

Multiple Channels under Fidelity Criteria*

BHU DEV SHARMA, VED PRIYA

Communication under fidelity criterion was introduced by Shannon. The problems concerning Multiple Channels have been the focus of recent interest. However fidelity criterion has not been considered in the studies for Multiple Channels. In this paper an attempt has been made to define rate distortion function for some special cases of Multiple Channels, viz., the Broadcast Channel, the Two-User Channel and the Multiple Access Channel. Basic equations for these channels are derived and convexity of the rate distortion functions is established. The investigations are then extended to the case of a general channel involving several sources and destinations. For the Two-User Channel and the Multiple Access Channel, examples have also been formulated.

1. INTRODUCTION

Areas of recent interest in Communication Theory are the transmission of information in a Multiple Access Channel introduced by Liao [4] and a Broadcast Channel introduced by Cover [2]. In a Multiple Access Channel several sources communicate with one receiver over a common channel. The message output from any source is assumed to be independent of the message outputs from other sources. Liao [4] defined capacity region and proved a coding theorem and its converse for such a discrete memoryless channel. Cover [2] introduced Broadcast Channels in which one source communicates with several receivers and obtained upper and lower bounds on the capacity region. Vander Meulen [8] obtained an inner bound to the General Broadcast Channels for the three communication situations and Sato [5] obtained an outer bound to the capacity region of Broadcast Channels. Shannon [7] was the first to introduce the idea of a Two-Way Communication Channel which involves sending information simultaneously in two directions over a Two-Way

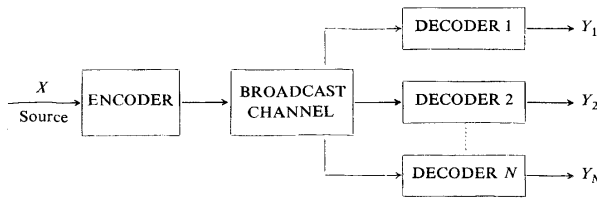
* This work was supported by a Research Fellowship awarded to the second author by the Council of Scientific and Industrial Research, New Delhi, India.

Channel and obtained inner and outer bounds to the capacity region of this channel. All these are cases of "Multiple Channels".

Study relating to communication under fidelity criteria which have been earlier modified for classical channels has not yet been extended for the case of Multiple Channels. In this paper we define rate distortion function for these channels. In Section 2 we determine the basic equations for Broadcast Channel. Convexity of the rate distortion function is established in Section 3. Section 4 deals with the basic equations for a Two-User Channel for which an example is formulated in Section 5. In Section 6, basic equations for a Multiple Access Channel are derived and an illustration for the same is presented in Section 7. The last section deals with a general model for M sources and N receivers which under certain conditions reduces to the cases studied in earlier sections.

2. BASIC EQUATIONS FOR BROADCAST CHANNEL

A general Broadcast Channel with N receivers is shown in the diagram below. There is one source which is denoted by X and there are N receivers which we denote by Y_1, Y_2, \dots, Y_N . The memoryless Broadcast Channel with one input and N outputs may be characterised by $(X, Q(y_1, y_2, \dots, y_N | x), Y_1 \times Y_2 \times \dots \times Y_N)$ where $Q(y_1, y_2, \dots, y_N | x)$ is the probability of receiving $y_1 \in Y_1, \dots, y_N \in Y_N$ when $x \in X$ is sent on the channel. Further $Q_i(y_i | x)$ is the transition probability of receiving $y_i \in Y_i$ by the i -th receiver ($i = 1, 2, \dots, N$) when $x \in X$ is transmitted.



Broadcast Channel

Now

$$Q_i(y_i | x) = \sum_{y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_N} Q(y_1, y_2, \dots, y_N | x) \quad (i = 1, 2, \dots, N)$$

and since the outputs are statistically independent, we have

$$Q(y_1, y_2, \dots, y_N | x) = \prod_{i=1}^N Q_i(y_i | x).$$

Further, let the distortion between the source letter $x \in X$ and the reproduced letter $y_i \in Y_i$ ($i = 1, 2, \dots, N$) in the i -th output be denoted by $\varrho_i(x, y_i)$, where as is usual

$$\varrho_i(x, y_i) \geq 0 \quad (i = 1, 2, \dots, N)$$

with equality iff $x = y_i$.

If $P(x)$ is taken to denote the input probability of $x \in X$, then the average distortion for the i -th output may as usual be defined as

$$\sum_{x, y_1, y_2, \dots, y_N} P(x) Q(y_1, y_2, \dots, y_N | x) \varrho_i(x, y_i) \quad (i = 1, 2, \dots, N).$$

If we communicate on the Broadcast Channel in such a way that the level of average distortion between the source and the i th reproduced letter does not exceed a given level D_i ($i = 1, 2, \dots, N$), then the rate distortion function $R_X(D_1, D_2, \dots, D_N)$ for the Broadcast Channel may be defined as

$$(1) \quad R_X(D_1, D_2, \dots, D_N) = \inf_{Q(y_1, \dots, y_N | x) \in Q_{D_1, D_2, \dots, D_N}} I(X; Y_1, Y_2, \dots, Y_N),$$

where

$$(2) \quad I(X; Y_1, Y_2, \dots, Y_N) = \sum_{x, y_1, \dots, y_N} P(x, y_1, y_2, \dots, y_N) \log \frac{Q(y_1, \dots, y_N | x)}{Q(y_1, \dots, y_N)}$$

is the ordinary Shannon's mutual information and

$$(3) \quad Q_{D_1, \dots, D_N} = \{Q(y_1, \dots, y_N | x) : \sum_{x, y_1, \dots, y_N} P(x) Q(y_1, \dots, y_N | x) \varrho_i(x, y_i) \leq D_i\}, \\ i = 1, 2, \dots, N.$$

The $I(X; Y_1, Y_2, \dots, Y_N)$ may be shown to be a convex \cup function with respect to transition probabilities $Q(y_1, y_2, \dots, y_N | x)$ as follows:

Consider two sets of transition probabilities $\{Q'(y_1, \dots, y_N | x)\}$ and $\{Q''(y_1, \dots, y_N | x)\}$ and a number $\lambda \in [0, 1]$ and let

$$Q^*(y_1, y_2, \dots, y_N | x) = \lambda Q'(y_1, y_2, \dots, y_N | x) + (1 - \lambda) Q''(y_1, y_2, \dots, y_N | x).$$

Then

$$(4) \quad I(Q^*(y_1, y_2, \dots, y_N | x)) = \sum_{x, y_1, \dots, y_N} P(x) [\lambda Q'(y_1, \dots, y_N | x) + (1 - \lambda) Q''(y_1, \dots, y_N | x)] \cdot \log \frac{\lambda Q'(y_1, y_2, \dots, y_N | x) + (1 - \lambda) Q''(y_1, y_2, \dots, y_N | x)}{\lambda Q'(y_1, y_2, \dots, y_N) + (1 - \lambda) Q''(y_1, y_2, \dots, y_N)}.$$

Now we use the inequality

$$\log \frac{a+b}{a} \leq \frac{a+b}{a} - 1, \quad a > 0, \quad b \geq 0,$$

i.e.

$$(5) \quad \log(a+b) \leq \log a + b/a$$

with equality iff $b = 0$. Let us set

$$(6) \quad a_1 = \frac{Q'(y_1, y_2, \dots, y_N | x)}{Q'(y_1, y_2, \dots, y_N)}, \quad a_2 = \frac{Q''(y_1, y_2, \dots, y_N | x)}{Q''(y_1, y_2, \dots, y_N)},$$

$$b_1 = \frac{(1-\lambda)[Q'(y_1, \dots, y_N) Q''(y_1, \dots, y_N | x) - Q''(y_1, \dots, y_N) Q'(y_1, \dots, y_N | x)]}{Q'(y_1, \dots, y_N) [\lambda Q'(y_1, \dots, y_N) + (1-\lambda) Q''(y_1, \dots, y_N)]},$$

$$b_2 = \frac{\lambda[Q''(y_1, \dots, y_N) Q'(y_1, \dots, y_N | x) - Q'(y_1, \dots, y_N) Q''(y_1, \dots, y_N | x)]}{Q''(y_1, \dots, y_N) [\lambda Q'(y_1, \dots, y_N) + (1-\lambda) Q''(y_1, \dots, y_N)]}.$$

Thus from (4), (5) and (6), we have

$$I(Q^*(y_1, y_2, \dots, y_N | x)) \leq \lambda \sum_{x, y_1, y_2, \dots, y_N} P(x) Q'(y_1, y_2, \dots, y_N | x) \cdot$$

$$\cdot \left[\log \frac{Q'(y_1, \dots, y_N | x)}{Q'(y_1, \dots, y_N)} + \right.$$

$$+ \frac{(1-\lambda)[Q''(y_1, \dots, y_N | x) Q'(y_1, \dots, y_N) - Q''(y_1, \dots, y_N) Q'(y_1, y_2, \dots, y_N | x)]}{Q'(y_1, \dots, y_N | x) [\lambda Q'(y_1, \dots, y_N) + (1-\lambda) Q''(y_1, \dots, y_N)]} \Big] +$$

$$+ (1-\lambda) \sum_{x, y_1, \dots, y_N} P(x) Q''(y_1, \dots, y_N | x) \cdot \left[\log \frac{Q''(y_1, \dots, y_N | x)}{Q''(y_1, \dots, y_N)} + \right.$$

$$+ \frac{\lambda[Q''(y_1, \dots, y_N) Q'(y_1, \dots, y_N | x) - Q'(y_1, \dots, y_N) Q''(y_1, \dots, y_N | x)]}{Q''(y_1, \dots, y_N | x) [\lambda Q'(y_1, \dots, y_N) + (1-\lambda) Q''(y_1, \dots, y_N)]} \Big] =$$

$$= \lambda I(Q'(y_1, \dots, y_N | x)) + (1-\lambda) I(Q''(y_1, \dots, y_N | x)).$$

Hence $I(X; Y_1, Y_2, \dots, Y_N)$ is a convex U function with respect to the transition probabilities $Q(y_1, y_2, \dots, y_N | x)$.

Thus our problem is to minimize $I(X; Y_1, Y_2, \dots, Y_N)$ subject to the constraints:

$$(7) \quad Q(y_1, y_2, \dots, y_N | x) \geq 0,$$

$$(8) \quad \sum_{y_1, y_2, \dots, y_N} Q(y_1, y_2, \dots, y_N | x) = 1$$

450 and

$$(9) \quad \sum_{x, y_1, y_2, \dots, y_N} P(x) Q(y_1, y_2, \dots, y_N | x) \varrho_i(x, y_i) = D_i \\ (i = 1, 2, \dots, N).$$

We solve this problem by Lagrange method of multipliers. Ignoring the constraint (7) temporarily, we form the augmented function

$$(10) \quad J(Q) = I(X; Y_1, Y_2, \dots, Y_N) - \sum_x \mu_x \sum_{y_1, y_2, \dots, y_N} Q(y_1, y_2, \dots, y_N | x) - \\ - \sum_{i=1}^N S_i \sum_{x, y_1, y_2, \dots, y_N} P(x) Q(y_1, y_2, \dots, y_N | x) \varrho_i(x, y_i),$$

where μ_x and S_i ($i = 1, 2, \dots, N$) are Lagrange multipliers. Taking $\log \lambda_x = \mu_x / P(x)$ and using (2) in (10), we may rewrite (10) as

$$J(Q) = \sum_{x, y_1, y_2, \dots, y_N} P(x) Q(y_1, \dots, y_N | x) \left[\log \frac{Q(y_1, \dots, y_N | x)}{Q(y_1, \dots, y_N) \lambda_x} - \sum_{i=1}^N S_i \varrho_i(x, y_i) \right].$$

Now, for stationary points, we have

$$\frac{dJ}{dQ(y_1, \dots, y_N | x)} = P(x) \left[\log \frac{Q(y_1, \dots, y_N | x)}{Q(y_1, \dots, y_N) \lambda_x} - \sum_{i=1}^N S_i \varrho_i(x, y_i) \right] = 0,$$

i.e.

$$(11) \quad Q(y_1, y_2, \dots, y_N | x) = \lambda_x Q(y_1, y_2, \dots, y_N) \exp \left[\sum_{i=1}^N S_i \varrho_i(x, y_i) \right].$$

Summing (11) over y_1, y_2, \dots, y_N and using (8), we get

$$(12) \quad \lambda_x = \left[\sum_{y_1, \dots, y_N} Q(y_1, y_2, \dots, y_N) \exp \left[\sum_{i=1}^N S_i \varrho_i(x, y_i) \right] \right]^{-1}.$$

Thus from (9) and (11), we have

$$(13) \quad D_i = \sum_{x, y_1, \dots, y_N} \varrho_i(x, y_i) P(x) \lambda_x Q(y_1, y_2, \dots, y_N) \exp \left[\sum_{i=1}^N S_i \varrho_i(x, y_i) \right] \\ (i = 1, 2, \dots, N)$$

and

$$I(X; Y_1, Y_2, \dots, Y_N) = \sum_{x, y_1, \dots, y_N} P(x) \lambda_x Q(y_1, y_2, \dots, y_N) \exp \left[\sum_{i=1}^N S_i \varrho_i(x, y_i) \right] \cdot \\ \cdot \left[\log \lambda_x + \sum_{i=1}^N S_i \varrho_i(x, y_i) \right].$$

Thus $R_X(D_1, D_2, \dots, D_N)$ which in view of convexity of $I(X; Y_1, Y_2, \dots, Y_N)$ is the minimum of $I(X; Y_1, Y_2, \dots, Y_N)$ has the parametric representation

$$(14) \quad R_X(D_1, D_2, \dots, D_N) = \sum_{i=1}^N S_i D_i + \sum_x P(x) \log \lambda_x,$$

where λ_x is given by (12). Expressions (13) and (14) give the required form of the basic equations for Broadcast Channel.

Now if for a particular value of S_i ($i = 1, 2, \dots, N$), the unconstrained solution procedure yields one or more $Q(y_1, \dots, y_N | x) \leq 0$ then the results can be formulated as in Berger [1, Lemma 1, p. 32].

3. CONVEXITY OF THE FUNCTION $R_X(D_1, D_2, \dots, D_N)$

In this section we will prove that $R_X(D_1, D_2, \dots, D_N)$ is a convex U function of (D_1, D_2, \dots, D_N) .

Let $Q'(y_1, y_2, \dots, y_N | x)$ and $Q''(y_1, y_2, \dots, y_N | x)$ achieve the points $(D'_1, D'_2, \dots, D'_N; R_X(D'_1, D'_2, \dots, D'_N))$ and $(D''_1, D''_2, \dots, D''_N; R_X(D''_1, D''_2, \dots, D''_N))$ respectively and let

$$Q^*(y_1, y_2, \dots, y_N | x) = \lambda Q'(y_1, y_2, \dots, y_N | x) + (1 - \lambda) Q''(y_1, y_2, \dots, y_N | x),$$

where $\lambda \in [0, 1]$. Now by definition

$$D_i(Q) = \sum_{x, y_1, y_2, \dots, y_N} P(x) Q(y_1, y_2, \dots, y_N | x) \varrho_i(x, y_i) \quad (i = 1, 2, \dots, N)$$

and in particular

$$D_i(Q^*(y_1, y_2, \dots, y_N | x)) = \lambda D'_i + (1 - \lambda) D''_i \quad (i = 1, 2, \dots, N).$$

This shows that $D_i(Q^*)$ for $i = 1, 2, \dots, N$ is a linear function of $Q^*(y_1, y_2, \dots, y_N | x)$ so that

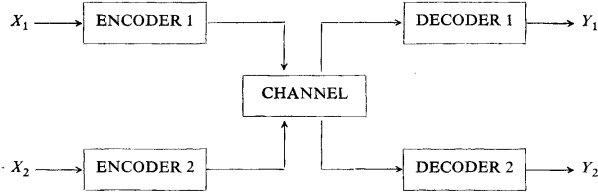
$$Q^*(y_1, y_2, \dots, y_N | x) \in Q_{\lambda D'_1 + (1-\lambda)D''_1, \dots, \lambda D'_N + (1-\lambda)D''_N}.$$

Next we have

$$\begin{aligned} R_X(\lambda D'_1 + (1 - \lambda) D''_1, \lambda D'_2 + (1 - \lambda) D''_2, \dots, \lambda D'_N + (1 - \lambda) D''_N) &\leq \\ &\leq I(Q^*(y_1, y_2, \dots, y_N | x)) \leq \lambda I(Q'(y_1, y_2, \dots, y_N | x)) + \\ &\quad + (1 - \lambda) I(Q''(y_1, y_2, \dots, y_N | x)) = \\ &= \lambda R_X(D'_1, D'_2, \dots, D'_N) + (1 - \lambda) R_X(D''_1, D''_2, \dots, D''_N). \end{aligned}$$

Hence $R_X(D_1, D_2, \dots, D_N)$ is a convex U function of (D_1, D_2, \dots, D_N) .

In this section we derive basic equations for Two-User Channels. We first define a Two-User Channel:



Two-User Channel

A channel with two sources and two receivers is called a Two-User Channel. We shall consider it to be discrete and memoryless. Let X_1, X_2 be the two sources and Y_1, Y_2 be the two receivers. A discrete memoryless Two-User Channel with two inputs and two outputs may be characterised by $(X_1 \times X_2, Q(y_1, y_2 | x_1, x_2), Y_1 \times Y_2)$. We shall denote by $Q_i(y_i | x_1, x_2)$ the transition probability of receiving $y_i \in Y_i$ on the i -th receiver ($i = 1, 2$) when $x_1 \in X_1$ and $x_2 \in X_2$ are transmitted, then since y_1 and y_2 are statistically independent given x_1 and x_2 , we have

$$Q(y_1, y_2 | x_1, x_2) = Q_1(y_1 | x_1, x_2) Q_2(y_2 | x_1, x_2).$$

Further the distortion between the source letter $x_i \in X_i$ and the reproduced letter $y_i \in Y_i$ is denoted by $\varrho_i(x_i, y_i)$ where as usual

$$\varrho_i(x_i, y_i) \geq 0 \quad (i = 1, 2)$$

with equality iff $x_i = y_i$.

If $P(x_1, x_2)$ is taken to denote input probability of $x_1 \in X_1$ and $x_2 \in X_2$ then the average distortion for the i -th output may be defined as

$$\sum_{x_1, x_2, y_1, y_2} P(x_1, x_2) Q(y_1, y_2 | x_1, x_2) \varrho_i(x_i, y_i) \quad (i = 1, 2).$$

Now if we communicate on the Two-User Channel in such a way that the average distortion between the i -th receiver and the i -th input does not exceed a given distortion level D_i ($i = 1, 2$) then the rate distortion function for the Two-User Channel may be defined as

$$(15) \quad R_{x_1, x_2}(D_1, D_2) = \inf_{Q(y_1, y_2 | x_1, x_2) \in Q_{D_1, D_2}} I(X_1, X_2; Y_1, Y_2),$$

where

$$(16) \quad I(X_1, X_2; Y_1, Y_2) = \sum_{x_1, x_2, y_1, y_2} P(x_1, x_2, y_1, y_2) \log \frac{Q(y_1, y_2 | x_1, x_2)}{Q(y_1, y_2)}$$

is the ordinary Shannon's mutual information and

$$(17) \quad Q_{D_i, D_2} = \\ = \{Q(y_1, y_2 | x_1, x_2) : \sum_{x_1, x_2, y_1, y_2} P(x_1, x_2) Q(y_1, y_2 | x_1, x_2) \varrho_i(x_i, y_i) \leq D_i\} \\ (i = 1, 2).$$

The $I(X_1, X_2; Y_1, Y_2)$ may be easily shown to be a convex \cup function of $Q(y_1, y_2 | x_1, x_2)$. Thus our problem is to minimize $I(X_1, X_2; Y_1, Y_2)$ subject to the constraints:

$$(18) \quad Q(y_1, y_2 | x_1, x_2) \geq 0,$$

$$(19) \quad \sum_{y_1, y_2} Q(y_1, y_2 | x_1, x_2) = 1$$

and

$$(20) \quad \sum_{x_1, x_2, y_1, y_2} P(x_1, x_2) Q(y_1, y_2 | x_1, x_2) \varrho_i(x_i, y_i) = D_i \quad (i = 1, 2).$$

As before we construct the augmented function (ignoring the constraints (18) temporarily)

$$J(Q) = I(X_1, X_2; Y_1, Y_2) - \sum_{x_1, x_2} \mu_{x_1, x_2} \sum_{y_1, y_2} Q(y_1, y_2 | x_1, x_2) - \\ - \sum_{i=1}^2 S_i \sum_{x_1, x_2, y_1, y_2} P(x_1, x_2) Q(y_1, y_2 | x_1, x_2) \varrho_i(x_i, y_i),$$

where μ_{x_1, x_2} and S_i 's ($i = 1, 2$) are Lagrange multipliers. Taking

$$\log \lambda_{x_1, x_2} = \frac{\mu_{x_1, x_2}}{P(x_1, x_2)},$$

for stationary points, we have

$$\frac{dJ}{dQ(y_1, y_2 | x_1, x_2)} = P(x) \left[\log \frac{Q(y_1, y_2 | x_1, x_2)}{Q(y_1, y_2) \lambda_{x_1, x_2}} - \sum_{i=1}^2 S_i \varrho_i(x_i, y_i) \right] = 0,$$

i.e.

$$(21) \quad Q(y_1, y_2 | x_1, x_2) = Q(y_1, y_2) \lambda_{x_1, x_2} \exp \left[\sum_{i=1}^2 S_i \varrho_i(x_i, y_i) \right].$$

454 Summing (21) over y_1, y_2 and using (19), we get

$$(22) \quad \lambda_{x_1, x_2} = \left(\sum_{y_1, y_2} Q(y_1, y_2) \exp \left[\sum_{i=1}^2 S_i q_i(x_i, y_i) \right] \right)^{-1}.$$

Thus we have

$$(23) \quad R_{X_1, X_2}(D_1, D_2) = \sum_{i=1}^2 S_i D_i + \sum_{x_1, x_2} P(x_1, x_2) \log \lambda_{x_1, x_2}$$

and

$$(24) \quad D_i = \sum_{x_1, x_2, y_1, y_2} q_i(x_i, y_i) P(x_1, x_2) \lambda_{x_1, x_2} Q(y_1, y_2) \exp \left[\sum_{i=1}^2 S_i q_i(x_i, y_i) \right] \\ (i = 1, 2)$$

as the required basic equations for Two-User Channels.

Now if for a particular value of S_i ($i = 1, 2$), one or more $Q(y_1, y_2 | x_1, x_2) \leq 0$ then as before the results can be formulated as in Berger [1, Lemma 1, p. 32].

We can prove that the function $R_{X_1, X_2}(D_1, D_2)$ is convex with respect to D_1 and D_2 . The proof can be developed on the lines adopted as in Section 3.

5. AN EXAMPLE OF A TWO-USER CHANNEL

Let us consider a discrete memoryless Two-User Channel with input alphabet sets $X_1 = \{1, 2\}$ and $X_2 = \{3, 4\}$ and output alphabet sets $Y_1 = \{1, 2\}$ and $Y_2 = \{3, 4\}$. Also let

$$q_{ik} = 1 - \delta_{ik} \quad \text{where} \quad \delta_{ik} = 1 \quad \text{for} \quad i = k, \\ = 0 \quad \text{for} \quad i \neq k,$$

so that

$$q_{ik} = 0 \quad \text{for} \quad i = k, \\ = 1 \quad \text{for} \quad i \neq k; \quad i, k = (1, 2); (3, 4).$$

Further we take the joint probability $P(x_1, x_2)$ to be given by

$$P(1, 3) = p_1; \quad P(2, 3) = p_3; \\ P(1, 4) = p_2; \quad P(2, 4) = p_4$$

and

$$\sum_{x_1, x_2} P(x_1, x_2) = 1.$$

Multiplying (21) by $P(x_1, x_2)$ and summing over x_1, x_2 , we get

$$(25) \quad \sum_{x_1, x_2} P(x_1, x_2) \lambda_{x_1, x_2} \exp \left[\sum_{i=1}^2 S_i q_i(x_i, y_i) \right] = 1.$$

Solving these simultaneous equations, we get

$$(26) \quad \lambda_{x_1, x_2} = \frac{1}{(1 + \alpha)(1 + \beta)P(x_1, x_2)}; \quad x_1 = 1, 2; \quad x_2 = 3, 4,$$

where

$$(27) \quad \alpha = \exp S_1 \quad \text{and} \quad \beta = \exp S_2.$$

Also from (22) we have

$$\frac{1}{\lambda_{x_1, x_2}} = \sum_{y_1, y_2} Q(y_1, y_2) \exp \left[\sum_{i=1}^2 S_i \varrho_i(x_i, y_i) \right]; \quad y_1 = 1, 2; \quad y_2 = 3, 4.$$

Solving these equations for $Q(y_1, y_2)$, we get

$$(28) \quad Q(1, 3) = \frac{p_1 - \alpha p_3 - \beta p_2 + \alpha \beta p_4}{(1 - \alpha)(1 - \beta)}, \quad Q(1, 4) = \frac{p_2 - \alpha p_4 - \beta p_1 + \alpha \beta p_3}{(1 - \alpha)(1 - \beta)},$$

$$Q(2, 3) = \frac{p_3 - \alpha p_1 - \beta p_4 + \alpha \beta p_2}{(1 - \alpha)(1 - \beta)}, \quad Q(2, 4) = \frac{p_4 - \alpha p_2 - \beta p_3 + \alpha \beta p_1}{(1 - \alpha)(1 - \beta)}.$$

On using (26), (27) and (28), equations (24) and (23) give

$$(29) \quad D_1 = \frac{\exp S_1}{1 + \exp S_1} = \frac{\alpha}{\alpha + 1}, \quad D_2 = \frac{\exp S_2}{1 + \exp S_2} = \frac{\beta}{\beta + 1}$$

and

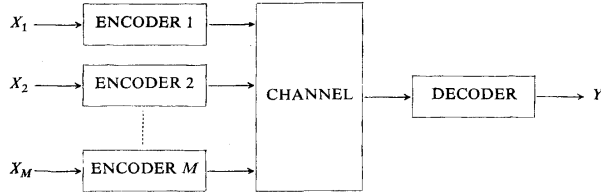
$$(30) \quad R_{X_1, X_2}(D_1, D_2) = H(p_1, p_2, p_3, p_4) + \frac{\alpha}{\alpha + 1} \log \alpha + \frac{\beta}{\beta + 1} \log \beta - \log(\alpha + 1)(\beta + 1).$$

Thus (29) and (30) determine the distortions and rate for the example considered.

6. MULTIPLE ACCESS CHANNEL

In this section we derive the basic equations for Multiple Access Channel. We consider a general Multiple Access Communication System with M sources communicating with one receiver over a common channel. The message output for one source is assumed to be independent from message outputs for other sources. A general Multiple Access Channel with M sources is shown in the diagram.

There are M sources which we denote by X_1, X_2, \dots, X_M and one receiver which we denote by Y . A Multiple Access Channel with M inputs and one output may be characterized by $(X_1 \times X_2 \times \dots \times X_M, Q(y | x_1, x_2, \dots, x_M), Y)$.



Multiple Access Channel

Further the distortion between the source letter $x_i \in X_i$ and the reproduced letter $y \in Y$ is denoted by $q_i(x_i, y)$ where as usual

$$q_i(x_i, y) \geq 0 \quad (i = 1, 2, \dots, M)$$

with equality iff $x_i = y$.

If $P(x_1, x_2, \dots, x_M)$ is taken to denote the input probability of $x_1 \in X_1, x_2 \in X_2, \dots, x_M \in X_M$ then the average distortion may as usual be defined as

$$\sum_{x_1, \dots, x_M, y} P(x_1, x_2, \dots, x_M) Q(y | x_1, x_2, \dots, x_M) q_i(x_i, y) \quad (i = 1, 2, \dots, M).$$

Now if we communicate on the Multiple Access Channel in such a way that the level of average distortion between the i -th source and the reproduced letter does not exceed a given level D_i ($i = 1, 2, \dots, M$) then the rate distortion function $R_{X_1, \dots, X_M}(D_1, \dots, D_M)$ for Multiple Access Channel may be defined as

$$(31) \quad R_{X_1, \dots, X_M}(D_1, \dots, D_M) = \inf_{Q(y | x_1, \dots, x_M) \in Q_{D_1, D_2, \dots, D_M}} I(X_1, X_2, \dots, X_M; Y),$$

where

$$(32) \quad I(X_1, X_2, \dots, X_M; Y) = \sum_{x_1, \dots, x_M, y} P(x_1, x_2, \dots, x_M, y) \log \frac{Q(y | x_1, x_2, \dots, x_M)}{Q(y)}$$

is the ordinary Shannon's mutual information and

$$(33) \quad Q_{D_1, \dots, D_M} = \{Q(y | x_1, x_2, \dots, x_M) : \sum_{x_1, \dots, x_M, y} P(x_1, \dots, x_M) \cdot Q(y | x_1, \dots, x_M) q_i(x_i, y) \leq D_i \quad (i = 1, 2, \dots, M)\}.$$

The $I(X_1, X_2, \dots, X_M; Y)$ may be easily shown to be a convex \cup function of $Q(y | x_1, x_2, \dots, x_M)$. Thus our problem is to minimize $I(X_1, X_2, \dots, X_M; Y)$ subject to the constraints:

$$(34) \quad Q(y | x_1, x_2, \dots, x_M) \geq 0,$$

$$(35) \quad \sum_y Q(y | x_1, x_2, \dots, x_M) = 1$$

and

$$(36) \quad \sum_{x_1, \dots, x_M, y} P(x_1, x_2, \dots, x_M) Q(y | x_1, x_2, \dots, x_M) \varrho_i(x_i, y) = D_i \\ (i = 1, 2, \dots, M).$$

We construct the augmented function (ignoring the constraints (34) temporarily)

$$J(Q) = I(X_1, \dots, X_M; Y) - \sum_{x_1, \dots, x_M} \mu_{x_1, x_2, \dots, x_M} \sum_y Q(y | x_1, x_2, \dots, x_M) - \\ - \sum_{i=1}^M S_i \sum_{x_1, \dots, x_M, y} P(x_1, x_2, \dots, x_M) Q(y | x_1, x_2, \dots, x_M) \varrho_i(x_i, y),$$

where μ_{x_1, \dots, x_M} and S_i 's ($i = 1, 2, \dots, M$) are Lagrange multipliers. Taking

$$\log \lambda_{x_1, \dots, x_M} = \frac{\mu_{x_1, x_2, \dots, x_M}}{P(x_1, x_2, \dots, x_M)},$$

we have for stationary points

$$\frac{dJ}{dQ(y | x_1, \dots, x_M)} = P(x_1, \dots, x_M) \left[\log \frac{Q(y | x_1, \dots, x_M)}{Q(y) \lambda_{x_1, \dots, x_M}} - \sum_{i=1}^M S_i \varrho_i(x_i, y) \right] = 0, \\ \text{i.e.}$$

$$(37) \quad Q(y | x_1, \dots, x_M) = Q(y) \lambda_{x_1, x_2, \dots, x_M} \exp \left[\sum_{i=1}^M S_i \varrho_i(x_i, y) \right].$$

Summing (37) over y and using (35), we get

$$(38) \quad \lambda_{x_1, x_2, \dots, x_M} = \left(\sum_y Q(y) \exp \left[\sum_{i=1}^M S_i \varrho_i(x_i, y) \right] \right)^{-1}.$$

Thus we have

$$R_{x_1, \dots, x_M}(D_1, \dots, D_M) = \sum_{i=1}^M S_i D_i + \sum_{x_1, \dots, x_M} P(x_1, \dots, x_M) \log \lambda_{x_1, \dots, x_M}$$

and

$$(40) \quad D_i = \sum_{x_1, \dots, x_M, y} \varrho_i(x_i, y) P(x_1, \dots, x_M) \lambda_{x_1, \dots, x_M} Q(y) \exp \left[\sum_{i=1}^M S_i \varrho_i(x_i, y) \right] \\ (i = 1, 2, \dots, M)$$

are the required basic equations for Multiple Access Channel.

Now if for a particular value of S_i ($i = 1, 2, \dots, M$) one or more $Q(y | x_1, x_2, \dots, x_M) \leq 0$ then as before the results can be formulated as in Berger [1, Lemma 1, p. 32].

7. AN EXAMPLE OF A MULTIPLE ACCESS CHANNEL

Let us consider a Multiple Access Channel with input alphabet sets $X_1 = \{1, 2\}$, $X_2 = \{2, 3\}$ and output alphabet set $Y = \{1, 2, 3, 4\}$. Also let

$$Q_{ik} = 1 - \delta_{ik} \quad \text{where} \quad \delta_{ik} = 1 \quad \text{for} \quad i = k, \\ = 0 \quad \text{for} \quad i \neq k,$$

so that

$$Q_{ik} = 0 \quad \text{for} \quad i = k \\ = 1 \quad \text{for} \quad i \neq k; \quad i = 1, 2, 3; \quad k = 1, 2, 3, 4.$$

Further we take the joint probability $P(x_1, x_2)$ to be given by

$$P(1, 2) = p_1; \quad P(2, 2) = p_3; \\ P(1, 3) = p_2; \quad P(2, 3) = p_4$$

and $\sum_{x_1, x_2} P(x_1, x_2) = 1$.

Now from (37), we have

$$\sum_{x_1, x_2} P(x_1, x_2) \lambda_{x_1, x_2} \exp \left[\sum_{i=1}^2 S_i Q_i(x_i, y) \right] = 1.$$

Solving these simultaneous equations for λ_{x_1, x_2} , we get

$$(41) \quad \lambda_{12} = \frac{\alpha\beta - 1}{\alpha\beta(1 - \alpha)(\beta - 1)p_1}; \quad \lambda_{22} = \frac{\alpha + \beta - 2\alpha\beta}{\alpha\beta(1 - \alpha)(\beta - 1)p_3}; \\ \lambda_{13} = \frac{1 - \alpha\beta}{\alpha\beta(1 - \alpha)(\beta - 1)p_2}; \quad \lambda_{23} = \frac{\alpha\beta - 1}{\alpha\beta(1 - \alpha)(\beta - 1)p_4},$$

where

$$(42) \quad \alpha = \exp S_1 \quad \text{and} \quad \beta = \exp S_2.$$

Also we have from (38),

$$\frac{1}{\lambda_{x_1, x_2}} = \sum_y Q(y) \exp \left[\sum_{i=1}^2 S_i Q_i(x_i, y) \right].$$

Solving these equations for $Q(y)$, we get

$$(43) \quad Q(1) = \alpha \left[\frac{\beta p_1 + p_2 + p_4}{\alpha\beta - 1} + \frac{\beta p_3}{2\alpha\beta - \alpha - \beta} \right]; \\ Q(2) = \alpha\beta \left[\frac{p_1 + p_2 + p_4}{\alpha\beta - 1} + \frac{p_3}{2\alpha\beta - \alpha - \beta} \right];$$

$$Q(3) = \beta \left[\frac{p_1 + p_2 + \alpha p_4}{\alpha\beta - 1} + \frac{\alpha p_3}{2\alpha\beta - \alpha - \beta} \right];$$

$$Q(4) = \frac{(\alpha + \beta + \alpha\beta) p_3}{\alpha + \beta - 2\alpha\beta} + \frac{1}{1 - \alpha\beta}.$$

$$\cdot [(1 + \beta + \alpha\beta) p_1 + (1 + \alpha + \beta) p_2 + (1 + \alpha + \alpha\beta) p_4].$$

Thus on using (41), (42) and (43), we have from (40) and (39)

$$(44) \quad D_1 = 1 - \frac{\alpha(\beta - 1)}{1 - \alpha} \left[\frac{p_1 + p_2 + p_4}{\alpha\beta - 1} + \frac{p_3}{2\alpha\beta - \alpha - \beta} \right];$$

$$D_2 = 1 - \frac{\beta(1 - \alpha)}{\beta - 1} \left[\frac{p_1 + p_2 + p_4}{\alpha\beta - 1} + \frac{p_3}{2\alpha\beta - \alpha - \beta} \right]$$

and

$$(45) \quad R_{X_1, X_2}(D_1, D_2) = S_1 D_1 + S_2 D_2 + H(p_1, p_2, p_3, p_4) + (p_1 + p_4) \log(\alpha\beta - 1) + p_2 \log(1 - \alpha\beta) + p_3 \log(\alpha + \beta - 2\alpha\beta) - \log \alpha\beta(1 - \alpha)(\beta - 1).$$

Equations (44) and (45) determine the distortions and rate for the example considered.

8. BASIC EQUATIONS FOR A GENERAL CASE

So far we have derived the basic equations for the cases of special interest. In this section we will study the general case having several inputs and several outputs. All the cases which were considered earlier become a special case of this general case. We consider a discrete memoryless channel with M inputs and N outputs. Let X_1, X_2, \dots, X_M represent M inputs and Y_1, Y_2, \dots, Y_N represent N outputs. We characterise the channel by $(X_1 \times X_2 \times \dots \times X_M, Q(y_1, y_2, \dots, y_N | x_1, x_2, \dots, x_M), Y_1 \times Y_2 \times \dots \times Y_N)$. The transition probability of receiving $y_i \in Y_i$ by the i -th receiver ($i = 1, 2, \dots, N$) when x_1, x_2, \dots, x_M are transmitted is represented by $Q_i(y_i | x_1, x_2, \dots, x_M)$ where

$$Q_i(y_i | x_1, x_2, \dots, x_M) = \sum_{y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_N} Q(y_1, y_2, \dots, y_N | x_1, x_2, \dots, x_M)$$

$$(i = 1, 2, \dots, N).$$

Then, since the outputs are statistically independent, we have

$$Q(y_1, \dots, y_N | x_1, \dots, x_M) = Q_1(y_1 | x_1, \dots, x_M) \dots Q_N(y_N | x_1, \dots, x_M).$$

Since there are M inputs and N outputs, so there will be $M \times N$ distortions. We will represent the distortion between the source letter $x_i \in X_i$ ($i = 1, 2, \dots, M$) and the reproduced letter $y_j \in Y_j$ ($j = 1, 2, \dots, N$) by $q_{ij}(x_i, y_j)$ where as usual

$$q_{ij}(x_i, y_j) \geq 0$$

with equality iff $x_i = y_j$ ($\forall i, j$).

If $P(x_1, x_2, \dots, x_M)$ is taken to denote the input probability of $x_i \in X_i$, $x_2 \in X_2, \dots, x_M \in X_M$ then the average distortion for the j -th output ($j = 1, 2, \dots, N$) may be as usual defined as

$$\sum_{x_1, \dots, x_M, y_1, \dots, y_N} P(x_1, \dots, x_M) Q(y_1, \dots, y_N | x_1, \dots, x_M) q_{ij}(x_i, y_j).$$

Now if we communicate on the channel in such a way that the level of average distortion between the i -th source and the j -th reproduced letter does not exceed a given level D_{ij} ($i = 1, 2, \dots, M$; $j = 1, 2, \dots, N$) then the rate distortion function $R_{X_1, \dots, X_M}(D_{11}, \dots, D_{MN})$ for this channel may be defined as

$$(46) \quad R_{X_1, \dots, X_M}(D_{11}, \dots, D_{MN}) = \inf_{Q(y_1, \dots, y_N | x_1, \dots, x_M) \in Q_{D_{11}, \dots, D_{MN}}} I(X_1, \dots, X_M; Y_1, \dots, Y_N),$$

where

$$(47) \quad I(X_1, \dots, X_M; Y_1, \dots, Y_N) = \sum_{x_1, \dots, x_M, y_1, \dots, y_N} P(x_1, \dots, x_M, y_1, \dots, y_N) \cdot \log \frac{Q(y_1, \dots, y_N | x_1, \dots, x_M)}{Q(y_1, \dots, y_N)}$$

is the ordinary Shannon's mutual information and

$$Q_{D_{11}, \dots, D_{MN}} = \{Q(y_1, \dots, y_N | x_1, \dots, x_M) : \sum_{x_1, \dots, x_M, y_1, \dots, y_N} P(x_1, \dots, x_M) \cdot Q(y_1, \dots, y_N | x_1, \dots, x_M) q_{ij}(x_i, y_j) \leq D_{ij}, \\ i = 1, 2, \dots, M; \quad j = 1, 2, \dots, N.\}$$

The $I(X_1, \dots, X_M; Y_1, \dots, Y_N)$ may be shown to be a convex U function of $Q(y_1, \dots, y_N | x_1, \dots, x_M)$. Thus our problem is to minimize $I(X_1, \dots, X_M; Y_1, \dots, Y_N)$ subject to the constraints:

$$(49) \quad Q(y_1, \dots, y_N | x_1, \dots, x_M) \geq 0,$$

$$(50) \quad \sum_{y_1, \dots, y_N} Q(y_1, \dots, y_N | x_1, \dots, x_M) = 1$$

and

$$(51) \quad \sum_{x_1, \dots, x_M, y_1, \dots, y_N} P(x_1, \dots, x_M) Q(y_1, \dots, y_N | x_1, \dots, x_M) q_{ij}(x_i, y_j) = D_{ij} \\ (i = 1, 2, \dots, M; \quad j = 1, 2, \dots, N).$$

We will use Lagrange's method of multipliers to solve this problem. Ignoring the constraints (49) temporarily we form the augmented function

$$J(Q) = I(X_1, \dots, X_M; Y_1, \dots, Y_N) - \sum_{x_1, \dots, x_M} \mu_{x_1, x_2, \dots, x_M} \cdot \\ \cdot \sum_{y_1, \dots, y_N} Q(y_1, \dots, y_N | x_1, \dots, x_M) - \sum_{i=1}^M \sum_{j=1}^N S_{ij} \sum_{x_1, \dots, x_M, y_1, \dots, y_N} P(x_1, \dots, x_M) \cdot \\ \cdot Q(y_1, \dots, y_N | x_1, \dots, x_M) \varrho_{ij}(x_i, y_j),$$

where $\mu_{x_1, x_2, \dots, x_M}$ and S_{ij} ($i = 1, 2, \dots, M; j = 1, 2, \dots, N$) are Lagrange multipliers. Taking

$$\log \lambda_{x_1, \dots, x_M} = \frac{\mu_{x_1, x_2, \dots, x_M}}{P(x_1, \dots, x_M)}$$

and proceeding as in Section 2, we have

$$(A) \quad R_{x_1, \dots, x_M}(D_{11}, \dots, D_{MN}) = \sum_{i=1}^M \sum_{j=1}^N S_{ij} D_{ij} + \\ + \sum_{x_1, \dots, x_M} P(x_1, \dots, x_M) \log \lambda_{x_1, \dots, x_M}$$

and

$$(B) \quad D_{ij} = \sum_{x_1, \dots, x_M, y_1, \dots, y_N} \varrho_{ij}(x_i, y_j) P(x_1, \dots, x_M) \lambda_{x_1, \dots, x_M} Q(y_1, \dots, y_N) \cdot \\ \cdot \exp \left[\sum_{i=1}^M \sum_{j=1}^N S_{ij} \varrho_{ij}(x_i, y_j) \right] \\ (i = 1, 2, \dots, M; j = 1, 2, \dots, N),$$

where

$$(52) \quad \lambda_{x_1, \dots, x_M} = \left[\sum_{y_1, \dots, y_N} Q(y_1, \dots, y_N) \exp \left[\sum_{i=1}^M \sum_{j=1}^N S_{ij} \varrho_{ij}(x_i, y_j) \right] \right]^{-1}.$$

Expressions (A) and (B) give the required form of the basic equations for this discrete memoryless channel in the general case.

Particular Cases

1. When $i = 1; j = 1, 2, \dots, N$, i.e. we come to the case when there is one source and N destinations. Then (A) and (B) reduce to

$$R_{x_1}(D_{11}, \dots, D_{1N}) = \sum_{j=1}^N S_{1j} D_{1j} + \sum_{x_1} P(x_1) \log \lambda_{x_1}$$

462 and

$$D_{1j} = \sum_{x_1, y_1, y_2, \dots, y_N} q_{1j}(x_1, y_j) P(x_1) \lambda_{x_1} Q(y_1, \dots, y_N) \cdot \exp \left[\sum_{j=1}^N S_{1j} q_{1j}(x_1, y_j) \right]$$

$$(j = 1, 2, \dots, N),$$

which are nothing but the basic equations for Broadcast Channel.

2. When $i = j = 1, 2$ and $q_{12}(x_1, y_2) = q_{21}(x_2, y_1) = 0$ i.e. when there are two sources and two receivers. Then (A) and (B) reduce to

$$R_{X_1, X_2}(D_{11}, D_{22}) = \sum_{i=1}^2 S_{ii} D_{ii} + \sum_{x_1, x_2} P(x_1, x_2) \log \lambda_{x_1, x_2}$$

and

$$D_{ii} = \sum_{x_i, x_2, y_1, y_2} q_{ii}(x_i, y_i) P(x_1, x_2) \lambda_{x_1, x_2} Q(y_1, y_2) \exp \left[\sum_{i=1}^2 S_{ii} q_{ii}(x_i, y_i) \right]$$

$$(i = 1, 2),$$

which are the basic equations for Two-User Channel. Thus our general model reduces to the case of Two-User Channels when $i = j = 1, 2$ and $q_{12}(x_1, y_2) = q_{21}(x_2, y_1) = 0$.

3. Let $i = 1, 2, \dots, M; j = 1$, i.e. let us consider the case when there are M sources and one receiver. Then in this case (A) and (B) reduce to

$$R_{X_1, \dots, X_M}(D_{11}, \dots, D_{M1}) = \sum_{i=1}^M S_{ii} D_{ii} + \sum_{x_1, \dots, x_M} P(x_1, \dots, x_M) \log \lambda_{x_1, \dots, x_M}$$

and

$$D_{ii} = \sum_{x_1, \dots, x_M, y_1} q_{ii}(x_i, y_1) P(x_1, \dots, x_M) \lambda_{x_1, \dots, x_M} Q(y_1) \exp \left[\sum_{i=1}^M S_{ii} q_{ii}(x_i, y_1) \right]$$

$$(i = 1, 2, \dots, M),$$

which are the basic equations for Multiple Access Channel. Thus our general model reduces to the case of Multiple Access Channel when $i = 1, 2, \dots, M$ and $j = 1$.

(Received December 18, 1978.)

REFERENCES

- [1] T. Berger: Rate Distortion Theory: A Mathematical Basis for Data Compression. Englewood Cliffs, Prentice Hall 1971.
- [2] T. M. Cover: Broadcast channels. IEEE Trans. Inf. Theory *IT-18*, Jan. 1972, 2-14.
- [3] P. M. Ebert: An extension of rate distortion theory to confusion matrices. IEEE Trans. Inf. Theory *IT-14*, Jan. 1968, 6-11.

- [4] H. Liao: Multiple access channels. Ph. D. Dissertation, Department of Elect. Engg., Univ. of Hawaii, Honolulu 1972.
- [5] H. Sato: An outer bound to the capacity region of broadcast channels. IEEE Trans. Inf. Theory *IT-24*, May 1978, 374–377.
- [6] H. Sato: Two-user communication channels. IEEE Trans. Inf. Theory *IT-23*, May 1977, 295–304.
- [7] C. E. Shannon: Two way communication channels. Proc. of the Fourth Berkeley Symp. on Prob. and Statist., Vol. 1, Berkeley, California, Univ. of California Press, 1961, 611–644.
- [8] E. C. Vander Meulen: Random coding theorems for the general discrete memoryless broadcast channel. IEEE Trans. Inf. Theory *IT-21*, March 1975, 180–190.

Prof. Bhu Dev Sharma, University of the West Indies, Department of Mathematics, St. Augustine, Trinidad (W. I.).

Ved Priya, University of Delhi, Department of Mathematics, Delhi - 110007, India.