The Annals of Probability
1974, Vol. 2, No. 5, 805-814

THE CAPACITY REGION OF A CHANNEL WITH TWO SENDERS AND TWO RECEIVERS

BY R. AHLSWEDE

The Ohio State University

A characterization of the capacity region of a two-way channel is given for the communication situation, in which both senders send independent messages simultaneously to both receivers and all senders and receivers are at different terminals.

0. Summary. Let $X = \{1, \cdots, a\}$ be the input alphabet for the sender S_X and let $Y = \{1, \cdots, b\}$ be the input alphabet for the sender S_Y. Denote by $\bar{X} = \{1, \cdots, \bar{a}\}$ the output alphabet for the receiver $R_{\bar{X}}$ and by $\bar{Y} = \{1, \cdots, \bar{b}\}$ the output alphabet for the receiver $R_{\bar{Y}}$. Let $w(\cdot, \cdot \mid \cdot, \cdot)$ be a nonnegative function, which is defined on $X * Y * \bar{X} * \bar{Y}$ and satisfies $\sum_{\bar{x} \in \bar{X}} \sum_{\bar{y} \in \bar{Y}} w(\bar{x}, \bar{y} \mid x, y) = 1$ for every $(x, y) \in X * Y$. Set $X^t = X$, $Y^t = Y$, $\bar{X}^t = \bar{X}$ and $\bar{Y}^t = \bar{Y}$ for $t = 1, 2, \cdots$. For every n; $n = 1, 2, \cdots$; the transmission probabilities of a two-way channel are defined by $P(\bar{x}_n, \bar{y}_n \mid x_n, y_n) = \prod_{t=1}^n w(\bar{x}^t, \bar{y}^t \mid x^t, y^t)$ for every $x_n = (x^1, \cdots, x^n) \in X_n = \prod_{t=1}^n X^t$, $y_n = (y^1, \cdots, y^n) \in Y_n = \prod_{t=1}^n Y^t$, $\bar{x}_n = (\bar{x}^1, \cdots, \bar{x}^n) \in \bar{X}_n = \prod_{t=1}^n \bar{X}^t$ and every $\bar{y}_n = (\bar{y}^1, \cdots, \bar{y}^n) \in \bar{Y}_n = \prod_{t=1}^n \bar{Y}^t$.

In Section 3 we determine the capacity region of this channel for the communication situation, in which both senders send messages simultaneously to both receivers and all senders and receivers are at different terminals. This result was announced in [1]. The proof is based on a new approach to the coding problem for a channel with two senders and one receiver. This approach is presented in Section 2. It seems to extend to channels with i senders and j receivers as long as all senders send messages to *all* receivers.

In case S_X sends messages to $R_{\bar{X}}$ and S_Y sends messages to $R_{\bar{Y}}$ only, no satisfactory characterization of the capacity region is known. We show by an example in Section 4 that the capacity region cannot be obtained by using "independent sources", contrary to the belief expressed in the bottom lines of page 636 in [4]. In Section 5 we generalize Shannon's random coding method. This generalized version may be of interest for coding problems of certain multi-way channels.

1. Basic definitions and statement of the problems. We continue using the notation introduced in [1]. There the present two-way channel was denoted by (P, T_{22}). The communication situation, in which S_X sends to $R_{\bar{X}}$ and S_Y sends to $R_{\bar{Y}}$, was called (P, T_{22}, I) and the communication situation, in which S_X and S_Y send to $R_{\bar{X}}$ *and* to $R_{\bar{Y}}$, was called (P, T_{22}, II). We used the abbreviation (P, T_{21}, I), if both senders send to $R_{\bar{X}}$ (or to $R_{\bar{Y}}$) only. For $A \subset \bar{X}_n$, $B \subset \bar{Y}_n$ and

Received November 2, 1972.
AMS 1970 *subject classifications.* Primary 94.10; Secondary 60.10.
Key words and phrases. Probabilistic coding theory, multi-way channels, capacity regions.

$(x_n, y_n) \in X_n * Y_n$ define

$$P(A \mid x_n, y_n) = \sum_{\bar{x}_n \in A} \sum_{\bar{y}_n \in \bar{Y}_n} P(\bar{x}_n, \bar{y}_n \mid x_n, y_n) \tag{1.1}$$

and

$$Q(B \mid x_n, y_n) = \sum_{\bar{x}_n \in \bar{X}_n} \sum_{\bar{y}_n \in B} P(\bar{x}_n, \bar{y}_n \mid x_n, y_n) . \tag{1.2}$$

The transmission probabilities for (P, T_{21}, I) are those defined in (1.1). They correspond to the transmission matrix $p(\cdot \mid \cdot, \cdot)$ given by

$$p(\bar{x} \mid x, y) = \sum_{\bar{y}} w(\bar{x}, \bar{y} \mid x, y) \qquad \text{for } x \in X, y \in Y, \bar{x} \in \bar{X} . \tag{1.3}$$

Similarly one defines $q(\cdot \mid \cdot, \cdot)$ by summing in (1.3) over $\bar{x}$. A code $(n, N_1, N_2, \lambda, I)$ for (P, T_{22}, I) is a system $\{(u_i, v_j, A_i, B_j) \mid i = 1, \cdots, N_1; j = 1, \cdots, N_2\}$, where $u_i \in X_n$, $v_j \in Y_n$, $A_i \subset \bar{X}_n$, $B_j \subset \bar{Y}_n$ (for $i = 1, \cdots, N_1; j = 1, \cdots, N_2$), $A_i \cap A_i' = B_j \cap B_j' = \emptyset$ for $i \neq i', j \neq j'$ and which satisfies

$$\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} [P(A_i^c \mid u_i, v_j) + Q(B_j^c \mid u_i, v_j)] \leqq \lambda . \tag{1.4}$$

A code $(n, N_1, N_2, \lambda, II)$ for (P, T_{22}, II) is a system $\{u_i, v_j, A_{ij}, B_{ij}) \mid i = 1, \cdots, N_1; j = 1, \cdots, N_2\}$ where $u_i \in X_n$, $v_j \in Y_n$, $A_{ij} \subset \bar{X}_n$, $B_{ij} \subset \bar{Y}_n$ $(i = 1, \cdots, N_1; j = 1, \cdots, N_2)$, $A_{ij} \cap A_{i'j'} = B_{ij} \cap B_{i'j'} = \emptyset$ for $(i, j) \neq (i', j')$, and which satisfies

$$\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} [P(A_{ij}^c \mid u_i, v_j) + Q(B_{ij}^c \mid u_i, v_j)] \leqq \lambda . \tag{1.5}$$

Finally, a code (n, N_1, N_2, λ) for (P, T_{21}, I) is a system $\{(u_i, v_j, A_{ij}) \mid i = 1, \cdots, N_1; j = 1, \cdots, N_2\}$, where $u_i \in X_n$, $v_j \in Y_n$, $A_{ij} \subset \bar{X}_n$ $(i = 1, \cdots, N_1; j = 1, \cdots, N_2)$, $A_{ij} \cap A_{i'j'} = \emptyset$ for $(i, j) \neq (i', j')$ and for which

$$\frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} P(A_{ij}^c \mid u_i, v_j) \leqq \lambda . \tag{1.6}$$

A pair of nonnegative real numbers (R_1, R_2) is called a pair of achievable rates for (P, T_{22}, II) if for any λ, $0 < \lambda < 1$, and any $\varepsilon > 0$ there exists a code $(n, \exp\{(R_1 - \varepsilon)n\}, \exp\{(R_2 - \varepsilon)n\}, \lambda, II)$ for all sufficiently large n. The capacity region $G(P, T_{22}, II)$ is the set of all pairs of achievable rates for (P, T_{22}, II). Analogously one defines $G(P, T_{22}, I)$ and $G(P, T_{21}, I)$. The objective of the present paper is to characterize the set $G(P, T_{22}, II)$. We also shall comment on the structure of $G(P, T_{22}, I)$. A look at (1.5) shows that the coding problem for (P, T_{22}, II) is actually a problem of simultaneous coding for two channels of type (P, T_{21}, I). We assume the reader's familiarity with Section 3 of [1], in which $G(P, T_{21}, I)$ was determined. However, the approach taken there seems to be not adaptive to simultaneous coding. In Section 2 we give a new characterization of $G(P, T_{21}, I)$, which is derived by a more canonical argument than the one given in [1] and which easily extends to the channel (P, T_{22}, II).

The following definitions are needed in order to state the results of the later Sections.

Let p be a probability distribution (p.d.) on X and let q be a p.d. on Y. We define the functions $R(p, q, \bar{X})$, $R_1(p, q, \bar{X})$, and $R_2(p, q, \bar{X})$ by

$$(1.7)\qquad R(p, q, \bar{X}) = \sum_{x,y,\bar{x}} p(x)q(y)p(\bar{x} \mid x, y) \log \frac{p(\bar{x} \mid x, y)}{\sum_{x,y} p(x)q(y)p(\bar{x} \mid x, y)}$$

$$(1.8)\qquad R_1(p, q, \bar{X}) = \sum_{x,y,\bar{x}} p(x)q(y)p(\bar{x} \mid x, y) \log \frac{p(\bar{x} \mid x, y)}{\sum_{x} p(x)p(\bar{x} \mid x, y)}$$

and

$$(1.9)\qquad R_2(p, q, \bar{X}) = \sum_{x,y,\bar{x}} p(x)q(y)p(\bar{x} \mid x, y) \log \frac{p(\bar{x} \mid x, y)}{\sum_{y} q(y)p(\bar{x} \mid x, y)} .$$

Analogously, we define functions $R(p, q, \bar{Y})$, $R_1(p, q, \bar{Y})$ and $R_2(p, q, \bar{Y})$. One easily verifies that all six functions are convex in p and in q. The following functions are needed in Section 4 only.

$$(1.10)\qquad R_{12}^1(p, q, \bar{X}) = \sum_{x,y,\bar{x}} p(x)q(y)p(\bar{x} \mid x, y) \log \frac{\sum_x p(x)p(\bar{x} \mid x, y)}{\sum_{x,y} p(x)q(y)p(\bar{x} \mid x, y)} .$$

$$(1.11)\qquad R_{21}^2(p, q, \bar{X}) = \sum_{x,y,\bar{x}} p(x)q(y)p(\bar{x} \mid x, y) \log \frac{\sum_y q(y)p(\bar{x} \mid x, y)}{\sum_{x,y} p(x)q(y)p(\bar{x} \mid x, y)} .$$

Similarly one defines functions $R_{12}^1(p, q, \bar{Y})$ and $R_{21}^2(p, q, \bar{Y})$.

For $t = 1, 2, \cdots$ let p^t be a p.d. on X^t and let q^t be a p.d. on Y^t. For $n = 1, 2, \cdots$ define probability distributions p_n on X_n and q_n on Y_n by

$$(1.12)\qquad p_n(x_n) = \prod_{t=1}^n p^t(x^t) \qquad \text{for } x_n = (x^1, \cdots, x^n) \in X_n \text{ and}$$

$$(1.13)\qquad q_n(y_n) = \prod_{t=1}^n q^t(y^t) \qquad \text{for } y_n = (y^1, \cdots, y^n) \in Y_n .$$

Finally, define the "information functions" I^1, I^2, I, J^1, J^2, and J by

$$(1.14)\qquad I^1(\bar{x}_n, x_n \mid y_n) = \log \frac{P(\bar{x}_n \mid x_n, y_n)}{\sum_{x_n} p_n(x_n)P(\bar{x}_n \mid x_n, y_n)} ;$$

$$(1.15)\qquad I^2(\bar{x}_n, y_n \mid x_n) = \log \frac{P(\bar{x}_n \mid x_n, y_n)}{\sum_{y_n} q_n(y_n)P(\bar{x} \mid x_n, y_n)} ;$$

$$(1.16)\qquad I(\bar{x}_n, x_n, y_n) = \log \frac{P(\bar{x}_n \mid x_n, y_n)}{\sum_{x_n, y_n} p_n(x_n)q_n(y_n)P(\bar{x}_n \mid x_n, y_n)} ;$$

$$(1.17)\qquad J^1(\bar{y}_n, x_n \mid y_n) = \log \frac{Q(\bar{y}_n \mid x_n, y_n)}{\sum_{x_n} p_n(x_n)Q(\bar{y}_n \mid x_n, y_n)} ;$$

$$(1.18)\qquad J^2(\bar{y}_n, y_n \mid x_n) = \log \frac{Q(\bar{y}_n \mid x_n, y_n)}{\sum_{y_n} q_n(y_n)Q(\bar{y}_n \mid x_n, y_n)} ;$$

$$(1.19)\qquad J(\bar{y}_n, x_n, y_n) = \log \frac{Q(\bar{y}_n \mid x_n, y_n)}{\sum_{x_n, y_n} p_n(x_n)q_n(y_n)Q(\bar{y}_n \mid x_n, y_n)} .$$

2. A new characterization of the capacity region $G(P, T_{21}, I)$. We shall need a result, which was derived by means of Fano's Lemma ([2]) in Section 3 of [1]. Let $\{(u_i, v_j, A_{ij}) \mid i = 1, \cdots, N_1; j = 1, \cdots, N_2\}$ be an (n, N_1, N_2, λ) code for

(P, T_{21}, I). It follows from the inequality (1.6) that one can pick—after renumbering—a sequence $u_1, \cdots, u_{N_1^*}$, with $N_1^* = [N_1/2]$, and a sequence $v_1, \cdots, v_{N_2^*}$, with $N_2^* = [N_2/2]$, such that

$$\frac{1}{N_2^*} \sum_{j=1}^{N_2^*} P(A_{ij}^c \mid u_i, v_j) \leqq 4\lambda \qquad \text{for } i = 1, \cdots, N_1^*, \text{ and} \tag{2.1}$$

$$\frac{1}{N_1^*} \sum_{i=1}^{N_1^*} P(A_{ij}^c \mid u_i, v_j) \leqq 4\lambda \qquad \text{for } j = 1, \cdots, N_2^*. \tag{2.2}$$

Write $u_i = (u_i^1, \cdots, u_i^n)$ for $i = 1, \cdots, N_1^*$ and $v_j = (v_j^1, \cdots, v_j^n)$ for $j = 1, \cdots, N_2^*$. We define the following probability distributions:

$$p^t(x) = N_1^{*-1}|\{i \mid u_i^t = x, i \in \{1, \cdots, N_1^*\}\}| \qquad \text{for } x \in X^t;\ t = 1, 2, \cdots, n; \tag{2.3}$$

and

$$q^t(y) = N_2^{*-1}|\{j \mid v_j^t = y, j \in \{1, \cdots, N_2^*\}\}| \qquad \text{for } y \in Y^t;\ t = 1, 2, \cdots, n. \tag{2.4}$$

For any λ, $0 < \lambda < \frac{1}{4}$, the following inequalities hold (see (3.35), (3.42) and (3.43) of [1]):

$$\log N_1 N_2 \leqq [\textstyle\sum_{t=1}^{n} R(p^t, q^t, \bar{X}) + 1](1 - 4\lambda)^{-1} + \log 4 \tag{2.5}$$

$$\log N_s \leqq [\textstyle\sum_{t=1}^{n} R_s(p^t, q^t, \bar{X}) + 1](1 - 4\lambda)^{-1} + \log 4 \qquad \text{for } s = 1, 2. \tag{2.6}$$

We give now a few more definitions needed to formulate the main result (Theorem 1) of this Section.

Define a set $F(\bar{X})$ by

$$F(\bar{X}) = \{(R_1(p, q, \bar{X}), R_2(p, q, \bar{X}), R(p, q, \bar{X})) \mid p \text{ p.d. on } X, q \text{ p.d. on } Y\}. \tag{2.7}$$

Let $F^*(\bar{X})$ be the convex hull of $F(\bar{X})$. Its elements are triples (R_1^*, R_2^*, R^*), which shall be denoted by $\mathbf{R}$. Finally, we define subsets $G(\mathbf{R}, \bar{X})$ and $G(\bar{X})$ of the Euclidean plane by

$$G(\mathbf{R}, \bar{X}) = \{(R_1, R_2) \mid \textstyle\sum_{s=1}^{2} R_s \leqq R^*, R_s \leqq R_s^* \text{ for } s = 1, 2\} \tag{2.8}$$

$$G(\bar{X}) = \textstyle\bigcup_{\mathbf{R} \in F^*(\bar{X})} G(\mathbf{R}, \bar{X}). \tag{2.9}$$

THEOREM 1.

$$G(P, T_{21}, I) = G(\bar{X}).$$

PROOF. The relationship $G(P, T_{21}, I) \subset G(\bar{X})$ is an immediate consequence of the inequalities (2.5) and (2.6).

We show now the converse relationship. As in [1] we again make use of Shannon's random coding method. The differences of the present approach as compared with the one given in [1] consist in the choice of the "source probabilities" and in the decoding employed. Here we admit "nonstationary sources" and we apply maximum likelihood decoding.

Let $U_1, \cdots, U_{N_1}$ be independent identically distributed random variables with values in X_n and distribution p_n, that is, $p_n(U_i = u_i) = p_n(u_i)$. Let $V_n, \cdots, V_{N_2}$ be identically distributed random variables with distribution q_n. The V_j's are assumed to be independent of each other and of the U_i's. Write $\hat{U}$ for $(U_1, \cdots, U_{N_1})$, $\hat{V}$ for $(V_1, \cdots, V_{N_2})$, $\hat{u}$ for $(u_1, \cdots, u_{N_1})$ and $\hat{v}$ for $(v_1, \cdots, v_{N_2})$. Denote the joint distribution of the U_i's by $\hat{p}_n$, the joint distribution of the V_j's by $\hat{q}_n$ and the joint distribution of all random variables by $\hat{p}_n \times \hat{q}_n$. For the outcome $(\hat{u}, \hat{v})$ of the random experiment $(\hat{U}, \hat{V}, \hat{p} \times \hat{q})$ we define decoding sets $A_{ij} = A_{ij}(\hat{u}, \hat{v})$ $(i = 1, \cdots, N_1;\ j = 1, \cdots, N_2)$ by

$$(2.10) \qquad A_{ij} = \{\bar{x}_n \mid P(\bar{x}_n \mid u_i, v_j) > P(\bar{x}_n \mid u_k, v_l) \text{ for } (k, l) \neq (i, j)\} .$$

The average error for the code $\{(u_i, v_j, A_{ij}) \mid i = 1, \cdots, N_1;\ j = 1, \cdots, N_2\}$ is given by

$$(2.11) \qquad \lambda(\hat{u}, \hat{v}) = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} P(A_{ij}^c \mid u_i, v_j) .$$

We want to give an upper bound on the expected error

$$E\lambda(\hat{U}, \hat{V}) = \sum_{\hat{u},\hat{v}} \hat{p}(\hat{u})\hat{q}(\hat{v})\lambda(\hat{u}, \hat{v}) .$$

Abbreviate $P(A_{ij}^c \mid u_i, v_j)$ as $\lambda_{ij}(\hat{u}, \hat{v})$. For reasons of symmetry one has $E\lambda(\hat{U}, \hat{V}) = E\lambda_{11}(\hat{U}, \hat{V})$. Denoting $p_n(u_1)q_n(v_1)P(\bar{x}_n \mid u_1, v_1)$ by $\tilde{P}(\bar{x}_n, u_1, v_1)$ we can write $E\lambda_{11}(\hat{U}, \hat{V})$ as

$$(2.12) \qquad \sum_{\bar{x}_n, u_1, v_1} \tilde{P}(\bar{x}_n, u_1, v_1)\hat{p} \times \hat{q} \times \{P(\bar{x}_n \mid u_1, v_1) \leqq P(\bar{x}_n \mid U_i, V_j) \text{ for some } (i, j) \neq (1, 1)\} .$$

This expression is smaller than the sum of the following terms:

$$(2.13) \qquad \sum_{\bar{x}_n, u_1, v_1} \tilde{P}(\bar{x}_n, u_1, v_1)\hat{p} \times \hat{q} \times \{P(\bar{x}_n \mid u_1, v_1) \leqq P(\bar{x}_n \mid U_i, v_1) \text{ for some } i \neq 1\} ;$$

$$(2.14) \qquad \sum_{\bar{x}_n, u_1, v_1} \tilde{P}(\bar{x}_n, u_1, v_1)\hat{p} \times \hat{q} \times \{P(\bar{x}_n \mid u_1, v_1) \leqq P(\bar{x}_n \mid u_1, V_j) \text{ for some } j \neq 1\} ;$$

$$(2.15) \qquad \sum_{\bar{x}_n, u_1, v_1} \tilde{P}(\bar{x}_n, u_1, v_1)\hat{p} \times \hat{q}\{P(\bar{x}_n \mid u_1, v_1) \leqq P(\bar{x}_n \mid U_i, V_j) \text{ for some } (i, j) : i \neq 1,\ j \neq 1\} .$$

In order to give a bound on the first term we introduce for any $\alpha > 0$ and for $v_1 \in Y_n$ the set $B_{v_1} = \{(\bar{x}_n, u_1) \mid I^1(\bar{x}_n, u_1 \mid v_1) > \log \alpha N_1\}$. Writing $\tilde{P}(\bar{x}_n, u_1, v_1)$ as $\tilde{P}_{v_1}(\bar{x}_n, u_1) \cdot q_n(v_1)$ and using the fact that the U_i's are identically distributed we obtain the bound

$$\sum_{v_1} q_n(v_1)[\tilde{P}_{v_1}(B_{v_1}^c) + N_1 \sum_{(\bar{x}_n, u_1) \in B_{v_1}} \tilde{P}_{v_1}(\bar{x}_n, u_1)\hat{p} \times \hat{q}\{P(\bar{x}_n \mid u_1, v_1) \leqq P(\bar{x}_n \mid U_2, V_1)\}] .$$

It follows from the definition of B_{v_1} that the last term in brackets is smaller than α^{-1}. Since $\sum_{\bar{x}_n, u_1, v_1} \tilde{P}(\bar{x}_n, u_1, v_1) \log P(\bar{x}_n \mid u_1, v_1) / \sum_{v_1} p_n(u_1)P(\bar{x}_n \mid u_1, v_1)$ equals $\sum_{t=1}^{n} R_1(p^t, q^t, \bar{X})$, $\sum_{v_1} q_n(v_1)\tilde{P}_{v_1}^2(B_{v_1}^c)$ can be made arbitrarily small by choosing αN_1 smaller than $\exp\{\sum_{t=1}^{n} R_1(p^t, q^t, \bar{X}) - kn^{\frac{1}{2}}\}$, where k is a suitable constant.

Thus, the whole term can be made arbitrarily small by choosing α sufficiently large. By the same argument we can establish a bound on (2.14). Using a set $B = \{(\bar{x}_n, u_1, v_1) \mid I(\bar{x}_n, u_1, v_1) > \log \alpha N_1 N_2\}$ the third term can be made small by choosing $\alpha N_1 N_2 \leqq \sum_{t=1}^n R(p^t, q^t, \bar{X}) - kn^{\frac{1}{2}}$. The estimate is exactly the one of the random coding method [3]. The proof is complete.

3. The capacity region $G(P, T_{22}, II)$. Let $\mathscr{P}$ be a finite set of pairs (p, q) and let μ be a p.d. on $\mathscr{P}$. With $(\mathscr{P}, \mu)$ we associate a vector $\mathbf{R}(\mathscr{P}, \mu)$, given by

$$(3.1) \qquad \mathbf{R}(\mathscr{P}, \mu) = \min\{\textstyle\sum_{p,q} \mu(p, q)[R_1(p, q, \bar{X}), R_2(p, q, \bar{X}), R(p, q, \bar{X})], \sum_{p,q} \mu(p, q)[R_1(p, q, \bar{Y}), R_2(p, q, \bar{Y}), R(p, q, \bar{Y})]\}\,.$$

(It is understood that the minimum is taken componentwise.)

Set $F(\bar{X}, \bar{Y}) = \{\mathbf{R} \mid \mathbf{R} = \mathbf{R}(\mathscr{P}, \mu)$ for some $(\mathscr{P}, \mu)\}$ and write its elements as $\mathbf{R} = (\tilde{R}_1, \tilde{R}_2, \tilde{R})$. Analogously to (2.8) and (2.9) we define now subsets $G(\mathbf{R})$ and G of the Euclidean plane by

$$(3.2) \qquad G(\mathbf{R}) = \{(R_1, R_2) \mid \textstyle\sum_{s=1}^2 R_s \leqq \tilde{R},\ R_s \leqq \tilde{R}_s \ \text{ for } s = 1, 2\}$$

and

$$(3.3) \qquad G = \textstyle\bigcup_{\mathbf{R} \in F(\bar{X}, \bar{Y})} G(\mathbf{R})\,.$$

THEOREM 2.

$$G(P, T_{22}, II) = G\,.$$

PROOF. We show first that $G(P, T_{22}, II) \subset G$. Let $\{(u_i, v_j, A_{ij}, B_{ij}) \mid i = 1, \cdots, N_1;\ j = 1, \cdots, N_2\}$ be an $(n, N_1, N_2, \lambda, II)$ code for (P, T_{22}, II). Replacing $P(A_{ij}^c \mid u_i, v_j)$ by $P(A_{ij}^c \mid u_i, v_j) + Q(B_{ij}^c \mid u_i, v_j)$ in the derivation which led to (2.1) and (2.2) we obtain from (1.5) that—after renumbering of the u_i's and the v_j's—the following inequalities hold:

$$(3.4) \qquad \frac{1}{N_2^*} \textstyle\sum_{j=1}^{N_2^*} P(A_{ij}^c \mid u_i, v_j) \leqq 4\lambda \qquad \text{for } i = 1, \cdots, N_1^* = \left[\frac{N_1}{2}\right].$$

$$(3.5) \qquad \frac{1}{N_2^*} \textstyle\sum_{j=1}^{N_2^*} Q(B_{ij}^c \mid u_i, v_j) \leqq 4\lambda \qquad \text{for } i = 1, \cdots, N_1^*\,.$$

$$(3.6) \qquad \frac{1}{N_1^*} \textstyle\sum_{i=1}^{N_1^*} P(A_{ij}^c \mid u_i, v_j) \leqq 4\lambda \qquad \text{for } j = 1, \cdots, N_2^* = \left[\frac{N_2}{2}\right].$$

$$(3.7) \qquad \frac{1}{N_1^*} \textstyle\sum_{i=1}^{N_1^*} Q(B_{ij}^c \mid u_i, v_j) \leqq 4\lambda \qquad \text{for } j = 1, \cdots, N_2^*\,.$$

For $t = 1, 2, \cdots, n$ define p^t and q^t as in (2.3) and (2.4). Using (2.5) and (2.6) for $P(\cdot \mid \cdot, \cdot)$ and for $Q(\cdot \mid \cdot, \cdot)$ we obtain the following inequalities:

$$(3.8) \qquad \log N_1 N_2 \leqq [\min\{\textstyle\sum_{t=1}^n R(p^t, q^t, \bar{X}), \sum_{t=1}^n R(p^t, q^t, \bar{Y})\} + 1] \times (1 - 4\lambda)^{-1} + \log 4\,.$$

$$(3.9) \qquad \log N_s \leqq [\min\{\textstyle\sum_{t=1}^n R_s(p^t, q^t, \bar{X}), \sum_{t=1}^n R_s(p^t, q^t, \bar{Y})\} + 1] \times (1 - 4\lambda)^{-1} + \log 4 \qquad \text{for } s = 1, 2\ (0 < \lambda < \tfrac{1}{4})\,.$$

The relationship $G(P, T_{22}, II) \subset G$ is a consequence of (3.8) and (3.9). In order to prove the converse relationship we select code words at random as in Section 2. $R_{\bar{X}}$ uses the decoding sets A_{ij} defined in (2.10) and $R_{\bar{Y}}$ uses decoding sets $B_{ij} = \{\bar{x}_n \mid Q(\bar{x}_n \mid u_i, v_j) > Q(\bar{x}_n \mid u_k, v_l) \text{ for } (k, l) \neq (i, j)\}$ $(i = 1, \cdots, N_1;\ j = 1, \cdots, N_2)$. The expected error probabilities in both decoding systems are small, if

$$\log N_s \leqq \min \{\textstyle\sum_{t=1}^n R_s(p^t, q^t, \bar{X}),\ \sum_{t=1}^n R_s(p^t, q^t, \bar{Y})\} - kn^{\frac{1}{2}} \qquad \text{for } s = 1, 2$$

and $\log N_1 N_2 \leqq \min \{\sum_{t=1}^n R(p^t, q^t, \bar{X}),\ \sum_{t=1}^n R(p^t, q^t, \bar{Y})\} - kn^{\frac{1}{2}}$.

Hence, there exist codes achieving all pairs of rates in G.

4. A remark about the capacity region $G(P, T_{22}, I)$. Lemma 1 of [1] gives a characterization of $G(P, T_{22}, I)$ in terms of dependent sources. This characterization is unsatisfactory, because it cannot be used to compute $G(P, T_{22}, I)$. Until now no satisfactory result exists. It was conjectured on page 636 of [4] that $G(P, T_{22}, I)$ equals the convex hull of the set $\{R_{21}^2(p, q, \bar{X}), R_{12}^1(p, q, \bar{Y}) \mid p \text{ p.d. on } X, q \text{ p.d. on } Y\}$. We give here a counter-example. First of all we notice that (P, T_{21}, I) can be viewed as a special case of (P, T_{22}, I). It is the case in which $\bar{X} = \bar{Y}$ and $w(\bar{x}, \bar{y} \mid x, y) = 0$ for $\bar{x} \neq \bar{y}$. Here both receivers receive always the same letters and can be identified. It suffices therefore to construct an example of a channel for which $G(P, T_{21}, I)$ is unequal to the convex hull of

$$\{(R_{21}^2(p, q, \bar{X}), R_{12}^1(p, q, \bar{X}) \mid p \ \text{ p.d. on } \ X, q \ \text{ p.d. on } \ Y\}.$$

Choose $X = Y = \{0, 1\}$, $\bar{X} = \{0, 1, 2\}$ and choose for $p(\cdot \mid \cdot, \cdot)$ the following matrix:

$$\begin{array}{c} \\ 00 \\ 01 \\ 10 \\ 11 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{array}.$$

One easily verifies that

$$R_{12}^1(p, q, \bar{X}) + R_1(p, q, \bar{X}) = R(p, q, \bar{X}), \tag{4.1}$$

$$R_{21}^2(p, q, \bar{X}) + R_2(p, q, \bar{X}) = R(p, q, \bar{X}), \tag{4.2}$$

and—by the well-known convexity properties of the rate function—also that

$$R_{12}^1(p, q, \bar{X}) \leqq R_2(p, q, \bar{X}), \tag{4.3}$$

$$R_{21}^2(p, q, \bar{X}) \leqq R_1(p, q, \bar{X}). \tag{4.4}$$

(4.1) and (4.3) imply that

$$R_1(p, q, \bar{X}) + R_2(p, q, \bar{X}) \geqq R(p, q, \bar{X}). \tag{4.5}$$

Theorem 1, (4.3), (4.4), and (4.5) yield that the sets in question are unequal, if the inequality $\max_{p,q} [R_{12}^1(p, q, \bar{X}) + R_{21}^2(p, q, \bar{X})] < \max_{p,q} R(p, q)$ holds. Straightforward computation shows that

$$R_1(p, q, \bar{X}) = H(p), \quad \text{where } H \text{ denotes the entropy function,} \tag{4.6}$$

and that

(4.7) $$R^2_{21}(p, q, \bar{X}) \leqq H(p)\,,$$ with equality holding exactly if $q(0)$ equals 0 or 1.

(4.1), (4.6) and (4.7) imply that

(4.8) $$R^1_{12}(p, q, \bar{X}) + R^2_{21}(p, q, \bar{X}) < R(p, q)$$ for all p and all q with $q(0) \neq 0, 1$.

With p^* and q^* as equal distribution on $\{0, 1\}$ one obtains $R(p^*, q^*) > \log 2$ and thus

(4.9) $$\max_{p,q} R(p, q) > \log 2\,.$$

Since $\max_{p,q:q(0)=0 \text{ or } 1} R(p, q) = \log 2$, we have for all (p, q) $R^1_{12}(p, q, \bar{X}) + R^2_{21}(p, q, \bar{X}) < \max_{p,q} R(p, q)$; as was to be shown.

5. A generalization of Shannon's random coding method. Let X^1, X^2, $\bar{X}^1$ and $\bar{X}^2$ be finite sets, $p^1(\cdot \mid \cdot)$ a transmission matrix from X^1 to $\bar{X}^1$ and $p^2(\cdot \mid \cdot)$ a transmission matrix from X^2 to $\bar{X}^2$. Set $X_2 = X^1 \times X^2$, $\bar{X}_2 = \bar{X}^1 \times \bar{X}^2$ and define P by

(5.1) $$P(\bar{x}_2 \mid x_2) = \prod_{t=1}^2 p^t(\bar{x}^t \mid x^t)\,; \qquad x_2 = (x^1, x^2) \in X_2\,, \quad \bar{x}_2 = (\bar{x}^1, \bar{x}^2) \in \bar{X}_2\,.$$

Furthermore, for $t = 1, 2$ let p^t be a p.d. on X^t, let $\bar{p}^t = p^t \cdot p^t(\cdot \mid \cdot)$ be a p.d. on $\bar{X}^t$, let $\tilde{p}^t = p^t * p^t(\cdot \mid \cdot)$ be a p.d. on $X^t * \bar{X}^t$ and finally set $p_2 = p^1 \times p^2$, $\bar{p}_2 = \bar{p}^1 \times \bar{p}^2$, and $\tilde{p}_2 = \tilde{p}^1 \times \tilde{p}^2$. For $x_2 \in X_2$, $\bar{x}_2 \in \bar{X}_2$ we define

(5.2) $$I(x_2, \bar{x}_2) = \log \frac{P(\bar{x}_2 \mid x_2)}{\bar{p}_2(\bar{x}_2)}$$

and for $x^2 \in X^2$, $\bar{x}^2 \in \bar{X}^2$ we set

(5.3) $$I(x^2, \bar{x}^2) = \log \frac{p^2(\bar{x}^2 \mid x^2)}{\bar{p}^2(\bar{x}^2)}\,.$$

Let $U_1, \cdots, U_{N_1}$ be independent identically distributed random variables with values in X^1 and with distribution p^1. Furthermore, let V_{ij}, $i = 1, \cdots, N_1, j = 1, \cdots, N_2$; be identically distributed random variables with values in X^2 and with distribution p^2. The V_{ij}'s are assumed to be independent of each other and of the U_i. Denote the joint distribution of all random variables by $\hat{p}$. Let $W_{ij} = (U_i, V_{ij})$ and define

(5.4) $$A_{ij} = \{\bar{x}_2 \mid P(\bar{x}_2 \mid W_{ij}) > P(\bar{x}_2 \mid W_{kl}) \text{ for all } (k, l) \neq (i, j)\}\,.$$

The system $\{W_{ij}, A_{ij}) \mid i = 1, \cdots, N_1; j = 1, \cdots, N_2\}$ is a code of length $N = N_1 \cdot N_2$ for the channel P. Its error probability is $N^{-1} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} P(A^c_{ij} \mid W_{ij})$. If we select the code words according to the random variables described above and apply maximum likelihood decoding, then we obtain an expected error probability

$$E\left[\frac{1}{N} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} P(A^c_{ij} \mid W_{ij})\right].$$

LEMMA (Generalized random coding).

$$E\left[\frac{1}{N}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} P(A_{ij}^c \mid W_{ij})\right] \leqq \tilde{P}_2\{(x_2, \bar{x}_2) \mid I(x_2, \bar{x}_2) \leqq \log \alpha N\}$$
$$+ \tilde{P}\{(x^2, \bar{x}^2) \mid I(x^2, \bar{x}^2) \leqq \log \alpha_2 N_2\}$$
$$+ \frac{1}{\alpha} + \frac{1}{\alpha_2} \qquad \textit{for } \alpha, \alpha_2 > 0 .$$

PROOF. For reasons of symmetry the expected error probability equals $EP(A_{11}^c \mid W_{11})$. Define sets B and B_2 by

$$B = \{(x_2, \bar{x}_2) \mid I(x_2, \bar{x}_2) > \log \alpha N\} \tag{5.5}$$

and

$$B_2 = \{(x^2, \bar{x}^2) \mid I(x^2, \bar{x}^2) > \log \alpha_2 N_2\} . \tag{5.6}$$

Since $EP(A_{11}^c \mid W_{11}) \leqq \sum_{\bar{x}_2, w_{11}} p_2(w_{11}) P(\bar{x}_2 \mid w_{11})[\hat{p}\{P(\bar{x}_2 \mid w_{11}) \leqq P(\bar{x}_2 \mid W_{ij})$ for some j and $i \neq 1\} + \hat{p}\{P(\bar{x}_2 \mid w_{11}) \leqq P(\bar{x}_2 \mid W_{ij})$ for some $j \neq 1\}$ we can conclude that

$$\begin{aligned} EP(A_{11}^c \mid W_{11}) &\leqq \tilde{p}_2(B^c) + \textstyle\sum_{(w_{11}, \bar{x}_2) \in B} \tilde{p}_2(w_{11}, \bar{x}_2) N \hat{p}\{P(\bar{x}_2 \mid w_{11}) \leqq P(\bar{x}_2 \mid W_{21})\} \\ &\quad + \tilde{p}^2(B_2^c) + \textstyle\sum_{(v_{11}, \bar{x}^2) \in B_2} \tilde{p}^2(v_{11}, \bar{x}^2) \\ &\quad \times N_2 \hat{p}\{p^2(\bar{x}^2 \mid v_{11}) \leqq p^2(\bar{x}^2 \mid V_{12})\} . \end{aligned} \tag{5.7}$$

For $(w_{11}, \bar{x}_2) \in B$ we have:

$$\begin{aligned} \hat{p}\{P(\bar{x}_2 \mid w_{11}) \leqq P(\bar{x}_2 \mid W_{21})\} &= p_2\{x_2 \mid P(\bar{x}_2 \mid w_{11}) \leqq P(\bar{x}_2 \mid x_2)\} \\ &= p_2\{x_2 \mid I(w_{11}, \bar{x}_2) \leqq I(x_2, \bar{x}_2)\} \\ &\leqq p_2\{x_2 \mid \log \alpha N \leqq I(x_2, \bar{x}_2)\} \\ &= P_2\left\{x_2 \mid \alpha N p_2(x_2) \leqq \frac{\tilde{p}^2(x_2, \bar{x}_2)}{\tilde{p}_2(\bar{x}_2)}\right\} \leqq \frac{1}{\alpha N} . \end{aligned}$$

By exactly the same arguments one obtains

$$P\{p^2(\bar{x}^2 \mid v_{11}) \leqq p^2(\bar{x}^2 \mid V_{12})\} \leqq \frac{1}{\alpha_2 N_2} .$$

These two bounds and (5.7) yield the statement of the lemma.

REMARK 1. One obtains Shannon's result by choosing $N_2 = 1$.

REMARK 2. Let us consider a discrete memoryless channel with alphabets X and $\bar{X}$.

Let $R(p)$ be the rate for the source probability p. Choose $X^1 = X_{n_1}$, $\bar{X}^1 = \bar{X}_{n_1}$, $X^2 = X_{n_2}$, $\bar{X}^2 = \bar{X}_{n_2}$, $X^1 * X^2 = X_n$ and $\bar{X}^1 * \bar{X}^2 = \bar{X}_n$, where $n = n_1 + n_2$. Application of the lemma yields the following result: for every λ, $0 < \lambda < 1$, for all n and $n_2 \leqq n$, and for every nonnegative integer N_2 satisfying $N_2 \leqq \exp\{R(p)n_2 + k(\lambda)n_2^{\frac{1}{2}}\}$ there exists a code $\{(u_i, v_{ij}, A_{ij}) \mid i = 1, \cdots, N_1; j = 1, \cdots, N_2\}$, (where $u_i \in X_{n_1}$, $v_{ij} \in X_{n_2}$, $A_{ij} \subset \bar{X}_n$) with an error probability smaller than λ and a length

$$N = N_1 N_2 \geqq \exp\{R(p)n - k(\lambda)n^{\frac{1}{2}}\} .$$

($k(\lambda)$ is a known function independent of n.) It seems to the author that the existence of such codes should be of interest for coding problems of certain multi-way channels, especially those involving time sharing.

REMARK 3. The lemma easily can be generalized to more than two components.

REFERENCES

[1] AHLSWEDE, R. (1971). Multi-way communication channels. *Second International Symposium on Information Theory*. Publishing House of the Hungarian Academy of Sciences. 23–52.

[2] FANO, R. M. (1952, 1954). Statistical theory of communication. MIT lecture notes.

[3] SHANNON, C. E. (1957). Certain results in coding theory for noisy channels. *Information and Control* 7 6–25.

[4] SHANNON, C. E. (1962). Two-way communication channels. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* **1** 611–644.

[5] SLEPIAN, D. and WOLF, J. K. ( ). A coding theorem for multiple access channels with correlated sources. To appear in *Bell System Tech. J.*

[6] ULREY, M. L. A coding theorem for a channel with s senders and r receivers. To appear in *Information and Control.*

[7] VAN DER MEULEN, E. C. (1973). On a problem by Ahlswede regarding the capacity region of certain multi-way channels. To appear in *Information and Control.*

DEPARTMENT OF MATHEMATICS
THE OHIO STATE UNIVERSITY
231 WEST 18TH AVENUE
COLUMBUS, OHIO 43210