KYBERNETIKA — VOLUME 8 (1972), NUMBER 2

# On Discrete Channels Decomposable into Memoryless Components

KAREL WINKELBAUER

The concept of $\varepsilon$-capacity as developed by the author in [6], is studied here for the case of discrete channels which may be decomposed into a finite number of memoryless components: the entire paper is devoted to the statement and proof of the theorem on the existence of $\varepsilon$-capacity for channels of the type described. The important difference between this paper and [6] is that we do not restrict ourselves here to regular cases (cf. [7]) as was done in [6].

To explain the result established in this paper, let us suppose that we are given a discrete channel having exactly two distinct components, the components to be memoryless channels in the usual sense. If we denote the component channels by $\nu^1$, $\nu^2$, and the composed channel by $\nu$, we can express the way in which the channel $\nu$ may be decomposed, symbolically by the relation $\nu = \xi\nu^1 + (1 - \xi)\,\nu^2$, where the probabilities $\xi$ and $1 - \xi$ represent the rates of influence of the components in the noise arising during transmission of information over the composed channel.

Let us denote the capacities of the component channels by $C^1$, $C^2$, and let the symbol $S_n(\varepsilon)$ designate the maximum length of those $n$-dimensional codes (in the sense of Wolfowitz; cf. [9]) which discern $n$-dimensional input sequences with probability of error not exceeding $\varepsilon$ $(0 < \varepsilon < 1)$. Up to now it is only known (as to the knowledge of the author) that the maximum length just mentioned lies in the interval

$$2^{n(V-\lambda)} < S_n(\varepsilon) < 2^{n(U+\lambda)}$$

for $\lambda$ positive and arbitrarily small, and for $n$ sufficiently large, where the bounds $V$, $U$ are given by

$$U = \max\left(C^1, C^2\right), \quad V = \left(\frac{1}{C^1} + \frac{1}{C^2}\right)^{-1};$$

the lower bound is due to Nedoma (cf. [4]), the upper bound being evident (cf. [6]).

In this paper we shall state and prove the theorem that, for every $\varepsilon$ except at most

$\varepsilon = \xi$, or $\varepsilon = 1 - \xi$, there is a number $C_\varepsilon$ ($\varepsilon$-capacity of the composed channel) such that 

$$2^{n(C_\varepsilon - \lambda)} < S_n(\varepsilon) < 2^{n(C_\varepsilon + \lambda)}$$

for $\lambda$ arbitrarily small and $n$ sufficiently large. Moreover, it may be shown that $C_\varepsilon$ equals exactly the maximum of $\varepsilon$-quantiles associated with the parameter family of transmission rates of the component channels fed by independent stationary inputs (playing the role of parameters), the transmission rates taken as random variables with probability distributions determined by the probability vector $(\xi, 1 - \xi)$; cf. the theorem in Sec. 1.

The theorem on the existence of $\varepsilon$-capacity will, of course, be stated for discrete channels with any finite number of memoryless components. Moreover, we shall first restrict ourselves to the non-singular case described as such that all the transition probabilities which determine the component memoryless channels are positive, the general case being postponed to another paper since it requires a refinement of the methods of proof used below.

CONTENTS



## 1. THE THEOREM

Let us begin with some remarks on terminology and notations. As usual, the letter $I$ designates the set of all integers. Given a finite non-empty set $M$, the symbol $M^I$ represents the space of doubly-infinite sequences $\zeta = \{\zeta_i\}_{i \in I} = \{\zeta_i\}_{i=-\infty}^{+\infty}$ of members $\zeta_i$ lying in $M$. Similarly, $M^n$ is the space of $n$-dimensional sequences $z = \{z_i\}_{i=0}^{n-1}$, where $z_i \in M$ $(i = 0, 1, \ldots, n - 1;\ n = 1, 2, \ldots)$. In accordance with [6] we shall set

$$[z] = \{\zeta : \zeta \in M^I,\ \{\zeta_i\}_{i=0}^{n-1} = z\} \quad \text{for} \quad z \in M^n . \tag{1.1}$$

If $T_M$ means the shift of the space $M^I$ defined by the property that $(T_M\zeta)_i = \zeta_{i+1}$ for $i \in I$, $\zeta \in M^I$, then the symbol $\mathbf{F}_M$ will be employed to denote the $\sigma$-algebra of subsets of $M^I$ which is generated by the class of all sets of the form $T_M^i[z]$ where $z \in M^n$, $i \in I$ $(n = 1, 2, \ldots)$. Finally, by a probability $M$-vector we shall mean a family $p = \{p(m)\}_{m \in M}$ of non-negative numbers $p(m)$ which add to one.

As in [6] we shall assume throughout the whole paper that we are given two (not necessarily distinct) finite non-empty sets $A$ and $B$ called *alphabets*, viz. *A the output* alphabet, and $B$ the *input* alphabet. Elements of $A^I$ and $A^n$, respectively, may be referred to as output sequences, and elements of $B^I$ and $B^n$, respectively, as input sequences.

In what follows we shall associate with any probability $B$-vector $p = \{p(b)\}_{b \in B}$ a probability measure $\mu^p$ uniquely determined on $\mathbf{F}_B$ by the condition that

$$(1.2) \qquad \mu^p(T_B^i[y]) = \prod_{j=0}^{n-1} p(y_j) \quad \text{for} \quad i \in I\,, \quad y \in B^n \quad (n = 1, 2, \ldots)\,.$$

Let us put (cf. (1.1))

$$(1.3) \qquad \mu[F] = \sum_{y \in F} \mu[y] \quad \text{for} \quad F \subset B^n\,, \quad \mu = \mu^p\,, \quad p \in P\,,$$

$$P = \{p : p \text{ is a probability } B\text{-vector}\}\,;$$

i.e. throughout the entire paper the set of all probability $B$-vectors is designated by $P$.

Following the terminology introduced in [3], we shall mean by a *discrete channel* a family $\nu = \{\nu_\eta\}_{\eta \in B^I}$ of probability measures $\nu_\eta$ defined on the measurable (sample) space $(A^I, \mathbf{F}_A)$ which satisfy conditions

$$(1.4) \qquad \nu_{T_B\eta}(T_A E) = \nu_\eta(E) \quad \text{for} \quad \eta \in B^I\,, \quad E \in \mathbf{F}_A\,,$$

$$\nu_{\eta'}[x] = \nu_\eta[x] \quad \text{for} \quad \eta, \eta' \in [y]\,, \quad x \in A^n\,, \quad y \in B^n \quad (n = 1, 2, \ldots)\,;$$

cf. (1.1). For such a channel we shall set

$$(1.5) \qquad \nu[E \mid y] = \sum_{x \in E} \nu[x \mid y]\,, \quad \nu[x \mid y] = \nu_\eta[x]$$

$$\text{where} \quad \eta \in [y]\,, \quad E \subset A^n\,, \quad x \in A^n\,, \quad y \in B^n\,.$$

Let us mention that the notion of discrete channel just stated coincides with that of stationary historyless channel (i.e. a stationary channel with zero past history; cf. [6]) as defined in [5].

In this paper we shall sometimes make use of the following notation: if $M$ is a finite set, then the symbol $\pi(M)$ means:

$$(1.6) \qquad \pi(M) = \text{the number of elements in } M\,;$$

cf. [5], [6]. The latter notation may especially be used to denote the length of a code. Let us recall that an *n-dimensional code* as defined by Wolfowitz in [9], Chapter 3, is a family $\{Q(y)\}_{y \in Y}$ of mutually disjoint sets $Q(y)$, $Q(y) \subset A^n$, with parameter set $Y \subset B^n$; the *length* of the code is defined as the number $\pi(Y)$.

Given a discrete channel $\nu$ and $\varepsilon$ $(0 < \varepsilon < 1)$, we shall say that a set $Y \subset B^n$ is *ε-discernible* (*weakly ε-discernible*) for $\nu$ if there is an $n$-dimensional code $\{Q(y)\}_{y \in Y}$ with the set of parameters equal to $Y$ and such that

$$\nu[Q(y) \mid y] > 1 - \varepsilon\,(\nu[Q(y) \mid y] \geqq 1 - \varepsilon) \quad \text{for all} \quad y \in Y\,.$$

Then the number $S_n(\varepsilon, \nu)$ [and $\bar{S}_n(\varepsilon, \nu)$, respectively] defined as the maximum 

(1.7) $$S_n(\varepsilon, \nu) = \max\{\pi(Y) : Y \subset B^n,\ Y \text{ is } \varepsilon\text{-discernible for } \nu\}$$

$$[\bar{S}_n(\varepsilon, \nu) = \max\{\pi(Y) : Y \subset B^n,\ Y \text{ weakly } \varepsilon\text{-discernible for } \nu\}]$$

represents the maximum length of those $n$-dimensional codes which discern $n$-dimensional input sequences with probability of error less than $\varepsilon$ [and $\leqq \varepsilon$, respectively]; as to the notion of $\varepsilon$-discernibility cf. also [2].

In the sequel a discrete channel $\nu$ is said to be *decomposable into* (a finite number of) *memoryless components* if there are a finite number, say $k$, of mutually distinct discrete channels $\nu^\alpha$ and positive numbers $\xi_\alpha$ ($\alpha = 1, 2, \ldots, k$) which add to one so that conditions (cf. (1.4), (1.5))

(1.8) $$\nu_\eta(E) = \sum_{\alpha=1}^{k} \xi_\alpha \nu_\eta^\alpha(E)\,, \quad \text{symbolically} \quad \nu = \Sigma \xi_\alpha \nu^\alpha\,;$$

$$\nu_\eta^\alpha[x] = \prod_{i=0}^{n-1} \nu^\alpha[x_i \mid \eta_i]\,; \quad \alpha = 1, 2, \ldots, k$$

are satisfied for any $E \in \mathbf{F}_A$, $\eta \in B^I$, $x \in A^n$ ($n = 1, 2, \ldots$); of course, $\nu^\alpha$ are memoryless channels in the usual sense (as to the notations used above cf. [8]). The latter channel will be called *non-singular* if

(1.9) $$\nu^\alpha[a \mid b] > 0 \quad \text{for all} \quad a \in A\,, \quad b \in B \quad (\alpha = 1, 2, \ldots, k)\,.$$

As mentioned in the introductory part, we want to deal in this paper only with non-singular channels which yield from the point of view of the problem to be solved the fewest technical difficulties. Nevertheless, the problem of the existence of $\varepsilon$-capacity has a positive answer also in singular cases, which will be shown in a paper to follow.

Given $\nu = \Sigma \xi_\alpha \nu^\alpha$, let us set

(1.10) $$\mathscr{R}_\alpha(p) = \mathscr{R}_\alpha(p; \nu) = \sum_{a \in A, b \in B} \nu^\alpha[a \mid b]\, p(b) \log_2 \frac{\nu^\alpha[a \mid b]}{\sum_{b' \in B} \nu^\alpha[a \mid b']\, p(b')}$$

$$\text{for} \quad p \in P \quad (\alpha = 1, 2, \ldots, k)\,.$$

Then the supremum of infima

(1.11) $$r(\theta, \nu) = \sup_{p \in P} \inf\{t : \xi\{\alpha : \mathscr{R}_\alpha(p; \nu) \leqq t\} \geqq \theta\}\,, \quad 0 < \theta \leqq 1\,,$$

represents the maximum of the (lower) $\theta$-quantiles of the family $\{\mathscr{R}_\alpha(p)\}_{p \in P}$ of random variables (cf. [7]) on the probability space $(\{1, 2, \ldots, k\}, \mathbf{A}, \xi)$, where

(1.12) $$\mathbf{A} = \{\mathscr{A} : \mathscr{A} \subset \{1, 2, \ldots, k\}\}\,, \quad \xi(\mathscr{A}) = \sum_{\alpha \in \mathscr{A}} \xi_\alpha \quad \text{for} \quad \mathscr{A} \in \mathbf{A}$$

(i.e. measure $\xi$ is determined by the probability vector $\{\xi_\alpha\}_{\alpha=1}^k$), and where $\mathscr{R}_\alpha(p)$ are the (transmission) rates of the (memoryless) components $\nu^\alpha$ taken with respect to inputs $\mu^p$ (cf. (1.2), (1.3)).

Denoting by $\Lambda_\nu$ the (necessarily finite; cf. also Lemma 3 in Sec. 3 below) set of discontinuity points of function $r(\theta, \nu)$, in symbols

$$\Lambda_\nu = \{\varepsilon : r(\varepsilon, \nu) < r(\varepsilon + 0, \nu)\} \subset \{\xi(\mathscr{A}) : \mathscr{A} \in \mathbf{A}\}\,, \tag{1.13}$$

we are prepared to state the main theorem of this paper.

**Theorem on ε-Capacity.** *Let ν be a discrete channel which is decomposable into memoryless components and non-singular* (i.e. of the form $\nu = \Sigma \xi_\alpha \nu^\alpha$; cf. (1.8), (1.9)). *Then the limit*

$$C_\varepsilon = C_\varepsilon(\nu) = \lim_n (1/n) \log_2 S_n(\varepsilon, \nu) \tag{1.14}$$

[cf. (1.7)] *exists at least for all* $\varepsilon \notin \Lambda_\nu$ $(0 < \varepsilon < 1)$, i.e. *except at most a finite number of* ε*'s of the form* $\varepsilon = \xi(\mathscr{A})$, $\mathscr{A} \in \mathbf{A}$ (cf. (1.12)), *and equals the supremal ε-quantile* $r(\varepsilon, \nu)$, *in symbols*:

$$C_\varepsilon(\nu) = r(\varepsilon, \nu)\,, \quad 0 < \varepsilon < 1\,, \quad \varepsilon \notin \Lambda_\nu \quad \text{[cf. (1.11)]}\,. \tag{1.15}$$

More precisely, *let* ε, $0 < \varepsilon < 1$, $\varepsilon \notin \Lambda_\nu$, *be arbitrary*: *then there are positive constants* $K_\varepsilon, K'_\varepsilon$ *such that for any n* $(n = 1, 2, \ldots)$

$$2^{nC_\varepsilon - K_\varepsilon' \sqrt{n}} < S_n(\varepsilon, \nu) \leqq \bar{S}_n(\varepsilon, \nu) < 2^{nC_\varepsilon + K_\varepsilon \sqrt{n}}\,. \tag{1.16}$$

*Hence*

$$C_\varepsilon(\nu) = \lim_n (1/n) \log_2 \bar{S}_n(\varepsilon, \nu) \quad \text{for} \quad \varepsilon \notin \Lambda_\nu\,. \tag{1.17}$$

*Remark.* Explicit expressions for the constants $K_\varepsilon, K'_\varepsilon$ are given in Sec. 3 below; cf. (3.13) and (3.14), respectively. The theorem may be restated in Wolfowitz's terms (cf. [9]) as follows: for ε such that $\varepsilon \notin \Lambda_\nu$ $(0 < \varepsilon < 1)$ there exist positive constants $K_\varepsilon, K'_\varepsilon$ having the property that for any $n$ there is a code $(n, N, \varepsilon)$ with $N > \exp_2(nC_\varepsilon - K'_\varepsilon \sqrt{n})$, and there does not exist a code $(n, N, \varepsilon)$ satisfying

$$N \geqq \exp_2(nC_\varepsilon + K_\varepsilon \sqrt{n})\,.$$

The subsequent two sections contain a group of lemmas and theorems the proof of the main theorem is based upon.

## 2. BASIC LEMMATA

In this section we shall prove two lemmas which will play a fundamental role in deriving the upper and the lower bounds for the maximum length of $n$-dimensional

$\varepsilon$-codes. To do it, we shall need some additional notations. First we shall set (cf. (1.6)):

(2.1) $$d = \max(\pi(A), \pi(B)), \quad N(a, b \mid x, y) = \pi\{i : x_i = a,\ y_i = b\},$$
$$N(a \mid x) = \pi\{i : x_i = a\} = \sum_{b \in B} N(a, b \mid x, y), \quad N(b \mid y) = \pi\{i : y_i = b\ (0 \leqq i < n)\}$$
$$\text{for} \quad a \in A, \quad b \in B, \quad x \in A^n, \quad y \in B^n \quad (n = 1, 2, \ldots).$$

Following Wolfowitz (cf. [9], Chapter 2), we shall call an $n$-dimensional input sequence $y \in B^n$ a *p-sequence* if

$$|N(b \mid y) - np(b)| \leqq 2d^{1/2}\sigma(b; p)\, n^{1/2} \quad \text{for all} \quad b \in B$$
$$\text{where} \quad \sigma(b; p) = [p(b)(1 - p(b))]^{1/2}, \quad p \in P \quad (\text{cf. } (1.3)).$$

Furthermore we shall make use of the subsequent notations:

(2.2) $$F_n(p) = \{y : y \in B^n,\ y \text{ is a } p\text{-sequence}\},$$
$$p_0 = \min\{p(b) : b \in B_p\}, \quad B_p = \{b : p(b) > 0\} \quad \text{for} \quad p \in P.$$

An application of Chebyshev's inequality yields the relation (cf. (1.2))

(2.3) $$\mu^p[F_n(p)] > \tfrac{3}{4} \quad \text{for} \quad p \in P \quad (n = 1, 2, \ldots).$$

If $\nu$ is a discrete channel, and if $p \in P$, we shall put (compare with (1.7))

(2.4) $$S_n^*(\varepsilon, \nu\mu^p) = \max\{\pi(Y) : Y \subset F_n(p),\ Y \text{ is } \varepsilon\text{-discernible for } \nu\}, \quad 0 < \varepsilon < 1;$$

(2.5) $$I_n(xy; \nu\mu) = (1/n)\log_2\left(\nu[x \mid y]/(\Sigma_{y' \in B^n} \nu[x \mid y']\,\mu[y'])\right)$$
$$\text{for} \quad x \in A^n, \quad y \in B^n \ (n = 1, 2, \ldots), \quad \mu = \mu^p \quad (\text{cf. } (1.5)).$$

Throughout the remainder of this section we shall suppose that we are given a non-singular decomposable discrete channel $\nu$ with $k$ distinct memoryless components $\nu^\alpha$, i.e. $\nu = \Sigma\xi_\alpha\nu^\alpha$, where $\{\xi_\alpha\}_{\alpha=1}^k$ is a probability vector having positive components (cf. (1.8)); let us set (cf. (1.9))

(2.6) $$\xi_0 = \min_{1 \leqq \alpha \leqq k} \xi_\alpha > 0,$$
$$w_0 = \min\{\nu^\alpha[a \mid b] : a \in A,\ b \in B,\ \alpha = 1, 2, \ldots, k\} > 0.$$

An $n$-dimensional output sequence $x \in A^n$ is said to be $\alpha$-*generated* by an input sequence $y \in B^n$ (compare with [9], Chapter 2) with error less than $\varepsilon'$ $(0 < \varepsilon' < 1)$ if (cf. (2.1))

$$|N(a, b \mid x, y) - \nu^\alpha[a \mid b]\, N(b \mid y)| \leqq$$
$$\leqq d(\varepsilon')^{-1/2}\,\sigma_\alpha(a \mid b)\,[N(b \mid y)]^{1/2} \quad \text{for all} \quad a \in A, \quad b \in B$$
$$\text{where} \quad \sigma_\alpha(a \mid b) = (\nu^\alpha[a \mid b]\,(1 - \nu^\alpha[a \mid b]))^{1/2}; \quad \alpha = 1, 2, \ldots, k.$$

 Putting for $\alpha = 1, 2, \ldots, k$, $0 < \varepsilon' < 1$, $y \in B^n$ $(n = 1, 2, \ldots)$

$$(2.7) \qquad \Gamma_n^\alpha(y; \varepsilon') = \{x : x \in A^n,\ x \text{ is } \alpha\text{-generated by } y \text{ with error } < \varepsilon'\},$$

we easily find by making use of Chebyshev's inequality that (cf. (1.5))

$$(2.8) \qquad \nu^\alpha[\Gamma_n^\alpha(y; \varepsilon') \mid y] > 1 - \varepsilon';$$

the latter fact justifies the above terminology. On the other hand, it follows directly from the definition of $p$-sequence and of generated input sequences that the number (cf. (2.1))

$$(2.9) \qquad (1/n)\, N(a, b \mid x, y) \quad \text{lies inside the bounds}$$

$$\nu^\alpha[a \mid b]\, p(b) \pm 4d^2(\varepsilon')^{-1/2}\, n^{-1/2}(\nu^\alpha[a \mid b]\, p(b))^{1/4}$$

$$\text{for any} \quad a \in A, \quad b \in B_p, \quad x \in \Gamma_n^\alpha(y; \varepsilon'), \quad y \in F_n(p), \quad p \in P, \quad 0 < \varepsilon' < 1,$$

$$\alpha = 1, 2, \ldots, k; \quad n = 1, 2, \ldots;$$

cf. (1.9), (2.2), (2.7).

In what follows we shall associate with any $\mathscr{A} \in \mathbf{A}$, $\mathscr{A} \neq \emptyset$ (cf. (1.12 ), the decomposable channel $\nu^{\mathscr{A}}$ defined symbolically by

$$(2.10) \qquad \nu^{\mathscr{A}} = \sum_{\alpha \in \mathscr{A}} [\xi_\alpha / \xi(\mathscr{A})]\, \nu^\alpha \quad (\text{cf. } (1.8)).$$

Let us put for $y \in B^n$, $0 < \varepsilon' < 1$

$$(2.11) \qquad \Gamma_n^{\mathscr{A}}(y; \varepsilon') = \bigcup_{\alpha \in \mathscr{A}} \Gamma_n^\alpha(y; \varepsilon') \quad [\text{cf. } (2.7)].$$

For the rest of this section we shall make the assumption that we are given a probability $B$-vector $p$; all the considerations and notations to follow will be in connection with the latter $p$ fixed; especially, we shall set (cf. (1.2))

$$(2.12) \qquad \mu = \mu^p; \quad \omega^{\mathscr{A}}[x] = \sum_{y \in B^n} \nu^{\mathscr{A}}[x \mid y]\, \mu[y],$$

$$\omega^{\mathscr{A}}[E] = \sum_{y \in E} \omega^{\mathscr{A}}[x] \quad \text{for} \quad E \subset A^n, \quad x \in A^n,$$

$$\mathscr{A} \in \mathbf{A}(\mathscr{A} \neq \emptyset); \quad \omega^\alpha = \omega^{\{\alpha\}}, \quad \alpha = 1, 2, \ldots, k.$$

It may be shown by making use of (2.9) [and (2.1)] that

$$(2.13) \qquad (1/n)\, N(a \mid x) \quad \text{belongs to the open interval having bounds}$$

$$\omega^\alpha[a] \pm 4d^3(\varepsilon')^{-1/2}\, n^{-1/2}(\omega^\alpha[a])^{1/4}$$

$$\text{for any } a \in A \text{ (cf. (1.9))}, \ x \in \Gamma_n^\alpha(y; \varepsilon'), \ 0 < \varepsilon' < 1, \ \alpha = 1, 2, \ldots, k$$

where $y$ is any $n$-dimensional $p$-sequence ($n = 1, 2, \ldots$); as to deriving the bounds cf. also [9], Chapter 2. 

Before stating the first of the lemmas mentioned above, let us remind a well-known inequality

$$\sum_{i=1}^{m} r_i \log \frac{r_i}{u_i} \geqq (r_1 + \ldots + r_m) \log \frac{r_1 + \ldots + r_m}{u_1 + \ldots + u_m} \tag{2.14}$$

valid for any $r_i \geqq 0$, $u_i \geqq 0$ ($i = 1, \ldots, m$) of which we shall make use in the proof of the lemma (cf. [1], Sec. 2).

A very important result known for memoryless channels upon which the second part of the proof of the first lemma is basing, may be reformulated in our terms as follows (cf. (1.10), (2.5)): If $y \in B^n$ is a $p$-sequence, and if $x$ is an $n$-dimensional output sequence $\alpha$-generated by $y$ with error $< \varepsilon'$ ($0 < \varepsilon' < 1$) then

(2.15) $I_\alpha(xy; v^\alpha \mu^p)$ lies between the bounds

$$\mathscr{R}_\alpha(p) \pm 32 d^4 (\varepsilon')^{-1/2} n^{-1/2} ; \quad x = 1, 2, \ldots, k \quad (n = 1, 2, \ldots) ;$$

the latter fact is a combination of the inequalities stated in Lemma 2.1.3 and Lemma 2.1.5 of [9] together with relations

$$\max_{0 < t \leqq 1} t(-\log_2 t) = \mathrm{e}^{-1} \log_2 \mathrm{e} \leqq 1 ,$$

where e is the base of natural logarithms.

**Lemma 1.** *Let $\varepsilon'$ ($0 < \varepsilon' < 1$), $\mathscr{A}(\mathscr{A} \in \mathbf{A}, \mathscr{A} \neq \emptyset)$, $n$ ($n = 1, 2, \ldots$) be arbitrary, let $\alpha \in \mathscr{A}$, and let $y$ be an n-dimensional p-sequence. Then the inequalities* (cf. (1.10), (2.5), (2.10))

$$\mathscr{R}_\alpha(p) - \lambda_n < I_n(xy; v^{\mathscr{A}} \mu^p) < \mathscr{R}_\alpha(p) + \lambda_n \tag{2.16}$$

*are satisfied for any n-dimensional output sequence x which is $\alpha$-generated by the input sequence y with error $< \varepsilon'$* (i.e. for any $x \in \Gamma_n^\alpha(y; \varepsilon')$; cf. (2.7)) *provided that*

$$\lambda_n = \lambda_n(\varepsilon') = (1/n) \log_2 (2/\xi_0) + K/\sqrt{n} , \tag{2.17}$$

$$K = K(\varepsilon') = 36 d^4 w_0^{-1} (\varepsilon')^{-1/2} \quad [\text{cf. } (2.6)]$$

*so that quantity $\lambda_n$ is independent of p, $\mathscr{A}$, $\alpha$, y, and x.*

**Corollary.** *If $\varepsilon'$ ($0 < \varepsilon' < 1$), $\mathscr{A}$ ($\mathscr{A} \in \mathbf{A}$, $\mathscr{A} \neq \emptyset$), $n$ ($n = 1, 2, \ldots$), $y \in F_n(p)$* [cf. (2.2)] *are chosen arbitrarily then*

$$\min_{\alpha \in \mathscr{A}} \mathscr{R}_\alpha(p) - \lambda_n < I_n(xy; v^{\mathscr{A}} \mu^p) < \max_{\alpha \in \mathscr{A}} \mathscr{R}_\alpha(p) + \lambda_n \tag{2.18}$$

$$\textit{for any} \quad x \in \Gamma_n^{\mathscr{A}}(y; \varepsilon') \quad [\text{cf. } (2.11)] .$$

Proof. Let us suppose that the quantities $\varepsilon'$, $\mathscr{A}$, $\alpha$, $n$, $y$, $x$ such that satisfy the assumptions of the lemma, are given and kept fixed during the proof. Then we shall make use of the abbreviations (cf. (2.12))

$$(1)\qquad I_n = I_n(xy; \nu^{\mathscr{A}}\mu^p)\,,\quad I_n^\alpha = I_n(xy; \nu^\alpha\mu^p)\,,\quad \Sigma_\beta^* = \Sigma_{\beta\in\mathscr{A}}\,,$$
$$\nu^{\mathscr{A}} = \nu^{\mathscr{A}}[x \mid y]\,,\quad \nu^\alpha = \nu^\alpha[x \mid y]\,,\quad \omega^{\mathscr{A}} = \omega^{\mathscr{A}}[x]\,,\quad \omega^\alpha = \omega^\alpha[x]\,,$$
$$\nu^\beta = \nu^\beta[x \mid y]\,,\quad \omega^\beta = \omega^\beta[x] \quad \text{for} \quad \beta \in \mathscr{A}\,.$$

According to (2.5), (2.10), (1.12), and (2.12) we obtain that

$$\exp_2(nI_n) = \frac{\nu^{\mathscr{A}}}{\omega^{\mathscr{A}}} = \exp_2(nI_n^\alpha)\,\frac{\xi(\mathscr{A})\,\nu^{\mathscr{A}}}{\xi_\alpha\nu^\alpha}\,\frac{\xi_\alpha\omega^\alpha}{\xi(\mathscr{A})\,\omega^{\mathscr{A}}}$$

so that the inequalities

$$(2)\qquad \exp_2(nI_n) \leqq \left[1 + \Sigma^*_{\beta\neq\alpha}(\xi_\beta/\xi_\alpha)\frac{\nu^\beta}{\nu^\alpha}\right]\exp_2(nI_n^\alpha)\,,$$
$$\exp_2(nI_n) \geqq \left[1 + \Sigma^*_{\beta\neq\alpha}(\xi_\beta/\xi_\alpha)\frac{\omega^\beta}{\omega^\alpha}\right]^{-1}\exp_2(nI_n^\alpha)$$

must hold.

Given $\beta \neq \alpha$ $(\beta \in \mathscr{A})$, let us assume that $\nu^\beta[a \mid b] \neq \nu^\alpha[a \mid b]$ for some $a \in A$, $b \in B_p$; cf. (2.2). Since

$$\frac{\nu^\beta}{\nu^\alpha} = \prod_{a\in A,b\in B}\left(\frac{\nu^\beta[a \mid b]}{\nu^\alpha[a \mid b]}\right)^{N(a,b|x,y)},$$

and since by assumption $x \in \Gamma_n^\alpha(y; \varepsilon')$ where $y \in F_n(p)$, we obtain by (2.9) that

$$-(1/n)\log_2\frac{\nu^\beta}{\nu^\alpha} > \sum_{b\in B}p(b)\sum_{a\in A}\nu^\alpha[a \mid b]\log_2\frac{\nu^\alpha[a \mid b]}{\nu^\beta[a \mid b]} -$$
$$- 4d^2(\varepsilon')^{-1/2}\,n^{-1/2}\sum_{a\in A,b\in B_p}(\nu^\alpha[a \mid b]\,p(b))^{1/4}\left|\log_2\frac{\nu^\alpha[a \mid b]}{\nu^\beta[a \mid b]}\right|.$$

On the one hand,

$$\sum_{a\in A}\nu^\alpha[a \mid b]\log_2\frac{\nu^\alpha[a \mid b]}{\nu^\beta[a \mid b]} \geqq 0 \quad \text{for all} \quad b \in B$$

according to (2.14), on the other hand,

$$|\log_2(\nu^\alpha[a \mid b]/\nu^\beta[a \mid b])| \leqq \log_2(1/w_0) \leqq 1/w_0$$

(cf. (1.9)) as follows from (2.6) so that (cf. (2.1)) 

(3) $$-(1/n)\log_2(\nu^\beta/\nu^\alpha) > -K_0 n^{-1/2}\,, \quad \text{i.e.} \quad \nu^\beta/\nu^\alpha < \exp_2(K_0 n^{1/2})$$
$$\text{for} \quad K_0 = 4d^4 w_0^{-1}(\varepsilon')^{-1/2}\,;$$

the latter inequalities remain true also in case $\nu^\beta[\ \mid b] = \nu^\alpha[\ \mid b]$ for all $b \in B_p$ because then $\nu^\beta = \nu^\alpha$ $(K_0 > 1)$.

Proceeding in a similar way in case $\omega^\beta \neq \omega^\alpha$, we get by making use of the relation

$$\frac{\omega^\beta}{\omega^\alpha} = \prod_{a\in A}\left(\frac{\omega^\beta[a]}{\omega^\alpha[a]}\right)^{N(a|x)}$$

and of (2.13) that

$$-\frac{1}{n}\log_2\frac{\omega^\beta}{\omega^\alpha} > \sum_{a\in A}\omega^\alpha[a]\log_2\frac{\omega^\alpha[a]}{\omega^\beta[a]} -$$
$$- 4d^3(\varepsilon')^{-1/2}\, n^{-1/2}\sum_{a\in A}(\omega^\alpha[a])^{1/4}\left|\log_2\frac{\omega^\alpha[a]}{\omega^\beta[a]}\right|.$$

Since $\omega^\beta[a] \geqq w_0$ for any $a \in A$, $\beta \in \mathscr{A}$ as follows from (2.12), (1.9), and (2.6), we may assert that

$$|\log_2(\omega^\alpha[a]/\omega^\beta[a])| \leqq \log_2\frac{1}{w_0} \leqq \frac{1}{w_0}\,.$$

So a similar argument as above (cf. (2.14)) yields the inequality

(4) $$-(1/n)\log_2(\omega^\beta/\omega^\alpha) > -K_0 n^{-1/2}\,, \quad \text{i.e.} \quad \omega^\beta/\omega^\alpha < \exp_2(K_0 n^{1/2})$$

remaining in power also for $\omega^\beta = \omega^\alpha$.

Now combining (2) with (3) and (4), respectively, we immediately find that

(5) $$\exp_2(nI_n) < [1 + (1/\xi_0)\,2^{K_0\sqrt{n}}]\exp_2(nI_n^\alpha)\,,$$
$$\exp_2(nI_n) > [1 + (1/\xi_0)\,2^{K_0\sqrt{n}}]^{-1}\exp_2(nI_n^\alpha)\,.$$

Since

$$1 + (1/\xi_0)\,2^{K_0\sqrt{n}} \leqq (2/\xi_0)\,2^{K_0\sqrt{n}}\,,$$

we obtain from (5) the inequalities

(6) $$I_n < I_n^\alpha + (1/n)\log_2(2/\xi_0) + K_0/\sqrt{n}\,,$$
$$I_n > I_n^\alpha - (1/n)\log_2(2/\xi_0) - K_0/\sqrt{n}\,.$$

The last step to do is to apply (2.15) and replace $I_n^\alpha$ by its bounds: by doing that we get the desired inequalities (2.16) where $\lambda_n$ satisfies (2.17), Q.E.D.

*Remark.* Let us emphasize that in general case, if (1.9) is not satisfied, the argument used in deriving (4) is not more true because we can assert only that $\omega^{\beta}[a] \geqq w_0 p_0$ [cf. (2.2)], where $w_0$ is the minimum taken over positive $\nu^{\beta}[a \mid b]$; hence, in general, the lower bound for $I_n$ [cf. (1) above] is (and must be) substantially dependent on the probability vector $p$.

**Lemma 2.** *Let* $\varepsilon$ $(0 < \varepsilon < 1)$, $\mathscr{A}$ $(\mathscr{A} \in \mathbf{A},\ \mathscr{A} \neq \emptyset)$, $n$ $(n = 1, 2, \ldots)$ *be arbitrary; then* (cf. (2.4), (2.10), (1.10))

$$(2.19)\quad (1/n)\log_2 S_n^*(\varepsilon, \nu^{\mathscr{A}}\mu^p) < \max_{\alpha\in\mathscr{A}} \mathscr{R}_\alpha(p) + (1/n)\log_2(1 - \varepsilon - \varepsilon')^{-1} + \lambda_n(\varepsilon')$$

$$\text{for}\quad 0 < \varepsilon' < 1 - \varepsilon\,,$$

$$(2.20)\qquad (1/n)\log_2 S_n^*(\varepsilon, \nu^{\mathscr{A}}\mu^p) > \min_{\alpha\in\mathscr{A}} \mathscr{R}_\alpha(p) - (1/n)\log_2 \frac{2}{\varepsilon - \varepsilon'} - \lambda_n(\varepsilon')$$

$$\text{for}\quad 0 < \varepsilon' < \varepsilon\,,$$

*where* $\lambda_n(\varepsilon')$ *is given by* (2.17). *Especially,*

$$(2.21)\ \frac{\varepsilon}{4}\exp_2\left(n\left[r' - \lambda_n\left(\frac{\varepsilon}{2}\right)\right]\right) < S_n^*(\varepsilon, \nu^{\mathscr{A}}\mu^p) < \frac{2}{1-\varepsilon}\exp_2\left(n\left[r + \lambda_n\left(\frac{1-\varepsilon}{2}\right)\right]\right)$$

$$\textit{where}\quad r' = \min_{\alpha\in\mathscr{A}} \mathscr{R}_\alpha(p)\,,\quad r = \max_{\alpha\in\mathscr{A}} \mathscr{R}_\alpha(p)\,.$$

Proof. Let us suppose $\varepsilon$, $\mathscr{A}$, $n$ to be given and fixed. Then (2.8) together with (2.10) imply that

$$(2.22)\qquad \nu^{\mathscr{A}}[\Gamma_n^{\mathscr{A}}(y; \varepsilon') \mid y] > 1 - \varepsilon'$$

for any $y \in B^n$, $0 < \varepsilon' < 1$. Given $\varepsilon'$, let us set

$$(1)\qquad \Gamma(y) = \Gamma_n^{\mathscr{A}}(y; \varepsilon')\,.$$

I. Let us assume first that $\varepsilon' < 1 - \varepsilon$. Let $\{Q(y)\}_{y\in Y}$ be an $n$-dimensional code such that $Y \subset F = F_n(p)$, $Y$ is $\varepsilon$-discernible for $\nu^{\mathscr{A}}$, and

$$(2)\qquad \pi(Y) = S = S_n^*(\varepsilon, \nu^{\mathscr{A}}\mu^p)\,;$$

the existence of such a code is guaranteed by (2.4) [cf. also (2.3)]. If $E(y) = Q(y) \cap \cap\, \Gamma(y)$, it follows from (2.22), (1), and from $\varepsilon$-discernibility of $Y$ that

$$(3)\qquad \nu^{\mathscr{A}}[E(y) \mid y] > 1 - \varepsilon - \varepsilon' \quad \text{for all}\quad y \in Y\,.$$

If $x \in E(y)$, $y \in Y$ then $x \in \Gamma(y)$, $y \in F$ (i.e. $x \in \Gamma_n^{\alpha}(y; \varepsilon')$ for some $\alpha \in \mathscr{A}$, $y \in F_n(p)$) so that Lemma 1 (cf. Corollary) may be applied: it follows from (2.18) and (2.5) that

$$\nu^{\mathscr{A}}[x \mid y] < 2^{n\varrho}\omega^{\mathscr{A}}[x]\,,\quad \varrho = r + \lambda_n \quad (\text{cf. } (2.21))\,,$$

$$\text{i.e.}\quad \varepsilon_1 < \nu^{\mathscr{A}}[E(y) \mid y] < 2^{n\varrho}\omega^{\mathscr{A}}[E(y)]\,,\quad y \in Y$$

for $\varepsilon_1 = 1 - \varepsilon - \varepsilon'$; cf. (3). Since $E(y)$ are mutually disjoint, we obtain from the latter inequalities according to (2) that

$$S\varepsilon_1 < 2^{n\varrho}\omega^{\mathscr{A}}[\bigcup_{y\in Y} E(y)] \leqq 2^{n\varrho}, \quad \text{i.e.} \quad S < (1/\varepsilon_1)\, 2^{n\varrho},$$

which shows the validity of (2.19).

II. Now we shall assume that $\varepsilon' < \varepsilon$. Let us construct a code $\{Q(y)\}_{y\in Y}$ such that

$$(4) \qquad Y = \{y^1, \ldots, y^S\} \subset F, \quad S = \pi(Y), \quad F = F_n(y),$$

$$v^{\mathscr{A}}[Q(y^j) \mid y^j] > 1 - \varepsilon, \quad Q(y^j) = \Gamma(y^j) - \bigcup_{t=1}^{j-1} \Gamma(y^t), \quad j = 1, \ldots, S,$$

$$v^{\mathscr{A}}[\Gamma(y) - \bigcup_{j=1}^{S} \Gamma(y^j) \mid y] \leqq 1 - \varepsilon \quad \text{for all} \quad y \in F,$$

where $\Gamma(y)$ is defined by (1); the existence of a non-empty set $Y$ satisfying conditions (4) follows from (2.22) because $\varepsilon' < \varepsilon$, and from (2.3) which shows that $F$ is not empty. An easy calculation yields the inequalities (cf. (2.12), (2.22))

$$(1 - \varepsilon')\,\mu[F] - \omega^{\mathscr{A}}[\bigcup_{j=1}^{S} Q(y^j)] < \sum_{y\in F} v^{\mathscr{A}}[\Gamma(y) - \bigcup_{j=1}^{S} Q(y^j) \mid y]\,\mu[y] \leqq (1 - \varepsilon)\,\mu[F],$$

$$\text{i.e.} \quad \tfrac{3}{4}(\varepsilon - \varepsilon') < (\varepsilon - \varepsilon')\,\mu[F] < \omega^{\mathscr{A}}[\bigcup_{j=1}^{S} Q(y^j)]$$

as follows from (2.3). Another application of the corollary to Lemma 1 will yield the inequalities

$$2^{n\varrho'}\omega^{\mathscr{A}}[Q(y^j)] < v^{\mathscr{A}}[Q(y^j) \mid y^j] \leqq 1, \quad j = 1, 2, \ldots, S,$$

$$\text{where} \quad \varrho' = r' - \lambda_n \quad [\text{cf. } (2.21)]$$

so that from the disjointness of $Q(y^j)$ and from above it will follow that

$$\tfrac{1}{2}(\varepsilon - \varepsilon') < \omega^{\mathscr{A}}[\bigcup_{j=1}^{S} Q(y^j)] < 2^{-n\varrho'} \,.\, S,$$

$$\text{i.e.} \quad \tfrac{1}{2}(\varepsilon - \varepsilon')\, 2^{n\varrho'} < S \leqq S_n^*(\varepsilon, v^{\mathscr{A}}\mu^P),$$

by which the validity of (2.20) is verified.

The inequalities given in (2.1) are easily obtained from (2.19) and (2.20) by setting $\varepsilon' = \frac{1}{2}(1 - \varepsilon)$ and $\varepsilon' = \frac{1}{2}\varepsilon$, respectively, which concludes the proof.

## 3. THE UPPER AND THE LOWER BOUNDS

In this section we shall establish the validity of all the assertions that are stated in the theorem on $\varepsilon$-capacity. We shall first deal with the problem of deriving the upper and the lower bounds for the maximum length of $n$-dimensional $\varepsilon$-codes for $\varepsilon$ given. In what follows we shall need some other notations, viz. (cf. (1.10), (1.11))

$$\mathbf{A}_\theta = \{\mathscr{A} : \mathscr{A} \in \mathbf{A},\ \xi(\mathscr{A}) \geqq \theta\}\,, \quad 0 < \theta \leqq 1 \quad [\text{cf. } (1.12)]\,; \tag{3.1}$$

$$r'(\theta, v) = \sup_{p\in P} \sup \{t : \xi\{\alpha : \mathscr{R}_\alpha(p; v) \geqq t\} \geqq 1 - \theta\}\,, \quad 0 \leqq \theta < 1\,, \tag{3.2}$$

the latter for $v$ decomposable into memoryless components. If $r'(\theta, v)$ is compared with $r(\theta, v)$, it is immediately seen that $r'(\theta, v)$ represents the maximum of the upper $\theta$-quantiles of the family $\{\mathscr{R}_\alpha(p)\}_{p\in P}$ taken with respect to the probability distribution $\xi$ for $v$ given; the fact that either of the supremal quantiles $r(\theta, v)$ and $r'(\theta, v)$ is a maximum, will be proved in the following

**Lemma 3.** *The quantile functions $r(\ ,v)$ and $r'(\ ,v)$ are both monotonically increasing, $r(\ ,v)$ is continuous from the left, and*

$$r'(\theta - 0, v) = r(\theta, v) \leqq r'(\theta, v) \quad \text{for} \quad 0 < \theta < 1\,.$$

*Both the quantile functions have the same set of discontinuities $\Lambda_v$* (cf. (1.13)),

$$r'(\varepsilon, v) = r(\varepsilon, v) \quad \text{for} \quad \varepsilon \notin \Lambda_v\,, \quad 0 < \varepsilon < 1\,, \tag{3.3}$$

*and may be expressed in the form* (cf. (3.1))

$$\begin{aligned} r(\theta, v) &= \max_{p\in P} \min_{\mathscr{A}\in \mathbf{A}_\theta} \max_{\alpha\in\mathscr{A}} \mathscr{R}_\alpha(p)\,, \qquad 0 < \theta \leqq 1\,; \\ r'(\theta, v) &= \max_{p\in P} \max_{\mathscr{A}\in \mathbf{A}_{1-\theta}} \min_{\alpha\in\mathscr{A}} \mathscr{R}_\alpha(p)\,, \quad 0 \leqq \theta < 1\,. \end{aligned} \tag{3.4}$$

Proof. The first part of the lemma up to (3.3) is an immediate consequence of the quantile character of both quantities; cf. Lemma 3.5 in [7], and more details in [6], Lemma 3.5. It remains to prove formulas (3.4). In either formula in (3.4) the maximum exists because $\mathscr{R}_\alpha(p)$ is a continuous function of parameter $p$ if $P$ is imbedded in $\pi(B)$-dimensional Euclidean space (cf. (1.10)). Then the first formula easily follows from definition (1.11) and from the relation

$$\{t : \xi\{\alpha : \mathscr{R}_\alpha(p) \leqq t\} \geqq \theta\} = \bigcup_{\mathscr{A}\in\mathbf{A}_\theta} \{t : \max_{\alpha\in\mathscr{A}} \mathscr{R}_\alpha(p) \leqq t\}\,.$$

Similarly, the second formula is a consequence of definition (3.2) and of a relation which is dual to the latter with $\theta$ replaced by $1 - \theta$.

**Theorem 1.** *If $v$ is a discrete channel which is decomposable and non-singular, $v = \Sigma\xi_\alpha v^\alpha$ with memoryless components, then for any $n$ $(n = 1, 2, \ldots)$ and $\varepsilon$ $(0 <$

$< \varepsilon < 1)$ *the inequality* [cf. (1.7), (1.11)]

$$(3.5) \qquad (1/n)\log_2 S_n(\varepsilon, v) < r(\theta, v) + \lambda_n(\varepsilon') + $$
$$+ (1/n)\log_2 \frac{1}{1 - (\varepsilon/\theta) - \varepsilon'} + d\,\frac{\log_2 (n+1)}{n}$$

*holds for every* $\theta$ *such that* $\varepsilon < \theta \leqq 1$, *and for any* $\varepsilon'$, $0 < \varepsilon' < 1 - (\varepsilon/\theta)$, *where* $\lambda_n(\varepsilon')$ *is defined by* (2.17); *especially,*

$$(3.6) \qquad S_n(\varepsilon, v) < \frac{2\theta}{\theta - \varepsilon}(n+1)^d_{\sharp} \exp_2\left(n\left[r(\theta, v) + \lambda_n\left(\frac{\theta - \varepsilon}{2\theta}\right)\right]\right)$$
$$\textit{for} \quad \varepsilon < \theta \leqq 1\,.$$

Proof. First we shall introduce some more notations which will be employed in the proof. Given $n$, let us set

$$(1) \qquad P_n = \{p : p \in P,\ np(b) \text{ is an integer for every } b \in B\}\,.$$

Given $\varepsilon$, let $\{Q(y)\}_{y \in Y}$ be an $n$-dimensional code such that $Y$ is $\varepsilon$-discernible for $v(Y \subset \subset B^n)$, and

$$(2) \qquad \pi(Y) = S = S_n(\varepsilon, v)\,;$$

such a code exists by definition (1.7). Setting (cf. (2.1))

$$(3) \qquad Y_p = \{y : y \in Y,\ N(b \mid y) = np(b) \text{ for every } b \in B\}\,, \quad p \in P_n\,,$$

we conclude from the disjointness of the family $Y_p$ $(p \in P_n)$ that

$$(4) \qquad S = \sum_{p \in P_n} \pi(Y_p)\,.$$

Given $\theta > \varepsilon$, let us fix $p \in P_n$ and take $\mathscr{A} \in \boldsymbol{A}_\theta$ such that (cf. (3.1))

$$(5) \qquad r = \max_{\alpha \in \mathscr{A}} \mathscr{R}_\alpha(p) = \min_{\mathscr{A}' \in \boldsymbol{A}_\theta} \max_{\alpha \in \mathscr{A}'} \mathscr{R}_\alpha(p)\,.$$

Then it follows from $\varepsilon$-discernibility of $Y$ for $v$ and from relation $\mathscr{A} \in \boldsymbol{A}_\theta$ which implies $\xi(\mathscr{A}) \geqq \theta$, that

$$v^{\mathscr{A}}[Q(y) \mid y] \geqq [\xi(\mathscr{A})]^{-1} (v[Q(y) \mid y] - \sum_{\alpha \notin \mathscr{A}} \xi_\alpha) >$$
$$> \frac{1}{\xi(\mathscr{A})}(1 - \varepsilon - [1 - \xi(\mathscr{A})]) = 1 - \frac{\varepsilon}{\xi(\mathscr{A})} \geqq 1 - \frac{\varepsilon}{\theta}$$

for any $y \in Y$, especially for any $y \in Y_p$. However, $Y_p \subset F_n(p)$ as follows from (3) and

(2.2), so we have found that (cf. (2.4))

$$\pi(Y_p) \leqq S_n^*(\varepsilon/\theta, v^{\mathscr{A}}\mu^p)\,.$$

Given $\varepsilon' < 1 - (\varepsilon/\theta)$ [$\varepsilon' > 0$], we obtain by making use of Lemma 2 and of the latter inequality that

$$\pi(Y_p) < (1/\varepsilon_1)\, 2^{n(r+\lambda)} \quad \text{where} \quad \varepsilon_1 = 1 - (\varepsilon/\theta) - \varepsilon'\,, \quad \lambda = \lambda_n(\varepsilon')\,;$$

cf. (2.19), (2.21), (5), (2.17). Lemma 3 shows that $r \leqq r(\theta, v)$ [cf. (3.4), (5)] so that we may assert that

$$\pi(Y_p) < (1/\varepsilon_1)\, 2^{n\varrho} \quad \text{for any} \quad p \in P_n \quad [\text{cf. } (1)]\,,$$
$$\text{where} \quad \varrho = r(\theta, v) + \lambda_n(\varepsilon')\,.$$

From here and from (4) we immediately conclude that

$$S < \frac{\pi(P_n)}{\varepsilon_1}\, 2^{n\varrho}\,.$$

As easily follows from definition (1), $\pi(P_n) \leqq (n+1)^d$ so that

$$S < (n+1)^d\, (1/\varepsilon_1)\, 2^{n\varrho}\,.$$

The latter inequality is, by (2), equivalent to the desired result stated in (3.5). The inequality in (3.6) is easily obtained by setting $\varepsilon' = (\theta - \varepsilon)/2\theta$, which is in accordance with (2.21). This concludes the proof.

**Theorem 2.** *If $v$ is a discrete channel which is decomposable and non-singular, $v = \Sigma \xi_\alpha v^\alpha$ with memoryless components, then for any $n$ ($n = 1, 2, \ldots$) and $\varepsilon$ ($0 < \varepsilon < 1$) the inequality* [cf. (1.7), (3.2)]

$$(3.7) \qquad (1/n)\log_2 S_n(\varepsilon, v) > r'(\theta', v) - \lambda_n(\varepsilon') - (1/n)\log_2 \frac{2}{\varepsilon - \theta' - \varepsilon'}$$

*holds for every $\theta'$ such that $0 \leqq \theta' < \varepsilon$, and for any $\varepsilon'$, $\varepsilon' < \varepsilon - \theta'$, where $\lambda_n(\varepsilon')$ is defined by* (2.17); *especially,*

$$(3.8) \quad S_n(\varepsilon, v) > \frac{\varepsilon - \theta'}{4} \exp_2\left(n\left[r'(\theta', v) - \lambda_n\left(\frac{\varepsilon - \theta'}{2}\right)\right]\right) \quad \textit{for} \quad 0 \leqq \theta' < \varepsilon\,.$$

Proof. Given $n$, $\varepsilon$, and $\theta' < \varepsilon(\theta' \geqq 0)$, let us choose $p \in P$ and $\mathscr{A} \in \boldsymbol{A}'$ such that $\boldsymbol{A}' = \boldsymbol{A}_{(1-\theta')}$ [cf. (3.1)],

$$(1) \qquad r' = \min_{\alpha \in \mathscr{A}} \mathscr{R}_\alpha(p) = \max_{p \in P, \mathscr{A}' \in \boldsymbol{A}'} \min_{\alpha \in \mathscr{A}} \mathscr{R}_\alpha(p) = r'(\theta', v)\,;$$

the existence of such a pair $p, \mathscr{A}$ is guaranteed by Lemma 3. Now let us construct an $n$-dimensional code $\{Q(y)\}_{y \in Y}$ having the property that $Y \subset F_n(p)$ [cf. (2.2)], $Y$ is $\varepsilon_2$-discernible for $v^{\mathscr{A}}$ (where $\varepsilon_2 = \varepsilon - \theta'$), and [cf. (2.4)]

$$\pi(Y) = S_n^*(\varepsilon_2, v^{\mathscr{A}}\mu^p)\,, \quad \varepsilon_2 = \varepsilon - \theta'\,. \tag{2}$$

Making use of that $\mathscr{A} \in \mathbf{A}'$ so that $\xi(\mathscr{A}) \geqq 1 - \theta'$, and of definition (2.10) of $v^{\mathscr{A}}$, we obtain the inequalities

$$v[Q(y) \mid y] \geqq \xi(\mathscr{A})\, v^{\mathscr{A}}[Q(y) \mid y] > (1 - \theta')(1 - \varepsilon_2) \geqq 1 - \varepsilon$$

valid for any every $y \in Y$; so we have found that [cf. (1.7)]

$$\pi(Y) \leqq S_n(\varepsilon, v)\,. \tag{3}$$

Now applying Lemma 2 to get a lower bound for $S_n^*(\varepsilon_2, v^{\mathscr{A}}\mu^p)$, we conclude by taking into account (1), (2), and (3) that

$$S_n(\varepsilon, v) > \frac{\varepsilon_2 - \varepsilon'}{2}\, 2^{n(r' - \lambda)} \quad \text{where} \quad \lambda = \lambda_n(\varepsilon')\,, \quad 0 < \varepsilon' < \varepsilon_2$$

[cf. (2.20), (2.21)], which gives an equivalent form of the desired inequality (3.7). The special case given in (3.8) corresponds to the lower bound stated in (2.21) with $\varepsilon$ replaced by $\varepsilon - \theta'$, Q.E.D.

As a consequence of the validity of the assertions stated in both theorems of this section we shall prove the following

**Corollary.** *Let $v$ be a discrete channel which is decomposable into memoryless components and non-singular,* i.e. $v = \Sigma \xi_\alpha v^\alpha$; *then for every $\varepsilon$ such that* $0 < \varepsilon < 1$, $\varepsilon \notin \Lambda_v$ (cf. (1.13)), *there is a positive constant $\delta_\varepsilon$ not exceeding $\varepsilon$ and having the property that, for any* $n = 1, 2, \ldots$, *the inequalities* (cf. (1.7), (1.11))

$$\tfrac{1}{4}\xi_0 \delta_\varepsilon \exp_2 (nr(\varepsilon, v) - n^{1/2} K(\delta_\varepsilon)) < S_n(\varepsilon, v) < \tag{3.9}$$

$$< \frac{2}{\xi_0 \delta_\varepsilon} \exp_2 (nr(\varepsilon, v) + n^{1/2}[d \log_2 \mathrm{e} + K(\delta_\varepsilon)])$$

*hold with $\xi_0$ expressed as minimum of $\xi_\alpha$ by* (2.6), *where* e *is the base of natural logarithms, and where $K$ as a function of parameter $\varepsilon'$ is expressed by formula* (2.17), i.e.

$$K(\delta) = \frac{36 d^4}{w_0 \sqrt{\delta}}\,. \tag{3.10}$$

*The constant $\delta_\varepsilon$ may be chosen as*

$$\delta_\varepsilon = \tfrac{1}{4} \min \{|\xi(\mathscr{A}) - \varepsilon| : \mathscr{A} \in \mathbf{A},\ \xi(\mathscr{A}) \neq \varepsilon\}\,. \tag{3.11}$$

Proof. Defining $\delta = \delta_\varepsilon$ by (3.11), we immediately find that $\delta \leqq \frac{1}{4} \min(\varepsilon, 1 - \varepsilon)$ so that Theorem 1 and Theorem 2 may be applied to $\theta = \varepsilon + 2\delta$ and $\theta' = \varepsilon - 2\delta$, respectively. Replacing $\lambda_n$ in formulas (3.6) and (3.8) by its expression according to (2.17), and making use of the inequality $n + 1 \leqq e^{\sqrt{n}}$, we obtain the bounds, for $n$ given arbitrarily, in the form

$$\tfrac{1}{4}\xi_0\delta \exp_2 \left[nr'(\varepsilon - 2\delta, \nu) - n^{1/2} K(\delta)\right] < S_n(\varepsilon, \nu) <$$
$$< \frac{2}{\xi_0\delta}(\varepsilon + 2\delta)\, e^{d\sqrt{n}} \exp_2\left[nr(\varepsilon + 2\delta) + n^{1/2} K\left(\frac{\delta}{\varepsilon + 2\delta}\right)\right].$$

On the other hand, it follows from the definition of $\delta$ that the open interval $(\varepsilon - 4\delta, \varepsilon + 4\delta)$ does not contain any $\xi(\mathscr{A}) \neq \varepsilon$ where $\mathscr{A} \in \boldsymbol{A}$; hence we conclude according to (3.1) that

$$\boldsymbol{A}_\theta = \boldsymbol{A}_\varepsilon \quad \text{for} \quad \varepsilon - 4\delta < \theta < \varepsilon + 4\delta\,,$$

because in the latter case $\xi(\mathscr{A}) \geqq \theta$ implies $\xi(\mathscr{A}) \geqq \varepsilon + 4\delta$. The relation just established together with those given in Lemma 3 show that [cf. (3.4), and also (3.3)]

$$r(\varepsilon - 4\delta + 0, \nu) = r(\varepsilon + 4\delta, \nu)\,, \tag{3.12}$$

and then $r'(\varepsilon - 2\delta, \nu) = r(\varepsilon, \nu) = r(\varepsilon + 2\delta, \nu)$. Now the desired inequalities (3.9) follow from those given above and from that $\varepsilon + 2\delta < 1$, and that $K$ is a decreasing function of its parameter.

*Remark.* The latter considerations also show that the inclusion given in (1.13) must be valid; cf. (3.12). In other words, if $\varepsilon$ does not belong to the set $\{\xi(\mathscr{A}) : \mathscr{A} \in \boldsymbol{A}\}$, then it is a continuity point of the quantile function $r(\ ,\nu)$, i.e. $r(\varepsilon + 0, \nu) = r(\varepsilon, \nu)$. Let us point out that it may happen that a continuity point of $r(\ ,\nu)$ is of the form $\varepsilon = \xi(\mathscr{A})$; for example, this is the case if $\nu$ is a channel with memoryless components which is regularly decomposable in the sense of [7], and which, at the same time, has the property that all the component channels possess the same capacity; then all $\varepsilon$ lying between 0 and 1 are points of continuity as follows from the main theorem of [6].

Proof of the theorem on $\varepsilon$-capacity. Under the assumptions of the theorem as stated in Section 1, we may apply the preceding corollary to any $\varepsilon \notin \Lambda_\nu$ and conclude that the limit $C_\varepsilon(\nu)$ expressed by (1.14) equals the supremal quantile $r(\varepsilon, \nu)$ as stated in (1.15).

To show the validity of (1.16), let us take $\varepsilon$ fixed and put $\Delta \doteq \delta_\varepsilon$ where $\delta_\varepsilon$ is defined by (3.11). Applying the corollary to $\varepsilon + \Delta$, we obtain the inequality

$$S_n(\varepsilon + \Delta, \nu) < \frac{8}{\xi_0\Delta} \exp_2\left(nr(\varepsilon, \nu) + n^{1/2}[d \log_2 e + K(\tfrac{1}{4}\Delta]\right)$$

because $\delta \geqq \frac{1}{4}\Delta$ for $\delta = \delta_{\varepsilon+\Delta}$ (cf. (3.11)), as follows from the definitions of both the latter constants, from (3.12) with $\delta$ replaced by $\Delta$, and from the fact that if $\varepsilon \neq \xi(\mathscr{A})$ for all $\mathscr{A} \in \mathbf{A}$,

$$4\delta = \min_{\mathscr{A}} |\xi(\mathscr{A}) - (\varepsilon + \Delta)| \geqq \min_{\mathscr{A}} |\xi(\mathscr{A}) - \varepsilon| - \Delta = 3\Delta ,$$

and that $\delta = \frac{1}{4}\Delta$ if $\varepsilon = \xi(\mathscr{A})$ for some $\mathscr{A} \in \mathbf{A}$. Setting

$$K_\varepsilon = K(\tfrac{1}{4}\delta_\varepsilon) + \log_2 \frac{1}{\xi_0 \delta_\varepsilon} + 5d , \tag{3.13}$$

$$K_\varepsilon' = K(\delta_\varepsilon) + \log_2 \frac{1}{\xi_0 \delta_\varepsilon} + 2 \quad [\text{cf. } (3.10), (3.11), (2.6)] ,$$

and making use of the inequalities

$$\bar{S}_n(\varepsilon, v) \leqq S_n(\varepsilon + \Delta, v) , \quad \log_2 \mathrm{e} \leqq 2 , \quad 3 \leqq 3d ,$$

we get from (3.9) and from above that

$$r(\varepsilon, v) - n^{-1/2} K_\varepsilon' < \frac{1}{n} \log_2 S_n(\varepsilon, v) \leqq \frac{1}{n} \log_2 \bar{S}_n(\varepsilon, v) < r(\varepsilon, v) + n^{-1/2} K_\varepsilon .$$

Since the latter inequalities are equivalent to those stated in (1.16), it is seen that the bounds given in (1.16) are functioning with constants $K_\varepsilon$, $K_\varepsilon'$ defined by (3.13). On the other hand, from the inequalities just established the assertion of (1.17) immediately follows, which concludes the proof of the theorem.

As a final remark, let us mention that a more comfortable expression for constants $K_\varepsilon$, $K_\varepsilon'$ is given by

$$K_\varepsilon = K_\varepsilon' = \frac{2^7 d^4}{\xi_0 w_0 \delta_\varepsilon} \quad [\text{cf. } (3.11), (2.6), (2.1)] . \tag{3.14}$$

The latter expression is easily obtained by enlarging the constants stated in (3.13) in accordance with the definition of $K(\delta)$ as expressed by (3.10).

(Received November 9, 1971.)

REFERENCES

[1] R. L. Dobrušin: A general formulation of Shannon's fundamental theorem in information theory (in Russian). Usp. matem. nauk *14* (1959), 3—104.

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

[3] J. Nedoma: Capacity of a discrete channel. Transact. First Prague Conf. on Infor. Theory etc. Prague 1957, 143—181.

[4] J. Nedoma: On non-ergodic channels. Transact. Second Prague Conf. on Inform. Theory etc. Prague 1960, 363—395.

[5] K. Winkelbauer: On discrete information sources. Transact. Third Prague Confer. on Information Theory etc. Prague 1964, 765—830.

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

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

[8] K. Winkelbauer: On the capacity of decomposable channels. Transact. Sixth Prague Conference on Information Theory etc. held in 1971 (in press).

[9] J. Wolfowitz: Coding theorems of information theory. Second ed. Berlin 1964.

VÝTAH

## Diskrétní kanály rozložitelné v bezpaměťové složky

KAREL WINKELBAUER

V práci se studuje pojem $\varepsilon$-kapacity, který byl zaveden autorem v článku [6], v případě diskrétních kanálů, jež lze rozložit na konečný počet bezpaměťových komponent. Celá práce je věnována formulaci a důkazu základního teorému o existenci $\varepsilon$-kapacity pro případ sdělovacích kanálů popsaného typu. Podstatný rozdíl mezi touto prací a pojednáním [6] je v tom, že se zde neomezujeme na případ kanálů regulárně rozložitelných (ve smyslu zavedeném v článku [7]), jak se činí v práci [6].

*Doc. RNDr. Mg. Mat. Karel Winkelbauer DrSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Vyšehradská 49, Praha 2.*