Z. Wahrscheinlichkeitstheorie verw. Gebiete 29, 331 – 344 (1974)
© by Springer-Verlag 1974

The Capacities of Certain Special Channels with Arbitrarily Varying Channel Probability Functions

N.S. Kambo and Samar Singh

In [5] Ahlswede and Wolfowitz have obtained the capacities of a.v.ch. with binary output in a number of cases, essentially with the aid of a lemma which relates the capacity of the a.v.ch. to that of a suitable ("underlying") d.m.c. A generalization of this lemma to a special kind of a.v.ch. with output alphabet $b>2$, has been given by Ahlswede (Lemma 1 of [1]) and used in [1] and [2] to prove the existence of the weak capacities of various channels under different conditions. We give a detailed proof of a weakened version of Ahlswede's lemma and show, in passing, that his lemma is incorrect. We then define certain special types of a.v.ch and, on the basis of the detailed analysis given by us earlier, we prove lemmas of a similar type for these a.v.ch. We are thus able to extend certain results given for binary output a.v.ch. in [4] and [5] to these special a.v.ch. for which $b>2$.

1. Preliminaries, Definitions and Introduction

Let $X=\{1,2,\ldots,a\}$ and $Y=\{1,2,\ldots,b\}$ respectively be the input and output alphabets of the channel and let S be any non-empty set. Further, let $X^t=X$, $Y^t=Y$ and $S^t=S$ for all $t=1,2,\ldots$, and let $X_n=\prod_{t=1}^{n} X^t$ and define Y_n and S_n similarly. Any member of X_n will be denoted by $x_n=(x^1,x^2,\ldots,x^n)$ and similarly $y_n=(y^1,\ldots,y^n)\in Y_n$ and $s_n\in S_n$. Then, let

$$\mathscr{C}=\{w(\cdot|\cdot|s)|s\in S\}$$

be any set of $a\times b$ stochastic matrices. To avoid unnecessary complications we assume throughout this paper that S (and hence $\mathscr{C}$) is a finite set. However, all our results can be easily generalized to the cases where S is arbitrary.

(1.1) We define the channel with arbitrarily varying channel probability functions (abbreviated a.v.ch.) determined by $\mathscr{C}$ as the sequence $\{\mathscr{C}_n|n=1,2,\ldots\}$ where for all n

$$\mathscr{C}_n=\{P(\cdot|\cdot|s_n)|s_n\in S_n\},$$

and for all $x_n\in X_n$, $y_n\in Y_n$, and $s_n\in S_n$ we have

$$P(y_n|x_n|s_n)=\prod_{t=1}^{n} w(y^t|x^t|s^t).$$

The channel $\{\mathscr{C}_n\}$ will also be referred to as the a.v.ch. $\mathscr{C}$, the sequence s_n will be called the channel sequence and the abbreviation c.p.f. will be used for "channel probability function".

(1.2) If $\mathscr{C}$ contains only one matrix $w(\cdot|\cdot)$, we call the channel a discrete memoryless channel (d.m.c.) and denote it by the sequence $\{P_n(\cdot|\cdot)|n=1,2,\ldots\}$ where

for each n
$$P_n(y_n|x_n)=\prod_{t=1}^{n} w(y^t|x^t)$$
for all $y_n\in Y_n$, $x_n\in X_n$. We shall also say "the d.m.c. $w(\cdot|\cdot)$" when we mean $\{P_n(\cdot|\cdot)\}$.

The first study of a.v.ch. was made in [7] and then they were also studied in [9, 4], and [5]. In [4] a thorough description of the problem was given and different types of communication situations were defined and discussed. Since, in the problem discussed here, side information available to the jammer and randomizations performed by him (see [4]) are not explicitly of any importance, we shall not mention these. However, we shall adopt the rest of the notation developed in [4] for describing the different problems. Thus, (λ_3, S^-, R^-) shall mean we are interested in the maximal error of "pure" codes (no randomizations in encoding or decoding) when neither the sender, nor the receiver have any side information regarding the channel sequence, and so on as in [4].

(1.3) A pure code, in the case (λ_3, S^-, R^-), is a system
$$\{(u_i, A_i)|i=1,2,\ldots,N\},$$
where $u_i\in X_n$ and $A_i\subseteq Y_n$ for all $i=1,\ldots,N$ and $A_i\cap A_j=\emptyset$ for all $i\neq j$.

(1.4) A (n,N,λ) pure code for the a.v.ch. $\mathscr{C}$ in the case (λ_3, S^-, R^-) is a (n,N) code which satisfies
$$P(A_i|u_i|s_n)\geqq 1-\lambda \qquad \text{for all } i=1,\ldots,N,$$
and for all $s_n\in S_n$.

(1.5) A (n,N) code $\{(u_i,A_i)|i=1,\ldots,N\}$ is a strict maximum likelihood code (s.m.l.c.) for the d.m.c. $w(\cdot|\cdot)$ iff
$$A_i=\{y_n|P_n(y_n|u_i)>\max_{j\neq i} P_n(y_n|u_j)\}$$
for all $i=1,2,\ldots,N$.

(1.6) We say that the d.m.c. $w(\cdot|\cdot)$ underlies the a.v.ch. $\mathscr{C}$ if for every $n=1,2,\ldots$, a (n, N) s.m.l.c. $\{(u_i A_i)\}$ for the d.m.c. $w(\cdot|\cdot)$ satisfies
$$P(A_i|u_i|s_n)\geqq P_n(A_i|u_i) \qquad \text{for all } i=1,\ldots,N$$
and for all $s_n\in S_n$.

(1.7) The d.m.c. $w(\cdot|\cdot)$ is basic to the a.v.ch. $\mathscr{C}$ if $w(\cdot|\cdot)$ underlies $\mathscr{C}$ and there exists $s'\in S$ such that $w(\cdot|\cdot|s')=w(\cdot|\cdot)$.

(1.8) A number, $\mathscr{C}$, is called the (strong) capacity of the a.v.ch. $\mathscr{C}$ in the case (λ_3, S^-, R^-), if for all n sufficiently large and all $\omega>0$, $\lambda\in[0,1)$ there exists a (n,N,λ) code for $\mathscr{C}$ with $N>2^{n(C-\omega)}$ and there does not exist such a code with $N>2^{n(C+\omega)}$.

An easy consequence of these definitions is the following

Theorem 1.1. *If the* d.m.c. $w(\cdot|\cdot)$ *is basic to the* a.v.ch. $\mathscr{C}$, *then the strong capacity of the* a.v.ch. $\mathscr{C}$ *exists and is given by the strong capacity of the* d.m.c. $w(\cdot|\cdot)$.

This theorem gives, in essence, the methods used in [5] to find the strong capacity of a.v.ch. with $b=2$ (binary output) in a number of cases. Crucial to this approach is the Lemma 1 of [5] which determines a suitable underlying d.m.c. for the a.v.ch. in question. Hence we are interested in generalizations of that lemma to cases where $b>2$ (see remarks following (1.4) in [5]). In [1] such a generalization has been given (Lemma 1 of [1]) for a special type of a.v.ch. We need the following definitions.

(1.9) If for any $\omega\in[0,1]$ we have a $a\times a$ square matrix $w(\cdot|\cdot)$ defined by

$$w(j|k)=\begin{cases}1-\omega & \text{if } j=k\\ \omega/(a-1) & \text{if } j\neq k,\end{cases}$$

then the d.m.c. $w(\cdot|\cdot)$ is called a (ω, a)-symmetric d.m.c. (In [1] this is called simply a a-ary symmetric channel.)

(1.10) If for any fixed $\omega\in[0,1]$ and $a=b$

$$\mathscr{C}=\{w(\cdot|\cdot)\,|\,w(k|k)\geqq 1-\omega \text{ for all } k=1,\ldots,a\},$$

then $\mathscr{C}$ is called a (ω, a)-symmetric a.v.ch. (In [1] this is called simply a a-ary symmetric a.v.ch.).

Now we can state Lemma 1 of [1] in our terminology as

Lemma 1.1. *Let $\mathscr{C}$ be a (ω, a)-symmetric* a.v.ch. *and $w(\cdot|\cdot)$ be a $(2\omega, a)$-symmetric* d.m.c. *Then, the* d.m.c. *$w(\cdot|\cdot)$ underlies the the* a.v.ch. *$\mathscr{C}$.*

This lemma has been used in [1] to prove the existence of the weak capacity of any a.v.ch. in the cases (λ_3, S^-, R^-) and (λ_3, S^-, R^+) (Theorems 2 and 3 in [1]). Again, in [2], this lemma has been used (as a special case where $\mathscr{C}$ is also a d.m.c.) to prove the existence of the "group code capacities" for d.m.c. and simultaneous channels (Theorems 1 and 2 in [2]). In view of these implications of this lemma, we shall spend some time in discussing it. We first give a detailed proof of a weaker version of this lemma. This proof enables one to see why the proof of Lemma 1.1, as given by Ahlswede in [1], is incorrect. We then make some remarks and briefly mention the impact of all this on the existence proofs given in [1] and [2]. The detailed proof of the weakened lemma further allows us to merely sketch the proofs of similar lemmas which are proved later on.

We then define certain special a.v.ch. for which useful generalizations of Lemma 1 of [5] can be proved. In all these a.v.ch. the "variation" of the c.p.f. is heavily constrained in some sense and we are therefore able to give strong capacities in the cases (λ_3, S^-, R^-), (λ_1, S^-, R^+) and (λ_1, S^+, R^-). The results given in the last two cases include another interesting result for the case (λ_3, S^-, R^+), where so far no capacities have been found, even for a.v.ch. with a binary output.

2. The Weakened Lemma

We give below a weakened version of Lemma 1.1 and a detailed proof, which serves to show why Lemma 1.1 is not correct. We shall assume in the rest of this paper that $b>2$, unless the contrary is explicitly stated.

Lemma 2.1. *For any given* $\omega\in[0,1]$, $d\geqq 1$ *and some alphabet size* a, *let* $\mathscr{C}$ *be a* (ω, a)*-symmetric* a.v.ch. *and* $w(\cdot|\cdot)$ *a* $(d\omega, a)$*-symmetric* d.m.c. *Further, let*

$$\{(u_i, A_i)\,|\,i=1,\dots,N\}$$

be a s.m.l.c. *for the* d.m.c. $w(\cdot|\cdot)$. *Then, provided*

(2.1) (i) $d\geqq a-1$ *and*

(ii) $\omega<(a-1)/ad$, *we have*

(2.2) $P(A_i|\,u_i\,|s_n)\geqq P_n(A_i|\,u_i)$ *for all* $i=1,\dots,N$ *and for all* $s_n\in S_n$. *Note that because of* (2.1)(ii) *we have* $1-d\omega>1/a>0$ *so that the* $(d\omega, a)$*-symmetric* d.m.c. *is always meaningful.*

Proof. We closely follow the proofs of Lemma 1 of [5] and Lemma 1 of [1]. For convenience we denote $w(\cdot|\cdot)$ by $w(\cdot|\cdot|s^*)$ and $P_n(\cdot|\cdot)$ by $P(\cdot|\cdot|s_n^*)$. Let $S'=S\cup\{s^*\}$. We prove (2.2) by the usual iterative argument. Let $s_n'\in S_n'$ and suppose that for some t the t-th component of s_n' is s^* i.e. $s'^t=s^*$. Let s_n'' be obtained from s_n' by replacing $s^*=s'^t$ by some $s\in S$. Then we prove that, for all $i=1,\dots,N$,

$$P(A_i|\,u_i\,|s_n'')\geqq P(A_i|\,u_i\,|s_n'). \tag{2.3}$$

Clearly (2.3) implies (2.2). As usual define for all $i=1,\dots,N$ and for all $j=1,\dots,a$,

$$\begin{aligned} {}_jA_i^t&=\{y_n|\,y_n\in A_i \text{ and } y^t=j\}\\ {}_jA_i^{*t}&=\{(y^1,\dots,y^{t-1},y^{t+1},\dots,y^n)\,|\\ &\qquad (y^1,\dots,y^{t-1},j,y^{t+1},\dots,y^n)\in{}_jA_i^t\}. \end{aligned} \tag{2.4}$$

Assume, without loss of generality (w.l.o.g.) that $u_i^t=1$. Then,

$$ {}_1A_i^{*t}\supseteq{}_jA_i^{*t} \qquad \text{for all } j=1,\dots,a. \tag{2.5}$$

Note that a violation of (2.1)(ii) implies either that $\omega=(a-1)/da$ so that $A_i=\emptyset$ for all i, or that $\omega>(a-1)/da$ so that the reverse inclusion to (2.5) holds. Also, in general,

$$ {}_jA_i^{*t}\neq{}_kA_i^{*t} \qquad \text{for } j\neq k. \tag{2.6}$$

Again, w.l.o.g. take $t=n$ and remember that the output alphabet is

$$Y=\{1,2,\dots,a\}. \tag{2.7}$$

Denote the complement with respect to Y of any $B\subseteq Y$ by B^c, and define for any $B\subseteq Y$ and $B\neq\emptyset$,

$$ {}_BA_i^{*n}=\{y_{n-1}=(y^1,\dots,y^{n-1})|\,y_{n-1}\in{}_jA_i^{*n}\ \ \forall j\in B \text{ and } y_{n-1}\notin{}_kA_i^{*n}\ \ \forall k\in B^c\}. \tag{2.8}$$

From (2.5) and (2.8) we see that if for some $B\subseteq Y$, $B\neq\emptyset$, we have $1\notin B$, then, necessarily,

$$ {}_BA_i^{*n}=\emptyset. \tag{2.9}$$

Therefore, define

$$\mathscr{A}_n=\{B\,|\,B\subseteq Y, B\neq\emptyset \text{ and } {}_BA_i^{*n}\neq\emptyset\}. \tag{2.10}$$

We remark here that if B, $C \subseteq Y$ and $B \neq C$, then

$$(2.11) \qquad {}_BA_i^{*n} \cap {}_CA_i^{*n} = \emptyset$$

and also that

$$(2.12) \qquad \bigcup_{\substack{\text{all}\\ B\in\mathscr{A}_n}} {}_BA_i^{*n} = \bigcup_{j=1}^{a} {}_jA_i^{*n}.$$

Then, remembering that $u_i^n = 1$ and $s'^n = s^*$, we can write

$$(2.13) \quad P(A_i|u_i|s'_n) = \sum_{B\in\mathscr{A}_n} P({}_BA_i^{*n}|(u_i^1, \dots, u_i^{n-1})|(s'^1, \dots, s'^{n-1}))\cdot\Big(\sum_{j\in B} w(j|1|s^*)\Big).$$

Similarly, since $s''^n = s \in S$, but $s''^t = s'^t$ for $t = 1, \dots, n-1$,

$$(2.14) \quad P(A_i|u_i|s''_n) = \sum_{B\in\mathscr{A}_n} P({}_BA_i^{*n}|(u_i^1, \dots, u_i^{n-1})|(s'^1, \dots, s'^{n-1}))\cdot\Big(\sum_{j\in B} w(j|1|s)\Big).$$

We prove below that

$$(2.15) \qquad \sum_{j\in B} w(j|1|s) \geqq \sum_{j\in B} w(j|1|s^*) \qquad \text{for all } B\in\mathscr{A}_n.$$

These last three equations imply (2.3) and hence (2.2). To prove (2.15), note that (2.9) implies that any $B\in\mathscr{A}_n$ contains 1 and m other members, $m \leqq a-1$. If $m = a-1$, then $B = Y$ and (2.15) holds with both sides equal to one. If $m < a-1$,

$$(2.16) \qquad \sum_{j\in B} w(j|1|s^*) = 1 - d\omega + md\omega/(a-1) \leqq 1-\omega,$$

the last step being true only if (2.1)(i) holds. Now, for any $s\in S$ we have $w(1|1|s) \geqq 1-\omega$ so that for any $B\in\mathscr{A}_n$

$$(2.17) \qquad \sum_{j\in B} w(j|1|s) \geqq 1-\omega.$$

Equality in (2.17) is possible, e.g. if $B = \{1, 3, \dots, a\}$ and $w(2|1|s) = \omega$. Hence the need for the last step in (2.16) and consequently for (2.1)(i). Now, (2.16) and (2.17) prove that (2.15) holds also when $m < a-1$, and the lemma is proved.

3. Miscellaneous Remarks

Remark 1. For binary channels ($a = b = 2$) Lemma 2.1 is true with $d \geqq 1$. In particular, with $d = 1$ this lemma reduces to Lemma 1 of [5] for the special case of the binary symmetric a.v.ch.

Remark 2. The inclusion (2.5) holds iff (2.1)(ii) holds. This is important for us later on (see Lemma 4.3). Using the remark following (2.5) it is easy to construct a counterexample to Theorem 1 of [1] (and hence to Lemma 1.1). According to Theorem 1 of [1], the capacity of the $(\frac{1}{2}, a)$-symmetric a.v.ch. $\geqq$ the capacity of the $(1, a)$-symmetric d.m.c. $> \log(a/(a-1)) > 0$. However, if we use the necessary and sufficient conditions for non-zero capacity for an a.v.ch. as given by Theorem 1 of [9], we find that the capacity of the $(\frac{1}{2}, a)$-symmetric a.v.ch. is zero.

Remark 3. Condition (2.1)(i) is needed if the last step in (2.16) is to hold for all possible $m<a-1$. This arises because of (2.6). If we had ${}_jA_i^{*t}=A_i^{*t}$ for all $j \neq k$, $j, k \neq 1$ (assuming $u_i^t=1$), then Lemma 1.1 would be true for any a and any $\omega<(a-1)/2a$. However, taking $n=4$ and just the two code words $u_1=(1,1,1,1)$ and $u_2=(2,2,2,2)$, we can construct the s.m.l. decoding region for u_1 and check that (2.6) is indeed true in this case.

Remark 4. In [1] and [2] the existence of weak capacity in various cases has been proved using Lemma 1.1 and the idea of concatenation. Because Lemma 1.1 is incorrect, these proofs are not valid. Further, the proofs cannot be carried through if we try to use Lemma 2.1 instead of Lemma 1.1, because of the conditions (2.1)(i) and (ii)[1].

Lemma 2.1 is of course an extension of Lemma 1 of [5]. Similar extensions for some special a.v.ch. will be used in the rest of this paper to find the strong capacities of these a.v.ch. in a number of communication situations[2].

4. Capacities of Certain Special a.v.ch. in the Case (λ_3, S^-, R^-)

Here we consider transmission over certain special a.v.ch. (defined below) using pure codes and maximal error, when neither the sender nor the receiver have any side information about the channel sequence (i.e. codes as defined in (1.3) and (1.4) are used). We now define the various a.v.ch. to be studied.

(4.1) The single parameter a-ary symmetric a.v.ch. is determined by the set of $a \times a$ stochastic matrices

$$\mathscr{C}_{(2)}=\{w_2(\cdot|\cdot|s)|s\in S\subseteq[0,1]\},$$

where for any $s\in[0,1]$ we have, for all $k=1,\ldots,a$,

$$w_2(j|k|s)=\begin{cases}1-s & \text{if } j=k\\ s/(a-1) & \text{for all } j\neq k.\end{cases}$$

Thus the a.v.ch. is determined by a set of (s,a)-symmetric matrices, indexed by the values of s.

(4.2) Let $z_a=(z^1,\ldots,z^a)$ be a point in the a-dimensional unit hypercube, I_a, i.e. $z^k\in[0,1]$ for all $k=1,\ldots,a$, and define for each $z_a\in I_a$ the $a\times b$ stochastic matrix

$$w_3(j|k|z_a)=\begin{cases}z^k & \text{if } j=k\\ 0 & \text{if } j\neq k, j\neq b=a+1\\ 1-z^k & \text{when } j=b=a+1.\end{cases}$$

Clearly such a matrix determines a general erasure-type d.m.c. The a-ary erasure a.v.ch. is determined by the collection

$$\mathscr{C}_{(3)}=\{w_3(\cdot|\cdot|s)|s\in S\subseteq I_a\}.$$

[1] We are informed by the referees, however, that the proofs can be fixed by using the idea of concatenation in conjunction with random coding.

[2] We are also informed by the referees that Lemma 1 of [5] has been extended by Ahlswede (see [3]) to the interesting Gaussian case.

(4.3) The matrix $w_4(\cdot|\cdot|s)$, defined for any $s\in[0,1]$ by

$$w_4(j|k|s)=\begin{cases}0 & \text{if } j>k\\ s(1-s)^{k-j} & \text{for all } j=2,\dots,k\\ (1-s)^{k-1} & \text{if } j=1,\end{cases}$$

is the c.p.f. of a "ladder" d.m.c. (see [6]). The collection

$$\mathscr{C}_{(4)}=\{w_4(\cdot|\cdot|s)|s\in S\subseteq[0,1]\},$$

will be called the a-ary ladder a.v.ch.

Remark. The "star" a.v.ch. can be defined as a generalization of the ladder a.v.ch. (see [6]), and all results proved in this paper for the ladder a.v.ch. can be generalized in a straightforward fashion for the star a.v.ch.

We shall now give detailed results for $\mathscr{C}_{(2)}$ only and merely state the relevant results for the other channels, pointing out differences in the proofs, if any.

(4.4) Let $\bar{S}$ denote the convex closure of S, i.e. $\bar{S}$ is the smallest closed convex set containing S. Then, we denote by $\overline{\mathscr{C}}_{(2)}$ the convex closure of $\mathscr{C}_{(2)}$, where

$$\overline{\mathscr{C}}_{(2)}=\{w_2(\cdot|\cdot|s)|s\in\bar{S}\}.$$

(4.5) For any $a\times b$ stochastic matrix $w(\cdot|\cdot|s)$ and any probability vector π_a (the input probabilities) define

$$\sigma_b(s)=(\sigma^1(s),\sigma^2(s),\dots,\sigma^b(s))$$

by

$$\sigma^j(s)=\sum_{k=1}^{a}\pi^k w(j|k|s),$$

and let

(4.6) $$H(\pi_a)=-\sum_{k=1}^{a}\pi^k\log\pi^k.$$

Then, we define the "rate", $R(\pi_a, w(\cdot|\cdot|s))$, by

(4.7) $$R(\pi_a, w(\cdot|\cdot|s))=H(\sigma_b(s))-\sum_{k=1}^{a}\pi^k H(w(\cdot|k|s)).$$

(4.8) It is easily verified that $R(\pi_a, w_2(\cdot|\cdot|s))$ is convex $\bigcap$ in π_a and convex $\bigcup$ in $s\in[0,1]$ so that by a well known theorem (see Lemma 4 of [5])

(4.9) $$\begin{aligned}C_2&=\max_{\pi_a}\min_{s\in\bar{S}}R(\pi_a, w_2(\cdot|\cdot|s))\\&=\min_{s\in\bar{S}}\max_{\pi_a}R(\pi_a, w_2(\cdot|\cdot|s))\geqq 0.\end{aligned}$$

Also, for any π_a we get

(4.10) $$R(\pi_a, w_2(\cdot|\cdot|(a-1)/a))=0.$$

(4.11) Further, for any fixed $s\in[0,1]$, $R(\pi_a, w_2(\cdot|\cdot|s))$ is maximized over π_a at $\pi_a^*=(1/a, 1/a, \dots, 1/a)$ yielding

(4.12) $$R^*(s)=\log a-s\log(a-1)-H(s,1-s),$$

so that

$$C_2 = \min_{s \in \bar{S}} R^*(s). \tag{4.13}$$

With these definitions we now state some lemmas and the main result for the a.v.ch. $\mathscr{C}_{(2)}$.

Lemma 4.1. *A* (n, N, λ) *code for* $\mathscr{C}_{(2)}$ *is a* (n, N, λ) *code for* $\bar{\mathscr{C}}_{(2)}$ *and conversely.*

Proof. The proof of this is so similar to that of Lemma 3 of [5] that we omit it.

Lemma 4.2. *Let* $\bar{S} \subseteq [0, (a-1)/a)$ *and define* $s'' = \max \bar{S}$. *Then, the* d.m.c. $w_2(\cdot|\cdot|s'')$ *underlies the* a.v.ch. $\bar{\mathscr{C}}_{(2)}$.

Proof. We prove this by showing that if $\{(u_i, A_i) | i = 1, \ldots, N\}$ is any s.m.l.c. for the d.m.c. $w_2(\cdot|\cdot|s'')$, then for all $i = 1, \ldots, N$, and for all $s_n \in \bar{S}_n$ we have

$$P(A_i | u_i | s''_n) \leqq P(A_i | u_i | s_n),$$

where $s''_n = (s'', s'', \ldots, s'')$. The proof of this relation is naturally similar to the proof of Lemma 2.1 except that the conditions (2.1) are not needed in this case. Thus, with the same notation, except that $w(\cdot|\cdot|s)$ is now replaced by $w_2(\cdot|\cdot|s)$ and $w(\cdot|\cdot|s^*)$ by $w_2(\cdot|\cdot|s'')$, we have instead of Eq. (2.16),

$$\sum_{j \in B} w_2(j|1|s) = 1 - s + ms/(a-1) \tag{4.14}$$

for all $s \in [0, 1]$. Hence, for any $s \leqq s''$ (i.e. for all $s \in \bar{S}$)

$$\sum_{j \in B} w_2(j|1|s) \geqq \sum_{j \in B} w_2(j|1|s''), \tag{4.15}$$

and the rest of the proof goes as in the proof of Lemma 2.1.

Lemma 4.3. *Let* $\bar{S} \subseteq ((a-1)/a, 1]$ *and define* $s' = \min \bar{S}$. *Then, the* d.m.c. $w_2(\cdot|\cdot|s')$ *underlies the* a.v.ch. $\bar{\mathscr{C}}_{(2)}$.

Proof. Again, we slightly modify the proof of Lemma 2.1 to obtain the required result. Thus, because now $s' > (a-1)/a$, (2.5) is replaced by the reverse inclusion viz.

$${}_1A_i^{*t} \subseteq {}_jA_i^{*t} \quad \text{for all } j = 2, \ldots, a. \tag{4.16}$$

And now, if for some $B \subseteq Y$, $B \neq \emptyset$ we have $1 \in B$, then $B = Y$. Thus any $B \in \mathscr{A}_n$ now contains either $m \leqq a-1$ elements *excluding* 1 or $B = Y$. Hence, for $B \neq Y$ we have

$$\sum_{j \in B} w_2(j|1|s) = ms/(a-1) \quad \text{for any } s \in \bar{S}, \tag{4.17}$$

and it follows that for all $B \in \mathscr{A}_n$ and for all $s \in \bar{S}$

$$\sum_{j \in B} w_2(j|1|s) \geqq \sum_{j \in B} w_2(j|1|s'), \tag{4.18}$$

so that the rest of the proof goes through as before.

Theorem 4.1. *The strong capacity of the* a.v.ch. $\mathscr{C}_{(2)}$ *in the case* (λ_3, S^-, R^-) *is given by* C_2.

Proof. Because of Lemma 4.1 we can consider $\bar{\mathscr{C}}_{(2)}$ instead of $\mathscr{C}_{(2)}$ (i.e. $\bar{S}$ instead of S). Then, we first note that for any $S \subseteq [0,1]$ we must have the following three mutually exclusive cases:

(i) $\bar{S} \subseteq [0,(a-1)/a)$,

(ii) $\bar{S} \subseteq ((a-1)/a, 1]$,

(iii) the point $(a-1)/a \in \bar{S}$. This includes the case when points of S lie in both, $[0,(a-1)/a)$ and $((a-1)/a, 1]$.

Suppose now that (i) is true. Then, from Lemma 4.2 we see that the d.m.c. $w_2(\cdot|\cdot|s'')$ underlies $\bar{\mathscr{C}}_{(2)}$. Also, clearly $s'' \in \bar{S}$ so that this d.m.c. is basic to $\bar{\mathscr{C}}_{(2)}$ and, by Theorem 1.1 the capacity of $\bar{\mathscr{C}}_{(2)}$ is $\max\limits_{\pi_a} R(\pi_a, w_2(\cdot|\cdot|s''))$ which is seen from (4.11) and (4.12) to be $R^*(s'')$. It remains to show that, in this case,

$$C_2 = \min_{s \in \bar{S}} R^*(s) = R^*(s''). \tag{4.19}$$

If we note from (4.8) that $R^*(s)$ is a convex $\bigcup$ function of s in $[0,1]$ with a unique minimum at $s=(a-1)/a$, then we see that it is non-increasing in the interval $[0,(a-1)/a)$ (see also (4.11) and (4.12)). This proves (4.19).

In case (ii) holds, we invoke Lemma 4.3 instead of 4.2 and the rest is similar to the argument given above.

If (iii) is true, then clearly C_2, as defined in (4.9), is zero and we must show that the capacity of $\bar{\mathscr{C}}_{(2)}$ in this case is indeed zero. This is easily done by noting that the matrix $w_2(\cdot|\cdot|(a-1)/a) \in \bar{\mathscr{C}}_{(2)}$ so that every code for $\bar{\mathscr{C}}_{(2)}$ is also a code for the d.m.c. determined by this matrix. The strong converse for this d.m.c., whose capacity is obviously zero, completes the proof.

With the same definition of $\bar{S}$ for the a.v.ch. $\mathscr{C}_{(4)}$ we see that Lemma 4.1 is true for this a.v.ch. also. Further, we have

Lemma 4.4. *If* $s' = \min \bar{S}$, *then the* d.m.c. $w_4(\cdot|\cdot|s')$ *underlies the* a.v.ch. $\mathscr{C}_{(4)}$.

Proof. The proof here is also similar to that of Lemma 2.1; we simply note that when $u_i^t = j$ we have the inclusion relations

$$_1A_i^{*t} \subseteq {}_2A_i^{*t} \subseteq \cdots \subseteq {}_jA_i^{*t}, \tag{4.20}$$

and further,

$$_kA_i^{*t} = 0 \qquad \text{for all } k > j. \tag{4.21}$$

The rest of the proof follows as usual.

Finally, we have a theorem analogous to Theorem 4.1 which can be proved in a similar fashion.

Theorem 4.2. *The capacity of the* a.v.ch. $\mathscr{C}_{(4)}$ *in the case* (λ_3, S^-, R^-) *is, with* $s' = \min \bar{S}$ *as usual,*

$$\mathscr{C}_4 = \max_{\pi_a} \min_{s \in \bar{S}} R(\pi_a, w_4(\cdot|\cdot|s)) = \max_{\pi_a} R(\pi_a, w_4(\cdot|\cdot|s')). \tag{4.22}$$

For the a.v.ch. $\mathscr{C}_{(3)}$ defined in (4.2) let (4.22) $Z^t=\{v\,|\,\text{there exists } z_a\in S \text{ such that } z^t=v\}$. Clearly $Z^t\subseteq[0,1]$ for all $t=1,\dots,a$. If we now define

(4.23) $$\bar{\bar{S}}=\prod_{t=1}^{a}\bar{Z}^t,$$

the row-convex closure of $\mathscr{C}_{(3)}$ (see (2.13) and (2.14) of [5]) is given by

(4.24) $$\bar{\bar{\mathscr{C}}}_{(3)}=\{w_3(\cdot|\cdot|s)\,|\,s\in\bar{\bar{S}}\subseteq I_a\}.$$

Now, if $s^*=z_a^*=(z^{*1},z^{*2},\dots,z^{*a})$, where

(4.25) $$z^{*t}=\min\bar{Z}^t \quad \text{for each } t=1,\dots,a,$$

then, in a manner similar to Theorems 4.1 and 4.2 we can prove the following:

Theorem 4.3. *The capacity of the* a.v.ch. $\mathscr{C}_{(3)}$ *in the case* (λ_3, S^-, R^-) *is given by*

$$\mathscr{C}_3=\max_{\pi_a}\min_{s\in\bar{\bar{S}}}R(\pi_a, w_3(\cdot|\cdot|s))=\max_{\pi_a}R(\pi_a, w_3(\cdot|\cdot|s^*)).$$

5. Capacities for Certain Special a.v.ch. in the Cases (λ_1, S^+, R^-) and (λ_1, S^-, R^+)

In this section we consider the possibility of randomized encoding and decoding and in all cases we are interested in the maximal error of the code. Further, in one case (denoted by S^+) the sender knows the c.p.f. governing the transmission of the t-th letter just before he sends it, and in the other case (denoted by R^+) the receiver knows the channel sequence governing a received word (sequence) before he decodes it. We give below upper and lower bounds to the capacity of $\mathscr{C}_{(2)}$ in these cases. Under certain conditions (which are given below) these bounds are equal and they give us the strong capacity of the channel. As a special case of these results, when $a=2$, we get the capacity of $\mathscr{C}_{(2)}$ for the cases (λ_3, S^+, R^-) and (λ_3, S^-, R^+), the last case mentioned being new. Further, a theorem of a general nature is proved and it gives the capacities of $\mathscr{C}_{(4)}$ and a restricted version of $\mathscr{C}_{(3)}$ in all cases of interest.

We first consider $\mathscr{C}_{(2)}$. Note that Lemma 4.2 states that if $\{(u_i, A_i)\,|\,i=1,\dots,N\}$ is a s.m.l.c. for the d.m.c. $w_2(\cdot|\cdot|s^*)$, for some $s^*<(a-1)/a$, then $\{(u_i, A_i)\}$ is also a code (n, N, λ) for the a.v.ch. determined by any set of matrices $w_2(\cdot|\cdot|s)$ with $s\leqq s^*$. Further, if the sender knows that a particular letter will be transmitted with the c.p.f. $w_2(\cdot|\cdot|s^t)$, where $s^t>(a-1)/a$, then he can randomize as follows: If the letter "i" was to be transmitted, he now transmits the result of a chance experiment whose outcome will be some letter "j" with probability $p(j)=w_2(j|i|1)$. Thus, in such cases the c.p.f. becomes, in effect,

(5.1) $$w_2(\cdot|\cdot|1)\cdot w_2(\cdot|\cdot|s^t)=w_2(\cdot|\cdot|1-s^t/(a-1)).$$

Here we note that since $s^t>(a-1)/a$, we get $1-s^t/(a-1)<(a-1)/a$. Thus by this randomization strategy the same code $\{(u_i, A_i)\}$ can also be used for an a.v.ch. determined by any set of matrices $w_2(\cdot|\cdot|s)$ with $s>(1-s^*)(a-1)$. With these facts in mind, we introduce the following definitions which are made clear with the help of the figure (Fig. 1).

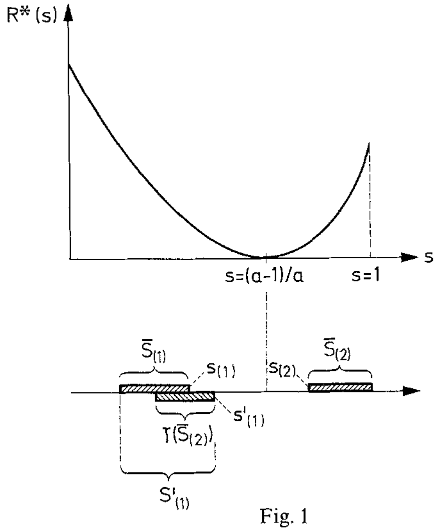

Fig. 1

(5.2) Define
$$S_{(1)}=S\cap[0,(a-1)/a],$$
$$S_{(2)}=S\cap[(a-1)/a,1]$$

and let T be a mapping defined by

(5.3) $$T(s)=1-s/(a-1)\quad\text{for all } s\in[0,1],$$

so that for any $S\subseteq[0,1]$

(5.4) $$T(S)=\{T(s)\,|\,s\in S\}.$$

Further, let

(5.5)
$$S'_{(1)}=\bar{S}_{(1)}\cup T(\bar{S}_{(2)}),$$
$$S'_{(2)}=\bar{S}_{(2)}\cup T(\bar{S}_{(1)}).$$

Note that $T(\bar{S}_{(i)})$ and hence $S'_{(i)}$ are closed sets for $i=1,2$. Finally, define

(5.6) $$s_{(1)}=\max\bar{S}_{(1)},\qquad s_{(2)}=\min\bar{S}_{(2)}$$

and similarly, $s'_{(1)}$ and $s'_{(2)}$. Also, we note that

(5.7) $$w_2(j|k|1)=\begin{cases}0 & \text{if } j=k\\ 1/(a-1) & \text{if } j\neq k\end{cases}$$

for all $k=1,\ldots,a$.

Theorem 5.1. *In the case* (λ_1, S^+, R^-) *the capacity of the* a.v.ch. $\mathscr{C}_{(2)}$ *is bounded below by*

$$C^-=\max_{i=1,2}R^*(s'_{(i)})$$

and bounded above by

$$C^+=\min_{i=1,2}R^*(s_{(i)}).$$

Proof. 1) First we prove that C^- is a lower bound to the capacity. Suppose that $C^-=R^*(s'_{(1)})$, the proof when $C^-=R^*(s'_{(2)})$ being (symmetrically) the same.

Let $\{(u_i, A_i)\}$ be a (n, N, λ) s.m.l.c. for the d.m.c. determined by $w_2(\cdot|\cdot|s'_{(1)})$. Using the randomization strategy given below the sender can always use $\{(u_i, A_i)\}$ as a (n, N, λ) code for $\mathscr{C}_{(2)}$. Suppose the i-th message is to be sent. Let u_i^t be the t-th letter of u_i and s^t be the t-th element of the channel sequence (i.e. the c.p.f. governing the transmission of the t-th letter will be $w_2(\cdot|\cdot|s^t)$). Then,

(i) if $s^t \in S_{(1)}$, u_i^t is sent;

(ii) if $s^t \in S_{(2)}$ the sender randomizes and sends letter "j" with probability $w_2(j|u_i^t|1)$. The resultant probability of receiving some letter "k", as we have seen, becomes $w_2(k|u_i^t|T(s^t))$. And from the Definition (5.6) of $s'_{(1)}$ we see that $T(s^t) \leqq s'_{(1)}$ for all possible $s^t \in S_{(2)}$. Hence, by Lemma 4.2, $\{(u_i, A_i)\}$ is also a code (n, N, λ) for $\mathscr{C}_{(2)}$ in the case (λ_1, S^+, R^-).

2) For the upper bound, we note that the existence of a (n, N, λ) code for the a.v.ch. $\mathscr{C}_{(2)}$ implies the existence of a (n, N, λ) code for the d.m.c. determined by any matrix in the collection $\mathscr{C}_{(2)}$. The smallest capacity of all these d.m.c. is C^+. This completes the proof of Theorem 5.1.

The question naturally arises as to whether these two bounds are ever equal. This is answered by the next theorem.

Theorem 5.2. *If, for any* $i=1,2$, *we have* $s'_{(i)} = s_{(i)}$, *then the strong capacity of the* a.v.ch. $\mathscr{C}_{(2)}$ *in the case* (λ_1, S^+, R^-) *is given by*

$$C^- = C^+ = R^*(s_{(i)}).$$

Proof. Suppose, without loss of generality, that $s'_{(1)} = s_{(1)}$. We have to show that

$$\max_{i=1,2} R^*(s'_{(i)}) = R^*(s'_{(1)}) = R^*(s_{(1)}) = \min_{i=1,2} R^*(s_{(i)}). \tag{5.8}$$

We do this by showing that

$$R^*(s_{(2)}) \geqq R^*(s_{(1)}) \tag{5.9}$$

and

$$R^*(s'_{(2)}) \leqq R^*(s'_{(1)}). \tag{5.10}$$

First note that $T(s_{(2)}) \in S'_{(1)}$ so $T(s_{(2)}) \leqq s'_{(1)}$. In the interval $[0, (a-1)/a]$, $R^*(s)$ is a decreasing function of s (see Fig. 1 and (4.8)) so we get

$$R^*(T(s_{(2)})) \geqq R^*(s'_{(1)}) = R^*(s_{(1)}). \tag{5.11}$$

But $w_2(\cdot|\cdot|T(s_{(2)}))$ is obtained from $w_2(\cdot|\cdot|s_{(2)})$ by multiplying on the left by $w_2(\cdot|\cdot|1)$. This can be considered as having two channels in cascade and it follows, from the "data processing theorem" (see for example Theorem 4.3.3 in [8]), that

$$R^*(s_{(2)}) \geqq R^*(T(s_{(2)})),$$

which with (5.11) proves (5.9).

Similarly, $T(s_{(1)}) \in S'_{(2)}$ so that $T(s_{(1)}) \geqq s'_{(2)}$ and here $R^*(s)$ is a non-decreasing function of s so that

$$R^*(T(s_{(1)})) \geqq R^*(s'_{(2)}). \tag{5.12}$$

And, again by the data processing theorem

$$R^*(T(s_{(1)}))\leqq R^*(s_{(1)})=R^*(s'_{(1)}),$$

which with (5.12) proves (5.10) and hence the theorem.

Because of the symmetry of these matrices, multiplication of any two is commutative. Hence, if the receiver knows channel sequence and the sender does not, the receiver can randomize before decoding in a manner similar to that described above and we would expect the same results to hold. More specifically, if $\{(u_i, A_i)\}$ is a (n, N, λ) code as in Theorem 5.1, then the sender transmits u_i when the i-th message is to be sent. The receiver, knowing the channel sequence, adopts the following strategy: If y^t is the t-th received letter and he knows that s^t is the t-th element of the channel sequence,

(i) if $s^t\in S_{(1)}$ he accepts y^t,

(ii) if $s^t\in S_{(2)}$ he randomizes over Y (the output alphabet) with probability $w_2(j|y^t|1)$ and accepts the resultant letter. Thus again the c.p.f. is, in effect, $w_2(\cdot|\cdot|T(s^t))$. It follows from the arguments used in Theorems 5.1 and 5.2 that

Theorem 5.3. *Theorems* 5.1 *and* 5.2 *are true for the* a.v.ch. $\mathscr{C}_{(2)}$ *if we change* (λ_1, S^+, R^-) *to* (λ_1, S^-, R^+).

It can be verified for the a.v.ch. $\mathscr{C}_{(2)}$ that we always have $s_{(1)}=s'_{(1)}$ if $a=2$. Further, the matrix $w_2(\cdot|\cdot|1)$ becomes $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ so that it is no longer necessary for the sender (receiver) to randomize when he knows the channel c.p.f. This leads to an interesting corollary to Theorems 5.2 and 5.3 above, which we state as a separate theorem.

Theorem 5.4. *Given* $a=2$ *so that*

$$\mathscr{C}_{(2)}=\left\{w_2(\cdot|\cdot|s)\,|\,w_2(\cdot|\cdot|s)=\begin{bmatrix}1-s & s\\ s & 1-s\end{bmatrix}\ \textit{for all}\ s\in S\right\},$$

where as usual $S\subseteq 0, 1$, *define*

$$w=\begin{bmatrix}1-s_0 & s_0\\ s_0 & 1-s_0\end{bmatrix}\quad \textit{where}\ 1-2s_0=\inf_{s\in S}|1-2s|.$$

Then, the capacity of the a.v.ch. $\mathscr{C}_{(2)}$ *in the cases* (λ_3, S^+, R^-) *and* (λ_3, S^-, R^+) *is the capacity of the* d.m.c. w.

Remark 1. In the above theorem, the case (λ_3, S^-, R^+) is of particular interest since it has not been treated for the general binary a.v.ch. in [5].

Remark 2. Theorems 5.1, 5.2, and 5.3 are essentially generalisations of Theorems 4.1 and 4.2 of [4], except for a difference in the method of proof which allows us to obtain Theorem 5.4 as a corollary. Theorem 5.4 states that Theorems 4.1 and 4.2 of [4] are also true when there is no randomization in encoding and decoding.

We now define a more restricted version of $\mathscr{C}_{(3)}$ by the collection

(5.13) $$\mathscr{C}_{(5)}=\{w_5(\cdot|\cdot|s)\,|\,s\in S\subseteq[0, 1]\},$$

where for all $k=1, \ldots, a$

$$w_5(j|k|s)=\begin{cases} s & \text{if } j=k \\ 0 & \text{if } j\neq k \text{ and } j\neq b=a+1 \\ 1-s & \text{if } j=b=a+1. \end{cases} \tag{5.14}$$

i.e. each matrix defines an erasure-type d.m.c. where the probability of erasure of each letter is the same.

The following theorem applies to $\mathscr{C}_{(4)}$ and $\mathscr{C}_{(5)}$.

Theorem 5.5. *If the* d.m.c. $w(\cdot|\cdot|)$ *is basic to the* a.v.ch. $\mathscr{C}$ *then the capacity of* $\mathscr{C}$ *in all the cases* $(\lambda_i, \cdot, \cdot, \cdot, \cdot)$ *and* $(\lambda_i, \cdot, \cdot, \cdot, \cdot)$, $i=1, 2, 3, 4$, *is given by the capacity of the* d.m.c. $w(\cdot|\cdot)$.

Proof. It can easily be verified that a (n, N, λ) code for the d.m.c. $w(\cdot|\cdot)$ can also be used as a (n, N, λ) code for the a.v.ch. $\mathscr{C}$ in each of the cases. Conversely, the existence of a (n, N, λ) code for the a.v.ch. in any of the cases implies the existence of a (n, N, λ) code for the d.m.c. $w(\cdot|\cdot)$. The same argument holds when we consider average errors.

Remark. For $r=4, 5$ we can see that the above theorem applies to the a.v.ch. $\mathscr{C}_{(r)}$, the d.m.c. $w_r(\cdot|\cdot|s')$ being the basic d.m.c. in each case ($s'=\min\bar{S}$ as usual).

References

1. Ahlswede, R.: A note on the existence of the weak capacity for channels with arbitrarily varying channel probability functions. Ann. Math. Statist. **41**, No. 3, 1027–1033 (1970)
2. Ahlswede, R.: Group Codes do not achieve Shannon's channel capacity for the general discrete memoryless channel. Ann. Math. Statist. **42**, No. 1, 224–240 (1971)
3. Ahlswede, R.: The capacity of a channel with arbitrarily varying Gaussian channel probability functions. To appear in the Transactions of the Sixth Prague Conference on Information Theory, Stat. Dec. Functions and Random Processes
4. Ahlswede, R., Wolfowitz, J.: Correlated decoding for channels with arbitrarily varying channel probability functions. Inform. Control **14**, No. 5, 457–473 (1969)
5. Ahlswede, R., Wolfowitz, J.: The capacity of channels with arbitrarily varying channel probability functions and binary output alphabet. Z. Wahrscheinlichkeitstheorie verw. Gebiete **15**, 186–194 (1970)
6. Berlekamp, E. R., Kleinrock, L.: Analysis of channels with unidirectional drift. JPL Space Programs Summary No. 37-39, Vol. IV, 226–230 (1966)
7. Blackwell, D., Breiman, L., Thomasian, A. J.: The capacities of certain channel classes under random coding. Ann. Math. Statist. **31**, 558–567 (1960)
8. Gallager, R. G.: Information Theory and Reliable Communication. New York: Wiley 1968
9. Kiefer, J., Wolfowitz, J.: Channels with arbitrarily verying channel probability functions. Inform. Control **5**, 44–54 (1962)
10. Wolfowitz, J.: Coding Theorems of Information Theory 2nd ed. Berlin-Heidelberg-New York: Springer 1964

N. S. Kambo and Samar Singh
Department of Mathematics
Indian Institute of Technology
Hauz Khas
New Delhi 110029
India

(Received April 4, 1973)