CODING AND DECODING FOR TIME-DISCRETE AMPLITUDE CONTINUOUS MEMORYLESS CHANNELS

by

JACOB ZIV

B.Sc., Technion -- Israel Institute of Technology (1954)
Ingenieur (E.E.), Technion -- Israel Institute of Technology (1955)
M.Sc., Technion -- Israel Institute of Technology (1957)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF SCIENCE
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February, 1962

Signature redacted

Signature of Author ....
Department of Electrical Engineering, January 8, 1962

Signature redacted

Certified by ....
Thesis Supervisor

Signature redacted

Accepted by ....
Chairman, Departmental Committee on Graduate Students

mRl

E. E.

Thesis

1962

Sc.D.

CODING AND DECODING FOR TIME-DISCRETE AMPLITUDE-CONTINUOUS MEMORYLESS CHANNELS

by

JACOB ZIV

Submitted to the Department of Electrical Engineering on January 8, 1962 in partial fulfillment of the requirements for the degree of Doctor of Science.

ABSTRACT

In this research we consider some aspects of the general problem of encoding and decoding for time-discrete, amplitude-continuous memoryless channels. The results are summarized below.

1. Signal Space Structure: A scheme for constructing a discrete signal space, for which sequential encoding-decoding methods are possible for the general continuous memoryless channel, is described in Chapter II. We consider random code selection from a finite ensemble. The engineering advantage is that each code word is sequentially generated from a small number of basic waveforms. The effects of these signal-space constraints on the average probability of error, for different signal power constraints, are also discussed.

2. Decoding Schemes: In Chapter III we discuss the application of sequential decoding to the continuous asymmetric channel. A new decoding scheme for convolutional codes, called successive decoding, is introduced in Chapter III. This new decoding scheme yields a bound on the average number of decoding computations for asymmetric channels that is tighter than has yet been obtained for sequential decoding. The corresponding probabilities of error of the two decoding schemes are also discussed in Chapter III.

3. Quantization at the Receiver: In Chapter IV, we consider the quantization at the receiver, and its effects on probability of error and receiver complexity.

Thesis Supervisor: John M. Wozencraft
Title: Associate Professor of Electrical Engineering

ACKNOWLEDGEMENT

The author acknowledges the encouragement and constructive criticism provided by Professor J.M. Wozencraft, the thesis supervisor, and Professors R.M. Fano and E.M. Hofstetter who served as thesis readers.

The original works of Professors J.M. Wozencraft, B. Reiffen and R.M. Fano provided the foundation for this research, and the author's debt to them is obvious. The author also wishes to acknowledge helpful discussions with Professors W.M. Siebert and R.G. Gallager and with Dr. T. Kailath of Jet Propulsion Laboratories. The author wishes to thank Miss Debby Ford who has patiently typed this thesis, and Melpar, Inc. for the use of their facilities for the multilith printing of this work.

The work presented in this thesis has been sponsored by a fellowship received from the Scientific Department, Israel Ministry of Defence.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

a	The number of input symbols per information digit
$A = \frac{\xi}{\sigma}$	Voltage signal-to-noise ratio
$A_{max} = \frac{\xi_{max}}{\sigma}$	Maximum signal-to-noise ratio
b	The number of branches emerging from each branching point in the convolutional tree code
C	Channel capacity
$D(u, v) = \ell n \frac{f(v)}{p(v\|u)}$	The "distance" between u and v
$d(x, y) = \ell n \frac{f(y)}{p(y\|x)}$	The "distance" between x and y
d	The dimensionality (number of samples) of each input symbol
E(R)	The optimum exponent of the upper bound to the probability of error (achieved through random coding)
$E_{\ell, d}(R)$	The exponent of the upper bound to the probability of error when the continuous input space is replaced by the discrete input set X_ℓ
f(y)	A probability-like function (Appendix A)
g(s), g(r, t)	Moment generating functions (Appendix A)
i	The number of source information digits per constraint-length (code word)
ℓ	The number of input symbols (vectors) in the discrete input space X_ℓ
m	The number of d-dimensional input symbols per constraint length (code word)
n	The number of samples (dimensions) per constraint length (code word)
$\overline{N}$	Average number of computations
P	The signal power
R	The rate of information per sample

R_{crit}	The critical rate above which E(R) is equal to the exponent of the lower bound to the probability of error
R_{comp}	The computational cut-off rate (Chapter III)
U	The set of all possible words of length n samples
u	The transmitted code word
u'	A member of U other than the transmitted message u
V	The set of all possible output sequences
v	The output sequence (a member of V)
X	The set of all possible d-dimensional input symbols
x	A transmitted symbol
x'	A member of X other than x
X_ℓ	The discrete input set that consists of ℓ d-dimensional vectors (symbols)
Y	The set of all possible output symbols
Ξ	The set of all possible input samples
ξ	A sample of the transmitted waveform u
ξ'	A sample of u'
H	The set of all possible output samples
η	A sample of the received sequence v
σ^2	The power of a Gaussian noise

CHAPTER I

INTRODUCTION

We intend to study some aspects of the problem of communication via a memoryless channel. A block diagram of a general communication system for such a channel is shown in Figure 1. The source consists of M equi-probable words of length T seconds each. The channel is of the following type: Once each $\frac{T}{n}$ seconds a real number is chosen at the transmitting point. This number is transmitted to the receiving point but is perturbed by noise, so that the ith real number ξ_i is received as η_i. Both ξ and η are members of continuous sets and therefore the channel is time discrete but amplitude continuous.

The channel is also memoryless in the sense that its statistics are given by a probability density $p(\eta_i | \xi_1, \xi_2, \ldots, \xi_i)$ such that

$$p(\eta_i | \xi_1, \xi_2, \ldots, \xi_i) = p(\eta_i | \xi_i) \qquad (1\text{-}1)$$

where

$$p(\eta_i | \xi_i) = p(\eta | \xi); \quad \xi = \xi_i, \ \eta = \eta_i \qquad (1\text{-}2)$$

A code word, or signal, of length n for such a channel is a sequence of n real numbers $(\xi_1, \ldots, \xi_n)$. This may be thought of geometrically as a point in n-dimensional Euclidean space. The type of channel we are studying here is, of course, closely related to a band limited channel (W cycles per seconds wide). For such a band limited channel we have $n = 2WT$.

The encoder maps the M messages into a set of M code words (signals).

The decoding system for such a code is a partitioning of the n dimensional output space into M subsets corresponding to the messages from 1 to M.

For a given coding and decoding system there is a definite probability of error for receiving a message. This is given by

FIGURE 1 COMMUNICATION SYSTEM
FOR MEMORYLESS CHANNELS

$$P_e = \frac{1}{M} \sum_{i=1}^{M} P_{e_i} \tag{1-3}$$

where P_{e_i} is the probability, if message i is sent, that it will be decoded as a message other than i.

The rate of information per sample is given by

$$R = \frac{1}{n} \ln M \tag{1-4}$$

We are interested in coding systems that, for a given rate R, minimize the probability of error, Pr.

In 1959, C.E. Shannon (1) studied coding and decoding systems for a time discrete but amplitude continuous channel with additive Gaussian noise, subject to the constraint that all code words were required to have exactly the same power. Upper and lower bounds were found for the probability of error when using optimal codes and optimal decoding systems. The lower bound followed from sphere packing arguments, and the upper bound was derived by using random coding arguments.

In random coding for such a Gaussian channel one considers the ensemble of codes obtained by placing M points randomly on a surface of a sphere of radius nP , (where nP is the power of each one of M signals, and n = 2WT where T is the time length of each signal and W is the bandwidth of the channel). More precisely, each point is placed independently of all other points with a probability measure proportional to surface area or, equivalently, to solid angle. Shannon's upper and lower bounds for the probability of error are very close together for signaling rates from some R_{crit} up to channel capacity C.

R.M. Fano (2) has recently studied the general discrete memoryless channel. In this case the signals are not constrained to have exactly the same power. If random coding is used, the upper and lower bounds for the probability of error are again very close together for all rates R above some R_{crit}.

The detection scheme that was used in both of these studies is an optimal one, that is, one which minimizes the probability of error for a given code. Such a scheme requires that the decoder compute an a posteriori probability measure, or a quantity equivalent to it, for each of (say) the M allowable code words.

In Fano's and Shannon's cases it can be shown that a lower bound on the probability of error has the form

$$P_e \geqslant K^* e^{-E^*(R)n} \qquad (1\text{-}5a)$$

where K^* is a constant independent of n. Similarly, when optimum random coding is used, the probability of error is upper bounded by:

$$P_e \leqslant Ke^{-E(R)n}; \qquad E(R) = E^*(R) \text{ for } R \geqslant R_{crit} \qquad (1\text{-}5b)$$

(In general, construction of a random code involves the selection of messages with some probability density P(u) from the set U of all possible messages. When P(u) is such that E(R) is maximized for the given rate R, the random code is called optimum.)

The behavior of $E^*(R)$ and E(R) as a function of R is given in Figure 2.

Fano's upper-bounding technique may be extended to include continuous channels, for all cases where the integrals involved exist. One such case is the gaussian channel. However, the lower bound is valid for discrete channels only. Therefore, as far as the continuous channel is concerned, the upper and lower bounds are not necessarily close together for rates $R \geqslant R_{crit}$.

The characteristics of many continuous physical channels, when quantized, are very close to the original ones if the quantization is fine enough. Thus, for such channels we have $E^*(R) = E(R)$ for $R \geqslant R_{crit}$.

We see from Figure 2, that the specification of an extremely small probability of error for a given rate R implies in general a significantly large value for the number of words M and for the number of decoding computations.

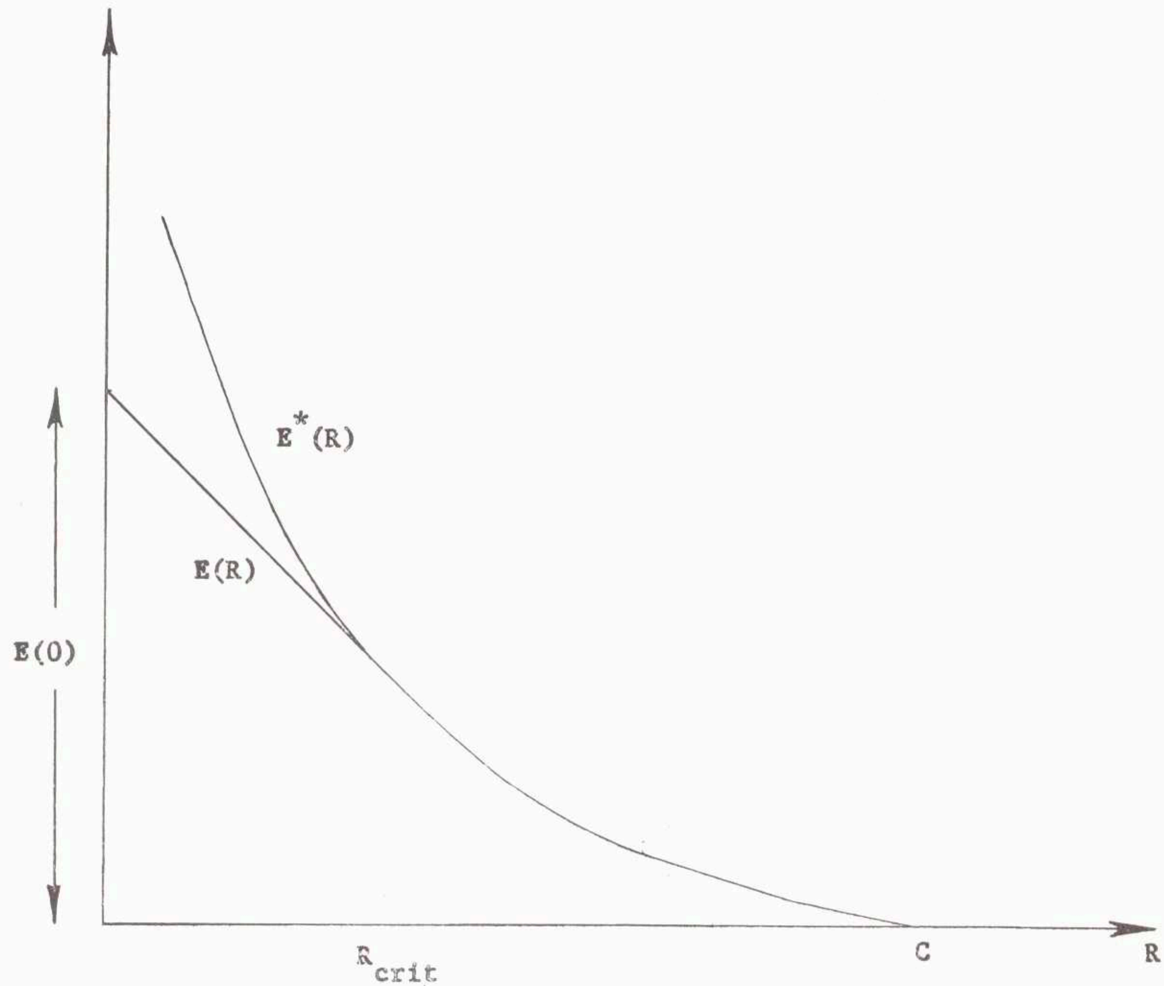


FIGURE 2 THE BEHAVIOR
OF $E^*(R)$ AND $E(R)$ AS A FUNCTION
OF R

J.L. Kelly (3) has derived a class of codes for continuous channels. These are block codes in which the (exponentially large) set of code words can be computed from a much smaller set of generators by a procedure analogous to group coding for discrete channels. Unfortunately, there seems to be no simple detection procedure for these codes. The receiver must generate each of the possible transmitted combinations and must then compare them with the received signal.

The sequential coding scheme of J.M. Wozencraft (4), extended by B. Reiffen (5), (6) is a code well suited to the purpose of reducing the number of coding and decoding computations. They have shown that, for a suitable sequential decoding scheme, the average number of decoding computations for channels which are symmetric at their output* is bounded by an algebraic function of n for all rates below some R_{comp}. Thus, the average number of decoding computations is not an exponential function of n as is the case when an optimal detection scheme is used.

In this research, we consider the following aspects of the general problem of encoding and decoding for time-discrete memoryless channels: 1. Signal space structure, 2. sequential decoding schemes, and 3. the effect of quantization at the receiver. Our results for each aspect are summarized below.

1. Signal Space Structure: A scheme for constructing a discrete signal space, in such a way as to make the application of sequential encoding-decoding possible for the general continuous memory-less channel, is described in Chapter II. In particular, whereas Shannon's work (1) considered code selection from an infinite ensemble, in this investigation the ensemble is a finite one. The engineering advantage is that each code word can be sequentially generated from a small set of basic waveforms. The effects of these signal spare constraints on the average probability of error, for different

---

*A channel with transition probability matrix $P(y|x)$ is symmetric at its output if the set of probabilities $P(y|x_1)$, $P(y|x_2)$, ... is the same for all output symbols y.

signal power constraints, are also discussed in Chapter II.

2. Sequential Decoding Schemes: In Chapter III we discuss the application of the sequential decoding scheme of Wozencraft and Reiffen to the continuous asymmetric channel. A lower bound on R_{comp} for such a channel is derived. The Wozencraft-Reiffen scheme provides a bound on the average number of computations that are needed to discard all the messages of the incorrect subset (5). No bound on the total number of decoding computations for asymmetric channels has heretofore been derived.

A new systematic decoding scheme for sequentially generated random codes is introduced in Chapter III. This decoding scheme, when averaged over the ensemble of code words, yields an average total number of computations that is upper-bounded by a quantity proportional to n^2, for all rates below some cut-off rate R_{comp}.

The corresponding probabilities of error of the two decoding schemes are also discussed in Chapter III.

3. Quantization at the Receiver: The purpose of introducing quantization at the receiver is to curtail the utilization of analogue devices. Due to the large number of computing operations which are carried out at the receiver, and the large flow of information to and from the memory, analogue devices may turn out to be more complicated and expensive than digital devices. In Chapter IV, the process of quantization at the receiver and its effect on the probability of error and the receiver complexity is discussed.

CHAPTER II

SIGNAL SPACE STRUCTURE

We proceed to introduce a structured signal space, and to investigate the effect of the particular structure on the probability of error.

2.1 The Basic Signal Space Structure

Let each code word of length n channel samples be constructed as a series of m elements, each of which has the same length d, as shown in Figure 3. Each one of the m elements is a member of a finite input space X_ℓ that consists of ℓ d-dimensional vectors ($d = \frac{n}{m}$), as shown in Figure 3. The advantage of such a structure is that a set of randomly constructed code words may be generated sequentially (4), (5), as discussed in Section 2.4.

Two cases will be considered

Case 1: The power of each of the n samples is less than or equal to P. (2-1)

Case 2: All code words have exactly the same power nP. (2-2)

The first case to be considered is that of Statement 2-1.

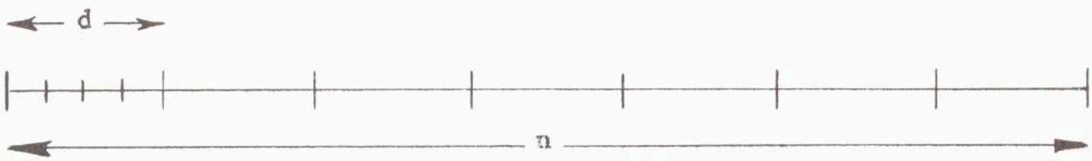
2.2 The Effect of the Signal Space Structure on the Average Probability of Error, Case 1

In order to evaluate the effect of a constrained input space on the probability of error, let us first consider the unrestricted channel.

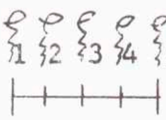
The constant memoryless channel is defined by the set of conditional probability densities $p(\eta \mid \xi)$, where ξ is the transmitted sample, and η is the corresponding channel output. The output η is considered to be a member of a continuous output ensemble H. By statement 2-1 we have

$$|\xi| \leqslant \sqrt{P} \qquad (2\text{-}1b)$$

Let us consider now the optimal unrestricted random code where each



$$m = \frac{n}{d}$$

Here, x  ; d = 5

$$d = \frac{n}{m}$$

FIGURE 3 CONSTRUCTION OF A
CODE WORD AS A SERIES OF ELEMENTS

particular message of length n is constructed by selecting the n samples independently at random with probability density $p(\xi)$ from a continuous ensemble Ξ. Then following Fano (2) it can be shown (Appendix A, Section 4) that the average probability of error over the ensemble of codes is bounded by

$$P_e \leqslant \begin{cases} 2e^{-nE(R)} & ; \quad R_{crit} \leqslant R < C \\ e^{-nE(R)} = e^{-n[E(0) - R]} & ; \quad 0 \leqslant R \leqslant R_{crit} \end{cases} \tag{2-3}$$

where $R = \frac{1}{n} \ln M$ is the rate per sample. $E(R)$ is the optimum exponent in the sense that it is equal, for large n and for $R \geqslant R_{crit}$, to the exponent of the lower bound to the average probability of error (Figure 2). For any given rate R, $p(\xi)$ is chosen so as to maximize $E(R)$ [i.e., to minimize P_e].

Let us now constraint each code word to be of the form shown in Figure 3, with the exception that we let the set X_e be replaced by a continuous ensemble with an infinite, rather than finite, number of members. We shall show that in this case, the exponent $E_d(R)$ of the upper bound to the average probability of error for such an input space can be made equal to the optimum exponent $E(R)$.

Theorem: Let us introduce a random code that is constructed in the following way: each code word of length n consists of m elements, where each element x is a d-dimensional vector

$$x = \xi_1, \xi_2, \ldots, \xi_d \tag{2-4}$$

selected independently at random with probability density $p(x)$ from the d-dimensional input ensemble X. Let the output event y that corresponds to x be

$$y = \eta_1, \eta_2, \ldots, \eta_d \tag{2-5}$$

y is a member of a d-dimensional output ensemble Y. The channel is defined by the set of conditional probabilities

$$p(y|x) = \prod_{i=1}^{d} p(\eta_i | \xi_i) \tag{2-6}$$

Also, let

$$p(x) = \prod_{i=1}^{d} p(\xi_i) \tag{2-7}$$

where $p(\xi_i) \equiv p(\xi)$, for all i, is the one dimensional probability density that yields the optimum exponent E(R). The average probability of error is then bounded by

$$P_e \geqslant \begin{cases} 2e^{-nE_d(R)} & ; R_{crit} \leq R < C \\ e^{-nE(R)} = e^{-n[E_d(0) - R]} & ; R \leqslant R_{crit} \end{cases} \tag{2-8}$$

where

$$E_d(R) \equiv E(R); \; E_d(0) = E(0) \tag{2-9}$$

<u>Proof</u>: The condition given by Eq. 2-7 is statistically equivalent to an independent, random selection of each one of the d samples of each element x. This corresponds to the construction of each code word by selecting each of the n samples independently at random with probability density $p(\xi)$ from the continuous space $\overleftrightarrow{\Xi}$, and therefore by Eqs. 2-7 and 2-3, yields the optimum exponent E(R).

<u>Q.E.D.</u>

The random code given by Eq. 2-6 is therefore an optimal random code, and yields the optimal exponent E(R).

We now proceed to evaluate the effect of replacing the continuous input space x by the discrete d-dimensional input space x , which consists of ℓ vectors. Consider a random code, for which the m elements of each word are picked at random with probability $\frac{1}{\ell}$ from the set X_ℓ of ℓ waveforms (vectors)

$$X_\ell = \left\{ x_k ; k = 1, \ldots, \right\} \tag{2-10}$$

The length or dimensionality of each x_k is d. Now let the set X_ℓ be generated in the following fashion: each vector x_k is picked at random with probability density $p(x_k)$ from the continuous ensemble X of all d-dimensional vectors matching the power constraint of Statement 2-1. The probability density $p(x_k)$ is given by

$$p(x_k) \equiv p(x) \qquad ; \qquad k = 1, \ldots, \ell \tag{2-11}$$

where p(x) is given by Eq. 2-7. Thus, we let $p(x_k)$ be identical with the probability density which was used for the generation of the optimal unrestricted random code. We can then state the following theorem.

<u>Theorem</u>: Let the general memoryless channel be represented by the set of probability densities $p(y|x)$. Given a set X_ℓ, let $E_{\ell,d}(R)$ be the exponent of the average probability of error over the ensemble of random codes constructed as above. Let $\overline{E_{\ell,d}(R)}$ be the expected value of $E_{\ell,d}(R)$ averaged over all possible sets X_ℓ.

Now define a tilted probability density for the product space XY

$$Q(x, y) = \frac{e^{sD(x,y)} p(x)\, p(y|x)}{\int_Y \int_X e^{sD(x,y)} p(x)\, p(y|x)\, dxdy} = \frac{p(x)\, p(y|x)^{1-s} f(y)^s}{\int_Y \int_X p(x)\, p(y|x)^{1-s} f(y)^s dxdy} \tag{2-12a}$$

where

$$f(y) = Q(y) = \frac{\left[\int_X p(x)\, p(y|x)^{1-s} dx \right]^{1/1-s}}{\int_Y \left[\int_X p(x)\, p(y|x)\, dx \right]^{1/1-s} dy} \quad ; \quad 0 \leq s \leq \frac{1}{2} \tag{2-12b}$$

$$Q(x|y) = \frac{Q(x,\ y)}{Q(y)} = \frac{p(x)\ p(y|x)^{1-s}}{\int_X p(x)\ p(y|x)^{1-s} dx} \quad ; \quad 0 \leqslant s \leqslant \frac{1}{2} \tag{2-13}$$

Then

$$1.\ \overline{E_{\ell,\ d}(R)} \geqslant E(R) - \frac{1}{d}\ \ln \frac{e^{F_1(R)} + \ell - 1}{\ell} \tag{2-14}$$

s and $F_1(R)$ are related parametrically to the rate R as shown below.

$$0 \leqslant F_1(R) = \ln \frac{\int_X \int_Y p(x)\ p(y|x)^{2(1-s)} Q(y)^{2s-1} dxdy}{\int_Y [\int_X p(x)\ p(y|x)^{1-s} dx]^2 Q(y)^{2s-1} dy} \quad ; \quad 0 \leqslant s \leqslant \frac{1}{2}$$

(2-15a)

$$R = \frac{1}{d} \int_X \int_Y Q(x,\ y)\ \ln \frac{Q(x|y)}{p(x)}\ dxdy \geqslant R_{crit} \tag{2-15b}$$

$R_{crit} = [R]_{s\ =\ 1/2}$

Also when $R \leqslant R_{crit}$

$$F_1(R) = F_1(R_{crit}) = dE(0) = -\ln \int_Y [\int_X p(x)\ p(y\ x)^{1/2} dx]^2 dy\ ;$$

$$s = \frac{1}{2} \tag{2-16}$$

$$2.\ \overline{E_{\ell,\ d}(R)} \geqslant E(R + \frac{1}{d}\ \ln \frac{e^{F_2(R)} + \ell - 1}{\ell}) \tag{2-17}$$

where $F_2(R)$ is related parametrically to the rate R by

$$0 \leq F_2(R) = \ln \frac{\int_X \int_Y p(x)\, p(y|x)^{2(1-s)} Q(y)^{2s-1} dx dy}{\int_Y [\int_X p(x)\, p(y|x)^{1-s}]^2\, Q(y)^{2s-1} dx dy} \quad ; 0 \leq s \leq \frac{1}{2}$$

(2-18a)

$$R = \frac{1}{d} \int_X \int_Y Q(x, y) \ln \frac{Q(x|y)}{p(x)} dx dy - \frac{1}{d} \ln \frac{e^{F_2(R)d} + \ell - 1}{\ell}$$

$$\geq R_{crit} - \frac{1}{d} \ln \frac{e^{E(0)d} + \ell - 1}{\ell} \tag{2-18b}$$

Also, when $R \leq R_{crit} - \frac{1}{d} \ln \frac{e^{E(0)d} + \ell - 1}{\ell}$

$$F_2(R) = F_2(R_{crit}) = E(0) = -\frac{1}{d} \ln \int_Y [\int_X p(x)\, p(y|x)^{1/2} dx]^2 dy \tag{2-19}$$

Proof: Given the set X_ℓ, each of the successive elements of a code word is generated by first picking the index k at random with probability $\frac{1}{\ell}$ and then taking x_k to be the element. Under these circumstances, by direct analogy to Appendix A, Eqs. A-46, A-41 and A-26 with the index k replacing the variable x, the average probability of error is bounded by

$$p(e|X_\ell) \leq e^{-nE^{(1)}_{\ell, d}(R)} + e^{-nE^{(2)}_{\ell, d}(R)} \tag{2-20}$$

where

$$E^{(1)}_{\ell, d}(R) = -R - \frac{1}{d} [\gamma_{\ell, d}(t, r) - r \frac{D_0}{m}] \tag{2-21}$$

$$E^{(2)}_{\ell,\, d}(R) = - \frac{1}{d} [\ \gamma_{\ell,\, d}(s) - sD_0] \qquad (2\text{-}22)$$

$$\gamma_{\ell,\, d}(t, r) = \ln g_{\ell,\, d}(t, r) \qquad (2\text{-}23a)$$

$$g_{\ell,\, d}(t, r) = \int_Y \sum_{k=1} \sum_{k=1} p(k)\ p(k')\ p(y|k)\ e^{(r-t)\ D(ky)\ +\ tD(k',\ y)} dy\ ; \quad r \leq 0\ ;\ t \leq 0$$

$$= \int_Y \sum_{i=1} \sum_{j=1} \frac{1}{\ell^2}\ p(y|x_i)\ e^{(r-t)\ D(x_i y)\ +\ tD(x_j y)} dy\ ; \quad r \leq 0\ ;\ t \leq 0 \qquad (2\text{-}23b)$$

$$D(ky) = D(x_k,\ y) = n \frac{f(y)}{p(y|x_k)} \qquad (2\text{-}24)$$

$f(y)$ is a positive function of y satisfying $\int_Y f(y)\ dy = 1$. D_0 is an arbitrary constant.

$$\gamma_{\ell,\, d}(s) = \ln g_{\ell,\, d}(s) \qquad (2\text{-}25a)$$

$$g_{\ell,\, d}(s) = \int_Y \sum_{k=1} p(k)\ p(y|k)\ e^{sD(ky)} dy\ ; \qquad 0 \leq s$$

$$= \int_Y \sum_{k=1} \frac{1}{\ }\ p(y|x_k)\ e^{sD(x_k,\ y)} dy\ ; \quad 0 \leq s \qquad (2\text{-}25b)$$

As in Eq. A-47, let D_0 be such that

$$E^{(1)}_{\ell,\, d}(R) = E^{(2)}_{\ell,\, d}(R) \qquad (2\text{-}26)$$

Inserting Eqs. 2-21 and 2-22 into Eq. 2-26 yields

$$- R - \frac{1}{d} [\ \gamma_{\ell,\, d}(t, r) - r \frac{D_0}{m}\ ] = - \frac{1}{d} [\ \gamma_{\ell,\, d}(s) - s \frac{D_0}{m}] \qquad (2\text{-}27)$$

Thus

$$\frac{1}{d} \frac{D_0}{m} = [\ \frac{1}{d}\ \gamma_{\ell,\, d}(s) - \frac{1}{d}\ \gamma_{\ell,\, d}(t, r) - R]\ \frac{1}{s-r} \qquad (2\text{-}28)$$

Inserting Eq. 2-28 into Eqs. 2-21 and 2-22 yields

$$E_{\ell,\, d}(R) = E^{(1)}_{\ell,\, d}(R) = E^{(2)}_{\ell,\, d}(R)$$

$$= - \frac{1}{s-r} \left[- \frac{r\, \gamma_{\ell,\, d}(s)}{d} + \frac{s\, \gamma_{\ell,\, d}(r,\, t)}{d} + sR \right] ;$$

$$0 \leq s,\ r \leq 0;\ t \leq 0 \qquad (2\text{-}29)$$

Inserting Eq. 2-29 into Eq. 2-20 yields

$$p(e|X_\ell) \leq 2e^{-nE_{\ell,\, d}(R)}$$

where $E_{\ell,\, d}(R)$ is given by Eq. 2-29.

We now proceed to evaluate a bound on the expected value of $E_{\ell,\, d}(R)$ when averaged over all possible sets X . The average value of $E_{\ell,\, d}(R)$ is by Eq. 2-29

$$\overline{E_{\ell,\, d}(R)} \geq - \frac{1}{s-r} \left[- \frac{r}{d}\, \overline{\gamma_{\ell,\, d}(s)} + \frac{s}{d}\, \overline{\gamma_{\ell,\, d}(r,\, t)} + sR \right] ;$$

$$0 \leq s,\ r \leq 0,\ t \leq 0 \qquad (2\text{-}30)$$

Ineq. 2-30 is not an equality since, in general, the parameters s, r and t should be chosen such as to maximize $E_{\ell,\, d}(R)$ of each individual input set X_ℓ , rather than being the same for <u>all</u> sets. From the convexity of the logarithmic function we have

$$- \overline{\ln x} \geq - \ln \overline{x} \qquad (2\text{-}31)$$

Inserting Eq. 2-31 into Eqs. 2-23a and 2-25a yields

$$- \overline{\gamma_{\ell,\, d}(s)} = - \overline{\ln g_{\ell,\, d}(s)} \geq - \ln \overline{g_{\ell,\, d}(s)} \qquad (2\text{-}32)$$

$$- \overline{\gamma_{\ell,\, d}(r,\, t)} \geq - \ln \overline{g_{\ell,\, d}(r,\, t)} \qquad (2\text{-}33)$$

Now, since $r \leq 0$, $s \geq 0$ we therefore have

$$\frac{r}{s-r} \leqslant 0 \; ; \quad - \frac{s}{s-r} \leqslant 0 \tag{2-34}$$

Inserting Ineqs. 2-32, 2-33 and 2-34 into Ineq. 2-30 yields

$$\overline{E_{\ell,\, d}(R)} \geqslant \frac{r}{s-r}\frac{1}{d}\, \ell n\, \overline{g_{\ell,\, d}(s)} - \frac{s}{s-r}\frac{1}{d}\, \ell n\, \overline{g_{\ell,\, d}(r,\, t)} - \frac{s}{s-r} R \tag{2-35}$$

From Eqs. 2-25b and 2-11 we have

$$\overline{g_{\ell,\, d}(s)} = \int_X p(x_k)\, g_{\ell,\, d}(s)\, dx_k = \frac{1}{\ell} \sum_{k=1} \int_Y \int_X p(x)\, p(y|x)\, e^{sD(x,y)} dxdy$$

where the index k has been dropped, since $p(x_k) = p(x)$. Thus

$$\overline{g_{\ell,\, d}(s)} = \int_Y \int_X p(x)\, p(y|x)\, e^{sD(xy)} dxdy \tag{2-36}$$

From Eq. 2-23b we have

$$\overline{g_{\ell,\, d}(t,\, r)} = \int_{X_i} \int_{X_j} p(x_i, x_j)\, g_{\ell,\, d}(t,\, r) dx_i dx_j \tag{2-37}$$

where, by construction

$$p(x_i,\, x_j) \begin{cases} = p(x_i)\, p(x_j) \; ; & i \neq j \\ = p(x_i)\, \delta\, (x_i - x_j) \; ; & i = j \end{cases} \tag{2-38}$$

and where, by Eq. 2-11

$$p(x_i) \equiv p(x), \qquad \text{for all } i \; . \tag{2-39}$$

Thus, from Eqs. 2-23b, 2-37, 2-38 and 2-39 we have

$$\overline{g_{\ell,\,d}(r,\,t)} = \overline{g_{\ell,\,d\;i\neq j}(r,t)} + \overline{g_{\ell,\,d\;i=j}(r,\,t)} \tag{2-41}$$

where

$$\overline{g_{\ell,\,d\;i\neq j}(r,\,t)} = \frac{1}{\ell^2}\sum_{\substack{i=1\\ i\neq j}}\sum_{j=1}\int_{X_i}\int_{X_j} p(x_i)\;p(x_j)\int_Y p(y|x_i)\;e^{(r-t)D(x_i,y)+tD(x_j,y)}\,dydx_idx_j$$

$$= \frac{1}{\ell^2}\sum_{i=1}\sum_{j=1}\int_Y\int_X\int_{X'} p(x)\;p(x')\;p(y|x)\;e^{(r-t)D(x,y)+tD(x',y)}\,dxdx'dy\;;$$

$$r \leq 0;\; t \leq 0 \tag{2-42}$$

and

$$\overline{g_{\ell,\,d\;i=j}(t,\,r)} = \frac{1}{\ell^2}\sum_{i=1}\int_Y\int_{X_i} p(x_i)\;p(y|x_i)\;e^{rD(x_i,y)}\,dxdy$$

$$= \frac{1}{\ell^2}\sum_{i=1}\int_Y\int_X p(x)\;p(y|x)\;e^{rD(x,y)}\,dxdy\;;\; r \leq 0 \tag{2-43}$$

Inserting Eq. 2-24 into Eqs. 2-42 and 2-43 yields

$$\overline{g_{\ell,\,d\;i\neq j}(r,\,t)} = \frac{\ell(\ell-1)}{\ell^2}\int_Y\int_X\int_{X'} p(x)\;p(x')\;p(y|x)^{1-r+t}\;p(y|x')^{-t}\;f(y)^{r}\,dxdx'dy$$

$$r \leq 0;\; t \leq 0 \tag{2-44}$$

$$\overline{g_{\ell,\,d\;i=j}(r,\,t)} = \frac{\ell}{\ell^2}\int_Y\int_X\int_{X'} p(x)\;p(y|x)^{1-r}f(y)^{r}\,dxdy;\quad r \leq 0 \tag{2-45}$$

In general, f(y) of Eq. 2-24 should be chosen so as to maximize $E_{\ell,\,d}(R)$

for each individual input set X_ℓ . However, we let f(y) be the same for all sets X_ℓ and equal to the f(y) that maximizes the exponent $E_d(R)$ corresponding to the unrestricted continuous set X. Thus inserting Eqs. A-52 and A-53 into Eqs. 2-44 and 2-36 yields

$$\overline{g_{\ell,\, d}(r,\, t)}_{i \neq j} = \frac{\ell-1}{\ell}\, e^{\gamma_d(r,\, t)} = \frac{\ell-1}{\ell}\, g_d(r,\, t) \tag{2-46}$$

$$\overline{g_{\ell,\, d}(r,\, t)}_{i = j} = \frac{1}{\ell}\, e^{\gamma_d(r)} = \frac{1}{\ell}\, g_d(r) \tag{2-47}$$

Thus, by Eq. 2-41

$$\overline{g_{\ell,\, d}(r,\, t)} = \frac{\ell-1}{\ell}\, g_d(r,\, t) + \frac{1}{\ell}\, g_d(r) \tag{2-48}$$

Also, by Eqs. A-52 and 2-36

$$\overline{g_{\ell,\, d}(s)} = g_{\ell,\, d}(s) \tag{2-49}$$

Inserting Eqs. 2-48 and 2-49 into Eq. 2-35 yields

$$\overline{E_{\ell,\, d}(R)} \geqslant \frac{r}{s-r}\,\frac{1}{d}\,\ell n\, g_d(s) - \frac{s}{s-r}\,\frac{1}{d}\,\ell n\Big[\frac{\ell-1}{\ell}\, g_d(r,\, t) + \frac{1}{\ell}\, g_d(r)\Big] - \frac{s}{s-r}\, R$$

$$= \frac{r}{s-r}\,\frac{1}{d}\,\ell n\, g_d(s) - \frac{s}{s-r}\,\frac{1}{d}\,\ell n\, g_d(r,\, t) - \frac{s}{s-r}\,\frac{1}{d}\,\ell n\Big[\frac{g_d(r)/g_d(r,t)+\ell-1}{\ell}\Big] - \frac{s}{s-r}\, R$$

Thus

$$\overline{E_{\ell,\, d}(R)} \geqslant \frac{r}{s-r}\,\frac{1}{d}\,\gamma_d(s) - \frac{s}{s-r}\,\frac{1}{d}\,\gamma_d(r,\, t) - \frac{s}{s-r}\,\frac{1}{d}\,\ell n\Big[\frac{e^{\gamma_d(r)-\gamma_d(r,t)}+\ell-1}{\ell}\Big] \tag{2-50}$$

Now, the exponent $E_d(R)$ that corresponds to the unconstrained d-dimensional continuous space X is given by Eq. A-49

$$E_d(R) = -\frac{1}{d}\,[\,\gamma_d(s) - s\,\frac{D_0}{m}\,] = -R + \frac{1}{d}\,[\,\gamma_d(r,t) - r\,\frac{D_0}{m}\,]$$

Eliminating D_0 yields

$$E_d(R) = \frac{r}{s-r}\,\frac{1}{d}\,\gamma_d(s) - \frac{s}{s-r}\,\frac{1}{d}\,\gamma_d(r,t) - \frac{r}{s-r}\,R \qquad (2\text{-}51)$$

Furthermore, $E_d(R)$ is maximized, as shown in Eqs. A-50, A-51, A-54, A-55 and A-56, by letting

$$f(y) = \frac{[\int_X p(x)\,p(y/x)^{1-s}dx\,]^{1/1-s}}{\int_Y [\int_X p(x)\,p(y|x)^{1-s}dx\,]^{1/1-s}dy} \qquad (2\text{-}52a)$$

$$r = 2s - 1\ ; \quad t = s - 1 \qquad (2\text{-}52b)$$

where, for $R \geqslant R_{crit}$ s is such that

$$R = \frac{1}{d}\,[(s-1)\,\gamma_d'(s) - \gamma_d(s)\,]\ ; \quad 0 \leqslant s \leqslant \frac{1}{2} \qquad (2\text{-}52c)$$

where

$$R_{crit} = [R]_{s\,=\,1/2} \qquad (2\text{-}52d)$$

If we let the parameters r, s, and t of Ineq. 2-50 be equal to those of Eq. 2-51 we have

$$\overline{E_{\ell,\,d}(R)} \geqslant E_d(R) - \frac{s}{s-r}\,\frac{1}{d}\,\ln\,[\,\frac{e^{\gamma_d(r) - \gamma_d(r,t)} + \ell - 1}{\ell}\,] \qquad (2\text{-}53)$$

The insertion of Eq. 2-52b yields

$$\overline{E_{\ell,\,d}(R)} \geqslant E_d(R) - \frac{s}{1-s}\,\frac{1}{d}\,\ln\,[\,\frac{e^{F_1(R)} + \ell - 1}{\ell}\,] \qquad (2\text{-}54)$$

where

$$F_1(R) = \gamma_d(2s - 1) - \gamma_d(2s - 1 ; \ s - 1)$$

Inserting Eqs. A-52 and A-53 into Eq. 2-55 yields

$$F_1(R) = \ln \int_Y \int_X p(x) \ p(y|x)^{2-2s} [f(y)]^{2s-1} dxdy$$

$$- \ln \int_Y \int_{X'} \int_X p(x) \ p(x') \ p(y|x)^{1-s} p(y|x')^{1-s} f(y)^{2s-1} dxdx'dy$$

Thus

$$F_1(R) = \ln \frac{\int_Y \int_X p(x) \ p(y|x)^{2(1-s)} f(y)^{2s-1} dxdy}{\int_Y [\int_X p(x) \ p(y|x)^{1-s} dx]^2 f(y)^{2s-1} dxdy} \tag{2-56}$$

where s and $F_1(R)$ are related parametrically to the rate R by Eq. (A-60c), for all rates above $R_{crit} = [R]_{s = 1/2}$.

As for rates below R_{crit}, we let

$$s = \frac{1}{2} ; \quad t = - \frac{1}{2} ; \quad r = 0 \tag{2-57}$$

Inserting Eq. 2-57 into Eqs. 2-54 and 2-55 yields, with the help of Eqs. A-69 and A-71

$$[F_1(R)]_{s = 1/2} = - \ln \int_Y [\int_X p(x) \ p(y|x)^{1/2} dx]^2 = dE_d(0) \tag{2-58a}$$

where

$$E_d(0) = [E_d(R)]_{R=0} \tag{2-58b}$$

$$\overline{E_{\ell, d}(R)} \geqslant E(R) - \frac{1}{d} \ln\left(\frac{e^{F_1(R)|_{s=1/2}} + \ell - 1}{\ell} \right) \tag{2-59a}$$

$$= E(R) - \frac{1}{d} \ln\left(\frac{e^{dE_d(0)} + \ell - 1}{\ell} \right) \tag{2-59b}$$

for $R \leqslant R_{crit}$. From Eqs. 2-11 and 2-9 we also have $E_d(R) \equiv E(R)$ for all rates, by construction. The proof of the first part of the theorem has therefore been completed.*

<u>Q.E.D.</u>

In order to prove the second part, let us rewrite Eq. 2-50 as

$$\overline{E_{\ell, d}(R)} \geqslant \frac{r}{s-r}\frac{1}{d}\gamma_d(s) - \frac{s}{s-r}\frac{1}{d}\gamma_d(r, t) - \frac{s}{s-r}\left[R + \frac{1}{d}\ln\frac{e^{\gamma_d(r)-\gamma_d(r, t)} + \ell - 1}{\ell}\right]$$

$$= \frac{r}{s-r}\frac{1}{d}\gamma_d(s) - \frac{s}{s-r}\frac{1}{d}\gamma_d(r, t) - \frac{s}{s-r}[R'] \tag{2-60}$$

where

$$R' = R + \frac{1}{d}\ln\frac{e^{\gamma_d(r) - \gamma_d(r, t)} + \ell - 1}{\ell} \tag{2-61}$$

Comparing Ineq. 2-60 with Eq. 2-51 yields

$$\overline{E_{\ell, d}(R)} \geqslant E_d(R') = E_d\left(R + \frac{1}{d}\ln\frac{e^{F_2(R)} + \ell - 1}{\ell}\right) \tag{2-62}$$

---

*A simplified proof for the region $R \leqslant R_{crit}$ is given in Ref. (7).

where, by Eqs. 2-52, 2-56, A-57, A-58, A-59, A-60c and A-60b, $F_2(R)$ is related parametrically to the rate R by

$$F_2(R) = \ln \frac{\int_Y \int_X p(x)\, p(y|x)^{2(1-s)} Q(y)^{2s-1} dxdy}{\int_Y [\int_X p(x)\, p(y|x)^{1-s} dx]^2 Q(y)^{2s-1} dxdy} \; ; \quad 0 \leqslant s \leqslant \frac{1}{2} \qquad (2\text{-}63)$$

$$R' = R + \frac{1}{d} \ln \frac{e^{F_2(R)} + \ell - 1}{\ell} \equiv \frac{1}{d} \int_X \int_Y Q(x, y) \ln \frac{Q(x|y)}{p(x)} \; ;$$

$$0 \leqslant s \leqslant \frac{1}{2} \qquad (2\text{-}64)$$

for all

$$R' \geqslant R'_{crit} = [R']_{s = 1/2} \qquad (2\text{-}65)$$

Ineq. 2-65 can be rewritten

$$R \geqslant R_{crit} - \frac{1}{d} \ln \frac{e^{F_2(R)|_{s=1/2}} + \ell - 1}{\ell}$$

$$= R_{crit} - \frac{1}{d} \ln \frac{e^{dE(0)} + \ell - 1}{\ell} \qquad (2\text{-}66)$$

As for rates below $R_{crit} - \frac{1}{d} \ln \frac{e^{dE(0)} + \ell - 1}{\ell}$, we let

$$s = \frac{1}{2}, \; t = -\frac{1}{2}, \; r = 0 \qquad (2\text{-}67)$$

Inserting Eq. 2-67 into Eqs. 2-62 and 2-63 yields

$$\overline{E_{\ell, d}(R)} \geqslant E_d\left(R + \frac{1}{d} \ln \frac{e^{E(0)d} + \ell - 1}{\ell} \right) ; \qquad R \leqslant R_{crit} \qquad (2\text{-}68)$$

From Eqs. 2-11 and 2-9 we have also that, by construction $E_d(R) \equiv E(R)$ for all rates. Thus, the proof of the second part of the theorem has been completed.

Q.E.D.

Discussion: We proceed now to discuss the bounds that were derived in the above theorem. We shall consider particularly the region $0 \leq R \leq R_{crit}$, which, as we shall see in Chapter 3, is of special interest. From Eqs. 2-14 and 2-16

$$\overline{E_{\ell, d}(R)} \geq E(R) - \frac{1}{d} \ln\left[\frac{e^{dE(0)} + \ell - 1}{\ell} \right] \quad \text{for } R \leq R_{crit} \tag{2-69}$$

From Eq. 2-3 we have

$$E(R) = E(0) - R \qquad \text{for } R \leq R_{crit} \tag{2-70}$$

Inserting Eq. 2-70 into Eq. 2-69 yields

$$\overline{E_{\ell, d}(R)} \geq E(0) - R - \frac{1}{d} \ln\left[\frac{e^{dE(0)} + \ell - 1}{\ell} \right] \quad \text{for } R \leq R_{crit} \tag{2-71}$$

Now, whenever

$$dE(0) \ll 1 \tag{2-72}$$

we have

$$\overline{E_{\ell, d}(R)} \mathrel{\overset{\sim}{\geq}} E(0) - R - \frac{1}{d} \ln\left(\frac{1 + dE(0) + \ell - 1}{\ell} \right)$$

$$\cong E(0) - R - \frac{1}{d} \ln\left(1 + \frac{dE(0)}{\ell} \right)$$

$$\cong E(0) - R - \frac{E(0)}{\ell}; \qquad \text{for } R \leq R_{crit}$$

Thus

$$\overline{E_{\ell,\, d}(R)} \gtrsim E(0)\ [\,1 - \frac{1}{\ell}\,] - R \qquad \text{for } R \leqslant R_{crit}$$

$$\text{and } dE(0) << 1 \qquad (2\text{-}73)$$

Comparing Eq. 2-73 with Eq. 2-70, we see that $\overline{E_{\ell,\, d}(0)}$ can be made to be practically equal to E(R) by using quite a small* number ℓ, of input symbols, whenever $E(0)d << 1$.

Ineq. 2-71 may be bounded by

$$\overline{E_{\ell,\, d}(R)} \geqslant E(0) - R - \frac{1}{d}\,\ell n\,[\,\frac{e^{E(0)d} + \ell}{\ell}\,] \qquad R \qquad R_{crit}$$

$$= E(0) - R - \frac{1}{d}\,\ell n\,[\,e^{E(0)d\, -\, \ell n\, \ell} + 1\,] \qquad (2\text{-}74)$$

---

*In cases where $|\xi|_{max} \to 0$ and $\frac{dp(\eta|\xi)}{d\xi}\Big|_{\xi=0} \neq 0$ so that $p(\eta|\xi)$ can be replaced by the first two terms of the Taylor series expansion, $p(\eta|\xi) = p(\eta|0) + \frac{dp(\eta|\xi)}{d\xi}\Big|_{\xi=0}$, it can be shown (by insertion into Eqs. A-74a-d) that

1. The optimum input space consists of two oppositely directed vectors, ξ_{max} and $-\xi_{max}$, for all rates $0 \leqslant R < C$

2. $$E(0) = \frac{1}{2}C = \frac{1}{4}\xi^2_{max} \int \frac{\left[\frac{dp(\eta|\xi)}{d\xi}\Big|_{\xi=0}\right]^2}{p(\eta|0)}\, d\eta$$

   where C is the channel capacity.

Thus, whenever

$$dE(0) >> 1 \tag{2-75}$$

we have from Ineq. 2-74

$$\overline{E_{\ell, d}(R)} \gtrapprox E(0) - R - \frac{1}{d} \ln [e^{E(0)d - \ln \ell}] \cong \frac{1}{d} \ln \ell - R \tag{2-76a}$$

when

$$\frac{1}{d} \ln \ell << E(0) \tag{2-76b}$$

and

$$\overline{E_{\ell, d}(R)} \gtrapprox E(0) - R - \frac{1}{d} \ln [e^{E(0)d - \ln \ell} + 1]$$

$$\cong E(0) - \frac{1}{d} e^{-[\ln \ell - E(0)d]} - R \tag{2-77a}$$

when

$$\frac{1}{d} \ln \ell >> E(0) \tag{2-77b}$$

Comparing Eq. 2-76 with Eq. 2-77 yields

$$E(R) \geqslant \overline{E_{\ell, d}(R)} \cong E(R) \ ; \qquad R \leqslant R_{crit}, \ dE(0) >> 1 \tag{2-78a}$$

or

$$E_{\ell, d}(R) \cong E(R) \tag{2-78b}$$

if

$$\frac{1}{d} \ln \ell \geqslant E(0) \qquad\qquad dE(0) >> 1 \tag{2-78c}$$

In the following section we shall discuss the construction of a semi-optimum finite input set X_ℓ for the Gaussian channel. [A semi-optimum input space is one that yields an exponent $E_{\ell, d}(R)$ which is practically equal to $E(R)$.] We shall show that the number of input vectors ℓ needed is approximately the same as that indicated by Eqs. 2-78 and 2-73. This therefore demonstrates the effectiveness of the bounds derived in this section.

2.3 Semi-Optimum Input Spaces for the White Gaussian Channel - Case 1 (2.1)

The white Gaussian channel is defined by the transition probability density

$$p(\eta | \xi) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(\eta-\xi)^2}{2\sigma^2}} \qquad (2\text{-}79)$$

where by Ineq. (2-1b)

$$|\xi| \leq |\xi|_{max} = \sqrt{P} \qquad (2\text{-}1b)$$

Let us define the "voltage signal to noise ratio" A, as

$$A = \frac{\xi}{\sigma} \qquad (2\text{-}80)$$

Inserting Eq. (2-1b) into 2-80 yields

$$A \leq A_{max} = \frac{|\xi|_{max}}{\sigma} = \frac{\sqrt{P}}{\sigma} \qquad (2\text{-}81)$$

We shall first discuss the case where

$$dA^2_{max} << 1 \qquad (2\text{-}82)$$

and proceed with the proof of the following theorem.

Theorem: Consider a white Gaussian Channel whose statistics are given by Eq. 2-79. Let the input signal power be constrained by Ineq. (2-1b) and by Ineq. 2-82. Let the input space consist of two d-dimensional oppositely-directed vectors. Then the exponent of the upper bound to the probability of error, $E_{2,\,d}(R)$ is asymptotically equal to the optimum exponent $E(R)$.

Proof: From Eq. 2-4 we have

$$x = \xi_1, \xi_2, \ldots, \xi_d \qquad (2\text{-}4)$$

The input set X_2 consists of two oppositely directed vectors. Let those two vectors be given by

$$x_1 = \xi_1^{\,1}, \xi_2^{\,1}, \ldots, \xi_d^{\,1} \tag{2-89a}$$

where

$$\xi_1^{\,1} = \xi_2^{\,1} = \ldots = \xi_d^{\,1} = \xi_{max} \tag{2-89b}$$

and

$$x_2 = \xi_1^{\,2}, \xi_2^{\,2}, \ldots, \xi_d^{\,2} \tag{2-90a}$$

where

$$\xi_1^{\,2} = \xi_2^{\,2} = \ldots = \xi_d^{\,2} = -\xi_{max} \tag{2-90b}$$

From Eqs. 2-5 and 2-6 we have

$$y = \eta_1, \eta_2, \ldots, \eta_d \tag{2-5}$$

$$p(y|x) = \prod_{i=1}^{d} p(\eta_i | \xi_i) \tag{2-6}$$

Inserting Eqs. 2-79, 2-89 and 2-90 into Eq. 2-6 yields

$$p(y|x_1) = \prod_{i=1}^{d} \frac{1}{(2\pi)^{d/2}\sigma^d}\, e^{-\frac{(\eta_i - \xi_{max})^2}{2\sigma^2}} \tag{2-91a}$$

$$p(y|x_2) = \prod_{i=1}^{d} \frac{1}{(2\pi)^{d/2}\sigma^d}\, e^{-\frac{(\eta_i + \xi_{max})^2}{2\sigma^2}} \tag{2-91b}$$

From Eqs. 2-29 and 2-30 we have

$$p(e| X_2) \leqslant 2e^{-nE_{2,d}(R)} \tag{2-92a}$$

where

$$E_{2,\ d}(R) = \frac{r}{s-r}\frac{1}{d}\gamma_{2,\ d}(s) - \frac{s}{s-r}\frac{1}{d}\gamma_{2,\ d}(r,\ t) - \frac{s}{s-r}R \tag{2-92b}$$

Let

$$r = 2t + 1\ ;\quad s = 1 + t\ ;\quad 0 \leqslant s \leqslant \frac{1}{2} \tag{2-93}$$

Also, let

$$f(y) = p(y|0) = \prod_{i=1}^{d} \frac{1}{(2\pi)^{d/2}\sigma^d} e^{-\frac{\eta_i^2}{2\sigma^2}} \tag{2-94}$$

and

$$p(x_1) = p(x_2) = \frac{1}{2} \tag{2-95}$$

Inserting Eqs. 2-93, 2-94 and 2-95 into Eqs. 2-23, 2-25 and 2-92b yields

$$E_{2,\ d}(R) = -\frac{1-2s}{1-s}\frac{1}{d}\gamma_{2,\ d}(s) - \frac{s}{1-s}\frac{1}{d}\gamma_{2,\ d}(2s-1\ ,\ s-1) - \frac{s}{1-s}R \tag{2-96}$$

$$\gamma_{2,\ d}(s) = \ln \sum_{i=1}^{2} \int_Y \frac{1}{2}\, p(y|x_i)^{1-s} p(y|0)^s dy \tag{2-97}$$

$$\gamma_{2,\ d}(2s-1,\ s-1) = \ln \sum_{i=1}^{2} \int_Y \frac{1}{4}\, p(y|x_i)^{1-s} p(y|x_j)^{1-s} p(y|0)^{2s-1} \tag{2-98}$$

Inserting Eqs. 2-91 and 2-94 into Eqs. 2-97 and 2-98 yields

$$\gamma_{2,\,d}(s) = \ln \frac{1}{2} \left\{ \left[\int_{\eta} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(1-s)\,(\eta - \xi_{max})^2 + \eta^2 s}{2\,\sigma^2}} d\eta \right]^d \right.$$

$$\left. + \left[\int_{\eta} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(1-s)(\eta + \xi_{max})^2 + \eta^2 s}{2\,\sigma^2}} d\eta \right]^d \right\}$$

$$= \ln \frac{1}{2} \left\{ 2e^{-\frac{\xi^2_{max}(1-s)sd}{2\,\sigma^2}} \right\}$$

$$= - \frac{\xi^2_{max} ds(1-s)}{2\,\sigma^2} ; \qquad 0 \leq s \leq \frac{1}{2} \tag{2-99}$$

$$\gamma_{2,\,d}(2s-1,\, s-1) = \ln \frac{1}{4} \left\{ \left[\int_{\eta} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{2(1-s)(\eta - \xi_{max})^2 + (2s-1)\eta^2}{2\,\sigma^2}} d\eta \right]^d \right.$$

$$+ \left[\int_{\eta} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{2(1-s)(\eta + \xi_{max})^2 + (2s-1)\eta^2}{2\,\sigma^2}} d\eta \right]^d$$

$$\left. + 2\left[\int_{\eta} \frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-(1-s)(\eta - \xi_{max})^2 - (1-s)(\eta + \xi_{max}) - (2s-1)\eta^2}{2\,\sigma^2}} d\eta \right]^d \right\}$$

$$= \ln \frac{1}{4} \left\{ 2e^{-\frac{\xi^2_{max} d[2(1-s) - 4(1-s)^2]}{2\,\sigma^2}} + 2e^{-\frac{2(1-s)\,\xi^2_{max} d}{2\,\sigma^2}} \right\};$$

$$0 \leq s \leq \frac{1}{2} \tag{2-100}$$

Now, since by Ineq. 2-82

$$dA^2_{max} = \frac{\xi^2_{max}}{\sigma^2} << 1$$

we have

$$\gamma_{2,\,d}(2s-1,\ s-1) \cong \ln \frac{1}{4}\ [4 - \frac{2\xi^2_{max}}{2\sigma^2}\ d[2(1-s)-r(1-s)^2 - \frac{2\xi^2_{max}}{2\sigma^2}\ d\ 2(1-s)]$$

$$= \ln\ [\ 1 - \frac{\xi^2_{max}}{2\sigma^2}\ 2ds(1-s)\ ]$$

$$\cong -\frac{\xi^2_{max}}{2\sigma^2}\ 2ds(1-s) = 2\,\gamma_{2,\,d}(s) \qquad (2\text{-}101)$$

Inserting Eqs. 2-99 and 2-101 into Eq. 2-96 yields

$$E_{2,\,d}(R) = +\ [\ \frac{1-2s}{1-s} + \frac{2s}{1-s}\ ]\ s(1-s)\ \frac{\xi^2_{max}}{2\sigma^2} - \frac{s}{1-s}\ R$$

$$= s\ \frac{\xi^2_{max}}{2\sigma^2} - \frac{s}{1-s}\ R\ ; \qquad 0 \leq s \leq \frac{1}{2}$$

$$= s\ \frac{A^2_{max}}{2} - \frac{s}{1-s}\ R\ ; \qquad 0 \leq s \leq \frac{1}{2} \qquad (2\text{-}102)$$

Maximizing $E_{2,\,d}(R)$ with respect to s yields

$$s = \frac{1}{2} \quad \text{for } R \leq R_{crit} = \frac{1}{8}\,A^2_{max}$$

$$s = 1 - \frac{\sqrt{2R}}{A} \quad \text{for } R \geq R_{crit} = \frac{1}{8}\,A^2_{max}$$

Thus:

$$E_{2,\ d}(R) \cong \frac{1}{4} A_{max}^{2} - R \ ; \qquad R \lesssim \frac{1}{8} A_{max}^{2} \qquad (2\text{-}103a)$$

$$E_{2,\ d}(R) \cong \frac{1}{2} A^{2} - 2A \sqrt{\frac{R}{2}} + R \ ; \qquad \frac{1}{8} A_{max}^{2} \leq R \lesssim \frac{1}{2} A_{max}^{2} \qquad (2\text{-}103b)$$

Comparing Eq. 2-103 with the results given in page 654 of Ref (1) yields*

$$E(R) \geqslant E_{2,d}(R) \geqslant E(R) \qquad (2\text{-}104)$$

Thus

$$E_{2,\ d}(R) \cong E(R) \qquad (2\text{-}105)$$

for

$$A_{max}^{2} d << 1$$

<u>Q.E.D.</u>

We proceed now to discuss cases where the condition of Ineq. 2-82 is no longer valid. The first step will be the evaluation of $E_{\ell,\ d}(0)$ for the white Gaussian channel. From Eqs. A-69 and A-71 we have

$$E_{\ell,\ d}(0) = -\frac{1}{d} \ln \sum_{X_\ell'} \sum_{X_\ell} p(x)\ p(x') \int_{Y} p(y|x)^{1/2} p(y|x')^{1/2} dy \qquad (2\text{-}106)$$

---

*The results of Ref. (1) are derived for the power constraint of Statement 2-2, and are valid also in cases where the average power is constrained to be equal to P.

The set of signals satisfying Statement 2-1 is included in the set of signals satisfying the above average power constraint. Thus, Shannon's exponent of the upper bound to the probability of error is larger than or equal to the optimum exponent E(R) which corresponds to the constraint of Statement 2-1.

Inserting Eqs. 2-6 and 2-79 into Eq. 2-106 yields

$$E_{\ell,d}(0) = -\frac{1}{d} \ln \sum_{X} \sum_{X'} p(x)\, p(x') \prod_{i=1}^{d} \int_{\eta} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(\eta_i - \xi_i)^2 - (\eta_i - \xi'_i)^2}{4\sigma^2}} d \tag{2-107a}$$

where

$$x = \xi_1, \xi_2, \ldots, \xi_d \tag{2-107b}$$

$$x' = \xi'_1, \xi'_2, \ldots, \xi'_d \tag{2-107c}$$

Thus

$$E_{\ell,d}(0) = -\frac{1}{d} \ln \sum_{X} \sum_{X'} p(x)\, p(x') \prod_{i=1}^{d} e^{-\frac{(\xi_i - \xi'_i)^2}{8\sigma^2}}$$

$$= -\frac{1}{d} \ln \sum_{X} \sum_{X'} p(x)\, p(x')\, e^{-\sum_{i=1}^{d} \frac{(\xi_i - \xi'_i)^2}{8\sigma^2}} \tag{2-108}$$

Let D be the geometrical distance between the two vectors x and x', given by

$$D^2 = [x - x']^2 = \sum_{i=1}^{d} (\xi_i - \xi'_j)^2 \tag{2-109}$$

Then, inserting Eq. 2-109 into Eq. 2-108 yields

$$E_{\ell,d}(0) = -\frac{1}{d} \ln \sum_{X} \sum_{X'} p(x)\, p(x')\, e^{-\frac{|x-x'|^2}{8\sigma^2}} \tag{2-110a}$$

or

$$E_{\ell, d}(0) = -\frac{1}{d} \ln \sum_{X} \sum_{X'} p(D)\, e^{-\frac{D^2}{8\sigma^2}} \qquad (2\text{-}110b)$$

where p(D) can be found from p(x) and p(x').

In the case where the input set X_2 consists of two oppositely directed vectors given by Eqs. 2-89, 2-90 and 2-95, we get from Eq. 2-108

$$E_{2, d}(0) = -\frac{1}{d} \ln \frac{1}{2} \left(1 + e^{-\frac{\xi_{max}^2 d}{2\sigma^2}} \right)$$

$$= -\frac{1}{d} \ln \frac{1}{2} \left(1 + e^{-\frac{A_{max}^2 d}{2}} \right) \qquad (2\text{-}111)$$

Again, for

$$\frac{A_{max}^2 d}{2} << 1 \qquad \text{we have} \qquad E_{2, d}(0) \cong \frac{A_{max}^2}{4}$$

as in Eq. (2-103a).

For higher values of peak signal-to-noise ratio we let d = 1. Then, by Eq. 2-111

$$E_{2, 1}(0) = -\ln \frac{1}{2} \left(1 + e^{-\frac{A_{max}^2}{2}} \right) \qquad (2\text{-}112)$$

$E_{2, 1}(0)$ together with $C_{2, 1}$, the rate for which $E_{2, 1}(R) = 0$, are given in Table 1. Also given in the same table are the channel capacity C* and the zero-rate exponent E(0),** that correspond to the power constraint of Statement 2-2. (The channel capacity C and the zero-rate exponent E(0) which

---

*The channel capacity is computed in Ref. (9).

**E(0) is computed in Ref. (2) and in Appendix B.

A_{max}	$E_{2,1}(0)$	$E(0)$	$\frac{E_{2,1}(0)}{E(0)}$	$C_{2,1}$	C	$C_{2,1}/C$
1	0.216	0.22	0.99	0.343	0.346	0.99
2	0.571	0.63	0.905	0.62	0.804	0.77
3	0.683	0.95	0.72	0.69	1.151	0.60
4	0.69	1.20	0.57	0.69	1.4	0.43

TABLE 1

TABULATION OF $E_{2,1}(0)$ AND $C_{2,1}$ vs. A_{max}

corresponds to Statement 2-1 are, as shown in the footnote on page 32, upper bounded by the C and E(0) that correspond to the power constraint of Statement 2-2.) From Table 1 we see that the replacing of the continuous input set by the discrete input set X_2, consisting of two oppositely directed vectors, has a negligible effect on the exponent of the probability of error as $A^2_{max} \lessapprox 1$.

Let us next consider the case where the input set X_ℓ consists of ℓ one-dimensional vectors as shown in Figure 4. The distance between each two adjacent vectors is

$$D_{min} = \frac{2\,\xi_{max}}{\ell - 1} \tag{2-113}$$

Let

$$p(x_i) = \frac{1}{\ell} \quad ; \quad i = 1, \ldots, \ell \tag{2-114}$$

Inserting Eqs. 2-113 and 2-114 into Eq. 2-110 yields

$$E_{2,\,1}(0) = -\ln \frac{1}{\ell^2} \left[\ell + 2\sum_{k=1}^{\ell} (\ell - k)\, e^{-\frac{(kD_{min})^2}{8\sigma^2}} \right] \tag{2-115}$$

Thus, since $4k < k^2 \; ; \; k \geq 2$ we have

$$E_{2,\,1}(0) \geq -\ln \frac{1}{\ell^2} \left[\ell + 2(\ell-1)\, e^{-\frac{D^2_{min}}{8\sigma^2}} + 2(\ell-2) \sum_{k=1}^{\ell} e^{-\frac{4kD^2_{min}}{8\sigma^2}} \right]$$

$$\geq -\ln \frac{1}{\ell^2} \left[\ell + 2(\ell-1)\, e^{-\frac{D^2_{min}}{8\sigma^2}} - 2(\ell-2) \frac{e^{-4kD^2_{min}/8\sigma^2}}{e^{-4D^2_{min}/8\sigma^2} - 1} \right]$$

$$\geq -\ln \frac{1}{\ell} \left\{ 1 + \left[2e^{-\frac{D^2_{min}}{8\sigma^2}} + 2 \frac{1}{e^{\frac{4D^2_{min}}{8\sigma^2}} - 1} \right] \right\} \tag{2-116}$$

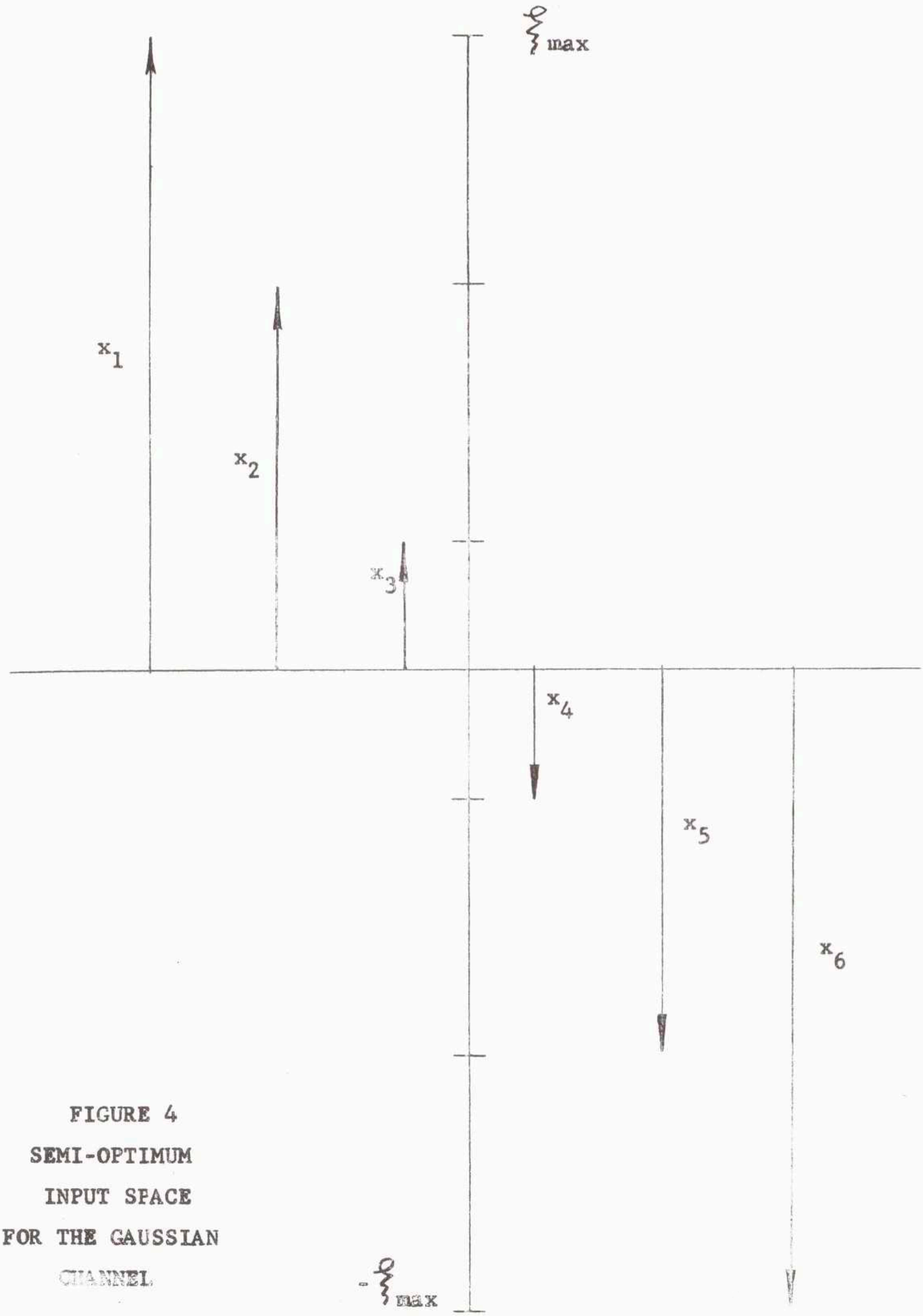


FIGURE 4
SEMI-OPTIMUM
INPUT SPACE
FOR THE GAUSSIAN
CHANNEL

Now define K by

$$\ell - 1 = \frac{A_{max}}{K} \quad ; \qquad 0 \leq K \tag{2-117}$$

Inserting Eqs. 2-117 and 2-81 into Eq. 2-113 yields

$$D_{min} = 2\sigma K \tag{2-118}$$

Inserting Eq. 2-118 into Eq. 2-116 then yields

$$E_{2,\,1}(0) \geq \ln(1 + KA_{max}) - \ln\left\{1 + 2e^{-\frac{K^2}{2}} + \frac{2}{e^{+2K^2}-1}\right\} \tag{2-119a}$$

If we choose ℓ so that $K \cong 1$, we have

$$E_{2,\,1}(0) = \ln(1 + A_{max}) - \ln 2{,}52 \tag{2-119b}$$

From Eqs. 2-119b and 2-117 we have, for $A_{max} >> 1$

$$E_{2,\,1}(0) \cong \ln A_{max} \tag{2-120a}$$

$$\ln \ell \cong \ln A_{max} = E_{2,\,1}(0) \tag{2-120b}$$

On the other hand, it can be shown (Appendix B) that

$$E(0) \lesssim \ln A_{max} \quad ; \qquad A_{max} >> 1 \tag{2-121}$$

Thus, by Eqs. 2-121 and 2-120, we have

$$E_{2,\,1}(R) \cong E(R) \quad ; \qquad R < R_{crit},\ A_{max} >> 1 \tag{2-122a}$$

if

$$\ln \ell = E(0) \qquad ; \qquad d = 1 \tag{2-122b}$$

Comparing Eqs. 2-73 and 2-78 with Eqs. 2-105 and 2-122 respectively yields that the lower bound on $E_{\ell, d}(R)$ derived in Section 2.2 is indeed a useful one.

2.4 The Effect of the Signal Space Structure on the Average Probability of Error, Case 2 (Statement 2-2)

In the case of a power constraint such as that of Statement 2-2, we consider the ensemble of codes obtained by placing M points on the surface of a sphere of radius $\sqrt{nP}$.

The requirement of Statement 2-2 can be met by making each of the m elements in our signal space have the same power dP (see Figure 3). (The power of each word is therefore mdP = nP and therefore Statement 2-2 is satisfied). This additional constraint produces an additional reduction in the value of $E_{\ell, d}(R)$ as compared with E(R). Even if we let the d-dimensional input space X_ℓ be an infinite set ($\ell = \infty$), the corresponding exponent $E_d(R)$ will, in general, be

$$E_d(R) \leq E(R) \qquad (2\text{-}123)$$

The discussion in this section will be limited to the white Gaussian channel and to rates below R_{crit}. Thus

$$E_d(R) = E_d(0) - R \quad ; \qquad R \leq R_{crit} \qquad (2\text{-}124)$$

Let

$$E_d(0) = E(0) - k_d(\overline{A^2})\, E(0) \qquad (2\text{-}125a)$$

where

$$\overline{A^2} = \frac{P}{\sigma^2} \qquad (2\text{-}125b)$$

Then by Eqs. 2-124 and 2-125 we have

$$E_d(R) = E(0) - k_d(\overline{A^2})\, E(0) - R \quad ; \qquad R \leq R_{crit} \qquad (2\text{-}126)$$

We shall now proceed to evaluate $k_d(\overline{A^2})$ as a function of $\overline{A^2}$ for different values of d.

The input space X is, by construction, a set of points on the surface of a d-dimensional sphere of radius $\sqrt{dP}$.

Let each point of the set X be placed at random and independently of all others with probability measure proportional to surface area or, equivalently, to solid angle. The probability $Pr(0 \leq \theta \leq \theta_1)$ that an angle between two vectors of the space X is less than or equal to θ_1, is therefore proportional to the solid angle of a cone in d-dimensions with half angle θ_1. This is obtained by summing the contributions due to ring-shaped elements of area (spherical surfaces in d-1 dimensions of radius sin θ and incremental width dθ as shown in Figure 5.) Thus, the solid angle of the cone is given by (1):

$$\Omega(\theta_1) = \frac{(d-1)\ \pi^{(d-1)/2}}{\Gamma\left(\frac{d+1}{2}\right)} \int_0^{\theta_1} (\sin\theta)^{d-2}\, d\theta \qquad (2\text{-}127)$$

Here we used the formula for the surface $s_d(r)$ of a sphere of radius r in d-dimensions,

$$s_d(r) = \pi^{d/2}\, r^{d-1} / \Gamma(d/2 + 1)$$

From Eq. 2-127 we have

$$Pr(0 \leq \theta \leq \theta_1) = \frac{\Omega(\theta_1)}{\Omega(\pi)} = \frac{(d-1)\ \pi^{(d-1)/2}}{\Gamma\left(\frac{d+1}{2}\right)}\ \frac{\int_0^{\theta_1} (\sin\theta)^{d-2} d\theta}{d\pi^{d/2} / \Gamma\left(\frac{d+1}{2}\right)}$$

$$= \frac{d-1}{d\sqrt{\pi}}\ \frac{\Gamma\left(\frac{d}{2}+1\right)}{\Gamma\left(\frac{d+1}{2}\right)} \int_0^{\theta_1} (\sin\theta)^{d-2} d\theta$$

$$= \frac{1}{\sqrt{\pi}}\ \frac{\Gamma\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d-1}{2}\right)} \int_0^{\theta_1} (\sin\theta)^{d-2} d\theta \qquad (2\text{-}128)$$

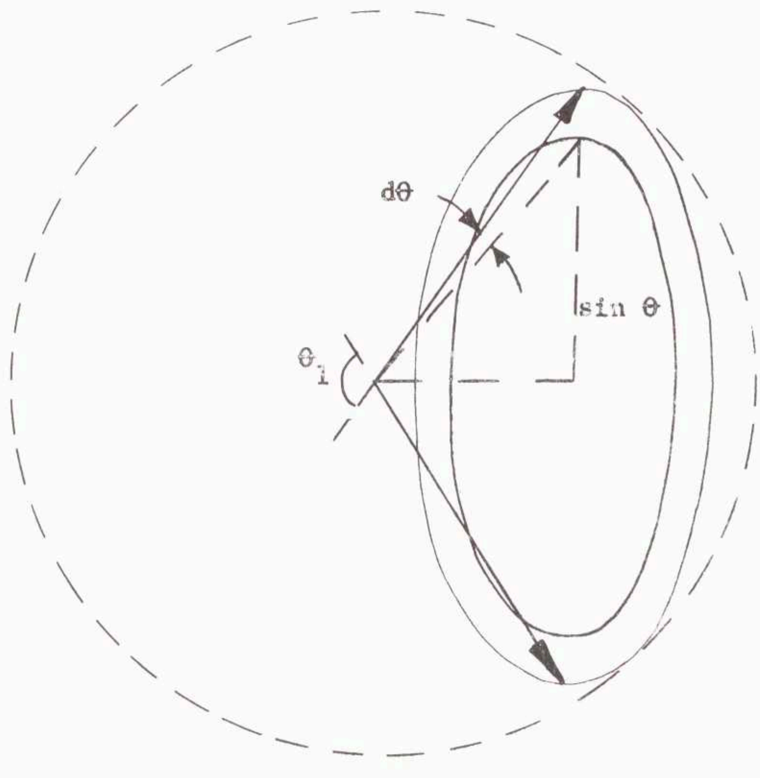


FIGURE 5 CAP CUTOUT BY A
CONE ON THE UNIT SPHERE

The probability density $p(\theta)$ is therefore given by

$$p(\theta) = \frac{d\Pr(0 \leq \theta \leq \theta_1)}{d\theta_1} = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} (\sin \theta)^{d-2} \tag{2-129}$$

Now, by Eq. 2-109, the geometrical distance between two vectors with an angle θ between them is (see Figure 5)

$$D^2 = 4(dP \sin^2 \frac{\theta}{2}) \tag{2-130}$$

Inserting Eqs. 2-130 and 2-129 into Eq. 2-110b yields

$$E_d(0) = -\frac{1}{d} \ln \int_0^{\pi} p(\theta)\, e^{-\frac{dP}{2\sigma^2} \sin^2 \frac{\theta}{2}} d\theta$$

$$= -\frac{1}{d} \ln \left\{ \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} \int_0^{\pi} e^{-\frac{dP}{2\sigma^2} \sin^2 \frac{\theta}{2}} (\sin \theta)^{d-2} d\theta \right\}. \tag{2-131}$$

Inserting Eq. 2-125b into 2-131 yields, for $d \geqslant 2$

$$E_d(0) = -\frac{1}{d} \ln \left\{ \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} \int_0^{\pi} e^{-\frac{d\overline{A^2}}{2} \sin^2 \frac{\theta}{2}} (\sin \theta)^{d-2} d\theta \right\} \tag{2-132a}$$

or

$$E_d(0) = -\frac{1}{d} \ln \left\{ \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} e^{-\frac{d\overline{A^2}}{4}} \int_0^{\pi} e^{\frac{d\overline{A^2}}{4} \cos \theta} \sin \theta^{d-2} d\theta \right\}. \tag{2-132b}$$

Equation 2-132 is valid for all $d \geqslant 2$. As for $d = 1$, it is clear that in order to satisfy the power constraint of Statement 2-2 the input space X must consist of two oppositely directed vectors with an amplitude of $\sqrt{P}$. Thus

$$A^2_{max} = \overline{A^2} = P \tag{2-133}$$

Inserting Eq. 2-133 into Eq. 2-111 yields

$$E_1(0) = -\ell n \left(\frac{1 + e^{-\overline{A^2}/2}}{2} \right) \tag{2-134}$$

In Appendix B we shown, for all d, that

$$E_d(0) \cong \frac{1}{4} A^2 = E(0) \ ; \quad A^2 d << 1 \ ; \quad d \geqslant 2 \tag{2-135a}$$

$$E_d(0) \cong \frac{d-1}{d} \frac{1}{2} \ell n \, A^2 \ ; \quad A^2 >> 1 \ ; \quad d \geqslant 2 \tag{2-135b}$$

Thus

$$E_d(0) \cong \frac{d-1}{d} E(0) \tag{2-135c}$$

Inserting Eqs. 2-134 and 2-135 into Eq. 2-125 yields, for <u>any</u> d

$$E_d(0) = E(0) - k_d(\overline{A^2}) \, E(0) \tag{2-136a}$$

where

$$k_d(\overline{A^2}) = 0 \ ; \qquad \overline{A^2} d << 1 \tag{2-136b}$$

$$k_d(\overline{A^2}) \cong \frac{1}{d} \ ; \qquad \overline{A^2} >> 1 \tag{2-136c}$$

The qualitative behaviour of $k_d(\overline{A^2})$ as a function of $\overline{A^2}$ and with d as a parameter is given in Figure 6.*

$k_1(\overline{A^2})$ and $k_3(\overline{A^2})$ are tabulated in Table 2.

---

*From Eq. 2-132a it is clear that $E_d(0)$ is a monotonic increasing function of $\overline{A^2}$.

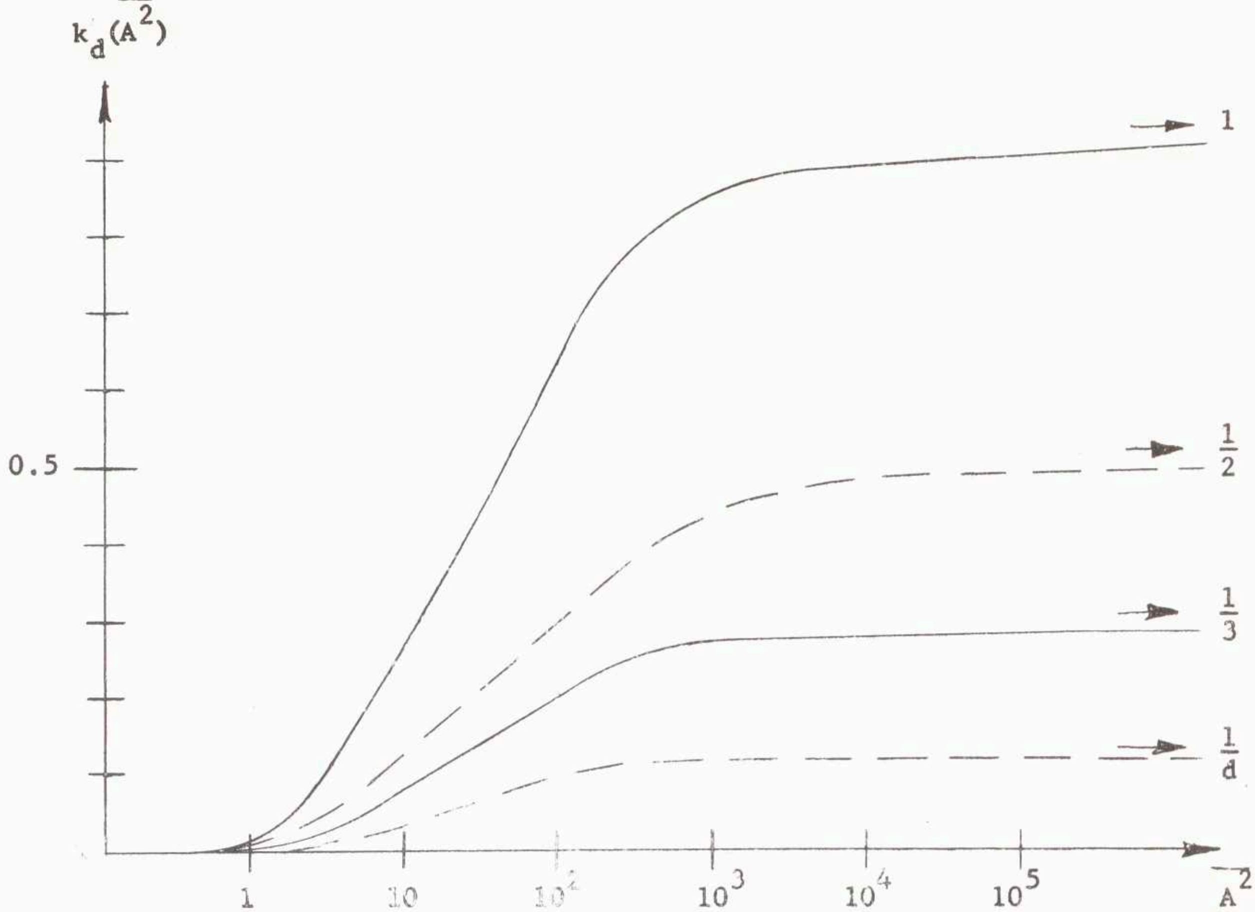


FIGURE 6

PLOT OF $k_d(\overline{A^2})$ vs. $\overline{A^2}$

A^2	$k_1(\overline{A^2})$	$k_3(\overline{A^2})$
1	0.01	
4	0.095	0.046
9	0.28	0.095
16	0.43	0.135
100	0.7	0.2
10^4	0.9	0.28

TABLE 2

TABULATION OF $k_1(\overline{A^2})$ AND $k_3(\overline{A^2})$ vs. $\overline{A^2}$

We proceed now to evaluate the effect of replacing the continuous d-dimensional input space X by a discrete d-dimensional input space X_ℓ, which consists of ℓ vectors.

Let each of the m elements be picked at random with probability $\frac{1}{\ell}$ from the set X_ℓ of ℓ vectors (waveforms), $X_\ell \equiv \{x_k: k = 1, \ldots, \ell\}$.

Let the set X_ℓ be generated in the following fashion: each vector x_k is picked at random with probability $p(x_k)$ from the ensemble X of all d-dimensional vectors matching the power constraint of Statement 2-2. The probability $p(x_k)$ is given by

$$p(x_k) = p(x)\Big|_{x = x_k} \quad ; k = 1, \ldots, \ell$$

where $p(x_k)$ is the same probability distribution which is used to generate $E_d(0)$. The following theorem can then be stated:

Theorem: Let $E_{\ell,\,d}(0)$ be the zero-rate exponent of the average probability of error for random codes constructed as above. Let $\overline{E_{\ell,\,d}(0)}$ be the expected value of $E_{\ell,\,d}(0)$ averaged over all possible sets X_ℓ. Then

$$\overline{E_{\ell,\,d}(0)} \geqslant E_d(0) - \frac{1}{d}\,\ell n\left(\frac{e^{dE_d(0)} + \ell - 1}{\ell}\right) \tag{2-137}$$

The proof is identical with that of the theorem of Section 2-2. Inserting Eq. 2-136 into Eq. 2-137 yields

$$\overline{E_{\ell,\,d}(0)} \geqslant E(0) - k_d(\overline{A^2})\,E(0) - \frac{1}{d}\,\ell n\left(\frac{e^{dE_d(0)} + \ell - 1}{\ell}\right) \tag{2-138}$$

Thus, there is a combined loss due to the two following independent constraints:

1. Constraining the power of each of the input vectors to be equal to dP ; the resulting loss is equal to $k_d(\overline{A^2})\,E(0)$.
2. Constraining the input space to consist of ℓ vectors only; the resulting loss is equal to $\frac{1}{d}\,\ell n\,\frac{e^{dE_d(0)} + \ell - 1}{\ell}$.

We now evaluate the effect of these constraints at high (and low) values of $\overline{A^2}$. From Eq. 2-137 we have

$$\overline{E_{\ell, d}(0)} \geqslant E_d(0) - \frac{1}{d}\ell n \left(\frac{e^{dE_d(0)} + \ell}{\ell}\right)$$

$$= E_d(0) - \frac{1}{d}\ell n(e^{dE_d(0) - \ell n \ell} + 1) \qquad (2\text{-}139)$$

Thus for $E(0) \cong \ell n \, \overline{A^2} >> 1$ we have

$$\overline{E_{\ell, d}(0)} \geqslant \frac{1}{d} \ell n \ell \quad ; \quad \frac{1}{d} \ell n \ell \ll E_d(0) \cong \frac{d-1}{d} \frac{\ell n \overline{A^2}}{2} \qquad (2\text{-}140a)$$

$$\overline{E_{\ell, d}(0)} \cong E_d(0) \quad ; \quad \frac{1}{d} \ell n \ell >> E_d(0) \cong \frac{d-1}{d} \frac{\ell n \overline{A^2}}{2} \qquad (2\text{-}140b)$$

On the other hand we have always

$$E_{\ell, d}(0) \leqslant E_d(0) \qquad (2\text{-}141)$$

Inserting Eq. 2-141 into Eq. 2-140b yields

$$E_{\ell, d}(0) \cong E_d(0) = \frac{d-1}{d} \frac{1}{2} \ell n \, \overline{A^2} \qquad (2\text{-}142a)$$

$$\text{for } \frac{1}{d} \ell n \ell >> E_d(0) \qquad (2\text{-}142b)$$

Whenever $\overline{A^2} d << 1$, an input space X_2 that consists of two oppositely directed vectors with an amplitude of $\sqrt{dP}$ yields the optimum exponent $E(R)$ for all rates $0 \leq R < C$, as shown in Section 2.3 of this chapter.

2.5 Convolutional Encoding

In the last three sections we have established a discrete signal space, generated from a d-dimensional input space which consists of ℓ input symbols. We have shown that a proper selection of ℓ and d yields an exponent $E_{\ell, d}(R)$ which is arbitrarily close to the optimum exponent $E(R)$.

We proceed now to describe an encoding scheme for mapping output sequences from an independent letter source into sequences of channel input symbols which are all members of the input set X_ℓ. We desire to do this encoding in a sequential manner so that sequential or other systematic decoding may be attempted at the receiver. By sequential encoding we mean that the channel symbol to be transmitted at any time is uniquely determined by the sequence of the output letters from the message source up to that time.

Decoding schemes for sequential codes will be discussed in the next chapter.

Let us consider a randomly selected code with a constraint length n, for which the size of M(w), the set of allowable messages at the length w input symbols, is an exponential function of the variable w.

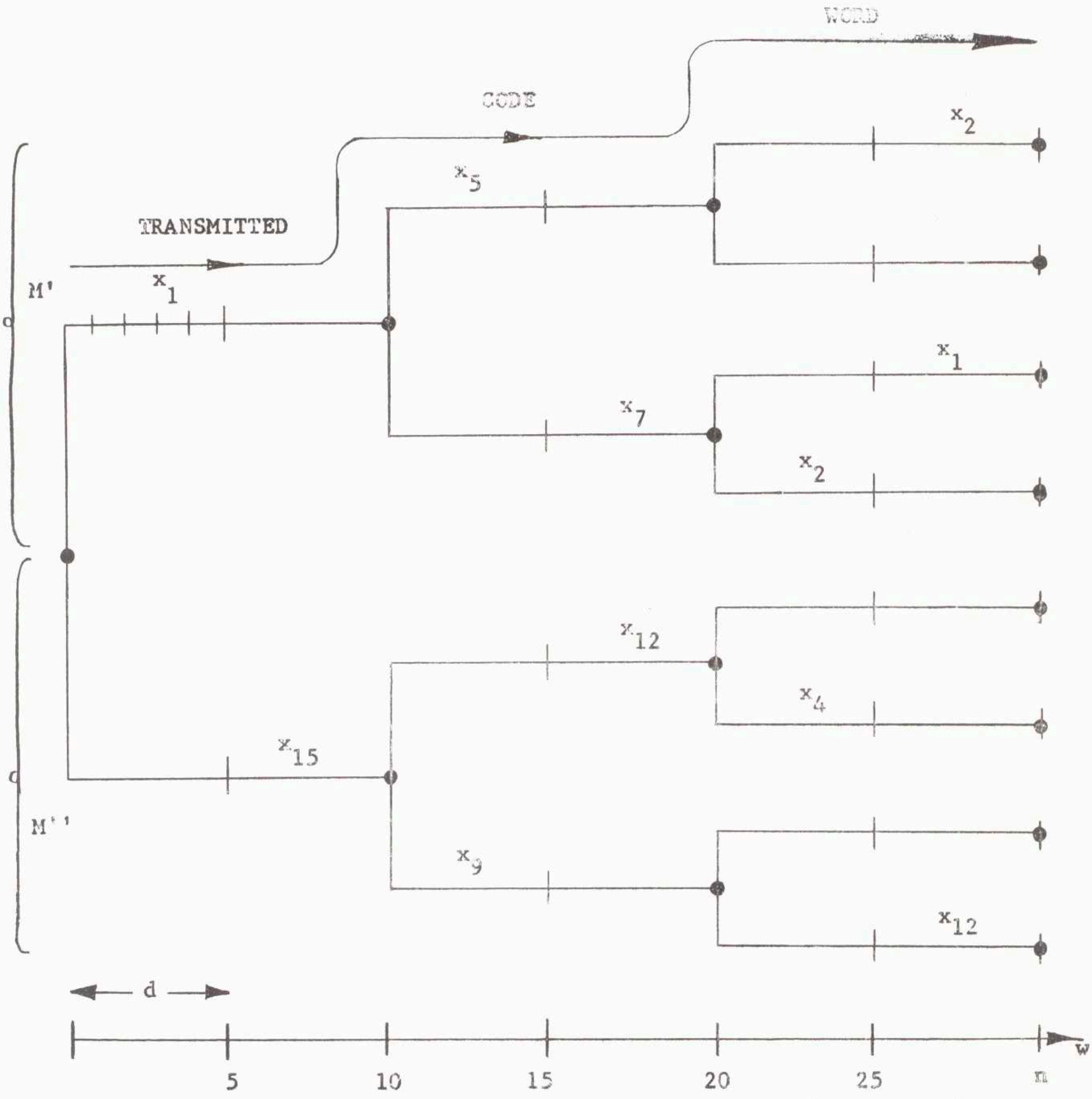
$$M(w) \leqslant A_1 e^{wRd} ; \qquad 1 \leqslant w \leqslant m \tag{2-143}$$

where A_1 is some small constant $\geqslant 1$.

A code structure that is consistent with Eq. 2-143 is a tree, as shown in Figure 7. There is one branch point for each information digit. Each information digit consists of "a" channel input symbols. All the input symbols are randomly selected from a d-dimensional input space X_ℓ which consists of ℓ vectors. From each branch point there diverges b branches. The constraint length is n samples and thus equal to m input symbols or i information digits where $i = \frac{m}{a}$.

The upper bound on the probability of error that was used in the previous sections and which is discussed in Appendix A, is based on random block codes, not on tree codes, to which we wish to apply them. The important feature of random block codes, as far as the average probability of error is concerned, is the fact that the M code words are statistically independent of each other, and that there is a choice of input symbol a priori probabilities which maximize the exponent in the upper bound expression.

In the case of a tree structure we shall seek in decoding to make a decision only about the first information digit. This digit divides the



$b = 2$; $n = 30$; $d = 5$; $a = 2$; $m = 6$; $\ell = 15$; $R = \frac{1}{30} \ln 8 = \frac{1}{10} \ln 2$

FIGURE 7 CONVOLUTIONAL TREE CODE

entire tree set M into two subsets: M' is the subset of all messages which start with the same digit as that of the transmitted message, and M'' is the subset of messages other than those of M'. It is clear that the messages in the set M' cannot be made to be statistically independent. However, each member of the incorrect subset M'' can be made to be statistically independent of the transmitted sequence which is a member of M'.

Reiffen (5) has described a way of generating such randomly selected tree codes where the messages of the incorrect subset M'' are statistically independent of the messages in the correct subset M'.

Thus, the probability of incorrect detection of the first information digit in a tree code is bounded by the same expressions as the probability of incorrect detection of a message encoded by a random block code.

Furthermore, these trees can be made infinite so that the above statistical characteristics are common to all information digits, which are supposed to be emitted from the information source in a continuous stream and a constant rate. These codes can be generated by a shift register (5), and the encoding complexity per information digit is proportional to m, where $m = \frac{n}{d}$ is the number of channel input symbols per constraint length.

Clearly, the encoding complexity is also a monotonically increasing function of ℓ, (the number of symbols in the input space X). Thus, let M_e be an encoding complexity measure, defined as

$$M_e = \ell m = n \frac{\ell}{d} \tag{2-144}$$

The decoding complexity for the two decoding schemes which are discussed in the next chapter, is shown to be proportional to m^{α}, $1 \leq \alpha \leq 2$, for all rates below some computational cut-off rate R_{comp}.

Clearly, the decoding complexity must by a monotonically increasing function of ℓ. Thus, let M_d be the decoding complexity measure defined as

$$M_d = \ell m^2 = n^2 \frac{\ell}{d^2} \tag{2-145}$$

In the next section we shall discuss the problem of minimizing M_e and M_d with respect to ℓ and d, for a given rate R, a given constraint length n and a suitably defined loss in the value of the exponent of the probability of error.

2.6 Optimization of ℓ and d

This discussion will be limited to rates below R_{crit}, and to cases where the power constraint of Statement 2-1 is valid. Let L be a loss factor, defined as

$$L = \frac{E(0) - \overline{E_{\ell,\, d}(0)}}{E(0)} \qquad (2\text{-}146)$$

Now, for rates below R_{crit} we have by Eq. A-70

$$E_{\ell,\, d}(R) = E_{\ell,\, d}(0) - R \, ; \qquad R \leqslant R_{crit}$$

Thus

$$\overline{E_{\ell,\, d}(R)} = E(0)\,(1 - L) - R \, ; \qquad R \leqslant R_{crit} \qquad (2\text{-}147)$$

Therefore specifying an acceptable $E_{\ell,\, d}(R)$ for any rate $R \leqslant R_{crit}$, corresponds to the specification of a proper loss factor L.

We proceed to discuss the minimization of M_ℓ and M_d with respect to ℓ and d, for a given acceptable loss factor L, and a given constraint-length n.

For $dE(0) << 1$ we have by Eq. 2-73

$$\frac{\overline{E_{\ell,\, d}(0)}}{E(0)} \geqslant (1 - \frac{1}{\ell}) \, ; \qquad dE(0) << 1 \, , \qquad R \leqslant R_{crit} \qquad (2\text{-}148)$$

Inserting Eq. 2-146 into Eq. 2-148 yields

$$\ell \leqslant \frac{1}{L} \, ; \qquad E(0)d << 1 \, ; \qquad R \leqslant R_{crit}$$

Thus, by Eqs. 2-144 and 2-145 we have

$$M_e \leqslant \frac{n}{Ld} ; \qquad E(0)d \ll 1 \qquad (2\text{-}149a)$$

$$M_d \leqslant \frac{n^2}{Ld^2} ; \qquad E(0)d \ll 1 \qquad (2\text{-}149b)$$

The lower bounds to M_e and M_d decrease when d increases.

Thus, d should be chosen as large as possible and the value of d that minimizes M_e and M_d is therefore outside the region of d for which $E(0)d \ll 1$. The choice of ℓ should be such as to yield the desired loss factor L. Also by Eq. 2-74

$$\overline{E_{\ell,\,d}(0)} \geqslant E(0) - \frac{1}{d} \ln \left(\frac{e^{E(0)d}}{\ell} + 1 \right) ; \quad R \leqslant R_{crit} \qquad (2\text{-}150)$$

This bound is effective whenever $\ell \gg 1$. This corresponds to the region $E(0)d \gg 1$. (In order to get a reasonably small L, ℓ should be much larger than unity if $E(0)d \gg 1$.) Inserting Eq. 2-150 into Eq. 2-146 yields

$$L = \frac{1}{dE(0)} \ln \left(\frac{e^{E(0)d}}{\ell} + 1 \right)$$

Thus

$$\ell = \frac{e^{E(0)d}}{e^{LE(0)d} - 1} \qquad (2\text{-}151)$$

Inserting Eq. 2-151 into Eqs. 2-144 and 2-145 yields

$$M_e \leqslant \frac{n}{d} \frac{e^{E(0)d}}{e^{LE(0)d} - 1} \qquad (2\text{-}152a)$$

$$M_d \leqslant \frac{n^2}{d^2} \frac{e^{E(0)d}}{e^{LE(0)d} - 1} \qquad (2\text{-}152b)$$

From Eq. 2-152a we have that the bound to M_e , has an extrimum point at

$$\frac{E(0)d - 1}{E(0)d(1 - L) - 1} = e^{LE(0)d} \tag{2-153}$$

if a solution exists. Thus, for $E(0)d >> 1$ this corresponds to

$$\frac{1}{1-L} = e^{LE(0)d}$$

or

$$dE(0) = \frac{1}{L} \ln \frac{1}{1-L}$$

Now, for reasonably small variables of the loss factor L we have

$$dE(0) \cong 1 \tag{2-154}$$

This point is outside the region of d for which $dE(0) >> 1$.

From Eq. 2-152b we have that the bound to M_d has a extrimum point at

$$\frac{E(0)d - 2}{E(0)d(1 - L) - 2} = e^{LE(0)d} \tag{2-155}$$

if a solution exists. This corresponds, for $E(0)d >> 1$, to

$$\frac{1}{1-L} = e^{LE(0)d}$$

or

$$dE(0) = \frac{1}{L} \ln \frac{1}{1-L}$$

For reasonably small variables of the loss factor L we therefore have

$$dE(0) \cong 1 \tag{2-156}$$

This point is outside the region of d for which $dE(0) >> 1$.

We may conclude therefore that the lower bounds to M_e and M_d are monotonically decreasing functions of d in the region $dE(0) << 1$ and are monotonically increasing functions of d in the region $dE(0) >> 1$.

<u>Both M_e and M_d are therefore minimized if:</u>

$$E(0)d \cong 1 \; ; \quad E(0) \leq 1, \quad R \leq R_{crit} \qquad (2\text{-}167a)$$

and since $d \geq 1$, if

$$d = 1 \; ; \qquad E(0) \gtrsim 1, \quad R \leq R_{crit} \qquad (2\text{-}167b)$$

The number ℓ is chosen to yield the desired loss factor L.

CHAPTER III

DECODING SCHEMES FOR CONVOLUTIONAL CODES

3.1 Introduction

Sequential decoding implies that we decode one information digit at a time.

The symbol s_i is to be decoded immediately after s_{i-1}. Thus the receiver has available the decoded set $(\ldots, s_{-1}, s_0)$ when it is about to decode s_1. We shall assume that these symbols have been decoded without an error. This assumption, although crucial to the decoding procedure, is not so restrictive as it may appear. This will be discussed later in this chapter.

We therefore restrict our attention to those s_t consistent with the previously decoded symbols.

3.2 Sequential Decoding (After Wozencraft and Reiffen)

Let u_w be the sequence that consists of the first w input symbols of the transmitted sequence, that diverges from the last information digit to be detected. Let u_w' be a member of the incorrect set M'' (see Section 2.5). u_w' therefore starts with an information digit other than that of the sequence u_w. Let v_w be the sequence that consists of the w output symbols that correspond to the transmitted segment u_w. Let

$$D_w(u, v) = \ln \frac{p(v_w)}{p(v_w | u_w)} \tag{3-1}$$

which we call the distance between u_w and v_w. Where

$$p(v_w) = \prod_{i=1}^{w} p(y_i) \tag{3-2}$$

$$p(v_w | u_w) = \prod_{i=1}^{w} p(y_i | x_i) \tag{3-3}$$

Let us define a constant $D_w^{\,j}$ given by

$$P(D_w(u, v) \geqslant D_w^{\,j}) \leqslant e^{-k_j} \tag{3-4a}$$

where k_j is some arbitrary positive constant which we call "probability criterion" and is a member of an ordered set

$$K = \left\{ k: \;\; k_j = k_{j-1} + \triangle \;; \quad k_{j_{max}} = E(R)n \right\} \tag{3-4b}$$

where $\triangle \geqslant 0$ is a constant.

Let us now consider the sequential decoding scheme in accordance with the following rules:

1. The decoding computer starts out to generate sequentially the entire tree set M (Section 2.5). As the computer proceeds, it discards any sequence u_w' of length w symbols ($1 \leqslant w \leqslant m$) for which the distance $D_w(u', v) \geqslant D_w^{\,1}$. ($D_w^{\,1}$ corresponds to the smallest "probability criterion" k_1).

2. As soon as the computer discovers any sequence is M that is retained through length m, it prints out the corresponding first information digit.

3. If the complete set M is discarded, the computer adopts the next larger criterion k_2, and its corresponding distance $D_w^{\,2}$; ($1 \leqslant w \leqslant m$).

4. The computer continues this procedure until <u>some</u> sequence in M is retained through length m. It then prints the corresponding first information digit.

When these rules are adopted, the decoder never uses a criterion K_j unless the correct subset M' (and, hence, the correct sequence u_w) is discarded for k_{j-1}. The probability that u_w is discarded depends on the channel noise only. By averaging both over all noise sequences and over the ensemble of all tree codes, we can bound the average number of computations, $\overline{N}$, required to eliminate the incorrect subset M''.

3.3 Determination of a Lower Bound to R_{comp} of Wozencraft-Reiffen Decoding Scheme (5), (6)

Let N(w) be the number of computations required to extend the search from w to w + 1. Using bars to denote averages

$$\overline{N} = \sum_{w} \overline{N(w)} \tag{3-5}$$

$\overline{N(w)}$ may be upper bounded in the following way: The number of incorrect messages of length w, M(w), is given by Eq. 2-143.

$$M(w) \leqslant A_1 e^{dRw} \tag{2-143}$$

The probability that an incorrect message is retained through length w + 1 when the criterion k_j is used is given by

$$\Pr[D_w(u', v) \leqslant D_w^{\,j} \,|\, j] \tag{3-6}$$

The criterion k_j is used whenever all sequences are discarded at some length λw ($\frac{1}{w} \leqslant \lambda \leqslant \frac{m}{w}$) with the criterion k_{j-1}.

Thus the probability Pr(j) of such an event is upper bounded by the probability of the correct sequence u being discarded at some length λw. Therefore

$$p(j) \leqslant \sum_{\lambda} \Pr(D_{\lambda w}(u, v) \geqslant D_{\lambda w}^{\,j-1}) \tag{3-7}$$

Thus, by Eqs. 2-143, 3-6 and 3-7

$$\overline{N(w)} \leqslant A_1 e^{dwR} \sum_{j} \Pr(D_w(u', v) \leqslant D_w^{\,j} | j) \;\; \Pr(j) \tag{3-8}$$

Inserting Ineq. 3-7 into Ineq. 3-8 yields

$$\overline{N(w)} \leqslant A_1 e^{dwR} \sum_{j,\lambda} \Pr(D_w(u', v) \leqslant D_w^{\,j} \; ; \; D_{\lambda w}(u, v) \geqslant D_{\lambda w}^{\,j-1}) \tag{3-9}$$

Inserting Eq. 3-9 into 3-5 yields

$$\overline{N} \leqslant \sum_{w\lambda, j} A_1 e^{wdR} \Pr[D_w(u', v) \leqslant D_w{}^j ; \quad D_{\lambda w}(u, v) \geqslant D_{\lambda w}{}^{j-1}] \qquad (3\text{-}10)$$

We would like to obtain an upper bound on

$$\Pr[D_w(u', v) \leqslant D_w{}^j ; \quad D_{\lambda w}(u, v) \geqslant D_{\lambda w}{}^{j-1}]$$

of the form

$$\Pr[D_w(u', v) \leqslant D_w{}^j ; \quad D_{\lambda w}(u, v) \geqslant D_{\lambda w}{}^{j-1}] \leqslant Be^{-R^*dw} \qquad (3\text{-}11)$$

where B is a constant that is independent of w and λ, and R^* is any positive number such that 3-11 is true. Inserting Eq. 3-11 into Eq. 3-10 yields

$$\overline{N} \leqslant \sum_{w, j, \lambda} Ke^{(R-R^*)wd} \qquad (3\text{-}12)$$

where $k = BA_1$.

The minimum value of R^*, over all w, λ and j is called "R_{comp}." Thus

$$R_{comp} = \min_{\lambda, w} \left\{ R^* \right\} \qquad (3\text{-}13)$$

Inserting Eq. 3-13 into 3-12 yields

$$\overline{N} \leqslant \sum_{w, j, \lambda} Ke^{-(R_{comp} - R)wd} \qquad (3\text{-}14)$$

For $R < R_{comp}$, the summation on w is a geometric series which may be upper bounded by a quantity independent of the constraint length m. The summation on λ contains m identical terms. The summation on j will contain a number of terms proportional to m. This follows from making the largest criterion used, $k_{j_{max}}$, equal to $E(R)n = E(R)md$. Thus for rates $R < R_{comp}$, $\overline{N}$ may be upper bounded by a quantity proportional to m^2. Reiffen (6) obtained an upper bound to R_{comp}

$$R_{comp} \leqslant E(0) \tag{3-15}$$

It has been shown (5) that $R_{comp} = E(0)$ whenever the channel is symmetricl.

We proceed now to evaluate a lower bound on R_{comp}.

From Eq. 3-1, 3-2 and 3-3 we have

$$D_{\lambda w}(u, v) = \sum_{i=1}^{\lambda w} d(x, y) = \sum_{i=1}^{\lambda w} \ln \frac{p(y_i)}{p(y_i|x_i)} \tag{3-16a}$$

$$D_w(u', v) = \sum_{i=1}^{w} d(x', y) = \sum_{i=1}^{w} \ln \frac{p(y_i)}{p(y_i|x_i')} \tag{3-16b}$$

Thus, by the use of Chernoff bounds (Appendix A, Section 3)

$$\Pr(D_{\lambda w}(u, v) \geqslant D_{\lambda w}{}^{j}) \leqslant e^{\lambda w(\gamma(s) - s\gamma'(s))} \tag{3-17a}$$

where, by Eqs. A-27, A-29 and 3-16a

$$\gamma(s) = \ln \sum_{X_\ell} \int_Y P(x)\, p(y|x)\, e^{sd(x, y)} dy$$

$$= \ln \sum_{X_\ell} \int_Y P(x)\, p(y|x)^{1-s} p(y)^s \, dy \; ; \qquad s \geqslant 0 \tag{3-17b}$$

and

$$\gamma'(s) = \frac{D_{\lambda w}{}^{j}}{\lambda w} \tag{3-17c}$$

Thus, by Ineqs. 3-4 and 3-17

$$\Pr[D_{\lambda w}(u, v) \geqslant D_{\lambda w}{}^{j}] \leqslant e^{\lambda w(\gamma(s) - s\gamma'(s))} = e^{-kj} \quad \text{for all } \lambda \tag{3-18a}$$

and

$$D_{\lambda w}{}^{j} = \lambda w \gamma'(s) \tag{3-18b}$$

where s is determined as the solution of the equation

$$\frac{k_j}{\lambda w} = s\gamma'(s) - \gamma(s) \tag{3-18c}$$

In the same way

$$\Pr[D_w(u', v) \leq D_w{}^{j}] \leq e^{w(\mu(t) - t\mu'(t))} \tag{3-19a}$$

where, by Section A.3 and Eq. 3-16 b

$$\mu(t) = \ln \sum_{X'_\ell} \int_Y P(x') \, p(y) \, e^{td(x', y)} dy$$

$$= \ln \sum_{X'_\ell} \int_Y P(x') \, p(y)^{1+t} p(y \ x')^{-t} dy dx' \quad ; \quad t \leq 0 \tag{3-19b}$$

and

$$\mu'(t) = \frac{D_{\lambda w}{}^{j}}{\lambda w} \tag{3-19c}$$

Now, returning to Ineq. 3-11

$$\Pr[D_w(u', v) \leq D_w{}^{j} \; ; \; D_{\lambda w}(y, v) \geq D_{\lambda w}{}^{j-1}] \leq \Pr[D_w(u', v) \leq D_w{}^{j}] \tag{3-20a}$$

Also

$$\Pr[D_w(u', v) \leq D_w{}^{j} \; ; D_{\lambda w}(u, v) \geq D_{\lambda w}{}^{j-1}] \leq \Pr[D_{\lambda w}(u, v) \geq D_{\lambda w}{}^{j-1}] \tag{3-20b}$$

Thus, by Ineqs. 3-20, 3-18 and 3-19

$$\Pr[D_w(u', v) \leqslant D_w{}^j \; ; \; D_{\lambda w}(u, v) \geqslant D_{\lambda w}{}^{j-1}]$$

$$\leqslant \min \left\{ \Pr[D_w(u', v) \leqslant D_w{}^j] \; ; \; \Pr[D_{\lambda w}(u', v) \geqslant D_{\lambda w}{}^{j-1}] \right.$$

$$\min \left\{ e^{\mu(t) - t\mu'(t)} \; ; \; e^{-k_{j-1}} \right\} \tag{3-21}$$

Now, by Eq. 3-4b) $k_j = k_{j-1} + \triangle$; $\triangle \geqslant 0$. Thus by Ineq. (3-18a)

$$e^{-k_{j-1}} = e^{\triangle} e^{-k_j} = e^{\triangle} e^{(\gamma(s) - s\gamma'(s))w} \tag{3-22a}$$

where

$$\gamma'(s) = \frac{D_w{}^j}{w} \tag{3-22b}$$

and

$$\gamma(s) - s\gamma'(s) = -\frac{k^j}{w} \tag{3-22c}$$

Therefore, inserting Eq. 3-22 into Ineq. 3-21 yields

$$\Pr[D_w(u', v) \leqslant D_w{}^j \; ; \; D_{\lambda w}(u, v) \geqslant D_{\lambda w}{}^{j-1}]$$

$$\leqslant e^{\triangle} \min \left\{ e^{w(\gamma(s) - s\gamma'(s))} ; \; e^{w(\mu(t) - t\mu'(t))} \right. \tag{3-23a}$$

Thus, if we choose

$$R^* = \max \left\{ -\gamma(s) + s\gamma'(s) \; ; \; -\mu(t) + t\mu'(t) \right\}$$

or

$$R^* \geqslant \frac{1}{2} \left\{ -\gamma(s) + s\gamma'(s) - \mu(t) + t\mu'(t) \right\} \tag{3-23b}$$

then Eq. 3-11 is valid.

Inserting Ineq. 3-23b into Eq. 3-13 yields

$$R_{comp} \geqslant \min \frac{1}{2} \left\{ - \gamma(s) + s\gamma'(s) - \mu(t) + t\mu'(t) \right\} \tag{3-24}$$

Now, by Eqs. 3-17 and 3-19

$$\gamma'(s) = \frac{\sum\limits_{X_\ell} \int\limits_{Y} P(x)\; p(y|x)^{1-s} p(y)^{s} \;\ell n \frac{p(y)}{p(y|x)}\, dy}{\sum\limits_{X_\ell} \int\limits_{Y} P(x)\; p(y|x)^{1-s} p(y)^{s}\, dy} \tag{3-25a}$$

$$\mu'(t) = \frac{\sum\limits_{X_\ell} \int\limits_{Y} P(x)\; p(y)^{1+t} p(y|x)^{-t} \;\ell n \frac{p(y)}{p(y|x)}\, dy}{\sum\limits_{X_\ell} \int\limits_{Y} P(x)\; p(y)^{1+t} p(y|x)^{-t}\, dy} \tag{3-25b}$$

If we let $t = s - 1$ we have

$$\mu'(t) = \gamma'(s) \; ; \; \mu(t) = \gamma(s) \tag{3-26}$$

Hence

$$R_{comp} \geqslant \min \frac{1}{2} \left[(2s-1)\gamma'(s) - 2 \quad (s) \right] \tag{3-27}$$

The minimum occurs at that s, for which

$$\left[(2s - 1)\, \gamma'(s) - 2\gamma(s) \right]' = 0$$

$$\text{which corresponds to } s = \frac{1}{2} \tag{3-28}$$

Also, $[(1 - 2s)\gamma'(s) - 2\gamma(s)]'' = 2\gamma''(\frac{1}{2}) \geqslant 0$, since $\gamma''(\frac{1}{2})$ is the variance (see Ref. (2)) of a random variable. Thus, $s = \frac{1}{2}$ is indeed a minimum point.

Inserting Eq. 3-28 into Ineq. 3-27 yields

$$R_{comp} \geqslant - \gamma\left(\frac{1}{2}\right) \tag{3-29}$$

Now, by Eq. 3-17b

$$\gamma\left(\frac{1}{2}\right) = \ln \int_Y \sum_{X_\ell} P(x)\ p(y/x)^{\frac{1}{2}}\ p(y)^{\frac{1}{2}} dy \tag{3-30a}$$

where

$$p(y) = \sum_X P(x)\ p(y\ x) \tag{3-30b}$$

Therefore

$$2\,\gamma\left(\frac{1}{2}\right) = \ln \left\{ \int_Y \sum_{X_\ell} p(x)\ p(y|x)^{\frac{1}{2}}\ p(y)^{\frac{1}{2}} dy \right\}^2$$

$$= \ln \left\{ \int_Y g(y)^{\frac{1}{2}} p(y)^{\frac{1}{2}} dy \right\}^2 \tag{3-31a}$$

where

$$g(y) = \left\{ \sum_{X_\ell} p(x)\ p(y/x)^{\frac{1}{2}} \right\}^2 \tag{3-31b}$$

By the Schwarz inequality

$$\left\{ \int_Y g(y)^{\frac{1}{2}}\ p(y)^{\frac{1}{2}} dy \right\}^2 \leqslant \int_Y g(y) dy \int_Y p(y) dy$$

$$= \int_Y g(y) dy \tag{3-32}$$

Inserting Ineq. 3-32 into Eq. 3-31 yields

$$2\gamma\left(\frac{1}{2}\right) \leqslant \ln \int_{Y} \left[\sum_{X_\ell} p(x)\, p(y\, x)^{\frac{1}{2}}\right]^2 dy \tag{3-33}$$

Inserting Ineq. 2-33 into Ineq. 3-29 yields

$$R_{comp} \geqslant -\frac{1}{2} \ln \int_{Y} \left[\sum_{X_\ell} p(x)\, p(y\, x)^{\frac{1}{2}}\right]^2 dy \tag{3-34}$$

Now, by Eqs. A-69 and A-71 in Appendix A,

$$-\ln \int_{Y} \left[\sum_{X_\ell} p(x)\, p(y|x)^{\frac{1}{2}}\right]^2 dy$$

is equal to the zero-rate exponent $E_{\ell,\, d}(0)$, for the given channel. By a proper selection of p(x) and the number of input symbols, $E_{\ell,\, d}(0)$ can be made arbitrarily close to the optimum zero-rate exponent E(0) (see Chapter 2). Thus

$$R_{comp} \geqslant \frac{1}{2} E_{\ell,\, d}(0) \tag{3-35}$$

and for semi-optimum input spaces

$$R_{comp} \geqslant \frac{1}{2} E(0) \tag{3-36}$$

In this section, we have been able to meaningfully bound the average number of computations necessary to discard the incorrect subset. The harder problem of bounding the computation on the correct subset has not been discussed. A modification of the decoding procedure above, adapted from a suggestion by Gallager [Ref (5), page 29] for binary symmetric channels, yields a bound on the total number of computations for any symmetric channel. However, no such bound for asymmetric channels have been yet discovered.

3.4 Upper Bound to the Probability of Error for the Sequential Decoding Scheme

Let us suppose we (conservatively) count as a decoding error the occurrence of either one of the following events.

1. The transmitted sequence u and the received sequence v are such that they fail to meet the largest criterion $k_{j_{max}}$. The probability of this event, over the ensemble, is less than $me^{-k_{j_{max}}}$.
2. Any element u' of the incorrect subset M'' together with the received v satisfies some $k_j < k_{j_{max}}$, when the j^{th} criterion is used.

An element of M'' picked at random, together with the received v_n has a probability of satisfying some k_j equal to

$$\sum_j Pr[D_m(u', v) \leq D_m^{\ j} \ ; \ k_j \text{ is used}]$$

Since the probability of a union of events is upper bounded by the sum of the probabilities of the individual events, the probability that any element of M'' together with the received signal v satisfies k_j is less than

$$e^{nR} \sum_j Pr[D_m(u', v) \leq D_m^{\ j} \ ; \ k_j \text{ is used}]$$

The two events stated above are not in general independent. However, the probability of their union is upper bounded by the sum of their probabilities. Thus the probability of error p_e may be bounded by

$$p_e \leq me^{-k_{j_{max}}} + e^{mdR} \sum_j Pr[D_m(u', v) \geq D_m^{\ j} \ ; \ k_j \text{ is used}] \qquad (3\text{-}37)$$

It has been shown in Ref. (5) that for channels which are symmetric at their output, the probability of error is bounded by

$p_e \leq me^{-E_{\ell, d}(R)\, n}$ where $E_{\ell, d}(R)$ is the optimum exponent for the given channel and the given input space. (See Appendix A.) We now proceed to evaluate Ineq. 3-37 for the general asymmetric memoryless channel. The event that k_j is used is included in the event that u' together with v will not satisfy the criterion k_{j-1}, or

$$D_m(u', v) \geqslant D_m^{j-1}$$

Thus

$$Pr[D_m(u', v) \leqslant D_m^{j} \; ; \; k_j \text{ is used}]$$

$$\leqslant Pr[D_m^{j} \geqslant D_m(u', v) \geqslant D_m^{j-1}] \qquad (3\text{-}38)$$

Inserting Ineq. 3-38 into Ineq. 3-37 yields

$$P_e \leqslant me^{-k_{j_{max}}} + e^{dmR} Pr[D_m(u', v) \leqslant D_m^{j_{max}}] \qquad (3\text{-}39)$$

Now, by Ineq. 3-4, D_m^j is chosen so as to make

$$Pr[D_m(u, v) \geqslant D_m^{j_{max}}] \leqslant e^{-k_{j_{max}}}$$

Also, by Ineq. 3-17

$$Pr[D_m(u, v) \geqslant D_m^{j_{max}}] \leqslant e^{m[\gamma(s) - s\gamma'(s)]} \; ; \qquad s \geqslant 0$$

where

$$\gamma'(s) = \frac{D_m^{j_{max}}}{m}$$

Thus, we let $-k_j = m[\gamma(s) - \gamma'(s)]$ and therefore

$$e^{-k_j} = e^{m[\gamma(s) - s\gamma'(s)]} \; ; \qquad s \geqslant 0 \qquad (3\text{-}40a)$$

where

$$\gamma'(s) = \frac{D_m^{j_{max}}}{m} \qquad (3\text{-}40b)$$

From Ineq. 3-19 we have

$$Pr[D_m(u', v) \leqslant D_m^{j_{max}}] \leqslant e^{m(\mu(t)-t\mu'(t))} \quad ; \qquad t \leqslant 0 \qquad (3\text{-}41a)$$

where

$$\mu'(t) = \frac{D_m^{j_{max}}}{m} \qquad (3\text{-}41b)$$

Inserting Ineq. 3-40 and 3-41 into Ineq. 3-39 yields

$$P_e \leqslant m[e^{m(\gamma(s)-s\gamma'(s))} + e^{m(dR+\mu(t)-t\mu'(t))}] \qquad (3\text{-}42a)$$

where

$$\mu'(t) = \gamma'(s) = D_m^{j_{max}} \qquad (3\text{-}42b)$$

By Eq. 3-26 we have that Eq. 3-42b is satisfied if we let $t = s - 1$. Thus by Eq. 3-26

$$P_e \leqslant m[e^{m(\gamma(s)-s\gamma'(s))} + e^{m(dR+\gamma(s)-(s-1)\gamma'(s))}] \qquad (3\text{-}43)$$

Making

$$\gamma(s) - s\gamma'(s) = dR + \gamma(s) - (s-1)\gamma'(s)$$

we get

$$P_e \leqslant 2me^{m(\gamma(s)-s\gamma'(s))} = 2me^{-nE_{sq}(R)} \qquad (3\text{-}44a)$$

where

$$- \frac{1}{d} \gamma'(s) = R \tag{3-44b}$$

and

$$E_{sq}(R) = \frac{1}{d} \left(\gamma(s) - s\gamma'(s) \right) \tag{3-44c}$$

The rate that makes $E_{sq}(R) = 0$ is the one that corresponds to $s = 0$, since $\left\{ \gamma(s) - s\gamma'(s) \right\} \Big|_{s=0} = 0$. By Eq. 3-44b

$$[R]_{s=0} = - \frac{1}{d} \gamma'(s) \Big|_{s=0}$$

Also, by Eq. 3-25a

$$- \frac{1}{d} \gamma'(s) \Big|_{s=0} = - \frac{1}{d} \sum_{X_\ell} \int_Y P(x)\ p(y|x)\ \ln \frac{p(y)}{p(y|x)}\ dy$$

Thus

$$E_{sq}(R) \geqslant 0 \ ; \qquad R \leqslant [R]_{s=0} \tag{3-45a}$$

where

$$[R]_{s=0} = \frac{1}{d} \sum_{X_\ell} \int_Y P(x)\ p(y|x) \ln \frac{p(y|x)}{p(y)}\ dy \tag{3-45b}$$

Comparing Eq. 3-45 with Eq. A-57 of Appendix A yields that $E_{sq}(R) \geqslant 0$ for the same region of rates as $E_{\ell,\, d}(R)$. Thus, if the input space X is semi-optimal, one can get an arbitrarily small probability of error for rates below the channel capacity C.

The zero-rate exponent $E_{sq}(0)$ is given by

$$E_{sq}(0) = -\gamma(s) + s\gamma'(s) = -\gamma(s) + (s-1)\gamma'(s) \qquad (3\text{-}46a)$$

where s is the solution of

$$\gamma'(s) = 0 \qquad (3\text{-}46b)$$

Thus

$$E_{sq}(0) \geqslant \min \frac{1}{2} \left\{ -2\gamma(s) + (2s-1)\gamma'(s) \right\} \qquad (3\text{-}47)$$

Following Eqs. 3-27 through 3-36 and substituting R_{comp} by $E_{sq}(0)$ we get

$$E_{sq}(0) \geqslant \frac{1}{2} E_{\ell,\, d}(0) \qquad (3\text{-}48)$$

and for semi-optimum input spaces

$$E_{sq}(0) \geqslant \frac{1}{2} E(0) \qquad (3\text{-}49)$$

3.5 New Successive Decoding Scheme for Memoryless Channels

In this section a new sequential decoding scheme for random convolutional codes is described. The average number of computations does not grow exponentially with n; for rates below some R^*_{comp}, the average number of computations is upper bounded by a quantity proportional to

$$m^{(1 + R/R^*_{comp})} \leqslant m^2$$

The computational cut-off rate R^*_{comp} of the new systematic decoding scheme is equal to the lower bound on R_{comp} for sequential decoding with asymmetric channels (see Section 3.3).

However, in the case of sequential decoding, R_{comp} is valid only for the incorrect subset of code words: the existence of R_{comp} for the correct subset has not yet been proven for asymmetric channels. The successive decoding scheme, which is different from other effective decoding schemes such as sequential decoding and low-density parity-check codes [Ref. (9)] yields a bound on the total average number of computations.

When this decoding scheme is averaged over a suitably defined ensemble of code words it has an average probability of error with an upper bound whose logarithm is $-nE_s(R)$. $E_s(R) \geqslant 0$ for rates below channel capacity if a semi-optimum input space is used.

A convolutional tree code is shown in Figure 7 and is discussed in Section 2.5.

Let us now consider the decoding procedure that consists of the following successive operations.

<u>Step 1</u>: Consider the set of b^{k_1} paths of k_1 information digits that diverge from the first node (branch point). Each path consists, therefore, of k_1a input symbols. The a posteriori probability of each one of the b^{k_1} paths, given the corresponding segment of v, is computed. The first branch of the path of length k_1a which, given v, yields the largest a posteriori probability is tentatively chosen to represent the corresponding first transmitted digit (see Figure 8).

Let us next consider the set of b^{k_1} paths of length k_1a symbols that diverges from the tentative decision of the previous step. The a posteriori probability of each one of these b^{k_1} paths, given the corresponding segment of the sequence v, is computed. The first branch of the link of length k_1a which, given v, yields the largest a posteriori probability is tentatively chosen to represent the second transmitted digit.

This procedure is continued until $i = \frac{m}{a}$ information digits have been tentatively detected.

The distance $D(u_1, v) = \ln \frac{p(v)}{p(v|u_1)}$ is then computed for the complete word u_1 of length m input symbols thus obtained.

If $D(u_1, v)$ is smaller than some preset threshold D_0, a firm decision is made that the first digit of u_1 represents the first encoded information digit.

If however, $D(u_1, v) \geqslant D_0$, the computation procedure is to proceed to Step 2.

<u>Step 2</u>: The decoding procedure of Step 2 is identical with that of Step 1, with the exception that the length k_1 (information digits for k_1a channel symbols) is replaced by

$$k_2 = k_1 + \triangle ; \qquad \triangle \text{ a positive integer} \tag{3-50}$$

Let u_2 be the detected word of Step 2. If $D(u_2, v) = \ln \frac{p(v)}{p(v|u_2)} < D_0$ a final decision is made, and the detection of the first information digit is completed. If $D(u_2, v) \geqslant D_0$ no termination occurs and the computation procedure is to then go to Step 3, and so on.

In general, for the j^{th} step we have

$$k_j = k_{j-1} + \triangle ; \ \triangle \text{ a positive integer} \tag{3-51}$$

and the detected word is u_j.

Following the detection of the first information digit, the whole procedure is repeated for the next information digit and so forth.

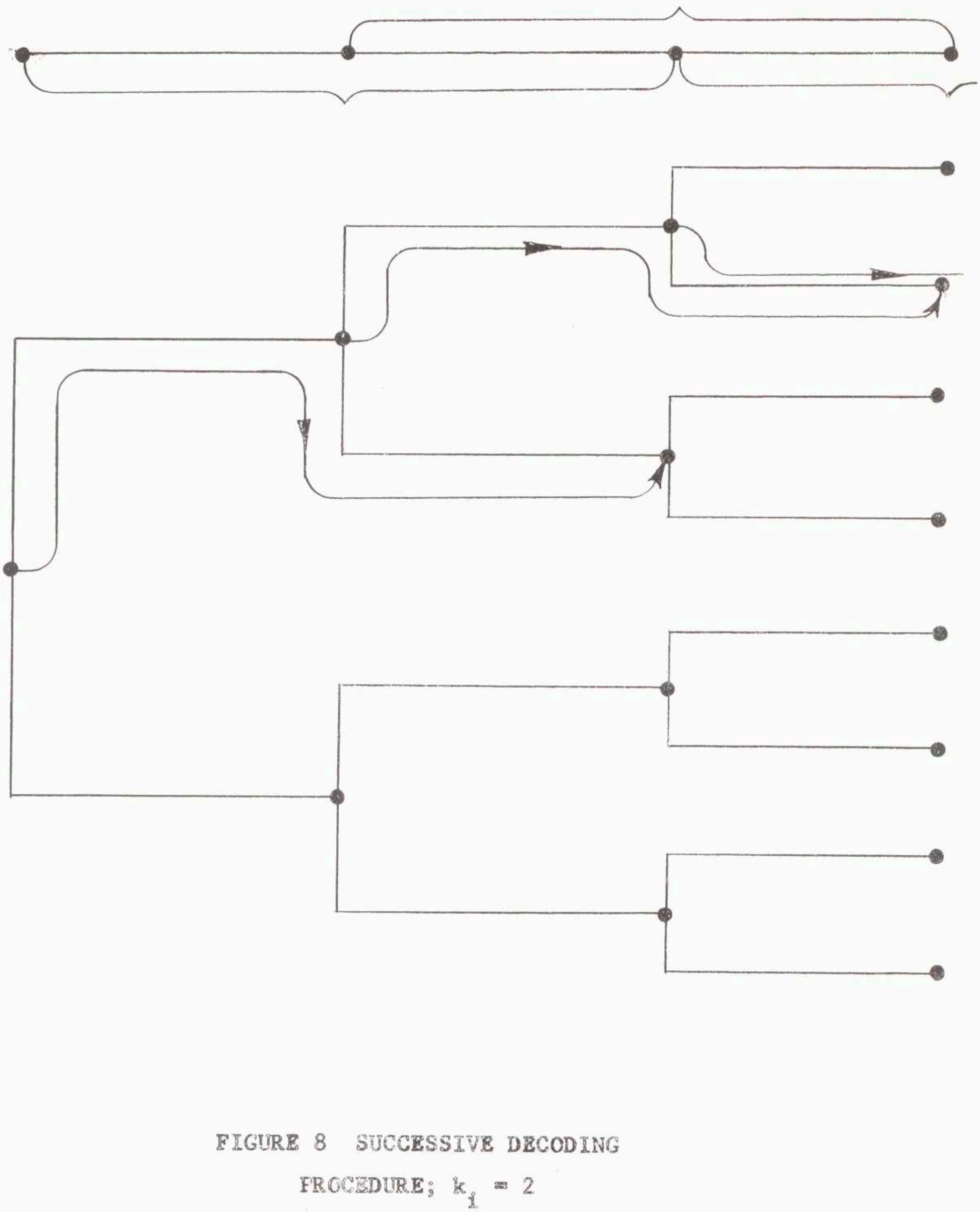


FIGURE 8 SUCCESSIVE DECODING
PROCEDURE; $k_i = 2$

3.6 The Average Number of Computations per Information Digit for the Successive Decoding Scheme

The number of computations that are involved in step j is bounded* by

$$N_j \leqslant mb^{k_j} \tag{3-52}$$

Let C_j be the condition that no termination occurs at step j. Step j will be used only if there are no terminations on all the j-1 previous steps. Thus the probability of step j being used is

$$P(j) = Pr(C_1, C_2, C_3, \ldots, C_{j-1}) \tag{3-53}$$

The average number of computation is given by

$$\overline{N} = N_1P(1) + N_2P(2) + \ldots + N_jP(j) + \ldots + N_{j_{max}}P(j_{max})$$

$$\leqslant \sum_{j=1} N_jP(j) \tag{3-54a}$$

where

$$P(1) \equiv 1 \tag{3-54b}$$

P(j) may be bounded by

$$P(j) = Pr(C_1, C_2, C_3, \ldots, C_{j-1}) \leqslant Pr(C_{j-1}) \tag{3-55}$$

Inserting Ineq. 3-55 and Eq. 3-52 into Eq. 3-54 yields

$$\overline{N} \leqslant N_1 + \sum_{j=2}^{\infty} N_j \, Pr(C_{j-1}) \leq mb^{k_1} + m\sum_{j=1}^{\infty} b^{k_j} \, Pr(C_{j-1}) \tag{3-56}$$

---

*We count as a computation the generation of one branch of a random tree code at the receiver. There are b^{k_j} paths (that consist of k_j branches), which have to be sequentially computed for each one of the i information digits ($i = \frac{m}{a}$). Thus

$$N_j \leqslant ib^{k_j} < iab^{k_j} = mb^{k_j}$$

Now let u_j be the code word detected at step j, and let u be the transmitted code word. Then

$$\Pr(C_j) = \Pr(D(u_j, v) \geqslant D_0)$$

$$= \Pr[D(u_j, v) \geqslant D_0;\ u_j = u] + \Pr[D(u_j, v) \geqslant D_0\ ;\ u_j \neq u]$$

$$= \Pr[D(u, v) \geqslant D_0\ ;\ u_j = u] + \Pr[D(u_j, v) \geqslant D_0\ ;\ u_j \neq u]$$

$$\Pr[D(u, v) \geqslant D_0] + \Pr[u_j \neq u] \tag{3-57}$$

We are free to choose the threshold D_0 so as to satisfy

$$\Pr[D(u, v) \geqslant D_0] \leqslant e^{-\frac{1}{2} E_{\ell, d}^{(0)} n} \tag{3-58}$$

Now, let e_{jr} be the event that the r^{th} information digit of u_j is not the same as the corresponding digit of the transmitted sequence u. Then

$$\Pr(u_j \neq u) = \Pr[\bigcup_{r=1}^{i} \{e_{jr}\}\ ] \tag{3-59}$$

The probability of a union of events is upper bounded by the sum of the probabilities of the individual events. Thus

$$\Pr(u_j \neq u) \leqslant \sum_{r=1}^{i} \Pr(e_{jr}) \tag{3-60}$$

There are $\frac{b-1}{b} b^{k_j}$ paths of length k_j information digits that diverge from the $(r-1)^{th}$ node of u, and which do not include the r^{th} information digit of u. Over the ensemble of random codes these $\frac{b-1}{b} b^{k_j}$ are statistically independent of the corresponding segment of the transmitted sequence u (see Section 2.5). The event e_{jr} occurs whenever the a posteriori probability of one of these $\frac{b-1}{b} b^{k_j}$ paths yields, given v, an a posteriori probability which is larger than that of the corresponding segment of u. Thus,

$Pr(e_{jr})$ is identical with the probability of error for randomly constructed block codes of length $k_j a$ channel input symbols. (All input symbols are members of the d-dimensional input space X_ℓ which consists of ℓ vectors). Bounds to the probability of error for such block codes are given in Appendix A. Thus, by Appendix A

$$Pr(e_{jr}) \leqslant 2e^{-E_{\ell,d}(R)\, k_j ad} \tag{3-61a}$$

where

$$R = \frac{1}{n} \ln b^{\frac{m}{a}} = \frac{1}{n} \ln b^{\frac{n}{ad}} = \frac{1}{ad} \ln b \tag{3-61b}$$

$$E_{\ell,d}(R) = E_{\ell,d}(0) - R \; ; \qquad R \leqslant R_{crit} \tag{3-61c}$$

$$E_{\ell,d}(R) \geqslant E_{\ell,d}(0) - R \; ; \qquad R \geqslant R_{crit} \tag{3-61d}$$

Inserting Ineq. 3-61 into Eq. 3-59 yields

$$Pr(u_j \neq u) \leqslant 2me^{-E_{\ell,d}(R)\, k_j ad} \tag{3-62}$$

Inserting Ineqs. 3-58 and 3-62 into Ineq. 3-57 yields

$$Pr(C_j) \leqslant e^{-\frac{1}{2} E_{\ell,d}(0) n} + 2me^{-E_{\ell,d}(R)\, k_j ad} \tag{3-63}$$

Now, by Ineq. 3-61 we have

$$\frac{1}{2} E_{\ell,d}(0) \leqslant E_{\ell,d}(R) \; ; \qquad R \leqslant \frac{1}{2} E_{\ell,d}(0) \tag{3-64}$$

Also

$$n = md = iad \geqslant k_j ad \tag{3-65}$$

Inserting Ineqs. 3-63 and 3-64 into Ineq. 3-63 yields

$$Pr(C_j) \leqslant 2m\, e^{-\frac{1}{2} \mathbf{E}_{\ell,\, d}(0)\, k_j ad} \quad ; \quad R \leqslant \frac{1}{2} \mathbf{E}_{\ell,\, d}(0) \qquad (3\text{-}66)$$

Inserting Ineq. 3-66 into Ineq. 3-56 yields

$$\overline{N} \leqslant m\, b^{k_1} + 2m^2 \sum_{j=2} b^{k_j} e^{-\frac{1}{2} \mathbf{E}_{\ell,\, d}(0)\, K_{j-1} ad} \; ; \; R \leqslant \frac{1}{2} \mathbf{E}_{\ell,\, d}(0) \qquad (3\text{-}67)$$

Inserting Ineq. 3-61b into Ineq. 3-67 yields

$$\overline{N} \leqslant 2 \left\{ m\, e^{Rk_1 ad} + m^2 \sum_{j=2} e^{[Rk_j - \frac{1}{2} \mathbf{E}_{\ell,\, d}(0)\, k_{j-1}] ad} \right\} ;$$

$$R \leqslant \frac{1}{2} \mathbf{E}_{\ell,\, d}(0) \qquad (3\text{-}68)$$

By Eq. 3-51 we have

$$\overline{N} \leqslant 2 \left\{ m\, e^{Rk_1 ad} + m^2\, e^{\triangle Rad} \sum_{j=1} e^{[R - \frac{1}{2} \mathbf{E}_{\ell,\, d}(0)]\, [k_1 + j \triangle]\, ad} \right. ;$$

$$R \leqslant \frac{1}{2} \mathbf{E}_{\ell,\, d}(0) \qquad (3\text{-}69)$$

Let R^*_{comp} be defined as

$$R^*_{comp} = \frac{1}{2} \mathbf{E}_{\ell,\, d}(0) \qquad (3\text{-}70)$$

Then, for all rates below R^*_{comp}, $R - \frac{1}{2} \mathbf{E}_{\ell,\, d}(0) < 0$, and therefore

$$\overline{N} \leqslant 2 \left\{ m\, e^{Rk_1 ad} + m^2 \frac{e^{[R - R^*_{comp}]\, k_1 ad}}{1 - e^{[R - R^*_{comp}] \triangle ad}}\, e^{\triangle Rad} \right\} ; \; R < R^*_{comp} = \frac{1}{2} \mathbf{E}_{\ell,\, d}(0)$$

$$(3\text{-}71)$$

The bound on the average number of computations given by Ineq. 3-71 is minimized if we let*

$$\triangle = \frac{1}{ad} \frac{\ln R/R^*_{comp}}{R/R^*_{comp} - 1} \tag{3-72a}$$

$$k_1 = \triangle + \frac{1}{ad} \ln m \tag{3-72b}$$

Eq. 3-72 can be satisfied only if both Eq. 3-72a and Eq. 3-72b yield positive integers.

Inserting Eqs. 3-72 yields

$$\bar{N} \leqslant \frac{B^{\frac{B}{1-B}}}{1-B} m^{1+B} \; ; \qquad R \leqslant R^*_{comp} = \frac{1}{2} E_{\ell,\, d}(0) \tag{3-73a}$$

where

$$B = R/R^*_{comp} \leqslant 1 \tag{3-73b}$$

3.7 The Average Probability of Error of the Successive Decoding Scheme

Let u be the transmitted sequence of length n samples. Let u' be a member of the incorrect subset M''. (The set M'' consists of M'' members.) As shown in Section 2.5, u' is statistically independent of u. The probability of error is then bounded by

$$P_e \leqslant Pr[D(u, v) \geqslant D_0] + M''Pr[D(u', v) < D_0]$$

$$\leqslant Pr[D(u, v) \geqslant D_0] + e^{mdR}Pr[D(u', v) < D_0] \tag{3-74}$$

Now

$$D(u, v) = \ln \frac{p(v)}{p(v|u)}$$

where

$$p(v) = \prod_{i=1}^{m} p(y_i)$$

and

$$p(v|u) = \prod_{i=1}^{m} p(y_i|x_i)$$

Thus, by the use of Chernoff Bounds (Appendix A)

$$\Pr[D(u, v) \geqslant D_0] \leqslant e^{m[\gamma(s) - s\gamma'(s)]} \qquad (3\text{-}75a)$$

where

$$\gamma(s) = \ln \sum_{X_\ell} \int_Y P(x)\ p(y\ x)^{1-s} p(y)^s dy\ ; \qquad s \geqslant 0 \qquad (3\text{-}75b)$$

$$\gamma'(s) = D_0 \qquad (3\text{-}75c)$$

Also

$$\Pr[D(u', v) < D_0] \leqslant e^{m[\mu(t) - t\,\mu'(t)]} \qquad (3\text{-}76a)$$

$$\mu(t) = \ln \sum_{X_\ell} \int_Y P(x)\ p(y|x)^{-t}\ p(y)^{1+t} dy \qquad (3\text{-}76b)$$

$$\mu'(t) = D_0 \qquad (3\text{-}76c)$$

If we let t = s - 1, we have by Eq. 3-26

$$\gamma(s) = \mu(t)\ ; \qquad \gamma'(s) = \mu'(t)$$

Inserting Ineqs. 3-75, 3-76 and Eq. 3-26 into Ineq. 3-74 yields

$$P_e \leqslant e^{m[\gamma(s) - s\gamma'(s)]} + e^{m[dR + \gamma(s) - (s-1)\gamma'(s)]} \qquad (3\text{-}77a)$$

where $\gamma'(s) = D_0$ (3-77b)

Now, comparing Ineq. 3-75 with Ineq. 3-58 yields

$$- \gamma(s) + s\gamma'(s) = \frac{1}{2} d E_{\ell, d}(0) \tag{3-78}$$

On the other hand we have by Ineqs. 3-27 through 3-35

$$- 2\gamma(s) + (2s - 1)\gamma'(s) \geqslant dE_{\ell, d}(0) \tag{3-79}$$

Thus, inserting Ineq. 3-79 into Ineq. 3-79 yields

$$-\gamma(s) + (s-1)\gamma'(s) \geqslant \frac{1}{2} dE_{\ell, d}(0) \tag{3-80}$$

Inserting Ineqs. 3-78 and 3-80 into Ineq. 3-77 yields

$$P_e \leqslant e^{-\frac{1}{2} n E_{\ell, d}(0)} + e^{-n[\frac{1}{2} E_{\ell, d}(0) - R]}$$

$$\leqslant 2e^{-[\frac{1}{2} E_{\ell, d}(0) - R]} = 2e^{-n[R^*_{comp} - R]} ; \quad R \leqslant R^*_{comp} \tag{3-81}$$

If the input space is semi-optimal we have by Chapter 2 that $E_{\ell, d}(0) \cong E(0)$. Thus

$$P_e \leqslant 2e^{-[\frac{1}{2} E(0) - R]n} ; \qquad R \leqslant R^*_{comp} = \frac{1}{2} E(0) \tag{3-82}$$

If, instead of setting D_0 as we did in Ineq. 3-58, we set it so as to make $\gamma(s) - s\gamma'(s) = dR + \gamma(s) - (s-1)\gamma'(s)$, where $\gamma'(s) = D_0$, we have by Ineq. 3-77 that

$$P_e \leqslant 2e^{-nE_s(R)}$$

where by Ineqs. 3-43 through 3-45 we have (for semi-optimal input spaces)

$$E_s(0) \geqslant \frac{1}{2} E(0)$$

$$E_s(R) > 0 ; \qquad R < C$$

However, following Eq. 3-58 to 3-70, it can be shown that the new setting of D_0 yields $R^*_{comp} \geqslant \frac{1}{4} E_{\ell, d}(0)$.

The fact that the successive decoding scheme yields a positive exponent for rates above R^*_{comp} does not imply that this scheme should be used for such rates, since the number of computations for $R \geqslant R^*_{comp}$ grows exponentially with m.

CHAPTER IV

QUANTIZATION AT THE RECEIVER

4.1 Introduction

The purpose of introducing quantization at the receiver is to avoid the utilization of analogue devices. Due to the large number of computing operations which are carried out at the receiver, and the large flow of information to and from the memory, analogue devices may turn out to be more complicated and expensive than digital devices.

The discussion is limited to the Gaussian channel and to rates below R_{crit}. The effect of quantization on the zero-rate exponent of the probability of error is discussed in the three following cases:

Case I: (See Figure 9.) The quantizer is connected to the output terminals of the channel.

Case II: (See Figure 9.) The logarithm of the a posteriori probability per input letter (i.e. $p(y|x_i)$; $i = 1, \ldots, \ell$) is computed and then quantized.

Case III: (See Figure 9.) The logarithm of the a posteriori probability per p input letters (i.e., $p(y^p|x_j^p)$; $j=1, \ldots, \ell^p$) is computed and then quantized. (x_j^p is the vector sum of p successive input-letters of one of the M code words; y^p is the vector sum of the p received outputs.

It was shown in Section 2.3 that whenever semi-optimum input spaces are used with white Gaussian channels, $E_{\ell, d}(0)$ is a function of A^2_{max}, the maximum signal-to-noise ratio.

In this chapter, the effect of quantization is expressed in terms of "quantization loss" L_q in the signal-to-noise ratio of the unquantized channel.

Let $E^q_{\ell, d}(0)$ be the zero-rate exponent of the quantized channel. Then, by Eq. A-70

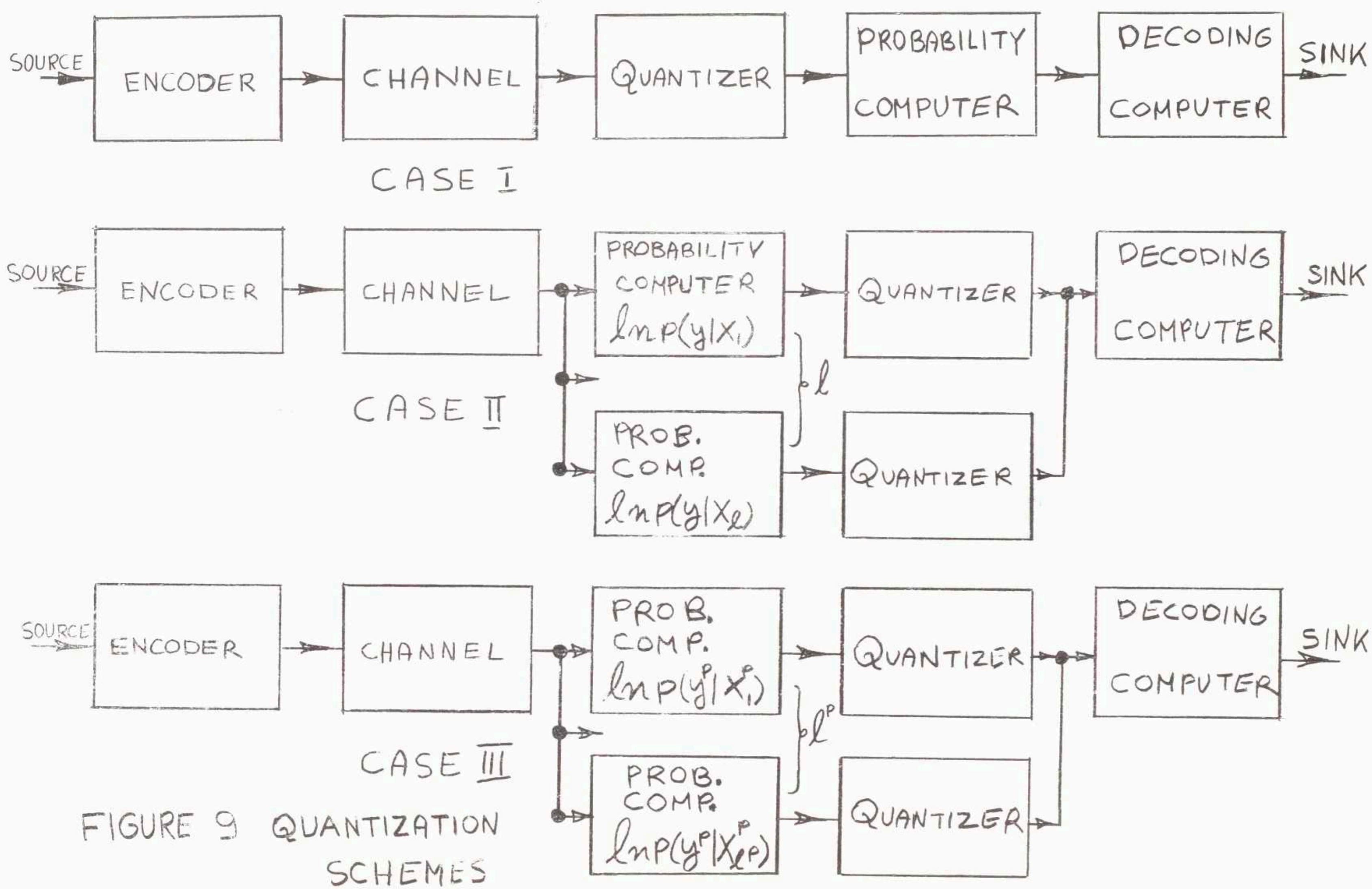


FIGURE 9 QUANTIZATION SCHEMES

$$E^q_{\ell, d}(R) = E^q_{\ell, d}(0) - R \; ; \qquad R \leqslant R_{crit}$$

Therefore specifying an acceptable $E^q_{\ell, d}(R)$ for any rate $R \leqslant R_{crit}$, corresponds to the specification of a proper loss factor L_q.

Let M_q be the number of quantization levels that are to be stored in the memory of the "Decision Computer" per one transmitted symbol.

Assuming that one of the two decoding schemes discussed in Chapter III is used, it is then convenient to define as the total decoding complexity measure* (including the quantizer).

$$M = M_d M_q \qquad \text{where } M_d \text{ is given by 2-145} \tag{4-1}$$

In Chapter 2 we have discussed the ways of minimizing M with respect to ℓ and d.

In this chapter we shall show that if semi-optimal input spaces are used with a white Gaussian channel, M_q of the quantization scheme of Case III (Figure 9) is always larger than that of Case II and therefore the quantization scheme of Case III should not be used.

Also, whenever $E(0) \cong \frac{1}{2} \ln A^2_{max} >> 1$, M_q of the quantization scheme of Case I (Figure 9) is smaller than that of Case II and therefore the quantization scheme of Case I should be used in such cases. On the other hand, whenever $E(0) << 1$ (or $A^2_{max} \ll 1$), M_q of Case II is smaller than that of Case I.

Furthermore, it will be shown that M, like M_d, is minimized if we let $d \cong \frac{1}{E(0)}$; $E(0) \ll 1$.

---

*We have assumed that the constraint length n, as well as the rate R and the signal power, are fixed.
The probability of error is then bounded by

$$P_e \leqslant 2e^{-E(R)n}$$

Now, given the acceptable probability of error $P_{e_a} \leqslant P_e$ one can find out what the acceptable exponent $E^q_{\ell, d}(R)$ is, and thus, what the acceptable quantization loss, L_q is. We shall therefore try to minimize M with respect to ℓ and d for a fixed n and a given quantization loss, L_q.

The results mentioned above are derived for the quantizer shown in Figure 10a which is equivalent to that of Figure 10b.

The interval Q (Figure 10b) is assumed to be large enough so that the limiter effect can be neglected as far as the effect on the exponent of the probability of error is concerned.

Thus, the quantizer of Figure 10b can be replaced by the one shown in Figure 10c. However, the actual number of quantization levels is not infinite as in Figure 10c, but rather is equal to $k = Q/q$ as in Figure 10b.

4.2 The Quantization Scheme of Case I (Figure 9)

The quantized zero-rate exponent $E^q_{\ell,\, d}(0)$ of Case I can be lower bounded by the zero-rate exponent of the following detection scheme:
The distance

$$d^q(x, y) = \frac{-2y^q x + x^2}{2\,\sigma^2} \tag{4-2}$$

is computed for each letter x_i of the tested code word.

Here y^q is the quantized vector of the channel output, y

$$y^q = \eta_1^q,\ \eta_2^q,\ \ldots,\ \eta_d^q \tag{4-3}$$

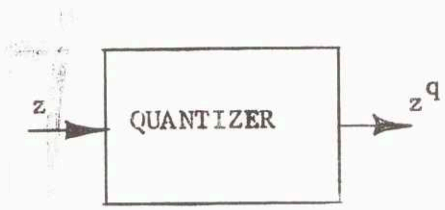
The distance

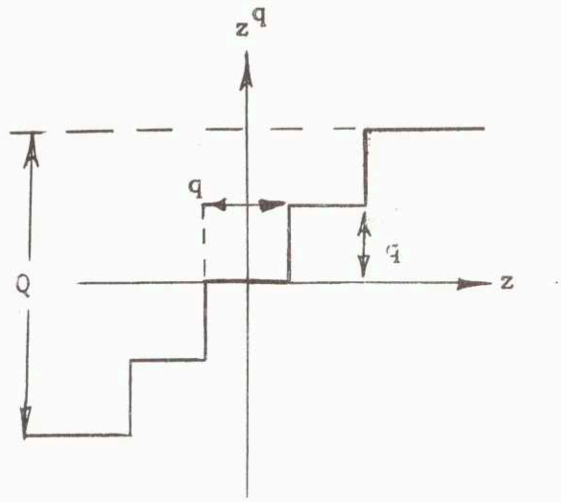
$$D^q(u, v) = \sum_{i=1}^{n} d_i^q(y_i, x_i)$$

is then computed.

The one code word that yields the smallest distance is chosen to represent the transmitted word. This detection procedure is optimal for the unquantized Gaussian variable y. However, y^q is not a Gaussian random variable and therefore $d^q(x, y)$ is not necessarily the best distance.

Thus, the above detection scheme will yield an exponent $E^*(0)$, which will be a lower bound on $E^q_{\ell,\, d}(0)$.

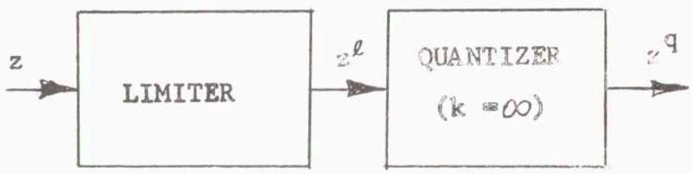


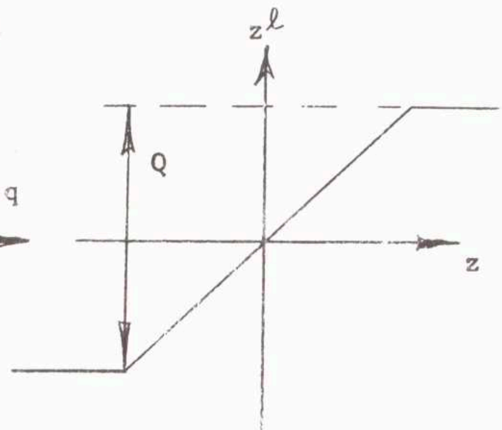


The number of quantization levels

$$k = \frac{Q}{q}$$

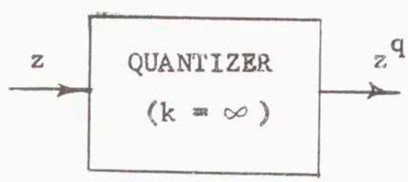
FIG. 10-A





The Quantizer ($k = \infty$) is given in Fig. 10-C

FIG. 10-B



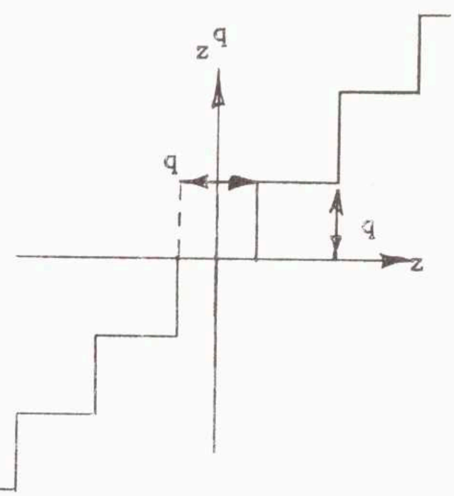


FIG. 10-C

FIGURE 10 QUANTIZERS AND THEIR TRANSFER CHARACTERISTICS

$$E^*(0) \leqslant E^q_{\ell, d}(0) \leqslant E_{\ell, d}(0) \quad (4\text{-}4)$$

The probability of error is bounded by

$$P_e \leqslant (M - 1)\ Pr[D(v^q, u') \leqslant D(v^q, u)] \quad (4\text{-}5a)$$

or

$$P_e \leqslant M\ Pr[D(v^q, u) - D(v^q, u') \geqslant 0] \quad (4\text{-}5b)$$

where u corresponds to the transmitted word and u' corresponds to some other code word.

$D^q(u, v)$ as well as $D^q(u', v)$ are sums of n independent random variables:

$$D^q(u, v) = \sum_{i=1}^{n} d_i^{\,q}(y_i, x_i) \quad (4\text{-}6a)$$

$$D^q(u', v) = \sum_{i=1}^{n} d_i^{\,q}(y_i, x_i') \quad (4\text{-}6b)$$

where x_i is the i^{th} transmitted letter and x_i' is the i^{th} letter of some other code word.

By the use of the Chernoff bounds (Appendix A, Section 3) it can be shown that

$$P_e \leqslant (M - 1)\ e^{-E^*(0)n} \leqslant e^{-n[E^*(0) - R]} \quad (4\text{-}7)$$

where, by Eq. A-65

$$-E^*(0) = \mu^*(s) = \frac{1}{d} \ln \sum_{Y^q} \sum_{X_\ell} \sum_{X_\ell'} P(x)\ P(x')\ p(y^q|x) e^{s[d^q(x,y) - d^q(x',y)]} \quad (4\text{-}7a)$$

$$;\quad s \geqslant 0 \quad (4\text{-}7b)$$

Now, let $s = \frac{1}{2}$. Then, by Eqs. 2-107a and 4-7

$$E^*(0) = -\frac{1}{d}\ln\sum_{X_\ell}\sum_{X_\ell'} P(x')P(x)e^{\frac{|x|^2-|x'|^2}{4\sigma^2}}\sum_{y^q} p(y^2|x)\, e^{\frac{y^q(x'-x)}{2\sigma^2}} \qquad (4\text{-}8a)$$

where

$$y^q x = \eta_1^q \xi_1 + \eta_2^q \xi_2 + \dots + \eta_d^q \xi_d \qquad (4\text{-}8b)$$

Thus, by Eq. 4-8, 4-4 and 2-107a

$$E^q_{\ell,d}(0) \geqslant -\frac{1}{d}\ln\sum_{X_\ell}\sum_{X_\ell'} P(x)P(x')\, e^{\frac{|x|^2-|x'|^2}{4\sigma^2}}\sum_{y^q} p(y^q|x)\, e^{\frac{y^q(x'-x)}{2\sigma^2}} \qquad (4\text{-}9a)$$

$$E^q_{\ ,d}(0) \leqslant -\frac{1}{d}\ln\sum_{X_\ell}\sum_{X_\ell'} P(x)P(x')\, e^{\frac{|x|^2-|x'|^2}{4\sigma^2}}\int_Y p(y|x)\, e^{\frac{y(x'-x)}{2\sigma^2}}\, dy = E_{\ell,d}(0)$$

$$(4\text{-}9b)$$

The complete information about the quantization effects is therefore carried by the term:

$$g^q(x, x') = \sum_{Y^q} p(y^q|x)\, e^{\frac{y^q(x'-x)}{2\sigma^2}} \qquad (4\text{-}10a)$$

when compared with the unquantized term:

$$g(x, x') = \int_Y p(y|x)\, e^{\frac{y(x'-x)}{2}}\, dy \qquad (4\text{-}10b)$$

The quantizer is a memoryless device; therefore, since the channel is memoryless as well, we have by Eq. 2-6

$$p(y^q|x) = p(\eta_1^q|\xi_1)\; p(\eta_2^q|\xi_2)\; \dots\; p(\eta_d^q|\xi_d)$$

Thus

$$g^q(x, x') = \prod_{i=1}^{d}\left\{\sum_{\eta_i^q} p(\eta_i^q|\xi_i)\, e^{\frac{\eta_i^q(\xi_i' - \xi_i)}{2\,\sigma^2}}\right\} \tag{4-11}$$

Two important signal-to-noise ratio conditions will be discussed

<u>Condition 1</u>: $\frac{\xi_{max}}{\sigma} = A_{max} \leqslant 1$ (4-12)

and at the same time $q \leq 2\,\sigma$

<u>Condition 2</u>: $\frac{\xi_{max}}{\sigma} = A_{max} > 1$ (4-13)

<u>a. Condition 1: $A_{max} \leq 1$ (Eq. 4-12)</u>

In this case we have that

$$\left|\frac{(\xi_i - \xi'_i)}{2\,\sigma}\right| \leqslant 1 \;;\; \text{for all } \xi \text{ and } \xi' \tag{4-14}$$

It is shown in Appendix C that whenever the quantizer of Figure 10c is used, and the input to the quantizer is a Gaussian random variable with a probability density such as in Eq. 2-79, we have that

$$\sum_{\eta_i^q} p(\eta_i^q|\xi_i)\, e^{\frac{\eta_i^q(\xi_i' - \xi_i)}{2\,\sigma^2}}$$

$$= \left[\int_{\eta_i} p(\eta_i|\xi_i)\, e^{\frac{\eta_i(\xi_i' - \xi_i)}{2\,\sigma^2}}\, d\eta_i\right] \frac{\operatorname{sh}\frac{(\xi_i' - \xi_i)q}{4\,\sigma^2}}{\frac{(\xi_i' - \xi_i)q}{4\,\sigma^2}} \;;\; \text{for } q < 2\,\sigma. \tag{4-15}$$

Inserting Eq. 4-15 into Eq. 4-11 yields:

$$g^q(x, x') = [\int_y p(y|x)\, e^{\frac{y(x'-x)}{2\sigma^2}} dy] \prod_{i=1}^{d} \frac{\operatorname{sh}\frac{(\xi_i' - \xi_i)q}{4\sigma^2}}{\frac{(\xi_i' - \xi_i)q}{4\sigma^2}} \tag{4-16}$$

Thus by Eqs. 4-16 and 4-10a

$$g^q(x', x) = g(x', x) \prod_{i=1}^{d} [\frac{\operatorname{sh}\frac{(\xi_i' - \xi_i)q}{4\sigma^2}}{\frac{(\xi_i' - \xi_i)q}{4\sigma^2}}] \tag{4-17}$$

Now

$$\frac{\operatorname{sh}x}{x} \leqslant e^{\frac{x^2}{6}} \tag{4-18a}$$

Also, for $x < 1$

$$\frac{\operatorname{sh}x}{x} \cong e^{\frac{x^2}{6}} \tag{4-18b}$$

Thus, by Eqs. 4-18a and 4-14

$$\frac{\operatorname{sh}\frac{(\xi_i' - \xi_i)q}{4\sigma^2}}{\frac{(\xi_i' - \xi_i)q}{4\sigma^2}} \lesssim e^{\frac{(\xi_i' - \xi_i)^2 q^2}{96\sigma^4}} \tag{4-18c}$$

Inserting Eq. 4-18c into Eq. 4-17 yields

$$g^q(x', x) \lesssim g(x', x) \prod_{i=1}^{d} e^{\frac{(\xi_i' - \xi_i)^2}{96\sigma^4}}$$

or

$$g^q(x', x) \lesssim g(x', x)\, e^{\sum_{i=1}^{d} (\xi_i' - \xi_i)^2 \frac{q^2}{96\sigma^4}} \qquad (4\text{-}19)$$

Therefore, by Eq. 2-4

$$g^q(x, x') \lesssim g(x, x')\, e^{\frac{q^2 |x-x'|^2}{96\sigma^4}} \qquad (4\text{-}20)$$

Replacing Eq. 4-10b by Eq. 4-20 and inserting Eq. 4-20 into Eq. 4-9a yields:

$$E^q_{\ell, d}(0) \geqslant -\frac{1}{d} \ln \sum_{X_\ell} \sum_{X'_\ell} P(x)\, p(x')\, e^{\frac{|x|^2 - |x'|^2}{4\sigma^2}}\, g(x, x')\, e^{\frac{q^2 |x - x'|^2}{96\sigma^4}} \qquad (4\text{-}21)$$

Inserting Eq. 4-10b into Eq. 4-21 yields:

$$E^q_{\ell, d}(0) \geqslant$$

$$-\frac{1}{d} \ln \sum_{X_\ell} \sum_{X'_\ell} P(x)\, P(x')\, e^{\frac{q^2 |x-x'|^2}{96\sigma^4}}\, e^{\frac{|x|^2 - |x'|^2}{4\sigma^2}} \int_y p(y|x)\, e^{\frac{y(x'-x)}{2\sigma^2}}\, dy \qquad (4\text{-}22)$$

Inserting Eq. 2-79 into Eq. 4-22 yields

$$E^q_{\ell, d}(0) \geqslant -\frac{1}{d} \ln \sum_{X_\ell} \sum_{X'_\ell} P(x)\, P(x')\, e^{\frac{q^2 |x-x'|^2}{96\sigma^4}}\, e^{-\frac{|x-x'|^2}{8\sigma^2}} \qquad (4\text{-}23)$$

$$E^q_{\ell, d}(0) \geqslant -\frac{1}{d} \ln \sum_{X_\ell} \sum_{X'_\ell} P(x')\, P(x)\, e^{-\frac{|x-x'|^2 (1 - q^2/12\sigma^2)}{8\sigma^2}} \qquad (4\text{-}24)$$

Now by Eq. 2-110a we have

$$E_{\ell,\, d}(0) = -\frac{1}{d} \ell n \sum_{X_\ell} \sum_{X'_\ell} P(x')\, P(x)\, e^{-\frac{|x-x'|^2}{8\sigma^2}} \qquad (4\text{-}25)$$

Comparing Eq. 4-23 with Eq. 4-25 yields that whenever the channel is in Condition 1 (eq. 4-12), the zero-rate exponent $E^q_{\ell,\, d}(0)$ of the quantized channel is lower-bounded by the zero-rate exponent of an unquantized Gaussian channel with an average noise power of:

$$\sigma_q^2 = \frac{\sigma^2}{1 - \frac{q^2}{12\sigma^2}} \qquad \text{for } q < 2\sigma \qquad (4\text{-}26a)$$

This result does not depend on the kind of input space which is used nor on its dimensionality, d. The effective signal-to-noise ratio of the quantized channel is given by

$$A_q^2 = \frac{\xi^2_{max}}{\sigma_q^2} = \frac{\xi^2_{max}}{\sigma^2} \left(1 - \frac{q^2}{12\sigma^2}\right)$$

Thus:

$$A_q^2 = A^2_{max} \left(1 - \frac{q^2}{12\sigma^2}\right) \qquad (4\text{-}26b)$$

Therefore, for a given quantization loss in the signal-to-noise ratio, let

$$q = \sqrt{12}\, L_q \sigma \qquad q < 2\sigma \qquad (4\text{-}27)$$

where L_q, the "quantization loss", is a constant that is determined by the acceptable loss in signal-to-noise ratio, as shown below:

$$\frac{A_q^2}{A^2_{max}} = 1 - L_q^2 \qquad (4\text{-}28)$$

The number of quantization levels is, as shown in Figure 10b, equal to

$$k = \frac{Q}{q} \tag{4-29}$$

It is quite clear* from the nature of the Gaussian probability density that if we let

$$\frac{Q}{2} = |\xi|_{max} + B\sigma \tag{4-30}$$

where B is a constant, then the effect of the limiter on $E_{\ell, d}(0)$ (shown in Figure 10b) is becoming negligible if B is large enough (in the order of 3).

Thus, inserting Eq. 4-30 into Eq. 4-29 yields

$$k = \frac{2\,\xi_{max} + 2B\sigma}{12\ L_q\,\sigma} \tag{4-31}$$

Now, if $\frac{\xi_{max}}{\sigma} = A << 1$

$$k = \frac{2B}{L_q\sqrt{12}} \tag{4-32}$$

The number of quantization levels for a given effective loss in signal-to-noise ratio is therefore independent of A_{max}, for $A_{max} << 1$. In the following section, the effective loss in signal-to-noise ratio for higher values of A, and the corresponding number of quantization levels k, are discussed.

b. Condition 2: $A_{max} \geq 1$ (Eq. 4-13)

In this condition we have that $A_{max} > 1$ and therefore

$$\left| \frac{\xi_i - \xi_i'}{2\sigma} \right| > 1$$

for some ξ and ξ'. Now, if $q \leq 2\sigma$, Eqs. 4-18c through 4-28 are valid.

*The following statement has been proved by the author.

The number of quantization levels is given by:

$$k = \frac{2\xi_{max} + 2B\sigma}{\sqrt{12}\,L\,\sigma}$$

or

$$k = \frac{2A_{max} + 2B}{\sqrt{12}\,L} \qquad (4\text{-}33a)$$

Thus, for $A_{max} \gg 1$

$$k = \frac{2A_{max}}{\sqrt{12}\,L} \quad ; \qquad q < 2\sigma \qquad (4\text{-}33b)$$

In this case, again, k does does not depend on the kind of input space which is used. There are many cases, however, where the assumption that $q < 2\sigma$ is unrealistic, since much larger quantization grain q can be used and still yield the acceptable loss L_q^2.

The effects of quantization in these cases depend heavily on the kind of input set which is used. This fact will be demonstrated by the following typical input sets.

1. The input set consists of two equiprobable oppositely directed vectors

$$x_1 = x \ ; \qquad x_2 = -x \qquad (4\text{-}34a)$$

where

$$P(x_1) = P(x_2) = \frac{1}{2} \qquad (4\text{-}34b)$$

As shown in Section 2.3, this input set is not optimal for $A_{max} > 1$. A semi-optimal input space for $A_{max} > 1$ is, as shown in Section 2.3, the following one:

2. The input set consists of ℓ equiprobable one-dimensional vectors. The distance between two adjacent vectors is $\frac{A_{max}}{\ell}$, as shown in Figure 4.

When the input set consists of two oppositely-directed vectors, we have by Eq. 2-111

$$E_{2,\,d}(0) = -\frac{1}{d}\ \ell n\ [\ \frac{1}{2} + \frac{1}{2}\ e^{-\frac{A_{max}^2 d}{2}}\ ] \qquad (2\text{-}111)$$

Also, by Eqs. 4-9a, 4-10a and 4-11

$$E^q_{2,\,d}(0) \geqslant -\frac{1}{d}\ \ell n \sum_{X_\ell}\sum_{X'_\ell} P(x)\ P(x')\ g^q(x,\ x')\ e^{\frac{|x|^2 - |x'|^2}{2\sigma^2}} \qquad (4\text{-}35a)$$

where, in this case,

$$g^q(x,\ x') = [\sum_{\eta^q} p(\eta^q|\xi)\ e^{\frac{\eta^q(\xi' - \xi)}{2\sigma^2}}]^d \qquad (4\text{-}35b)$$

since by Eq. 4-34, $\xi_i = \xi_j = \xi$ and $\xi_i' = \xi_j' = \pm\,\xi$ for all $i = 1, \ldots, d$ and $j = 1, \ldots, d$.

Thus, by Eq. 4-35

$$g^q(x,\ x') = 1 \qquad ; \qquad x' = x \qquad (4\text{-}36a)$$

$$g^q(x,\ x') = [\sum_{\eta^q} p(\eta^q|\xi)\ e^{\frac{\eta^q(\xi' - \xi)}{2\sigma^2}}]^d \quad ; \quad x \neq x' = -x \qquad (4\text{-}36b)$$

Now for $x' \neq x$,

$$\frac{|\xi' - \xi|}{2\sigma} = \frac{|2\xi|_{max}}{2\sigma} > 1, \text{ since } A_{max} > 1$$

It is shown in Appendix C that in such a case

$$\sum_{\eta^q} p(\eta^q|\xi)\ e^{\frac{\eta^q(\xi' - \xi)}{2\sigma^2}} \leq [\sum_{\eta} p(\eta|\xi)\ e^{\frac{(\xi' - \xi)\eta}{2\sigma^2}}]\ e^{\frac{|q(\xi' - \xi)|}{4\sigma^2}} \qquad (4\text{-}37)$$

Thus, inserting Eq. 4-37 into Eq. 4-36 yields

$$g^q(x, x') \leqslant \left[\sum_{\eta} p(\eta \mid \xi)\, e^{\frac{(\xi' - \xi)\eta}{2\sigma^2}}\right]^d e^{\frac{|q(\xi' - \xi)|d}{4\sigma^2}}$$

or

$$g^q(x, x') \leqslant \int_y p(y|x)\, e^{\frac{y(x'-x)}{2\sigma^2}} dy \cdot e^{\frac{|q(\xi' - \xi)|d}{4}} \qquad (4\text{-}38)$$

Thus, by Eqs. 4-38 and 4-10b we have

$$g^q(x, x') \leqslant g(x, x')\, e^{\frac{|q(\xi' - \xi)|d}{4}}$$

Thus

$$g^q(x, x') = g(x', x) = 1 \; ; \qquad x' = x \qquad (4\text{-}39a)$$

and

$$g^q(x, x') = g\,(x, x')\, e^{\frac{qAd}{2\sigma}} \; ; \qquad x' \neq x \qquad (4\text{-}39b)$$

Thus, inserting Eqs. 4-39 and 4-10b into Eq. 4-9a yields, together with Eq. 4-34

$$E^q_{2,\, d}(0) \geqslant -\frac{1}{d}\, \ln\left(\frac{1}{2} + \frac{1}{2}\, e^{-\frac{A^2 d}{2}}\, e^{\frac{qAd}{2\sigma}}\right) \qquad (4\text{-}40)$$

The zero-rate exponent of the unquantized channel is given by Eq. 2-111. Let

$$A_q^{\;2} = A^2 - \frac{q}{\sigma} A \qquad (4\text{-}41)$$

Inserting Eq. 4-41 into Eq. 4-40 yields

$$E^q_{2,\,d}(0) \geqslant -\frac{1}{d} \ln \left(\frac{1}{2} + \frac{1}{2} e^{-\frac{A_q^2 d}{2}} \right) \tag{4-42}$$

Thus, comparing Eq. 4-42 with Eq. 2-111 yields that the zero-rate exponent $E^q_{2,\,d}(0)$ of the quantized channel may be lower bounded by the zero-rate exponent of the unquantized channel (with the same input set) if the original signal-to-noise ratio A^2 is replaced by $A_q^{\ 2}$, which is given by Eq. 4-41.

Let

$$q = L_q^2 \, \xi_{max} = L_q^{\ 2} \xi \ ; \ \xi_{max} = \xi \tag{4-43}$$

where L_q is the "quantization loss" factor determined by the acceptable loss in the effective signal-to-noise ratio as shown below

$$\frac{A_q^{\ 2}}{A^2} = \frac{-(L_q^{\ 2} \xi / (\sigma A) + A^2}{A^2} = 1 - L_q^{\ 2} \ ; \qquad A = \frac{\xi}{\sigma} \tag{4-44}$$

Inserting Eqs. 4-43 and 4-30 into Eq. 4-29 yields

$$k = \frac{2\xi + 2B\sigma}{L_q^{\ 2} \xi} = \frac{2 + 2\ B/A}{L_q^{\ 2}}$$

Thus, for $A \gg 1$

$$k = \frac{2}{L_q^{\ 2}} \tag{4-45}$$

The number of quantization levels, for a given quantization loss in signal-to-noise ratio, is therefore independent of A_{max} for $A_{max} \gg 1$. Comparing Eq. 4-45 with Eq. 4-33b yields that for reasonably small L_q, the number of quantization levels needed for a given loss in signal-to-noise ratio is higher for $A_{max} \gg 1$ than it is for $A_{max} \ll 1$.

The binary input set does not yield the optimum zero-rate exponent since more than two letters are needed for $A_{max} > 1$. It was shown in Section 2-3 that an input-set that consists of ℓ one-dimensional vectors, yields a zero-rate exponent which is very close to the optimum one if the distance between two adjacent vectors is:

$$\frac{2\,\xi_{max}}{\ell} = 2\sigma \quad \text{or} \quad A_{max} = \ell$$

The zero-rate exponent of this input set is given by Eq. 2-115

$$E_{\ell,\,1}(0) = -\ln\left(\frac{1}{\ell} + 2\,\frac{\ell-1}{\ell^2}\, e^{-\frac{4A_{max}^2}{8\,\ell^2}} + 2\,\frac{-2}{2}\, e^{-\frac{16A_{max}^2}{8\,\ell^2}} + \dots\right)$$

Now since $\frac{A_{max}}{\ell} = 1$, we get from Eq. 2-116

$$E_{\ell,\,1}(0) \cong -\ln\left(\frac{1}{\ell} + 2\,\frac{\ell-1}{\ell^2}\, e^{-\frac{4A_{max}^2}{8\,\ell^2}}\right)$$

In other words, only adjacent vectors with a distance $|\xi' - \xi| = 2\sigma$ are considered.

For all such vectors we have

$$\frac{|\xi' - \xi|}{2\sigma} = 1 \tag{4-46}$$

Following Eqs. 4-14 through 4-31 yields that the number of quantization levels is

$$k = \frac{2\,\xi_{max} + 2B\sigma}{\sqrt{12}\,L_q\,\sigma}$$

or

$$k = \frac{2A_{max} + 2B}{\sqrt{12}\,L_q} \tag{4-47a}$$

Thus

$$k \cong \frac{2A_{max}}{\sqrt{12}\, L_q} \quad ; \qquad A_{max} >> 1 \qquad (4\text{-}47b)$$

The number of quantization levels in this case is therefore increasing with the signal-to-noise ratio.

<u>Summary</u>

The zero-rate exponent $E^q_{\ell,\, d}(0)$ of the quantized channel of Case I (Figure 9) may be lower bounded by the zero-rate exponent of the unquantized channel with the same input space if the signal-to-noise ratio, A^2, is replaced by $A_q^{\ 2}$, where

$$\frac{A_q^{\ 2}}{A^2} = 1 - L_q^{\ 2}$$

The quantization loss L^2 is a function of the number of quantization levels, k. The number of quantization levels for a given loss L_q^2 is constant for all $A << 1$, for all input sets. However, the number of quantization levels does depend on the input space whenever $A_{max} > 1$. Two typical input sets were introduced. The first input set consisted of two letters only, while the second input set was large enough to yield an $E_{\ell,\, d}(0)$ which is close to the optimum exponent $E(0)$.

It has been shown for both input spaces discussed in this section that the number of quantization levels for a given loss L_q is higher for $A_{max} << 1$ than it is for $A_{max} >> 1$. In the case of the semi-optimal input space shown in Figure 4, the number of quantization levels is increasing linearly with A_{max} (for $A_{max} >> 1$). The results are summarized in Table 3.

The quantization scheme of Case II (Figure 9) will now be discussed.

Input Space	A^2 Signal-to-noise Ratio	k No. of Quantization Levels	q Quantization Grain
All	$A \ll 1$	$\frac{2B}{L_q\sqrt{12}}$	$q = \sqrt{12}\, L_q \sigma$
Binary	$A \gg 1$	$\frac{2}{L_q^2}$ $\frac{2A}{L_q\sqrt{12}}$ for $q < 2\sigma$	$q = L_q^2 \xi_{max}$ $q = 12\, L \sigma$
Optimal	$A_{max} \gg 1$	$\frac{2A_{max}}{L_q\sqrt{12}}$	$q = \sqrt{12}\, L_q \sigma$

TABLE 3

QUANTIZATION SCHEME OF CASE I -- RESULTS

4.3 The Quantization Scheme of Case II (Figure 9)

In this case the logarithm of the a posteriori probability per input letter is computed and then quantized. The a posteriori probability per input letter is by Eq. 2-79*

$$p(y|x) = \frac{1}{(2\pi)^{d/2}\sigma^d} e^{-\frac{y^2 - 2xy + x^2}{2\sigma^2}} \qquad (4\text{-}48)$$

Thus

$$\ell n\, p(y|x) = \frac{2xy - x^2}{2\sigma^2} - \frac{y^2}{2\sigma^2} + \ell n\, (2\pi)^{\frac{d}{2}}(\sigma^2)^{\frac{d}{2}} \qquad (4\text{-}49)$$

The only part of Eq. 4-49 that carries information about x is

$$d(x, y) = \frac{|x|^2 - 2yx}{2\sigma^2} \qquad (4\text{-}50)$$

Thus, the computation of $\ell n\, p(y|x)$ may be replaced by the somewhat simpler computation of d(x, y) with no increase in the probability of error. The decoding scheme for the unquantized channel is discussed in Appendix A, Section A.2, with d(x, y) of Eq. 4-50 replacing d(x, y) of Eq. A-18. The corresponding probability of error is bounded by Section A-3

$$P_e \leq e^{-[E_{\ell, d}(0) - R]n} \quad ; \qquad R \leq R_{crit} \qquad (4\text{-}51a)$$

where $E_{\ell, d}(0)$ is given by

$$E_{\ell, d}(0) = -\frac{1}{d}\, \ell n \sum_{X_\ell}\sum_{X'_\ell} P(x)\, P(x') \int_y p(y|x)^{\frac{1}{2}}\, e^{[d(x, y)-d(x',y)]}\, dy \qquad (4\text{-}51b)$$

or

$$E_{\ell, d}(0) = -\frac{1}{d}\, \ell n \sum_{X_\ell}\sum_{X'_\ell} P(x)\, P(x')\, g(x, x') \qquad (4\text{-}52a)$$

---

* $xy = \xi_1\eta_1 + \xi_2\eta_2 + \dots + \xi_d\eta_d$

where

$$g(x', x) = \int_{y} p(y|x)\, e^{\frac{1}{2}[d(x, y) - d(x', y)]}\, dy \qquad (4\text{-}52b)$$

Now, comparing Eqs. 4-52 and 4-50 with Eq. 4-9b yields

$$g(x', x) = e^{-\frac{(x-x')^2}{8\sigma^2}} \qquad (4\text{-}52c)$$

Let the input to the quantizer be $d(x, y)$ given by Eq. 4-50 and let the output be $d^q(x, y)$.* The zero-rate exponent $E^q_{\ell, d}(0)$ of the quantized channel in Case II, can be lower bounded by the zero-rate exponent $E^*(0)$ of the following detection scheme: The distance $d_i(x_i, y_i)$, given by Eq. 4-50 is computed for each letter x_i of the tested code-word and then quantized to yield $d_i^q(x, y)$ the quantized version of $d_i(x, y)$. The distance

$$D^q(u, v) = \sum_{i=1}^{m} d_i^q(x_i, y_i) \qquad (4\text{-}53)$$

is then computed. The one code word that yields the smallest distance $D^q(u, v)$ is then chosen to represent the transmitted code word.

Thus

$$E^*(0) \leq E^q_{\ell, d}(0) \leq E_{\ell, d}(0) \qquad (4\text{-}54)$$

Following Eqs. A-65 and A-70 of Appendix A, we have

$$E^*(0) = -\frac{1}{d} \ln \sum_{X_\ell} \sum_{X'_\ell} P(x)\, P(x') \int_{Y} p(y|x)\, e^{[d^q(x', y) - d^q(x, y)]} dy \; ; \quad t \leq 0 \qquad (4\text{-}55a)$$

*As shown by Eq. 4-50 the quantity $x^2/2\sigma^2$ is added to $-2yx/2\sigma^2$ at the input to the quantizer rather than at its output. If each $\frac{1}{2}x_i^2/\sigma^2$ is equal to one of the k quantization levels exactly, one can add the quantity $\frac{1}{2}x_i^2/\sigma^2$ at the output to the quantizer, and the bounds will still be the same as those derived in the text.

and if we let $t = -\frac{1}{2}$

$$E^*(0) = -\frac{1}{d}\ \ln \sum_{X_\ell}\sum_{X'_\ell} P(x)\ p(x')\ g^q(x', x) \tag{4-55b}$$

where

$$g^q(x, x') = \int_Y p(y|x)\ e^{\frac{1}{2}\left[d^q(x, y) - d^q(x', y)\right]} dy \tag{4-55c}$$

Now, by Eq. 4-50

$$d(x,y) = \frac{|x|^2 - 2yx}{2\sigma^2} \tag{4-50}$$

where y is a d-dimensional Gaussian vector which, for a given x, consists of d independent Gaussian variables, each of which has a mean power of σ^2. Thus, d(x, y) is a Gaussian variable with an average variance of

$$\sigma_{d(x,\ y)} = \sqrt{\overline{(d\ (x,\ y)\ )^2} - (\overline{d\ (x,\ y)}\ )^2} = \sqrt{\frac{4\sigma^2\ |x|^2}{4\ \sigma^4}} = \frac{|x|}{\sigma} \tag{4-56}$$

Now, let

$$\begin{aligned} \frac{d(x,\ y)}{\sigma_{d(x,\ y)}} &= \frac{d(x,\ y)}{\frac{|x|}{\sigma}} = z \\ \frac{d(x',\ y)}{\sigma_{d(x',\ y)}} &= \frac{d(x',\ y)}{\frac{|x'|}{\sigma}} = z' \end{aligned} \tag{4-57}$$

Thus, by Eqs. 4-56 and 4-57, z and z' are normalized random Gaussian variables with a unit variance.

Inserting Eq. 4-57 into Eq. 4-55c yields

$$g^q(x, x') = \iint_{Z\,Z'} p(zz'|x,x')\, e^{\frac{1}{2}\left[\left(\frac{z\,|x|}{\sigma}\right)^q - \left(\frac{z'\,|x'|}{\sigma}\right)^q\right]} dz dz' \qquad (4\text{-}58)$$

since the product space Z Z' is identical with the space y for given x and x'. Now

$$\left(\frac{z\,|x|}{\sigma} \right)^q = d^q(x, y) = d(x, y) + n_q$$

where n_q is the "quantization noise". Thus

$$\left(\frac{|x|}{\sigma} z \right)^q = \frac{z\,|x|}{\sigma} + \left(\frac{n_q}{\frac{x}{\sigma}} \right) \frac{|x|}{\sigma}$$

or

$$\left(\frac{|x|}{\sigma} z \right)^q = \frac{|x|}{\sigma} (z^q) = \frac{|x|}{\sigma} \left[z + \frac{n_q}{\frac{|x|}{\sigma}} \right] \qquad (4\text{-}59a)$$

where

$$z^q = z + \frac{n_q}{\frac{|x|}{\sigma}} = \frac{d^q(x, y)}{\frac{x}{\sigma}} \qquad (4\text{-}59b)$$

Thus, z^q is equivalent to the output of a quantizer with z as an input and with a quantization grain that is equal to

$$q_z = \frac{q}{\frac{|x|}{\sigma}} \qquad (4\text{-}60)$$

where q is the quantization grain of the quantizer of d(x, y). Inserting Eq. 4-59 into Eq. 4-58 yields

$$g^q(x, x') = \int_{z'^q}\int_{z^q} p(z^q, z'^q | x, x')\, e^{\frac{1}{2}\left[\frac{|x|}{\sigma} z^q - \frac{|x'|}{\sigma} z'^q\right]} dz^q dz'^q \quad (4\text{-}61)$$

Both z and z' are Gaussian random variables, governed by the joint probability density:

$$p(z, z' | x, x') = \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left[\frac{-(z-\bar{z})^2 + 2\rho(z-\bar{z})(z'-\bar{z}') - (z'-z')^2}{2(1-\rho^2)}\right]$$

$$\text{where} \quad \rho = \overline{(z-\bar{z})\,(z'-\bar{z}')} \quad (4\text{-}62)$$

It is shown in Appendix C that for such a joint probability density as in Eq. 4-62, we have that

1. When $|\rho| = 1$ $(x = ax')$

$$g^q(x, x') = [g(x, x')]^{\frac{\frac{sh(|x| - |x'|)q_z}{4\sigma}}{\frac{(|x| - |x'|)}{4\sigma}q_z}} ; \qquad \text{for } q_z < 2 \quad (4\text{-}63)$$

Where, by Eq. 4-58

$$g(x, x') = \int_Z\int_{Z'} p(z, z' | x, x')\, e^{\frac{1}{2}\left[\frac{|x|}{\sigma}z - \frac{|x'|}{\sigma}z'\right]} dz dz' = \int_Y p(y|x)\, e^{\frac{1}{2}[d(x,y) - d(x',y)]} dy$$

(4-64a)

Also, it is assumed that

$$q^z = \frac{q}{\frac{|x|}{\sigma}} = \frac{q'}{\frac{|x'|}{\sigma}} = q_z' \quad (4\text{-}65)$$

where q is the quantization grain of the quantizer of d(x, y), and q' is the quantization grain of the quantizer of d(x', y). In other words, it is assumed that both d(x, y) and d(x', y) have the same <u>normalized</u>* quantization grain.

2. When $|\rho| < 1$

$$g^q(x, x') = [g(x, x')] \left(\frac{\frac{sh\,|x|\,q_z}{4\sigma}}{\frac{x\,q_z}{4\sigma}} \right) \left(\frac{\frac{sh\,|x'|\,q'_z}{4\sigma}}{\frac{x'\,q_z}{4\sigma}} \right) \text{ for } q_z < 2(1 - \rho^2) \tag{4-66}$$

3. When $|\rho| = 1$ and $q_z > 2$ we have that

$$g^q(x, x') \leqslant g(x, x')\, e^{\frac{(|x| - |x'|)q_z}{4\sigma}} \quad \text{for } q_z = q_z' \tag{4-67}$$

4. When $|\rho| < 1$ and $q_z > 2(1 - \rho^2)$ we have that

$$g^q(x, x') \leqslant g(x', x)\, e^{\frac{|x|q_z + |x'|q_z}{4\sigma}} \tag{4-68}$$

Studying Eqs. 4-66 through 4-68 yields that the effects of quantizations depend on the kind of input space which is used. The effect of quantization for three important input spaces will be discussed.

<u>a. The Binary Input Space</u>

The binary input space consists of two oppositely directed vectors

$$x_1 = x \;; \qquad x_2 = -x \tag{4-34a}$$

where

$$P(x_1) = P(x_2) = \frac{1}{2} \tag{4-34b}$$

This corresponds to $= -1$

---

*The quantization grain of each of one of the ℓ quantizers of Case II (Figure 9) is assumed to be proportional to the variance of the Gaussian variable $d_i(x, y)$ fed into that quantizer. The ℓ quantizers are therefore <u>not</u> identical.

The first signal-to-noise condition to be considered is

$$\sqrt{A^2 d} = \frac{|x|}{\sigma} \quad 1 ; \qquad q < 2 \tag{4-69}$$

By Eq. 4-63 we have

$$g^q(x, x') = [g(x, x')]^{\frac{\frac{sh|x| q_z}{2\sigma}}{\frac{|x| q_z}{2\sigma}}} \tag{4-70a}$$

and

$$g^q(x, x') = g(x, x') = 1 ; \qquad x = x' \tag{4-70b}$$

Now, inserting Eqs. 4-72 and 4-55c into Eq. 4-70 yields

$$E^q_{\ell,\,d}(0) \geqslant -\frac{1}{d} \ell n \left[\frac{1}{2} + \frac{1}{2} e^{-\frac{x^2}{2\sigma^2} \frac{\frac{sh|x| q_z}{2\sigma}}{\frac{|x| q_z}{2\sigma}}} \right. \quad ; q < 2 \tag{4-71}$$

Inserting Eq. 4-65 into Eq. 4-71 yields

$$E^q_{\ell,\,d}(0) \geqslant -\frac{1}{d} \ell n \left[\frac{1}{2} + \frac{1}{2} e^{-\frac{x^2}{2\sigma^2} \frac{sh\frac{q}{2}}{\frac{q}{2}}} \right] ; \qquad q < 2 \tag{4-72}$$

Inserting Eq. 4-18a into Eq. 4-72 yields

$$E^q_{\ell,\,d}(0) \geqslant -\frac{1}{d} \ell n \left[\frac{1}{2} + \frac{1}{2} e^{-\frac{x^2}{2\sigma^2}} e^{-\frac{q^2}{24}} \right] ; \qquad q < 2 \tag{4-73}$$

let

$$q = 12 L_q \frac{|x|}{\sigma} \tag{4-74}$$

Inserting Eq. 4-74 into Eq. 4-73 yields

$$E^q_{\ell,\,d}(0) \geq -\frac{1}{d}\,\ell n\,\left[\frac{1}{2}+\frac{1}{2}\,e^{-\frac{A^2 d}{2}(1-L_q^{\,2})}\right] \qquad (4\text{-}75)$$

as compared with Eq. 2-111

$$E_{\ell,\,d}(0) = -\frac{1}{d}\,\ell n\,\left[\frac{1}{2}+\frac{1}{2}\,e^{-\frac{A^2 d}{2}}\right] \qquad (2\text{-}111)$$

Thus, the zero-rate exponent $E^q_{\ell,\,d}(0)$ of the quantized channel is lower bounded by the zero-rate exponent of an unquantized channel with an effective signal-to-noise ratio $A_q^{\,2}$ given by $A_q^{\,2} = A^2(1 - L_q^{\,2})$ or $A_q^{\,2}/A^2 = 1 - L_q^{\,2}$ where L_q is the "quantization loss" factor.

Now the mean value of the Gaussian variable d(x, y) is, in general, different from zero. Thus, Eq. 4-30, which was derived for the Gaussian variable y which has a zero mean, is replaced by

$$Q = \overline{d(x', y)}_{max} - \overline{d(x', y)}_{min} + 2B\sqrt{\overline{d^2(x', y)} - (\overline{d(x', y)})^2} \qquad (4\text{-}76)$$

Now

$$y = \varphi + x \qquad (4\text{-}77)$$

where φ is a Gaussian vector that consists of d independent Gaussian variables with zero mean and a variance σ. Thus, by Eq. 4-50

$$\overline{d(x', y)}_{max} = \frac{|x|^2 + 2\,|x|\,|x|}{2\,\sigma^2} = \frac{3\,|x|^2}{2\,\sigma^2} \qquad (4\text{-}78a)$$

$$\overline{d(x', y)}_{min} = \frac{|x|^2 - 2\,|x|\,|x|}{2\,\sigma^2} = \frac{|x|^2}{2\,\sigma^2} \qquad (4\text{-}78b)$$

Inserting Eqs. 4-78 and 4-56 into Eq. 4-76 yields

$$Q = \frac{2\,|x|^2}{\sigma^2} + 2B\,\frac{|x|}{\sigma} \qquad (4\text{-}79)$$

Inserting Eqs. 4-79 and 4-74 into Eq. 4-29 yields

$$k = \frac{2\frac{|x|}{\sigma} + 2B}{\sqrt{12}\, L_q} \tag{4-80}$$

Thus for $A\sqrt{d} = \frac{|x|}{\sigma} << 1$

$$k = \frac{2B}{\sqrt{12}\, L_q} \tag{4-81}$$

Equation 4-75 is valid also in cases where $\sqrt{A^2 d} = \frac{|x|}{\sigma} > 1$ as long as $q_z < 2$ (or: $q < \frac{2|x|}{\sigma}$). Thus, by Eq. 4-80

$$k = \frac{2\frac{|x|}{\sigma}}{\sqrt{12}\, L} = \frac{2A\sqrt{d}}{\sqrt{12}\, L} \quad ; \quad A\sqrt{d} = \frac{|x|}{\sigma} >> 1,\ q < \frac{2|x|}{\sigma} \tag{4-82}$$

However, there are cases where much larger grain may be used. In such cases, where $q > \frac{2|x|}{\sigma}$ and $\frac{|x|}{\sigma} >> 1$, Equation 4-67 should be used. Therefore, by Eq. 4-67

$$g^q(x, x') = 1 \ ; \qquad x' = x \tag{4-83a}$$

and

$$g^q(x, x') = g(x, x')\, e^{\frac{|x| q_z}{2\sigma}} \ ; \qquad x' = -x \tag{4-83b}$$

Inserting Eq. 4-60 into Eq. 4-83 yields

$$g^q(x, x') = 1 \ ; \qquad x = x' \tag{4-84a}$$

and

$$g^q(x, x') = g(x, x')\, e^{\frac{q}{2}} \ ; \qquad x' = -x \tag{4-84b}$$

Inserting Eqs. 4-84 and 4-34b into Eq. 4-55b yields

$$E^q_{\ell,\, d}(0) \geqslant -\frac{1}{d}\ \ell n \left(\frac{1}{2} + \frac{1}{2}\, e^{-\frac{|x|^2}{2\sigma^2}}\, e^{\frac{q}{2}} \right) \tag{4-85}$$

Let

$$q = L_q^{\ 2}\, \frac{|x|^2}{\sigma^2} = L_q^{\ 2} A^2 d \tag{4-86}$$

then

$$E^q_{\ell,\, d}(0) \geqslant -\frac{1}{d}\ \ell n \left(\frac{1}{2} + \frac{1}{2}\, e^{-\frac{A^2 d}{2\sigma^2}(1 - L_q^{\ 2})} \right) \tag{4-87}$$

as compared with Eq. 2-111

$$E_{\ell,\, d}(0) \geqslant -\frac{1}{d}\ \ell n \left(\frac{1}{2} + \frac{1}{2}\, e^{-\frac{A^2 d}{2\sigma^2}} \right) \tag{2-111}$$

The number of quantization levels, for a given loss of signal-to-noise ratio is determined by inserting Eqs. 4-86 and 4-79 into Eq. 4-29 which yields

$$k = \frac{2\,\frac{|x|^2}{\sigma^2} + 2B\,\frac{|x|}{\sigma}}{L_q^2\, \frac{|x|^2}{\sigma^2}}$$

Thus, for $A\sqrt{d} = \frac{|x|}{\sigma} >> 1$ we have

$$k = \frac{2}{L_q^2} \tag{4-88}$$

<u>b. Orthogonal Input Set</u>

The binary input space is an optimal one, for $A < 1$, as shown in Section 2.3. Another optimal input space for $A \ll \frac{1}{d}$ is the orthogonal input space. In this case

$$x_i\, x_j = 0\ ; \qquad i \neq j \tag{4-89a}$$

$$x_i\, x_i = x_j\, x_j = |x|^2\ ; \qquad i = j \tag{4-89b}$$

for all $i = 1, \ldots, \ell$; $\quad j = 1, \ldots, \ell$.

Inserting Eq. 4-89 into Eq. 2-110a yields

$$E_{\ell,\, d}(0) = -\frac{1}{d}\ \ell n\left(\frac{1}{\ell} + \frac{\ell-1}{\ell}\, e^{-\frac{|x|^2}{4\sigma^2}}\right) \tag{4-90}$$

Now, since the input signals are orthogonal it can be shown that $\mathcal{S} = 0$. Following Eqs. 4-69 through 4-88, with Eq. 4-66 replacing 4-70a, Eq. 4-90 replacing 2-111 and Eq. 4-78 replaced by

$$\overline{d(x',\, y)}_{max} = \frac{3|x|^2}{2\sigma^2}\ ; \qquad \overline{d(x',\, y)}_{min} = -\frac{|x|^2}{2\sigma^2} \tag{4-91}$$

it can be shown that the number of quantization levels is

$$k = \frac{2B}{12\ L_q}\ ; \ A\sqrt{d} = \frac{|x|}{\sigma} << 1 \tag{4-92a}$$

$$k = \frac{2A\sqrt{d}}{12\ L_q}\ ; \ A\sqrt{d} = \frac{|x|}{\sigma} >> 1\ ;\ q < \frac{2|x|}{\sigma} \tag{4-92b}$$

$$k = \frac{2}{L_q^{\ 2}}\ ; \quad A\sqrt{d} = \frac{|x|}{\sigma} >> 1\ ;\ q > \frac{2|x|}{\sigma} \tag{4-92c}$$

c. Optimal Input Space

Both the binary and the orthogonal input spaces are non optimal for $A >> 1$. An input set which is a semi-optimal one for $A >> 1$ is shown in Figure 4 (Section 2.3). Now, if $d = 1$, it can be shown that $E^q_{\ell,\,1}(0)$ of the quantization scheme of Case II is equal to that of Case I.

The results of this section are summarized in Table 4. From Table 4 we may conclude that in Case II, as in Case I, the number of quantization levels for a given "quantization loss" is increasing with the signal-to-noise ratio which, in this case, is equal to $A^2 d/\sigma^2$.

4.4 The Quantization Scheme of Case III (Figure 9)

In this case the logarithm of the a posteriori probability per p input letters is computed and then quantized.

Let x^p be the vector sum of p input symbols. One can regard the vector sum x^p as a member of a new input space with "dp" dimensions. Equations 4-48 to 4-68 are therefore valid in Case III, once x is replaced by x^p.

Now, it has been demonstrated that, in both Case I and Case II, the number of quantization levels is increasing with the signal-to-noise ratio. If the signal-to-noise ratio in Case II is $A^2 d$, the signal-to-noise ratio in Case III is then $A^2 dp$.

Thus, given a quantization loss L_q and given an input space X_ℓ

$$k_{\text{Case II}} \leqslant k_{\text{Case III}} \tag{4-93}$$

4.5 Conclusions

Let M_q be the number of digits that are to be stored in the memory of the decision computer per each transmitted symbol.

Let M_{qI} and k_I be M_q and k of Case I.

Let M_{qII} and k_{II} be M_q and k of Case II.

Let M_{qIII} and k_{III} be M_q and k of Case III.

We therefore have

$$M_{qI} = k_I d \tag{4-94a}$$

$$M_{qII} = k_{II} \quad \text{for a binary input space (since only one "matched filter" should be used for both signals)} \tag{4-94b}$$

$$M_{qII} = k_{II}\ell \quad \text{(for any input space other than binary)} \tag{4-94c}$$

$$M_{qIII} = \frac{1}{p} k_{III} \ell^p \tag{4-94d}$$

Input Space	$A^2 = \frac{\lvert x \rvert^2}{d \sigma^2}$ Signal-to-noise ratio	k No. of Quantization Levels	q Quantization Grain
Binary	$A^2 d << 1$	$\frac{2B}{L\sqrt{12}}$	$\sqrt{12}\, A \sqrt{d}$
Binary	$A^2 d >> 1$	$\frac{2}{L^2}$; $q > 2A\sqrt{d}$ $\frac{2A\sqrt{d}}{\sqrt{12}\, L}$; $q < 2A\sqrt{d}$	$L^2 A^2 d$ $\sqrt{12}\, A\sqrt{d}$
Orthogonal	$A^2 d << 1$	$\frac{2B}{L\sqrt{12}}$	$\sqrt{12}\, A\sqrt{d}$
Orthogonal	$A^2 d >> 1$	$\frac{2}{L^2}$; $q > 2A\sqrt{d}$ $\frac{2A\sqrt{d}}{\sqrt{12}\, L}$; $q < 2A\sqrt{d}$	$\frac{L^2 A^2 d}{2}$ $\sqrt{12}\, A\sqrt{d}$
d = 1	See Table 10.1		

TABLE 4

QUANTIZATION SCHEME OF CASE II -- RESULTS

Inserting Ineq. 4-93 into Eq. 4-94d yields

$$M_{qIII} \geqslant \frac{1}{p} k_{II} \ell^{p} \tag{4-95}$$

Now $\frac{1}{p}\ell^{p} \geqslant \ell \quad ; \quad \ell \geqslant 2$ Thus

$$M_{qIII} \geqslant k_{II} \ell \; ; \; \ell \geqslant 2 \tag{4-96}$$

Comparing Ineq. 4-96 with Ineqs. 4-94b and 4-94c yields

$$M_{qIII} \geqslant M_{qII} \tag{4-97}$$

Thus, we may conclude that the quantization scheme of Case III should not be used.

Comparing Table 3 with Table 4 yields

$$k_I = k_{II} \quad ; \quad d = 1$$

Thus, by Eq. 4-94

$$M_{qI} \leqslant M_{qII} \; ; \quad d = 1 \tag{4-98}$$

We may therefore conclude that the quantization scheme of Case I should be used whenever $d = 1$.

From Tables 3 and 4 we have, that in the case of the binary input space, $k_I = k_{II}$ for $A^2 d \ll 1$. Thus, we have by Eq. 4-94

$$M_{qI} = k_I d \quad k_{II} = M_{qII}$$

or

$$M_{qI} > M_{qII}; \quad \text{binary input space;} \quad A^2 d \ll 1 \tag{4-99}$$

We may therefore conclude that whenever the signal-to-noise ratio is low enough ($A^2 d \ll 1$), the quantization scheme of Case II should be used.

As shown in Table 4, the number of quantization levels for a given L_q is not a function of d (as long as $A^2 d << 1$). Thus, the complexity measure, M, defined in Section 4.1, is like M_d, minimized by letting

$$d \cong \frac{1}{E(0)} \quad ; \qquad E(0) \cong \frac{1}{4} A^2 << \frac{1}{d}$$

From Section 2.3 and 2.5 it is clear that the binary input space is the best semi-optimal input space (for $A^2 d << 1$) since it yields the optimum exponent while the number of input vectors is kept as small as possible (i.e., $\ell = 2$).

If $E(0) \cong \ln A_{max} >> 1$, we have by Section 2.2 that $\frac{1}{d} \ell n \ell \cong E(0) >> 1$. Thus

$$\ell > \ell n \ell >> d \tag{4-100}$$

On the other hand one should expect k_{II} to be larger than k_I since the signal-to-noise ratio $A^2 d$ of Case II is larger than that of Case I (which is A^2), if $d > 1$. Thus

$$k_I \leq k_{II} \quad ; \qquad d > 1 \tag{4-101}$$

Inserting Ineqs. 4-100 and 4-101 into 4-94a and 4-94c yields

$$M_{qI} < M_{qII} \quad ; \qquad d > 1 \quad ; \qquad E(0) >> 1$$

<u>We thus conclude that whenever $E(0) \cong \ell n A_{max} >> 1$ and $d > 1$, the quantization scheme of Case I should be used.</u>

If an orthogonal set is used and at the same time $A^2 d << 1$, we have from Tables 3 and 4 that $k_I = k_{II}$. Thus, by Eq. 4-94

$$M_{qI} > M_{qII}; \qquad A^2 d << 1 \; ; \qquad \ell < d$$

4.6 Evaluation of $E^q_{\ell,d}(0)$ for the Gaussian Channel with a Binary Input Space ($\ell = 2$)

In the earlier sections of this chapter, methods were derived to lower bound $E^q_{\ell,d}(0)$.

In this section the exact value of $E^q_{\ell,d}(0)$ is evaluated for a binary input space (see Eq. 2-89). Let us first discuss the case where $d = 1$ and the output of the channel is quantized as follows (Case I; $k = 2$)

$$\begin{aligned} &\text{For all } y \geqslant 0 \; ; \qquad y^q = 1 \\ &\text{For all } y < 0 \; ; \qquad y^q = -1 \end{aligned} \tag{4-102}$$

where y^q is the output of the quantizer. The channel is converted into a binary symmetric channel, and is described by the following probabilities

$$P(x_1) = P(x_2) = \frac{1}{2} \quad ; \qquad x_1 = \xi_{max}, \; x_2 = -\xi_{max} \tag{4-103a}$$

$$P(1|x_1) = P(-1|x_2) = \int_0^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-x_1)}{2\sigma^2}} dy \tag{4-103b}$$

$$P(1|x_2) = P(2|x_1) = \int_0^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y+x_1)^2}{2\sigma^2}} dy \tag{4-103c}$$

By Eqs. A-71 and A-69

$$E^q_{2,1}(0) = -\ln \sum_{Y^q} \sum_{X_2} \sum_{X_2'} P(x)\, P(x')\, P(y^q|x)^{\frac{1}{2}}\, p(y^q|x')^{\frac{1}{2}} \tag{4-104}$$

Inserting Eq. 4-103 into Eq. 4-104 yields

$$E^q_{2,1}(0) = -\ln\left\{\frac{1}{2} + \frac{1}{2}\left[\int_0^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-x_1)^2}{2\sigma^2}} dy \int_0^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y+x_1)^2}{2\sigma^2}} dy\right]^{\frac{1}{2}}\right. \tag{4-105}$$

Now

$$\int_0^\infty \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}} dy \cong \frac{1}{2} + \frac{1}{\sqrt{2\pi}\,\sigma} x \quad ; \quad \frac{x}{\sigma} << 1 \qquad (4\text{-}106)$$

Inserting Eq. 4-106 into Eq. 4-105 yields

$$E^q_{2,\,1}(0) \cong - \ln \left\{ \frac{1}{2} + \frac{1}{2} \left(1 - \frac{2}{\pi} \frac{x_1^2}{2\sigma^2}\right) \right.$$

$$\cong \frac{2}{\pi} \frac{x_1^2}{4\sigma^2} = \frac{2}{\pi} \frac{\xi_1^2}{4\sigma^2} = \frac{2}{\pi} \frac{A^2_{max}}{4} \quad ; \quad A_{max} << 1 \; ; \; k = 2 \qquad (4\text{-}107)$$

Thus, by Eq. 2-103a we have

$$\frac{E^q_{2,\,1}(0)}{E_{2,\,1}(0)} = \frac{2}{\pi} \quad ; \quad A_{max} << 1 \; ; \quad k = 2 \qquad (4\text{-}108a)$$

and

$$L_q^{\;2} = 1 - \frac{2}{\pi} \quad ; \quad A_{max} << 1 \; ; \; k = 2 \qquad (4\text{-}108b)$$

Also

$$\int_0^\infty \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}} dy \cong 1 \; ; \; \frac{|x|}{\sigma} >> 1 \; , \; x > 0 \qquad (4\text{-}109a)$$

$$\int_0^\infty \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}} dy \cong e^{-\frac{x^2}{2\sigma^2}} \quad ; \quad \frac{|x|}{\sigma} >> 1, \quad x < 0 \qquad (4\text{-}109b)$$

Inserting Eq. 4-109 into 4-105 yields

$$E^q_{2,\,1}(0) \cong -\ \ln\left(\frac{1 + e^{-\frac{x_1^2}{4\sigma^2}}}{2}\right)$$

$$= -\ \ln\left(\frac{1 + e^{-\frac{A^2_{max}}{4}}}{2}\right)\ ; \qquad k = 2,\ A_{max} >> 1 \qquad (4\text{-}110)$$

Comparing Eq. 2-112 with Eq. 4-110 yields

$$L_q^{\ 2} = \frac{1}{2}\ ; \qquad A_{max} >> 1\ ;\ k = 2 \qquad (4\text{-}111)$$

If three quantization levels are used, (k = 3), it can be shown that

$$\frac{E^q_{2,\,1}(0)}{E^q_{2,\,1}(0)} = 0.81\ ; \qquad A_{max} << 1,\ k = 3 \qquad (4\text{-}112a)$$

$$L_q^{\ 2} = 0.19\ ; \qquad A_{max} << 1,\ k = 3 \qquad (4\text{-}112b)$$

If four quantization levels are used, (k = 4), it can be shown that

$$\frac{E^q_{2,\,1}(0)}{E_{2,\,1}(0)} \approx 0.86\ ; \qquad A_{max} << 1,\ k = 4 \qquad (4\text{-}113a)$$

$$L_q^{\ 2} = 0.14\ ; \qquad A_{max} << 1,\ k = 4 \qquad (4\text{-}113b)$$

Eqs. 4-107, 4-108, 4-111, 4-112 and 4-113 are valid also for the quantization scheme of Case II, if A^2_{max} is replaced by $A^2_{max}d$.

CHAPTER V

CONCLUDING REMARKS

The important features of this research are as follows:

1. It presents a method of sending data over a time discrete, amplitude-continuous memoryless channel with a probability of error which, for $R > R_{crit}$, has an exponent that can be made arbitrarily close to the optimum exponent $E(R)$. This is achieved by using a discrete input space.

2. It presents a decoding scheme with a probability of error no more than a quantity proportional to $\exp\left[-n\left(\frac{1}{2}E(0) - R\right)\right]$ and an average number of computations no more than a quantity proportional to m^2. The number of channel input symbols is roughly equal to $\ell n\, E(0)$ when $E(0) >> 1$, and is very small when $E(0) << 1$ (for the Gaussian channel we have that $\ell = 2$). The dimensionality of each input symbol is $d = 1$, when $E(0) >> 1$ and is equal to $d \cong \frac{1}{E(0)}$ whenever $E(0) << 1$.

3. It presents a method of estimating the effects of quantization at the receiver, for the white Gaussian channel. It was shown that the quantization scheme of Case I is to be used whenever $A_{max}^2 >> 1$. The quantization scheme of Case II is the one to be used whenever $A_{max}^2 << 1$.

Suggestions for Future Research

A method has been suggested (11) for adapting coding and decoding schemes for memoryless channels to channels with memory converted into memoryless channels by means of "scrambling" the transmitted messages. Extension of the results of this thesis to channels with memory, using scrambling or more sophisticated methods, would be of great interest.

Another very important and attractive extension would be the investigation of communication systems with a feedback channel. One should expect a further decrease in the decoding complexity and, probably, a smaller probability of error if feedback is used.

APPENDIX A

BOUNDS ON THE AVERAGE PROBABILITY OF ERROR-SUMMARY

A.1 Definitions

Following Fano (2), we shall discuss in this appendix a general technique for evaluating bounds on the probability of decoding error when a set of M equiprobable messages are encoded into sequences of m channel input events.

Let us consider a memoryless channel that is defined by a set of conditional probability densities $p(\eta \mid \xi)$, where ξ is the transmitted sample and η is the corresponding channel output ($p(\eta \mid \xi)$ is a probability distribution if η is discrete). We consider the case where each input event x is a d dimensional vector, and is a member of the (continuous) input space X.

The vector x is given by $x = \xi_1, \xi_2, \ldots, \xi_d$.

The corresponding d dimensional output vector y is a member of the d dimensional continuous space Y, with $y = \eta_1, \eta_2, \ldots, \eta_d$. The number of dimensions d is given by $d = \frac{n}{m}$, where n is the number of samples per message. The channel statistics are therefore given by

$$p(y|x) = \prod_{i=1}^{d} p(\eta_i | \xi_i) \text{ where } p(\eta_i | \xi_i) = p(\eta | \xi); \; \xi_i = \xi, \; \eta = \eta_i$$

The m^{th} power of this channel is defined as a channel with input space U consisting of all possible sequences u of m events belonging to X, and with output space V consisting of all possible sequences of m events belonging to Y. The i^{th} event of the sequence u will be indicated with y^i. Thus,

$$u = x^1, x^2, x^3, \ldots, x^m, \quad v = y^1, y^2, y^3, \ldots, y^m \tag{A-1}$$

where x^i may be any point of the input space X, and y^i may be any point of the output space Y.

Since the channel is constant and memoryless the conditional probability density $p(v|u)$ for the m^{th} power channel is given by

$$p(v|u) = \prod_{i=1}^{m} p(y^i|x^i) \tag{A-2}$$

where

$$p(y^i|x^i) = p(y|x); \quad y^i = y, \; x^i = x \tag{A-3}$$

We shall assume in what follows that the message space consists of M equiprobable messages $m_1, m_2, \ldots, m_M$.

A.2 Random Encoding for Memoryless Channels

In the case of random encoding we consider the case where the input sequences assigned to messages are selected independently at random with probability density $p(u)$, if U is a continuous space, or with probability distribution $p(u)$ if U is discrete. The average probability of error corresponding to any such random assignment of input sequences to messages depends, of course, on the probability density $p(u)$. We shall set

$$p(u) = \prod_{i=1}^{m} p(x^i) \tag{A-4}$$

where

$$p(x^i) = p(x); \quad x^i = x \tag{A-5}$$

$p(x)$ is an arbitrary probability density whenever X is continuous, and is an arbitrary probability distribution whenever X is discrete. Eq. A-4 is equivalent to saying that the input sequence corresponding to each particular message is constructed by selecting its component events independently at random with probability (density) $p(x)$.

We shall assume, unless mentioned otherwise, that the channel output is decoded according to the maximum likelihood criterion: that is, that any particular output sequence v is decoded into the message m_i that maximizes the conditional probability (density) $p(v|m_i)$. Since messages are, by assumption, equiprobable, this decoding criterion is equivalent to maximizing the a posteriori probability $p(m_i|v)$, which in turn results in the minimization of the probability of error.

Let us assume that a particular message has been transmitted, and indicate by u the corresponding input sequence and by v the resulting output sequence. According to the specified decoding criterion, an error can occur only if one of the other M-1 messages is represented by an input sequence u' for which

$$p(v|u') \geqslant p(v|u) \tag{A-6}$$

Let F(v) be an arbitrary positive function of v satisfying the condition

$$\int F(v)\, dv = 1 \tag{A-7}$$

or

$$\sum F(v) = 1 \tag{A-7a}$$

if v is discrete.

Also define

$$D(u, v) = \ln \frac{F(v)}{p(v|u)} \tag{A-8}$$

as the "distance" between u and v. In terms of this measure of distance the condition expressed by Eq. A-6 becomes

$$D(u', v) \leqslant D(u, v) \tag{A-9}$$

For any arbitrary constant D_0, the average probability of error then satisfies the inequality

$$P_e \leqslant MP_1 + P_2 \tag{A-10}$$

where

$$P_1 = \Pr[D(u, v) \leqslant D_0, D(u', v) \leqslant D(u, v)] \tag{A-11}$$

and

$$P_2 = \Pr[D(u, v) > D_0] \tag{A-12}$$

The bound of Eq. A-10 corresponds to the following decoding scheme: $D(u, v)$ of Eq. A-8 is computed for each one of the M sequences of the input space U and the one given output sequence v. The only distances $D(u, v)$ that are taken into further consideration are those for which $D(u, v) \leqslant D_0$, where D_0 is an arbitrary constant. The one sequence u, out of all the sequences for which $D(u, v) \leqslant D_0$, that yields the smallest distance $D(u, v)$ is chosen to represent the transmitted signal. If no such sequence u exists, an error will occur.

If the above decoding procedure is carried out with an arbitrary distance function of u and v, $D^q(u, v)$, other than the $D(u, v)$ of Eq. A-8, then the average probability of error satisfies the inequality

$$P_e \leqslant MP_1 + P_2 \tag{A-13}$$

where

$$P_1 = \Pr[D^q(u, v) \leqslant D_0; D^q(u', v) \leqslant D_q(u, v)] \tag{A-14}$$

$$P_2 = \Pr[D^q(u, v) > D_0] \tag{A-15}$$

However, one would expect the bound of Ineq. A-13 to be larger than that of Eq. A-10, if $D_q(u, v)$ is not a monotonic function of the a posteriori probability $p(u|v)$.

A.3 Upper Bounds on P_1 and P_2 by Means of Chernoff Bounds

The m events constituting the sequence u assigned to a particular message are selected independently at random with the same probability p(x). If we let

$$F(v) = \prod_{i=1}^{m} f(y^i) \tag{A-16}$$

where

$$[f(y^i)]_{y^i = y} \equiv f(y); \quad \int f(y)dy = 1$$

or

$$\sum f(y) = 1, \tag{A-17}$$

when y is discrete,

it then follows from Eqs. A-2, A-3, and A-16 that the random variable D(u, v) defined by Eq. A-8 is the sum of m statistically independent, equally distributed, random variables:

$$D(u, v) = \sum_{i=1}^{m} d(x_i, y_i) \tag{A-18a}$$

where

$$d(x^i, y^i) \equiv d(x, y) = \ln \frac{f(y)}{p(y\ x)} \quad ; x^i = x; \ y^i = y \tag{A-18b}$$

In cases where an arbitrary distance $D^q(uv)$ other than D(uv) of Eq. A-8 is used, the discussion will be limited to such distances $D^q(uv)$ which may be represented as a sum of m statistically independent, equally distributed, random variables.

$$D^q(u, v) = \sum_{i=1}^{m} d^q(x^i, y^i) \tag{A-19}$$

where

$$d^q(x^i, y^i) = d^q(x, y); \qquad x^i = x, \; y^i = y \tag{A-20}$$

The moment generating function (m.g.f.) of the random variable D(u,v), is:

$$G(s) = \int_{D(u,v)} p[D(u, v)] \, e^{sD(u, v)} \, dD(u, v) \tag{A-21}$$

where p[D(u,v)] is the probability density of D(u,v). Thus

$$G(s) = \int_u \int_v p(u) \, p(v|u) \, e^{sD(u, v)} \, dudv \tag{A-22}$$

From Eqs. A-18, A-4, A-5, A-2, and A-3, we get

$$G(s) = \prod_{i=1}^{m} \int_x \int_y p(x) \, p(y|x) \, e^{sd(x, y)} dxdy = [g_d(s)]^m \tag{A-23}$$

where

$$g_d(s) = \int_x \int_y p(x) \, p(y|x) \, e^{sd(xy)} dxdy \tag{A-24}$$

In the case of P_2 of Eq. A-12 we are interested in the probability that D(u,v) is greater than some value D_0. For all values of D(uv) for which $D(uv) \geqslant D_0$

$$e^{sD(uv)} \geqslant e^{sD_0}, \qquad \text{for } s \geqslant 0.$$

Using this fact, we may rewrite Eq. A-21 as:

$$G(s) \geqslant e^{sD_0} \int_{D(uv) > D_0} p[D(uv)]\, dD(uv) \tag{A-25}$$

Using Eq. A-23,

$$\Pr[D(uv) > D_0] \leqslant e^{m\,\gamma_d(s) - sD_0} \tag{A-26}$$

where

$$\gamma_d(s) = \ln g_d(s) = \ln \int_X \int_Y p(x)\; p(y|x)\; e^{sd(xy)} dxdy \tag{A-27}$$

Equation (A-26) is valid for all $s \geqslant 0$. We may choose s such that the exponent is minimized. Differentiation with respect to s and setting the result equal to zero, yields

$$\Pr[D(u, v) > D_0] \leqslant e^{m[\gamma_d(s) - s\,\gamma_d'(s)]}; \qquad s \geqslant 0 \tag{A-28}$$

where s is the solution to

$$\gamma_d'(s) = \frac{d\,\gamma_d(s)}{ds} = \frac{D_0}{m} \tag{A-29}$$

In the same way

$$\Pr[D^q(u, v) > D_0] \leqslant e^{m\,\gamma_d^q(s) - sD_0} \tag{A-30}$$

where

$$\gamma_d^q(s) = \ln g_d^q(s) = \ln \int_X \int_Y p(x)\; p(y|x)\; e^{sD^q(x,\, y)} dxdy \tag{A-31}$$

The exponent of Eq. A-30 is minimized if we choose s such that

$$\gamma_d^{q'}(s) = \frac{d\gamma_d^{q}(s)}{ds} = \frac{D_0}{m} \tag{A-32}$$

Thus

$$\Pr[D^q(u, v) > D_0] \leq e^{m[\gamma_d^{q}(s) - s\gamma_d^{q'}(s)]} \tag{A-33}$$

In the case of P_1 we desire an upper bound to the probability

$$P_1 = \Pr[D(u, v) \leq D_0, D(u', v) \leq D(u, v)] \tag{A-11}$$

For this purpose let us identify the point uu'v of the product space UU'V with the point a of a space A, the probability (density) p(uu'v)= p(u)p(v')p(v|u) with the probability (density) p(a), the random variable D(u,v) with $\emptyset(a)$, and the random variable D(u', v) - D(u, v) with the random variable $\theta(a)$. Inserting $\theta(a)$ and $\emptyset(a)$ into Eq. A-11 yields

$$P_1 = \Pr[\emptyset(a) \leq D_0, \theta(a) \leq 0] \tag{A-34}$$

Let us form the m.g.f. of the pair $(\emptyset(a), \theta(a))$.

$$G(r, t) = \int p(a)\, e^{r\emptyset(a)+t\theta(a)}\, da \tag{A-35}$$

Now, for all values of $\{a: \emptyset(a) \leq D_0; \theta(a) \leq 0\}$,

$$e^{r\emptyset(a)+t\theta(a)} \geq e^{rD_0} \qquad \text{for } r \leq 0;\ t \leq 0$$

Using this fact, Eq. A-35 may be rewritten as

$$G(r, t) \geq e^{rD_0} \int_{\{a:\ \emptyset(a) \leq D_0;\ \theta(a) \leq 0\}} p(a)\,da$$

or

$$G(r, t) \geqslant e^{rD_0} \Pr[\phi(a) \leqslant D_0, \theta(a) \leqslant 0] \qquad \text{(A-36)}$$

Thus

$$p_1 = \Pr[\phi(a) \leqslant D_0, \theta(a) \leqslant 0] \leqslant G(r, t)\, e^{-rD_0}, \quad r \leqslant 0;\ t \leqslant 0 \qquad \text{(A-37)}$$

Now

$$\phi(a) = D(u, v), \ \theta(a) = D(u', v) - D(u, v),$$

and

$p(a) = p(uu'v) = p(u)p(u')p(v|u)$. Thus

from Eqs. A-18, A-4, A-5, A-2, A-3 and A-35 we get

$$G(r, t) = [g_d(r, t)]^m \qquad \text{(A-38)}$$

where

$$g_d(r, t) = \iiint_{Y\,X'\,X} p(x)p(x')p(y|x)\, e^{(r-t)d(x, y)+td(x', y)} dxdx'dy \qquad \text{(A-39)}$$

Inserting Eq. A-38 into Eq. A-36 yields

$$p_1 = \Pr[D\ u, v) \leqslant D_0, D(u', v) \leqslant D(u, v)] \leqslant e^{m\gamma_d(r, t)-rD_0};$$

$$r \leqslant 0;\ t \leqslant 0 \qquad \text{(A-40)}$$

where

$$\gamma_d(r, t) = \ln g_d(r, t)$$

$$= \ln \iiint_{Y\,X\,X'} p(x)p(x')p(y|x)\, e^{(r-t)d(xy)+td(x', y)} dxdx'dy \qquad \text{(A-41)}$$

We may choose r and t such that the exponent of the right-hand side of Ineq. A-40 is minimized. Differentiating with respect to r and setting the result equal to zero and then repeating the same procedure with respect to t, we obtain

$$P_1 = \Pr[D(uv) \leqslant D_0,\ D(u'v) \leqslant D(uv)] \leqslant e^{m[\gamma_d(r,\ t) - r\,\gamma'_{d_r}(r,\ t)]} \tag{A-42}$$

where

$$\gamma_{d_r}{}'(r,\ t) = \frac{\partial \gamma_d(r,\ t)}{\partial\ r} = \frac{D_0}{m} \tag{A-43}$$

and

$$\gamma_{d_t}{}'(r,\ t) = \frac{\partial \gamma(r,\ t)}{\partial\ t} = 0 \tag{A-44}$$

In the same way

$$\Pr[D^q(uv) \leqslant D_0,\ D^q(u'v) \leqslant D_q(uv)] \leqslant e^{m[\gamma_d{}^q(r,\ t) - rD_0/m]} \tag{A-45a}$$

where

$$\gamma_d{}^q(r,\ t) = \ell n\ g_d{}^q(r,\ t) = \ell n \int_Y\int_X p(x)p(y|x)\ e^{(r-t)d_q(xy)+td_q(x'y)}\ dxdy \tag{A-45b}$$

Inserting Ineqs. A-40, and A-30 into Ineq. A-13 yields:

$$P_e \leqslant e^{m[\gamma_d(s) - sD_0/m]} + e^{m[n/m\ R + \gamma_d(r,\ t)\ -\ rD_0/m]} \tag{A-46a}$$

where R, the rate of information per sample, is given by

$$R = \frac{1}{n}\ \ell nM \tag{A-46b}$$

From Eqs. A-11 and A-12 we have that the two probabilities, p_1 and p_2, vary monotonically with D_0 in the opposite directions. Thus, the right-hand side of Ineq. A-45 is approximately minimized by the value of D_0 for which the two exponents are equal to each other. Therefore, let D_0 be such that

$$\gamma_d(s) - s\frac{D_0}{m} = \frac{n}{m}R + \gamma_d(r, t) - r\frac{D_0}{m} \qquad \text{(A-47)}$$

The insertion of Eqs. A-47, A-44, and A-43 into A-46a yields

$$P_e \leqslant 2e^{m[\gamma_d(s) - sD_0/m]} = 2e^{-nE_d(R)} \qquad \text{(A-48)}$$

where

1. $$E_d(R) = -\frac{m}{n}[\gamma_d(s) - s\frac{D_0}{m}] = -R + \frac{m}{n}(\gamma_d(r, t) - r\frac{D_0}{m}) \qquad \text{(A-49)}$$

2. $$\gamma_{d_s}'(s) = \gamma_{d_r}'(r, t) = \frac{D_0}{m} \qquad s \geqslant 0,\ t \leqslant 0,\ r \leqslant 0 \qquad \text{(A-50)}$$

3. $$\gamma_{d_t}'(r, t) = 0 \qquad r \leqslant 0;\ t \leqslant 0 \qquad \text{(A-51)}$$

Now, from Eqs. A-25 and A-18 we have

$$\gamma_d(s) = \ln \int_Y \int_X p(x)\ p(y|x)\ e^{sd(xy)} dxdy$$

$$= \ln \int_Y \int_X p(x)\ p(y|x)^{1-s} f(y)^s dxdy \qquad \text{(A-52)}$$

Also, from Eqs. A-39 and A-18 we have

$$\gamma_d(r, t) = \ln \int_y \int_{x'} \int_x p(x)\ p(x')\ p(y|x)^{1-r\ t} p(y|x')^{-t} f(y)^r \qquad \text{(A-53)}$$

It can be shown (Ref. (2), pp. 324-332) that

1. Eq. A-44 is satisfied if we let

$$1 - r + t = -t \qquad ; \quad r \leqslant 0, \quad t \leqslant 0$$

or

$$r = 1 + 2t \qquad ; \quad t \leqslant -\frac{1}{2} \tag{A-54}$$

2. Eq. A-50 is satisfied by letting

$$f(y) = \frac{\left[\int_X p(x)\; p(y|x)^{1-s} dx\right]^{1/1-s}}{\int_Y \left[\int_X p(x)\; p(y|x)^{1-s} dx\right]^{1/1-s} dy} \tag{A-55a}$$

and

$$s = 1 + t \qquad ; \; 0 \leqslant s \leqslant \frac{1}{2} \tag{A-55b}$$

3. Eq. A-49 is then satisfied if we let

$$R = \frac{1}{d}\left[(s-1)\,\gamma_d{}'(s) - \gamma_d(s)\right] \qquad ; \quad 0 \leqslant s \leqslant \frac{1}{2} \tag{A-56a}$$

We should notice, however, that Eqs. A-44, A-50 and A-49 are satisfied if, and only if, R is such as to make $0 \leqslant s \leqslant \frac{1}{2}$. It can be shown (Ref. (2), pp. 324-332) that this corresponds to the region

$$R_{crit} \leqslant R \leqslant I \tag{A-56b}$$

where

$$I = \frac{1}{d} \int_Y \int_X p(x)\; p(y|x)\; \ln \frac{p(y|x)}{p(y)}\, dxdy = [R]_{s=0} \tag{A-56c}$$

and

$$R_{crit} = [R]_{s = 1/2} \qquad \text{(A-56d)}$$

Let us now define for the product space XY the tilted probability (density)

$$Q(x, y) = \frac{e^{sD(x, y)} p(x)\ p(y|x)}{\int_Y \int_X e^{sD(x, y)} p(x)\ p(y|x)\ dxdy}$$

$$= \frac{p(x)\ p(y|x)^{1-s} f^s(y)}{\int_Y \int_X p(x)\ p(y|x)^{1-s}\ f^s(y)\ dxdy} \qquad \text{(A-57)}$$

where

$$Q(y) = f(y) = \frac{[\int_X p(x)\ p(y|x)^{1-s} dx]^{1/1-s}}{\int_Y [\int_X p(x)\ p(y|x)^{1-s} dx]^{1/1-s} dy} \quad ; 0 \leqslant s \leqslant \frac{1}{2} \qquad \text{(A-58)}$$

$$Q(x|y) = \frac{Q(x, y)}{Q(y)} \quad \frac{p(x)\ p(y|x)^{1-s}}{\int_X p(x)\ p(y|x)^{1-s} dx} \quad ; 0 \leqslant s \leqslant \frac{1}{2} \qquad \text{(A-59)}$$

Inserting Eqs. A-52, A-53, A-54, A-56, A-57, A-59

$$P_e \leqslant 2e^{-nE(R)} \quad ; R_{crit} \leqslant R \leqslant I \qquad \text{(A-60a)}$$

where the exponent E(R) is related parametrically to the transmission rate per sample R, for $R_c \leqslant R \leqslant I$, by

$$0 \leqslant E(R) = \frac{1}{d} \iint_{Y\,X} Q(x, y) \ln \frac{Q(x, y)}{p(x)\ p(y|x)} dxdy \tag{A-60b}$$

$$I \leqslant R \leqslant \frac{1}{d} \iint_{X\,Y} Q(x, y) \ln \frac{Q(x|y)}{P(x)} \geqslant R_{crit} \quad ; 0 \leqslant s \leqslant \frac{1}{2} \tag{A-60c}$$

$$R_{crit} = [R]_{s = 1/2}; \; I = [R]_{s = 0} = \frac{1}{d} \iint_{X\,Y} p(x)\ p(y|x) \ln \frac{p(y|x)}{p(y)} \tag{A-60d}$$

Whenever $R < R_{crit}$, there does not exist a D_0 that simultaneously satisfies Eqs. A-49, A-50 and A-51. However, the average probability of error may always, for any rate, be bounded by

$$P_e \leqslant MPr[D(u'v) \leqslant D(uv)] \tag{A-61}$$

This is equivalent to setting $D_0 =$ in Eqs. A-11 and A-12. Thus,

$$P_1 = Pr[D(u'v) \leqslant D(uv)] \quad ; P_2 = 0 \tag{A-61}$$

In the same way:

$$P_e \leqslant MP_1 = MPr[D_q(u'v) \leqslant D_q(uv)] \tag{A-63}$$

The evaluation of P_1 under these conditions proceeds as before, except for setting $r = 0$ in Ineqs. A-42 and A-45a. Therefore

$$P_e \leqslant e^{m[n/m\,R + \gamma_d(0, t)]} \quad ; t \leqslant 0 \tag{A-64}$$

where

$$\gamma_d(0,\ t) = \gamma\,(t)$$

$$= \ln \iiint\limits_{Y\ X\ X'} p(x')\ p(x)\ p(y|x)\ e^{t[d(x'y)\ -\ d(xy)]} dx'dxdy\ ;$$

$$t \leq 0 \qquad \text{(A-64a)}$$

and

$$P_e \leq e^{m[n/m\ R\ +\ \gamma_d{}^q(0,\ t)]} \quad ; \qquad t \leq 0 \qquad \text{(A-65)}$$

where

$$\gamma_d{}^q(0,\ t) = \gamma_d{}^q(t)$$

$$= \ln \iiint\limits_{Y\ X\ X'} p(x)\ p(x')\ p(y|x)\ e^{t[d^q(x'y)\ -\ d^q(xy)]} dx'dxdy$$

$$t \leq 0 \qquad \text{(A-65a)}$$

Thus

$$\gamma_d(o,\ t) = \ln \iint\limits_{X\ Y} p(x)\ p(y|x)^{1-t} p(y|x')^t\ ; \quad t \leq 0 \qquad \text{(A-66)}$$

$\gamma_d(0,\ t)$ may be minimized by choosing a proper t. Differentiation with respect to t and setting the result equal to zero, yields

$$t = -\frac{1}{2} \qquad \text{(A-67)}$$

$$\gamma_d(0,-\tfrac{1}{2}) = \gamma_d(\tfrac{1}{2}) = \ln \iiint\limits_{Y\ X\ X'} p(x)\ p(x')\ p(y|x)^{\frac{1}{2}}\ p(y|x')^{\frac{1}{2}} \qquad \text{(A-68)}$$

or

$$\gamma_d(0, \tfrac{1}{2}) = \ln \int_Y [\int p(x)\; p(y|x)^{1/2} dx]^2 \qquad \text{(A-69)}$$

The insertion of Eq. A-67 into Eq. A-64 yields

$$P_e \leqslant e^{-n[E_d(0) - R]} \qquad \text{(A-70)}$$

where

$$E_d(0) = -\frac{1}{d}\, \gamma_d(0, \tfrac{1}{2}) \qquad \text{(A-71)}$$

From Eq. A-60 we have that, for $R = R_{crit}$, $s = \frac{1}{2}$, $t = -\frac{1}{2}$ and $r = 0$. Thus, by Eq. A-49

$$E_d(R)\Big|_{R_{crit}} = - R_{crit} + \frac{1}{d}\, \gamma_d(0, \tfrac{1}{2}) = E_d(0) - R_{crit} \qquad \text{(A-72)}$$

and the exponentials of Eq. A-70 and Eq. A-49 are indeed identical for $R = R_{crit}$.

It can also be shown that $dE_d(R)/dR\Big|_{R_{crit}} = -1$ so that the derivatives of the two exponents with respect to R are also the same at $R = R_{crit}$.

The average probability of error can therefore be bounded by

$$P_e \leqslant \begin{cases} e^{-n[E_d(0) - R]} & ; \; R \leqslant R_{crit} \\ 2e^{-n[E_d(R)]} & ; \; R_{crit} \leqslant R \leqslant I \end{cases} \qquad \text{(A-73a)}$$

where

$$E_d(0) - R_{crit} = E_d(R)\Big|_{R_{crit}} \qquad \text{(A-73b)}$$

and

$$\left.\frac{d[E_d(0) - R]}{dR}\right|_{R_{crit}} = \left.\frac{dE_d(R)}{dR}\right|_{R_{crit}} = -1 \qquad \text{(A-73c)}$$

A.4 Optimum Upper Bounds for the Average Probability of Error

The upper bound of Eq. A-73 may be optimized by choosing $p(x)$ such that, for a given rate R the exponent of the bound is minimized.

Let $d = 1$ and $m = n$. x^i is then identical with ξ^i, where ξ^i is the i^{th} input sample which is a member of the (continuous) one-dimensional space Ξ. y^i is identical with η^i, where η^i is the i^{th} output sample which is a member of the (continuous) one-dimensional space H.

It can be shown (Ref. (2), pp. 332-340) that there exists an optimum probability (density) $p(x) \equiv p(\xi)$ defined on Ξ that minimizes the upper bound to the average probability of error so that, for large n and for $R \geqslant R_{crit}$, it becomes exponentially equal to the lower bound on the probability of error.

The characteristics of many continuous physical channels, when quantized and thus converted into a discrete channel, are very close to the original ones if the quantization is fine enough. Thus, for such continuous channels there exists one random code with an optimum probability density $p(x) = p(\xi)$ that yields an exponent $E(R)$ which is equal to the exponent of the lower bound of the average probability of error, for n very large and for $R \geqslant R_{crit}$.

APPENDIX B

EVALUATION OF $E_d(0)$ OF SECTION 2.3, FOR CASES WHERE EITHER $A^2 d \ll 1$ OR $A^2 > 1$

In this appendix we shall evaluate lower bounds to the exponent $E_d(0)$ which is given by Eq. 2-132 and is equal to

$$E_d(0) = -\frac{1}{n} \ln \left\{ \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} \int_0^{\pi} e^{-\frac{dA^2}{2} \sin^2 \frac{\theta}{2}} \sin \theta^{d-2} d\theta \right\} \qquad (2\text{-}132)$$

When d = 2 we have

$$E_2(0) = -\frac{1}{2} \ln \left\{ \frac{1}{\pi} e^{-\frac{A^2}{2}} \int_0^{\pi} e^{\frac{A^2}{2} \cos \theta} d\theta \right\}$$

$$= -\frac{1}{2} \ln \left\{ e^{-\frac{A^2}{2}} \frac{1}{2\pi} \int_0^{2\pi} e^{\frac{A^2}{2} \cos \theta} d\theta \right\}$$

Thus

$$E_2(0) = -\frac{1}{2} \ln \left\{ e^{-\frac{A^2}{2}} I_0 \left(\frac{A^2}{2} \right) \right\} \qquad (B\text{-}1)$$

For $\frac{A^2}{2} \ll 1$ we have

$$E_2(0) \cong -\frac{1}{2} \ln e^{-\frac{A^2}{2}} \cong \frac{A^2}{4} \qquad (B\text{-}2)$$

For $\frac{A^2}{2} >> 1$ we have

$$E_2(0) \cong -\frac{1}{2} \ln \left\{ e^{-\frac{A^2}{2}} \frac{1}{\sqrt{2\pi \frac{A^2}{2}}} e^{+\frac{A^2}{2}} \right\}$$

$$\cong \frac{1}{4} \ln A^2 + \frac{1}{4} \ln \pi \cong \frac{1}{2} \ln (A^2)^{1/2} \tag{B-3}$$

Now, for d = 3 we have from Eq. 2-132

$$E_3(0) = -\frac{1}{3} \ln \left\{ \frac{1}{2} e^{-\frac{3}{4} A^2} \int_0^{\pi} e^{\frac{3}{4} A^2 \cos \theta} \sin \theta \, d\theta \right\}$$

$$= -\frac{1}{3} \ln \left\{ \frac{1}{2} e^{-\frac{3}{4} A^2} \frac{4}{3A^2} \int_{\pi}^{0} \frac{d}{d\theta} e^{\frac{3}{4} A^2 \cos \theta} d\theta \right\}$$

Thus

$$E_3(0) = -\frac{1}{3} \ln \left\{ e^{-\frac{3}{4} A^2} \frac{\mathrm{sh}\, \frac{3}{4} A^2}{\frac{3}{4} A^2} \right\} \tag{B-4}$$

For $\frac{3A^2}{4} << 1$ we have

$$E_3(0) \cong -\frac{1}{3} \ln e^{-\frac{3}{4} A^2} = \frac{A^2}{4} \tag{B-5}$$

For $\frac{3}{4} A^2 >> 1$ we have

$$E_3(0) \cong \frac{1}{3} \ln \frac{3}{4} A^2 = \frac{1}{3} \ln A^2 + \frac{1}{3} \ln \frac{3}{4}$$

$$\cong \frac{1}{3} \ln A^2 = \frac{1}{2} \ln (A^2)^{2/3} \tag{B-6}$$

In general, for an d $\geqslant$ 3, we have

$$E_d(0) = \frac{1}{4} A^2 \; ; \qquad A^2 d \ll 1 \tag{B-7}$$

$$E_d(0) \cong \frac{1}{2} \ln (A^2)^{\frac{d-1}{d}} = \frac{d-1}{d} E(0) \; ; \qquad A^2 \gg 1 \tag{B-8a}$$

where

$$E(0) = E_d(0)\Big|_{d=n \gg 1} \cong \frac{1}{2} \ln(A^2) \; ; \quad A^2 \gg 1 \tag{B-8b}$$

Proof: From Eq. 2-132 we have, for $A^2 d \ll 1$

$$E_d(0) \cong -\frac{1}{d} \ln\left\{ \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} (1 - \frac{dA^2}{4}) \int_0^{2\pi} \left(1 - \frac{dA^2}{4}\right) \cos\theta \quad \sin\theta^{d-2} \, d\theta \right\}$$

$$= -\frac{1}{d} \ln \left(1 - \frac{dA^2}{4}\right) \cong \frac{A^2}{4}$$

Thus

$$E_d(0) \cong \frac{A^2}{4} \; ; \qquad A^2 d \ll 1$$

Q.E.D.

We now proceed to prove Eqs. 2-135b and 2-135c. Let $x = \sin^2 \frac{\theta}{2}$ and insert x into Eq. 2-132. We then have

$$E_d(0) = -\frac{1}{d} \ln \left\{ \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{d-1}{2})} 2^{d-3} \, I \right\} \tag{B-9a}$$

where

$$I = \int_0^1 e^{-\frac{dA^2}{2} x} \, x^{\frac{d-3}{2}} (1-x)^{\frac{d-3}{2}} \, dx \tag{B-9b}$$

Now

$$1 - x \leq e^{-x} \tag{B-10}$$

Inserting Eq. B-10 into Eq. B-9b yields

$$I \leq \int_0^1 e^{-\left[\frac{A^2 d}{d-3} + 1 \right] \frac{d-3}{2} x} \; x^{\frac{d-3}{2}} \, dx$$

$$= \frac{\Gamma\left(\frac{d-1}{2} \right)}{\left[\frac{d-3}{2} \left(\frac{A^2 d}{d-3} + 1 \right) \right]^{\frac{d-1}{2}}} \tag{B-11}$$

Inserting Eq. B-11 into B-9a yields

$$E_d(0) \geq - \frac{1}{d} \ln \left\{ \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{d}{2} \right)}{\left(\frac{d}{2} \right)^{\frac{d-1}{2}}} \; 2^{d-3} + \frac{1}{d} \ln \left[\frac{d-3}{d} + A^2 \right]^{\frac{d-1}{2}} \right\}$$

The first term on the r.h.s. of Ineq. 2-139 is bounded by

$$- \frac{1}{d} \ln \; \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{d}{2} \right)}{\left(\frac{d}{2} \right)^{\frac{d-1}{2}}} \; 2^{d-3} \leq - \frac{1}{d} \ln 2^{d-3}$$

$$= - \frac{d-3}{d} \ln 2$$

Thus, for $A^2 >> 1$ we have

$$E_d(0) \gtrsim \frac{1}{d} \ln \left(A^2 \right)^{\frac{d-1}{2}} = \frac{d-1}{d} \frac{1}{2} \ln A^2 \tag{B-12}$$

Now, let

$$E(0) = E_{d=n}(0)\Big|_{n>>1} \qquad \text{(B-13)}$$

Inserting d = n >> 1 into Ineq. B-12 yields

$$E(0) \gtrsim \frac{1}{2} \ln A^2$$

From the convexity of the exponent E(R), when plotted as a function of R we have

$$E(0) \leq C \qquad \text{(B-14)}$$

where C is the channel capacity and is given by (Ref. 1)

$$C \cong \frac{1}{2} \ln A^2 ; \qquad A^2 >> 1$$

Thus, by Ineqs. B-13 and B-14

$$E(0) \cong C \cong \frac{1}{2} \ln A^2 ; \qquad A^2 >> 1 \qquad \text{(B-15)}$$

Inserting Eq. B-15 into Ineq. B-12 yields

$$E_d(0) \cong \frac{d-1}{d} \frac{1}{2} \ln A^2 \cong \frac{d-1}{d} E(0)$$

APPENDIX C

MOMENT GENERATING FUNCTIONS OF QUANTIZED GAUSSIAN VARIABLES

(AFTER B. WIDROW, REF. (8))

A quantizer is defined as a non-linear operator having the input-output relation shown in Figure 10c. An input lying somewhere within a quantization "box" of width q will yield an output corresponding to the center of the box (i.e., the output is rounded off to the center of the box).

Let the input z be a random variable. The probability density distribution of z, p(z), is given.

The moment generating function (m.g.f.) of the input signal is therefore

$$g(s) = \int_{Z} p(z)\, e^{-sz} dz \qquad \text{(C-1)}$$

Our attention is devoted to the m.g.f. of the quantized signal z^q, given by

$$g^q(s) = \int_{Z^q} p(z^q)\, e^{-sz^q}\, dz^q \qquad \text{(C-2)}$$

where $p(z^q)$ is the probability density of the output of the quantizer, z^q. $p(z^q)$ consists of a series of impulses. Each impulse must have an area equal to the area under the probability density p(z) within the bound of the "box" of width q, in which the impulse is centered. Thus, the probability density $p(z^q)$ of the quantizer output consists of "area samples" of the input probability density p(z). The quantizer may be thought of as an area sampler acting upon the "signal," the probability density p(z).

Thus, $p(z^q)$ may be constructed by sampling the difference $\phi(z + \frac{q}{2}) - \phi(z - \frac{q}{2})$ where $\phi(z)$ is the input probability distribution given by

$$\emptyset(z) = \int_{-\infty}^{z} p(z)\, dz \tag{C-3}$$

This is equivalent to first modifying p(z) by a linear "filter" whose transfer function is

$$\frac{e^{\frac{sq}{2}} - e^{-\frac{sq}{2}}}{s} = q\,\frac{\text{sh}\,\frac{qs}{2}}{\frac{qs}{2}} \tag{C-4}$$

and then impulse-modulating it to give $p(z^q)$.

Using "△" notation to indicate sampling we get

$$g^q(s) = \left[\, g(s)\, \frac{\text{sh}\,\frac{qs}{2}}{\frac{qs}{2}} \,\right]^{\triangle} = F^{\triangle}(s) \tag{C-5a}$$

where

$$F(s) = g(s)\; q\, \frac{\text{sh}\,\frac{qs}{2}}{q\,\frac{s}{2}} \tag{C-5b}$$

Now, let the function F(s) be the transform of a function f(z)

$$F(s) = \int_{Z} f(z)\, e^{-sz} dz \tag{C-6}$$

Then

$$F^{\triangle}(s) = \int_{Z} f^{\triangle}(z)\, e^{-sz}\, dz \tag{C-7}$$

where f (z) is the sampled version of f(z).

Thus

$$f^{\triangle}(z) = f(z)\ c(z) \tag{C-8}$$

where c(z) is a train of impulses, q amplitude-units apart. A Fourier analysis may be made of the impulse train c(z). The form of the exponential Fourier series will be

$$c(z) = \frac{1}{q} \sum_{k=-\infty}^{\infty} e^{ik\Omega z} \ ; \qquad \Omega = \frac{2\pi}{q} \tag{C-9}$$

Inserting Eq. C-9 into Eq. C-7 yields

$$F^{\triangle}(s) = \frac{1}{q} \sum_{k=-\infty}^{\infty} F(s - ik\Omega) \tag{C-10}$$

Inserting Eq. C-10 into Eq. C-5 yields

$$g^{q}(s) = \sum_{k=-\infty}^{\infty} g(s - ik\Omega) \frac{\operatorname{sh} [q \frac{(s-ik\Omega)}{2}}{q \frac{(s-ik\Omega)}{2}} \tag{C-11}$$

Now, if the input is a Gaussian variable governed by the probability density

$$p(z|x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(z-x)^2}{2\sigma^2}} \tag{C-12}$$

it can be shown that

$$g^{q}(s) \cong g(s) \frac{\operatorname{sh} \frac{qs}{2}}{\frac{qs}{2}} \ ; \qquad q < 2\sigma \tag{C-13}$$

where

$$g(s) = e^{\frac{s^2\sigma^2}{2} + xs} \tag{C-14}$$

Let z and z' be two random Gaussian variables governed by the following probability density

$$p(z, z') = \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left[\frac{-(z-\bar{z})^2 + 2\rho(z-\bar{z})(z'-\bar{z}') - (z'-\bar{z}')^2}{2(1-\rho^2)} \right] \qquad \text{(C-15a)}$$

where

$$\rho = \overline{(z - \bar{z})\,(z' - \bar{z}')} \qquad \text{(C-15b)}$$

Let the corresponding m.g.f. be given by

$$g(r, t) = \iint_{z'\,z} p(z, z')\, e^{rz + tz'} dz dz' \qquad \text{(C-16)}$$

Now let z and z' be quantized by the quantizer of Figure 10c, to yield z^q and z'^q. Thus

$$g^q(r, t) = \iint_{z^q z'^q} p(z^q, z'^q)\, e^{-rz^q - tz'^q} dz^q dz'^q \qquad \text{(C-17)}$$

It can then be shown (as has been shown in the derivation of Eq. (C-13)) that

$$g^q(r, t) = g(r, t)\, \frac{\operatorname{sh}\frac{qr}{2}}{\frac{qr}{2}}\, \frac{\operatorname{sh}\frac{qt}{2}}{\frac{qt}{2}} ; \qquad q < 2(1 - \rho^2) \qquad \text{(C-18)}$$

Also if z = z' (ρ = 1) we have

$$g^q(r, t) = g(r, t)\, \frac{\operatorname{sh}\frac{q}{2}(r + t)}{\frac{q}{2}(r + t)} ; \qquad q < 2 \qquad \text{(C-19)}$$

Now, if z = - z' (ρ = -1) we have

$$g^q(r, t) = g(r, t) \frac{\operatorname{sh} \frac{q}{2} (r - t)}{\frac{q}{2} (r - t)} ; \qquad q < 2 \tag{C-20}$$

We now proceed to derive upper bounds to g(s) and g(r, t) to be used whenever the quantization grain q is large, so that Eqs. C-13, C-17, C-18 and C-19 are not valid any more.

Let $z^q = z + n_q(z)$ where $n_q(z)$ is the "quantization noise." Thus, by Eq. C-2

$$g^q(s) = \int_Z p(z)\, e^{-s(z + n_q(z))}\, dz$$

Now, $|n_q(z)| \leqslant \frac{q}{2}$. Therefore

$$g^q(s) \leqslant \int_Z p(z)\, e^{-sy}\, dz\, e^{|s \frac{q}{2}|}$$

$$= g(s)\, e^{|s \frac{q}{2}|} \tag{C-21}$$

In the same way, let

$$z^q = z + n_q(z) ; \qquad z'^q = z' + n_q(z')$$

Thus, by Eq. C-17

$$g^q(r, t) = \int_Z \int_{Z'} p(z, z')\, e^{-r(z + n_q(z)) - t(z' + n_q(z'))}\, dz' dz$$

Now

$$|n_q(z)| \leqslant \frac{q}{2} ; \qquad |n_q(z')| \leqslant \frac{q}{2}$$

Thus

$$g^{q}(r,\ t) \leqslant \int_{Z}\int_{Z'} p(z,\ z')\ e^{-rz\ -\ tz'} dz'dz\ e^{\left|r\ \frac{q}{2}\right| + \left|t\ \frac{q}{2}\right|}$$

$$= g(r,\ t)\ e^{\left|r\ \frac{q}{2}\right| + \left|t\ \frac{q}{2}\right|} \qquad \text{(C-22)}$$

BIOGRAPHICAL NOTE

Jacob Ziv was born in Tiberias, Israel, on November 27, 1931. He attended primary school in Ra'anana, Israel, secondary school in Tel-Aviv, Israel, and received his engineering education at the Technion -- Israel Institute of Technology, Haifa, Israel. While a graduate student at Technion, Mr. Ziv was a Teaching Assistant in Electrical Engineering. Since September, 1955, he has been in the employ of the Scientific Department, Israel Ministry of Defence, working in the fields of transistor circuitry and communications.

Mr. Ziv has been attending the Graduate School of the Massachusetts Institute of Technology as a full-time student since September, 1959, on a fellowship granted by the Scientific Department, Israel Ministry of Defence.

Mr. Ziv is a member of the Institute of Radio Engineers.

Married to Shoshana (née Salomon), they have a son Noam Abraham (1961).

REFERENCES

(1) C.E. Shannon, "Probability of Error for Optimal Codes in a Gaussian Channel," The Bell System Tech. Journal, May, 1959.

(2) R.M. Fano, Transmission of Information, M.I.T. Press, John Wiley and Sons, Inc., Chapter 9, Equations: 9.116, 9.247, 9.248, 9.249.

(3) J.L. Kelly, "A Class of Codes for Signaling on a Noisy Continuous Channel," I.R.E. Transactions on Information Theory, March, 1960.

(4) J.M. Wozencraft, "Sequential Decoding for Reliable Communications," M.I.T., R.L.E. Tech. Report 325, August, 1957

(5) B. Reiffen, "Sequential Encoding and Decoding for the Discrete Memoryless Channel," M.I.T., R.L.E. Tech. Report 379, August, 1960.

(6) B. Reiffen, "Determination of R cutoff for the Asymmetric Channel," Lincoln Laboratory Report No. 25 G-0021, 1961.

(7) J. Ziv, "Encoding for Time-discrete Memoryless Channels," R.L.E. Quarterly Report, April, 1961.

(8) B. Widrow, "A Study of Rough Amplitude Quantization by Means of Nyquist Sampling," I.R.E. Transactions of P.G.C.T., Volume CT-3, No. 4, December, 1954.

(9) F.J. Bloom et al, "Improvement of Binary Transmission by Null-Zone Reception," I.R.E. Proceedings, Volume No. 7, July, 1957

(10) R.M. Gallager, "Low-Density Parity Check Codes," Sc.D. Thesis, Department of Electrical Engineering, M.I.T., September, 1960.

(11) P. Elias, "Coding for Time-Varying Channels," paper presented at URSI General Assembly, London, September 6-14, 1960.