1. Introduction. We shall study the distribution of periodic points for a class of diffeomorphisms defined by Smale [16, §1.6].

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 \( U \cap \bigcup_{n \neq 0} f^n(U) = \emptyset \); 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 \( \Lambda \) of \( M \) is hyperbolic under \( f \) if the tangent bundle of \( M \) restricted to \( \Lambda \), \( T_\Lambda(M) \), 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).

If 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 \( \Omega(f) = \Omega_1 \cup \cdots \cup \Omega_m \) where the \( \Omega_i \) are disjoint closed \( f \)-invariant sets and \( f|\Omega_i \) is topologically transitive (the \( \Omega_i \) 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 \( \mu_f \) on \( \Omega_i \). In the algebraic case \( \mu_f \) turns out to be Haar measure.

We show that \( \mu_f \) is ergodic, positive on open sets and zero on points (unless \( \Omega_i \) 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 \( \mu_f \) is the unique invariant normalized Borel measure on \( \Omega_i \) 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 \( \mu_f \). First we decompose \( \Omega_f = X_1 \cup \cdots \cup X_m \) 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...
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 $f: \Omega_i \to \Omega_i$ is $C$-dense. What we want is a measure $\mu_f$ 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 $\Omega_i$ (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 $g: M \to M$ 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 $\delta > 0$,

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

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 $\epsilon(\delta) > 0$ such that $W^s_\delta(x) \cap W^s_\delta(z) \neq \emptyset$ whenever $d(x, z) \leq \epsilon(\delta)$.

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

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

and

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

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 $f: X \to X$ is a homeomorphism satisfying Axiom A*.
(2.2) Easy facts. (i) \( f^n W^u(x) = W^u(f^n(x)) \).
(ii) For \( n \geq 0 \), \( f^{-n} W^s_\delta(x) \subseteq W^s(f^{-n}(x)) \).
(iii) If \( y \in W^s_\delta(x) \), then \( W^s_\delta(y) \subseteq W^s_{\delta_1 + \delta_2}(x) \).
(iv) Let \( f^m(x) = x \) and \( \delta \leq \delta^* \). Then \( f^{m(k+1)} W^s_\delta(x) \supseteq f^{mk} W^s_\delta(x) \) and (by A3)
\[ W^u(x) = \bigcup_{k=0}^{\infty} f^{mk} W^s_\delta(x). \]

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

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

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

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

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

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 \( f^*: X_i \rightarrow X_i \) is \( C \)-dense when \( g^*(i) = i \).

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

(a) \( X(p) \) is open.

Proof. Let \( a = \epsilon(\delta^*) \). 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 \( M = \text{ord} p \cdot \text{ord} q \). By A2 choose \( z \in W^u_M(x) \cap W^u_M(q) \). Then \( z \in W^u(p) \) and
\[ d(f^kM(z), q) = d(f^kM(z), f^kM(q)) \rightarrow 0 \quad \text{as} \quad k \rightarrow +\infty. \]
Since \( f^kM W^u(p) \subseteq W^u(p) \), we get \( q \in \text{Cl} (W^u(p)) = X(p) \). (Note: We use 2.1 without explicit mention.)

(b) \( X(p) = X(q) \) or \( X(p) \cap X(q) = \emptyset \).

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 \( M = \text{ord} p \cdot \text{ord} q \). Then as \( k \rightarrow +\infty, f^{-kM}(w) \rightarrow q \).
But \( f^{-M} X(p) = X(p) \) since \( f^{-M} W^u(p) = W^u(p) \). Thus \( q \in \text{Cl} (X(p)) = X(p) \). By (a) we have \( X(p) \supset W^u_\delta(q) \).

Since
\[ W^u(q) \subseteq \bigcup_{k=0}^{\infty} f^{kM} W^u_\delta(q). \]
and \( f^m X(p) = X(p) \), 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 \( f(p_i) \in X(p_i) \). That \( f \) is a homeomorphism and (c) below show that \( g \) is a permutation.

(c) \( f(X_i) = X(p_i) \).

**Proof.** As \( f \) is a homeomorphism, \( f X(p_i) = X(f(p_i)) \) follows from \( f W^u(p_i) = W^u(f(p_i)) \). Since \( f(p_i) \in X(f(p_i)) \cap X(p_i) \), \( X(f(p_i)) = X(p_i) \) by (b).

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

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

(2.8) Lemma. Let \( f : X \to X \) be C-dense and \( \alpha > 0 \). Then there is an \( N \) such that \( f^n W^a_\alpha(x) \cap W^b_\alpha(y) \neq \emptyset \) whenever \( x, y \in X \) and \( m \geq N \).

**Proof.** Set \( \delta = \min \{ \delta_*, \delta_* \epsilon(\delta) \} \) and choose \( p_0, \ldots , p_r \) periodic such that every \( x \in X \) is within \( \epsilon(\epsilon) \) of some \( p_k \). Let \( t_k \) be the period of \( p_k \). By 2.2 and \( \text{Cl}(W^u(p_0)) = X \), there is a \( m_k \) such that every \( y \in X \) is within \( \epsilon(\delta) \) of \( f^m x \).

(2.9) Definitions. Let \( \text{Per}_n(U) = \{ x \in U : f^n(x) = x \} \), \( N_n(U) = \text{card}(\text{Per}_n(U)) \), and \( \text{N}_n(f) = N_n(X) \).

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

Finally, for \( \epsilon > 0 \), let

\[ U(s, \epsilon) = \{ x \in X : d(f^t(x), P_t(t) < \epsilon \text{ for all } t \in \text{Tim}(s)) \} \]

(2.10) Theorem. Suppose \( f : X \to X \) is C-dense and \( \epsilon > 0 \). There is an \( M(\epsilon) \) such that \( U(s, \epsilon) \neq \emptyset \) whenever \( s \) is an \( M(\epsilon) \)-delayed \( f \)-specification. In fact \( M(\epsilon) \) can be chosen so that \( \text{Per}_n U(s, \epsilon) \neq \emptyset \) for all \( d \geq M(\epsilon) + L(s) \).
Proof. We tend $s$ to a new specification $s'$ as follows. Let $a_1$ be the smallest integer in $\text{Tim}(s)$. 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 $\beta = \frac{1}{3} \min \{\psi(\frac{1}{2} \epsilon), \epsilon, \delta^8\}$ ($\psi$ defined in 2.5) and $\alpha = \beta/3c$; let $N$ be the integer given by 2.8 for this $\alpha$. Choose $M = M(\epsilon) \geq \max \{N, R(\frac{1}{2} \epsilon)\}$ ($R$ defined in 2.5) large enough so that $\sum_{j=0}^{\infty} 2^{M_j} < 2$. Assume $d \geq M(\epsilon) + L(s)$; 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 $1 \leq k \leq m$) recursively as follows. Suppose $z_k$ has been chosen for some $1 \leq k < m$. As $s^1$ is $M$-delayed, $a_{k+1} - b_k \geq M \geq N$ and so by 2.8 there exists a point

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

Set $z_{k+1} = f^{a_k + 1}(v)$; then $f^{a_k}(z_{k+1}) \in W^{a}_{\alpha}(f^{a_k}(z_k))$ and $f^{a_k + 1}(z_{k+1}) \in W^{a}_{\alpha}(P_{s}(a_{k+1}))$.

By induction on $r$ we show that

$$f^{a_k}(z_{k+r}) \in W^{a}_{c_1 + \cdots + c_{a_k + 1}}(f^{a_k}(z_k)).$$

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

$$f^{a_k}(z_{k+r+1}) \in W^{a}_{c_1 + \cdots + c_{a_k + 1}}(f^{a_k}(z_{k+r})).$$

and so our induction is done.

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

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

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

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

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

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

$$d(f^{t}(z), f^{t}(z^*)) \leq \frac{1}{2} \epsilon \quad \text{for all } 0 \leq t \leq d.$$

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

$$d(f^{t}(z^t), f^{t}(z_m)) \leq \frac{1}{2} \epsilon \quad \text{for all } a_1 \leq t \leq a_1 + d.$$
Applying the triangle inequality to this and \( z_m \in U(s^1, \beta) \),
\[
z^1 \in U(s^1, \beta + \frac{1}{2} \epsilon) \subseteq U(s^1, \epsilon) \subseteq U(s, \epsilon);
\]
also \( z^1 \in \text{Per}_d (X) \).

(2.11) \textbf{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) \textbf{Definition.} For \( \epsilon > 0 \), \( E \subset X \) is an \((n, \epsilon)\)-separated set if for any distinct \( x, y \in E \) there is a \( t \) for which \( 0 \leq t < n \) and \( d(f^t(x), f^t(y)) > \epsilon \). We let \( N(n, \epsilon) \) denote the maximum cardinality of an \((n, \epsilon)\)-separated set.

(3.2) \textbf{Lemma.} (i) If \( \epsilon \leq \delta^* \), then \( N(n, \epsilon) \leq N_d(f) \).

(ii) If \( \epsilon \leq \alpha \), then \( N(n, \alpha) \leq N(n, \epsilon) \); for any \( \epsilon > 0 \) there is an \( m_\epsilon \) such that \( N(n, \epsilon) \leq N(n + m_\epsilon, \delta^*) \) for all \( n \geq 0 \).

(iii) \( N(\bigcap n_i, \epsilon) \leq \prod N(n_i, \frac{\epsilon}{2}) \).

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

(ii) The first statement is obvious; if \( E \) is an \((n, \epsilon)\)-separated set, then \( \bigcup_{t=0}^{n} E \) is an \((n + 2D(\epsilon), \delta^*)\)-separated set (use 2.4).

(iii) We prove the following stronger statement for later use: Suppose \( E \subset X \) and \( n_i, m_i (1 \leq i \leq s) \) are integers \( (n_i \geq 0) \) 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 \( d(f^t(x), f^t(y)) > \epsilon \); then \( \text{card} (E) \leq \prod_{i=1}^{s} N(n_i, \frac{\epsilon}{2}) \).

\textbf{Proof.} Choose \( R_i \subset X \) so that \( f^m R_i \) is a maximal \((m_i, \frac{\epsilon}{2})\)-separated set. Construct a map \( g = \bigcap g_i: E \to \bigcap R_i \) by requiring that \( d(f^t(x), f^t(g(x))) \leq \frac{\epsilon}{2} \) for all \( t \in [m_i, m_i + n_i) \). Such a \( g_i(x) \) exists by the maximality of \( f^m R_i \)—otherwise \( f^m(R(x)) \) would be an \((n, \frac{\epsilon}{2})\)-separated set.

If \( g(x) = g(y) \) the triangle inequality would give us \( d(f^t(x), f^t(y)) \leq \epsilon \) 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 \( d(P_c(t), P_{s^1}(t)) > p \) for some \( t \in \text{Tim} (s) \cap \text{Tim} (s^1) \); 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 \( U(A, \epsilon) = \bigcup_{s \in A} U(s, \epsilon) \).

3.3 \textbf{Lemma.} (i) If \( s \) and \( s^1 \) are \( p\)-separated, then \( U(s, \frac{1}{2}p) \cap U(s^1, \frac{1}{2}p) = \emptyset \).

(ii) If \( A \) is a \( 2\epsilon\)-separated \( S\)-set, \( \tau(A) \) is \( M(\epsilon)\)-delayed, and \( d \geq L(A) + M(\epsilon) \), then \( N_d(U(A, \epsilon)) \geq \text{card} (A) \).

\textbf{Proof.} (i) Trivial. (ii) Follows from (i) and 2.10.

Two specifications \( s \) and \( s^1 \) are disjoint if \( \text{Tim} (s) \cap \text{Tim} (s^1) = \emptyset \). In this case we define a new specification \( s \land s^1 \) by \( \tau(s \land s^1) = \tau(s) \cup \tau(s^1) \) and
\[
P_{s \land s^1}(t) = P_s(t) \quad \text{for } t \in \text{Tim} (s),
\]
\[
P_{s \land s^1}(t) = P_{s^1}(t) \quad \text{for } t \in \text{Tim} (s^1).
\]
Notice that $U(s \land s^1, e) = U(s, e) \cap U(s^1, e)$. We call a $G$-time $\tau$ an $m$-time if $\text{card } \tau = m$; $s$ is an $m$-specification if $\tau(s)$ is an $m$-time.

(3.4) **Lemma.** If $\tau$ is an $n$-delayed $m$-time and $N \geq L(\tau)$, there is a $\tau^1$ such that

(a) $\text{Tim} (\tau) \cap \text{Tim} (\tau^1) = \emptyset$,

(b) $\tau \cup \tau^1$ is $n$-delayed,

(c) $L(\tau \cup \tau^1) \leq N$, and

(d) $T(\tau^1) \geq N - 2mn - T(\tau)$.

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

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

This determines a $G$-time $\tau$ which satisfies our condition.

(3.5) **Remark.** $\tau^1$ could be empty.

(3.6) **Lemma.** If $\tau$ is a time specification and $e > 0$, there is an $e$-separated $S$-set $A$ with $\tau(A) = \tau$ and $\text{card } (A) \geq N(T(\tau), 2e)$.

**Proof.** Let $\tau = \{ I_1, \ldots, I_m \}$ and $\tau_k = \{ I_k \}$ for $1 \leq k \leq m$. Let $A_k$ be an $e$-separated $S$-set with $\tau(A_k) = \tau_k$ and $\text{card } (A_k) = N(T(\tau_k), e)$. Then

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

is $e$-separated with $\tau(A) = \tau_1 \land \cdots \land \tau_m = \tau$ and $\text{card } (A) = \prod N(T(\tau_k), e) \geq N(\sum T(\tau_k), 2e) = N(T(\tau), 2e)$ by 3.2(iii).

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

$$K(m, e) \text{card } (B) \leq N(d, 8e) \leq N(T(\tau(B)), 4e)$$

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

**Proof.** Let $N = d - M(e) \geq L(\tau(B))$. Let $\tau = \tau(B)$ and choose $\tau^1$ as in Lemma 3.4. By Lemma 3.5 let $A$ be a $2e$-separated $S$-set with $\tau(A) = \tau^1$ and $\text{card } (A) \geq N(T(\tau^1), 4e)$. Now $A \land B$ is a $2e$-separated $S$-set with $M(e)$-delayed time $\tau \land \tau^1$; $d \geq N + M(e) \geq L(\tau \land \tau^1) + M(e)$. Hence, by 3.3(iii), we have

$$N(d(U(A \land B, e)) \geq \text{card } (A \land B) = \text{card } (A) \text{card } (B).$$

Since $U(B, e) \geq U(A \land B, e)$,

$$N(d(U(B, e)) \geq \text{card } (A) \text{card } (B).$$

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

$$\text{card } A \geq \max \{ 1, N(N - 2mM(e) - T(\tau), 4e) \} = W$$

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

$$N(d, 8e) \leq W \cdot N((2m + 1)M(e), 4e)N(T(\tau), 4e)$$
N_d(U(B, \epsilon)) \geq \text{card } (B) \cdot W \\
\geq \frac{K(m, \epsilon) \text{card } (B) \cdot N(d, 3 \epsilon)}{N(T(\tau), 4 \epsilon)}

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

(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 \( V \neq \emptyset \) is open.

(iii) There is a \( K^* > 0 \) such that \( \varphi(U) \geq K^* \theta(V) \) whenever \( U \) and \( V \) are open in \( X \) and \( U \supseteq V \).

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

(v) There are \( m_0 \) and \( S > 0 \) such that, if \( m \geq m_0 \) and \( U \subset X \) satisfies \( \text{diam } f^n(U) \leq \delta^* \) for all \( 0 \leq k < m \), then \( \theta(U) \leq 1/SN_n(f) \).

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

\[ N_d(f) \geq N_d(V) \geq K(1, \epsilon) N(d, 8\epsilon)/N(1, 4\epsilon) \]

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

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

Trivially, \( U(B_n, \frac{1}{4}\delta^*) \subset U \); so by the theorem

\[ N_d(U) \geq K(1, \frac{1}{4}\delta^*)N_n(V)N(d, \delta^*)/N(n, \frac{1}{4}\delta^*) \]

for \( d \geq n + M(\frac{1}{4}\delta^*) \). By (i) above there is an \( n_0 \) and a \( K_1 \) such that \( N(n, \frac{1}{4}\delta^*) \leq K_1 N_n(f) \) when \( n \geq n_0 \); also \( N(d, \delta^*) \geq N_d(f) \). Thus for \( n \geq n_0 \) and \( d \geq n + M(\frac{1}{4}\delta^*) \) we have

\[ N_d(U)/N_d(f) \geq K^* N_n(V)/N_n(f) \]

where \( K^* = K(1, \frac{1}{4}\delta^*)/K_1 > 0 \). Then \( \varphi(U) \geq K^* \theta(V) \).

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

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
and card $B = N(m-m_0, \frac{1}{2}\delta^*)$. Now $A \wedge B$ is $\frac{1}{2}\delta^*$-separated with $M(\frac{1}{2}\delta^*)$-delayed time.

By 3.3(ii) we have

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

By Proposition 3.2(iii) we have

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

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

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

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

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

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

(3.11) Proposition. For each $e$ with $0 < e < \frac{1}{2}\delta^*$ there are constants $c_e > 0$ and $0 < \tau_e < 1$ for which the following holds. If $A \subseteq X$, $0 \leq k_1 < k_2 < \cdots < k_m$ are integers and $w_{k_1}, \ldots, w_{k_m} \in X$ satisfy $f^k(A) \cap B_{e}(w_{k_r}) = \emptyset$ for $r = 1, \ldots, m$, then $N(n, e, A) \leq c_e\tau^mN(n, e)$ for all $n > k_m$.

Proof. Let $M = M(\frac{1}{2}e)$ 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 $q \geq m/(2M+1)$. Let $n > k_m$ and $E_n \subseteq A$ be an $(n, e)$-separated set. For each $I = J = \{j_1, \ldots, j_q\}$ and each $x \in E_n$ we define the specification $s(x, I)$ by requiring that it be an $M$-delayed specification with

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

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

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

$$p(x, I) \subseteq U(s(x, I), \frac{1}{2}e) \cap \text{Per}_d(X).$$

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 $e$-separated; for if $j_1 \in I_1 \setminus I_2$, then $j_1 \in \text{Tim } s(x, I_1) \cap \text{Tim } s(y, I_2)$ and

$$d(P_{s(x, I_1)}(j_1), P_{s(y, I_2)}(j_1)) = d(w_{j_1}, f^t(y)) > e.$$ 

By lemma (i) we have $p(x, I_1) \neq p(y, I_2)$; thus $I_1 \neq I_2$ implies $F_1 \cap F_2 = \emptyset$.

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

$$t \in [0, n]\text{((Tim } s(x, I)\setminus I) = \bigcup_{j_t\in I} [j_t - M, j_t + M].}$$
By the proof of 3.2(iii), \( \{x \in E_n : p(x, I) = z\} \) has at most \( g^{\text{card} I} \) elements where \( g = N(2M + 1, \frac{1}{2}e) \). Thus \( F_I \) has at least \( \text{card} E_n \cdot g^{\text{card} I} \) elements.

As the \( F_I \)'s are disjoint

\[
N_d(f) \geq \sum_{i \leq j} \text{card} F_i \geq \sum_{i \leq j} \frac{1}{g^{\text{card} I}} \text{card} E_n
\]

\[
\geq \sum_{r = 0}^{\text{card} J} \left( \begin{array}{c} \text{card} J \\ r \end{array} \right) \frac{1}{g^r} \text{card} E_n = \left( 1 + \frac{1}{g} \right)^{\text{card} J} \text{card} E.
\]

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

\[
N_d(f) = N_{n+m}(f) \leq N(n + M, 2e) \leq N(n, e)N(M, e).
\]

Also \( \text{card} J = q \geq m/(2M + 1) \). Thus

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

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

\[
h(f, \mathcal{A}) = \lim_{n \to \infty} \frac{1}{n} \log M_n(f, \mathcal{A})
\]

exists and the topological entropy of \( f \) is defined by

\[
h(f) = \sup \ h(f, \mathcal{A}).
\]

(The above definitions and 4.1 and 4.2 below do not depend on our standing hypothesis that \( f \) satisfies Axiom \( \text{A}^* \); 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 \( C \neq X \neq D \).

(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 \( C \neq \emptyset \) with \( f^{-1}(C) \subset C \) and \( \overline{C} \neq X \). Let \( B \neq \emptyset \) be open with \( \overline{B} \subset C \) and set \( D = X \setminus \overline{B} \). Then \( \{C, D\} \) is as above. Let

\[
E_n = \{(C, \ldots, C, D, \ldots, D) : i + j = n, i, j \geq 0\}.
\]

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 \( 0 \leq k < n \) or there is a largest \( k \), denoted \( k(x) \), such that \( 0 \leq k < n \) and \( f^k(x) \notin D \).
In the latter case \( f^k(x) \in C \) and so \( f^m(x) \in C \) for all \( m \leq k(x) \) as \( f^{-1}(C) \subset C \); \( f^m(x) \in D \) for \( m > k(x) \). As \( \text{card } E_n = n + 1 \), \( M_n(f; \{ C, D \}) \leq n + 1 \) and \( h(C, D) = 0 \) — a contradiction.

Suppose \( \{ C, D \} \) is an open cover of \( Y \) with \( C \neq \overline{Y} \neq D \). 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 \( B_\varepsilon(p) \subset C \setminus D \) and \( B_\varepsilon(q) \subset D \setminus C \). Let \( M(\varepsilon) \) be the integer given by 2.10; set \( N = M(\varepsilon) + 1 \). Then \( \tau_n = \{ \{ kN \} : 0 \leq k < n \} \) 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 \( P_n(kN) = a_k \). 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 \( F_n(x) = (F_n^0(x), \ldots, F_n^{n-1}(x)) \in E_n \) be such that \( f^j(x) \in F_j(x) \) for \( 0 \leq j < nN \). Suppose \( (a_0, \ldots, a_{n-1}) \neq (b_0, \ldots, b_{n-1}) \); say \( a_k = p \) and \( b_k = q \). Then
\[
f^{kN}(x_n(a_0, \ldots, a_{n-1})) \in B_k(p) \subset C \setminus D
\]
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 \( E_n \geq 2^n \) and \( M_n(f; \{ C, D \}) \geq 2^n \); thus
\[
h(f; \{ C, D \}) \geq \lim_{n \to \infty} \frac{1}{nN} \log 2^n = \frac{1}{N} \log 2 > 0.
\]

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

Let \( Q \) be a maximal \( (n, \beta) \)-separated set. For \( z \in Q \) choose \( B(z) = (A_0(z), \ldots, A_{n-1}(z)) \) with \( A_k(z) \in \mathcal{A} \) and
\[
A_k(z) \supset \text{Cl } (B_k(f^k(z))) \quad \text{for all } 0 \leq k < n.
\]
We claim $E_n = \{B(z) : z \in Q\}$ is an $n$-cover for $(f, \mathcal{A})$. For each $x \in X$ there is a $z_x \in Q$ for which $d(f^k(x), f^k(z_x)) \leq \beta$ for all $0 \leq k < n$; otherwise $Q \cup \{x\}$ would be an $(n, \beta)$-separated set bigger than $Q$. Since $f^k(x) \in A_k(z_x)$, $E_n$ is an $n$-cover. We have shown $M_n(f, \mathcal{A}) \leq N_n(\beta)$.

Let $E$ be an $n$-cover for $(f, \mathcal{A})$ and $R$ an $(n, \delta^*)$-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 $0 \leq k < n$. If $g(x) = g(y)$, then $A_k(x) = A_k(y)$ and $d(f^k(x), f^k(y)) \leq \text{diam} A_k(x) < \delta^*$ for $0 \leq k < n$; $x = y$ as $R$ is an $(n, \delta^*)$-separated set. As $g : R \to E$ is injective, $\text{card } E \geq \text{card } R$ and $M_n(f, \mathcal{A}) \geq N(n, \delta^*) \geq N_n(f)$.

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

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

(4.6) **Remark.** Let

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

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 $\delta$ is an expansive constant; if $\varepsilon \leq \delta$, then 3.2(ii) goes through, i.e.

$$N(n, \delta) \leq N(n, \varepsilon) \leq N(n + m, \delta)$$

for some $m_\varepsilon$, 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, $h(f|A) = \gamma_f(\varepsilon)$ for $\varepsilon \leq \delta^*$. Choose $w \in X \setminus A$ and $\varepsilon > 0$ so small that $A \cap B_w(w) = \varnothing$. Recall 3.11, $N(n, \varepsilon, A) \leq c_\varepsilon \tau^n$, for $n > m$ where $\tau < 1$. Then

$$\gamma_f(\varepsilon) = \lim sup \frac{1}{n} \log N(n, \varepsilon, A)$$

$$\leq \lim_{n \to \infty} \frac{1}{n} \log c_\varepsilon \tau^{n-1} \leq \log \tau + \gamma_f(\varepsilon) = \log \tau + h(f) < h(f).$$

5. **Construction of a measure.** Let $\psi$ be a countable base for the topology of $X$ which is closed under finite union. Assume $\omega : \psi \to R$ satisfies, for $B \in \psi$, $\omega(B) \geq 0$, $\omega(X) = 1$,

$$\omega(B_1) \geq \omega(B_2) \text{ when } B_1 \supseteq B_2,$$

$$\omega(B_1 \cup \cdots \cup B_n) \leq \sum \omega(B_i),$$
and
\[ \omega(B_1 \cup B_2) = \omega(B_1) + \omega(B_2) \quad \text{when } B_1 \cap B_2 = \emptyset. \]

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

(5.1) Lemma. If \( U \subseteq \bigcup_{i=1}^{n} U_i \), then \( m(U) \leq \sum m(U_i) \). If \( U \cap V = \emptyset \), then \( m(U \cup V) = m(U) + m(V) \).

Proof. Let \( B \in \psi \) with \( \overline{B} \subseteq U \). By compactness let \( U_1, \ldots, U_n \) cover \( B \). For \( x \in \overline{B} \) choose \( B_x \in \psi \) so that \( B_x \subseteq U_i \) for some \( i \) satisfying \( 1 \leq i \leq n \). Let \( B_{x_1}, \ldots, B_{x_r} \) cover \( B \) and set \( A_i = \bigcup \{ B_{x_j} : B_{x_j} \subseteq U_i \} \). Then
\[ \omega(B) \leq \omega \left( \bigcup_{i=1}^{n} A_i \right) \leq \sum_{i=1}^{n} \omega(A_i) \leq \sum_{i=1}^{n} m(U_i). \]

Now vary \( B \).

By the first part of the lemma, \( m(U \cup V) \leq m(U) + m(V) \). Suppose \( B_1, B_2 \in \psi \) with \( \overline{B}_1 \subseteq U \) and \( \overline{B}_2 \subseteq V \). Then \( \overline{C} (B_1 \cup B_2) \subseteq U \cup V \) and \( \overline{B}_1 \cap \overline{B}_2 = \emptyset \); so
\[ m(U \cup V) \geq \omega(B_1 \cup B_2) = \omega(B_1) + \omega(B_2). \]

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

For any \( E \subseteq X \) we define
\[ m(E) = \inf \{ m(U) : U \supseteq E, U \text{ open} \}. \]

One sees easily that this definition agrees with the earlier one on open sets and that \( m(K) = \inf \{ \omega(B) : B \supseteq K, B \in \psi \} \) when \( K \) is closed. We let
\[ \mathcal{M} = \{ E \subseteq X : m(E) = \sup \{ m(K) : K \subseteq E, K \text{ closed} \} \}. \]

With standard arguments we get

(5.2) Proposition. \( \mathcal{M} = \mathcal{M}_{\psi, \omega} \) is a \( \sigma \)-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 \( \omega_1 : \psi_1 \to \mathbb{R} \) and \( \omega_2 : \psi_2 \to \mathbb{R} \) are as above and there is a \( K > 0 \) such that \( \omega_2(B_2) \geq K \omega_1(B_1) \) when \( B_2 \supseteq B_1 \) and \( \omega_1(B_1) \geq K \omega_2(B_2) \) when \( B_1 \supseteq B_2 \), then \( m_{\psi_1, \omega_1} = m_{\psi_2, \omega_2} \) and \( K m_{\psi_1, \omega_1} \leq m_{\psi_2, \omega_2} \leq (1/K) m_{\psi_1, \omega_1} \).

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

Suppose \( E \in \mathcal{M}_{\psi_1, \omega_1} \). Letting \( K_n \subseteq E \) be compact with \( m_{\psi_1, \omega_1}(K_n) \geq m_{\psi_1, \omega_1}(E) - 1/n \) we see that \( E = E_1 \cup \bigcup_{n=1}^{\infty} K_n \) where \( E_1 \subseteq F \) 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 \( m_{\psi_2, \omega_2} \) is complete. As \( \psi_2, \omega_2 \)}
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 $\omega$'s when we are given a homeomorphism $f: X \to X$ which is C-dense. Let $\psi$ 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 $\mu_n$ be the measure obtained by giving each point of $\text{Per}_n(X)$ 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 $\Psi_1$ and $\Psi_2$ respectively. By 3.9(iii) there is a $K^*>0$ such that, if $B_1 \supset B_2$, then

$$\alpha_{(n_k)}(B_1) \geq R(B_1) \geq K^*R(B_2) \geq \alpha_{(m_k)}(B_2).$$

5.3 gives equivalence.

If $U \neq \emptyset$ is open, then $U \supset B \neq \emptyset$ for some $B \in \Psi$. Then, using 3.9(ii), $\mu_{f,(n_k)}(\emptyset) \geq R(B) \geq R(B_2) > 0$. Suppose $x \in X$ but $X \neq \{x\}$. Let

$$U_m = \{ y \in X : d(f^k(y), f^k(x)) < \frac{1}{k}K^* \mbox{ for } 0 \leq k \leq m \}. $$

Let $B_m \in \Psi$ with $x \in B_m \subset U_m$. Then $\mu_{f,(n_k)}(\{x\}) \leq \alpha_{(n_k)}(B_m) = \alpha_{(m_k)}(B_m)$. By 3.9(b) there are $m_0$ and $S>0$ with $\theta(U_m) \leq 1/sN_m(f)$ for all $m \geq m_0$. 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 $\Psi$, $\alpha_{(n_k)}$ 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 $\text{Per}_n(X)$). Hence

$$f\mu_{f,(n_k)} = m\Psi, \alpha_{(n_k)} = m\Psi, \alpha_{(n_k)} = m\Psi, \alpha_{(n_k)} = \mu_{f,(n_k)}.$$

(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 $f(X_1)=X_{i+1}$ ($X_{m+1}=X_1$) and $f^m$: $X_1 \to X_1$ C-dense. From an invariant measure $\mu$ for $f^m$: $X_1 \to X_1$ we get one $\mu'$ for $f$: $X \to X$ by defining $\mu'(f^mE) = \mu(E)/m$ 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 $\mu'$ is ergodic if and only if $\mu$ 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 \( f: X \to X \). We shall study the \( C \)-dense case and this will give us results also for the general transitive case.


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

\[
\liminf_{n \to \infty} \mu(E \cap f^{-n}F) \geq R \mu(E) \mu(F).
\]

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 \( I(c_1^r, \ldots, c_r^r) \to \infty \) as \( n \to \infty \), then

\[
\liminf_{n \to \infty} \mu(f^{-c_1^r}E_1 \cap \cdots \cap f^{-c_r^r}E_r) \geq R_r \mu(E_1) \cdots \mu(E_r).
\]

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

(6.2) Theorem. If \( f: X \to X \) is \( C \)-dense, then \( f \) is partially mixing in all orders with respect to each \( \mu = \mu_{f^r},(n_k) \).

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

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

\[
U(B, \alpha) \subseteq \bigcap_{i=1}^r f^{-c_i^r}B_\varepsilon(E_i) \subseteq \bigcap_{i=1}^r f^{-c_i^r}V_i.
\]

By 3.7, we get

\[
N_d(\bigcap_{i=1}^r f^{-c_i^r}V_i) \geq N_d(U(B, \alpha)) \geq K(r, \alpha) \text{card}(B) \frac{N(d, \delta^*)}{N(n_k, \frac{1}{r} \delta^*)}
\]

for \( d \) sufficiently large. Now

\[
N(d, \delta^*) \geq N_d(f), \quad \text{card}(B) = \prod N_{n_k}(V_i)
\]

and, using 3.2(iii),

\[
N(n_k, \frac{1}{r} \delta^* \gamma) \leq N_{n_k}(f)^r / S_r.
\]

Combining all these,

\[
\frac{N_d(\bigcap_{i=1}^r f^{-c_i^r}V_i)}{N_d(f)} \geq R_r \prod \frac{N_{n_k}(V_i)}{N_{n_k}(f)}
\]
where \( R_r = K(r, a)S^r > 0 \). Letting \( d \to \infty \),

\[
\varphi(\bigcap f^{-c^t}_i V_i) = \liminf_{d \to \infty} \frac{N_d(\bigcap f^{-c^t}_i V_i)}{N_d(f)} \geq R_r \prod \frac{N_{n_k}(V_i)}{N_{n_k}(f)}.
\]

This being true for all big \( n \),

\[
\liminf_{n \to \infty} \varphi(\bigcap f^{-c^t}_i V_i) \geq R_r \prod \frac{N_{n_k}(V_i)}{N_{n_k}(f)}.
\]

Letting \( n_k \to \infty \),

\[
\liminf_{n \to \infty} \varphi(\bigcap f^{-c^t}_i V_i) \geq R_r \prod \alpha(n_k)(V_i) \geq R_r \prod \mu(E_i).
\]

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

\[
\bigcap f^{-c^t}_i V_i \supset \overline{\bigcap f^{-c^t}_i V_i}.
\]

Choose \( B \in \Psi \) so that

\[
\bigcap f^{-c^t}_i V_i \supset \overline{B} \supset \bigcap f^{-c^t}_i V_i.
\]

Then

\[
\mu(\bigcap f^{-c^t}_i V_i) \geq \alpha(n_k)(B) \geq \varphi(\bigcap f^{-c^t}_i V_i)
\]

and

\[
\liminf_{n \to \infty} \mu(\bigcap f^{-c^t}_i V_i) \geq R_r \prod \mu(E_i).
\]

Now

\[
\mu(\bigcap f^{-c^t}_i E_i) \geq \mu(\bigcap f^{-c^t}_i V_i) - \sum \mu(V_i \setminus E_i).
\]

Letting \( \mu(V_i \setminus E_i) \to 0 \) we get

\[
\liminf_{n \to \infty} \mu(\bigcap f^{-c^t}_i E_i) \geq R_r \prod \mu(E_i).
\]

For any \( E_i^* \in \mathcal{M} \) consider \( E_i \in E_i^* \) closed. Then

\[
\liminf_{n \to \infty} \mu(\bigcap f^{-c^t}_i E_i^*) \geq \liminf_{n \to \infty} \mu(\bigcap f^{-c^t}_i E_i) \geq R_r \prod \mu'(E_i).
\]

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 \( \mu^* \) on \( X \) corresponding to \( \mu_{f^n, (n_k)} \) 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 $\mu_f$.

**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
\[
\lim_{n \to \infty} \left( \frac{N_n(K)}{N_n(f)} \right) = 0.
\]
If $U$ is open with $\mu_f(\partial U)=0$, then
\[
\lim_{n \to \infty} \left( \frac{N_n(U)}{N_n(f)} \right) = \mu_f(U).
\]
**Proof.** Suppose $\{m_i\}$ is an increasing sequence of integers so that either
\[
N_{m_i}(K)/N_{m_i}(f) \to a > 0 \quad \text{or} \quad N_{m_i}(U)/N_{m_i}(f) \to b \neq \mu_f(U).
\]
Let $\psi$ be a countable base closed under finite union and $\{n_k\}$ a subsequence of $\{m_i\}$ so that $\mu_{f,(n_k)}$ is defined with $\psi$.

Suppose $N_{m_i}(K)/N_{m_i}(f) \to a > 0$. If $B \supseteq K$, $B \in \psi$, then
\[
\alpha_{(n_k)}(B) = \lim \frac{N_{n_k}(B)}{N_{n_k}(f)} \geq \lim \frac{N_{n_k}(K)}{N_{n_k}(f)} = a.
\]
It follows that $\mu_f(K)=\inf \alpha_{(n_k)}(B) \geq a > 0$, a contradiction. Suppose $N_{m_i}(U)/N_{m_i}(f) \to b \neq \mu_f(U)$. For $B \supset U$, $B \in \psi$ we have $\alpha_{(n_k)}(B) \geq b$; hence $\mu_f(U) = \mu_{f,(n_k)}(U) \geq b$.

For $B \subseteq U$, $B \in \psi$, we have $\alpha_{(n_k)}(B) \leq b$; hence $\mu_f(U) \leq b$. As $\mu_f(\partial U)=0$, $b \geq \mu_f(U) = \mu_f(U) = b$ and so $\mu_f(U) = b$, a contradiction.

(6.7) **Corollary.** Let $f: X \to X$ be $C$-dense. Then, for any $F \in C(X)$,
\[
\frac{1}{N_n(f)} \sum_{x \in \text{Per}_n(f)} F(x) \to \int F \, d\mu_f
\]
as $n \to \infty$. (We say that $\mu_f$ 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 $e>0$. Choose $-b = a_0 < a_1 < \cdots < a_r = b$ with $a_{i+1} - a_i < e$, $\mu_f(\{x : F(x)=a_i\})=0$ and $F(x)=a_i$ for no periodic point $x$.

Let $U_i = \{x : a_{i-1} < F(x) < a_i\}$. Choose $N(e)$ so big that
\[
|N_n(U_i)/N_n(f) - \mu_f(U_i)| < e/b
\]
for all $n \geq N(e)$ and each $i$. This is possible since $F(\partial U_i) \subseteq \{a_{i-1}, a_i\}$ and so $\mu_f(\partial U_i)=0$ by construction; hence 6.6 applies to $U_i$. We also have
\[
\left| \frac{N_n(f)^{-1}}{x \in \text{Per}_n(f)} \sum_{x \in \text{Per}_n(f)} F(x) - \sum_{i=1}^r a_i(N_n(U_i)/N_n(f)) \right| \leq e.
\]
Putting our above two inequalities together one sees that
\[
\left| \frac{N_n(f)^{-1}}{x \in \text{Per}_n(f)} \sum_{x \in \text{Per}_n(f)} F(x) - \sum_{i=1}^r a_i\mu_f(U_i) \right| \leq 2e.
\]
Since \( |\int F d\mu_f - \sum a_i \mu_f(U_i)| \leq \epsilon \), we finally get
\[
\left| \int F d\mu_f - N_n(f)^{-1} \sum_{x \in \text{Per}_n(f)} F(x) \right| \leq 3\epsilon
\]
for all \( n \geq N(\epsilon) \).

7. The algebraic case. Suppose \( f: G \to G \) is an automorphism of an \( n \)-dimensional torus \( G \). \( f \) is a hyperbolic if \( Df: T_0G \to T_xG \) 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 \( E \subseteq G \) is closed. Let \( \mu_f = \mu_{f,(n_0)} \) be defined via the base \( \Psi \). Consider \( B \in \Psi \) with \( B \supseteq E + g \). There are \( B^1 \in \Psi \) and open \( V \) such that \( B^1 \supseteq E \), \( g \in V \) and \( B^1 + V \subseteq B \). By 3.9(ii) there is an \( N \) such that \( N_n(V) > 0 \) for all \( n \geq N \). For \( n_k \geq N \) and \( g_{n_k} \in \text{Per}_{n_k}(V) \) we have \( g_{n_k} + \text{Per}_{n_k}(B^1) \subseteq B \). If \( x \in \text{Per}_{n_k}(B^1) \), 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 \( g_{n_k} + x \in \text{Per}_{n_k}(B) \). Thus \( N_{n_k}(B) \supseteq N_{n_k}(B^1) \) for \( n_k \geq N \) and \( \alpha_{(n_k)}(B) \geq \alpha_{(n_k)}(B^1) \geq \mu_{f,(n_k)}(E) \). Varying \( B, \mu_{f,(n_k)}(g + E) \geq \mu_{f,(n_k)}(E) \). Using \( -g \) instead of \( g \), \( g_{n_k} + x \in \text{Per}_{n_k}(B) \). Thus \( \mu_f(E) = \mu_f(g + E) \) for all \( g \in G \) and \( E \) closed; it follows that \( \mu_f \) is Haar measure.

Now let \( G \) be a torus acting freely on a compact metric space \( X \) (i.e. \( g_1 x = g_2 x \) implies \( g_1 = g_2 \)) and let \( \mu \) be normalized Haar measure on \( G \). Let \( \pi: X \to X_G = X/G \) 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
\[
\int_X F(gx_1) \, d\mu = \int_X F(gx_2) \, d\mu
\]
for \( x_1 = g_1 x_2 \) for some \( g_1 \in G \) and then \( F(gx_1) = F(g_1 gx_2) \) 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 \( S: X \to X \) is a homeomorphism and \( \sigma: G \to G \) 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 \( \sigma \)-extension of \( (S_G, m_G) \).

(7.2) Proposition. Let \( (S, m) \) be a \( \sigma \)-extension of \( (S_G, m_G) \) with \( \sigma \) 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 \( \epsilon > 0 \). Choose \( x_1, \ldots, x_\epsilon \in X \) such that for each \( x \in X \).
there is an \( x_i \) such that \( |F(gx) - F(gx_i)| \leq \varepsilon/3 \) for all \( g \in G \). Since \( \mu \) is derived from \( \sigma \) by periodic points (see 6.7), there is an \( N(\varepsilon) \) such that
\[
N_n(\sigma)^{-1} \sum_{g \in \text{Per}_n(\sigma)} F(gx) - \int_G F(gx) \, d\mu \leq \varepsilon/3
\]
for any \( n \geq N(\varepsilon) \). Combining the above inequalities we get
\[
N_n(\sigma)^{-1} \sum_{g \in \text{Per}_n(\sigma)} F(gx) - \int_G F(gx) \, d\mu \leq \varepsilon
\]
for any \( x \in X \) and any \( n \geq N(\varepsilon) \).

Recall that \( \int_X F \, dm = \int_{X_0} H_P \, dm_0 \) where \( H_P(\pi(x)) = \int_G F(gx) \, d\mu \). As \( m_0 \) is derived from \( S_0 \) by periodic points there is an \( M \geq N(\varepsilon) \) such that
\[
\left| \int_{X_0} H_P \, dm_0 - N_n(S_0)^{-1} \sum_{g \in \text{Per}_n(S_0)} H_P(y) \right| \leq \varepsilon
\]
for any \( n \geq M \). At this stage of the proof we need the following.

**Lemma.** If \( S^n(y) = y \), 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 \( (\sigma^n - 1)g = -g_1 \). Since \( \sigma^n \) is hyperbolic, there is such a \( g \). Let \( x = gz \). By this lemma for \( y \in \text{Per}_n(S_0) \) choose \( x_y \in \pi^{-1}(y) \cap \text{Per}_n(S) \).

Then
\[
H_P(y) - N_n(\sigma)^{-1} \sum_{g \in \text{Per}_n(\sigma)} F(gx_y) \leq \varepsilon.
\]

Now \( gx_y \in \text{Per}_n(S) \) 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 \( g \in \text{Per}_n(\sigma) \). Thus
\[
\text{Per}_n(S) = \{ gx_y : g \in \text{Per}_n(\sigma), y \in \text{Per}_n(S_0) \}
\]
(for clearly \( z \in \text{Per}_n(S) \) implies \( \pi(z) \in \text{Per}_n(S_0) \)). Thus
\[
N_n(S_0)^{-1} \sum_{g \in \text{Per}_n(S_0)} N_n(\sigma)^{-1} \sum_{y \in \text{Per}_n(S)} F(gx_y) = N_n(S)^{-1} \sum_{z \in \text{Per}_n(S)} F(z).
\]
Hence, as \( \int_X F \, dm = \int_{X_0} H_P \, dm_0 \), we have
\[
\left| \int F \, dm - N_n(S)^{-1} \sum_{z \in \text{Per}_n(S)} F(z) \right| \leq 2\varepsilon
\]
for all \( n \geq M \).

Suppose \( f : N/\Gamma \to N/\Gamma \) is a hyperbolic automorphism of a nilmanifold (one can see [13] or [16] for the definition). Then \( N/\Gamma \) 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.
Theorem. If $f$ is a hyperbolic automorphism of a nilmanifold, then $\mu_f = m$.

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

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 $\sigma$. 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 $\mu_f$. We refer the reader to [5] for a definition of measure theoretic entropy.

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 $U_i \supset A_i$ and $X = \bigcup_{i=1}^r A_i$. In [8] L. Goodwyn shows that for any $f$-invariant normalized Borel measure $\rho$ on $X$ (and $f : X \to X$ any continuous map) we have $h_\rho(f) \leq h(f)$. We complete our proof by showing the partition $\beta = \{A_1, \ldots, A_r\}$ satisfies $h_{\mu_f}(f, \beta) \geq h(f)$. For any $1 \leq i_0, \ldots, i_{m-1} \leq r$ 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 $\theta(V) \leq 1/SN_m(f)$ for all $m \geq m_0$. Then $\mu_f(D) \leq \theta(V) \leq 1/SN_m(f)$. Define the function

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

where $D(i_0, \ldots, i_{m-1})$ is the characteristic function of $D$. For $m \geq m_0$ we have

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

By definition

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

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

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

References


University of Warwick, Coventry, England