Acta Universitatis Carolinae. Mathematica et Physica

Miloslav Nosál

Stationary non-ergodic sources

Acta Universitatis Carolinae. Mathematica et Physica, Vol. 3 (1962), No. 1, 39--54

Persistent URL: `http://dml.cz/dmlcz/142146`

1962 ACTA UNIVERSITATIS CAROLINAE MATHEMATICA ET PHYSICA NO. 1 PAG. 39—54

STATIONARY NON-ERGODIC SOURCES

MILOSLAV NOSÁL

Department of Mathematical Statistics at the Charles University, Praha

[Received 17. VIII. 1961

In this paper the special type of stationary non-ergodic sources is studied and fundamental theorems on transmission for this type of sources are proved. This work proceed from Winkelbauer's work [5].

INFORMATION SOURCES

Throughout the whole paper the set of all natural numbers will be denoted by N and the set of all integers will be denoted by I. If S is a finite non-empty set then the symbol S^I designates the set of all mappings of I into S, i. e. the set of all sequences of elements in S which are infinite to both sides; if $x \in S^I$ then we shall write x_i instead of $x(i)$ for any $i \in I$. The σ-algebra of subsets of S^I generated by the class of all sets of the form

$$\{x : x \in S^I,\ x_i = a\},\ i \in I,\ a \in S$$

will be denoted by $\mathbf{F}_S$. If μ is a probability measure on the σ-algebra $\mathbf{F}_S$, then the couple $[S, \mu]$ will be called a information source or shortly source. Let T_S be the transformation of S^I defined by the equation

$$(T_S x)_i = x_{i+1},\ x \in S^I,\ i \in I$$

If $\mu(T_S E) = \mu(E)$ holds for every $E \in \mathbf{F}_S$, then the source $[S, \mu]$ is said to be stationary. If further

$$E \in \mathbf{F}_S,\ T_S E = E,\ \mu(E) > 0 \Rightarrow \mu(E) = 1$$

holds, then the source $[S, \mu]$ is said to be ergodic.

LEMMA 1. Let $[S, \mu^{(i)}]$, $i = 1, 2, \ldots, k$ be k different ergodic sources. Then the measures $\mu^{(i)}$ and $\mu^{(j)}$ $(i \neq j)$ are singular and there exists the measurable partition $\mathscr{E} = (E_1, E_2, \ldots, E_k)$ of S^I such that

$$\mu^{(i)}(E_j) = \delta_{ij},\quad i, j = 1, 2, \ldots, k \text{ (Kronecker's symbol).}$$

Note: The sources $[S, \mu]$ and $[S, \nu]$ are said to be different if there exists a set $E \in \mathbf{F}_S$, such that $\mu(E) \neq \nu(E)$.

Proof: It follows from Birkhoff's ergodic theorem, that the following relation holds for any $\mu^{(i)}$-integrable, $i = 1, 2, \ldots, k$, function $f(x)$ *on* S^I:

$$(1) \qquad \mu^{(i)}\left\{x : x \in S^I, \ \frac{1}{n}\sum_{l=0}^{n-1} f(T_S^l x) \to \int f d\mu^{(i)}\right\} = 1, \quad i = 1, 2, \ldots, k.$$

Since the sources $[S, \mu^{(i)}]$ are different, there exist the sets $E_{ij} \in \mathbf{F}_S$, $i, j = 1, 2, \ldots, k$, $i < j$ such that

$$\mu^{(i)}(E_{ij}) \neq \mu^{(j)}(E_{ij}).$$

Now we put in (1) $f_{ij} = \varkappa_{E_{ij}}$ (characteristic function of the set E_{ij}).

Then $$\int f_{ij} d\mu^{(i)} = \mu^{(i)}(E_{ij}), \quad \int f_{ij} d\mu^{(j)} = \mu^{(j)}(E_{ij})$$

Further, the sets

$$F_{ij} = \left\{x : x \in S^I, \ \frac{1}{n}\sum_{l=0}^{n-1} f_{ij}(T_S^l x) \to \mu^{(i)}(E_{ij})\right\}$$

$$F_{ji} = \left\{x : x \in S^I, \ \frac{1}{n}\sum_{l=0}^{n-1} f_{ij}(T_S^l x) \to \mu^{(j)}(E_{ij})\right\}$$

are disjoint for any pair $i, j = 1, 2, \ldots, k$, $i < j$ and

$$\mu^{(i)}(F_{ij}) = sj_1, \quad \mu^{(j)}(F_{ij}) = 0, \quad i, j = 1, 2, \ldots, k, \ i \neq j.$$

If now we put

$$E_i' = \bigcap_{\substack{j=1 \\ j \neq i}}^{k} F_{ij}, \qquad i = 1, 2, \ldots, k$$

then it is clear that

$$\mu^{(i)}(E_i') = 1, \qquad i = 1, 2, \ldots, k$$

Further

$$\mu^{(j)}(E_i') = \mu^{(j)}\Big(\bigcap_{\substack{l=1 \\ l \neq i}}^{k} F_{il}\Big) \leqq \mu^{(j)}(F_{ij}) = 0, \qquad i, j = 1, 2, \ldots, k, \ i \neq j$$

and

$$E_i' \cap E_j' = \bigcap_{\substack{l=1 \\ l \neq i}}^{k} F_{il} \cap \bigcap_{\substack{m=1 \\ m \neq j}}^{k} F_{jm} \subset F_{ij} \cap F_{ji} = \theta \text{ (empty set)}, \quad i, j = 1, 2, \ldots, k, \ i \neq j$$

If we put

$$E_i' = E_i, \ i = 1, 2, \ldots, k-1, \qquad E_k = S^I - \bigcup_{i=1}^{k-1} E_i,$$

then the partition $\mathscr{E} = (E_1, E_2, \ldots, E_k)$ has the property from the assertion of Lemma 1. q.e.d.

LEMMA 2. Let $[S, \mu^{(i)}]$, $i = 1, 2, \ldots, k$ $(k > 1)$ be k different ergodic sources. Let $\alpha_1, \alpha_2 \ldots, \alpha_k$ be k real numbers such that

$$\sum_{i=1}^{k} \alpha_i = 1, \qquad \alpha_i > 0, \qquad i = 1, 2, \ldots, k$$

Further we put

$$\mu = \sum_{i=1}^{k} \alpha_i \mu^{(i)}$$

Then the source $[S, \mu]$ is stationary but non-ergodic.

Proof: If $E \in \mathbf{F}_S$ then obviously

$$\mu(T_S E) = \sum_{i=1}^{k} \alpha_i \mu^{(i)}(T_S E) = \sum_{i=1}^{k} \alpha_i \mu^{(i)}(E) = \mu(E)$$

In accordance with LEMMA 1 there exists the partition $\mathscr{E} = (E_1, E_2, \ldots, E_k)$ such that

$$\mu^{(i)}(E_j) = \delta_{ij}, \qquad i, j = 1, 2, \ldots, k.$$

Now we shall investigate the set $E = \bigcup_{i=-\infty}^{\infty} T_S^i E_1$.

$$\mu^{(1)}(E) = \mu^{(1)}\Big(\bigcup_{i=-\infty}^{\infty} T_S^i E_1\Big) \geq \mu^{(1)}(E_1) \quad \Rightarrow \quad \mu^{(1)}(E) = 1$$

$$\mu^{(j)}(E) = \mu^{(j)}\Big(\bigcup_{i=-\infty}^{\infty} T_S^i E_1\Big) \leq \sum_{i=-\infty}^{\infty} \mu^{(j)}(T_S^i E_1) = \sum_{i=-\infty}^{\infty} \mu^{(j)}(E_1) = 0, \qquad j = 2, 3, \ldots, \text{k}.$$

It is clear that

$$T_S E = E.$$

But

$$\mu(E) = \sum_{i=1}^{k} \alpha_i \mu^{(i)}(E) = \alpha_1 \qquad \text{q. e. d.}$$

Let $[S, \mu]$ be a source. For any $n \in N$, let S^n be the set of all "ordered n-tuples" of the elements from S. We define the probability measure on the σ-algebra of all subsets of S^n by the relation

$$\mu_n(E) = \mu\{x : x \in S^I, (x_0, x_1, \ldots, x_{n-1}) \in E\}, \qquad E \subset S^n$$

For every $n \in N$ we denote

$$H_n = -\sum_{\mathbf{x} \in S^n} \mu_n(\mathbf{x}) \log \mu_n(\mathbf{x})$$

where logarithm is taken to the base 2. If a real number h exists, such that sequence

$$\frac{1}{n} H_n, \qquad n = 1, 2, \ldots$$

converges to h, then the number h is called the entropy of the source $[S, \mu]$ and is denoted by $H[S, \mu]$. (It is easy to show that the entropy of a stationary source exists.)

Let $[S, \mu]$ be an information source and for every $n \in N$ let π_n be the mapping of the set $\{1, 2, \ldots, s^n\}$ (where s is the number of elements in S i. e. $s =$ card S) on S^n so that

$$\mu_n(\pi_n(j)) \geq \mu_n(\pi_n(j + 1)).$$

If ε is a real number, $0 < \varepsilon < 1$, then the number $L_n([S, \mu], \varepsilon)$, where

$$L_n([S, \mu], \varepsilon) = \min \left\{k : \sum_{j=1}^{k} \mu_n(\pi_n(j)) > 1 - \varepsilon\right\}$$

is called the n-dimensional ε-length of the source $[S, \mu]$.

Lemma 3. Let $[S, \mu^{(i)}]$, $i = 1, 2, \ldots, k$ be k sources. Let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be k real numbers such that

$$\alpha_i \geq 0,\ i = 1, 2, \ldots, k, \quad \sum_{i=1}^{k} \alpha_i = 1.$$

Further let

$$\mu = \sum_{i=1}^{k} \alpha_i \mu^{(i)}.$$

Then for every $n \in N$, $0 < \varepsilon < 1$ the following inequality holds

$$L_n([S, \mu], \varepsilon) \leq \sum_{i=1}^{k} L_n([S, \mu^{(i)}], \varepsilon).$$

Proof: Let $\pi_{jn}, j = 1, 2, \ldots, k, n \in N$ be the mapping of the set $\{1, 2, \ldots, s^n\}$ (where $s =$ card S) on S^n such that

$$\mu_n^{(j)}(\pi_{jn}(l)) \geq \mu_n^{(j)}(\pi_{jn}(l + 1)).$$

Then for any $n \in N$, $0 < \varepsilon < 1$

$$\mathrm{L}_n[(S, \mu^{(j)}], \varepsilon) = \min\left\{m : \sum_{l=1}^{m} \mu_n^{(j)}(\pi_{jn}(l)) > 1 - \varepsilon\right\}, j = 1, 2, \ldots, \mathrm{k}.$$

Now we form the sets

$$M_j = \{\mathrm{x} : \mathrm{x} \in S^n,\ \pi_{jn}(l) = \mathrm{x},\ l = 1, 2, \ldots,\ L_n([S, \mu^{(j)}], \varepsilon)\},\ j = 1, 2, \ldots, k.$$

Obviously

$$\text{card } M_j = L_n([S, \mu^{(j)}], \varepsilon),$$

$$\mu^{(j)}(M_j) > 1 - \varepsilon, \quad j = 1, 2, \ldots, k$$

Let

$$M = \bigcup_{j=1}^{k} Mj.$$

$$\mu_n(M) = \sum_{j=1}^{k} \alpha_j \mu_n^{(j)}(M) \geq \sum_{j=1}^{k} \alpha_j \mu_n^{(j)}(M_j) > 1 - \varepsilon$$

In accordance with LEMMA 3, § 8 in [5]

$$\text{card } M \geq \mathrm{L}_n([S, \mu], \varepsilon).$$

But

$$\text{card } M \leq \sum_{j=1}^{k} \text{card } M_j = \sum_{j=1}^{k} L_n([S, \mu^{(j)}], \varepsilon).$$

Then

$$L_n([S, \mu], \varepsilon) \leq \sum_{j=1}^{k} L_n([S, \mu^{(j)}], \varepsilon), \; n \in N, \; 0 < \varepsilon < 1 \qquad \text{q.e.d.}$$

COMMUNICATION CHANNELS

Let A and B be a finite non-empty sets and let ν be a real valued function on the Cartesian product $\mathrm{F}_B \times A^I$ satisfying the following properties:

a). For any $x \in A^I$, the set function $\nu(., x)$ is the probability measure on the σ-algebra F_B.

b). For any set $E \in \mathrm{F}_B$, the point function $\nu(E, .)$ is F_A-measurable. Then the triple $[A, \nu, B]$ is called the communication channel. If further next condition is satisfied:

c). $\nu(T_B E, T_A x) = \nu(E, x)$ for any $E \in \mathrm{F}_B$, $x \in A^I$,
then the channel $[A, \nu, B]$ is said stationary.

Let $[A, \nu, B]$ be a channel. If there exists a non-negative integer m such that, for any

$$x \in A^I, \; x' \in A^I, \; n \in N, \; i \in I, \; \mathrm{b} \in B^n$$

the relations

$$x_{i+j} = x'_{i+j}, \qquad j = -m, -m+1, \ldots, -1, 0, 1, \ldots, n-1$$

imply that

$$\nu(V_{in}(\mathrm{b}), x') = \nu(V_{in}(\mathrm{b}), x)$$

where

$$V_{in}(\mathrm{b}) = \{y : y \in B^I, \quad y_{i+j} = b_j, \; j = 0, 1, \ldots, n-1\},$$

then the channel $[A, \nu, B]$ is called a channel with finite past history; the least non-negative integer m, for which the above conditon is satisfied will be called the length of the finite past history. We emphasize that the concept of the channel with finite past history, introduced and investigated first in [5], is more general than the concept of the channel with finite memory.

Let $[A, \nu, B]$ be a channel with finite past history of length m and let s be an integer, $s \geq m$. If $n \in N$, then for any $\mathbf{u} \in A^{S+n}$ we define the probability measure $\nu_n(.|\mathbf{u})$ on the σ-algebra of all subsets of B^n by the relation

$$\nu_n(E|\mathbf{u}) = \nu(\{y : y \in B^I, (y_0, y_1, \ldots, y_{n-1}) \in E\}, x), \text{ where } x \in V_{-s, n+s}(\mathbf{u}).$$

Let $[A, \nu, B]$ be a channel with finite past history of length m and let $n \in N$, $0 < \varepsilon < 1$. The set of all mappings of B^n into A^{m+n} will be denoted by $K(B, A, m + n)$. Then the number $S_n([A, \nu, B], \varepsilon)$, where

$$S_n([A, \nu, B], \varepsilon) = \max \{\text{card}\{\mathbf{x} : \mathbf{x} \in A^{m+n}\ \nu(\psi^{-1}(\mathbf{x})|\mathbf{x}) > 1 - \varepsilon\} :$$
$$: \psi \in K(B, A, n, m + n)\}$$

is called an n-dimensional ε-size of the channel $[A, \nu, B]$. Further we define the number $C[A, \nu, B]$ by the relation

$$C[A, \nu, B] = \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log S_n([A, \nu, B], \varepsilon)$$

(the existence of the limit is established). This number is called a capacity of the channel $[A, \nu, B]$.

TRANSMISSION

Let S be a finite non-empty set. If n is a natural number then a non-negative finite real-valued function w_n on the Cartesian product $S^n \times S^n$ is called an n-dimensional weight function on S. An n-dimensional weight function w_n on S is called regular if, for any $\mathbf{z} \in S^n$

$$w_n(\mathbf{z}, \mathbf{z}) = 0.$$

The sequence $w = \{w_n\}$, where for any $n \in N$ w_n is an n-dimensional weight function on S will be called a weight function on S. If a weight function w consists of a regular n-dimensional weight functions only, then w is said to be a regular weight function on S. A weight function w on S is said to be bounded by a real number t if

$$w_n(\mathbf{z}, \mathbf{z}') \leq t \text{ for } \mathbf{z}, \mathbf{z}' \in S^n, n \in N\,.$$

If w is a weight function on S, we shall define the fuction v_n on the space $S^I \times S^I$ by the relation

$$v_n(x, x') = w_n((x_0, x_1, \ldots, x_{n-1}), (x'_0, x'_1, \ldots, x'_{n-1})), x, x' \in S^I\,.$$

If we define, for $n \in N$, the n-dimensional weight function w^e_n on S by the equation

$$w^e_n(\mathbf{x}, \mathbf{x}') = 0 \text{ for } \mathbf{x} = \mathbf{x}', \mathbf{x}, \mathbf{x}' \in S^n$$

$$w^e_n(\mathbf{x}, \mathbf{x}') = 1 \text{ for } \mathbf{x} \neq \mathbf{x}', \mathbf{x}, \mathbf{x}' \in S^n$$

then the sequence $w^e = \{w^e_n\}$ will be called the error weight function. The errror weight function is regular and bounded by the number 1.

If we define, for $n \in N$, the n-dimensional weight function w^f_n on S by the relation

$$w^f_n(\mathbf{x}, \mathbf{x}') = \frac{1}{n} \sum_{i=1}^{n} w^f_1(x_i, x_i'), \ \mathbf{x}, \mathbf{x}' \in S^n,$$

then the sequence $w^f = \{w^f_n\}$ will be called the frequency weight function. The frequency weight function is regular and bounded by the number 1.

It is clear that

$$(2) \qquad (n + s)\, w^f_{n+s}((\mathbf{u}, \mathbf{v}), (\mathbf{u}', \mathbf{v}')) = n . w^f_n(\mathbf{u}, \mathbf{u}') + s . w^f_s(\mathbf{v}, \mathbf{v}'),$$
$$\mathbf{u}, \mathbf{u}' \in S^n, \ \mathbf{v}, \mathbf{v}' \in S^s.$$

Let A and B be two finite sets and $n, s \in N$. A mapping $\varkappa$ from A^n into B^s is called a code of type (n, s) from A into B. If the mapping $\varkappa$ is one-to-one, then the code $\varkappa$ is also said to be one-to-one. We denote the set of all codes from A into B of type (n, s) by $K(A, B, n, s)$.

If $m \in I$, $m \geq 0$, $n, s \in N$, $m < n$, $\varkappa \in K(A, B, n, s)$ and if

$$\varkappa(\mathbf{x}', \mathbf{x}) = \varkappa(\mathbf{x}'', \mathbf{x}), \ \mathbf{x} \in A^{n-m}$$

for any $\mathbf{x}''$, $\mathbf{x}' \in A^m$, then we shall say that the code $\varkappa$ does not distinguish the first m symbols. Then the code $\varkappa' \in K(A, B, n - m, s)$ uniquely defined by

$$\varkappa'(\mathbf{x}) = \varkappa(\mathbf{x}', \mathbf{x}), \ \mathbf{x}' \in A^m, \ \mathbf{x} \in A^{n-m}$$

will be called the reduced form of the code $\varkappa$.

Let $\varkappa$ be a code of type (n, s) from A into B, where $n > 0$. We define the transformation τ of A^I into B^I by the equation

$$((\tau x)_{ks}, (\tau x)_{ks+1}, \ldots, (\tau x)_{ks+s-1}) = \varkappa(x_{kn}, x_{kn+1}, \ldots, x_{kn+n-1}), \ k \in I, \ x \in A^I .$$

The transformation τ is said to be associated with the code $\varkappa$.

Let $[S, \mu]$ be an information source, let $[A, \nu, B]$ be a communication channel and let w be a weight function on S. Further let $n, s \in N$, $\varkappa \in K(S, A, n, s)$, $\delta \in K(B, S, s, n)$. We shall define the risk $r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w)$ of length n with respect to the weight function w on S, corresponding to the transmission of the source $[S, \mu]$ over the channel $[A, \nu, B]$ with the encoder $\varkappa$ and decoder δ by the equation

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w) = \iint v_n(x, \varrho y)\, d\nu(y, \tau x)\, d\mu(x),$$

where τ and ϱ are the transformations associated with the codes $\varkappa$ and δ. (It is easy to show that the risk is defined by the above equation.) The minimum risk $r_{n,s}([S, \mu], [A, \nu, B]; w)$ will be defined by the relation

$$r_{n,s}([S, \mu], [A, \nu, B]; w) = \min \{r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w) :$$
$$: \varkappa \in K(S, A, n, s), \delta \in K(B, S, s, n)\} .$$

The risks with respect to w^e or w^f will be called the probability of error or the average frequency of errors, respectively.

It is clear that if $[A, \nu, B]$ is a channel with finite past history of length m, $[S, \mu]$ is a source, w is a weight function on S, $n, s \in N$, $\varkappa \in K(S, A, n, s)$ $\delta \in K(B, S, s, n)$ then

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w) = \sum_{\mathbf{x} \in S^n} \sum_{\mathbf{y} \in B^s} w_n(\mathbf{x}, \delta\mathbf{y})\, \nu_s(\mathbf{y} \mid \varkappa\mathbf{x}) \mu_n(\mathbf{x}) .$$

If even $[S, \mu]$ and $[A, \nu, B]$ are stationary, if δ does not distinguish the first m symboles and if δ' is a reduced form of the code δ, then clearly

$$(3) \qquad r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w] = \sum_{\mathbf{x} \in S^n} \sum_{\mathbf{y} \in B^{s-m}} w_n (\mathbf{x}, \delta'\mathbf{y})\, \nu_{s-m}(\mathbf{y} \mid \varkappa\mathbf{x}) \mu_n(\mathbf{x}) .$$

We shall say that the source $[S, \mu]$ is transmissible over the channel $[A, \nu, B]$ (with respect to the weight function w on S) if, for every positive real number ε, there is a natural number n such that

$$r_{n,}{}^{n}([S, \mu], [A, \nu, B]; w) < \varepsilon$$

We shall say that the source $[S, \mu]$ is strictly transmissible over the channel $[A, \nu, B]$ (with respect to the weight function w on S) if, for every positive real number ε, there is a positive integer n_0 such that, for any natural number $n \geq n_0$ the inequality

$$r^{n}{}_{,n}([S, \mu], [A, \nu, B]; w) < \varepsilon$$

holds.

We shall say that the source $[S, \mu]$ is not transmissible over the channel $[A, \nu, B]$ (with respect to the weight function w on S) if there exists a positive real number ε such that

$$\inf_{n \in N} r_{n,n}([S, \mu], [A, \nu, B]; w) \geq \varepsilon .$$

LEMMA 4. Let $[S, \mu^{(i)}]$ be sources, $i = 1, 2, \ldots, k$. Let $\alpha, \alpha_2, \ldots, \alpha_k$ be non-negative real numbers such that

$$\sum_{i=1}^{k} \alpha_i = 1 .$$

Let

$$\mu = \sum_{i=1}^{k} \alpha_i \mu^{(i)} .$$

Further let $[A, \nu, B]$ be a stationary channel with finite past history, w a weight function on S, $n, s \in N$.

Then

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w) = \sum_{i=1}^{k} \alpha_i r_n ([S, \mu^{(i)}], [A, \nu, B], \varkappa, \delta; w)$$

$$r_{n,s}([S, \mu], [A, \nu, B]; w) \geq \sum_{i=1}^{k} \alpha_i r_{n,s}([S, \mu^{(i)}], [A, \nu, B]; w).$$

Proof: By definition

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w) = \iint v_n(x, \varrho y)\, d\nu(y, \tau x)\, d\mu(x),$$

where τ is the transformation associated with the code $\varkappa$ and ϱ is the transformation associated with the code δ. Let $f(x)$ be a function on S^I defined by the relation

$$f(x) = \int v_n(x, \varrho y)\, d\nu(y, \tau x).$$

It is clear that $f(x)$ is a simple function and cosequently is integrable. Obviously

$$\int f(x)\, d\mu(x) = \sum_{i=1}^{k} \alpha_i \int f(x) d\mu^{(i)}(x)$$

and then

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w) = \sum_{i=1}^{k} \alpha_i r_n([S, \mu^{(i)}], [A, \nu, B], \varkappa, \delta; w).$$

Further

$$\inf \{r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w) : \varkappa \in K(S, A, n, s), \delta \in K(B, S, s, n)\} =$$
$$= r_{n,s}([S, \mu], [A, \nu, B]; w) \geq$$
$$\geq \sum_{i=1}^{k} \alpha_i \inf \{r_n([S, \mu^{(i)}], [A, \nu, B], \varkappa, \delta; w) : \varkappa \in K(S, A, n, s), \delta \in K(B, S, s, n)\} =$$
$$= \sum_{i=1}^{k} \alpha_i r_{n,s}([S, \mu^{(i)}], [A, \nu, B]; w).$$

q. e. d.

Now we can proceed to the proofs of the fundamental theorems.

THEOREM 1. Let $[A, \nu, B]$ be a stationary channel with finite past history. Let $[S, \mu^{(i)}]$, $i = 1, 2, \ldots, k$ be different ergodic sources such that

$$\max_{1 \leq i \leq k} H[S, \mu^{(i)}] < C[A, \nu, B].$$

Let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be positive real numbers for which $\sum_{i=1}^{k} \alpha_i = 1$.

Further let

$$\mu = \sum_{i=1}^{k} \alpha_i \mu^{(i)}$$

and let w be a regular bounded weight function on S. Then the source $[S, \mu]$ is strictly transmissible over the channel $[A, \nu, B]$ (with respect to the weight function w on S). Especially the source $[S, \mu]$ is strictly transmissible over the channel $[A, \nu, B]$ with respect to the error weight function and with respect to the frequency weight function.

Proof: In accordance with Theorem 4, § 16 in [5] it is sufficient to prove that the source $[S, \mu]$ is strictly transmissible with respect to the error weight function.

We denote

$$H = \max_{1 \le i \le k} H[S, \mu^{(i)}] \quad .$$

According to LEMMA 3 if

$$0 < \varepsilon < 1, \ n \in N$$

then

$$L_n([S, \mu], \varepsilon) \le \sum_{i=1}^{k} L_n([S, \mu^{(i)}], \varepsilon) .$$

Since the sources $[S, \mu^{(i)}]$, $i = 1, 2, \ldots, k$ are ergodic then according to Theorem 4, § 7 in [5] and Lemma 5, § 8 in [5]

$$\frac{1}{n} \log L_n([S, \mu^{(i)}], \varepsilon) \to H[S, \mu^{(i)}], \ i = 1, 2, \ldots, k$$

and consequently for any $\varepsilon > 0$ there exists n_{i0} natural such that for $n \ge n_{i0}$

$$\frac{1}{n} \log L_n([S, \mu^{(i)}), \varepsilon] < H[S, \mu^{(i)}] + \varepsilon < H + \varepsilon , \ i = 1, 2, \ldots, k .$$

Now we denote $n_0 = \max_{1 \le i \le k} n_{i0}$. Obviously for any natural number $n \ge n_0$ is

$$L_n([S, \mu^{(i)}], \varepsilon) < 2^{n(H+\varepsilon)}, \quad i = 1, 2, \ldots, k .$$

Consequently for any $\varepsilon > 0$ there exists a natural number n_0 such that for every natural number $n \ge n_0$ is

(4) $$L_n([S, \mu], \varepsilon) < k \, . \, 2^{n(H+\varepsilon)} .$$

Let a real number ε, $0 < \varepsilon < 1$ be given. We choose $\lambda > 0$ and $0 < \varepsilon' < 1$ so that

$$H - C[A, \nu, B] + \varepsilon' + \lambda = \beta < 0 , \quad 2\varepsilon' < \varepsilon .$$

Let m be the length of finite past history of the channel [A, ν, B]. We put

$$h = 2^{-m(C[A, \nu, B] - \lambda)} .$$

According to (4) there exists a natural number n_1 such that for any $n \ge n_1$

$$L_n([S, \mu], \varepsilon') < k \, . \, 2^{n(H+\varepsilon')} .$$

According to Theorem 1, § 17 in [5] there exists a natural number n_2 such that for any $n \ge n_2$

$$S_{n-m}([A, \nu, B] \varepsilon') > 2^{(n-m)(C[A, \nu, B] - \lambda)} = h \, . \, 2^{n(C[A, \nu, B] - \lambda)} .$$

We denote $n_3 = \max(n_1, n_2)$. Then for any natural number $n \geq n_3$

$$\frac{L_n([S, \mu], \varepsilon')}{S_{n-m}([A, \nu, B], \varepsilon')} < \frac{k \,.\, 2^{n(H+\varepsilon)}}{h \,.\, 2^{n(C[A,\nu,B]-\lambda)}} = \frac{k}{h} 2^{n\beta}$$

Since $\beta < 0$ there exists a natural number n_0 such that for any $n > n_0$

$$L_n([S, \mu], \varepsilon') < S_{n-m}([A, \nu, B], \varepsilon') .$$

According to Theorem 7, § 17 in [5] there exists a one-to-one code $\varkappa \in K(S, A, n, n)$ and code $\delta \in K(B, S, n, n)$ so that

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w^f) < 2\varepsilon' < \varepsilon \qquad \text{q.e.d.}$$

THEOREM 2. Let $[A, \nu, B]$ be a stationary channel with finite past history such that $C[A, \nu, B] > 0$. Let $[S, \mu^{(i)}]$, $i = 1, 2, \ldots, k$ be ergodic sources such that

$$\max_{1 \leq i \leq k} H[S, \mu^{(i)}] = C[A, \nu, B] .$$

Let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be real positive numbers such that $\sum_{i=1}^{k} \alpha_i = 1$.

Further let

$$\mu = \sum_{i=1}^{k} \alpha_i \mu^{(i)} .$$

Then the source $[S, \mu]$ is strictly transmissible over the channel $[A, \nu, B]$ with respect to the frequency weight function w^f.

Proof: First of all we prove then the source $[S, \mu]$ is transmissible over the channel $[A, \nu, B]$ with respect to the frequency weight function w^f. In accordance with Theorem 3, § 16 in [5] the source $[S, \mu]$ shall be strictly transmissible over the channel $[A, \nu, B]$ with respect to the frequency weight function w^f.

Let $\varepsilon' > 0$ be given. We must choose a natural number n and codes $\varkappa \in K(S, A, n, n)$, $\delta \in K(B, S, n, n)$ such that

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w^f) < \varepsilon' .$$

We denote $H = \max_{1 \leq i \leq k} H[S, \mu^{(i)}]$ and let m be a lenght of finite past history of the channel $[A, \nu, B]$.

Now we choose $0 < \varepsilon < 1$ and $\lambda > 0$ such that

$$3\varepsilon > \varepsilon' , \quad \lambda < H, \quad \frac{2\lambda}{\lambda + H} < \frac{1}{4} \varepsilon' .$$

In accordance with Theorem 1, § 17 in [5] there exists a natural number p_0 such that for any $p \geq p_0$

$$S_p([A, \nu, B], \varepsilon) > 2^{p(H-\lambda)}$$

Further in accordance with (4) there exists a natural number q_0 such that for any $q \geq q_0$

$$L_q([S, \mu], \varepsilon) < k \,.\, 2^{q(H+\lambda)}$$

holds.

Now we choose fast p and q such that $p \geq p_0$, $q \geq q_0$, $q = \left[p \frac{H - \lambda}{H + \lambda} - \right.$ $\left. - \frac{\log k}{H + \lambda}\right]$.
Then

$$q \leq p \frac{H - \lambda}{H + \lambda} - \frac{\log k}{H + \lambda}$$

and cosequently

$$k \, . \, 2^{q(H+\lambda)} \leq 2^{p(H-\lambda)}$$

Since $p \geq p_0$, $q \geq q_0$ we have

$$L_q([S, \mu], \varepsilon) < S_p([A, \nu, B], \varepsilon) \, .$$

According to Theorem 7, § 17 in [5] there are codes $\varkappa' \in K(S, A, q, m + p)$, $\delta' \in K(B, S, m + p, q)$ where δ' does not distinguish the first m symbols such that

$$r_q([S, \mu], [A, \nu, B], \varkappa', \delta'; w^e) < 2\varepsilon \, .$$

Now we put $n = m + p$. Obviously $q \leq n$.
Further we define a code $\varkappa \in K(S, A, n, n)$ by the equation

$$\varkappa(\mathbf{z}) = \varkappa' \, (z_1, z_2, \ldots, z_q), \; \mathbf{z} = (z_1, z_2, \ldots, z_q, \ldots z_n) \in S^n \, .$$

and a code $\delta \in k(B, S, n, n)$ by the equation

$$\delta(\mathbf{x}) = (\delta'(\mathbf{x}), \mathbf{u}), \; \mathbf{x} \in B^n$$

where $\mathbf{u} \in S^{n-q}$ is a firm vector. We show that the codes $\varkappa$ and δ are suitable for our purpose.

The code δ does not distinguish the first m symbols. Let δ'' be the reduced form of the code δ. Obviously $\delta'' \in \varkappa(B, S, p, n)$.
In accordance with (2) and with (3)

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w^f) = \sum_{\mathbf{z} \in S^n} \sum_{\mathbf{x} \in B^p} w^f_n(\mathbf{z}, \delta''\mathbf{x}) \nu_p(\mathbf{x} \mid \varkappa\mathbf{z}) \mu_n(\mathbf{z}) =$$

$$= \sum_{\mathbf{z} \in S^q} \sum_{\mathbf{z}' \in S^{n-q}} \sum_{\mathbf{x} \in B^p} \left(\frac{q}{n} w^f_q(\mathbf{z}, \delta^*\mathbf{x}) + \frac{n-q}{n} w^f_{n-q}(\mathbf{z}', \mathbf{u}) \right) \nu_p(\mathbf{x} \mid \varkappa'\mathbf{z}) \mu_n(\mathbf{z}, \mathbf{z}')$$

where δ^* is the reduced form of the code δ'.

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w^f) = \frac{q}{n} \sum_{\mathbf{z} \in S^q} \sum_{\mathbf{x} \in B^p} w^f_q(\mathbf{z}, \delta^*\mathbf{x}) \nu_p(\mathbf{x} \mid \varkappa'\mathbf{z}) \mu_q(\mathbf{z}) +$$

$$+ \frac{n-q}{n} \sum_{\mathbf{z}' \in S^{n-q}} w^f_{n-q}(\mathbf{z}, \mathbf{u}) \mu_{n-q}(\mathbf{z}') \leq$$

$$\leq \frac{q}{n} \sum_{\mathbf{z} \in S^q} \sum_{\mathbf{x} \in B^p} w^e_q(\mathbf{z}, \delta^*\mathbf{x}) \nu_p(\mathbf{x} \mid \varkappa'\mathbf{z}) \mu_q(\mathbf{z}) + \frac{n-q}{n} =$$

$$= \frac{q}{n} \sum_{z \in S^q} \sum_{x \in B^n} w_q^e(z, \delta' x) \nu_n(x \mid \varkappa' z) \mu_q(z) + \frac{n-q}{n} =$$

$$= \frac{q}{n} r_q([S, \mu], [A, \nu, B], \varkappa', \delta'; w^e) + \frac{n-q}{n}$$

But $\frac{n-q}{n} = \frac{m+p-q}{m+p} < \frac{1}{3} \varepsilon'$, $\frac{q}{n} < 1$, $r_q([S, \mu], [A, \nu, B], \varkappa', \delta'; w^e) < 2\varepsilon$ and consequently

$$r_n([S, \mu], [A, \nu, B], \varkappa, \delta; w^f) < 2\varepsilon + \frac{1}{3} \varepsilon' < \varepsilon'$$

q. e. d.

THEOREM 3. Let $[A, \nu, B]$ be a stationary channel with finite past history. Let $[S, \mu^{(i)}]$, $i = 1, 2, \ldots, k$ be argodic sources such that

$$\max_{1 \leq i \leq k} H[S, \mu^{(i)}] > C[A, \nu, B] .$$

Let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be positive real numbers such that $\sum_{i=1}^{k} \alpha_i = 1$.

Further let

$$\mu = \sum_{i=1}^{k} \alpha_i \mu^{(i)}$$

Then the source $[S, \mu]$ is not transmissible over the channel $[A, \nu, B]$ neither with respect to the error weight function w^e nor with respect to the frequency weight function w^f.

Proof: Let $[S, \mu^{(l)}]$ be the source for which $H[S, \mu^{(l)}] = \max_{1 \leq i \leq k} H[S, \mu^{(i)}]$. Since $H[S, \mu^{(l)}] > C[A, \nu, B]$, there exists (according to Theorem 7, § 19 in [5]) $\varepsilon' > 0$ such that

$$\inf_{n \in N} r_{n,n}([S, \mu^{(l)}], [A, \nu, B]; w^e) \geq \varepsilon' .$$

By assumption $\alpha_l > 0$. We denote $\varepsilon = \alpha_l \varepsilon' > 0$.
According to LEMMA 4 we have

$$r_{n,n}([S, \mu], [A, \nu, B]; w^e) \geq \sum_{i=1}^{k} \alpha_i r_{n,n}([S, \mu^{(i)}], [A, \nu, B]; w^e)$$

Then

$$\inf_{n \in N} r_{n,n}([S, \mu], [A, \nu, B]; w^e) \geq \sum_{i=1}^{k} \alpha_i \inf r_{n,n}([S, \mu^{(i)}], [A, \nu, B]; w^e) \geq$$

$$\geq \alpha_l \inf_{n \in N} r_{n,n}([S, \mu^{(l)}], [A, \nu, B]; w^e) \geq \varepsilon .$$

Second part of the proof we obtain if we write w^f instead w^e. q. e. d.

THEOREM 4. Let ε and K be arbitrary real numbers, $\varepsilon > 0$. Then there exists a stationary source $[S, \mu]$ with the following property:

1. $$H[S, \mu] = \varepsilon .$$

2. If $[A, \nu, B]$ is any stationary channel with finite past history such that

$$C[A, \nu, B] \leq K$$

then the source $[S, \mu]$ is not transmissible over the channel $[A, \nu, B]$ neither with respect to the error weight function w^e nor with respect to the frequency weight function w^f.

Proof: If $K < \varepsilon$ then the assertion is the consequence of Theorem 3, § 19 and of Lemma 7, § 6 and Lemma 3, § 7 in [5]. Then let be $K \geq \varepsilon$. We choose real numbers t_1, t_2 such that

$$0 < t_1 < \varepsilon , \quad K < t_2 < + \infty .$$

According to Lemma 7, § 6 and Lemma 3, § 7 in [5] there exist ergodic sources $[S, \mu^{(1)}]$ and $[S, \mu^{(2)}]$ such that

$$H[S, \mu^{(1)}] = t_1 \quad H[S, \mu^{(2)}] = t_2 .$$

Now we choose $\alpha > 0$ and $\beta > 0$ such that $\alpha t_1, + \beta t_2 = \varepsilon$, $\alpha + \beta = 1$.
Further we define the source $[S, \mu]$ by the equation

$$\mu = \alpha \mu^{(1)} + \beta \mu^{(2)}.$$

According to LEMMA 2 the source $[S, \mu]$ is stationary and according to Theorem 6, § 6 in (5)

$$H[S, \mu] = \alpha t_1 + \beta t_2 = \varepsilon .$$

Since

$$\max_{1 \leq i \leq 2} H[S, \mu^{(i)}] = H[S, \mu^{(2)}] = t_2 > K \geq C[A, \nu, B] ,$$

then the source $[S, \mu]$ is not transmissible over the channel $[A, \nu, B]$ neither with respect to the error weight function w^e nor with respect to the frequency weight function w^f (the consequence of THEOREM 6).

q. e. d.

COROLLARY. Let ε be a positive real number, let n be a natural number. Then there exists a stationary source $(S, \mu]$ with the following property:

1. $$H[S, \mu] = \varepsilon .$$

2. If $[A, \nu, B]$ is any stationary channel with finite past history where the numbers of elements both in A and B do not exceed n, then the source $[S, \mu]$ is not transmissible over the channel $[A, \nu, B]$ neither with respect to the error weight function w^e nor with respect to the frequency weight function w^f.

Proof: If $\varepsilon > 0$ and n are given, we put

$$K = \log n .$$

If $[A, \nu, B]$ is a channel for which the numbers of elements both in A and B do not exceed n, then

$$C[A, \nu, B] < K$$

(see Theorem 4, § 17 in [5].)
The assertion is the consequence of Theorem 4.

Lemma 1 and 2 show that a non-trivial linear combination of the different ergodic sources is not the ergodic source and thus the problem of the transmission is not solved in this case. Theorems 1, 2 and 3 solve this problem. Theorem 4 shows that the entropy of the stationary ergodic source is not so important for transmission as in the ergodic case. It would be interesting if it is possible to solve a general stationary case by a similar method (i. e. by means of the aproximation of stationary source by the linear combination of ergodic sources).

Резюме

В этой работе изучается специальный тип стационарных неэргодических источников информации и доказываются некоторые основные теоремы о переносе. Работа основывается на исчерпывающей работе [5] Карла Винкелвауера. Основные результаты следующие:

Пусть $[A, \nu, B]$ стационарный канал с конечным прошлым, пусть $[S, \mu^{(i)}]$, $i = 1, 2, \ldots, k$ эргодические источники, пусть $\alpha_1, \alpha_2, \ldots, \alpha_k$ положительные действительные числа также что $\sum_{i=1}^{k} \alpha_i = 1$, пусть

$$\mu = \sum_{i=1}^{k} \alpha_i \mu^{(i)} .$$

Если $k > 1$ и составные источники разные, то источник $[S, \mu]$ есть стационарный неэргодический источник. Если

$$\max_{1 \leq i \leq k} H[S, \mu^{(i)}] < C[A, \nu, B]$$

($H\ [S, \mu^{(i)}]$ энтропия i-того источника $[S, \mu^{(i)}]$ и $C[A, \nu, B]$ пропускная способность канала $[A, \nu, B)]$) и если w регулярная органиченная функция потерь, то источник $[S, \mu]$ стриктно переносительный каналом $[A, \nu, B]$ относительно w. Специально источник $[S, \mu]$ стриктно переносительный каналом $[A, \nu, B]$ относительно ошибочной функции потерь и относительно частотной функции потерь (т. е. с произвольно малой вероятностей ошибки и с произвольно малой средней частотой ошибок). Если $C[A, \nu, B] > 0$ и если $\max_{1 \leq i \leq k} H[S, \mu^{(i)}] = C[A, \nu, B]$ то источник $[S, \mu]$ стриктно переносительный каналом (A, ν, B) относительно частотной функции потерь.

Если

$$\max_{1 \leq i \leq k} H[S, \mu^{(i)}] > C[A, \nu, B]$$

то источник (S, μ) непереносительный каналом $[A, \nu, B]$ относительно ошибочной функции потерь и относительно частотной функции потерь.

Интересно следующее утверждение:

Пусть ε и K произвольные действительные числа, $\varepsilon > 0$. Тогда существует стационарный источник $[S, \mu]$ так что

1. $$H[S, \mu] = \varepsilon.$$

2. Если $[A, \nu, B]$ произвольный стационарный канал с конечным прошлым, для которого

$$C[A, \nu, B] \leq K,$$

то источник $[S, \mu]$ непереносительный каналом $[A, \nu, B]$ относительно ошибочной функции потерь и относительно частотной функции потерь.

REFERENCES

[1] P. R. Halmos: Measure theory, New York, 1950.

[2] P. R. Halmos: Lectures on ergodic theory, The mathematical society of Japan, Tokyo, 1956.

[3] J. Nedoma: The Capacity of a Discrete Channel, Transactions of the First Prague Conference, pg. 143—181, NČSAV, Praha 1957.

[4] K. Takano: On the basic theorems of information theory, Ann. of the Inst Stat. Math., Tokyo, Vol. 9, No 2.

[5] K. Winkelbauer: Communication Channels With Finite Past History, Transactions of the Second Prague Conference, pg. 685—832, NČSAV, Praha 1960.

M. Nosál
Matematický ústav ČSAV
Žitná 25,
Praha 1