Kybernetika

Štefan Šujan

Channels with additive asymptotically mean stationary noise

Kybernetika, Vol. 17 (1981), No. 1, 1--15

Persistent URL: http://dml.cz/dmlcz/124379

KYBERNETIKA — VOLUME 17 (1981), NUMBER 1

CHANNELS WITH ADDITIVE ASYMPTOTICALLY MEAN STATIONARY NOISE

ŠTEFAN ŠUJAN

The paper is devoted to the study of problems connected with the Group Coding Theorem and its Converse for channels decomposable into components with additive asymptotically mean stationary and ergodic noise for codes associated with finite factor groups of the (countable) group alphabet.

1. INTRODUCTION

Parthasarathy [8] first touched the question on the asymptotic behavior of the maximum length of the n-dimensional ε-codes for separate values of the error probability $\varepsilon \in (0, 1)$. He answered the question for channels with additive stationary noise. These channels were shown to be regularly decomposable by Winkelbauer [15]. Consequently, his Theorem on ε-Capacities applies well and results in a connection between the behavior of the ε-codes and the information quantiles, respectively. As known, the concepts of information quantiles and information quantile capacity (see [15 – 19, 21] and [5]) arose from the attempts to overcome the problems pointed out by Nedoma [6] who demonstrated an example of a stationary channel whose Shannon capacity [9] strictly majorized the operational capacity (defined as optimum over actual deterministic codes; cf. [22]). An excellent survey on channel coding may be found in [3].

In this paper we consider channels decomposable into components with additive asymptotically mean stationary (a.m.s.) and ergodic noise (cf. [2] for the basic facts about a.m.s. measures). Thus, the results of [8] and partly of [15] are extended to a class of non-stationary channels. The technique of handling with an infinite alphabet is an adaptation of the techniques introduced first in [10, 11] and extended further in [12] and [20].

2. PRELIMINARIES AND RESULTS

Let X designate, generically, a countable discrete topological space — the alphabet. The corresponding space of messages is

$$X^I = \{z = \{z_i\}_{i \in I} : z_i \in X \text{ for } i \in I\}$$

(I = integers). The σ-algebra $\mathscr{F}_X$ generated by the countable class $\mathscr{V}_X$ of all elementary cylinders in X^I is but the σ-algebra of all Borel subsets of X^I with respect to its natural product topology. The shift transformation

$$(T_X z)_i = z_{i+1} \quad \text{for} \quad z \in X^I, \quad i \in I$$

is a Borel automorphism, that is, T_X is an invertible $\mathscr{F}_X$-measurable map. Let $\mathbf{P}_X$ denote the set of all probability measures on $(X^I, \mathscr{F}_X)$. Following [2] a measure $\mu \in \mathbf{P}_X$ is said to be *a.m.s.* (with respect to T_X) if the limits

$$\bar{\mu}(F) = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \mu(T_X^{-j} F) \tag{2.1}$$

exist for all $F \in \mathscr{F}_X$; in this case $\bar{\mu}$ is a T_X-invariant probability measure on $\mathscr{F}_X$ and is called the *stationary mean* of μ. We let $\mathbf{S}_X(\mathbf{M}_X)$ designate the set of all a.m.s. (of all T_X-invariant) elements of $\mathbf{P}_X$. If $\mu \in \mathbf{S}_X$ then $\mu = \bar{\mu}$ on the σ-algebra $\mathscr{I}_X$ of all T_X-invariant events so that it is natural to call μ *ergodic* if $\bar{\mu}$ is ergodic, i.e., if $\mu(F) = \bar{\mu}(F) \in \{0, 1\}$ for all $F \in \mathscr{I}_X$. For the sake of brevity we use the symbols $\mathbf{S}_X^*$ and $\mathbf{M}_X^*$ to denote the sets of all T_X-ergodic elements in $\mathbf{S}_X$ and in $\mathbf{M}_X$, respectively.

A channel with the input alphabet B and the output alphabet A will be denoted by $[B, \nu, A]$ or simply by the symbol ν. By definition, ν is a parametric class, $\nu = \{\nu(\cdot \mid y);\ y \in B^I\}$ with $\nu(\cdot \mid y) \in \mathbf{P}_A$ and such that the maps $y \mapsto \nu(G \mid y)$ are $\mathscr{F}_B$-measurable for all $G \in \mathscr{F}_A$. Given an input source $\mu \in \mathbf{P}_B$ the joint input/output distribution of the channel $[B, \nu, A]$ is a measure $\mu\nu \in \mathbf{P}_{B \times A}$ defined by the properties that

$$\mu\,\nu(E) = \int \nu(E_y \mid y)\, \mu(\mathrm{d}y)\,; \quad E \in \mathscr{F}_{B \times A}\,, \tag{2.2}$$
$$E_y = \{x \in A^I : (y, x) \in E\}\,.$$

Note that $T_{B \times A}$ is defined component-wise:

$$(T_{B \times A}(y, x))_i = ((T_B y)_i, (T_A x)_i)\,, \quad i \in I\,. \tag{2.3}$$

As well-known, a channel $[B, \nu, A]$ is said to be stationary if

$$\nu(T_A G \mid T_B y) = \nu(G \mid y) \quad \text{for} \quad y \in B^I\,, \quad G \in \mathscr{F}_A$$

and stationary and ergodic if, moreover,

$$(\mu \in \mathbf{M}_B^*) \Rightarrow (\mu\nu \in \mathbf{M}_{B \times A}^*)\,.$$

Fontana, Gray and Kieffer [1] called a channel $[B, \nu, A]$ *asymptotically mean stationary* (a.m.s.) if

$$(\mu \in \mathbf{S}_B) \Rightarrow (\mu\nu \in \mathbf{S}_{B \times A})$$

and *a.m.s. and ergodic* if, moreover,

$$(\mu \in \mathbf{S}_B^*) \Rightarrow (\mu\nu \in \mathbf{S}_{B \times A}^*) .$$

The a.m.s. property of channels is defined with the aid of the input sources so that it is a little bit complicated to obtain an input independent definition of the stationary mean for the channels. In order not to overcomplicate the exposition we take the properties established in [1, Theorem 3] as a definition. Accordingly, the stationary channel $[B, \bar{\nu}, A]$ is said to be the *stationary mean* of a channel $[B, \nu, A]$ if

$$\overline{\mu\nu} = \mu\bar{\nu} \quad \text{(cf. (2.1) and (2.2))} ; \tag{2.4}$$

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} \nu(T_A^{-j}G \mid T_B^{-j}y) = \bar{\nu}(G \mid y) \quad \mu\text{-a.e.} ; \tag{2.5}$$

$$\nu(\cdot \mid y) \ll \bar{\nu}(\cdot \mid y) \quad \mu\text{-a.e.} \tag{2.6}$$

for every $\mu \in \mathbf{M}_B$. In the quoted theorem it is proved that the stationary mean exists for every a.m.s. channel. Moreover, if $\bar{\nu}$ is also ergodic then ν is a.m.s. and ergodic (cf. [1, Lemma 4]).

Following [15] a channel $[B, \nu, A]$ is said to be *decomposable* if there are a measurable one-parameter family $\{[B, \nu^\lambda, A]; \lambda \in \Lambda\}$ of channels with the same alphabets with the parameters lying in the measurable space $(\Lambda, \mathscr{L})$ and a probability measure γ on $(\Lambda, \mathscr{L})$ such that

$$\nu(G \mid y) = \int_\Lambda \nu^\lambda(G \mid y)\, \gamma(\mathrm{d}\lambda) ; \quad y \in B^I , \quad G \in \mathscr{F}_A . \tag{2.7}$$

If $\Lambda = \{\lambda\}$ and $\gamma\{\lambda\} = 1$ then the channel ν is said to be indecomposable.

In what follows we suppose that $B = A$ is a countable Abelian (additively written) group and A^I as well as the finite powers A^n $(n \in N = \{1, 2, \ldots\})$ assume the natural direct sum structures. We prefer, however, to use distinct symbols to keep clear the distinction between the input and the output. A channel $[B, \nu, A]$ is said to be a channel with additive noise if there is a measure $\varkappa \in \mathbf{P}_B$ such that

$$\nu(G \mid y) = \varkappa(G - y) ; \quad y \in B^I , \quad G \in \mathscr{F}_A , \tag{2.8}$$

where $G - y$ designates the group theoretic difference in $A^I = B^I$. The class of channels of our interest consists of all channels decomposable into components with additive a.m.s. and ergodic noise. That is, channels are considered expressible in the form (2.7) such that each component ν^λ is determined by (2.8) with the corresponding measure $\varkappa^\lambda \in \mathbf{S}_B^*$. Our first results relate these channels to the a.m.s. ones:

Proposition 1. Let ν be a channel with additive a.m.s. noise determined by a measure $\varkappa \in \mathbf{S}_B$. Then ν is a.m.s. and its stationary mean $\bar{\nu}$ is defined by

$$\bar{\nu}(G \mid y) = \bar{\varkappa}(G - y)\,; \quad y \in B^I\,, \quad G \in \mathscr{F}_A\,, \tag{2.9}$$

where $\bar{\varkappa} \in \mathbf{M}_B$ is the stationary mean of $\varkappa$.

Corollary 1. Let ν be a channel with additive a.m.s. and ergodic noise. Then ν is an a.m.s. and ergodic channel.

Proposition 2. Any channel decomposable into components with additive a.m.s. noise is an a.m.s. channel.

Corollary 2. Let (2.7) be the decomposition of the channel ν into components with additive a.m.s. noise. Then its stationary mean $\bar{\nu}$ is a channel decomposable into components with additive stationary noise via the formulae

$$\bar{\nu}(G \mid y) = \int_A \bar{\nu}^\lambda(G \mid y)\,\gamma(\mathrm{d}\lambda)\,; \quad y \in B^I\,, \quad G \in \mathscr{F}_A\,, \tag{2.10}$$

where $\{\bar{\nu}^\lambda(\,\cdot \mid y);\ y \in B^I\}$ designates the stationary mean of the channel ν^λ, $\lambda \in \Lambda$.

In particular, if the components are channels with additive a.m.s. and ergodic noise then the stationary mean of the composed channel is decomposable into components with additive ergodic noise, i.e., a channel decomposable into ergodic components in the sense of [15].

As in [20] and [12] we shall deal with a restricted class of codes, namely the codes associated with finite factor groups of the group alphabet. Let X denote, generically, a free finitely generated Abelian group (it is assumed that $B = A = X$ in what follows). We let $\mathscr{X} = \mathscr{X}(X)$ designate the class of all finite factor groups of X. There is a sequence $\{\eta_n\}_{n\in N} \subset \mathscr{X}$ such that, for any $\xi \in \mathscr{X}$, there is n_0 with the property $\xi \leqq \eta_n$ whenever $n \geqq n_0$ (i.e., $\{\eta_n\}_{n\in N}$ is cofinal in $\mathscr{X}$ with respect to the partial ordering $\eta \leqq \xi$ defined by the property that ξ is a divisor of η; cf. [20] and [12]). Let $\mu \in \mathbf{P}_X$. As shown in [11], any $\xi \in \mathscr{X}$ induces a measure $\mu_\xi \in \mathbf{P}_\xi$. Moreover, if $\mu \in \mathbf{M}_X$ (if $\mu \in \mathbf{M}_X^*$) then $\mu_\xi \in \mathbf{M}_\xi(\mu_\xi \in \mathbf{M}_\xi^*)$ [11]. Similarly, $\mu \in \mathbf{S}_X(\mu \in \mathbf{S}_X^*)$ entails $\mu_\xi \in \mathbf{S}_\xi(\mu_\xi \in \mathbf{S}_\xi^*)$; cf. [13].

In order we can employ the method of finite partitions for our channels we have to ensure that both the input and the output symbols of the codes possess the same alphabet. Thus we are led to the following concept of a code [20]: a disjoint parameter family $\mathscr{Q} = \{Q_Y;\ Y \in \mathscr{Y}\}$ of classes Q_Y of subsets in X^n is said to be an *n-dimensional group code* if there is a group $\xi \in \mathscr{X}(X)$ such that

$$(Y \in \mathscr{Y}) \Rightarrow [(Y \in \xi^n) \text{ et } (Q_Y \subset \xi^n)]\,. \tag{2.11}$$

Alternatively, we call such a code also a ξ-code. The length of the code $\mathscr{Q}$ is defined as

$$l(\mathscr{Q}) = \operatorname{card}(\mathscr{Y}) \tag{2.12}$$

(one can easily check that $l(\mathscr{Q})$ is upperbounded by $r(\xi^n)$, the order of the group ξ^n). Let ν be a channel decomposable into components with additive a.m.s. and ergodic noise, the component ν^λ being determined by $\varkappa^\lambda \in \mathbf{S}_B^*$ via (2.8). Let, for $\xi \in \mathscr{Z}(X)$,

$$[V] = \{z \in \xi^I : (z_0, \ldots, z_{n-1}) \in V\}, \quad n \in N.$$

The maximum (n-dimensional) error probability of the code $\mathscr{Q}$ with respect to the channel ν is defined as

$$e_n(\mathscr{Q}) = 1 - \min_{Y \in \mathscr{Y}} \int_\Lambda \varkappa_\xi^\lambda[Q_Y - Y]\, \gamma(\mathrm{d}\lambda), \tag{2.13}$$

where $Q_Y - Y$ stands for the group theoretic difference (in ξ^n) and $\varkappa_\xi^\lambda \in \mathbf{S}_\xi^*$ is the measure induced from $\varkappa^\lambda \in \mathbf{S}_B^*$ by the factor group $\xi \in \mathscr{Z}(X)$.

The operational meaning of capacity may be expressed by means of the Channel Coding Theorem and its Converse (cf. [22]). In other words, we are interested in the asymptotic behavior as $n \to \infty$ of the sequences

$$\left\{\frac{1}{n} \log_2 S_n(\varepsilon, \xi);\ n \in N\right\}; \quad 0 < \varepsilon < 1, \quad \xi \in \mathscr{Z}, \tag{2.14}$$

where $S_n(\varepsilon, \xi)$ denotes the maximum length of n-dimensional ξ-codes with the error probability e_n less than ε.

In Section 4 below we shall use a general construction (as explained in [12, Section 4]) in order to prove the existence of an extended real-valued random variable $C(\varkappa^\lambda)$ on $(\Lambda, \mathscr{L}, \gamma)$ and define the *information quantile capacity* C^* as the limit

$$C^* = \lim_{\vartheta \to 0} c(\vartheta) \tag{2.15}$$

where

$$c(\vartheta) = \inf\{t : \gamma\{\lambda \in \Lambda : C(\varkappa^\lambda) \leqq t\} \geqq \vartheta\} \tag{2.16}$$

for $0 < \vartheta < 1$. Due to our interest in fixed error probabilities the quantiles $c(\vartheta)$ themselves naturally enter the coding theorems:

Theorem 1. If $0 < \vartheta < \varepsilon < 1$, and if $t < c(\vartheta)$ then there is a finite factor group ξ of the alphabet such that

$$\liminf_n \frac{1}{n} \log_2 S_n(\varepsilon, \xi) \geqq t. \tag{2.17}$$

Dually, we have

Theorem 2. If $0 < \varepsilon < \vartheta < 1$, and if ξ is any finite factor group of the alphabet then

$$\limsup_n \frac{1}{n} \log_2 S_n(\varepsilon, \xi) \leqq c(\vartheta). \tag{2.18}$$

In particular, if we restrict ourselves to the continuity points of $c(\cdot)$, the above theorems together with the fact

$$(\eta \leqq \xi) \Rightarrow (S_n(\varepsilon, \eta) \leqq S_n(\varepsilon, \xi)) \tag{2.19}$$

give the main result:

Group Coding Theorem for Decomposable Channels. Let ν be a channel decomposable into components with additive a.m.s. and ergodic noise. Let ε be a continuity point of the quantile function defined in (2.16).

I. *Coding Theorem*

$$\forall\{c' < c(\varepsilon)\}\ \exists\{\xi_0 \in \mathscr{X}(X)\}\ \exists\{n_0 \geqq 1\}\ \forall\{\xi \in \mathscr{X}(X);\ \xi \geqq \xi_0\}$$
$$\forall\{n \geqq n_0\}\ \exists\{n\text{-dimensional } \xi\text{-code } \mathscr{Q}\} : (e_n(\mathscr{Q}) < \varepsilon) \text{ et } (l(\mathscr{Q}) > 2^{nc'})\,.$$

II. *Converse*

$$\forall\{c'' > c(\varepsilon)\}\ \forall\{\xi \in \mathscr{X}(X)\}\ \exists\{n_0 \geqq 1\}\ \forall\{n \geqq n_0\}$$
$$\forall\{n\text{-dimensional } \xi\text{-code } \mathscr{Q}\} : (e_n(\mathscr{Q}) < \varepsilon) \Rightarrow (l(\mathscr{Q}) < 2^{nc''})\,.$$

Corollary 3. Let C^* be the information quantile capacity of the channel ν. Then the assertion of the Coding Theorem is valid for all real $c' < C^*$. If, in addition, ν is indecomposable, and if $c'' > C^*$ (in case $C^* < \infty$) then

$$\forall(\xi \in \mathscr{X}(X)\}\ \exists\{n_0 \geqq 1\}\ \forall\{n \geqq n_0\}$$
$$(e_n(\mathscr{Q}) < \varepsilon) \Rightarrow (l(\mathscr{Q}) < 2^{nc''})\,.$$

Our next results concern the trasmission rate capacities. First of all, the following auxiliary result is necessary:

Proposition 3. Let $\mu \in \boldsymbol{M}_B$ and let ν be an a.m.s. channel. Then the output distribution $\mu\nu^0$ (i.e., the A^I marginal of $\mu\nu$) is a.m.s.

Given a decomposable channel ν of the type considered and given $\xi \in \mathscr{X}(X)$ we define a family $\nu_\xi = \{\nu_\xi(\cdot \mid y);\ y \in \xi^I\}$ by the properties

$$\nu_\xi(G \mid y) = \int_\Lambda \varkappa_\xi^\lambda(G - y)\,\gamma(\mathrm{d}\lambda)\,; \quad y \in \xi^I\,, \quad G \in \mathscr{F}_\xi\,. \tag{2.20}$$

Proposition 4. Let ν be a channel decomposable into components with additive a.m.s. and ergodic noise. Then the family $\nu_\xi(\xi \in \mathscr{X}(X)$ arbitrary) is a channel (i.e., a measurable family of probability measures) decomposable into components with additive a.m.s. and ergodic noise.

In light of Propositions 2 and 3 we have two a.m.s. measures, $(\mu\nu_\xi)^0$ and $\mu\nu_\xi$, associated with any $\mu \in \mathbf{M}_\xi$ and $\xi \in \mathscr{X}(X)$. Since all these measures determine finite-alphabet a.m.s. sources, their respective entropy rates, $h(\mu)$, $h(\mu\nu_\xi)$ and $h((\mu\nu_\xi)^0)$, are all well-defined [13]. Let

$$I(\mu\nu_\xi) = h(\mu) + h((\mu\nu_\xi)^0) - h(\mu\nu_\xi) \tag{2.21}$$

denote the transmission rate. We define the stationary capacity $C_s(\xi)$ of the channel ν_ξ by

$$C_s(\xi) = \sup \{I(\mu\nu_\xi) : \mu \in \mathbf{M}_\xi\} \tag{2.22}$$

and the ergodic capacity $C_e(\xi)$ by replacing $\mathbf{M}_\xi$ in (2.22) by $\mathbf{M}_\xi^*$. For finite-alphabet channels there is no need of restriction to the group codes. In particular, we aim to prove the Theorem on ε-Capacities [15] for such channels. This will be a consequence of the next theorem.

Theorem 3. The channel ν_ξ derived from a channel decomposable into components with additive a.m.s. and ergodic noise is regularly decomposable in the sense of [15].

Finally, we have the following result connecting the stationary and ergodic capacities with the information quantile capacity:

Theorem 4. Let ν be a channel decomposable into components with additive a.m.s. and ergodic noise. Let

$$C_s = \sup_{\xi \in \mathscr{X}} C_s(\xi)\,, \quad C_e = \sup_{\xi \in \mathscr{X}} C_e(\xi)\,. \tag{2.23}$$

Then

$$C_s = C_e\,, \tag{2.24}$$

$$(\nu \text{ indecomposable}) \Rightarrow (C_s = C_e = C^*)\,. \tag{2.25}$$

3. STRUCTURE OF CHANNELS

In this section we prove the assertions related to the properties of the channels themselves.

Proof of Proposition 1. By [1, Corollary 1] we have to consider only stationary test sources when proving a channel a.m.s. So let $\mu \in \mathbf{M}_B$ be arbitrary. We shall prove that the joint input/ouput measure $\mu\nu$ (cf. (2.2)) is a.m.s. by proving that the limits

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} \mu\nu(T_{B\times A}^{-j}(F \times G)) \tag{3.1}$$

exist for all rectangles $F \times G$; $F \in \mathscr{F}_B$, $G \in \mathscr{F}_A$ (as to the symbol $T_{B\times A}$ cf. (2.3)).

Now

$$\mu\nu(T_{B\times A}^{-j}(F \times G)) = \mu\nu(T_B^{-j}F \times T_A^{-j}G) =$$

$$= \int_{T_B^{-j}F} \nu(T_A^{-j}G \mid y)\,\mu(\mathrm{d}y) = \int_{T_B^{-j}F} \varkappa(T_A^{-j}G - y)\,\mu(\mathrm{d}y) =$$

$$= \int_F \varkappa(T_A^{-j}G - T_A^{-j}y)\,\mu(\mathrm{d}y) = \int_F \varkappa(T_A^{-j}(G - y))\,\mu(\mathrm{d}y)\,,$$

where we used the invariance of μ and the fact that $T_B = T_A$ is one-to-one. Routine arguments including the Bounded Convergence Theorem give that the limit (3.1) coincides with

$$\int_F \bar{\varkappa}(G - y)\,\mu(\mathrm{d}y)\,;$$

the numbers $\bar{\varkappa}(G - y)$ being well-defined because $\varkappa$ is a.m.s. This shows that ν is a.m.s. Now since $\bar{\varkappa} \in \boldsymbol{M}_B$, the relations

$$\tilde{\nu}(G \mid y) = \bar{\varkappa}(G - y)\,; \quad y \in B^I\,, \quad G \in \mathscr{F}_A$$

define a stationary channel. One can easily check that $\tilde{\nu}(\cdot \mid y) = \bar{\nu}(\cdot \mid y)$ μ-a.e. for every $\mu \in \boldsymbol{M}_B$, where $\bar{\nu}$ is the stationary mean of the a.m.s. channel ν so that the proof is complete. □

Proof of Corollary 1. If $\varkappa \in \boldsymbol{S}_B^*$ then $\bar{\varkappa} \in \boldsymbol{M}_B^*$. By the preceding proposition, $\bar{\nu}$ is a channel with additive stationary and ergodic noise, hence $\bar{\nu}$ is stationary and ergodic (see [8] or [15]). By [1, Lemma 4], ν is a.m.s. and ergodic. □

Proof of Proposition 2. Let the components ν^λ be determined by a.m.s. measures $\varkappa^\lambda$, $\lambda \in \Lambda$. By [1, Corollary 3], in order to prove ν a.m.s. it suffices to prove that the limits

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} \nu(T_A^{-j}G \mid T_B^{-j}y)$$

exist for all $y \in B^I$ and $G \in \mathscr{F}_A$. But this assertion holds true for ν replaced by ν^λ ($\lambda \in \Lambda$). For fixed G, j, and y, the map $\lambda \mapsto \varkappa^\lambda(T_A^{-j}(G - y))$ is $\mathscr{L}$-measurable so that the Bounded Convergence Theorem applies to conclude the proof. □

To the proof of Corollary 2 note the following. It is an easy exercise to show that (2.10) is valid for all y except some set of invariant measure zero, i.e., except some $F \in \mathscr{F}_B$ with the property that $\mu(F) = 0$ for all $\mu \in \boldsymbol{M}_B$. The obtained $\bar{\nu}$ can be then modified on that exceptional set so that (2.10) becomes valid for all $y \in B^I$ (set [1] for the details about such a.e. modifications). The proof of Propositon 3 is immediate.

Proof of Proposition 4. By assumption,

$$\nu(G \mid y) = \int_\Lambda \varkappa^\lambda(G - y)\,\gamma(\mathrm{d}\lambda)\,, \quad y \in B^I\,, \quad G \in \mathscr{F}_A\,,$$

where $\varkappa^\lambda \in \mathbf{S}_B^*$ for all $\lambda \in \Lambda$. Let $\varkappa_\xi^\lambda \in \mathbf{S}_\xi^*$ be the induced measure on $\mathscr{F}_\xi$ ($\xi \in \mathscr{Z}$). By definition,

$$\nu_\xi(G \mid y) = \int_\Lambda \varkappa_\xi^\lambda(G - y)\,\gamma(\mathrm{d}\lambda)\,, \quad G \in \mathscr{F}_\xi\,.$$

This integral representation immediately results in the desired measurability of the maps $y \mapsto \nu_\xi(G \mid y)$, $G \in \mathscr{F}_\xi$. □

Let us close this section with the following

Lemma 2. Let ν be a channel decomposable into components with additive a.m.s. and ergodic noise. Then ν is a channel with additive a.m.s. noise determined by the measure $\varkappa \in \mathbf{S}_B$ defined by the property that

$$\varkappa(\cdot) = \int_\Lambda \varkappa^\lambda(\cdot)\,\gamma(\mathrm{d}\lambda)\,. \tag{3.2}$$

Proof. By definition, $\lambda \mapsto \varkappa^\lambda(G)$ is $\mathscr{L}$-measurable and uniformly bounded in λ for any fixed $G \in \mathscr{F}_A$. Hence, the integral in (3.2) is well defined and the a.m.s. property of the probability measure $\varkappa$ follows easily by a direct verification. Finally,

$$\nu(G \mid y) = \int_\Lambda \nu^\lambda(G \mid y)\,\gamma(\mathrm{d}\lambda) = \int_\Lambda \varkappa^\lambda(G - y)\,\gamma(\mathrm{d}\lambda) = \varkappa(G - y)\,;$$
$$G \in \mathscr{F}_A\,, \quad y \in B^I\,.$$
□

Let R_X denote the set of all regular points in X^I and let $\mu_z \in \mathbf{M}_X^*$ be the measure uniquely determined by $z \in R_X$ [14]. As well-known, $R_X \in \mathscr{I}_X$, μ_z is a T_X-invariant function of the variable z and $\mu(R_X) = 1$ for all $\mu \in \mathbf{M}_X$. In particular, since any $\varkappa \in \mathbf{S}_X$ coincides with its stationary mean $\bar{\varkappa} \in \mathbf{M}_X$ on $\mathscr{I}_X$, $\varkappa(R_X) = 1$ for all $\varkappa \in \mathbf{S}_X$. In addition,

$$\varkappa(F) = \int_{R_X} \mu_z(F)\,\varkappa(\mathrm{d}z)\,, \quad F \in \mathscr{I}_X\,, \quad \varkappa \in \mathbf{S}_X\,. \tag{3.3}$$

However, $\varkappa$ and $\bar{\varkappa}$ may differ significantly on $\mathscr{F}_X \setminus \mathscr{I}_X$ so that the converse of Lemma 2 fails to hold. In other words, the relation (3.2) does not coincide with the ergodic decomposition as known in Krylov-Bogolyubov theory (cf. [14]).

4. INFORMATION QUANTILES AND CAPACITY

Let $\varkappa \in \mathbf{S}_X^*$. Then $\bar{\varkappa} \in \mathbf{M}_X^*$ and, as one can easily check, $\overline{\varkappa_\xi} = (\bar{\varkappa})_\xi \in \mathbf{M}_\xi^*$. For brevity, we shall write $\bar{\varkappa}_\xi$ for $(\bar{\varkappa})_\xi$. Due to the invertibility of the shift T_ξ in ξ^I, $\varkappa_\xi \ll \bar{\varkappa}_\xi$. In other words, we are in the setup investigated by Jacobs [4]. Let

$$[z_1, \ldots, z_n] = [\{z_1, \ldots, z_n\}] = \{y \in \xi^I : y_i = z_{i+1},\ i = 0, 1, \ldots, n - 1\}\,. \tag{4.1}$$

We let $h(\bar{\varkappa}_\xi)$ denote the entropy rate of $\bar{\varkappa}_\xi$. By the McMillan's theorem of Jacobs we know that the sequence $\{-n^{-1} \log_2 \varkappa_\xi[z_1, \ldots, z_n];\ n \in N\}$ converges in $L^1(\varkappa_\xi)$ to the limit $h(\bar{\varkappa}_\xi)$. As in [13] we identify that limit with the entropy rate of the a.m.s. measure $\varkappa_\xi$ and we use the notation

$$h(\xi, \varkappa) = h(\varkappa_\xi) = h(\bar{\varkappa}_\xi) . \tag{4.2}$$

As pointed out in [10] and [12] we can use

$$h(\bar{\varkappa}) = \lim_{\xi \in \mathscr{X}} h(\xi, \varkappa) \tag{4.3}$$

as the definition of the entropy rate of $\varkappa \in \mathbf{S}_X$ even if X is infinite so that Jacobs' arguments fail to work. This definition has been justified in [13, Section 3]. Let

$$C(\xi, \varkappa^\lambda) = \log_2 r(\xi) - h(\xi, \varkappa^\lambda) , \quad \lambda \in \Lambda . \tag{4.4}$$

Following [12, Section 4] or [20] we can prove the existence of an extended real-valued random variable C on the probability space $(\Lambda, \mathscr{L}, \gamma)$ such that

$$C(\varkappa^\lambda) = \sup_{\xi \in \mathscr{X}} C(\xi, \varkappa^\lambda) = \lim_n C(\eta_n, \varkappa^\lambda) \tag{4.5}$$

for all $\lambda \in \Lambda$, where $\{\eta_n;\ n \in N\} \subset \mathscr{X}$ is the above mentioned cofinal sequence. Consequently, the definitions (2.16) and (2.15) are justified.

5. PROOFS OF THE CODING THEOREMS

For $\varkappa \in \mathbf{S}_X$, $\xi \in \mathscr{X}(X)$, and $0 < \varepsilon < 1$ put

$$L_\varkappa(\varepsilon, \xi^n) = \min \{\operatorname{card}(\Delta) : \Delta \subset \xi^n,\ \varkappa_\xi[\Delta] > 1 - \varepsilon\} . \tag{5.1}$$

As shown in [13], if $\varkappa \in \mathbf{S}_X^*$ then

$$\lim_n \frac{1}{n} \log_2 L_\varkappa(\varepsilon, \xi^n) = h(\xi, \varkappa) \tag{5.2}$$

for all $\varepsilon \in (0, 1)$. Further, recall from [20] the basic relations between the quantities $S_n(\varepsilon, \xi)$ and $L_\varkappa(\varepsilon, \xi^n)$:

$$(\varepsilon - \varepsilon')\, r(\xi^n) < S_n(\varepsilon, \xi)\, L_\varkappa(\varepsilon', \xi^n) ; \tag{5.3}$$

$$S_n(\varepsilon, \xi)\, L_\varkappa(\varepsilon, \xi^n) \leqq r(\xi^n) ; \quad \xi \in \mathscr{X} , \quad 0 < \varepsilon' < \varepsilon < 1 . \tag{5.4}$$

To the proof note that all these quantities relate to finite probability vectors so that the relations are valid also without the stationary assumption. The proofs of the coding theorems (Theorems 1 and 2) are based on the ideas first developed in [14] and then frequently adapted (see e.g. [20] and [12]). Therefore, only rather sketchy proofs are given.

Proof of Theorem 1. It follows from (2.16) that

$$\gamma\{\lambda \in \Lambda : C(\varkappa^\lambda) > t\} > 1 - \vartheta .$$

Hence, for some $n \geqq 1$,

$$\gamma\{\lambda \in \Lambda : C(\eta_n, \varkappa^\lambda) > t\} > 1 - \vartheta$$

(cf. (4.5)). Let $d = \log_2 r(\xi) - t$ for $\xi = \eta_n$ with the foregoing property. Then, according to (4.4),

$$\gamma\{\lambda \in \Lambda : h(\xi, \varkappa^\lambda) < d\} > 1 - \vartheta$$

so that $\gamma(D_\Lambda) > 1 - \vartheta$ if $D_\Lambda = \{\lambda \in \Lambda : h(\xi, \varkappa^\lambda) < d\}$. As $\mu \mid \mathscr{I}_X = \bar{\mu} \mid \mathscr{I}_X$ for $\mu \in \mathbf{S}_X$, we conclude that

$$\forall\{\lambda \in D_\Lambda\}\ \varkappa_\xi^\lambda(D) = 1 ,$$

where $D = \{z \in R_\xi : h(\mu_z) < d\}$. Therefore, if

$$\varkappa^{(1)} = \gamma(D_\Lambda)^{-1} \int_{D_\Lambda} \varkappa^\lambda\, \gamma(\mathrm{d}\lambda)$$

then $\varkappa_\xi^{(1)}(D) = 1$. From now on the proof follows the one given in [14] until we get

$$\limsup_n \frac{1}{n} \log_2 L_\varkappa(\varepsilon, \xi^n) \leqq d$$

for all $\varepsilon > \vartheta$. Now let $\vartheta < \varepsilon' < \varepsilon$. Then, according to (5.3), we get

$$\liminf_n \frac{1}{n} \log_2 S_n(\varepsilon, \xi) \geqq$$

$$\geqq \log_2 r(\xi) - \limsup_n \frac{1}{n} \log_2 L_\varkappa(\varepsilon', \xi^n) \geqq$$

$$\geqq \log_2 r(\xi) - d = t . \qquad \square$$

Proof of Theorem 2. Let $t = c(\vartheta)$ so that

$$\gamma\{\lambda \in \Lambda : C(\xi, \varkappa^\lambda) \leqq t\} \geqq \vartheta .$$

Let $D_\Lambda = \{\lambda \in \Lambda : h(\xi, \varkappa^\lambda) \geqq d\}$, $d = \log_2 r(\xi) - t$. Then $\gamma(D_\Lambda) \geqq \vartheta > 0$ so that we can define

$$\gamma_2(L) = \gamma(L \cap D_\Lambda)/\gamma(D_\Lambda) , \quad L \in \mathscr{L} .$$

Let $\varkappa^{(2)}$ denote the corresponding composition of the measures $\varkappa^\lambda$. Then

$$\varkappa_\xi^{(2)}\{z \in R_\xi : h(\mu_z) \geqq d\} = 1$$

so that, using [13, Lemma 2], [14, Lemma 8] and (5.4), we get the conclusion. $\square$

The direct part of the Group Coding Theorem follows from Theorem 1 by taking into account the assertion (2.19). The Converse is just a restatement of Theorem 2. Corollary 3 is trivial.

Proof of Theorem 3. Let ν_ξ be determined by (2.20), whence $\varkappa_\xi^\lambda \in \mathbf{S}_\xi^*$ for all $\lambda \in \Lambda$. By [15, Theorem 5] and Corollary 2, the stationary mean $\bar{\nu}_\xi$ of the channel ν_ξ

is regularly decomposable. Now we use the fact that the regularity condition relates merely to the distribution functions

$$F(t) = \gamma\{\lambda \in \Lambda : C(\xi, \varkappa^\lambda) \leqq t\},$$
$$G(t) = \gamma\{\lambda \in \Lambda : I(\mu\nu_\xi^\lambda) \leqq t\}, \quad t \text{ real}.$$

(cf. [15, (1.25), (1.26)] or [16]). Now $C(\xi, \varkappa^\lambda) \leqq t$ if and only if $h(\varkappa_\xi^\lambda) \geqq \log_2 r(\xi) - t$ (see (4.4)), hence

$$F(t) = \gamma\{\lambda \in \Lambda : h(\varkappa_\xi^\lambda) \geqq \log_2 r(\xi) - t\} =$$
$$= \varkappa_\xi\{z \in R_\xi : h(\mu_z) \geqq \log_2 r(\xi) - t\} =$$
$$= \bar{\varkappa}_\xi\{z \in R_\xi : h(\mu_z) \geqq \log_2 r(\xi) - t\} = \bar{F}(t).$$

Going backwards we get

$$\bar{F}(t) = \gamma\{\lambda \in \Lambda : C(\xi, \bar{\varkappa}^\lambda) \leqq t\}, \quad t \text{ real}.$$

If $\mu \in \mathbf{M}_\xi^*$ then both $(\mu\bar{\nu}_\xi^\lambda)^0$ and $\mu\bar{\nu}_\xi^\lambda$ are stationary and ergodic finite-alphabet sources, for all $\lambda \in \Lambda$. In particular, there is a $T_{\xi\times\xi}$-invariant function $I(z)$, $z \in R_{\xi\times\xi}$ such that

$$\mu\bar{\nu}_\xi^\lambda\{z \in R_{\xi\times\xi} : I(z) = I(\mu\bar{\nu}_\xi^\lambda)\} = 1.$$

But $I(\mu\nu_\xi^\lambda) = I(\mu\bar{\nu}_\xi^\lambda)$ so that

$$\mu\nu_\xi^\lambda\{z \in R_{\xi\times\xi} : I(z) = I(\mu\nu_\xi^\lambda)\} = 1.$$

Consequently, $G(t) = \bar{G}(t)$, where $\bar{G}(t)$ has analogous meaning as $\bar{F}(t)$ above, and this proves the regularity condition. In fact, the distribution functions $\bar{F}$ and $\bar{G}$ correspond to the regularly decomposable stationary channel $\bar{\nu}_\xi$. □

Corollary 5. Let the channel ν be decomposable into components with additive a.m.s. and ergodic noise, and let $\xi \in \mathscr{X}(X)$. The limit

$$C_\varepsilon(\xi) = \lim_n \frac{1}{n} \log_2 S_n(\varepsilon, \xi) \tag{5.5}$$

exists for all $\varepsilon \in (0, 1)$, except some countable set, and

$$C_\varepsilon(\xi) = \inf\{t : \gamma\{\lambda : C(\xi, \varkappa^\lambda) \leqq t\} \geqq \varepsilon\}. \tag{5.6}$$

This follows, on account of Theorem 3, from the main Theorem on ε-Capacity in [15]. The corollary suggests an alternative definition of the information quantiles. Actually, let

$$\bar{c}(\vartheta) = \sup\{C_\vartheta(\xi) : \xi \in \mathscr{X}\} \text{ (cf. (5.6))} \tag{5.7}$$

and

$$\bar{C} = \lim_{\vartheta\to 0} \bar{c}(\vartheta). \tag{5.8}$$

Proposition 6. Let ν be a channel decomposable into components with additive a.m.s. and ergodic noise. Then

$$c(\vartheta) = \bar{c}(\vartheta), \quad 0 < \vartheta < 1, \tag{5.9}$$

so that, in particular, $C^* = \bar{C}$.

Proof. Throughout the proof the symbols lim $\uparrow$ and lim $\downarrow$ will denote monotone limits. Let $t \geqq c(\vartheta)$. By the definition of $c(\vartheta)$ and with the aid of (4.5) we get

$$\gamma\{\lambda \in \Lambda : \lim_n \uparrow C(\eta_n, \varkappa^\lambda) \leqq t\} \geqq \vartheta$$

so that

$$\lim_n \downarrow \gamma\{\lambda \in \Lambda : C(\eta_n, \varkappa^\lambda) \leqq t\} \geqq \vartheta .$$

Therefore $t \geqq C_\vartheta(\eta_n)$ for all $n \geqq 1$. As

$$\lim_n \uparrow C_\vartheta(\eta_n) = \sup_{\xi \in \mathscr{X}} C_\vartheta(\xi) = \bar{c}(\vartheta)$$

we get, due to the arbitrariness of t, the relation $c(\vartheta) \geqq \bar{c}(\vartheta)$. The converse inequality can be obtained by a similar reasoning and the details are left to the reader. □

Corollary 6. The information quantile capacity C^* of a channel decomposable into components with additive a.m.s. and ergodic noise is expressed by the formula

$$C^* = \lim_{\varepsilon \to 0} \sup_{\xi \in \mathscr{X}} \lim_n \frac{1}{n} \log_2 S_n(\varepsilon, \xi) . \tag{5.10}$$

Proof of Theorem 4. It suffices to prove

$$C_s(\xi) = C_e(\xi), \quad \xi \in \mathscr{X}(X) . \tag{5.11}$$

Using the notations introduced above in the proof of Theorem 3 we get immediately from [4] that

$$I(\mu\nu_\xi) = \int_{R_{\xi \times \xi}} I(z)\, \mu\nu_\xi(\mathrm{d}z)$$

for any $\mu \in \mathbf{M}_\xi$ and any $\xi \in \mathscr{X}(X)$, respectively. If $z \in R_{\xi \times \xi}$ then $z = (y, x)$, where $y \in R_\xi$, $x \in R_\xi$ and the marginal distributions of μ_z coincide with μ_y and μ_x, respectively (cf. [15]). Hence

$$I(\mu\nu_\xi) = \int_{R_\xi} I(\mu_z \nu_\xi)\, \mu(\mathrm{d}z) , \quad \mu \in \mathbf{M}_\xi . \tag{5.12}$$

But (5.12) implies (5.11) as shown in detail by Parthasarathy [7]. It remains to prove (2.25). But the proof is easy as the capacities $C_s(\xi)$ and $C_e(\xi)$ of the channel ν_ξ coincide with those of the stationary channel $\bar{\nu}_\xi$. □

Thus we have proved all assertions formulated in the previous sections. On the other hand, there are several other definitions of capacity not yet mentioned. Therefore we conclude the paper by the following remarks. A measure $\mu \in \mathbf{P}_\xi$ is said to be block-stationary if there is $n \in N$ such that μ is T_ξ^n-invariant. Let $\mathbf{M}_\xi^b$ designate the set of all block-stationary measure on ξ^I. Let

$$C = \sup_{\xi \in \mathscr{X}} \sup \{I(\mu\nu_\xi) : \mu \in \mathbf{M}_\xi^b\} . \tag{5.13}$$

At the same time, we can consider the quantiles of the transmission rate function $I(z)$ instead of the quantiles of the capacity function so that we get

$$\tilde{C} = \sup_{\xi \in \mathscr{X}} \lim_{\vartheta \to 0} \tilde{C}_\vartheta(\xi) , \tag{5.14}$$

where

$$\tilde{C}_\vartheta(\xi) = \sup_{\mu \in \mathbf{M}_\xi^b} \sup \{t : \mu\nu_\xi\{I(z) \leqq t\} < \vartheta\} .$$

Lemma 3. $$C \geqq \tilde{C} .$$

For the proof see [3, Lemma 1]. The proof works also in the a.m.s. setting due to the invariance of $I(z)$.

Corollary 7. Let ν be an indecomposable channel, i.e., a channel with additive a.m.s. and ergodic noise. Then

$$C^* = C_s = C_e = \bar{C} = C = \tilde{C} = c(\vartheta) = \bar{c}(\vartheta) \tag{5.15}$$

for all $\vartheta \in (0, 1)$.

(Received February 19, 1980.)

REFERENCES

[1] R. J. Fontana, R. M. Gray, J. C. Kieffer: Asymptotically mean stationary channels (1979 — preprint).

[2] R. M. Gray, J. C. Kieffer: Asymptotically mean stationary measures (1979 — submitted to Ann. of Prob.).

[3] R. M. Gray, D. S. Ornstein: Block coding for discrete stationary $\bar{d}$-continuous noisy channels. IEEE Trans. Inform. Theory *IT-25* (1979), 292—306.

[4] K. Jacobs: Die Übertragung diskreter Informationen durch periodische und fastperiodische Kanäle. Math. Annalen *137* (1959), 125—135.

[5] J. C. Kieffer: A general formula for the capacity of stationary nonanticipatory channels. Inform. and Control *26* (1971), 381—391.

[6] J. Nedoma: The capacity of a discrete channel. Trans. 1st Prague Conf. Inform. Theory etc., NČSAV, Prague 1957, 143—181.

[7] K. R. Parthasarathy: On the integral representation of the rate of transmission of a stationary channel. Ill. J. Math. *2* (1961), 299—305.

[8] K. R. Parthasarathy: Effective entropy rate and transmission of information through channels with additive random noise. Sankhya *A 25* (1963), 75—84.

[9] C. E. Shannon: A mathematical theory of communication. Bell. Syst. Techn. J. *27* (1948), 379—423, 623—656.

[10] Š. Šujan: On the integral representation of the entropy rate. Studia Sci. Math. Hung. *11* (1976), 25—36.

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

[12] Š. Šujan: Epsilon-rates, epsilon-quantiles, and group coding theorems for finitely additive information sources. Kybernetika *16* (1980), 105—119.

[13] Š. Šujan: Existence of asymptotic rate for asymptotically mean stationary sources with countable alphabets. 3rd Czechoslovak-Soviet-Hungarian Seminar on Inform. Theory, Liblice 1980, 201—206.

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

[15] K. Winkelbauer: On the coding theorem for decomposable channels I, II. Kybernetika *7* (1971), 109—123, 230—255.

[16] K. Winkelbauer: On the regularity condition for decomposable communication channels. Kybernetika *7* (1971), 314—327.

[17] K. Winkelbauer: On discrete channels decomposable into memoryless components. Kybernetika *8* (1972), 114—132.

[18] K. Winkelbauer: On the capacity of decomposable channels. Trans. 6th Prague Conf. Inform. Theory etc., Academia, Prague 1973, 903—914.

[19] K. Winkelbauer: Information channels with memoryless components. Trans. 7th Prague Conf. Inform. Theory etc., Academia, Prague 1978, 559—576.

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

[21] K. Winkelbauer: Discrete communication channels decomposable into finite-memory components. In: Contributions to Statistics (Jaroslav Hájek Memorial Volume, J. Jurečková, ed.), Academia, Prague 1979, 277—306.

[22] J. Wolfowitz: Coding Theorems of Information Theory. 2nd edition. Springer-Verlag, Berlin—Göttingen—New York 1964.

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