Entropy and the Complexity of Trajectories of a Dynamical System

A. A. Brudno

Dedicated to the memory of
V. M. Alekseev

Trudy Moskovskogo Matematicheskogo Obshchestva, 1982, Vol. 44. UDC 513.88

Contents

Introduction

1. Statement of the problem

The classical theory of dynamical systems (DS) has usually been concerned with *flows*, that is, systems with continuous time. However, since Poincaré it has been well known that it is generally sufficient to confine oneself to the study of systems with discrete time, i.e. not flows but *mappings*. Accordingly, by a DS (X, T) in this paper we mean one of two objects:

1. A continuous mapping $T\colon X \to X$ of a compact topological space X into itself (a *topological* DS).

2. A single-valued mapping $T\colon X \to X$ of a space X equipped with a σ-algebra of measurable sets, measurable with respect to that σ-algebra (a *metric* DS).

Interest in the problems of the theory of dynamical systems, especially heightened in the last 10–20 years, has been stimulated by the fact that the results and methods of this theory find direct application in problems of classical mechanics, statistical physics, and in a number of applied disciplines (in particular, in information theory). At present the consideration of various characteristics of a dynamical system associated with the concept of *entropy* introduced by A. N. Kolmogorov [10] is especially popular. This entropy is called *metric*; the authors of [18] introduced the notion of *topological* entropy.

Many facts of the entropy aspect of the theory of dynamical systems allow one to regard the entropy of a dynamical system as a certain measure of its complexity.

On the other hand, every dynamical system consists of individual trajectories of generally different nature and complexity. For example, trajectories may be periodic and wandering, recurrent, regular, and so on. It is intuitively clear that the global complexity of the system (i.e. its entropy) should be related in some way to the complexity of its individual trajectories.

In order to give precise meaning to this assertion, one must first of all give a mathematically rigorous definition of the notion of "complexity of a trajectory".

In the present paper[1] such a definition is based, on the one hand, on the ideas of the method of *symbolic dynamics* (a classical method going back to Hadamard and Morse), and, on the other hand, makes use of the concept of *algorithmic complexity of a finite object* introduced by A. N. Kolmogorov in his paper [10] devoted to a new foundation of probability theory and information theory.

Using a combination of these two notions, one associates with every trajectory of a topological dynamical system a nonnegative real number $K(x,T)$ characterizing the degree of complexity of the behavior of this trajectory in the phase space of the system.

2. The main results of the paper

The main results of the paper consist in establishing connections between the introduced notion of complexity of a trajectory and such traditional notions of the theory of dynamical systems as metric and topological entropy, decomposition into ergodic sets, and others. In particular, the following theorems are proved in the paper:

Theorem 3.1. *For every ergodic measure* μ*,*

$$K(x,T) = h_\mu(T)$$

for μ*-almost all* $x \in X$*.*

Theorem 3.2. *For every point* $x \in X$*,*

$$K(x,T) \le h_T(x).$$

Here $h_\mu(T)$ is the metric entropy (with respect to the measure μ) of the DS (X,T), and $h_T(x)$ is the entropy of a point in the sense of Kamae [24] (that is, $h_T(x) = \sup_{\mu \in V_T(x)} h_\mu(T)$, where $V_T(x)$ is the set of individual measures corresponding to the point x in the framework of the Krylov–Bogolyubov theory).

In particular, by Theorem 3.2 the complexity of any trajectory does not exceed the topological entropy of the system.

3. Historical survey

We briefly dwell on results bearing in some degree on our subject. Apparently the first problem of this kind arose in information theory, and the first result in this direction was Shannon's result on the complexity of coding messages of a discrete ergodic source. However, Shannon's results, like several others, were essentially incomplete, primarily because of the absence of a mathematically rigorous definition of the notion of "complexity of a message". Such a definition was given, as already mentioned, in the language of algorithm theory by A. N. Kolmogorov in 1965. The survey paper [7] gives many examples showing the fruitfulness of considering

[1] A brief presentation of some results of the present paper in somewhat different notation is contained in the note [5].

Kolmogorov-type complexity characteristics for probability theory and information theory. In addition, that paper contains several assertions relating entropy and complexity characteristics within those disciplines. In particular, the method of proof of Theorem 5.1 of [7], due to A. N. Kolmogorov, is essentially used in our paper, and an assertion of L. A. Levin announced in [7] served as one of the stimuli for the appearance of the present paper, and can be interpreted as a special case of the situation occurring for metric DS of general type considered in [5].

As for the theory of dynamical systems itself, in it the study of various characteristics and properties of individual trajectories has always been extremely popular. It suffices to mention such classical theorems as the Poincaré–Carathéodory theorem on recurrence, the Birkhoff–Khinchin ergodic theorem, the Shannon–McMillan–Breiman theorem, and others. By now a certain number of characteristics of an individual trajectory have accumulated, connected in one way or another with the complexity of behavior of that trajectory in the phase space of the system. A comparative analysis of these characteristics is given in §4 of the present paper, where, in particular, the question of the connection of complexity of a trajectory with the notion of randomness in the sense of Martin-Löf, generally accepted at present, is examined.

In conclusion I consider it my duty to express my deep gratitude to V. M. Alekseev under whose direction this work was carried out. I am also grateful to E. I. Dinaburg, Ya. G. Sinai, A. M. Stëpin, A. T. Tagi-Zade and M. V. Yakobson for useful remarks.

§0. Basic concepts and facts used in the paper

1

Both in the metric and in the topological case, the space X is assumed to be equipped with a σ-algebra $\mathfrak{M}$ consisting of measurable (respectively Borel) sets, invariant under the transformation T. In the set $M(X)$ of all normalized measures defined on the σ-algebra $\mathfrak{M}$ we denote by $M(X,T)$ and $EM(X,T)$ the subsets formed by the invariant and the ergodic measures of the DS (X,T). Following [12], we shall call a set $A \in \mathfrak{M}$ a *set of invariant measure zero* if $\mu(A) = 0$ for every $\mu \in M(X,T)$.

A quadruple $(X, \mathfrak{M}, T, \mu)$ with $\mu \in M(X,T)$ defines a dynamical system in the sense of ergodic theory.

At present there is a known number of survey papers devoted to the entropy aspect of the theory of dynamical systems; the terminology and notation are well established. We therefore find it possible to omit the definitions and description of the basic properties of the familiar entropy characteristics (we follow [20], without specific references), and to touch only briefly on the method of symbolic dynamics. The main attention will be devoted to questions related to the consideration of the notion of algorithmic complexity in the sense of Kolmogorov, relatively recently introduced and not traditional for the theory of DS.

Let $\mathfrak{A} = \{A_i\}_{i\in\mathscr{L}}$ be a covering of the space X, and

$$\Lambda_{\mathscr{L}} = \mathscr{L}^{\mathbb{Z}^+} = \{\omega = \{\omega(i)\}_{i\in\mathbb{Z}^+} : \ \omega(i) \in \mathscr{L} \text{ for all } i \in \mathbb{Z}^+\}.$$

The shift transformation $\sigma \colon \Lambda_{\mathscr{L}} \to \Lambda_{\mathscr{L}}$ is defined by

$$\widetilde{\omega} = \sigma(\omega) \iff \widetilde{\omega}(i) = \omega(i+1) \text{ for all } i \in \mathbb{Z}^+.$$

The mapping $\varphi_{\mathfrak{A}} \colon X \to \Lambda_{\mathscr{L}}$ is constructed as follows:

$$\omega = \{\omega(i)\}_{i\in\mathbb{Z}^+} \in \varphi_{\mathfrak{A}}(x) \iff T^i x \in A_{\omega(i)} \text{ for all } i \in \mathbb{Z}^+.$$

The method of symbolic dynamics in the above or somewhat different form has been used by many authors; in particular it is set out in detail in the monograph of V. M. Alekseev [1]. In particular, $\sigma(\Omega_{\mathfrak{A}}) \subseteq \Omega_{\mathfrak{A}}$, where $\Omega_{\mathfrak{A}}$ is the image of the space X under the mapping $\varphi_{\mathfrak{A}}$, i.e.

$$\Omega_{\mathfrak{A}} = \Big\{\omega \in \Lambda_{\mathscr{L}} : \bigcap_{i\in\mathbb{Z}^+} T^{-i}A_{\omega(i)} \neq \varnothing\Big\}.$$

The DS $(\Omega_{\mathfrak{A}}, \sigma)$ is called the *symbolic model* of the DS (X,T) corresponding to the cover $\mathfrak{A}$. A pair (Ω, σ), where Ω is an arbitrary invariant Borel[2] subset of $\Lambda_{\mathscr{L}}$, is called a *symbolic dynamical system*. In what follows we shall consider not only elements $\omega \in \Omega$ but also their finite segments, putting by definition:

$$\begin{aligned} \omega^{\langle K,N\rangle} &= \{\omega(i)\}_{i=K}^{N-1} = \omega(K)\omega(K+1)\ldots\omega(N-1);\\ \omega^N &= \omega^{\langle 0,N\rangle};\\ \Omega^N &= \{\omega^N :\ \omega \in \Omega\}. \end{aligned}$$

In addition, we denote by

$$C(\omega^{\langle K,N\rangle}) = \{\widetilde{\omega} = \{\widetilde{\omega}(i)\}_{i\in\mathbb{Z}^+} :\ \widetilde{\omega}(i) = \omega(i) \text{ for } K \le i \le N-1\}$$

the *cylinders* in the space $\Lambda_{\mathscr{L}}$.

If a partition[3] $\beta = \{B_i\}_{i\in\mathscr{L}}$ is measurable, then the additive function $\widetilde{\mu}$ defined on the family of cylinders

$$\{C(\omega^{\langle K,N\rangle}) :\ \omega \in \Lambda_{\mathscr{L}};\ K,N \in \mathbb{Z}^+,\ K<N\}$$

by

$$\widetilde{\mu}(C(\omega^{\langle K,N\rangle})) = \mu\Big(\bigcap_{i=K}^{N-1} T^{-i}A_{\omega(i)}\Big)$$

can be extended (by the classical Kolmogorov theorem) to a measure $\widetilde{\mu}_\beta$ on the space $\Lambda_{\mathscr{L}}$. Moreover, if $\mu \in M(X,T)$ ($\mu \in EM(X,T)$), then $\widetilde{\mu}_\beta \in M(\Lambda_{\mathscr{L}}, \sigma)$ ($\widetilde{\mu}_\beta \in EM(\Lambda_{\mathscr{L}}, \sigma)$).

2

We now turn to the description of the notion of *algorithmic complexity* of a finite object. Let A be a computable function (algorithm) defined on some subset of the space of all finite words (programs) over the alphabet $\{0,1\}$ and taking values in the set of finite words over some finite alphabet $\mathscr{L}$. We denote by $l(p)$ the length of a word p, i.e. the number of letters in it. The complexity $K_A(s)$ of a word s with respect to the function A is defined by

$$K_A(s) = \begin{cases} \inf_{p\colon A(p)=s} l(p), & \\ \infty, & \text{if } A(p) \neq s \text{ for all } p. \end{cases}$$

In [10] it is shown that there exists a computable function B such that for any other computable function A and any word s,

$$K_B(s) \le K_A(s) + C_A,$$

where C_A is a constant depending on A (but not on s).

Such a function B will be called *asymptotically optimal* (a. opt. f.). In the same paper it is proposed to fix one such function and denote complexity with respect to it simply by $K(s)$.

In the present paper an analogous characteristic is considered for words of infinite length (sequences).

[2]It is assumed that $\Lambda_{\mathscr{L}}$ is equipped with the Tikhonov topology.

[3]A cover $\mathfrak{A} = \{A_i\}_{i\in\mathscr{L}}$ is called a *partition* if $A_i \cap A_j = \varnothing$ for all $i \neq j$.

Definition 0.1. The *complexity* of a sequence $\omega = \{\omega(i)\}_{i\in\mathbb{Z}^+}$ is the number

$$K(\omega) = \limsup_{n\to\infty} \frac{K(\omega^n)}{n}.$$

The following simple but technically important statement follows directly from this definition.

Proposition 0.2. *For any computable function A and any infinite word ω,*

$$K(\omega) \le K_A(\omega);$$

moreover, if A is an a. opt. f., then equality holds.

In what follows we shall often use the natural correspondence between words over the alphabet $\{0,1\}$ and integers that arises from the binary representation of the latter. Thus, for example,

$$0 \leftrightarrow 0, \quad 1 \leftrightarrow 1, \quad 2 \leftrightarrow 10, \quad 3 \leftrightarrow 11, \ \ldots$$

If $p = p(1)p(2)\ldots p(k)$ and $q = q(1)q(2)\ldots q(l)$ are words, then by pq and $\bar{p}$ we denote respectively the words

$$pq = p(1)\ldots p(k)q(1)\ldots q(l),$$
$$\bar{p} = p(1)p(1)\ldots p(k)p(k)\,01.$$

The latter notation will help us in the sequel to distinguish the words p and q separately within the word $\bar{p}q$.

The following simple relations hold:

$$l(pq) = l(p) + l(q);$$
$$l(\bar{p}) = 2l(p) + 2;$$
$$\operatorname{card}\{p : l(p) = n\} = 2^n;$$
$$\operatorname{card}\{p : l(p) \le r\} \le 2^{[r]+1}.$$

Here and below, $[r]$ denotes the integer part of the number r.

§1 Symbolic dynamical systems

Throughout this section we shall consider symbolic DS (Ω, σ), with the alphabet $\mathscr{L}$ $(\Omega \subseteq \Lambda_{\mathscr{L}})$ assumed to be finite. Our main goal is the proof of the following theorem.

Theorem 1.1. *If $\mu \in EM(\Omega, \sigma)$, then*

$$K(\omega) = h_\mu(\sigma)$$

for μ-almost all $\omega \in \Omega$.

The proof of this theorem breaks up into a series of successive lemmas. The first of them is well known.

Lemma 1.2. *If $\mathscr{L} = \{1, 2, \ldots, k\}$, $A_i = \{\omega : \omega(0) = i\}$ and $\mathfrak{A} = \{A_1, A_2, \ldots, A_k\}$, then $h_\mu(\sigma \mid \Omega) = h_\mu(\sigma, \mathfrak{A} \mid \Omega)$ for any measure $\mu \in M(\Omega, \sigma)$.*

Analogously, if (Ω, σ) is a topological DS, then

$$h(\sigma \mid \Omega) = h(\sigma, \mathfrak{A} \mid \Omega).$$

Lemma 1.3. *If $\widetilde{\omega} = \sigma^k\omega$ for some $k \in \mathbb{Z}^+$, then $K(\widetilde{\omega}) = K(\omega)$.*

Proof. Let A be an arbitrary a. opt. f. Define a computable function $A' = A'(p)$ whose value on words of the form $p = \bar{p}_1\bar{p}_2p_3$ is obtained in three stages:

1. Write out the word $A(p_3)$.
2. Interpret the word p_2 as the number of symbols on the left that must be discarded in the word $A(p_3)$.
3. Prepend the word $A(p_1)$ to the left of what remains after step 2.

The resulting word is taken as $A'(p)$.

Now let a word p be such that $A(p) = \omega^{n+k}$. Put $p_1 = \varnothing$, let $p_2 = k$ stand for the number k, and $p_n = \bar{p}_1\bar{p}_2p_n$. Then $A'(\widetilde{p_n}) = \widetilde{\omega}^n$ and

$$K_{A'}(\widetilde{\omega}^n) \leq 2(l(k)+1) + 6 + K_A(\omega^{n+k}),$$

i.e.

$$K(\widetilde{\omega}) \leq K_{A'}(\widetilde{\omega}) \leq \lim_{n\to\infty}\Big(\frac{C}{n} + \frac{K_A(\omega^{n+k})}{n+k}\cdot\frac{n+k}{n}\Big) \leq K_A(\omega) = K(\omega).$$

The reverse inequality is obtained analogously, taking $\widetilde{p_n} = \bar{p}_1\bar{p}_2p_n$ with $A(p_1) = \omega^k$, $p_2 = 0$ and $A(p_n) = \widetilde{\omega}^{n-k}$. ☐

Lemma 1.4 (Corollary of 1.3)**.** *The sets $\bar{A}_t = \{\omega : K(\omega) = t\}$, $\check{A}_t = \{\omega : K(\omega) < t\}$, $\hat{A}_t = \{\omega : K(\omega) > t\}$ are invariant under the shift transformation σ.*

Lemma 1.5. *The sets $\bar{A}_t$, $\check{A}_t$, $\hat{A}_t$ are measurable.*

Proof. We show measurability of the set $\check{A}_t$ (the others are handled analogously). Indeed,

$$\check{A}_t = \{\omega : K(\omega) < t\} = \bigcup_{k=1}^{\infty}\bigcup_{N=1}^{\infty}\bigcap_{n>N}\{\omega : K(\omega^n) < (t - 1/k)n\}.$$

But for fixed k and n the set in braces is a union of a finite number of cylinders, and hence is measurable; therefore so is the set $\check{A}_t$. ☐

Lemma 1.6. *If $\mu \in EM(\Omega,\sigma)$, then $K(\omega) \geq h_\mu(\sigma)$ for μ-almost all $\omega \in \Omega$.*

Proof. Suppose the contrary. Put

$$Q = \{\omega : K(\omega) < h_\mu(\sigma)\}, \quad \mu Q \neq 0.$$

By Lemmas 1.3 and 1.4 the set Q is invariant and measurable; consequently $\mu Q = 1$ (by ergodicity of μ). We have the decomposition $Q = \bigcup_{r\in\mathbb{Z}^+} Q_r$, where

$$Q_r = \{\omega : K(\omega) < h_\mu(\sigma) - 1/r\}.$$

Since the sets Q_r are also invariant and measurable, and moreover

$$Q_1 \subseteq Q_2 \subseteq \dots \quad \text{and} \quad \bigcup_{r\in\mathbb{Z}^+} Q_r = Q,$$

there exists R such that $\mu Q_R = 1$. In turn,

$$Q_R = \bigcup_{k=1}^{\infty} Q_{R,k},$$

where $Q_{R,k} = \{\omega : K(\omega^l) < (h_\mu(\sigma) - 1/R)\, l \text{ for all } l > k\}$. Analogously there exists K such that for $k \geq K$, $\mu Q_{R,k} > 1 - \delta$ ($\delta > 0$ arbitrary).

Further, let $\varepsilon < \min(1/R, 1 - \delta)$, and let $N = N(\varepsilon)$ and the sets A_n and B_n satisfy the condition of the well-known consequence of the Shannon–McMillan–Breiman theorem (see e.g. [2]). In other words, for all $n > N$ we have $\Omega = A_n \cup B_n$, with $\mu(A_n) < \varepsilon$, and for any $\omega \in B_n$

$$2^{-n(h_\mu(\sigma)+\varepsilon)} \leq \mu(C(\omega^n)) \leq 2^{-n(h_\mu(\sigma)-\varepsilon)}.$$

Put

$$Q^A_{R,k} = Q_{R,k} \cap A_k, \quad Q^B_{R,k} = Q_{R,k} \cap B_k.$$

Since $Q^A_{R,k} \subseteq A_k$, for all $k > \max(N(\varepsilon), K)$,

$$\mu Q^A_{R,k} \leq \mu A_k < \varepsilon \quad \text{and} \quad \mu Q^B_{R,k} > 1 - \delta - \varepsilon > 0. \tag{$*$}$$

On the other hand, if $\omega \in Q^B_{R,k}$, then $K(\omega^k) \leq k(h_\mu(\sigma) - 1/R)$. Hence,

$$\operatorname{card}\{\omega^k : \omega \in Q^B_{R,k}\} \leq 2^{k(h_\mu(\sigma)-1/R)+1},$$

and

$$\mu(Q^B_{R,k}) \leq 2^{k(h_\mu(\sigma)-1/R)+1} \cdot 2^{-k(h_\mu(\sigma)-\varepsilon)} \leq 2^{k(\varepsilon-1/R)+1}.$$

Therefore $\lim_{k\to\infty} \mu(Q^B_{R,k}) = 0$, contradicting ($*$). The proof is complete. □

Lemma 1.7. *Let $\beta = \{B_1, \ldots, B_M\}$ be a finite measurable partition; $r, k \in \mathbb{Z}^+$. Put*

$$p^{r,k}_m(\omega, n) = \frac{\sum_{i=0}^{[(n-r)/k]} \chi_{B_m}(\sigma^{ik+r}\omega)}{[n/k]},$$

where χ_{B_m} is the characteristic function of the set B_m. Then for any $k \in \mathbb{Z}^+$ and any measure $\mu \in EM(\Omega, \sigma)$ the following two statements are simultaneously true for μ-almost all $\omega \in \Omega$:

(a) *For all $1 \leq m \leq M$ there exists*

$$p^{r,k}_m(\omega) = \lim_{n\to\infty} p^{r,k}_m(\omega, n).$$

(b) *There exists $r < k$ such that*

$$-\sum_{m=1}^{M} p^{r,k}_m(\omega) \log p^{r,k}_m(\omega) \leq H_\mu(\beta).$$

Proof. Indeed, since $\mu \in EM(\Omega, \sigma)$, we have $\mu \in M(\Omega, \sigma^k)$, and the validity of (a) follows directly from the Birkhoff–Khinchin ergodic theorem. From the same theorem it follows that for μ-almost all $\omega \in \Omega$,

$$p^{0,1}_m(\omega) = \mu(B_m) \quad \text{for all } 1 \leq m \leq M.$$

But then the validity of (b) for the same ω follows from the concavity of entropy and the equality

$$-\frac{1}{k} \sum_{r=0}^{k-1} p^{r,k}_m(\omega) = p^{0,1}_m(\omega).$$

□

Lemma 1.8 (obvious). *If $t_n = kn + r$ is an arithmetic progression, then*

$$K(\omega) = \limsup_{n\to\infty} \frac{K(\omega^{t_n})}{t_n}.$$

Lemma 1.9. *If $\mu \in EM(\Omega, \sigma)$, then*

$$K(\omega) \le h_\mu(\sigma)$$

for μ-almost all $\omega \in \Omega$.

Proof. The proof uses a simple modification of a construction used by Kolmogorov (see [7, Theorem 5.3]). We shall assume that all words ω^k ($\omega \in \Omega$, k arbitrary) are ordered, e.g. lexicographically. Now fix some k and represent an arbitrary word ω^n ($n = ik + r$, $r < k$) in the form

$$\omega^n = \omega_0^r \omega_1^k \omega_2^k \dots \omega_i^k.$$

Denote by $s_j = s_j(\omega^n)$ the number of times the j-th word of the set Ω^k occurs among the words ω_s^k ($1 \le s \le i$). The set of numbers $\{s_j = s_j(\omega^n)\}_{j=1}^M$, where $M = \text{card}\{\Omega^k\}$, is called the *frequency set* of the word ω^n. Moreover, we denote by $h(\omega^n)$ the logarithm of the number of all words of length n having the same frequency set as ω^n, i.e.

$$h(\omega^n) = \log \frac{i!}{s_1! \dots s_M!},$$

and by $N(\omega^n)$ the index of ω^n among these words.

Finally, it is clear that there exist $R \in \mathbb{Z}^+$ and a computable function $A_0 = A_0(p)$ such that $K_{A_0}(\omega^t) < R$ for any ω if $t < k$. Then the word ω^n is uniquely recovered from the set of numbers

$$\{k, i, s_1, \dots, s_M, r, A_0^{-1}(\omega_0^r), N(\omega^n)\},$$

or what is the same, from the binary word

$$p = \bar{k}\,\bar{i}\,\bar{s}_1 \dots \bar{s}_M\,\bar{r}\,\overline{A_0^{-1}(\omega_0^r)}\,\log N(\omega^n).$$

We shall call the computable function A^* effecting this encoding of the word ω^n into the word p the *frequency-encoding function.* Thus,

$$K_{A^*}(\omega^n) \le 2\big(l(k) + l(i) + \sum_{m=1}^{M} l(s_m) + l(r) + R\big) + h(\omega^n),$$

or, since

$$i = \left[\frac{n}{k}\right] \quad \text{and} \quad l(s_m) \le l(i) \le [\log i] + 1$$

for all m, we have

$$K_{A^*}(\omega^n) \le h(\omega^n) + o(n).$$

Applying Stirling's formula

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1 + o(1)),$$

we get

$$h(\omega^n) \le -i \sum_{m=1}^{M} \frac{s_m(\omega^n)}{i} \log\Big(\frac{s_m(\omega^n)}{i}\Big);$$

in the notation of Lemma 1.7,

$$h(\omega^n) \le -i \sum_{m=1}^{M} p_m^{r,k}(\omega, n) \log p_m^{r,k}(\omega, n).$$

But then, by Lemmas 1.7 and 1.8, for μ-almost all ω,

$$K(\omega) \le K_{A^*}(\omega) \le \frac{H_\mu(\beta_k)}{k}, \tag{1}$$

where $\beta_k = \{B_1, \dots, B_M\}$ is the partition into all cylinders of length k, i.e. $\beta_k = \{C(\omega^k)\}_{\omega \in \Omega}$. The validity of Lemma 1.9 now easily follows from the arbitrariness of k in inequality (1) and Lemma 1.2. The proof is complete. □

Theorem 1.1 obviously follows from Lemmas 1.6 and 1.9.

The statement of Theorem 1.1, besides its independent interest, in combination with known facts of the entropy theory of DS, allows us to answer a number of questions that remain open in the purely topological treatment of symbolic DS.

Proposition 1.10. *We have earlier shown (see Theorem 5.4 of [1]) that for any* $\omega \in \Omega$,

$$K(\omega) \le H(\omega) \le h(\sigma \mid \Omega),$$

where $H(\omega)$ *is the combinatorial entropy of* ω. *If now* μ_M *is the measure of maximal entropy (such a measure always exists for a symbolic DS), then for* μ_M*-almost all* $\omega \in \Omega$,

$$K(\omega) = H(\omega) = h(\sigma \mid \Omega).$$

Proposition 1.11. *There exist minimal DS for which*

$$K(\omega) \not\equiv \text{const}.$$

The validity of this assertion follows from the existence of minimal DS (for which $H(\omega) \equiv h(\sigma \mid \Omega)$) admitting invariant measures with different entropies (see [20]) and the assertion of Theorem 1.1.

Further development of these questions can be found in §4 of this article.

It seems interesting to consider the question of a *constructive description* of sequences whose complexity does not exceed the entropy. An answer to this question within the framework of restrictions adopted in information theory is given in the following assertion. Note that its proof is a "translation" into the complexity language of the proof of Shannon's classical result (see [15, Theorem 4]).

Proposition 1.12. *If* $\mu \in EM(\Omega, \sigma)$ *is a computable measure (see the definition in [7]) and* ω_0 *"satisfies" the Shannon–McMillan–Breiman theorem, that is,*

$$\lim_{n\to\infty} -(1/n) \log \mu(C(\omega_0^n)) = h_\mu(\sigma),$$

then $K(\omega_0) \le h_\mu(\sigma)$.

Proof. Let the function $A = A(p)$ be defined on words $p = \bar{p}_1 p_2$ so that if $p_1 = n$, $p_2 = k$, then $A(p) = \omega^n$ if and only if the cylinder $C(\omega^n)$ has index k among all cylinders $C(\omega^n)$ ordered (by decreasing magnitude) according to their measures. Note that computability of A follows from computability of the measure μ. Suppose $K(\omega_0) > h' > h_\mu(\sigma)$. Then, since $K_A(\omega_0) \ge K(\omega_0)$, there exists a sequence $\{n_i\}_{i\in\mathbb{Z}^+}$ such that for all n_i

$$K_A(\omega_0^{n_i}) > n_i h'.$$

Since

$$K_A(\omega_0^{n_i}) \le 2\log n_i + \log(N_\mu(\omega_0^{n_i})),$$

where $N_\mu(\omega_0^{n_i})$ is the index of the cylinder $C(\omega_0^{n_i})$ according to its measure, we have

$$\lim_{i\to\infty} \frac{\log(N_\mu(\omega_0^{n_i}))}{n_i} > h'.$$

Set

$$\Phi(\omega_0^{n_i}) = \{\omega^{n_i} : N_\mu(\omega^{n_i}) < N_\mu(\omega_0^{n_i})\}$$

and put

$$S_{n_i} = \sum_{\omega^{n_i}\in\Phi(\omega_0^{n_i})} \mu(C(\omega^{n_i})).$$

In view of the normalization of μ,

$$\mu(S_{n_i}) \le 1 \quad \text{for all } i \in \mathbb{Z}^+.$$

On the other hand,

$$\lim_{i\to\infty} \frac{\log S_{n_i}}{n_i} \ge \lim_{i\to\infty} \frac{\log([2^{n_i h'}]\cdot \mu(C(\omega_0^{n_i})))}{n_i} \ge h' - h_\mu(\sigma) > 0.$$

The proof of Proposition 1.12 is complete. □

§2 Properties of the quantity $K(x, T \mid \mathfrak{A})$ — the complexity of the trajectory of a point x relative to a cover $\mathfrak{A}$

In this section, for an arbitrary DS, the notion of *complexity of a trajectory relative to a cover* is introduced and studied. We note that the results of this section, besides their auxiliary character (they are essentially used in §3 — the main section of the present work), are also of independent interest, allowing one, in particular, to estimate the complexity of the classical Sturmian trajectories of Hedlund–Morse.

Definition 2.1. Let $\mathfrak{A} = \{A_1, \dots, A_k\}$ be a finite cover of X. The quantity

$$K(x, T \mid \mathfrak{A}) = \limsup_{n\to\infty} (1/n) \min_{\omega\in\varphi_{\mathfrak{A}}(x)} K(\omega^n)$$

will be called the *complexity of the trajectory of the point $x \in X$ relative to the cover $\mathfrak{A}$*.

The quantity $K(x, T \mid \mathfrak{A})$ possesses the following entropy-like properties:

Proposition 2.2. $K(x, T \mid \mathfrak{A}_1) \le K(x, T \mid \mathfrak{A}_2)$ *if* $\mathfrak{A}_2 \ge \mathfrak{A}_1$.

Proposition 2.3. $K(x, T \mid \mathfrak{A}_1 \vee \mathfrak{A}_2) \le K(x, T \mid \mathfrak{A}_1) + K(x, T \mid \mathfrak{A}_2)$.

Proposition 2.4. $K(x, T \mid \mathfrak{A}^k) = K(x, T \mid \mathfrak{A})$ *for any* $k \in \mathbb{Z}^+$.

We prove Property 2.2 (the proofs of the other properties are analogous).

Let $\mathfrak{A}_1 = \{A_1, \dots, A_k\}$, $\mathfrak{A}_2 = \{A'_1, \dots, A'_l\}$ and $\mathfrak{A}_1 \le \mathfrak{A}_2$.

Clearly there exists a computable function $A_0\colon \{1, \dots, l\} \to \{1, \dots, k\}$ such that if $A_0(j) = i$, then $A'_j \subseteq A_i$. If now A is an arbitrary a. opt. f., then $\widetilde{A} = A_0 \circ A$ is computable and, moreover, if $A(p) \in \{(\varphi_{\mathfrak{A}_2}(x))^n\}$, then $\widetilde{A}(p) \in \{(\varphi_{\mathfrak{A}_1}(x))^n\}$. From this, by the arbitrariness of n, we obtain Property 2.2.

The following simple but extremely useful statement holds.

Proposition 2.5. $K(x, T \mid \mathfrak{A}) \leq \inf_{\omega \in \varphi_{\mathfrak{A}}(x)} K(\omega)$*; in particular, if* $\mathfrak{A}$ *is a partition, then* $K(x, T \mid \mathfrak{A}) = K(\varphi_{\mathfrak{A}}(x))$.

Analogously to the fact that for almost all sequences of a symbolic DS the complexity coincided with the entropy, for the complexity of a trajectory relative to a partition we have:

Lemma 2.6. *Let* β *be a finite measurable partition and* $\mu \in EM(X, T)$*. Then for* μ*-almost all* $x \in X$,

$$K(x, T \mid \beta) = h_\mu(T, \beta).$$

Proof. As already noted in §0, if $\mu \in EM(X, T)$, then $\widetilde{\mu}_\beta \in EM(\Omega_\beta, \sigma)$; it is well known that in that case $h_\mu(T, \beta) = h_{\widetilde{\mu_\beta}}(\sigma \mid \Omega_\beta)$. But by Theorem 1.1

$$\widetilde{\mu}(\{\omega \in \Omega_\beta : K(\omega) = h_\mu(T, \beta)\}) = 1.$$

Put $Q = \varphi_\beta^{-1}(\Omega_\beta)$. For $x \in Q$,

$$K(x, T \mid \beta) = K(\varphi_\beta(x)) = h_\mu(T, \beta).$$

Hence $\mu Q = \widetilde{\mu}_\beta \Omega_\beta = 1$. □

Corollary 2.7. *If* $\mu \in EM(X, T)$ *and* β *is a generator for the measure* μ*, then for* μ*-almost all* $x \in X$,

$$K(x, T \mid \beta) = h_\mu(T);$$

and for any partition β *the set*

$$Q = \{x : K(x, T \mid \beta) > \sup_{\mu \in M(X,T)} h_\mu(T)\}$$

has invariant measure zero.

Throughout the remainder of §2 the quantity $\sup_{x \in X} K(x, T \mid \mathfrak{A})$ is investigated.

The first of the statements relating to this circle of questions looks somewhat unexpected.

Proposition 2.8. *For any non-periodic point* $x_0 \in X$ *and any number* $N \in \mathbb{Z}^+$ *there exists a finite measurable partition* β *such that*

$$K(x_0, T \mid \beta) = \log N.$$

Proof. Let $\omega_0 \in \Lambda_{N=\{0,1,\ldots,N-1\}}$ and $K(\omega_0) = \log N$ (such sequences form a set of full measure by Theorem 1.1). For each i $(0 \leq i \leq N-1)$ form the set

$$B_i = \{x \in X : \text{ there exists } k \in \mathbb{Z}^+ \text{ such that } T^k x_0 = x \text{ and } \omega_0(k) = i\}$$

and put

$$B_N = X \setminus \bigcup_{i=0}^{N-1} B_i.$$

It is easy to see that the partition $\beta = \{B_1, \ldots, B_N\}$ is measurable and $\varphi_\beta(x_0) = \omega_0$, i.e. $K(x_0, T \mid \beta) = \log N$. □

Note that it is precisely Proposition 2.8 that prevents one from introducing the concept $K(x, T)$ — the simple (without a cover) complexity of a trajectory — in the traditional entropy-theoretic manner, by putting

$$K(x, T) = \sup K(x, T \mid \beta),$$

where sup is taken over all finite measurable partitions of X.

Another, also entropy-theoretic, approach to introducing the quantity $K(x, T)$ for metric DS is presented in [5] (there this quantity is denoted by $K^0(x, T)$).

In what follows we shall assume that (X, T) is a topological dynamical system.

Proposition 2.9. *If $u = \{U_1, \ldots, U_l\}$ is a finite open cover, then for all $x \in X$,*

$$K(x, T \mid u) \leq h(T, u).$$

Proof. Indeed, let $h' > h(T, u)$; then there exist $N \in \mathbb{Z}^+$ and a cover $\widetilde{u}$ (a subcover of the cover u^N) such that $\log \operatorname{card}\{\widetilde{u}\} < h' \cdot N$. Since $\widetilde{u}$ is a cover, there exists $\omega \in \varphi_u(x)$ such that for all $k \in \mathbb{Z}^+$,

$$U_{\omega(kN)} \cap T^{-1} U_{\omega(kN+1)} \cap \ldots \cap T^{-N+1} U_{\omega(kN+N-1)} \in \widetilde{u}.$$

Hence, by Lemma 1.8, Proposition 2.5 and the properties of complexity, we obtain

$$K(x, T \mid u) \leq \frac{1}{N} \limsup_{k \to \infty} \frac{K(\omega^{N \cdot k})}{k} \leq \frac{\log \operatorname{card} \widetilde{u}}{N} < h'.$$

□

Returning to the case of a partition, we present two estimates for the quantity $\sup_{x \in X} K(x, T \mid \mathfrak{A})$, based respectively on a topological and a metric estimate of the massivity of the set

$$W(\mathfrak{A}) = \{x \in \bar{A}_i \cap \bar{A}_j;\ i \neq j\}$$

— the boundary of the cover $\mathfrak{A}$.

Proposition 2.10. *Let $S_{\mathfrak{A}}$ be the dynamical analogue of the index of a cover considered in [4] (there it is denoted simply S). The estimate*

$$S_{\mathfrak{A}} \geq \sup_{x \in X} K(x, T \mid \mathfrak{A}) - h(T)$$

holds.

Proposition 2.11. *If the set $W(\mathfrak{A})$ has invariant measure zero, then $K(x, T \mid \mathfrak{A}) \leq h(T)$.*

Proof. In [22] it is shown that under the hypotheses of Proposition 2.11 there exists a measure $\nu \in M(X, T)$ such that $h(\sigma \mid \Omega_{\mathfrak{A}}) = h_\nu(T, \mathfrak{A})$. But then by the Dinaburg–Goodwyn theorem and the inequality $K(\omega) \leq h(\sigma \mid \Omega)$, we have

$$K(\omega) \leq h(T) \quad \text{for all } \omega \in \Omega_{\mathfrak{A}}.$$

But if $\omega \in \Omega_{\mathfrak{A}}$, then *a fortiori* $\omega \in \Omega_{\mathfrak{A}}$, i.e.

$$K(x, T \mid \mathfrak{A}) = K(\varphi_{\mathfrak{A}}(x)) \leq h(T).$$

□

With the help of Proposition 2.11 one can estimate the complexity of the classical Sturmian trajectories of Hedlund–Morse (see [26]). Indeed, let $X = [0, 1)$ be the unit circle, and $T\colon X \to X$ the rotation by an angle α (i.e. $x \mapsto x + \alpha$; α irrational). It is well known that such a DS is strictly ergodic; its unique invariant measure m is the usual Lebesgue measure, and $h_m(T) = 0$. Further, denote by $\beta = \{B_1, \ldots, B_k\}$ the partition determined by the condition $B_i = [r_i, r_{i+1})$, where

$$0 = r_1 < r_2 < \ldots < r_{k+1} = 1.$$

The elements of the space Ω_β are called *Sturmian trajectories*. Since $m(W(\beta)) = m(\{r_0, r_1, \ldots, r_{k+1}\}) = 0$, by 2.11 we obtain

Corollary 2.12. *For all $x \in X$, $K(x, T \mid \beta) = K(\varphi_\beta(x)) = 0$, i.e. the complexity of Sturmian trajectories equals zero.*

§3 Complexity of trajectories of a topological dynamical system

In this section, with each trajectory of a topological dynamical system we shall associate a number $K(x,T)$ characterizing the degree of complexity of the behavior of this trajectory in the phase space of the system. In so doing, two main statements of the present paper will be established, namely

Theorem 3.1. *If $\mu \in EM(X,T)$, then*

$$K(x,T) = h_\mu(T)$$

for μ-almost all $x \in X$.

Theorem 3.2. *For any point $x \in X$,*

$$K(x,T) \le h_T(x),$$

where $h_T(x)$ is the entropy of the point x in the sense of Kamae [24].

1. Definition

The quantity

$$K(x,T) = \sup_u K(x,T \mid u),$$

where the sup is taken over all finite open covers of the space X, will be called the *complexity of the trajectory of the point x*.

Remark. *Although the present definition makes sense for arbitrary topological DS, most of the results of §3 are obtained under the assumption that X is a metric compactum.*

All covers considered below are assumed to be open and finite.

In my view, the following statement (not used in the sequel) is curious:

Proposition 3.3 (proof omitted). *Let the DS (X,T) and (Y,S) be topologically conjugate, with $\pi\colon X \to Y$ the conjugating homeomorphism and $\pi(x) = y$. Then*

$$K(x,T) = K(y,S).$$

From Properties 2.2, 2.4, 2.6 and well-known properties of topological DS we obtain:

Proposition 3.4. *If $\lim_{i\to\infty} d(u_i) = 0$ for a family of covers, then*

$$K(x,T) = \sup_{i\in\mathbb{Z}^+} K(x,T \mid u_i).$$

Corollary 3.5. *If a cover u is a* topological generator *of the DS (X,T), then*

$$K(x,T) = K(x,T \mid u),$$

and in particular, for a symbolic DS,

$$K(\omega,\sigma) = K(\omega).$$

Proposition 3.6. *For all $x \in X$,*

$$K(x,T) \le h(T).$$

Remark. *Sharpening Proposition 3.6, one can show that* $K(x,T) \le h(T \mid \overline{O(x)})$*, where* $\overline{O(x)}$ *is the closure of the trajectory of the point* x*. A substantially sharper estimate of the complexity of a concrete trajectory is given in Theorem 3.2.*

We also note that consideration of the simplest examples of DS not possessing a measure of maximal entropy shows that there may not exist trajectories of maximal complexity, i.e. it is not excluded that $K(x,T) < h(T)$ *for all* $x \in X$*.*

Our immediate goal is the proof of Theorem 3.1.

We present auxiliary results.

Definition 3.7. A set $A \in \mathfrak{M}$ will be called *μ-continuous* if $\mu(\nu A) = 0$, where νA denotes the boundary of A.

A cover $u = \{U_1, \dots, U_k\}$ will be called μ-continuous if all the sets U_i $(1 \le i \le k)$ are μ-continuous.

Lemma 3.8. *For any cover* u *there exists a* μ*-continuous cover* v *such that* $v \ge u$*.*

Proof. Let $\delta(u)$ be the Lebesgue number of the cover u. For each point $x \in X$ there exists a neighborhood $O(x)$ such that $d(O(x)) < \delta$ and the set $O(x)$ will be μ-continuous. For the cover $0 = \{O(x)\}_{x \in X}$ of the compactum X by such sets, there is a finite subcover v, which is the required one. □

Corollary 3.9. *From Property 2.2 and Lemma 3.8 it follows that*

$$K(x,T) = \sup K(x,T \mid u),$$

where sup *is taken only over* μ*-continuous covers* u *of the compactum* X*.*

Lemma 3.10. *For any cover* u *and any* $\varepsilon > 0$ *there exists a cover* v *such that:*

1. $v \ge u$*;*
2. v *is* μ*-continuous;*
3. $\mu(W(v)) < \varepsilon$*, where, as before,*

$$W(v) = \{x \in \bigcup_{i \ne j} \bar{V}_i \cap \bar{V}_j\}.$$

Proof. By Lemma 3.8 there exists a μ-continuous cover $v' \ge u$. We now describe an iterative procedure by which we obtain the required cover v from v'. Let $M = \operatorname{card} v'$ (i.e. $v' = \{V_1, \dots, V_M\}$) and $\varepsilon_0 < \varepsilon / M^2$. At each step of the iteration we choose a pair of sets such that $\mu(V_i^k \cap V_j^k) > \varepsilon_0$ and construct a cover $v^{k+1} = \{V_1^{k+1}, \dots, V_M^{k+1}\}$, where $V_s^{k+1} = V_s^k$ for $s \ne j$ and $V_j^{k+1} \subseteq V_j^k$, with:

1. V_j^{k+1} is μ-continuous;
2. $\mu(V_i^{k+1} \cap V_j^{k+1}) < \varepsilon_0$.

It is clear that the iterative process so defined ends in a finite number of steps, and the resulting cover v is the required one.

We describe a step of the iteration. By the μ-continuity of the sets V_i^k and V_j^k, there exists a μ-continuous open set O_{ij} such that

$$O_{ij} \supseteq \nu(V_i^k \cap V_j^k) \quad \text{and} \quad \mu(O_{ij}) < \varepsilon_0.$$

Put

$$V_j^{k+1} = (V_j^k \setminus V_i^k) \cup (O_{ij} \cap V_j^k).$$

It is easy to see that the set V_j^{k+1} is open, μ-continuous, and moreover,

$$\mu(V_j^{k+1} \cap V_i^k) \leq \mu(O_{ij}) < \varepsilon_0.$$

□

In what follows, for a cover $u = \{U_1, \ldots, U_M\}$ we denote by $\beta(u)$ the partition of the compactum X into sets B_i $(1 \leq i \leq M)$ defined by

$$\begin{aligned} B_1 &= U_1; \\ B_2 &= U_2 \setminus U_1; \\ &\vdots \\ B_M &= U_M \setminus \bigcup_{i=1}^{M-1} U_i. \end{aligned}$$

Lemma 3.11 (obvious)**.** *If a cover u is μ-continuous, then so are the covers*

$$\beta(u),\ T^{-1}(u),\ u^N,\ \beta^N(u).$$

We shall need the following standard construction from the theory of dynamical systems.

To every point $x \in X$ and number $n \in \mathbb{Z}^+$ we associate a measure $\mu_{x,n}$ defined by

$$\mu_{x,n} = \frac{1}{n}(\delta_x + \delta_{Tx} + \ldots + \delta_{T^{n-1}x}),$$

where δ_x is the measure concentrated at the point x.

It is well known (see [12, 20]) that for any $x \in X$ the set $V_T(x)$ — the limit set (in the weak topology) of the family of measures $\{\mu_{x,n}\}_{n \in \mathbb{Z}^+}$ — possesses two properties, namely:

1. $V_T(x) \neq \varnothing$;
2. $V_T(x) \subseteq M(X,T)$.

By an *ergodic decomposition* of the dynamical system (X,T) one means a representation of X in the following form:

$$X = \bigcup_{\mu \in EM(X,T)} M_\mu \cup N,$$

where

1. $x \in M_\mu \iff V_T(x) = \{\mu\}$;
2. $\mu(M_\mu) = 1$ for any $\mu \in EM(X,T)$;
3. the set N is of invariant measure zero.

In the next two lemmas the relationship between the concept of μ-continuity of a set and the decomposition of the space X into ergodic sets is established. The first of them is well known.

Lemma 3.12. *If a sequence of measures μ_i converges (in the weak topology) to a measure μ ($\mu_i \Rightarrow \mu$), then for every μ-continuous set B,*

$$\lim_{i \to \infty} \mu_i(B) = \mu(B).$$

Lemma 3.13. *Let $\mu \in EM(X,T)$, $\bar{\mu}$ its upper measure, B a μ-continuous set with $\mu(B) < \tau$. Then for every $\varepsilon > 0$ there exists an integer L such that the set*

$$P_L = \{x : \mu_{x,l}(B) < \tau \textit{ for } l > L\}$$

satisfies $\mu P_L > 1 - \varepsilon$.

Proof. Indeed, for all $x \in M_\mu$, by Lemma 3.12, $\lim_{l\to\infty} \mu_{x,l}(B) = \mu(B)$. Thus for each $x \in M_\mu$ there exists $L(x)$ such that

$$\mu_{x,l}(B) < \tau \quad \text{for } l > L(x).$$

Denote $P_k = \{x : x \in M_\mu \text{ and } L(x) < k\}$; then

$$P_1 \subseteq P_2 \subseteq \dots \quad \text{and} \quad \bigcup_{k=1}^{\infty} P_k = M_\mu.$$

Now the existence of a number L with the required properties follows from a well-known measure-theoretic theorem. □

The proof and role of Lemma 3.14 are analogous to the proof and place of Lemmas 1.4 and 1.5 in §1.

Lemma 3.14. *The sets $X_{t,u} = \{x : K(x,T \mid u) < t\}$ are invariant and measurable for all $t \in \mathbb{R}^+$.*

Lemma 3.15. *Let $\mu \in EM(X,T)$. For μ-almost all $x \in X$,*

$$K(x,T) \geq h_\mu(T).$$

Proof. Indeed, suppose there exists $h' < h_\mu(T)$ such that $\mu(X_{h'} = \{x : K(x,T) < h'\}) \neq 0$. Consequently, for any cover u and any $\varepsilon > 0$, by arguments analogous to the proof of Lemma 1.10, we obtain the existence of a number $k_0 \in \mathbb{Z}^+$ such that $\mu(Z_{k_0,h'}) > 1 - \varepsilon$, where

$$Z_{k_0,h'} = \{x : \min_{\omega^k \in \{(\psi_u(x))^k\}} K(\omega^k) < k \cdot h' \text{ for } k > k_0\}.$$

For each $k \in \mathbb{Z}^+$ define a mapping $\psi_{u,k} \colon X \to \Omega_u^k$ as follows:

$$\omega_0^k \in \psi_{u,k}(x) \iff \omega_0^k \in \Omega_u^k \text{ and } K(\omega_0^k) = \min_{\omega \in \{\psi_u(x)\}} K(\omega^k).$$

Note that for each fixed $k > k_0$,

$$\operatorname{card}\{\psi_{u,k}(Z_{k,h'})\} < 2^{kh'+1}. \tag{1}$$

Without loss of generality, we may assume that:

1. the cover u is μ-continuous;
2. $\mu(W(u)) < \tau$ ($\tau > 0$, arbitrary);
3. $h_\mu(T, \beta(u)) > h' + \varepsilon$ ($h' + \varepsilon < h_\mu(T)$).

Then by Lemma 3.13 there exists L such that $\mu P_L > 1-\varepsilon$. Finally, by virtue of the consequence of the Shannon–McMillan–Breiman theorem, $\Omega_{\beta(u)} = A_n \cup B_n$ and for $n > N(\varepsilon)$:

1. $\mu(B_n) > 1 - \varepsilon$,

2. $\mu(C(\omega^n)) \geq 2^{-n(h_\mu(T,\beta(u))+\varepsilon)}$ for $\omega \in B_n$.

Of course,

$$\text{card}\{\omega : \omega \in B_n\} \leq 2^{n(h_\mu(T,\beta(u))+\varepsilon)}.$$

Put $S_n = Z_{n,h'} \cap P_n \cap \varphi^{-1}_{\beta(u)}(B_n)$.

For $n > R = \max(k_0, L, N(\varepsilon))$ we have $\mu(S_n) > 1 - 3\varepsilon$. Since $S_n \subseteq Z_{n,h'}$, for $n > R$,

$$\text{card}\{\Phi_{u,n} = \{\psi_{u,n}(x) : x \in S_n\}\} < 2^{nh'+1}. \tag{2}$$

On the other hand, if $x \in S_n$, then $x \in \varphi^{-1}_{\beta(u)}(B_n)$ and for $n > R$

$$\text{card}\{(\varphi_{\beta(u)}(x))^n : x \in S_n\} \geq (1 - 3\varepsilon) \cdot 2^{n(h_\mu(T,\beta(u))-\varepsilon)}. \tag{3}$$

Put $\Theta_u(x, n) = \text{card}\{(\varphi_u(x))^n\}$; then from (3) we obtain:

$$\sum_{x \in S_n} \Theta_{\beta(u)}(x, n) \geq (1 - 3\varepsilon) \cdot 2^{n(h_\mu(T,\beta(u))-\varepsilon)}.$$

On the other hand, from (2)

$$\sum_{x \in S_n} \Theta_{\beta(u)}(x, n) = \sum_{\omega^n \in \Phi_{u,n}} \sum_{\substack{x \in \psi^{-1}_{u,n}(\omega^n) \\ x \in S_n}} \Theta_{\beta(u)}(x, n) \leq 2^{nh'+1} \cdot \max_{\omega^n \in \Phi_{u,n}} \sum_{\substack{x \in \psi^{-1}_{u,n}(\omega^n) \\ x \in S_n}} \Theta_{\beta(u)}(x, n).$$

Thus, for each $n > R$ there exists $\omega^n \in \Phi_{u,n}$ such that

$$G(\omega^n) = \sum_{\substack{x \in \psi^{-1}_{u,n}(\omega^n) \\ x \in S_n}} \Theta_{\beta(u)}(x, n) \geq (1 - 3\varepsilon) \cdot 2^{n(h_\mu(T,\beta(u))-\varepsilon-h')-1}.$$

Consequently,

$$\lim_{n\to\infty} \frac{\log(\max G(\omega^n))}{n} \geq h_\mu(T, \beta(u)) - \varepsilon - h' > 0. \tag{4}$$

However, if $n > R$, $x \in S_n$, and moreover $\omega^n \in \{\psi_{u,n}(x)\}$, $\widetilde{\omega}^n = \{(\varphi_{\beta(u)}(x))^n\}$, then the words

$$\omega^n = \omega(0) \ldots \omega(n - 1) \quad \text{and} \quad \widetilde{\omega}_n = \widetilde{\omega}(0) \ldots \widetilde{\omega}(n - 1)$$

may differ in not more than $n\tau$ places, since $x \in P_n$.

Consequently, for any $\omega^n \in \Phi_{u,n}$,

$$G(\omega^n) = \sum_{\substack{x \in \psi^{-1}_{u,n}(\omega^n) \\ x \in S_n}} \Theta_{\beta(u)}(x, n) \leq \binom{n}{n\tau} \cdot M^{n\tau};$$

in this formula $M = \text{card}\, u$. Applying Stirling's formula, we obtain for sufficiently large n

$$\frac{\log G(\omega^n)}{n} \leq \tau\Big(\log \frac{1 - \tau}{\tau} + \log M\Big) + o(1) \tag{5}$$

also for any $\widetilde{\omega}_n \in \Phi_{u,n}$.

But since

$$\lim_{\tau\to 0} \tau\Big(\log \frac{1 - \tau}{\tau} + \log M\Big) = 0,$$

(5) contradicts (4). The proof of Lemma 3.15 is complete. □

Lemma 3.16. *Let $\mu \in EM(X, T)$. Then $K(x, T) \leq h_\mu(T)$ for μ-almost all $x \in X$.*

Proof. Let $\{u_i\}_{i\in\mathbb{Z}^+}$ be a countable family of covers with $\lim_{i\to\infty} d(u_i) = 0$. Then for all $x \in X$, by Properties 3.4 and 2.5,

$$K(x,T) = \sup_{i\in\mathbb{Z}^+} K(x,T \mid u_i) \le \sup_{i\in\mathbb{Z}^+} K(x,T \mid \beta(u_i))$$

and hence, for μ-almost all $x \in X$, by Lemma 2.6,

$$K(x,T) \le \sup_{i\in\mathbb{Z}^+} h_\mu(T,\beta(u_i)) \le h_\mu(T).$$

□

Theorem 3.1 follows directly from Lemmas 3.15 and 3.16.

Of independent interest is

Proposition 3.17. *The partition of the space X into level sets of the function $K(x,T)$ is measurable and invariant.*

Proof. For a family of covers $\{u_i\}_{i\in\mathbb{Z}^+}$ ($\lim_{i\to\infty} d(u_i) = 0$), denote

$$\bar{R}^t_{k,i} = \{x : K(x,T \mid u_i) \ge t - 1/k\},$$
$$\underline{R}^t_{k,i} = \{x : K(x,T \mid u_i) \le t + 1/k\}.$$

By Lemma 3.14 the sets $\bar{R}^t_{k,i}$ and $\underline{R}^t_{k,i}$ are invariant and measurable. Consequently, so are the sets

$$\bar{R}^t = \bigcap_{k=1}^{\infty}\bigcup_{i=1}^{\infty} \bar{R}^t_{k,i} \quad \text{and} \quad \underline{R}^t = \bigcap_{k=1}^{\infty}\bigcap_{i=1}^{\infty} \underline{R}^t_{k,i},$$

and so is $\bar{X}_t = \{x : K(x,T) = t\} = \bar{R}^t \cap \underline{R}^t$. □

Remark. *Later, in Proposition 3.20, the relationship of the decomposition of the space X into the sets $\bar{X}_t$ with the ergodic decomposition of the DS (X,T) is established.*

2. Estimating the complexity of a concrete trajectory

Definition 3.18 (T. Kamae [24])**.** The number

$$h_T(x) = \sup_{\mu\in V_T(x)} h_\mu(T)$$

will be called the *entropy of the point x*; if $h_T(x) = 0$, the point x is called *deterministic.*

Theorem 3.2. *For any point $x \in X$,*

$$K(x,T) \le h_T(x).$$

Proof. Suppose the contrary: let $x_0 \in X$ and h' be such that

$$K(x_0,T) > h' > h_T(x_0).$$

Then it is not hard to show that there exist: a cover u, a sequence $\{n_k\}$ ($\lim_{k\to\infty} n_k = \infty$), and a measure $\mu_0 \in V_T(x_0)$ such that simultaneously

1. $\limsup_{n_k\to\infty}(1/n_k)\min_{\omega^{n_k}\in\{(\psi_u(x_0))^{n_k}\}} K(\omega^{n_k}) > h'$;

2. $\mu_{x_0,n_k} \Rightarrow \mu_0$;

3. the cover u is μ_0-continuous;

4. the sequence $\{n_k\}$ has the special form $n_k = t_k \cdot k_0$, where $t_k \in \mathbb{Z}^+$, and k_0 is such that $H_{\mu_0}(\beta^{k_0}(u)) < k_0 \cdot h'$.

From 1) it follows that

$$\lim_{n_k\to\infty} \frac{K((\varphi_{\beta(u)}(x_0))^{n_k})}{n_k} > h', \tag{1}$$

where $\beta(u)$, as before, denotes the partition obtained from the cover u.

On the other hand, putting

$$p_B^{r,k}(x_0,n) = \frac{\sum_{i=0}^{[(n-r)/k]} \chi_B(T^{ik+r}x_0)}{[n/k]},$$

from 2) and 3), by Lemma 3.12, for any $B \in \beta^{k_0}(u)$,

$$\lim_{n_k\to\infty} p_B^{0,1}(x_0,n_k) = \mu_0(B). \tag{2}$$

And since for all n_k of our form

$$\frac{1}{k_0}\sum_{r=0}^{k_0-1} p_B^{r,k_0}(x_0,n_k) = p_B^{0,1}(x_0,n_k), \tag{3}$$

by the continuity and concavity of entropy and relations (2) and (3), we conclude that for any sufficiently large n_k there exists $r = r(n_k)$ such that

$$\begin{aligned} -\sum_{B\in\beta^{k_0}(u)} p_B^{r,k_0}(x_0,n_k)\log p_B^{r,k_0}(x_0,n_k) &\le \Big(h'\cdot\frac{n_k}{k_0}\Big)k_0 + o(n_k) \\ &\le h'\cdot n_k + o(n_k). \end{aligned}$$

But then, for the same n_k, for the frequency-encoding function A^* constructed in the proof of Theorem 3.1, we obtain:

$$\limsup_{n_k\to\infty} \frac{K_{A^*}((\varphi_{\beta(u)}(x_0))^{n_k})}{n_k} \le h',$$

which contradicts (1). □

Corollary 3.19. (a) *If $x \in M_\mu$, then $K(x,T) \le h_\mu(T)$.*

(b) *If x is deterministic, then $K(x,T) = 0$.*

Remarks. *1. The inequality in Theorem 3.2 can be strict (see §4, Example 4.3.2).*

2. The estimate of $K(x,T)$ given by Theorem 3.2 is sharper than the estimate of Proposition 3.6, since by the Dinaburg–Goodwyn theorem $h_T(x) \le h(T \mid \overline{O(x)})$; the latter inequality can also be strict (see §4, Example 4.2.2).

In conclusion of this section we note that Theorems 3.1 and 3.2 indicate the existence of a close connection between the decomposition of the space into level sets of the function $K(x,T)$ and the classical decomposition of a DS into ergodic sets.

Proposition 3.20. *Let*

$$\overline{W}_t = \Big\{\bigcup_{\mu\in EM(X,T)} M_\mu : h_\mu(T) \le t\Big\}, \quad \overline{X}_t = \{x : K(x,T) \le t\};$$

then $\overline{X}_t \supseteq \overline{W}_t$, and the set $\overline{X}_t \setminus \overline{W}_t$ is of invariant measure zero (but possibly non-empty — see Example 4.3.2).

§4 On various approaches to the notion of a typical (complex, random) trajectory

It seems beyond doubt that at the present time many researchers working in the most diverse parts of the theory of dynamical systems have run into "complexity" questions. However, the naturally arising diversity of complexity-related characteristics in the absence of a universal measure of complexity makes the exchange of ideas and results between various aspects of the theory of dynamical systems difficult.

In the present section the author has set himself the goal of bringing together and giving, as far as possible, a uniform analysis of some "complexity" characteristics of the theory of dynamical systems. As before, we shall not draw distinctions between the notions of point, trajectory, and sequence where this does not lead to misunderstanding.

The first question on the typicality of a trajectory was considered by A. Poincaré. The set of wandering points which he singled out turned out to be a set of the first category and of invariant measure zero. After the appearance of works by Birkhoff, Krylov and Bogolyubov, an approach took shape according to which the investigation of certain typicality properties of a trajectory for a DS of general type reduces to the analysis of analogous characteristics for minimal (topological typicality) or uniquely ergodic (typicality in the measure-theoretic sense) dynamical systems. A dynamical system possessing both these properties is called *strictly ergodic*.[4]

4.1. Minimality, unique ergodicity and topological entropy

4.1.1

All trajectories of a minimal system are recurrent in the sense of Birkhoff (almost periodic). A minimal system need not be uniquely ergodic. Furstenberg [21] constructed an example of a minimal system with $h(T) > 0$.

4.1.2

Within the combinatorial approach to problems of ergodic theory, many authors (see [17] and references in [8]) have given constructive examples of minimal or strictly ergodic systems. The final result, due to Grillenberger [23], is the following: for any number $h < \log N$ there exists a constructively describable sequence $\omega \in \Lambda_N$ such that the DS $(\overline{O}(\omega), \sigma)$ is strictly ergodic and $h(\sigma \mid \overline{O}(\omega)) = h$.

4.2. Regularity, entropy of a point in the sense of Kamae, combinatorial entropy of a trajectory

4.2.1

In the case of a DS of general type, the analogue of the notion of unique ergodicity is the notion of *μ-regularity*. By definition, a point x is called μ-regular if $V_T(x) = \{\mu\}$. Of course, $h_T(x) = h_\mu(T)$ for every μ-regular point x. For an ergodic measure, Bowen [19] showed[5] that $h(T \mid M_\mu) = h_\mu(T)$.

[4] In terminology we follow [20]; in [12] it is somewhat different.

[5] Here the quantity $h(T \mid M_\mu)$ is understood in the sense of [19].

4.2.2

By the Dinaburg–Goodwyn theorem,

$$h_T(x) \leq h(T \mid \overline{O}(x));$$

thus the Kamae entropy does not exceed the combinatorial entropy. The example of the sequence

$$\omega_0 = p_0\, 0\, p_1\, 0\, 0\, p_2\, 0\, 0\, 0\, \ldots\, p_k \underbrace{0 \ldots 0}_{2^k},$$

where p_k is the k-th binary word, provides an example of a point of the symbolic DS (Λ_2, σ) for which this inequality is strict. Indeed, it is not hard to check that $h_\sigma(\omega_0) = 0$. On the other hand, $\overline{O}(\omega_0) = \Lambda_2$ and hence $h(\sigma \mid \overline{O}(\omega_0)) = \log 2$.

At the same time we have obtained an example of a trajectory that is quasi-random in the sense of Alekseev [1] and yet deterministic.

For any trajectory of a strictly ergodic DS,

$$h(T \mid \overline{O}(x)) = h_T(x) = h(T).$$

In conclusion we note that the notion of a μ-regular point first appeared and is used fruitfully in number theory. An ergodic treatment of a number of problems in analytic number theory can be found in the monograph of A. G. Postnikov [13].

4.3. Algorithmic complexity of a trajectory

4.3.1

By the remark to Proposition 3.6,

$$K(x, T) \leq h(T \mid \overline{O}(x))$$

for any $x \in X$, i.e. the complexity of a trajectory does not exceed its combinatorial entropy; this inequality can be strict (see Example 4.3.2 below). Thus, if a DS is not "quasi-random" in the sense of Alekseev (i.e. if $h(T) = 0$), then the complexity of every trajectory equals zero.

In conclusion we note that in the case of a symbolic DS the complexity and combinatorial entropy are related by (A. A. Brudno, see [1, Theorem 5.4])

$$h(\sigma \mid \overline{O}(\omega)) = \limsup_{n\to\infty} (1/n) K((\sigma^k \omega)^n).$$

4.3.2

By Theorem 3.2 the complexity of a μ-regular trajectory does not exceed the entropy with respect to the measure μ, i.e. $K(x, T) \leq h_\mu(T)$. We show that this inequality can be strict. Indeed, let

$$\omega_1 = 0.1.01.10.11.000.\ldots,$$

i.e. all finite binary words are written out in a row. It is well known that $\omega_1 \in M_m$. Here m is the uniform Bernoulli measure, hence $h_\sigma(\omega_1) = \log 2$. On the other hand, it is not hard to construct a computable function A_0 that recovers the binary word ω_1^n from the number n. We obtain:

$$K(\omega_1) \leq \lim_{n\to\infty} \frac{K_{A_0}(\omega_1^n)}{n} \leq \lim_{n\to\infty} \frac{\log n}{n} = 0.$$

Thus, there exist trajectories which are non-deterministic in the sense of Kamae with zero complexity, and the set $\overline{X}_t \setminus \overline{W}_t$ from Proposition 3.20 may be non-empty.

4.3.3

A stronger result also holds. Since the construction of the sequence ω in the example of Grillenberger mentioned above is constructive, $K(\omega) = 0$. By the arbitrariness of the choice of the number h in this example we see that a trajectory need not be typical, in the sense of complexity, even for the strictly ergodic system it generates, and that the function $K(x, T)$ need not be constant on such a system.

It is appropriate to introduce the following

Definition 4.1. A point $x \in X$ will be called *μ-complex* if $\mu \in V_T(x)$ and $K(x, T) = h_\mu(T)$.

4.4. Complexity and randomness

4.4.1

Here we shall consider the connection of the notion of complexity of a trajectory, introduced in 4.3.2, with the notion of randomness in the sense of Martin-Löf, generally accepted at present. We briefly dwell on the definition of this notion (considering, for simplicity, the space Λ_N).

A *test* is a family of sets $c = \{C_i\}_{i \in \mathbb{Z}^+}$ constructively describable and satisfying the conditions:

1. $C_0 \supseteq C_1 \supseteq \ldots$;
2. each C_i is sequentially open;
3. $\mu(C_i) \le 2^{-i}$ for each $i \in \mathbb{Z}^+$.

The set $R_c = \bigcap_{i \in \mathbb{Z}^+} C_i$ by definition consists of the sequences which do not pass the test c. In [25] Martin-Löf showed the existence of a universal test c_0 such that $R_c \subseteq R_{c_0}$ for any test c. Sequences $\omega \notin R$ are proposed to be called *μ-random.*

In [11] it is shown that under some natural assumptions all probability-theoretic laws hold for sequences random in this sense. Therefore the equality $K(\omega) = h_\mu(\sigma)$ is (by Theorem 3.1) a necessary condition for randomness of ω with respect to the measure μ. We show that this condition is not sufficient. Let m be the uniform Bernoulli measure on Λ_2, and let the point ω_0 be m-complex and m-random. Define a sequence $\widetilde{\omega} = \{\widetilde{\omega}(i)\}_{i \in \mathbb{Z}^+}$ as follows:

$$\widetilde{\omega}_0(i) = \begin{cases} 0, & \text{if } i = 2^k \ (k = 0, 1, 2, \ldots) \\ \omega_0(i), & \text{otherwise.} \end{cases}$$

Obviously $V_\sigma(\widetilde{\omega}_0) = V_\sigma(\omega_0)$. Furthermore, let $A' = A'(p)$ be an a. opt. f. whose value on words $p = \bar{p}_1 p_2$ is obtained by successively inserting the letters of the word p_1 into the word $A(p_2)$ at the positions numbered 2^i $(i = 1, 2, \ldots, l(p_1))$. Then for all n,

$$K_{A'}(\widetilde{\omega}_0^n) \le K_{A'}(\omega_0^n) + 2 + 2 \log n.$$

Therefore,

$$K(\widetilde{\omega}_0) \ge K(\omega_0) = \log 2, \quad \text{i.e.} \quad K(\widetilde{\omega}_0) = \log 2.$$

Thus $\widetilde{\omega}_0$ is also m-complex. However, it is not difficult to verify that the family of sets

$$c = \{C_i\}_{i \in \mathbb{Z}^+}, \quad \text{where } C_i = \{\omega \in \Lambda_2 : \omega(2^k) = 0 \text{ for } k = 0, 1, \ldots, i-1\},$$

is a test. Since the sequence $\widetilde{\omega}_0 \in R_c$, it cannot be m-random.

Conclusion

In contrast to the preceding material, the conclusion does not contain mathematically rigorous results. Its goal is to state, at an intuitive level, the author's attempt to find a general approach to the consideration of various characteristics of a dynamical system that are, in some degree, related to questions of complexity. The considerations adduced below, which may seem disputable, allow one to regard all results of the present work as a single whole, connecting the theory of algorithms with the theory of dynamical systems.

The main connecting link is the following

Thesis. *Every entropy characteristic of a dynamical system is some algorithmic function effecting the encoding of a finite segment of a trajectory of the dynamical system into a binary word. In so doing, the numerical expression of this characteristic is the asymptotics of the amount of information (measured in bits/symbol) necessary for recovering, with the aid of this function, any trajectory among one or another class of trajectories.*

We give some considerations in support of this thesis (restricting ourselves — which is not very essential — to the case of a symbolic system).

In Proposition 1.12 a function constructed by Shannon is given which performs the encoding of μ-almost any sufficiently long trajectory using $h_\mu(\sigma)$ bits/symbol. The restriction "almost any" here arises naturally from the insensitivity of the metric entropy to the behavior of the system on a set of measure zero. The second restriction — on the length of the trajectory — is an inevitable consequence of any algorithmic approach, since at fixed lengths one would have to take into account too many particularities which do not affect the asymptotic behavior.

In the general case (without restrictions on the measure) contained in Proposition 1.12, the same result is obtained if one uses the frequency-encoding function A^* constructed in Lemma 1.9.

Thus, $h_\mu(\sigma)$ bits/symbol is the amount of information necessary and sufficient for recovering, by means of the function A^*, almost any trajectory of the DS (Ω, σ).

We turn now to topological entropy. The computable function (which we shall denote by A^T) constructed in the proof of Lemma 2 of [3] allows one to assert in turn that the amount of information necessary and sufficient for recovering any trajectory of the system by means of the function A^T is $h(T)$ bits/symbol.

It is clear that in the case of entropy characteristics of *individual* trajectories one must recover only that trajectory itself. One can say that a simple modification of A^* will lead us to the entropy of a point in the sense of Kamae (as is done, for example, in the proof of Theorem 3.2), and the combinatorial entropy of a trajectory is obtained by a certain transformation of the function A^T.

Finally, the very definition of algorithmic complexity of a trajectory corresponds to an asymptotically optimal function; Schnorr in [16] showed how, following this idea but *not* averaging over the length of the trajectory, one obtains sequences random in the sense of Martin-Löf.

We can now explain the nature of Examples 4.2.2, 4.3.2 and 4.4.1. The first of them is an example of a dense trajectory with pathological statistical properties. Hence this trajectory, typical for the function A^T and consequently quasi-random, is trivial from the point of view of the function A^* and consequently deterministic (since $h_\sigma(\omega_0) = 0$).

The sequence ω_1 in Example 4.3.2, by its statistical properties, does not differ from other regular sequences; therefore $h_\sigma(\omega_1) = \log 2$ (non-determinacy). However it is describable if one knows binary arithmetic, and therefore its complexity with respect to an asymptotically optimal function equals zero. In Example 4.4.1 it is shown that averaging over the length of a trajectory, which does not affect the μ-complexity of the trajectory, may affect its membership in the set of μ-random trajectories.

Received 1.XII.1978

References

[1] Alekseev V. M. *Symbolic dynamics.* Trudy of the XI Summer Mathematical School. Institute of Mathematics, AN UkrSSR, Kiev, 1976.

[2] Billingsley P. *Ergodic theory and information.* Moscow: Mir, 1963.

[3] Brudno A. A. Topological entropy and complexity in the sense of A. N. Kolmogorov. — Uspekhi Mat. Nauk, 1974, 29, no. 6, 157–158.

[4] Brudno A. A. Symbolic dynamics and entropy. — Uspekhi Mat. Nauk, 1977, 32, no. 3, 180.

[5] Brudno A. A. On the complexity of trajectories of a dynamical system. — Uspekhi Mat. Nauk, 1978, 76, no. 1, 207–208.

[6] Dinaburg E. I. Connection between various entropy characteristics of dynamical systems. — Izv. AN SSSR, Ser. Mat., 1971, 35, no. 2, 324–366.

[7] Zvonkin A. K., Levin L. A. The complexity of finite objects and the foundations of the notions of information and randomness by means of the theory of algorithms. — Uspekhi Mat. Nauk, 1970, 25, no. 6, 85–127.

[8] Katok A. B., Sinai Ya. G., Stepin A. M. Theory of dynamical systems and general groups of transformations with invariant measure. — In: Itogi Nauki, Mat. Analiz, vol. 13. Moscow, 1967, 129–262.

[9] Kolmogorov A. N. A new metric invariant of transitive automorphisms of Lebesgue spaces. — Dokl. AN SSSR, 1958, 119, no. 5, 861–864.

[10] Kolmogorov A. N. Three approaches to the definition of the concept of "amount of information". — Problemy Peredachi Informatsii, 1965, 5, no. 3, 3–7.

[11] Levin L. A. Laws of conservation (non-increase of information) and questions of the foundations of probability theory. — Problemy Peredachi Informatsii, 1976, 74, no. 3, 30–35.

[12] Oxtoby J. Ergodic sets. — Uspekhi Mat. Nauk, 1953, 8, no. 3, 75–97.

[13] Postnikov A. G. Ergodic problems in the theory of congruences and the theory of diophantine approximations. — Trudy MIAN, vol. 82, Moscow, 1966.

[14] Rokhlin V. A. Lectures on the entropy theory of measure-preserving transformations. — Uspekhi Mat. Nauk, 1967, 22, no. 5, 3–56.

[15] Shannon C. *Papers on information theory and cybernetics.* Moscow: IL, 1963.

[16] Schnorr C. P. A unified approach to the definition of random sequences. — In: Complexity of computations and algorithms. Moscow: Mir, 1974, 370–387.

[17] Jacobs K. Machine-generated 0–1 sequences. — In: Turing machines and recursive functions. Moscow: Mir, 1972, 216–247.

[18] Adler R. L., Konheim A. G., McAndrew M. H. Topological entropy. — Trans. Amer. Math. Soc., 1967, 114, no. 2, 309–319.

[19] Bowen R. Topological entropy for noncompact sets. — Trans. Amer. Math. Soc., 1973, vol. 184, 125–136.

[20] Denker M., Grillenberger C., Sigmund K. *Ergodic theory on compact sets.* — Lecture Notes in Mathematics, 527, 1976.

[21] Furstenberg H. Disjointness in ergodic theory, minimal sets and a problem in diophantine approximation. — Math. Syst. Theory, 1967, 1, 1–49.

[22] Goodwyn L. W. Comparing topological entropy with measure-theoretic entropy. — Amer. J. Math., 94, no. 2 (1972), 366–388.

[23] Grillenberger C. Constructions of strictly ergodic systems, I. Given entropy. — Z. Wahrscheinlichkeitstheorie und Verw. Geb., 1973, 25, no. 4, 323–334.

[24] Kamae T. Normal numbers and ergodic theory. — Lecture Notes in Mathematics, 550, 1976.

[25] Martin-Löf P. The definition of random sequences. — Information and Control, 1966, 9, 619–682.

[26] Morse M., Hedlund G. A. Symbolic dynamics II, Sturmian trajectories. — Amer. J. Math., 1940, 62, no. 1, 1–42.