KYBERNETIKA — VOLUME 16 (1980), NUMBER 2

Epsilon-Rates, Epsilon-Quantiles, and Group Coding Theorems for Finitely Additive Information Sources

ŠTEFAN ŠUJAN

Finitely additive information sources are investigated with a countably infinite alphabet having the structure of a free finitely generated abelian group. The epsilon-rates that determine the optimum encoding rates for group codes associated with finite factor groups of the alphabet are related with the epsilon-quantiles of the corresponding entropy functions. The resulting group coding theorems are formulated.

1. INTRODUCTION

As well-known, the entropy fails to be an effective measure of uncertainty for stationary non-ergodic information sources [9, 12, 13]. Moreover, the optimum encoding rates for such sources depend on the error probability [10]. In the present paper we analyze the conditions under which the ideas of [10] work in the more general case of finitely additive sources [9]. Accordingly, much place is devoted to the concept of entropy in the finitely additive setting that has been introduced in [9] rather formally.

2. BASIC NOTATIONS

Throughout the paper the symbols CA and FA will be used as abbreviations for "countably additive" and "finitely additive", respectively. CA is assumed to be a special case of FA.

Let A be a countable set. A set $E \subset A^I$ (I = integers) is said to be a finite-dimensional cylinder (in symbols, $E \in \mathscr{K}_A$) iff there are $J \subset I$ with $0 < \operatorname{card}(J) < \infty$ and $C \subset A^J$ such that

$$E = \{z \in A^I : (z_j)_{j \in J} \in C\} .$$

Especially, if $C = \{\bar{x}\}$, $\bar{x} \in A^J$, E is said to be an elementary cylinder (in symbols, $E \in \mathscr{V}_A$). As well-known, A^I is a Polish space in its product topology derived from the discrete one in A. The sigma-field $\mathscr{F}_A = \sigma\mathscr{K}_A (= \sigma\mathscr{V}_A)$ consists precisely of all Borel subsets of A^I. We shall use the notations:

$$(1) \qquad [C] = \{z \in A^I : (z_0, \ldots, z_{n-1}) \in C\} \quad \text{for} \quad C \subset A^n ;$$

$$(2) \qquad [\bar{x}] = [\{\bar{x}\}] \quad \text{for} \quad \bar{x} \in A^n, n \in N = \{1, 2, \ldots\} .$$

Let S_A denote the shift in A^I:

$$(3) \qquad (S_A z)_i = z_{i+1} \quad \text{for} \quad z \in A^I \quad \text{and} \quad i \in I .$$

Any S_A-invariant FA probability on $\mathscr{K}_A$ is said to be a *source*. The set M_A of all sources is non-empty, convex, and contains extreme points (= ergodic FA probabilities, cf. [4] and [9]). Let $\mathscr{B}_A$ be the sigma-field on the set W_A of all ergodic elements of M_A defined by

$$(4) \qquad \mathscr{B}_A = \sigma\{\{\mu \in W_A : \mu(E) \leqq t\} : E \in \mathscr{K}_A, 0 \leqq t \leqq 1\} .$$

To every source $m \in M_A$ there corresponds a unique CA probability m_0 on $\mathscr{B}_A$ such that

$$(5) \qquad m(E) = \int_{W_A} \mu(E)\, m_0(\mathrm{d}\mu) , \qquad E \in \mathscr{K}_A .$$

For the proofs of these and other results from FA ergodic theory see [4] and [9]. Let R_A denote the set of all regular points in A^I [5], [13]. Let μ_z denote the S_A-ergodic CA probability on $\mathscr{F}_A$ determined uniquely by $z \in R_A$ [13]. We can identify $\operatorname{ext}(M_A \cap \mathrm{CA}) = W_A \cap \mathrm{CA}$ with R_A and $\mathscr{B}_A$ with $R_A \cap \mathscr{F}_A = \{R_A \cap E : E \in \mathscr{F}_A\}$, respectively. Thus, for $m \in M_A \cap \mathrm{CA}$, (5) becomes the usual ergodic decomposition formula

$$(6) \qquad m(E) = \int_{R_A} \mu_z(E)\, m(\mathrm{d}z) , \qquad E \in \mathscr{F}_A .$$

Finally, notice that $M_A = M_A \cap \mathrm{CA}$ in case $\operatorname{card}(A) < \infty$.

3. THE NOTION OF ENTROPY

Let $Z(\mathscr{F}_A)$ designate the set of all finite partitions ζ of A^I such that $\zeta \subset \mathscr{F}_A$. Let $\eta \leqq \zeta$ mean ζ refines η. The partial ordering $\leqq$ gives rise to a lattice structure in $Z(\mathscr{F}_A)$. Let, for $\zeta, \eta \in Z(\mathscr{F}_A)$, $\zeta \vee \eta$ denote the (least with respect to $\leqq$) common refinement of ζ and η. The set $Z(\mathscr{F}_A)$ is filtered to the right by means of $\leqq$. Analogous conclusions apply to all lattices of partitions met in the sequel. If $m \in M_A \cap \mathrm{CA}$ then define

$$(7) \qquad h_m(\zeta) = \sum_{C \in \zeta} m(C) \left|\log m(C)\right| \quad \text{for} \quad \zeta \in Z(\mathscr{F}_A).$$

The base of the logarithm is fixed but unspecified. Let exp denote the corresponding exponential. Let $h_m(S_A, \zeta)$ denote the entropy of the shift S_A with respect to ζ. Since m is S_A-invariant we have

$$(8) \qquad (1/n)\, h_m\left(\bigvee_{i=0}^{n-1} S_A^{-i}\zeta\right) \downarrow h_m(S_A, \zeta)$$

as $n \to \infty$. The quantity

$$(9) \qquad h_m(S_A) = \sup \{h_m(S_A, \zeta) : \zeta \in Z(\mathscr{F}_A)\}$$

was proposed by Sinai [7] as the definition of the entropy of the shift. In our terminology, $h_m(S_A) = h(m)$ – the entropy of the source $m \in M_A \cap \mathrm{CA}$.

Let $Z(\mathscr{K}_A) = \{\zeta \in Z(\mathscr{F}_A) : \zeta \subset \mathscr{K}_A\}$. Further, let $Z(A)$ denote the lattice of all finite partitions of the alphabet A. $Z(A)$ can be thought as a sublattice of $Z(\mathscr{K}_A)$:

$$\xi \mapsto [\xi] = \{[C] : C \in \xi\} \quad \text{for} \quad \xi \in Z(A) \quad (\text{cf. } 1)).$$

Let

$$(10) \qquad \xi^n = \{X \subset A^n : X = X_1 \times \ldots \times X_n, X_i \in \xi, 1 \leqq i \leqq n\}.$$

The above correspondence yields the assignment

$$\xi^n \mapsto \bigvee_{i=0}^{n-1} S_A^{i-}[\xi]$$

so that $h_m(S_A, [\xi])$ (cf. (8)) is well-defined. Let $\gamma_A = \{\{a\} : a \in A\}$. The general arguments due to Sinai [7] imply the following way of the computation of $h(m)$:

$$(11) \qquad h(m) = \lim_{n \to \infty} h_m(S_A, [\xi_n])$$

where $\{\xi_n, n \in N\} \subset Z(A)$ is any sequence such that $\xi_n \leqq \xi_{n+1} (n \in N)$ and $\lim_{n \to \infty} \xi_n = \gamma_A$. The formula is valid even if $h_m(\gamma_A) = +\infty$ (cf. (7)). For the proof see [7] or [6]. It involves the following two identifications:

$$(12) \qquad h(m) = \sup \{h_m(S_A, \zeta) : \zeta \in Z(\mathscr{K}_A)\};$$

$$(13) \qquad h(m) = \sup \{h_m(S_A, [\xi]) : \xi \in Z(A)\}.$$

The reduction from $Z(\mathscr{K}_A)$ to $Z(A)$ is a consequence of the structure of the space A^I and not of the properties of the source m. Therefore (13) was accepted as the definition of entropy in case when $m \in M_A - (M_A \cap \mathrm{CA})$ in [9].

On the other hand, the entropy in the CA case has the following desirable and important property. Given $m \in M_A \cap \mathrm{CA}$, $h(\mu_z)$ as the function of variable z is an almost everywhere mod m defined random variable on $(A^I, \mathscr{F}_A)$. Actually, $h_{\mu_z}(S_A, [\xi])$ can be easily shown to possess this property and the rest follows from (11). Let

$$h(\mu) = \sup \{h_\mu(S_A, [\xi]) : \xi \in Z(A)\} \tag{14}$$

for $\mu \in W_A$. Is $h(\cdot)$ a $\mathscr{B}_A$-measurable function on W_A? The next example shows that (11) fails to work in the FA case.

Example. Let $A = N$, let

$$\xi_n = \{\{1\}, \ldots, \{n\}, N - \{1, \ldots, n\}\} \quad \text{for} \quad n \in N .$$

Clearly $\xi_n \leqq \xi_{n+1}$ and $\lim_{n \to \infty} \xi_n = \gamma_N$. Now let m be a (memoryless) FA source given as the product of its one-dimensional marginals $m_i = m_0$ $(i \in I)$, where m_0 is a FA probability chosen below. The infinite products are well-defined even in the FA case (cf. [2], Chapter III). Let m_0 correspond to the model of a randomly chosen natural number (cf. [1] for the basic ideas and [9] and [11] for a rigorous treatment of this model). As well-known,

$$h_m(S_A, [\xi]) = h_{m_0}(\xi)$$

so that

$$h(m) = \sup \{h_{m_0}(\xi) : \xi \in Z(A)\} .$$

Let η_n be the partition of N into residue classes mod n. Then $h_{m_0}(\eta_n) = \log n$ so that $h(m) = +\infty$ and this properly reflects our idea of infinitely many equally likely cases. On the other hand,

$$m_0\{1\} = \ldots = m_0\{n\} = 0 , \quad m_0(N - \{1, \ldots, n\}) = 1$$

so that $h_{m_0}(\xi_n) = 0$ for any $n \in N$. Hence, (11) does not apply.

Since $Z(A)$ cannot be reduced to a sequence, the $\mathscr{B}_A$-measurability of $h_\mu(S_A, [\xi])$ does not entail the $\mathscr{B}_A$-measurability of $h(\mu)$ on W_A, in general. Of course, it will suffice if $Z(A)$ will contain a confinal sequence. But this is apparently not true. Following [14] we add some algebraic properties to A.

Let A be a free abelian finitely generated group. Let $Z_0(A)$ denote the lattice of all finite factor groups of A. For $\xi \in Z_0(A)$ by G_ξ is denoted the kernel of the natural homomorphism $A \to \xi$. We say that $\xi \in Z_0(A)$ is a divisor of $\eta \in Z_0(A)$ iff $G_\xi \subset G_\eta$, and write $\eta \leqq \xi$ because, if ξ and η are considered as partitions, $\eta \leqq \xi$ means simply that ξ refines η. As shown in [14] a sequence $(\eta_n, n \in N)$ exists in $Z_0(A)$ such that $\eta_{n+1} \geqq \eta_n (n \in N)$ and to every $\xi \in Z_0(A)$ there is n_0 with $\eta_{n_0} \geqq \xi$ (hence, $\eta_n \geqq \xi$ for all $n \geqq n_0$). In other words, $(\eta_n, n \in N)$ is a sequence cofinal with $Z_0(A)$.

Definition 1. Let A be a free finitely generated abelian group. The *entropy* $h(m)$ of a source $m \in M_A$ is defined by

$$(15) \qquad h(m) = \sup \{h_m(S_A, [\xi]) : \xi \in Z_0(A)\} ,$$

where $Z_0(A)$ is the lattice of all finite factor groups of A.

4. THE GENERAL CONSTRUCTION OF RANDOM VARIABLES

By rephrasing the properties of a cofinal sequence in $Z_0(A)$ and by taking some elementary properties of measurable functions into account we get the following general principle that will be used to construct the relevant entropy functions below.

Proposition 1. Let $(f_\xi;\ \xi \in Z_0(A))$ be a net of $\mathscr{B}_A$-measurable non-negative functions on W_A. If, for any fixed $\mu \in W_A$, the net $(f_\xi(\mu);\ \xi \in Z_0(A))$ is monotonically increasing then its limit $f(\mu) = \sup \{f_\xi(\mu) : \xi \in Z_0(A)\}$ is a non-negative number (possibly $+\infty$) for which

$$f_{\eta_n}(\mu) \uparrow f(\mu) \quad \text{as} \quad n \to \infty .$$

Consequently, $f = (f(\mu),\ \mu \in W_A)$ is a non-negative, extended real-valued, $\mathscr{B}_A$-measurable function on W_A.

Now let us concern the reduction from (12) to (13). Now the lattice $Z(\mathscr{K}_A)$ has to be replaced by $Z_0(\mathscr{K}_A)$ that contains precisely all $\zeta \in Z(\mathscr{K}_A)$ such that there are $n \in N$ and a finite factor group η of A^n with $\zeta = [\eta] = \{[Y] : Y \in \eta\}$ (cf. (1)).

Proposition 2. Let η be an arbitrary finite factor group of A^n $(n \in N)$. Then there exists $\xi \in Z_0(A)$ such that $\xi^n \geqq \eta$ (cf. (10) for the symbol ξ^n).

Proof. Due to the direct sum structure of A^n it suffices to explain the idea in case $A = I$ and $n = 2$. So let η be a finite factor group of $I \oplus I$. Then there are (linearly independent) generators a_1 and a_2 of the group $I \oplus I$ and non-negative integers q_1 and q_2 such that G_η is the free abelian group generated by the set $\{q_1 a_1, q_2 a_2\}$ (cf. e.g.) [3]. At least one of the numbers q_1 and q_2 is positive and at least one of them is greater than 1. If both are positive then we can always choose them in such a way that q_2 is divisible by q_1 – this is obvious if $q_1 = 1$. If both q_1 and q_2 are greater than 1 then we get finite factor groups with at least four elements. We illustrate only the first two possibilities that are less straightforward:

$$G = \{(0, 2i) : i \in I\} \text{ – the group generated by the set } \{0 \,.\, (1, 0), 2 \,.\, (0, 1)\} ;$$

$$G = \{(x, y) \in I^2 : \exists i \in I,\ x + y = 2i\} \text{ – the group generated by the set } \{1 \,.\, (-1, 1), 2 \,.\, (1, 0)\} .$$

Consequently, the things can always be arranged so that

$$G_\eta \supset q_2(I \oplus I) = \{q_2 x : x \in I \oplus I\} = \{(q_2 i, q_2 j) : i, j \in I\} .$$

Hence, $I \oplus I/q_2(I \oplus I) \geqq \eta$. Let $\xi = I/q_2 I$. Since $q_2 > 1$, $\xi \in Z_0(A)$ [14]. Now a straightforward verification yields

$$G_{\xi^2} = \{(x, y) \in I^2 : \exists i, j \in I,\ x = q_2 i,\ y = q_{2j}\} = q_2(I \oplus I) .$$

Thus $G_{\xi^2} \subset G_\eta$ so that $\xi^2 \geqq \eta$, QED.

5. THE BASIC QUANTITIES

Let us start with the concept of a finite memory source. Let $\xi \in Z_0(A)$. We can consider ξ as a new alphabet and define a mapping $T_\xi : A^I \to \xi^I$ by the property that

$$(T_\xi z)_i = X \quad \text{iff} \quad S_A^i z \in [X] , \quad X \in \xi .$$

Then $T_\xi^{-1} \mathscr{K}_\xi \subset \mathscr{K}_A$ so that, given $m \in M_A$, we can define a CA source m_ξ on $\mathscr{F}_\xi = \sigma \mathscr{K}_\xi$ by means of the relations

$$m_\xi(E) = m(T_\xi^{-1} E) , \quad E \in \mathscr{K}_\xi . \tag{16}$$

A source $m \in M_A$ is said to be Markov of order k ($k \geqq 0$ integer) if, for any $\xi \in Z_0(A)$, m_ξ corresponds to an S_ξ-stationary finite Markov chain of order k with the state space ξ and $m_\xi(E) > 0$ for all $E \in \mathscr{V}_\xi$. A source $m \in M_A$ is of finite memory ($m \in M_A(FM)$) iff there is $k \geqq 0$ such that m is Markov of order k (cf. [9] for the examples).

Following the relation (8) and using the fact that $[\xi]^n = [\xi^n]$ if $[\xi]^n$ is interpreted as $\bigvee\{S_A^{-i}[\xi]: i = 0, \ldots, n-1\}$ we easily see that

$$h_m(S_A, [\xi]) = h(m_\xi) \quad \text{(cf. 16))} .$$

Let

$$f_\xi^{(1)}(\mu) = h(\mu_\xi) ; \quad \xi \in Z_0(A) , \quad \mu \in W_A . \tag{17}$$

For any pair of sources $(m, \tilde{m}) \in M_A \times M_A(FM)$ we define

$$k(m_\xi, \tilde{m}_\xi) = -\lim_{n \to \infty} (1/n) \int \log \tilde{m}_\xi[z_1, \ldots, z_n]\, m_\xi(\mathrm{d}z) . \tag{18}$$

Let

$$f_\xi^{(2)}(\mu) = k(\mu_\xi, \tilde{m}_\xi) ; \quad \xi \in Z_0(A) , \quad \mu \in W_A . \tag{19}$$

Finally, let

$$d(m_\xi, \tilde{m}_\xi) = k(m_\xi, \tilde{m}_\xi) - h(m_\xi) = \lim_{n \to \infty} (1/n) \int \log \frac{m_\xi[z_1, \ldots, z_n]}{\tilde{m}_\xi[z_1, \ldots, z_n]}\, m_\xi(\mathrm{d}z) . \tag{20}$$

We denote

$$f_\xi^{(3)}(\mu) = d(\mu_\xi, \tilde{m}_\xi)\,; \quad \xi \in Z_0(A)\,, \quad \mu \in W_A\,. \tag{21}$$

Let

$$f^{(i)}(\mu) = \sup_{\xi \in Z_0(A)} f_\xi^{(i)}(\mu)\,, \quad i = 1, 2, 3\,. \tag{22}$$

Remark 1. We can formally define also the quantities $f^{(i)}(m)$ ($i = 1, 2, 3$; $m \in M_A$). These quantities were called in [9] the *entropy rate* ($i = 1$), the *K-entropy* ($i = 2$), and the *I-entropy* ($i = 3$), respectively. We can prove that

$$f^{(i)}(m) = \int f^{(i)}(\mu)\, m_0(\mathrm{d}\mu) \quad (i = 1, 2, 3)$$

(cf. the relation (5); the proof of the above integral representation formula is given in [8] for the case $i = 1$, the proofs in the remaining cases follow the same idea). These quantities, however, differ from those ones introduced in [9] because we are using a different lattice of partitions.

Proposition 3. The functions $f^{(i)}$ defined by (22) are non-negative extended real-valued random variables on the probability space $(W_A, \mathscr{B}_A, m_0)$ for any $m \in M_A$ ($i = 1, 2, 3$).

Proof. In order an application of Proposition 1 be justified we must show

(I) The functions $f_\xi^{(i)}$ ($i = 1, 2, 3$) are non-negative and $\mathscr{B}_A$-measurable on W_A for any $\xi \in Z_0(A)$.

(II) Given $\mu \in W_A$, the net $(f_\xi^{(i)}(\mu),\ \xi \in Z_0(A))$ is monotonically increasing ($i = 1, 2, 3$).

(I) *Measurability*. Case $i = 1$. We know that

$$f_\xi^{(i)}(\mu) = h(\mu_\xi) = h_\mu(S_A, [\xi]) = \lim_{n\to\infty} (1/n)\, h_\mu([\xi]^n) = \\ = \lim_{n\to\infty} (1/n) \sum_{C\in[\xi]^n} \mu(C)\, |\log \mu(C)|\,.$$

Now

$$\{\mu \in W_A : f_\xi^{(1)}(\mu) \leqq t\} = \\ = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{\mu \in W_A : \sum_{C\in[\xi]^n} \mu(C)\, |\log \mu(C)| \leqq nt\} \in \mathscr{B}_A\,.$$

Case $i = 2$. By definition

$$k(\mu_\xi, \tilde{m}_\xi) = \lim_{n\to\infty} (1/n) \int |\log \tilde{m}_\xi[z_1, \ldots, z_n]|\, \mu_\xi(\mathrm{d}z) =$$

$$= \lim_{n\to\infty} (1/n) \sum_{\bar{x}\in[\xi]^n} \mu_\xi[\bar{x}] \left|\log \tilde{m}_\xi[\bar{x}]\right| .$$

Since $\mu_\xi[\bar{x}]$ is measurable as a function of the variable μ, the sum is also measurable so that $k(\mu_\xi, \tilde{m}_\xi)$ is $\mathscr{B}_A$-measurable as well. By (20), $d(\mu_\xi, \tilde{m}_\xi)$ is a $\mathscr{B}_A$-measurable function of the variable $\mu \in W_A$, too.

(II) *Monotonicity*. We shall treat the three cases separately because through the proofs new quantities will arise of importance in the sequel.

Case $i = 1$. Let $m \in M_A$, $\xi \in Z_0(A)$, $n \in N$, and $0 < \varepsilon < 1$, respectively. Let

$$L_n(\varepsilon, m_\xi) = \min \{\text{card}(\Delta) : \Delta \subset \xi^n, m_\xi[\Delta] > 1 - \varepsilon\} . \tag{23}$$

If $m \in M_A \cap \text{CA}$ then

$$L_n(\varepsilon, m_{\gamma_A}) = L_n(\varepsilon, m)$$

in the notation of [12] and [13]. If $\mu \in W_A$ then μ_ξ is an S_ξ-ergodic source on $(\xi^I, \mathscr{F}_\xi)$ and $f_\xi^{(1)}(\mu) < \log \text{card}(\xi) < \infty$ so that Lemma, p. 10 of [14] applies and we get

$$(1/n) \log L_n(\varepsilon, \mu_\xi) \to f_\xi^{(1)}(\mu)$$

as $n \to \infty$ for any ε, $0 < \varepsilon < 1$. For fixed ε, $\eta \precsim \xi$ implies $L_n(\varepsilon, \mu_\xi) \geqq L_n(\varepsilon, \mu_\eta)$ $(\xi, \eta \in Z_0(A),\ \mu \in W_A,\ n \in N)$ so that the desired monotonicity follows.

Case $i = 2$. Let $(m, \tilde{m}) \in M_A \times M_A(FM)$, $\xi \in Z_0(A)$, $n \in N$, and $0 < \varepsilon < 1$, respectively. Let

$$L_n(\varepsilon, m_\xi, \tilde{m}_\xi) = \min \{ \sum_{\bar{x}\in\Delta} m_\xi[\bar{x}]/\tilde{m}_\xi[\bar{x}] : \Delta \subset \xi^n, m_\xi[\Delta] > 1 - \varepsilon\} . \tag{24}$$

If card $(A) < \infty$ then

$$L_n(\varepsilon, m_{\gamma_A}, \tilde{m}_{\gamma_A}) = L_n(\varepsilon, m, \tilde{m})$$

in the notation of [9]. Especially, if $\mu \in W_A$, then

$$(1/n) \log L_n(\varepsilon, \mu_\xi, \tilde{m}_\xi) \to f_\xi^{(2)}(\mu)$$

as $n \to \infty$ for any $0 < \varepsilon < 1$ (cf. [9], Corollary 21.11). For fixed ε, $\eta \precsim \xi$ implies $L_n(\varepsilon, \mu_\eta, \tilde{m}_\eta) \leqq L_n(\varepsilon, \mu_\xi, \tilde{m}_\xi)$ $(\xi, \eta \in Z_0(A),\ \mu \in W_A,\ n \in N)$. Hence the net $(f_\xi^{(2)}(\mu),\ \xi \in Z_0(A))$ is monotonically increasing for any $\mu \in W_A$.

Case $i = 3$. Let $(m, \tilde{m}) \in M_A \times M_A(FM)$, $\xi \in Z_0(A)$, $n \in N$, and $0 < \varepsilon < 1$, respectively. Let

$$I_n(\varepsilon, m_\xi, \tilde{m}_\xi) = \min \{\tilde{m}_\xi[\Delta] : \Delta \subset \xi^n, m_\xi[\Delta] > 1 - \varepsilon\} . \tag{25}$$

If card $(A) < \infty$ then

$$I_n(\varepsilon, m_{\gamma_A}, \tilde{m}_{\gamma_A}) = I_n(\varepsilon, m, \tilde{m})$$

in the notation of [9]. Especially, if $\mu \in W_A$, then

$$(-1/n) \log I_n(\varepsilon, \mu_\xi, \tilde{m}_\xi) \to f_\xi^{(3)}(\mu)$$

as $n \to \infty$ for all $0 < \varepsilon < 1$ (cf. [9], Corollary 22.11). Let $\eta \leqq \xi$ ($\xi, \eta \in Z_0(A)$). Then there is $k \in N$ such that card $(\xi) =$ card $(\eta) + k$. The inequality

$$-\log I_n(\varepsilon, \mu_\eta, \tilde{m}_\eta) \leqq -\log I_n(\varepsilon, \mu_\xi, \tilde{m}_\xi)$$

follows by induction on k from Lemma 23.3 of [9], QED.

6. EPSILON-RATES AND - QUANTILES

Let $m \in M_A$, $\tilde{m} \in M_A(FM)$, $\xi \in Z_0(A)$, respectively. Let

$$V_\varepsilon(m_\xi) = \lim_{n \to \infty} (1/n) \log L_n(\varepsilon, m_\xi) .$$

Definition 2. The quantity

$$V_\varepsilon(m) = \sup \{V_\varepsilon(m_\xi) : \xi \in Z_0(A)\} \tag{26}$$

is said to be the *epsilon-rate* of the source m.
Similarly, let

$$V_\varepsilon(m_\xi, \tilde{m}_\xi) = \lim_{n \to \infty} (1/n) \log L_n(\varepsilon, m_\xi, \tilde{m}_\xi) ,$$

$$I_\varepsilon(m_\xi, \tilde{m}_\xi) = \lim_{n \to \infty} (-1/n) \log I_n(\varepsilon, m_\xi, \tilde{m}_\xi) ,$$

respectively.

Definition 2 (continued). The quantity

$$V_\varepsilon(m, \tilde{m}) = \sup \{V_\varepsilon(m_\xi, \tilde{m}_\xi) : \xi \in Z_0(A)\} \tag{27}$$

is said to be the *epsilon-K-rate* of the pair $(m, \tilde{m})$ of sources. The quantity

$$I_\varepsilon(m, \tilde{m}) = \sup \{I_\varepsilon(m_\xi, \tilde{m}_\xi) : \xi \in Z_0(A)\} \tag{28}$$

is said to be the *epsilon-I-rate* of the pair $(m, \tilde{m})$ of sources.

Let f be a non-negative, extended real-valued, $\mathscr{B}_A$-measurable function defined on W_A. The (lower) *epsilon-quantile* of f with respect to a source $m \in M_A$ is defined as

$$Q(\varepsilon, m, f) = \inf \{t : m_0\{\mu \in W_A : f(\mu) \leqq t\} \geqq \varepsilon\} , \tag{29}$$

where m_0 corresponds to $m \in M_A$ by (5). The function $Q(., m, f)$ is left-continuous in $0 < \varepsilon < 1$ so that it is defined also for $\varepsilon = 1$ and

$$Q(1, m, f) = \inf\{t : m_0\{\mu \in W_A : f(\mu) \leqq t\} = 1\} = \operatorname*{ess.\,sup}_{\mu \in W_A[m_0]} f(\mu) . \tag{30}$$

Proposition 4. Let $m \in M_A$ and $\tilde{m} \in M_A(FM)$, respectively. Let $f^{(i)}$ $(i = 1, 2, 3)$ be the random variables defined by (22). Then

$$Q(1, m, f^{(1)}) = \lim_{\varepsilon \downarrow 0} V_\varepsilon(m) ,$$
$$Q(1, m, f^{(2)}) = \lim_{\varepsilon \downarrow 0} V_\varepsilon(m, \tilde{m}) ,$$
$$Q(1, m, f^{(3)}) = \lim_{\varepsilon \downarrow 0} I_\varepsilon(m, \tilde{m}) ,$$

Proof. In case $i = 1$ see [9], Theorem 15.3. In cases $i = 2, 3$ cf. [9], (23.1) and (23.3). Even the proofs are identical, Proposition 4 represents a different assertion; see Remark 1.

As we already know,

$$f_\xi^{(i)}(\mu) = f^{(i)}(\mu_\xi) \quad \text{for} \quad \xi \in Z_0(A) \quad \text{and} \quad \mu \in W_A .$$

Since μ_ξ is S_ξ-ergodic,

$$\mu_\xi\{z \in R_\xi : \mu_z = \mu_\xi\} = 1 \tag{31}$$

(cf. [12], [13]). Consequently, we can define the functions $f^{(i)}$ on R_ξ $(\xi \in Z_0(A))$ by the properties

$$f^{(i)}(z) = f^{(i)}(\mu_z) = f_\xi^{(i)}(\mu) \quad (i = 1, 2, 3) , \tag{32}$$

where $\mu \in W_A$ is such that, for μ_ξ, the relation (31) takes place. By definitions (cf. [13], Lemma 5 and (1.8) in case $i = 1$; [9], (21.1) and (21.3) in the remaining cases) the functions $f^{(i)}$ are almost everywhere mod m_ξ defined (for any $m \in M_A$), $\mathscr{F}_\xi$-measurable and S_ξ-invariant.

Lemma 1. Let $m \in M_A$, $\tilde{m} \in M_A(FM)$ and $\xi \in Z_0(A)$, respectively. For any real number t we have

$$m_0\{\mu \in W_A : f_\xi^{(i)}(\mu) \leqq t\} = m_\xi\{z \in R_\xi : f^{(i)}(z) \leqq t\}$$

for $i = 1, 2, 3$.
For the proof see [8] or [14].

Theorem 1. Let $0 < \delta < \varepsilon < 1$. Then for any finite factor group ξ of the alphabet A we have

$$\limsup_{n \to \infty} (1/n) \log L_n(\varepsilon, m_\xi) \leqq Q(1 - \delta, m, f^{(1)}) ;$$

$$\limsup_{n\to\infty} (1/n) \log L_n(\varepsilon, m_\xi, \tilde{m}_\xi) \leqq Q(1-\delta, m, f^{(2)}) ;$$

$$\limsup_{n\to\infty} (-1/n) \log I_n(\varepsilon, m_\xi, \tilde{m}_\xi) \leqq Q(1-\delta, m, f^{(3)}) .$$

Proof. Case $i = 1$. Let $t = Q(1-\delta, m, f^{(1)})$. By definition,

$$m_0\{\mu \in W_A : f^{(1)}(\mu) \leqq t\} \geqq 1 - \delta$$

so that

$$m_0\{\mu \in W_A : f^{(1)}_\xi(\mu) \leqq t\} \geqq 1 - \delta$$

for any $\xi \in Z_0(A)$. By Lemma 1,

$$m_\xi\{z \in R_\xi : f^{(1)}(z) \leqq t\} \geqq 1 - \delta .$$

Let $D = \{z \in R_\xi : f^{(1)}(z) \leqq t\}$. Then $D \in \mathscr{F}_\xi$ and D is S_ξ-invariant so that the relations

$$m'(E) = m_\xi(E \cap D)/m_\xi(D) , \quad E \in \mathscr{F}_\xi$$

define a new source $m' \in M_\xi = M_\xi \cap \mathrm{CA}$. Let $m'' \in M_\xi$ be such that

$$m_\xi(E) = (1-\alpha)\, m'(E) + \alpha\, m''(E) ,$$

where $1 - \alpha = m_\xi(D)$. Then $1 - \alpha \geqq 1 - \delta > 1 - \varepsilon$ so that $\alpha < \varepsilon < 1$. Consequently, Lemma 7 of [13] applies:

$$\limsup_{n\to\infty} (1/n) \log L_n(\varepsilon', m_\xi) \leqq \limsup_{n\to\infty} (1/n) \log L_n(\varepsilon' - \alpha, m')$$

for $\alpha < \varepsilon' < 1$. At the same time, $m'(D) = 1$ so that Lemma I of [13] applies to m':

$$\limsup_{n\to\infty} (1/n) \log L_n(\varepsilon', m') \leqq t , \quad 0 < \varepsilon' < 1 .$$

Now let $\varepsilon > \delta$. Since $\delta \geqq \alpha$, we have $\varepsilon > \alpha$ and so both obtained inequalities work and give, by the definition of t, the desired assertion in case $i = 1$.
Case $i = 2$. Replace Lemma 7 of [13] by a similar assertion for $L_n(\varepsilon, m_\xi, \tilde{m}_\xi)$ (the elementary proof is omitted). Use [9], Lemma 21.7 in place of [13], Lemma I.
Case $i = 3$. The proof is the same using now Lemmas 22.4 and 22.7 of [9], respectively, QED.

Theorem 2. Let $0 < \varepsilon < \delta < 1$, let $t < Q(1-\delta, m, f^{(i)})$ $(i = 1, 2, 3)$. Then there exists a finite factor group of the alphabet A such that

$$\liminf_{n\to\infty} (1/n) \log L_n(\varepsilon, m_\xi) \geqq t \quad (i = 1) ;$$

$$\liminf_{n\to\infty} (1/n) \log L_n(\varepsilon, m_\xi, \tilde{m}_\xi) \geqq t \quad (i = 2) ;$$

$$\liminf_{n\to\infty} (-1/n) \log I_n(\varepsilon, m_\xi, \tilde{m}_\xi) \geqq t \quad (i = 3)\,.$$

Proof. Case $i = 1$. Let $t < Q(1 - \delta, m, f^{(1)})$. Then

$$m_0\{\mu \in W_A : f^{(1)}(\mu) \leqq t\} \leqq 1 - \delta$$

so that

$$m_0\{\mu \in W_A : f^{(1)}(\mu) > t\} \geqq \delta\,.$$

Since $f^{(1)}(\mu) = \sup\,\{f_{\xi_A}^{(1)}(\mu) : \xi \in Z_0(A)\}$, we can find $\xi \in Z_0(A)$ such that

$$m_0\{\mu \in W_A : f_\xi^{(1)}(\mu) > t\} = m_\xi\{z \in R_\xi : f^{(1)}(z) > t\} \geqq \delta\,.$$

Let $D = \{z \in R_\xi : f^{(1)}(z) > t\}$. Then $D \in \mathscr{F}_\xi$ and D is S_ξ-invariant so that the relations

$$m'(E) = m_\xi(E \cap D)/m_\xi(D)\,, \quad E \in \mathscr{F}_\xi$$

define a new source $m' \in M_\xi$. Let $m'' \in M_\xi$ be such that

$$m_\xi(E) = \alpha\, m'(E) + (1 - \alpha)\, m''(E)\,,$$

where $\alpha = m_\xi(D)$. Then $\alpha > \delta > \varepsilon > 0$, so that [13], Lemma 8 applies and we get

$$\liminf_{n\to\infty} (1/n) \log L_n(\varepsilon', m_\xi) \geqq \liminf_{n\to\infty} (1/n) \log L_n(\varepsilon'/\alpha, m')$$

for $0 < \varepsilon' < \alpha$. At the same time, $m'(D) = 1$, and so Lemma II of [13] applies for m':

$$\liminf_{n\to\infty} (1/n) \log L_n(\varepsilon', m') \geqq t\,, \quad 0 < \varepsilon' < 1\,.$$

Now let $\varepsilon < \delta$. Since $\delta < \alpha$, we have $\varepsilon < \alpha$ so that for any ε, $0 < \varepsilon < \delta$, both inequalities work and together yield the desired assertion in case $i = 1$.
Case $i = 2$. Repeat the proof of case $i = 1$, Lemma 8 of [13] being replaced by a similar assertion for the quantity $L_n(\varepsilon, m_\xi, \tilde{m}_\xi)$ and Lemma II of [13] being replaced by [9], Lemma 21.8.
Case $i = 3$. The same using now lemmas 22.5 and 22.8 of [9], QED.

Corollary 1. Let $1 - \varepsilon$ $(0 < \varepsilon < 1)$ be a continuity point of $Q(., m, f^{(i)})$ $(i = 1, 2, 3)$. Then $Q(1 - \varepsilon, m, f^{(i)})$ equals the corresponding epsilon rate (cf. (26) and the relations following it).
The proof follows immediately from Theorems 1 and 2.

7. GROUP CODES AND CODING THEOREMS

In the preceding sections we introduced different criteria by means of which we characterized the sets $\Delta \subset \xi^n$ exhausting the space of messages up to a prescribed error probability ε:

$$m_\xi[\Delta] > 1 - \varepsilon .$$

Following the usual notions of group coding theory we shall call any set Δ with the above properties an n-dimensional ε-code associated with the finite factor group ξ of the group alphabet A. Then we can restate the results of Section 6 in the language of coding assertions:

Theorem 3. Let $m \in M_A$, $\tilde{m} \in M_A(FM)$. Let $1 - \varepsilon$ be a continuity point of the quantile function $Q(., m, f^{(i)})$ $(i = 1, 2, 3)$.

I. Let $t' > Q(1 - \varepsilon, m, f^{(i)})$ $(i = 1, 2, 3)$. Then for every finite factor group ξ of the alphabet A there is a natural number n_0 such that, for $n \geqq n_0$, there is an n-dimensional ε-code Δ associated with ξ such that

$$\left\{\begin{array}{l} \operatorname{card}(\Delta) < \exp(nt') \quad (i = 1); \\ \sum\limits_{\bar{x} \in \Delta} m_\xi[\bar{x}]/\tilde{m}_\xi[\bar{x}] < \exp(nt') \quad (i = 2); \\ \tilde{m}_\xi[\Delta] > \exp(-nt') \quad (i = 3), \end{array}\right. \tag{33}$$

respectively.

II. Let $t'' < Q(1 - \varepsilon, m, f^{(i)})$ $(i = 1, 2, 3)$. Then there are a finite factor group ξ_0 of the alphabet A and a natural number n_0 such that, for any divisor ξ of ξ_0 and for any $n \geqq n_0$ every n-dimensional *ε-code* Δ associated with ξ satisfies

$$\left\{\begin{array}{l} \operatorname{card}(\Delta) > \exp(nt'') \quad (i = 1); \\ \sum\limits_{\bar{x} \in \Delta} m_\xi[\bar{x}]/\tilde{m}_\xi[\bar{x}] > \exp(nt'') \quad (i = 2); \\ \tilde{m}_\xi[\Delta] < \exp(-nt'') \quad (i = 3), \end{array}\right. \tag{34}$$

respectively.

The assertion I represents the direct part of the coding theorem while II represents its converse. The dependence on ε in II is essential. Nevertheless, if $m = \mu \in W_A$ then we can easily deduce the next assertion that represents the strong converse of the coding theorem.

Corollary 2. Let us suppose that the general assumptions of Theorem 3 are satisfied. If $m = \mu \in W_A$ and if $t'' < Q(1, m, f^{(i)})$ $(i = 1, 2, 3)$ then assertion II of Theorem 3 is valid for any ε, $0 < \varepsilon < 1$, if n is large enough.

The proof follows from the fact that, given $\xi \in Z_0(A)$, the functions $f^{(i)}$ on R_ξ are constant mod μ_ξ (since they are S_ξ-invariant and μ_ξ is S_ξ-ergodic, respectively).

It is intuitively clear that the limits of the epsilonrates as $\varepsilon \downarrow 0$ should provide us with an asymptotic characterization of the n-dimensional ε-codes uniformly in $0 < \varepsilon < 1$. Formally, we have the following.

Corollary 3. Let us suppose that the general assumptions of Theorem 3 are satisfied. If $t' > Q(1, m, f^{(i)})$ (in case when $Q(1, m, f^{(i)}) < \infty$) then the assertion I of Theorem 3 is valid for any ε, $0 < \varepsilon < 1$, provided n is sufficiently large.
Indeed, if $t' > Q(1, m, f^{(i)})$ then for all $0 < \varepsilon < 1$, $t' > Q(1 - \varepsilon, m, f^{(i)})$ so that Theorem 3, I applies.

One can easily see that if m is an indecomposable source then

$$V_\varepsilon(m) = Q(1 - \varepsilon, m, f^{(1)}) = h(m),$$

the entropy of the source m which, according to [8], can be expressed as the mean

$$h(m) = \int_{W_A} f^{(1)}(\mu)\, m_0(\mathrm{d}\mu).$$

Since the above relations do not depend on ε, we have

$$Q(1, m, f^{(1)}) = h(m). \tag{35}$$

In order to make our investigations complete we have to specify that subclass of M_A for which (35) is valid. As shown in [9], Theorem 18.2 this is the subclass consisting precisely of all so-called strongly stable sources. For the details as well as for non-trivial examples of such sources the reader has to refer to [9], Section 18.

ACKNOWLEDGMENT

A preliminary version of the manuscript was prepared during the author's stay at Accademia Nazionale dei Lincei in Rome. Thanks are due to the Accademia, in particular, to Centro Linceo Interdisciplinare di Scienze Matematiche e Loro Applicazioni for the kind hospitality. I am gratefully indebted to prof. B. de Finetti and others for their encouraging interest.

(Received December 21, 1978.)

REFERENCES

[1] B. de Finetti: Probability, Induction and Statistics. J. Wiley, London—New York—Toronto 1972.

[2] N. Dunford, J. T. Schwartz: Linear Operators. Part I. Interscience, New York 1957.

[3] I. D. McDonald: The Theory of Groups. Clarendon Press, Oxford 1968.

[4] R. A. Olshen: Representing finitely additive invariant probabilities. Ann. Math. Stat. *39* (1968), 2131—2135.

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

[6] K. R. Parthasarathy: A note on McMillan's theorem for countable alphabets. Trans. 3rd Prague Conf. on Inf. Theory etc., NČSAV, Prague 1964, 541—543.

[7] Ja. G. Sinai: On the notion of entropy of a dynamical system (in Russian). Doklady AN SSSR *124* (1959), 768—771.

[8] Š. Šujan: Integral representation of the entropy rate. Studia Sci. Math. Hungarica *11* (1976) 25—36.

[9] Š. Šujan: A generalized coding problem for discrete information sources. Kybernetika *13* (1977), Supplement, 95 pp.

[10] Š. Šujan: Epsilon-Quantiles, Epsilon-Rates, and Source Coding Theorems (submitted to EIK).

[11] Š. Šujan: Ergodic theory and coding theorems for finitely additive invariant probabilities. Lecture held in October 1978 at Accademia Nazionale dei Lincei (text available upon request from the Accademia).

[12] K. Winkelbauer: On discrete information sources. Trans. 3rd Prague Conf. on Inform. Theory etc., NČSAV, Prague 1964, 765—830.

[13] K. Winkelbauer: On the asymptotic rate of non-ergodic information sources. Kybernetika *6* (1970), 127—148.

[14] K. Winkelbauer: Non-smooth channels with additive random noise. Trans. 8th Prague Conf. on Inform. Theory etc. Vol. B, 365—381. Academia, Prague 1978.

RNDr. Štefan Šujan, CSc., Ústav merania a meracej techniky SAV (Institute of Measurement and Measuring Technique, Slovak Academy of Sciences), 885 27 Bratislava. Czechoslovakia.