Kybernetika

Karel Winkelbauer

On the asymptotic rate of non-ergodic information sources

Kybernetika, Vol. 6 (1970), No. 2, (127)--148

Persistent URL: http://dml.cz/dmlcz/124906

KYBERNETIKA ČÍSLO 2, ROČNÍK 6/1970

On the Asymptotic Rate of Non-Ergodic Information Sources

KAREL WINKELBAUER

In this paper the proofs of the main theorems on the existence and explicit representation of asymptotic rate, as introduced by the author in [15] for non-ergodic stationary sources, are newly given, in general, for countable alphabets and in such a manner that the methods of proofs do not exceed the frame of the theory of invariant measures.

Before proceeding to the formulation of the problem we are treating in this paper, we shall remind some simple facts upon which the concepts given in the sequel are based. Let us imagine such a situation when a receiver is expecting one of the messages z belonging to a finite set Z of messages which are possible to come under the given situation. The uncertainty of the situation described is evidently the greater, the larger is the number of all a priori possible messages; let us denote the latter number, i.e. the number of messages in Z, by $|Z|$. Originally the quantity of information which is needed to remove this uncertainty, is numerically expressed by the number $\log |Z|$, where the logarithm is taken to the base 2; in other words, $\log |Z|$ represents the quantity of information expressed in bits which is contained in any message z in Z reaching the receiver. Consequently, the number $\log |Z|$ as a measure of uncertainty may be called the *logarithmic uncertainty* of the set Z.

C. E. Shannon extended the original concept of quantity of information to cases in which the uncertainty of the situation considered is caused by random factors. In such a situation the receiver gets a message $z^{(j)}$ when a random event E_j occurs with known probability p_j. Assume that $|Z| = k$,

$$Z = \{z^{(1)}, z^{(2)}, \ldots, z^{(k)}\}\,; \quad p_1 \geqq p_2 \geqq \ldots \geqq p_k\,; \quad \sum_{j=1}^{k} p_j = 1\,.$$

Let ε be a very small positive number such that, from the viewpoint of the receiver, any event composed from the events E_j considered above and having its probability less than ε is so unlikely to occur that its appearing is not expected and, therefore, not taken into account by the receiver. If such a composed event with probability $<\varepsilon$ is equivalent to occurring of one of the events

$$E_{l+1}, E_{l+2}, \dots, E_k$$

so that

$$\sum_{i=l+1}^{k} p_i < \varepsilon, \quad \text{or} \quad \sum_{i=1}^{l} p_i > 1 - \varepsilon \; (1 \leqq l \leqq k),$$

then the receiver may take as practically certain that he receives a message $z^{(j)}$ with index $j \leqq l$. The quantity of information which is necessary for removing the uncertainty contained in the situation described now cannot be for ε sufficiently small greater than the number $\log l$, which is the above measure for uncertainty corresponding to the reduced set of messages; cf. the pioneer work of Shannon [13].

Let us suppose that l is the smallest number that satisfies the inequalities given above, which means that, moreover,

$$\sum_{i=1}^{l-1} p_i \leqq 1 - \varepsilon.$$

Let us denote the smallest l by $L(\varepsilon)$. The number $\log L(\varepsilon)$ represents the minimum number of bits of imformation that is needed to remove the uncertainty of the situation considered, provided that we admit an error the probability of which is less than ε. In the author's paper [15] is proposed to use the quantity $\log L(\varepsilon)$ as a measure of uncertainty at the ε-level of the error admitted; the number $\log L(\varepsilon)$ may be called the logarithmic uncertainty *at the ε-level* of the set Z together with the probability distribution p_j.

One of the fundamental problems of information theory is the question how to characterize the quantity of information which is contained in a source of messages the statistical properties of which are known. Shannon and McMillan have shown that the best characterization of information quantity for stationary sources discrete in time and having finite alphabets is the entropy rate, provided that the sources under consideration satisfy the condition of ergodicity. The author has shown in [15] that the entropy rate does not express the effective level of uncertainty if the stationary sources under consideration are not ergodic; the new measure of uncertainty the author has treated in [15] is based upon the concept of logarithmic ε-uncertainty as described above. This new quantity was called by the author the asymptotic rate and shown to be effective and coinciding with entropy rate on the closer class of ergodic sources.

A stationary information source discrete in time the statistical properties of which are described by a probability distribution μ, produces at discrete moments of time letters belonging to a discrete alphabet, say A, the probability of the source production being independent of time. If A^n means the set of all messages composed of n letters, then the number

$$\text{(I)} \qquad \frac{1}{n} \log L_n(\varepsilon)$$

represents the logarithmic uncertainty at ε-level of the set A^n related to one letter of n-dimensional messages. The author has shown in [15] (*theorem on the existence of asymptotic rate*) that the sequence (I) converges to a limit, say H_ε, where H_ε monotonically increases for ε decreasing to zero to a limit H, which is then called the *asymptotic rate* of the source under consideration. As to the main properties of the quantity H, they may be deduced as corollaries from an explicit formula derived in [15] (*theorem on explicit definition of asymptotic rate*): the formula states that the asymptotic rate of a stationary source equals the essential supremum of the entropy rates of its ergodic components; in symbols:

$$H = \text{ess. sup } \mathsf{H}(\mu_z) ,$$

where $\mathsf{H}(\mu_z)$ designates the entropy rate of the ergodic component μ_z of the source described by the probability distribution μ. Especially, if the source contains only a finite number of ergodic components, i.e.

$$\mu = \sum_{j=1}^{m} \xi_j \mu_j , \quad \xi_j > 0 , \quad j = 1, 2, \ldots, m ,$$

the asymptotic rate of the source is expressed as the maximum

$$H = \max_{j=1,2,\ldots,m} \mathsf{H}(\mu_j) .$$

The methods of the proofs of both the main theorems on asymptotic rate quoted in the preceding text that are used in [15] are based upon considering together with information sources also communication channels which are to transmit the information produced.

Undoubtedly, communication channels are a more complicated object for investigation compared with information sources which represent in case of stationarity nothing else than invariant measures in the space of messages. The problem arises whether it is in principle possible to base the proofs of the main theorems on the existence and explicit representation of the asymptotic rate exclusively on methods which do not exceed the frame of the theory of invariant measures.

The problem is studied in the following four sections of this paper; the first two are of preparatory character, developing the main tool needed in the investigation. In the third section are derived two basic lemmas necessary for the proofs of the main theorems quoted above, which are stated in the last section.

1. RESTATEMENT OF McMILLAN'S THEOREM FOR DISCRETE ALPHABETS

In this paper we shall follow the terminology and notations used by the author in [15]. Throughout the entire paper we shall study only discrete information sources with discrete alphabets, i.e. such information sources that produce at discrete moments of time letters belonging to a finite or denumerable alphabet. Without a loss of generality we may assume that letters are represented by natural numbers.

The set of all natural numbers is denoted in the sequel by N; the set of all integers, as usual, by I. Since the symbol B^A designates the class of all mappings of a set A into a set B, the set N^I contains all doubly-infinite sequences of natural numbers, and z_i means the i-the member of sequence $z \in N^I$ $(i \in I)$. The basic space under consideration is the measurable space $(N^I, \mathbf{F}^I)$, where $\mathbf{F}$ means the class of all subsets of the set N, and where $\mathbf{F}^I$ is the σ-algebra of subsets of N^I generated by the class of one-dimensional cylinders, i.e. sets of the form

$$\{z : z \in N^I, z_i \in E\}, \quad i \in I, \quad E \in \mathbf{F}.$$

A discrete information source, or briefly a *source*, is then defined as a probability measure given on the σ-algebra $\mathbf{F}^I$.

The coordinate-shift transformation will be denoted by T; it is defined by the property that $(Tz)_i = z_{i+1}$, $z \in N^I$, $i \in I$. A source μ is a *stationary* one if measure μ is invariant with respect to the transformation T, i.e. if $\mu = \mu T^{-1}$ (cf. [1], § 39). We shall say that a source μ has a *finite alphabet* if

$$(\{1, 2, \ldots, k\}^I) = 1 \tag{1.1}$$

for some natural number k. The *entropy rate* of a stationary source is defined as the limit

$$\mathsf{H}(\mu) = - \lim_n (1/n) \int \log \mu[z_0, z_1, \ldots, z_{n-1}] \, \mathrm{d}\mu(z), \tag{1.2}$$

where

$$[z_0, z_1, \ldots, z_{n-1}] = \bigcap_{i=0}^{n-1} \{x : x \in N^I, x_i = z_i\}, \quad z \in N^I; \tag{1.3}$$

all the logarithms in this paper will be to the base 2. The existence of the latter limit is a consequence of the invariance of measure μ (cf. [2] or [4]). As well-known, if μ is a stationary source satisfying condition (1.1) of finiteness of its alphabet, then its entropy rate fulfils the inequality

$$\mathsf{H}(\mu) \leqq \log k\,. \tag{1.4}$$

Throughout this paper we shall consider only stationary sources; nevertheless, the methods of proofs given in the sequel are valid without change also for periodic sources. For convenience, we shall denote the class of all stationary sources by $\mathcal{M}_{st}$.

Our considerations will be based on the well-known McMillan's theorem which was proved by its author for stationary sources with finite alphabets. Since we are dealing more generally with sources with countable alphabets, we shall need a generalization of the theorem mentioned for the case of infinite alphabets. A general version of McMillan's theorem for infinite alphabets was first proved by A. Perez in [10] (cf. also the treatment [11] by the same author); the methods of Perez's proofs are based on the investigation of his concerning the relations between martingales and generalized entropies (in [9]). However, we shall need a special formulation of Perez's version of McMillan's theorem which we shall state and prove in this section.

First we shall define the concepts that are necessary in the statement of the theorem. If J is a non-empty subset of I, let $\mathbf{F}^J$ be the σ-algebra of sets in space N^I which is generated by the class of one-dimensional cylinders with coordinates in J, i.e. all sets of the form $\{z : z_i \in E\}$, $i \in J$, $E \in \mathbf{F}$. For the sake of simplicity, we shall put

$$\mathbf{F}^- = \mathbf{F}^{\{i : i \in I, i \leqq -1\}}\,.$$

If $\mu \in \mathcal{M}_{st}$, $a \in N$, let us denote by $\hat{g}_\mu(\cdot, a)$ the conditional probability of the event $\{z : z_0 = a\}$ under the condition $\mathbf{F}^-$ which is uniquely determined by the equations

$$\int_E \hat{g}_\mu(z, a)\,\mathrm{d}\mu(z) = \mu(E \cap \{z : z_0 = a\})\,, \quad E \in \mathbf{F}^- \tag{1.5}$$

modulo measure μ restricted to the class $\mathbf{F}^-$. With the aid of the latter conditional probability we shall define the function g_μ by the equation

$$g_\mu(z) = \hat{g}_\mu(z, z_0)\,, \quad z \in N^I \quad (\mu \in \mathcal{M}_{st})\,; \tag{1.6}$$

z_0 is, of course, the zeroth coordinate of point z.

Now we shall assume that we are given a stationary source μ with finite entropy, i.e. $\mathsf{H}(\mu) < +\infty$. In the following text we shall show that then the entropy rate of the source μ may be expressed by the formula

$$\mathsf{H}(\mu) = -\int \log g_\mu(z)\,\mathrm{d}\mu(z)\,. \tag{1.7}$$

From the finiteness of entropy rate and from (1.7) it follows that, according to the individual ergodic theorem, there exists the limit

(1.8) $$h_\mu(z) = -\lim_n (1/n) \sum_{j=0}^{n-1} \log g_\mu(T^j z), \quad z \in N^I[\mu]$$

(as to the symbol $[\mu]$, cf. [1], § 30) satisfying the equality

(1.9) $$\int h_\mu(z)\, d\mu(z) = \mathsf{H}(\mu).$$

Now we are able to state Perez-McMillan's theorem in the form:

Theorem 1. *If μ is a stationary source with finite entropy, i.e. $\mathsf{H}(\mu) < +\infty$, then the sequence $-(1/n)\log \mu[z_0, z_1, \ldots, z_{n-1}]$ converges in the mean to the function $h_\mu(z)$ with respect to μ; in symbols:*

$$\int |-(1/n) \log \mu[z_0, z_1, \ldots, z_{n-1}] - h_\mu(z)|\, d\mu(z) \to 0 \quad \text{for} \quad n \to \infty .$$

Proof. In the entire proof, in which we shall show also the validity of the formula (1.7), we shall employ the notations given in [10]. Throughout the proof we assume that we are given a source $\mu \in \mathscr{M}_{st}$ with finite entropy.

I. We shall set $\mathfrak{X}^- = \mathbf{F}^{\{i: i \in I, i \leqq 0\}}$. Let us define on $\mathfrak{X}^-$ the measure λ by the property that

(1) $$\lambda(E \cap F) = \mu(E) \cdot \mu(F), \quad E \in \mathbf{F}^-, \quad F \in \mathbf{F}^{\{0\}}.$$

We will show that measure μ restricted to the class $\mathfrak{X}^-$ is absolutely continuous with respect to λ; in symbols:

(2) $$\mu < < \lambda(\mathfrak{X}^-).$$

Let $E \in \mathfrak{X}^-$, $\lambda(E) = 0$. The set E may be expressed in the form

$$E = \bigcup_{a \in N} (E_a \cap [a]),$$

$$E_a = \{z: \{z_i\}_{i \leqq -1} = \{x_i\}_{i \leqq -1} \quad \text{for some} \quad x \in E \cap [a]\},$$

where we have put $[a] = \{z: z_0 = a\}$. Since $E_a \in \mathbf{F}^-$ and

$$0 = \lambda(E_a \cap [a]) = \mu(E_a) \cdot \mu[a]$$

(cf. (1)), we obtain that $\mu(E_a \cap [a]) = 0$. The latter equality implies that $\mu(E) = 0$; hence (2) holds.

Let g be the (Radon-Nikodym) density

$$g = \frac{d\mu}{d\lambda}(\mathfrak{X}^-) . \tag{3}$$

It is an immediate consequence of (1.6) that (cf. (1.3))

$$g(z) = \frac{g_\mu(z)}{\mu[z_0]}, \quad z \in N^I[\mu, \mathfrak{X}^-] . \tag{4}$$

Let $E \in F^-$, $a \in N$, $\mu[a] > 0$. Using (1.5) and (1), we obtain from Fubini's theorem that

$$\int_{E \cap [a]} \frac{g_\mu(z)}{\mu[a]} d\lambda(z) = \int_{[a]} \left[\int_E \frac{\hat{g}_\mu(z, a)}{\mu[a]} d\mu(z)\right] d\mu(x) =$$

$$= \mu(E \cap [a]) = \int_{E \cap [a]} g(z) \, d\lambda(z) .$$

If $E \in \mathfrak{X}^-$, then the above expression of E with the aid of sections E_a will yield from the latter relations the equality

$$\int_E \frac{g_\mu(z)}{\mu[z_0]} d\lambda(z) = \int_E g(z) \, d\lambda(z) .$$

Hence it follows (4) as a consequence of (2).

II. Let w be the source that is uniquely determined by the condition that

$$w\{z: z_j = a_j (i \leqq j < i + n)\} = \prod_{j=i}^{i+n-1} \mu[a_j] ; \quad a_j \in N , \quad i \in I , \quad n \in N . \tag{5}$$

According to Theorem 3 stated in [10] and the relation (2) there exists the generalized entropy H_w of the source μ with respect to measure w which may be expressed in the form

$$H_w = - \int \log g(z) \, d\mu(z) = \lim_n \frac{1}{n} H_w(\mu, \mathfrak{X}_0^n) , \tag{6}$$

$$\mathfrak{X}_0^n = F^{\{i: i \in I, 0 \leqq i < n\}} ,$$

where

$$H_w(\mu, \mathfrak{X}_0^n) = - \int f_n \log f_n \, dw = - \int \log f_n \, d\mu ; \tag{7}$$

the function f_n represents the (Radon-Nikodym) density

$$f_n = \frac{d\mu}{dw}(\mathfrak{X}_0^n); \quad n = 1, 2, \ldots \tag{8}$$

 The latter function may be expressed in the form

(9) $$f_n(z) = \frac{\mu[z_0, z_1, \ldots, z_{n-1}]}{\mu[z_0]\,\mu[z_1]\ldots\mu[z_{n-1}]}, \quad z \in N^I[\mu, \mathfrak{X}_0^n];$$

this formula is an immediate consequence of (8) and (5). Let us set for the sake of brevity

(10) $$\mathsf{H}_n = -\sum_{z\in N^n}\mu[z]\log\mu[z] = -\int\log\mu[z_0, \ldots, z_{n-1}]\,\mathrm{d}\mu(z).$$

Since $\mathsf{H}_n \leqq \mathsf{H}_{n+1}$, $\mathsf{H}_n \leqq n\mathsf{H}_1$ (cf., for example, [17], § 6), it is clear that $\mathsf{H}(\mu) < +\infty$ holds if and only if $\mathsf{H}_n < +\infty$ for $n = 1, 2, \ldots$ Consequently, it follows from the finiteness of the entropy rate of the source μ that the difference $\mathsf{H}_n - n\mathsf{H}_1$ is defined, and is finite and non-positive. By an easy calculation we obtain from (7), (9), (10) the equality

$$H_w(\mu, \mathfrak{X}_0^n) = \mathsf{H}_n - n\mathsf{H}_1$$

so that the generalized entropy may be expressed by the formula

(11) $$H_w = \mathsf{H}(\mu) - \mathsf{H}_1;$$

hence H_w is finite and non-positive. The relations (6), (10), and (4) immediately yield the equality

$$H_w + \mathsf{H}_1 = -\int\log g_\mu(z)\,\mathrm{d}\mu(z).$$

We have shown the validity of formula (1.7), as follows from (11).

III. As we have found, the finiteness of the entropy rate $\mathsf{H}(\mu)$ implies the finiteness of the generalized entropy H_w. Applying now the individual ergodic theorem to the function g, we have that there is the limit

(12) $$h(z) = -\lim_n \frac{1}{n}\sum_{j=0}^{n-1}\log g(T^j z), \quad z \in N^I[\mu]$$

satisfying the integral relation

(13) $$\int h(z)\,\mathrm{d}\mu(z) = H_w.$$

Since the quantity H_1 is finite as well, we may apply to the function $\log\mu[z_0]$ the mean ergodic theorem (cf. (10) for $n = 1$) so that we get that the sequence $(1/n)\,.\,\sum_{j=0}^{n-1}\log\mu[z_j]$ converges in the mean (with respect to μ) to the limit $h(z) - h_\mu(z)$ as follows from (1.8) and (12).

Owing to (2), (5), and (11), the assumptions of Theorem 4 stated in [10] are fulfilled so that Perez's version of McMillan's theorem is valid: it says that the sequence $-(1/n)\log f_n$ (cf. (8) and (9)) converges in the mean (with respect to μ) to the function h. According to (9) we have

$$(14)\qquad -\frac{1}{n}\log f_n(z) = -\frac{1}{n}\log \mu[z_0, z_1, \ldots, z_{n-1}] + \\ + \frac{1}{n}\sum_{j=0}^{n-1}\log \mu[z_j]\,, \quad z \in N^I[\mu]\,.$$

It is easy to see that the difference

$$-\frac{1}{n}\log f_n(z) - \frac{1}{n}\sum_{j=0}^{n-1}\log \mu[z_j]$$

converges in the mean to the limit $h_\mu(z)$; from here and from (14) the validity of the assertion of the theorem immediately follows, Q.E.D.

2. ON THE ERGODIC THEORY FOR DISCRETE ALPHABETS

This section is devoted to a closer investigation of the function h_μ which appears in our statement of McMillan's theorem (cf. definition (1.8) and Theorem 1). The investigation will be based on the ergodic theory of invariant measures in the basic space $(N^I, \mathbf{F}^I)$.

For convenience, we shall denote by V_n $(n \in N)$ the class of (elementary) $(2n + 1)$-dimensional cylinders in coordinates running from $-n$ to n, i.e. all sets expressible in the form

$$(2.1)\qquad \bigcap_{i=-n}^{n} \{z\colon z \in N^I, z_i = a_i\}\,, \quad (a_{-n}, \ldots, a_n) \in N^{2n+1}\,.$$

Obviously the class V_n is a countable partition of the basic space. The union of all the classes V_n will be designated in what follows by V: $\mathsf{V} = \bigcup_{n\in N} \mathsf{V}_n$.

A point z in the basic space N^I will be called *quasi-ergular* if and only if there is an invariant measure μ_z, i.e. $\mu_z \in \mathcal{M}_{\mathrm{st}}$ (a stationary source), having the property that

$$(2.2)\qquad \mu_z(E) = \lim_n (1/n)\sum_{j=0}^{n-1}\chi_E(T^j z) \quad \text{for each} \quad E \in \mathsf{V}\,;$$

here χ_E means the characteristic function of the set E. It is evident that the measure μ_z is uniquely determined by the quasi-regular point z and by the condition (2.2). We shall designate the set of all quasi-regular points in the space N^I by Q.

Remark. The latter definition of quasi-regularity is a correction to the definition (8.4) given in the author's paper [15], which is not sufficient. If we define the set Q_0 by the equation

$$Q_0 = \bigcap_{E \in \mathsf{V}} \{z: z \in N^I, \overline{\lim_n}\, \chi_E^{(n)}(z) = \underline{\lim_n}\, \chi_E^{(n)}(z)\}, \tag{2.3}$$

where we have put

$$\chi_E^{(n)}(z) = (1/n) \sum_{j=0}^{n-1} \chi_E(T^j z), \tag{2.4}$$

then we have the set inclusion $Q_0 \supset Q$, but the equality $Q_0 = Q$ is not valid: it is easy to show that, for example, the point z with coordinates

$$z_i = 1 \quad \text{for} \quad i \leqq 0\,; \quad z_{2i} = 1\,, \quad z_{2i+1} = i \quad \text{for} \quad i > 0 \quad (i \in I)$$

belongs to Q_0, but it does not belong to Q. The following lemma shows that the concept of quasi-regularity as defined above may be used to develop the ergodic theory.

Lemma 1. *The set Q of quasi-regular points is measurable, i.e. $Q \in \mathbf{F}^I$, and $\mu(Q) = 1$ for every $\mu \in \mathcal{M}_{\mathrm{st}}$.*

Proof. It follows from the countability of the class V and from the individual ergodic theorem that the set Q_0 defined by (2.3) is measurable and of the property that $\mu(Q_0) = 1$ for any $\mu \in \mathcal{M}_{\mathrm{st}}$. It may easily be shown that the set function μ_z defined by (2.2) on V for such a point $z \in Q_0$ that satisfies the relations

$$\sum_{E \in \mathsf{V}_n} \lim_n \chi_E^{(n)}(z) = 1 \quad \text{for every} \quad n \in N\,,$$

uniquely determines an invariant measure on $\mathbf{F}^I$ (cf. the well-known Kolmogorov's theorem, e.g., in [1], § 49; the invariance is, of course, an obvious consequence of the definition of the set function μ_z). This fact enables us to express the set Q in the form

$$Q = \bigcap_{n=1}^{\infty} \{z: z \in Q_0, \sum_{E \in \mathsf{V}_n} \lim_n \chi_E^{(n)}(z) = 1\}\,. \tag{2.5}$$

Hence Q is a measurable set.

Given $0 < \varepsilon < 1$ and $\mu \in \mathcal{M}_{\mathrm{st}}$, let $F_k(\varepsilon)$ be a finite union of sets belonging to the partition V_k $(k \in N)$ which satisfies the inequality

$$\mu(F_k(\varepsilon)) > 1 - (\varepsilon\,.\,2^{-k})^2\,. \tag{1}$$

Let us set (cf. (2.4))

$$Q_k(\varepsilon) = \{z: z \in Q_0, \lim_n \chi_{F_k(\varepsilon)}^{(n)}(z) > 1 - \varepsilon\,.\,2^{-k}\}\,. \tag{2}$$

Since $\mu(Q_0) = 1$, it follows from the individual ergodic theorem applied to the characteristic function of the set $F_k(\varepsilon)$ that

$$\int_{Q_0} \lim_n \chi^{(n)}_{F_k(\varepsilon)} \,\mathrm{d}\mu = \mu(F_k(\varepsilon)) . \tag{3}$$

From the relations (1) and (3) we obtain that

$$\mu(Q_k(\varepsilon)) > 1 - \varepsilon \,.\, 2^{-k} ; \tag{4}$$

if the contrary were true, it would be

$$\int_{Q_0} \lim_n \varkappa_{F_k(\varepsilon)} \,\mathrm{d}\mu \leqq \mu(Q_k(\varepsilon)) + \left(1 - \varepsilon \,.\, 2^{-k}\right) \mu(Q_k'(\varepsilon)) =$$

$$= 1 - \frac{\varepsilon}{2^k}\, \mu(Q_k'(\varepsilon)) \leqq 1 - \left(\varepsilon \,.\, 2^{-k}\right)^2 ,$$

which is the desired contradiction. If we set

$$K_\varepsilon = \bigcap_{k=1}^{\infty} Q_k(\varepsilon) \subset Q_0 \tag{5}$$

(cf. (2)), we immediately get from (4) that $\mu(K_\varepsilon) > 1 - \varepsilon$. Let

$$K = \bigcup_{s=1}^{\infty} K_{(1/s)} ; \quad \text{hence} \quad \mu(K) = 1 . \tag{6}$$

Now we shall show that $Q \supset K$ so that (6) will imply that $\mu(Q) = 1$, which is to be proved. Let $z \in K$ so that $z \in K_{(1/s)}$ for some natural number s. Writing $\varepsilon = 1/s$ and assuming on the contrary that z does not belong to Q, we obtain from the set inclusion (5) the inequality (cf. (2.5))

$$\sum_{F \in \mathsf{V}_k} \lim_n \chi^{(n)}_F(z) \leqq 1 - \delta < 1$$

for some $k \in N$ and $0 < \delta < 1$. Let us choose $m \geqq k$ such that $\varepsilon \,.\, 2^{-m} \leqq \delta$. Since $z \in Q_m(\varepsilon)$, as follows from (5), it must be

$$\sum_{E \in \mathsf{V}_m} \lim_n \varkappa^{(n)}_E(z) \geqq \lim_n \chi^{(n)}_{F_m(\varepsilon)}(z) > 1 - \varepsilon \,.\, 2^{-m} \geqq 1 - \delta .$$

On the other hand it holds for every $F \in \mathsf{V}_k$ that

$$\sum_{(E \in \mathsf{V}_m, E \subset F)} \lim_n \chi^{(n)}_E(z) \leqq \lim_n \chi^{(n)}_F(z) ;$$

it is because $m \geqq k$ and because V_m is a subpartition of V_k. Summarizing the preceding facts, we deduce that

$$\sum_{F \in \mathsf{V}_k} \lim_n \chi_F^{(n)}(z) > 1 - \delta ,$$

which is the desired contradiction. This proves the assertion of the lemma.

A stationary source μ is, by definition, *ergodic* if it satisfies the condition of ergodicity that

$$TE = E \quad \text{and} \quad \mu(E) > 0 \quad \text{implies} \quad \mu(E) = 1 \tag{2.6}$$

for $E \in \mathsf{F}^I$. A point z in the basic space N^I is called *regular* if it is quasi-regular and if the measure μ_z uniquely determined by (2.2) is an ergodic source.

Now we shall summarize the fundamental results of ergodic theory that will be needed in the sequel, in three lemmas. In the lemmas the set of all regular points in the space N^I is denoted by R.

Lemma 2. *The set R of regular points is measurable, i.e. $R \in \mathsf{F}^I$, and $\mu(R) = 1$ for every $\mu \in \mathcal{M}_{st}$.*

Lemma 3. *For every ergodic source μ the set of those regular points z for which $\mu_z = \mu$, is measurable and is of μ-measure one; in symbols:*

$$\mu\{z: z \in R, \mu_z = \mu\} = 1 .$$

Lemma 4. *For any $E \in \mathsf{F}^I$, $\mu_z(E)$ is a measurable function of variable z on R, and*

$$\mu(E) = \int_R \mu_z(E) \, \mathrm{d}\mu(z) \quad \text{for every} \quad \mu \in \mathcal{M}_{st} .$$

More generally, for any non-negative measurable function f on the space (N^I, F^I) the integral $\int f \, \mathrm{d}\mu_z$ is a measurable function of variable z on R, and

$$\mu(E) = \int_R \mu_z(E) \, \mathrm{d}\mu(z) \quad \text{for every} \quad \mu \in \mathcal{M}_{st} .$$

Proofs of the lemmas are based on Lemma 1 and are given in [15] (cf. also the systematic survey in [5], and the original contribution [3] where are studied compact dynamical systems only).

Now we are prepared to investigate the function h_μ. The desired property we shall need in the sequel is stated in the following lemma. The method of the proof of the lemma is due to K. R. Parthasarathy (cf. [6] and [8], where the sources investigated have finite alphabets).

Lemma 5. *If $\mu \in \mathscr{M}_{st}$, $\mathsf{H}(\mu) < +\infty$, then*

$$\mu\{z: z \in R,\, h_\mu(z) = \mathsf{H}(\mu_z)\} = 1\,;$$

in words, $h_\mu(z)$ equals with probability one the entropy of ergodic component μ_z of any given stationary source μ with finite entropy.

Proof. For $a \in N$, let

$$(1) \qquad E_a = \{z: z \in R,\, \mu_z\{x: x \in N^I,\, \hat{g}_\mu(x, a) = \hat{g}_{\mu_z}(x, a)\} = 1\}\,;$$

cf. (1.5). The set E_a is measurable by Lemma 2 and 4. According to Theorem 2.6 in [6] (the proof is valid also for countable alphabets) it holds that

$$(2) \qquad \mu(E_a) = 1 \quad \text{for every} \quad \mu \in \mathscr{M}_{st}\,.$$

From (2) it follows that for

$$(3) \qquad E = \bigcap_{a \in N} E_a\,, \quad \mu(E) = 1 \quad \text{for every} \quad \mu \in \mathscr{M}_{st}\,.$$

Let μ be a stationary source with finite entropy: $\mathsf{H}(\mu) < +\infty$. Owing to (1) we have for any $z \in E$

$$\begin{aligned} 1 &\geqq \mu_z\{x: g_\mu(x) = g_{\mu_z}(x)\} = \\ &= \mu_z\Big(\bigcup_{a \in N}\{x: x_0 = a,\, \hat{g}_\mu(x, a) = \hat{g}_{\mu_z}(x, a)\}\Big) = \\ &= \sum_{a \in N} \mu_z\{x: x_0 = a,\, \hat{g}_\mu(x, a) = \hat{g}_{\mu_z}(x, a)\} \geqq \sum_{a \in N} \mu_z[a] = 1\,. \end{aligned}$$

Hence it follows that

$$(4) \qquad \mu_z\{x: x \in N^I,\, g_\mu(x) = g_{\mu_z}(x)\} = 1\,, \quad z \in E\,.$$

As a consequence of Theorem 8.2 stated in [15] we obtain that, provided the entropy rate $\mathsf{H}(\mu)$ is finite,

$$\mu\{z: z \in R,\, \mathsf{H}(\mu_z) < +\infty\} = 1\,.$$

Putting

$$(5) \qquad F = E \cap \{z: z \in R,\, \mathsf{H}(\mu_z) < +\infty\}\,,$$

we deduce from (3) the equality

$$(6) \qquad \mu(F) = 1\,.$$

Now let $z \in F$. Relations (4) and (5) yield the equality

$$(7) \qquad \mu_z\{x: x \in N^I\,, \quad h_\mu(x) = h_{\mu_z}(x)\} = 1 \quad (z \in F)\,.$$

If $z \in R$, $\mathsf{H}(\mu_z) < +\infty$, it follows from the ergodicity of μ_z and from McMillan's theorem as given in the preceding section (cf. Theorem 1) that

$$h_{\mu_z}(x) = \mathsf{H}(\mu_z)\,, \quad x \in N^I[\mu_z]$$

(the latter fact is a consequence of the invariance of h_μ). From (7) and from the latter relation we deduce that

$$\mu_z\{x\colon x \in N^I, h_\mu(x) = \mathsf{H}(\mu_z)\} = 1\,, \quad z \in F\,.$$

Applying Lemma 3, we can write the latter equality in the form

$$\mu_z\{x : x \in R, \mu_x = \mu_z, h_\mu(x) = \mathsf{H}(\mu_x\} = 1\,, \quad z \in F\,.$$

Integrating with respect to μ and using Lemma 2 and Lemma 4, we immediately obtain that (cf. (6))

$$\int_F \mu_z\{x : x \in R, h_\mu(x) = \mathsf{H}(\mu_x)\}\, \mathrm{d}\mu(z) =$$

$$= \mu\{x\colon x \in R, h_\mu(x) = \mathsf{H}(\mu_x)\} = 1\,,$$

which proves the assertion of the lemma.

The following theorem is an immediate corollary to both Lemma 5 and Theorem 1 of the preceding section; it constitutes the version of McMillan's theorem for countable alphabets that will be used as the main tool in the proofs of the basic lemmas stated in the following section.

Theorem 2. *If μ is a stationary source with finite entropy, i.e.* $\mathsf{H}(\mu) < +\infty$, *then the sequence* $-(1/n) \log \mu[z_0, z_1, \ldots, z_{n-1}]$ *convergences in the mean (with respect to μ) to the function* $\mathsf{H}(\mu_z)$, *i.e. to the entropy rate of the ergodic component μ_z of the source given.*

Let us point out that we shall use in the sequel only the fact that the sequence $-(1/n) \log \mu[z_0, z_1, \ldots, z_{n-1}]$ converges in probability to $\mathsf{H}(\mu_z)$ with respect to the given stationary source μ.

3. BASIC LEMMAS

After the preparations made in the preceding sections, we can proceed to the main programme of this paper.

Throughout the remainder of this paper the symbol $|M|$ means for a finite set M the number of elements in M, and for an infinite set we define $|M| = +\infty$. If $\mu \in \mathcal{M}_{st}$, $n \in N$, we put

$$\mu_n(E) = \mu\{z\colon z \in N^I, (z_0, z_1, \ldots, z_{n-1}) \in E\}\,, \quad E \subset N^n\,, \tag{3.1}$$

and define the *n-dimensional ε-length* of the source μ, in symbols $L_n(\varepsilon, \mu)$, by the relation

$$(3.2) \qquad L_n(\varepsilon, \mu) = \min \{|E|: E \subset N^n, \mu_n(E) > 1 - \varepsilon\}, \quad 0 < \varepsilon < 1 .$$

The number $L_n(\varepsilon, \mu)$ represents the minimum number of n-tuples composed of letters of our universal alphabet N whose total probability exceeds $1 - \varepsilon$.

Lemma I. *If $\mu \in \mathcal{M}_{st}$, and if c is a finite real number, then the assumption that*

$$(1) \qquad \mu\{z: z \in R, \mathsf{H}(\mu_z) \leqq c\} = 1$$

implies the inequality

$$(2) \qquad \overline{\lim_n}\, (1/n) \log L_n(\varepsilon, \mu) \leqq c \quad \text{for} \quad 0 < \varepsilon < 1 .$$

Proof. If μ is a source satisfying the assumptions of the lemma, it follows from Theorem 8.2 given in [15] that the entropy rate $\mathsf{H}(\mu) \leqq c < +\infty$. Then it is possible to apply Theorem 2 of the preceding section; we obtain that the sequence $-(1/n) \times \times \log \mu[z_0, \ldots, z_{n-1}]$ convergences to $\mathsf{H}(\mu_z)$ in probability modulo μ. Let $0 < \varepsilon < 1$ and $\delta > 0$. Then there is an index n_0 such that for any $n \geqq n_0$ the inequality

$$\{z: z \in R, -(1/n) \log \mu[z_0, \ldots, z_{n-1}] \leqq \mathsf{H}(\mu_z) + \delta\} > 1 - \varepsilon .$$

holds. Given $n \geqq n_0$, we deduce from (1) and from the preceding inequality that the set

$$(3) \qquad E_n = \{z: z \in N^n, \mu_n(z) \geqq 2^{-n(c+\delta)}\}$$

has the property that $\mu_n(E_n) > 1 - \varepsilon$ (cf. (3.1)); it is because $\mathsf{H}(\mu_z) + \delta \leqq c + \delta$ for $z \in R[\mu]$. From the latter fact and from the definition (3.2) we obtain that $|E_n| \geqq l_n = L_n(\varepsilon, \mu)$. On the other hand, owing to (3.1) and (3), we have

$$1 \geqq \mu_n(E_n) = \sum_{z \in E_n} \mu_n(z) \geqq |E_n| \cdot 2^{-n(c+\delta)} .$$

Combining the inequalities which were found, we get that

$$l_n \leqq 2^{n(c+\delta)}, \quad \text{i.e.} \quad (1/n) \log L_n(\varepsilon, \mu) \leqq c + \delta .$$

The latter result implies the desired inequality (2) because of the arbitrariness of δ, Q.E.D.

The second basic lemma is a dual version of the first. There is a difficulty in the proof of this dual version that we can use Theorem 2 only in case the entropy rate of the source considered is finite; if not, we must approximate the source by a source with finite entropy. This is possible to do, as pointed out by Parthasarathy in [7] where an approximation is based on results given in [12] and [14].

Let us define the measurable transformation of the basic space (N^I, F^I), denoted in the sequel as τ_k $(k \in N)$, by the relations

$$(3.3) \qquad (\tau_k z)_i = z_i \quad \text{for} \quad z_i \leqq k\,, \quad (\tau_k z)_i = k + 1 \quad \text{for} \quad z_i > k \quad (i \in I)\,.$$

If μ is a stationary source, then the measure $\mu\tau_k^{-1}$ is again a stationary source which has a finite alphabet and finite entropy rate satisfying the inequality

$$(3.4) \qquad \mathsf{H}(\mu\tau_k^{-1}) \leqq \log(k + 1)\,,$$

as follows from (1.4).

Lemma 6. *If $\mu \in \mathcal{M}_{st}$, then the sequence of entropy rates $\mathsf{H}(\mu\tau_k^{-1})$, $k = 1, 2, \ldots$, monotonically increases to the entropy rate $\mathsf{H}(\mu)$.*

Proof. The assertion of the lemma coincides with Theorem 1 stated in [7].

Lemma II. *If $\mu \in \mathcal{M}_{st}$, and if c is a finite real, number, then the assumption that*

$$(1) \qquad \mu\{z\colon z \in R,\ \mathsf{H}(\mu_z) \geqq c\} = 1$$

implies the inequality

$$(2) \qquad \varliminf_n (1/n) \log L_n(\varepsilon, \mu) \geqq c \quad \text{for} \quad 0 < \varepsilon < 1\,.$$

Proof. Let μ be a source satisfying the assumptions of the lemma. Given ε, $0 < \varepsilon < 1$, choose ε' such that

$$(3) \qquad \varepsilon + 2\varepsilon' < 1\,, \quad \varepsilon' \leqq \tfrac{1}{2}\varepsilon\,.$$

Let $\delta > 0$. Choose $\delta' > 0$ such that

$$(4) \qquad 4\delta' < \delta\,.$$

If we apply Lemma 6 to the source μ_z for $z \in R$, we obtain that (cf. (3.3)) the sequence $\mathsf{H}(\mu_z\tau_k^{-1})$ monotonically increases to $\mathsf{H}(\mu_z)$ for $k \to \infty$; hence it follows the set inclusion

$$\{z\colon z \in R,\ \mathsf{H}(\mu_z) > c - \delta'\} = \bigcup_{k=1}^{\infty} \{z\colon z \in R,\ \mathsf{H}(\mu_z\tau_k^{-1}) > c - \delta'\}\,.$$

Then we deduce from (1) that there is k_0 such that for $k \geqq k_0$

$$(5) \qquad \mu\{z\colon z \in R,\ \mathsf{H}(\mu_z\tau_k^{-1}) > c - \delta'\} > 1 - \varepsilon'\,.$$

It is easy to see that (cf. (2.2))

$$\mu_z\tau_k^{-1} = \mu_{\tau_k z} \quad \text{for} \quad z \in R,\ \tau_k z \in R \quad (k \in N)\,.$$

Hence we get according to (5) that

(6) $$\mu\tau_k^{-1}\{z: z \in R, \mathsf{H}(\mu_z) > c - \delta'\} > 1 - \varepsilon' \quad \text{for} \quad k \geqq k_0,$$

since $\mu(R \cap \tau_k^{-1}(R)) = 1$.

Now let us choose $k \geqq k_0$ fixed. Applying Theorem 2 to the stationary source $\mu\tau_k^{-1} = \mu'$, we have that the sequence $-(1/n) \log \mu'[z_0, \ldots, z_{n-1}]$ converges in probability to $H(\mu_z)$ with respect to μ'; cf. (3.4). Consequently, there is n_0 such that (cf. (3))

(7) $$n_0 \geqq \frac{-2 \log (1 - \varepsilon - 2\varepsilon')}{\delta},$$

and that, for any $n \geqq n_0$,

(8) $$\mu'\left\{z: z \in R, -\frac{1}{n} \log \mu'[z_0, \ldots, z_{n-1}] \geqq \mathsf{H}(\mu_z) - \delta'\right\} > 1 - \varepsilon'.$$

Combining the relations (6) and (8) for $\mu' = \mu\tau_k^{-1}$, we obtain the inequality

(9) $$\mu'\left\{z: z \in N^I, -\frac{1}{n} \log \mu'[z_0, \ldots, z_{n-1}] > c - 2\delta'\right\} > 1 - 2\varepsilon'$$

for $n \geqq n_0$. Putting for a given $n \geqq n_0$

$$E_n = \{z: z \in N^n, \mu_n'(z) < 2^{-n(c-2\delta')}\},$$

we deduce from (9) the relation

(10) $$\mu_n'(E_n) > 1 - 2\varepsilon'.$$

Let $F_n \subset N^n$ be such that

(11) $$|F_n| = L_n(\varepsilon, \mu'), \quad \mu_n'(F_n) > 1 - \varepsilon;$$

the existence of such a set is guaranteed by (3.2). It follows from Lemma 1.4 given in [15] that

(12) $$L_n(\varepsilon, \mu') = L_n(\varepsilon, \mu\tau_k^{-1}) \leqq L_n(\varepsilon, \mu) = l_n.$$

From the inequalities stated in (10), (11), and (12) we deduce the relations

$$1 - \varepsilon - 2\varepsilon' < \mu(E_n \cap F_n) \leqq |E_n \cap F_n| \,.\, 2^{-n(c-2\delta')} \leqq$$
$$\leqq |F_n| \,.\, 2^{-n(c-2\delta')} \leqq l_n \,.\, 2^{-n(c-2\delta')}.$$

Owing to (7), we obtain from the latter inequalities that

$$\log l_n > n(c - 2\delta') - \tfrac{1}{2}n\delta.$$

 Finally, using (4), we deduce the relation

$$\frac{1}{n} \log L_n(\varepsilon, \mu) > c - \delta ,$$

which implies the desired inequality (2) because of the fact that δ was chosen arbitrarily. This proves the lemma.

To derive the main theorems, we shall make use of the following two lemmas which are a direct consequence of the definition (3.2).

Lemma 7. *If* $0 \leqq \xi < 1$, *and if* μ, μ_1, and μ_2 are stationary sources such that

$$\mu(E) = (1 - \xi)\, \mu_1(E) + \xi\, \mu_2(E) , \quad E \in \mathbf{F}^I ,$$

then

$$\overline{\lim_n} \frac{1}{n} \log L_n(\varepsilon, \mu) \leqq \overline{\lim_n} \frac{1}{n} \log L_n(\varepsilon - \xi, \mu_1) \quad \text{for} \quad \xi < \varepsilon < 1 .$$

Proof. The lemma is an immediate corollary to Lemma 1.2 given in [15].

Lemma 8. *If* $0 < \xi' \leqq 1$, *and if* μ', μ'_1, *and* μ'_2 *are stationary sources such that*

$$\mu'(E) = \xi'\, \mu'_1(E) + (1 - \xi')\, \mu'_2(E) , \quad E \in \mathbf{F}^I ,$$

then

$$\underline{\lim_n} \frac{1}{n} \log L_n(\varepsilon, \mu') \geqq \underline{\lim_n} \frac{1}{n} \log L_n\left(\frac{\varepsilon}{\xi'}, \mu'_1\right) \quad \text{for} \quad 0 < \varepsilon < \xi' .$$

Proof. The assertion of the lemma coincides with Lemma 5.2 stated in [15] for the special case of stationary sources.

4. MAIN THEOREMS ON ASYMPTOTIC RATE

In this section we shall apply the lemmas of the preceding section to prove the main theorems on the existence and explicit representation of the asymptotic rate.

Theorem I. *Let* μ *be a stationary source. Then it holds the inequality*

$$\overline{\lim_n} \frac{1}{n} \log L_n(\varepsilon_1, \mu) \leqq \underline{\lim_n} \frac{1}{n} \log L_n(\varepsilon_2, \mu) \quad \text{for} \quad 0 < \varepsilon_2 < \varepsilon_1 < 1 ;$$

consequently, the limit

$$\lim_n \frac{1}{n} \log L_n(\varepsilon, \mu) = H_\varepsilon(\mu)$$

exists except at most a countable set of numbers ε. The function $H_\varepsilon(\mu)$ monotonically increases for $\varepsilon \searrow 0$ to a limit, which will be denoted by $H(\mu)$ and called the asymptotic rate of the source μ.

Proof. It is sufficient to prove the above inequality. In the proof we shall repeat the method used in [15], but applying the basic lemmas of the preceding section which are based on the ergodic theory of invariant measures only.

Let $\mu \in \mathcal{M}_{st}$. Let λ_1 and λ_2 be chosen arbitrarily such that $0 < \lambda_2 < \lambda_1 < 1$. If c is a non-negative extended real number, let us set

$$Z_1(c) = \{z: z \in R, \mathsf{H}(\mu_z) \leqq c\}, \tag{1}$$

$$Z_2(c) = \{z: z \in R, \mathsf{H}(\mu_z) \geqq c\}. \tag{2}$$

Define the numbers c_1 and c_2 by

$$c_1 = \inf\{c: \mu(Z_1(c)) \geqq 1 - \lambda_1\},$$

$$c_2 = \sup\{c: \mu(Z_2(c)) \geqq \lambda_2\}.$$

Then we have

$$1 - \xi = \mu(Z_1(c_1)) \geqq 1 - \lambda_1 > 0, \tag{3}$$

$$\xi' = \mu(Z_2(c_2)) = \geqq \lambda_2 > 0. \tag{4}$$

Let μ_1, μ_2 and μ_1', μ_2' be the sources which are defined by the equations

$$\mu_1(E) = (1 - \xi)^{-1} \mu(E \cap Z_1(c_1)), \tag{5}$$

$$\mu_1'(E) = (\xi')^{-1} \mu(E \cap Z_2(c_2)), \quad E \in \mathbf{F}^I; \tag{6}$$

$$\mu_2 \quad = \xi^{-1}[\mu - (1 - \xi)\mu_1] \quad \text{if} \quad \xi > 0,$$

$$\mu_2' \quad = (1 - \xi')^{-1}[\mu - \xi'\mu_1'] \quad \text{if} \quad \xi' < 1.$$

The stationarity of the sources just defined follows from the equality $\mu_{Tz} = \mu_z$ $(z \in R)$; let us mention that the latter fact guarantees the invariance of the sets (1) and (2). If $c_1 < +\infty$, then the source μ_1 satisfies the assumptions of the first basic lemma for $c = c_1$, as follows from (5), (3), and (1); we have

$$\overline{\lim_n}\, (1/n) \log L_n(\varepsilon, \mu_1) \leqq c_1, \quad 0 < \varepsilon < 1; \tag{7}$$

the latter inequality remains valid also for $c_1 = +\infty$. Now applying Lemma 7 and the inequality $\xi \leqq \lambda_1$ (cf. (3)), we deduce from (7) the inequality

$$\overline{\lim_n}\, (1/n) \log L_n(\varepsilon, \mu) \leqq c_1 \quad \text{for} \quad \lambda_1 < \varepsilon < 1. \tag{8}$$

If $c_2 < +\infty$, then the source μ_1' satisfies the assumptions of the second basic lemma, as follows from (6), (4), and (2); hence we obtain that

$$\underline{\lim}_n (1/n) \log L_n(\varepsilon, \mu_1') \geqq c_2 \,, \quad 0 < \varepsilon < 1 \,; \tag{9}$$

the latter inequality remains valid for $c_2 = +\infty$ since in such a case the source μ_1' satisfies the assumptions of Lemma II for any finite c so that the given lower limit cannot be finite. By an application of Lemma 8 and the inequality $\xi' \geqq \lambda_2$ (cf. (4)) we deduce from (9) the relation

$$\underline{\lim}_n (1/n) \log L_n(\varepsilon, \mu) \geqq c_2 \quad \text{for} \quad 0 < \varepsilon < \lambda_2 \,. \tag{10}$$

From the definitions of c_1 and c_2 it follows that

$$\mu\{z: z \in R, \mathsf{H}(\mu_z) > c_2\} \leqq \lambda_2 < \lambda_1 \leqq \mu\{z: z \in R, \mathsf{H}(\mu_z) \geqq c_1\} \,;$$

hence the inequality $c_2 \geqq c_1$ must hold. The latter inequality together with the inequalities (8) and (10) imply the first assertion of the theorem, Q.E.D.

Corollary. *If μ is an ergodic source, then the equalities*

$$\lim_n (1/n) \log L_n(\varepsilon, \mu) = \mathsf{H}(\mu) = H(\mu)$$

are valid for any ε, $0 < \varepsilon < 1$.

Proof. It follows from Lemma 3 that the numbers c_1 and c_2 both equal to the entropy rate $\mathsf{H}(\mu)$ independently of which λ_1 and λ_2 were chosen.

Let us make the convention that the essential supremum of a measurable function f on a measurable set Z with respect to a probability measure μ on $\boldsymbol{F}^I$ will be denoted as

$$\operatorname*{ess.\,sup}_{z \in Z[\mu]} f(z) = \inf\{t: \mu\{z: z \in Z, f(z) \leqq t\} = 1\} \,.$$

Theorem II. *The asymptotic rate of a stationary source μ equals the essential supremum of the entropy rates of its ergodic components:*

$$H(\mu) = \operatorname*{ess.\,sup}_{z \in R[\mu]} \mathsf{H}(\mu_z) \,.$$

Proof. Let us put $h = \text{ess. sup}\, \mathsf{H}(\mu_z)$. Let us assume that $H(\mu) > h$, i.e. $H(\mu) > c > h$ for some c. From the inequality $h < c$ we deduce that the source satisfies the assumptions of Lemma I for c, which yields the contradictory inequality $H(\mu) \leqq c$. Then the inequality $H(\mu) \leqq h$ must be valid.

Assume that $H(\mu) < h$; i.e. $H(\mu) < c_2 < h$ for some real number c_2. From here and from the definition of the essential supremum we obtain that $\mu(Z_2(c_2)) = \xi' > 0$ (cf. (2) above). Defining the sources μ_1' and μ_2' formally in the same way as in the proof of Theorem I (cf. (6)), we deduce from Lemma II applied to μ'_1 an inequality of the form (9), which yields together with Lemma 8 the contradictory inequality $H(\mu) \geqq c_2$. This proves the theorem.

As a final remark, let us mention that an axiomatic definition of the asymptotic rate was given by the author in his paper [16].

(Received December 1st, 1969.)

REFERENCES

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

[2] A. I. Khinchin: Mathematical foundations of information theory. Dover Publications, New York 1957.

[3] N. Kryloff - N. Bogoliouboff: La théorie générale de la mesure dans son application à l'étude des systèmes dynamique de la mécanique non linéaire. Ann. of Math. *38* (1937), 65—113.

[4] B. McMillan: The basic theorems of information theory. Ann. Math. Stat. *24* (1953), 196—219.

[5] J. C. Oxtoby: Ergodic sets. Bull. Amer. Math. Soc. *58* (1952), 116—136.

[6] K. R. Parthasarathy: On the integral representation of the rate of transmission of a stationary channel. Illinois Journ. of Math. *5* (1961), 2, 299—305.

[7] K. R. Parthasarathy: A note on McMillan's theorem for countable alphabets. Transact. Third Prague Conf. on Inform. Theory etc., Prague 1964, 541—543.

[8] K. R. Parthasarathy: Effective entropy rate and transmission of information through channels with additive random noise. Manuscript, 1962.

[9] A. Perez: Notions généralizées d'incertitude, d'entropie et d'information. Transact. First Prague Conf. on Inform. Theory etc., Prague 1957, 183—208.

[10] A. Perez: Sur la théorie de l'information dans le cas d'un alphabet abstrait. Transact. First Prague Conf. on Inform. Theory etc., Prague 1957, 209—243.

[11] A. Perez: Extensions of Shannon-McMillan's limit theorem to more general stochastic processes. Transact. Third Prague Conf. on Inform. Theory etc., Prague 1964, 545—574.

[12] V. A. Rohlin: New progress in the theory of transformation with invariant measure (In Russian). Usp. matem. nauk *15* (1960), 3—26.

[13] C. E. Shannon: A mathematical theory of communication. Bell System Tech. Journ. *27* (1948), Part I, 379—423.

[14] I. G. Sinaĭ: On the flows with finite entropy (In Russian). Dokl. Akad. Nauk 125 (1959), *6*, 1200—1202.

[15] K. Winkelbauer: On discrete information sources. Transact. Third Prague Conf. on Inform. Theory etc., Prague 1964, 765—830.

[16] K. Winkelbauer: Axiomatic definition of channel capacity and entropy rate. Transact. Fourth Prague Conf. on Inform. Theory etc., Prague 1967, 661—705.

[17] K. Winkelbauer: Channels with finite past history. Transact. Second Prague Conf. on Inform. Theory etc., Prague 1960, 685—831.

O asymptotické neurčitosti neergodických zdrojů informace

KAREL WINKELBAUER

Autor ukázal ve své práci [15], že entropie není vhodnou informační charakteristikou pro případ neergodických stacionárních zdrojů informace, a proto zavedl novou charakteristiku informační vydatnosti, která je založena na pojmu logaritmické ε-neurčitosti a je nazvána asymptotickou neurčitostí stacionárního diskrétního zdroje informace. Již v práci [15] bylo ukázáno, že asymptotická neurčitost, a nikoli entropie, má všechny potřebné vlastnosti efektivní míry množství informace.

Důkazy uvedených důležitých vlastností asymptotické neurčitosti opřel autor v citované práci o obecnou teorii přenosu informace diskrétními sdělovacími kanály. Avšak stacionární zdroj informace představuje z matematického hlediska invariantní míru, tedy matematický objekt značně jednodušší, než jsou sdělovací kanály. Vznikl problém, zda je principiálně možné založit důkazy o existenci a explicitním vyjádření asymptotické neurčitosti na metodách nevybočujících z rámce teorie invariantních měr.

Tento problém je řešen v této stati, a to v kladném smyslu. Ukazuje se, že důkazy hlavních vět o asymptotické neurčitosti lze opřít o jemnější prostředky teorie invariantních měr, jimiž jsou ergodická teorie a upřesněná verse McMillanovy věty.

Doc. RNDr. Mg. Mat. Karel Winkelbauer, DrSc., Ústav teorie informace a automatizace ČSAV, Vyšehradská 49, Praha 2.