z) \leqq \varepsilon(\delta)$.

A3. There are $\delta^* > 0$, $0 < \lambda < 1$ and $c \geqq 1$ such that for all $n \geqq 0$,

$$d(f^n(x), f^n(y)) \leqq c\lambda^n d(x, y) \quad \text{if } y \in W^s_{\delta^*}(x)$$

and

$$d(f^{-n}(x), f^{-n}(y)) \leqq c\lambda^n d(x, y) \quad \text{if } y \in W^u_{\delta^*}(x).$$

The above statements are about f and do not refer to g or M. Any homeomorphism f of a compact metric space (X, d) we shall say *satisfies Axiom* A* provided that A1, A2, and A3 hold.

(2.1) *Standing hypothesis.* We shall assume throughout the remainder of the paper that $f\colon X \to X$ is a homeomorphism satisfying Axiom A*.

(2.2) *Easy facts.* (i) $f^n W^u(x) = W^u(f^n(x))$.

(ii) For $n \geqq 0, f^{-n} W^u_\delta(x) \subseteq W^u_\delta(f^{-n}(x))$.

(iii) If $y \in W_{\delta_1}(x)$, then $W_{\delta_2}(y) \subseteq W_{\delta_1+\delta_2}(x)$.

(iv) Let $f^m(x) = x$ and $\delta \leqq \delta^*$. Then $f^{m(k+1)} W^u_\delta(x) \supseteq f^{mk} W^u_\delta(x)$ and (by A3)

$$W^u(x) = \bigcup_{k=0}^{\infty} f^{mk} W^u_\delta(x).$$

The following fact is due to S. Smale and M. Shub:

(2.3) LEMMA [6]. *δ^* is an expansive constant for f (i.e. if $x \neq y$, then $d(f^n(x), f^n(y)) > \delta^*$ for some $n \in Z$).*

(2.4) LEMMA. *For any $\varepsilon > 0$ there is a $D(\varepsilon)$ so that $d(x, y) < \varepsilon$ whenever $d(f^n(x), f^n(y)) \leqq \delta^*$ for all $|n| \leqq D(\varepsilon)$.*

Proof. This is a property of expansive homeomorphisms [18].

(2.5) *Periodic point construction.* For any $\varepsilon > 0$ there are $\psi(\varepsilon) > 0$ and $R(\varepsilon)$ such that, if $m \geqq R(\varepsilon)$ and $d(f^m(y), y) \leqq \psi(\varepsilon)$, there is a point $z \in X$ with $f^m(z) = z$ and $d(f^k(z), f^k(y)) \leqq \varepsilon$ for all $0 \leqq k \leqq m$.

Proof. This is a translation of [6, Proposition 3.5] using [6, 3.4(h)].

(2.6) DEFINITION. f (satisfying Axiom A*) is C-dense if $W^u(p)$ is dense in X for every periodic point $p \in X$.

We permute ideas of Smale [16, pp. 780–782] to obtain

(2.7) C-DENSITY DECOMPOSITION THEOREM. *$X = X_1 \cup \cdots \cup X_m$ where the X_i are disjoint closed sets, $f(X_i) = X_{g(i)}$ where g is a permutation of $(1, \ldots, m)$, and $f^r: X_i \to X_i$ is C-dense when $g^r(i) = i$.*

Proof. For p a periodic point let $X(p) = \mathrm{Cl}\,(W^u(p))$.

(a) *$X(p)$ is open.*

Proof. Let $a = \varepsilon(\delta^*)$. We show that

$$X(p) \supset B_a(X(p)) = \{y \in X : d(y, X(p)) < a\}.$$

Since $X(p)$ is closed, it suffices to show that periodic $q \in B_a(X(p))$ are in $X(p)$ because of A1. Let $x \in W^u(p)$ with $d(x, q) < a$ and set $M = \mathrm{ord}\, p \cdot \mathrm{ord}\, q$. By A2 choose $z \in W^u_{\delta^*}(x) \cap W^s_{\delta^*}(q)$. Then $z \in W^u(p)$ and

$$d(f^{kM}(z), q) = d(f^{kM}(z), f^{kM}(q)) \to 0 \quad \text{as } k \to +\infty.$$

Since $f^{kM} W^u(p) \subset W^u(p)$, we get $q \in \mathrm{Cl}\,(W^u(p)) = X(p)$. (Note: We use 2.1 without explicit mention.)

(b) *$X(p) = X(q)$ or $X(p) \cap X(q) = \varnothing$.*

Proof. Suppose $z \in X(p) \cap X(q)$. By (a) $X(p)$ is a neighborhood of z and so there is a $w \in W^u(q) \cap X(p)$. Let $M = \mathrm{ord}_f\, p \cdot \mathrm{ord}_f\, q$. Then as $k \to +\infty, f^{-kM}(w) \to q$. But $f^{-M} X(p) = X(p)$ since $f^{-M} W^u(p) = W^u(p)$. Thus $q \in \mathrm{Cl}\,(X(p)) = X(p)$. By (a) we have $X(p) \supset W^u_a(q)$. Since

$$W^u(q) \subset \bigcup_{k=0}^{\infty} f^{kM} W^u_a(q)$$

and $f^{kM}X(p)=X(p)$, we get $W^u(q)\subset X(p)$. Hence $X(q)\subset X(p)$. Symmetrically $X(p)\subset X(q)$.

Now by compactness, let $X=X(p_1)\cup\cdots\cup X(p_m)$ with $X(p_i)\neq X(p_j)$ for $i\neq j$. Set $X_i=X(p_i)$ and define g by $f(p_i)\in X_{g(i)}$. That f is a homeomorphism and (c) below show that g is a permutation.

(c) $f(X_i)=X_{g(i)}$.

Proof. As f is a homeomorphism, $fX(p_i)=X(f(p_i))$ follows from $fW^u(p_i)=W^u(f(p_i))$. Since $f(p_i)\in X(f(p_i))\cap X(p_{g(i)})$, $X(f(p_i))=X(p_{g(i)})$ by (b).

(d) *If* $g^r(i)=i$, *then* $f^r\colon X_i\to X_i$ *is C-dense.*

Proof. Suppose $p\in X_i$ is periodic. It is an easy exercise to check that $W^u_f(p)=W^u_{f^r}(p)$. Note that $f^r\colon X\to X$ satisfies Axiom A* whenever $f\colon X\to X$ does.

(2.8) LEMMA. *Let* $f\colon X\to X$ *be C-dense and* $\alpha>0$. *Then there is an N such that* $f^mW^u_\alpha(x)\cap W^s_\alpha(y)\neq\varnothing$ *whenever* $x, y\in X$ *and* $m\geqq N$.

Proof. Set $\delta=\min\{\delta^*, \frac{1}{2}\alpha, \frac{1}{4}\varepsilon(\frac{1}{2}\alpha)\}$ and choose $p_i, \ldots, p_r$ periodic such that every $x\in X$ is within $\frac{1}{2}\varepsilon(\frac{1}{2}\alpha)$ of some p_k. Let t_k be the period of p_k. By 2.2 and $\mathrm{Cl}\,(W^u(p_k))=X$, there is an m_k such that every $y\in X$ is within $\varepsilon(\delta)$ of $f^{mt_k}W^u_\delta(p_k)$ for $m\geqq m_k$. Let $N=(m_1t_1)\cdots(m_rt_r)$. Then $d(y, f^NW^u_\delta(p_k))\leqq\varepsilon(\delta)$ for all k and all $y\in X$.

Suppose $x, y\in X$. Then $d(x, p_j)<\frac{1}{2}\varepsilon(\frac{1}{2}\alpha)$ for some j and $d(y, z)\leqq\varepsilon(\delta)$ for some $z\in f^NW^u_\delta(p_j)$. Let $w\in W^u_\delta(z)\cap W^s_\delta(y)$. Then $f^{-N}(w)\in W^u_\delta(f^{-N}(z))\subset W^u_{2\delta}(p_j)$ and $d(f^{-N}(w), p_j)\leqq\frac{1}{2}\varepsilon(\frac{1}{2}\alpha)$; thus $d(f^{-N}(w), x)\leqq\varepsilon(\frac{1}{2}\alpha)$ and there is a $v\in W^s_{\alpha/2}(f^{-N}(w))\cap W^u_{\alpha/2}(x)$. Then $f^N(v)\in f^NW^u_\alpha(x)$ and $f^N(v)\in W^s_{\alpha/2}(w)\subset W^s_\alpha(y)$. Therefore $f^NW^u_\alpha(x)\cap W^s_\alpha(y)\neq\varnothing$, $\forall x, y\in X$. If $m\geqq N$, then

$$f^mW^u_\alpha(x)\cap W^s_\alpha(y)\supset f^NW^u_\alpha(f^{m-N}(x))\cap W^s_\alpha(y)\neq\varnothing.$$

(2.9) DEFINITIONS. Let $\mathrm{Per}_n\,(U)=\{x\in U : f^n(x)=x\}$, $N_n(U)=\mathrm{card}\,(\mathrm{Per}_n\,(U))$, and $N_n(f)=N_n(X)$.

A *G-time* is a finite collection $\tau=\{I_1, \ldots, I_m\}$ of disjoint (finite) intervals of integers. We let $\mathrm{Tim}\,(\tau)=\bigcup_{I\in\tau}I$, $T(\tau)=\mathrm{card}\,(\mathrm{Tim}\,(\tau))$, and $L(\tau)$ be the length of the shortest interval containing $\mathrm{Tim}\,(\tau)$. A map $P\colon \mathrm{Tim}\,(\tau)\to X$ is (f, τ)*-admissible* if $f^{t_2-t_1}$, $P(t_1)=P(t_2)$ whenever t_1, $t_2\in I\in\tau$ (i.e. $P(I)$ is part of an f-orbit). A *specification* is a pair $s=(\tau, P)$ with τ a G-time and P an (f, τ)-admissible map; set $L(s)=L(\tau)$ and $\mathrm{Tim}\,(s)=\mathrm{Tim}\,(\tau)$; we also write sometimes $\tau=\tau(s)$ or $P=P_s$. For $n\geqq 0$ we say that τ is *n-delayed* if there is an interval of length at least n between every pair of invervals belonging to τ; s is n-delayed if $\tau(s)$ is. Notice that while $\mathrm{Tim}\,(\tau)$ does not determine τ, it does if τ is n-delayed with $n>0$.

Finally, for $\varepsilon>0$, let

$$U(s, \varepsilon)=\{x\in X : d(f^t(x), P_s(t)<\varepsilon \text{ for all } t\in\mathrm{Tim}\,(s)\}.$$

(2.10) THEOREM. *Suppose* $f\colon X\to X$ *is C-dense and* $\varepsilon>0$. *There is an* $M(\varepsilon)$ *such that* $U(s, \varepsilon)\neq\varnothing$ *whenever s is an* $M(\varepsilon)$*-delayed f-specification. In fact* $M(\varepsilon)$ *can be chosen so that* $\mathrm{Per}_d\;U(s, \varepsilon)\neq\varnothing$ *for all* $d\geqq M(\varepsilon)+L(s)$.

Proof. We tend s to a new specification s' as follows. Let a_1 be the smallest integer in Tim (s). Set $\tau(s')=\tau(s)\cup\{\{a_1+d\}\}$ and define $P_{s'}$ by $P_{s'}(a_1+d)=P_s(a_1)$ and $P_{s'}|\text{Tim}(s)=P_s$.

Set $\beta=\frac{1}{2}\min\{\psi(\frac{1}{2}\varepsilon), \varepsilon, \delta^*\}$ (ψ defined in 2.5) and $\alpha=\beta/3c$; let N be the integer given by 2.8 for this α. Choose $M=M(\varepsilon)\geqq\max\{N, R(\frac{1}{2}\varepsilon)\}$ (R defined in 2.5) large enough so that $\sum_{j=0}^{\infty}\lambda^{Mj}<2$. Assume $d\geqq M(\varepsilon)+L(s)$; then s' is M-delayed.

Let $I_1=[a_1, b_1]$, $I_2=[a_2, b_2]$, $\ldots$, $I_m=[a_m, b_m]=\{a_1+d\}$ be the members of $\tau(s')$ in their natural order. We set $z_1=x_1$ and define z_k (for $1\leqq k\leqq m$) recursively as follows. Suppose z_k has been chosen for some $1\leqq k<m$. As s^1 is M-delayed, $a_{k+1}-b_k>M\geqq N$ and so by 2.8 there exists a point

$$v\in f^{a_{k+1}-b_k}W^u_\alpha(f^{b_k}(z_k))\cap W^s_\alpha(P_{s^1}(a_{k+1})).$$

Set $z_{k+1}=f^{-a_{k+1}}(v)$; then $f^{b_k}(z_{k+1})\in W^u_\alpha(f^{b_k}(z_k))$ and $f^{a_{k+1}}(z_{k+1})\in W^s_\alpha(P_{s'}(a_{k+1}))$.

By induction on r we show that

$$f^{b_k}(z_{k+r})\in W^u_{c\alpha+c\alpha\lambda^M+\cdots+c\alpha\lambda^{M(r-1)}}(f^{b_k}(z_k)).$$

For $r=1$, this was seen above (since $c\geqq 1$). Assume the statement is true for some $r\geqq 1$. Since s^1 is M-delayed; $b_{k+r}-b_k\geqq rM$; because $f^{b_{k+r}}(z_{k+r+1})\in W^u_\alpha(f^{b_{k+r}}(z_{k+r}))$ we get

$$(*)\qquad f^{b_k}(z_{k+r+1})\in W^u_{c\alpha\lambda^{Mr}}(f^{b_k}(z_{k+r})).$$

(Here we use A3: If $x\in W^u_\alpha(y)$, then $d(f^{-n}(x), f^{-n}(y))\leqq c\alpha\lambda^n$ for $n\geqq 0$ and so $f^{-m}(x)\in W^u_{\alpha\lambda^m}(f^{-m}(y))$ for $m\geqq 0$.) Applying (*) and our inductive hypothesis, it follows that (see 2.2(ii))

$$f^{b_k}(z_{k+r+1})\in W^u_{c\alpha+\cdots+c\alpha\lambda^{Mr}}(f^{b_k}(z_k))$$

and so our induction is done.

Since $\sum_{j=0}^{\infty}\lambda^{Mj}<2$ and $\alpha=\beta/3c$ we have $f^{b_k}(z_m)\in W^u_{2\beta/3}(f^{b_k}(z_k))$ and $d(f^t(z_m), f^t(z_k))<2\beta/3$ for any $t\in I_k$ and any $k\in[1, m]$. Since $f^{a_k}(z_k)\in W^s_\alpha(P_{s^1}(a_k))$ (by the definition of the z_k's) we have

$$\beta/3\ \geqq\ \alpha\ \geqq\ d(f^t(z_k), f^{t-a_k}(P_{s^1}(a_k)))\ =\ d(f^t(z_k), P_{s^1}(t))$$

for any $t\in I_k$. Combining inequalities,

$$d(f^t(z_m), P_{s^1}(t))\ <\ \beta\quad\text{for all } t\in\text{Tim}(s^1).$$

Thus $z_m\in U(s^1, \beta)$.

Now let $z^*=f^{a_1}(z_m)$. Then $z^*, f^d(z^*)\in B_\beta(P_s(a_1))$, and so $d(z^*, f^d(z^*))\leqq\psi(\frac{1}{2}\varepsilon)$. Now $d>M(\varepsilon)\geqq R(\frac{1}{2}\varepsilon)$ and by 2.5 there is a $z\in\text{Per}_d(X)$ with

$$d(f^t(z), f^t(z^*))\ \leqq\ \tfrac{1}{2}\varepsilon\quad\text{for all } 0\leqq t\leqq d.$$

Letting $z^1=f^{-a_1}(z)$ we get

$$d(f^t(z^1), f^t(z_m))\ \leqq\ \tfrac{1}{2}\varepsilon\quad\text{for all } a_1\ \leqq\ t\ \leqq\ a_1+d.$$

Applying the triangle inequality to this and $z_m \in U(s^1, \beta)$,

$$z^1 \in U(s^1, \beta+\tfrac{1}{2}\varepsilon) \subseteq U(s^1, \varepsilon) \subseteq U(s, \varepsilon);$$

also $z^1 \in \mathrm{Per}_d(X)$.

(2.11) REMARK. The above theorem is a statement about the freedom one has in specifying the approximate orbit of a periodic point. The remainder of this paper shall be derived from this freedom (together with expansiveness).

3. **Counting.** Throughout this section $f: X \to X$ is a C-dense map.

(3.1) DEFINITION. For $\varepsilon>0$, $E\subset X$ is an *(n, ε)-separated* set if for any distinct $x, y \in E$ there is a t for which $0 \leqq t<n$ and $d(f^t(x), f^t(y))>\varepsilon$. We let $N(n, \varepsilon)$ denote the maximum cardinality of an (n, ε)-separated set.

(3.2) LEMMA. (i) *If $\varepsilon \leqq \delta^*$, then $N(n, \varepsilon) \geqq N_n(f)$.*

(ii) *If $\varepsilon \leqq \alpha$, then $N(n, \alpha) \leqq N(n, \varepsilon)$; for any $\varepsilon>0$ there is an m_ε such that $N(n, \varepsilon) \leqq N(n+m_\varepsilon, \delta^*)$ for all $n \geqq 0$.*

(iii) $N(\sum n_i, \varepsilon) \leqq \prod N(n_i, \tfrac{1}{2}\varepsilon)$.

Proof. (i) By 2.3 ε is an expansive constant; i.e. if $p \neq q$, then $d(f^t(p), f^t(q))>\varepsilon$ for some t. If $p, q \in \mathrm{Per}_n(X)$, then t can be chosen so that $0 \leqq t<n$; i.e. $\mathrm{Per}_n(X)$ is (n, ε)-separated.

(ii) The first statement is obvious; if E is an (n, ε)-separated set, then $f^{-D(\varepsilon)}E$ is an $(n+2D(\varepsilon), \delta^*)$-separated set (use 2.4).

(iii) We prove the following stronger statement for later use: Suppose $E\subset X$ and n_i, m_i $(1 \leqq i \leqq s)$ are integers $(n_i>0)$ such that, when $x, y \in E$ and $x \neq y$, there is a $t \in \bigcup_{i=1}^s [m_i, m_i+n_i)$ for which $d(f^t(x), f^t(y))>\varepsilon$; then $\mathrm{card}(E) \leqq \prod_{i=1}^s N(n_i, \tfrac{1}{2}\varepsilon)$.

Proof. Choose $R_i \subset X$ so that $f^{m_i}R_i$ is a maximal $(m_{i_1}, \tfrac{1}{2}\varepsilon)$-separated set. Construct a map $g=\prod g_i: E \to \prod R_i$ by requiring that $d(f^t(x), f^t(g_i(x))) \leqq \tfrac{1}{2}\varepsilon$ for all $t \in [m_i, m_i+n_i)$. Such a $g_i(x)$ exists by the maximality of $f^{m_i}R_i$—otherwise $f^{m_i}(Ru\{x\})$ would be an $(n, \tfrac{1}{2}\varepsilon)$-separated set.

If $g(x)=g(y)$ the triangle inequality would give us $d(f^t(x), f^t(y)) \leqq \varepsilon$ for all $t \in \bigcup [m_i, m_i+n_i)$; thus g is injective and we are done.

Two specifications s and s^1 are *p-separated* if $d(P_s(t), P_{s^1}(t))>p$ for some $t \in \mathrm{Tim}(s) \cap \mathrm{Tim}(s^1)$; a set of specifications is p-separated if every two members are. An *S-set A* is a set of specifications with the same G-time; let $\tau(A)$ denote this common G-time, $T(A)=T(\tau(A))$, $L(A)=L(\tau(A))$, and $U(A, \varepsilon)=\bigcup_{s\in A} U(s, \varepsilon)$.

3.3 LEMMA. (i) *If s and s^1 are p-separated, then $U(s, \tfrac{1}{2}p) \cap U(s^1, \tfrac{1}{2}p)=\varnothing$.*

(ii) *If A is a 2ε-separated S-set, $\tau(A)$ is $M(\varepsilon)$-delayed, and $d \geqq L(A)+M(\varepsilon)$, then $N_d(U(A, \varepsilon)) \geqq \mathrm{card}(A)$.*

Proof. (i) Trivial. (ii) Follows from (i) and 2.10.

Two specifications s and s^1 are disjoint if $\mathrm{Tim}(s) \cap \mathrm{Tim}(s^1)=\varnothing$. In this case we define a new specification $s \wedge s^1$ by $\tau(s \wedge s^1)=\tau(s) \cup \tau(s^1)$ and

$$\begin{aligned} P_{s\wedge s^1}(t) &= P_s(t) \quad \text{for } t \in \mathrm{Tim}(s), \\ &= P_{s^1}(t) \quad \text{for } t \in \mathrm{Tim}(s^1). \end{aligned}$$

Notice that $U(s\wedge s^1, \varepsilon)=U(s, \varepsilon)\cap U(s^1, \varepsilon)$. We call a G-time τ an *m-time* if card $\tau=m$; s is an m-specification if $\tau(s)$ is an m-time.

(3.4) LEMMA. *If τ is an n-delayed m-time and $N\geqq L(\tau)$, there is a τ^1 such that*
(a) $\operatorname{Tim}(\tau)\cap\operatorname{Tim}(\tau^1)=\varnothing$,
(b) *$\tau\cup\tau^1$ is n-delayed,*
(c) *$L(\tau\cup\tau^1)\leqq N$, and*
(d) *$T(\tau^1)\geqq N-2mn-T(\tau)$.*

Proof. Let a_1 be the smallest integer in $\operatorname{Tim}(\tau)$. Set

$$\operatorname{Tim}(\tau^1) = \{t\in[a_1, a_1+N) : |t-r| > n \quad \text{for all } r\in\operatorname{Tim}(\tau)\}.$$

This determines a G-time τ which satisfies our condition.

(3.5) REMARK. τ^1 could be empty.

(3.6) LEMMA. *If τ is a time specification and $\varepsilon>0$, there is an ε-separated S-set A with $\tau(A)=\tau$ and* card $(A)\geqq N(T(\tau), 2\varepsilon)$.

Proof. Let $\tau=\{I_1, \dots, I_m\}$ and $\tau_k=\{I_k\}$ for $1\leqq k\leqq m$. Let A_k be an ε-separated S-set with $\tau(A_k)=\tau_k$ and card $(A_k)=N(T(\tau_k), \varepsilon)$. Then

$$A = A_1\wedge\cdots\wedge A_m = \{s_1\wedge\cdots\wedge s_m : s_k\in A_k, 1\leqq k\leqq m\}$$

is ε-separated with $\tau(A)=\tau_1\wedge\cdots\wedge\tau_m=\tau$ and card $(A)=\prod N(T(\tau_k), \varepsilon)\geqq N(\sum T(\tau_k), 2\varepsilon)=N(T(\tau), 2\varepsilon)$ by 3.2(iii).

(3.7) THEOREM. *Suppose B is a 2ε-separated S-set with $\tau(B)$ an $M(\varepsilon)$-delayed m-time. Then*

$$N_d(U(B, \varepsilon)) \geqq \frac{K(m, \varepsilon)\operatorname{card}(B)\,N(d, 8\varepsilon)}{N(T(\tau(B)), 4\varepsilon)}$$

for all $d\geqq L(\tau(B))+M(\varepsilon)$ where $K(m, \varepsilon)>0$ depends only on m and $\varepsilon>0$.

Proof. Let $N=d-M(\varepsilon)\geqq L(\tau(B))$. Let $\tau=\tau(B)$ and choose τ^1 as in Lemma 3.4. By Lemma 3.5 let A be a 2ε-separated S-set with $\tau(A)=\tau^1$ and card $(A)\geqq N(T(\tau^1), 4\varepsilon)$. Now $A\wedge B$ is a 2ε-separated S-set with $M(\varepsilon)$-delayed time $\tau\wedge\tau^1$; $d\geqq N+M(\varepsilon)\geqq L(\tau\wedge\tau^1)+M(\varepsilon)$. Hence, by 3.3(ii), we have

$$N_d(U(A\wedge B, \varepsilon)) \geqq \operatorname{card}(A\wedge B) = \operatorname{card}(A)\operatorname{card}(B).$$

Since $U(B, \varepsilon)\geqq U(A\wedge B, \varepsilon)$,

$$N_d(U(B, \varepsilon)) \geqq \operatorname{card}(A)\operatorname{card}(B).$$

Now $T(\tau^1)\geqq\max\{0, N-2mM(\varepsilon)-T(\tau)\}$ (see Remark 3.5). Thus

$$\operatorname{card} A \geqq \max\{1, N(N-2mM(\varepsilon)-T(\tau), 4\varepsilon)\} = W$$

(taking 1 in case $N-2mM(\varepsilon)-T(\tau)\leqq 0$). Recalling that $N=d-M(\varepsilon)$ and 3.2(iii) we get

$$N(d, 8\varepsilon) \leqq W\cdot N((2m+1)M(\varepsilon), 4\varepsilon)N(T(\tau), 4\varepsilon)$$

(the inequality is good in the exceptional case we have been noting). Thus

$$N_d(U(B, \varepsilon)) \geqq \text{card}\,(B) : W$$
$$\geqq \frac{K(m, \varepsilon)\,\text{card}\,(B)\,N(d, \delta\varepsilon)}{N(T(\tau), 4\varepsilon)}$$

where $K(m, \varepsilon) = N((2m+1)M(\varepsilon), 4\varepsilon)^{-1}$.

(3.8) DEFINITION. For $U \subset X$ let

$$\varphi(U) = \liminf_{n\to\infty} \frac{N_n(U)}{N_n(f)} \quad \text{and} \quad \theta(U) = \limsup_{n\to\infty} \frac{N_n(U)}{N_n(f)}.$$

(3.9) COROLLARY. (i) *For any* $\alpha > 0$

$$\liminf_{d\to\infty} \frac{N_d(f)}{N(d, \alpha)} > 0.$$

(ii) $\varphi(V) > 0$ *when* $V \neq \varnothing$ *is open.*

(iii) *There is a* $K^* > 0$ *such that* $\varphi(U) \geqq K^*\theta(V)$ *whenever* U *and* V *are open in* X *and* $U \supset \overline{V}$.

(iv) *There are* m_0 *and* $S > 0$ *such that* $N_{m+n}(f) \geqq SN(m, \delta^*)N(n, \delta^*) \geqq SN_m(f)N_n(f)$ *provided that* $m \geqq m_0$.

(v) *There are* m_0 *and* $S > 0$ *such that, if* $m \geqq m_0$ *and* $U \subset X$ *satisfies* $\operatorname{diam} f^k(U) \leqq \delta^*$ *for all* $0 \leqq k < m$, *then* $\theta(U) \leqq 1/SN_m(f)$.

Proof. (i) and (ii). Let $x \in V$ and choose $\varepsilon > 0$ so small that $B_\varepsilon(x) \subset V$ and $8\varepsilon \leqq \min\{\alpha, \delta^*\}$. Let s be given by $\tau(s) = \{\{0\}\}$ and $P_s(0) = x$; $B = \{s\}$. Then $V \supset U(s, \varepsilon)$ and by the theorem

$$N_d(f) \geqq N_d(V) \geqq K(1, \varepsilon)\,N(d, 8\varepsilon)/N(1, 4\varepsilon)$$

for $d \geqq 1 + M(\varepsilon)$. As $N(d, 8\varepsilon) \geqq N(d, \alpha)$, (i) follows immediately. As $N(d, 8\varepsilon) \geqq N(d, \delta^*) \geqq N_d(f)$, so does (ii).

(iii) Choose $\varepsilon > 0$ so that $U \supset B_\varepsilon(V)$ and let $D(\varepsilon)$ be given as in 2.4. Consider $n > 2D(\varepsilon)$. For each $p \in \text{Per}_n(V)$ form the 1-specification $s(p)$ with $\tau(s(p)) = \{[-D(\varepsilon), n - D(\varepsilon))\}$ and $P_{s(p)}(f) = f^t(p)$. $B_n = \{s(p) : p \in \text{Per}_n(V)\}$ is δ^*-separated (see the proof of 3.2(iii)). By the definition of ε and $D(\varepsilon)$ we have $U(B_n, \delta^*) \subset U$.

Trivially, $U(B_n, \frac{1}{8}\delta^*) \subset U$; so by the theorem

$$N_d(U) \geqq K(1, \tfrac{1}{8}\delta^*)N_n(V)N(d, \delta^*)/N(n, \tfrac{1}{2}\delta^*)$$

for $d \geqq n + M(\frac{1}{8}\delta^*)$. By (i) above there is an n_0 and a K_1 such that $N(n, \frac{1}{2}\delta^*) \leqq K_1 N_n(f)$ when $n \geqq n_0$; also $N(d, \delta^*) \geqq N_d(f)$. Thus for $n \geqq n_0$ and $d \geqq n + M(\frac{1}{8}\delta^*)$ we have

$$N_d(U)/N_d(f) \geqq K^* N_n(V)/N_n(f)$$

where $K^* = K(1, \frac{1}{8}\delta^*)/K_1 > 0$. Then $\varphi(U) \geqq K^*\theta(V)$.

(iv) Set $m_0 = 2M(\frac{1}{4}\delta^*)$. Let A be a $\frac{1}{2}\delta^*$-separated S-set with $\tau(A) = \{[0, n)\}$ and card $A = N(n, \frac{1}{2}\delta^*)$; B a $\frac{1}{2}\delta^*$-separated S-set with $\tau(B) = \{[n + M(\frac{1}{4}\delta^*), n + m$

$-M(\frac{1}{4}\delta^*))\}$ and card $B=N(m-m_0, \frac{1}{4}\delta^*)$. Now $A \wedge B$ is $\frac{1}{2}\delta^*$-separated with $M(\frac{1}{4}\delta^*)$-delayed time.

By 3.3(ii) we have

$$N_{n+m}(f) \geqq \text{card}\,(A \wedge B) = N(n, \tfrac{1}{2}\delta^*)N(m-m_0, \tfrac{1}{2}\delta^*).$$

By Proposition 3.2(iii) we have

$$N(m, \delta^*) \leqq N(m-m_0, \tfrac{1}{2}\delta^*)N(m_0, \tfrac{1}{2}\delta^*).$$

Taking $S=N(m_0, \frac{1}{2}\delta^*)^{-1}$, $N_{n+m}(f) \geqq SN(n, \delta^*)N(m, \delta^*)$.

(v) Let m_0 and S be as above. Since $\text{Per}_{n+m}\,(U)$ is an $(n+m, \delta^*)$-separated set and diam $f^k(U) \leqq \delta^*$ for $0 \leqq k < m$, $f^m\,\text{Per}_{n+m}\,(U)$ is an (n, δ^*)-separated set; thus $N_{n+m}(U) \leqq N(n, \delta^*)$. By (iv) we have, since $m \geqq m_0$, $N_{n+m}(f) \geqq SN(n, \delta^*)N(m, \delta^*)$ and so

$$N_{n+m}(U)/N_{n+m}(f) \leqq 1/SN_m(f).$$

Letting $n \to \infty$, $\theta(U) \leqq 1/SN_m(f)$.

(3.10) Definition. For $A \subset X$ let $N(n, \varepsilon, A)$ be the largest cardinality of an (n, ε)-separated set contained in A.

(3.11) Proposition. *For each ε with $0<\varepsilon<\frac{1}{2}\delta^*$ there are constants $c_\varepsilon>0$ and $0<\tau_\varepsilon<1$ for which the following holds. If $A \subset X$, $0 \leqq k_1 < k_2 < \cdots < k_m$, are integers and $w_{k_1}, \ldots, w_{k_m} \in X$ satisfy $f^{k_r}(A) \cap B_\varepsilon(w_{k_r}) = \varnothing$ for $r=1, \ldots, m$, then $N(n, \varepsilon, A) \leqq c_\varepsilon \tau^m N(n, \varepsilon)$ for all $n>k_m$.*

Proof. Let $M=M(\frac{1}{2}\varepsilon)$ as in 2.10. Let $j_1<j_2<\cdots<j_q$ be a subsequence of $k_1<\cdots<k_m$ such that $j_{i+1}-j_i>2M$ and $q \geqq m/(2M+1)$. Let $n>k_m$ and $E_n \subset A$ be an (n, ε)-separated set. For each $I \subset J=\{j_1, \ldots, j_q\}$ and each $x \in E_n$ we define the specification $s(x, I)$ by requiring that it be an M-delayed specification with

$$\text{Tim}\, s(x, I) = \Big([0, n) \setminus \bigcup_{j_i \in I} [j_i-M, j_i+M]\Big) \cup I,$$

$$P_{s(x,I)}(t) = f^t(x) \quad \text{for } t \notin I \quad \text{and} \quad P_{s(x,I)}(j_i) = w_{j_i} \quad \text{for } j_i \in I.$$

Set $d=n+m$. By Theorem 2.10 choose

$$p(x, I) \in U(s(x, I), \tfrac{1}{2}\varepsilon) \cap \text{Per}_d\,(X).$$

Let $F_I=\{p(x, I) : x \in E_n\}$. If $I_1 \neq I_2$ and $x, y \in E_n$, then $s(x, I_1)$ and $s(y, I_2)$ are ε-separated; for if $j_i \in I_1 \setminus I_2$, then $j_i \in \text{Tim}\, s(x, I_1) \cap \text{Tim}\, s(y, I_2)$ and

$$d(P_{s(x,I_1)}(j_i), P_{s(y,I_2)}(j_i)) = d(w_{j_i}, f^{j_i}(y)) > \varepsilon.$$

By lemma (i) we have $p(x, I_1) \neq p(y, I_2)$; thus $I_1 \neq I_2$ implies $F_{I_1} \cap F_{I_2} = \varnothing$.

Suppose $z=p(x, I)=p(y, I)$ and $x \neq y$. For $t \in \text{Tim}\, s(x, I) \setminus I$, we have $P_{s(x,I)}(t) = f^t(x)$ and $P_{s(y,I)}(t)=f^t(y)$; so $d(f^t(z), f^t(x))<\frac{1}{2}\varepsilon$ and $d(f^t(z), f^t(y))<\frac{1}{2}\varepsilon$, hence $d(f^t(x), f^t(y))<\varepsilon$. Since $x, y \in E_n$, an (n, ε)-separated set, we must have $d(f^t(x), f^t(y))>\varepsilon$ for some

$$t \in [0, n) \setminus (\text{Tim}\, s(x, I) \setminus I) = \bigcup_{j_i \in I} [j_i-M, j_i+M].$$

By the proof of 3.2(iii), $\{x \in E_n : p(x, I) = z\}$ has at most $g^{\text{card} I}$ elements where $g = N(2M+1, \frac{1}{2}\varepsilon)$. Thus F_I has at least card $E_n \backslash g^{\text{card} I}$ elements.

As the F_I's are disjoint

$$N_d(f) \geqq \sum_{I \subset J} \text{card } F_I \geqq \sum_{I \subset J} \frac{1}{g^{\text{card} I}} \text{card } E_n$$
$$\geqq \sum_{r=0}^{\text{card} J} \binom{\text{card } J}{r} \frac{1}{g^r} \text{card } E_n = \left(1 + \frac{1}{g}\right)^{\text{card} J} \text{card } E.$$

Since $2\varepsilon < \delta^*$, by 3.2(i) and 3.2(iii)

$$N_d(f) = N_{n+m}(f) \leqq N(n+M, 2\varepsilon) \leqq N(n, \varepsilon) N(M, \varepsilon).$$

Also card $J = q \geqq m/(2M+1)$. Thus

$$N(n, \varepsilon, A) = \text{card } E_n \leqq \frac{N(M, \varepsilon)}{[(1+1/g)^{1/2M+1}]^m} N(n, \varepsilon).$$

4. **Topological entropy.** Suppose $\mathscr{A}$ is a finite open cover of X. $E \subset \mathscr{A} \times \cdots \times \mathscr{A}$ (n-times) is an *n-cover* for $(f, \mathscr{A})$ if for every $z \in X$ there is an $(A_0, \ldots, A_{n-1}) \in E$ such that $f^k(x) \in A_k$ for all $0 \leqq k < n$. Let $M_n(f, \mathscr{A})$ denote the minimum cardinality of an n-cover for $(f, \mathscr{A})$. Then (see Adler, Konheim and McAndrew [1]) the limit

$$h(f, \mathscr{A}) = \lim_{n \to \infty} \frac{1}{n} \log M_n(f, \mathscr{A})$$

exists and the *topological entropy* of f is defined by

$$h(f) = \sup_{\mathscr{A}} h(f, \mathscr{A}).$$

(The above definitions and 4.1 and 4.2 below do not depend on our standing hypothesis that f satisfies Axiom A*; they work for any continuous map of a compact Hausdorff space.)

(4.1) Definition. $f: X \to X$ has *completely positive topological entropy* (c.p.t.e.) if $h(f, \{C, D\}) > 0$ whenever $\{C, D\}$ is an open cover of X with $\bar{C} \neq \bar{X} \neq \bar{D}$.

(4.2) Proposition. *Suppose $f: X \to X$ has c.p.t.e. Then $h(f) > 0$ unless X is a single point, and it is topologically transitive. If $g: Y \to Y$ and $h: X \to Y$ are continuous maps with h surjective and $g \circ h = h \circ f$, then g has c.p.t.e.*

Proof. Unless X is a single point an open cover $\{C, D\}$ as in 4.1 can be found and so $h(f) > 0$.

If f is not transitive, then there is an open set $C \neq \varnothing$ with $f^{-1}(C) \subset C$ and $\bar{C} \neq X$. Let $B \neq \varnothing$ be open with $\bar{B} \subset C$ and set $D = X \backslash \bar{B}$. Then $\{C, D\}$ is as above. Let

$$E_n = \{(\underbrace{C, \ldots, C}_{i \text{ times}}, \underbrace{D, \ldots, D}_{j \text{ times}}) : i + j = n, i, j \geqq 0\}.$$

We claim E_n is an n-cover for $(f, \{C, D\})$. For, if $x \in X$, then either $f^k(x) \in D$ for all $0 \leqq k < n$ or there is a largest k, denoted $k(x)$, such that $0 \leqq k < n$ and $f^k(x) \notin D$.

In the latter case $f^{k(x)}(x) \in C$ and so $f^m(x) \in C$ for all $m \leqq k(x)$ as $f^{-1}(C) \subset C$; $f^m(x) \in D$ for $m > k(x)$. As card $E_n = n+1$, $M_n(f, \{C, D\}) \leqq n+1$ and $h(\{C, D\}) = 0$—a contradiction.

Suppose $\{C, D\}$ is an open cover of Y with $\bar{C} \neq \bar{Y} \neq \bar{D}$. Then $\{h^{-1}(C), h^{-1}(D)\}$ satisfies the condition of 4.1 also. h and h^{-1} induce a bijection between n-covers for $(f, \{h^{-1}(C), h^{-1}(D)\})$ and $(g, \{C, D\}) = h(f_1\{h^{-1}(C), h^{-1}(D)\}) > 0$.

(4.3) THEOREM. *If $f: X \to X$ is C-dense, then f has c.p.t.e.*

Proof. Let $\{C, D\}$ be a cover as in 4.1. Choose $\varepsilon > 0$ and $p, q \in X$ such that $B_\varepsilon(p) \subset C \backslash D$ and $B_\varepsilon(q) \subset D \backslash C$. Let $M(\varepsilon)$ be the integer given by 2.10; set $N = M(\varepsilon) + 1$. Then $\tau_n = \{\{kN\} : 0 \leqq k < n\}$ is $M(\varepsilon)$-delayed.

For $(a_0, \ldots, a_{n-1}) \in \prod_{k=0}^{n-1} \{p, q\}$ define a specification $s = s_n(a_0, \ldots, a_{n-1})$ by $\tau(s) = \tau_n$ and $P_s(kN) = a_k$. By 2.10 choose points

$$x_n(a_0, \ldots, a_{n-1}) \in U(s_n(a_0, \ldots, a_{n-1}), \varepsilon).$$

Let E_n be an nN-cover for $(f, \{C, D\})$; for $x \in X$ let $F_n(x) = (F_n^0(x), \ldots, F_n^{nN-1}(x)) \in E_n$ be such that $f^j(x) \in F_n^j(x)$ for $0 \leqq j < nN$. Suppose $(a_0, \ldots, a_{n-1}) \neq (b_0, \ldots, b_{n-1})$; say $a_k = p$ and $b_k = q$. Then

$$f^{kN}(x_n(a_0, \ldots, a_{n-1})) \in B_\varepsilon(p) \leqq C \backslash D$$

and so $F_n^{kN}(x_n(a_0, \ldots, a_{n-1})) = C$; similarly $F_n^{kN}(x_n(b_0, \ldots, b_{n-1})) = D$ and so $F_n(x_n(b_0, \ldots, b_{n-1})) \neq F_n(x_n(a_0, \ldots, a_{n-1}))$. It follows that card $E_n \geqq 2^n$ and $M_{nN}(f, \{C, D\}) \geqq 2^n$; thus

$$h(f, \{C, D\}) \geqq \lim \frac{1}{nN} \log 2^n = \frac{1}{N} \log 2 > 0.$$

(4.4) REMARK. Now $f: X \to X$ satisfying Axiom A* could not be topologically transitive unless the permutation g in its C-dense decomposition (2.7) is a cycle, i.e. if the decomposition $X = X_1 \cup \cdots \cup X_m$ satisfies $X = \bigcup f^k X_1$; with 4.2 and 4.3 one sees that this is a sufficient condition for transitivity. It is now clear how 2.7 is just another version of Smale's Spectral Decomposition [16, p. 777]. We also see that $h(f) > 0$ unless X is finite; this result was proved before in [6]. The following is an improvement of the main result of [6].

(4.5) THEOREM. *If $f: X \to X$ is C-dense, then*

$$h(f) = \lim_{n \to \infty} \frac{1}{n} \log N_n(f).$$

Proof. Let $\mathscr{A}$ be a finite open cover of X with diam $(A) < \delta^*$ for all $A \in \mathscr{A}$ and let $\beta > 0$ be a Lebesgue number for $\mathscr{A}$ (i.e. every closed β-ball $B_\beta(x)$ lies inside some member of $\mathscr{A}$).

Let Q be a maximal (n, β)-separated set. For $z \in Q$ choose $B(z) = (A_0(z), \ldots, A_{n-1}(z))$ with $A_k(z) \in \mathscr{A}$ and

$$A_k(z) \supset \mathrm{Cl}\,(B_\beta(f^k(z))) \quad \text{for all } 0 \leqq k < n.$$

We claim $E_n=\{B(z) : z\in Q\}$ is an n-cover for $(f, \mathscr{A})$. For each $x\in X$ there is a $z_x\in Q$ for which $d(f^k(x), f^k(z_k))\leqq\beta$ for all $0\leqq k<n$; otherwise $Q\cup\{x\}$ would be an (n, β)-separated set bigger than Q. Since $f^k(x)\in A_k(z_x)$, E_n is an n-cover. We have shown $M_n(f, \mathscr{A})\leqq N(n, \beta)$.

Let E be an n-cover for $(f, \mathscr{A})$ and R an (n, δ^*)-set. For $x\in R$ choose $g(x)=(A_0(x), \ldots, A_{n-1}(x))\in E$ such that $f^k(x)\in A_k(x)$ for all $0\leqq k<n$. If $g(x)=g(y)$, then $A_k(x)=A_k(y)$ and $d(f^k(x), f^k(y))\leqq \text{diam } A_k(x)<\delta^*$ for $0\leqq k<n$; $x=y$ as R is an (n, δ^*)-separated set. As $g: R\to E$ is injective, card $E\geqq$ card R and $M_n(f, \mathscr{A})\geqq N(n, \delta^*)\geqq N_n(f)$.

By 3.9(i) there is an $S>0$ and n_0 such that $N_n(f)\geqq SN(n, \beta)$ for $n\geqq n_0$. Hence $SM_n(f, \mathscr{A})\leqq N_n(f)\leqq M_n(f, \mathscr{A})$ for all $n\geqq n_0$. Since $(1/n)\log M_n(f, \mathscr{A})$ approaches the limit $h(f, \mathscr{A})$, so does $(1/n)\log N_n(f)$. As this is true for every $\mathscr{A}$ with diam $\mathscr{A}<\delta^*$ and in calculating $h(f)$ we need only consider $h(f, \mathscr{A})$ with $\mathscr{A}$ having small diameter,

$$h(f) = h(f, \mathscr{A}) = \lim_{n\to\infty} \frac{1}{n}\log N_n(f).$$

(4.6) REMARK. Let

$$\gamma_f(\varepsilon) = \limsup \frac{1}{n}\log N(n, \varepsilon).$$

The proof above shows that, for any map f a compact metric space, $h(f)=\lim_{\varepsilon\to 0}\gamma_f(\varepsilon)$. Suppose f is a homeomorphism and δ is an expansive constant; if $\varepsilon\leqq\delta$, then 3.2(ii) goes through, i.e.

$$N(n, \delta) \leqq N(n, \varepsilon) \leqq N(n+m_\varepsilon, \delta)$$

for some m_ε, and so $\gamma_f(\varepsilon)=\gamma_f(\delta)$. In this case we have $\gamma_f(\delta)=h(f)$.

(4.7) THEOREM. *Suppose $f: X\to X$ is C-dense and $A\subset X$ is closed with $\varnothing\neq A\neq X$ and $f(A)=A$. Then $h(f|A)<h(f)$.*

Proof. By the remark above, $h(f|A)=\gamma_{f|A}(\varepsilon)$ for $\varepsilon\leqq\delta^*$. Choose $w\in X\backslash A$ and $\varepsilon>0$ so small that $A\cap B_\varepsilon(w)=\varnothing$. Recall 3.11, $N(n, \varepsilon, A)\leqq c_\varepsilon\tau_\varepsilon^m$, for $n>m$ where $\tau_\varepsilon<1$. Then

$$\begin{aligned}\gamma_{f|A}(\varepsilon) &= \limsup_{n\to\infty}\frac{1}{n}\log N(n, \varepsilon, A)\\ &\leqq \lim_{n\to\infty}\frac{1}{n}\log c_\varepsilon\tau_\varepsilon^{n-1}N(n, \varepsilon)\\ &\leqq \log\tau_\varepsilon+\gamma_f(\varepsilon) = \log\tau_\varepsilon+h(f) < h(f).\end{aligned}$$

5. **Construction of a measure.** Let ψ be a countable base for the topology of X which is closed under finite union. Assume $\omega: \psi\to R$ satisfies, for $B\in\psi$,

$$\omega(B)\geqq 0, \qquad \omega(X) = 1,$$
$$\omega(B_1)\geqq\omega(B_2) \quad \text{when } B_1\supset B_2,$$
$$\omega(B_1\cup\cdots\cup B_n)\leqq\sum\omega(B_i),$$

and

$$\omega(B_1 \cup B_2) = \omega(B_1)+\omega(B_2) \quad \text{when } \bar{B}_1 \cap \bar{B}_2 = \varnothing.$$

For U open in X define $m(U)=\sup\{\omega(B): \bar{B}\subset U \text{ and } B \in \psi\}$.

(5.1) LEMMA. *If $U\subset\bigcup_{i=1}^{\infty} U_i$, then $m(U)\leqq\sum m(U_i)$. If $U\cap V=\varnothing$, then $m(U\cup V)=m(U)+m(V)$.*

Proof. Let $B\in\psi$ with $\bar{B}\subset U$. By compactness let $U_1, \dots, U_n$ cover B. For $x\in\bar{B}$ choose $B_x\in\psi$ so that $\bar{B}_x\subset U_i$ for some i satisfying $1\leqq i\leqq n$. Let $B_{x_1}, \dots, B_{x_r}$ cover $\bar{B}$ and set $A_i=\bigcup\{B_{x_j} : \bar{B}_{x_j}\subset U_i\}$. Then

$$\omega(B) \leqq \omega\left(\bigcup_{i=1}^{n} A_i\right) \leqq \sum_{i=1}^{n} \omega(A_i) \leqq \sum_{i=1}^{n} m(U_i).$$

Now vary B.

By the first part of the lemma, $m(U\cup V)\leqq m(U)+m(V)$. Suppose $B_1, B_2\in\psi$ with $\bar{B}_1\subset U$ and $\bar{B}_2\subset V$. Then $\mathrm{Cl}\,(B_1\cup B_2)\subset U\cup V$ and $\bar{B}_1\cap\bar{B}_2=\varnothing$; so

$$m(U\cup V) \geqq \omega(B_1\cup B_2) = \omega(B_1)+\omega(B_2).$$

Varying the B_i we obtain $m(U\cup V)\geqq m(U)+m(V)$.

For any $E\subset X$ we define

$$m(E) = \inf\{m(U) : U\supset E, U \text{ open}\}.$$

One sees easily that this definition agrees with the earlier one on open sets and that $m(K)=\inf\{\omega(B) : B\supset K, B\in\psi\}$ when K is closed. We let

$$\mathcal{M} = \{E\subset X : m(E) = \sup\{m(K) : K\subset E, K \text{ closed}\}\}.$$

With standard arguments we get

(5.2) PROPOSITION. *$\mathcal{M}=\mathcal{M}_{\psi,\omega}$ is a σ-field containing the Borel sets of X and $m=m_{\psi,\omega}$ is a complete normalized regular measure on $\mathcal{M}$.*

Proof. One can, for example, use 5.1 and imitate the proof of the Riesz Representation Theorem given in Rudin [19, p. 40].

(5.3) LEMMA. *If $\omega_1: \psi_1\to R$ and $\omega_2: \psi_2\to R$ are as above and there is a $K>0$ such that $\omega_2(B_2)\geqq K\omega_1(B_1)$ when $B_2\supset\bar{B}_1$ and $\omega_1(B_1)\geqq K\omega_2(B_2)$ when $B_1\supset\bar{B}_2$, then $\mathcal{M}_{\psi_1,\omega_1}=\mathcal{M}_{\psi_2,\omega_2}$ and $Km_{\psi_1,\omega_1}\leqq m_{\psi_2,\omega_2}\leqq(1/K)m_{\psi_1,\omega_1}$.*

Proof. For U open and $\bar{B}_1\subset U$ with $B_1\in\psi_1$ we can find $B_2\in\psi_2$ such that $\bar{B}_1\subset B_2\subset\bar{B}_2\subset U$. Hence $m_{\psi_2,\omega_2}(U)\geqq\omega_2(B_2)\geqq K\omega_1(B_1)$. Varying B_1, $m_{\psi_2,\omega_2}(U)\geqq Km_{\psi_1,\omega_1}(U)$. Similarly $m_{\psi_1,\omega_1}(U)\geqq Km_{\psi_2,\omega_2}(U)$. These inequalities extend to any $E\subset X$.

Suppose $E\in\mathcal{M}_{\psi_1,\omega_1}$. Letting $K_n\subset E$ be compact with $m_{\psi_1,\omega_1}(K_n)\geqq m_{\psi_1,\omega_1}(E)-1/n$ we see that $E=E_1\cup\bigcup_{n=1}^{\infty} K_n$ where $E_1\subset F$ for some Borel set F with $m_{\psi_1,\omega_1}(F)=0$. Then $m_{\psi_1,\omega_1}(F)=0$ also and $E_1\in\mathcal{M}_{\psi_2,\omega_2}$ since m_{ψ_2,ω_2} is complete. As ψ_2, ω_2

contains Borel sets, we finally see that $E \in \mathcal{M}_{\psi_2,\omega_2}$. The proof of $\mathcal{M}_{\psi_1,\omega_1} \subset \mathcal{M}_{\psi_2,\omega_2}$ is the same.

We will now see how to define some ω's when we are given a homeomorphism $f: X \to X$ which is C-dense. Let ψ be any base as above. By diagonalization we can find increasing sequences of integers $\{n_k\}$ such that

$$\omega(B) = \alpha_{\{n_k\}}(B) = \lim_k \frac{N_{n_k}(B)}{N_{n_k}(f)}$$

exists for every $B \in \psi$. The measure we obtain we denote by $\mu_{f,\{n_k\}}$. Lemma 5.3 (with $K=1$) shows us that the measure does not depend on the base used.

Let μ_n be the measure obtained by giving each point of $\mathrm{Per}_n(X)$ measure $1/N_n(f)$. Then $\mu_{n_k} \to \mu_{f,\{n_k\}}$ weakly (see Corollary 6.7).

(5.4) THEOREM. *Suppose $f: X \to X$ is C-dense. The measures $\mu_{f,\{n_k\}}$ are all equivalent in the sense of* 5.3. *They are positive on nonempty open sets and $\mu_{f,\{n_k\}}(\{x\})=0$ unless $X=\{x\}$. f is an automorphism of $(\mathcal{M}, \mu_{f,\{n_k\}})$.*

Proof. Let $\mu_{f,\{n_k\}}$ and $\mu_{f,\{m_k\}}$ be defined using bases Ψ_1 and Ψ_2 respectively. By 3.9(iii) there is a $K^*>0$ such that, if $B_1 \supset \bar{B}_2$, then

$$\alpha_{\{n_k\}}(B_1) \geqq \varphi(B_1) \geqq K^*\theta(B_2) \geqq \alpha_{\{n_k\}}(B_2).$$

5.3 gives equivalence.

If $U \neq \varnothing$ is open, then $U \supset \bar{B} \neq \varnothing$ for some $B \in \Psi$. Then, using 3.9(ii), $\mu_{f,\{n_k\}}(U) \geqq \alpha_{\{n_k\}}(B) \geqq \varphi(B) > 0$. Suppose $x \in X$ but $X \neq \{x\}$. Let

$$U_m = \{y \in X : d(f^k(y), f^k(x)) < \tfrac{1}{2}\delta^* \text{ for } 0 \leqq k \subset m\}.$$

Let $B_m \in \Psi$ with $x \in B_m \subset U_m$. Then $\mu_{f,\{n_k\}}(\{x\}) \leqq \alpha_{\{n_k\}}(B_m) \leqq \theta(U_m)$. By 3.9(b) there are m_0 and $S>0$ with $\theta(U_m) \leqq 1/sN_m(f)$ for all $m \geqq m_0$. By 4.3 and 4.2

$$h(f) = \lim \frac{1}{m} \log N_m(f) > 0.$$

Thus $N_m(f) \to \infty$, $\theta(U_m) \to 0$ and $\mu_{f,\{n_k\}}(\{x\})=0$.

Now Ψ, $\alpha_{\{n_k\}}$ and $f\Psi$, $\alpha_{\{n_k\}}$ clearly satisfy the hypotheses of 5.3 with $K=1$ (by the obvious and crucial fact that f permutes $\mathrm{Per}_n(X)$). Hence

$$f\mu_{f,\{n_k\}} = fm_{\Psi,\alpha_{\{n_k\}}} = m_{f\Psi,\alpha_{\{n_k\}}} = m_{\Psi,\alpha_{\{n_k\}}} = \mu_{f,\{n_k\}}.$$

(5.5) REMARK. Above we assumed $f: X \to X$ is C-dense. Suppose $f: X \to X$ satisfying Axiom A* is only assumed to be topologically transitive. Then $X = X_1 \cup \cdots \cup X_m$ with $f(X_i) = X_{i+1}$ ($X_{m+1} = X_1$) and $f^m: X_1 \to X_1$ C-dense. From an invariant measure μ for $f^m: X_1 \to X_1$ we get one μ' for $f: X \to X$ by defining $\mu'(f^n E) = \mu(E)/m$ for $E \subset X_1$ measurable. This gives a bijection between invariant Borel measures for $f^m: X_1 \to X_1$ and $f: X \to X$. One sees that μ' is ergodic if and only if μ is, $h(f^m|X_1) = mh(f)$ and $h_\mu(f^m|X_1) = mh_{\mu'}(f)$. The measures defined above, in terms of periodic points of $f^m|X$, correspond to measures on X defined in terms of periodic points of $f\colon X \to X$. We shall study the C-dense case and this will give us results also for the general transitive case.

6. **Ergodicity and equality of measures.**

(6.1) DEFINITION. f is said to be *partially mixing* with respect to the f-invariant measure μ if there is an $R>0$ such that for any $E, F \in \mathcal{M}$,

$$\liminf_{n\to\infty} \mu(E \cap f^{-n}F) \geqq R\mu(E)\mu(F).$$

If $c_1 < c_2 < \cdots < c_r$ are integers, set $I(c_1, \ldots, c_r) = \min_i (c_{i+1} - c_i)$. f is *partially mixing in order r* if there is an $R_r > 0$ such that, if $E_1, \ldots, E_r \in \mathcal{M}$ and $I(c_1^n, \ldots, c_r^n) \to \infty$ as $n \to \infty$, then

$$\liminf_{n\to\infty} \mu(f^{-c_1^n}E_1 \cap \cdots \cap f^{-c_r^n}E_r) \geqq R_r\mu(E_1)\cdots\mu(E_r).$$

Notice that partially mixing is a stronger condition than ergodicity or weak mixing.

(6.2) THEOREM. *If $f\colon X \to X$ is C-dense, then f is partially mixing in all orders with respect to each $\mu = \mu_{f,\{n_k\}}$.*

Proof. Let $I(c_1^n, \ldots, c_r^n) \to \infty$. Let $\alpha = \frac{1}{8}\delta^*$; by 3.9(i) choose n_0 and $S>0$ so that $N_n(f) \geqq SN(n, 2\alpha)$ for all $n \geqq n_0$.

Suppose $E_1, \ldots, E_r$ are closed and $V_i \supset E_i$ with $V_i \in \Psi$. Choose $\varepsilon > 0$ so that $B_\varepsilon(E_i) \subset V_i$. Choose k large enough so that $n_k > 2D(\varepsilon)$ (see 2.4) and n so that $I(c_1^n, \ldots, c_r^n) > M(\alpha) + n_k$. Let $\tau_i = \{[c_i^n - D(\varepsilon), c_i^n + n_k - D(\varepsilon))\}$ and for $x \in \mathrm{Per}_{n_k}(V_i)$ define the specification s_x by $\tau(s_x) = \tau_i$ and $P_{s_x}(t) = f^{t-c_i^n}(x)$; let $A_i = \{s_x : x \in \mathrm{Per}_{n_k}(V_i)\}$. One notes now that $B = A_1 \wedge \cdots \wedge A_r$ is an 8α-separated s-set which is $M(\alpha)$-delayed. Also, by 2.4, we get

$$U(B, \alpha) \subset \bigcap_{i=1}^r f^{-c_i^n}B_\varepsilon(E_i) \subset \bigcap_{i=1}^r f^{-c_i^n}V_i.$$

By 3.7, we get

$$N_d\left(\bigcap f^{-c_i^n}V_i\right) \geqq N_d(U(B, \alpha)) \geqq \frac{K(r, \alpha)\,\mathrm{card}\,(B)\,N(d, \delta^*)}{N(rn_k, \frac{1}{2}\delta^*)}$$

for d sufficiently large. Now

$$N(d, \delta^*) \geqq N_d(f), \quad \mathrm{card}\,(B) = \prod N_{n_k}(V_i)$$

and, using 3.2(iii),

$$N(rn_k, \tfrac{1}{2}\delta^*) \leqq N(n_k, \tfrac{1}{4}\delta^*)^r \leqq N_{n_k}(f)^r/S^r.$$

Combining all these,

$$\frac{N_d(\bigcap f^{-c_i^n}V_i)}{N_d(f)} \geqq R_r \prod \frac{N_{n_k}(V_i)}{N_{n_k}(f)}$$

where $R_r = K(r, \alpha)S^r > 0$. Letting $d \to \infty$,

$$\varphi(\bigcap f^{-c_i^n} V_i) = \liminf_{d\to\infty} \frac{N_d(\bigcap f^{-c_i^n} V_i)}{N_d(f)} \geqq R_r \prod \frac{N_{n_k}(V_i)}{N_{n_k}(f)}.$$

This being true for all big n,

$$\liminf_{n\to\infty} \varphi(\bigcap f^{-c_i^n} V_i) \geqq R_r \prod \frac{N_{n_k}(V_i)}{N_{n_k}(f)}.$$

Letting $n_k \to \infty$,

$$\liminf_{n\to\infty} \varphi(\bigcap f^{-c_i^n} V_i) \geqq R_r \prod \alpha_{\{n_k\}}(V_i) \geqq R_r \prod \mu(E_i).$$

Now suppose $V_i^1 \supset E_i$ open and choose the V_i above so that $V_i^1 \supset \overline{V}_i$. Then

$$\bigcap_i f^{-c_i^n} V_i^1 \supset \mathrm{Cl}\left(\bigcap_i f^{-c_i^n} V_i\right).$$

Choose $B \in \Psi$ so that

$$\bigcap f^{-c_i^n} V_i^1 \supset \overline{B} \supset \bigcap f^{-c_i^n} V_i.$$

Then

$$\mu(\bigcap f^{-c_i^n} V_i^1) \geqq \alpha_{\{n_k\}}(B) \geqq \varphi(\bigcap f^{-c_i^n} V_i)$$

and

$$\liminf_{n\to\infty} \mu(\bigcap f^{-c_i^n} V_i^1) \geqq R_r \prod \mu(E_i).$$

Now

$$\mu(\bigcap f^{-c_i^n} E_i) \geqq \mu(\bigcap f^{-c_i^n} V_i^1) - \sum \mu(V_i^1 \backslash E_i).$$

Letting $\mu(V_i \backslash E_i) \to 0$ we get

$$\liminf_{n\to\infty} \mu(\bigcap f^{-c_i^n} E_i) \geqq R_r \prod \mu(E_i).$$

For any $E_i^* \in \mathscr{M}$ consider $E_i \in E_i^*$ closed. Then

$$\liminf_{n\to\infty} \mu(\bigcap f^{-c_i^n} E_i^*) \geqq \liminf_{n\to\infty} \mu(\bigcap f^{-c_i^n} E_i) \geqq R_r \prod \mu'(E_i).$$

Now let $\mu(E_i) \to \mu(E_i^*)$.

(6.3) Corollary. *Suppose $f: X \to X$ satisfying Axiom* A* *is topologically transitive. Then the measure μ^* on X corresponding to $\mu_{f^m,\{n_k\}}$ on one of its C-dense factors is ergodic under f.*

Proof. See Remark 5.5.

The following standard fact was pointed out to us by W. Parry.

(6.4) Lemma. *Suppose $f: X \to X$ is an ergodic automorphism of two equivalent normalised Borel measures m_1 and m_2. Then $m_1 = m_2$.*

Proof. Let dm_1/dm_2 denote the Radon-Nikodym derivative. It is f-invariant, hence a constant (clearly 1) by ergodicity.

(6.5) THEOREM. *Let $f: X \to X$ be C-dense. Then all the $\mu_{f,\{n_k\}}$ have a common value μ_f.*

Proof. 5.4, 6.2, and 6.4.

(6.6) THEOREM. *Let $f: X \to X$ be C-dense. If K is closed and $\mu_f(K)=0$, then*

$$\lim_{n\to\infty} (N_n(K)/N_n(f)) = 0.$$

If U is open with $\mu_f(\partial U)=0$, then $\lim (N_n(U)/N_n(f))=\mu_f(U)$.

Proof. Suppose $\{m_j\}$ is an increasing sequence of integers so that either

$$N_{m_j}(K)/N_{m_j}(f) \to a > 0 \quad \text{or} \quad N_{m_j}(U)/N_{m_j}(f) \to b \neq \mu_f(U).$$

Let ψ be a countable base closed under finite union and $\{n_k\}$ a subsequence of $\{m_j\}$ so that $\mu_{f,\{n_k\}}$ is defined with ψ.

Suppose $N_{m_j}(K)/N_{m_j}(f) \to a>0$. If $B\supset K$, $B\in\psi$, then

$$\alpha_{\{n_k\}}(B) = \lim \frac{N_{n_k}(B)}{N_{n_k}(f)} \geqq \lim \frac{N_{n_k}(K)}{N_{n_k}(f)} = a.$$

It follows that $\mu_f(K)=\inf \alpha_{\{n_k\}}(B)\geqq a>0$, a contradiction. Suppose $N_{m_j}(U)/N_{m_j}(f) \to b\neq\mu_f(U)$. For $B\supset \overline{U}$, $B\in\psi$ we have $\alpha_{\{n_k\}}(B)\geqq b$; hence $\mu_f(\overline{U})=\mu_{f,\{n_k\}}(\overline{U})\geqq b$. For $\overline{B}\subset U$, $B\in\psi$, we have $\alpha_{\{n_k\}}(B)\leqq b$; hence $\mu_f(U)\leqq b$. As $\mu_f(\partial U)=0$, $b\geqq\mu_f(U) =\mu_f(\overline{U})=b$ and so $\mu_f(U)=b$, a contradiction.

(6.7) COROLLARY. *Let $f: X \to X$ be C-dense. Then, for any $F\in C(X)$,*

$$\frac{1}{N_n(f)} \sum_{x\in \mathrm{Per}_n(f)} F(x) \to \int F\, d\mu_f$$

as $n\to\infty$. (We say that μ_f is derived from f by periodic points to mean the above statement.)

Proof. Choose b such that $-b<F(x)<b$ for all $x\in X$. Let $\varepsilon>0$. Choose $-b =a_0<a_1<\cdots<a_r=b$ with $a_{i+1}-a_i<\varepsilon$, $\mu_f(\{x : F(x)=a_i\})=0$ and $F(x)=a_i$ for no periodic point x.

Let $U_i=\{x : a_{i-1}<F(x)<a_i\}$. Choose $N(\varepsilon)$ so big that

$$|(N_n(U_i)/N_n(f))-\mu_f(U_i)| < \varepsilon/b$$

for all $n\geqq N(\varepsilon)$ and each i. This is possible since $F(\partial U_i)\subset\{a_{i-1}, a_i\}$ and so $\mu_f(\partial U_i)=0$ by construction; hence 6.6 applies to U_i. We also have

$$\left| N_n(f)^{-1} \sum_{x\in \mathrm{Per}_n(f)} F(x) - \sum_{i=1}^{r} a_i(N_n(U_i)/N_n(f)) \right| \leqq \varepsilon.$$

Putting our above two inequalities together one sees that

$$\left| N_n(f)^{-1} \sum_{x\in \mathrm{Per}_n(f)} F(x) - \sum a_i\mu_f(U_i) \right| \leqq 2\varepsilon.$$

Since $|\int F\,d\mu_f - \sum a_i\mu_f(U_i)| \leqq \varepsilon$, we finally get

$$\left|\int F\,d\mu_f - N_n(f)^{-1}\sum_{x\in \mathrm{Per}_n(f)} F(x)\right| \leqq 3\varepsilon$$

for all $n \geqq N(\varepsilon)$.

7. **The algebraic case.** Suppose $f\colon G \to G$ is an automorphism of an n-dimensional torus G. f is a *hyperbolic* if $Df\colon T_eG \to T_eG$ has no eigenvalues on the unit circle. Then (see [16]) f satisfies Axiom A* and is C-dense because G is connected (using 2.7). f of course preserves the normalized Haar measure m on G.

(7.1) PROPOSITION. *If f is a hyperbolic automorphism of a torus, then $\mu_f = m$.*

Proof. Suppose $g \in G$ and $E \subset G$ is closed. Let $\mu_f = \mu_{f,\{n_k\}}$ be defined via the base Ψ. Consider $B \in \Psi$ with $B \supset E+g$. There are $B^1 \in \Psi$ and open V such that $B^1 \supset E$, $g \in V$ and $B^1 + V \subset B$. By 3.9(ii) there is an N such that $N_n(V) > 0$ for all $n \geqq N$. For $n_k \geqq N$ and $g_{n_k} \in \mathrm{Per}_{n_k}(V)$ we have $g_{n_k} + \mathrm{Per}_{n_k}(B^1) \subset B$. If $x \in \mathrm{Per}_{n_k}(B^1)$, then as f is a group automorphism $f^{n_k}(g_{n_k}+x) = f^{n_k}(g_{n_k}) + f^{n_k}(x) = g_{n_k} + x$; so $g_{n_k} + x \in \mathrm{Per}_{n_k}(B)$. Thus $N_{n_k}(B) \geqq N_{n_k}(B^1)$ for $n_k \geqq N$ and $\alpha_{\{n_k\}}(B) \geqq \alpha_{\{n_k\}}(B^1) \geqq \mu_{f,\{n_k\}}(E)$. Varying B, $\mu_{f,\{n_k\}}(g+E) \geqq \mu_{f,\{n_k\}}(E)$. Using $-g$ instead of g, $\mu_{f,\{n_k\}}(g+E) \leqq \mu_{f,\{n_k\}}(E)$. Thus $\mu_f(E) = \mu_f(g+E)$ for all $g \in G$ and E closed; it follows that μ_f is Haar measure.

Now let G be a torus acting freely on a compact metric space X (i.e. $g_1x = g_2x$ implies $g_1 = g_2$) and let μ be normalized Haar measure on G. Let $\pi\colon X \to X_G = X/G$ be the projection map. Now suppose X_G has a normalized Borel measure m_G. Suppose $F \in C(X)$. If $\pi(x_1) = \pi(x_2) = y$, then

$$\int_G F(gx_1)\,d\mu = \int_G F(gx_2)\,d\mu$$

for $x_1 = g_1x_2$ for some $g_1 \in G$ and then $F(gx_1) = F(g_1gx_2)$ is obtained from $F(gx_2)$ (as a function on G) by translating the variable. Denote this common value by $H_F(y)$; $H_F \in C(X_G)$. Define a measure m on X by

$$\int_X F\,dm = \int_{X_G} H_F\,dm_G.$$

Now suppose $S\colon X \to X$ is a homeomorphism and $\sigma\colon G \to G$ an automorphism such that $S(gx) = \sigma(g)S(x)$. Then S induces a homeomorphism S_G of X_G such that $\pi \circ S = S_G \circ \pi$. If S_G preserves m_G, then S preserves m and we say (S, m) is a σ-*extension* of (S_G, m_G).

(7.2) PROPOSITION. *Let (S, m) be a σ-extension of (S_G, m_G) with σ a hyperbolic automorphism of the torus. If m_G is derived from S_G by periodic points, then m is derived from S by periodic points.*

Proof. Let $F \in C(X)$ and $\varepsilon > 0$. Choose $x_1, \ldots, x_s \in X$ such that for each $x \in X$

there is an x_i such that $|F(gx)-F(gx_i)| \leqq \varepsilon/3$ for all $g \in G$. Since μ is derived from σ by periodic points (see 6.7), there is an $N(\varepsilon)$ such that

$$\left| N_n(\sigma)^{-1} \sum_{g \in \mathrm{Per}_n(\sigma)} F(gx_i) - \int_G F(gx_i)\, d\mu \right| \leqq \varepsilon/3$$

for any $n \geqq N(\varepsilon)$. Combining the above inequalities we get

$$\left| N_n(\sigma)^{-1} \sum_{g \in \mathrm{Per}_n(\sigma)} F(gx) - \int_G F(gx)\, d\mu \right| \leqq \varepsilon$$

for any $x \in X$ and any $n \geqq N(\varepsilon)$.

Recall that $\int_X F\, dm = \int_{X_G} H_F\, dm_G$ where $H_F(\pi(x)) = \int_G F(gx)\, d\mu$. As m_G is derived from S_G by periodic points there is an $M \geqq N(\varepsilon)$ such that

$$\left| \int_{X_G} H_F\, dm_G - N_n(S_G)^{-1} \sum_{y \in \mathrm{Per}_n(S_G)} H_F(y) \right| \leqq \varepsilon$$

for any $n \geqq M$. At this stage of the proof we need the following.

Lemma. *If $S_G^n(y)=y$, then $S^n(x)=x$ for some $x \in \pi^{-1}(y)$.*

Proof. Let $z \in \pi^{-1}(y)$. Then $S^n(z)=g_1 z$ for some $g_1 \in G$, $S^n(gz)=\sigma^n(g)g_1 z$. We want to solve $S^n(gz)=gz$ or $g=\sigma^n(g)g_1$. In additive notation $(\sigma^n - I)g = -g_1$. Since σ^n is hyperbolic, there is such a g. Let $x=gz$. By this lemma for $y \in \mathrm{Per}_n\,(S_G)$ choose $x_y \in \pi^{-1}(y) \cap \mathrm{Per}_n\,(S)$. Then

$$\left| H_F(y) - N_n(\sigma)^{-1} \sum_{g \in \mathrm{Per}_n(\sigma)} F(gx_y) \right| \leqq \varepsilon.$$

Now $gx_y \in \mathrm{Per}_n\,(S)$ if and only if $\sigma^n(g)x_y = \sigma^n(g)S^n(x_y) = S^n(gx_y) = gx_y$, i.e. if and only if $g \in \mathrm{Per}_n\,(\sigma)$. Thus

$$\mathrm{Per}_n\,(S) = \{gx_y : g \in \mathrm{Per}_n\,(\sigma),\, y \in \mathrm{Per}_n\,(S_G)\}$$

(for clearly $z \in \mathrm{Per}_n\,(S)$ implies $\pi(z) \in \mathrm{Per}_n\,(S_G)$). Thus

$$N_n(S_G)^{-1} \sum_{y \in \mathrm{Per}_n(S_G)} N_n(\sigma)^{-1} \sum_{g \in \mathrm{Per}_n(\sigma)} F(gx_y) = N_n(S)^{-1} \sum_{z \in \mathrm{Per}_n(S)} F(z).$$

Hence, as $\int_X F\, dm = \int_{X_G} H_F\, dm_G$, we have

$$\left| \int F\, dm - N_n(S)^{-1} \sum_{z \in \mathrm{Per}_n(S)} F(z) \right| \leqq 2\varepsilon$$

for all $n \geqq M$.

Suppose $f\colon N/\Gamma \to N/\Gamma$ is a hyperbolic automorphism of a nilmanifold (one can see [13] or [16] for the definition). Then N/Γ has a unique normalized Borel measure m which is invariant under the action of N; m is f-invariant. It is well known that (f, m) is obtained through a succession of extensions via hyperbolic toral automorphisms with a single point as the initial base space. By 7.2 we have that m is derived from f by periodic points.

(7.3) THEOREM. *If f is a hyperbolic automorphism of a nilmanifold, then $\mu_f = m$.*

Proof. f satisfies Axiom A* and is C-dense since N/Γ is connected (by 2.7). 6.7 says that μ_f is derived from f by periodic points. At most one measure can be derived from f by periodic points.

(7.4) REMARK. Conversations with W. Parry, S. Smale, and P. Walters were helpful in finding a proof for 7.3. Parry in particular pointed out how the periodic points of S are related to those of S_G and σ. Hyperbolic automorphisms of nilmanifolds thus distribute their periodic points uniformly with respect to the usual measure. For this particular case §§6 and 8 yield already known facts (see [2] or [13] for example).

8. **The entropy of μ_f.** We refer the reader to [5] for a definition of measure theoretic entropy.

(8.1) Suppose $f: X \to X$ satisfying Axiom A* is topologically transitive. Then $h_{\mu_f}(f) = h(f)$.

Proof. By 5.5 we may assume f is C-dense. Cover X by open sets $U_1, \ldots, U_r$ with diam $U_i < \delta^*$. Choose disjoint Borel sets $A_1, \ldots, A_r$ such that $U_i \supset \bar{A}_i$ and $X = \bigcup_{i=1}^r A_i$. In [8] L. Goodwyn shows that for any f-invariant normalized Borel measure ρ on X (and $f: X \to X$ any continuous map) we have $h_\rho(f) \leqq h(f)$. We complete our proof by showing the partition $\beta = \{A_1, \ldots, A_r\}$ satisfies $h_{\mu_f}(f, \beta) \geqq h(f)$. For any $1 \leqq i_0, \ldots, i_{m-1} \leqq r$ consider the sets

$$V = \bigcap_{k=0}^{m-1} f^{-k} U_{i_k} \supset \bigcap_{k=0}^{m-1} f^{-k} A_{i_k} = D(i_0, \ldots, i_{m-1}).$$

By 3.9(v) there are m_0 and $S > 0$ such that $\theta(V) \leqq 1/SN_m(f)$ for all $m \geqq m_0$. Then $\mu_f(D) \leqq \theta(V) \leqq 1/SN_m(f)$. Define the function

$$h_m = \frac{1}{m} \sum_{(i_0, \ldots, i_{m-1})} (-\log \mu_f(D))\chi_D$$

where χ_D is the characteristic function of D. For $m \geqq m_0$ we have

$$-\log \mu_f(D) \geqq \log S + \log N_m(f).$$

By definition

$$\int_{h_m} d\mu_f \to h_{\mu_f}(f, \beta)$$

as $n \to \infty$. Hence, using 4.5,

$$h_{\mu_f}(f, \beta) \geqq \lim \frac{1}{m} [\log N_m(f) + \log S] = h(f).$$

REFERENCES

1. R. L. Adler, A. G. Konheim and M. H. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc. **114** (1965), 309–319. MR **30** #5291.

2. D. V. Anosov and Ja. G. Sinaĭ, *Certain smooth ergodic systems*, Uspehi Mat. Nauk **22** (1967), no. 5 (137), 107–172 = Russian Math. Surveys **22** (1967), 103–167. MR **37** #370.

3. A. Avez, *Propriétés ergodiques des endomorphismes dilatants des variétés compactes*, C. R. Acad. Sci. Paris Sér. A-B **266** (1968), A610–A612. MR **37** #6944.

4. K. Berg, *Convolution of invariant measures, maximal entropy*, (to appear).

5. P. Billingsley, *Ergodic theory and information*, Wiley, New York, 1965. MR **33** #254.

6. R. Bowen, *Topological entropy and Axiom* A, Proc. Sympos. Pure Math., vol. 14, Amer. Math. Soc., Providence, R. I., 1970.

7. ———, *Markov partitions for Axiom* A *diffeomorphisms*, Amer. J. Math. (to appear).

8. L. W. Goodwyn, *Topological entropy bounds measure-theoretic entropy*, Proc. Amer. Math. Soc. **23** (1969), 679–688.

9. M. Hirsch, *Expanding maps and transformation groups*, Proc. Sympos. Pure Math., vol. 14, Amer. Math. Soc., Providence, R. I., 1970.

10. M. Hirsch and C. Pugh, *Stable manifolds and hyperbolic sets*, Proc. Sympos. Pure Math., vol. 14, Amer. Math. Soc., Providence, R. I., 1970.

11. K. Krzyżewski and W. Schlenk, *On invariant measures for expanding differentiable mappings*, Studia Math. **33** (1969), 83–92.

12. W. Parry, *Intrinsic Markov chains*, Trans. Amer. Math. Soc. **112** (1964), 55–66. MR **28** #4579.

13. ——, *Ergodic properties of affine transformations and flows on nilmanifolds*, Amer. J. Math. **91** (1969), 757–777.

14. Ja. G. Sinaĭ, *Markov partitions and Y-diffeomorphisms*, Funkcional. Anal. i Priložen. **2** (1968), 64–89 = Functional Anal. Appl. **2** (1968), 61–82. MR **38** #1361.

15. ———, *Classical dynamic systems with countably-multiple Lebesgue spectrum*. II, Izv. Akad. Nauk SSSR Ser. Mat. **30** (1966), 15–68; English transl., Amer. Math. Soc. Transl. (2) **68** (1968), 34–88. MR **33** #5847.

16. S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817. MR **37** #3598.

17. R. F. Williams, *Classification of one dimensional attractors*, Proc. Sympos. Pure Math., vol. 14, Amer. Math. Soc., Providence, R. I., 1970.

18. W. H. Gottschalk and G. A. Hedlund, *Topological dynamics*, Amer. Math. Soc. Colloq. Publ., vol. 36, Amer. Math. Soc., Providence, R. I., 1955. MR **17**, 650.

19. W. Rudin, *Real and complex analysis*, McGraw-Hill, New York, 1966. MR **35** #1420.

University of Warwick,
Coventry, England