

MSC 2010: 37A05, 37B10

# On the entropy of symbolic image of a dynamical system<sup>1</sup>

G. S. Osipenko\*, N. B. Ampilova\*\*

\*Sebastopol Branch of Lomonosov Moscow State University, Sebastopol 299001

*E-mail:* george.osipenko@mail.ru

\*\*Saint-Petersburg State University, Saint-Petersburg 199034.

*E-mail:* n.ampilova@spbu.ru

**Abstract.** We consider the estimation of the entropy of a discrete dynamical system by using a symbolic image that is a directed graph constructed by means of a finite covering of phase space. Trajectories of the system are coded by paths on the graph. Flows on a symbolic image approximate invariant measures of the system. The maximal metric entropy of a symbolic image is estimated by the logarithm of the maximal eigenvalue of the symbolic image adjacency matrix. There is the flow on which this entropy is achieved. The invariant measure of the maximal entropy is estimated by using this flow.

**Keywords:** symbolic image, dynamical systems, invariant measure, entropy.

## 1. Symbolic image of a dynamical system

Let  $f: M \rightarrow M$  be a homeomorphism of a compact manifold  $M$  generating a discrete dynamical system

$$x_{n+1} = f(x_n), \quad (1.1)$$

and  $\rho(x, y)$  be a distance on  $M$ . In what follows we use the concept of symbolic image of a dynamical system [16], which brings together symbolic dynamics [1, 13] and numerical methods [9]. Let  $C = \{M(1), \dots, M(n)\}$  be a finite closed covering of a manifold  $M$ . The set  $M(i)$  is called cell with index  $i$ .

**Definition 1.** [15] Symbolic image of the dynamical system (1.1) for a covering  $C$  is a directed graph  $G$  with vertices  $\{i\}$  corresponding to cells  $\{M(i)\}$ . The vertices  $i$  and  $j$  are connected by the edge  $i \rightarrow j$  iff

$$f(M(i)) \cap M(j) \neq \emptyset.$$

Symbolic image is a tool for a space discretization and graphic representation of the dynamic of a system under study, which allows the obtaining useful information about the global structure of the system dynamics. Symbolic image depends on a covering  $C$ . The existence of an edge  $i \rightarrow j$  guaranties the existence of a point  $x \in M(i)$  such that

---

<sup>1</sup>This research was carried out with the financial support of the Russian Foundation for Basic Research (grant A no. 19-01-00388)

$f(x) \in M(j)$ . In other words, an edge  $i \rightarrow j$  is the trace of the mapping  $x \rightarrow f(x)$ , where  $x \in M(i)$ ,  $f(x) \in M(j)$ . If there isn't an edge  $i \rightarrow j$  on  $G$  then there are not the points  $x \in M(i)$  such that  $f(x) \in M(j)$ .

We do not place special restrictions on a covering  $C$ , but basing on the theorem about the triangulation of a compact manifold [18] we may without loss of generality assume that cells  $M(i)$  are polyhedrons intersecting on their boundary disks. In practice  $M$  is a compact in  $R^d$ , and  $M(i)$  are cubes or parallelepipeds. Let  $C$  be a covering of  $M$  by polyhedrons intersecting on their boundary disks. In what follows we also use a measurable partition  $C^*$ , such that a boundary disk belongs only one of adjoining cells. We assume that cells-polyhedrons are closures of their interiors.

**Definition 2.** A vertex of a symbolic image  $G$  is said to be recurrent if there is a periodic path passing through it. The set of recurrent vertices is denoted by  $RV$ . The recurrent vertices  $i$  and  $j$  are called equivalent if there exists a periodic path passing through  $i$  and  $j$ .

Thus, the set of recurrent vertices  $RV$  is split into equivalence classes  $\{H_k\}$ . In the graph theory such classes are called strong connectivity components.

Let

$$diam M(i) = \max(\rho(x, y) : x, y \in M(i))$$

be the diameter of a cell  $M(i)$  and  $d = diam(C)$  be the maximum of the diameters of the cells. The number  $d$  is called the diameter of the covering  $C$ .

A directed graph  $G$  is uniquely defined by its adjacency matrix (matrix of admissible transitions)  $\Pi$ . The matrix  $\Pi = (\pi_{ij})$  has sizes  $n \times n$ , where  $n$  is the number of vertices of  $G$ , and  $\pi_{ij} = 1$  iff there exists the edge  $i \rightarrow j$ , else  $\pi_{ij} = 0$ . Hence an  $i$ -th row in  $\Pi$  corresponds to the vertex  $i$  (cell  $M(i)$ ), and on the place  $j$  in this row there is 1 or 0 depending on the existence (or nonexistence) of nonempty intersection  $f(M(i))$  and  $M(j)$ . Matrix of admissible transitions depends on the numbering of vertices (cells of the covering), so that a change of numeration leads to a change of matrix  $\Pi$ . Note that there exists a numeration transforming the matrix of admissible transitions to a canonical form.

**Proposition 1.** [1] Vertices of a symbolic image  $G$  may be numbered such that the adjacency matrix has the form

$$\Pi = \begin{pmatrix} (\Pi_1) & \dots & \dots & \dots & \dots \\ & \ddots & & & \\ 0 & & (\Pi_k) & \dots & \dots \\ & \ddots & & \ddots & \\ 0 & & 0 & & (\Pi_s) \end{pmatrix},$$

where every diagonal block  $\Pi_k$  corresponds either an equivalence class  $H_k$  of recurrent vertices or a nonrecurrent vertex and consists of one zero. Under diagonal blocks are only zeroes (upper triangular matrix)

## 2. Entropy

In 1865 R. Clausius [4] introduced the most important in thermodynamics concept entropy. To explain the irreversibility of macroscopic states L. Boltzmann in 1872 [3] first introduced statistical approach in thermodynamics : he proposed to describe a state of a system by using its microstates. The Boltzmann entropy  $S$  is statistical entropy for the equiprobable distribution of a system over  $P$  states, it is defined as

$$S = k \log(P),$$

where  $k$  is the Boltzmann constant.

In 1948 C. Shannon [19, 20] introduced the notion of capacity ( $C$ ) for an information channel as follows

$$C = \lim_{T \rightarrow \infty} \frac{\log_2 N(T)}{T},$$

where  $N(T)$  is the number of admissible signals for the time  $T$ . He also defined information entropy as follows

$$h = - \sum_i p_i \log_2 p_i,$$

where  $p_i$  is a probability of  $i$ -th signal (message),  $i \in 1, \dots, n$ , and  $n$  is the number of signals. A. N. Kolmogorov in 1958 [11] introduced entropy in the theory of dynamical systems. Entropy is a fine invariant of a dynamical system, it may be interpreted as a measure of the system chaoticity. Comprehensive information on entropy in dynamical systems is given in [6, 10]. It turns out that entropy characteristics may be obtained both for a system described analytically and for its phase portraits. The application of such characteristics to digital image analysis is given in [2].

*Motivation* Consider a discrete dynamical system  $x_{n+1} = f(x_n)$  on a compact manifold  $M$ , where  $f : M \rightarrow M$  is a homeomorphism. Let  $C = \{M(1), \dots, M(n)\}$  be a finite covering of  $M$  and the sequence  $\{x_k = f^k(x), k = 0, \dots, N-1\}$  be the  $N$ -length part of the trajectory of a point  $x$ . The covering  $C$  generates a coding of this part via a finite sequence  $\xi(x) = \{i_k, k = 0, \dots, N-1\}$ , where  $x_k \in M(i_k)$ . In other words,  $i_k$  is the index of the cell from  $C$  which contains the point  $x_k = f^k(x)$ . Generally speaking, the mapping  $x_k \rightarrow i_k$  is multivalued. The sequence  $\xi = \{i_k\}$  is said to be (admissible) encoding of the trajectory  $\{x_k = f^k(x)\}$  with respect to the covering  $C$ . Assume that we know all admissible  $N$ -length encodings, and there is a need to predict subsequent  $p$ -length encodings, i. e. to find admissible encodings of length  $N+p$ .

Let the number of admissible encodings  $K(N)$  grows exponentially depending on  $N$ . We estimate the rate of growth of encodings by the number

$$h = \lim_{N \rightarrow +\infty} \frac{\log_b K(N)}{N}, \quad (2.1)$$

where  $b$  may be any real number greater than 1. The bases  $b = 2$  (following to Shannon) or  $b = e$  are in common use. The existence of the limit in (2.1) follows from the Polya lemma [1, 13].

**Lemma 1** ([13], p.103). *If a sequence of non-negative numbers  $a_n$  satisfies the inequality*

$$a_{n+m} \leq a_n + a_m,$$

*then there exists  $\lim_{n \rightarrow \infty} \frac{a_n}{n}$ .*

For the number of admissible encodings we have

$$K(n+m) \leq K(n)K(m),$$

hence for the sequence  $a_n = \log_b K(n)$  there exists the limit (2.1).

Thus, for the number of encodings  $K(N)$  we obtain the estimation

$$K(N) \sim Bb^{hN},$$

where  $B$  is a constant. If  $h \neq 0$  then

$$\frac{K(N+p)}{K(N)} \sim b^{hp}.$$

This relation means that for any  $N$  the uncertainty of future encodings grows with the exponent  $hp$  regardless the knowledge of previous encodings.

If the growth of the number of different encodings is not exponential (i. e.  $h = 0$ ), for example as

$$K(N) \sim BN^A,$$

where  $A$  is a positive number (may be large), then

$$\frac{K(N+p)}{K(N)} \sim \left(1 + \frac{p}{N}\right)^A \rightarrow 1,$$

when  $N \rightarrow \infty$ . In other words, the uncertainty of the future decreases when the length  $N$  of known encodings increases.

Thus, if the growth of the number of different encodings is exponential, the uncertainty does not depend on  $N$ , in other case it decreases as  $N$  increases. Value  $h$  may be interpreted as a characteristics of uncertainty (chaoticity) of the dynamic of the system considered.

*Topological entropy.* Let  $f$  be a continuous mapping defined on a manifold  $M$  and  $C = \{M(1), \dots, M(n)\}$  be an open covering of  $M$ . For an integer positive number  $N$  consider a sequence

$$\omega = \omega_1 \omega_2 \cdots \omega_N,$$

where  $\omega_k$  is a number from 1 to  $n$ . Construct the intersection of the form

$$|M(\omega) = M(\omega_1) \cap f^{-1}(M(\omega_2)) \cap \cdots \cap f^{-N+1}(M(\omega_N)), \quad (2.2)$$

which is an open set. The admissible encoding  $\omega$  corresponds to the nonempty intersection  $M(\omega)$ , i.e. there exists  $x \in M(\omega_1)$  such that  $f^k(x) \in M(\omega_{k+1})$ . The

sequence  $\omega$  codes the segment of the trajectory  $\{f^k(x), k = 1, 2, \dots, N\}$ . Consider all the admissible  $N$ -length encodings  $\{\omega\}$  and the collection of sets  $C^N = \{M(\omega)\}$ , which is an open covering. Choose in  $C^N$  a minimal by the number of elements (denoted further by  $|C_N|$ ) finite subcovering  $C_N$ .

Then according to the Polya lemma there exists the limit

$$H(C) = \lim_{N \rightarrow \infty} \frac{\log |C_N|}{N}. \quad (2.3)$$

**Definition 3.** The number

$$h(f) = \sup_C H(C),$$

where supremum is taken over all open coverings  $C$ , is called *topological entropy* of the mapping  $f: M \rightarrow M$ .

It is easy to see that there is little point in using this definition for practical calculation of the entropy, so we consider some other methods.

A covering  $C_2$  is said to be *refined* in a covering  $C_1$ , if any  $A \in C_2$  lies in a set  $B \in C_1$ . A sequence of open coverings  $C_n$  is called *exhaustive* if for any open covering  $C$  there exists the number  $n^*$  such that the covering  $C_n$  is refined in  $C$  for  $n \geq n^*$ .

**Proposition 2** ([1] p. 122). (1) If  $C_n$  is a sequence of open coverings with diameters

$$d_n = \max_{A \in C_n} \text{diam} A$$

tending to zero, then  $C_n$  is an exhaustive sequence. (2) The entropy of the mapping  $f$  is calculated as follows

$$h(f) = \lim_{n \rightarrow \infty} H(C_n).$$

Consider coverings  $C_1$  and  $C_2$ , and construct for each of them nonempty intersections of the form (2.2). Denote the obtained collections of sets by  $C_1^N$  and  $C_2^N$  respectively. In each collection choose the minimal (by the number of elements) subcovering, and denote them  $C_{N1}$  and  $C_{N2}$ .

**Proposition 3.** If  $C_2$  is refined in  $C_1$ , then

$$|C_{N1}| \leq |C_{N2}|,$$

where  $|C_{Ni}|$  denotes the number of elements in the set considered.

*Proof.* If  $A_1 \subset B_1$  and  $A_2 \subset B_2$ , then  $A_1 \cap f^{-1}(A_2)$  lies in  $B_1 \cap f^{-1}(B_2)$ . By the same way one can prove that elements of  $C_2^N$  are in corresponding elements of  $C_1^N$ . Consider  $C_{N1}$  and  $C_{N2}$ . Take from  $C_1^N$  all the elements which contain corresponding elements of  $C_{N2}$ , and form the covering  $C_{N1}^*$ .

Then  $|C_{N1}| \leq |C_{N1}^*| \leq |C_{N2}|$ , because  $C_{N1}$  is a minimal subcovering for  $C_1^N$ . The proposition is proved.

**Corollary 1.** *Assume that  $C_2$  is refined in  $C_1$ , and the numbers  $H(C_1)$ ,  $H(C_2)$  are calculated by (2.3). Then*

$$H(C_1) \leq H(C_2).$$

### 3. Entropy of a symbolic image

Let  $G$  be a graph with the adjacency matrix  $\Pi$ . Denote by  $b_n$  the number of admissible  $n$ -length paths on  $G$ .

**Definition 4.** The number

$$h(G) = \lim_{n \rightarrow \infty} \frac{\ln b_n}{n}$$

is called the *entropy of the graph  $G$* .

Remember that if  $(\Pi^n)$  is the power of  $\Pi$  then the element  $(\Pi^n)_{ij}$  equals the number of admissible  $n$ -length paths from  $i$  to  $j$  and

$$b_n = \sum_{ij} (\Pi^n)_{ij}.$$

**Subdivision process.** We will apply a process of adaptive subdivision to coverings and construct a sequence of symbolic images. At first, let us consider a main step of the process a subdivision of covering. Let  $C = \{M(i)\}$  be a covering of  $M$  and  $G$  be the symbolic image with respect to  $C$ . Suppose a new covering  $NC$  is a subdivision of  $C$ . It is convenient to designate cells of the new covering as  $m(i, k)$ . This means that each cell  $M(i)$  is subdivided by cells  $m(i, k)$ ,  $k = 1, 2, \dots$  which form a subdivision of the cell  $M(i)$ , i.e.,

$$\bigcup_k m(i, k) = M(i).$$

Denote by  $NG$  the symbolic image for the new covering  $NC$ . The vertices of the new symbolic image are denoted by  $(i, k)$ . It is possible that some cells will not be subdivided, i.e.  $m(i, 1) = M(i)$ , and the vertex  $i$  in  $G$  is the vertex  $(i, 1)$  in  $NG$ . The subdivision just described generates a natural mapping  $s$  from  $NG$  onto  $G$  which takes the vertices  $(i, k)$  onto the vertex  $i$ . From  $f(m(i, k)) \cap m(j, l) \neq \emptyset$  it follows that  $f(M(i)) \cap M(j) \neq \emptyset$ , so the edge  $(i, k) \rightarrow (j, l)$  is mapped onto the edge  $i \rightarrow j$ . Hence, the mapping  $s$  takes the directed graph  $NG$  onto the directed graph  $G$ .

**Theorem 1.** *Let  $C_n$  be a sequence of subdivisions of a closed covering, such that cells are polyhedrons intersecting on boundary disks, and diameters  $d_n$  of  $C_n$  tend to zero. Denote by  $G_n$  the symbolic images constructed for a mapping  $f : M \rightarrow M$  in accordance with the sequence  $C_n$ . Then for the entropy of  $f$  the following inequality holds*

$$h(f) \leq \lim_{n \rightarrow \infty} h(G_n).$$

*Proof.* Let  $C_n = \{M(i)\}$  be a closed covering from the sequence described above, and  $G_n$  be a symbolic image of the mapping  $f$  constructed according to  $C_n$ . Consider the set  $P_N$  of encodings of  $N$ -length segments of trajectories. It is obvious that  $P_N$  is not greater than the number of admissible  $N$ -length paths (denoted by  $b_N$ ) on  $G_n$ .

Hence the number  $|C^N|$  of nonempty intersections of the form

$$M(\omega) = M(\omega_1) \cap f^{-1}(M(\omega_2)) \cap \dots \cap f^{-N+1}(M(\omega_N))$$

is not greater than  $b_N$ . Thus,

$$H(C_n) \leq h(G_n).$$

If  $C$  is an open covering then there exists the number  $n^*$  such that the covering  $C_n$  is refined in  $C$  for  $n \geq n^*$ . Then in accordance with Proposition 3 we have

$$H(C) \leq H(C_n) \leq h(G_n).$$

Now consider an exhausting sequence of open coverings  $\{\tilde{C}_m\}$ . Let  $n(m)$  be the number  $n^*$  constructed for the covering  $\tilde{C}_m$ , then

$$H(\tilde{C}_m) \leq H(C_{n(m)}) \leq h(G_{n(m)}),$$

where  $n \geq n(m)$ . If  $m \rightarrow \infty$ , then according to Proposition 2 we have

$$h(f) = \lim_{m \rightarrow \infty} H(\tilde{C}_m) \leq \lim_{n \rightarrow \infty} h(G_n).$$

□

Remember that a matrix  $A(n \times n)$  is called decomposable if it admits an invariant subspace with dimension less than  $n$ , and a matrix  $A$  is called nonnegative (positive) if it has nonnegative (positive) elements. If for a nonnegative matrix  $A$  there is an integer  $s > 0$  such that all the elements of  $A^s$  are positive, then  $A$  is called primitive. In particular the matrix of admissible transitions  $\Pi$  is nonnegative. It is nondecomposable if the symbolic image consists of one class of equivalent recurrent vertices.

**Theorem 2.** (*Perron-Frobenius*) [7, 13]

- If  $A$  is a decomposable nonnegative matrix then it has an eigenvector  $e$  with positive coordinates and the eigenvalue  $\lambda$  with multiplicity 1, and the other eigenvalues  $\mu$  satisfy the inequality  $|\mu| \leq \lambda$ .
- If  $A$  is a decomposable nonnegative matrix and  $|\mu| < \lambda$ , then  $A$  is primitive.
- If  $A$  is a decomposable nonnegative matrix and it has  $h > 1$  eigenvalues which are equal in modulus  $\lambda$ , then  $A$  is not primitive, and by an agreed renumeration of rows and columns it may be transformed to the form

$$\begin{pmatrix} 0 & A_{12} & 0 & \cdots & 0 \\ 0 & 0 & A_{23} & \cdots & 0 \\ & \ddots & & \ddots & \\ 0 & 0 & 0 & \cdots & A_{h-1h} \\ A_{h1} & 0 & 0 & \cdots & 0 \end{pmatrix},$$

where  $A_{ij}$  are square blocks, and  $A^h$  consists of  $h$  primitive blocks.

**Theorem 3.** *The entropy of the graph  $G$  is equal to the logarithm of the maximal eigenvalue of the adjacency matrix*

$$h(G) = \ln \lambda.$$

*Proof.* 1. Consider the case when  $G$  consists of one class of equivalent recurrent vertices. Let  $e$  be a positive eigenvector for the maximal eigenvalue  $\lambda$ , i.e.

$$\Pi e = \lambda e.$$

In the coordinate form we have

$$\sum_j (\Pi^n)_{ij} e_j = \lambda^n e_i. \quad (3.1)$$

Let  $c = \min\{e_i\}$  and  $d = \max\{e_i\}$ . In accordance with the Perron-Frobenius theorem  $c > 0$ . Then the following inequalities hold

$$c \sum_j (\Pi^n)_{ij} \leq \sum_j (\Pi^n)_{ij} e_j \leq d \lambda^n.$$

It follows that

$$\sum_j (\Pi^n)_{ij} \leq \frac{d}{c} \lambda^n$$

for any  $i$ . Summing by  $i$  we obtain

$$b_n = \sum_{ij} (\Pi^n)_{ij} \leq \frac{dr}{c} \lambda^n,$$

where  $r$  is the number of rows in the matrix  $\Pi$ . It follows from (3.1) that

$$c \lambda^n \leq \lambda^n e_i = \sum_j (\Pi^n)_{ij} e_j \leq d \sum_j (\Pi^n)_{ij} \leq d \sum_{ij} (\Pi^n)_{ij}.$$

Hence we have the estimation

$$\frac{c}{d} \lambda^n \leq \sum_{ij} (\Pi^n)_{ij} = b_n,$$

and the inequalities

$$\frac{c}{d} \lambda^n \leq b_n \leq \frac{dr}{c} \lambda^n.$$

The required equality follows from it.



2. Consider the case when there are several classes of equivalent recurrent vertices, i.e. the matrix  $\Pi$  is decomposable. According to Proposition 1, by a renumbering of vertices this matrix may be transformed to the form

$$\Pi = \begin{pmatrix} (\Pi_1) & \cdots & \cdots & \cdots & \cdots \\ & \ddots & & & \\ 0 & & (\Pi_k) & \cdots & \cdots \\ & & \ddots & \ddots & \\ 0 & & 0 & & (\Pi_s) \end{pmatrix},$$

where each diagonal block  $\Pi_k$  corresponds to either one of the classes of equivalent recurrent vertices  $H_k$  or some non-recurrent vertex, and consists of one zero. Under diagonal blocks are zeroes. Each class  $H_k$  has the entropy

$$h(H_k) = \ln \lambda_k,$$

where  $\lambda_k$  is the maximal eigenvalue of  $\Pi_k$ . By the definition the entropy of a symbolic image  $G$  equals

$$h(G) = \lim_{n \rightarrow \infty} \frac{\ln b_n}{n}.$$

Consider an admissible  $n$ -length path  $\omega$ . Assume that on  $G$  there are  $s$  classes of equivalence  $H_k$ . The path  $\omega$  passes both through the vertices from  $H_k$  and non-recurrent vertices not belonging to these classes. Denote by  $\omega_k$  the parts of  $\omega$  which lie in the class  $H_k$ . If we delete from  $\omega$  all the  $\omega_k$  it will contain only different paths  $\sigma_l$  passing through non-recurrent vertices. Thus,  $\omega$  is the sum  $\omega_k$  and  $\sigma_l$ . Combine all the paths  $\sigma_l$  into a sequence  $\sigma$ , which is not necessarily is an admissible path. Let  $K$  be the number of non-recurrent vertices in  $G$ . The sequence  $\sigma$  contains nonrecurrent vertices without repetition. Hence the number of the sequences  $\sigma$  is not greater than the number of permutations of  $K$  elements, i.e.  $K!$ .

Denote by  $n(k)$  the length of the path  $\omega_k$  from  $H_k$ . Then  $n(1) + n(2) + \cdots + n(s) \leq n$ . According to item 1 for every class  $H_k$  there is a number  $d$  such that the number  $b_k(n(k))$  of different  $n(k)$ -length paths  $\omega_k$  is estimated as follows

$$b_k(n(k)) \leq d\lambda_k^{n(k)} \leq d\lambda^{n(k)},$$

where  $\lambda = \max \lambda_k$ . Then for the number of different paths  $\omega_k$  lying in  $\omega$  we have the estimation

$$\prod_k b_k(n(k)) \leq d^s \lambda^{n(1)+n(2)+\cdots+n(s)} \leq d^s \lambda^n.$$

Summing the above results we have

$$b_n \leq K! d^s \lambda^n.$$

Thus, we obtain an upper estimation for the entropy of  $G$ :

$$h(G) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \ln(K! d^s \lambda^n) = \ln \lambda.$$

Prove the opposite inequality. Note that the number of admissible paths on  $G$  is greater than the number of admissible paths in a class  $H_k$ . Then  $h(G) \geq h(H_k) = \ln \lambda_k$  for any  $k$ , which gives the low estimation

$$h(G) \geq \ln \lambda.$$

Hence we have

$$h(G) = \ln \lambda,$$

and the proof is completed.

#### 4. Flows on a symbolic image

Let  $f: M \rightarrow M$  be a homeomorphism of a compact manifold  $M$ . A measure  $\mu$  defined on  $M$  is said to be  $f$ -invariant, if for any measurable set  $A \subset M$  the equality

$$\mu(f^{-1}(A)) = \mu(A) = \mu(f(A))$$

holds. In what follows we assume that all measures considered are the Borel ones. The Krylov-Bogoliubov theorem [12, 10] guaranties the existence of an invariant measure  $\mu$  which is normed on  $M$ :  $\mu(M) = 1$ . Denote by  $\mathcal{M}(f)$  the set of all  $f$ -invariant normed measures. This set is a convex closed compact in weak topology (see [14], p.511). The convergence  $\mu_n \rightarrow \mu$  in this topology means that

$$\int_M \phi d\mu_n \rightarrow \int_M \phi d\mu$$

for any continuous function  $\phi: M \rightarrow R$ .

To understand how a distribution of a measure may appear on a symbolic image, consider the following observation. Assume that there exists a  $f$ -invariant normed measure  $\mu$  on  $M$ , and the cells of a covering  $C$  are polyhedrons intersecting by boundary disks. Construct a measurable partition  $C^* = \{M^*(i)\}$  such that a boundary disk belongs to one of adjoining cells. Then, to every edge  $i \rightarrow j$  of a symbolic image  $G$  we can assign the measure

$$m_{ij} = \mu(M^*(i) \cap f^{-1}(M^*(j))) = \mu(f(M^*(i)) \cap M^*(j)), \quad (4.1)$$

where the last equality follows from the invariance of  $\mu$ . Besides that, the invariance of  $\mu$  leads to the equalities

$$\begin{aligned} \sum_k m_{ki} &= \sum_k \mu(f(M^*(k)) \cap M^*(i)) = \mu(M^*(i)) = \\ &= \sum_j \mu(M^*(i) \cap f^{-1}(M^*(j))) = \sum_j m_{ij}. \end{aligned}$$

The value  $\sum_k m_{ki}$  is called the flow incoming in the vertex  $i$ , and the  $\sum_j m_{ij}$  is called the flow outgoing from  $i$ . The equality

$$\sum_k m_{ki} = \sum_j m_{ij} \quad (4.2)$$

may be interpreted as Kirchoff's law: for any vertex the incoming flow equals the outgoing one. Furthermore, we have

$$\sum_{ij} m_{ij} = \mu(M) = 1. \quad (4.3)$$

It means that the distribution  $m_{ij}$  is normed (probabilistic). Thus, a  $f$ -invariant measure  $\mu$  generates on a symbolic image a distribution  $m_{ij}$  which satisfies the conditions (4.2) and (4.3). The above reasoning leads to the following definition.

**Definition 5.** Let  $G$  be a directed graph. The distribution  $\{m_{ij}\}$  on edges  $\{i \rightarrow j\}$  such that

- $m_{ij} \geq 0$ ;
- $\sum_{ij} m_{ij} = 1$ ;
- for any vertex  $i$

$$\sum_k m_{ki} = \sum_j m_{ij}$$

is called flow on  $G$ .

The last property may be called the invariance of a flow. The norming condition may be written as  $m(G) = 1$ , where the measure of  $G$  is the sum of measures of all edges. Sometimes in the graph theory for such a distribution the term "closed flow" is used.

For the flow  $\{m_{ij}\}$  on  $G$  we may define the measure of a vertex  $i$  as

$$m_i = \sum_k m_{ki} = \sum_j m_{ij}.$$

Then  $\sum_i m_i = m(G) = 1$ .

Thus, a  $f$ -invariant measure generates a flow on a symbolic image. Now we consider the inverse construction. Let on a symbolic image  $G$  a flow  $m = \{m_{ij}\}$  be given, then we can construct the measure  $\mu$  on  $M$  as follows

$$\mu(A) = \sum_i \frac{m_i(v(A \cap M(i)))}{v(M(i))}, \quad (4.4)$$

where  $A$  is a measurable set. Here  $v$  is a normed on  $M$  Lebesgue's measure, and on the assumption  $v(M(i)) \neq 0$ . In this case the measure of a cell  $M(i)$  coincides with

the measure of the vertex  $i$ :  $\mu(M(i)) = m_i$ . As  $v$  is the Lebesgue measure, the measure of boundary disks is equal to zero and the measure of a cell does not depend on the measure of its boundary. In general, the constructed measure  $\mu$  is not  $f$ -invariant. But it is an approximation to an invariant measure in the sense that  $\mu$  converges in weak topology to an invariant measure if the diameter of the covering tends to zero.

**Theorem 4.** [17] *Let on a sequence of symbolic images  $\{G_t\}$  of a homeomorphism  $f$  a sequence of flows  $\{m^t\}$  be defined, and the maximal diameter  $d_t$  of coverings tends to zero when  $t \rightarrow \infty$ . Then*

- *there exists the subsequence of measures  $\mu_{t_k}$  (constructed according to (4.4)) which converges in weak topology to a  $f$ -invariant measure  $\mu$ ;*
- *if a subsequence of measures  $\mu_{t_i}$  converges in weak topology to a measure  $\mu^*$ , then  $\mu^*$  is  $f$ -invariant.*

**Theorem 5.** [17] *For any neighborhood (in weak topology)  $U$  of the set  $\mathcal{M}(f)$  there is a positive number  $d_0$  such that for any covering  $C$  with the diameter  $d < d_0$  and any flow  $m$  on a symbolic image  $G$  with respect to  $C$ , the measure  $\mu$  constructed according to (4.4) by  $m$ , lies in  $U$ .*

## 5. Metric entropy

Let  $\mu$  be a normed invariant measure of a homeomorphism  $f : M \rightarrow M$  and  $C = \{M_1, M_2, \dots, M_m\}$  a measurable partition of the manifold  $M$ .

**Definition 6.** The entropy of the partition  $C$  is defined as

$$H(C) = - \sum_i \mu(M_i) \ln \mu(M_i).$$

Construct a covering  $C^N$  which consists of nonempty intersections of the form

$$A_{i_1} \cap f^{-1}(A_{i_2}) \cap \dots \cap f^{-N+1}(M_{i_N}).$$

If such an intersection is nonempty then the sequence  $i_1, i_2, \dots, i_N$  is admissible coding with respect to the covering  $C$ .

The metric entropy of  $f$  for the covering  $C$  is defined as

$$H(f, C) = \lim_{N \rightarrow \infty} \frac{1}{N} H(C^N).$$

The existence of the limit follows from the Polya lemma.

**Definition 7.** The entropy of  $f$  for an invariant measure  $\mu$  is defined as

$$h(f, \mu) = \sup_C H(f, C),$$

where sup is taken over all measurable finite partitions.

The connection between topological and metric entropy is given by the following theorem.

**Theorem 6.** [5, 8] *The topological entropy of a homeomorphism  $f$  is the least upper bound of metric entropies*

$$h(f) = \sup_{\mu} h(f, \mu).$$

## 6. Stochastic Markov chains

Stochastic Markov chain [7, 13] is defined by a set of states of a system  $\{i = 1, 2, \dots, n\}$  and the matrix of transition probabilities  $P_{ij}$  from a state  $i$  to state  $j$ . Such a matrix is called stochastic if it satisfies the conditions  $P_{ij} \geq 0$  and  $\sum_j P_{ij} = 1$  for every  $i$ . A probabilistic distribution  $p = (p_1, p_2, \dots, p_n)$ ,  $\sum_i p_i = 1$  is said to be stationary if

$$(p_1, p_2, \dots, p_n) \begin{pmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \cdot & \cdot & \dots & \cdot \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{pmatrix} = (p_1, p_2, \dots, p_n),$$

i.e.  $p$  is a left eigenvector of  $P$ .

We show that there is a one-to-one correspondence between a Markov chain and a flow on a graph in which vertices correspond to the states with positive measure.

Let  $m = \{m_{ij}\}$  be a flow on a graph  $G$ . The measure of a vertex  $i$  equals  $m_i = \sum_j m_{ij} = \sum_k m_{ki}$ . If  $m_i \neq 0$  then the vertex  $\{i\}$  is necessary recurrent. It is easy to verify that any flow  $m = \{m_{ij}\}$  on  $G$  generates a stochastic Markov chain in which the states are vertices with nonzero measures, and the transition probabilities from  $i$  to  $j$  are calculated as

$$P_{ij} = \frac{m_{ij}}{m_i}.$$

In this case the stochastic matrix  $P = (m_{ij}/m_i)$  has the stationary distribution coinciding with the distribution of the measure  $m$  over the vertices  $(m_1, m_2, \dots, m_n)$ . This follows from the equality

$$(m_1, m_2, \dots, m_n) \begin{pmatrix} \frac{m_{11}}{m_1} & \frac{m_{12}}{m_1} & \dots & \frac{m_{1n}}{m_1} \\ \frac{m_{21}}{m_2} & \frac{m_{22}}{m_2} & \dots & \frac{m_{2n}}{m_2} \\ \cdot & \cdot & \dots & \cdot \\ \frac{m_{n1}}{m_n} & \frac{m_{n2}}{m_n} & \dots & \frac{m_{nn}}{m_n} \end{pmatrix} = (m_1, m_2, \dots, m_n).$$

Thus, any flow  $m = \{m_{ij}\}$  on a graph  $G$  generates a stochastic Markov chain for which the distribution of the measure  $(m_i)$  on vertices is stationary.

Now we prove the inverse fact: for any stochastic matrix  $P = (P_{ij})$  and its stationary distribution  $p = (p_i)$  there exists a flow  $m = \{m_{ij}\}$  on a graph  $G$  for which the

distribution of the measure on vertices coincides with the stationary distribution, i.e.  $m_i = p_i$ .

Actually, let  $P$  be a stochastic matrix and  $pP = p$ . Consider a graph  $G$  which has  $n$  vertices  $\{i\}$ , and the edge  $i \rightarrow j$  there exists if  $P_{ij} > 0$ . Construct the distribution on edges  $m_{ij} = P_{ij}p_i$  and show that the distribution is a flow on  $G$ . As  $P$  is stochastic then  $\sum_j P_{ij} = 1$  for any  $i$ . Hence

$$\sum_j m_{ij} = \sum_j P_{ij}p_i = p_i \sum_j P_{ij} = p_i.$$

As  $pP = p$  then  $\sum_k p_k P_{ki} = p_i$ , so

$$\sum_k m_{ki} = \sum_k p_k P_{ki} = p_i = \sum_j m_{ij},$$

i.e. for the distribution  $m_{ij}$  the Kirchoff law holds. Moreover,  $\sum_{ij} m_{ij} = \sum_i p_i = 1$ .

From the above it follows that the construction of a flow on a graph results in obtaining a Markov chain.

## 7. Flow entropy

The developed technique may be applied to estimate metric entropy. Let for a mapping  $f$  and a covering  $C$  a symbolic image  $G$  and a flow  $m = \{m_{ij}\}$  be constructed. Any flow  $m$  may be considered as the approximation to an invariant measure  $\mu$ , if the diameter of  $C$  is small enough. The flow  $m$  on  $G$  generates the Markov chain in which the states coincide with vertices of  $G$ , and transition probabilities are defined as

$$p_{ij} = \frac{m_{ij}}{m_i}.$$

The matrix  $P = (p_{ij})$  has the stationary distribution  $(m_1, m_2, \dots, m_n)$  for which entropy is calculated by the formula (see [13], p. 443)

$$h_m = - \sum_i m_i \sum_j p_{ij} \ln p_{ij}.$$

Substituting  $p_{ij} = m_{ij}/m_i$  we obtain

$$\begin{aligned} h_m &= - \sum_i m_i \sum_j \frac{m_{ij}}{m_i} \ln \left( \frac{m_{ij}}{m_i} \right) = - \sum_{ij} m_{ij} \ln \left( \frac{m_{ij}}{m_i} \right) = \\ &= - \sum_{ij} m_{ij} \ln m_{ij} + \sum_{ij} m_{ij} \ln m_i = - \sum_{ij} m_{ij} \ln m_{ij} + \sum_i m_i \ln m_i. \end{aligned}$$

By this means entropy can be calculated by the flow  $m_{ij}$  as

$$h_m = - \sum_{ij} m_{ij} \ln m_{ij} + \sum_i m_i \ln m_i. \quad (7.1)$$

The last equality allows estimating the entropy of  $f$  for the invariant measure  $\mu$ , where the flow  $m$  is an approximation of  $\mu$ .

## 8. Flow with maximal entropy

Let  $\Pi$  be the matrix of admissible transitions for a graph  $G$ . Our objective is to construct the flow which has maximal entropy among all the flows on  $G$ . As any flow is grouped on a component of recurrent vertices, it may be thought that  $G$  consists from one component.

**Theorem 7.** *There is a flow  $m$  on  $G$  such that:*

$$h_m = h(G) = \ln \lambda.$$

*Proof.* 1. Eigenvalues of any real matrix  $A = (a_{ij})$  coincide with the eigenvalues of the transposed (conjugate) matrix  $A^*$ . Really, as  $\det A = \det A^*$ , then

$$\det(A - \lambda E) = \det(A - \lambda E)^* = \det(A^* - \bar{\lambda} E).$$

Hence to an eigenvalue  $\lambda$  of  $A$  corresponds the conjunctive eigenvalue  $\bar{\lambda}$  of  $A^*$ . The roots of a real characteristic polynomial are either real or complex-conjugate, hence the eigenvalues of the matrices  $A$  and  $A^*$  coincide.

2. Let  $A$  be the matrix of admissible transitions of a graph  $G$  and  $\lambda$  be the maximal eigenvalue from the Perron-Frobenius theorem. Then for  $A$  there exists a left eigenvector  $e$  with nonnegative coordinates  $e_i$ ,  $\sum_i e_i = 1$ , such that

$$eA = \lambda e, \quad A^*e = \lambda e.$$

Hence for every  $i$  we have

$$\sum_j a_{ji} e_j = \lambda e_i, \quad (8.1)$$

which leads to the equality

$$\sum_j \frac{a_{ji} e_j}{\lambda e_i} = 1$$

for every  $i$ . Hence a matrix of the form

$$P = \left( p_{ij} = \frac{a_{ji} e_j}{\lambda e_i} \right)$$

is the stochastic matrix for which vector  $e$  is a stationary distribution:

$$eP = e.$$

The distribution on edges  $i \rightarrow j$  defined by

$$m_{ij} = p_{ij}e_i = \frac{a_{ji}e_j}{\lambda}$$

is the flow  $m$  on the graph  $G$  such that the measure  $m_i$  of the vertex  $i$  equals  $e_i$ . The entropy of  $m$  is calculated by the formula

$$h_m = - \sum_{ij} m_{ij} \ln m_{ij} + \sum_i m_i \ln m_i.$$

Hence

$$h_m = - \sum_{ij} \frac{a_{ji}e_j}{\lambda} \ln \frac{a_{ji}e_j}{\lambda} + \sum_i e_i \ln e_i.$$

Here we assume that  $0 \ln 0 = 0$ . That means that the sum is taken over  $i, j$  for which  $a_{ij} = 1$ . Thus we obtain

$$\begin{aligned} h_m &= - \sum_{ij} \frac{a_{ji}e_j}{\lambda} (\ln a_{ij} + \ln e_i - \ln \lambda) + \sum_i e_i \ln e_i = \\ &= \sum_i \left( \sum_j \frac{a_{ji}e_j}{\lambda} \right) \ln \lambda - \sum_i \left( \sum_j \frac{a_{ji}e_j}{\lambda} \right) \ln e_i + \sum_i e_i \ln e_i = \\ &= \ln \lambda \sum_i e_i - \sum_i e_i \ln e_i + \sum_i e_i \ln e_i = \ln \lambda. \end{aligned}$$

□

## References

1. *Alekseev, V.* Symbolic dynamics, 11 mathematical school, Institut matematiki AN SSSR, Kiev, 1976. (in Russian)
2. *Ampilova, N., Soloviev, I.* Entropies in investigation of dynamical systems and their application to digital image analysis. Journal of Measurements in Engineering, 6, No. 2, 107-118 (2018). <https://doi.org/10.21595/jme.2018.19891>
3. *Boltzmann, L.* Sitzungsber. Kaiserl. Akad. Wiss. 66 (2) 275 (1872).
4. *Clausius, R.* Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie. Annalen der Physik und Chemie 125, No.7, 353-400 (1865).
5. *Dinaburg, E. I.* On the relations among various entropy characteristics of dynamical systems. Math. USSR, Izv. 5(1971), 337-378 (1972).
6. *Downarowicz, T.* Entropy in dynamical systems. Cambridge: CUP, 2011.
7. *Gantmacher, F. R.* The theory of matrices. Vol. 1, 2. Transl. from the Russian by K. A. Hirsch. Reprint of the 1959 translation. Providence, RI: AMS Chelsea Publishing, 1998.
8. *Goodman, T.* Relating topological entropy and measure theoretic entropy. Bulletin of London Math. Soc. 3, No. 2, 176-180 (1971).



9. *Hsu, C. S.* Cell-to-Cell Mapping, N.Y.: Springer-Verlag, 1987.
10. *Katok, A., Hasselblat, B.* Introduction to the Modern Theory of Dynamical Systems, Cambridge: Cambridge University Press, 1995.
11. *Kolmogorov, A.* New metric invariants of transitive dynamical systems and automorphisms of the Lebesgue spaces, Doklady AN SSSR, 119, No.5, 861–864 (1958).
12. *Kryloff, N., Bogoluboff, N.* La theorie generale de la mesure dans son application a l'etude des systemes dynamiques de la mecanique non lineaire, Ann. Math., 38, No.1, 65–113 (1937).
13. *Lind, D., Marcus, B.* An introduction to symbolic dynamics and coding. Cambridge: Cambridge University Press, 1995.
14. *Nemytskii, V. V. and Stepanov, V. V.* Qualitative Theory of Dynamical Systems, Princeton University Press, 1960.
15. *Osipenko, G.* On symbolic image of a dynamical system. Kraevie zadachi (Sb. trudov), Perm (pp. 101–105) 1983. (in Russian)
16. *Osipenko, G.* Dynamical systems, Graphs, and Algorithms. Lectures Notes in Mathematics, 1889, Berlin: Springer, 2007.
17. *Osipenko, G.* Symbolic images and invariant measures of dynamical systems. Ergodic Theory and Dynamical Systems 30, 1217–1237 (2010).
18. *Prasolov, V. V.* Elements of Combinatorial and Differential Topology. Graduate Studies in Mathematics, Vol. 74, 2006.
19. *Shannon, C.* A Mathematical Theory of Communication. Bell System Technical Journal, 27, No.3, 379–423 (July 1948). doi:10.1002/j.1538-7305.1948.tb01338.x. hdl:11858/00-001M-0000-002C-4314-2.
20. *Shannon, C.* A Mathematical Theory of Communication. Bell System Technical Journal, 27, No.4, 623–666 (October 1948). doi:10.1002/j.1538-7305.1948.tb00917.x. hdl:11858/00-001M-0000-002C-4314-2.

Получена 02.06.2019