ON THE PREVALENCE OF ZERO ENTROPY

BY
KARL SIGMUND

ABSTRACT

For a residual subset of the space of discrete stationary stochastic processes (under the weak topology for measures) the entropy is 0. The processes with arbitrarily large entropy are dense. For a residual subset of subshifts of the shift on a finite alphabet, the topological entropy is 0.

Let $(X, \mathscr{B}, m)$ be a Lebesgue measure space and $\mathscr{T}$ the space of measure preserving transformations under the weak topology. In [6], Rohlin proved that for a residual set in $\mathscr{T}$, the entropy vanishes. In this sense, a measure preserving transformation has "in general" entropy zero.

There are two additional situations in which once again entropy is generically equal to zero. The first part of this note shows that in the space of real valued discrete stationary stochastic processes under the weak topology, the processes with zero entropy form a residual set. (Here a transformation is fixed and the measure varies.) In the second part, one considers the finite alphabet $\mathscr{S}$ and the space $\mathscr{S}^Z$ of bilateral sequences in $\mathscr{S}$. The space of subshifts of $\mathscr{S}^Z$ (that is, the space of compact shift invariant subsets of $\mathscr{S}^Z$) becomes a compact metric space under the Hausdorff metric. The set of subshifts with topological entropy zero is again a residual set.

It is a pleasure to thank Benjamin Weiss for helpful conversations, and especially for suggesting Theorem 3.

1. On the entropy of stationary stochastic processes

Let Z denote the integers, R the real line and R^Z the set of bilateral sequences in R with the product topology. Let $T: R^Z \to R^Z$ denote the shift transformation given by $(Tx)_k = x_{k+1}$, where x_k is the kth coordinate of $x \in R^Z (-\infty < k < \infty)$.

Received June 5, 1971

Let $\mathfrak{M}$ denote the space of (Borel) probability measures on R^Z which are invariant under T. Let $\mathfrak{M}$ be given the weak topology for measures: in this topology, a sequence $\mu_n \in \mathfrak{M}$ converges to $\mu \in \mathfrak{M}$ iff $\int f d\mu_n$ converges to $\int f d\mu$ for all continuous bounded real valued functions f on R^Z. Since R^Z is a complete separable metrizable space, $\mathfrak{M}$ is a complete separable metrizable space (see [5], th. 6.2 and 6.5).

A real valued discrete stationary stochastic process is a triple (R^Z, T, μ) where μ is in $\mathfrak{M}$.

A collection $\mathscr{C} = \{C_1, \cdots, C_r\}$ of finitely many subsets of R^Z is called a Borel partition of R^Z if the C_i are pairwise disjoint Borel sets whose union is R^Z. For $\mu \in \mathfrak{M}$ one writes

$$h_\mu(\mathscr{C}) = - \sum_{i=1}^{r} \mu(C_i) \log \mu(C_i)$$

If $\mathscr{C}$ is a Borel partition, one writes $\mathscr{C}^n$ for the Borel partition whose elements are $\bigcap_{k=0}^{n-1} T^{-k} C_{i_k}$ (with $C_{i_k} \in \mathscr{C}$) and defines

$$h_\mu(T,\mathscr{C}) = \lim \frac{1}{n} h_\mu(\mathscr{C}^n)$$

(this limit always exists) and

$$h_\mu(T) = \sup_{\mathscr{C}} h_\mu(T,\mathscr{C}).$$

$h_\mu(T)$ is called the entropy of the stochastic process (R^Z, T, μ). For fundamental theorems on entropy see [7].

We shall prove:

THEOREM 1. *For a dense G_δ of measures $\mu \in \mathfrak{M}$, one has $h_\mu(T) = 0$.*

THEOREM 2. *For any $K > 0$, the set of measures μ such that $h_\mu(T) \geqq K$ is dense in $\mathfrak{M}$.*

For the proof of Theorem 1 (which is similar to th. 6 in [8]), we define for every $n > 0$ a class ξ_n of subsets of R^Z which will be a Borel partition of R^Z into μ-continuity sets—not for all, but for "most" μ of $\mathfrak{M}$.

Let J_n denote the collection of integers j with $|j| \leqq n \cdot 2^n$. For any $(2n+1)$-tuple $(j_{-n}, \cdots, j_n)$ of elements of J_n, set

$$A(j_{-n}, \cdots, j_n) = \{x = (x_i) \in R^Z : j_k \cdot 2^{-n} \leqq x_k < (j_k + 1) \cdot 2^{-n} \text{ for all } k \in [-n, n]\}.$$

Let ξ_n be the class consisting of all sets of this form and the set $A(n)$ which is the complement of their union. ξ_n is a Borel partition of R^Z. Define

$$F(n) = \{x = (x_i) \in R^Z : x_k = j_k \cdot 2^{-n} \text{ for some } k \in [-n, n] \text{ and some } j_k \in J_n\}.$$

LEMMA 1. *$F(n) \subset F(n+1)$. The boundaries of elements of ξ_n are contained in $F(n)$. The boundaries of elements of the form $\bigcap_{k=-m}^{m} T^{-k}B_{i_k}$ (with $B_{i_k} \in \xi_n$) are contained in $\bigcup_{n=1}^{\infty} F(n)$.*

Let $\mathcal{N}$ denote the set of $\mu \in \mathfrak{M}$ such that $\mu(F(n)) = 0$ for $n = 1, 2, \cdots$.

LEMMA 2. *$\mathfrak{N}$ is a G_δ in $\mathfrak{M}$.*

PROOF. Let $C(r,n) = \{\mu \in \mathfrak{M}: \mu(F(n)) \geqq 1/r\}$. Suppose μ_k is a sequence in $C(r,n)$ converging to $\mu \in \mathfrak{M}$. Since $F(n)$ is closed, it follows from [5], th.6.1, that $\mu(F(n)) \geqq \limsup \mu_k(F(n)) \geqq 1/r$, and hence $\mu \in C(r,n)$. Thus $C(r,n)$ is closed, $\bigcup_{r=1}^{\infty}\bigcup_{n=1}^{\infty} C(r,n)$ is an F_σ and its complement is a G_δ in $\mathfrak{M}$.

Let $x \in R^Z$ be a periodic point of period s, and denote by μ_x the invariant measure concentrated on the periodic orbit of x: it has mass $1/s$ at the points $x, Tx, \cdots, T^{s-1}x$. Measures of this form will be called p.o.-measures. By a simple modification of [3], th. 2, one obtains

LEMMA 3. *Every $\mu \in \mathfrak{M}$ can be approximated by p.o.-measures belonging to $\mathfrak{N}$.*

LEMMA 4. *$\mathfrak{N}$ is a residual set in $\mathfrak{M}$.*

LEMMA 5. *The set $\mathfrak{N}_0 = \{\mu \in \mathfrak{N}: h_\mu(T) = 0\}$ is a dense G_δ in $\mathfrak{N}$.*

PROOF. (1) The denseness of $\mathfrak{N}_0$ follows from Lemma 3 since p.o.-measure have entropy zero.

(2) The sequence ξ_n is a nondecreasing sequence of partitions, and one has $\bigvee_{n=1}^{\infty}\xi_n = \varepsilon$, where ε is the partition of R^Z into distinct points. It follows by [7] th. 9.5 that $h_\mu(T, \xi_n) \uparrow h_\mu(T)$. Hence

$$\mathfrak{N}_0 = \{\mu \in \mathfrak{N}: h_\mu(T,\xi_n) = 0 \text{ for } n = 1, 2 \cdots\} = \bigcap_{n=1}^{\infty} \bigcap_{r=1}^{\infty} \{\mu \in \mathfrak{N}: h_\mu(T,\xi_n) < 1/r\}$$

Since by [7], th. 7.4.:

$$h_\mu(T,\xi_n) = \liminf \frac{1}{2m+1} h_\mu\left(\bigvee_{i=-m}^{m} T^{-i}\xi_n\right), \text{ one obtains}$$

$$\mathfrak{N}_0 = \bigcap_{n=1}^{\infty} \bigcap_{r=1}^{\infty} \bigcap_{k=1}^{\infty} \bigcup_{m=k}^{\infty} \left\{\mu \in \mathfrak{N}: \frac{1}{2m+1} h_\mu\left(\bigvee_{i=-m}^{m} T^{-i}\xi_n\right) < 1/r\right\}.$$

Let A be an element of $\bigvee_{i=-m}^{m} T^{-i}\xi_n$. By Lemma 1 and the definition of $\mathfrak{N}$, $\mu(\partial A) = 0$ holds for all $\mu \in \mathfrak{N}$. Hence, if μ_k is a sequence in $\mathfrak{N}$ converging weakly to $\mu \in \mathfrak{N}$, it follows by [5], th. 6.1, that $\mu_k(A)$ converges to $\mu(A)$ and thus that $h_{\mu_k}(\bigvee_{i=-m}^{m} T^{-i}\xi_n)$ converges to $h_\mu(\bigvee_{i=-m}^{m} T^{-i}\xi_n)$.

Therefore, the sets

$$\left\{\mu \in \mathfrak{N}: \frac{1}{2m+1} h_\mu \left(\bigvee_{i=-m}^{m} T^{-i}\xi_n \right) \geqq 1/r \right\}$$

are closed in $\mathfrak{N}$, and $\mathfrak{N}_0$ is a G_δ in $\mathfrak{N}$.

Theorem 1 now follows from Lemma 4 and Lemma 5.

For the proof of Theorem 2, we approximate p.o.-measures by Markov chains with finitely many states, as in Parthasarathy's proof that strongly mixing processes are dense (see [4], th. 1). In order to bound the entropy of the Markov chain from below, however, it is convenient to consider processes with values in a compact space.

Let $\bar{R}$ be the extended real line and $\bar{R}^Z$ the set of bilateral sequences in $\bar{R}$ with the product topology. Let T (as before) denote the shift and let $\overline{\mathfrak{M}}$ be the space of T-invariant probability measures on $\bar{R}^Z$ with the weak topology.

LEMMA 6. *Let there be given a $\mu \in \overline{\mathfrak{M}}$ such that $\mu(R^Z) = 1$ and an $s_0 > 0$. Then μ can be approximated by p.o.-measures $\mu_x \in \overline{\mathfrak{M}}$ where $x = (x_i)$ is in R^Z, x has period $s \geqq s_0$ and $x_i \neq x_j$ if $i \neq j$, $1 \leqq i,j \leqq s$.*

This lemma can be proved by another easy modification of Oxtoby's proof of th. 2 and remark 3.4 in [3].

LEMMA 7. *Let $\mu \in \overline{\mathfrak{M}}$ with $\mu(R^Z) = 1$ and $K > 0$ be given. Then every neighborhood V of μ contains a $\rho \in \overline{\mathfrak{M}}$ with $\rho(R^Z) = 1$ and $h_\rho(T) \geqq K$.*

PROOF. One can assume that V is of the form

$$V_\mu(f_1 \cdots, f_r; \delta) = \{\nu \in \overline{\mathfrak{M}}: \left| \int f_j d\mu - \int f_j d\nu \right| < \delta \text{ for } j = 1, \cdots, r\}$$

where $\delta > 0$ and the f_j are in $C(\bar{R}^Z)$ (the set of continuous real valued functions on $\bar{R}^Z$ with sup norm). One can further assume that there is an N such that, for $j = 1, \cdots, r$, one has $f_j(x) = f_j(y)$ if $x_i = y_i$ for $-N \leqq i \leqq N$. Indeed, since $\bar{R}$ is compact, functions of this kind are dense in $C(\bar{R}^Z)$.

Let M be an upper bound for $\{|f_j(x)|: x \in \bar{R}^Z, j = 1, \cdots, r\}$. Choose ε such that

$$\varepsilon < (4M)^{-1} 2^{-2N} \delta$$

and

$$1 - (1 - \varepsilon)^{2N+1} < (4M)^{-1} \delta.$$

Choose $s_0 > 0$ such that

$$\varepsilon \log(s_0 - 1) - \varepsilon \log \varepsilon - (1 - \varepsilon) \log(1 - \varepsilon) \geqq K.$$

By Lemma 6, there is an $x = (x_i) \in R^Z$ with period $s \geqq s_0$ such that the $x_1, \cdots, x_s$, are different from each other and $\mu_x \in V_\mu(f_1, \cdots f_r; \delta/2)$.

Let ρ be the measure on $\bar{R}^Z$ given by the Markov chain whose states are $x_1, \ldots, x_s$ whose initial probabilities are given by the s-tuple $(1/s, \ldots, 1/s)$ and whose transition probabilities are given by the $s \times s$-matrix (p_{ij}) where

$$p_{s,1} = 1 - \varepsilon$$

$$p_{i\,i+1} = 1 - \varepsilon \text{ for } 1 \leqq i < s$$

$$p_{i\,j} = \frac{\varepsilon}{s-1} \text{ otherwise.}$$

The entropy for a stochastic process given by a Markov chain with initial probabilities π_i $(1 \leqq i \leqq s)$ and transition matrix p_{ij} $(1 \leqq i,j \leqq s)$ is given by $\sum_{i,j} \pi_i \, p_{ij} \cdot \log p_{ij}$. Therefore, one obtains

$$h_\rho(T) = \varepsilon \log(s-1) - \varepsilon \log \varepsilon - (1-\varepsilon)\log(1-\varepsilon) \geqq K.$$

In order to evaluate $\left| \int f_j d\mu_x - \int f_j d\rho \right|$, one remarks that the integrals depend only on the values of f_j on thin $(2N+1)$-cylinders of the form $(a_{-N}, a_{-N+1}, \cdots, a_N)$, where $a_i \in \{x_1, \cdots, x_s\}$ for $-N \leqq i \leqq N$. The s^{2N+1} cylinders of this form split into two groups P and Q. P consists of those $(2N+1)$-cylinders which contain an element of the orbit of x. P has s elements and

$$\left| \int_P f_j d\mu_x - \int_P f_j d\rho \right| \leqq M \cdot (1 - (1-\varepsilon)^{2N+1}) < \delta/4$$

The second group, Q, splits in turn into $Q_1, \cdots, Q_{2N}$, where Q_p is the set of those $(2N+1)$-cylinders for which it happens at exactly p places k that a_{k+1} is not the "natural follower" of a_k, in the sense that, if $a_k = x_l$ and $a_{k+1} = x_m$, one has not $m = l + 1 \pmod s$. Q_p has $s \binom{2N}{p} (s-1)^p$ elements. Hence

$$\left| \int_{Q_p} f_j d\rho \right| \leqq s \binom{2N}{p} (s-1)^p \cdot M \cdot \frac{1}{s} \left(\frac{\varepsilon}{s-1} \right)^p \leqq M \varepsilon \binom{2N}{p}$$

Since $\displaystyle\int_Q f_j d\mu_x = 0$, it follows that

$$\left| \int_Q f_j d\mu_x - \int_Q f_j d\rho \right| \leqq M \cdot \varepsilon \cdot 2^{2N} < \delta/4$$

Therefore, one has $\rho \in V_{\mu_x}(f_1, \cdots, f_r; \delta/2)$ and hence $\rho \in V_\mu(f_1 \cdots, f_r; \delta)$. This proves the lemma.

The proof of Theorem 2 now follows easily: if $\mu' \in \mathfrak{M}$ is given, it can be considered as an element μ of $\overline{\mathfrak{M}}$ satisfying $\mu(R^Z)=1$. By the preceding lemma, there is a sequence ρ_n in $\overline{\mathfrak{M}}$ converging weakly to μ, with $\rho_n(R^Z) = 1$ and $h_{\rho_n}(T) \geqq K$ for $n = 1,2\cdots$. A standard argument shows that the measures ρ_n' induced by ρ_n on R^Z converge weakly to μ' in $\mathfrak{M}$ (see, for example, [3] prop. 6.1).

2. On the topological entropy of subshifts

Let $\mathscr{S}$ be a finite set with discrete topology, $\mathscr{S}^Z$ the set of bilateral sequences in $\mathscr{S}$ with the product topology, $T: \mathscr{S}^Z \to \mathscr{S}^Z$ the shift map. As metric for the compact space $\mathscr{S}^Z$ one can choose

$$d(x,y) = \sum_{k \in Z} 2^{-|k|}\delta(x_k, y_k)$$

where δ is the Kronecker function on $\mathscr{S}$.

The Hausdorff metric $\hat{d}$ on the space of closed subsets of the metric space $(\mathscr{S}^Z, d)$ is given by $\hat{d}(A,B) = \max[\max_{x \in A}\min_{y \in B} d(x,y), \max_{x \in B}\min_{y \in A} d(x,y)]$ for A, B closed in $\mathscr{S}^Z$. By a well-known theorem of Hausdorff, the compactness of the space of closed subsets of $\mathscr{S}^Z$ follows from the compactness of $\mathscr{S}^Y$.

Let $\mathscr{L}(\mathscr{S})$ denote the space of all subshifts of $\mathscr{S}^Z$, i.e., the space of all closed shift invariant subsets. Under the topology induced by the Hausdorff metric, $\mathscr{L}(\mathscr{S})$ becomes a compact metric space.

To the element Λ of $\mathscr{L}(\mathscr{S})$ corresponds the dynamical system consisting of the compact set Λ and the homeomorphism obtained by restricting T to Λ. A measure for the size of Λ is therefore given by the topological entropy $h(\Lambda)$ as defined in [1]. Let $H_m(\Lambda)$ denote the number of different m-blocks occuring in Λ. (The m-block $(b_1, \cdots, b_m)$ occurs in Λ if there is an $x=(x_k)$ in Λ with $x_i = b_i$, $i = 1, \cdots, m$). Since $H_{m+n}(\Lambda) \leqq H_m(\Lambda) \cdot H_n(\Lambda)$, one easily sees that $\lim 1/n \log H_m(\Lambda)$ always exists and is equal to $\inf 1/n \log H_n(\Lambda)$. It is well known that this limit is just $h(\Lambda)$. (Cf., for example, [2]).

THEOREM 3. *For a dense G_s of subshifts $\Lambda \in \mathscr{L}(\mathscr{S})$, one has $h(\Lambda) = 0$.*

For the proof, one needs a characterization of ε-balls in $\mathscr{L}(\mathscr{S})$. For given $\Lambda_0 \in \mathscr{L}(\mathscr{S})$ and $N = 1,2,\cdots$, let $K(N, \Lambda_0)$ denote the set of all $\Lambda \in \mathscr{L}(\mathscr{S})$ which have exactly the same N-blocks as Λ_0.

LEMMA 8. *The $K(N, \Lambda_0)$ form a basis for the neighborhoods of $\Lambda_0 \in \mathscr{L}(\mathscr{S})$.*

PROOF. If Λ is in $K(2N+1, \Lambda_0)$, one can find by the shift invariance of Λ and Λ_0 for each $x \in \Lambda_0$ a $y \in \Lambda$ with the same centered $(2N+1)$-block. Hence, $x_i = y_i$

for $-N \leqq i \leqq N$, and therefore $d(x, y) \leqq 2^{-N}$. Similarly for each $y \in \Lambda$, there is an $x \in \Lambda_0$ with $d(x, y) \leqq 2^{-N}$. Hence

$$K(2N+1, \Lambda_0) = \{\Lambda \in \mathscr{L}(\mathscr{S}) \colon \bar{d}(\Lambda, \Lambda_0) \leqq 2^{-N}\}$$

LEMMA 9. *The function* $h\colon \mathscr{L}(\mathscr{S}) \to R^+$ *is upper semicontinuous.*

PROOF. Let $a > h(\Lambda_0)$ be given. There exists an N such that $1/(2N+1) \times \log H_{2N+1}(\Lambda_0) < a$. If $\bar{d}(\Lambda, \Lambda_0) \leqq 2^{-N}$, then Λ has the same $(2N+1)$-blocks as Λ_0 and hence $H_{2N+1}(\Lambda_0) = H_{2N+1}(\Lambda)$. Since $h(\Lambda) = \inf 1/n \log H_n(\Lambda)$, it follows that $h(\Lambda) < a$.

LEMMA 10. *The set of* $\Lambda \in \mathscr{L}(\mathscr{S})$ *with* $h(\Lambda) = 0$ *is a* G_δ.

PROOF. By the upper semicontinuity of h, the set $\{\Lambda\colon h(\Lambda) \geqq 1/k\}$ is closed. Thus the set of all subshifts with positive entropy, which can be written

$$\bigcup_{k=1}^{\infty} \{\Lambda \in \mathscr{L}(\mathscr{S})\colon h(\Lambda) \geqq 1/k\}$$

is an F_σ.

LEMMA 11. *The set of* $\Lambda \in \mathscr{L}(\mathscr{S})$ *with* $h(\Lambda) = 0$ *is dense in* $\mathscr{L}(\mathscr{S})$.

PROOF. Let $\varepsilon > 0$ and $\Lambda_0 \in \mathscr{L}(\mathscr{S})$ be given. Choose an integer N such that $2^{-N} < \varepsilon$ and let $\mathscr{A} = \{A_1, \cdots, A_l\}$ be the collection of all $(2N+1)$-blocks occuring in Λ_0. The block $A_j = (a_1^{(j)}, \cdots, a_{2N+1}^{(j)})$ is called a successor of the block $A_i = (a_1^{(i)}, \cdots, a_{2N+1}^{(i)})$ if $a_k^{(j)} = a_{k+1}^{(i)}$ for $2 \leqq k \leqq 2N$; A_i is called predecessor of A_j. Clearly each $A_i \in \mathscr{A}$ has at least one successor and one predecessor in $\mathscr{A}$.

Starting with $A = (a_1, .., a_{2N+1}) \in \mathscr{A}$, one can build a bilateral sequence $b(A)$ with elements in $\mathscr{S}$ as follows. Let $A_{r_1} \in \mathscr{A}$ be a successor of A, $A_{r_2} \in \mathscr{A}$ a successor of A_{r_1} etc. There is an $h \leqq l$ such that $A_{r_j} = A_{r_h}$ for some $j \leqq h$. The right extension of A will consist of the block

$$a_1\, a_2 \cdots a_{2N+1}\, a_{2N+1}^{(r_1)}\, a_{2N+1}^{(r_2)} \cdots a_{2N+1}^{(r_{j-1})}$$

followed by repeated juxtapositions of the block $a_{2N+1}^{(r_j)}\, a_{2N+1}^{(r_{j+1})} \cdots a_{2N+1}^{(r_{h-1})}$. The left extension of A is obtained by concatenating predecessors in a similar way.

All $(2N+1)$-blocks of the resulting sequence $b(A)$ are elements of $\mathscr{A}$. Since both tail ends of $b(A)$ are periodic, one easily sees that the subshift $\mathscr{O}(A)$ consisting of the closure of the orbit of $b(A)$ has topological entropy zero. Therefore, the subshift $\bar{\Lambda} = \bigcup_{A \in \mathscr{A}} \mathscr{O}(\mathscr{A})$ has entropy zero, and the $(2N+1)$-blocks of $\bar{\Lambda}$ are exactly those of $\mathscr{A}$, i.e., those of Λ_0. Hence $\bar{d}(\bar{\Lambda}, \Lambda_0) \leqq 2^{-N} < \varepsilon$, and the lemma is proved.

Theorem 3 follows from Lemma 10 and Lemma 11.

REMARK. The set of $\Lambda \in \mathscr{L}(\mathscr{S})$ with $h(\Lambda) > 0$ is not dense in $\mathscr{L}(\mathscr{S})$. Indeed, the subshift Λ_0 given by the orbit of a periodic point in $\mathscr{S}^Z$ with minimal period p has entropy 0, and cannot be approximated since $K(p, \Lambda_0)$ consists of Λ_0 alone. However, one can show that the subshifts with positive topological entropy are dense in the set $\overline{\mathscr{L}}(\mathscr{S}) \subset \mathscr{L}(\mathscr{S})$ consisting of all non-isolated points.

REMARK. Theorem 3 can be considered as a topological analogue of Theorem 1, with topological entropy replacing metric entropy. It seems doubtful whether there exists for topological entropy a counterpart of Rohlin's theorem, which says that metric automorphisms have generically entropy zero (and automorphisms with positive entropy are dense) in the set of measure preserving transformations with the weak topology (see [6]). Indeed, the space of homeomorphisms of the torus (with the C^0 topology) contains an open set of transformations with positive entropy. (This is an immediate consequence of [9]). On the other hand, all homeomorphisms of S^1 have topological entropy zero (see [1]).

REFERENCES

1. R. Adler, A. Konheim and M. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc. **114,** 309–319 (1965).

2. W. Parry, *Intrinsic Markov chains*, Trans. Amer. Math. Soc. **112,** 55–66 (1964).

3. J. Oxtoby, *On two theorems of Parthasarathy and Kakutani concerning the shift transformation*, Proceedings of the International Symposium on Ergodic Theory, p. 203–215, New York: Academic Press, 1963.

4. K. R. Parthasarathy, *A note on mixing processes*, Sankhya, Ser. A, **24,** 331–332 (1962).

5. K. R. Parthasarathy, *Probability Measures on Metric Spaces*, London-New York, Academic Press, 1967.

6. V. A. Rohlin, *On the entropy of a metric automorphism*, Dokl. Akad. Nauk. **128,** 980–983 (1959).

7. V. A. Rohlin, *Lectures on entropy theory*, Uspehi Mat. Nauk. **22,** 3–56 (1967).

8. K. Sigmund, *Generic properties for invariant measures of axiom A diffeomorphisms*, Invent. Math. **10,** 99–110 (1970).

9. P. Walters, *Anosov diffeomorphisms are topologically stable*, Topology **9,** 71–78 (1970).

HEBREW UNIVERSITY OF JERUSALEM
AND
UNIVERSITY OF VIENNA