Proceedings of the Colloquium on Information Theory
organized by the Bolyai Mathematical Society, Debrecen, (Hungary), 1967

Certain results in coding theory for compound channels I.

by

R. Ahlswede

Department of Mathematics, The Ohio State University,
Columbus, USA

0. ABSTRACT

In [2] we proved a coding theorem and the weak converse of the coding theorem for averaged channels under different assumptions on the time structure and the output alphabet of the channel, and we gave explicit formulas for the weak capacity. The strong converse of the coding theorem does not hold; therefore, it is of interest to know the λ-capacity $C_1(\lambda) = \lim_{n\to\infty} \frac{1}{n} \log N(n,\lambda)$ [9]. We could not decide, if $C_1(\lambda)$ exists for every λ $(0<\lambda<1)$; however, we give upper and lower bounds for $\limsup_{n\to\infty} N(n,\lambda)$ and $\liminf_{n\to\infty} N(n,\lambda)$, (Chapt. I, 2).

In Chapter I, 5 we proved a coding theorem and the weak converse for stationary semicontinuous averaged channels, where the average is taken with respect to a general probability distribution. (We obtained this result in [2], Chapter 7, remark 1, by a different method only under an additional restriction.) The new method (Lemma 3) applies to averaged channels with side information (cf. [8], [13]).

In Chapter II we introduce a new compound channel:

the sender can choose for the transmission of a code word the individual channel over which he wants to transmit.

Acknowledgment.

The author wishes to thank Professor Wolfowitz for stimulating discussions and remarks.

I. AVERAGED CHANNELS WHERE EITHER THE SENDER OR THE RECEIVER KNOWS THE INDIVIDUAL CHANNEL WHICH GOVERNS THE TRANSMISSION

1. INTRODUCTION AND DEFINITIONS

Simultaneous discrete memoryless channels, where the individual channel which governs the transmission is known to the sender or receiver were discussed for the first time by Wolfowitz [8], [11]. Kesten gave an extension to the semicontinuous case [6]. Compound channels where the channel probability function (c. p. f.) for each letter is stochastically determined were introduced by Shannon [13]. Wolfowitz proved strong converses [11]. In [2] we proved a coding theorem and its weak converse for averaged channels under different assumptions on the time structure (stationary, almost periodic, nonstationary) and the output alphabet (finite, infinite) of the channel. We introduce now averaged channels with side information. First let us repeat the definition of a general averaged channel.

Let $X^t = \{1, \dots a\}$ for $t = 1, 2, \dots$ and $(X'^t, \mathcal{L}'^t) = (X', \mathcal{L}')$ for $t = 1, 2, \dots$ where X' is an arbitrary set and $\mathcal{L}'$ is a σ-algebra of subsets in X'.

Furthermore, let $S = \{s, \dots\}$ be a nonempty (index) set, $(S, \mathcal{M}, q)$ a normed measure space and let $F^t(\cdot|1|s), \dots F^t(\cdot|a|s)$ be probability distributions (p. d.) on $(X'^t, \mathcal{L}'^t)$ $(t \in N,\ s \in S)$.

For each $x_n = (x^1, \dots, x^n) \in X_n = \prod_{t=1}^{n} X^t$ we define a p. d. on

$$\left(X'_n = \prod_{t=1}^{n} X'^t, \mathcal{L}'_n = \prod_{t=1}^{n} \mathcal{L}'^t\right) \text{ by } F_n(\cdot|x_n|s) = \prod_{t=1}^{n} F^t(\cdot|x^t|s) .$$

The sequence of kernels $(F_n(\cdot|\cdot|s))$ $n = 1, 2, \dots$ forms a semicontinuous (in general nonstationary) channel without memory. [In case $X'^t = X'$ is finite, the kernels $F^t(\cdot|\cdot|s)$ are given by stochastic matrices $w^t(k|i|s) = F^t(\{k\}|i|s)$ $(i \in X, k \in X')$. We speak then of a discrete channel without memory]. Thus we have assigned to each $s \in S$ a semicontinuous channel. If we are interested in the simultaneous behaviour of all these channels, then we call this indexed set of channels a simultaneous channel (semicontinuous, without memory). The set $\{F_n(\cdot|\cdot|s) | s \in S)$ designed by S_n is called a simultaneous channel in the discrete time-interval $\langle 1, n \rangle$. (cf. [8], [6], [1])

If $F_n(A|x_n|s)$ is a measurable function on $(S, \mathcal{M}, q)$ for each $A \in \mathcal{L}'_n, x_n \in X_n$, then we can define an averaged channel by

$$P_n(A|x_n) = \int_S F_n(A|x_n|s)\, dq(s)$$

for $A \in \mathcal{L}'_n, x_n \in X_n, n = 1, 2, \ldots$.

A more intuitive description of this channel can be given as follows: at the beginning of the transmission of each word of length n an independent random experiment is performed according to $(S, \mathcal{M}, q)$ with probability $q(s)$ that the outcome of the experiment be $s \in S$. If s is the outcome of the experiment the word (of length n) is transmitted according to $F_n(\cdot|\cdot|s)$.

The definition of a code depends on the knowledge of the channel $F_n(\cdot|\cdot|s)$ by the sender and or receiver. If neither knows the channel over which the message is transmitted, a (n, N, λ) code for the compound channel is defined as a set $\{(u_1, A_1), \ldots (u_N, A_N)\}$, where $u_i \in X_n$, $A_i \in \mathcal{L}'_n$ for $i = 1, \ldots N$, $A_i \cap A_j = \phi$ for $i \neq j$, such that

(1) $$\int_S F_n(A_i | u_i | s)\, dq(s) \geq 1 - \lambda \qquad (i = 1, \ldots N)$$

The u_i and A_i do not depend on s . (cf. [2])

Paragraph 2 is concerned with existence problems of the λ-capacity of this channel.(cf. [9])
If only the sender knows the channel of transmission, the u_i's but not the A_i may depend on s . A (n,N,λ) code $\{(u_1(s), A_1), \ldots, (u_N(s), A_n)\}$ must now satisfy

(2) $$\int_S F_n(A_i | u_i(s) | s)\, dq(s) \geq 1 - \lambda \qquad (i = 1, \ldots N)$$

If only the receiver knows the channel, the A_i but not the u_i may depend on s and A_i in (1) is replaced by $A_i(s)$. Finally if both sender and receiver know the channel, u_i and A_i may depend on s .

We put

$N_1(n,\lambda)$ = maximal N for which a (n,N,λ) code exists if neither sender nor receiver knows the channel over which the message is transmitted.

$N_i(n,\lambda)$ for $i = 2,3,4$ are the maximal N for which a (n,N,λ) code exists respectively if the sender only $(i=2)$, the receiver only $(i=3)$, the sender and receiver $(i=4)$ know the channel.

We designate the different compound channels by $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4$. In this chapter we prove a coding theorem and the weak converse for $\mathcal{C}_i$ $(i = 1, \dots 4)$ in the stationary case. (I, 4, 5) (For results in the nonstationary case cf. [2], [5])

In § 3 we show that in general memory need not "increase capacity" (cf. [10]).

2. A REMARK ON THE λ-CAPACITY OF $\mathcal{C}_1$.

According to Wolfowitz [9] $\lim_{n\to\infty} \frac{1}{n} \log N_1(n,\lambda)$ is, if the limit exists, the λ-capacity $C_1(\lambda)$ of the channel $\mathcal{C}_1$. It is known [5], [9], [2] that $\mathcal{C}_1(\lambda)$ cannot be constant, because the strong converse of the coding theorem does not hold in general We need the

LEMMA 1:

If

$\frac{1}{N}\sum_{i=1}^{N} P_n(A_i \mid u_i) \geq 1-\lambda$ for $i = 1,2,\dots N$ and $1 > \gamma, \beta > 0$

such that $\gamma\beta > \lambda$, then

$$q\left\{ s \,\middle|\, \frac{1}{N}\sum_{i=1}^{N} F_n(A_i \mid u_i \mid s) \geq 1-\gamma \right\} \geq 1-\beta$$

PROOF:

Assume $q\left\{ s \,\middle|\, \frac{1}{N}\sum_{i=1}^{N} F_n(A_i \mid u_i \mid s) \geq 1-\gamma \right\} < 1-\beta$, then we have, if we write $f(s)$ instead of $\frac{1}{N}\sum_{i=1}^{N} F_n(A_i \mid u_i \mid s)$:

$$\int_S f(s)dq(s) = \int_{f<1-\gamma} f(s)dq(s) + \int_{f\geq 1-\gamma} f(s)\,dq(s) \leq$$

$$\leq (1-\gamma)q\{s|f<1-\gamma\}+q\{s|f\geq 1-\gamma\}, \quad \text{since} \quad f(s)\leq 1.$$

$$= (1-\gamma)(1-q\{s|f\geq 1-\gamma\}) + q\{s|f\geq 1-\gamma\} =$$

$$= 1-\gamma+\gamma q\{s|f\geq 1-\gamma\} \leq$$

$$\leq 1-\gamma+\gamma(1-\beta) = 1-\gamma\beta < 1-\lambda,$$

in contradiction to the proposition.
This proves the Lemma.

For a p. d. p on X let $R(p, F(\cdot|\cdot|s))$ be the rate of the channel $F(\cdot|\cdot|s)$.

THEOREM 1

Let X' be finite.

a) $$_+C_1(\lambda) := \inf_{\gamma\beta>\lambda} \frac{1}{1-\gamma} \sup_p \sup_{\{S': q(S')\geq 1-\beta\}} \inf_{s\in S'} R(p, F(\cdot|\cdot|s)) \geq$$

$$\geq \limsup_{n\to\infty} \frac{1}{n} \log N_1(n,\lambda)$$

b) $$_-C_1(\lambda) := \sup_{0<\varepsilon<1} \sup_p \sup_{S': q(S')>1-\varepsilon\lambda} \inf_{s\in S'} R(p, F(\cdot|\cdot|s)) \leq$$

$$\leq \liminf_{n\to\infty} \frac{1}{n} \log N_1(n,\lambda)$$

PROOF:

Choose ε, S' with $q(S') > 1 - \varepsilon\lambda$ such that

$$\left| \sup_{p} \inf_{s \in S'} R(p, F(\cdot|\cdot|s)) - {}_{-}C_1(\lambda) \right| \leq \frac{\delta}{2},$$

then η such that $(1-\varepsilon\lambda)(1-\eta) = 1-\lambda$.

A η-code for the simultaneous channel $S_n = \{ F_n(\cdot|\cdot|s) \mid s \in S' \}$ is a λ-code for P_n.

b) follows now from the coding theorem for S_n.

A λ-code for P_n is an averaged λ-code for P_n:

$$\frac{1}{N} \sum_{i=1}^{N} P_n(A_i | u_i) \geq 1 - \lambda .$$

From Lemma 1 follows that for every pair (γ, β) with $0 < \gamma, \beta < 1$, $\gamma\beta > \lambda$ there exists a subset S'' of S with $q(S'') \geq 1-\beta$ such that

$$\text{(3)} \qquad \frac{1}{N} \sum_{i=1}^{N} F_n(A_i | u_i | s) \geq 1 - \gamma \qquad (s \in S'')$$

Applying Fano's Lemma for averaged errors we get

$$\frac{1}{n} \log N_1(n,\lambda) \leq \frac{1}{1-\gamma} \sup_{p} \inf_{s \in S''} R(p, F(\cdot|\cdot|s))$$

and furthermore

$$\frac{1}{n} \log N_1(n,\lambda) \leq {}_{+}C_1(\lambda),$$

this proves a)

It is an easy consequence of the definition for ${}_{+}C_1(\lambda)$ and ${}_{-}C_1(\lambda)$ that ${}_{+}C_1(\lambda)$, ${}_{-}C_1(\lambda)$ are monoton increasing in λ, that $\lim_{\lambda \to 0} {}_{+}C_1(\lambda)$, $\lim_{\lambda \to 0} {}_{-}C_1(\lambda)$ exist and that $\lim_{\lambda \to 0} {}_{+}C_1(\lambda) = \lim_{\lambda \to 0} {}_{-}C_1(\lambda)$. Let us denote this limit by C_1. Then we have as a consequence of Theorem 1

COROLLARY 1. C_1 is the weak capacity for $\mathcal{C}_1$. In [2] we gave a different proof. (Theorem 5, chapt. 7.)

REMARK 1. For an individual channel it is unessential whether we work with a λ-code or with an averaged λ-code, (cf. [11], ch. 3.1 Lemma 3.11) however, it makes a difference for simultaneous channels. If we use averaged λ-codes for simultaneous discrete memoryless channels, then the strong converse of the coding theorem does not hold.

EXAMPLE 1.

$$X = X' = \{1,2,\dots 5\}, \quad S = \{1,2\}$$

$$(w(j|i|1))_{i,j=1,\dots,5} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

$$(w(j|i|2))_{i,j=1,\dots,5} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The capacity of the simultaneous channel is given by

$$\max_p \min_{s=1,2} R(p, w(\cdot|\cdot|s)) = \max_p \min_{s=1,2} \sum_{i,j=1}^{5} p_i w(j|i|s) \log \frac{w(j|i|s)}{\sum_{k=1}^{5} p_k w(j|k|s)}$$

$$R(p, w(\cdot|\cdot|1)) = (p_3 + p_4 + p_5) \log \frac{1}{p_3 + p_4 + p_5} + \\ + p_1 \log \frac{1}{p_1} + p_2 \log \frac{1}{p_2}$$

$$R(p, w(\cdot|\cdot|2)) = (p_1 + p_2 + p_3) \log \frac{1}{p_1 + p_2 + p_3} + \\ + p_4 \log \frac{1}{p_4} + p_5 \log \frac{1}{p_5}$$

(4) $$\max_p R(p,w(\cdot|\cdot|1)) = \log 3$$

The maximum is attained for $p=(\frac{1}{3},\frac{1}{3},p_3,p_4,p_5)$ and no other p.d.'s.

(5) $$\max_p R(p,w(\cdot|\cdot|2)) = \log 3$$, the maximum is attained

for $p=(p_1,p_2,p_3,\frac{1}{3},\frac{1}{3})$ and no other p.d.'s.

From (4), (5), it follows that

(6) $$\max_p \min_{s=1,2} R(p,w(\cdot|\cdot|s)) < \log 3$$

Consider the sets

$$V_n = \{x_n \mid x_n = (x^1,\dots,x^n) \in X_n,\ x^t \in \{3,4,5\}\}$$

$$W_n = \{x_n \mid x_n = (x^1,\dots,x^n) \in X_n,\ x^t \in \{1,2,3\}\}$$

Define the code $\{(\bar{u}_i, A_i) \mid \bar{u}_i \in V_n \cup W_n,\ A_i = \{x'_n \mid x'^t = t\text{-th}$ component of $\bar{u}_i\}\}$

The length of this code is

$$N = 2\cdot 3^n - 1 > 3^n$$

For $\lambda \geq 1/2$ the code is an averaged simultaneous λ-code:

$$\frac{1}{N}\sum_{i=1}^{N} F_n(A_i|\bar{u}_i|s) \geq \frac{1}{2} \geq 1-\lambda \quad (s=1,2)$$

If we denote the maximal length of an averaged simultaneous λ-code in $\langle 1,n \rangle$ by $N_a(n,\lambda)$ then we have

$$N_a(n,\lambda) > 3^n = e^{\log 3\cdot n}$$

However, it follows from Fano's Lemma that

$$\max_{p} \min_{s=1,2} R(p, w(\cdot|\cdot|s)) < \log 3$$

is the weak capacity for our simultaneous channel with averaged error. The strong converse of the coding theorem does not hold. In special cases we can give a sharper estimate than that given by Theorem 1.

EXAMPLE 2. Given $X = X' = \{1, \dots a\}$ and the stochastic matrices $w(\cdot|\cdot|1), w(\cdot|\cdot|2)$ with a rows and a columns. For $s = 1, 2$ we define the discrete memoryless channel

$$(P_n(\cdot|\cdot|s))_{n=1,2,\dots} \quad \text{by} \quad P_n(x'_n | x_n | s) = \prod_{t=1}^{n} w(x'^t | x^t | s) \quad \text{for}$$

all $x_n \in X_n$, $x'_n \in X'_n$, $n = 1, 2, \dots$ and the averaged channel $\mathcal{C}_1$ by

$$P_n(x'_n | x_n) = \frac{1}{2} P_n(x'_n | x_n | 1) + \frac{1}{2} P_n(x'_n | x_n | 2) \quad (x_n \in X_n,\ x'_n \in X'_n,\ n = 1, 2, \dots).$$

We get

(7)
$$C_1(\lambda) = \max_{s=1,2} \max_{p} R(p, w(\cdot|\cdot|s)) = \bar{\bar{C}} \quad \text{for} \quad \lambda > \frac{1}{2}$$

and

(8)
$$C_1(\lambda) = \max_{p} \inf_{s=1,2} R(p, w(\cdot|\cdot|s)) = \bar{C} \quad \text{for} \quad \lambda < \frac{1}{2}$$

and

(9)
$$\bar{\bar{C}} \geq \limsup_{n} \frac{1}{n} \log N(n, \tfrac{1}{2}) \geq \liminf_{n \to \infty} \frac{1}{n} \log N(n, \tfrac{1}{2}) \geq \bar{C}$$

for $\lambda = \frac{1}{2}$.

[$C_1(\lambda) \geq \max_{s=1,2} \max_{p} R(p, w(\cdot|\cdot|s)) = \bar{\bar{C}}$ for $\lambda > \frac{1}{2}$ follows from the coding theorem for an individual channel. It remains to show that $C_1(\lambda) \leq \bar{\bar{C}}$:

a λ-code $\{(u_i, A_i) / i = 1, \dots N\}$ for

$\mathcal{C}_1$ has the property that either

(10) $P_n(A_i|u_i|1) \geq 1-\lambda$ or $P_n(A_i|u_i|2) \geq 1-\lambda$ $(i = 1, \dots N)$.

Therefore:

$$N(n,\lambda) \leq 2e^{\bar{\bar{C}}n + k(\lambda)\sqrt{n}} \quad \text{and} \quad C_1(\lambda) \leq \bar{\bar{C}}.$$

(8) is trivial. (9) is a consequence of (7), (8).]

It is possible that $C_1(\lambda)$ exists for $\lambda = \frac{1}{2}$ and is unequal to $\bar{\bar{C}}$ and to $\bar{C}$. Choose for example $w(\cdot|\cdot|1)$ such that $0 < C_{zero\ error} < \max_p R(p, w(\cdot|\cdot|1))$ ([12]) and

$$w(1|i|2) = 1 \quad \text{for} \quad i = 1, \dots a$$

$$w(j|i|2) = 0 \quad \text{for} \quad j \neq 1,\ i = 1, \dots a.$$

Then we get

$$\bar{\bar{C}} > C_{zero\ error} = \lim_{n\to\infty} \frac{1}{n} \log N(n, \tfrac{1}{2}) > \bar{C} = 0.$$

In general we have

$$\bar{\bar{C}} \geq \limsup_{n\to\infty} \log N(n, \tfrac{1}{2}) \geq \liminf_{n\to\infty} \log N(n, \tfrac{1}{2}) > \max(\bar{C}, \bar{\bar{C}}_{zero\ error}),$$

where $\bar{\bar{C}}_{zero\ error}$ is the maximum of the zero error capacities of $w(\cdot|\cdot|1)$, $w(\cdot|\cdot|2)$.

A formula for $C(\frac{1}{2})$ would imply a formula for $C_{zero\ error}$, which is unknown [12]. But even the existence of $\lim_{n\to\infty} \frac{1}{n} \log N(n, \frac{1}{2})$ is not obvious. This seems to be a difficult problem. It is easy to construct channels for which $C(\lambda)$ has countable many jumbs but for which the weak capacity $C = \lim_{\lambda\to 0} C(\lambda)$ exists, as was shown in [9]. Probably there exist even channels for which $C(\lambda)$ does not exist for all λ but for which C still exists.

REMARK 2. In case S is finite we can give a sharper estimate than a) in

THEOREM 1:

$$\inf_{0<\eta<1} \sup_{p} \sup_{S' : q(S') \geq 1 - \frac{\lambda}{\eta}} \inf_{s \in S'} R(p, F(\cdot|\cdot|s)) \geq$$

$$\geq \limsup_{n\to\infty} \frac{1}{n} \log N_1(n,\lambda)$$

This can be proved by extending the argument used under (10) for $S = \{1,2\}$ to the general finite case.

3. A REMARK ON THE PAPER OF WOLFOWITZ: "MEMORY INCREASE CAPACITY"

Given a stochastic matrix $w(i|j)$, $i = 1, \dots a$, $j = 1, \dots a$

We define the channel O without memory:

$$P_n(y_n|x_n) = \prod_{t=1}^{n} w(y^t|x^t), \quad x_n \in X_n, \; y_n \in X'_n, \quad n = 1,2,\dots$$

Let $(P_n(\cdot|\cdot|M))$ $n = 1,2,\dots$ be any channel with the property

$$p^t(y^t|x^t|M) = w(y^t|x^t), \quad x^t \in X^t, \; y^t \in X'^t, \quad t = 1,2,\dots$$

Thus, the two channels are directly comparable. Wolfowitz [10] proved: suppose that in channel M the power of the memory ([11], ch. 6. 7) between blocks of letters separated by d letters approaches zero as $d \to \infty$, uniformly in the blocks, then the capacity of channel M is not less than that of channel O.

It follows from Dobrushin's inequality [4] that

$$C_M = \limsup_{n\to\infty} \frac{1}{n} \sup_{p_n} R(p_n, P_n(\cdot|\cdot|M)) \geq C_O .$$

Wolfowitz's result holds, iff C_M is capacity. But C need not be the capacity of channel M. [cf. [2], ch. 3, remark 3.]

Averaged channels are channels with memory. They give us examples of channels, where memory decreases capacity.

EXAMPLE 3. Let $w(\cdot|\cdot|1)$, $w(\cdot|\cdot|2)$ be stochastic matrices with

$$\max_p R(p, w(\cdot|\cdot|2)) = 0$$

$$\max_p R(p, w(\cdot|\cdot|1)) > 0 .$$

Let $w(\cdot|\cdot) = q_1 w(\cdot|\cdot|1) + q_2 w(\cdot|\cdot|2)$.

For q_1 sufficiently near to 1 :

$$\max_p R(p, w) > 0 ,$$

but the weak capacity of

$$(P_n(\cdot|\cdot|M))_{n=1,\dots} = (\sum_{s=1}^{2} q_s P_n(\cdot|\cdot|s))_{n=1,2,\dots}$$

is $\max_p \inf_{s=1,2} R(p, w(\cdot|\cdot|s)) = 0$.

In general we have:

$$C_M \geq \max_p \sum_{s=1,2} q_s R(p, w(\cdot|\cdot|s)) \qquad \text{[cf. [2], ch. 3. remark 3]}$$

$$C_0 = \max_p R(p, q_1 w(\cdot|\cdot|1) + q_2 w(\cdot|\cdot|2))$$

the weak capacity of $M = \max_p \inf_{s=1,2} R(p, w(\cdot|\cdot|s))$.

If $\max_p \inf_{s=1,2} R(p, w(\cdot|\cdot|s)) < C_0$, then the memory decreases capacity.

If $\max_p \inf_{s=1,2} R(p, w(\cdot|\cdot|s)) \geq C_0$ then memory increases capacity.

EXAMPLE 4.

$$w(\cdot|\cdot|1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$w(\cdot|\cdot|2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$q_1 = q_2 = \frac{1}{2}$$

$$w(\cdot|\cdot) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\max_{p} \inf_{s=1,2} R(p, w(\cdot|\cdot|s)) = \log 2$$

$$C_0 = 0$$

Even if the strong converse holds for channel M, the capacity need not be greater than C_0. (For the definition of the general discrete channel see [11], ch. 5)

EXAMPLE 5. $X = X' = \{1,2\}$

$$P_n(1\,1\,1\,.\,.\,1 \mid 1\,1\,1\,.\,.\,1) = 1$$

$$P_n(0\,0\,0\,.\,.\,0 \mid 0\,0\,0\,.\,.\,0) = 1$$

$$P_n(y_n|x_n) = \frac{1}{2^n} \text{ iff } x_n \neq (1,1,1\ldots 1)$$
$$\neq (0,0,0\ldots 0)$$

$(n = 1,2,\ldots)$, $x_n \in X_n$, $y_n \in X'_n$.

$$w(\cdot|\cdot) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The strong capacity of $(P_n(\cdot|\cdot))_{n=1,2,\ldots}$ is 0 and $C_0 = \log 2$.

4. DISCRETE AVERAGED CHANNELS WHERE EITHER THE SENDER OR THE RECEIVER KNOWS THE c. p. f.

Assume $X'^t = \{1, \dots a\}$ $t = 1, 2, \dots;$ $S = \{1, 2, \dots\}$ $q = (q_1, \dots)$, $q_i > 0$

p. d. on S.

The discrete averaged channel is defined by

$$(P_n(A \mid x_n))\ n = 1, 2, \dots = \left(\sum_{s=1}^{\infty} q_s F_n(A \mid x_n \mid s)\right)\ n = 1, 2, \dots$$

$$A \subset X'_n,\ x_n \in X_n$$

LEMMA 2: For $q = (q_1, \dots)$ define

$$\eta_k = \inf_{\varkappa = 1, \dots k} q_\varkappa > 0$$

If $\{(u_1(s), A'_1(s)), \dots (u_N(s), A'_N(s))\}$ is a set of pairs,

where $u_i(s) \in X_n$, $A'_i(s) \in \mathcal{L}'_n$ for $i = 1, \dots N$, $s \in S$,

$A'_i(s) \cap A'_j(s) = \phi$ for $i \neq j$, $s \in S$

and furthermore

(11)
$$\sum_{s \in S} q_s P_n(A'_i(s) \mid u_i(s) \mid s) \geq 1 - \frac{\eta_k}{2}$$

then

(12) $P_n(A'_i(s) \mid u_i(s) \mid s) \geq \frac{1}{2} \eta_k$ for $s = 1, \dots k;$ $i = 1, \dots N$.

PROOF: Define $\varepsilon_k = \sum_{\varkappa = k+1}^{\infty} q_\varkappa$

From (11) we conclude

$$\sum_{s=1}^{k} q_s P_n(A'_i(s) \mid u_i(s) \mid s) \geq 1 - \varepsilon_k - \frac{1}{2} \eta_k,$$ since

(13) $$\sum_{s=1}^{k} q_s P_n(A'_i(s) \mid u_i(s) \mid s) \geq \sum_{s=1}^{\infty} q_s P_n(A'_i(s) \mid u_i(s) \mid s) - \varepsilon_k \geq$$

$$\geq 1 - \left(\frac{1}{2}\eta_k + \varepsilon_k\right).$$

But

(14) $$\sum_{s=1}^{k} q_s P_n(A'_i(s) \mid u_i(s) \mid s) - q_{s'} P_n(A'_i(s') \mid u_i(s') \mid s') \leq 1 - \varepsilon_k - \eta_k$$

for $s' = 1, \dots k$.

From (13), (14) we have

$$q_s P_n(A'_i(s) \mid u_i(s) \mid s) \geq \frac{1}{2}\eta_k \quad \text{for} \quad s = 1, \dots k;\ i = 1, \dots N$$

and therefore

$$P_n(A'_i(s) \mid u_i(s) \mid s) > \frac{1}{2}\eta_k \qquad \text{for} \qquad s = 1, \dots k;\ i = 1, \dots N$$

(cf. [2] proof of th. 2)

REMARK 3. The proof goes through verbatim for the semi-continuous case. Averages with respect to general p. d. can be treated in the same way as in [2], § 7. However, in § 5 we give a different proof, which covers all these cases.

THEOREM 2. (Cod. th. and weak converse for $\mathcal{C}_3$)

Let $C_3 = \max_{p} \inf_{s \in S} R(p, w(\cdot|\cdot|s))$.

Then the following estimates hold:

a) Given $0 < \lambda < 1, \delta > 0$, then there exists an $n_0 = n_0(\lambda, \delta)$ such that $N_3(n, \lambda) > e^{(C_3 - \delta)n}$ for $n \geq n_0$.

b) Given $\delta > 0$, then there exists a λ and an $n_0 = n_0(\lambda, \delta)$ such that $N_3(n,\lambda) < e^{(C_3+\delta)n}$ for $n \geq n_0$.

PROOF: a) A λ-code for the simultaneous channel $\{P_n(\cdot|\cdot|s) | s \in S\}$ is a λ-code for $\mathcal{C}_3$.

b) For the given $\delta > 0$ choose k such that $|C_k - C| \leq \frac{\delta}{2}$, then choose $\lambda = \frac{\eta_k}{2}$.

It follows from Lemma 2, that for a λ-code

$$\{(u_i, A_i'(s)) | i = 1, \dots N; s \in S\} \text{ for } \mathcal{C}_3 .$$

$$P_n(A_i'(s) | u_i | s) \geq \frac{1}{2}\eta_k \qquad s = 1, \dots k; \ i = 1, \dots N$$

holds.
Statement b) follows now by usual arguments. ([2], Theorem 4.5.2.)

THEOREM 3. (Coding theorem and weak converse for $\mathcal{C}_2$ and $\mathcal{C}_4$) Let $C_2 = C_4 = \inf_{s \in S} \max_p R(p, w(\cdot|\cdot|s))$.

Then the following estimates hold:

a) Given $0 < \lambda < 1$, $\delta > 0$, there exists an $n_0 = n_0(\lambda, \delta)$ such that $N_j(n,\lambda) > e^{(C_j - \delta)n}$ for $n \geq n_0$ $(j = 2,4)$.

b) Given $\delta > 0$, there exists a λ and an $n_0 = n_0(\lambda, \delta)$ such that $N_j(n,\lambda) < e^{(C_j + \delta)n}$ for $n \geq n_0$ $(j = 2,4)$.

PROOF: a) follows from Theorem 4.5.3 in [11].
b) By the same arguments as in the proof for

Theorem b) we get

$$P_n(A_i'(s) | u_i(s) | s) \geq \frac{1}{2}\eta_k \qquad s = 1, \dots k; \ i = 1, \dots N .$$

Applying the strong converse of the coding theorem for individual channels we get b).

REMARK 4. The strong converse of the coding theorem does not hold for $\mathcal{C}_2$, $\mathcal{C}_3$ and $\mathcal{C}_4$:

EXAMPLE 6 ([2], ch. 3, remark 1) Choose $w(\cdot|\cdot|s)$, $s=1,2$ such that

$(P_n(\cdot|\cdot|1))_{n=1,2,\dots}$ has capacity 0 and $(P_n(\cdot|\cdot|2))_{n=1,2,\dots}$ has capacity $C(2)>0$. Then a fortiori $C_2 = C_3 = C_4 = 0$

A λ-code for $\frac{1}{2} P_n(\cdot|\cdot|2)$ is a λ-code for P_n. For $\lambda > \frac{1}{2}$ we get

$N_i(n,\lambda) > e^{C(2)n - k(\lambda)\sqrt{n}}$ for $(i = 2,3,4)$.

That $\mathcal{C}_3$ does not have a strong capacity was earlier shown be Wolfowitz ([11], ch. 7. 7)

5. The general case

We return to the case, where the individual channels $(F_n(\cdot|\cdot|s))$ $n=1,2,\dots$ are semicontinuous and q is a general p. d. (as described in § 1).

LEMMA 3

If $$\frac{1}{N}\sum_{i=1}^{N}\int_S P_n(A_i(s)|u_i(s)|s)\,dq(s) \geq 1-\lambda$$

for $i=1,\dots N$ and $1>\gamma,\beta>0$ such that $\gamma\beta>0$

then

$$q\{s|\frac{1}{N}\sum_{i=1}^{N} P_n(A_i(s)|u_i(s)|s) \geq 1-\lambda\} \geq 1-\beta$$

PROOF. Define $f^*(s) = \frac{1}{N}\sum_{i=1}^{N} P_n(A_i(s)|u_i(s)|s)$,

then the proof of Lemma 1 turns over verbatim.

THEOREM 4. (Coding theorem and weak converse for $\mathcal{C}_1$)

Let $$C_1 = \inf_{\alpha>0}\ \sup_{\{S'\subset S:\,q(S')\geq 1-\alpha\}}\ \lim_{n\to\infty}\frac{1}{n}\sup_{\mathcal{D}_n\in\mathcal{Z}_n}\sup_{p_n}\inf_{s\in S'} R(p_n|\mathcal{D}_n|F_n(\cdot|\cdot|s)),$$

where $\mathcal{D}_n = \langle D_1, \dots, D_b \rangle$ is a partition of X'_n in finitely many elements of $\mathcal{L}'_n$ and, $\mathcal{Z}_n$ is the set of all such finite partitions, and $R(p_n | \mathcal{D}_n | F_n(\cdot|\cdot|s)) =$

$$\sum_{i=1,\dots b} \sum_{x_n \in X_n} p_n(x_n) F_n(D_i | x_n | s) \log \frac{F_n(D_i | x_n | s)}{\sum_{y_n \in X_n} p_n(y_n) F_n(D_i | y_n | s)}$$

then the following estimates hold:

a) Given $0 < \lambda < 1$, $\delta > 0$, than there exists an $n_0 = n_0(\lambda, \delta)$, such that $N_1(n, \lambda) > e^{(C_1 - \delta)n}$ for $n \geq n_0$.

b) Given $\delta > 0$, then there exists a λ and an $n_0 = n_0(\lambda, \delta)$, such that $N_1(n, \lambda) < e^{(C_1 - \delta)n}$ for $n \geq n_0$.

PROOF. a) Define

$$C_1(\alpha) = \sup_{\{S' \subset S : q(S') \geq 1-\alpha\}} \lim_{n \to \infty} \frac{1}{n} \sup_{\mathcal{D}_n} \sup_{p_n} \inf_{s \in S'} R(p_n | \mathcal{D}_n | F_n(\cdot|\cdot|s)) .$$

Given $\lambda, \delta > 0$, choose $\alpha < \lambda$ and S' such that $q(S') \geq 1 - \alpha$ and

$$\left| \lim_{n \to \infty} \frac{1}{n} \sup_{\mathcal{D}_n} \sup_{p_n} \inf_{s \in S'} R(p | \mathcal{D}_n | F_n(\cdot|\cdot|s)) - C_1(\alpha) \right| \leq \frac{\delta}{2}$$

Define $\lambda' = \frac{\lambda - \alpha}{1 - \alpha}$. A λ'-code for the semicontinuous simultaneous channel $\{F_n(\cdot|\cdot|s) \mid s \in S'\}$ is a λ-code for $\mathcal{C}_1$, because $(1-\lambda')(1-\alpha) =$ $= 1 - (1-\alpha)\lambda' - \alpha = 1 - \lambda$.

The coding theorem for semicontinuous simultaneous channels ([6], th. 1) gives us

$$N_1(n, \lambda) \geq e^{(C_1(\alpha) - \frac{\delta}{2})n - k(\lambda')\sqrt{n}} \geq e^{(C_1 - \delta)n}$$

for n sufficiently large.

b) Choose α such that

(15) $|C_1(\alpha) - C_1| \leq \frac{\delta}{3}$, then β such that $\log a \left(\frac{1}{1-\beta} - 1\right) \leq \frac{\delta}{3}$

and finally λ such that $\alpha\beta > \lambda$.

By Lemma 3 there exists a set S with $q(S) \geq 1-\alpha$ and

$\frac{1}{N}\sum_{i=1}^{N} F_n(A_i | u_i | s) \geq 1-\beta$ for $s \in S$. From Fano's Lemmas we obtain

$$N_1(n,\lambda) \leq e^{\frac{C_1(\alpha)}{1-\beta} n} .$$

If we use (15) we get statement b).

THEOREM 5 (Coding theorem and weak converse for $\mathscr{C}_3$)

Let $C_3 = \inf_{\alpha>0} \sup_{p} \sup_{S' \subset S : q(S') \geq 1-\alpha} \inf_{s \in S'} R(p, F_n(\cdot|\cdot|s))$

then the following estimates hold:

a) Given $0<\lambda<1, \delta>0$, then there exists a $n_0 = n_0(\lambda,\delta)$, such that

$$N_3(n,\lambda) > e^{(C_3-\delta)n} \quad \text{for} \quad n \geq n_0$$

b) Given $\delta > 0$, then there exists a λ and an $n_0 = n_0(\lambda,\delta)$, such that

$$N_3(n,\lambda) < e^{(C_3+\delta)n} \quad \text{for} \quad n \geq n_0$$

PROOF: a) Define

$$C_3(\alpha) = \sup_{p} \sup_{\{S' | S' \subset S, q(S') \geq 1-\alpha\}} \inf_{s \in S'} R(p, F(\cdot|\cdot|s))$$

Given $\lambda, \delta > 0$, choose $\alpha < \lambda$ and S' such that $q(S') \geq 1-\alpha$ and

$$\left| \sup_{p} \inf_{s \in S'} R(p, F(\cdot|\cdot|s)) - C_3(\alpha) \right| \leq \frac{\delta}{2}$$

Define $\lambda' = \frac{\lambda-\alpha}{1-\alpha}$. A λ'-code for the compound channel (with receiver knowledge)

$S'_n(R) = \{F_n(\cdot|\cdot|s) \mid s \in S'\}$ is a λ-code for $\mathscr{C}_3$, because

$(1-\lambda')(1-\alpha) = 1-\lambda$.

The coding theorem for $S_n'(R)$ ([6], th. 4) gives us

$$N_3(n,\lambda) \geq N_R(n,\lambda') \geq e^{(C_3(\alpha)-\frac{\delta}{2})n-k(\lambda')\sqrt{n}} \geq$$

$$\geq e^{(C_3-\delta)n} \quad \text{for } n \text{ sufficiently large.}$$

b) Choose α such that

(16) $\quad |C_3(\alpha)-C_3| \leq \frac{\delta}{3}$, then β such that $\log a\left(\frac{1}{1-\beta}-1\right) \leq \frac{\delta}{3}$

and finally λ such that $\alpha\beta > \lambda$

By Lemma 3 there exists a set S with $q(S) \geq 1-\alpha$ and

$$\frac{1}{N}\sum_{i=1}^{N} F_n(A_i(s)|u_i|s) \geq 1-\beta \quad \text{for} \quad s \in S .$$

From Fano's Lemma we get

$$N_3(n,\lambda) \leq e^{\frac{C_3(\alpha)}{1-\beta}n}$$

If we use (16) we get statement b)

THEOREM 6 (Coding theorem and weak converse for $\mathcal{C}_2$ and $\mathcal{C}_4$.)

Let

$$C_2 = C_4 = \inf_{\alpha>0} \sup_{S'\subset S: q(S')\geq 1-\alpha} \lim_{n\to\infty} \frac{1}{n} \sup_{\mathcal{D}_n} \inf_{s\in S'} \sup_{p_n} R(p_n,\mathcal{D}_n,F_n(\cdot|\cdot|s))$$

(the existence of the limit was shown in [6]), then the following estimates hold:

a) Given $0<\lambda<1, \delta>0$, then there exists an $n_0 = n_0(\lambda,\delta)$, such that

$$N_j(n,\lambda) > e^{(C_j-\delta)n} \quad \text{for} \quad n \geq n_0 \quad (j=2,4)$$

b) Given $\delta>0$, then there exists a λ and an $n_0 = n_0(\lambda,\delta)$, such that

$$N_j(n,\lambda) < e^{(C_j+\delta)n} \qquad \text{for} \qquad n \ge n_0 \qquad (j = 2,4)$$

PROOF: a) It follows from the definition of $\mathcal{C}_2$ and $\mathcal{C}_4$ that

(17) $N_2(n,\lambda) \ge N_4(n,\lambda)$ therefore it is enough to prove a) for $\mathcal{C}_2$.

Define

$$C_2(\alpha) = \sup_{\{S' \mid S' \subset S : q(S') \ge 1-\alpha\}} \lim_{n\to\infty} \frac{1}{n} \sup_{\mathcal{D}_n} \inf_{s \in S'} \sup_{p_n} R(p_n \mid \mathcal{D}_n \mid F_n(\cdot\mid\cdot\mid s))$$

Given $\lambda, \delta > 0$, choose $\alpha < \lambda$ and S' such that $q(S') \ge 1-\alpha$ and

$$\left| \lim_{n\to\infty} \frac{1}{n} \sup_{\mathcal{D}_n} \inf_{s\in S'} \sup_{p_n} R(p_n \mid \mathcal{D}_n \mid F_n(\cdot\mid\cdot\mid s)) - C_2(\alpha) \right| \le \frac{\delta}{2} .$$

Define $\lambda' = \frac{\lambda-\alpha}{1-\alpha}$. A λ'-code for the simultaneous channel with sender knowledge $S_n'(S) = \{F_n(\cdot\mid\cdot\mid s) \mid s \in S'\}$ is a λ-code for $\mathcal{C}_2$.

The coding theorem for $S_n'(S)$ ([6], th. 3) gives us

$N_2(n,\lambda) \ge N_s(n,\lambda') \ge e^{(C_2(\alpha)-\frac{\delta}{2})n}$ for n sufficiently large and therefore

$N_2(n,\lambda) \ge e^{(C_2-\delta)n}$ for n sufficiently large.

b) Choose α such that

(18) $\quad |C_4(\alpha) - C_4| \le \frac{\delta}{3}$ then β such that

$\log a\left(\frac{1}{1-\beta} - 1\right) \le \frac{\delta}{3}$ and finally λ such that $\alpha\beta > \lambda$.

By Lemma 3 there exists a set S with $q(S) \ge 1-\alpha$ and

$$\frac{1}{N} \sum_{i=1}^{N} F_n(A_i(s) \mid u_i(s) \mid s) \ge 1-\beta \qquad \text{for} \qquad s \in S .$$

From Fano's Lemma and (17) we get

$$N_2(n,\lambda) \le N_4(n,\lambda) \le e^{\frac{C_4(\alpha)}{1-\beta} n}$$

. If we use (18) we get statement b).

II. A NEW CHANNEL

Let S be an arbitrary (index)-set and to each $s \in S$ assigned a semicontinuous nonstationary channel without memory, $(F_n(\cdot|\cdot|s))$ $n = 1, 2, \dots$.

In the theory of simultaneous channels [8], [1] one uses the following definition of a λ-code: a λ-code (N, n, λ) is a set of pairs $\{(u_i, A_i | i = 1, \dots N\}$ with $u_i \in X_n, A_i \in \mathcal{L}_n'$ for $i = 1, \dots N$. $A_i \cap A_j = \phi$ for $i \neq j$ and with

(19) $$\inf_{s \in S} F_n(A_i | u_i | s) \ge 1 - \lambda$$

This describes the situation, where neither the sender nor the receiver knows the individual channel which governs the transmission of a code word u_i.

If the sender can choose for the transmission of a code word u_i the channel over which he wants to transmit, then we have in the code definition (19) to exchange by

(20) $$\sup_{s \in S} P_n(A_i | u_i | s) \ge 1 - \lambda$$

We denote the described compound channel by $\mathcal{W}$.

1. THE STATIONARY DISCRETE CASE

Assume that the channels $(F_n(\cdot|\cdot|s))$ $n = 1, 2, \dots$ are discrete and stationary.

THEOREM 7 (Coding theorem and strong converse for $\mathcal{W}$.)

Let $C = \sup_{s \in S} \max_{p} R(p, w(\cdot|\cdot|s))$

Given $\delta > 0,\ 0 < \lambda < 1$, then there exists an n_0 such that for $n \geq n_0$

a) $N(n,\lambda) > e^{(C-\delta)n}$

b) $N(n,\lambda) < e^{(C+\delta)n}$

PROOF: Part a) follows from the coding theorem for an individual channel.

For the proof of part b) we need the

LEMMA. [[11], 4.2.2. p. 35], [1]

Let b be greater than 0. There exists a nullsequence of positive real numbers

$\{a_n\}_{n=1,2,\dots}$ with the property:

let $n \in N,\ A \subset X'_n,\ x_n \in S_n,\ s, s^* \in S$ such that

$F_n(A|x_n|s) > b$ and

$$|w(i|j|s^*) - w(i|j|s)| \leq \frac{a}{2^{\sqrt{n}}} \qquad (1 \leq i, j \leq a)$$

then
$$\left| \frac{F_n(A|x_n|s^*)}{F_n(A|x_n|s)} - 1 \right| < a_n .$$

S can be written as a finite union of disjoint sets

$S_r(r=1,\dots(\frac{2^{\sqrt{n}}}{a})^{a^2} = R)$, such that

$$|w(j|i|s) - w(j|i|s^*)| \leq a \cdot 2^{-\sqrt{n}} \quad (i,j=1,\dots a) \text{ for } s, s' \in S_r .$$

Let now $\{(u_i, A_i) | i=1,\dots N\}$ be a code (N,n,λ) of maximal length.

To every u_i corresponds an individual channel $P_n(\cdot|\cdot|s_i)$ and therefore a matrix $w(\cdot|\cdot|s_i)$. Let $\{(u_{i_r}, A_{i_r}) | i_r = 1, \dots N_r\}$

be the subcode which corresponds to S_r . It follows from the Lemma that for a $\lambda' > \lambda$ there exists an n_0 such that for $n \geq n_0$

$$P_n(A_{i_r} | u_{i_r} | s) \geq 1 - \lambda' \qquad \text{for all} \qquad s \in S_r \quad (r = 1, \dots R)$$

The strong converse of the coding theorem for an individual channel gives us

$$N(n,\lambda) = \sum_{r=1}^{R} N_r(n,\lambda) \leq R e^{Cn + k(\lambda')\sqrt{n}} \leq e^{(C+\delta)n}$$

for n sufficiently large.

REMARKS: In the semicontinuous case we have in general no compactness property which leads to the Lemma. The coding problem for channel $\mathcal{W}$ is than equivalent to the coding problem of an individual channel with the input alphabet

$X = \bigcup_{s \in S} X_s \cdot X_s = \{(1,s) \dots (a,s)\}$ for $s \in S$, where the code words are restricted to have as components elements from one set X_s .

2. THE c. p. f. VARIES FROM LETTER TO LETTER

Given a set of kernels $\{F(\cdot|\cdot|s) | s \in S\}$. For every n -tupel $s_n = (s^1, \dots s^n)$, $s^i \in S$ we can define the product kernel

$$F_n(\cdot|\cdot|s_n) = \prod_{t=1}^{n} F(\cdot|\cdot|s^t) .$$

Consider now the class

$$\mathcal{C}_n = \{F_n(\cdot|\cdot|s_n) \mid s_n = (s^1, \dots s^n),\ s^i \in S\}$$

A simultaneous λ-code for $\mathcal{C}_n$ is a set of pairs $\{(u_i, A_i) | i = 1, \dots N\}$,

$u_i \in X_n$, $A_i \in \mathcal{L}'_n$, $A_i \cap A_j = \phi$ for $i \neq j$ with

$$F_n(A_i | u_i | s_n) \geq 1 - \lambda \qquad \text{for all} \qquad s_n = (s^1, \dots s^n);\ i = 1, \dots N .$$

It is an unsolved problem to estimate the maximal length of such a λ-code [7].

If we use the code definition (20) then the sequence $(\mathcal{C}_n)$ $n = 1, 2, \dots$ defines a compound channel $\overline{\mathcal{W}}$.

$$\sup_{s_n} F_n(A_i | u_i | s_n) > 1 - \lambda \qquad \text{for } i = 1, \dots N \text{ is equivalent with:}$$

there exist $s_n^1, \dots s_n^N$ such that

$$\text{(21)} \qquad F_n(A_i | u_i | s_n^i) > 1 - \lambda \qquad \text{for} \qquad i = 1, \dots N.$$

Define a new channel $\mathcal{U}$ in the following way;

input alphabet $X \times S$

output alphabet $(X', \mathcal{L}')$

$$(x_n, s_n) : \{(x^1, s^1) \dots (x^n, s^n)\}$$

$$F_n(A | (x_n, s_n)) = F_n(A | x_n | s_n) \quad \text{for}$$

$$A \in \mathcal{L}'_n, \; (x_n, s_n) \in (X \times S) \times \dots \times (X \times S) \qquad n = 1, 2, \dots .$$

A λ-code for channel $\overline{\mathcal{W}}$ is a λ-code for $\mathcal{U}$ and vice versa. We have reduced the coding problem for $\overline{\mathcal{W}}$ to a known situation: $\mathcal{U}$ is a stationary memoryless channel with general input and output alphabets and can be treated by the methods developed in [3].

REFERENCES

[1] R. AHLSWEDE: Beiträge zur Shannonschen Informationstheorie im Falle nichtstationärer Kanäle. Zeitschrift f. Wahrscheinlichkeitstheorie u. verw. Geb.

[2] R. AHLSWEDE: The Weak Capacity of Averaged Channels Zeitschrift f. Wahrscheinlichkeitstheorie u. verw. Geb.

[3] U. AUGUSTIN: Gedächtnisfreie Kanäle für diskrete Zeit. Zeitschrift f. Wahrscheinlichkeitstheorie u. verw. Geb.

[4] R. L. DOBRUSHIN: Arbeiten zur Informationstheorie IV, VEB Deutscher Verlag der Wissenschaften Berlin 1963

[5] K. JACOBS: Almost periodic channels, Colloquium on Combinatorial Methods in Probability Theory, Aarhus, pp. 118-126 (1962)

[6] H. KESTEN: Some remarks on the capacity of compound channels in the semicontinuous case. Information and Control 4, 169-184. (1961)

[7] J. KIEFER, J. WOLFOWITZ: Channels with Arbitrarily Varying Channel Probability Functions Inf. and Control 5, 169-184 (1962)

[8] J. WOLFOWITZ: Simultaneous Channels, Arch. Rat. Mech. Analysis 4, No. 4, 371-386 (1960)

[9] J. WOLFOWITZ: On Channels without Capacity, Information and Control 6, No. 1, 49-54 (1963)

[10] J. WOLFOWITZ: Memory increases capacity, Colloquium on Information Theory, Kossuth Lajos University, Debrecen 19-24 Sept. 1967.

[11] J. WOLFOWITZ: Coding Theorems of Information Theory, Erg. d. Math. u. ihrer Grenzgebiete (1961), Springerverlag.

[12] C. E. SHANNON: On the zero-error capacity of a noisy channel. Trans. IRE, PGIT, Sept. 1956, 8-19.

[13] C. E. SHANNON: Channels with side information at the transmitter. IBMJ Research and Development 2, No. 4, 289-293 (1958).