STATISTICAL MECHANICS ON A COMPACT SET WITH $Z'$ ACTION SATISFYING EXPANSIVENESS AND SPECIFICATION

BY

DAVID RUELLE

ABSTRACT. We consider a compact set $\Omega$ with a homeomorphism (or more generally a $Z'$ action) such that expansiveness and Bowen's specification condition hold. The entropy is a function on invariant probability measures. The pressure (a concept borrowed from statistical mechanics) is defined as function on $C(\Omega)$—the real continuous functions on $\Omega$. The entropy and pressure are shown to be dual in a certain sense, and this duality is investigated.

0. Introduction. Invariant measures for an Anosov diffeomorphism have been studied by Sinai [16], [17]. More generally, Bowen [2], [3] has considered invariant measures on basic sets for an Axiom A diffeomorphism. The problems encountered are strongly reminiscent of those of statistical mechanics (for a classical lattice system—see [14, Chapter 7]). In fact Sinai [18] has explicitly used techniques of statistical mechanics to show that an Anosov diffeomorphism does not in general have a smooth invariant measure.

In this paper, we rewrite a part of the general theory of statistical mechanics for the case of a compact set $\Omega$ satisfying expansiveness and the specification property of Bowen [2]. Instead of a $Z$ action we consider a $Z'$ action as is usual in lattice statistical mechanics, where $Q = FZ'$ ($F$: a finite set). This rewriting gives a more general and intrinsic formulation of (part of) statistical mechanics; it presents a number of technical problems, but the basic ideas are contained in the papers of Gallavotti, Lanford, Miracle, Robinson, and Ruelle [7], [11], [12], [13], etc. The ideas of Bowen [2] and Goodwyn [8] on the relation between topological and measure-theoretical entropy are also used.

We describe now some of our results in the case of a homeomorphism $T$ of a metrizable compact set $\Omega$ satisfying expansiveness and specification (see §1).

Let $\Pi_a = \{x \in \Omega: T^a x = \{x\}\}$, and let $\mathcal{C}(\Omega)$ be the Banach space of real continuous functions on $\Omega$. The pressure $P$ is a continuous convex function on $\mathcal{C}(\Omega)$ defined by

$$P(\varphi) = \lim_{n \to \infty} \frac{1}{n} \log Z(\varphi, a),$$

$$Z(\varphi, a) = \sum_{x \in \Pi_a} \exp \sum_{m=1}^a \varphi(T^m x)$$

(§2). Let $I$ be the set of probability measures on $\Omega$, invariant under $T$ with the vague topology. The (measure theoretic) entropy $s$ is an affine upper semicontinuous function on $I$ defined in the usual way (§4). The following variational principle holds (§5)

Received by the editors June 20, 1972.

Those $\mu$ for which the maximum is reached in (0.1) form a nonempty set $I_\varphi$. $I_\varphi$ is a Choquet simplex and consists of precisely those $\mu \in I$ such that

$$P(\varphi + \psi) - P(\varphi) \geq \mu(\psi), \quad \forall \psi \in C(\Omega).$$

Let $\mu_{\varphi,a}$ be the measure on $\Omega$ which is carried by $\Pi_a$ and gives $x \in \Pi_a$ the mass

$$\mu_{\varphi,a}(\{x\}) = Z(\varphi,a)^{-1}\exp \sum_{m=1}^{a} \varphi(T^m x).$$

Then, any limit point of $\mu_{\varphi,a}$ as $a \to \infty$ is in $I_\varphi$ (§3). There is a residual subset $D$ of $C(\Omega)$ such that $I_\varphi$ consists of one single point $\mu_\varphi$ if $\varphi \in D$. In that case

$$\lim_{a \to \infty} \mu_{\varphi,a} = \mu_\varphi.$$

Miscellaneous properties of invariant states are reviewed in §6.

I am indebted to J. Robbin for acquainting me with Bowen’s papers, starting the present work.

1. Notation and assumptions. We denote by $|S|$ the cardinal of the set $S$. If $m = (m_1, \ldots, m_v) \in \mathbb{Z}^v$, $v \geq 1$, we let $||m|| = \sup_i |m_i|$. Given integers $a_1, \ldots, a_v > 0$, we define $\Lambda(a) = \{m \in \mathbb{Z}^v : 0 \leq m_i < a_i\}$. If $(\Lambda_\alpha)$ is a directed family of finite subsets of $\mathbb{Z}^v$, $\Lambda_\alpha \uparrow \infty$ means $|\Lambda_\alpha| \to \infty$ and $|\Lambda_\alpha + F|/|\Lambda_\alpha| \to 1$ for every finite $F \subset \mathbb{Z}^v$. In particular $\Lambda(a) \uparrow \infty$ when $a \to \infty$ (i.e. when $a_1, \ldots, a_v \to \infty$).

Let $\mathbb{Z}^v$ act by homeomorphisms on the compact set $\Omega$. We suppose that $\Omega$ is metrizable with metric $d$. $C(\Omega)$ is the space of real continuous functions on $\Omega$ with the sup norm. On the space $C(\Omega)^*$ of real measures on $\Omega$, we put the vague topology. We denote by $\delta_x$ the unit mass at $x$.

The following assumptions are made.(1)

1.1. Expansiveness. There exists $\delta^* > 0$ such that

$$(d(mx,my) \leq \delta^* \text{ for all } m \in \mathbb{Z}^v) \Rightarrow (x = y).$$

1.2. Weak specification. Given $\delta > 0$ there exists $p(\delta) > 0$ such that for any families $(\Lambda_i)_{i \in I}$, $(x_i)_{i \in I}$ satisfying

if $i \neq j$, the distance of $\Lambda_i, \Lambda_j$

(i)

(as subsets of $\mathbb{Z}^v$, with the distance $||\cdot||$) is $> p(\delta)$,

there is $x \in X$ such that

$$d(m_i x, m_i x_i) < \delta, \quad \forall i \in I, \forall m_i \in \Lambda_i.$$

1.3. Strong specification. Let $\mathbb{Z}^v(a)$ be the subgroup of $\mathbb{Z}^v$ with generators $(a_1,0,\ldots,0), \ldots, (0,\ldots,a_v)$, and let $\Pi_a = \{x \in \Omega : \mathbb{Z}^v(a)x = \{x\}\}$. For any

(1) Cf. Bowen [2].
families \((\Lambda_i)_{i \in \mathcal{I}}, (x_i)_{i \in \mathcal{I}}\) satisfying

\[ \Lambda_i \subset \Lambda(a) \text{ for all } i \text{ and, if } i \neq j, \]
the distance of \(\Lambda_i + Z'(a)\) and \(\Lambda_j\) is \(> p(\delta)\),

there is \(x \in \Pi_\mathcal{I}\) such that

\[ d(m_i x, m_j x) < \delta, \text{ all } i \in \mathcal{I}, \text{ all } m_i, m_j \in \Lambda_i. \]

It is easily seen that strong specification implies weak specification. If \(\Omega\) is a basic set for an Axiom A diffeomorphism \((\nu = 1)\), it is known that expansiveness [19] holds, and that (strong) specification [2] holds for some iterate of the diffeomorphism.

We note that expansiveness has the following easy consequence.

1.4. Proposition [9]. If \(0 < \delta\) there exists \(q(\delta)\) such that \((d(m x, m y) < \delta^* \text{ if } |m| < q(\delta)) \Rightarrow (d(x, y) < \delta)\).

2. Partition functions and pressure.

2.1. Definitions. Let \(\delta > 0\); \(E \subset \Omega\) is \((\delta, \Lambda)\)-separated if \((x, y \in E, \text{ and } d(m x, m y) < \delta \text{ for all } m \in \Lambda) \Rightarrow (x = y)\). Let \(\varphi \in \mathcal{C}(\Omega)\). Given \(\delta > 0\) and a finite \(\Lambda \subset \mathbb{Z}^d\), or given \(a = (a_1, \ldots, a_d)\) we introduce the partition functions

\[ Z(\varphi, \delta, \Lambda) = \max_E \sum_{x \in E} \exp \sum_{m \in \Lambda} \varphi(m x) \]

where the max is taken over all \((\delta, \Lambda)\)-separated sets, or

\[ Z(\varphi, a) = \sum_{x \in \Pi_\mathcal{I}} \exp \sum_{m \in \Lambda(a)} \varphi(m x). \]

We write

\[ P(\varphi, \delta, \Lambda) = \left(1/|\Lambda|\right) \log Z(\varphi, \delta, \Lambda), \]
\[ P(\varphi, a) = \left(1/|\Lambda(a)|\right) \log Z(\varphi, a). \]

2.2. Theorem. If \(0 < \delta < \delta^*\), the following limits exist:

\[ \lim_{\Lambda \to \infty} P(\varphi, \delta, \Lambda) = P(\varphi), \]
\[ \lim_{a \to \infty} P(\varphi, a) = P(\varphi), \]

and define a finite-valued convex function \(P\) on \(\mathcal{C}(X)\). Furthermore

\[ |P(\varphi) - P(\psi)| \leq ||\varphi - \psi|| \]

and if \(\tau_t \psi(x) = \psi(m x), t \in \mathbb{R},\)

\[ P(\varphi + \tau_t \psi - \psi + t) = P(\varphi) + t. \]
P is called the pressure.

Let \( \epsilon > 0 \); we choose \( \delta' > 0 \) so small that \( \delta + 2\delta' \leq \delta^\ast \) and

\[
(\delta(x, y) < \delta') \Rightarrow |\varphi(x) - \varphi(y)| < \epsilon;
\]

then take \( p(\delta') \) according to 1.2. Given \( a \), write \( b = (a_1 + p(\delta'), \ldots, a_r + p(\delta')) \).

We consider the partition \( (\Lambda(b) + r)_{r \in \mathbb{Z}^r} \) of \( \mathbb{Z}^r \). For a finite \( \Lambda \subset \mathbb{Z}^r \), let \( R = \{ r : \Lambda(b) + r \subset \Lambda \} \). Using specification we obtain

\[
Z(\varphi, \delta, \Lambda) \geq \text{exp}\left(-|\Lambda(b)|\epsilon\right) \exp\left(-|\Lambda(\delta)|\epsilon\right) \exp\left(-|\Lambda(a)|\epsilon\right)
\]

Since \( \Pi_a \) is \((\delta^\ast, \Lambda(a))\)-separated by expansiveness, we have also

\[
Z(\varphi, \delta^\ast, \Lambda(a)) \geq Z(\varphi, a).
\]

If \( \Lambda \uparrow \infty \) we have \( |R|/|\Lambda(b)| \rightarrow 1 \), and therefore (2.10) and (2.11) yield

\[
\lim_{\Lambda \uparrow \infty} \inf P(\varphi, \delta, \Lambda) \geq \frac{|\Lambda(\delta)|}{|\Lambda(b)|} \cdot [P(\varphi, a) - \epsilon] - \left(1 - \frac{|\Lambda(\delta)|}{|\Lambda(b)|}\right)\epsilon
\]

Suppose now that \( \delta' < \frac{1}{2} \delta \), and let \( N \) be the cardinal of a finite cover of \( \Omega \) by sets of diameter \( < \delta \). Let \( F \) be a \((\delta', \Lambda(b))\)-separated set such that

\[
Z(\varphi, \delta', \Lambda(b)) = \sum_{y \in F} \exp \sum_{m \in \Lambda(b)} \varphi(my).
\]

Given \( x \in E \) and \( r \in R \) we choose \( y \in F \) such that \( d((r + m)x, my) < \delta' \), for all \( m \in \Lambda(b) \). The mapping \((x, r) \rightarrow y\) defines an injection \( E \rightarrow FR \), and therefore

\[
Z(\varphi, \delta, \Lambda) \leq \left[Z(\varphi, \delta', \Lambda(b)) \exp(|\Lambda(b)|\epsilon)\text{exp}(|\Lambda(\delta)|\epsilon)\text{exp}(-|\Lambda(c)|\epsilon)\right]
\]

Taking \( c = (b_1 + p(\delta'), \ldots, b_r + p(\delta')) \), strong specification gives

\[
Z(\varphi, \delta', \Lambda(b)) \exp(-|\Lambda(\delta)|\epsilon) \exp(-|\Lambda(c)|\epsilon) \leq Z(\varphi, c).
\]

From (2.13) and (2.14) we obtain

\[
\lim_{\Lambda \uparrow \infty} \sup P(\varphi, \delta, \Lambda) \leq \frac{|\Lambda(c)|}{|\Lambda(b)|} P(\varphi, c) + 2\epsilon + \left(1 - \frac{|\Lambda(c)|}{|\Lambda(b)|}\right)\epsilon
\]

Letting \( a \rightarrow \infty \) in (2.12) and (2.15) we obtain (2.5) and (2.6).

The finiteness of \( P(\varphi) \) follows from \( \exp(-|\Lambda|\epsilon) \leq Z(\varphi, \delta, \Lambda) \leq N|\Lambda|\epsilon \exp(|\Lambda|\epsilon) \).

The other properties follow from Lemma 2.3 below.

2.3. Lemma. \( P(\varphi, \delta, \Lambda) \) is a convex function of \( \varphi \). Furthermore \( |P(\varphi, \delta, \Lambda) - P(\psi, \delta, \Lambda)| \leq ||\varphi - \psi|| \) and \( P(\varphi + t, \delta, \Lambda) = P(\varphi, \delta, \Lambda) + t \), if \( t \in \mathbb{R} \). Similar
properties hold for $P(q, a)$, and also $P(q + \tau_m \psi - \psi, a) = P(q, a)$.

We have $P(q, \delta, A) = \max_E P(q)$ where

$$P(q) = \frac{1}{|A|} \log Z(q), \quad Z(q) = \sum_{x \in E} \exp \sum_{m \in A} \varphi(m x),$$

$$\frac{d}{dt} P(q + \psi) = \frac{1}{Z(q + \psi)} \sum_{x} \left[ \sum_{m} \psi(m x) \right] \exp \sum_{m} \varphi(m x) + \psi(m x).$$

Therefore

$$\frac{d^2}{dt^2} P(q + \psi) \bigg|_{t=0} = \frac{1}{Z^2} \sum_{x} \sum_{y} \frac{1}{2} \left[ \sum_{m} \psi(m x) - \sum_{m} \psi(m y) \right]^2 \exp \sum_{m} [\varphi(m x) + \varphi(m y)] \geq 0.$$

On the other hand $|dP(q + \psi)/dt| \leq \|\psi\|$; hence

$$|P(q) - P(\psi)| \leq \sup_{0 \leq t \leq 1} \left| \frac{d}{dt} P(q + t(\psi - \psi)) \right| \leq \|\psi - \psi\|.$$

Finally $Z(q + \psi) = e^{\psi A} Z(q), Z(q + \psi - \psi, a) = Z(q, a).

2.4. Remark. Let $\Sigma$ be the subgroup of $Z'$ with linearly independent generators $s_1, \ldots, s_\sigma$ and define $\Lambda(\Sigma) = \{ m \in Z' : m = \sum_i t_i s_i \text{ with } t_i \text{ real}, 0 \leq t_i < 1 \}$. If a suitable extension of the strong specification property holds, one can prove

$$P(q) = \lim_{\Lambda(\Sigma) \to \infty} \frac{1}{|\Lambda(\Sigma)|} \log \sum_{x \in \Pi_x} \exp \sum_{m \in \Lambda(\Sigma)} \varphi(m x),$$

where $\Pi_x = \{ x : \Sigma x = \{ x \} \}$.

On the other hand, except for (2.6), Theorem 2.2 can be proved without the strong specification property (but assuming expansiveness and weak specification).

3. Equilibrium states.

3.1. Definition. Let $\mu_{q,a}$ be the measure on $\Omega$ which is carried by $\Pi_a$ and gives $x \in \Pi_a$ the mass

$$\mu_{q,a}(\{ x \}) = Z(q, a)^{-1} \exp \sum_{m \in \Lambda(a)} \varphi(m x).$$

3.2. Theorem. (a) Let $I_q \subset C(\Omega)^*$ be the set of measures $\mu$ such that

$$P(q + \psi) \geq P(q) + \mu(\psi)$$

for all $\psi$ (equilibrium states for $q$). Then $I_q$ is nonempty and there is a residual $(2)$ set $D \subset C(\Omega)$ such that $I_q$ consists of a single point $\mu_q$ if $q \in D$.

$(2)$ I.e. $D$ is a countable intersection of dense open subsets of $C(\Omega)$; in particular $D$ is dense in $C(\Omega)$ by Baire's theorem.
(b) $I_\varphi$ is convex, (vaguely) compact, and consists of $\mathbb{Z}^d$-invariant probability measures.

(c) The probability measure $\mu_{\varphi, a}$ is $\mathbb{Z}^d$-invariant, and

\begin{equation}
\mu_{\varphi, a}(\psi) = dP(\varphi + t \psi, a)/dt \big|_{t=0}.
\end{equation}

(d) If $\mu$ is a (vague) limit point of the $(\mu_{\varphi, a})$ when $a \to \infty$, then $\mu \in I_\varphi$. In particular, if $\varphi \in D$,

\begin{equation}
\lim_{a \to \infty} \mu_{\varphi, a} = \mu_{\varphi}.
\end{equation}

(e) If $\mathcal{B}$ is dense in $C(\Omega)$ and is a separable Banach space with respect to a norm $||\cdot|| \geq ||\cdot||$, then $D \cap \mathcal{B}$ is residual in $\mathcal{B}$.

(a) holds for any convex continuous function $P$ on a separable Banach space (see Dunford-Schwartz [6, Theorem V.9.8]). This proves also (e).

Let $\mu$ satisfy (3.2). Then by (2.8),

\[ 0 = P(\varphi + \tau_m \psi - \psi) - P(\varphi) \geq \mu(\tau_m \psi - \psi) \geq -[P(\varphi - \tau_m \psi + \psi) - P(\varphi)] = 0 \]

so that $\mu$ is $\mathbb{Z}^d$-invariant. Using (2.7) and (2.8) we obtain also $\pm \mu(\psi) \leq P(\varphi \pm \psi) - P(\varphi) \leq ||\psi||$ and $\mu(1) = -\mu(-1) \geq -[P(\varphi - 1) - P(\varphi)] = 1$. Therefore $||\mu|| \leq 1, \mu(1) \geq 1$ which implies that $\mu \geq 0, ||\mu|| = 1$, i.e. $\mu$ is a probability measure. Clearly, $I_\varphi$ is convex and compact, and (b) is thus proved.

(c) follows readily from the definitions. From (3.3) and the convexity of $P(\varphi, a)$ (Lemma 2.3), we obtain

\[ P(\varphi + \psi, a) \geq P(\varphi, a) + \mu_{\varphi, a}(\psi). \]

If $\mu_{\varphi, a} \to \mu$ this yields (3.2), proving (d).

4. Entropy.(3)

4.1. Definitions. Let $\mathcal{A} = \{A_i\}_{i \in \mathcal{I}}$ be a finite Borel partition of $\Omega$, and $\Lambda$ a finite subset of $\mathbb{Z}^d$. We denote by $\mathcal{A}^\Lambda$ the partition of $\Omega$ consisting of the sets $A(k) = \cap_{m \in \Lambda} (-m)A_{k(m)}$ indexed by maps $k : \Lambda \to \mathcal{I}$. We write

\begin{equation}
S(\mu, \mathcal{A}) = -\sum_i \mu(A_i) \log \mu(A_i).
\end{equation}

Let $I$ be the (convex compact) set of $\mathbb{Z}^d$-invariant probability measures on $\Omega$.

4.2. Theorem. If $\mathcal{A}$ consists of sets with diameter $\leq \delta^*$, and $\mu \in I$, then

\begin{equation}
\lim_{\mathcal{A} \to \infty} \frac{1}{|\Lambda|} S(\mu, \mathcal{A}^\Lambda) = \inf_{\Lambda} \frac{1}{|\Lambda|} S(\mu, \mathcal{A}^\Lambda) = s(\mu).
\end{equation}

(3) See also J.-P. Conze, Entropie d'un groupe abélien de transformations [Z. Wahrscheinlichkeitstheorie Verw. Gebiete 25 (1972), 11–30].
This limit is finite \( \geq 0 \), and independent of \( \mathcal{A} \). Furthermore, \( s \) is affine upper semicontinuous on \( I \); \( s \) is called the entropy.

\( S(\mu, \mathcal{A}^\Lambda) \) is an increasing function of \( \Lambda \), and satisfies the strong subadditivity property

\[
(4.3) \quad S(\mu, \mathcal{A}^\Lambda_{1} \cup \mathcal{A}^\Lambda_{2}) + S(\mu, \mathcal{A}^\Lambda_{1} \cap \mathcal{A}^\Lambda_{2}) \leq S(\mu, \mathcal{A}^\Lambda_{1}) + S(\mu, \mathcal{A}^\Lambda_{2}).
\]

[These are well-known properties. The increase follows from increase of the logarithm. To prove strong subadditivity we write \( S(\mu, \mathcal{A}^\Lambda) = S_\Lambda \), and use the inequality \(-\log(1/t) \leq t - 1\), then

\[
S_{\Lambda_1 \cup \Lambda_2} + S_{\Lambda_1 \cap \Lambda_2} - S_{\Lambda_1} - S_{\Lambda_2} = -\sum_{k : \Lambda_1 \cap \Lambda_2 \rightarrow \mathcal{A}} \sum_{k' : \Lambda_1 \cap \Lambda_2 \rightarrow \mathcal{A}} \sum_{k'' : \Lambda_1 \rightarrow \mathcal{A}} \mu(A(k, k', k'')) \log \frac{\mu(A(k, k')) \mu(A(k'))}{\mu(A(k', k'')) \mu(A(k''))} \leq \sum_{k' k''} \mu(A(k, k', k'')) \left[ \frac{\mu(A(k, k')) \mu(A(k'))}{\mu(A(k', k'')) \mu(A(k''))} - 1 \right] = \sum_{k' k''} \mu(A(k, k')) \mu(A(k')) - \sum_{k' k''} \mu(A(k, k', k'')) = \sum_{k' k''} \mu(A(k, k')) - 1 = 0.]

\]

If \( \Lambda_1 \cap \Lambda_2 = \emptyset \), \( 4.3 \) becomes subadditivity:

\[
S(\mu, \mathcal{A}^\Lambda_{1} \cup \mathcal{A}^\Lambda_{2}) \leq S(\mu, \mathcal{A}^\Lambda_{1}) + S(\mu, \mathcal{A}^\Lambda_{2}).
\]

Since \( \mu \in I \) we have also \( S(\mu, \mathcal{A}^\Lambda) = S(\mu, \mathcal{A}^{\Lambda+}) \) and therefore,(4)

\[
(4.4) \quad \lim_{a \rightarrow \infty} \frac{1}{|\Lambda(a)|} S(\mu, \mathcal{A}^{\Lambda(a)}) = \inf_{a} \frac{1}{|\Lambda(a)|} S(\mu, \mathcal{A}^{\Lambda(a)}) = s.
\]

Given \( \epsilon > 0 \), choose \( a \) such that \( |\Lambda(a)|^{-1} S(\mu, \mathcal{A}^{\Lambda(a)}) \leq s + \epsilon \). Consider the partition \( (\Lambda(a) + r)_{r \in \mathcal{Z}^*(a)} \) of \( \mathcal{Z}^* \), and let \( R = \{ r \in \mathcal{Z}^*(a) : (\Lambda(a) + r) \cap \Lambda \neq \emptyset \} \).

If \( \Lambda_+ = \bigcup_{r \in R} (\Lambda(a) + r) \) we have by increase and subadditivity

\[
S(\mu, \mathcal{A}^\Lambda) \leq S(\mu, \mathcal{A}^{\Lambda+}) \leq |R| S(\mu, \mathcal{A}^{\Lambda(a)}) \leq |R| |\Lambda(a)|(s + \epsilon).
\]

But \( |R| |\Lambda(a)| / |\Lambda| \rightarrow 1 \) when \( \Lambda \uparrow \infty \), and therefore

\[
(4.5) \quad \limsup_{\Lambda \uparrow \infty} |\Lambda|^{-1} S(\mu, \mathcal{A}^\Lambda) \leq s + \epsilon.
\]

Strong subadditivity shows that

\[
(4.6) \quad S(\mu, \mathcal{A}^{\Lambda \cup \{m\}}) - S(\mu, \mathcal{A}^\Lambda) \geq S(\mu, \mathcal{A}^{\Lambda \cup \{m\}}) - S(\mu, \mathcal{A}^\Lambda)
\]

when \( m \notin \Lambda' \supset \Lambda \). This permits an estimate of the increase in the entropy for a set \( \Lambda \) to which points are added successively in lexicographic order. In

\[(4) \quad \text{See for instance [14, Proposition 7.2.4].}\]
particular if $\Lambda$ is fixed and one takes for $\Lambda'$ the sets successively obtained in the lexicographic construction of a large $\Lambda(a)$, (4.6) holds for most $\Lambda'$. Therefore

$$S(\mu, A^{\Lambda \cup \{m\}}) - S(\mu, A^\Lambda) \geq \lim_{a \to \infty} |\Lambda(a)|^{-1} S(\mu, A^{\Lambda(a)}) = s$$

and hence

$$(4.7) \quad S(\mu, A^\Lambda) \geq |\Lambda|s$$

for all $\Lambda$; (4.2) follows from (4.5) and (4.7).

Let $x \in \Omega$ and for each $m \in \Lambda$, let $B_m$ be the union of those $A_i$ which contain $x$ in their closure. Then $B_\Lambda = \bigcap_{m \in \Lambda} (-m)B_m$ contains $x$ in its interior and is a union of elements of $A^\Lambda$. If $y \in B_\Lambda$ and $\Lambda = \{m: |m| < q(\delta)\}$, then $d(x, y) < \delta$ (see (1.4)). Therefore the $\sigma$-field generated by the $A^\Lambda$ is the Borel $\sigma$-field. The Kolmogorov-Sinai theorem (see [20, 5.5]) holds for the group $\mathbb{Z}^\ast$ and implies that the limit (4.2) is independent of $\mu$ (it is clearly finite $\geq 0$).

If $\mu, \mu' \in I$, and $0 < \alpha < 1$, the following inequalities are standard:

$$\alpha S(\mu, A) + (1 - \alpha) S(\mu', A) \leq S(\alpha \mu + (1 - \alpha) \mu', A) \leq \alpha S(\mu, A) + (1 - \alpha) S(\mu', A) + \log 2.$$  

(4.8)

[Writing $\mu_i = \mu(A_i), \mu'_i = \mu(A'_i)$ we have indeed, using the convexity of $t \log t$ and the increase of $\log t$,

$$- \sum_i [\alpha \mu_i \log \mu_i + (1 - \alpha) \mu'_i \log \mu'_i]$$

$$\leq - \sum_i [\alpha \mu_i + (1 - \alpha) \mu'_i] \log [\alpha \mu_i + (1 - \alpha) \mu'_i]$$

$$\leq - \sum_i [\alpha \mu_i \log \alpha \mu_i + (1 - \alpha) \mu'_i \log (1 - \alpha) \mu'_i]$$

$$= - \sum_i [\alpha \mu_i \log \mu_i + (1 - \alpha) \mu'_i \log \mu'_i] - \alpha \log \alpha - (1 - \alpha) \log 1 - (1 - \alpha) \log (1 - \alpha)$$

$$\leq - \sum_i [\alpha \mu_i \log \mu_i + (1 - \alpha) \mu'_i \log \mu'_i] + \log 2.$$]

(4.8) implies that $s$ is affine.

To prove that $s$ is upper semicontinuous at $\mu$, choose $A$ such that the boundaries of the $A_i$ have $\mu$-measure zero. [If $x \in \Omega$ one can choose $\delta \leq \frac{1}{4} \delta^*$ such that the boundary of the sphere of radius $\delta$ centered at $x$ has $\mu$-measure 0. Take a finite covering of $\Omega$ by such spheres and let $A$ be generated by this covering.] The boundaries of the $A(k) \in A^\Lambda$ have also measure 0, hence

$$\lim_{\mu' \to \mu} \mu'(A(k)) = \mu(A(k)), \quad \lim_{\mu' \to \mu} S(\mu', A^\Lambda) = S(\mu, A^\Lambda),$$

and $s$ is upper semicontinuous as inf of continuous functions.

4.3. Remarks. (a) Theorem 4.2 reduces to the usual definition of the measure theoretic entropy for $\nu = 1$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
(b) The condition that the diameters of the $A_i$ are $\leq \delta^*$ can be replaced by the weaker condition that the $\mathcal{A}^\Lambda$ generate the Borel $\sigma$-field (see the proof).

(c) The proof of Theorem 4.2 assumes expansiveness, but specification is not used.

5. Variational principle.

5.1. Theorem. For all $\varphi \in \mathcal{C}(\Omega)$,

\begin{equation}
P(\varphi) = \max_{\mu \in I} [s(\mu) + \mu(\varphi)]
\end{equation}

and the maximum is reached precisely on $I$. For all $\mu \in I$,

\[ s(\mu) = \inf_{\varphi} [P(\varphi) - \mu(\varphi)]. \]

Let $\varphi \in \mathcal{C}(\Omega)$ and $\mu \in I$ be given. Since $\Omega$ is metrizable compact, there exists a finite set \{\psi_1, \ldots, \psi_t\} of elements of $\mathcal{C}(\Omega)$ such that if $|\psi_l(x) - \psi_i(y)| < 1$ for $l = 1, \ldots, t$, then $d(x, y) < \delta^*$. Given $\epsilon > 0$ and $a$ we construct a partition $\mathcal{B} = (B_j)_{j \in \mathcal{A}}$ consisting of sets of the form $B_j = \{x: u_{lm} \leq \psi(m x) < v_{lm}$ and $u_{lm} \leq \varphi(m x) < v_{lm}^n \text{ for all } i, l, \text{ and } m \in \Lambda(a)\}$. By suitable choice of the $u_{lm}$, $v_{lm}$, $u_{lm}^n$, $v_{lm}^n$ we can achieve that

(a) the diameter of each set $(-m)B_n$, for $m \in \Lambda(a)$, is $\leq \delta^*$;
(b) if $B_i, B_j$ are adjacent (i.e. $\overline{B_i} \cap \overline{B_j} \neq \emptyset$) and $x \in B_i, y \in B_j$, then

\[ |\varphi(m x) - \varphi(m y)| < \epsilon/2 \text{ for all } m \in \Lambda(a); \]

(c) each $x \in X$ is contained in the closure of at most $(t + 1)|\Lambda(a)| + 1$ sets $B_i$.

Because of (c) there exists $\delta, 0 < \delta < \delta^*$, such that for each $x$ there are at most $(t + 1)|\Lambda(a)| + 1$ sets $B_i$ with distance $< \delta$ to $x$, and these sets are all adjacent to that containing $x$.

Let $R$ be a subset of $\mathcal{Z}^*(a)$, then

\begin{equation}
\inf_{R} \frac{1}{|R|} S(\mu, \mathcal{B}^R) = |\Lambda(a)|s(\mu).
\end{equation}

To see this notice that the $\mathcal{B}^R$ generate the Borel $\sigma$-field (by (a) above), and apply Remark 4.3(b) with $\mathcal{Z}^*$ replaced by $\mathcal{Z}^*(a)$. It follows that the left-hand side of (5.3) is not changed if $\mathcal{B}$ is replaced by $\mathcal{A}^\Lambda(a)$, and (5.3) follows. If $E$ is a maximal $(\delta, R)$-separated set, for each $k: R \rightarrow \mathcal{A}$ such that $B(k) \neq \emptyset$, one can choose $x \in B(k)$ and then $x_k \in E$ such that $d(r x_k, r x) < \delta$, all $r \in R$. By the choice of $\delta$, $r x_k$ is in a set $B_i$ adjacent to $B_{k(r)}$. Therefore, by (b),

\[ \left| \sum_{m \in \Lambda(a)} \varphi((r + m)x_k) - \sum_{m \in \Lambda(a)} \varphi(mx) \right| < |\Lambda(a)|\epsilon/2 \]

(4) The $B_i$ may be viewed as $(t + 1)|\Lambda(a)|$-dimensional rectangles and they can be adjusted so that at most $(t + 1)|\Lambda(a)| + 1$ meet at a corner. This idea is used by Goodwyn [8].
for all $v \in B_{k(r)}$. Choose $y_i \in B_i$ for each $i \in \mathcal{I}$, then

$$
\frac{1}{|R|} \sum_{k: R \to \mathcal{I}} \mu(B(k)) \sum_{r \in \mathcal{R}} \sum_{m \in \Lambda(a)} \varphi((r + m)x_k)
\geq \frac{1}{|R|} \sum_{r \in \mathcal{R}} \sum_{i \in \mathcal{I}} \sum_{k: k(r) = i} \mu(B(k)) \left[ \sum_{m \in \Lambda(a)} \varphi(my_i) - |\Lambda(a)| \epsilon/2 \right]
\geq \frac{1}{|R|} \sum_{i \in \mathcal{I}} \mu(B_i) \sum_{m \in \Lambda(a)} \varphi(my_i) - |\Lambda(a)| \epsilon/2
= \sum_{i \in \mathcal{I}} \mu(B_i) \sum_{m \in \Lambda(a)} \varphi(my_i) - |\Lambda(a)| \epsilon/2
\geq |\Lambda(a)| (\mu(\varphi) - \epsilon).

(5.4)

Notice that each $x_k \in E$ comes from at most $[(t + 1)|\Lambda(a)| + 1]|R|$ different $k$'s. Using this, and also (5.3), (5.4) and the concavity of the log, we obtain

$$
|\Lambda(a)| (s(\mu) + \mu(\varphi) - \epsilon)
\leq \frac{1}{|R|} \sum_{k} \mu(B(k)) \left[ -\log \mu(B(k)) + \sum_{r \in \mathcal{R}} \sum_{m \in \Lambda(a)} \varphi((r + m)x_k) \right]
= \frac{1}{|R|} \sum_{k} \mu(B(k)) \log \left( \exp \left( \sum_{r \in \mathcal{R}} \sum_{m \in \Lambda(a)} \varphi((r + m)x_k) \right)/\mu(B(k)) \right)
\leq \frac{1}{|R|} \log \sum_{k} \exp \left( \sum_{r \in \mathcal{R}} \sum_{m \in \Lambda(a)} \varphi((r + m)x_k) \right)
\leq \frac{1}{|R|} \log \left[ (t + 1)|\Lambda(a)| + 1 \right] \sum_{x \in E} \exp \sum_{r \in \mathcal{R}} \sum_{m \in \Lambda(a)} \varphi((r + m)x).$

If $\Lambda = \bigcup_{r \in \mathcal{R}} (\Lambda(a) + r)$ then $E$ is $(\delta, \Lambda)$-separated; therefore

$$
\frac{1}{|R|} \log \sum_{x \in E} \exp \sum_{r \in \mathcal{R}} \sum_{m \in \Lambda(a)} \varphi((r + m)x) \leq |\Lambda(a)| P(\varphi(\delta, \Lambda)).
$$

so that

$$
s(\mu) + \mu(\varphi) - \epsilon \leq P(\varphi(\delta, \Lambda)) + (1/|\Lambda(a)|) \log [(t + 1)|\Lambda(a)| + 1].$

By taking $|\Lambda(a)|$ large then letting $\Lambda \uparrow \infty$, this yields

(5.5) 

$$
s(\mu) + \mu(\varphi) \leq P(\varphi).
$$

We show now that equality holds in (5.5) for some $\mu$. Let $\langle u \rangle = (2^n, \ldots, 2^n)$ and let $\mu$ be a limit of the sequence $\mu_{\langle u \rangle}$. Choose now a partition of $\mathcal{A}$ consisting of sets with diameter $< \delta^*$, and with boundaries of $\mu$-measure 0. Given $\epsilon > 0$, there exists $u$ such that $s(\mu) + \epsilon/2 > (1/|\Lambda(\langle u \rangle)|) S(\mu, \mathcal{A}(\langle u \rangle))$ and since $\mu_{\langle u \rangle}(A(k)) \to \mu(A(k))$ when $v \to \infty$, one can choose $V \geq u$ such that if $v \geq V$,
where we have used the subadditivity of $\Lambda \to S(\mu, A^\Lambda)$, and then expansiveness.

Using the definition of $\mu_{\Psi, \phi}$ we obtain

$$s(\mu) + \epsilon > -\frac{1}{|\Lambda(\langle \nu \rangle)|} \sum_{x \in \Pi(\phi)} \mu_{\Psi, \phi}(\{x\}) \left[ \sum_{m \in \Lambda(\langle \nu \rangle)} \Psi(m x) - \log Z(\Psi, \langle \nu \rangle) \right]$$

and the desired result follows by letting $\mu_{\Psi, \phi} \to \mu$. We have thus proved (5.1).

Let $J_\phi = \{ \mu \in I: s(\mu) + \mu(\phi) = P(\phi) \}$; $J_\phi$ is the set where the affine upper semicontinuous function $\mu \to s(\mu) + \mu(\phi)$ reaches its maximum; hence $J_\phi$ is convex and compact. If $\mu \in J_\phi$, we have

$$P(\phi + \psi) \geq s(\mu) + \mu(\phi + \psi) = s(\mu) + \mu(\phi) + \mu(\psi)$$

$$= P(\phi) + \mu(\psi);$$

hence $\mu \in J_\phi$. Therefore $J_\phi \subset I_\phi$. If $J_\phi$ were different from $I_\phi$ one could find $\psi \in C(\Omega)$ such that

$$\sup_{\mu \in J_\phi} \mu(\psi) \neq \sup_{\mu \in I_\phi} \mu(\psi).$$

Let $\mu_n \in J_{\phi + \psi/n}$ and $\mu \in I_\phi$, we have

$$\mu(\psi) = n\mu(\psi/n) \leq n[P(\phi + \psi/n) - P(\phi)]$$

$$\leq n[P(\phi + \psi/n) - s(\mu_n) - \mu_n(\phi)]$$

$$= n[\mu_n(\phi + \psi/n) - \mu_n(\phi)] = \mu_n(\phi).$$

If $\mu^*$ is a limit point of the sequence $(\mu_n)$, then $\mu^* \in J_\phi$ (by upper semicontinuity of $s$), and therefore $\mu(\psi) \leq \mu^*(\psi)$ for all $\mu \in I_\phi$, in contradiction with (5.6). We have thus shown that $J_\phi = I_\phi$.

We want now to prove (5.2). We already know by (5.5) that $s(\mu) \leq P(\phi) - \mu(\phi)$ and it remains to show that by proper choice of $\phi$ the right-hand side becomes as close as desired to $s(\mu)$. Let $C = \{ (\mu, t) \in C(\Omega)^* \times R: \mu \in I$ and $0 \leq t \leq s(\mu) \}$. Since $s$ is affine upper semicontinuous, $C$ is convex and compact. Given $\mu^* \in I$ and $u > s(\mu^*)$ there exist (because $C$ is convex and compact) $\phi \in C(\Omega)$ and $c \in R$ such that

$$-\mu^*(\phi) + c = u, \quad -\mu(\phi) + c > s(\mu), \quad \text{for all } \mu \in I;$$
hence $-\mu(\varphi) + u + \mu^*(\varphi) > s(\mu)$ and we have, if $\mu \in L_\varphi$,
\begin{align*}
0 & \leq P(\varphi) - s(\mu^*) - \mu^*(\varphi) \\
& = s(\mu) + \mu(\varphi) - s(\mu^*) - \mu^*(\varphi) \\
& < u - s(\mu^*).
\end{align*}

The right-hand side is arbitrarily small and (5.2) follows.

5.2. Remark. If $\Omega$ is a basic set for an Axiom A diffeomorphism it is known [3] that $0 \in D$, i.e., the maximum of $s(\mu)$ is reached for just one $\mu \in I$. Further results on $D$ have been obtained for Anosov diffeomorphisms using methods of statistical mechanics [18].

6. The sets of invariant states. In this section we study the set $I$ of all $\mathcal{Z}'$-invariant probability measures and its relations with the $L_\varphi$.

6.1. Proposition. For each $\varphi \in \mathcal{C}(\Omega)$, $L_\varphi$ is a Choquet simplex, and a face (see [4]) of the simplex $I$.

It is well known that the set $I$ of invariant probability measures is a simplex.(6) If $\mu \in L_\varphi$, let $m_\mu$ be the unique probability measure on $I$, carried by the extremal points of $I$, and with resultant $\mu$. Writing $\hat{\varphi}(\nu) = \nu(\varphi)$, we have (see [4])
\begin{align*}
m_\mu(s + \hat{\varphi}) &= s(\mu) + \mu(\varphi) = P(\varphi);
\end{align*}

hence the support of $m_\mu$ is contained in $\{\nu \in I : s(\nu) + \nu(\varphi) = P(\varphi)\} = L_\varphi$. This shows that $L_\varphi$ is a simplex and a face of $I$.

6.2. Proposition. Suppose that $\mathcal{B}$ is dense in $\mathcal{C}(\Omega)$ and is a separable Banach space with respect to a norm $|||\cdot||| \geq ||\cdot||$. If $\varphi \in \mathcal{B}$, then $L_\varphi$ is the closed convex hull of the set of $\mu$ such that
\begin{align*}
\mu &= \lim_{n \to \infty} \mu_{\varphi(n)}, \quad \lim_{n \to \infty} |||\varphi(n) - \varphi||| = 0, \quad \varphi(n) \in D \cap \mathcal{B},
\end{align*}

where $D$ is defined in Theorem 3.2(a). This applies in particular with $\mathcal{B} = \mathcal{C}(\Omega)$.

We have $P(\varphi(n) + \psi) \geq P(\varphi(n)) + \mu_{\varphi(n)}(\psi)$ for all $\psi$, hence $P(\varphi + \psi) \geq P(\varphi) + \mu(\psi)$ so that $\mu \in L_\varphi$ if $\mu$ is of the above form.

Suppose now that $L_\varphi$ were not in the closed convex hull of those $\mu$. There would then exist $\psi \in \mathcal{B}$ such that
\begin{align*}
\sup_{\nu \in L_\varphi} \nu(\psi) > \sup_{\mu} \mu(\psi).
\end{align*}

Let $\varphi(n) = \varphi + \psi/n + \chi_n \in D \cap \mathcal{B}$; then, by convexity of $P$, if $\nu \in L_\varphi$,
\begin{align*}
\nu(\psi/n + \chi_n) & \leq \mu_{\varphi(n)}(\psi/n + \chi_n).
\end{align*}

(6) See for instance Jacobs [10, p. 162].
Using Theorem 3.2(e) we may take $|||x||| < 1/n^2$; we have thus

$$\nu(\psi) - 1/n \leq \mu_{\psi_0}(\psi) + 1/n,$$

and if $\mu^*$ is a limit point of $(\mu_{\psi_0})$, $\nu(\psi) \leq \mu^*(\psi)$ in contradiction with (6.1).

6.3. Proposition. The set of measures $\mu$ on $\Omega$ such that

$$\mu(\varphi) \leq P(\varphi) \quad \text{for all } \varphi \in \mathcal{C}(\Omega)$$

is $I$.

If $\mu \in I$ we have $\mu(\varphi) \leq P(\varphi) - s(\mu) \leq P(\varphi)$ because $s \geq 0$. Let now (6.2) hold for some $\mu \in \mathcal{C}(\Omega)^*$. By (2.8) we have

$$\mu(\varphi) - \mu(\tau_{m}\varphi) = t^{-1}\mu(\tau_{p} - t\tau_{m}\varphi) \leq t^{-1}P(\tau_{p} - t\tau_{m}\varphi) = t^{-1}P(0).$$

Letting $t \to \infty$ gives $\mu(\varphi) - \mu(\tau_{m}\varphi) \leq 0$. Replacing $\varphi$ by $-\varphi$ yields $\mu(\varphi) = \mu(\tau_{m}\varphi)$. Therefore $\mu$ is $Z^*$ invariant. Using now (2.7) and (2.8) we find

$$\pm \mu(\varphi) = \lim_{t \to \pm \infty} \frac{1}{|t|} \mu(\tau_{p} \varphi) \leq \lim_{t \to \pm \infty} \frac{1}{|t|} P(\tau_{p} \varphi)$$

$$\leq \lim_{t \to \pm \infty} \frac{1}{|t|} [P(0) + \|\tau_{p} \varphi\|] = \|\varphi\|$$

so that $||\mu|| \leq 1$. Furthermore (2.8) shows that, for all $t$, $t\mu(1) = \mu(t) \leq P(0) + t$, so that $\mu(1) = 1$. Since $||\mu|| = 1$ and $\mu(1) = 1$, $\mu$ is a probability measure.

6.4. Proposition. The set

$$\mathcal{M}_p = \{ -\frac{1}{|A(\mu)|} \sum_{m \in A(\mu)} \delta_{m_x} : x \in \Pi_\mu \}$$

is dense in $I$.

A vague neighbourhood of $\mu \in I$ is given by $\{ \nu : ||\nu - \mu||_\psi < \varepsilon \text{ for } i = 1, \ldots, n \}$ where $||\nu - \mu||_\psi = ||\nu(\varphi_i) - \mu(\varphi_i)||$ and $\varphi_1, \ldots, \varphi_n \in \mathcal{C}(\Omega)$, $\varepsilon > 0$.

We assume without loss of generality that $||\varphi_i|| \leq 1$ for $i = 1, \ldots, n$.

Given $\varepsilon > 0$, we choose $\delta > 0$ such that $d(x, y) < \delta$ implies $|\varphi_i(x) - \varphi_i(y)| < \varepsilon$ for $i = 1, \ldots, n$.

Let $p(\delta)$ be given by $1.2, N > p(\delta)/\varepsilon$ and $a = (N, \ldots, N), b = (N + p(\delta), \ldots, N + p(\delta))$. By the density of measures with finite support we can choose $c_a > 0, x_a \in \Omega$ such that

$$\sum_a c_a = 1, \quad \left\| \sum_a c_a \delta_{x_a} - \mu \right\|_{\tau_{p} \varphi} < \varepsilon.$$
for $i = 1, \ldots, n$, and $m \in \Lambda(b)$. We have thus
\[
\left\| \sum_a c_a \delta_{mxa} - \mu \right\|_{\psi_i} < \epsilon \quad \text{for } m \in \Lambda(b);
\]
hence
\[
(6.3) \quad \left\| \frac{1}{|\Lambda(b)|} \sum_{m \in \Lambda(b)} \sum_a c_a \delta_{mxa} - \mu \right\|_{\psi_i} < \epsilon.
\]
By 1.3, we can choose $y_a \in \Pi_b$ such that $|\varphi_i(mxa) - \varphi_i(my_a)| < \epsilon$ for $m \in \Lambda(a)$, and we have $|\varphi_i(mxa) - \varphi_i(my_a)| < 2$ for $m \in \Lambda(b) \setminus \Lambda(a)$; hence
\[
(6.4) \quad \left\| \sum_{m \in \Lambda(b)} \sum_a \frac{c_a}{|\Lambda(b)|} \delta_{my_a} - \sum_{m \in \Lambda(b)} \sum_a \frac{c_a}{|\Lambda(b)|} \delta_{mxa} \right\|_{\psi_i} < \epsilon \frac{|\Lambda(a)|}{|\Lambda(b)|} + 2 \frac{|\Lambda(b)| - |\Lambda(a)|}{|\Lambda(b)|} < \epsilon + 2(1 + \epsilon)^* - 2.
\]
We can now find integers $P, M_\alpha > 0$ such that $\sum_a M_\alpha = P^*$ and
\[
(6.5) \quad \left\| \sum_a \frac{M_\alpha}{|\Lambda(b)|P^*} \sum_{m \in \Lambda(b)} \delta_{my_a} - \sum_{m \in \Lambda(b)} \sum_a \frac{c_a}{|\Lambda(b)|} \delta_{mxa} \right\|_{\psi_i} < \epsilon.
\]
Let $c = ((N + p(\delta))P, \ldots, (N + p(\delta))P)$. By application of (1.3), there exists $y \in \Pi_c$ such that when $\tilde{m}$ varies over $\Lambda(c)$, $my_\alpha$ takes $M_\alpha$ times a value close to $my_a$ for each $\alpha$ and each $m \in \Lambda(a)$. Close means $d(my_\alpha, my_a) < \delta$. Then
\[
(6.6) \quad \left\| \frac{1}{|\Lambda(c)|} \sum_{\tilde{m} \in \Lambda(c)} \delta_{my_\alpha} - \frac{1}{|\Lambda(b)|P^*} \sum_a M_\alpha \sum_{m \in \Lambda(b)} \delta_{mxa} \right\|_{\psi_i} \leq \epsilon \frac{|\Lambda(a)|}{|\Lambda(b)|} + 2 \frac{|\Lambda(b)| - |\Lambda(a)|}{|\Lambda(b)|} < \epsilon + 2(1 + \epsilon)^* - 2.
\]
Finally, (6.3), (6.4), (6.5), (6.6) give
\[
\left\| \frac{1}{|\Lambda(c)|} \sum_{\tilde{m} \in \Lambda(c)} \delta_{my_\alpha} - \mu \right\|_{\psi_i} < 4\epsilon + 4(1 + \epsilon)^* - 4,
\]
proving the proposition.

6.5. Proposition. (8) (a) The set of ergodic measures (extremal points of $I$) is residual in $I$.

(b) The set of measures with zero entropy is residual in $I$.

Since $\mathcal{M}_\mu$ is dense (Proposition 6.4) and consists of ergodic measures with zero entropy, it suffices to show that the set of ergodic measures and the set of measures with zero entropy are $G_\delta$ (i.e. countable intersections of open sets). For ergodic measures this is well known (see [4]); for measures with zero entropy, it follows from the fact that the entropy is upper semicontinuous.

(8) See Sigmund [15] where other residual sets are also discussed.
Added in proof. A proof of the variational principle (0, 1) has been obtained without the expansiveness and specification assumptions by P. Walters (preprint).

REFERENCES


INSTITUT DES HAUTES ETUDES SCIENTIFIQUES, BURES-SUR-YVETTE, FRANCE