COMMUNICATION THROUGH CHANNELS IN CASCADE

by

CHARLES A. DESOER

Ingénieur Radio-électricien

University of Liège

(1949)

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Science

at the

Massachusetts Institute of Technology

(1953)

Signature of the Author ______________________________

DEPARTMENT OF ELECTRICAL ENGINEERING, January, 1953.

Certified by ______________________________

Thesis Supervisor

____

Chairman, Departmental Committee on Graduate Students

COMMUNICATION THROUGH CHANNELS IN CASCADE

by

CHARLES AUGUSTE DESOER

Submitted to the Department of Electrical Engineering on January 12, 1953, in partial fulfillment of the requirements for the degree of Doctor of Science.

ABSTRACT

The present work is a theoretical discussion of communication through noisy channels in cascade. The point of view adopted for that discussion is that of information theory. After a general discussion of channels in cascade, the dependence of the cascade performance on two factors is studied in detail by considering suitable examples. These factors are, respectively, the delay allowed at the intermediate station and the intermediate station transfer characteristic. In the course of these discussions, a technique for constructing a double and a triple error correcting code is indicated. This technique is generalized and forms the basis of a constructive proof of Shannon's theorem in the case of the binary channel.

Thesis Supervisor: Robert M. Fano.

Title: Associate Professor of Electrical Engineering.

ACKNOWLEDGMENT

The writer wishes to express his gratitude to Professor R. M. Fano for his helpful suggestions in the supervision of this work. He is also grateful to Professors J. B. Wiesner and W. B. Davenport for their suggestions. He would like to thank Miss Elizabeth J. Campbell for carrying out some calculations.

TABLE OF CONTENTS

Chapter I

INTRODUCTION

1.1 Historical Remarks

The purpose of this section is to draw the attention of the reader to some major contributions, the results of which are repeatedly used in this thesis. For a detailed history of information theory, the reader is referred to the literature.(1,2)

During the last two decades, a large number of new modulation methods were developed. We may mention frequency modulation, phase modulation and the family of pulse modulation methods such as P.A.M., P.D.M., P.P.M, and P.C.M. This sudden wealth of design possibilities led to a reexamination of the fundamental aspects of the communication problem, and, as it is usual in science, the answer was found in a more abstract approach. A major step was achieved when Norbert Wiener pointed out that the communication problem is essentially statistical in nature. He also defined, for a particular situation, a measure of the rate of transmission of information. In fact, Hartley, in a much earlier paper pointed out that the measure of information should involve the logarithmic function. Another fundamental contribution was that of C. E. Shannon whose 1948 paper presented a complete theory and the derivation of a number of basic theorems among which the "fundamental theorem" is the most important and by far the most interesting. For ease of reference let us state it here:

"Let a channel have the capacity C and a source the entropy per second H. If $H \leqslant C$, there exists a coding system such that the output of the source can be transmitted over the channel with an arbitrarily small frequency of error. If $H > C$, it is possible to encode the source so that the equivocation is less than $H - C + \epsilon$ where ϵ is arbitrarily small. There is no method of encoding which gives an equivocation less than $H - C$."(6,7)

It should be stressed that the proof of this theorem is non-constructive.

In the last few years the interest in the theory grew larger and now many papers have been and are being published. Many concepts have been made clear and some problems have been solved. No paper, however, has yet dealt with the problem of channels in cascade from the information theory point of view which is the purpose of the present work.

1.2 Terminology

In information theory, the terminology is still somewhat fluid. It is therefore important to start by defining carefully some of the terms which will occur repeatedly.

For simplicity, we assume that the purpose of a communication system is to reproduce as closely as possible a message generated at some other point. The message is defined as a sequence of symbols. We assume furthermore that the messages consist of a sequence of statistically independent symbols.

In order to transmit a symbol or a group of symbols, the transmitter controls the evolution in time of a suitable physical phenomenon. The evolution in time corresponding to a particular symbol (or group of symbols) is completely described by a function of time, which is called a signal. For bandlimited channels, a signal may be completely described by 2TW equidistant samples,[8] where T is the duration of the signal and W the bandwidth. There is a one-to-one correspondence between the symbols and the signals at the transmitter. In general, the transmitted signal is modified by some kind of random disturbance which is referred to as noise. If the transmitted symbols form a finite set and if the channel's output symbols (by the channel's output we mean the output of the receiver; in other words, the channel includes the receiver) form also a finite set, the channel is said to be discrete. It should be pointed out that, in many discrete channels, the received signals (that is the signals, distorted by noise, as they enter the receiver) form an infinite set but the receiver operates on them in such a way that the channel's output is discrete, that is consists of symbols belonging to a finite set. This is the case of a teletypewriter system for example. If the channel's output symbols form an infinite set, that is the output alphabet is infinite, the channel is said to be continuous.

1.3 Channels in Cascade

Cascaded channels are very often used in practice.

Their use is made necessary because, as in microwave links, the electromagnetic waves do not follow the curvature of the earth or, as in coaxial cables, because the attenuation suffered by the signal becomes prohibitive when the distance becomes large. The designer is then forced to break up the channel AB into a cascade of channels AP_1, P_1P_2, $\cdots$ $P_{n-1}B$. We shall call i^{th} "intermediate station" the assemblage of the i^{th} channel receiver and the $(i + 1)^{th}$ channel transmitter.

The large number of microwave links recently built enhances the desirability of a discussion of channels in cascade from the point of view of information theory. Designers know that in cascaded channels it is important to use modulation systems exhibiting noise reducing properties such as F.M. and P.C.M.

From the information theory point of view, there is a very important difference between the problem of transmitting information through a single channel and that of transmitting information through a cascade of channels. In the first case, the transmitter has all the information to be transmitted; whereas, in the second case, (except for the first transmitter) the information which is available to each transmitter (to be precise information about what was transmitted by the first transmitter) is no more in the form of a symbol but rather in the form of a set of a-posteriori probabilities. We should therefore expect to find that the manner in which the intermediate station operates will be very important for the performance of the cascade.

1.4 The Present Work

(a) Purpose

As stated earlier, the purpose of the work presented in the following chapters is a theoretical discussion of the problem of communication through noisy channels in cascade, and the point of view adopted for that discussion is that of information theory.

(b) Results

The investigation was divided in three parts corresponding respectively to Chapters II, III and IV. In Chapter II, the problem of cascaded noisy channels is discussed in general terms. It is shown that the channel capacity of the cascade is smaller than the channel capacity of any of the cascaded channels. As an illustration of the theory, a cascade of P.C.M. channels is compared to a cascade of continuous channels. The results are best summarized by Fig. II.1 and Fig. II.2.

In Chapter III, we try to find out how much the system performance can be improved by increasing the delay allowed at each intermediate station. In all cases under discussion the intermediate station either retransmits the signal having the largest a-posteriori probability or retransmits the received signal as it is. The discussion is carried out in two cases: continuous channels affected by gaussian additive noise and binary channels. In both cases, the gain in performance is very important. Perhaps the most interesting result of Chapter III is the constructive proof of Shannon's fundamental theorem for the binary channel.

In Chapter IV we optimize the intermediate station transfer characteristic, the allowed delay and the average retransmitted power being kept constant. The formal treatment leads to equations that are not soluble in general. However in the case of gaussian additive noise and for sample by sample retransmission at the intermediate station, it is shown that the optimum input probability density is gaussian and that the received sample should be retransmitted as it is by the intermediate station. The simple, but very important, case of a binary channel in which the noise is gaussian and additive is considered next (still assuming that a sample by sample retransmission is required at the intermediate station). For simplicity, the probability of error of the equivalent channel is minimized in this case. The difference between a maximum a-posteriori probability detector and as "optimum" detector (that is a detector which would extract all the information contained in the received signal) is computed numerically for a simple case.

1.5 General Assumptions

For emphasis it is convenient to state at this stage the general assumptions made throughout the thesis.

The message to be transmitted consists of a sequence of statistically independent symbols. Everything happens as if the symbols were independent random selections from a specified ensemble.

Each channel under consideration is noisy and the noise statistics are known in each particular case.

The noise in a particular channel is independent of all the noise disturbances in the other channels.

The noise is independent of the signal and affects each sample of the signal inc endently of the way it affected the previous samples.

The majority of the channels considered in the following chapters will be built according to a model to be described presently.(12)

The transmitter includes a storage device, a selector and a transmitter. The storage device memorizes the M signals--an alphabet of M symbols is assumed--which are functions of time of duration T. The selector is the element which, according to the symbol that has to be transmitted, selects the associated signal and feeds it to the transmitter.

In the majority of cases the receiver of any channel consists of a computing element and a comparator. The computer determines for each received signal the a-posteriori probabilities that it was caused by the various possible transmitted signals. The computer must therefore have in store all the signal-functions and the relevant statistical characteristics of the noise. In many cases the comparator selects the symbol which has the largest a-posteriori probability. To describe this type of receiver operation we use the expression "maximum a-posteriori probability operation." In some cases, the intermediate station retransmits the signal as it is received, so that its role is simply that of raising the power level of the signal. In such cases, the intermediate station will be referred to as a "repeater." Finally there will be cases where a "transfer characteristic" determines the signal to be retransmitted in terms of the particular received signal.

Chapter II

CHANNELS IN CASCADE

In this chapter, the formalism needed for dealing with channels in cascade is developed. In particular it is shown that, provided the transition probability matrices of the cascaded channels are non-singular, the channel capacity of the cascade may be equal to that of one of the channels only if all others are noiseless. Finally a cascade of P.C.M. channels is compared to a cascade of continuous channels connected by repeaters.

2.1 Equivalent Channel

It is often convenient to consider the cascade of channels as a unit, that is, to think of the cascade only in terms of its input and output. This unit will be called the equivalent channel. More precisely, the equivalent channel is the channel which has statistical properties identical to those of the cascade, at least as far as its input-output relations are concerned.

At this point it should be stressed that the statistical properties of the equivalent channel depends very much on the assumed operation of the intermediate stations. Many examples will be presented later showing that a change in the operation of the intermediate station produces very drastic changes in the performance of the equivalent channel. From the point of view adopted here, as long as the operations of the intermediate stations are not specified, the cascade of channels is not yet completely defined.

2.2 Discrete Channels in Cascade

Consider a cascade of n discrete channels. Since each of these channels must transmit the same message, we assume that they have a common alphabet of M symbols. In each channel, appropriate signals are associated to each symbol. We assume that in a particular channel, all signals have the same duration, say T_i in the i^{th} channel. We assume that each intermediate station operates as a "maximum a-posteriori probability detector." Under these conditions, in addition to the propagation time, a delay at least equal to T_i will occur in the i^{th} channel because the receiver must have received the complete signal before being able to compute the a-posteriori probabilities.

For each channel, on the basis of the noise statistics and the decoding procedure, it is possible, in principle at least, to obtain the transmition probabilities, that is, the probability that a particular symbol, say σ_i , being transmitted, some other symbol, say σ_j , will be received. Let this probability, for the k^{th} channel, be represented by

$$p^{(k)}(\sigma_j|\sigma_i)$$

As there are M^2 such probabilities, let them be arranged in a square matrix P_k. More precisely, let $p^{(k)}(\sigma_j \mid \sigma_i)$ belong to the i^{th} row and the j^{th} column. Thus all the elements of a particular row represent the probabilities that the various symbols be received when a particular symbol is transmitted.

We define the operation of the intermediate stations as follows: as soon as a symbol, say σ_i, is received at the output of the k^{th} channel, it is immediately retransmitted by the $(k + 1)^{th}$ channel; this statement holds for $k = 1, 2, \cdots, n - 1$.

The equivalent channel has all its properties defined by its transition probability matrix which is obtainable, by the following:

> Theorem: the transition probability matrix of the equivalent channel is equal to the product of the transition probability matrices of each channel of the cascade; the order of the factor matrices is identical to the order of the channels in the cascade.

The transition probability matrix P of the equivalent channel will be known once all its elements are known. In order to determine the element $p(\sigma_j \mid \sigma_i)$ of the i^{th} row and the j^{th} column we consider the compound event defined as the joint occurrence of the following events: knowing that σ_i is sent by the 1^{st} channel transmitter,

σ_{i_1} is received and retransmitted by the 1^{st} intermediate station

σ_{i_2} is received and retransmitted by the 2^{nd} intermediate station

$\vdots$

$\sigma_{i_{n-1}}$ is received and retransmitted by the $(n - 1)^{st}$ intermediate station and finally σ_j is received by the last receiver.

Because of the assumed independence of the noise in each channel, the probability of the joint event is equal to the product of the probabilities of all individual transitions, hence it is equal to

$$p^{(1)}(\sigma_{i_1}|\sigma_i)\ p^{(2)}(\sigma_{i_2}|\sigma_{i_1}) \cdots\cdots p^{(n-1)}(\sigma_{i_{n-1}}|\sigma_{i_{n-2}})\ p^{(n)}(\sigma_j|\sigma_{i_{n-1}})$$

Consider all sequences of numbers $(i, i_1, i_2, \ldots, i_{n-1}, j)$ where i and j are fixed and the i_k's $(k = 1, 2, \ldots, n-1)$ ranging over all integers from one to M. To each one of these sequences corresponds a compound event and in each case the symbol σ_i is transmitted and the symbol σ_j is received. As these compound events are mutually exclusive and form an exhaustive set, the probability that σ_j is received when σ_i is transmitted is given by the sum of the probabilities of each one of these events(13,14), thus

$$p(\sigma_j|\sigma_i) = \sum_{i_1, i_2, \ldots i_{n-1}=1}^{M} p^{(1)}(\sigma_{i_1}|\sigma_i)\ p^{(2)}(\sigma_{i_2}|\sigma_{i_1}) \cdots\cdots p^{(n-1)}(\sigma_{i_{n-1}}|\sigma_{i_{n-2}})\ p^{(n)}(\sigma_j|\sigma_{i_{n-1}})$$

(1)

If we remember that $p^{(k)}(\sigma_{i_k} \mid \sigma_{i_{k-1}})$ is the element of the i^{th}_{k-1} row and i^{th}_{k} column of the k^{th} channel transition probability matrix, we recognize that the sums (1) represent the elements of a product of matrices, namely

$$P = P^{(1)} . P^{(2)} . \quad . \quad . \quad . \ P^{(n)} \tag{2}$$

It should be stressed that the proof of the theorem did not require any assumptions on the noise characteristics of any channel. The theorem would still be true if the actual signals used to represent a particular symbol are different in each channel. But it should be kept in mind that the assumed intermediate station operation is essential for the validity of the theorem.

In general the matrices $P^{(i)}$ do not commute, thus we state the following:

> Theorem: in general the characteristics of the equivalent channel depend on the order of the channels in the cascade.

In this connection, it is useful to recall the following matrix property: if two matrices are hermitian (that is, if $a_{ij} = a^*_{ji}$) a necessary and sufficient condition that they shall be reducible to the diagonal form by the <u>same</u> collineatory transformation is that they commute. Thus if the matrices $P^{(k)}$ are symmetrical and commutable, they may be all diagonalized by the same transformation. The elements of the product matrix, in diagonal form, are equal to the product of the characteristic values of the factor matrices. As a result, the problem of finding the product of the matrices $P^{(k)}$ is reduced to that of finding their characteristic values. This method will be found useful later on.

2.3 Channel Capacity of the Equivalent Channel

From an information theoretical point of view, the most interesting characteristic of the equivalent channel is its channel capacity. Simple relations between the equivalent channel capacity and those of the individual channels do not seem to exist. But the equivalent channel of the cascade defined in section 2.2 has a capacity limited by the following

> Theorem: The channel capacity of the equivalent channel is always smaller or equal to the smallest channel capacity of the cascaded channels. When the transition-probability-matrices of all channels are non-singular, the equal sign holds only if all but one of the channels are noiseless. An example will show that if one of the matrices is singular the equal sign may hold although all channels are noisy.

To prove this theorem we need only to investigate the case of two channels in cascade, for an obvious recurrence

rence reasoning will extend the result to <u>n</u> channels in cascade.

Let C_1, (C_2 respectively) be the channel capacity of channel 1 (2 respectively); let C_e be the channel capacity of the equivalent channel.

Consider first the case of $C_2 < C_1$. Let us prove the absurdity of the hypothesis $C_e > C_2$. If it were so, the rate at which information (about the input of channel 1) could be received through channel 2 would be larger than C_2. Let R_{12} be the rate at which information (about the input of channel 1) can be received through channel 2. Let R_{22} be the rate at which information (about the input of channel 2) can be received through channel 2. Then it is clear that

$$R_{12} \leqslant R_{22}$$

and if our assumption

$$C_e > C_2$$

were valid, then R_{12} could be made arbitrarily close to C_e. Thus we would have

$$R_{12} > C_2$$

which would imply that

$$R_{22} > C_2$$

which has been shown to be impossible. Hence we must have $C_e \leqslant C_2$.

If we had assumed $C_1 < C_2$, the proof would be along the same lines.

For the second part of the theorem, we make the additional assumption that the transition probability matrices of each channel are non-singular. In particular it will be so if all the diagonal elements of the matrices are larger than $\frac{1}{2}$, for a theorem of J. Hadamard[16] states that if the elements of a matrix $[p_{ij}]$ are such that, for all i's,

$$|p_{ii}| > \sum_{j \neq i} |p_{ij}|$$

then the determinant of the matrix is positive.

<u>First case $C_2 \leq C_1$.</u>

The assumption $C_e = C_2$ requires that the optimum input probability $p(\sigma_x)$ of the equivalent channel must be transformed, in going through channel 1, into the optimum input probability of channel 2. For if it were not the case, we would have

$$R_{22} < C_2$$

and since

$$R_{12} \leqslant R_{22}$$

this would imply

$$R_{12} < C_2$$

which would contradict the assumption $C_e = C_2$.

Thus, for both the equivalent channel and channel 2, the output probability distribution will be identical and the entropy of the output symbol, say σ_z , will be the same in both cases.

Since

$$C_e = H(\sigma_z) - H(\sigma_z | \sigma_x)$$

and

$$C_2 = H(\sigma_z) - H(\sigma_z | \sigma_y)$$

where σ_y is the output symbol of channel 1 and, therefore it is also the input symbol of channel 2, we conclude that

$$H(\sigma_z | \sigma_y) = H(\sigma_z | \sigma_x) \tag{4}$$

By definition we have

$$H(\sigma_z | \sigma_y) = -\sum_{x\,y} p(\sigma_x)\, p^{(1)}(\sigma_y | \sigma_x) \sum_z p^{(2)}(\sigma_z | \sigma_y) \log p^{(2)}(\sigma_z | \sigma_y)$$

and, using the previous theorem, we also have

$$H(\sigma_z | \sigma_x) = -\sum_x p(\sigma_x) \sum_z \left[\sum_y p^{(1)}(\sigma_y | \sigma_x)\, p^{(2)}(\sigma_z | \sigma_y) \right] \log \left[\sum_y p^{(1)}(\sigma_y | \sigma_x)\, p^{(2)}(\sigma_z | \sigma_y) \right]$$

As the function

$$F(u) = -u \log u \tag{5}$$

is a function of u which is convex upward, we have[17]

$$F\left(\sum_i q_i u_i\right) \geqslant \sum_i q_i F(u_i) \tag{6}$$

provided the non negative weighting factors g_1 satisfy the relation

$$\sum_i q_i = 1$$

The equal sign in the inequality occurs only if all the u_i's are equal or if all but one of the g_i's are zero. This theorem allows us to write, using the notation defined by (5),

$$H(\sigma_z|\sigma_y) = \sum_{x,z} p(\sigma_x)\left\{\sum_y p^{(1)}(\sigma_y|\sigma_x)\ F\left(p^{(2)}(\sigma_z|\sigma_y)\right)\right\}$$

$$\leqslant \sum_{x,z} p(\sigma_x)\ \ F\left[\sum_y p^{(1)}(\sigma_y|\sigma_x)\, p^{(2)}(\sigma_z|\sigma_y)\right]$$

or

$$H(\sigma_z|\sigma_y) \leqslant H(\sigma_z|\sigma_x)$$

As, in the case under consideration, the equal sign holds, (see Eq. (4)), either all the terms $p^{(2)}(\sigma_z|\sigma_y)$ are equal or, for each x, all but one of the set $\left\{p^{(1)}(\sigma_y|\sigma_x)\right\}_{y=1,2,\ldots M}$ are equal to zero. The first possibility is to be discarded for it would imply that the input and the output of channel 2 are independent. Thus we conclude that $\left[p^{(1)}(\sigma_y|\sigma_x)\right]$ is a unit matrix (more precisely, it can be changed into a unit matrix by a suitable reordering of its rows and columns) hence the channel is noiseless.

q.e.d.

<u>Second case $C_1 < C_2$.</u>

We have to show that if $C_e = C_1$, channel 2 is noiseless.

From the results of the preliminary discussion, it is clear that the optimum input probability of the equivalent channel is identical to that of channel 1. As a result, in both cases, the entropy of the input symbol σ_x is the same.

Since

$$C_e = H(\sigma_x) - H(\sigma_x \mid \sigma_z)$$

and

$$C_1 = H(\sigma_x) - H(\sigma_x \mid \sigma_y)$$

then

$$H(\sigma_x \mid \sigma_z) = H(\sigma_x \mid \sigma_y) \tag{7}$$

By the theorem on total probability, we have

$$r(\sigma_x \mid \sigma_z) = \sum_y r^{(2)}(\sigma_y \mid \sigma_z)\, r^{(1)}(\sigma_x \mid \sigma_y)$$

where we use the letter r to distinguish, from the transition probabilities, the conditional probabilities of the input given the output. Thus, using inequality (6), we obtain

$$\begin{aligned} H(\sigma_x \mid \sigma_y) &= \sum_{y,z} p(\sigma_z)\, r^{(2)}(\sigma_y \mid \sigma_z) \sum_x F\left[r^{(1)}(\sigma_x \mid \sigma_y)\right] \\ &\leqslant \sum_{x,z} p(\sigma_z)\, F\left[\sum_y r^{(2)}(\sigma_y \mid \sigma_z)\, r^{(1)}(\sigma_x \mid \sigma_y)\right] \end{aligned}$$

that is

$$H(\sigma_x \mid \sigma_y) \leqslant H(\sigma_x \mid \sigma_z)$$

We know from (7) that the equal sign holds. Therefore, in the light of the previous discussion, the only possibility left is that the matrix $\left[r^{(2)}(\sigma_y \mid \sigma_z)\right]$ is a unit matrix, (again, here, some reordering of the rows or columns might be necessary). In addition Bayes' theorem states that

$$r(\sigma_y \mid \sigma_z) = \frac{p(\sigma_y)\, p^{(2)}(\sigma_z \mid \sigma_y)}{\sum_{y'} p(\sigma_{y'})\, p^{(2)}(\sigma_z \mid \sigma_{y'})}$$

Therefore the matrix $\left[p^{(2)}(\sigma_z \mid \sigma_y)\right]$ is also a unit matrix. Thus the second channel is noiseless.

q.e.d.

The following example shows the necessity of the assumption that the transition probability matrices are non-singular.

Consider two channels, I and II, having the respective transition probability matrices

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

The channel capacities are respectively $C_1 = 1.32$ bits/symbol and $C_2 = 1$ bit/symbol.

It can be easily verified that if the input symbols of the cascade are equally probable, the rate of reception of information through the cascade is equal to 1 bit/symbol, that is equal to the channel capacity of channel 2, although channel 1 is noisy.

The theorem just proved is, of course, in accordance with our intuitive feeling which is that each time a signal goes through a noisy channel the equivocation must be increased. It supports also the empirical notion that in a communication system consisting of cascaded channels, for a specified quality of transmission through the system, each channel must satisfy more rigorous requirements than the system itself.

A very obvious consequence of Shannon's fundamental theorem is that if, in contrast with what was assumed in section 2.2, the intermediate stations were allowed an infinite delay before retransmitting any signal, the rate of reception of information through the whole cascade could become arbitrarily close to the smallest channel capacity of the cascaded channels.

2.4 Cascade of Repeaters

The type of intermediate station operation assumed in section 2.2, caused in each channel, an additional delay equal to the length of the signal used. In certain cases, this cumulative delay may be undesirable. It is therefore of interest to consider a case where this delay is reduced to a minimum. In particular we wish to consider here the case where the signals are retransmitted exactly as they are received.

Let us assume that all channels are bandlimited and have the same bandwidth W. Thus the signals are completely defined by a sequence of equidistant samples taken at a rate of 2W samples per second. For simplicity let us assume that the intermediate stations operate as repeaters, that is retransmit the signal sample by sample exactly as it has been received.

Thus in order to obtain the input-output statistics of the cascade we need only to consider the signal one sample at a time.

The samples x of the first transmitter belong to an ensemble completely specified by the probability density p(x). The sample x will travel down the first channel and, because of the noise, will be received as y_1 by the first intermediate station, as y_2 by the second intermediate station, and finally, as y_n by the last receiver.

Each channel is represented by a conditional probability density; for the k^{th} channel $p^{(k)}(y_k|y_{k-1})$ gives the probability distribution of the samples y_k, received by the k^{th} intermediate station, on the condition that y_{k-1} was received at and transmitted by the preceding station. Again we use the concept of equivalent channel which, in this case, has the sample $\underline{x}$ as input and the sample $\underline{y}_n$ as output. It will be completely defined by the transition probability density $p(y_n|x)$.

The results of the discrete case may be immediately extended to the continuous case: thus we obtain

$$p(y_n|x) = \int dy_1 \int dy_2 \cdots \int dy_{n-1}\, p^{(1)}(y_1|x)\, p^{(2)}(y_2|y_1) \cdots p^{(n)}(y_n|y_{n-1}) \quad (8)$$

where the integrations are carried out over the whole range of the variables.

This result is based on the assumption of the independence of the noise in different channels and in successive samples but is otherwise absolutely general.

Special case of additive noise

In a large number of applications, though not always, the noise may be represented as a random variable added to the signal.

Under these conditions we may write

$$p^{(k)}(y_k | y_{k-1}) = f^{(k)}(y_k - y_{k-1}) \qquad (k = 1, 2, \ldots n)$$

Substituting into Eq. (8), we see that $p(y_n|x)$ is the result of $\underline{n}$ successive convolutions and therefore, also

$$p(y_n | x) = f(y_n - x)$$

These results may be expressed in a more elegant form. Let $\phi^{(k)}(t)$ be the "characteristic function" of the distribution $f^{(k)}(u)$. It is defined(18) as

$$\phi^{(k)}(t) = \int_{-\infty}^{\infty} f^{(k)}(u)\, e^{iut}\, du$$

It immediately follows that

$$\phi(t) = \phi^{(1)}(t)\, \phi^{(2)}(t) \ldots \phi^{(n)}(t)$$

where $\phi(t)$ is the characteristic function of the distribution $f(u)$ relative to the equivalent channel.

We therefore state the following:

<u>Theorem</u>: If the noise in each channel is independent and additive, the characteristic function of the noise for the equivalent channel is equal to the product of the characteristic functions for each individual channel.

It is evident that the properties of the equivalent channel are independent of the order of the channels in the cascade.

In this connection it is worth recalling that[18] the mean square deviation of the sum of n independent random variables is equal to the sum of the mean square deviations of each random variable.

2.6 Pulse Code Modulation in Cascaded Channels

By pulse code modulation we mean a coding method in which the signals consist of a succession of pulses of standard shape and of either polarity more precisely a pulse code modulation of order k has an alphabet of 2^k symbols, each symbol being represented by a particular sequence of <u>k</u> pulses.

Since we assumed that the noise affects each pulse independently of the way it affected the previous pulses and since in a P.C.M. system of order k the sign of a pulse is independent of the sign of all preceding pulses, the amount of information obtainable from a symbol of a k-order code is k times the amount of information obtainable from a single pulse.

If we assume that for a single pulse the transition probability matrix is

$$\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

the amount of information obtainable from a symbol of a k-order code is then[7,10]

$$k\left[1-f(p)\right]$$

where

$$f(p) = -p\log_2 p - (1-p)\log_2(1-p)$$

If we consider a cascade of two channels with the respective probabilities of error p_1, p_2 it is easily recognized that the equivalent channel probability matrix may be written as

$$\begin{bmatrix} 1-p_e & p_e \\ p_e & 1-p_e \end{bmatrix}$$

where p_e is given by

$$1-2p_e = (1-2p_1)(1-2p_2)$$

In the case of a cascade of n channels, in which the i^{th} channel has the probability of error p_i, we would have

$$1-2p_e = \prod_{i=1}^{n}(1-2p_i) \qquad (9)$$

2.7 A Cascade of Repeaters and a P.C.M. System

2.71 General assumptions.

In this section we compare the behaviour of cascaded continuous channels and cascaded P.C.M. channels operating with the same average transmitter power. The noise power spectrum is the same in both cases. The conditions that have to be imposed in order to obtain a meaningful comparison are not obvious, therefore we consider two cases: in the first, the two systems have a common average transmitter power and a common bandwidth and in the second, the bandwidth of the P.C.M. system is increased so that a single channel of either system has about the same channel capacity.

For simplicity, we assume that the noise is gaussian and has a flat spectrum and that it is additive to the signal. In this connection it might be worth while to point out the shot noise and the resistance noise have been shown[19,20] to be gaussianly distributed and to have a flat spectrum at least up to frequencies higher than any yet of importance in communication work.

Let N_0 be the noise power per cycle, so that with a bandlimited channel of bandwidth W, the noise power is N_0W. Let S be the average signal power received.

2.72 Cascade of continuous channels.

The noise in each channel (of bandwidth W) is gaussian and additive to the signal as specified in section 2.71. We assume that each intermediate station operates as a repeater, i.e., it retransmits a sample identical to that received. If n identical channels are so cascaded and if the noise power per cycle is N_o in each channel, the noise power per cycle in the equivalent channel is nN_o. Therefore the maximum amount of information receivable through the cascade is

$$\frac{1}{2} \log_2 \left(1 + \frac{S}{nN_oW}\right) \quad \text{bits per sample.} \tag{10}$$

2.73 <u>The Cascade of P.C.M. Channels</u>

The average transmitter power will be S as for the continuous channels. If the integer k is the order of the code, the bandwidth is chosen to be kW, so that the rate at which the continuous channel transmits its samples is equal to the rate at which the k-order P.C.M. symbols are transmitted. Thus the signal to noise ratio becomes $\frac{S}{kN_oW}$ for each channel. The noise samples will have a mean square deviation $N = N_okW$ and a probability density

$$\frac{e^{-\frac{n^2}{2N}}}{\sqrt{2\pi N}}$$

provided we select units such that the amplitudes of the transmitted pulses are $\pm\sqrt{S}$. The probability of error p is then:

$$p = \int_{\sqrt{S}}^{\infty} \frac{e^{-\frac{n^2}{2N}}}{\sqrt{2\pi N}}\, dn = \int_{\sqrt{\frac{S}{N}}}^{\infty} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\, dz \tag{11}$$

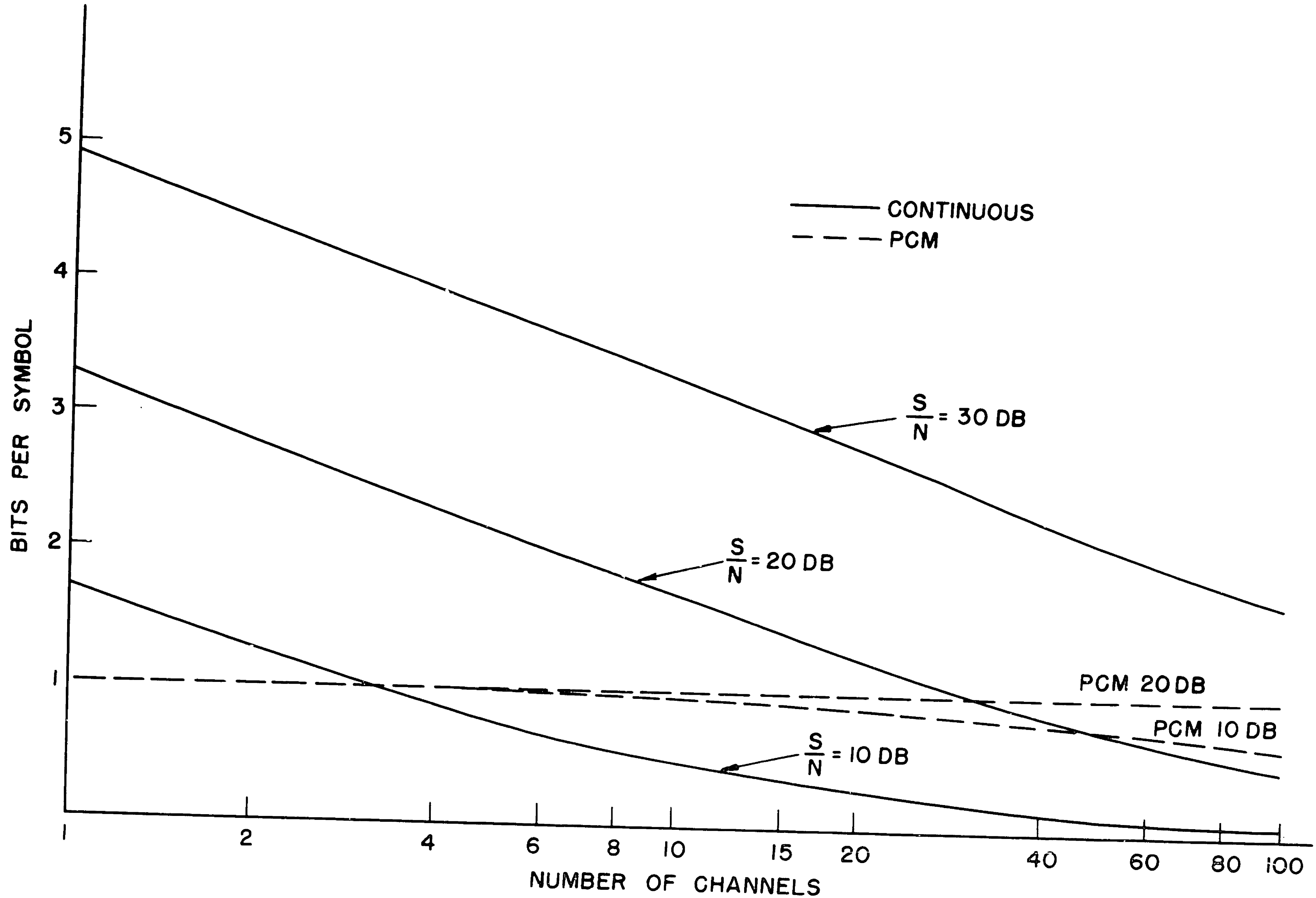

CONTINUOUS
PCM
$\frac{S}{N}$ = 30 DB
$\frac{S}{N}$ = 20 DB
$\frac{S}{N}$ = 10 DB
PCM 20 DB
PCM 10 DB
BITS PER SYMBOL
NUMBER OF CHANNELS
1
2
3
4
5
1
2
4
6
8
10
15
20
40
60
80
100

FIG. II, 1

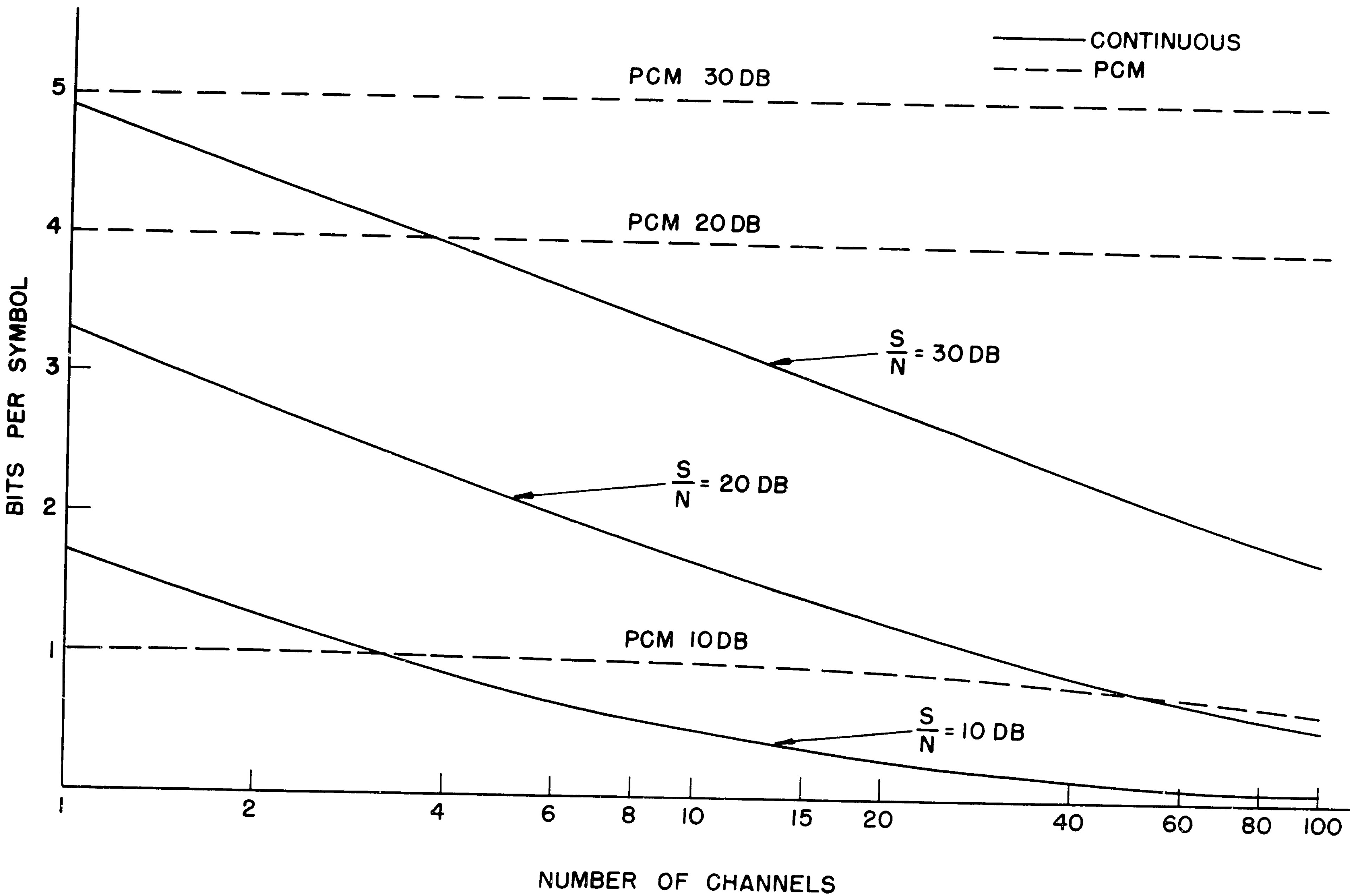

CONTINUOUS
PCM
PCM 30 DB
PCM 20 DB
PCM 10 DB
S/N = 30 DB
S/N = 20 DB
S/N = 10 DB
BITS PER SYMBOL
1
2
3
4
5
1
2
4
6
8
10
15
20
40
60
80
100
NUMBER OF CHANNELS

FIG. II, 2

The channel capacity of the equivalent channel is

$$k\left[1 - f(p_e)\right] \qquad \text{bits per symbol}$$

and p_e is given by Eq. (9).

2.74 Comparison

Case I. Both systems have the same bandwidth W, therefore $k = 1$. The numerical results are presented in Fig. II,1. As long as $\frac{S}{N}$ is equal to 20 db or higher the P.C.M. cascade has, for all practical purposes, a channel capacity of one bit per symbol. Indeed when $\frac{S}{N} = 20$ db the parameter P of a channel is equal to $7.66\ 10^{-24}$. For $\frac{S}{N} = 10$ db, the decrease in the channel capacity becomes appreciable already for $n = 20$. The channel capacity of the continuous case decreases appreciably as n increases as expected from Eq. (10).

Case II. The order k of the P.C.M. system is selected so that a single channel of either system has about the same channel capacity. (The average transmitter power and the noise power spectrum are the same in both cases.) The results are presented in Fig. II,2.

In the writer's opinion the superior performance of the P.C.M. can only be ascribed to the sample by sample requantization of the signal. In the P.C.M., the detector carries out a ruthless elimination of noise. In some rare instances, the noise sample is so large that the detector is misled. The point is that as long as these instances are very infrequent there is only a very slight loss in the quality of the system as more and more channels are cascaded.

CHAPTER III

THE INFLUENCE OF DELAY AT THE INTERMEDIATE STATION

3.0 Introduction

The examples of the previous chapter indicate without any doubt that the operation of the intermediate station is a very important factor in the system performance. For example, if, in a cascade of P.C.M. channels, the intermediate stations did not requantize the samples but retransmitted them as they were received, it is clear that the probability p_e, relative to the equivalent channel, would have been much larger than that given by Eq. (II,9) and consequently the system performance would have been very much poorer. The intermediate station may operate on one sample at a time or on groups of samples, in the latter case the signal will experience a certain amount of delay. Intuitively we feel that the larger these groups of samples, the greater will be the improvement in the performance of the system, provided suitable signals are used. Under these conditions, if delay is allowed at the intermediate station, the set of a-posteriori probabilities obtained after decoding will usually be very peaked. As a consequence, if the symbol which has the largest a-posteriori probability is retransmitted, the intermediate station retransmits with a relatively small amount of information (namely that necessary to specify that symbol) a relatively good description of the set of a-posteriori probabilities. If on the other hand, the intermediate station retransmits the received signal, exactly in the form in which it has been received, it essentially retransmits data from which the whole set of a-posteriori probabilities may be obtained. This procedure corresponds to retransmitting a large amount of information (usually, it is infinite) and the corresponding rate of re-transmission is, usually, much larger than the channel capacity. As a result a large fraction of the retransmitted information will be lost and, at the second receiver, the set of a-posteriori probabilities will convey much less information (about what was originally transmitted) than the set of a-posteriori probabilities that would have been obtained if the signal of maximum a-posteriori probability (at the intermediate station) would have been retransmitted.

In fact, the problem of representing in a convenient form, information conveyed by a set of a-posteriori probabilities is still unsolved. However it is possible that some future advanced in the theory will, in some cases, show how to represent, by a selection from a finite set, the information contained in a set of a-posteriori probabilities.

Thus in the present state of the theory it appears that, in a cascade of channels, the per-unit equivocation, in each channel, must be kept as small as possible. And the "suitable" signals are those signals which allow information to be transmitted in the channel at a high rate while keeping the per-unit equivocation smaller than a prescribed amount. This is the coding problem which must be faced each time one has to communicate information through noise. This problem will not be solved here. Only two types of channel are considered: the first is a continuous channel in which the noise is gaussianly distributed, additive to the signal and has a flat spectrum, and the second is the usual binary channel.

In the case of the continuous channel, two sets of signals are indicated and the most efficient one is exclusively used in the discussion.

In the binary channel case, a coding procedure is constructed on the general idea of error correcting codes[21,22]. These codes are the only ones considered in the discussion.

In both cases, the codes are proved to be optimum in the limit of very long signals.

3.1 The Continuous Channel

3.11 Definition of the channel.

Consider a channel of bandwidth W in which the noise is gaussian distributed, additive to the signal and, as usual independent of the signal. In addition, let the noise spectrum be flat and the average noise power be N. For convenience let $N_o = N/W$.

The signals used are of duration T and have an energy ST so that S is the average signal power. Since the channel is bandlimited, we may represent the signals by a sequence of 2TW samples. The signals may be thought of as vectors in a 2TW dimensional space.[8] For all practical purposes, the scalar product of two such vectors is equal to the cross-correlation (without delay) of the corresponding time functions.* In this representation, the noise samples are gaussian random variables of zero mean and of mean square deviation equal to N.

This type of channel has already been given considerable attention, both because it is a good model for many channels encountered in practice and also because it is convenient to discuss mathematically.[8,26,27,28,29] Shannon discussed the problem from a geometrical point of view.[8]

He showed that the transition probability from one point in signal space to another point depends only on the distance, say d, between these two points. On the other hand, as the average power of the signal is fixed, the signal points lie on the surface of a hypersphere and, consequently, in the expression of d^2, the only term which can vary is the double product term, that is the double scalar product of the two signal vectors. Thus to obtain the transition probabilities from one point to another or to obtain (by using Bayes'

* These two quantities are not rigorously equal. This is related to the well-known fact that a function of time cannot at the same time be bandlimited and be different from zero only in a finite time interval. This question is completely discussed in reference 29.

theorem) the a-posteriori probabilities we need only to carry out the cross-correlations (without delay) between the received signal and all the possible transmitted signals.[12,27]

It appears then that the signal points should be chosen as far apart from each other as possible. Therefore, it is expected that a highly symmetrical configuration of points in signal space might constitute an efficient set of signals. It is natural therefore to investigate the regular polytopes as possible configurations of signal points.

3.12 <u>Signals based on regular polytope configurations</u>.

For channels defined in section 3.11, M. J. Golay[26] has shown that, for a fixed average transmitter power, a P.P.M. system will achieve the maximum rate of reception with a vanishingly small per-unit equivocation in the limit of infinitely large bandwidths and infinitely large signals. This result may be extended by the same technique to the much larger class of orthogonal signals.[29] In this case, the received signal is cross-correlated with all the $M = 2TW$ signals of the transmitter's alphabet and the probability P_e, that the signal to which corresponds the largest cross-correlation coefficient is not the actually transmitted signal, satisfies the inequality

$$P_e < 2\psi(\beta r) \tag{1}$$

where

$$r = \sqrt{\frac{2TS}{N_0}} \tag{2}$$

$$\beta r = \frac{r}{2}\left[1 - 2\,\frac{\log(M-1)}{r^2}\right] \tag{3}$$

and
$$\psi(x) = \int_x^\infty \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,dt \tag{4}$$

When $\beta r \gg 1$, we shall often use the first term of the asymptotic expansion of $\psi(\beta r)$ and write

$$P_e < 2\,\frac{e^{-\frac{\beta^2 r^2}{2}}}{\sqrt{2\pi}\,\beta r} \tag{5}$$

Considering (1) and (3), it is clear that P_e will go to zero, in the limit of $T \to \infty$, only if we have

$$2\,\frac{\log(M-1)}{r^2} < 1 \tag{6}$$

If this inequality is satisfied, then, in the limit of $T \to \infty$, the rate of transmission of information (assuming that all signals have equal a-priori probabilities) will be smaller than the channel capacity.

Let us reformulate Golay's results in a slightly different way in order to make easier a discussion of the asymptotic behaviour of other sets of signals.

The assumptions are

(1) the channel under consideration is defined in section 3.11

(2) the number of signals, M, satisfies the inequality (6)

(3) the cross-correlation coefficients (C_1, C_2, $\cdots C_M$) of the received signal with the M possible transmitted signals are such that

$$c_i = m_i \qquad (i = 1, 2, \ldots, t-1, t+1, \ldots M)$$

$$c_t = \sqrt{\frac{2TS}{N_0}} + m_t$$

where the subscript t refers to the actually transmitted signal, and the numbers m_1, m_2, $\cdots\cdot m_t$, $\cdots\cdot m_M$ are gaussian random variables of unit dispersion

(4) the output of the channel is the signal which has the largest cross-correlation coefficient with the received signal.

Then when $T \to \infty$ the probability of error and, therefore the per-unit equivocation goes to zero. If we make the additional assumption that, in the limit of $T \to \infty$, $2\frac{\log M}{\kappa^2}$ is arbitrarily close to unity, then the rate of <u>transmission</u> of information is arbitrarily close to the channel capacity and as the per-unit equivocation is zero (in the limit), the rate of <u>reception</u> of information is arbitrarily close to the channel capacity.

In n-dimensional space, when $n \geqslant 5$, there are only three kinds of regular polytopes; the simplest is the regular simplex.[23] It has $n + 1$ vertices $S^{(1)}$, $S^{(2)}$, ... $S^{(n+1)}$ joined by $\frac{n(n+1)}{2}$ edges so that any vertex is connected to all other vertices by an edge of the polytope. In two dimensions, the regular simplex is the equilateral triangle, in three dimensions the regular simplex is the regular tetrahedron.

Suppose we choose as signal points the vertices of a regular simplex in the 2TW dimensional space, thus $n = 2TW$. Let the signals be $\vec{S^{(1)}}$, $\vec{S^{(2)}}$, $\vec{S^{(n+1)}}$.

Since S is the average signal power we must have

$$\vec{S^{(k)}} \cdot \vec{S^{(k)}} = 2TWS \qquad (k = 1, 2, \ldots, n+1)$$

and for $j \neq k$

$$\vec{S^{(k)}}\,\vec{S^{(j)}} = -S$$

since for any regular simplex[23]

$$\frac{\vec{S^{(j)}}\;\vec{S^{(k)}}}{\vec{S^{(k)}}\;\vec{S^{(k)}}} = -\frac{1}{n} \qquad \text{if } k \neq j.$$

For a particular orientation of the polytope, the coordinates of the k^{th} vertex, i.e., the samples of the k^{th} signals, are[23]

$$S^{(k)}_{2r-1} = \sqrt{2S}\,\cos\frac{2rk\pi}{n+1}$$

$$S^{(k)}_{2r} = \sqrt{2S}\,\sin\frac{2rk\pi}{n+1} \qquad \left(r = 1, 2, \ldots \left[\tfrac{n}{2}\right]\right)$$

$$S^{(k)}_{n} = (-1)^k\sqrt{S} \qquad \text{if } n \text{ is odd}$$

Let $\vec{r}$ be the received signal, then according to the assumption of section 3.11, we may write

$$\vec{r} = \overrightarrow{S^{(t)}} + \vec{n}$$

where $\overrightarrow{S^{(t)}}$ is the transmitted signal and $\vec{n}$ the noise vector. The components of $\vec{n}$ are gaussian random variables of probability density

$$\frac{e^{-\frac{u^2}{2N}}}{\sqrt{2\pi N}}$$

Let us introduce the matrix $[D]$ defined by its elements

$$d_{ik} = \frac{S_k^{(i)}}{\sqrt{2TWSN}} \qquad (i, k = 1, 2, \ldots M)$$

Suppose that the detector carries out the cross-correlations (without delay) corresponding to the product

$$[D] \cdot r] = c]$$

Then

$$c_i = \frac{1}{\sqrt{2TWSN}} \left[\overrightarrow{S^{(i)}} \cdot \overrightarrow{S^{(t)}} + \overrightarrow{S^{(i)}} \cdot \vec{n} \right] = -\sqrt{\frac{S}{2TWN}} + m_i \qquad (i \neq t)$$

$$c_t = \frac{1}{\sqrt{2TWSN}} \left[\overrightarrow{S^{(t)}} \cdot \overrightarrow{S^{(t)}} + \overrightarrow{S^{(t)}} \cdot \vec{n} \right] = \sqrt{\frac{2TS}{N_0}} + m_t$$

where m_i and m_t are gaussian random variables of unit dispersion.

As T increases indefinitely $\sqrt{\frac{S}{2TWN}}$ becomes vanishingly small. Then, it is clear that, as T and W increase indefinitely so that $2 \frac{\log M}{n^2}$ is very close to unity (although smaller than unity), the per-unit equivocation goes to zero and the rate of reception of information will become very close to the maximum rate.

In n-dimensional space, the next regular polytope is the "regular crosspolytope" which has 2n vertices. In two dimensions the regular crosspolytope is the square, in three dimensions, it is the regular octahedron. Any vertex $B^{(k)}$ (where k $1, 2, \cdots, 2n$) is joined to all other vertices except one, denoted by $B^{(k \pm n)}$, (where the $+$ sign holds for $k < n + 1$, and the $-$ sign for $k > n$) by an edge of the polytope, and, as can be easily verified in the case of the square and of the octahedron, we have in general

$$\overline{B^{(k)} B^{(k \pm n)}}^{\,2} = 2a^2 \tag{7}$$

where a is the length of the edge of the regular polytope. If we consider the vectors joining the center of the polytope, say O, to the vertices, we obtain a set of 2n vectors $\overrightarrow{OB^{(1)}}$, $\overrightarrow{OB^{(2)}}$, ... $\overrightarrow{OB^{(2n)}}$. It can be verified that each vector is orthogonal to all others but one; more precisely, $\overrightarrow{OB^{(k)}}$ is orthogonal to all vectors but $\overrightarrow{OB^{(k \pm n)}}$. It follows from (7) that $\overrightarrow{OB^{(k)}}$ and $\overrightarrow{OB^{(k \pm n)}}$ are directly opposite. Thus the set of vectors $\overrightarrow{OB^{(1)}}$, $\overrightarrow{OB^{(2)}}$, ... $\overrightarrow{OB^{(2n)}}$ consists of n mutually orthogonal vectors $\overrightarrow{OB^{(1)}}$, $\overrightarrow{OB^{(2)}}$, ... $\overrightarrow{OB^{(n)}}$ and their opposites.

Let us consider then a matrix $[B]$ defined by its element

$$b_{ik} = \frac{b_k^{(i)}}{\sqrt{2TWSN}}$$

Let $\vec{r}$ be the received signal and $\overrightarrow{b^{(t)}}$ be the actually transmitted signal. Suppose that at the receiver, the computer element carries out the cross-correlations corresponding to the product

$$[B] \cdot r] = c]$$

Then

$$c_i = \frac{1}{\sqrt{2TWSN}} \left(\overrightarrow{b^{(i)}} \cdot \overrightarrow{b^{(t)}} + \overrightarrow{b^{(i)}} \cdot \vec{n} \right) = m_i \tag{8}$$

$$c_t = \frac{1}{\sqrt{2TWSN}} \left(\overrightarrow{b^{(t)}} \cdot \overrightarrow{b^{(t)}} + \overrightarrow{b^{(t)}} \cdot \vec{n} \right) = \sqrt{\frac{2TS}{N_0}} + m_t \,. \tag{9}$$

$$c_{t+n} = \frac{1}{\sqrt{2TWSN}} \left(\overrightarrow{b^{(t)}} \cdot \overrightarrow{b^{(t\pm n)}} + \overrightarrow{b^{(t\pm n)}} \cdot \vec{n} \right) = -\sqrt{\frac{2TS}{N_0}} + m_{t\pm n} \tag{10}$$

where m_i, m_t and $m_{t\pm n}$ are gaussian random variables of unit dispersion.

On the basis of the previous discussion we conclude that: when T and W increase indefinitely so that $2 \log \frac{4TW}{r^2}$

becomes arbitrarily close to one, the rate of reception of information is arbitrarily close to the channel capacity.

In the discussion that follows only crosspolytope type signals will be used.

For completeness it should be pointed out that the third kind of regular polytope is of no interest to us. This regular polytope has, in n-dimensional space, 2^n vertices which, for a particular orientation of the coordinates system, might have ($\pm$ 1, $\pm$ 1,··· $\pm$1) as coordinates. It is obvious then that the minimum distance between two vertices is independent of the number of dimensions <u>n</u>. Thus the probability of error will not go to zero as $n \to \infty$.

3.13 <u>Transition probability matrix of a channel</u>.

We consider the channel, defined in section 3.11, in which we use signals of the crosspolytope type. We further assume that, in the receiver, the cross-correlations specified in section 3.12 are performed and that the output of the receiver is the signal which has the largest cross-correlation coefficient with the received signal.

We shall use the approximate value for the probability of error given by Eq. (1). But, in order to obtain the transition probability matrix, we must look into the problem in more detail because we are now interested in the relative frequency of the various possible ways in which an error may occur.

It has not been possible to arrive at exact expressions for the elements of the transition probability matrix. It should be stressed, however, that, from a practical point of view, only those cases where the probability of error is small are of interest and that it is even more so if the channels will eventually belong to a cascade. Indeed it is well known that for a given quality of overall transmission the requirements on each channel become more severe as the number of channels increase.

In view of Eqs. (1), (2) and (3) it is then reasonable to assume $\beta r \gg 1$, which implies also $r \gg 1$. It is clear that, since $r \gg 1$, the probability that $c_{t \pm n}$ will be the largest number of the set c_i $(i = 1, 2 \cdots M)$ is very much smaller than the probability that c_k, $(k \neq t$ and $k \neq t \pm n)$, be the largest number of the set c_i ; this follows immediately from the Eqs. (8), (9) and (10). Moreover these relations show that the probability that c_k, $(k \neq t$ and $k \neq t \pm n)$, be the largest number of the set c_i is independent of k. Therefore the transition probability matrix may be approximated by the following M by M matrix:

$$\begin{bmatrix}
1-a & p & \cdots & p & 0 & p & \cdots & p \\
p & 1-a & & & p & 0 & & \\
\vdots & & \ddots & & \vdots & & \ddots & \vdots \\
p & & \cdots & 1-a & p & & \cdots & 0 \\
0 & p & \cdots & p & 1-a & p & & p \\
p & 0 & & & p & 1-a & & \\
\vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\
p & & \cdots & 0 & p & & \cdots & 1-a
\end{bmatrix} \qquad (11)$$

where a = (M - 2)p.

By symmetry, the parameter has the approximate value

$$p \simeq \frac{2\,\psi(\beta r)}{M-2} \tag{12}$$

3.14 Transition Probability Matrix of the Equivalent Channel.

Consider a cascade of n identical channels of the type defined in section 3.11. Each one of them is supposed to be operated as described in the previous section, thus at each intermediate station the symbol most likely to have caused the received signal is the one which is retransmitted. Each channel is then described by a matrix such as that given by (11).

The equivalent-channel transition-probability matrix is equal to the product of the transition-probability-matrices of the individual channels. It is easily seen that the two diagonals of zeros, present in each factor, will not be present in the product. In order to obtain simple formulas, let us make a slight approximation: let us replace in each matrix the zeros by a "p." This essentially replaces each channel by a channel of slightly lower quality. The form of the new matrices is left intact when one of them is multiplied by any other of the same form.

In order to find the product of the matrices we only need to determine the value p_e of the parameter of the equivalent-channel transition-probability matrix. As these matrices are symmetrical and commutable, we need only to determine their characteristic values. It is shown in Appendix III-A that the characteristic values of a matrix $[T]$ defined by its element

$$t_{ik} = \left[1-(M-1)p\right]\delta_{ik} + p \qquad (i, k = 1, 2, \dots M)$$

are 1 and 1 - Mp with the respective multiplicities 1 and M - 1.

It follows that the equation for p_e is

$$1 - Mp_e = \prod_{i=1}^{n} (1 - Mp_i) \tag{13}$$

where p_i is the parameter of the i^{th} channel.

In the special case of a cascade of identical channels we have

$$Mp_e = 1 - (1 - Mp)^n \tag{14^a}$$

or in series form

$$Mp_e = \binom{n}{1} Mp - \binom{n}{2} M^2p^2 + \binom{n}{3} M^3p^3 - \cdots \tag{14^b}$$

3.15 <u>Capacity of the channel</u>.

The symmetry of the transition probability matrix (11) requires that the input probability of the symbols which will maximize the rate of reception of information is uniform. Thus the channel capacity is

$$C = \log M + (M-2)p \log p + \left[1-(M-2)p\right] \log \left[1-(M-2)p\right]$$

It should be stressed that this expression is approximate since it is based on the expression (11) of the transition probability matrix which itself is approximate. Often it is more convenient to consider the equivocation

$$I_E = -(M-2)\,p \log p \;-\left[1-(M-2)p\right] \log\left[1-(M-2)p\right] \tag{15}$$

and if $Mp \ll 1$ we have, approximately,

$$I_E \simeq (M-2)\,p \log\left(\frac{e}{p}\right) \tag{16}$$

From a design point of view it is worth noting that in view of the relative insensitivity of the logarithm function on variations of its argument, roughly speaking, I_E is unchanged provided $(M - 2)p$ is kept constant.

3.16 Threshold phenomenon.

In order to be able to discuss the performance of the system when we change various parameters, such as the signal-to-noise-ratio, the length of the signal and the number of cascaded channels, we introduce a parameter μ which will be referred to as the safety factor. It is defined by the relation

$$\log_e M = \frac{ST}{\mu N_0} \tag{17}$$

That it plays the role of a safety factor is made clear once it is remembered that the signals used may achieve, in the limit, the maximum rate of reception of information only if

$$\log_e M < \frac{ST}{N_0}$$

Therefore μ measures the ratio between the maximum allowable noise power and the actual noise power. For sufficiently large bandwidths, the safety factor is approximately equal to the ratio between the channel capacity and the rate of transmission. It is to be noted that once μ and M are known, the other parameters of the channel are specified. The probability that one signal will be received in error is approximately given according to Eq. (5), by

$$P_e \simeq 2\,\frac{e^{-\frac{\beta^2 r^2}{2}}}{\sqrt{2\pi}\,\beta r} \tag{18}$$

In terms of μ and M, we have

$$\beta r = \left(1-\frac{1}{\mu}\right)\sqrt{\tfrac{1}{2}\mu\log M}$$

The sensitivity of P_e for the variations of μ is by definition

$$\mathcal{S} = \frac{\frac{dP_e}{P_e}}{\frac{d\mu}{\mu}} = -\frac{\mu}{4}\left(1-\frac{1}{\mu}\right)^2\log M - \frac{\log M}{2}\left(1-\frac{1}{\mu}\right) - \frac{\mu+1}{2\mu\left(1-\frac{1}{\mu}\right)} \tag{19}$$

The first term of (19) is $\frac{\beta^2 r^2}{2}$ so that, for any reasonably good channel, it is already of the order of 10 or more.

Rewriting (19) we get

$$\mathcal{S} = -\frac{\beta^2 r^2}{2} - \frac{\beta^2 r^2}{\mu\left(1-\frac{1}{\mu}\right)} - \frac{\mu+1}{2\mu\left(1-\frac{1}{\mu}\right)}$$

Thus the behaviour of $\mathcal{S}$ as a function of μ falls into two broad classes:

for large μ $\quad \mathcal{S} \simeq -\frac{\beta^2 r^2}{2}$

for μ close to one $\quad \mathcal{S} \simeq -\frac{\beta^2 r^2 + 1}{\left(1-\frac{1}{\mu}\right)}$

Thus if we consider different channels having the same βr we see that when μ becomes close to unity, they are very sensitive to variations in μ . Remembering Eqs. (12) and (16) we may write

$$I_E \simeq P_e \log\left(\frac{e}{P}\right)$$

and noting that the variations of the logarithmic factor are much less important than those of P_e we state that:

> For a given amount of equivocation, the sensitivity of the equivocation on variations in the safety factor μ becomes very large as μ approaches unity.

This is the well-known threshold phenomenon which is more pronounced the more complicated the coding system is and which manifests itself as the collapse of the system performance when the noise power reaches a certain critical value.

3.17 <u>The importance of the delay at intermediate stations.</u>

The fact that the system performance experiences only a slight decrease when the number <u>n</u> of cascaded channels increase, as shown by Eq. (14^b), is obtained at the cost of an increased delay. The delay between transmission and reception of the symbol is increased by at least nT where T is the duration of the signals used. On the other hand, if the delay must be kept minimum, each intermediate station must retransmit each received sample as soon as it is received; in other words the intermediate station cannot wait for a time T to decode the signal completely. Thus we shall compare the pure repeater type of system with a system in which the signals are completely decoded before transmission. In both systems, the same signals are sent by the first transmitter and the operation of the receiver of the last channel is also the same.

Thus we shall compare the pure repeater type of system with a system in which the signals are completely decoded before retransmission.

It is clear that the cause of any difference of performance between the two systems is closely related to the previously discussed sensitivity of the performance on the safety factor. Indeed, in the case of pure repeater operation, the noise encountered in each channel will add itself to the already distorted sample. As a result, everything happens as if there were only one channel in which the noise power were n times the noise power of the individual channels. In other words the safety factor of the equivalent channel is n times smaller than that of the individual channels. From the previous discussion, we expect the quality of the cascade of n repeaters to collapse as soon as n approaches the safety factor μ of the individual channels.

In order to emphasize numerically the difference in performance, the following tables give the probability that the finally received signal is in error.

In the first table a very large value of μ is taken to illustrate the importance of a complete detection of the signal at each intermediate station and to show that, in the case of repeaters, the quality of the cascade deteriorates very rapidly as the number of cascaded channels increase.

Table I

$\mu = 100$	Complete detection at each intermediate station			Repeaters	
	n=1	n=10	n=50	n=10	n=50
M = 10	$8 \; 10^{-25}$	$8 \; 10^{-24}$	$8 \; 10^{-23}$	$2.3 \; 10^{-3}$	.45
M = 100	$4.7 \; 10^{-51}$	$4.7 \; 10^{-50}$	$4.7 \; 10^{-49}$	$1.6 \; 10^{-5}$	.28
M = 1000	$1.7 \; 10^{-75}$	$1.7 \; 10^{-74}$	$1.7 \; 10^{-73}$	$1.3 \; 10^{-7}$	.13

In the second table, some less extravagant cases are presented which still exhibit the same type of behaviour.

Table II

$\mu = 20$	Complete detection at each intermediate station			Repeaters
	n=1	n=5	n=10	n=5
M = 4	$4.52 \; 10^{-4}$	$2.75 \; 10^{-3}$	$4.5 \; 10^{-3}$	.30
M = 10	$5.33 \; 10^{-6}$	$2.66 \; 10^{-5}$	$5.33 \; 10^{-5}$	.135
M = 100	$1.14 \; 10^{-10}$	$5.7 \; 10^{-10}$	$1.14 \; 10^{-9}$	.0236
M = 1000	$2.82 \; 10^{-15}$	$1.41 \; 10^{-14}$	$2.82 \; 10^{-14}$	$5.81 \; 10^{-3}$

3.2 The Discrete Case.

To discuss exhaustively the influence of delay in discrete channels is by itself a vast problem. It was decided, therefore, to consider exclusively the case of the binary channel. This decision was made for convenience and because it is felt that the binary channel is the most representative of all discrete channels.

In order to evaluate the gain in performance of a cascade when some delay is allowed at each intermediate station we must first find sets of signals which, by their nature, have some noise combatting properties. A new coding method has been devised and is described in section 3.21. In the next section it is shown that those signals provide the means for a constructive proof of Shannon's theorem. In section 3.23 it is shown how this coding method may be used for single, double and triple error correction.

3.21 Principle of the codes.

Having restricted ourselves to the binary channel, our signals will consist of sequences of binary digits. Thus the received signal will differ from the transmitted signal by some "errors." This suggests that we approach the problem of coding from the error correction point of view.[21,22] In other words, the kind of signals we are interested in are those which, by the constraints imposed on them, permit the correction of the errors, provided the number of these errors is not larger than some maximum number. M. J. Golay[21] and R. W. Hamming[22] have indicated a procedure by which a single error correcting code may be obtained. Our new method allows us to construct error correcting codes that may take care of several errors.

The problem is not solved directly: we start by solving it under restricted conditions; then a method is indicated by which this restriction may be removed.

Let us formulate the restricted problem. We suppose that the information source supplies the message in the form of k binary digits which we represent by $S_1 S_2 \cdots S_k$. (This sequence of binary digits, which will be referred to as the "sequence S," may be any one of the 2^k possible sequences of that type.) The problem is to find a sequence of ℓ binary digits (which will be referred to as the "checking sequence" or C-sequence") $C_1\ C_2 \cdots C_\ell$ to be associated to the sequence S so that, on the basis of the received sequence $S_1^r, S_2^r, \cdots S_k^r$ and of the checking sequence $C_1\ C_2 \cdots C_\ell$, we may correct all errors of the sequence S, provided the number of these errors is not larger than the integer "a."

This problem is artificial in the sense that it assumes the C-sequence to be available at the receiver, whereas in practice the code will be transmitted together with the sequence S and is therefore usually subject to errors.

In general terms, the method of solution of the restricted problem may be described as follows:

(a) A generalized matrix is defined and is used to compute the binary digits $C_1C_2\cdots C_\ell$ from the digits of the sequence S.

(b) It is assumed that, at the receiver, the same computation is carried out on the <u>received</u> sequence, that is, the sequence $S^r = (S_1^r, S_2^r\cdots S_k^r)$. The result of the computation is a set of binary digits denoted by $C_1^r, C_2^r\cdots C_\ell^r$.

(c) The comparison of the sets of binary digits $C = (C_1, C_2,\cdots C_\ell)$ and $C^r = (C_1^r, C_2^r\cdots C_\ell^r)$ provides enough information to obtain the sequence S from the <u>received</u> sequence S^r, provided the sequence S did not suffer more than "a" errors.

Let us consider <u>the double error case</u>.

In this case, we define a generalized matrix $A_{\alpha\beta h}$ where α and β range over all integers from 1 to k, and h ranges over all integers from 1 to ℓ. As will be shown later, the elements of the matrix $A_{\alpha\beta h}$ have to be either equal to zero or equal to one.

It is convenient, at this stage, to define a simplified notation. If we consider a particular value of α, say i, and a particular value of β, say j, then we may consider the sequence of binary digits

$$A_{ij1},\ A_{ij2},\ A_{ij3},\cdots\quad A_{ij\ell}$$

which is the binary representation of some number, say Q.

For simplicity, we denote this sequence by $\{A_{ijk}\}_{k=1,2,\dots,\ell}$ and we say that it "represents" the number Q.

With each pair of numbers (i, j), (where i and j are integers no larger than k, and $i < j$) we associate a number in such a way that the correspondence is one-to-one. For convenience, we assume that these numbers range from $k+1$ to $k+\binom{k}{2}$.

All the elements of the generalized matrix $A_{\alpha\beta k}$ are then defined by the following set of conditions:

D_1: For $i < j$, the sequence of binary digits

$$\{A_{ijk}\}_{k=1,2,\dots\ell}$$

"represents" the number associated to the pair (i, j).

D_2: The sequence of binary digits

$$\{A_{iik}\}_{k=1,2,\dots\ell}$$

represents the number <u>i</u>.

D_3: For $i < j$, the binary digit A_{jik} is defined by the congruence

$$A_{jik} + A_{iik} + A_{jjk} \equiv 0 \pmod 2 \qquad (k=1,2,\dots\ell)$$

As a consequence of these definitions it appears that ℓ may be chosen as the least integer such that

$$2^{\ell} > \binom{k}{1} + \binom{k}{2}$$

Let us show that if we define the C_k's and the C_k^h's by the congruences

$$C_k \equiv \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} A_{\alpha\beta k} S_\alpha S_\beta \quad (\text{mod } 2) \quad (k=1,2,\dots,\ell) \tag{20}$$

$$C_k^r \equiv \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} A_{\alpha\beta k} S_\alpha^r S_\beta^r \quad (\text{mod } 2) \quad (k=1,2,\dots \ell) \tag{21}$$

we have a double error correcting code.

(a) Suppose a single error occurred at the i^{th} position; then the received sequence is defined by

$$S_\alpha^r \equiv S_\alpha + \delta_{i\alpha} \quad (\text{mod } 2) \quad (\alpha=1,2,\dots,k) \tag{22}$$

where $\delta_{i\alpha}$ is the usual Kronecker symbol, that is $\delta_{i\alpha}=1$ if $i=\alpha$ and $\delta_{i\alpha}=0$ if $i \neq \alpha$.

Let us consider the difference $C_k^r - C_k$

$$C_k^r - C_k \equiv \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} A_{\alpha\beta k} (S_\alpha^r - S_\alpha)(S_\beta^r - S_\beta) \quad (\text{mod } 2) \quad (k=1,2,\dots \ell)$$

Taking into account (22) we get

$$C_k^r - C_k \equiv \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} A_{\alpha\beta k}\, \delta_{i\alpha} \delta_{i\beta} \quad (\text{mod } 2)$$

$$\equiv A_{iik} \quad (\text{mod } 2) \quad (k=1,2,\dots \ell)$$

And, according to D_2, the numbers A_{iik} define the position $\underline{i}$.

(b) Suppose two errors occurred respectively at the i^{th} and at the j^{th} position, where $i < j$. The received sequence is then defined by

$$S_\alpha^r \equiv S_\alpha + \delta_{i\alpha} + \delta_{j\alpha} \quad (\text{mod } 2) \quad (\alpha=1,2,\dots k) \tag{23}$$

Computing the differences $C_h^r - C_h$, we get successively

$$C_h^r - C_h \equiv \sum_\alpha \sum_\beta A_{\alpha\beta h} (S_\alpha^r - S_\alpha)(S_\beta^r - S_\beta) \pmod 2$$
$$\equiv \sum_\alpha \sum_\beta A_{\alpha\beta h} (\delta_{i\alpha} + \delta_{j\alpha})(\delta_{i\beta} + \delta_{j\beta}) \pmod 2$$
$$\equiv A_{iih} + A_{ijh} + A_{jih} + A_{ijh} \pmod 2$$

and using D_3 the last congruence becomes:

$$C_h^r - C_h \equiv A_{ijh} \pmod 2 \quad (h = 1, 2, \ldots l)$$

Referring to D1 we see that the sequence $C_h^r - C_h$ defines uniquely the error positions, namely <u>i</u> and <u>j</u>.

Let us consider the <u>triple error correcting case</u>.

First let us introduce a one-to-one correspondence between numbers, on the one hand, and all pairs (i,j) (such that $i < j$) and all triples (i,j,m), (such that $i < j < m$), on the other hand. Of course i,j,m are integers no larger than k. For convenience we assume that these numbers range from k + 1 and $\binom{k}{1} + \binom{k}{2} + \binom{k}{3}$.

All the elements of the generalized matrix $A_{\alpha\beta\gamma h}$ (where $\alpha, \beta, \gamma = 1, 2, \ldots k$ and $h = 1, 2 \ldots l$) are then defined by the following set of conditions:

D_1: The sequence of binary digits

$$\{A_{iiih}\}_{h=1,2,\ldots l}$$

represents the number <u>i</u>.

D_2: For $i < j$, the sequence $\{A_{iijk}\}_{k=1,2,\ldots \ell}$ represents the number associated to the pair (i,j).

D_3: For $i < j < m$, the sequence $\{A_{ijmk}\}_{k=1,2,\ldots \ell}$ represents the number associated to the triple (i,j,m).

D_4: For $i < j$, A_{jjik} is defined by the congruence

$$A_{jjik} + A_{iiik} + A_{jjjk} \equiv 0 \pmod 2 \qquad (k = 1,2,\ldots \ell)$$

D_5: For $i < j < m$, A_{mjik} is defined by the congruence

$$\begin{aligned} A_{mjik} + [A_{iiik} &+ A_{iijk} + A_{iimk} \\ &+ A_{jjik} + A_{jjjk} + A_{jjmk} \\ &+ A_{mmik} + A_{mmjk} + A_{mmmk}] \equiv 0 \\ &\pmod 2 \qquad (k = 1,2,\ldots \ell) \end{aligned}$$

D_6: All elements not yet defined are set equal to zero.

It is clear that ℓ may be taken as the least integer such that

$$2^\ell > k + \binom{k}{2} + \binom{k}{3}$$

Now we wish to prove that if we define the C_h's and the C_h^r's by congruences analogous to (20) and (21), namely

$$C_k \equiv \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} \sum_{\gamma=1}^{k} A_{\alpha\beta\gamma k} S_\alpha S_\beta S_\gamma \pmod 2 \qquad (24)$$

and

$$C_k^r \equiv \sum_{\alpha=1}^{k} \sum_{\beta=1}^{k} \sum_{\gamma=1}^{k} A_{\alpha\beta\gamma k} S_\alpha^r S_\beta^r S_\gamma^r \pmod 2 \qquad (25)$$

then we actually have a triple error correcting code.

(a) Suppose a single error occurred at the i^{th} position. Then the Eq. (22) holds and we obtain easily

$$C_h^r - C_h \equiv A_{iiih} \quad (\text{mod } 2) \quad (h = 1, 2, \ldots l)$$

If we refer to D_1 we see that the $(C_h^r - C_h)$'s define uniquely the i^{th} position.

(b) Suppose two errors occurred, at the i^{th} and the j^{th} positions, respectively. Let, as usual $i \quad j$. Eq. (23) holds in this case and if we compute $C_h^r - C_h$ we obtain

$$C_h^r - C_h \equiv \sum_\alpha \sum_\beta \sum_\gamma A_{\alpha\beta\gamma h} (\delta_{i\alpha} + \delta_{j\alpha})(\delta_{i\beta} + \delta_{j\beta})(\delta_{i\gamma} + \delta_{j\gamma}) \quad (\text{mod } 2)$$

or

$$C_h^r - C_h \equiv A_{iiih} + A_{iijh} + A_{jjih} + A_{jjjh} \quad (\text{mod } 2)$$

where we used the sifting property of the Kronecker symbol and the fact that many of the sifted terms are equal to zero according to D_6.

Remembering D_4 we get

$$C_h^r - C_h \equiv A_{iijh} \quad (\text{mod } 2) \quad (h = 1, 2, \ldots l).$$

If we refer to D_2, we see that the $(C_h^r - C_h)$'s define uniquely the positions $\underline{i}$ and $\underline{j}$.

(c) Suppose that three errors occurred, at the i^{th}, j^{th} and m^{th} positions. Let, as usual, $i < j < m$. The sequence S^r is given in terms of the sequence S by the congruences:

$$S^r_\alpha \equiv S_\alpha + \delta_{i\alpha} + \delta_{j\alpha} + \delta_{m\alpha} \pmod 2 \quad (\alpha = 1, 2, \ldots k) \tag{25}$$

If we compute $C^r_h - C_h$, using Eqs. (24), (25) and (26) we obtain

$$C^r_h - C_h \equiv \sum_\alpha \sum_\beta \sum_\gamma A_{\alpha\beta\gamma h} (\delta_{i\alpha} + \delta_{j\alpha} + \delta_{m\alpha})(\delta_{i\beta} + \delta_{j\beta} + \delta_{m\beta})(\delta_{i\gamma} + \delta_{j\gamma} + \delta_{m\gamma})$$
$$\pmod 2 \quad (h = 1, 2, \ldots l)$$

or

$$\begin{aligned} C^r_h - C_h \equiv\ & A_{iiih} + A_{iijh} + A_{iimh} \\ & + A_{jjih} + A_{jjjh} + A_{jjmh} \\ & + A_{mmih} + A_{mmjh} + A_{mmmh} \\ & + A_{mjih} + A_{ijmh} \pmod 2 \end{aligned}$$

where we used the sifting property of the Kronecker symbol and the fact that many of the sifted terms are equal to zero according to D_6.

Remembering D_5, we get

$$C^r_h - C_h \equiv A_{ijmh} \pmod 2 \quad (h = 1, 2, \ldots l)$$

If we refer to D_3, we see that the $(C^r_h - C_h)$'s define uniquely the error positions $\underline{i}$, $\underline{j}$ and $\underline{m}$.

q.e.d.

These two examples show very clearly how to construct an a-error correcting code.

First we create a one-to-one correspondence between numbers, on the one hand, and all singles $\underline{i}$, all pairs (i,j), all triples (i,j,m), ... all a-uple $(i,j, \cdots g)$ on the other hand; we assume that the integers $i,j,m \cdots g$ are not larger than k and for all the pairs $i < j$, for all the triples $i < j < m$, $\cdots$, for all the a-uples $i < j < m < \cdots < g$. For convenience we assume that the numbers used in the one-to-one correspondence range from 1 to $1 + \binom{k}{1} + \cdots + \binom{k}{a}$.

All the elements of the generalized matrix $A_{\alpha\beta \ldots \lambda h}$ (where the $\underline{a}$ subscripts $\alpha, \beta, \ldots \lambda$ range from 1 to k and h ranges from 1 to ℓ) are then defined by the following set of conditions:

D_1: The sequences of binary digits

$$\left\{ A_{iii \ldots ih} \right\}_h$$

$$\left\{ A_{iii \ldots ijh} \right\}_h \qquad (\text{where } i < j)$$

$$\left\{ A_{iii \ldots ijmh} \right\}_h \qquad (\text{where } i < j < m)$$

$$\vdots$$

$$\left\{ A_{ijm \ldots gh} \right\}_h \qquad (\text{where } i < j < m \cdots < g)$$

represent the numbers associated with the single $\underline{i}$, the pair (i,j), the triple (i,j,m), $\cdots$, the a-uple $(i,j,m, \cdots g)$ respectively.

D_2: All the elements of the matrix $A_{\alpha\beta \ldots \lambda h}$ not defined in D_1, are subjected to the only constraint that the equations defining the $(C_h^r - C_h)$'s, namely,

$$C_h^r - C_h \equiv \sum_{\alpha} \sum_{\beta} \cdots \sum_{\lambda} A_{\alpha\beta\ldots\lambda h} (S_\alpha^r - S_\alpha)(S_\beta^r - S_\beta) \ldots\ldots (S_\lambda^r - S_\lambda) \pmod 2$$

must respectively become

$$\left.\begin{aligned} C_h^r - C_h &\equiv A_{ii\ldots\ldots ih} \\ C_h^r - C_h &\equiv A_{ii\ldots\ldots ijh} \\ &\vdots \\ C_h^r - C_h &\equiv A_{ijm\ldots gh} \end{aligned}\right\} \begin{array}{l} \pmod 2 \\ (h = 1, 2, \ldots l) \end{array}$$

in the case of simple, double, triple, ... a-uple errors.

In this case, it is clear that l need not be larger than the least integer such that

$$2^l > \binom{k}{1} + \binom{k}{2} + \cdots + \binom{k}{a}.$$

The proof that the procedure just described provides an a-error correcting code is entirely analogous to that of the triple error correcting case but will not be given here.

Thus the restricted problem stated at the beginning of this section is completely solved. In the next section it is shown how the methods developed here may be used to achieve as closely as we wish the maximum rate of reception of information, in the asymptotic case of $k \to \infty$.

3.22 <u>Constructive proof of Shannon's fundamental theorem in the binary case</u>.

By binary channel we mean a discrete channel having

$$\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

as a transition probability matrix.

It is well known that the channel capacity of such a channel is $1 - f(p)$

whereas in (II,11) $f(x) = -x \log_2 x - (1-x) \log_2 (1-x)$

Suppose k is very large, then according to the law of large numbers,(14) the probability that the number of errors, say e, does not fulfill the condition

$$\left| \frac{e}{k} - p \right| < \varepsilon$$

where ε is a positive arbitrarily small number, goes to zero when $k \to \infty$.

Thus if we provide error correction for errors the total number of which is between $k(p - \varepsilon)$ and $k(p + \varepsilon)$, then, in the limit, the signal will be almost always correctly received. The number ℓ of redundant digits is the smallest integer ℓ such that

$$2^{\ell} > \binom{k}{k(p-\varepsilon)} + \binom{k}{k(p-\varepsilon)+1} + \cdots + \binom{k}{k(p+\varepsilon)}$$

Let $p' = p + \varepsilon$. The integer ℓ', defined as the smallest integer satisfying

$$2^{\ell'} > \binom{k}{k\,p'} (2\varepsilon k + 1)$$

will never be smaller than ℓ ; in other words ℓ' is an upper bound for ℓ .

For very large k, using Stirling's formula, the last inequality becomes

$$\ell' > k\,\beta(p') + \log_2(1 + 2ck)$$

and as $k \to \infty$

$$\frac{\ell'}{k} = \beta(p')$$

Thus to correct all errors in the very long message of k digits, we must transmit without errors a correcting signal $k\,\beta(p')$ digits long. We may go on repeating this process, say N times; N is bounded above by the condition that $k\,\beta^N(p')$ be large enough for the law of large numbers to be applicable.

Let us evaluate the probability that some of the first N correcting signals will fail, assuming that the $(N + 1)^{th}$ correcting signal is correctly received. This will happen when the number of errors e_λ in some one of them (whose length is for the time being represented by λ) does not fulfill the condition

$$\left|\frac{e_\lambda}{\lambda} - p\right| < \epsilon$$

The probability that this condition is not fulfilled is given by[14]

$$\Pr\left\{\left|\frac{e_\lambda}{\lambda} - p\right| > \epsilon\right\} \simeq \sqrt{\frac{2}{\pi}}\;\frac{e^{-\frac{\epsilon^2\lambda}{2pq}}}{\epsilon\sqrt{\frac{\lambda}{pq}}} \tag{27}$$

when λ is large.

For k sufficiently large, the right-hand side of (27) is very small, thus, neglecting second order terms, the probability P_e that the number of errors lies outside the prescribed intervals is

$$P_e \simeq \sqrt{\frac{2}{\pi}} \sum_{\lambda} \frac{e^{-\frac{\epsilon^2 \lambda}{2pq}}}{\epsilon \sqrt{\frac{\lambda}{pq}}} \tag{28}$$

where the summation is carried out over $\lambda = k$, $\lambda = k\rho(p')$,..... .., $\lambda = k\rho^N(p')$.

Since $\rho(p') < 1$, in the sum (28) the last term is the largest, therefore P_e has an upper bound given by

$$\overline{P}_e = N\sqrt{\frac{2}{\pi}} \frac{e^{-\frac{\epsilon^2 k \rho^N}{2pq}}}{\epsilon\sqrt{\frac{k\rho^N}{pq}}} \tag{29}$$

Up to now we have assumed that the correcting signal of length $k\rho^{N+1}$ was received without errors.

Suppose that to insure the correct reception of this last correcting signal, we repeat it $2\alpha + 1$ times. It is easy to show that the probability that this correcting signal still has an error, is bounded above by (cf. Appendix III.B.)

$$\frac{2}{\sqrt{\pi}} \alpha^{3/2} (4pq)^{\alpha} k\rho^{N+1}$$

Suppose we select α so that

$$\frac{2}{\sqrt{\pi}} \alpha^{3/2} (4pq)^{\alpha} k\rho^{N+1} = N\sqrt{\frac{2}{\pi}} \frac{e^{-\frac{\epsilon^2 k\rho^N}{2pq}}}{\epsilon\left(\frac{k\rho^N}{pq}\right)^{1/2}} \tag{30}$$

Equation (30) essentially requires that the upper bound on the probability of error of the last error correcting signal (of length $k\rho^{N+1}$) be equal to the upper bound of P_e, given by (29).

Taking the logarithm of both sides of (30), we get:

$$\alpha \log(4pq) + \frac{3}{2}\log\alpha + \log k\rho^{N+1} = -\frac{\epsilon^2 k\rho^N}{2pq} - \log\left(N\epsilon\sqrt{\frac{k\rho^N}{pq}}\right)$$

thus, for large $k\rho^N$,

$$\alpha\,|\log 4pq|\,\frac{2pq}{\epsilon^2 k\rho^N} \simeq 1$$

That is, as $k\rho^N$ goes to infinity, α is given by

$$\alpha \simeq \frac{\epsilon^2 k\rho^N}{2pq\,|\log 4pq|}$$

Thus as $k\rho^N$ goes to infinity, the length L of the signal and all the correcting signals is given by

$$L = k\left\{\frac{1-\rho^{N+2}}{1-\rho} + 2\,\frac{\epsilon^2 k\rho^N}{2pq\,|\log 4pq|}\,\rho^{N+1}\right\} \tag{31}$$

and the probability of error is smaller than $2\overline{P}_e$.

Suppose we choose to have N depend on k in such a way that

$$\rho^N \propto k^{-\frac{2}{3}}$$

Then $k\rho^N \propto k^{1/3}$ while $k\rho^{2N} \propto k^{-\frac{1}{3}}$.

Thus as $k \to \infty$, we see that, however small ϵ is, $\overline{P}_e \to 0$ (see Eq. (29)) and from (31) we get

$$\lim_{k\to\infty} L = \frac{k}{1-\rho(p')}$$

That is, in the limit, to transmit k bits we need only $\frac{k}{1-\mathcal{H}(p')}$ digits. In other words as $k \to \infty$, the probability of error goes to zero and the rate of transmission is

$$1-\mathcal{H}(p') \qquad \text{bits per digit.}$$

q.e.d.

3.23 The use of error correcting codes.

From a practical point of view it is, of course, impossible to use extremely long codes, not only because they introduce a delay (which, in a cascade of channels, will be multiplied many times) but also because they would require an impractically large amount of equipment. In this respect it should be stressed that the binary channel has an important advantage over the continuous channel, namely, that all the operations of coding are binary and thus are likely to be performed by simpler, cheaper and more rugged equipment.

First let us consider the single error correcting code. This case is interesting because the artificial restriction imposed on the coding problem in section 3.21 is easily removed. In fact the construction of single error correcting codes is well known,[21,22] nevertheless, it is of interest to obtain them as a particular case of our more general method. As in section 3.21, we assume that the information source provides a sequence of k binary digits S_1, $S_2, \cdots S_k$. On the basis of this sequence of digits, we shall compute ℓ' additional digits C_1, $C_2, \ldots, C_{\ell'}$, where ℓ' is the least integer such that

$$2^{\ell'} > k + \ell'$$

In order to obtain the C_h's ($h = 1, 2 \cdots, \ell'$) we shall define a matrix $A_{\alpha h}$ ($\alpha = 1, 2 \cdots, k$ and $h = 1, 2 \cdots, \ell'$) the elements of which are either equal to zero or equal to one.

Let B be the set of integers ranging from 1 to k but from which all the powers of 2 (that is 2^0, 2^1, 2^2, $\cdots, 2^{\ell'-1}$) have been removed. The set B contains only <u>k</u> integral numbers. The matrix $A_{\alpha h}$ is defined by the condition that each of the sequences of binary digits $\{A_{ih}\}_{h=1,2,\ldots,\ell'}$ represents a number of the set B in such a way that the correspondences between the sequences and the numbers are one-to-one.

The C_h's are computed as follows:

$$C_h + \sum_{\alpha=1}^{k} A_{\alpha h} S_\alpha \equiv 0 \quad (\mathrm{mod}\,2) \quad (h=1,2,\ldots,\ell') \tag{32}$$

The sequence S and the C_h's are then transmitted. Suppose that we receive S_1^r, S_2^r, $\cdots$ S_k^r, $C_1^r \cdots C_{\ell'}^r$. Then we compute the binary digits D_h by the congruence

$$C_h^r + \sum_{\alpha=1}^{k} A_{\alpha h} S_\alpha^r \equiv D_h \quad (\mathrm{mod}\,2) \quad (h=1,2,\ldots \ell')$$

If the error occurred at the i^{th} position of the sequence S then, referring to Eq. (32), we see that

$$D_h \equiv A_{ih} \qquad (h=1,2,\ldots \ell') \tag{33}$$

and, according to the definition of $A_{\alpha h}$, the relations (33) define uniquely the position $\underline{i}$.

If the error occurred at the j^{th} position of the sequence C, we would have

$$D_h = \delta_{jh} \qquad (h = 1, 2, \ldots, \ell')$$

which obviously defines the j^{th} position.

The two cases are differentiated by the fact that the sequence of D's given by (33) contains at least two ones. This is obvious if we remember that the set of integers B does not contain any power of two.

An obvious way to extend the error correction scheme would be to use the following method, which is discussed for the double error correcting case.

Let ℓ' be the least integer such that

$$2^{\ell'} > \binom{k+\ell'}{1} + \binom{k+\ell'}{2}$$

Suppose we define the redundant digits S_{k+1}, S_{k+2}, $\cdots$ $S_{k+\ell'}$ by the set of congruences:

$$\sum_{\alpha=1}^{k+\ell'} \sum_{\beta=1}^{k+\ell'} A_{\alpha\beta h} S_\alpha S_\beta \equiv 0 \pmod 2 \qquad (h = 1, 2, \ldots \ell') \quad (34)$$

where the elements $A_{\alpha\beta h}$ are defined as in section 3.21.

It is almost obvious, by now, that such a scheme provides double error correction for all cases provided that the system of simultaneous congruences (34) admits a solution. Examples have shown that this is not necessarily the case. To illustrate the difficulty let us consider two examples. The congruence

$$x^2 + y^2 + x + y \equiv 1 \pmod{2}$$

has no solution.

The system of congruences

$$x^2 + xy + y^2 + x + y \equiv 1 \pmod{2}$$
$$x^2 + xy \qquad + y \equiv 1 \pmod{2}$$

has no solution, although each of the equations has a solution.

Nevertheless the results obtained by solving the restricted problem may be used to devise schemes which provide double error correction, triple error correction ··· . The schemes that will be proposed have been obtained by trial and error and have been selected from many other workable schemes. These schemes are certainly not optimum but the writer believes that, probably for some range of values of k, they may turn out to be reasonably close to the optimum.

As usual let us call S the sequence of k binary digits supposed to be put out by the information source. For double error correction case, it is proposed to use as transmitted signal S, D_1, D_2, P_1 and P_2.

Where S stands for the k signal digits

D_1 stands for the digits of a double error correcting scheme applied to S, using Eq. (20).

D_2 stands for the digits obtained by the same procedure but applied to D_1

P_1 & P_2 stands for parity checks on D_1 and D_2 respectively.

For the triple error correction case, it is proposed to transmit the sequences S, T_1, T_2, D_1, D_2, P_1, P_2

where T_1 consists of the digits of a triple error correcting scheme applied to S computed by using Eq. (24)

T_2 consists of the digits of a triple error correcting scheme applied to T_1

D_1 is a double error correcting scheme applied to T_2

D_2 is a double error correcting scheme applied to D_1

P_1 & P_2 are parity checks on D_1 and D_2 respectively.

These coding schemes are used as follows: The transmitted signal consists of a succession of sequences of digits such that each sequence is deducible logically from some preceding one. The receiver verifies whether all these relations between the proper <u>received</u> sequences agree or not. For the two coding schemes proposed it can be verified that any combination of errors (provided their number is no larger than the maximum number of errors for which the code is designed) will create between the different sequences of the received signal some discordances on the basis of which the errors can be located and corrected.

For completeness, we mention here that the proposed schemes will be satisfactory only after a trivial change is made in the definition of the generalized matrices $A_{\alpha\beta k}$ and

$A_{\alpha\beta\gamma\ell}$ the question is discussed in Appendix III.C.

The method used to justify the codes presented here is indicated in Appendix III.D.

3.24 <u>The influence of the delay</u>.

It was not found possible to determine the transition probability matrix of a binary channel in which the proposed correcting codes are used. Thus the comparison is carried out on a probability-of-error basis.

In order to transmit k bits of information, we use $k' = k + \ell$ digits and if we use an a-error correcting code, the probability that the transmitted symbol is misinterpreted at the first receiver is given by

$$P = \sum_{\lambda = a+1}^{k'} \binom{k'}{\lambda} p^{\lambda} q^{k'-\lambda} \tag{35}$$

In practice, only the first term need be taken, thus

$$P \simeq \binom{k'}{a+1} p^{a+1} q^{k'-a-1}$$

and the probability that the symbol is in error, after having gone through <u>n</u> channels is approximately given by:

$$P_{e,a} \simeq 1 - (1-P)^{n} \tag{36}$$

which, if $nP \ll 1$, may be written as

$$P_{e,a} \simeq nP - \binom{n}{2} P^2 + \binom{n}{3} P^3 - \cdots$$

and if it is legitimate to take into account only the first term of (35) and (36), then

$$P_{e,a} \simeq n \binom{k'}{a+1} p^{a+1} q^{k'-a-1} \tag{37}$$

On the other hand, if a digit per digit transmission is carried out, the probability of error per symbol is:

$$P_p = 1 - (1 - p_e)^k \tag{38}$$

where p_e is given by Eq. (II,10). If the latter equation is expanded in series, we obtain after simplifications:

$$p_e \simeq np - 2\binom{n}{2}p^2 + 2^2\binom{n}{3}p^3 - \cdots$$

which combined with the expansion of (38), becomes

$$P_p \simeq nkp\left[1 - (n-1)p + \frac{k-1}{2}np\right] + nkp\,O(kn^2p^2) \tag{39}$$

It must be remembered that the reduction of the probability of error, as indicated by Eqs. (37) and (39) is achieved at the cost of three items:

(1) The rate is reduced: we need k' digits instead of k digits. However when k is fairly large, a being in practice only a few units, the relative difference between k' and k is small.

(2) The delay is increased by $nk'\frac{1}{2W}$ seconds where W is the common bandwidth of the cascaded channels.

(3) The amount of equipment is increased.

The formulas given in the discussion above may be illustrated by the following numerical examples.

<u>Example I</u> $n = 100$ $k = 100$

p	P_p	$P_{e,1}$	$P_{e,2}$	$P_{e,3}$
10^{-4}	.67	$5.67\ 10^{-3}$	$2.95\ 10^{-5}$	$2.59\ 10^{-9}$
10^{-5}	.105	$5.67\ 10^{-5}$	$2.95\ 10^{-8}$	$2.59\ 10^{-13}$
10^{-6}	.01	$5.67\ 10^{-7}$	$2.95\ 10^{-11}$	$2.59\ 10^{-17}$

Example II n 100 k 1000

p	P_p	$P_{e.1}$	$P_{e,2}$	$P_{e,3}$
10^{-4}	.730	.399	.0179	$5.77\ 10^{-4}$
10^{-5}	.630	.050	1.8 10	$5.77\ 10^{-8}$
10^{-6}	.088	$5.1\ 10^{-4}$	$1.81\ 10^{-8}$	$5.77\ 10^{-12}$
10^{-7}	.00995	$5.1\ 10^{-6}$	$1.81\ 10^{-11}$	$5.77\ 10^{-16}$

The fact that these coding procedures may actually lead to the maximum efficiency may be intuitively felt by considering the case of $p = 10^{-5}$ in the second table. A three-error correcting check produces, at the cost of a few percent increase in signal length a probability of error per message through the whole cascade nearly 200 times smaller than the probability that a single pulse is misinterpreted after going through a single channel.

It might be of interest to point out that, in the case of $p = 10^{-7}$, if the samples were repeated as they are received (that is without requantization) the probability of error of a single pulse after a couple of channels would have been already reduced to approximately 10^{-4} and after 100 channels to .37 (in those conditions the probability that a group of 100 digits is without errors is of the order of 10^{-20}!)

The threshold phenomenon[24] is also clearly exhibited in both tables: it is immediately perceived if the first and last columns are read simultaneously. Mathematically, Eq. (37) makes this threshold phenomenon obvious, and, of course the larger is "a" the more pronounced is the threshold phenomenon.

3.25 Further considerations on error correcting codes.

The use of error correcting codes increases the length of the signals. It might be of interest to consider what happens if the bandwidth is increased in such a way that the rate at which information is sent remains constant. As usual we assume that the noise is gaussian and additive, for simplicity, we also assume that its power spectrum is flat at least throughout the frequency band of interest. As a result, the noise power is increased in proportion to the increase in bandwidth.

As the channel capacity of the continuous channel, affected by gaussian additive noise, increases as W increases (the signal power S remaining constant) it might at first appear that the performance of the system under consideration should also improve as the bandwidth increases. It is found that this is not always the case. This is to be expected, since we violate the conditions required for maximum rate of received information (for the continuous channel) in at least two aspects: 1) the input probability distribution should be gaussian and 2) the detection should be done by cross-correlation. In the case under consideration, the input samples are restricted to take, with equal probability, the values ± 1 and the received signal is detected pulse by pulse.

We consider two examples, both involving cascades of 100 channels ($n = 100$).

Example I. The messages to be transmitted are coded by blocks of 40 bits at a time ($k = 40$). The signals require 46, 58 and 76 pulses for the single, double and triple error correcting codes respectively. The probabilities of error are given in the following table.

p	P_p	$P_{e,1}$	$P_{e,2}$	$P_{e,3}$
10^{-8}	$4 \; 10^{-5}$	$8 \; 10^{-10}$	$1.05 \; 10^{-11}$	$1.03 \; 10^{-10}$
10^{-6}	$4 \; 10^{-3}$	$3.2 \; 10^{-6}$	$7.7 \; 10^{-7}$	$1.92 \; 10^{-6}$
10^{-5}	$4.2 \; 10^{-2}$	$3.3 \; 10^{-4}$	$4.45 \; 10^{-5}$	$2.62 \; 10^{-4}$

Example II. The signals require 88, 101 and 124 pulses for the single, double and triple error correcting codes respectively. The probabilities of error are tabulated hereafter.

$n = 100$ $\qquad\qquad$ $k = 80$

p	P_p	$P_{e,1}$	$P_{e,2}$	$P_{e,3}$
10^{-8}	$8 \; 10^{-5}$	$3.45 \; 10^{-10}$	$1.07 \; 10^{-12}$	$8.1 \; 10^{-14}$
10^{-6}	$8 \; 10^{-3}$	$3.45 \; 10^{-6}$	$1.6 \; 10^{-8}$	$1.0 \; 10^{-8}$
10^{-5}	$8.3 \; 10^{-2}$	$1.53 \; 10^{-4}$	$2.1 \; 10^{-6}$	$1.1 \; 10^{-6}$

As k becomes larger, the increase in the number of pulses becomes relatively smaller, (for example if $k = 1000$ triple error correction is provided by an increase of 6% in length) and therefore the increase in bandwidth has less pronounced effects. Nevertheless it should be borne in mind that for $k = 100$ the increase in bandwidth has important effects and should not be neglected.

CHAPTER IV

THE OPERATION OF THE INTERMEDIATE STATION AS A DESIGN PROBLEM

4.0 Introduction

In this chapter we attempt to optimize the operation of the intermediate station. For that purpose it is convenient to define a new term. We shall call "intermediate station transfer characteristic," or for short, "transfer characteristic," the function which relates the output signal to the input signal of the intermediate station. In other words, the transfer characteristic describes mathematically what was usually called the "operation of the intermediate station." When the intermediate station operates as a repeater, i.e., retransmits the received signal as it is, the transfer characteristic is an identity operator. When the intermediate station retransmits the signal having the largest a-posteriori probability (of having been the originally transmitted one) the corresponding transfer characteristic will be called maximum a-posteriori transfer characteristic (abbreviated M.A.P.T.C.).

In the first section the criterion of design is stated and discussed. In section 2 the equations determining the optimum transfer characteristic in the general case are derived formally for a cascade of two channels. In order to obtain a soluble set of equations, the problem is, then, slightly modified and restricted to a sample by sample retransmission at the intermediate station. Under this condition, the optimum input probability distribution and the optimum transfer characteristic are obtained for the gaussian additive noise case: it is shown that the linear transfer characteristic is optimum. Next the same problem is considered in the case where the transmitter sends identical pulses of either polarity. In order to obtain soluble equations the criterion of design is modified and the transfer characteristic minimizing the probability of error is obtained numerically. The equation defining this transfer characteristic is also obtained by a simple heuristic reasoning. The difference between a maximum a-posteriori probability detector and an "optimum" detector (that is a detector which would extract all the information contained in the received signal) is computed numerically for a simple case.

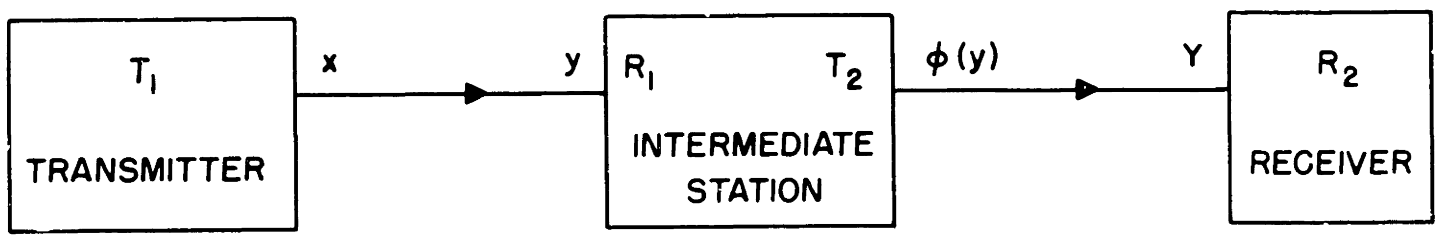
4.1 The Criterion of Design

At first sight, it might appear that the criterion of design should require the maximization of the rate of reception of information. This point of view, however, implies an unwarranted idealization: in most practical situations, we are not only interested in getting as much information (about the transmitted signals) as possible but we also require that the information received should contain most of the information transmitted, in other words we require the per-unit equivocation to be small. This is caused by the fact (already pointed out in Chapter III) that, at present, we do not know how to handle efficiently information represented by a set of a-posteriori probabilities. When the information received is represented by the member of the set having the largest a-posteriori probability, it appears that the primary factor of importance is the per-unit equivocation.

Thus the criterion that we shall use is the minimization of the per-unit equivocation which is equivalent to maximizing the information received when the information transmitted is kept constant. Of course the obtainable per-unit equivocation depends on the relative magnitude of the rate of transmission and the channel capacity in the sense that a reduction of the rate of transmission of information will reduce the per-unit equivocation.

For simplicity, we consider exclusively a cascade of two channels (see Fig. IV,1). The transmitted signal x is received by the intermediate station receiver as y. The latter signal is retransmitted by the intermediate station as a signal X which is finally received at R_2 as Y. The problem is then: given an adequate ensemble of signals x, find the intermediate station transfer characteristic which will maximize the information received. Let the amount of information (about x) supplied by Y be indicated by I(x,Y).

The quantity I(x,Y) is obtained by averaging over the ensembles of signals x and Y. In particular we may imagine that it has been obtained by averaging $I(xY \mid y_1)$, (the information about x provided by Y, when a particular y, say y_1, has been received by R_1) over the ensemble of all signals y_1. Once the



T_1
TRANSMITTER
x
y
R_1
T_2
INTERMEDIATE
STATION
$\phi(y)$
Y
R_2
RECEIVER

FIG. IV, 1

transfer characteristic is chosen, the quantity $I(xY \mid y_1)$ may be computed for any y_1 and may be considered to provide a measure for the performance of the system in that particular case. In other words, $I(xY \mid y_1)$ may be considered as a measure of the effectiveness of the "strategy" adopted; here the strategy under evaluation is the transfer characteristic. $I(xY \mid y_1)$ will therefore be referred to as the performance factor. There is no reason to believe that this performance factor has any basic significance other than that its average is equal to $I(xY)$. As a matter of fact, it is not used directly, in what follows. However, it has been found of great use in the derivation of the results that follow and for that reason it is mentioned here. It can easily be obtained from the following expressions:

$$I(x.Y/y_i) = -\sum_{x,Y} P(x.Y/y_i) \log \frac{p(x)}{r(x/Y)}$$

or
$$I(x.Y/y_i) = -\sum_{x,Y} P(x.Y/y_i) \log \frac{q_2(Y)}{t(Y/x)} \tag{1}$$

where $P(xY / y_1)$ is the probability of the pair xY when y_1 is the signal received by R_1.

If the signals x and/or the signals Y range over a continuous domain, the sums are replaced by integrals without difficulty since the integrand would then be invariant with respect to any changes of scales of either x or Y.

It is of interest to point out that, in some cases, whatever the transfer characteristic is, the performance factor $I(xY/y_i)$ will be negative for some y_i's. Consider the following example: Suppose that the input signals x have all equal a-priori probabilities and that a y exists, say y_0, such that the conditional probabilities $r(x/y_0)$ are all equal. Thus when y_0 is received by R_1, the intermediate station has received no information (about x) since the sets of probabilities $p(x)$ and $r(x/y_0)$ are identical. As a result the optimum signal that R_2 could receive from T_2 is the one that would mean "your guess is just as good as mine." Even if such a signal were transmitted by T_2, the signal will be distorted by noise and in some cases, maybe very rare, it will be transformed into some other symbol which will mislead R_2. Hence sometimes R_2 receives no information (about x) and at other times it receives some misleading information. Thus the average, for that particular y, will be negative.

4.2 The Equations Specifying the Optimum Transfer Characteristic

Suppose that both channels are bandlimited (their common bandwidth is W) and that they are affected by a continuous type of noise, in that, even if their input signals form a finite set, the received signals will form an infinite set. We assume that the alphabet, at the transmitter T_1, consists of M symbols represented by M signal-vectors $\vec{S_1}$, $\vec{S_2} \cdots \vec{S_M}$. Let $\vec{y}$ be the signal received at R_1 and $\vec{\varphi}(\vec{y})$ be the signal retransmitted by T_2. Thus the vector-function $\vec{\varphi}(\vec{y})$ completely describes the intermediate station operation and is the unknown of the present problem.

From the statistical properties of the noise, we can obtain the transition probability densities

$$p^{(1)}(\vec{y}/\vec{S_i}) \qquad \text{and} \qquad p^{(2)}(\vec{Y}/\vec{\varphi}(\vec{y}))$$

of the first and second channel, respectively.

By the theorem on total probabilities, the equivalent channel transition probability density is

$$t(\vec{Y}/\vec{S_i}) = \iint\cdots\int_D d\vec{y}\; p^{(1)}(\vec{y}/\vec{S_i})\; p^{(2)}\left[\vec{Y}/\vec{\varphi}(\vec{y})\right] \tag{2}$$

where the integration is carried out over the domain D of the signal space in which $\vec{y}$ may happen to be.

Using the following well-known expression for the information received

$$I = H(\vec{Y}) - H(\vec{Y}/\vec{S}) \tag{3}$$

we obtain

$$I = \iint\cdots\int_{D_2} d\vec{Y} \sum_i P(\vec{S_i})\; t(\vec{Y}/\vec{S_i})\; \log t(\vec{Y}/\vec{S_i})$$

$$-\iint\cdots\int_{D_2} d\vec{Y}\left[\sum_\ell P(\vec{S_\ell})\, t(\vec{Y}/\vec{S_\ell})\right] \log\left[\sum_\ell P(\vec{S_\ell})\, t(\vec{Y}/\vec{S_\ell})\right] \tag{4}$$

where $t(\vec{Y}/\vec{S_1})$ is given by equation (2), and D_2 is the domain of $\vec{Y}$. Thus the problem is to find the vector-function $\vec{\varphi}(\vec{y})$

which maximizes the amount of information I while fulfilling the power constraint imposed on the transmitter T_2:

$$\sum_{i=1}^{M} P(\vec{S}_i) \iint \cdots \int p^{(1)}(\vec{y}|\vec{S}_i) \; |\vec{\varphi}(\vec{y})|^2 d\vec{y} = P_2 \tag{5}$$

The necessary conditions for maximum I may be written, using Lagrange's method, (see Appendix IV,A)

$$\frac{\partial I}{\partial \varphi_\alpha} - \frac{1}{\lambda} \varphi_\alpha(\vec{y}) \sum_{i=1}^{M} P(\vec{S}_i) \; p^{(1)}(\vec{y}|\vec{S}_i) = 0 \tag{6}$$

where $\alpha = 1, 2 \cdots K$; K being the number of samples in a signal.

$\varphi_\alpha(\vec{y})$ is the α^{th} component of the vector (y)

λ^{-1} is the Lagrangian multiplier.

$$\frac{\partial I}{\partial \varphi_\alpha} = \iint \cdots \int_{D_2} d\vec{Y} \sum_{i=1}^{M} P(\vec{S}_i) \log r(\vec{S}_i|\vec{Y}) \; \frac{\partial p^{(2)}(\vec{Y}|\vec{\varphi}(\vec{y}))}{\partial \varphi_\alpha} \; p^{(1)}(\vec{y}|\vec{S}_i) \tag{7}$$

where $$r(\vec{S}_i|\vec{Y}) = \frac{P(\vec{S}_i)\, t(\vec{Y}|\vec{S}_i)}{\sum_{l=1}^{M} P(\vec{S}_l)\, t(\vec{Y}|\vec{S}_l)} \tag{8}$$

If we write

$$q_1(\vec{y}) = \sum_{i=1}^{M} P(\vec{S}_i) \; p^{(1)}(\vec{y}|\vec{S}_i)$$

using (7) we may rewrite (6) into

$$\varphi_\alpha(\vec{y}) = \frac{\lambda}{q_1(\vec{y})} \sum_{i=1}^{M} P(\vec{S}_i) \; p^{(1)}(\vec{y}|\vec{S}_i) \iint \cdots \int d\vec{Y} \frac{\partial p^{(2)}(\vec{Y}|\vec{\varphi})}{\partial \varphi_\alpha} \log r(\vec{S}_i|\vec{Y}) \tag{9}$$

$(\alpha = 1, 2, \cdots K)$

This set of equations defines the optimum transfer characteristic. Thus in order to obtain an optimum design we should solve the system of K integral equations given by (9). An exact solution is very nearly hopeless because of the rather involved character of the equations, indeed the integrand of (9) is itself a functional of the unknown functions as it is easily seen by referring to Eq. (8) and Eq. (2). Thus we may hope to be able to solve the Eq. (9) only in a few very special cases.

4.3 Particular Case: Sample by Sample Transmission Through Additive Noise

Let us consider the following case: (1) no delay is allowed at the intermediate station, thus the signal must be retransmitted sample by sample; (2) the noise is, in both channels, additive to the signal, and (3) the noise probability density, say $n_1(t)$, is the same in both channels and is an even function of t. Let us formulate the problem as follows: using two transmitters, T_1 and T_2, of fixed average power, find the optimum input probability density p(x) and the optimum transfer characteristic $\varphi(y)$. In other words, we have to determine the functions p(x) and $\varphi(y)$ which maximize the amount of information (about x) supplied by Y at R_2. This problem may be properly considered as the determination of the channel capacity because the solution of the problem will specify the transmitted signals only by their amplitude probability density.

The average amount of information (about x) supplied by Y , say I(x,Y), is given by

$$I(x,Y) = -\int q_2(Y) \log q_2(Y)\, dY + \int p(x)\, dx \int t(Y/x) \log t(Y/x)\, dY \quad (10)$$

where $q_2(Y)$ is the probability density of the sample Y (at R_2),

t(Y x) is the transition probability from x to Y.

The limits of integration have been omitted because it is understood that the integration interval must include all points where the integrand is different from zero.

It is easy to see, by direct application of the theorem on total probability that

$$t(Y/x) = \int dy\; n_1(y-x)\; n_1[Y - \varphi(y)]$$

We also have

$$q_2(Y) = \int p(x)\, t(Y/x)\, dx$$

Thus $t(Y/x)$ is a functional of $\varphi(y)$ and $q_2(Y)$ is itself a functional depending on both $\varphi(y)$ and $p(x)$. Referring to Eq. (10) we see that I(x,Y) is a functional of t(Y x) and $q_2(Y)$.

The unknown functions p(x) and $\varphi(y)$ must maximize I(x,Y) while fulfilling the following constraints:

$$\int p(x)\, dx = 1 \quad (11)$$

$$\int x^2 p(x)\, dx = P_1 \quad (12)$$

$$\int q_1(y)\, [\varphi(y)]^2\, dy = P_2 \quad (13)$$

where P_1 and P_2 are, respectively, the average powers of transmitters T_1 and T_2.

In order to obtain the necessary conditions for maximum we introduce small continuous variations $\delta\varphi(y)$ and $\delta p(x)$. If we let

$$n_2(t) = \frac{dn_1(t)}{dt}$$

we obtain for the first variation of $t(Y|x)$

$$\delta t(Y|x) = -\int n_1(y-x)\, n_2[Y-\varphi(y)]\, \delta\varphi(y)\, dy \tag{14}$$

Similarly the first variation of $q_2(Y)$ is

$$\delta q_2(Y) = \int dx\, t(Y|x)\, \delta p(x) - \int dx \int dy\, p(x)\, n_1(y-x)\, n_2[Y-\varphi(y)]\, \delta\varphi(y) \tag{15}$$

Using Eq. (14) and Eq. (15), the first variation of $I(xY)$ is easily obtained:

$$\begin{aligned}\delta I = &-\int dY \log q_2(Y)\left[\int dy \int dx\, n_1(y-x)\, n_1[Y-\varphi(y)]\,\delta p(x) - \int dy\, q_1(y)\, n_2[Y-\varphi]\,\delta\varphi(y)\right] \\ &+\int dY \int dx\, t(Y|x) \log t(Y|x)\, \delta p(x) \\ &-\int dx \int dY\, p(x) \log t(Y|x) \int dy\, n_2[Y-\varphi(y)]\, n_1(y-x)\, \delta\varphi(y)\end{aligned} \tag{16}$$

The necessary conditions are directly obtained from (16) by application of the fundamental lemma of the calculus of variation.[30] But in the application of this lemma, we must remember that the unknown probability density p(x) must, in addition to satisfying the constraints (11) and (12), be non-negative. It is expedient then to replace p(x) by the square of a (real) function p (x). Hence $p(x) = p'^2(x)$

$$\text{and } \delta p(x) = 2\, p'(x)\, \delta p'(x)$$

It is then found that the necessary conditions for maximum take the form of a set of three equations:

$$\int dx \int dY\, p(x)\, n_1(y-x)\, n_2(Y-\varphi) \log \frac{q_2(Y)}{t(Y|x)} = \lambda\, q_1(y)\, \varphi(y) \tag{17}$$

$$\int t(Y|x) \log \frac{q_2(Y)}{t(Y|x)}\, dY = \mu x^2 + \nu \tag{18}$$

$$p(x) = 0 \tag{19}$$

Eq. (17) must be satisfied for all values of y; for any x, either Eq. (18) or Eq. (19) must be satisfied. The constants ν, μ and λ are the Lagrangian multipliers corresponding to the constraints (11), (12) and (13).

4.4 Gaussian Additive Noise

We have already pointed out the importance of gaussian additive noise. So let us assume that, in both channels, the noise probability density is

$$n_1(t) = \frac{e^{-\frac{t^2}{2N}}}{\sqrt{2\pi N}} \tag{20}$$

where N is the average noise power.

In order to solve the Eqs. (17), (18) and (19) in this case we have only one method available: by physical reasoning guess a possible solution and check whether it satisfies the equations.

Let us recall that the entropy H(y) together with the information (about x) received by R_1, will be a maximum if and only if p(x) is gaussian.[6,7] As the noise in the second channel is also gaussian, it seems natural that the input of the second channel should also be gaussian. For, in that case, R_2 receives as much information about y as possible under the constraint that the average power of T_2 is constant. Thus a linear transfer characteristic is required for only if $\varphi(y)$ is linear in y, can both y and $\varphi(y)$ have a gaussian distribution.

At first sight, one might wonder how it is possible that a linear transfer characteristic may be optimum, for a linear transfer characteristic implies that some signals, although very rare, are retransmitted with a very large amount of energy. This conjecture is not valid because the performance factor is equal to:

$$\log_e \sqrt{\frac{1+2N}{2N}} + \frac{y^2-(1+N)}{2(1+2N)}$$

where we assumed $P_1 = 1$ and $P_2 = 1 + N$ to simplify the notation. This shows that as y becomes very large, the average amount of information that R_2 receives about x becomes approximately proportional to y^2. As, on the other hand, the energy is also proportional to y^2, the linear characteristic seems quite natural, since for large y's the (energy) expense becomes proportional to the (information) return.

To test this plausibility reasoning, we must substitute, into the Eqs. (17) and (18), the assumed solution:

$$p(x) = \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}$$

$$\varphi(y) = y$$

These relations imply

$$q_1(y) = \frac{e^{-\frac{y^2}{2(1+N)}}}{\sqrt{2\pi(1+N)}}$$

and

$$q_2(Y) = \frac{e^{-\frac{Y^2}{2(1+2N)}}}{\sqrt{2\pi(1+2N)}}$$

To simplify the notation let us define R and C_2 such that

$$\log_e q_2(Y) = -\frac{Y^2}{2R} + c_2$$

* The manipulations would remain essentially the same if we had taken

If in Eq. (17) we let

$$\log \frac{q_2(Y)}{t(Y|x)} = \log q_2(Y) - \log t(Y|x) \tag{21}$$

it is easy to show that the contribution of the 1st term of (21) is

$$\frac{1}{R}\, q_1(y)\, \varphi(y) \tag{22}$$

Let $$\log t(Y|x) = -\frac{(Y-x)^2}{4N} + C_3$$

where C_3 is a constant independent of x or Y.

Integration by parts (with respect to Y) of the 2nd term produces an integrand of the form

$$\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\ \frac{e^{-\frac{(y-x)^2}{2N}}}{\sqrt{2\pi N}}\ \frac{e^{-\frac{(Y-y)^2}{2N}}}{\sqrt{2\pi N}}\ \frac{-(Y-x)}{2N}$$

which after integration with respect to Y, gives

$$\frac{1}{2}\frac{\partial}{\partial y}\left\{ \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\ \frac{e^{-\frac{(y-x)^2}{2N}}}{\sqrt{2\pi N}} \right\}$$

Inverting the order of differentiation and of integration we finally get

$$-\frac{y}{1+N}\, q_1(y) \tag{23}$$

Thus the left-hand member of Eq. (17), which is the sum of expression (22) and (23) is proportional to the product $\varphi(y)\, q_1(y)$, for all y, as it is required by (17).

The check of Eq. (18) is immediate.

Thus, it has been shown that the necessary conditions for the maximum amount of received information are satisfied by the gaussian distributed input and the linear transfer characteristic $\varphi(y) = ky$.

4.5 <u>The Discrete Case</u>

Consider a two-channel system such as the one represented in Fig. IV,1. Suppose that the transmitter T_1 sends pulses of unit amplitude and of either polarity, each type of pulse having the same probability. Suppose that the intermediate station is required to retransmit the samples as soon as they are received. The problem is to find, under these conditions, the optimum transfer characteristic $\varphi(y)$ of the intermediate station.

The equation for the optimum $\varphi(y)$ may be obtained from Eq. (17) provided we take into account that

$$p(x) = \frac{1}{2}\left[\delta(x-1) + \delta(x+1)\right] \tag{24}$$

where $\delta(x)$ is the usual Dirac or impulse function.

If this substitution is carried out, the following equation is obtained for $\varphi(y)$

$$\frac{1}{2} n_1(y-1) \int n_2[Y-\varphi(y)] \log \frac{q_2(Y)}{t(Y|1)} \, dY$$
$$+ \frac{1}{2} n_1(y+1) \int n_2[Y-\varphi(y)] \log \frac{q_2(Y)}{t(Y|-1)} \, dY = \lambda q_1(y) \varphi(y). \tag{25}$$

The direct solution of this equation is well nigh impossible. Nor was it found possible to devise an approximate method which would lead to a solution within a reasonable amount of time.

On the basis of the results of Chapter II, it is clear that the performance of the system, assuming $\varphi(y) = ky$ as a transfer characteristic (where the constant k is adjusted to fit the power constraint) is certainly worse than that obtained with a maximum a-posteriori probability transfer characteristic. It is shown in Appendix IV.B, that the latter transfer characteristic is not optimum either. This proof requires only very general assumptions on the probability density $n_1(t)$.

Nevertheless it is felt that the problem under consideration is of sufficient interest to create the need for an even approximate determination of the optimum $\varphi(y)$. In order to obtain a simpler equation for $\varphi(y)$, let us assume that the final receiver R_2 operates as a maximum a-posteriori probability detector, that is, its output consists of the sample most likely to have caused the received sample.

In addition to the assumption that $n_1(t)$ is even, let us assume that $n_1(t)$ is a decreasing function of t, for positive t. As the symmetry of the problem requires that $\varphi(y)$ be odd, it follows that when the received sample Y, at R_2, is positive (resp. negative) the output of R_2 will be $+1$ (resp. -1). The probability that the output of R_2 is in error is then a functional of $\varphi(y)$ given by

$$p = \frac{1}{2}\int_{-\infty}^{0} dY \int_{-\infty}^{+\infty} dy\, n_1(y-1)\, n_1[Y-\varphi] + \frac{1}{2}\int_{0}^{\infty} dY \int_{-\infty}^{+\infty} dy\, n_1(y+1)\, n_1[Y-\varphi] \qquad (26)$$

Taking into account the average power constraint on $\varphi(y)$ we obtain the following equation for $\varphi(y)$

$$n_1[\varphi(y)]\,\frac{n_1(y+1) - n_1(y-1)}{n_1(y+1) + n_1(y-1)} = \lambda\,\varphi(y) \tag{27}$$

If the Lagrangian multiplier λ were known, the transfer characteristic $\varphi(y)$ would be implicitely defined by (27). Equation (27) can be solved numerically by assuming a particular value of λ and adjusting the λ by successive approximations until the solution $\varphi(y)$ satisfies the power requirement.

The optimization problem is an important problem because, if it were solved, it would indicate the most that can be achieved, by the system under consideration. As we have seen in section 4.2, the problem, when treated formally, leads to an unsoluble system of equations. Apparently, the difficulty comes from the fact that, in this treatment, at each step of the derivation, all the characteristics of the system under consideration are taken into account. On the other hand, it seems reasonable to assume that if, by introducing certain approximations, one could separate, even partially, the different factors of the problem, one would obtain an approximation leading to more readily solved equations.

This kind of thinking led to a heuristic approach of the problem. In the particular case under consideration it leads to the exact form of Eq. (27). As it is felt that this is more than a mere coincidence, this heuristic derivation is given here.

It is intuitively clear that the optimum transfer characteristic should depend on the following three factors:

(1) A sample of amplitude y received at R_1, has a "value" which is a function of y.

(2) The usefulness (to the last receiver R_2) of a retransmitted sample of amplitude $\varphi(y)$ is a function of $\varphi(y)$.

(3) The intermediate station transmitter T_2 has a fixed average power.

Since we wish to derive heuristically the condition resulting from the minimization of the probability of error, we should use only probability concepts. Suppose y is received and $\varphi(y)$ is retransmitted, let us find a function of y and $\varphi(y)$, say $F[y, \varphi(y)]$, which will represent the average value, to the last receiver R_2, of the sample retransmitted as $\varphi(y)$.

If the sample y received at R_1 is positive and if, as a consequence, it is assumed that + 1 was transmitted by T_1, the probability of error p(y) is given by

$$p(y) = \frac{n_1(y+1)}{n_1(y-1) + n_1(y+1)} \qquad \text{for } y > 0$$

Since the channel preceding the intermediate station has a binary input let us consider the quantity 1 - 2p(y) whose form is identical to the quantity of interest in the analysis of cascaded binary channels, cf. Eq. II,9. If $p(y) = \frac{1}{2}$, the received sample y is of no information value and $1 - 2p(y) = 0$ If $p(y) = 0$, the received sample has the maximum information value and $1 - 2p(y) = 1$. Thus we might expect that $1 - 2p(y)$ occurs as a factor in $F[y, \varphi(y)]$. It seems reasonable to further assume that the second factor must describe the effectiveness of the retransmitted sample $\varphi(y)$ from the point of view of the last receiver. A natural choice would be the probability $p_c[\varphi(y)]$ that the retransmitted sample $\varphi(y)$ will be correctly interpreted by the last receiver. Thus in our case

$$p_c[\varphi(y)] = \frac{1}{2} + n[\varphi(y)] \tag{28}$$

where

$$n(z) = \int_0^z n_1(t)\,dt$$

Thus we write

$$F[y, \varphi(y)] \sim [1 - 2p(y)]\, p_c[\varphi(y)] \tag{29}$$

Since we are interested in optimizing the average behaviour of the communication system, we must obviously consider the average value of $F[y, \varphi(y)]$, the averaging being carried out over all y's.

Thus the problem is then to find the $\varphi(y)$ which maximizes

$$\langle F(y, \varphi(y)) \rangle_{ave} \tag{30}$$

subject to the condition that

$$\langle [\varphi(y)]^2 \rangle_{ave} = P \tag{31}$$

Geometrically, in terms of a Hilbert space in which $\varphi(y)$ is a point, the condition (31) represents a surface to which the point $\varphi(y)$ is constrained. The problem is then to

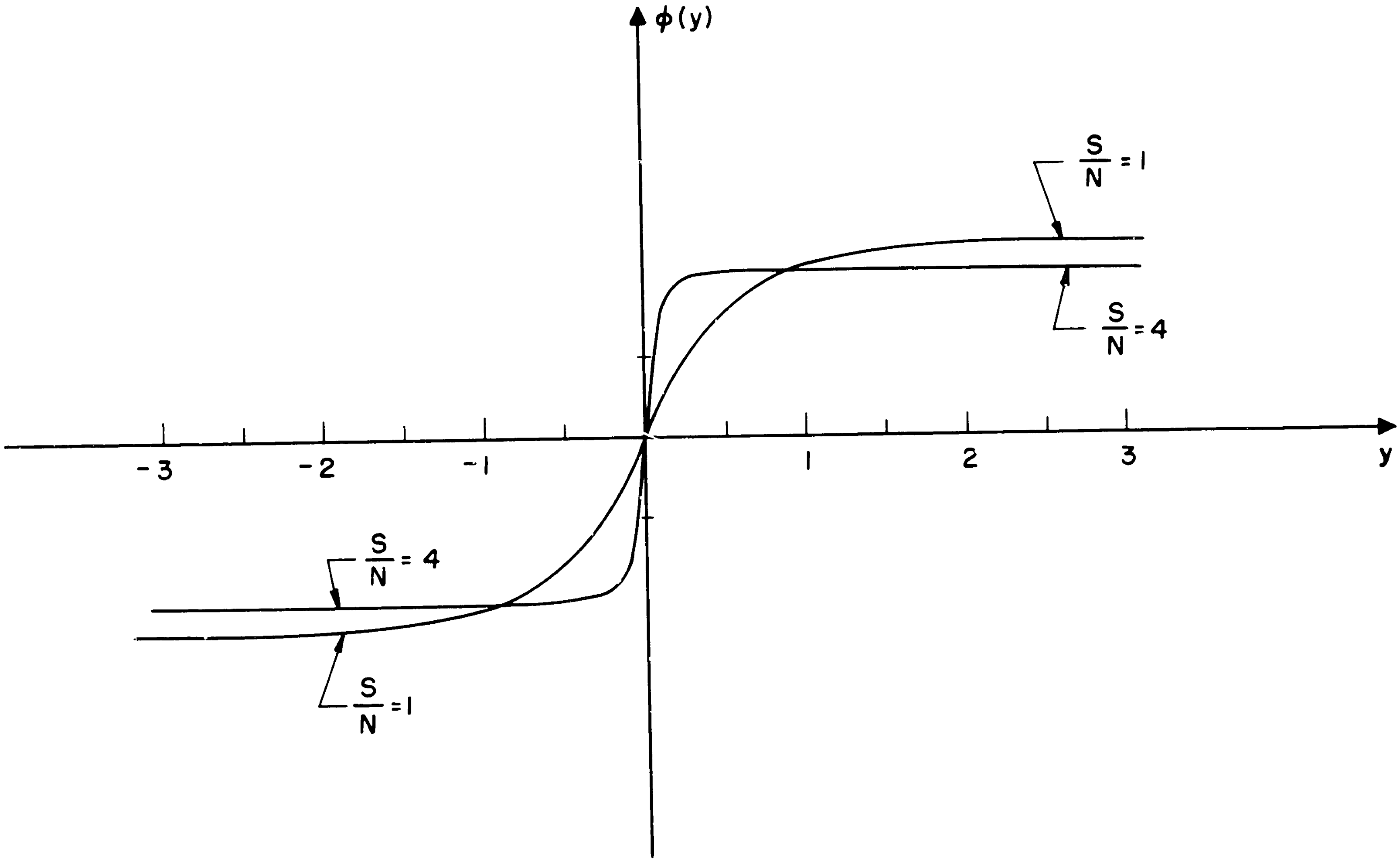


FIG. IV, 2

find a point on that surface for which the scalar (30) is maximum. At that point, the surface (31) and the surface

$$< F[y, \varphi(y)] >_{ave} = C^t$$

will have a common normal. Hence at that point, we shall have

$$\varphi(y) \sim \frac{\partial}{\partial \varphi} F[y, \varphi(y)]$$

If we take into account Eqs. (28) and (29), we obtain

$$\varphi(y) \sim \frac{n_1(y-1) - n_1(y+1)}{n_1(y-1) + n_1(y+1)} \; n_1[\varphi(y)]$$

which is identical to Eq. (27).

4.6 Special Case of Gaussian Noise.

Let the noise be gaussian and additive. Let N be the average noise power, then the noise probability density is given by Eq. (20). Taking this into account Eq. (27) becomes

$$\frac{e^{-\frac{[\varphi(y)]^2}{2N}}}{\sqrt{2\pi N}} \tanh \frac{y}{N} = \lambda \, \varphi(y) \tag{32}$$

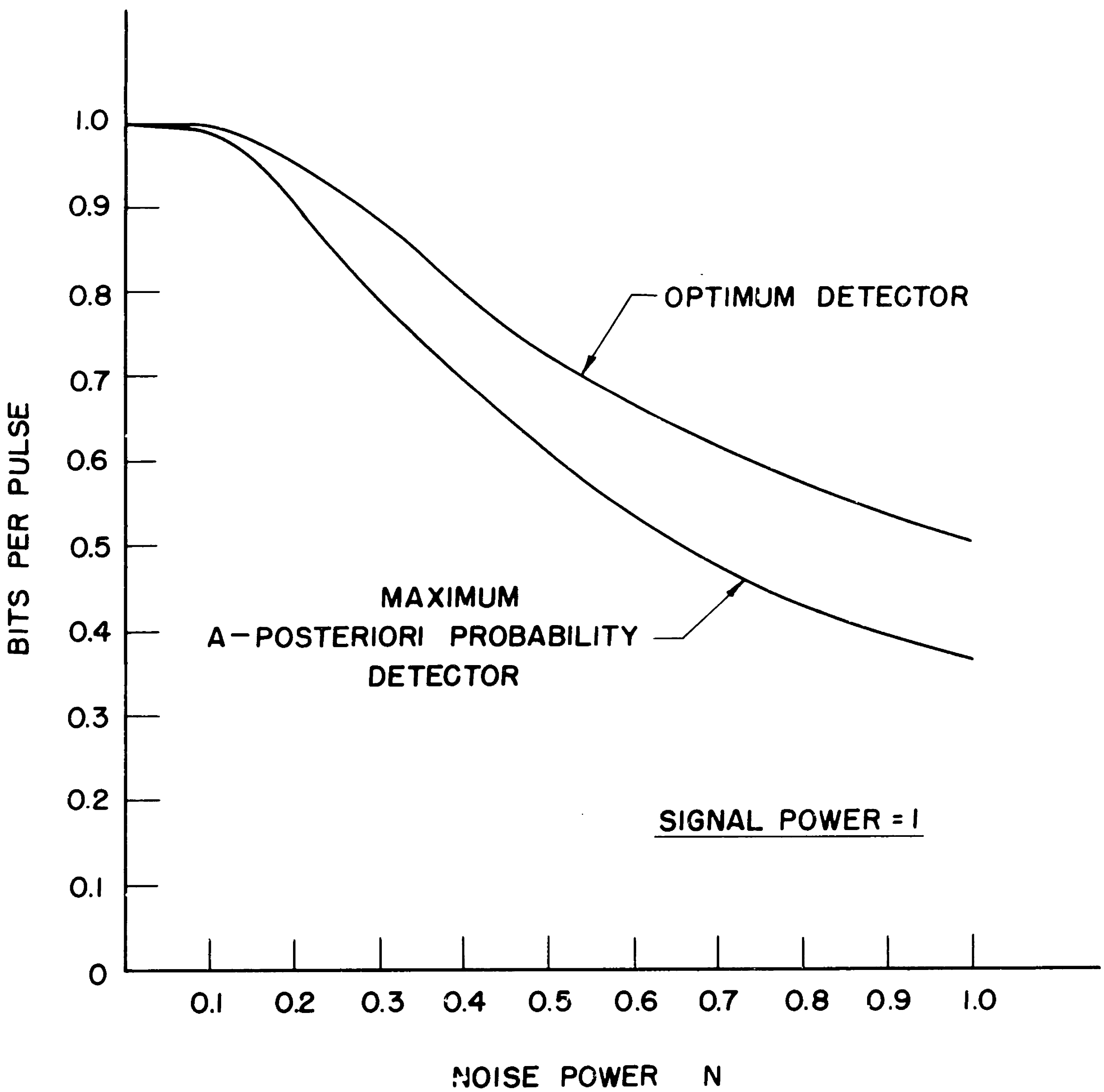
This equation has been solved numerically for $\frac{S}{N} = 1$ and $\frac{S}{N} = 4$. The solutions are presented in Fig. IV,2. We used them to compute the probability of error P_e and the per-unit equivocation E. For purposes of comparison, the probability of error P_e' and the equivocation E' have been computed on the basis of the maximum a-posteriori probability transfer characteristic (for short M.A.P.T.C.).

	P_e	P_e'	E	E'
$\frac{S}{N}=1$	.257	.267	.825	.850
$\frac{S}{N}=4$	.0432	.0445	.257	.262

It should be stressed that as the signal to noise ratio becomes large, the solution of (32) resembles more and more the M.A.P.T.C. and the transition region of the solution of Eq. (32) gets smaller and smaller.

The results indicated by the table above are of interest because they give the largest decrease in the probability of error that can be achieved under the condition of sample by sample retransmission. They imply, therefore, that any other strategy, such as, for example, requantizing the received sample y to a larger number of levels, will not lead to an appreciable improvement in the system, once the signal to noise ratio is larger than, say, 4. In fact some of these possibilities have been investigated by the writer and the results were found to be within the bounds indicated by the table above.

Essentially, the equation for $\varphi(y)$ was obtained in a soluble form at the cost of minimizing the probability of error instead of maximizing the information contained in the received sample. It would be therefore of interest to evaluate the difference between the information content of the input-



1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
BITS PER PULSE
OPTIMUM DETECTOR
MAXIMUM
A-POSTERIORI PROBABILITY
DETECTOR
SIGNAL POWER = 1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
NOISE POWER N

FIG. IV, 3

signal and the output-signal of the receiver defined above. This could be done only in the following simple case: The system consists of a single channel perturbed by gaussian additive noise, its input consists of samples of amplitude ± 1, the receiver operates as a maximum a-posteriori probability detector. Thus the information per pulse (about what was transmitted) contained in the detector's output is, in bits,

$$I_M = 1 - \mathcal{f}(p)$$

where

$$p = \psi\left(\frac{1}{\sqrt{N}}\right)$$

The amount of information per sample contained in the received signal and that, by definition, would be contained in the output of an "optimum" detector is given by

$$I_0 = -\int_{-\infty}^{+\infty} \tfrac{1}{2}\left[p(y|1) + p(y|-1)\right] \log \tfrac{1}{2}\left[p(y|1) + p(y|-1)\right] dy$$
$$+ \tfrac{1}{2}\int_{-\infty}^{+\infty} p(y|1) \log p(y|1)\, dy + \tfrac{1}{2}\int_{-\infty}^{\infty} p(y|-1) \log p(y|-1)\, dy \tag{33}$$

where

$$p(y|\pm 1) = \frac{e^{-\frac{(y \mp 1)^2}{2N}}}{\sqrt{2\pi N}} \tag{34}$$

The results are presented on Fig. IV,3 and the details of the derivation are presented in Appendix IV,C. These results are in accordance with the intuitive feeling in that, for large signal to noise ratios, the relative difference in the information content is small and that it becomes quite appreciable when the signal to noise ratio approaches unity.

4.7 Concluding Remarks

Ordinarily the intuitive feeling which guides the expert is built up by the experience of many simple cases. In the domain which is the object of this work only a few cases have been treated. Therefore any conclusion must be considered tentative and is made with the aim of communicating a way of thinking rather than summarizing, in a few bold sentences, the basic nature of the problem.

The characteristic difference between the problem of communication through channels in cascade and that of communication through a single channel is that, in the latter case, the transmitter possesses the complete knowledge of what it should transmit. Whereas in the cascade, each intermediate station has only partial information about what it would like to transmit. In fact, the information available to the intermediate station is in the form of a set of a-posteriori probabilities.

The amount of (selective) information required to specify this set of probabilities is infinite. Even if the probabilities were specified only approximately, it is usually very much greater than the amount of information (about what has been transmitted by the first transmitter) supplied by the received signal. As a result, the intermediate station must retransmit one or a few of the characteristics of the set of a-posteriori probabilities. A convenient characteristic to retransmit is the member of the set having the largest probability. This corresponds to the maximum a-posteriori probability transfer characteristic. In this particular case, it appears that the important factor is the per-unit equivocation of the channel (or of the cascade of channels) which precedes the intermediate station under consideration. When the per-unit equivocation is small, the sum of the probabilities of all the other members of the set is small, so that the specification of the member having the largest probability conveys nearly all the information contained in the received signal. When the per-unit equivocation is appreciable, the specification of that member indicates only one of the many characteristics of the set of a-posteriori probabilities. This way of thinking makes it clear that, in the cases where the per-unit equivocation (per channel) is appreciable, the performance of the cascade should deteriorate rapidly as the number of cascaded channels increases. It also makes obvious the reason why such techniques as the requantization of pulses at each intermediate station or the complete detection of the signals at each intermediate station play such an important role in the performance of the cascade.

Appendix III.A

The characteristic values of the M by M matrix.

$$\begin{bmatrix} b & p & \cdots & p \\ p & b & \cdots & p \\ \vdots & & & \vdots \\ p & \cdots & & b \end{bmatrix}$$

where $b = 1 - (M - 1)p$, are respectively 1 and $1 - M_p$.

The characteristic values are solution of the determinental equation

$$\begin{vmatrix} c & p & \cdots & p \\ p & c & \cdots & p \\ \vdots & & & \vdots \\ p & \cdots & & c \end{vmatrix} = 0$$

where $c = b - \lambda$

Subtracting the last column from the 1st, 2nd, ... $(M - 1)^{th}$ column we get

$$\begin{vmatrix} (c - p) & 0 & 0 & \\ 0 & (c - p) & 0 & \\ \vdots & & \ddots & \\ & & & (c - p)p \\ (p - c) & (p - c) & & (p - b)c \end{vmatrix} = 0$$

Adding the 1st, 2nd, ... $(M - 1)^{th}$ row to the last row we get

$$\begin{vmatrix} (c - p) & 0 & \cdots & p \\ 0 & (c - p) & & p \\ \vdots & & \ddots & \\ \vdots & & \ddots & p \\ 0 & 0 & 0 & 1 - \lambda \end{vmatrix} = (1-\lambda)(b-\lambda)^{M-1} = 0$$

or

$$\left[(1-Mp)-\lambda\right]^{M-1}(1-\lambda) = 0$$

q.e.d.

Appendix III.B

Consider a signal of ℓ binary pulses, the whole signal being repeated $2\alpha + 1$ times. The probability that a particular pulse of the signal will be misinterpreted is:

$$p'_e = \sum_{\beta=\alpha+1}^{2\alpha+1} \binom{2\alpha+1}{\beta} p^{\beta} q^{(2\alpha+1-\beta)}$$

The probability that an error will occur somewhere in the signal is

$$P'_e = 1-(1-p'_e)^{\ell} \simeq \ell p'_e - \binom{\ell}{2} p_e'^{2} + \cdots$$

The probability p'_e is a decreasing function of α, and $p'_e \to 0$ as $\alpha \to \infty$ thus, for sufficiently large α, $\ell p'_e < 1$ and the first term of the binomial expansion is an upper bound to P'_e,

$$P'_e < \ell p'_e$$

But

$$p'_e < \alpha \binom{2\alpha+1}{\alpha} p^{\alpha} q^{\alpha+1}$$

By Stirling's formula, $\binom{2\alpha+1}{\alpha} \simeq 2^{2\alpha} \frac{2\alpha+1}{\sqrt{\pi\alpha}} \simeq \frac{2}{\sqrt{\pi}} 4^{\alpha} \sqrt{\alpha}$

hence

$$p'_e < \alpha^{3/2} 4^{\alpha} p^{\alpha} q^{\alpha+1} \frac{2}{\sqrt{\pi}} < \frac{2}{\sqrt{\pi}} (4pq)^{\alpha} \alpha^{3/2}$$

and

$$P'_e < \frac{2}{\sqrt{\pi}} \ell (4pq)^{\alpha} \alpha^{3/2}$$

Appendix III.C

The double and the triple error correcting checks, described in section 3.23, should be modified in a trivial way in order to meet the following objection. For simplicity, this objection will be formulated in detail for the double error correcting case.

Consider a particular combination of two errors, one affecting the pulse sequence S and the other affecting D_1 such that the resulting sequences S^r and D_1^r agree with each other. Let us remember that the sequence D_1 is obtained from S by carrying out the operations specified by Dq. (20). It is clear that such a situation can occur only if

(a) the error affecting S occurs in a position to which is associated a number, the binary representation of which contains only a single one.

(b) this digit, just mentioned, is the one affected by the 2nd error, that is, the error affecting D_1.

Essentially the 2nd error erases the trace of the 1st one. These occurrences will obviously be avoided if to single errors are associated numbers the binary representations of which contain at least two ones.

We shall now show that if $k > 2$, we can always associate to single errors, numbers the binary representations of which contain a single one.

The number of these numbers is ℓ_1, if ℓ_1 is the number of pulses contained in D_1. On the other hand ℓ_1 is is defined as the least integer such that

$$2^{\ell_1} > k + \binom{k}{2} \tag{C1}$$

Thus, in order to fulfill our supplementary conditi -

we need to have in addition

$$2^{\ell_1} - \ell_1 > k \tag{C2}$$

since there are k possible single errors in S. It is obvious that, for large k, C_1 implies C_2. It can be verified that it is indeed so except for the case of k = 2.

A similar reasoning will show that for the triple error correcting case we must impose the following requirements:

(a) single errors should be associated to numbers the binary representations of which contain at least three ones.

(b) double errors should be associated to numbers the binary representations of which contain at least two ones.

We shall show that once k > 3, we can always fulfill these additional requirements.

Indeed the triple error correcting code T associated to the sequence S has a number of pulses ℓ defined as the least number ℓ such that

$$2^{\ell} > k + \binom{k}{2} + \binom{k}{3} \tag{C3}$$

Condition <u>a</u> requires

$$2^{\ell} - \ell - \binom{\ell}{2} > k \tag{C4}$$

Condition <u>b</u> requires

$$2^{\ell} - \ell > k + \binom{k}{2} \tag{C5}$$

Again it is obvious that for large k, (C4) and (C5) are implied by (C3). It can be verified numerically that it is also the case for small k provided k > 3.

Appendix III.D

The aim of this appendix is to show how the codes presented in the text may be justified. We shall reason only on the triple error correcting case.

The proof is carried out by considering all possible cases. To consider them all here would be very long, especially in view of the fact that the reasoning used falls into a few definite patterns. We shall therefore examine here a few typical cases.

(a) Suppose that three errors occurred in the sequence T_1; hence the received sequence T_1^r differs from T_1 by three digits. The received signal is then S T_1^r T_2 D_1 D_2 P_1 P_2. As stated in the text, the receiver uses this signal to verify whether all the relations between the proper received sequences agree or not. In the present case, there are discordances between S and T_1^r, on the one hand, and T_1^r and T_2 on the other. The pairs $T_2 - D_1$, $D_1 - D_2$, $D_1 - P_1$ and $D_2 - P_2$ are found to agree. We must remember that the code is designed to correct all errors provided their total number is $\leqslant 3$. Thus we constantly assume in the reasoning here that the number of errors which did occur is $\leqslant 3$. From the discordances, it is concluded that there is at least one error in the first three sequences S, T_1^r, T_2.

Thus there can be at most two errors in the last five sequences T_2, D_1, D_2, P_1 and P_2. A moment of reflection will show that no two errors could have affected these sequences and at the same time produce the agreements between the above mentioned pairs. Hence T_2 is free from any errors and is used to correct T_1^r. The corrected sequence obtained from T_1^r is found, in this case, to agree with S, from which it is deduced that S was correctly received.

(b) Suppose a pair of errors occurred in S and a single error affected T_2. The received signal is then of the form $S^r\ T_1 T_2^r D_1 D_2 P_1 P_2$. The receiver notes the following agreements $D_1 - D_2$, $D_1 - P_1$, $D_2 - P_2$ and the following discordances $S^r - T_1$, $T_1 - T_2^r$, $T_2^r - D_1$. In order to obtain these three disagreements at least two sequences must contain some errors. Thus, at most, a single error could have affected the last three sequences, D_2, P_1 and P_2. It is obvious, then, that D_2 is free from errors and so is D_1 (on the basis of the agreement $D_1 - D_2$). D_1 may be used to correct T_2, for, indeed, it is known that all errors did not occur in the same sequence, thus T_2 is affected by at most two errors. In the present case, the corrected T_2 will agree with T_1, which in its turn, will be used to correct S.

(c) Suppose one error affected T_1, another D_2 and the last P_2. The received sequence is then of the form S, T_1^r, T_2, D_1, D_2^r, P_1, P_2^r. The receiver notes the following agreements $T_2 - D_1$, $D_1 - P_1$, $D_2^r - P_2^r$ and the following discordances $S - T_1^r$, $T_1^r - T_2$, $D_1 - D_2^r$. This last discordance indicates that at least one error must affect one of the D's. The other two discordances indicate that at least one error affects the first three sequences. From the first conclusion and the fact that P_1 (resp. P_2^r) agrees with D_1 (resp. D_2^r) it follows that one of the P's is in error. Thus there are at least three errors and since we need only consider the cases where not more than three errors occurred, we conclude that a single error affects the group $ST_1^rT_2$. Remembering that this single error causes the discordances $S - T_1^r$ and $T_1^r - T_2$ it follows that the error affects T_1, hence S is free from any error.

Obviously the cases in which several errors affect a single sequence are very easily dealt with because the errors are easily located. The cases where each one of several sequences are affected by a single error require subtler reasoning but essentially the technique is the same as in the case C. In order to convince the reader we shall consider a second situation of this type.

(d) Suppose one error affected S, another T_2 and the last one D_2. The received sequence is then of the form S^r, T_1, T_2^r, D_1, D_2^r, P_1, P_2. The receiver notes the following agreement $D_1 - P_1$ and the following discordances $S^r - T_1$, $T_1 - T_2^r$, $T_2^r - D_1$, $D_1 - D_2^r$, $D_2^r - P_2$. From the first three discordances at least two sequences of the set S, T_1, T_2, D_1 must be in error. In addition, from the last discordance, some error must affect either D_2 or P_2, hence at most two errors (in two different sequences) must have affected the set S, T_1, T_2, D_1. Thus (if the total number of errors is $\leqslant 3$, the only case we are interested in) P_1 is correct and from the agreement P_1 - D_1 we conclude that D_1 is also correct. Since we know that all three errors did not affect the same sequence, the double error correcting code, D_1 will suffice to obtain the correct S from the received sequence.

Using the same method to discuss all other possible cases, it may be shown that the proposed code allows the correct S to be extracted from the received sequences provided they were not affected by more than three errors.

Appendix IV.A

Let $\vec{\varphi}_0(\vec{y})$ be the optimum transfer characteristic. Consider a continuous bounded vector function $\vec{\eta}(\vec{y})$ and a real number ϵ such that, for small enough ϵ's, $\epsilon\,\vec{\eta}(\vec{y})$ is for all $\vec{y}$'s very small.

If we replace $\vec{\varphi}_0$ by $\vec{\varphi}_0 + \epsilon\,\vec{\eta}$ in the expression for $t(\vec{Y} \mid \vec{S}_1)$ we obtain the transition probability density corresponding to the new transfer characteristic. This probability density is a function of ϵ. Let us expand the integrand in MacLaurin's series, neglecting terms higher than the 1st order, thus

$$p^{(2)}[\vec{Y} \mid \vec{\varphi}_0 + \epsilon\vec{\eta}] = p^{(2)}[\vec{Y} \mid \vec{\varphi}_0] + \epsilon \sum_{\alpha=1}^{K} \eta_\alpha \frac{\partial p^{(2)}(\vec{Y} \mid \vec{\varphi}_0)}{\partial \varphi_{0\alpha}}$$

where η_α is the α^{th} component of $\vec{\eta}$

and $\varphi_{0\alpha}$ is the α^{th} component of $\vec{\varphi}_0$

The variation of $t(\vec{Y} \mid \vec{S}_1)$ is then, using (IV,2),

$$\delta t(\vec{Y} \mid \vec{S}_i) = \epsilon \sum_{\alpha} \iint \cdots \int d\vec{y}\; \eta_\alpha(\vec{y}) \frac{\partial p^{(2)}(\vec{Y} \mid \vec{\varphi}_0)}{\partial \varphi_{0\alpha}}\; p^{(1)}(\vec{Y} \mid \vec{S}_i) \qquad \text{(A.1)}$$

The variation of the information received is

$$\delta I = \iint \cdots \int d\vec{Y} \sum_{i=1}^{M} P(\vec{S}_i)\left[1 + \log t(\vec{Y} \mid \vec{S}_i)\right] \delta t(\vec{Y} \mid \vec{S}_i)$$

$$- \iint \cdots \int d\vec{Y} \left[1 + \log \sum_{\ell=1}^{M} P(\vec{S}_\ell)\, t(\vec{Y} \mid \vec{S}_\ell)\right] \sum_{i=1}^{M} P(\vec{S}_i)\, \delta t(\vec{Y} \mid \vec{S}_i)$$

or

$$\delta I = \iint\cdots\int \sum_{i=1}^{M} P(\vec{S}_i) \log \frac{t(\vec{Y}\,|\,\vec{S}_i)}{\sum_{\ell} P(\vec{S}_\ell)\, t(\vec{Y}\,|\,\vec{S}_\ell)}\, \delta t(\vec{Y}\,|\,\vec{S}_i)\, d\vec{Y}$$

If we substitute in the last equation $\delta t(\vec{Y}\,|\,\vec{S}_1)$ by its value according to Eq. (B.1), and if we use the fundamental lemma of variation calculus,[30] the equations for the optimum $\vec{\varphi}_0(\vec{y})$ would be

$$\frac{\partial I}{\partial \varphi_{0\alpha}} = 0 \qquad (\alpha = 1, 2, \ldots K)$$

if $\vec{\varphi}_0(\vec{y})$ had not to fulfill any constraint,

Where

$$\frac{\partial I}{\partial \varphi_{0\alpha}} = \iint\cdots\int d\vec{Y} \sum_{i} P(\vec{S}_i) \log \frac{t(\vec{Y}\,|\,\vec{S}_i)}{\sum_{\ell} P(\vec{S}_\ell)\, t(\vec{Y}\,|\,\vec{S}_\ell)}\; \frac{\partial p^{(1)}(\vec{Y}\,|\,\vec{\varphi}_0)}{\partial \varphi_{0\alpha}}\; p^{(1)}(\vec{y}\,|\,\vec{S}_i) \qquad \text{(A.2)}$$

Remembering that

$$\iint\cdots\int d\vec{Y}\; \delta t(\vec{Y}\,|\,\vec{S}_i) = 0 \qquad (i = 1, 2, \ldots M)$$

it is clear then that expression (B.2) is equivalent to (IV.7).

If, as in the text, the optimum vector function $\vec{\varphi}_0(\vec{y})$ must satisfy the power constraint (IV.5), using Lagrange's method one obtains immediately Eq. (IV.6).

Appendix IV.B

The aim of this appendix is to show that the maximum a-posteriori probability characteristic is not optimum. To do so we consider a modified transfer characteristic which, for $\Delta = 0$, reduces to the preceding one. It is shown that for infinitely small Δ, the information received is larger than that obtained in the case $\Delta = 0$.

Both transfer characteristics are represented in Fig. IV.B. The size of the modified transfer characteristic is obtained from the condition that the average power of the intermediate station should remain unchanged. Thus, for small Δ 's, the retransmitted sample will be $1 + n_1(1)\Delta$.

The transfer probability density of the equivalent channel for the case of maximum a-posteriori probability detection is given by:

$$t(Y|1) = \left[\tfrac{1}{2} + n(1)\right] n_1(Y-1) + \left[\tfrac{1}{2} - n(1)\right] n_1(Y+1)$$

where

$$n(y) = \int_0^y n_1(t)\,dt$$

In order to obtain the transfer probability density of the equivalent channel for the case of the modified characteristic we note first that the second channel is used as a three level pulse system. The transition probability matrix of the first channel is

$$\begin{bmatrix} \frac{1}{2} + n(1-\Delta) & n(1+\Delta) - n(1-\Delta) & \frac{1}{2} - n(1+\Delta) \\ \frac{1}{2} - n(1+\Delta) & n(1+\Delta) - n(1-\Delta) & \frac{1}{2} + n(1-\Delta) \end{bmatrix}$$

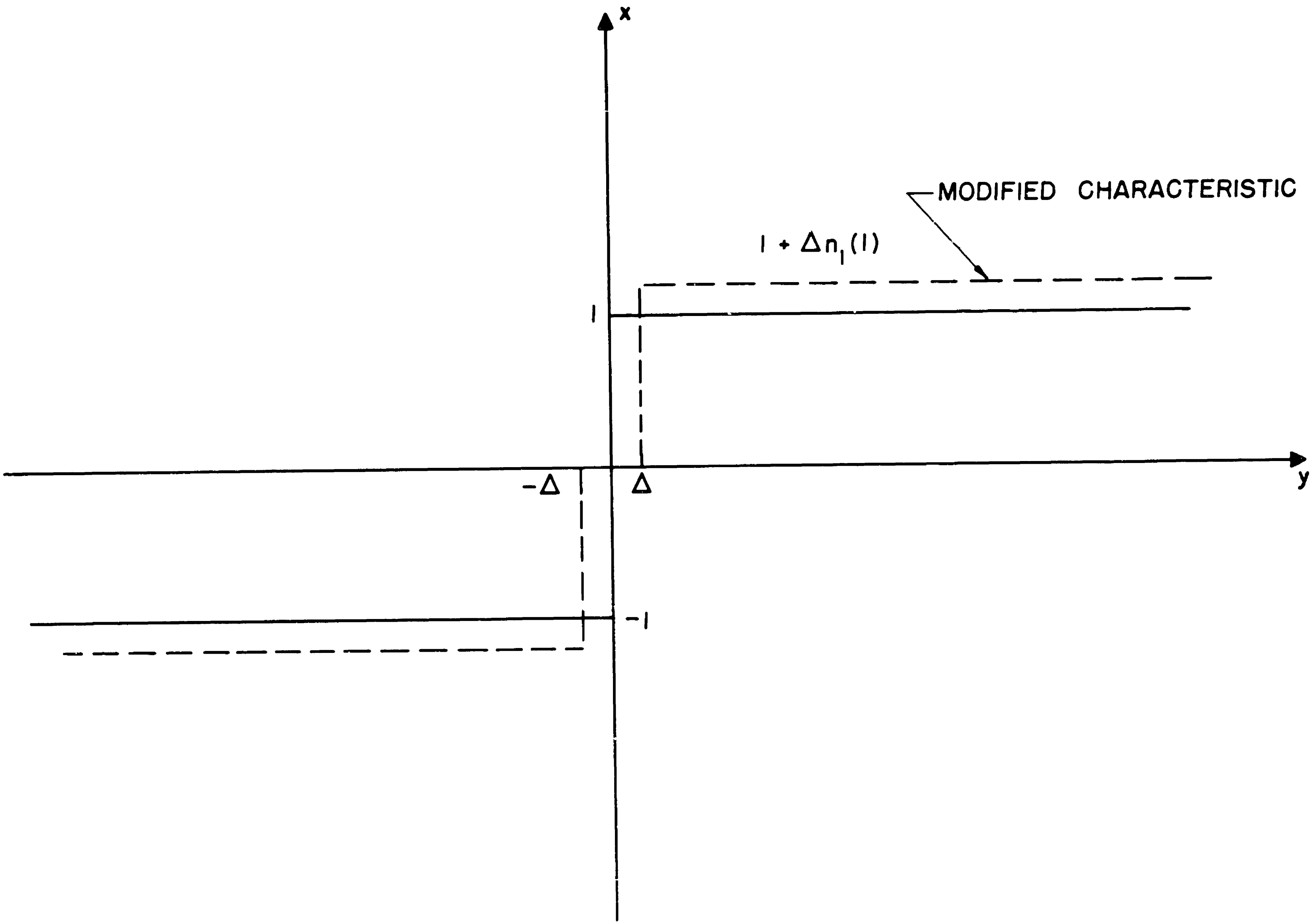


FIG. IV ,4

The transition probability density of the second channel is the column matrix

$$\begin{bmatrix} n_1[Y-1-\Delta n_1(1)] \\ n_1(Y) \\ n_1[Y+1+\Delta n_1(1)] \end{bmatrix}$$

The equivalent-channel-probability density is given by the product of the two matrices, thus we obtain respectively $t_m(Y|1)$ and $t_m(Y|-1)$.

We have

$$t_m(Y|1) = \left[\tfrac{1}{2} + n(1)\right] n_1(Y-1) + \left[\tfrac{1}{2} - n(1)\right] n_1(Y+1)$$

$$+\Delta n_1(1)\left\{ 2n_1(Y) - n_1(Y-1) - n_1(Y+1) - \left[\tfrac{1}{2}+n(1)\right] n_2(Y-1) + \left[\tfrac{1}{2}-n(1)\right] n_2(Y+1)\right\}$$

where we neglected the second order terms in Δ .

Thus when Δ changes from zero to an infinitely small value, $t(Y|1)$ changes by

$$\delta t(Y|1) = t_m(Y|1) - t(Y|1)$$

The change in the density of Y is

$$\delta q_2(Y) = \Delta n_1(1)\left\{ 2n_1(Y) - n_1(Y-1) - n_1(Y+1) + \tfrac{1}{2} n_2(Y+1) - \tfrac{1}{2} n_2(Y-1)\right\}$$

The change δI in the average amount of information received is:

$$\delta I = -\int \delta q_2(Y) \log q_2(Y)\, dY + \tfrac{1}{2}\int \delta t(Y|1) \log t(Y|1)\, dY + \tfrac{1}{2}\int \delta t(Y|-1) \log t(Y|-1)\, dY$$

and by substitution:

$$\delta I = n_1(1)\ \Delta \int F(Y)\,dY$$

where

$$F(Y) = \left[2\,n_1(Y) - n_1(Y+1) - n_1(Y-1)\right]\left\{ -\log\left[\tfrac{1}{2}\,n_1(Y-1) + \tfrac{1}{2}\,n_1(Y+1)\right]\right.$$

$$\left. + \tfrac{1}{2}\log\left\{\left[\tfrac{1}{2} + n(1)\right] n_1(Y-1) \quad + \left[\tfrac{1}{2} - n(1)\right] n_1(Y+1)\right\}\right\}$$

$$-\left[-\tfrac{1}{2}\,n_2(Y-1) + \tfrac{1}{2}\,n_2(Y+1)\right] \log\left[\tfrac{1}{2}\,n_1(Y-1) + \tfrac{1}{2}\,n_1(Y+1)\right]$$

$$+\tfrac{1}{2}\left\{-\left[\tfrac{1}{2} + n(1)\right] n_2(Y-1) + \left[\tfrac{1}{2} - n(1)\right] n_2(Y+1)\right\} \log\left\{\left[\tfrac{1}{2} + n(1)\right] n_1(Y-1) + \left[\tfrac{1}{2} - n(1)\right] n_1(Y+1)\right\}$$

$$+\tfrac{1}{2}\left\{-\left[\tfrac{1}{2} - n(1)\right] n_2(Y-1) + \left[\tfrac{1}{2} + n(1)\right] n_2(Y+1)\right\} \log\left\{\left[\tfrac{1}{2} - n(1)\right] n_1(Y-1) + \left[\tfrac{1}{2} + n(1)\right] n_1(Y+1)\right\}$$

The last three terms, when integrated, may be recognized to be $\frac{1}{\Delta n_1(1)}$ times the gain of information received when the retransmitted amplitudes are raised from ± 1 to $\pm\left[1 + \Delta n_1(1)\right]$. As it is clear that this must produce a gain in information received, these three terms make a positive contribution to the integral (1).

The second factor in the first term in F(Y) may be written as

$$\frac{1}{2}\log\left\{1 - 4\,n_1^2(1)\,\frac{n_1(Y-1) - n_1(Y+1)}{n_1(Y-1) + n_1(Y+1)}\right\}$$

and consequently the contribution of the first term may be written as

$$\int_{-\infty}^{\infty} [n_1(Y) - n_1(Y-1)] \log\left\{ 1 - 4 n_1^2(1) \left[\frac{n_1(Y-1) - n_1(Y+1)}{n_1(Y-1) + n_1(Y+1)} \right]^2 \right\} dY \qquad (2)$$

Under the condition that $\frac{n_1(Y-1)}{n_1(Y+1)}$ is a non-decreasing function of Y for $Y > 0$, we can show that this last integral is positive. Let us note that the condition just stated is satisfied when $n_1(t) \sim e^{-\frac{t^2}{2N}}$ or when $n_1(t) \sim e^{-\frac{|t|}{\lambda}}$.

The logarithmic term in the last integral is an even function of Y which has a maximum at $Y = 0$, is constantly decreasing for $Y > 0$ and as $Y \to \infty$ it reaches the value

$$\log\left[1 - 4 n_1^2(1)\right] = A$$

If we write the logarithmic factor of the integrand of (2) as $A + f(Y)$, where $f(Y)$ is positive and even, we get for the integral (2)

$$2\int_{-\infty}^{\infty} f(Y)\, n_1(Y)\, dY - 2\int_{-\infty}^{\infty} f(Y)\, n_1(Y-1)\, dY$$

which, from the properties of $n_1(Y)$ and $f(Y)$, is positive.

q.e.d.

Appendix IV.C

The purpose of this appendix is to determine the numerical value of I_o, as defined by Eq. (IV.33).

It is convenient here to use, in expression (33), natural logarithms instead of logarithms to the base 2, the result is then written as $I_o^{(e)}$.
Using (34), we get

$$\log_e \tfrac{1}{2}\left[p(y|1) + p(y|-1)\right] = \log \frac{1}{2\sqrt{2\pi N}} - \frac{(y-1)^2}{2N} + \log_e \left(1 + e^{-\frac{2y}{N}}\right) \qquad \text{(C.1)}$$

$$= \log_e \frac{1}{2\sqrt{2\pi N}} - \frac{(y+1)^2}{2N} + \log_e \left(1 + e^{\frac{2y}{N}}\right) \qquad \text{(C.2)}$$

The first term of (33) is itself a sum of two terms X_1 and X_2 where

$$X_1 = -\frac{1}{2}\int_{-\infty}^{+\infty} \frac{e^{-\frac{(y-1)^2}{2N}}}{\sqrt{2\pi N}} \log_e \frac{1}{2}\left[p(y|1) + p(y|-1)\right] dy$$

$$X_2 = -\frac{1}{2}\int_{-\infty}^{\infty} \frac{e^{-\frac{(y+1)^2}{2N}}}{\sqrt{2\pi N}} \log_e \frac{1}{2}\left[p(y|1) + p(y|-1)\right] dy$$

Since $p(y|1) + p(y|-1)$ is even it is evident that $X_1 = X_2$.
In order to compute X_1, for positive y's we use Eq. (C.1) and for negative y's Eq. (C.2), thus

$$X_1 = -\tfrac{1}{2}\int_0^{\infty}\frac{dy}{\sqrt{N}}\ \frac{e^{-\frac{(y-1)^2}{2N}}}{\sqrt{2\pi}}\left\{-\log 2\sqrt{2\pi N} - \frac{(y-1)^2}{2N} + \log\left(1+e^{-\frac{2y}{N}}\right)\right\}$$

$$-\tfrac{1}{2}\int_{-\infty}^{0}\frac{dy}{\sqrt{N}}\ \frac{e^{-\frac{(y-1)^2}{2N}}}{\sqrt{2\pi}}\left\{-\log 2\sqrt{2\pi N} - \frac{(y+1)^2}{2N} + \log\left(1+e^{\frac{2y}{N}}\right)\right\}$$

Hence, by simple transformations,

$$X_1 = \tfrac{1}{2}\log 2\sqrt{2\pi N} + \tfrac{1}{4} - \frac{e^{-\frac{1}{2N}}}{\sqrt{2\pi N}} + \tfrac{1}{N}\left[1 - \Phi\left(\tfrac{1}{\sqrt{N}}\right)\right]$$

$$-\tfrac{1}{2}\int_0^{\infty}\frac{e^{-\frac{(y-1)^2}{2N}}}{\sqrt{2\pi N}}\log\left(1+e^{-\frac{2y}{N}}\right)dy - \tfrac{1}{2}\int_{-\infty}^{0}\frac{e^{-\frac{(y-1)^2}{2N}}}{\sqrt{2\pi N}}\log\left(1+e^{\frac{2y}{N}}\right)dy \qquad (C.3)$$

Now for $y \geqslant 0$

$$\log\left(1+e^{-\frac{2y}{N}}\right) = -\sum_{1}^{\infty}\frac{(-1)^k}{k}\, e^{-\frac{2ky}{N}}$$

and since

$$\int_0^{\infty}\frac{e^{-\frac{(y-1)^2}{2N}}}{\sqrt{2\pi N}}\, e^{-\frac{2ky}{N}}\,dy = e^{-\frac{(2k-1)^2-1}{2N}}\left[1 - \Phi\left(\frac{2k-1}{\sqrt{N}}\right)\right]$$

where

$$\Phi(z) = \int_{-\infty}^{z}\frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,dt$$

the last two terms of (C.3) become respectively,

$$\frac{1}{2} e^{-\frac{1}{2N}} \sum_{k=1}^{\infty} \frac{(-1)^k}{k} e^{\frac{(2k-1)^2}{2N}} \left[1 - \Phi\left(\frac{2k-1}{\sqrt{N}}\right)\right] \tag{C.4}$$

$$\frac{1}{2} e^{-\frac{1}{2N}} \sum_{k=1}^{\infty} \frac{(-1)^k}{k} e^{\frac{(2k+1)^2}{2N}} \left[1 - \Phi\left(\frac{2k+1}{\sqrt{N}}\right)\right] \tag{C.5}$$

If we remember that the contribution of the last two terms of (33) is $-\log\sqrt{2\pi e N}$ and if we combine (C.4) and (C.5) we finally get

$$I_0^{(e)} = \log_e 2 + \left(\frac{2}{N} - 1\right)\left[1 - \Phi\left(\frac{1}{\sqrt{N}}\right)\right] - 2\frac{e^{-\frac{1}{2N}}}{\sqrt{2\pi N}}$$

$$+ e^{-\frac{1}{2N}} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k(k+1)} e^{\frac{(2k+1)^2}{2N}} \left[1 - \Phi\left(\frac{2k+1}{\sqrt{N}}\right)\right]$$

If we use the asymptotic expansion of $1 - \Phi(x)$ we get

$$I_0^{(e)} = \log_e 2 + \left(\frac{2}{N} - 1\right)\left[1 - \Phi\left(\frac{1}{\sqrt{N}}\right)\right] - 2\,\frac{e^{-\frac{1}{2N}}}{\sqrt{2\pi N}}$$

$$- N\,\frac{e^{-\frac{1}{2N}}}{\sqrt{2\pi N}}\left[S_1 - N S_3 + 1\cdot 3\, N^2 S_5 - 1\cdot 3\cdot 5\, N^3 S_7 + \cdots\right]$$

where

$$S_\alpha = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k(k+1)(2k+1)^\alpha} \qquad (\alpha = 1, 3, 5, \ldots)$$

$S_1 = .082\ 74$

$S_3 = .003\ 577$

$S_5 = .000\ 061\ 9$

$S_7 = .000\ 006\ 2$

It is of interest to compare the asymptotic values of I_M and I_0. Expressing them both in bits we have

$$I_M \sim 1 - \frac{1}{2} \frac{e^{-\frac{1}{2N}}}{\sqrt{2\pi N}} \log_2 e$$

$$I_0 \sim 1 - .583 \frac{e^{-\frac{1}{2N}}}{\sqrt{2\pi N}} \log_2 e$$

Thus in the optimum detector case the equivocation is roughly N times the equivocation of the M.A.P. detector case.

Biographical Note

Charles Auguste Desoer was born in Ixelles (Belgium) on January 11, 1926. When he was three he moved to Verviers (Belgium), a textile town whose mills are very much like those familiar to New Englanders. He attended the public schools of that city and graduated from High School receiving the "Prix Spécial du Gouvernement." Thanks to the courage and the imagination of the school staff the German occupant never caught up with him and he, therefore, escaped slave labor in Germany. In 1944, he volunteered for active service in the Belgian Army. After demobilization (1945) he attended the University of Liège from which he graduated, in 1949, receiving the degree of "Ingénieur Radio-Electricien." He started graduate work at the Massachusetts Institute of Technology in fall 1949. He became a Research Assistant at the Research Laboratory of Electronics in February 1951. He married Claudine P. Osterrieth in July 1951.

Bibliography

1. M. LEIFER and W. F. SCHREIBER, "Communication Theory," in Advances in Electronics, Vol. III, p. 306-348, Academic Press Inc., Publishers, New York, 1951.
2. E. C. CHERRY, A History of the Theory of Information, Proc. Inst. of Electrical Engineers (London), Part III, September 1951.
3. NORBERT WIENER, "The Extrapolation, Interpolation and Smoothing of Stationary Time Series," N.D.R.C. Report, M.I.T. (February 1, 1942); or in book form: same title, Technology Press.
4. NORBERT WIENER, "Cybernetics," The Technology Press, John Wiley & Sons, Inc., New York, 1948.
5. R. V. HARTLEY, "The Transmission of Information," Bell System Technical Journal, July 1928.
6. C. SHANNON, "The Mathematical Theory of Communications," Bell System Technical Journal 27, p. 384, July and October, 1948.
7. C. SHANNON and W. WEAVER, "The Mathematical Theory of Communications," The University of Illinois Press, 1949.
8. C. SHANNON, "Communication in the Presence of Noise," Proc. Inst. Radio Engineers, January 1949.
9. B. M. OLIVER, J. R. PIERCE and C. SHANNON, "The Philosophy of P.C.M." Proc. Inst. Radio Engineers, November 1948.

10. R. M. FANO, "The Transmission of Information I and II," Technical Report #65 (March 1949) and #149 (February 1950), Research Laboratory of Electronics, M.I.T.

11. H. F. MAYER, "The Principles of Pulse Code Modulation" in "Advances in Electronics," Vol. III, Academic Press Inc., Publishers, New York, 1951.

12. R. M. FANO, Information Theory Seminar (unpublished notes) Spring term 1951.

13. A. N. KOLMOGOROV, "Foundation of the Theory of Probability," Chelsea Publishing Company, New York, 1950.

14. W. FELLER, "An Introduction to Probability Theory and its Applications," Vol. I, John Wiley & Sons, New York and Chapman and Hall, Ltd., London.

15. H. JEFFREYS and B. S. JEFFREYS, "Methods of Mathematical Physics," 2nd Edition, Cambridge University Press, 1950.

16. J. HADAMARD, "Leçons sur la Propagation des Ondes," Herman & C^{ie}, Paris, 1903.

17. G. H. HARDY, J. E. LITTLEWOOD, G. POLYA, "Inequalities," Cambridge University Press, 1934.

18. H. CRAMER, "Mathematical Methods of Statistics," Princeton University Press, 1946.

19. J. L. LAWSON and G. E. UHLENBECK, "Threshold Signals," Massachusetts Institute of Technology Radiation Laboratory Series, Vol. 24, McGraw Hill Book Co., Inc., New York 1950.

20. H. HURWITZ and M. KAC, "Statistical Analysis of Certain Random Functions," Ann. of Math. Stat. 15, p. 173, (1944).

21. M. J. GOLAY, "Notes on Digital Cooling," Proc. Inst. of Radio Engineers 37, 6, 657 (June 1949).

22. R. W. HAMMING, "Error Detecting and Error Correcting Codes," Bell System Technical Journal 29, 147, (April 1950).

23. H. S. M. COXETER, "Regular Polytopes," Methuen and Co., Ltd., London 1948.

24. S. GOLDMAN, "Some Fundamental Consideration Concerning Noise Reduction and Range in Radar and Communications," Proc. Inst. of Radio Engineers, 36, 5, 584 (1948).

25. G. BIRKHOFF and S. MacLANE, "A Survey of Modern Algebra," The McMillan Company, New York, 1948.

26. M. J. GOLAY, "Note on the Theoretical Efficiency of Information Reception with P.P.M.," Proc. Inst. of Radio Engineers 37, 9, 1031 (September 1943).

27. P. M. WOODWARD and I. L. DAVIS, "A Theory of Radar Information," Phil. Mag. (October 1950).

28. P. M. WOODWARD and I. L. DAVIS, "Information Theory and Inverse Probability in Telecommunication," Proc. Inst. of Electrical Engineers (London) 59, III, 58, March 1952.

29. R. M. FANO, "Information Theory," Lecture notes, Spring term, 1952.

30. R. COURANT and D. HILBERT, "MATHEMATISCHE PHYSIK I," Interscience Publishers Inc., New York, 1943.

31. H. E. SINGLETON, "Theory of Nonlinear Transducers," Technical Report 160, Research Laboratory of Electronics, M.I.T.