

STATISTICAL MECHANICS ON A COMPACT SET WITH Z^V ACTION
SATISFYING EXPANSIVENESS AND SPECIFICATION

David RUELLE

I.H.E.S.

91, BURES-sur-YVETTE, France

Février 1972

[37]

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 Ω satisfying expansiveness and the specification property of Bowen [2]. Instead of a \mathbb{Z} action we consider a \mathbb{Z}^v action as is usual in lattice statistical mechanics, where $\Omega = F^{\mathbb{Z}^v}$ (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 Ω satisfying expansiveness and specification (see Section 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 Ω . The pressure P is a continuous convex function on $\mathcal{C}(\Omega)$ defined by

$$P(\varphi) = \lim_{a \rightarrow \infty} \frac{1}{a} \log Z(\varphi, a)$$
$$Z(\varphi, a) = \sum_{x \in \Pi_a} \exp \sum_{m=1}^a \varphi(T^m x)$$

(Sections 2 and 7). Let I be the set of probability measures on Ω , invariant under T with the vague topology. The (measure theoretic) entropy s is an affine upper semi-continuous function on I defined in the usual way (Section 4). The following variational principle holds (Section 5)

$$P(\varphi) = \max_{\mu \in I} [s(\mu) + \mu(\varphi)] \quad (0.1)$$

$$s(\mu) = \inf_{\varphi \in \mathcal{C}(\Omega)} [P(\varphi) - \mu(\varphi)]$$

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

$$P(\varphi + \psi) - P(\varphi) \geq \mu(\psi) \quad , \quad \text{all } \psi \in \mathcal{C}(\Omega) \quad .$$

Let $\mu_{\varphi, a}$ be the measure on Ω which is carried by Π_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 $a \rightarrow \infty$ is in I_φ (Section 3). There is a residual subset D of $\mathcal{C}(\Omega)$ such that I_φ consists of one single point μ_φ if $\varphi \in D$. In that case

$$\lim_{a \rightarrow \infty} \mu_{\varphi, a} = \mu_\varphi$$

Miscellaneous properties of invariant states are reviewed in Section 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 $m = (m_1, \dots, m_\nu) \in \mathbb{Z}^\nu$, $\nu \geq 1$, we let $\|m\| = \sup_i |m_i|$. Given integers $a_1, \dots, a_\nu > 0$ we define

$$\Lambda(a) = \{m \in \mathbb{Z}^\nu : 0 \leq m_i < a_i\}$$

If (Λ_α) is a directed family of finite subsets of \mathbb{Z}^ν , $\Lambda_\alpha \uparrow \infty$ means $|\Lambda_\alpha| \rightarrow \infty$ and $|\Lambda_\alpha + F| / |\Lambda_\alpha| \rightarrow 1$ for every finite $F \subset \mathbb{Z}^\nu$. In particular $\Lambda(a) \uparrow \infty$ when $a \rightarrow \infty$ (i.e. when $a_1, \dots, a_\nu \rightarrow \infty$).

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

The following assumptions are made ^{*)}

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

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

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

(i) if $i \neq j$, the distance of Λ_i, Λ_j (as subsets of \mathbb{Z}^ν , with the distance $\|\cdot\|$) is $> p(\delta)$

there is $x \in X$ such that

$$d(m_i x, m_i x_i) < \delta, \text{ all } i \in \mathcal{J}, \text{ all } m_i \in \Lambda_i$$

*) Cf. Bowen [2]

1.3. Strong specification. Let $Z^{\nu}(a)$ be the subgroup of Z^{ν} with generators $(a_1, 0, \dots, 0), \dots, (0, \dots, a_{\nu})$, and let $\Pi_a = \{x \in \Omega : Z^{\nu}(a) x = \{x\}\}$. For any families $(\Lambda_i)_{i \in \mathcal{J}}$, $(x_i)_{i \in \mathcal{J}}$ satisfying

(ii) $\Lambda_i \subset \Lambda(a)$ for all i and, if $i \neq j$, the distance of

$$\Lambda_i + Z^{\nu}(a) \text{ and } \Lambda_j \text{ is } > p(\delta)$$

there is $x \in \Pi_a$ such that

$$d(m_i x, m_i x_i) < \delta, \text{ all } i \in \mathcal{J}, \text{ all } m_i \in \Lambda_i$$

It is easily seen that strong specification implies weak specification. If Ω 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

$$(d(mx, my) \leq \delta^* \text{ if } |m| < q(\delta)) \Rightarrow (d(x, y) < \delta) .$$

2. Partition functions and pressure.

2.1. Definitions Let $\delta > 0$; $E \subset \Omega$ is (δ, Λ) -separated if

$$(x, y \in E, \text{ and } d(mx, my) < \delta \text{ for all } m \in \Lambda) \Rightarrow (x = y)$$

Let $\varphi \in C(\Omega)$. Given $\delta > 0$ and a finite $\Lambda \subset \mathbb{Z}^V$, or given $a = (a_1, \dots, a_V)$ we introduce the partition functions

$$Z(\varphi, \delta, \Lambda) = \max_E \sum_{x \in E} \exp \sum_{m \in \Lambda} \varphi(mx) \quad (2.1)$$

where the max is taken over all (δ, Λ) separated sets, or

$$Z(\varphi, a) = \sum_{x \in \Pi_a} \exp \sum_{m \in \Lambda(a)} \varphi(mx) \quad (2.2)$$

We write

$$P(\varphi, \delta, \Lambda) = \frac{1}{|\Lambda|} \log Z(\varphi, \delta, \Lambda) \quad (2.3)$$

$$P(\varphi, a) = \frac{1}{|\Lambda(a)|} \log Z(\varphi, a) \quad (2.4)$$

2.2. Theorem. If $0 < \delta < \delta^*$, the following limits exist

$$\lim_{\Lambda \uparrow \infty} P(\varphi, \delta, \Lambda) = P(\varphi) \quad (2.5)$$

$$\lim_{a \rightarrow \infty} P(\varphi, a) = P(\varphi) \quad (2.6)$$

and define a finite-valued convex function P on $C(X)$. Furthermore

$$|P(\varphi) - P(\psi)| \leq \|\varphi - \psi\| \quad (2.7)$$

and if $\tau_m \psi(x) = \psi(mx)$, $t \in \mathbb{R}$,

$$P(\varphi + \tau_m \psi - \psi + t) = P(\varphi) + t \quad (2.8)$$

P is called the pressure

Let $\epsilon > 0$, we choose $\delta' > 0$ so small that $\delta + 2\delta' \leq \delta^*$ and

$$(d(x, y) < \delta') \Rightarrow |\varphi(x) - \varphi(y)| < \epsilon \quad (2.9)$$

then take $p(\delta')$ according to 1.2. Given a , write

$b = (a_1 + p(\delta'), \dots, a_n + p(\delta'))$. We consider the partition $(\Lambda(b) + r)_{r \in \mathbb{Z}^{\nu}(b)}$ of \mathbb{Z}^{ν} . For a finite $\Lambda \subset \mathbb{Z}^{\nu}$, let $R = \{r : \Lambda(b) + r \subset \Lambda\}$. Using specification we obtain

$$\begin{aligned} & Z(\varphi, \delta, \Lambda) \\ \geq & [Z(\varphi, \delta + 2\delta', \Lambda(a)) e^{-|\Lambda(a)| \epsilon} e^{-(|\Lambda(b)| - |\Lambda(a)|) \|\varphi\|} |R| e^{-(|\Lambda| - |R|) |\Lambda(b)| \|\varphi\|}] \quad (2.10) \end{aligned}$$

Since Π_a is $(\delta^*, \Lambda(a))$ separated by expansiveness, we have also

$$Z(\varphi, \delta^*, \Lambda(a)) \geq Z(\varphi, a) \quad (2.11)$$

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

$$\liminf_{\Lambda \uparrow \infty} P(\varphi, \delta, \Lambda) \geq \frac{|\Lambda(a)|}{|\Lambda(b)|} \cdot [P(\varphi, a) - \epsilon] \cdot \left(1 - \frac{|\Lambda(a)|}{|\Lambda(b)|}\right) \|\varphi\| \quad (2.12)$$

Suppose now that $\delta' < \frac{1}{2} \delta$, and let N be the cardinal of a finite cover of Ω 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

$$d((r+m)x, my) < \delta', \text{ for all } m \in \Lambda(b)$$

the mapping $x, r \rightarrow y$ defines an injection $A \rightarrow B^R$, and therefore

$$Z(\varphi, \delta, \Lambda) \leq [Z(\varphi, \delta', \Lambda(b)) e^{|\Lambda(b)| \epsilon}] |R| \left(\frac{1}{N} e^{\|\varphi\|}\right)^{|R|} (|\Lambda| - |R|) |\Lambda(b)| \quad (2.13)$$

Taking $c = (b_1 + p(\delta'), \dots, b_v + p(\delta'))$, strong specification gives

$$Z(\varphi, \delta', \Lambda(b)) e^{-|\Lambda(b)|\varepsilon} e^{-(|\Lambda(c)| - |\Lambda(b)|)\|\varphi\|} \leq Z(\varphi, c) \quad (2.14)$$

From (2.13) and (2.14) we obtain

$$\limsup_{\Lambda \uparrow \infty} P(\varphi, \delta, \Lambda) \leq \frac{|\Lambda(c)|}{|\Lambda(b)|} P(\varphi, c) + 2\varepsilon + \left(\frac{|\Lambda(c)|}{|\Lambda(b)|} - 1\right)\|\varphi\| \quad (2.15)$$

Letting $\varepsilon \rightarrow 0$ in (2.12) and (2.15) we obtain (2.5) and (2.6).

The finiteness of $P(\varphi)$ follows from

$$e^{-|\Lambda|\|\varphi\|} \leq Z(\varphi, \delta, \Lambda) \leq N^{|\Lambda|} e^{|\Lambda|\|\varphi\|}$$

The other properties follow from Lemma 2.3 below.

2.3 Lemma. $P(\varphi, \delta, \Lambda)$ is a convex function of φ , Furthermore

$$|P(\varphi, \delta, \Lambda) - P(\psi, \delta, \Lambda)| \leq \|\varphi - \psi\|$$

and

$$P(\varphi + t, \delta, \Lambda) = P(\varphi, \delta, \Lambda) + t, \text{ if } t \in \mathbb{R}$$

similar properties hold for $P(\varphi, a)$, and also

$$P(\varphi + \frac{1}{m}(\psi - \varphi), a) = P(\varphi, a)$$

We have $P(\varphi, \delta, \Lambda) = \max_E p(\varphi)$ where

$$p(\varphi) = \frac{1}{|\Lambda|} \log Z(\varphi)$$

$$Z(\varphi) = \sum_{x \in E} \exp \sum_{m \in \Lambda} \varphi(mx)$$

$$\frac{d}{dt} p(\varphi + t\psi) = \frac{1}{Z(\varphi + t\psi)} \sum_x \frac{1}{|\Lambda|} \left[\sum_{m'} \psi(m'x) \right] \exp \sum_m [\varphi(mx) + t\psi(mx)]$$

Therefore

$$\begin{aligned}
 & \left| \Lambda \left| \frac{d^2}{dt^2} p(\varphi + t\psi) \right|_{t=0} \right. \\
 &= \frac{1}{Z^2} \sum_x \sum_y \frac{1}{2} \left[\sum_m \psi(m, x) - \sum_m \psi(m, y) \right]^2 \exp \sum_m [\varphi(m, x) + \varphi(m, y)] \geq 0
 \end{aligned}$$

On the other hand

$$\left| \frac{d}{dt} p(\varphi + t\psi) \right| \leq \|\psi\|$$

hence

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

Finally

$$\begin{aligned}
 Z(\varphi + t) &= e^{|\Lambda|t} Z(\varphi) \\
 Z(\varphi + \tau \frac{\psi}{m} - \psi, a) &= Z(\varphi, a)
 \end{aligned}$$

2.4. Remark. Let Σ be the subgroup of \mathbb{Z}^v with linearly independent generators s_1, \dots, s_v , and define

$$\Lambda(\Sigma) = \{m \in \mathbb{Z}^v : m = \sum_{i=1}^v t_i s_i \text{ with } t_i \text{ real, } 0 \leq t_i < 1\}$$

If a suitable extension of the strong specification property holds, one can prove

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

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

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

$$\mu_{\varphi, a}(\{x\}) = Z(\varphi, a)^{-1} \exp \sum_{m \in \Lambda(a)} \varphi(m x) \quad (3.1)$$

3.2. Theorem (a) Let $I_\varphi \subset \mathcal{C}(\Omega)^*$ be the set of measures μ such that

$$P(\varphi + \psi) \geq P(\varphi) + \mu(\psi) \quad (3.2)$$

for all ψ (equilibrium states for φ), Then I_φ is non empty and there is a residual ^{*} set $D \subset \mathcal{C}(\Omega)$ such that I_φ consists of a single point μ_φ if $\varphi \in D$.

(b) I_φ is convex, (vaguely) compact, and consists of \mathbb{Z}^V invariant probability measures.

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

$$\mu_{\varphi, a}(\psi) = \left. \frac{d}{dt} P(\varphi + t\psi, a) \right|_{t=0} \quad (3.3)$$

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

$$\lim_{a \rightarrow \infty} \mu_{\varphi, a} = \mu_\varphi \quad (3.4)$$

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

^{*}) i.e. D is a countable intersection of dense open subsets of $\mathcal{C}(\Omega)$, in particular D is dense in $\mathcal{C}(\Omega)$ by Baire's theorem.

Let μ 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 μ is \mathbb{Z}^V 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. μ is a probability measure. Clearly, I_φ is convex and compact, and (b) is thus proved.

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

$$P(\varphi + \psi, a) \geq P(\varphi, a) + \mu_{\varphi, a}(\psi)$$

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

4. Entropy *)

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

$$S(\mu, G) = - \sum_i \mu(A_i) \log \mu(A_i) \quad (4.1)$$

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

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

$$\lim_{\Lambda \uparrow \infty} \frac{1}{|\Lambda|} S(\mu, G^\Lambda) = \inf_{\Lambda} \frac{1}{|\Lambda|} S(\mu, G^\Lambda) = s(\mu) \quad (4.2)$$

This limit is finite ≥ 0 , and independent of G . Furthermore, s is affine upper semi-continuous on I ; s is called the entropy.

$S(\mu, G^\Lambda)$ is an increasing function of Λ , and satisfies the strong subadditivity property

$$S(\mu, G^{\Lambda_1 \cup \Lambda_2}) + S(\mu, G^{\Lambda_1 \cap \Lambda_2}) \leq S(\mu, G^{\Lambda_1}) + S(\mu, G^{\Lambda_2}) \quad (4.3)$$

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

$$\begin{aligned} & S_{\Lambda_1 \cup \Lambda_2} + S_{\Lambda_1 \cap \Lambda_2} - S_{\Lambda_1} - S_{\Lambda_2} \\ &= - \sum_{k: \Lambda_1 \cap \Lambda_2 \rightarrow \mathcal{G}} \sum_{k': \Lambda_1 \setminus \Lambda_2 \rightarrow \mathcal{G}} \sum_{k'': \Lambda_2 \setminus \Lambda_1 \rightarrow \mathcal{G}} \\ & \mu(A(k, k', k'')) \log \frac{\mu(A(k, k', k'')) \mu(A(k))}{\mu(A(k, k')) \mu(A(k, k''))} \end{aligned}$$

*) see also I.-P. Conze Entropie d'un groupe abélien de transformations, Preprint

$$\leq \sum_{kk'k''} \mu(A(k, k', k'')) \left[\frac{\mu(A(k, k'))\mu(A(k, k''))}{\mu(A(k, k', k''))\mu(A(k))} - 1 \right]$$

$$= \sum_{kk'} \frac{\mu(A(k, k'))}{\mu(A(k))} \sum_{k''} \mu(A(k, k'')) - \sum_{kk'k''} \mu(A(k, k', k''))$$

$$= \sum_{kk'} \mu(A(k, k')) - 1 = 0$$

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

$$S(\mu, \mathbb{G}^{\Lambda_1 \cup \Lambda_2}) \leq S(\mu, \mathbb{G}^{\Lambda_1}) + S(\mu, \mathbb{G}^{\Lambda_2})$$

Since $\mu \in \mathcal{I}$ we have also

$$S(\mu, \mathbb{G}^{\Lambda}) = S(\mu, \mathbb{G}^{\Lambda+m})$$

and therefore ^{*)}

$$\lim_{a \rightarrow \infty} \frac{1}{|\Lambda(a)|} S(\mu, \mathbb{G}^{\Lambda(a)}) = \inf_a \frac{1}{|\Lambda(a)|} S(\mu, \mathbb{G}^{\Lambda(a)}) = s \quad (4.4)$$

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

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

$$S(\mu, \mathbb{G}^{\Lambda}) \leq S(\mu, \mathbb{G}^{\Lambda_+}) \leq |R| S(\mu, \mathbb{G}^{\Lambda(a)}) \leq |R| |\Lambda(a)| (s + \epsilon)$$

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

$$\limsup_{\Lambda \uparrow \infty} |\Lambda|^{-1} S(\mu, \mathbb{G}^{\Lambda}) \leq s + \epsilon \quad (4.5)$$

Strong subadditivity shows that

$$S(\mu, \mathbb{G}^{\Lambda \cup \{m\}}) - S(\mu, \mathbb{G}^{\Lambda}) \geq S(\mu, \mathbb{G}^{\Lambda' \cup \{m\}}) - S(\mu, \mathbb{G}^{\Lambda'}) \quad (4.6)$$

when $m \notin \Lambda' \supset \Lambda$. This permits an estimate of the increase in the entropy

*) See for instance [14] Proposition 7.2.4.

for a set Λ to which points are added successively in lexicographic order. In particular if Λ is fixed and one takes for Λ' the sets successively obtained in the lexicographic construction of a large $\Lambda(a)$, (4.6) holds for most Λ' . Therefore

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

and hence

$$S(\mu, G^\Lambda) \geq |\Lambda|s \tag{4.7}$$

for all Λ ; (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 G^Λ . If $y \in B_\Lambda$ and $\Lambda = \{m : |m| < q(\delta)\}$, then $d(x, y) < \delta$ (see 1.4). Therefore the σ -field generated by the G^Λ is the Borel σ -field. The Kolmogorov-Sinai theorem (see [20] 5.5) holds for the group \mathbb{Z}^v and implies that the limit (4.2) is independent of G (it is clearly finite ≥ 0).

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

$$\alpha S(\mu, G) + (1-\alpha)S(\mu', G) \leq S(\alpha\mu + (1-\alpha)\mu', G) \leq \alpha S(\mu, G) + (1-\alpha)S(\mu', G) + \log 2 \tag{4.8}$$

[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 $\log t$

$$\begin{aligned} &= - \sum_i [\alpha \mu_i \log \mu_i + (1-\alpha)\mu'_i \log \mu'_i] \\ &\leq - \sum_i [\alpha \mu_i + (1-\alpha)\mu'_i] \log [\alpha \mu_i + (1-\alpha)\mu'_i] \end{aligned}$$

$$\begin{aligned} &\leq - \sum_i [\alpha \mu_i \log \alpha \mu_i + (1-\alpha) \mu'_i \log(1-\alpha) \mu'_i] \\ &= - \sum_i [\alpha \mu_i \log \mu_i + (1-\alpha) \mu'_i \log \mu'_i] - \alpha \log \alpha - (1-\alpha) \log(1-\alpha) \\ &\leq - \sum_i [\alpha \mu_i \log \mu_i + (1-\alpha) \mu'_i \log \mu'_i] + \log 2 \quad]. \end{aligned}$$

(4.8) implies that s is affine.

To prove that s is upper semi-continuous at μ , choose \mathcal{G} such that the boundaries of the A_i have μ -measure zero.

[If $x \in \Omega$ one can choose $\delta \leq \frac{1}{2} \delta^*$ such that the boundary of the sphere of radius δ centered at x has μ -measure 0. Take a finite covering of Ω by such spheres and let \mathcal{G} be generated by this covering]. The boundaries of the $A(k) \in \mathcal{G}^\wedge$ have also measure 0, hence

$$\lim_{\mu' \rightarrow \mu} \mu'(A(k)) = \mu(A(k))$$

$$\lim_{\mu' \rightarrow \mu} S(\mu', \mathcal{G}^\wedge) = S(\mu, \mathcal{G}^\wedge)$$

and s is upper semi-continuous as inf of continuous functions.

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

(b) The condition that the diameters of the A_i are $\leq \delta^*$ can be replaced by the weaker condition that the \mathcal{G}^\wedge generate the Borel σ -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 $\varphi \in \mathcal{C}(\Omega)$,

$$P(\varphi) = \max_{\mu \in I} [s(\mu) + \mu(\varphi)] \quad (5.1)$$

and the maximum is reached precisely on I_φ . 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 Ω is metrizable compact, there exists a finite set $\{\psi_1, \dots, \psi_t\}$ of elements of $\mathcal{C}(\Omega)$ such that if $|\psi_\ell(x) - \psi_\ell(y)| < 1$ for $\ell = 1, \dots, t$, then $d(x, y) < \delta^*$. Given $\epsilon > 0$ and a we construct a partition $\mathcal{B} = (B_i)_{i \in \mathcal{J}}$ consisting of sets of the form

$$B_i = \{x : u_{i\ell m} \leq \psi_\ell(mx) < v_{i\ell m} \text{ and } u'_{im} \leq \varphi(mx) < v'_{im} \text{ for all } i, \ell, \text{ and } m \in \Lambda(a)\} .$$

By suitable choice of the $u_{i\ell m}$, $v_{i\ell m}$, u'_{im} , v'_{im} we can achieve that

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

$$|\varphi(mx) - \varphi(my)| < \epsilon/2 \quad \text{for all } m \in \Lambda(a)$$

- (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 δ , $0 < \delta < \delta^*$ 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 .

*) 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].

Let R be a subset of $Z^V(a)$, then

$$\inf_R \frac{1}{|R|} S(\mu, \mathcal{B}^R) = |\Lambda(a)| s(\mu) \quad (5.3)$$

To see this notice that the \mathcal{B}^R generate the Borel σ -field (by (a) above), and apply Remark 4.3(b) with Z^V replaced by $Z^V(a)$. It follows that the left-hand side of (5.3) is not changed if \mathcal{B} is replaced by $\mathcal{G}^{\Lambda(a)}$, and (5.3) follows. If E is a maximal (δ, R) separated set, for each $k \in R \rightarrow \mathcal{J}$ such that $B(k) \neq \emptyset$, one can choose $x \in B(k)$ and then $x_k \in E$ such that

$$d(r x_k, r x) < \delta, \quad \text{all } r \in R$$

By the choice of δ , $r x_k$ is in a set B_i adjacent to $B_{k(r)}$. Therefore, by (b),

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

for all $y \in B_{k(r)}$. Choose $y_i \in B_i$ for each $i \in \mathcal{J}$, then

$$\begin{aligned} & \frac{1}{|R|} \sum_{k:R \rightarrow \mathcal{J}} \mu(B(k)) \sum_{r \in R} \sum_{m \in \Lambda(a)} \varphi(r+m)x_k \\ & \geq \frac{1}{|R|} \sum_{r \in R} \sum_{i \in \mathcal{J}} \sum_{k:k(r)=i} \mu(B(k)) \left[\sum_{m \in \Lambda(a)} \varphi(my_i) - |\Lambda(a)| \epsilon/2 \right] \\ & = \frac{1}{|R|} \sum_{r \in R} \sum_{i \in \mathcal{J}} \mu(B_i) \sum_{m \in \Lambda(a)} \varphi(my_i) - |\Lambda(a)| \epsilon/2 \\ & = \sum_{i \in \mathcal{J}} \mu(B_i) \sum_{m \in \Lambda(a)} \varphi(my_i) - |\Lambda(a)| \epsilon/2 \\ & \geq |\Lambda(a)| (\mu(\varphi) - \epsilon) \end{aligned} \quad (5.4)$$

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

$$\begin{aligned}
 & |\Lambda(a)| (s(\mu) + \mu(\varphi) - \varepsilon) \\
 \leq & \frac{1}{|R|} \sum_k \mu(B(k)) [-\log \mu(B(k)) + \sum_{r \in R} \sum_{m \in \Lambda(a)} \varphi((r+m)x_k)] \\
 = & \frac{1}{|R|} \sum_k \mu(B(k)) \log \frac{\exp \sum_r \sum_m \varphi((r+m)x_k)}{\mu(B(k))} \\
 \leq & \frac{1}{|R|} \log \sum_k \exp \sum_r \sum_m \varphi((r+m)x_k) \\
 \leq & \frac{1}{|R|} \log [(t+1) |\Lambda(a)| + 1] \sum_{x \in E} \exp \sum_r \sum_m \varphi((r+m)x)
 \end{aligned}$$

If $\Lambda = \bigcup_{r \in R} (\Lambda(a) + r)$ then E is (δ, Λ) separated, therefore

$$\begin{aligned}
 & \frac{1}{|R|} \log \sum_{x \in E} \exp \sum_r \sum_m \varphi((r+m)x) \\
 \leq & |\Lambda(a)| P(\varphi, \delta, \Lambda)
 \end{aligned}$$

so that

$$s(\mu) + \mu(\varphi) - \varepsilon \leq P(\varphi, \delta, \Lambda) + \frac{1}{|\Lambda(a)|} \log [(t+1) |\Lambda(a)| + 1]$$

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

$$s(\mu) + \mu(\varphi) \leq P(\varphi) \tag{5.5}$$

We show now that equality holds in (5.5) for some μ . Let $\langle u \rangle = (2^u, \dots, 2^u)$ and let μ be a limit of the sequence $\mu_{\varphi, \langle u \rangle}$. Choose now a partition of G consisting of sets with diameter $< \delta^*$, and with boundaries of μ -measure 0. Given $\varepsilon > 0$, there exists u such that

$$s(\mu) + \varepsilon/2 > \frac{1}{|\Lambda(\langle u \rangle)|} s(\mu, G^{\Lambda(\langle u \rangle)})$$

and since $\mu_{\varphi, \langle v \rangle}(\Lambda(k)) \rightarrow \mu(\Lambda(k))$ when $v \rightarrow \infty$, one can choose $V \geq u$ such that if $v \geq V$,

$$\begin{aligned}
 s(\mu) + \epsilon &> \frac{1}{|\Lambda(\langle u \rangle)|} s(\mu_{\varphi, \langle v \rangle}, \mathbb{Q}^{\Lambda(\langle u \rangle)}), \\
 &\geq \frac{1}{|\Lambda(\langle v \rangle)|} s(\mu_{\varphi, \langle v \rangle}, \mathbb{Q}^{\Lambda(\langle v \rangle)}), \\
 &\geq \frac{1}{|\Lambda(\langle v \rangle)|} \sum_{x \in \Pi_{\langle v \rangle}} \mu_{\varphi, \langle v \rangle}(\{x\}) \log \mu_{\varphi, \langle v \rangle}(\{x\})
 \end{aligned}$$

where we have used the subadditivity of $\Lambda \rightarrow S(\mu, \mathbb{Q}^{\Lambda})$, and then expansiveness. Using the definition of $\mu_{\varphi, \langle v \rangle}$ we obtain

$$\begin{aligned}
 s(\mu) + \epsilon &> - \frac{1}{|\Lambda(\langle v \rangle)|} \sum_{x \in \Pi_{\langle v \rangle}} \mu_{\varphi, \langle v \rangle}(\{x\}) \left[\sum_{m \in \Lambda(\langle v \rangle)} \varphi(mx) - \log Z(\varphi, \langle v \rangle) \right] \\
 &= - \mu_{\varphi, \langle v \rangle}(\varphi) + \frac{1}{|\Lambda(\langle v \rangle)|} \log Z(\varphi, \langle v \rangle)
 \end{aligned}$$

and the desired result follows by letting $\mu_{\varphi, \langle v \rangle} \rightarrow \mu$. We have thus proved (5.1).

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

$$\begin{aligned}
 P(\varphi + \psi) &\geq s(\mu) + \mu(\varphi + \psi) = s(\mu) + \mu(\varphi) + \mu(\psi) \\
 &= P(\varphi) + \mu(\psi)
 \end{aligned}$$

hence $\mu \in I_{\varphi}$. Therefore $J_{\varphi} \subset I_{\varphi}$. If J_{φ} were different from I_{φ} one could find $\psi \in \mathcal{C}(\bar{\Omega})$ such that

$$\sup_{\mu \in I_{\varphi}} \mu(\psi) > \sup_{\mu \in J_{\varphi}} \mu(\psi) \tag{5.6}$$

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

$$\begin{aligned} \mu(\psi) &= n \mu(\psi/n) \leq n [P(\varphi + \psi/n) - P(\varphi)] \\ &\leq n [P(\varphi + \psi/n) - s(\mu_n) - \mu_n(\varphi)] \\ &= n [\mu_n(\varphi + \psi/n) - \mu_n(\varphi)] = \mu_n(\psi) \end{aligned}$$

If μ^* is a limit point of the sequence (μ_n) , then $\mu^* \in J_\varphi$ (by upper semi-continuity of s), and therefore

$$\mu(\psi) \leq \mu^*(\psi)$$

for all $\mu \in I_\varphi$, in contradiction with (5.6). We have thus shown that $J_\varphi = I_\varphi$.

We want now to prove (5.2). We already know by (5.5) that

$$s(\mu) \leq P(\varphi) - \mu(\varphi)$$

and it remains to show that by proper choice of φ the right hand side becomes as close as desired to $s(\mu)$. Let

$$C = \{(\mu, t) \in \mathcal{C}(\Omega)^* \times \mathbb{R} : \mu \in I \text{ and } 0 \leq t \leq s(\mu)\}$$

Since s is affine upper semi-continuous, C is convex and compact. Given $\mu^* \in I$ and $u > s(\mu^*)$ there exist (because C is convex and compact) $\varphi \in \mathcal{C}(\Omega)$ and $c \in \mathbb{R}$ such that

$$\begin{aligned} -\mu^*(\varphi) + c &= u \\ -\mu(\varphi) + c &\geq s(\mu) \quad \text{for all } \mu \in I \end{aligned}$$

hence

$$-\mu(\varphi) + u + \mu^*(\varphi) \geq s(\mu)$$

and we have, if $\mu \in I_\varphi$,

$$\begin{aligned} 0 &\leq P(\varphi) - s(\mu^*) - \mu^*(\varphi) \\ &= s(\mu) + \mu(\varphi) - s(\mu^*) - \mu^*(\varphi) \\ &< u - s(\mu^*) \end{aligned}$$

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

5.2. Remark. If Ω 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 \mathbb{Z}^V -invariant probability measures and its relations with the I_φ .

6.1. Proposition. For each $\varphi \in \mathcal{C}(\Omega)$, I_φ 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^{*)}. If $\mu \in I_\varphi$, let m_μ be the unique probability measure on I , carried by the extremal points of I , and with resultant μ writing $\hat{\varphi}(v) = v(\varphi)$, we have (see [4])

$$m_\mu(s + \hat{\varphi}) = s(\mu) + \mu(\varphi) = P(\varphi)$$

hence the support of m_μ is contained in $\{v \in I : s(v) + v(\varphi) = P(\varphi)\} = I_\varphi$. This shows that I_φ 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 I_φ is the closed convex hull of the set of μ such that

$$\mu = \lim_{n \rightarrow \infty} \mu_{\varphi(n)}, \quad \lim_{n \rightarrow \infty} \|\varphi(n) - \varphi\| = 0, \quad \varphi(n) \in D \cap \mathcal{B}$$

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 ψ , hence

*) See for instance Jacobs [10] p 162.

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

so that $\mu \in I_\varphi$ if μ is of the above form.

Suppose now that I_φ were not in the closed convex hull of those μ . There would then exist $\psi \in \mathcal{B}$ such that

$$\sup_{\nu \in I_\varphi} \nu(\psi) > \sup_{\mu} \mu(\psi) \quad (6.1)$$

Let $\varphi(n) = \varphi + \frac{1}{n}\psi + \chi_n \in D \cap \mathcal{B}$ then, by convexity of P , if $\nu \in I_\varphi$,

$$\nu\left(\frac{1}{n}\psi + \chi_n\right) \leq \mu_{\varphi(n)}\left(\frac{1}{n}\psi + \chi_n\right)$$

Using Theorem 3.2(e) we may take $\|\chi_n\| < \frac{1}{n^2}$, we have thus

$$\nu(\psi) - \frac{1}{n} \leq \mu_{\varphi(n)}(\psi) + \frac{1}{n}$$

and, if μ^* is a limit point of $(\mu_{\varphi(n)})$,

$$\nu(\psi) \leq \mu^*(\psi)$$

in contradiction with (6.1).

6.3. Proposition. The set of measures μ on Ω such that

$$\mu(\varphi) \leq P(\varphi) \quad \text{for all } \varphi \in \mathcal{C}(\Omega) \quad (6.2)$$

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

$$\mu(\varphi) - \mu(\tau_m \varphi) = \frac{1}{t} \mu(t\varphi - t\tau_m \varphi) \leq \frac{1}{t} P(t\varphi - t\tau_m \varphi) = \frac{1}{t} P(0)$$

Letting $t \rightarrow \infty$ gives $\mu(\varphi) - \mu(\tau_m \varphi) \leq 0$. Replacing φ by $-\varphi$ yields

$\mu(\varphi) = \mu(\tau_m \varphi)$. Therefore μ is \mathbb{Z}^V invariant. Using now (2.7) and (2.8)

we find

$$\begin{aligned} \underline{\pm} \mu(\varphi) &= \lim_{t \rightarrow \underline{\pm} \infty} \frac{1}{|t|} \mu(t\varphi) \leq \lim_{t \rightarrow \underline{\pm} \infty} \frac{1}{|t|} P(t\varphi) \\ &\leq \lim_{t \rightarrow \underline{\pm} \infty} \frac{1}{|t|} [P(0) + \|t\varphi\|] = \|\varphi\| \end{aligned}$$

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 $\mu(1) = 1$, μ is a probability measure.

6.4. Proposition ^{*)}. The set $m_p = \bigcup_a \left\{ \frac{1}{|\Lambda(a)|} \sum_{m \in \Lambda(a)} \delta_{mx} : x \in \Pi_a \right\}$ is dense in I ,

A vague neighbourhood of $\mu \in I$ is given by

$$\{v \in I : \|v - \mu\|_{\varphi_i} < \epsilon \text{ for } i = 1, \dots, n\}$$

where

$$\|v - \mu\|_{\varphi_i} = |v(\varphi_i) - \mu(\varphi_i)|$$

and $\varphi_1, \dots, \varphi_n \in \mathcal{C}(\Omega)$, $\epsilon > 0$. We assume without loss of generality that $\|\varphi_i\| \leq 1$ for $i=1, \dots, n$.

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

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

*) Sigmund [15] has proved this result by a somewhat different method for $v=1$.

$$\sum_{\alpha} c_{\alpha} = 1$$

$$\left\| \sum_{\alpha} c_{\alpha} \delta_{x_{\alpha}} - \mu \right\|_{\tau_m} \varphi_i < \varepsilon$$

for $i=1, \dots, n$, and $m \in \Lambda(b)$. We have thus

$$\left\| \sum_{\alpha} c_{\alpha} \delta_{mx_{\alpha}} - \mu \right\|_{\varphi_i} < \varepsilon \quad \text{for } m \in \Lambda(b)$$

hence

$$\left\| \frac{1}{|\Lambda(b)|} \sum_{m \in \Lambda(b)} \sum_{\alpha} c_{\alpha} \delta_{mx_{\alpha}} - \mu \right\|_{\varphi_i} < \varepsilon \quad (6.3)$$

By 1.3, we can choose $y_{\alpha} \in \Pi_b$ such that

$$|\varphi_i(m x_{\alpha}) - \varphi_i(m y_{\alpha})| < \varepsilon \quad \text{for } m \in \Lambda(a)$$

and we have

$$|\varphi_i(m x_{\alpha}) - \varphi_i(m y_{\alpha})| \leq 2 \quad \text{for } m \in \Lambda(b) \setminus \Lambda(a)$$

hence

$$\left\| \sum_{m \in \Lambda(b)} \sum_{\alpha} \frac{c_{\alpha}}{|\Lambda(b)|} \delta_{my_{\alpha}} - \sum_{m \in \Lambda(b)} \sum_{\alpha} \frac{c_{\alpha}}{|\Lambda(b)|} \delta_{mx_{\alpha}} \right\|_{\varphi_i}$$

$$< \varepsilon \frac{|\Lambda(a)|}{|\Lambda(b)|} + 2 \frac{|\Lambda(b)| - |\Lambda(a)|}{|\Lambda(b)|} < \varepsilon + 2(1+\varepsilon)^{\nu} - 2 \quad (6.4)$$

We can now find integers $P, M_{\alpha} > 0$ such that $\sum_{\alpha} M_{\alpha} = P^{\nu}$ and

$$\left\| \sum_{\alpha} \frac{M_{\alpha}}{|\Lambda(b)| P^{\nu}} \sum_{m \in \Lambda(b)} \delta_{my_{\alpha}} - \sum_{m \in \Lambda(b)} \sum_{\alpha} \frac{c_{\alpha}}{|\Lambda(b)|} \delta_{my_{\alpha}} \right\|_{\varphi_i} < \varepsilon \quad (6.5)$$

Let $c = ((N + p(\delta))P, \dots, (N + p(\delta))P)$. By application of 1.3, there exists $y \in \Pi_c$ such that when \tilde{m} varies over $\Lambda(c)$, $m y$ takes M_{α} times a value close to my_{α} for each α and each $m \in \Lambda(a)$. Close means $d(\tilde{m} y, my_{\alpha}) < \delta$. Then

$$\begin{aligned} & \left\| \frac{1}{|\Lambda(c)|} \sum_{m \in \Lambda(c)} \delta_{my} - \frac{1}{|\Lambda(b)| P^\nu} \sum_{\alpha} M_{\alpha} \sum_{m \in \Lambda(b)} \delta_{my} \alpha \right\| \varphi_i \\ & \leq \varepsilon \frac{|\Lambda(a)|}{|\Lambda(b)|} + 2 \frac{|\Lambda(b)| - |\Lambda(a)|}{|\Lambda(b)|} < \varepsilon + 2(1+\varepsilon)^\nu - 2 \end{aligned} \quad (6.6)$$

Finally, (6.3), (6.4), (6.5), (6.6) give

$$\left\| \frac{1}{|\Lambda(c)|} \sum_{m \in \Lambda(c)} \delta_{my} - \mu \right\| \varphi_i < 4 \varepsilon + 4(1+\varepsilon)^\nu - 4$$

proving the proposition.

6.5. Proposition ^{*}). (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 m_p 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_δ (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 semi-continuous.

* See Sigmund [15] where other residual sets are also discussed.

7. New definition of the pressure.

7.1. Definition. Given $\varphi \in C(\bar{\Omega})$ and a finite Borel partition $\mathcal{G} = (A_i)_{i \in \mathcal{J}}$ of $\bar{\Omega}$, we introduce the partition function

$$Z(\varphi, \mathcal{G}, \Lambda) = \sum_{k: \Lambda \rightarrow \mathcal{J}}^* \exp \sum_{m \in \Lambda} \varphi(mx_k) \quad (7.1)$$

where \sum^* is extended over those k such that $A(k) \neq \emptyset$ and $x_k \in A(k)$ (use the axiom of choice!). Let also μ_Λ be the measure carried by the x_k and such that

$$\mu_\Lambda(\{x_k\}) = Z(\varphi, \mathcal{G}, \Lambda)^{-1} \exp \sum_{m \in \Lambda} \varphi(mx_k)$$

7.2. Theorem. If the finite partition \mathcal{G} consists of Borel sets with diameter $< \delta^*$, then

$$\lim_{\Lambda \uparrow \infty} \frac{1}{|\Lambda|} \log Z(\varphi, \mathcal{G}, \Lambda) = P(\varphi) \quad (7.2)$$

We write

$$P(\varphi, \mathcal{G}, \Lambda) = \frac{1}{|\Lambda|} \log Z(\varphi, \mathcal{G}, \Lambda)$$

and, since this quantity also depends on the choice of the x_k , let P^+ (resp. P^-) be its upper (resp. lower) bound with respect to this choice. If $m \in \Lambda$ and all the points at distance $< q(\delta)$ of m are also in Λ , then $m \in A(k)$ has diameter $< \delta$ by 1.4. From this it follows readily that

$$\lim_{\Lambda \uparrow \infty} (P^+ - P^-) = 0 \quad (7.3)$$

Let the diameters of the A_i be $< \delta_0 < \delta^*$. If E is (δ_0, Λ) separated each $A(k)$ will contain at most one $x \in E$, therefore

$$P^+ \geq P(\varphi, \delta_0, \Lambda) \tag{7.4}$$

From (7.3) and (7.4) we obtain

$$\liminf_{\Lambda \uparrow \infty} P(\varphi, G, \Lambda) \geq P(\varphi) \tag{7.5}$$

We have

$$\begin{aligned} P(\varphi, G, \Lambda) &= \frac{1}{|\Lambda|} \log Z(\varphi, G, \Lambda) \\ &= - \frac{1}{|\Lambda|} \sum_k \mu_\Lambda(\{x_k\}) \log \mu_\Lambda(\{x_k\}) + \frac{1}{|\Lambda|} \sum_k \mu_\Lambda(\{x_k\}) \sum_m \varphi(m, x_k) \end{aligned}$$

Defining the probability measure $\bar{\mu}_\Lambda$ by

$$\bar{\mu}_\Lambda = \frac{1}{|\Lambda|} \sum_{m \in \Lambda} m \mu_\Lambda$$

(average over translations by $m \in \Lambda$) we find thus

$$P(\varphi, G, \Lambda) = \frac{1}{|\Lambda|} S(\mu_\Lambda, G^\Lambda) + \bar{\mu}_\Lambda(\varphi) \tag{7.6}$$

If $m \in \Lambda(a)$ we have, by subadditivity of the entropy,

$$S(\mu_\Lambda, G^\Lambda) \leq \sum_r^* S(\mu_\Lambda, G^{\Lambda(a)+r+m}) + \sigma_m$$

where \sum_r^* extends over those $r \in \mathbb{Z}^V(a)$ such that $\Lambda(a) + r + m \subset \Lambda$ and $\sigma_m / |\Lambda| \rightarrow 0$ when $\Lambda \uparrow \infty$. Therefore

$$\begin{aligned} S(\mu_\Lambda, G^\Lambda) &\leq \frac{1}{|\Lambda(a)|} \sum_{m \in \Lambda(a)} \sum_r^* S(\mu_\Lambda, G^{\Lambda(a)+r+m}) + \sigma \\ &= \frac{1}{|\Lambda(a)|} \sum_{m \in \Lambda'} S(\mu_\Lambda, G^{\Lambda(a)+m}) + \sigma \end{aligned}$$

where $\sigma/|\Lambda| \rightarrow 0$ and $\Lambda' = \{m \in \mathbb{Z}^V : \Lambda(a) + m \subset \Lambda\}$. Using (4.8) we have

$$s(\mu_\Lambda, \mathcal{G}^\Lambda) \leq \frac{|\Lambda'|}{|\Lambda(a)|} s(\mu', \mathcal{G}^{\Lambda(a)}) + \sigma \quad (7.7)$$

where

$$\mu' = \frac{1}{|\Lambda'|} \sum_{m \in \Lambda'} m \mu_\Lambda$$

If μ is a limit point of $(\bar{\mu}_\Lambda)$ for $\Lambda \uparrow \infty$ then $\mu \in I$ and, since

$\bar{\mu}_\Lambda - \mu' \rightarrow 0$, (7.7) gives

$$\lim \frac{1}{|\Lambda|} s(\mu_\Lambda, \mathcal{G}^\Lambda) \leq \frac{1}{|\Lambda(a)|} s(\mu, \mathcal{G}^{\Lambda(a)}) \quad (7.8)$$

From (7.6) and (7.8) we obtain,

$$\lim P(\varphi, \mathcal{G}, \Lambda) \leq \frac{1}{|\Lambda(a)|} s(\mu, \mathcal{G}^\Lambda) + \mu(\varphi)$$

for all a , and therefore

$$\limsup_{\Lambda \uparrow \infty} P(\varphi, \mathcal{G}, \Lambda) \leq \sup_{\mu \in I} (s(\mu) + \mu(\varphi)) = P(\varphi)$$

This, together with (7.5), proves the theorem.

7.3. Remark. If \mathcal{G} is a finite Borel covering of Ω rather than a partition, let

$$Z(\varphi, \mathcal{G}, \Lambda) = \inf_{K, (x_k)} \sum_{k \in K} \exp \sum_{m \in \Lambda} \varphi(m x_k) \quad (7.9)$$

where K is a set of maps $k : \Lambda \rightarrow \mathcal{J}$ such that the $A(k)$ cover Ω , and $x_k \in A(k)$. With this definition, Theorem 7.2 still holds. By refining \mathcal{G} to

a partition, we find indeed

$$\limsup_{\Lambda \uparrow \infty} \frac{1}{|\Lambda|} \log Z(\varphi, \mathcal{G}, \Lambda) \leq P(\varphi)$$

The converse inequality follows from an immediate adaptation of the first part of the proof of Theorem 7.2.

Taking for \mathcal{G} an open covering, we see from (7.9) that $P(0)$ is just the topological entropy [1]. Under our assumptions, the topological entropy can thus be computed from Borel coverings (or partitions) rather than open coverings. Notice also that if we set $\varphi = 0$ and $\nu = 1$ in (5.1) this variational principle appears now as a special case of the fact [5] [8] that the topological entropy is the sup of the measure theoretic entropy.

References.

- [1] R.L. ADLER, A.G. KONHEIM and M.H. Mc ANDREW. - Topological entropy. Trans. Amer. Math. Soc. 114, 309-319 (1965).
- [2] R. BOWEN. - Periodic points and measures for axiom A diffeomorphisms. Trans. Amer. Math. Soc. 154, 377-397 (1971).
- [3] R. BOWEN. - Markov partitions for axiom A diffeomorphisms. Amer. J. Math. 92, 725-747 (1970).
- [4] G. CHOQUET et P.A. MEYER. - Existence et unicité des représentations intégrales dans les convexes compacts quelconques. Ann. Inst. Fourier, 13, 139-154 (1963).
- [5] E.I. DINABURG. - The relation between topological entropy and metric entropy. D.A.N. SSSR, 190, 19-22 (1970), Engl. Transl. Soviet Math. 11, 13-16 (1970).
- [6] N. DUNFORD and J.T. SCHWARTZ. - Linear operators, I. General theory. Interscience, New York, 1958.
- [7] G. GALLAVOTTI and S. MIRACLE-SOLE. - Statistical mechanics of lattice systems. Commun. math. Phys. 5, 317-323 (1967).
- [8] L. GOODWYN. - Topological entropy bounds measure theoretic entropy. Proc. Amer. Math. Soc. 23, 679-688 (1969).
- [9] W.H. GOTTSCHALK and G.A. HEDLUND. - Topological dynamics Amer. Math. Soc. Colloq. Publ., vol 36, Providence, R.I., 1955.
- [10] K. JACOBS. - Lecture notes on ergodic theory. Aarhus Universitet, Aarhus, 1963.
- [11] O.E. LANFORD and D.W. ROBINSON. - Statistical mechanics of quantum spin systems III. Commun. math. Phys. 9, 327-338 (1968).
- [12] D.W. ROBINSON and D. RUELLE. - Mean entropy of states in classical statistical mechanics. Commun. math. Phys. 5, 288-300 (1967).
- [13] D. RUELLE. - A variational formulation of equilibrium statistical mechanics and the Gibbs phase rule. Commun. math. Phys. 5, 324-329 (1967).
- [14] D. RUELLE. - Statistical mechanics. Rigorous results. Benjamin, New York, 1969.