TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 184, October 1973

TOPOLOGICAL ENTROPY FOR NONCOMPACT SETS

BY

RUFUS BOWEN([1])

ABSTRACT. For $f: X \to X$ continuous and $Y \subset X$ a topological entropy $h(f, Y)$ is defined. For X compact one obtains results generalizing known theorems about entropy for compact Y and about Hausdorff dimension for certain $Y \subset X = S^1$. A notion of entropy-conjugacy is proposed for homeomorphisms.

The topological entropy of a continuous map on a compact space was defined by Adler, Konheim and McAndrew [1]. In the present paper we will define entropy for subsets of compact spaces in a way which resembles Hausdorff dimension. This will be used to generalize known results about the Hausdorff dimension of the quasiregular points of certain measures and to define a notion of conjugacy that is a cross between the topological and measure theoretic ones.

In [5] we gave a definition of entropy for uniformly continuous maps on metric spaces. That definition was motivated by different examples (linear maps on R^n and calculating entropy on T^n) and it sometimes differs from the definition given here.

We wish to thank Karl Sigmund who pointed us in the direction this paper takes and Ben Weiss who helped us formulate §4.

1. **The definition.** Let $f: X \to X$ be continuous and $Y \subset X$. The topological entropy $h(f, Y)$ will be defined much like Hausdorff dimension, with the "size" of a set reflecting how f acts on it rather than its diameter. Let $\mathfrak{A}$ be a finite open cover of X. We write $E \prec \mathfrak{A}$ if E is contained in some member of $\mathfrak{A}$ and $\{E_i\} \prec \mathfrak{A}$ if every $E_i \prec \mathfrak{A}$. Let $n_{f,\mathfrak{A}}(E)$ be the biggest nonnegative integer such that

$$f^k E \prec \mathfrak{A} \text{ for all } k \in [0, n_{f,\mathfrak{A}}(E));$$

$n_{f,\mathfrak{A}}(E) = 0$ if $E \not\prec \mathfrak{A}$ and $n_{f,\mathfrak{A}}(E) = +\infty$ if all $f^k E \prec \mathfrak{A}$. Now set

$$D_{\mathfrak{A}}(E) = \exp(-n_{f,\mathfrak{A}}(E)) \quad \text{and} \quad D_{\mathfrak{A}}(\mathcal{E}, \lambda) = \sum_{i=1}^{\infty} D_{\mathfrak{A}}(E_i)^{\lambda}$$

Received by the editors November 15, 1972.
AMS (MOS) subject classifications (1970). Primary 54H20, 28A65.
Key words and phrases. Entropy, Hausdorff dimension, invariant measure, generic points.
(1) Partially supported by NSF Grant GP-14519.

for $\mathcal{E} = \{E_i\}_{i=1}^{\infty}$ and $\lambda \in R$. We define a measure $m_{\mathfrak{A},\lambda}$ by

$$m_{\mathfrak{A},\lambda}(Y) = \lim_{\epsilon \to 0} \inf\left\{D_{\mathfrak{A}}(\mathcal{E}, \lambda): \bigcup E_i \supset Y \text{ and } D_{\mathfrak{A}}(E_i) < \epsilon\right\}.$$

Notice that $m_{\mathfrak{A},\lambda}(Y) \leq m_{\mathfrak{A},\lambda'}(Y)$ for $\lambda > \lambda'$ and $m_{\mathfrak{A},\lambda}(Y) \notin \{0, +\infty\}$ for at most one λ. Define

$$h_{\mathfrak{A}}(f, Y) = \inf\{\lambda: m_{\mathfrak{A},\lambda}(Y) = 0\} \quad \text{and finally} \quad h(f, Y) = \sup_{\mathfrak{A}} h_{\mathfrak{A}}(f, Y)$$

where $\mathfrak{A}$ ranges over all finite open covers of X. For $Y = X$ we write $h(f) = h(f, X)$.

Remark. The number $h(f, Y) = h_X(f, Y)$ depends very much on which space X we consider the domain of f. For instance, $f(x) = x + 1$ defines a homeomorphism of R which can be extended to a homeomorphism of S^1. By Proposition 1 below $h_{S^1}(f, S^1)$ is just the usual entropy of the homeomorphism $f: S^1 \to S^1$ and thus equals 0 [1, p. 315]; for $Y \subset S^1$ we have $0 \leq h_{S^1}(f, Y) \leq h_{S^1}(f, S^1)$ and so $h_{S^1}(f, Y) = 0$. On the other hand suppose $Y = \bigcup_{n=-\infty}^{+\infty}(n + A)$ where $A \subset (0, 1)$ is a Cantor set. Since Y is closed in R, one can prove $h_Y(f, Y) = h_R(f, Y)$. For any homeomorphism $g: A \to A$, $\pi: Y \to A$ defined by $\pi(n + a) = g^n(a)$ displays g as a quotient of $f \mid Y$. From this one can conclude that $h(g) \leq h(f \mid Y)$; as $h(g)$ can be made large, $h(f \mid Y) = +\infty$. Then $h_R(f, Y) = +\infty$ but $h_{S^1}(f, Y) = 0$. This example was suggested to us by L. Goodwyn.

Proposition 1. *If X is compact, then $h(f)$ equals the usual topological entropy.*

Proof. First let us recall the usual definition of entropy for compact X [1]. Let $\mathfrak{A}_{f,n} = \{A_{i_0} \cap f^{-1}A_{i_0} \cap \cdots \cap f^{-n+1}A_{i_{n-1}}: A_{i_k} \in \mathfrak{A}\}$ for an open cover $\mathfrak{A}$ of X. If $N(\mathfrak{B})$ denotes the smallest cardinality of any subcover of the open $\mathfrak{B}$, then

$$\underline{h}(f, \mathfrak{A}) = \lim_{n \to \infty} \frac{1}{n} \log N(\mathfrak{A}_{f,n})$$

exists and the topological entropy is defined by

$$\underline{h}(f) = \sup_{\mathfrak{A}} \underline{h}(f, \mathfrak{A})$$

where $\mathfrak{A}$ runs over all finite open covers of X. Letting $\mathcal{E}_n$ be a subcover with $N(\mathfrak{A}_{f,n})$ members

$$D_{\mathfrak{A}}(\mathcal{E}_n, \lambda) \leq N(\mathfrak{A}_{f,n})e^{-n\lambda}$$

and

$$m_{\mathfrak{A},\lambda}(X) \leq \lim_{n \to \infty} [\exp(-\lambda + n^{-1} \log N(\mathfrak{A}_{f,n}))]^n.$$

For $\lambda > \underline{h}(f, \mathcal{C})$ we get $m_{\mathcal{C},\lambda}(X) = 0$. Hence $h_{\mathcal{C}}(f, X) \le \underline{h}(f, \mathcal{C})$.

We prove $h_{\mathcal{C}}(f, X) \ge \underline{h}(f, \mathcal{C})$ by showing $\underline{h}(f, \mathcal{C}) \le \lambda$ whenever $m_{\mathcal{C},\lambda}(X) = 0$. For such a λ there is a countable covering $\mathcal{E} = \{E_i\}$ of X so that $D_{\mathcal{C}}(\mathcal{E}, \lambda) < 1$. If $n_{f,\mathcal{C}}(E_i) < \infty$, we may assume E_i is open (there is an open $F_i \supset E_i$ with $D_{\mathcal{C}}(F_i) = D_{\mathcal{C}}(E_i)$). The E_i's with $n_{f,\mathcal{C}}(E_i) = \infty$ may be replaced by open sets so that $D_{\mathcal{C}}(\mathcal{E}, \lambda)$ is still less than 1 (though it may increase). As X is compact, the open cover $\mathcal{E}$ now has a finite subcover $\mathcal{D} = \{D_1, \cdots, D_m\}$. Then

$$\sum_{s=1}^{\infty} \sum_{j_1,\cdots,j_s} \exp(-\lambda n_{f,\mathcal{C}}(D_{j_1}, \cdots, D_{j_s})) = \sum_{k=1}^{\infty} D_{\mathcal{C}}(\mathcal{D}, \lambda)^k < \infty$$

where $n_{f,\mathcal{C}}(D_{j_1}, \cdots, D_{j_s}) = \sum_{r=1}^{s} n_{f,\mathcal{C}}(D_{j_r})$.

Let

$$C(D_{j_1}, \cdots, D_{j_s}) = \{x \in X: f^{t_r}x \in D_{j_r} \text{ for each } r \in [1, s]$$
$$\text{where } t_r = n_{f,\mathcal{C}}(D_{j_1}) + \cdots + n_{f,\mathcal{C}}(D_{j_{r-1}})\}.$$

Then $C(D_{j_1}, \cdots, D_{j_s}) \prec \mathcal{C}_{f,n}$ for $n \le n_{f,\mathcal{C}}(D_{j_1}, \cdots, D_{j_s})$. If $M = \max_i n_{f,\mathcal{C}}(D_i)$, then $\{C(D_{j_1}, \cdots, D_{j_s}): s \ge 1,\ n_{f,\mathcal{C}}(D_{j_1}, \cdots, D_{j_s}) \in [n, n + M)\}$ is a cover of X subordinate to $\mathcal{C}_{f,n}$. Hence

$$N(\mathcal{C}_{f,n})e^{-\lambda n}$$
$$\le e^{M\lambda} \sum \{\exp(-\lambda n_{f,\mathcal{C}}(D_{j_1}, \cdots, D_{j_s})): n_{f,\mathcal{C}}(D_{j_1}, \cdots, D_{j_s}) \in [n, n + M)\}^{\lambda}.$$

As the right side is bounded in n, $h(f, \mathcal{C}) \le \lambda$.

This proof is almost identical with Furstenberg [10, Proposition III.1] and resembles the proof of a well-known theorem of information theory [19].

We now state (without proof) some basic facts.

Proposition 2. (a) *If $f_1: X_1 \to X_1$ and $f_2: X_2 \to X_2$ are topologically conjugate (i.e., there is a homeomorphism $\pi: X_1 \to X_2$ with $\pi f_1 = f_2 \pi$), then $h(f_1, Y_1) = h(f_2, \pi(Y_1))$ for $Y_1 \subset X_1$.*

(b) $h(f, f(Y)) = h(f, Y)$.

(c) $h(f, \bigcup_{i=1}^{\infty} Y_i) = \sup_i h(f, Y_i)$.

(d) $h(f^m, Y) = mh(f, Y)$ *for* $m > 0$.

We now give an example which motivated this paper. Define $f: S^1 \to S^1$ by $f(z) = z^n$. If $Y \subset S^1$ is closed and $f(Y) \subset Y$ then the Hausdorff dimension of Y satisfies $hd(Y) = h(f \mid Y)/\log n$. This was proved by Furstenberg [10, Proposition III.1]. For an ergodic f-invariant probability measure μ on S^1, it is known that (Colebrook [7]; see also [3] and [9]) $hd(G(\mu)) = h_{\mu}(f)/\log n$ where $G(\mu)$ denotes the set of generic points of μ. The above two formulas suggest that one might have $h_\mu(f) = h(f, G(\mu))$ if the right side is correctly defined for the noncompact set $G(\mu)$. The intermediate Hausdorff dimension of course motivated our definition of entropy; Theorem 3 shows that the hoped for formula holds for any continuous map on a compact metric space. We mention that another aspect of Colebrook's paper [7] has been generalized by K. Sigmund [20].

2. **Goodwyn's theorem.** In this section we will generalize a theorem of Goodwyn [13]. For a continuous map $f: X \to X$ let $M(f)$ be the set of all f-invariant Borel probability measures on X. We refer the reader to [4] or [14] for the definition of $h_\mu(f)$.

Theorem 1. *Let $f: X \to X$ be a continuous map of a compact metric space and $\mu \in M(f)$. If $Y \subset X$ and $\mu(Y) = 1$, then $h_\mu(f) \leq h(f, Y)$.*

Lemma 1. *Let α be a finite Borel partition of X such that every $x \in X$ is in the closures of at most M sets of α. Then*

$$h_\mu(f, \alpha) \leq h(f, Y) + \log M.$$

Proof. For each $x \in X$ let $I_n(x) = -\log \mu(A)$ where $A \in \alpha_{f,n}$ contains x. The Shannon-McMillian-Breiman theorem [14] says that for some μ-integrable function $I(x)$ one has $I_n(x)/n \to I(x)$ a.e. and $a = \int I(x)d\mu = h_\mu(f, \alpha)$. For $\delta > 0$ the set $Y_\delta = \{y \in Y: I(y) \geq a - \delta\}$ has positive measure. By Egorov's theorem there is an N so that

$$Y_{\delta,N} = \{y \in Y_\delta: I_n(y)/n \geq a - 2\delta \ \forall n \geq N\}$$

has positive measure.

Let $\mathfrak{B}$ be a finite open cover of X each member of which intersects at most M members of α. Suppose $\mathcal{E} = \{E_i\}$ covers Y and $D_{\mathfrak{B}}(E_i) \leq e^{-N}$. If $\beta \in \alpha_{f,n_{f,\mathfrak{A}}(E_i)}$ intersects $Y_{\delta,N}$, then $\mu(\beta) \leq \exp((-a + 2\delta)n_{f,\mathfrak{A}}(E_i))$. Since $E_i \cap Y_{\delta,N}$ is covered by at most $M^{n_{f,\mathfrak{A}}(E_i)}$ such β's,

$$\mu(E_i \cap Y_{\delta,N}) \leq \exp(n_{f,\mathfrak{A}}(E_i)(\log M - a + 2\delta)).$$

For $\lambda = -\log M + a - 2\delta$ we have

$$D_{\mathfrak{A}}(\mathcal{E}, \lambda) = \sum_i \exp(-\lambda n_{f,\mathfrak{A}}(E_i)) \geq \sum_i \mu(E_i \cap Y_{\delta,N}) \geq \mu(Y_{\delta,N}).$$

Letting $\mathcal{E}$ vary, $m_{\mathfrak{A},\lambda}(Y) \geq \mu(Y_{\delta,N}) > 0$. Hence $h(f, Y) \geq h_{\mathfrak{A}}(f, Y) \geq \lambda = -\log M + a - 2\delta$. Letting $\delta \to 0$ we have our result.

Lemma 2. *Let $\mathfrak{A}$ be a finite open cover of X. For each $n > 0$ there is a finite Borel partition α_n of X such that $f^k\alpha_n \prec \mathfrak{A}$ for all $k \in [0, n)$ and at most n* card $\mathfrak{A}$ *sets in α_n can have a point in all their closures.*

Proof. This idea for this lemma is from Goodwyn [13] and the statement as above is in [15]. Let $\mathfrak{A} = \{A_1, \cdots, A_m\}$ and $g_1, \cdots, g_m$ be a partition of unity subordinate to $\mathfrak{A}$. Then $G = (g_1, \cdots, g_m)\colon X \to s_{m-1} \subset R^m$ where s_{m-1} is an $m-1$ dimensional simplex. Now $\{U_1, \cdots, U_m\}$ is an open cover of s_{m-1} where $U_i = \{\underline{x} \in s_{m-1}\colon x_i > 0\}$ and $G^{-1}U_i \subset A_i$. As $(s_{m-1})^n$ is $nm-n$ dimensional, there is a finite Borel partition α_n^* of s_{m-1}^n with at most nm members having a point in all their closures and such that each member of α_n^* lies in some $U_{i_1} \times \cdots \times U_{i_n}$. Then $\alpha_n = L^{-1}\alpha_n^*$ works where $L = (G, G\circ f, \cdots, G\circ f^{n-1})\colon X \to s_{m-1}^n$.

Lemma 3. *Given a finite Borel partition β and $\epsilon > 0$ there is an open cover $\mathfrak{A}$ so that $H_\mu(\beta|\alpha) < \epsilon$ whenever α is a finite Borel partition with $\alpha \prec \mathfrak{A}$.*

Proof. Let $\beta = \{B_1, \cdots, B_m\}$. There is a $\delta > 0$ so that the following is true:

$H_\mu(\beta\,|\,\alpha) < \epsilon$ if there is a Borel partition $\{C_1, \cdots, C_m\}$ with each C_i a union of members of α and $\sum_{i\neq j} P(B_i \cap C_j) < \delta$

(see [4, Theorem 6.2]). Choose compact sets $K_i \subset B_i$ so that $\mu(B_i \backslash K_i) < \delta/m$. Let $\mathfrak{A}$ be an open cover each member of which intersects at most one K_i. For $\alpha \prec \mathfrak{A}$ put $A \in \alpha$ in C_i if $A \cap K_i \neq \varnothing$, and in any C_j if $A \cap \bigcup_j K_j = \varnothing$. Then $C_j \cap K_i = \varnothing$ for $i \neq j$ and so

$$\sum_{i\neq j} P(B_i \cap C_j) \leq \sum_i P(B_i \backslash K_i) < \delta.$$

Proof of Theorem 1. Let β be a finite Borel partition of X and $\epsilon > 0$. Let $\mathfrak{A}$ be as in Lemma 3 and α_n as in Lemma 2. Then

$$\begin{aligned} h_\mu(f, \beta) &= n^{-1}h_\mu(f^n, \beta_{f,n}) \leq n^{-1}h_\mu(f^n, \alpha_n) + n^{-1}H_\mu(\beta_{f,n}|\alpha_n) \\ &\leq n^{-1}[h(f^n, Y) + \log(n \operatorname{card} \mathfrak{A})] + n^{-1}\sum_{k=0}^{n-1} H_\mu(f^{-k}\beta\,|\,\alpha_n) \\ &\leq h(f, Y) + n^{-1}\log(n \operatorname{card} \mathfrak{A}) + n^{-1}\sum_{k=0}^{n-1} H_\mu(\beta\,|\,f^k\alpha_n) \\ &\leq h(f, Y) + n^{-1}\log(n \operatorname{card} \mathfrak{A}) + \epsilon. \end{aligned}$$

Here we used Lemmas 1 and 2 and some general facts:

$$h_\mu(f, \eta) \leq h_\mu(f, \xi) + H_\mu(\eta \mid \xi),$$

$$H_\mu(\eta \vee \gamma \mid \xi) \leq H_\mu(\eta \mid \xi) + H(\gamma \mid \xi),$$

$$H_\mu(f^{-1}\eta \mid f^{-1}\xi) = H_\mu(\eta \mid \xi).$$

Proofs of these are in [4] and [14]. Finally, let $n \to \infty$ and then let $\epsilon \to 0$. The proof is finished.

3. **Generic points.** For X a compact metric space, the set $M(X)$ of all Borel probability measures on X with the weak topology is a compact metrizable space [18]. $\mu_n \to \mu$ implies that for $V \supset K$ with V open and K compact one has $\liminf \mu_n(V) \geq \mu(K)$. For $x \in X$ let μ_x denote the unit measure concentrated on x. If a continuous $f: X \to X$ is given, define

$$\mu_{x,n} = n^{-1}(\mu_x + \mu_{fx} + \cdots + \mu_{f^{n-1}x}).$$

Let $V_f(x)$ be the set of all limit points in $M(X)$ of the sequence $\mu_{x,n}$. Then $V_f(x) \neq \emptyset$ and one checks that $V_f(x) \subset M(f)$. x is a *generic point* for μ if $V_f(x) = \{\mu\}$. Our main result is that $h(f, G(\mu)) = h_\mu(f)$ for μ ergodic where $G(\mu)$ is the set of generic points for μ.

$p = (p_1, \cdots, p_N)$ is an *N-distribution* if $\Sigma_1^N p_i = 1$ and $p_i \geq 0$; we set $H(p) = -\Sigma_i p_i \log p_i$. If $a = (a_1, \cdots, a_m) \in \{1, \cdots, N\}^m$, then $\operatorname{dist} a = (p_1, \cdots, p_N)$ where $p_i = m^{-1}$ (number of j with $a_j = i$). If p and q are N-distributions, then $|p - q| = \max_i |p_i - q_i|$.

Lemma 4. *Let*

$$R(N, m, t) = \{a \in \{1, \cdots, N\}^m: H(\operatorname{dist} a) \leq t\}.$$

Then, fixing N and t,

$$\limsup_{m \to \infty} \frac{1}{m} \log \operatorname{card} R(N, m, t) \leq t.$$

Proof. For an N-distribution q and $\alpha \in (0, 1)$ consider $R_m(q) = \{a \in \{1, \cdots, N\}^m: |q - \operatorname{dist} a| < \alpha\}$. Let μ be the measure on $\Sigma_N = \{1, \cdots, N\}^Z$ for the Bernoulli shift with distribution $q' = (1 - \alpha)q + \alpha(1/N, \cdots, 1/N)$. Each $a \in R_m(q)$ corresponds to a cylinder set $C_a \subset \Sigma_N$. Since $|q - \operatorname{dist} a| < \alpha$, the number of i's occurring in a is at most $(q_i + \alpha)m$. As the symbol i has probability $q_i' = (1 - \alpha)q_i + \alpha/N$,

$$\mu(C_a) \geq \prod_{i=1}^{N} q_i'^{(q_i+\alpha)m}.$$

Since the C_a's are disjoint and have total μ-measure at most 1,

$$1 \geq \text{card } R_m(q) \prod_i q_i'^{(q_i+\alpha)m}.$$

Taking logarithms we get

$$\frac{1}{m} \log \text{card } R_m(q) \leq \sum_i -(q_i + \alpha) \log q_i'$$

$$\leq H(q') + \sum_i (|q_i' - q_i| + \alpha)|\log q_i'|.$$

As $q_i' \geq \alpha/N$, $|\log q_i'| \leq \log N - \log \alpha$; also $|q_i' - q_i| = |\alpha/N - \alpha q_i| \leq 2\alpha$. So

$$m^{-1} \log \text{card } R_m(q) \leq H(q') + 3\alpha N(\log N - \log \alpha).$$

Now $H(q)$ is uniformly continuous in q and $\log \alpha \to 0$ as $\alpha \to 0$. Hence given any $\epsilon > 0$, for small α one has

$$m^{-1} \log \text{card } R_m(q) \leq H(q) + \epsilon$$

for all m and q.

Once an α is chosen one can find a finite set Q of N-distributions so that

(a) $H(q) \leq t$ for $q \in Q$ and

(b) if $H(q^*) \leq t$, then $|q^* - q| < \alpha$ for some $q \in Q$.

Then $R(N, m, t) \subset \bigcup_{q \in Q} R_m(q)$.

$$m^{-1} \log \text{card } R(N, m, t) \leq m^{-1} \log \text{card } Q + (t + \epsilon).$$

Letting $m \to \infty$ and then $\epsilon \to 0$ we get our result.

Now suppose $\beta = \{B_1, \cdots, B_N\}$ is a cover of X. An *n-choice for* x (with respect to β and f) is a $\underline{B} = (B_{i_0}, \cdots, B_{i_{n-1}}) \in \beta^n$ with $f^k(x) \in B_{i_k}$ for $k \in [0, n)$. An n-choice gives an N-distribution $q(\underline{B}) = \text{dist}(i_0, \cdots, i_{n-1})$. The set of such distributions for the various n-choices for x we denote by $\text{Dist}_\beta(x, n)$.

Lemma 5. *Suppose $f: X \to X$ is a continuous map of a topological space, $\mathfrak{B}$ an open cover of X, β a finite cover of X and M a positive integer so that $f^k\beta \prec \mathfrak{B}$ for all $k \in [0, M)$. For $t \geq 0$ define*

$$Q(t, \beta) = \left\{x \in X: \liminf_{n \to \infty} (\inf\{H(q) : q \in \text{Dist}_\beta(x, n)\}) \leq t\right\}.$$

Then $h_{\mathfrak{B}}(f, Q(t, \beta)) \leq t/M$.

Proof. Let $N = \text{card } \beta$ and $\epsilon > 0$. By Lemma 4 there is an m_ϵ so that

$$\text{card } R(N, m, t + \epsilon) \leq e^{m(t+2\epsilon)}$$

for all $m \geq m_\epsilon$. As $\text{Dist}_\beta(x, n)$ depends only slightly on the last few $f^j(x)$ when n is large and $H(q)$ is continuous in q, one has

$$\liminf_{m\to\infty}(\inf\{H(q): q \in \mathrm{Dist}_\beta(x, mM)\}) \le t$$

for $x \in Q(t, \beta)$. Let $\underline{B}_n(x) = (B_{i_0}, \cdots, B_{i_{n-1}})$ be an n-choice with distribution $q(x, n)$ minimizing $H(q)$ over $\mathrm{Dist}_\beta(x, n)$. For $k \in [0, M)$ define

$$q_k(x, m) = \mathrm{dist}\{i_{k+rM}: r \in [0, m)\}.$$

Then $q(x, mM) = (1/M)\Sigma_k q_k(x, m)$. By the concavity of $H(q)$ in q one has $H(q_k(x, m)) \le H(q(x, mM))$ for some k (depending on x and m).

Fix now any $m_0 \ge m_\epsilon$. For $m \ge m_0$ and $k \in [0, M)$ define

$$S(m, k) = \{x \in X: H(q_k(x, m)) \le t + \epsilon\}.$$

Then $Q(t, \beta) \subset \bigcup\{S(m, k): m \ge m_0,\ k \in [0, M)\}$. Assume $x \in S(m, k)$; $a(x) = (B_{i_k}, B_{i_{k+M}}, \cdots, B_{i_{k+(m-1)M}})$ is in $R(N, m, t + \epsilon)$. Define

$$A_k(x, m) = \{y \in X: f^j y \in B_{i_j} \quad \text{for} \quad j \in [0, k) \text{ and}$$

$$f^{k+rM} y \in B_{i_{k+rm}} \quad \text{for} \quad r \in [0, m)\}.$$

Now $f^j A_k(x, m)$ is contained in some member of β for each $j \in [0, mM)$. Hence $D_{\mathfrak{B}} A_k(x, m) \le e^{-mM}$. Let $\mathcal{E}(m_0) = \{A_k(x, m): x \in S(m, k),\ m \ge m_0,\ k \in [0, M)\}$.

Then $\mathcal{E}(m_0)$ covers $Q(t, \beta)$. Since there are at most $(\mathrm{card}\ \beta)^k \cdot \mathrm{card}\ R(N, m, t + \epsilon)$ different $A_k(x, m)$ with $x \in S(m, k)$,

$$D_{\mathfrak{B}}(\mathcal{E}(m_0), (t + 3\epsilon)/M) \le \sum_{\substack{k \in [0, M) \\ m \ge m_0}} (\mathrm{card}\ \beta)^k\ \mathrm{card}\ R(N, m, t + \epsilon) e^{-m(t+3\epsilon)}$$

$$\le (\mathrm{card}\ \beta)^{M-1} \sum_{m \ge m_0} e^{-m\epsilon}.$$

As this quantity approaches 0 as $m_0 \to \infty$, $m_{\mathfrak{B},(t+3\epsilon)/M}(Q(t, \beta)) = 0$ and $h_{\mathfrak{B}}(f, Q(t, \beta)) \le (t + 3\epsilon)/M$. Now let $\epsilon \to 0$.

Theorem 2. *Let $f: X \to X$ be a continuous map on a compact metric space. Set*

$$QR(t) = \{x \in X: \exists \mu \in V_f(x) \quad \textit{with} \quad h_\mu(f) \le t\}.$$

Then $h(f, QR(t)) \le t$.

Proof. Let $\mathfrak{B}$ be a finite open cover of X and α a Borel partition of X with the closures of members of α contained in members of $\mathfrak{B}$. Fix $\epsilon > 0$ and let

$$W_\epsilon(M) = \{x \in X: \exists \mu \in V_f(x) \quad \text{with} \quad (1/M)H_\mu(\alpha_{f,M}) < t + \epsilon\}.$$

If $h_\mu(f) \le t$, then

$$\lim_{M\to\infty} \frac{1}{M} H_\mu(\alpha_{f,M}) = h_\mu(f, \alpha) \le h_\mu(f)$$

implies that $(1/M)H_\mu(\alpha_{f,M}) < t + \epsilon$ for some M. Hence $QR(t) \subset \bigcup_M W_\epsilon(M)$.

Now fix an M and let $\alpha_{f,M} = \{E_1, \cdots, E_N\}$. Pick $U_i \supset E_i$ open so that $f^k U_i \prec \mathfrak{B}$ for $k \in [0, M)$; set $\beta = \{U_1, \cdots, U_N\}$. We will show $W_\epsilon(M) \subset Q(M(t + 2\epsilon), \beta)$. Consider $x \in W_\epsilon(M)$ and $\mu \in V_f(x)$ with $(1/M)H(\alpha_{f,M}) < t + \epsilon$. Let $q' = (\mu(E_1), \cdots, \mu(E_n))$ and pick $\delta > 0$ so that

$$|q - q'| \le \delta \quad \text{implies} \quad H(q) \le M(t + 2\epsilon).$$

Now choose compact $K_i \subset E_i$ so that $\mu(E_i \setminus K_i) < \delta/2N$ and disjoint open V_i's with $U_i \supset V_i \supset K_i$. Let $\underline{B}_n(x) \in \beta^n$ be an n-choice for x so that $B_{i_k} = U_j$ whenever $f^k x \in V_j$. Since $\mu \in V_f(x)$, $\mu_{x,n_j} \to \mu$ for some $n_j \to \infty$. For large j one has

$$\mu_{x,n_j}(V_i) \ge \mu(K_i) - \delta/2N$$

for all i. If $q^j = \operatorname{dist} \underline{B}_{n_j}(x) = (q^j, \cdots, q^j_N)$, it follows that $q^j_i \ge \mu(K_i) - \delta/2N \ge \mu(E_i) - \delta/N$. We get $|q^j - q'| \le \delta$ and $H(q^j) \le M(t + 2\epsilon)$. Hence $x \in Q(M(t + 2\epsilon), \beta)$.

Lemma 5 now gives us $h_{\mathfrak{B}}(f, W_\epsilon(M)) \le t + 2\epsilon$. By Proposition 2(d) we get $h_{\mathfrak{B}}(f, QR(t)) \le t + 2\epsilon$. Letting $\epsilon \to 0$ and varying $\mathfrak{B}$ we are done.

Corollary. *Let $f: X \to X$ be a continuous map of a compact metric space. Then*

$$h(f) = \sup_{\mu \in M(f)} h_\mu(f).$$

Proof. Let $t = \sup_\mu h_\mu(f)$. As $V_f(x) \ne \emptyset$ for $x \in X$, $X \subset QR(t)$ and $h(f) = h(f, X) \le t$. On the other hand $h(f) \ge t$ by Goodwyn's theorem (Theorem 1).

Remark. This result is already known; see [8] for the finite dimensional metric case and [12] for compact Hausdorff spaces.

Theorem 3. *Let f be a continuous map on a compact metric space and $\mu \in M(f)$ be ergodic. Let $G(\mu)$ be the set of generic points of μ, i.e.,*

$$G(\mu) = \{x: V_f(x) = \{\mu\}\}.$$

Then $h(f, G(\mu)) = h_\mu(f)$.

Proof. By the ergodic theorem, one has $\mu(G(\mu)) = 1$. Theorem 1 then gives $H(f, G(\mu)) \ge h_\mu(f)$. As $G(\mu) \subset QR(h_\mu(f))$, Theorem 2 gives the reverse inequality.

4. A type of conjugacy. We will call two homeomorphisms $f: X \to X$ and $g: Y \to Y$ *entropy conjugate* if there are $X' \subset X$ and $Y' \subset Y$ such that

(i) X' and Y' are Borel sets,

(ii) $f(X') \subset X'$, $g(Y') \subset Y'$,

(iii) $h(f, X \backslash X') < h(f)$, $h(g, Y \backslash Y') < h(g)$, and

(iv) $f|X'$ and $g|Y'$ are topologically conjugate.

Unfortunately this does not seem to be an equivalence relation.

Proposition 3. *If f and g are entropy-conjugate homeomorphisms of compact metric spaces, then $h(f) = h(g)$.*

Proof. Suppose $\mu \in M(f)$ and $h_\mu(f) > h(f, X\backslash X')$. Since μ is f-invariant and $f(X') \subset X'$, one can find $B \subset X \backslash X'$ with $\mu(B) = \mu(X\backslash X')$ and $f(B) = B$. By Theorem 1, $\mu(B) < 1$. Define $\mu_{X'}(E) = \mu(E \cap X')/\mu(X')$. Then $\mu_{X'} = \mu$ (if $\mu(X') = 1$) or $\mu = \mu(X')\mu_{X'} + \mu(B)\mu_B$. In the second case $\mu_{X'}, \mu_B \in M(f)$ and

$$h_\mu(f) = \mu(X')h_{\mu_{X'}}(f) + \mu(B)h_{\mu_B}(f).$$

By Theorem 1 we have $h_{\mu_B}(f) \leq h(f, X\backslash X') < h_\mu(f)$ and so $h_{\mu_{X'}}(f) \geq h_\mu(f)$. If $\mu_{X'} = \mu$, we of course also have $h_{\mu_{X'}}(f) \geq h_\mu(f)$. Since $\mu_{X'}(X') = 1$, the topological conjugacy of $f \mid X'$ and $g \mid Y'$ gives us a measure ν on Y' with (g, ν) conjugate to $(f, \mu_{X'})$; in particular

$$h_\nu(g) = h_{\mu_{X'}}(f) \geq h_\mu(f).$$

By Goodwyn's theorem $h(g) \geq h_\mu(f)$. Using the Dinaburg-Goodman theorem (corollary to Theorem 2) one can make $h_\mu(f)$ arbitrarily close to $h(f)$ (and so satisfy $h_\mu(f) > h(f, X\backslash X')$). One gets $h(g) \geq h(f)$. By symmetry one likewise has $h(g) \leq h(f)$.

There is a natural class of homeomorphisms for which the converse of Proposition 3 may hold. Let $\Sigma_n = \Pi_Z \{1, \cdots, n\}$ and define the shift $\sigma_n: \Sigma_n \to \Sigma_n$ by

$$(\sigma_n x)_i = x_{i+1} \quad \text{for} \quad x = (x_i).$$

σ_n is a homeomorphism of the compact metrizable space Σ_n. For A an $n \times n$ matrix of 0's and 1's define

$$\Sigma_n(A) = \{(X_i) \in \Sigma_n: A_{x_i x_{i+1}} = 1 \ \forall i\}.$$

Then $\sigma_n \mid \Sigma_n(A)$ is a homeomorphism of a compact space.

Conjecture. Suppose $\sigma_n \mid \Sigma_n(A)$ and $\sigma_m \mid \Sigma_m(B)$ are topologically mixing and have the same topological entropy. Then they are entropy conjugate.

This conjecture is related to the symbolic dynamics of diffeomorphisms [2],

[16], [6]. From [6] it follows that the nonwandering set of an Axiom A diffeomorphism is entropy conjugate to some $\Sigma_n(A)$ (called a subshift of finite type). The codings of [2] show that the conjecture is true for the subshifts of finite type that arise from hyperbolic automorphisms of T^2. The codings were used in [2] to prove that entropy classifies such maps on T^2 up to measure theoretic conjugacy; Friedman and Ornstein [11] now supplant these codes for this purpose. The notion of entropy conjugacy attempts to clarify the topological content of the Adler-Weiss codings (see problem 3 of [21]).

Proposition 4. *Suppose f and g are entropy-conjugate homeomorphisms of compact metric spaces. Then f is intrinsically ergodic iff g is.*

Proof. Intrinsic ergodicity [17] means there is a unique $\mu \in M(f)$ with $h_\mu(f) = h(f)$. The proof is like that of Proposition 3.

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. R. L. Adler and B. Weiss, *Similarity of automorphisms of the torus*, Mem. Amer. Math. Soc. No. 98 (1970). MR 41 #1966.

3. P. Billingsley, *Hausdorff dimension in probability theory*, Illinois J. Math. 4 (1960), 187–209. MR 24 #A1750.

4. ———, *Ergodic theory and information*, Wiley, New York, 1965. MR 33 #254.

5. R. Bowen, *Entropy for group endo-morphisms and homogeneous spaces*, Trans. Amer. Math. Soc. 153 (1971), 401–414. MR 43 #469.

6. ———, *Markov partitions for Axiom A diffeomorphisms*, Amer. J. Math. 92 (1970), 725–747. MR 43 #2740.

7. C. M. Colebrook, *The Hausdorff dimension of certain sets of non-normal numbers*, Michigan Math. J. 17 (1970), 103–116. MR 41 #5321.

8. E. I. Dinaburg, *The relation between topological entropy and metric entropy*, Dokl. Akad. Nauk SSSR 190 (1970), 19–22 = Soviet Math. Dokl. 11 (1970), 13–16. MR 41 #425.

9. H. G. Eggleston, *The fractional dimension of a set defined by decimal properties*, Quart. J. Math. Oxford Ser 20 (1949), 31–36. MR 11, 88.

10. H. Furstenberg, *Disjointness in ergodic theory, minimal sets and a problem in Diophantine approximation*, Math. Systems Theory 1 (1967), 1–49. MR 35 #4369.

11. N. A. Friedman and D. S. Ornstein, *On isomorphism of weak Bernoulli transformations*, Advances in Math. 5 (1970), 365–394. MR 43 #478c.

12. T. N. T. Goodman, *Relating topological entropy and measure entropy*, Bull. London Math. Soc. 3 (1971), 176–180. MR 44 #6934.

13. L. W. Goodwyn, *Topological entropy bounds measure theoretic entropy*, Proc. Amer. Math. Soc. 23 (1969), 679–688. MR 40 #299.

14. W. Parry, *Entropy and generators in ergodic theory*, Benjamin, New York, 1969. MR 41 #7071.

15. D. Ruelle, *Statistical mechanics on a compact set with Z^ν action satisfying expansiveness and specification* (preprint).

16. Ja. G. Sinaĭ, *Markov partitions and Y-diffeomorphisms*, Funkcional. Anal. i Priložen. 2 (1968), no. 1, 64–89 = Functional Anal. Appl. 2 (1968), 61–82. MR 38 #1361.

17. B. Weiss, *Intrinsically ergodic systems*, Bull. Amer. Math. Soc. 76 (1970), 1266–1269. MR 42 #1978.

18. K. Parathasarathy, *Probability measures on metric spaces*, Academic Press, New York, 1967.

19. R. Ash, *Information theory*, Interscience Tracts in Pure and Appl. Math., no. 19, Interscience, New York, 1965. pp. 35–36. MR 37 #5049.

20. K. Sigmund, *On dynamical systems with the specification property* (to appear).

21. B. Weiss, *The isomorphism problem in ergodic theory*, Bull. Amer. Math. Soc. 78 (1972), 668–684.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720