APPLICATION OF STATISTICAL DECISION THEORY TO THE DETECTION OF RADIOTELEGRAPH SIGNALS

BY

DAVID JOEL GOODMAN

Department of Electrical Engineering
Imperial College of Science and Technology
University of London

May 1967

A thesis
submitted for the degree of
Doctor of Philosophy

ABSTRACT

Because a radiotelegraph signal detector makes decisions in the face of uncertain evidence, its operation may be viewed in the context of statistical decision theory. An important determinant of a detector's structure and performance is its fidelity criterion. Although a unit-cost (minimum-error) criterion is most common, the serious effects of teleprinter-control-symbol errors on message legibility suggest a more complicated cost function to be appropriate.

A telegraph system has been simulated on a digital computer in order to demonstrate the character of messages produced by four detectors. The outputs of the three symbol detectors are clearly more legible than those of the binary-element detector. Among the symbol detectors, the one whose cost matrix takes into account the effects of control-symbol errors generates the most useful messages but its extensive computational requirements do not, for most practical purposes, appear to be justified by the improvements offered.

From a practical point of view, the most promising symbol detector is the one which operates accord-

ing to the maximum-likelihood decision rule. It has a particularly simple configuration when symbols are represented by the four-out-of-seven code and its performance characteristics have been calculated under eight propagation conditions. A comparison with the performance of a conventional binary detector shows the introduction of a maximum-likelihood detector to be justified and the effects of different fading conditions indicate the merits of certain modifications of message-multiplex arrangements.

In two-way feedback systems, the use of maximum-likelihood detectors would admit the introduction of a predecision feedback scheme which promises to control the error rate over a wider range of channel conditions than the method of ARQ signalling.

TABLE OF CONTENTS

CHAPTER 1: THE SCOPE OF THE THESIS

The introduction of improved detection techniques would serve the international high-frequency radio-telegraph network in two respects. Not only would the quality of messages received over presently-used channels be enhanced, but additional channels whose transmission conditions are too severe to permit reliable detection by existing methods, could be added to the network. Channel congestion in the hf spectrum (25)* would thereby be reduced.

Although it is generally appreciated that better results could be achieved if telegraph symbols were detected directly, conventional detectors are designed to operate independently on the binary code elements which represent the symbols. (The nomenclature is illustrated in Figure 1.1.) A principal obstacle to the implementation of symbol-detection techniques has been the complexity of the equipment required for the storage and processing of information. In recent years, however, the cost of computational equipment has been reduced and in many cases extensive computational facilities have been installed in telegraph networks for message-switching purposes. It is thus becoming more

---

* Numbers in parentheses refer to bibliography entries.

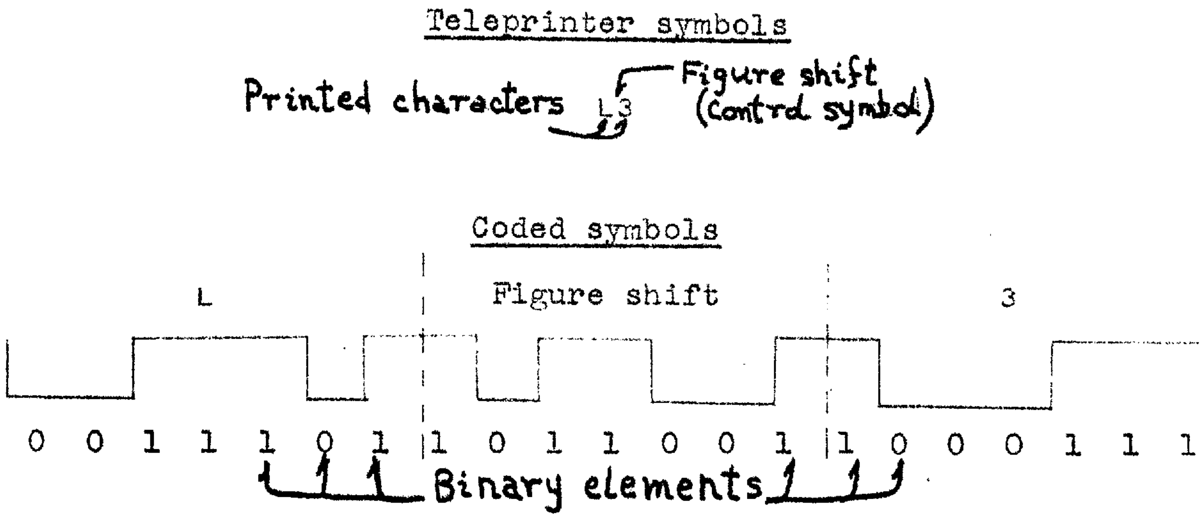


FIGURE 1.1 Guide to nomenclature

feasible to implement complicated data processing operations at a telegraph station, which suggests that an investigation of sophisticated detection methods would now be of practical value to the design engineer. This thesis offers an operational description of various detectors of telegraph symbols and investigates the performance improvements which would accrue from their introduction to existing telegraph systems.

1.1 The role of decision theory

The purpose of a telegraph detector is to produce printed messages which are as similar as possible to the messages introduced at the transmitting end of a telegraph link. On the basis of the uncertain evidence contained in a randomly perturbed waveform, the detector must decide which symbols to include in an output message. The principles of decision making in the face of uncertain evidence about an objective phenomenon constitute the subject matter of the mathematical theory of statistical decision functions. A decision function is a rule for making decisions which, on the average, result in minimum cost to the decision maker. The cost of a decision is a numerical measure of the consequences of its correctness; in telegraphy the cost of deciding upon ~~generating~~ an output symbol depends on the identity of the transmitted symbol. If the two are identical, the cost is zero. The set of costs for all possible transmitted and detected symbolts comprises a measure of the dissimilarity of transmitted and received messages. Minimizing the average cost of detection coincides with maximizing the similarity of the utilities of the two messages.

The design objective of conventional telegraph detectors is the maximization of the probability that a detected symbol is identical to the transmitted symbol. This minimum-error fidelity criterion corresponds in decision-theory terminology to a unit-cost matrix, a cost function which assigns identical costs to all incorrect decisions. Although this criterion facilitates detector design, it does not properly reflect the effects upon message quality of the various possible decisions. The electromechanical teleprinter, which is the terminal instrument of much of the world's telegraphy, generates two different types of symbols: printed characters and control symbols. Errors which involve control symbols have, in general, a much more serious effect upon the legibility of a received message than errors which cause one printed character to be generated in place of another. The non-uniform effects of the various detection errors imply a more complicated cost function than the unit-cost matrix. Statistical decision theory may be applied to the problem of designing a detector which operates under the constraint of any cost function.

1.2 Measurement and computation in detection.

A decision rule specifies the information processing operations of a telegraph detector. At a receiving station, there are two types of information available relevant to the identity of transmitted symbols. There is prior information relating to the telegraph code and statistical constraints upon a source message; there is the statistical information contained in the received signal. In order to apply the decision rule, the detector must extract from the received signal a representation of this latter type of information.

This representation takes the form of a set of data which comprise the uncertain evidence of the decision problem. The signal-processing operations involved in acquiring this data constitute the measurement aspect of detection. The computational aspect includes the information processing operations on the extracted data and the prior information. The computational role of a detector is defined by its decision rule.

The decision rules of practical detectors relate primarily to the binary code elements which represent the telegraph symbols. The detectors are designed to make an independent decision about the likely identity of each of the transmitted code elements. Although the computations performed at a binary detector are of an elementary nature, the signal measurements are often achieved by means of sophisticated techniques which apply many of the principles of statistical communication theory (74). The problems of aerial design, diversity reception and optimum filtering are the traditional subject matter of analytic and design studies in the field of telegraph detection (4, 24, 36). The principal contribution of this thesis is an investigation of the advantages of augmenting highly refined measurement techniques with the relatively complex computational procedures demanded by detectors of telegraph symbols.

1.3 The telegraph system.

Although certain of the detection principles introduced in subsequent chapters are relevant to digital communication systems in general, they are presented from the point of view of their applicability to the world's high frequency radiotelegraph network. Except for the variable-speed system discussed in Chapter 7, the detectors formulated in this thesis are assumed to operate on signals emanating from the typical transmitting station whose configuration is specified in Chapter 3.

The telegraph messages to be communicated are produced by means of a teleprinter with 27 printed characters and 4 control symbols. Each symbol is represented by a fixed number of binary code elements which are transformed into a sequence of sinusoidal signal elements of limited duration. The telegraph code assumed in most of the numerical calculations is the synchronous seven-unit fixed-ratio code in which each symbol is represented by a sequence of binary elements consisting of four "ones" and three "zeros". For purposes of comparison, performance measures of a system using the five-element start/stop code are given in a few instances. In many systems, several messages are transmitted simultaneously by means of frequency-division and time-division multiplex arrangements. In

Chapter 3, the concept of a generalized multiplex organization is introduced and the advantages of specific configurations are investigated in Chapter 6.

The transmitted signals are subjected, in the channel, to the additive perturbation of white gaussian noise. The multiplicative effects which are investigated are constant attenuation and four types of Rayleigh-distributed amplitude fading. A specific fading condition is characterized by the correlations among the amplitudes of the seven elements of each symbol and by the correlation between the amplitudes of the two sinusoids assigned to a single element. With respect to a conventional multiplex arrangement these two characteristics of the fading may be identified with time and frequency correlations of the amplitude stochastic process.

1.4 The detectors.

Various telegraph detectors are considered in this thesis with the aim of comparing their decision rules, i.e., computational procedures. To ensure that only the computational aspects of the detectors account for their performance differences, all of the detectors are designed to operate on the same set of signal measurements. Two data relevant to each signal element are extracted from a received waveform by means of a pair of matched filters. It is assumed that these filters operate with precise information regarding the frequencies and time spans of the signal elements but no knowledge of their individual amplitudes. The conditions at the receiver of precise information of the phase of each element, i.e., coherent reception, and complete uncertainty regarding the phase, i.e., non-coherent reception, are both considered.

An output symbol of a given detector depends upon the 14 measurements (in the case of the seven-unit code representation) extracted by the matched filters and upon the detector's decision rule. In Chapter 5, three symbol detectors with varying computational requirements are formulated. The most complex is the minimum-risk detector, designed under the constraint of a cost function representing the different consequences of the various types of detection errors. This detector operates with prior information of the statistical constraints upon source messages. This information is also required by the minimum-error detector whose fidelity criterion is the unit-cost matrix. The simplest of the symbol detectors, the maximum-likelihood detector, operates under a unit-cost matrix and with no prior information relating to message statistics. When symbols are represented by the seven-unit fixed-ratio code, this device has a particularly simple structure. It is thus the most practical of the symbol detectors and its properties are investigated in detail. In order to provide a basis for comparing the symbol-detection techniques with conventional methods, a corresponding characteristic of a binary detector using the same signal measurements accompanies each characterization of a symbol detector.

1.5 Feedback communication systems

If a feedback channel from the receiving station to the transmitter of telegraph messages exists, the attainable performance of the link is better than that of a strictly one-way system. In two-way telegraph systems, message information transmitted in each direction may be accompanied by control information related to the signals transmitted in the opposite direction. In this event each of the two one-way channels acts as a feedback link for the other. A practical mechanism for using this feedback facility is the technique of ARQ signalling which is based upon the redundancy of the seven-unit code. The ARQ system uses the control information transmitted over a return path to instigate the retransmission of symbols which are detected as erasures, i.e., redundant code sequences, by a binary detector. When a symbol detector is employed, no erasures are generated and information transmitted over a feedback path may be used to control the behaviour of the transmitter in a manner which does not necessarily involve retransmissions. In particular, the feedback information may modify the time duration of the signal elements and thus the system transmission speed. This variable-speed operation offers error-control opportunities which in some conditions are superior to those provided by the ARQ mechanism. The two types of feedback system are compared and the possibility of implementing a variable-speed mechanism is discussed.

1.6 Organization of the thesis.

The results presented in this thesis derive from two methods of investigating the effects of introducing symbol-detection techniques to a practical telegraph system. The first investigation is the computer simulation of Chapter 5 which demonstrates the messages produced by four detectors when a specific text is transmitted. Of the simulated detectors, the most promising from a practical viewpoint is the maximum-likelihood detector. The possibility of implementing this device is investigated in Chapter 6 by means of a presentation and critical discussion of the detector's performance characteristics. While Chapter 6 is confined to a consideration of one-way telegraph systems, Chapter 7 discusses the applicability of a maximum-likelihood detector to a two-way system incorporating feedback signalling facilities. In Chapters 5, 6,and 7, the performance of each symbol detector is compared with that of a conventional binary detector. The maximum-likelihood error probabilities presented in Chapter 6 have been calculated from formulas derived in Chapter 8, which contains a mathematical analysis of the error immunity of fixed-ratio codes.

The earlier chapters of the thesis review the established theory and practice which serve as the basis of the original material. Chapter 2 surveys the principles of statistical decision theory and presents the Bayes decision rule as a computational algorithim. The detectors of the later chapters operate according to this rule. The nature of the signals available to the detectors and their relation to the transmitted messages depend on the specification of the elements of a transmitting station and the model of the channel presented in Chapter 3. The established methods of signal measurement and binary detection are described in Chapter 4.

1.7 Acknowledgements

This thesis is the result of a research project which was supported by a Government grant. The work was performed under the helpful supervision of Professor Colin Cherry and it was assisted by the advice and cooperation of Mr A.D.V. Ridlington and Mr J. Heyes. The author expresses his thanks to his colleague at Imperial College, Brian Travers, whose advice, assistance and criticism have been invaluable.

CHAPTER 2: STATISTICAL DECISION THEORY.

Kotel'nikov, in presenting his doctoral dissertation in Moscow in 1947 was the earliest writer to recognise the relevance of a statistical-decision-theory approach to signal detection. He proposed a minimum-error fidelity criterion for digital communications and on the basis of a mathematical analysis of random noise (40) calculated the optimum performance characteristics of various systems. Kotel'nikov's work was not generally available in the West until 1960, when an English translation of his dissertation was published (27).

The development of detection theory in Britain and the United States in the 1950's was largely based on the statistical formulation of Woodward and Davies (72, 73) whose approach to the problem was suggested by Shannon's mathematical theory of communication (48, 49). Woodward in his book (73), describes the essential function of a detector as the extraction of wanted and the rejection of unwanted information from a received signal. He equates this procedure with the derivation of a statistically sufficient (71) representation of the signal. The procedure of information extraction is identified in Section 1.3 of this thesis as the measurement aspect of detection; Woodward and Davies demonstrate that the measurements required by their theory may be obtained by means of a correlation operation. In the case of binary signalling, this operation is performed on a received signal by a pair of filters matched (59,56) to the binary waveforms. Helstrom (23) extended the theory of binary detection by deriving the optimum performance characteristics associated with certain classes of modulation waveforms. Performance characteristics of binary detectors of fading radio signals were presented by Law (28) and Turin (55).

Concurrent with these developments in binary-detection theory was the formulation by Middleton and Van Meter (58,33) of a general reception problem as an application of Wald's theory of statistical decision functions (69). This formulation merged the current trends in statistical inference and communication theory. In their general expositions, Middleton and Van Meter considered two classes of reception problem which they call detection and extraction. In the context of decision theory these problems come under the headings of null-hypothesis testing and parameter estimation, respectively; in communication theory they are referred to as detection of a signal in noise and optimum filtering. A third category, which includes telegraph-symbol detection, is multiple-alternative detection (or multiple-hypothesis testing [66]). It is treated in a separate paper (34) by Middleton as a special case of the general formulation.

This chapter presents the aspects of decision theory which pertain directly to the telegraph-detection problem solved in Chapter 5. In the intervening two chapters, the elements of a telegraph system and the measurement aspect of detection are formulated in order to demonstrate the quantitative relationships used in applying decision theory to telegraphy. In first presenting the principles of decision theory and then demonstrating their relevance to a communication problem, this thesis follows a plan similar to that adopted in the papers of Middleton and Van Meter. The present work differs from these papers by being primarily concerned with practical aspects of a specific system; the formulation of decision theory is more restricted but the practical applications are presented in more detail.

In the following section, the development of decision theory from a generalized problem in statistical inference and the theory's important elements are discussed heuristically. The relevant concepts are defined rigorously in Section 2.2.

2.1 Background of the theory.

The mathematical theory of statistical decision functions was developed by Wald between 1939 (66) and 1950 (69). His contribution to the subject began with a formulation of a generalized statistical-inference problem. This formulation included as special cases the three traditional problems of statistical inference which had previously been considered separately: hypothesis testing, point parameter estimation and confidence-interval estimation. In each of these problems, the statistician seeks to test the data generated under the constraint of a probability distribution in order to classify the parameters of the distribution. An hypothesis test assigns the parameters to one of two pre-specified categories; (i.e., it makes a binary decision); a point estimate is a numerical value assigned to a continuously variable parameter; and a confidence-interval estimate consists of two numbers defining an interval which is likely to include the unknown parameter. The success of a particular inference depends on the true values of the parameters being classified. Average success is measured by two error probabilities in the case of hypothesis testing, by the mean square error of a point estimate, or by the length of a confidence interval together with its associated probability of error.*

Wald's formulation broadened the scope of statistical inference in two respects: it admitted more types of parameter classification and it admitted terminal statements which do not refer directly to the parameters of a distribution (for example - a statement relating to the independence of two marginal distributions). The formulation unified the subject by introducing a single measure of success - the cost function, which depends on the decision (i.e., the final statement about the distribution) and the true state of the distribution. The average cost is the risk, a function of the true distribution and the rule for selecting a decision. The solution to the generalized inference problem is the specification of a decision rule, a method of examining data generated by the distribution in order to come to a decision.

The cost function may be defined in any manner whatsoever prior to the solution of the problem. It is the fidelity criterion of the problem expressing the relative importance of the various possible outcomes. The cost function may be considered a contraint imposed on the statistician by his client, the party that bears the consequences of correct or incorrect

---

*A brief but valuable introduction to the principles of inference is given by Wald in the published version of a series of tutorial lectures given in 1942 (67). A more rigorous treatment is found in (71).

decisions. The statistician must find a decision rule consistent with the client's set of values. Given a cost function, a set of possible distribution functions and a set of decisions, the statistician first seeks to determine the set of admissible decision rules. A rule is admissible if there is no other rule with lower risks for all possible distributions. The final rule is selected from the set of admissible ones according to the configurations of the admissible risk functions. Figure 2.1 shows three such functions plotted against

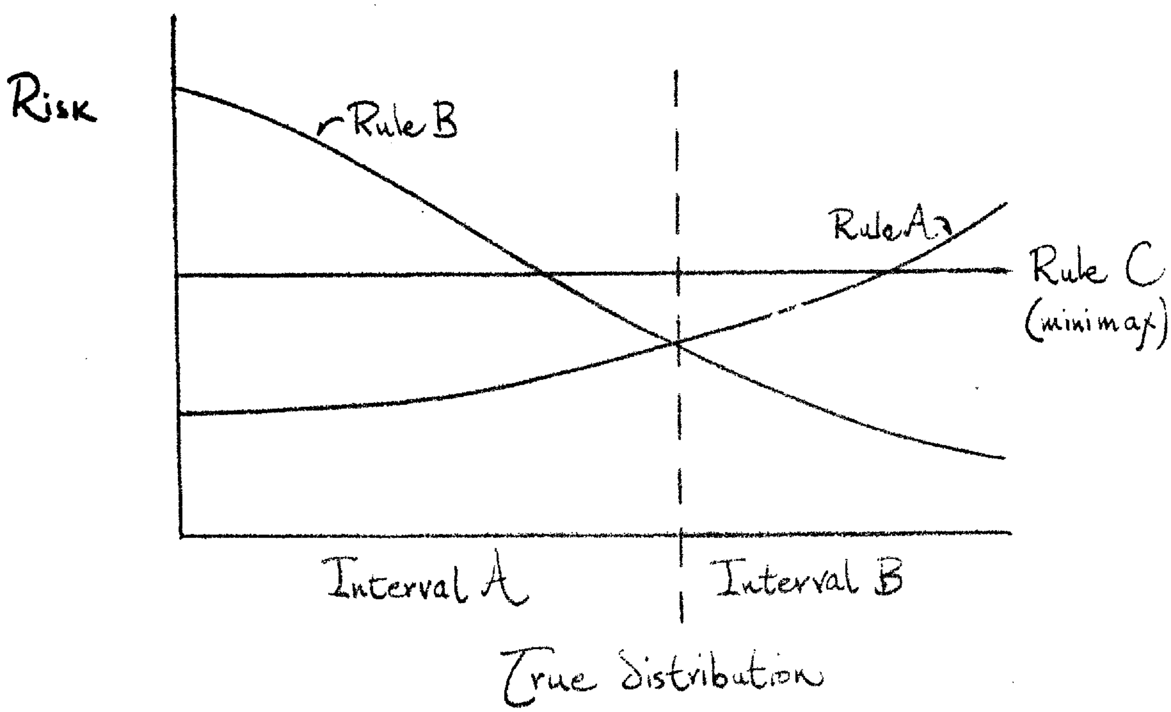


FIGURE 2.1 Risk functions of three decision rules.

the possible distributions which are represented as points on the abscissa line segment. If only these three rules are admissible, rule A is desirable if the distribution is likely to be in interval A (where rule A has the lowest risk), and similarly rule B is preferable if the distribution is likely to fall in interval B. If, however, the true distribution is equally likely to have any one of its possible forms, rule C may well be the most desirable of the three. If the prior probability function of the possible distributions is known, a rigorous procedure exists for selecting one from the set of admissible decision rules. If the prior probabilities are unknown the statistician often refers to the beliefs or temperament of his client. The beliefs may lead to the formulation of a "subjective" prior probability function; the temperament may suggest the choice of a more or less conservative rule. Rule C is the most conservative of the three shown in that its risk is minimal if it is assumed that for any decision rule, the least favourable distribution will prevail. Rule C is the minimax rule. Its maximum risk is minimal.

The foregoing discussion suggests that a wide range of situations may be encompassed by Wald's formulation. By identifying both the objective aspects of a problem and the role of a person's set of values, his temperament and beliefs, decision theory has served as a vehicle for the application of mathematical logic to processes which were previously considered to be entirely subjective. Examples of two important but diverse areas of application of decision theory are business decisions (46) and theories of human perception (53). Although expositions of communication applications have concentrated on analysis of physical phenomena, the opportunity exists to incorporate into the cost function a consideration of the effects of detected symbols on the legibility of a message by its human recipient. This opportunity is investigated in Chapter 5.

The present treatment of the inference problem has thus far stated the importance of finding admissible decision rules and indicated the factors which enter into a choice among these rules. Wald's 1950 treatise solves the problem he formulated in 1939 by demonstrating the principles governing the existence and character of admissible decision rules. In particular, he proved in the complete class theorem that the class of admissible decision rules is identical to the class of Bayes decision rules. In developing his proof of this identity, Wald recognised the similarity of the inference problem to the mathematical model of a zero-sum two-person game (9).. He was then able to apply theorems of game theory to the proof of the complete class theorem.

In addition to solving the problems posed by the 1939 formulation, the decision theory presented by Wald in 1950 incorporates another statistical discipline, that of sequential analysis (68). In a sequential problem, an examination of the measured data leads to one of two types of decision: a terminal decision or a decision to continue experimentation. The decision procedure ends with a terminal decision, the type treated in the non-sequential problem; each terminal decision carries a cost which depends on the true state of the distribution function of the data. The sequential problem offers the opportunity to extend the decision procedure through the acquisition of additional data. When this data is analysed, the procedure may be further iterated (by a decision to continue experimentation) or terminated with a terminal decision. The additional data is observed at a cost which must be added to the eventual cost of the terminal decision. Wald presents the principles of choosing a decision procedure which minimizes the average total cost. The telegraph detectors formulated in Chapter 5 are based on a non-sequential decision model. The signal relating to each symbol of a source message is transmitted only once. The scope of this work may be extended, by means of an application of sequential decision theory, to the design of detectors for a feedback system based on repetitions.

In many treatments of decision theory, the phenomenon which determines the cost of a particular terminal decision is referred to as the "state of nature" rather than the true distribution of the data. Although the newer term has a more general connotation, the two are indistinguishable operationally. Besides incorporating sequential analysis, decision theory has a broader scope than the generalized inference problem in that it admits terminal decisions which are not explicit statements about the state of nature. The cost function expresses the only essential relationship between the states of nature and the possible decisions. Developments in decision theory since 1950 have been mainly concerned with methodologies related to specific applications (29, 30). All of the newer work leans heavily on Wald for its mathematical content.

2.2 The formal decision problem

The remainder of this chapter consists of a formal presentation of the decision theory problem applied in Chapter 5 to telegraph detection. In anticipation of this application, the scope of the problem is limited to a non-sequential situation in which the set of decisions and the set of states of nature are both finite.

2.2.1 Elements of the problem

Statistical decision theory treats the problem of decision making in the face of uncertain evidence about a state of nature. Application of the theory leads to a rule for choosing decisions which, on the average, result in minimum costs. These statements suggest six elements of a decision problem:

1) A set of possible states of nature

$X = X_1, X_2, \ldots, X_M$

2) A set of possible decisions

$Y = Y_1, Y_2, \ldots, Y_N$

3) The decisions and the states of nature are related by the MxN cost matrix [c]. The elements of [c], c_{ij}, express the desirability of various decisions; the lower the value of c_{ij}, the more desirable the choice of Y_j when X_i is the state of nature.

4) A set of data

$\overline{\alpha} = \alpha_1, \alpha_2, \ldots \alpha_m$

which comprises the uncertain evidence about X_i, the true state of nature. $\overline{\alpha}$ may be considered a point in the multidimensional data space A.

5) The relation of $\overline{\alpha}$ to each possible state of nature is given by the set of conditional probability density functions defined on the data space A:

$$f_X(\overline{\alpha}) = f_1(\overline{\alpha}), f_2(\overline{\alpha}), \ldots, f_M(\overline{\alpha}).$$

Each of these functions is conditional on a member of X. For a given set of data, the set of numbers $f_X(\overline{\alpha})$ may be considered as a function defined on X, known as the likelihood function.

6) A class of decision functions, Δ, with members δ representing all possible solutions to the problem. Each decision function consists of the N numbers

$$\delta(Y_1,\overline{\alpha}),\ \delta(Y_2,\overline{\alpha}),\ \ldots,\ \delta(Y_N,\overline{\alpha})$$

where $\delta(Y_j,\overline{\alpha})$ is the probability of deciding Y_j when the data $\overline{\alpha}$ are observed. In telegraphy it is generally the case that for each $\overline{\alpha}$

$$\begin{aligned} &\delta(Y_{j^*},\overline{\alpha}) = 1 \\ \text{and}\quad &\delta(Y_j,\overline{\alpha}) = 0 \quad \text{for all } j \neq j^* \end{aligned} \qquad (2.1)$$

Each set of data is thus associated with a unique decision. A decision function which conforms to Equation 2.1 is called a pure decision rule. A mixed rule is one for which $\delta(Y_j,\overline{\alpha}) > 0$ for more than one value of j and a given $\overline{\alpha}$. Under such a rule each occurrence of $\overline{\alpha}$ results in a decision generated at random under the constraint of $\delta(Y_j,\overline{\alpha})$. When mixed rules are included in a mathematical analysis of decision theory, the set of all admissible decision rules is found to have properties essential to the proof of important theorems. A mixed rule is an important stategy in a game where it is imperative that an opponent be unable to predict a player's action in every possible situation.

The solution to a specific problem is the adoption of a decision function governing the choice of a decision for every possible set of data $\overline{\alpha}$. Table 2.1 lists the elements of the decision problem and associates them with the corresponding elements of a telegraph system.

Table 2.1

Decision theory term	Notation	Telegraphy term
1) State of nature	X_i	Symbol of source message
2) Decision	Y_j	Symbol of detected message
3) Cost matrix	$[c]$	Fidelity criterion
4) Data	$\overline{\alpha}$	Signal measurements
5) Set of conditional probability functions	$f_X(\overline{\alpha})$	Symbol likelihood function
6) Decision rule	$\delta(Y_j,\overline{\alpha})$	Set of computations which result in choice of Y_{j*}

2.2.2 The risk function

The average cost of decision making is the risk, a function of the true state of nature and the rule which governs the choice of a decision. If δ is the adopted decision rule and X_i the state of nature the average cost of all trials in which $\overline{\alpha}$ appears is

$$\sum_{j=1}^{N} \delta(Y_j,\overline{\alpha})c_{ij} \quad . \tag{2.2}$$

The risk of X_i is defined as the average of 2.2 taken over all possible data points:

$$r(X_i,\delta) = \int_A f_i(\overline{\alpha}) \sum_{j=1}^{N} \delta(Y_j,\overline{\alpha})c_{ij} \, d\overline{\alpha} \tag{2.3}$$

This function plays a central role in an evaluation of possible decision rules. The concept of admissibility is defined in the following system of comparison:

1) δ_1 is <u>equivalent</u> to δ_2 if

$r(X_i,\delta_1) = r(X_i,\delta_2)$ for all X_i.

2) δ_1 is <u>uniformly better</u> than δ_2 if they are not equivalent and

$r(X_i,\delta_1) \leq r(X_i,\delta_2)$ for all X_i.

3) δ_1 is <u>admissible</u> if there is no member of Δ which is uniformly better than δ_1.

4) δ_1 is the <u>best</u> decision rule if it is the only admissible member of Δ.

In general there is no best decision rule and the statistician seeks the set of admissible rules and a rationale for adopting one of its members.

2.2.3 Prior probabilities and expected risk

If the decision procedure is to be repeated many times, the expected risk of the selected decisions depends on

$$p = p(X_1),\ p(X_2),\ \ldots,\ p(X_M)$$

the relative frequencies with which the states of nature occur. For a given decision problem defined by X, Y, [c], $f_X(\overline{\alpha})$ and Δ, the expected risk is a function of δ and p:

$$R(p,\delta) = \sum_{i=1}^{M} p(X_i)r(X_i,\delta) \qquad (2.4)$$

For any p a decision rule δ_p in Δ exists such that

$$R(p,\delta_p) \leq R(p,\delta) \quad \text{for all } \delta \text{ in } \Delta. \qquad (2.5)$$

This decision rule which has minimal expected risk when p is the set of relative frequencies is defined as a "Bayes decision rule relative to p".

The significance of the set p in a practical situation is a controversial aspect of many applications of decision theory. The validity of 2.4 as a formula for average risk does not depend on any "randomness" of the mechanism which generates the states of nature. From the point of view of game theory, this mechanism is anything but random. It is influenced by an opponent's judgment of a player's decision rule and if the opponent is astute, the states of nature will be distributed in a manner which favors the higher-risk states. For this reason a game-theory oriented statistician does not concern himself with the relative frequencies and their expected risk in choosing a decision rule. He considers instead the maximum risks of admissible rules and seeks a minimax rule (defined in Section 2.2.4) -- which is an admissible rule with minimal maximum risk.

In many applications, including telegraphy, the prior probabilities do not depend on the decision rule and the minimax rule is likely to be inappropriately conservative. In such cases it is usually desirable to apply the Bayes rule relative to a set of prior probabilities which reflect the available information about how the states of nature are generated. In only a minority of cases is it possible to specify the M-1 independent constraints necessary to completely determine a prior probability function. Often (in business situations, especially) the prior information exists as the well-founded beliefs of a person experienced in dealing with the phenomena involved. These beliefs may be elicited by the statistician in a manner which admits the specification of a set of relative frequencies.

It may also be the case that the prior information takes the form of a limited set of constraints which impose restrictions on the prior probability function without completely determining it. For this situation, the principle of maximum entropy provides a rationale for choosing a set of prior probabilities. The entropy of the function p is defined as

$$-\sum_{i=1}^{M} p(X_i)\log p(X_i) \quad . \qquad (2.6)$$

It has been demonstrated that a prior probability function for which 2.6 is minimal subject to the constraints of the problem summarizes the available information about the states of nature in a "least-prejudiced" (a concept explained in (54)) manner. The maximum entropy principle is a generalization of Bayes' axiom.

2.2.4 Bayes decision rules, minimax rules

The importance of the set of all Bayes decision rules is demonstrated by the complete class theorem which states that all admissible rules are members of this set. Furthermore if $p(X_i) > 0$ for all i, (which means that only the states of nature that occur with non-zero probability are included in X), all Bayes rules are admissible. The decision problem thus reduces to the problem of choosing a prior probability function which in turn determines a Bayes rule. The preceding section describes two approaches to the choice of this function.

In one approach a prior probability function is sought which reflects the available information about how the states of nature are generated. This function determines a "representative Bayes rule" whose implementation completes the solution.

The other approach to the problem seeks maximum protection against the generation of unfavourable states of nature. This approach leads to the following definition: δ_μ is a minimax rule if

$$\max_i \; r\,(X_i, \delta_\mu) \leq \max_i \; r\,(X_i, \delta) \qquad (2.7)$$

for all δ in Δ. The maximum risk of δ_μ is minimal.

A minimx rule is clearly admissible and therefore a Bayes decision rule relative to a prior probability function

$$\mu = \mu(X_1), \mu(X_2), \ldots, \mu(X_M) .$$

Wald demonstrates that the expected risk of the minimax rule is maximal over the set of Bayes decision rules. That is,

$$R(\mu,\delta_\mu) \geq R(p,\delta_p) \qquad (2.8)$$

for any Bayes rule, δ_p, and for this reason, μ is known as the least favorable prior probability function. If a function μ is derived which satisfies 2.8, the minimax rule may be readily implemented according to the procedure presented in the next two sections. Another guide to the implementation of a minimax rule is the theorem which states that if δ_μ is a Bayes rule, it is a minimax rule if and only if its risk, $r(X_i,\delta_\mu)$ is constant over all X_i (for which $\mu(X_i) > 0$). On the basis of this property, Weiss (70) presents a linear programming method for finding δ_μ.

2.2.5 Practical implementation

For the selection of an individual decision, the quantities available to the statistician are a specific set of data $\overline{\alpha}$, the cost matrix, [c], and a set of prior probabilities p, obtained in a manner described in the preceding sections. Mathematical analysis demonstrates that a Bayes rule relative to p should dictate the choice of a decision. The problem will be solved, therefore, by a method of processing the available quantities in order to find a decision which conforms to such a Bayes rule.

The derivation of a computational procedure begins with the calculation of the posterior probability function relative to p. The conditional probabilities $f_i(\overline{\alpha})$ are used in this calculation which is performed according to Bayes' theorem. (It is due to this fact that the decision rule has been given Bayes' name.) For a given set of data, the posterior probability of X_i is:

$$P_{\overline{\alpha}}(X_i) = \frac{p(X_i)f_i(\overline{\alpha})}{\sum_{k=1}^{M} p(X_k)f_k(\overline{\alpha})} \tag{2.9}$$

$P_{\overline{\alpha}}(X_i)$ is the relative frequency of X_i measured over repeated occurrences of $\overline{\alpha}$. If Y_j is decided every time $\overline{\alpha}$ occurs, the average cost of observing $\overline{\alpha}$ is the "posterior risk of Y_j given $\overline{\alpha}$",

$$r_{\overline{\alpha}}(Y_j) = \sum_{i=1}^{M} P_{\overline{\alpha}}(X_i)c_{ij} \tag{2.10}$$

The significance of this function is expressed by the following theorem:

> A decision rule $\delta_p(Y_j,\overline{\alpha})$ is a Bayes rule relative to p if and only if it generates decisions Y_{j*} such that $r_{\overline{\alpha}}(Y_{j*}) \leq r_{\overline{\alpha}}(Y_j)$ for all Y_j in Y.

The theorem states, in effect, that minimization of the posterior risk for each $\overline{\alpha}$ results in minimal expected risk of decision making. The validity of this statement may be demonstrated by defining

$$t(Y_j) = \sum_{i=1}^{M} p(X_i) f_i(\overline{\alpha}) c_{ij} \qquad (2.11)$$

a function proportional to $r_{\overline{\alpha}}(Y_j)$ because the denominator of 2.9 is a constant for a given $\overline{\alpha}$. If Equations 2.3, 2.4 and 2.11 are combined, the expected risk may be expressed in terms of $t(Y_j)$ as:

$$R(p,\delta) = \int_A \sum_{j=1}^{N} t(Y_j)\delta(Y_j,\overline{\alpha})d\overline{\alpha} \; . \qquad (2.12)$$

In order to minimize 2.12 a decision rule δ_p must be specified with the property that it admits only minimal values of $t(Y_j)$ to the integrand. The appropriate specification is, therefore

$$\delta_p(Y_j,\overline{\alpha}) = 0 \quad \text{for all } Y_j \text{ such that}$$

$$t(Y_j) > \min_k t(Y_k)$$

$$\sum_j \delta_p(Y_j,\overline{\alpha}) = 1 \quad \text{for all } Y_j \text{ such that}$$

$$t(Y_j) = \min_k t(Y_k).$$

Because $t(Y_j)$ and $r_{\overline{\alpha}}(Y_j)$ are proportional, either of these functions may be used in this specification or in the above theorem. The specification and the theorem are identical statements.

2.2.6 Computational algorithm

A Bayes decision rule relative to p thus dictates the choice of a decision Y_{j*} for which $t(Y_j)$ is minimal. Since $t(Y_j)$ may be computed for each set of measurements, the solution to the decision problem may be expressed as the following procedure:

Having measured $\overline{\alpha}$,

1) Calculate the M numbers $p(X_i)f_i(\overline{\alpha})$,

2) Multiply this 1×M matrix by the M×N cost matrix to derive the N numbers

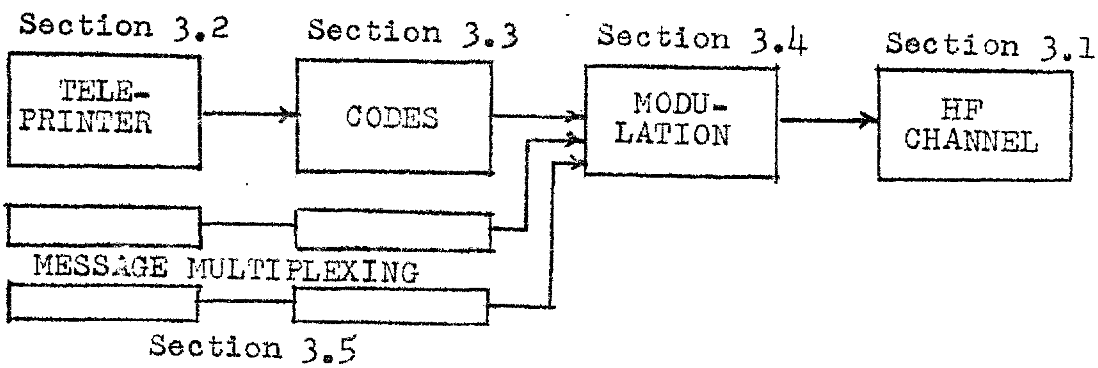
$$t(Y_j) = \sum_{i=1}^{M} p(X_i)f_i(\overline{\alpha})c_{ij} \qquad j=1,2,\ldots, N$$

3) Choose Y_{j*} where j* is the subscript of a minimal $t(Y_j)$.

If the minimax Bayes rule is adopted, p is identical to the least favorable prior probability function μ. If a representative Bayes rule is adopted p is a prior probability function which expresses the available information about the mechanism which generates the states of nature.

CHAPTER 3. THE TRANSMITTING STATION AND THE CHANNEL.

Before the Bayes decision procedure can be applied to telegraph detection the relationship between the received signals and the source message must be known. In the present chapter the two determinants of this relationship - the configuration of the transmitting station and the nature of the channel - are discussed. The elements of a typical telegraph system are described according to the following plan:



The remainder of the thesis investigates the detection of messages transmitted by this system. In the next chapter, the measurement of signals appearing at the output of the channel is considered and the conditional probability functions which relate the measurements to transmitted symbols are presented.

3.1 Ionospheric signal propagation.

The fundamental constraint upon the performance of a high frequency radiotelegraph system is the ionospheric propagation medium. While the electrostatic properties of the ionosphere make the reflection of radio signals, and hence long-distance transmissions possible, its magnetic properties and the variability of ionospheric phenomena severely limit the capabilities of practical systems. Although the existence of the ionosphere and its basic physical characteristics were discovered in the early decades of the present century, its fine structure and the mechanisms which cause certain propagation phenomena remain controversial topics in the subject of ionosphere physics.

3.1.1 Physical phenomena (35).

The ionosphere consists of several regions or "layers" of ionized gases at heights ranging from 60 to 450 km above the earth. The presence there of charged particles (free electrons have the greatest effect) causes a decrease in dielectric constant relative to an uncharged region so that an electromagnetic wave incident on one of the layers is refracted away from the vertical direction. If the density of free electrons increases with height to some critical value, the wave is progressively bent toward the horizontal and is finally redirected toward the earth. This process is illustrated by the ray diagram of a single "hop" shown in Figure 3.1.

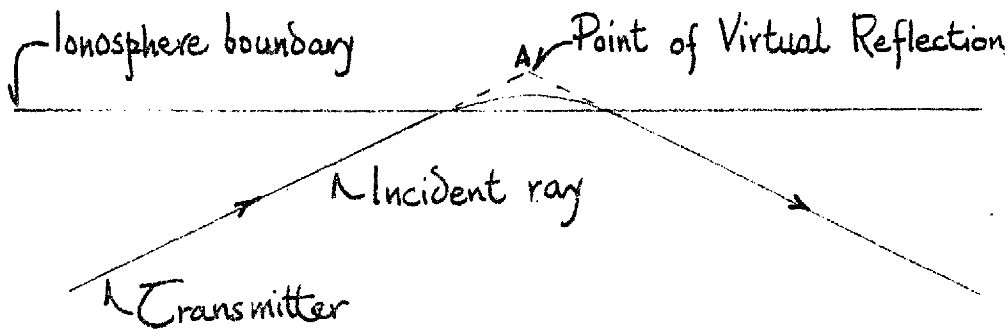


FIGURE 3.1 Single-hop transmission

It suggests that the wave is "virtually reflected" from point A in the ionosphere. The critical electron density, is directly proportional to $(f \cos\theta)^2$ where f is the frequency of the radiation and θ its angle of incidence.

The geometry of a propagation path and the electron density at a given time and place in the ionosphere determine a "maximum usable frequency" (MUF) of the path. Radiation at frequencies greater than the MUF penetrates the inosphere and does not return to the earth's surface. The MUF does not in general exceed 30 MHz which is usally considered to be the upper limit of the high frequency spectrum. Although signals at all frequencies below the MUF are reflected from the ionosphere, attenuation due to the absorption of energy by particle collision is an increasing function of wave length and thus imposes a "lowest usable frequency" (LUF) upon each possible route.

Because the ionization of atmospheric gases is caused by solar radiation, the electron densities and therefore the MUF, LUF and most other ionospheric phenomena vary with geographical position and, at a given point, with time of day, season and with the eleven-year cycle of solar activity. These variations are significant to the operators of high frequency communication systems and in order to chart them, several ionospheric observatories publish, from month to month, measurements and predictions of propagation conditions over most regions of the earth. These publications provide the data most often used by system planners and operators in selecting carrier frequencies and times of transmission. In recent years, equipment and techniques for "oblique ionospheric sounding" have been developed(18) in order to provide accurate and timely propagation data relevant to a specific route.

Because the span of an ionospheric hop is no longer than a few hundred miles, intercontinental radio communication depends on reflection of down-coming radiation from the surface of the earth and propagation of signals from transmitter to receiver by way of several hops between earth and sky. This multihop transmission is usually characterized by the phenomenon of multipath propagation illustrated in Figure 3.2.

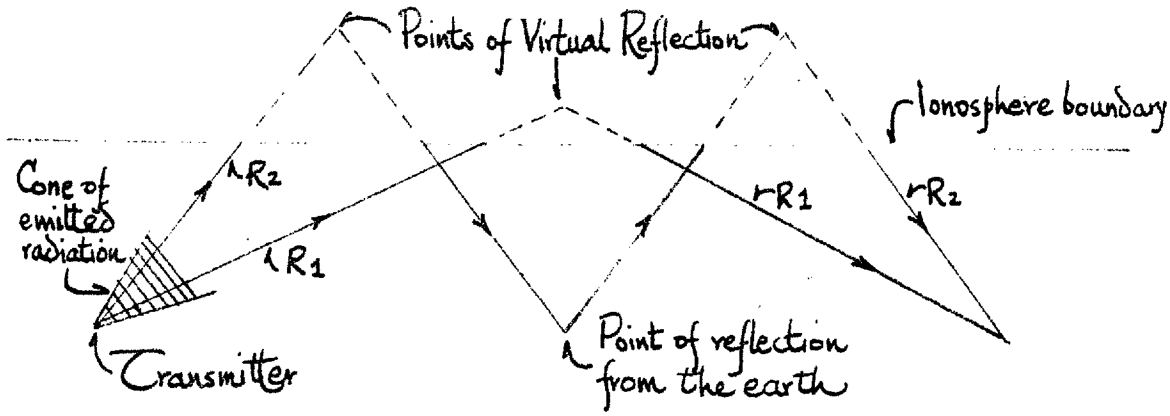


FIGURE 3.2 Double-path propagation

The transmitter emits radiation in a cone containing rays R1 and R2 which are reflected from different points in the inosphere.

The rays are propagated over different paths but, due to the overall geometry, finally converge at the receiver. Their different path lengths cause the two components of the received signal to be offset in time with respect to one another. If they reach the receiver with similar amplitudes, their resultant exhibits a fluctuating interference pattern which depends upon the relative delay of the two paths. In practice the multipath phenomenon is far more complicated than the one illustrated in Figure 3.2. Several paths usually exist and their delays and attenuations vary in time so that the amplitude of the resultant exhibits an apparently random pattern of fluctuation.

The earth's magnetic field produces a phenomenon similar in effect to multipath propagation. Magnetic forces cause each reflected ray to be resolved into "ordinary" and "extraordinary" magneto-ionic components characterized by different path-propagation delays. In general the strength of one of these components is negligible compared with that of the other but there are also circumstances in which they are attenuated to approximately the same extent and themselves produce interference patterns over a single ionospheric hop.

In addition to the interference effects which result from multimode and multipath propagation, there are usually rapid fluctuations in the strength of radio signals received over a single path and associated with only one magneto-ionic mode (38,39). In order to explain this phenomenon, various hypotheses regarding the fine structure of the ionosphere have been proposed. These hypotheses are influenced by the additional observation of spatial interference patterns, measurable by receiving the same signal at various points on the ground. One model suggests a moving, "corrugated" ionosphere characterized by ionospheric winds and an inhomogeneous distribution of electrons in a horizontal plane. This structure would produce, in addition to the reflection of energy from the ionosphere, a diffraction effect through which a single incident ray is returned in a cone of radiation containing component waves with various Doppler shifts relative to the incident signal. The effect observable on the earth is an amplitude fluctuation at a single point and a diffraction pattern distributed over a region. Other explanations of this variation in the strength of a single-ray received signal are based on models of the ionosphere which postulate the existence of discrete scattering centres from which signals are returned to earth. If these centres were to form and decay at various times or else move about in the ionosphere the observed diffraction phenomena would occur.

3.1.2 Characterization of signal fading.

Although the physical causes of the various ionospheric propagation phenomena are of general scientific interest, the communication engineer is primarily concerned with their effects on specific classes of radio signals. One effect of multipath propagation on digital communications is intersymbol interference which occurs when two signal paths of appreciable strength have significantly different propagation times. At a given instant, the received signal contains path components associated with different signal elements. This interference between signal elements has the same effect as a decrease in the system signal-to-noise ratio. This overlapping of successive signal elements is a principal factor in the limitation of signalling speeds of hf digital systems to the order of 100 bauds (signal elements/second). At this transmission speed, the difference in path delays may approach 5 milliseconds before the telegraph distortion reaches the level of 50%, a condition illustrated schematically in Figure 3.3.

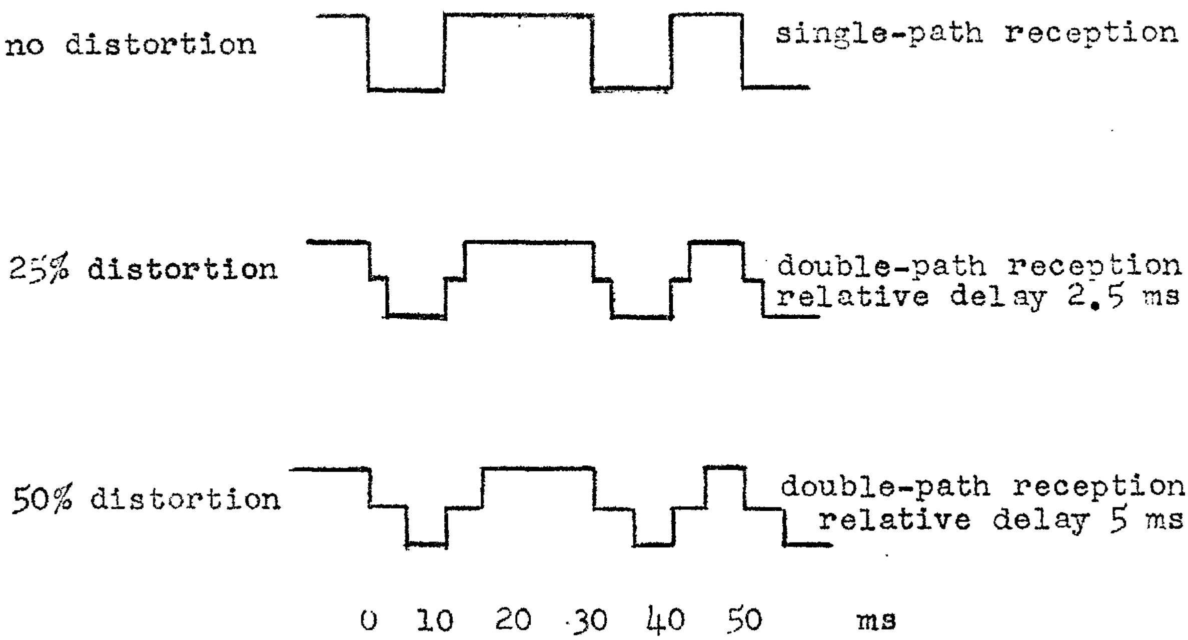


FIGURE 3.3 Distortion effects of multipath propagation

This level is often considered to be the maximum tolerable by practical signal measuring equipment (20).

Signal fading is another effect of both multiray propagation and diffraction-type phenomena. These propagation conditions, when they produce signal components with relative delays of the order of the period of the radio frequency carrier, cause spurious amplitude and phase modulations of the transmitted signal. If a steady sinusoid of frequency $\omega/2\pi$ is transmitted through an hf system, the received signal has a waveform

$$A \cos (\omega t + \theta)$$

where A and θ are the amplitude and phase modulations respectively. Owing to the complex mechanisms which produce the fading, these variables must be treated statistically - as random time functions of a stochastic process. Their magnitude variations are characterized by probability density functions and their time structures by correlation functions or power density spectra.

Because ionospheric propagation conditions exhibit extremely wide variability even with respect to a specific route and carrier frequency (11), it is impossible to precisely specify representative stochastic processes for the purpose of computing measures of system performance under all possible circumstances. There are, nevertheless, many instances in the design and analysis of an hf system when it is desirable to make some assumptions about the statistical properties of A and θ if only as first approximations to be verified in the field. It is this approach which is adopted in Chapters 6, 7, and 8 in which two detection methods are compared. For this purpose eight propagation conditions are considered and although none is strictly typical of a physical situation, they have been chosen to cover broadly the range of conditions under which most radiotelegraph systems operate.

In each of the eight cases, A and θ are assumed to be constant over the time duration of an individual signal element. There are two different conditions imposed upon the phase angle. At a coherent receiver the value of θ associated with each signal element is a constant, or if it varies, it does so in a manner which permits its measurement for each signal element. At a non-coherent receiver, the value of θ is uniformly distributed over a range of 2π radians and statistically independent from element to element. In this case the phase angle contains no information relevant to the identity of the transmitted signals. These conditions of complete coherence and non-coherence are the extremes of possible phase angle conditions. It is often the case that a receiver may operate in a state of partial coherence (55, 63) such that the phase angle is not known precisely but has a probability distribution which shows some values to be more likely than others. For both of the phase angle conditions, a non-fading amplitude(constant A) and four types of amplitude fading are considered.

In each case, the amplitude of a fading signal is assumed to be Rayleigh-distributed with probability density function

$$p(A) = \frac{2A}{S} \exp\left(-\frac{A^2}{S}\right) \text{ for } A \geq 0 , \qquad (3.1)$$

where S is the mean value of A^2. This distribution conforms to the model of a received signal as the resultant of a large number of randomly scattered components. A more general model includes a steady non-fading signal of amplitude A_1 added to the scattered components so that the probability density function becomes

$$r(A) = \frac{2A}{S} \exp\left(-\frac{A^2 + A_1^2}{S}\right) I_o\left(\frac{2AA_1}{S}\right) , \qquad (3.2)$$

the Rice probability density function in which $I_o(x)$ is the modified Bessel function of zero order. Equation 3.2 may be considered a family of density functions with parameter $u = A_1/S^{\frac{1}{2}}$, the ratio of steady component to rms fading component. When u = 0, 3.2 reduces to the Rayleigh density function of 3.1 As $u \to \infty$, A approaches the constant A_1, corresponding to the non-fading case. The analyses in the later chapters will therefore treat the two extremes of the Rice distribution of signal amplitude.

Owing to the temporal variation of ionospheric phenomena, the mean square level, S, of a fading signal is a function of time and the amplitude stochastic process is, therefore, inherently non-stationary. For practical purposes, however, the value of S may be considered a constant over a period of 15 minutes to one hour, the transmission time of several thousand symbols. A condition of short-term stationarity thus exists and a 15-minute time average of A^2 may be used as the value of S in Equation 3.1. The other important characteristic of the time structure of the amplitude fading is the manner in which the random variable, A, assumes new values over a sequence of signal elements. Like the amplitude and phase conditions, the time structures considered in this thesis are simplifications of theoretically (39) and empirically (32) derived characterizations. Although a gaussian autocorrelation function has been derived from physical assumptions and also been fitted to measured data, it is assumed in the present study either that the signal amplitude is constant over all the elements of a symbol (slow fading) or that it has a new, statistically independent value for each signal element (fast fading) (64). Whether the fading is slow or fast, the amplitudes associated with different symbols are assumed to be statistically independent. The autocorrelation function implied by each of these conditions is, therefore, rectangular in shape. With respect to a particular message, the applicable time structure depends not only on the nature of the physical variation of the signal amplitude but also on the organization of the transmitting terminal. If several messages are time-division multiplexed, the adaption of an element-multiplex arrangement (see Section 3.5) separates in time the elements of each symbol and thus causes their amplitudes to be less correlated than they would be in a symbol-multiplex arrangement. For a given channel therefore, it may be that the fast-fading model is applicable to element-multiplexed messages and the slow-fading model to symbol-multiplexed messages.

In addition to the amplitude and time structures of the fading an important characteristic is its frequency selectivity. The nature of the interference mechanisms which produce the fading suggest that when two sinusoids at different frequencies are transmitted over the same path they will be received with amplitudes which differ at a given instant, even though the patterns of amplitude fluctuation are identical. The telegraph system studied in this thesis conveys binary coded information by means of sinusoidal signal elements which are distinguished by their frequencies. The correlation between amplitudes of signals received simultaneously at the two frequencies, a decreasing function of the frequency difference, is a measure of the frequency selectivity of the system. In this case, too, two extremes are considered. For "flat fading" the amplitudes are identical while for "frequency selective fading" they are statistically independent.

The additive noise is assumed to be white and gaussian, a reasonable representation of atmospheric effects and of thermal sources. It fails, however, to characterize "man-made"noise - in particular, noise which results from reception of extraneous signals, sometimes a limiting constraint on system quality.

Table 3.1 is a summary of the amplitude and phase characteristics whose effects are discussed in later chapters. The two phase angle conditions are denoted coherent reception and non-coherent reception and in addition to steady signals, four types of amplitude fading are considered. Two of them, however, fast flat fading and fast selective fading will be seen to have the same effect upon message detectability so that there are eight effectively different transmission conditions.

TABLE 3.1

SUMMARY OF PROPAGATION CONDITIONS

	Coherent reception		Non-coherent reception	
No fading	Case 1		Case 5	
Slow fading	Case 2	Case 3	Case 6	Case 7
Fast fading	Case 4		Case 8	
	Flat fading	Selective fading		Flat fading

3.2 The message source.

The basic terminal instrument of much of the world's telegraphy is an electromechanical teleprinter with an alphabet of 31 symbols. Twenty-seven are printed characters and four are control symbols, namely:

1) "carriage return" which resets the printer to the beginning of its line of print,
2) "line feed" which advances the paper roller so that printing continues on a new line,
3) "letter shift" which sets the printer to operate in one of its two printing modes and
4) "figure shift" which sets the printer to its other mode.

Except for the space symbol which is a printing symbol common to both modes of operation, the character printed at any time is determined by the identity of the most recent shift symbol. The characters in the letters mode are the 26 letters of the English language while the figures mode contains numerals, punctuation marks and special characters. Fig. 3.4 illustrates the keyboard of the Teletype Model 32 Page Printer Set.

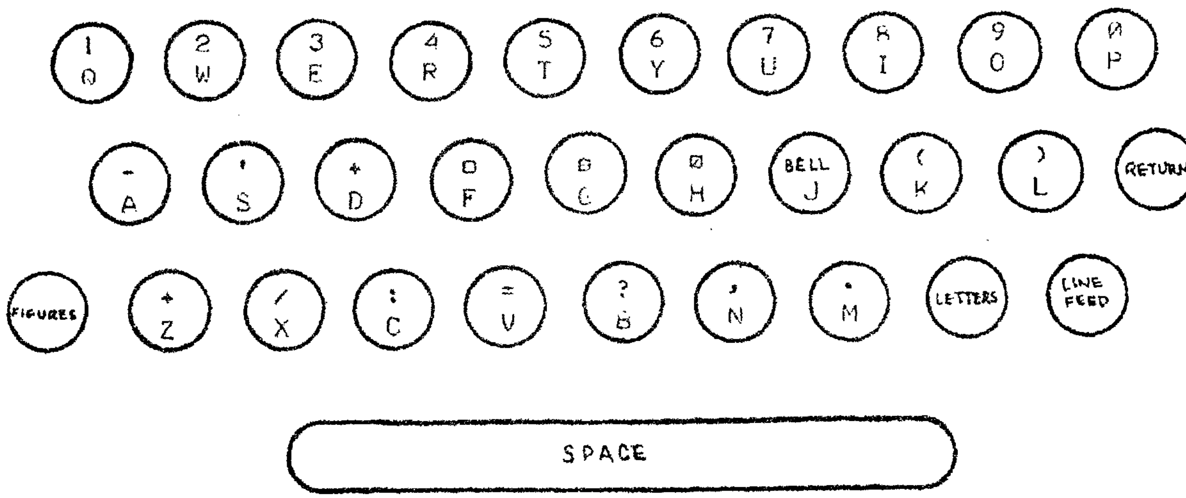


FIGURE 3.4 Teleprinter keyboard

Although the meaning of a telegraph message is conveyed by the pattern of characters appearing on a printed page, the ability of a machine to produce a legible message depends to a very great extent upon its reception of the correct control symbols. This is demonstrated in Chapter 5 which shows how entire lines of print may be lost if a carriage return or line feed is erroneously inserted in or deleted from a message and that if a printer is set to the wrong mode of operation, a sequence of several characters becomes illegible. An important aspect of a decision-theory approach to detection is the specification of a cost matrix which properly reflects the effects of the various types of symbol errors upon message quality.

In addition to its printing alphabet, the teleprinter has an electrical alphabet which is used for communication with other teleprinters. Each telegraph symbol is associated with a sequence of binary code elements which appear as d.c. currents in the machine. Although a d.c. telegraph signal may be transmitted directly from one teleprinter to another, telegraph networks usually convey information by means of voice-frequency or radio-frequency carriers. In the case of radiotelegraphy, in particular, there is likely to be a binary code conversion between teleprinter and modulator and also an alteration in the timing of the symbol sequence. To facilitate these transformations and for reasons of network organization, the electrical representation of a telegraph message is usually recorded prior to modulation and radio transmission and again after detection. The most common storage medium is perforated paper tape, but with the current trend toward computer-controlled message-switching centres, magnetic media are finding wider application. A stored message serves as a buffer between the teleprinters and the remainder of the system. In a system with a message storage facility, the telegraph message may be considered with respect to its symbolic content alone, irrespective of the rate and rhythm of teleprinter operation. Although recording often takes place at a centralized telegraph station, the feeder link between teleprinter and recorder is ignored in this thesis and only radio transmission of a pre-recorded telegraph message is considered.

3.3 Telegraph codes.

Of the codes which have been devised to meet the requirements of various transmission media and telegraph equipment (19,45), there are two which are especially relevant to modern radiotelegraphy. Both are fixed-length binary codes which differ in the number of code elements representing each telegraph symbol and in the manner in which the elements of successive symbols are framed.

The five-unit start/stop code is the electrical alphabet of the teleprinter. Each symbol is represented by five binary elements and thus the electrical alphabet consists of 32 symbols. Thirty-one correspond to the printed characters and control symbols and the other one is associated with an idle state of the mechanical parts of the teleprinter. In a signal sequence which conveys information by means of the five-unit code, the elements of each symbol are preceded by a start element whose electrical waveform resembles a "one" and followed by a stop element which resembles a "zero". In some systems, the stop element has the duration of $1\frac{1}{2}$

information elements while in other cases the durations of the two types of element are the same. In either event the five-unit code is uneconomical in terms of signalling speed and power since energy is transmitted during 7 or $7\frac{1}{2}$ time units per symbol of which only five time units are used for the conveyance of telegraph information. The five-unit code is generally used in feeder links and in some cases throughout an entire radio system because of its direct relationship with the mechanical operation of a teleprinter.

A more efficient code is the widely used seven-unit synchronous code. It was first proposed for application in "error-correcting" systems with a return path signalling facility (see Section 4.4), but it also holds distinct advantages for strictly one-way telegraph links. The most important characteristic of the seven-unit code is its fixed-ratio property: four elements of the code word for each symbol are "ones" and the other three are "zeros". Thirty-five code combinations meet this specification; thirty-two correspond to teleprinter symbols and three are available for special signalling roles within the system.

The fixed-ratio constraint provides symbol-framing information so that start and stop elements are unnecessary. All seven elements of each symbol convey telegraph information and the signalling speed of one telegraph symbol per seven time units is at least as high as that of a five-unit code system. The principal advantage of the seven-unit code is the considerable information redundancy it offers. There are 93 combinations of seven binary elements which fail to satisfy the four-out-of-seven requirement. These include all those which are caused by a single inversion of a code element within a symbol and many which result from multiple inversions.

Conversion to seven-unit from five-unit code format and the inverse transformation are examples of operations normally performed at a network's centralized transmitting and receiving terminals. Information is often represented in a radio link by the seven-unit code and by the five-unit code in the cables linking terminals and teleprinters. Specific codes recommended (by the CCITT) for use in international telegraphy are shown in Table 3.2. The principles of radio-signal detection introduced later in this thesis are illustrated by reference to telegraph systems which use this seven-unit code over their radio links.

3.4 Modulation.

The design of the modulator for a radiotelegraph system is influenced by the anticipated character of the hf transmission channel, a quality objective for

Table 3.2

TELEGRAPH CODES

FIVE-UNIT TELEPRINTER CODE	SYMBOLS LETTERS MODE	SYMBOLS FIGURES MODE	SEVEN-UNIT FIXED-RATIO CODE
00011	A	-	1100101
11001	B	?	1100110
01110	C	:	0110011
01001	D	+	1100011
00001	E	3	1000111
01101	F	□	1101100
11010	G	⊟	0011110
10100	H	⧄	0101101
00110	I	8	0001111
01011	J	bell	1011100
01111	K	(	1110100
10010	L	)	0011101
11100	M	.	0101110
01100	N	,	0101011
11000	O	9	0111001
10110	P	0	0110101
10111	Q	1	1110010
01010	R	4	0011011
00101	S	'	1010101
10000	T	5	0111010
00111	U	7	1001101
11110	V	=	0110110
10011	W	2	1011010
11101	X	/	1101001
10101	Y	6	1101010
10001	Z	+	1001110
00100	space		0010111
01000	car. ret.		0111100
00010	line feed		0100111
11011	figures		1011001
11111	letters		1110001
00000	idle		1111000
	RQ symbol		1001011
	symbol α		1010110
	symbol β		1010011

the received messages and economic considerations reflected in contraints upon bandwidth, signalling speed and radiated power. Although any two disjoint waveform classes may be specified as the modulation of a binary-coded system, it is generally the case that telegraph information is conveyed by keying between two sinusoidal waveforms distinguished by amplitude, frequency or phase.

In a communication system which conveys binary-coded information over a channel perturbed by white gaussian noise, the quality of received messages depends upon the energies of the two signal elements and their cross correlation (23) rather than their detailed waveforms. If x_o (t) and x_1 (t) are the binary signal elements, the quality increases with $S_0 + S_1 - 2\,S_{01}$ where

$$S_n = \int_0^T [x_n(t)]^2\,dt; \qquad n = 0,1 \qquad (3.3)$$

defines the energy of a signal element,T is the element duration and

$$S_{01} = \int_0^T x_0(t)x_1(t)dt \qquad (3.4)$$

is a measure of the correlation of x_0 (t) and x_1 (t) Since transmitters are power limited, it follows that $S_0 + S_1$ is maximized if peak power is radiated over the T seconds duration of both signal elements. This implies $S_0 = S_1 = S$ and precludes a system of amplitude modulation such as on/off keying in which $x_o(t) = 0$ and $x_1(t)$ is a sine wave of maximum power.

For modulations with equal-energy signal elements, the most negative cross-correlation, $S_{01} = -S$, is achieved when $x_0(t) = -x_1(t)$ as is the case in a system of 180^o phase-shift keying. This modulation offers the best possible noise immunity and also has very modest bandwidth requirements. Its weakness is its susceptibility to detection errors caused by the spurious phase modulation associated with hf signal fading (65). Because of these properties, phase shift keying is often used in cable telegraphy but is infrequently adopted in radiotelegraph systems. The binary signal elements of most radiotelegraph systems are sinusoids at two different frequencies. There are two means of achieving this type of modulation. One method involves frequency modulation of a single oscillator by a d.c.telegraph signal (this is frequency shift keying or FSK) and the other method involves switching the modulator output between two independent hf oscillators (two-tone keying). In the case of FSK, the output phase angle is continuous at signal-element transitions while for two-tone keying, there is a phase discontinuity when the output is switched from one frequency to the other. In both cases $S_{01}=0$

so that the degradations due to gaussian noise are identical. The systems differ in their bandwidth requirements (10) and FSK is generally used with narrow frequency shifts while systems of two tone keying have relatively wide spearations between the two signal frequencies. In this latter case, any fading tends to be frequency selective and the received signal may be treated as the reciprocal on/off transmissions offering a frequency-diversity advantage (3). (See Section 4.2).

The binary error rates of practical systems employing several types of modulation are presented by Ridout and Wheeler (41).

3.5 Multiplex arrangements.

The organization of the international hf telegraph network is greatly influenced by the constraints imposed by the ionospheric propagation medium, the properties of the electromechanical teleprinter and the practice of allocating transmission channels in 3 kHz independent sideband units. The signal fading characteristics impose an upper limit of around 100 bauds on the signalling speed which implies an approximately commensurate bandwidth per telegraph channel. In practice, the frequency difference in a 100-baud FSK system is often 85 Hz and the allocated signal bandwidth is 170 Hz. The figures for typical two-tone systems are 170 Hz and 340 Hz respectively. With these bandwidths twelve FSK signals or six two tone signals are frequency-division multiplexed (FDM) in a 3 kHz transmission channel. The methods of frequency allocation are illustrated in Figure 3.5 which shows that in two-tone systems the frequency pairs associated with individual messages are interleaved within the 3 kHz channel in an attempt to ensure that the fading is frequency selective with respect to each message. Because teleprinters are usually built to operate at 50 bauds, the practice has grown of time-division multiplexing (TDM) two "full-speed" or four "half-speed" messages to form a 100 baud signal. A practical time-division scheme consists of either a symbol multiplex, an element multiplex, or sometimes in the case of four messages, a hybrid arrangement. Three schemes are illustrated in Figure 3.6, which shows the time slot allocations for four TDM messages represented by the seven-unit code.

Incorporating various FDM and TDM arrangements, the international telegraph network contains many 3 kHz channels carrying 12 to 48 messages simultaneously. A typical organization of a 12-message multiplex package is shown in Figure 3.7. Each message is assigned to seven time slots and two frequencies which together may be considered as 14 rectangles in a time-frequency plane. All "ones" in a specific message cause the transmission at the appropriate times

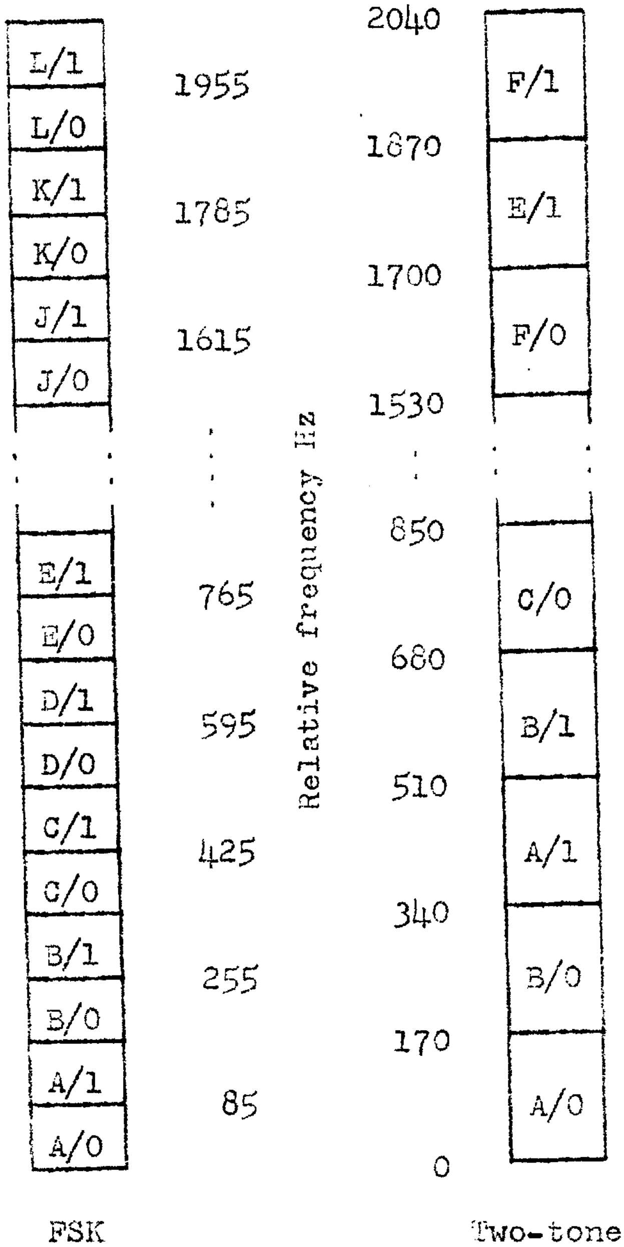


FIGURE 3.5 Frequency division multiplex of messages A,B,C,... conveyed by two-tone or FSK signals.

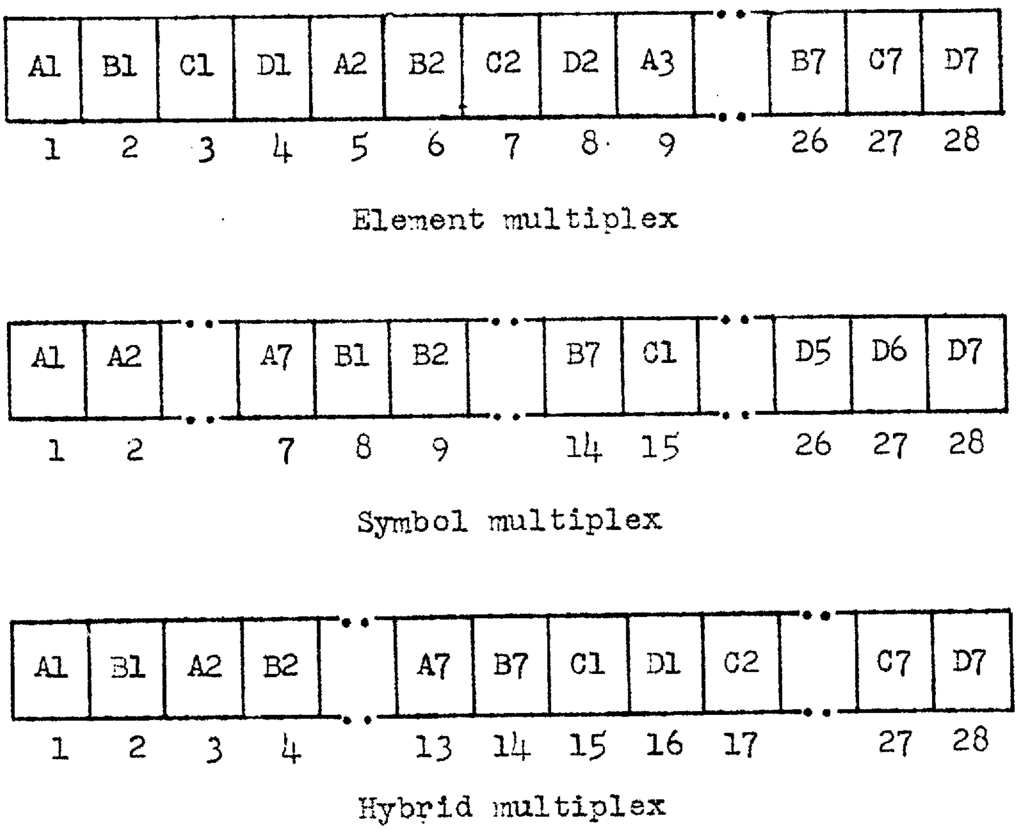


FIGURE 3.6 Three time division multiplex arrangements for messages A,B,C,D, represented by seven-unit code.

of a sine wave at the frequency assigned to "one" for that message. The same telegraph information may be conveyed, however, by means of any arbitrary assignment of 14 time-frequency rectangles to each of the twelve messages. For each message, the rectangles are grouped in seven pairs and one member of each pair is associated with a value "one" of a specific code element and the other member with the value "zero".

A new multiplex system may be implemented by simply rewiring parts of conventional transmitting and re-

Frequencies							
12	F1/1	L1/1	F2/1	L2/1	...	F7/1	L7/1
11	E1/1	K1/1	E2/1	K2/1	...	E7/1	K7/1
10	F1/0	L1/0	F2/0	L2/0	...	F7/0	L7/0
⋮	⋮	⋮	⋮	⋮		⋮	⋮
5	C1/0	I1/0	C2/0	I2/0	...	C7/0	I7/0
4	B1/1	H1/1	B2/1	H2/1	...	B7/1	H7/1
3	A1/1	G1/1	A2/1	G2/1	...	A7/1	G7/1
2	B1/0	H1/0	B2/0	H2/0	...	B7/0	H7/0
1	A1/0	G1/0	A2/0	G2/0	...	A7/0	G7/0
	1	2	3	4		13	14

Time slots

FIGURE 3.7 Combined FDM-TDM arrangement. Ak/0 and Ak/1 denote rectangles assigned to kth element of symbols of message A ceiving terminals. As in the conventional modulation-multiplex arrangements, each code element is communicated by means of a sinusoidal signal element with characteristic time and frequency co-ordinate so that for the entire multiplex package the signal measurement requirements are identical to those presented in Section 4.1. The recognition of this flexibility permits, however, the design of multiplex arrangements with the aim of ensuring the best fading speed and selectivity for each message. Principles of time and frequency assignment are discussed in Chapter 6 on the basis of the performance characteristics associated with the various propagation conditions.

CHAPTER 4 : SIGNAL MEASUREMENT AND BINARY DETECTION

The material presented in Chapter 3 implies a formal model of the transmitting end of the telegraph system and the radio channel. The transmitted signals are related to the symbols of the source message according to the specified properties of the teleprinter, the code and the modulator. The assumption of white gaussian noise and one of the eight propagation conditions determines the statistical relationship between the transmitted signal and the received signal.

4.1 Signal measurement

The decision procedure of Section 2.2.6 requires the calculation, for each i, of $f_i(\overline{\alpha})$, the probability density that the uncertain evidence results when X_i is the state of nature. In telegraphy the evidence is a received waveform, y(t), from which the data $\overline{\alpha}$ must be extracted. y(t) consists of the fading, noisy versions of the seven signal elements and because there is no intersymbol interference and the noise is white and gaussian, the additive perturbations of all of the signal elements are statistically independent. The received signal may be separated, therefore, into seven waveform components:

$$y_1(t),\quad y_2(t),\quad \ldots,\quad y_7(t),$$

each of which carries the information relevant to one code element. For a given y(t), the 31 numbers $f_i(\overline{\alpha})$ comprise the symbol likelihood function which may be calculated from the 14 element likelihoods,

$$\lambda_o^{(1)},\quad \lambda_1^{(1)},\quad \lambda_o^{(2)},\quad \lambda_1^{(2)},\ \ldots\ \lambda_o^{(7)},\quad \lambda_1^{(7)},$$

where $\lambda_n^{(k)}$ is the probability density of $y_k(t)$ when the kth code element of the transmitted symbol has value n (either 0 or 1). If the code representation of X_i is

$$b_{i1}\ b_{i2}\ \ldots\ b_{i7}$$

the likelihood of X_i is the product of the seven appropriate element likelihoods

$$f_i(\overline{\alpha}) = \prod_{k=1}^{7}\left[\lambda_o^{(k)}(1-b_{ik})+\lambda_1^{(k)}\ b_{ik}\right] \tag{4.1}$$

In a two tone or FSK system with signalling speed of 1/T bauds, the signal component of $y_k(t)$ is

$$A_n\cos(\omega_n\tau_k + \theta_n) \tag{4.2}$$

where the time variable τ_k is defined such that the kth signal element is received over the time interval $\tau_k = 0$ to $\tau_k = T$. The value of ω_n identifies the transmitted code element while the characteristics of A_n, the amplitude, and θ_n, the phase, depend upon the fading and coherence conditions. The likelihood of n is the probability density of a noise sample whose shape is

$$y_k(\tau_k) = A_n \cos(\omega_n \tau_k + \theta_n) .$$

The likelihood may be derived (see page 82 of ref. 73), on the basis of an orthogonal expansion of $y_k(t)$, as

$$\lambda_n^{(k)} = \eta_k \exp\left(-\frac{S_n}{N}\right) \exp\left[\frac{2A_n}{N} \int_0^T y_k(\tau_k) \cos(\omega_n \tau_k + \theta_n) d\tau_k\right] \quad \text{... (4.3)}$$

in which η_k is a constant independent of n, N $\frac{\text{watts}}{\text{Hz}}$ is the power spectral density of the noise and S_n, the energy of the signal component, is

$$S_n = \int_0^T A_n^2 \cos^2(\omega_n \tau_k + \theta_n) d\tau_k$$

which for common values of ω_n is equal to or very nearly approximated by

$$S_n = \frac{A_n^2 T}{2} \quad \text{... (4.4)}$$

The signal-to-noise ratio of the element is expressed as the ratio of its energy to the noise power per unit bandwidth, or

$$\rho_n = \frac{A_n^2 T}{4N} \quad \text{... (4.5)}$$

At a coherent receiver, θ_n is known for each k and n and Equation 4.3 may be used directly to calculate the 14 element likelihoods. $y_k(t)$ appears in the formula only within the integral

$$a_n = \int_0^T y_k(\tau_k)\cos(\omega_n\tau_k + \theta_n)\, d\tau_k \qquad \ldots \qquad (4.6)$$

so that a_n may be taken as the basic signal measurement required by the decision procedure. It may be acquired physically as the output at $\tau_k = T$ of the filter matched to the signal element n, and having y(t) as its input over the interval $0 \leq \tau_k < T$. This is a linear filter with inpulse response (56):

$$h_n(t) = \cos\,[\omega_n(T-t) + \theta_n]\,. \qquad \ldots \qquad (4.7)$$

When a_n has been measured the element likelihood may be calculated as:

$$\lambda_n^{(k)} = \eta_k \exp(-\text{[illegible]}\rho_n)\, \exp\left(\frac{\text{[illegible]}A_n}{N}\, a_n\right). \qquad \ldots \qquad (4.8)$$

At a non-coherent receiver, on the other hand, θ_n is a stray parameter and the appropriate likelihood formula is found by averaging 4.3 over the possible values of θ_n. Thus (4)

$$\lambda_n^{(k)} = \eta_k \exp(-\text{[illegible]}\rho_n) I_o \left(\frac{\text{[illegible]}A_n}{N}\, c_n\right) \qquad \ldots \qquad (4.9)$$

in which the signal measurement is c_n, the envelope at $\tau_k = T$ of the output of the matched filter. c_n is non-negative and its square may be expressed as the following function of $y_k(t)$:

$$c_n^2 = \left[\int_0^T y_k(\tau_k)\cos\omega_n\tau_k d\tau_k\right]^2 + \left[\int_0^T y_k(\tau_k)\sin\omega_n\tau_k d\tau_k\right]^2$$

The data, $\overline{\alpha}$, consists, therefore, of 14 numbers; a_o and a_1 or c_o and c_1 for each of the seven elements. $\overline{\alpha}$ comprises a set of sufficient statistics in that it summarizes the information in the received waveform relevant to the identification of a transmitted symbol (73). The specification of the filters for measuring $\overline{\alpha}$ and the formulas (4.8 or 4.9 followed by 4.1) for computing $f_i(\overline{\alpha})$ provide the practical basis of an application of the decision procedure in Section 2.2.6 to the detection of telegraph symbols.

The matched filter defined by Equation 4.7 may be realized physically as a lossless resonant circuit tuned to ω_n (17,42) and excited with y(t) over the time interval $0 \leq \tau_k < T$. If the output at $\tau_k = T$ is to depend only on a single element, no energy due to previous elements may remain at $\tau_k = 0$, a condition achieved by quenching the filter at this instant. In a system of multiplexed messages the measurements may be obtained from a set of quenched resonators each tuned to one of the system's characteristic frequencies. All of the filters are quenched at the beginning of each characteristic time interval. Each output, or its envelope, at the end of the interval is associated with the frequency and time co-ordinate assigned to an element of a particular message in the multiplex scheme.

4.2 Diversity operation

Section 3.1.1 describes the diffraction pattern observable when energy reflected from the ionosphere is received simultaneously at several aerials located in an array on the ground. At any instant the signal strengths measured at different aerials differ from one another even though their temporal patterns are similar statistically. In this situation, the signal strength, although weak for a significant fraction of time at each aerial, is likely at any instant to be relatively strong at at least one aerial. Typical signal amplitudes, received at two aerials of a diversity system, are illustrated in Figure 4.1.

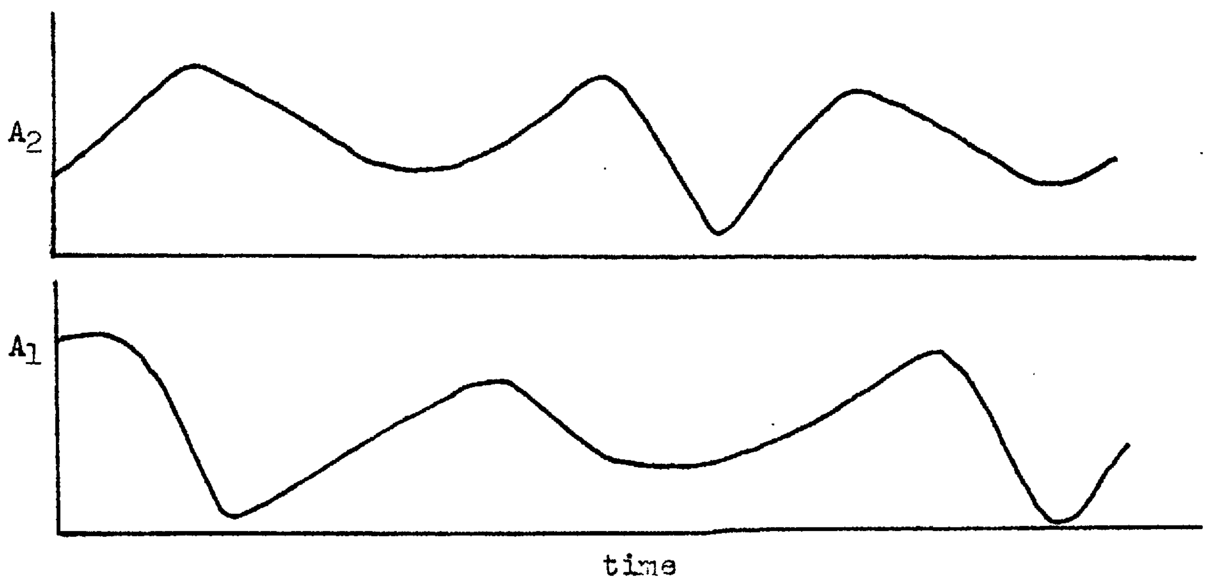


FIGURE 4.1 Amplitude records in two diversity branches

The installation of a multiple-aerial system provides an important means of combating the harmful effects of fading. In a system of spaced-aerial diversity reception the aerial array is designed to provide diversity branch signals whose amplitude fluctuations are at any instant statistically independent (5). In some diversity systems, the detector measures the branch signals individually and employs a computational procedure to combine the measurements (28). In other cases, however, the branch signals are combined prior to measurement and the detector obtains its data by operating on the resultant signal.

The three standard methods of diversity combination are all linear techniques (12) in that the resultant signal y(t) is the weighted sum of each of the R branch signals $z_r(t)$:

$$y(t) = \sum_{r=1}^{R} w_r z_r(t) \qquad \ldots (4.10)$$

If a steady sinusoid has been transmitted and there is phase equalisation of all of the diversity branches, the branch signals may be written:

$$z_r(t) = A_r \cos(\omega t + \theta) + \text{noise}$$

so that the resultant is also a sinusoid:

$$y(t) = A\cos(\omega t + \theta) + \text{noise} \qquad \ldots \quad (4.11)$$

whose signal-to-noise ratio characterizes the merits of a particular combination technique.

In the simplest of the three combination methods, the branch signals are simply added, i.e. $w_r = 1$ for all branches. This equal-gain combination results in a high signal-to-noise ratio when the noise powers in all branches are similar. A second technique, maximal-ratio combination, (26) offers optimum performance (regardless of the relationship among the branch-signal noise powers) in that it results in the highest signal-to-noise ratio obtainable by linear combination. In this case

$$w_r = \frac{A_r}{N_r} \qquad \ldots \quad (4.12)$$

where N_r is the noise power spectral density of branch r. The complication of measuring this ratio, however, often precludes the adoption of the maximal-ratio method. The third technique, optimal selection, demands a choice of a single branch signal for use as y(t). Thus, $w_r = 1$ for $r = r'$ and $w_r = 0$ for all other branches, where r' is chosen on the basis of a signal-to-noise ratio assessment. For equal noise powers in all branches, this method has a resultant with the lowest signal-to-noise ratio of the three (since useful information in the rejected branches is ignored) but it offers a practical advantage since no phase equalisation is required. The probability density function of A and the mean value of the signal-to-noise ratio obtained by all three combination methods are indicated in Table 4.1.

TABLE 4.1

DIVERSITY COMBINATION DATA (12)*

COMBINATION METHOD	AMPLITUDE PROBABILITY DENSITY OF RESULTANT	SIGNAL-TO-NOISE RATIO OF RESULTANT
SINGLE BRANCH (NO DIVERSITY)	$\frac{2A}{S} \exp(-\frac{A^2}{S})$	ρ
SELECTION FROM R BRANCHES	$\left[\frac{2A}{S}\exp(-\frac{A^2}{S})\right]R\left[1-\exp(-\frac{A^2}{S})\right]^{R-1}$	$\rho\sum_{r=1}^{R}\frac{1}{r}$
MAXIMAL RATIO COMBINATION	$\left[\frac{2A}{S}\exp(-\frac{A^2}{S})\right](\frac{A^2}{S})^{R-1}$	ρR
EQUAL GAIN COMBINATION	No closed-form solution available; distribution has been tabulated (31a)	$\rho(.215+.785R)$

*Data applies to R statistically independent branches. All branches have the same mean square amplitude, S, and the same noise power. The signal-to-noise ratio of each individual branch is ρ.

In addition to the diversity reception provided by spaced aerials, it is possible to receive two uncorrelated branch signals from a pair of orthogonally polarized aerials at the same point (22). Polarization diversity may be used in place of or in addition to spaced-aerial diversity and the two methods are similar in that they may both be implemented at a receiving station independently of the organization of the transmitter.

Time-diversity operation and frequency diversity, on the other hand, are methods which must be co-ordinated at both ends of a system. Their advantages derive from the time and frequency selectivity of the fading and they are implemented by means of the transmission of branch signals which are identical except for frequency or time translations with respect to one another. If branch-signal combination is to be performed in a time or frequency-diversity system the branch signals must all be translated to the same time or frequency range. Separate branch measurements are generally more appropriate to these types of system. A two-tone or FSK signal may, in some cases, be considered as an on/off transmission over two frequency-diversity branches (3). The measurements obtained by a pair of matched filters correspond to separate measurements in each branch. The weighting of a_n or c_n by A_n/N, specified in Equation 4.8 or 4.9 is equivalent to the optimal combination of branch signals (Equation 4.12) prior to measurement of their resultant.

The foregoing discussion of branch-signal combination has been based on the transmission of the steady sinusoid of Equation 4.11. When the combination procedures are applied to a sequence of binary signal elements, the measurements specified in Section 4.1 may be performed on the resultant signal. Within this resultant, however, the element amplitudes are not Rayleigh distributed. In the analysis of diversity-system performance, therefore one of the probability density functions of Table 4.1 must be applied in place of Equation 3.1. In this manner the performance characteristics, which in this thesis apply only to single-branch reception, may be extended to cover various types of diversity operation as well. Such extensions are omitted from the analytic and numerical results reported in Chapters 6, 7, and 8 because the purpose here is to emphasize the effects on performance of modifications in the computational role of a detector. On the other hand, the opportunity for improving the quality of the measurements by means of diversity reception must be considered in any practical application of any of these results.

4.3 Detection of binary elements

The detectors of most telegraph systems make decisions about individual binary code elements. They perform, in effect, exercises in statistical hypothesis testing and represent solutions to a very elementary decision-theory problem -- one which includes two states of nature and two possible decisions. In terms of the model of Section 2.2.1, the states of nature, X_1 and X_2, are identified with the transmission of "one" and "zero" respectively and Y_1 and Y_2 are decisions to generate the corresponding elements at the receiver. The likelihood functions are λ_1, identical with $f_1(\overline{\alpha})$ and λ_2 which corresponds to $f_2(\overline{\alpha})$. They are expressed in Equation 4.8 or 4.9 as functions of the signal measurements and channel characteristics. There are only two prior probabilities and the cost matrix has four elements.

The binary decision procedure is derived by applying Section 2.2.6 and consists of a comparison of the two numbers:

$$t(Y_1) = \lambda_1 \, p\,(X_1)\, c_{11} + \lambda_o \, p\,(X_2)\, c_{21}$$

$$\text{and} \quad t(Y_2) = \lambda_1 \, p\,(X_1)\, c_{12} + \lambda_o \, p\,(X_2)\, c_{22} \quad \ldots (4.13)$$

Decision Y_1 is adopted when

$$t(Y_1) < t(Y_2)$$

which requires

$$\frac{\lambda_1}{\lambda_o} = L > \frac{p(X_2)}{p(X_1)} \frac{(c_{21}-c_{22})}{(c_{12}-c_{11})} \quad .. \qquad (4.14)$$

Equation 4.14 demonstrates the well known principle of hypothesis testing that a binary decision procedure requires a test of L, the likelihood ratio, against a threshold. This test weighs the information conveyed by the received signal, summarized by L, with the expression which summarizes the prior information and the fidelity criterion. A high value of the threshold in 4.14 indicates that "one" is detected only when the signal contains very strong evidence that this is the correct decision. A high-valued threshold may derive from a high $p(X_2)$, indicating a strong belief by the designer that "zero" is normally sent, or it may be due to a high cost of mistakenly detecting "one".

The quality objective most commonly adopted in the design of a binary detector is minimization of the total number of detection errors. If each error is assigned unit cost, the elements of the cost matrix are:

$$c_{11} = c_{22} = 0; \qquad c_{12} = c_{21} = 1$$

so that the likelihood ratio must be compared with $p(X_2)/p(X_1)$. Because the same element would be detected if any monotonic function of L were compared with the same function of $p(X_2)/p(X_1)$, the binary decision rule generates Y_1 when

$$\hat{L} = \log L > \log \frac{p(X_2)}{p(X_1)} \quad \ldots \quad (4.15)$$

The test appropriate at a coherent receiver is derived by substituting 4.8 into 4.15, which becomes

$$\frac{A_1}{N} a_1 - \frac{A_o}{N} a_o - (\rho_1 - \rho_o) > \log \frac{p(X_2)}{p(X_1)} \quad (4.16)$$

The implementation of this test requires knowledge of the channel conditions as indicated by the values of A_1 and A_o and the noise power density. If there is no fading or flat fading, the amplitudes of the two elements are equal and the test reduces to

$$a_1 - a_o > \frac{N}{A} \log \frac{p(X_2)}{p(X_1)} \quad (4.17)$$

This inequality is also the decision criterion of the system which is not designed to measure the signal amplitudes separately and therefore assumes them equal, even though the fading may be selective.

Although the prior probabilities associated with the seven-unit code are automatically $p(X_1) = 4/7$, $p(X_2) = 3/7$, they are both implicitly assumed to equal $\frac{1}{2}$ in the design of most practical detectors. This assumption further simplifies the binary decision procedure to a test of the inequality

$$a_1 > a_o \, . \qquad \cdots \quad (4.18)$$

If 4.18 is satisfied, Y_1 is generated. In this case the computational role of the detector is reduced to a simple comparison of the two measurements. The degradation in performance due to the application of 4.18 instead of 4.17 is insignificant except when A/N is so low that both tests result in extremely high error rates (28). The decision rule specified in 4.18 is a maximum-likelihood rule in that its decision is determined by the value of i for which $f_i(\overline{\alpha})$ is maximal. It is a Bayes rule only when the right side of 4.14 is unity, i.e., when the costs and prior probabilities suggest no prior disposition toward the detection of one element in preference to the other.

At a non-coherent receiver, the simplifying assumptions are even more attractive. If the minimum-error criterion is adopted, the rigorous application of 4.15 demands a test of the complicated inequality,

$$\log I_o \left(\frac{\text{[illegible]}A_1}{N} c_1\right) - \log I_o\left(\frac{\text{[illegible]}A_o}{N} c_o\right) - \text{[illegible]}(\rho_1 - \rho_o) > \log \frac{p(X_2)}{p(X_1)} \qquad \cdots \quad (4.19)$$

which is derived by substituting 4.9 into 4.15.* If

---

* The approximation of $\log I_o(x) - \log I_o(y)$ by $\frac{1}{4}(x^2 - y^2)$ is discussed in Section 8.5.

equal prior probabilities and equal signal amplitudes are assumed, the test at a non-coherent receiver also reduces to a comparison of the signal measurements:

$$c_1 > c_o \quad \ldots \quad (4.20)$$

The generation of telegraph symbols by binary detectors is discussed in Section 4.4 which describes the ARQ signalling technique often associated with binary detection of symbols represented by the seven-unit code. Chapters 5, 6, and 7 compare the performances of binary detectors with various types of symbol detector. Because the existence of a binary detector is normally assumed in telegraph studies and the computational requirements of this device are so modest -- in general only a comparison of measurements is necessary -- most advances in theory and technique of telegraph detection have been applicable to the measurement role rather than the computational role of a detector. When, however, telegraph detection is considered to be a procedure leading to decisions about entire elements, the computational aspect of detection becomes significant. It is this aspect which is emphasized in subsequent chapters.

4.4 ARQ Operation

In a radiotelegraph system incorporating the seven-unit code, a binary detector generates a symbol whenever the seven detected elements corresponding to a transmitted symbol obey the fixed-ratio constraint of the code. The redundancy of the code often prevents the output of an incorrect symbol even though all of the elements have not been correctly detected. Any single binary error in the code sequence of a symbol and many multiple-error patterns result in a detected sequence which corresponds to one of the 93 redundant combinations. When this event occurs, it may be said that the channel has erased the transmitted symbol and the detector may be designed to generate a special erasure symbol to indicate this event to the message reader. Instead of this action, or addition to it, the detector may also initiate a special sequence of system operations in order to derive additional information relevant to the identity of the transmitted symbol.

In an ARQ system (15, 16) the detection of an erasure automatically causes the retransmission of the seven elements of the affected symbol. This technique is widely applied in two-way point-to-point radio systems. If an erasure is detected at terminal B a special RQ symbol is inserted in the message transmitted from B to A. When this RQ symbol is received at A, the erased symbol is retransmitted to B. The process is repeated until an acceptable seven-element sequence is received at B. By means of a similar set of operations, the detection of erasures at A results in retransmission of symbols from B to A.

Because the path delay of an hf radio link is generally of the order of magnitude of the symbol duration, practical ARQ systems are organized to retransmit a block of several symbols in response to each erasure. In most systems each repetition block consists of four symbols; over some very long routes, however, retransmissions occur in groups of eight symbols. When an erasure is detected at station B, the output of telegraph symbols is suspended and a repetition block consisting of four (or eight) RQ symbols is transmitted from B to A. When the first RQ symbol is detected at A, a repetition block is transmitted to B. This block consists of the RQ symbol, the erased symbol and the two (or six) symbols following the erased symbol in the message.

The correct detection at B of the returned RQ symbol permits the receiver to resume its normal operation. If the RQ symbol is not detected when it is expected at B another repetition block is initiated.

With this organization the system operates effectively when erasures are detected at both terminal points.

An ARQ system adapts its operation to changing channel conditions and in effect adjusts its transmission speed (the number of output symbols per unit time) in order to maintain a low symbol error rate. Adaptive operation is especially desirable when the channel is subject to fading so that the signal-to-noise ratio changes frequently. The performance at a non-coherent receiver of steady signals is demonstrated in Figure 4.2, which shows the proportion of incorrect symbols in the message detected at Station B as a function of signal-to-noise ratio. For comparison, the error rate of a constant-speed system using a five-unit code is also shown. Although the error rate of the ARQ system is effectively controlled over a wide range of signal-to-noise ratios, the transmission speeds in both directions of transmission decrease as the signal-to-noise ratio decreases. This effect is demonstrated in Figure 4.3 which shows, for each channel, the relative speed of transmission, defined as the ratio of output symbols to transmitted symbols. It is assumed that there is no noise in the B-to-A channel so that all of the extra transmissions result from erasures detected at B (57).

Figure 4.3 indicates that the output rates in both directions are controlled by the conditions over the worse of the two one-way paths. The ARQ system is efficient when the signal-to-noise ratios over both paths are high so that few repetitions are necessary. It is also effective when both are so low that accurate communication is impossible and the transmission virtually stops. When, on the other hand, one signal-to-noise ratio is high and the other low or both have intermediate values, the deterioration of transmission speeds would seem to inhibit excessively efficient utilization of the information available in one or both of the received signals. The efficiency of the ARQ technique is investigated in Chapter 7 by means of a comparison of ARQ signalling with a variable-speed mechanism which may be introduced with the maximum-likelihood detector discussed in Chapter 6. The ARQ technique is not applicable to systems in which no return path is available as in strictly one-way transmission or in broadcasts of telegraph information from a single transmitter to several receivers. The detection of symbols in these circumstances is described in Chapters 5 and 6.

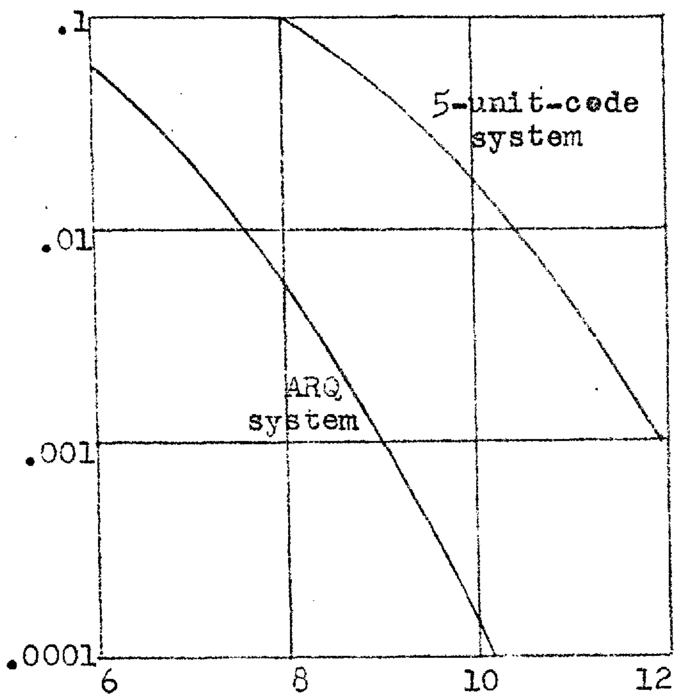


FIGURE 4.2 Symbol error rate v. signal-to-noise ratio (db)

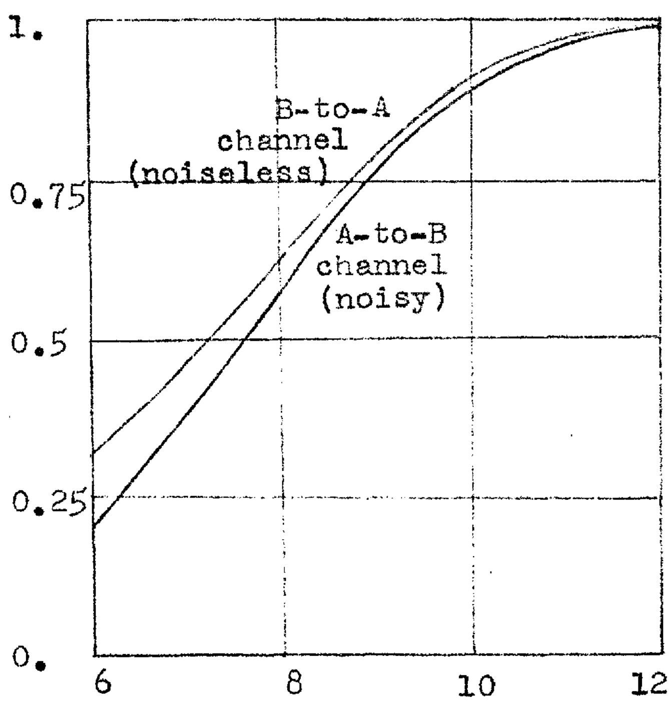


FIGURE 4.3 Relative speed v. signal-to-noise ratio (db)

CHAPTER 5: SIMULATED TELEGRAPH SYSTEM

Section 4.1 describes the measurement role of a detector of signals generated at the transmitting station and propagated through the channel formulated in Chapter 3. The set of numbers comprising the symbol likelihood function may be computed according to Equation 4.1 and either 4.8 or 4.9. By means of operations on these numbers and any specified cost matrix and prior probability function, the Bayes decision procedure of Section 2.2.6 may be readily implemented. The likelihood function summarizes the relevant information conveyed by the received signal; its form depends on the physical configuration of the system. The costs and prior probabilities represent the quality objective of the system and the designer's prior knowledge of the statistics of the source message. In the present chapter, three symbol detectors which are constrained by different sets of costs and prior probabilities are formulated. Their permance and that of the binary detector of Section 4.3 are investigated by means of a computer simulation.

5.1 Cost matrix

The assignment of detection costs is the most arbitrary aspect of an application of decision theory to telegraphy. The cost incurred when one symbol is transmitted and another detected is likely to depend on message content and purpose and thus vary from message to message and from user to user of the telegraph system. Within a given message, furthermore, the cost of a particular error depends on its context so that, for example, the detection of R instead of D is more costly if it changes DECEIVE to RECEIVE than if it causes RETECT to be printed instead of DETECT. This example suggests that detection might be better organized on the basis of symbol sequences than individual symbols. In principle, this extension of the scope of the detector would improve the quality of a system. In practice, however, the problem of assigning costs would be no less difficult than it is in the case of symbol detection and the computational requirements of the detector would increase enormously.

Detection of individual symbols seems a reasonable objective because they are the basic constituents of a telegraph message. Although it is impossible to specify a set of numbers which represents the cost of every error in every possible message, the nature of the teleprinter suggests an ordering of costs which provides a starting point for the specification of an entire matrix. The appearance and legibility of a received message depends to a great extent on the detected control symbols (45). Incorrect insertions and omissions of carriage-return and line-feed symbols jeopardize entire lines of text. In the following example one line has overprinted another because a line feed was omitted:

MHEIDATSONOOFOFWA EONE EENDSSOMBBRINS.LTHELCARRIAGESES

If carriage return is omitted, the carriage is likely to reach the end of its line and repeatedly overprint the same symbol position in this manner,

IF A CARRIAGE RETURN IS OMITTED, THE MARGIN AT THE END OF THE■

When shift characters are not correctly detected, long blocks of a message may be printed in the wrong mode and thereby render the message unacceptable.

These examples suggest that higher costs should be assigned to errors involving control symbols than to errors which cause one printed character to be substituted for another. In the detection of a specific text, the effects of control-symbol errors depend on the format of the printed source message. In the computer simulation, the message format has been expressed as the set of the following four numbers:

k_1, the average number of printed characters per line;

k_2, the average number of printed characters (excluding spaces) per letter-shift block;

k_3, the average number of printed characters (excluding spaces) per figure-shift block and

k_4, the average number of spaces in the margin at the right of each line of print. A cost matrix, expressed in terms of k_1 through k_4, has been adopted.

In the formulation of the detection problem, Y_i has been defined as the decision to generate the teleprinter symbol X_i, so that c_{ii} is the cost of detecting the transmitted symbol. This cost is assigned the value zero. Errors involving only printed characters are assigned unit costs so that $c_{ij} = 1$ if X_i anx Y_j are different printed characters. When one printed character is transmitted and another received, the legibility of other characters in the output text is unaffected. This type of error may often be corrected by a human reader on the basis of his knowledge of grammar and spelling rules. Written English is very redundant (14).

The other errors involve control symbols and their costs are given in Table 5.1 as functions of k_1 through k_4. The table entries are based on a consideration of the consequences of the possible errors. When, for example, a printed character is detected in

Table 5.1

Cost Matrix

Received Symbol Y_j

Transmitted Symbol X_i

	Printed Character	Carriage Return	Line Feed	Letter Shift	Figure Shift
Printed Character	27 x 27 unit-cost submatrix	$\frac{1}{2}k_1$	1	$\frac{k_2+\frac{1}{2}k_3^2}{k_2+k_3}$	$\frac{k_3+\frac{1}{2}k_2^2}{k_2+k_3}$
Carriage Return	k_1-k_4	0	k_1-k_4	$k_1-k_4+\frac{\frac{1}{2}k_3^2}{k_2+k_3}$	$k_1-k_4+\frac{\frac{1}{2}k_2^2}{k_2+k_3}$
Line Feed	$2k_1$	$2k_1$	0	$\mathrm{Max}(2k_1, \frac{\frac{1}{2}k_3^2}{k_2+k_3})$	$\mathrm{Max}(2k_1, \frac{\frac{1}{2}k_2^2}{k_2+k_3})$
Letter Shift	k_2	$\mathrm{Max}(k_2, \frac{1}{2}k_1)$	k_2	0	k_2
Figure Shift	k_3	$\mathrm{Max}(k_3, \frac{1}{2}k_1)$	k_3	k_3	0

place of a carriage return, subsequent characters are printed in the margin of the current line until this margin is exhausted. The last position on the line is then overprinted until a carriage return is detected. If this error is isolated, an entire line is lost except for the characters which fill the margin so that, on the average, k_1-k_4 characters are rendered illegible. Other errors may heighten or lessen the effects of this one but second-order effects have been ignored in the assignment of the costs listed in Table 5.1. Appendix A states the rationale of each entry.

In addition to the constraint on detection of an elaborate cost matrix other measures may be adopted in order to protect system quality against the harmful effects of control-symbol errors. The telegraph code may be redesigned (45, 50) to reduce the likelihood of the most harmful errors or the mechanical operation of the teleprinter may be modified to prevent overprinting of characters. In a particular application, economic and other practical criteria determine the choice of a protection method.

5.2 Four detectors

Each of the four simulated detectors conforms to the Bayes decision rule of Section 2.2.6. Three of them are symbol detectors with different cost matrices and prior probability functions and the fourth is the binary detector described in Section 4.3. All of the detectors operate on the same set of measurements.

5.2.1 Minimum-risk detector

The minimum-risk detector is constrained by the cost matrix of Table 5.1 and a prior probability function which accurately represents the relative frequencies of the symbols of the source message.

5.2.2 Minimum-error detector

The fidelity criterion of the minimum-error detector is the maximization of the number of correctly detected symbols. The cost function of this device is the unit-cost matrix defined as

$$c_{ij} = 0 \quad \text{if } i = j$$

$$\text{and} \quad c_{ij} = 1 \quad \text{if } i \neq j . \tag{5.1}$$

For this detector, $t(Y_j)$, defined in Equation 2.11, may be written

$$t(Y_j) = \sum_{i=1}^{31} p(X_i)f_i(\overline{\alpha}) - p(X_j)f_j(\overline{\alpha}) \tag{5.2}$$

in which the first term is independent of j. The value of j for which $t(Y_j)$ is minimal is, therefore, the one for which

$$f_j p(X_j) \tag{5.3}$$

is maximal. This quentity is proportional to the posterior probability of X_j, defined in Equation 2.9, so that the minimum-error detector generates symbols with maximal posterior probability. Its decision represents a solution to the detection problem as it is often formulated in the context of statistical communication theory.

5.2.3 Maximum-likelihood detector

The third symbol detector performs according to the maximum-likelihood decision rule in that its output is a symbol for which $f_j(\overline{\alpha})$ is maximal. The operation of the maximum-likelihood detector conforms to the Bayes decision procedure relative to the uniform prior probability function ($p(X_i) = {}^1/31$ for all i) when the unit-cost fidelity criterion is imposed.

5.2.4 Binary detector

Because the simulated telegraph system includes no return path for ARQ signalling, the binary detector generates a special printed character when an erasure, i.e., a redundant code combination, is detected. For each signal element, the decision of this detector depends only on whether L, the binary likelihood ratio, defined in Equation 4.14, exceeds unity. The simulated binary detector thus operates according to a maximum-likelihood decision rule.

5.3 Plan of the simulation

The input and output of the computer program are paper-tape recordings of messages represented by the five-unit code of the Teletype Model 32 Page Printer. It is on this machine that source messages are produced and output messages converted to printed texts. Within the computer, the symbols of the source message are represented by the seven-unit fixed-ratio code used over radio links. The effects of the channel noise and fading are reflected in the statistical properties of the signal measurements. For each transmitted element, the simulated signal measurements are derived by means of transformations of a sequence of pseudo-random numbers. All of the detectors formulated in Section 5.2 operate on the same set of measurements so that differences in their output messages are due to their different computational procedures (i.e., decision rules) rather than differences in information extracted from the received signal.

In addition to simulating the communication of a telegraph message, the computer also monitors the performance of each detector and prints out the number of errors and total cost of each output message. While these measures provide one indication of the relative merits of the decision rules, the ultimate figure of merit of each detector is the usefulness of its message to a human reader. The relevance of the cost matrix of Section 5.1 must be judged on the basis of a human assessment of the messages produced by the minimum-risk detector. If this detector holds an advantage over the minimum-error device, its lower-cost messages must be more legible, in spite of their higher error rates, than the messages produced by the minimum-error detector.

5.4 Characterization of channel effects

The randomness of the hf channel causes the signal measurements to fluctuate in a manner which is characterized by a pair of conditional probability density functions. The parameters of these functions reflect the signal and noise levels of the channel. At a coherent receiver, the two measurements per element have independent normal distributions with the same variance. The output of the filter matched to the transmitted element has a mean value proportional to the square root of the signal-to-noise ratio.

The mean value of the other filter output is zero. At a non-coherent receiver the measurements are characterized by independent Rice distributions. The parameter, u, defined in Section 3.1.2, is proportional to the square root of signal-to-noise ratio in the probability distribution of the envelope of the filter matched to the transmitted element. The corresponding parameter of the other measurement is zero so that this measurement is Rayleigh distributed. The four conditional density functions are listed in Table 5.2 in terms of the normalized measurements α_o, α_1, γ_o, and γ_1 defined as

$$\alpha_n = (NT)^{-\frac{1}{2}} a_n$$

$$\text{and} \quad \gamma_n = (NT)^{-\frac{1}{2}} c_n, \quad \text{for } n = 0,\ 1. \qquad (5.4)$$

The introduction of the normalizing factor, $(NT)^{-\frac{1}{2}}$, a system constant, admits a concise expression of the probability densities.

When the signal amplitudes are subject to fading the signal-to-noise ratios are themselves sample values of a random variable. The nature of the fading determines the manner in which independent sample values are associated with the different measurements pertaining to a given symbol. If the fading is slow and flat, the same value of signal-to-noise ratio is the parameter

Table 5.2

Probability Distributions of Signal Measurements

Notation

The letter g denotes a conditional probability density function of a signal measurement at a coherent receiver; h is a probability density at a non-coherent receiver. At each type of receiver there are four relevant functions:

1. $g_o(\alpha_o)$ or $h_o(\gamma_o)$, the probability density of α_o or γ_o when "zero" is transmitted,

2. $g_o{}^*(\alpha_1)$ or $h_o{}^*(\gamma_1)$, the probability density of α_1 or γ_1 when "zero" is transmitted,

3. $g_1{}^*(\alpha_o)$ or $h_1{}^*(\gamma_o)$, the probability density of α_o or γ_o when "one" is transmitted,

4. $g_1(\alpha_1)$ or $h_1(\gamma_1)$, the probability density of α_1 or γ_1 when "one" is transmitted.

Formulas

$$g_n(\alpha_n) = \frac{1}{\sqrt{\pi}} \exp[-(\alpha_n - \sqrt{\rho_n})^2]$$

$$g^*_{1-n}(\alpha_n) = \frac{1}{\sqrt{\pi}} \exp(-\alpha_n{}^2)$$

$$h_n(\gamma_n) = 2\gamma_n \exp(-\rho_n - \gamma_n{}^2) I_o(2\gamma_n\sqrt{\rho_n})$$

$$h^*_{1-n}(\gamma_n) = 2\gamma_n \exp(-\gamma_n{}^2)$$

of all 14 measurements. In the case of frequency-selective fading statistically independent signal-to-noise ratios pertain to the two matched filters. For fast fading, a different signal-to-noise ratio is applicable to each time interval. For any type of fading the signal-to-noise ratios of successive message symbols are statistically independent.

Each of the simulated signal-to-noise ratios is proportional to the square of a Rayleigh-distributed amplitude. The average signal-to-noise ratio is defined in terms of the mean square amplitude, S, as

$$\rho = \frac{ST}{4N} \quad . \tag{5.5}$$

The individual values of ρ_o and ρ_1 are related through ρ to the standardized Rayleigh-distributed variables v_o and v_1. Thus

$$\rho_n = \tfrac{1}{2}v_n^{\,2}, \quad \text{for } n = 0,\ 1 \tag{5.6}$$

and the probability density function of v_n is

$$p(v_n) = v_n \exp\left(-\tfrac{1}{2}v_n^{\,2}\right) \quad . \tag{5.7}$$

5.5 The simulated measurements

The multiplicative and additive channel effects are simulated by the generation of independent normal, Rice or Rayleigh random deviates to represent the signal-to-noise ratios and measurements associated with each signal element. Random deviates of each type are obtained by means of arithmetic operations on members of a sequence of pseudo-random numbers, uniformly distributed over the unit interval. The following congruence formula is used to generate the sequence of pseudo-random numbers (75):

$$U_{i+1} = 2^{-35} \left[(129\, U_i + 1) \text{ modulo } 2^{35}\right] . \quad (5.8)$$

The number sequence

$$U = U_o,\ U_1,\ U_2,\ \ldots$$

may begin with any 35-bit binary fraction. The sequence contains all such fractions and is cyclic with a period of 2^{35} numbers. The numbers generated by Equation 5.8 satisfy several statistical tests of randomness (75).

A set of Rayleigh random deviates, V, with mean square of unity may be obtained from U by means of the transformation

$$V_i = \sqrt{-2 \log U_i} \quad . \qquad (5.9)$$

The numbers

$$V_i' = V_i \sqrt{S} \qquad (5.10)$$

are Rayleigh distributed with a mean square of S.

The projections on a pair of orthogonal axes of a random vector with Rayleigh distributed magnitude and uniformly distributed phase are independent normal random variables. A sequence of normal random deviates, Z, may therefore be generated by the transformation:

$$Z_{2i} = V_{2i} \sin 2\pi\, U_{2i+1}$$

$$Z_{2i+1} = V_{2i} \cos 2\pi\, U_{2i+1}\,, \quad i=0,1,2,\ldots \qquad (5.11)$$

The mean value of Z is zero; its variance is unity. The numbers

$$Z_i' = \mu + \sigma Z_i \qquad (5.12)$$

have mean value μ and variance σ^2.

The resultant of a constant vector and a random vector with Rayleigh-distributed amplitude and uniformly distributed phase is a vector with Rice-distributed magnitude. The numbers

$$R_i = \sqrt{Z_{2i}{}^2 + (Z_{2i+1} + u)^2}\,, \quad i=0,1,2,\ldots (5.13)$$

are characterized by the probability density function

$$R_i \exp\,[-\tfrac{1}{2}(u^2+R_i{}^2)]\; I_o\,(uR_i).$$

If $u = u'/\sqrt{S}$, the numbers

$$R_i' = R_i\sqrt{S} \qquad (5.14)$$

are distributed according to the general Rice probability density

$$\frac{R_i'}{S} \exp\left[-\frac{(u')^2 + (R_i')^2}{2S}\right] I_o\left(\frac{u'R_i'}{S}\right) .$$

Equations 5.9 to 5.14 specify the transformations of the pseudo-random numbers into the random deviates which simulate the signal-to-noise ratios of a fading signal and the measurements associated with each element.

5.6 Computational requirements of the detectors

It is demonstrated in this section that the computational requirements of the four detectors of Section 5.2 may be met by arithmetic operations which are less complicated than those suggested by the formulas of Section 2.2.6. In the computer simulation, it is assumed that the measure of channel conditions available to the detectors consists of a single amplitude statistic, A, and the noise power density N. In the case of steady-signal propagation, A is the exact amplitude of the two sinusoids assigned to each signal element. When the signal is fading, $A = \sqrt{S}$, the rms value of the fading amplitudes. The simulated detectors operate on the normalized measurements α_o and α_1 or γ_o and γ_1 defined in Equation 5.4. The formulas

4.8 and 4.9 for the element likelihoods require the weighting of a_n and c_n by $2A/N$. The corresponding coefficient of α_n and γ_n is $2\sqrt{\rho}$, where ρ, the signal-to-noise ratio is defined as

$$\rho = \frac{A^2 T}{4N} .$$

5.6.1 The binary detector

At the binary detector, the seven code elements of the detected symbol are generated independently. For a particular element the output of the detector is "one" if $\alpha_1 > \alpha_o$ or $\gamma_1 > \gamma_o$; otherwise the output is "zero". A teleprinter symbol is generated whenever the sequence of seven output elements is a member of the code alphabet. When a redundant sequence is detected, the output is an erasure which appears in the printed text as the special punctuation symbol ə .

5.6.2 The symbol likelihood function

The decisions of the three symbol detectors depend upon the 31 numbers which comprise the symbol likelihood function $f_X(\overline{\alpha})$. The formula for each likelihood, $f_i(\overline{\alpha})$ is given in Equation 4.1 as the product of seven of the fourteen element likelihoods. The identity of the seven numbers which must be multiplied to give $f_i(\overline{\alpha})$ depends on the code representation of X_i. At a coherent receiver, the element likelihoods are given by Equation 4.8 and the symbol likelihoods may be expressed as

$$f_i(\overline{\alpha}) = \exp(-7\rho)\left[\prod_{k=1}^{7} \eta_k\right]\exp\left\{2\sqrt{\rho}\sum_{k=1}^{7}\left[(1-b_{ik})\alpha_{ok} + b_{ik}\alpha_{1k}\right]\right\} \qquad (5.15)$$

where the second subscript k has been added to α_o and α_1 to denote the measurements pertaining to the kth element. The decisions of the three symbol detectors depend on the relative values of the posterior risks, the posterior probabilities and the symbol likelihoods. All of these numbers are proportional to

$$\eta = \exp(-7\rho)\left[\prod_{k=1}^{7} \eta_k\right] \qquad (5.16)$$

the factor of $f_i(\overline{\alpha})$ which is independent of i. It follows that this factor may be ignored in the three decision procedures. Instead of calculating $f_X(\overline{\alpha})$ explicitly, the detectors at a coherent receiver may operate instead on the set of numbers

$$x_i = \exp\left\{2\sqrt{\rho}\sum_{k=1}^{7}\left[(1-b_{ik})\alpha_{ok} + b_{ik}\alpha_{1k}\right]\right\} . \qquad (5.17)$$

At a non-coherent receiver, the element likelihoods are given by Equation 4.9 and each symbol likelihood is

$$f_i(\overline{\alpha}) = \eta \prod_{k=1}^{7} I_o\left\{2\sqrt{\rho}\left[(1-b_{ik})\gamma_{ok}+b_{ik}\gamma_{1k}\right]\right\} . \qquad (5.18)$$

The product of seven modified Bessel functions appearing in 5.18 represents a formidable computational requirement. For practical purposes, the rigors of 5.18 may be reduced considerably by the approximation

$$\log I_o(\chi) \approx \frac{1}{4}\chi^2 \qquad (5.19)$$

discussed in Section 8.5. If 5.19 is accepted the symbol likelihoods at a non-coherent receiver become functions of the squared measurements. Thus

$$f_i(\overline{\alpha}) = \eta \exp\{\rho \sum_{k=1}^{7} [(1-b_{ik})\gamma_{ok}^2 + b_{ik}\gamma_{1k}^2]\} \quad .(5.20)$$

The quantity η may be factored from this formula too so that the detectors at a non-coherent receiver need only consider the 31 numbers

$$y_i = \exp\{\rho \sum_{k=1}^{7} [(1-b_{ik})\gamma_{ok}^2 + b_{ik}\gamma_{1k}^2]\} \quad . (5.21)$$

5.6.3 <u>The maximum-likelihood detector</u>

The decision of this detector is Y_{j*} where j* is the subscript of a maximal member of $f_i(\overline{\alpha})$, i.e., a maximal x_i or y_i. The symbol for which either of these statistics is maximal is the one for which its logarithm is maximal so that the maximum-likelihood detector needs only to compute 31 numbers proportional to $\log x_i$ or $\log y_i$. These numbers may be derived from Equations 5.17 and 5.21:

$$x_i^* = \sum_{k=1}^{7} [(1-b_{ik})\alpha_{ok} + b_{ik}\alpha_{1k}] \qquad (5.22)$$

at a coherent receiver and

$$y_i^* = \sum_{k=1}^{7} [(1-b_{ik})\gamma_{ok}^2 + b_{ik}\gamma_{1k}^2] \qquad (5.23)$$

at a non-coherent receiver. No measure of the channel conditions is necessary for determining the output of the maximum-likelihood detector. The output is a symbol for which the appropriate sum of measurements (at a coherent receiver) or the sum of their squares (at a non-coherent receiver) is maximal.

5.6.4 The minimum-error detector

The output of this detector is a symbol for which 5.3 is maximal. The logarithm of this quantity is a monotonic increasing function of

$$\log p(X_i) + \log x_i \qquad (5.24)$$

at a coherent receiver and

$$\log p(X_i) + \log y_i \qquad (5.25)$$

at a non-coherent receiver. The symbol generated by the minimum-error detector is one for which 5.24 or 5.25 is maximal. This symbol may be identified by the detector if it computes a set of numbers proportional to 5.24 or 5.25. An appropriate set of numbers is

$$x_i{}^* + \frac{1}{2\sqrt{\rho}} \log p(X_i) \tag{5.26}$$

at a coherent receiver or

$$y_i{}^* + \frac{1}{\rho} \log p(X_i) \tag{5.27}$$

at a non-coherent receiver. As the signal-to-noise ratio increases, the detector attaches relatively more importance to $x_i{}^*$ or $y_i{}^*$, which reflects the information in the received signal, than it does to $p(X_i)$, which reflects the prior information. At high signal-to-noise ratios the outputs of the maximum-likelihood detector and the minimum-error detector are very likely to be identical.

5.6.5 <u>The minimum-risk detector</u>

When the detection process is constrained by a complicated cost function, a set of numbers related to the 31 posterior risks must be calculated. The posterior risk of Y_j is proportional to

$$r_j = \sum_{i=1}^{31} p(X_i) c_{ij} x_i \tag{5.28}$$

at a coherent receiver and

$$r_j' = \sum_{i=1}^{31} p(X_i) c_{ij} y_i \tag{5.29}$$

at a non-coherent receiver. Each number r_j or $r_j{}'$ is a sum of 31 terms where each term contains a sum of seven measurements or a sum of seven squares of measurements. The output symbol is associated with a minimal value of 5.28 or 5.29.

In the special case of the cost function of Section 5.1, the minimum-risk detection procedure may be simplified because of the nature of the costs related to the 27 printed characters. It will be demonstrated that a printed character with minimal posterior risk is one having maximal posterior probability. If the symbols $X_1, X_2, \ldots, X_{27}$ are the printed characters and $X_{28}, \ldots, X_{31}$ are the control symbols, the function $t(Y_j)$, defined in Section 2.2.5, may be written as

$$t(Y_j) = \sum_{i=1}^{27} f_i(\overline{\alpha})p(X_i)c_{ij} + \sum_{i=28}^{31} f_i(\overline{\alpha})p(X_i)c_{ij} \quad . \tag{5.30}$$

If $1 \leq j \leq 27$, the costs in the first summation of 5.30 are given by Equation 5.1 and this summation may be expressed as

$$\sum_{i=1}^{27} f_i(\overline{\alpha})p(X_i)$$ [redacted] $$- f_j p(X_j)$$

in which the summation is independent of j and the other term is proportional to the posterior probability of X_j. The second summation of 5.30 is also independent of j because the cost of detecting a printed character when a control symbol is transmitted depends only on the identity of the control symbol, i.e., the subscript i. Equation 5.30 may thus be written as

$$t(Y_j) = \text{const} - f_j(\overline{\alpha})p(X_j), \quad j=1,2,\ldots,27 \quad (5.31)$$

which demonstrates that Y_j for which $t(Y_j)$ is minimal is the printed character associated with maximal posterior probability.

The minimum-risk detector may therefore operate as a minimum-error device with respect to the 27 printed characters and then perform the general minimum-risk procedure with respect to the four control symbols and a printed character whose posterior probability is maximal.

5.7 Example of a simulated transmission

The previous sections of the present chapter have specified: 1) a cost function based on four message-format parameters, 2) definitions of four detectors, 3) a means of simulating the transmission of a message and 4) the computational requirements of the detectors. The implementation of the minimum-risk and minimum-error detectors requires, in addition to these specifications, a numerical representation of the prior information in the form of a set of message-format parameters and a set of prior probabilities. In a practical situation, this implementation may be impeded by the difficulty of selecting sets of numbers which adequately represent the prior information of every message which may be transmitted. In some applications, however, the message structure may be constrained to conform approximately to a known prior-information representation or the parameters of each message may be obtained at the receiver.

The simulation demonstrated in this section has the aim of investigating the effects of using this information when it is available in idealized circumstances. The detectors are provided, therefore, with precise knowledge of the message structure by means of a computer count of symbol frequencies and format parameters.* These quantities provide the basis of a cost matrix and a prior probability function which are precisely relevant to the transmitted message.

The output messages presented in this section result from the simulated transmission over a non-fading channel of an English text containing 1069 symbols. Coherent reception has been assumed and the simulation repeated for signal-to-noise ratios between 2.0 and 6.0. In each case the same set of pseudo-random numbers has been used to simulate the effects of channel

* In the example presented, this computer count was based on the message actually transmitted.

noise. The physical analogue of this procedure is the occurrence of the same noise waveform regardless of the channel attenuation. For a given signal-to-noise ratio, all four detectors operate on the same set of signal measurements so that the differences in their output messages result from the effects of the different computational procedures.

Appendix B contains the original message and the outputs produced by the four detectors operating at five signal-to-noise ratios. Following these messages is a measure of the simulation results in the form of a table of costs and error rates. In the present discussion, some of the properties of the detectors are demonstrated by reference to the detection results corresponding to the following single line of the source message:

TELEPRINTER USE IS THE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE

Segment of transmitted message

The influence of the prior information on the detector outputs varies inversely with signal-to-noise ratio. At the lowest simulated signal-to-noise ratio, 2.0, the information in the received signal is least reliable and the detection results reflect to the greatest extent the effects of the computational procedures. At this signal-to-noise ratio the above line appears in the maximum-likelihood message as:

TELEPRIYTER USE HS YHE FLV -13) .3,5 ?8,+4/. AT SEVEGAL STAGES N THE

Maximum-likelihood detector $\rho=2.0$

Of the 72 symbols (68 printed characters and four shifts) in this line, 11 or just over 15% (compared with 16% in the message as a whole) are incorrect. The most serious error is the numeral 1 detected instead of a letter shift after the character "-". This error has caused the next 14 characters to be printed in the wrong mode and thus appear nonsensical* in spite of the fact that all but three of them are correct.

The 72 symbols generated by the minimum-error detector,

TECEPRIYTER USE HS NHE FLV -13) 83.5 ?8,346. AT SEVERAL STAGES IN THE

Minimum-error detector $\rho=2.0$

have a similar appearance to those in the maximum-likelihood message. If the two devices reach different decisions about a given symbol, the minimum-error output is always a symbol which occurs more frequently in the source message than the symbol generated by the maximum-likelihood detector. Thus in the word SEVERAL, the R which occurs with relative frequency .047 was correctly detected by the minimum-error detector. The maximum-likelihood detector, on the other hand, generated SEVEGAL; the relative frequency

* An exception to this statement is found in the case of a trained reader who may readily translate these characters to the letters mode.

of G is .010. In this case the knowledge of the relative frequencies prevented the error which resulted from an interpretation of the received signals alone. On the other hand, the L in TELEPRINTER was correctly detected by the maximum-likelihood detector. The minimum-error detector generated the more frequently occurring O instead. Minimum-error detection thus results in gains and losses in accuracy relative to maximum-likelihood detection. On the average the gains outnumber the losses because the most frequent symbols are favored. In the illustrated line of the minimum-error message, there are 10 errors, one fewer than in the corresponding segment of the maximum-likelihood message. The 14% error rate of this line of the minimum-error message is approximately the same as the error rate of the entire message.

An interesting error in the illustrated segment of the maximum-likelihood message is the line-feed symbol detected in place of the I in IN. This error is absent from the output of minimum-error detector and as a result the message segment of this detector has the more correct appearance to a reader. This error, however, provides an insight into a deficiency of the minimum-error detector. Because the control symbols occur relatively infrequently in the source message, the minimum-error detector is inherently biased against generating control symbols. Although this bias prevented the error of the maximum-likelihood detector in deciding line feed instead of I, a more serious error was committed by the minimum-error detector at the end of the illustrated line. After detecting carriage return, this detector failed to generate the subsequent line feed and the following line overprinted the illustrated segment. The appearance in the output message of these two lines of print is thus:

Minimum-error detector ρ=2.0 (in context of entire message)

All of the characters on the two lines are mutilated.

The configuration of its cost function insures that the minimum-risk detector operates most reliably in detecting control symbols. There is only one character lost when line feed is detected in place of a printed character while the omission of a line feed brings a serious penalty. The minimum-risk detector therefore tends to accept the low-cost error of spuriously inserting a line feed and avoid the high-cost error of omitting one. The result is an output message with more line-feed symbols than there are in the source message. The minimum-risk version of the illustrated line thus has two extra line feeds:

```
TEOEPRIYTER USE HS NHE FLV -EL IE
                                   T BINERY. AT SEVERL STAGES
                                                              N THE
```

<u>Minimum-risk detector ρ=2.0</u>

This line is followed by a correctly detected carriage-return-line-feed combination so that the overprinting which disfigures the minimum-error message is absent from the minimum-risk output.

The minimum-risk detector is also biased toward the detection of letter shift because most of the printed characters are in the letters mode. Thus the word BINARY, illegible in the other output messages, is easily discerned (even though the A is replaced by E) in the minimum-risk output. Although the characters corresponding to the word ELEMENT are printed in the correct mode, the word is unrecognizable because of the three detection errors among its seven letters. The tendency of the minimum-risk detector to generate letter shift has caused this symbol to appear in the output message in place of the A of SEVERAL. This error was avoided by the other detectors but it carries a low cost and is not serious in its context.

Twelve symbols, i.e. 16%, of this minimum-risk output line have been incorrectly detected. In the entire message the same percentage of the minimum-risk symbols are errors. In detecting this line, the minimum-risk detector, in spite of the fact that its error rate is highest among the three symbol detectors, has generated more useful information than either of the other two symbol detectors. The least legible of the three illustrated output lines is the one produced by the minimum-error detector. Although it contains the fewest errors, this line is illegible because it is overprinted by its successor.

The message of the binary detector conveys less information than any of the symbol-detector messages even though only 4% of its symbols are incorrect. The proportion of erasures is 44% and consequently many of the control symbols are omitted so that the printed characters are often overprinted or printed in the wrong mode. The illustrated line would thus appear as:

5eeee48656e4e7'3 e' 6e3 e)=e-e3)ee3ee 7e,e4e.ee-5e'ee3ee)e'e-e3' eeee8

Binary detector p=2.0

except for the fact that the first four lines of the message (in which this one is included) all overprint one another with the result:

ØØYØØ

Binary detector ρ=2.0 (in context of entire message)

All of the foregoing illustrations are examples of detection results at a signal-to-noise ratio of 2.0. As the signal amplitude is increased, longer blocks of each output message become legible and the four messages become more similar. For signal-to-noise ratios of 3.0 and higher, the illustrated line is no longer overprinted in the minimum-error message and the characters on this line are identical to the corresponding outputs of the minimum-risk detector. At the signal-to-noise ratios of 3.0 and 4.0, the same two errors appear in the output line of the two detectors. The line is detected as:

TELEPRIYTER USE IS NHE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE

Minimum-error and minimum-risk detectors ρ=3.0, ρ=4.0

Both errors are absent when the signal-to-noise ratio is 5.0 or higher. With the signal-to-noise ratio equal to 3.0, 4.0, and 5.0, the maximum-likelihood line segment contains three, two and one error respectively:

TELEPRIYTER USE IS YHE FLVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE

<u>Maximum-likelihood detector ρ=3.0</u>

TELEPRIYTER USE IS YHE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE

<u>Maximum-likelihood detector ρ=4.0</u>

TELEPRIYTER USE IS THE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE

<u>Maximum-likelihood detector ρ=5.0</u>

The corresponding segments of the binary-detector outputs as they appear in the complete messages are:

[illegible]AGUBE$W THE

<u>Binary detector ρ=3.0</u> (in context of entire message)

TELEPRI■TER USE IS ■HE FIVE-ELEMENT BIN■R■. AT S■VERAL STAGES IN THE

<u>Binary detector ρ=4.0</u>

TELEPRI■TER USE IS ■HE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE

<u>Binary detector ρ=5.0</u>

There are no incorrect symbols in the illustrated line of the binary-detector output when the signal-to-noise ratio is 4.0 or higher.

Among the complete output texts in Appendix B, those generated by the three symbol detectors at a signal-to-noise ratio of 6.0 are all clearly legible throughout. The only errors in the three messages are errors involving printed characters. The minimum-risk and minimum-error messages are identical with three errors each and the maximum-likelihood message contains four errors. The binary-detector output contains only two incorrect symbols but two of the 56 erasures replace line feeds and as a result four of the message's 17 lines are mutilated due to overprinting. With the signal-to-noise ratio reduced to 5.0, the minimum-error and maximum-likelihood detectors both omitted a line feed and thus produced messages with two illegible lines. The seven errors generated by the minimum-risk detector at the 5.0 signal-to-noise ratio involve only printed characters so that this entire message remains legible. At this signal-to-noise ratio the minimum-risk procedure has protected the system against the control-symbol error which disfigures the outputs of the other detectors.

With the signal-to-noise ratio further decreased, all of the detectors committed errors involving control symbols. As the channel conditions deteriorate, the messages differ increasingly in detail from one another. The legibility of all of them decreases and the differences in message quality tend to become indistinguishable.

5.8 Evaluation of results

The quality objective of the minimum-risk detector is maximization of the legibility of output texts by their human readers. The cost matrix of Section 5.1 has been formulated in order to reflect the effects of detection errors on message legibility. Because the costs were specified intuitively, it is not certain that the total cost of a text is an accurate measure of its usefulness to a reader. If the relative merits of the four detectors are to be assessed rigorously on the basis of the legibility criterion, a program of operational ~~subjective~~ testing must be undertaken. The tests would involve the presentation of a set of output messages to potential users of the system and the measurement of each reader's comprehension of the texts. Although such a test program is beyond the scope of the work reported in this thesis, an informal examination of the messages of Appendix B suggests certain preliminary conclusions about the relative merits of the detectors.

The advantage of the minimum-risk detector over the other symbol detectors is most evident in the messages produced at the 5.0 signal-to-noise ratio. In this condition only the minimum-risk output is free of control-symbol errors. The signal-to-noise ratio must be 6.0 for this type of error to be eliminated from the other output texts. In this single example, a minimum-risk detection has thus reduced by about 1 db the signal-to-noise ratio necessary for the elimination of control-symbol errors. At lower signal-to-noise ratios the advantage of minimum-risk detection is less evident to a casual reader and it is under these conditions that the subjective tests are necessary to distinguish among detector merits. It has been observed in general that the outputs of all the detectors collectively are more valuable to a reader than any one message. When the received signal is unreliable, the reader uses his knowledge to select one symbol from the alternatives offered by the four detectors.

From a practical point of view, the simulated transmissions show the minimum-risk detector to be less promising than might have been expected. The 1 db reduction in the signal-to-noise ratio required to prevent control-symbol errors is not very significant and an inspection of the lower-quality messages does not immediately indicate that the improvement in legibility offered by minimum-risk detection justifies its complex computational facilities. In most practical situations a mechanical modification of the receiving teleprinter in order to prevent overprinting of symbols would probably offer a more economical means of protection against control-symbol errors than the introduction of a minimum-risk detector.

Among the other detectors, it is evident that at any of the signal-to-noise ratios, the binary detector offers the least acceptable results. Although this device generates the fewest incorrect symbols, the high proportion of erasures in its output (5% even at the 6.0 signal-to-noise ratio) leads to a high cost because of the control symbols affected. The outputs of the minimum-error and maximum-likelihood detectors appear quite similar and in fact their costs are nearly equal at all of the signal-to-noise ratios.

The unjustified complexity of minimum-risk detection, the inferior performance of the binary detector and the similarity in effects of minimum-error and maximum-likelihood detection combine to suggest that for practical systems, the most promising of the simulated detectors is the maximum-likelihood device with its relatively simple computational requirements. The properties of this detector are investigated in detail in the next three chapters where it is shown that the fixed-ratio character of the telegraph code implies a particularly simple computational role for a maximum-likelihood detector.

CHAPTER 6 : MAXIMUM-LIKELIHOOD DETECTION

Of the three symbol detectors formulated in Chapter 5, the simplest is the maximum-likelihood detector. Its decisions depend only on the relative values of the likelihoods $f_i(\overline{\alpha})$, which together represent the relevant information conveyed by the received waveform. No cost matrix or prior probability function need be specified in the design of a maximum-likelihood detector. It is a Bayes detector relative to the uniform prior probability function if the costs of all errors are equal. Thus the maximum-likelihood detector offers optimal error performance when all symbols are transmitted with equal frequency.

Maximum-likelihood detection is particularly relevant to systems using a fixed-ratio code because this code constraint admits a marked simplification of the computational requirements of the detector. The likelihoods need not be computed explicitly; instead, output symbols may be generated on the basis of the relative values of the seven binary likelihood ratios defined in Chapter 4 (6, 44, 51). In the present chapter, the configuration of the maximum-likelihood detector is described and compared with that of the binary detector of practical one-way telegraph systems.

The performance characteristics of the two systems are compared and the practical applicability of the maximum-likelihood detector is discussed. Chapter 7 considers the applicability to two-way systems and Chapter 8 contains a formal derivation of the properties of the maximum-likelihood detector.

6.1 The decision rule

Although the maximum-likelihood decision rule is defined (in Section 5.2) in terms of the set of conditional probability densities $f_i(\overline{\alpha})$, the explicit computation of these numbers is not essential to the implementation of the rule. When telegraph symbols are represented by the seven-unit fixed-ratio code, the most likely symbol Y_{j*} may be determined on the basis of the relative values of the seven binary likelihood ratios L_k ($k = 1, 2, \ldots, 7$), defined in Section 4.3. The maximum-likelihood decision rule may be stated as:

> Y_{j*} is the symbol whose code sequence has a "one" in the code positions corresponding to the four highest values of L_k and a "zero" in the other three positions.

The code word for Y_{j*} is, therefore, one of the 35 seven-unit combinations which satisfy the four-out-of-seven constraint. Thirty-two of these combinations represent teleprinter symbols and the other three may be used for signalling within the telegraph network. In general, however, one or more of these extra combinations is redundant, a situation which admits the possibility of the channel erasing a transmitted symbol by generating a set of data which results in the detection of a redundant code sequence. This type of erasure does not occur very often and in the performance characteristics presented in this chapter and in Chapter 7, it is counted as an error which causes the wrong teleprinter symbol to be generated. The detector specified in this section determines, therefore, the most likely of the 35 code words which satisfy the fixed-ratio constraint. It does not take advantage of the possibility that as many as three of them may be redundant.

6.2 Data processing

The performance of the maximum-likelihood detector is unchanged if $\hat{L}_k$, a strictly monotonic function of the likelihood ratio, is substituted for L_k in the decision rule specified in Section 6.1. The same symbols are detected when this modified decision rule is applied because the ordering of the seven values of $\hat{L}_k$ is identical to the ordering of the L_k. The detector whose performance characteristics are presented in Section 6.4 does not measure the signal amplitudes A_o

and A_1, but instead assumes them equal. In this event, the formulas in Section 4.1 for $\lambda_o^{(k)}$ and $\lambda_1^{(k)}$ lead to the adoption of the function

$$\hat{L}_k = \frac{N}{\text{[illegible]}A} \log L_k \qquad (6.1a)$$

at a coherent receiver and

$$\hat{L}_k = 4\left(\frac{N}{\text{[illegible]}A}\right)^2 \log L_k \qquad (6.1b)$$

at a non-coherent receiver.

At a coherent receiver, Equation 6.1a becomes

$$\hat{L}_k = a_1 - a_o \qquad (6.2)$$

and the computation required by the detector includes the subtraction, for each element, of one measurement from the other and the storage of the result. When all seven differences have been recorded the detector must determine the four highest of the stored numbers and generate a symbol according to the rule specified in Section 6.1.

The situation at a non-coherent receiver is more complicated. In this case,

$$\hat{L}_k = 4\left(\frac{N}{\text{[illegible]}A}\right)^2 \log \frac{I_o\left(\frac{\text{[illegible]}A}{N} c_1\right)}{I_o\left(\frac{\text{[illegible]}A}{N} c_o\right)} \qquad (6.3)$$

a function which is not easily computed. A practical solution to the problem posed by the complexity of 6.3 is the approximation of this formula by

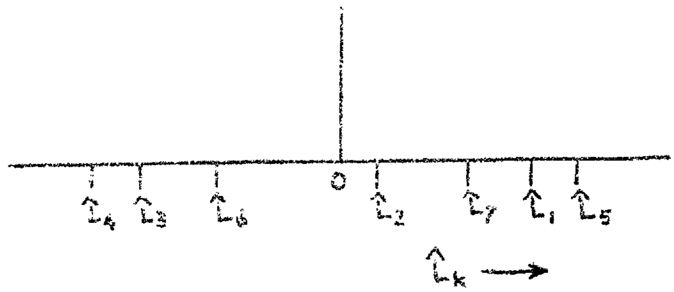
$$\hat{L}_k = c_1^2 - c_o^2 \qquad (6.4)$$

This is not, unfortunately, a strictly monotonic function of L_k and for certain sets of signal measurements, the ordering of the seven values of 6.4 is not precisely the same as the ordering of the L_k. On the other hand, the maximum-likelihood detector is most susceptible to error when A/N is low and c_o and c_1 are nearly equal for two or more elements. In this event the measurements are likely to have low values and it is under this condition that 6.4 most nearly approximates a monotonic function of L_k. For this reason and because of its simple computational requirement relative to 6.3, the square-law formula of Equation 6.4 is accepted as the basis of a practical maximum-likelihood detector at a non-coherent receiver. It is for this device that the performance characteristics of Section 6.5 have been calculated. The validity of approximating Equation 6.3 by 6.5 is discussed in more detail in Section 8.5.

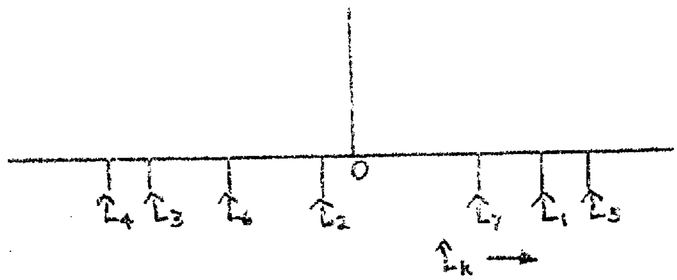
6.3 Relation to binary detection

The binary detectors described in Section 4.3 make an independent decision about each of the transmitted signal elements. The decision of the simplest binary detector depends only on an assessment which of the two signal measurements is greater. The operation of this device may be characterized as a test of the $\hat{L}_k$ defined in Equations 6.2 and 6.4 (for coherent and non-coherent reception respectively). A "one" is detected when $\hat{L}_k > 0$ and a "zero" otherwise. For a set of seven binary decisions to result in the detection of a telegraph symbol, it is necessary that four of the $\hat{L}_k$ be positive and the other three zero or negative. When a set of measurements meets this requirement, the symbol generated by the binary detector is identical to the one derived by a maximum-likelihood detector operating on the same data. When the measurements do not satisfy this condition, the binary detector generates an erasure while the maximum-likelihood device continues to detect acceptable code sequences.

Figure 6.1 illustrates two sets of $\hat{L}_k$ which may result from the transmission of a telegraph symbol. In a), $\hat{L}_1$, $\hat{L}_2$, $\hat{L}_5$ and $\hat{L}_7$ are all positive and $\hat{L}_3$, $\hat{L}_4$ and $\hat{L}_6$ negative, so that both detectors generate



a) Both detectors generate 1100101 (A or -)



b) Binary detector generates 1000101 (erasure)
Maximum-likelihood detector generates 1100101 (A or -)

FIGURE 6.1 Comparison of detection methods

1100101 (the character A or -). In b), the values of $\hat{L}_k$ have the same order but in this case $\hat{L}_2$ is negative. The binary detector generates 1000101 -- an erasure -- while the output of the maximum-likelihood detector is again 1100101.

Viewed as a black box the maximum-likelihood detector is identical to the simpler binary detector when four and only four of the $\hat{L}_k$ are positive. If fewer than four are positive, the maximum-likelihood detector may be considered to be a device which inverts the mostdoubtful "zeros" of the binary detector's output code sequence. These are the "zeros" for which $\hat{L}_k$ is least negative. When more than four "ones" are generated by the binary detector, the maximum-likelihood detector inverts those for which $\hat{L}_k$ is least positive.

This functional description of the maximum-likelihood detector implies that not all of the computations specified in Section 6.2 are required for the detection of every symbol. The seven pairs of measurements (or the set of $\hat{L}_k$) must be stored but the explicit calculation of $\hat{L}_k$ and the determination of the highest four numbers in a set of seven is necessary only when the binary detector generates an erasure. This part-time requirement for the more complex computations suggests the possibility of time-sharing one computational facility among several detectors.

6.4 Quality-assessment mechanism

Although the maximum-likelihood detector performs a relatively complicated computation which, in effect, translates the erasures generated by a binary detector into telegraph symbols, it does not necessarily follow that the resulting messages are better than those producted by the binary detector alone. The symbols which are erased by a binary detector are the ones most susceptible to error at a maximum-likelihood detector. If too many erasures were to be translated incorrectly, the value to a reader of the maximum-likelihood message would be lower than that of the message produced by the binary detector. The pattern of erasures, furthermore, in the binary detector's message is an indication to the message reader of the reliability of the printed characters. The performance characteristics presented in Section 6.5 demonstrate that the number of incorrect symbols in the output of a binary detector increases with the number of erasures. At a maximum-likelihood detector only telegraph symbols are generated; there are no erasures to indicate the reliability of the printed characters.

In spite of these mitigating factors, the simulated messages in Appendix B indicate, in terms of both the tabulated costs and an ~~subjective~~ operational assessment of the output messages, that messages produced by the maximum-likelihood detector are more legible than those generated by the binary detector. The deficiency of the binary detector derives from the fact that the detection of an erasure in place of a control symbol is, in general, as harmful as the detection of an incorrect teleprinter symbol. A measure of the relative merits of the two detectors is proposed in Equation 6.5 and discussed in Section 6.6 in terms of calculated performance curves.

Beyond any advantage indicated by this measure and the messages of Appendix B, the relative value of a maximum-likelihood detector may be enhanced by extending its scope to include a "quality-assessment" mechanism. This mechanism would add to the maximum-likelihood output messages the useful information provided by the erasure symbols in a binary detector's message. A detector incorporating the quality-assessment mechanism would provide, with each output symbol, a binary assessment of the symbol's reliability. The assessment may be based on the relationship between binary detection and maximum-likelihood detection. In this event a symbol would be specified as "relatively certain" if the same symbol was generated by a binary detector and "relatively uncertain" if an erasure was generated. A more general assessment may be based on the actual values of the $\hat{L}_k$. If the difference between the fourth and fifth highest $\hat{L}_k$ exceeds some threshold, the output symbol is relatively certain; otherwise it is relatively uncertain.*

---

* This is the assessment proposed by Barrow (6) for a generalised erasure mechanism in which the maximum-likelihood detector would generate an erasure when the threshold is not exceeded. This erasure would be treated as if it occurred in a binary detector -- either printed in the message or used to initiate an ARQ procedure.

In practice the quality assessment could be indicated to a message reader by the manner in which a symbol is printed. A teleprinter with a two-color typewriter ribbon may be modified to print the relatively uncertain symbols in red and the others in black. The number of red characters is an indication of the overall reliability of the message. A high proportion of relatively uncertain symbols indicates a low signal-to-noise ratio and thus suggests a higher error rate -- even among the relatively certain symbols. For any signal-to-noise ratio, the set of red characters contains proportionally more errors than the set of black ones so that the specific characters printed in red are those which are most likely to be incorrect.

The maximum-likelihood detector with a quality-assessment facility represents a solution to an extended decision problem (50) in which the set of decisions (output symbols) contains more members than there are states of nature (message symbols). The relatively certain and relatively uncertain outputs are sets of different symbols and if different costs are assigned to them, the scope of the minimum-risk detector may be extended to generate messages which provide a quality assessment with each symbol. Owing to the greatly increased amount of computation that a quality-assessment facility would entail it is less attractive as an extension of a minimum-risk detector than it is in the case of maximum-likelihood detection where only a nominal increase in computation is required.

Messages which include a quality assessment contain more information about the received signals than messages which do not. With this facility they approach to some extent the outputs of a sufficient detector (73) which presents to a reader the posterior probabilities of all possible symbols rather than a decision to detect one of them. The output of a sufficient detector offers a reader more information than he can effectively use. The quality-assessment mechanism makes a fraction of this information available and presents it in a concise form which may be readily interpreted.

The practicability of the quality-assessment proposal is limited by the fact that its introduction to a system would require a modification of receiving teleprinters and changes in the signalling format used over the feeder lines between radio terminal and teleprinters. In this respect it differs from the other detection schemes thus far proposed. On the other hand, it is similar to the other schemes in that it requires no alteration of the teleprinter and signalling format at the transmitting end of the system.

The quality-assessment facility may thus be offered to message readers who are willing to pay for its extra information. Its introduction does not require the general reorganization of an entire telegraph network.

6.5 Performance characteristics.

The curves presented in this section demonstrate the error immunity of binary detectors and maximum-likelihood detectors under the eight propagation conditions listed in Table 3.1. Each of the figures 6.3 through 6.10 applies to one of the conditions and shows three error rates as functions of signal-to-noise ratio. Two of the curves indicate the performance of a binary detector and the third pertains to maximum-likelihood detection. At a binary detector, there are three possible results of the transmission of a telegraph symbol:

1) it is detected correctly,
2) an incorrect symbol is detected,
3) an erasure is detected.

The probabilities of these results are given as Curves A and B. The function $B(\rho)$ is the probability of result 2) -- the detection of an incorrect symbol. $A(\rho)$ is the probability that the correct symbol is not detected. This is the probability that result 1) does not occur which implies that either 2) or 3) does. Since all seven binary elements must be correctly detected in order that the correct symbol be generated, Curve A is the probability of at least one binary error in the block of seven elements.

The symbols produced by the binary detector also appear in the message generated by the maximum-likelihood detector. Because this device may be considered as one which correctly translates some of the binary detector's erasures into telegraph symbols and incorrectly translates the others, Curve B is a lower bound on the maximum-likelihood error rate and Curve A an upper bound. The probability of a maximum-likelihood error is $C(\rho)$ which always falls between $A(\rho)$ and $B(\rho)$. The maximum-likelihood detector generates a greater number of correct symbols than the binary detector does; it also generates more incorrect symbols.

Figure 6.2 demonstrates the characteristic shapes of the three curves and indicates various probabilities which may be determined from them. $A(\rho) - B(\rho)$ is the probability of an erasure at the binary detector, and $A(\rho) - C(\rho)$ is the probability of an erasure which is correctly translated by the maximum-likelihood detector. $C(\rho) - B(\rho)$ is the probability of an erasure which is incorrectly translated. The value of the maximum-

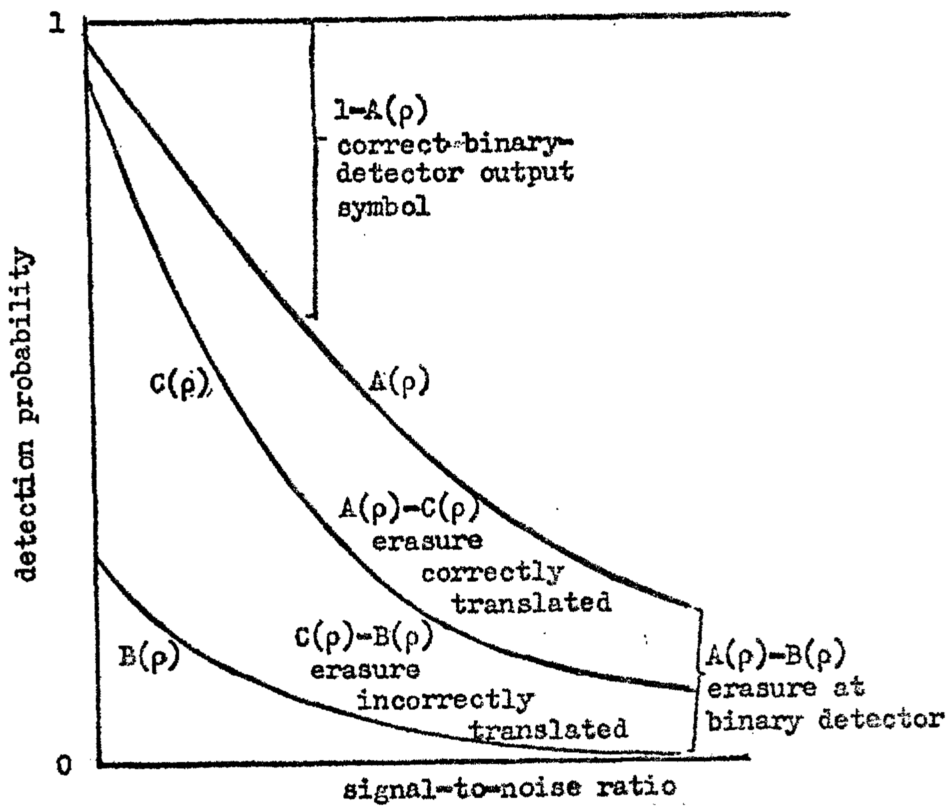


FIGURE 6.2 Binary detector and maximum-likelihood detector characteristics

likelihood detector relative to that of a binary detector depends on the proportion of the binary detector's errors which are correctly translated. This proportion,

$$D(\rho) = \frac{A(\rho) - C(\rho)}{A(\rho) - B(\rho)} \tag{6.5}$$

is a figure of merit of the maximum likelihood detector. It is plotted in Figures 6.14 and 6.15 for the various propagation conditions. The values of C(ρ) plotted in this section have been obtained by numerical integration of the error-rate formulas derived in Chapter 8.*

* Allowing for the minor differences in the initial assumptions, the curves of Figures 6.3 through 6.10 show good agreement with the simulation results tabulated in Appendix B.

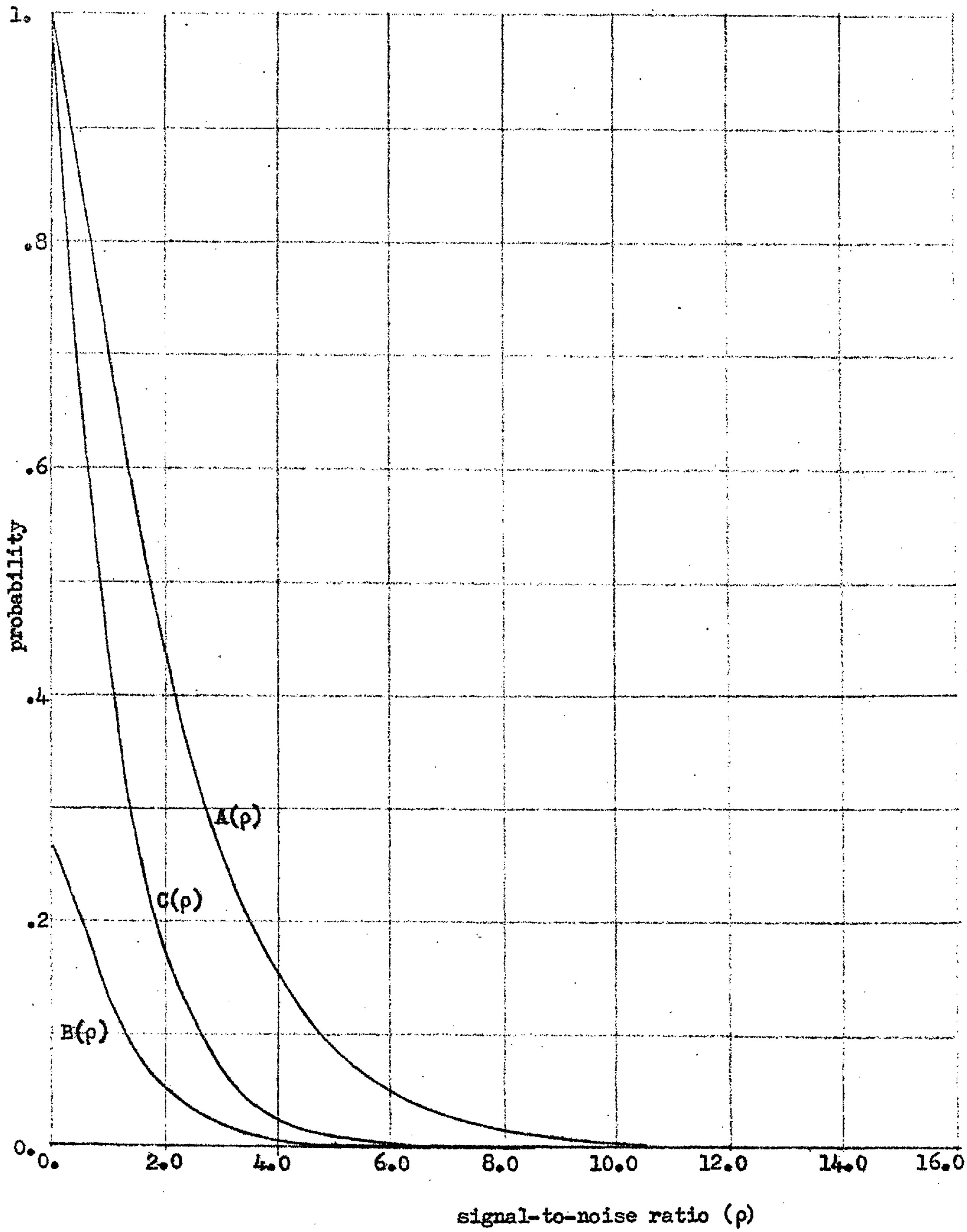


FIGURE 6.3 Case 1: Coherent receiver, no fading

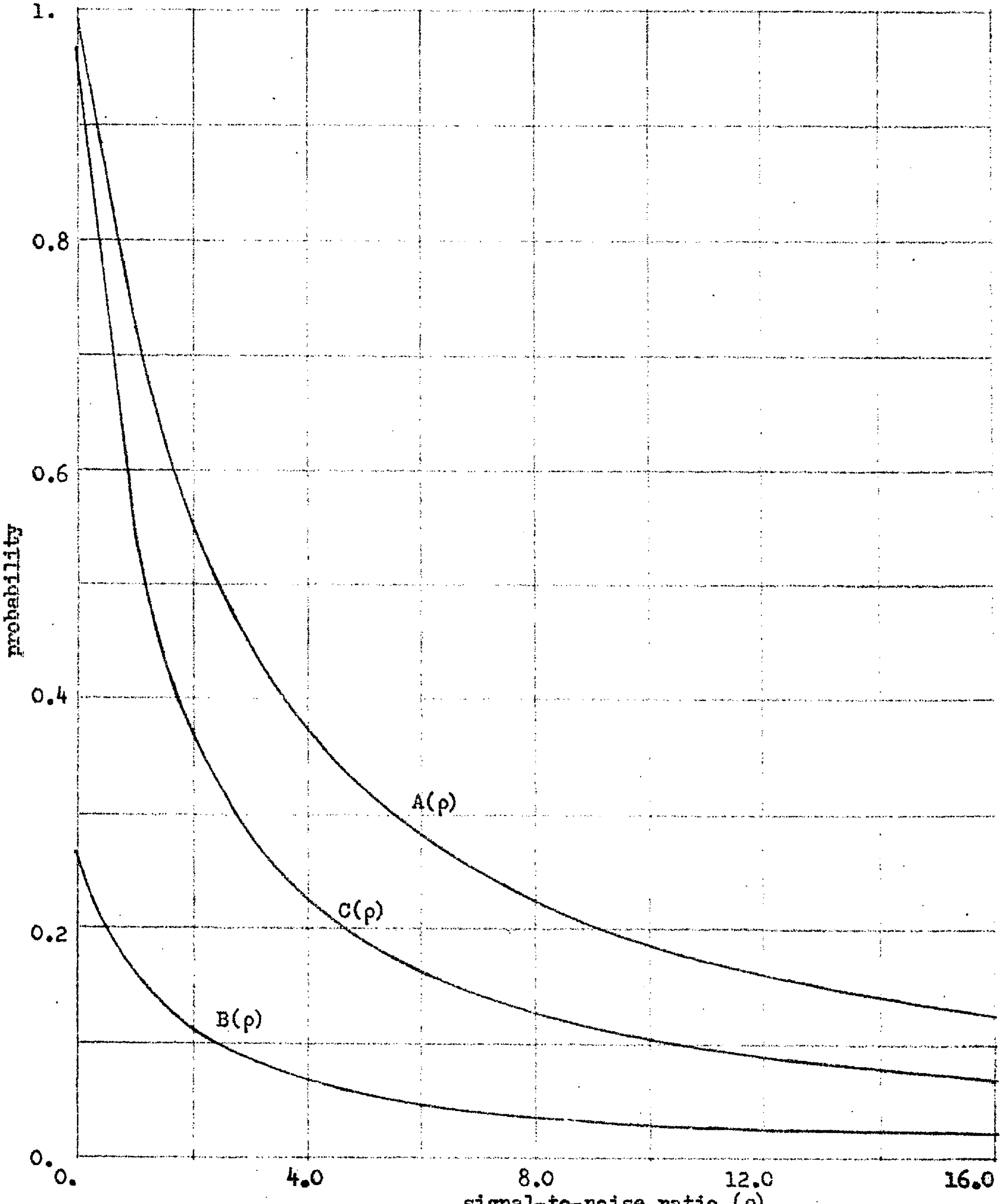


FIGURE 6.4 Case 2: Coherent receiver, slow flat fading

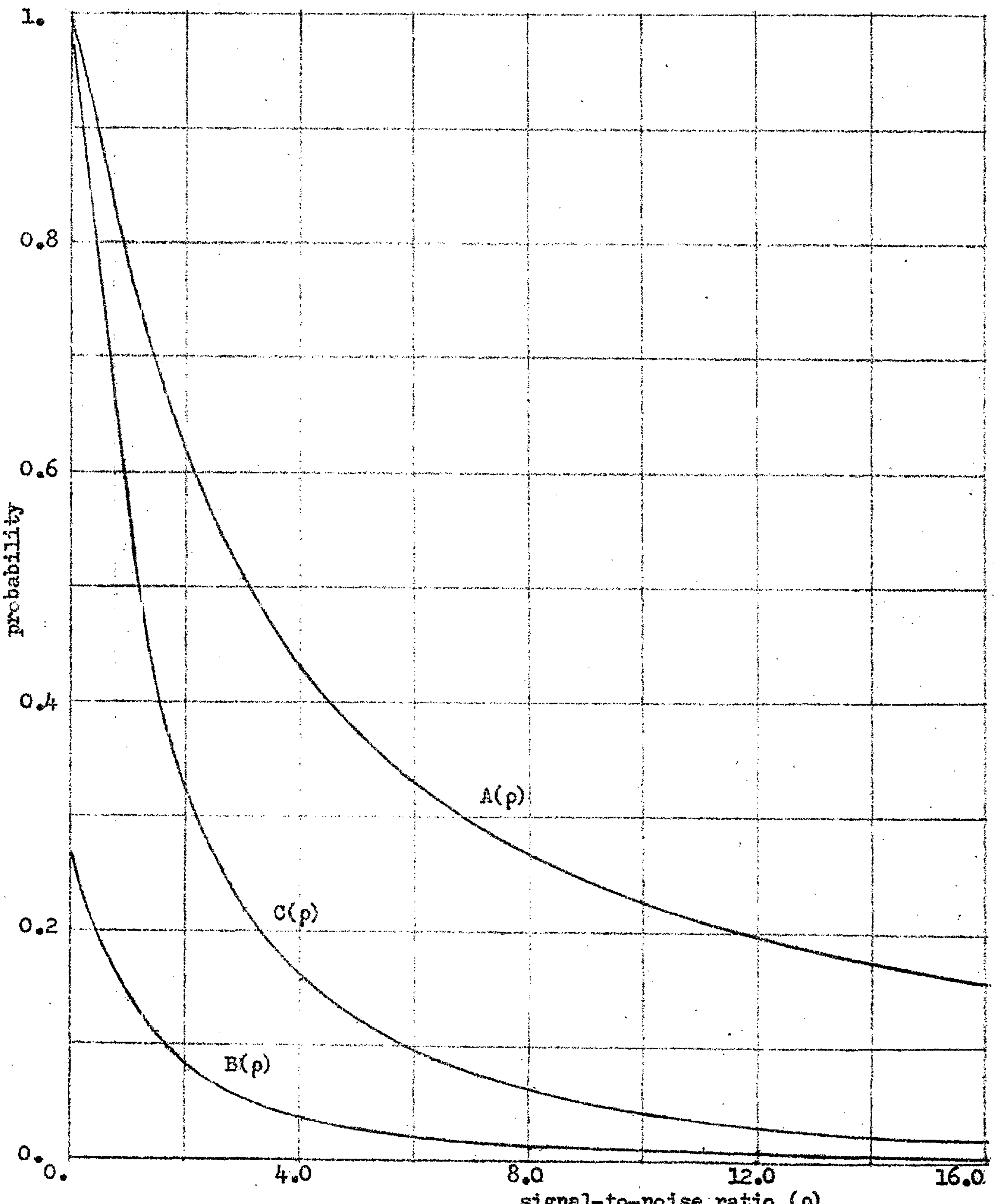


FIGURE 6.5 Case 3: Coherent receiver, slow selective fading

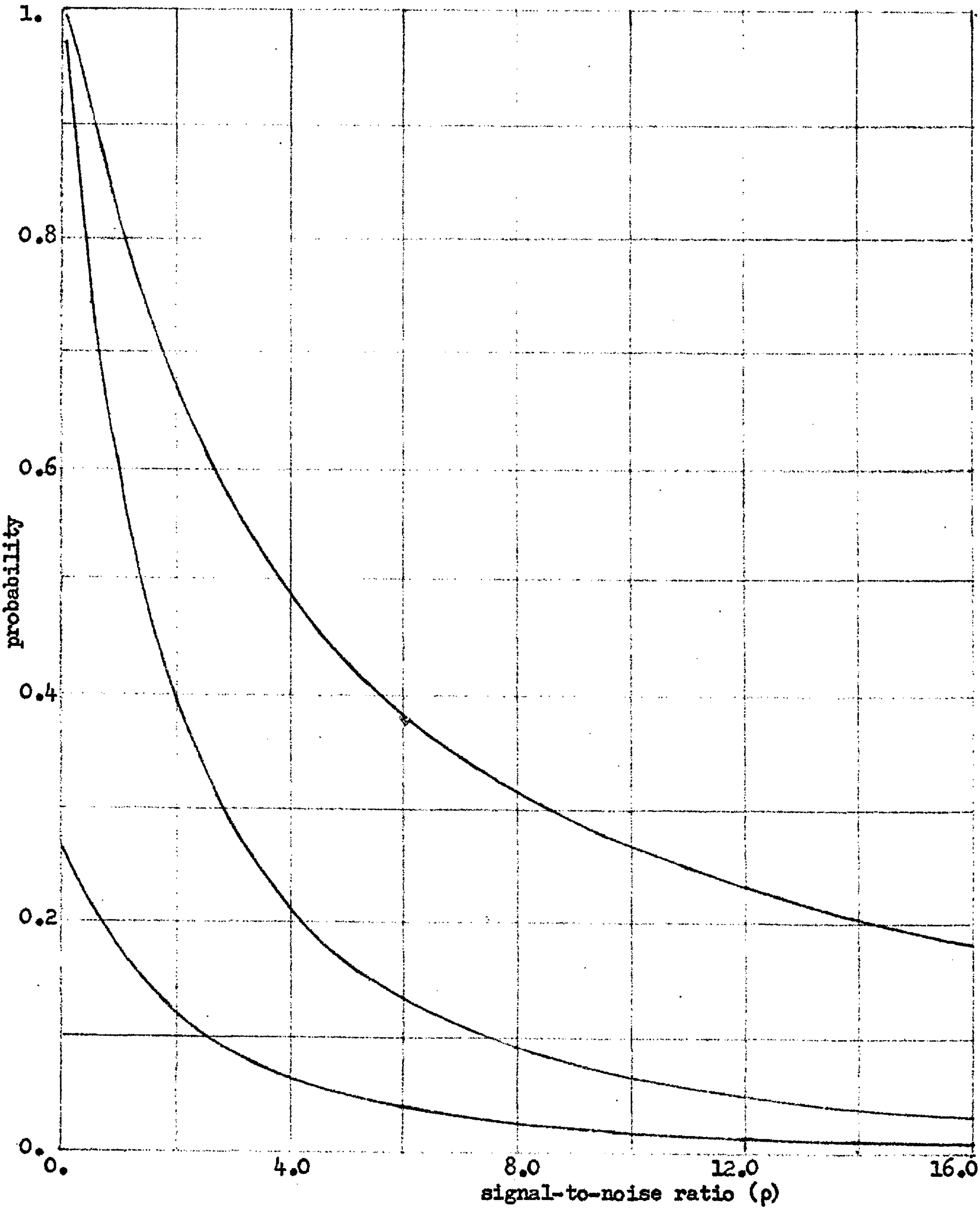


FIGURE 6.6 Case 4: Coherent receiver, fast fading

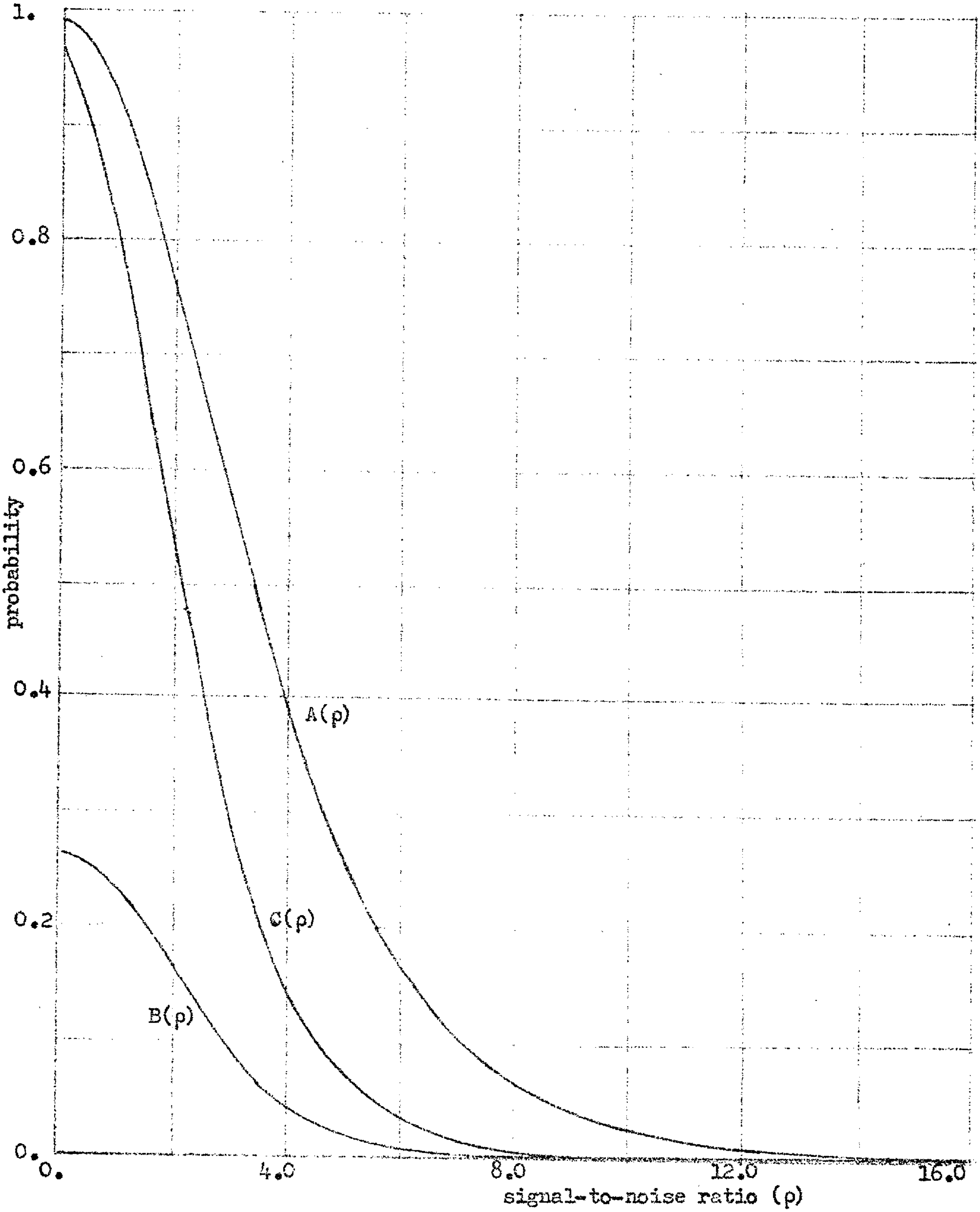


FIGURE 6.7 Case 5: Non-coherent receiver, no fading

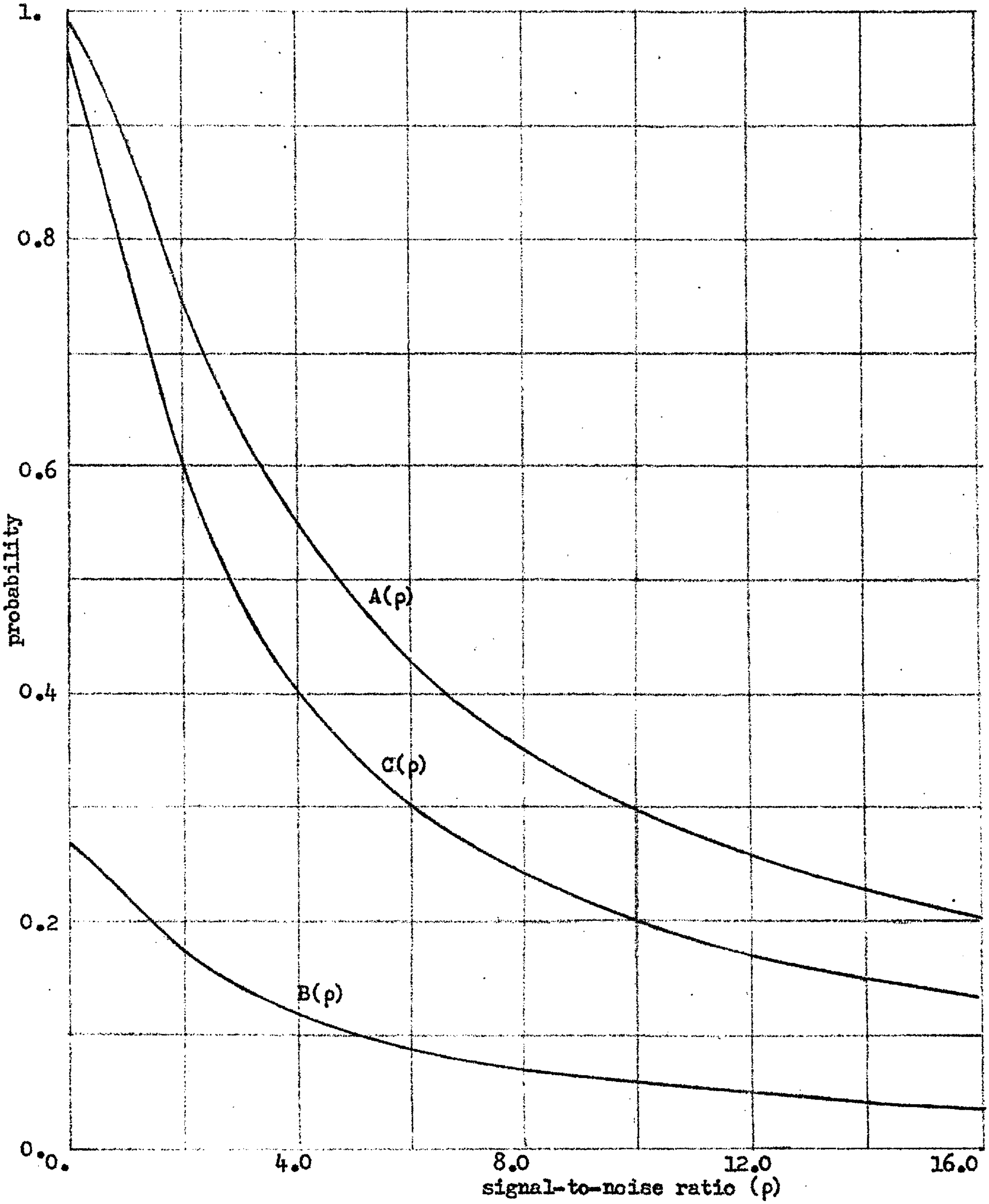


FIGURE 6.8 Case 6: Non-coherent receiver, slow flat fading

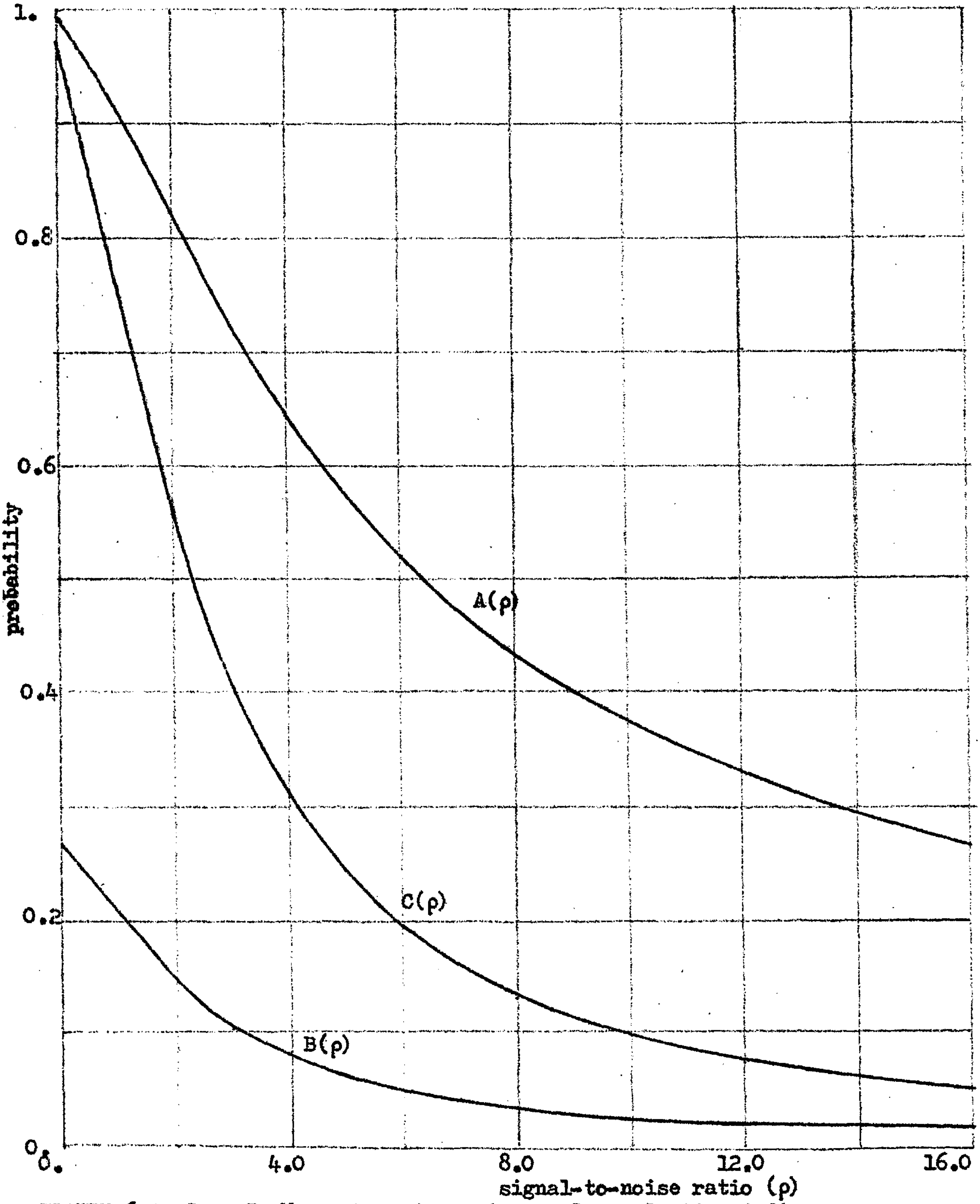


FIGURE 6.9 Case 7: Non-coherent receiver, slow selective fading

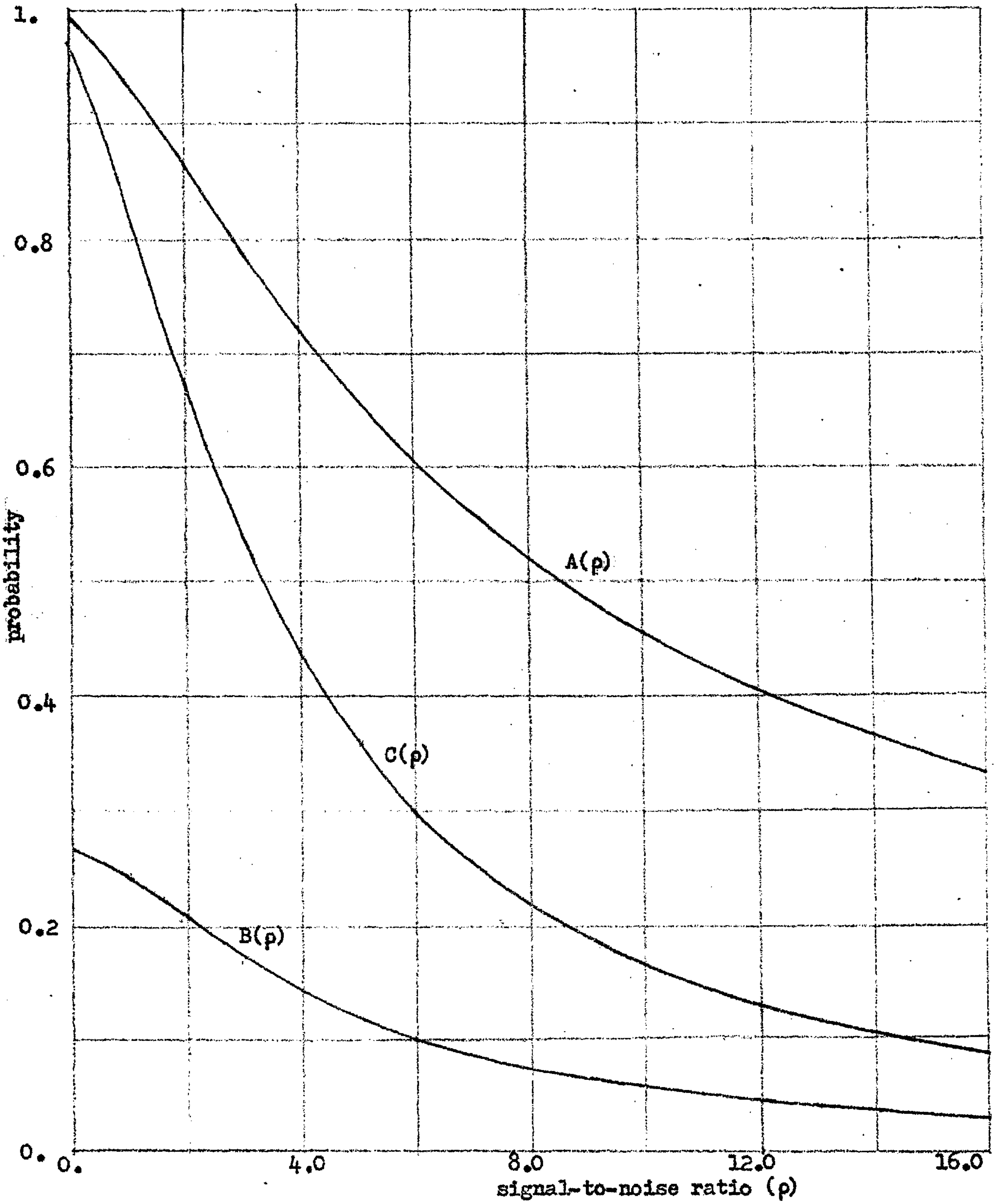


FIGURE 6.10 Case 8 Non-coherent receiver, fast fading

6.6 Evaluation of the characteristics

The curves presented in Figures 6.3 to 6.10 characterize radiotelegraph systems whose properties vary within three categories: the type of phase information available at the receiver, the type of amplitude fading in the channel and the type of detector at the receiving end of the radio link. There are 16 situations covered by the characteristics and in order to compare them it is convenient to consider the variations within each category and to assess their effects upon the error rates of detected messages. The discussion of the present section applies to one-way systems. Chapter 7 evaluates the performance of systems which include return-path signalling facilities.

6.6.1 Phase information

Figure 6.11 illustrates, for each fading condition, the maximum-likelihood error rates at a coherent and non-coherent receiver. They are plotted against the average signal-to-noise ratio expressed as 10 log ρ decibels. The two curves which apply to a given type of fading have similar shapes and are approximately parallel. In order to compensate for the absence of phase information, the signal-to-noise ratio associated with a given error rate is, over most of the range shown, about 3 db higher at a non-coherent receiver

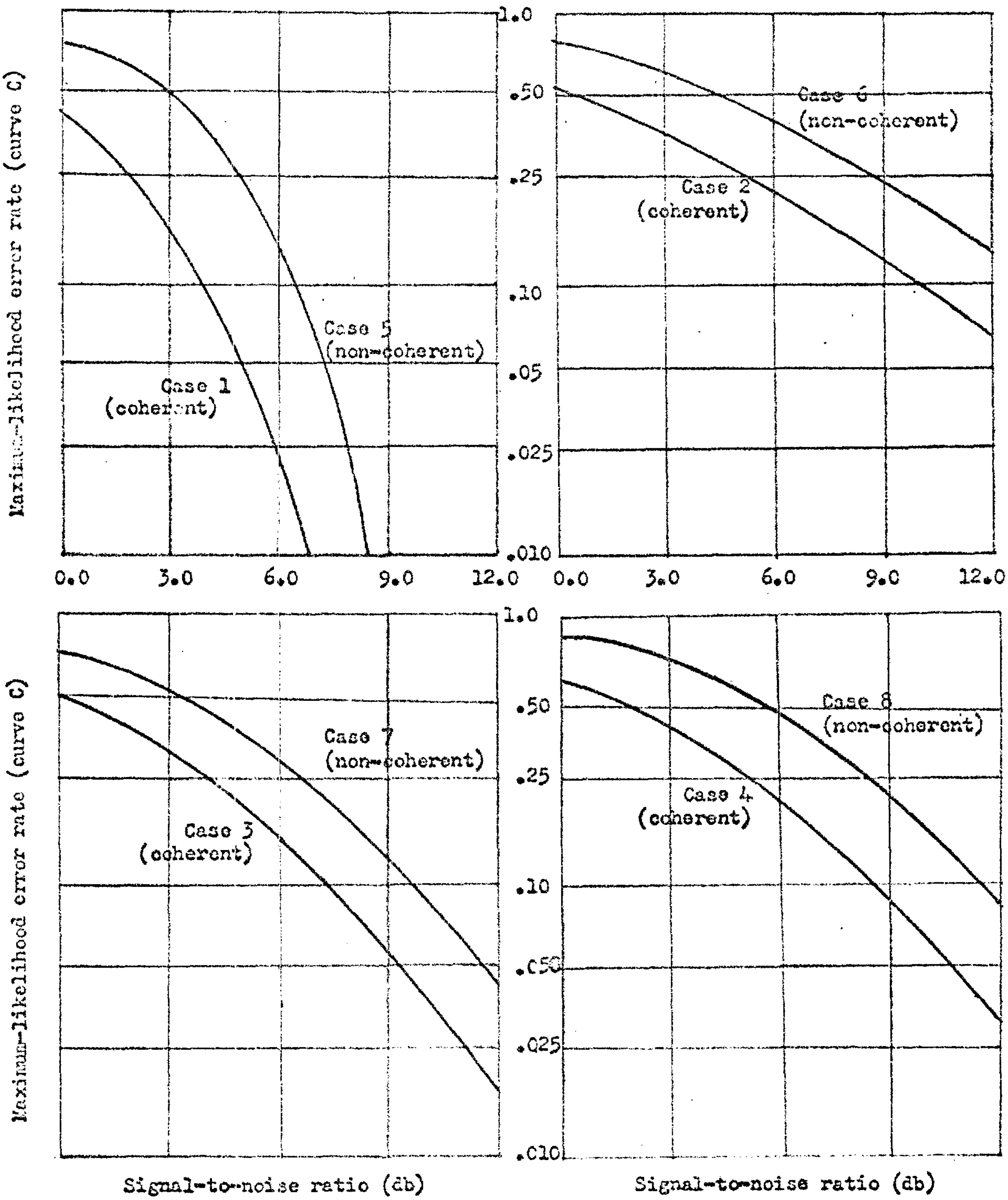


FIGURE 6.11 Comparison of coherent and non-coherent reception

than it is at a coherent receiver. The relationship between coherent and non-coherent reception is similar in the case of binary detection.

6.6.2 Amplitude fading

In certain applications, the type of fading to which a message is subjected may be designed by the engineer when he assigns signal elements to the time-frequency coordinates in a multiplex package. An assessment of the effect of the type of fading on system error rates thus has direct practical implications. In the steady-signal conditions all of the transmitted signal elements have the same signal-to-noise ratio. When there is fading, on the other hand, the signal-to-noise ratio of each element is a random variable whose average value, ρ, is the independent variable of the characteristic curves. Except when ρ is very low, the fluctuation of the signal-to-noise ratio about this average results in significantly higher error rates than those achieved when every element has a signal-to-noise ratio of ρ. Unfortunately, the steady signal condition, which generally offers better performance than any of the fading conditions, is the one type of amplitude variation which cannot be achieved by multiplex design when the attenuation of the propagation medium fluctuates randomly.

Because amplitude fluctuations almost always occur, multiplex design decisions must be based on the relative effects of the three types of fading. Each type has two labels. One label alludes to amplitude correlations in time (slow or fast) and the other to correlations in frequency (flat or selective).
The nomenclature is descriptive of the physical effects of the channel upon messages in a conventional multiplex arrangement in which each message has two characteristic frequencies for all seven elements and each element is assigned to a characteristic time interval. The identity of the element is denoted by the transmission of one of the two tones at the appropriate time. In a more general multiplex scheme, on the other hand, the two states of each signal element may be assigned to any pair of time-frequency regions. The identity of the element is signalled by the transmission of energy in one of the regions and the absence of energy from the other. The types of fading which may affect a message in a general multiplex configuration are distinguished by the correlation of the two amplitudes associated with a given element and by the correlations between the amplitudes of the different elements in a symbol. The fading conditions may therefore be specified more

generally in terms of "within-element" selectivity and "between-element" selectivity. The fading called frequency selective with respect to conventional multiplexing is, in the more general terminology, selective within elements and flat fading is correlated within elements. Fast fading is selective between elements and slow fading correlated between elements.

Figure 6.11 demonstrates clearly that for a given signal-to-noise ratio, slow flat fading has the most harmful effect upon the maximum-likelihood error rate. It follows, therefore, that the multiplex scheme of a system which includes a maximum-likelihood detector should be designed with the aim of achieving at least one of the two types of selectivity. If a binary detector is employed, the relationship between completely non-selective fading and the other two types is more complicated. When the fading is selective, the elements with low signal-to noise ratios-- that is, those most susceptible to error -- are distributed over many symbols instead of being concentrated in a few. Because a binary detector must correctly interpret all seven elements in order to generate the correct symbol, this dispersal of unreliable elements implies that the probability of correct detection is lower when the fading is selective than it is for completely non-selective fading. The binary error rate is independent of the type of fading so that when the fading is selective, the more frequent occurrence of symbols with at least one error implies a less frequent occurrence of multiple errors. Because multiple binary errors are prerequisites of incorrect output symbols, the probability of an incorrect output is lower for selective fading than it is for non-selective fading.

Completely non-selective fading, among the three types considered, results in the greatest number of correct symbols and the greatest number of incorrect symbols at the output of a binary detector. It therefore causes the fewest erasures. Thus Curve $A(\rho)$ in Figure 6.4 is lower than the corresponding function in either 6.5 and 6.6 and $B(\rho)$ is higher. Similar relationships prevail in Figures 6.8, 6.9 and 6.10 which pertain to fading signals at a non-coherent receiver. These binary-detector characteristics are, therefore, equivocal with respect to the desirability of selective fading. When a binary detector is employed, the design decision to aim for selective or correlated fading must be based upon the fidelity criterion of the message reader -- in particular, the relative importance he attaches to the objectives of obtaining correct symbols and avoiding incorrect symbols.

With respect to binary detection, within-element selectivity is preferable to the between-element type. When the fading is selective only within elements, the elements with low signal-to-noise ratios are distributed over fewer symbols and the probability of a correct symbol at the output of a binary detector is higher than it is in the condition of between-element selectivity. The probability of an incorrect symbol is lower because the "ones" and "zeros" of a symbol have different signal-to-noise ratios and there is, in each symbol, a bias toward one type of binary error. Both types of error must occur if an incorrect symbol is to be generated and this event is least likely when the fading is selective within elements. Thus both A(ρ) and B(ρ) in Figure 6.5 are lower than the corresponding curves in 6.6 and each of these binary characteristics in 6.9 is lower than its counterpart in 6.10.

For systems which include a maximum-likelihood detector, the effects of the two types of selective fading may be compared by means of the error-rate curves presented in Figure 6.12. Within-element selectivity, which offers better results in the case of binary detection, is also the more desirable type when a maximum-likelihood detector is employed.

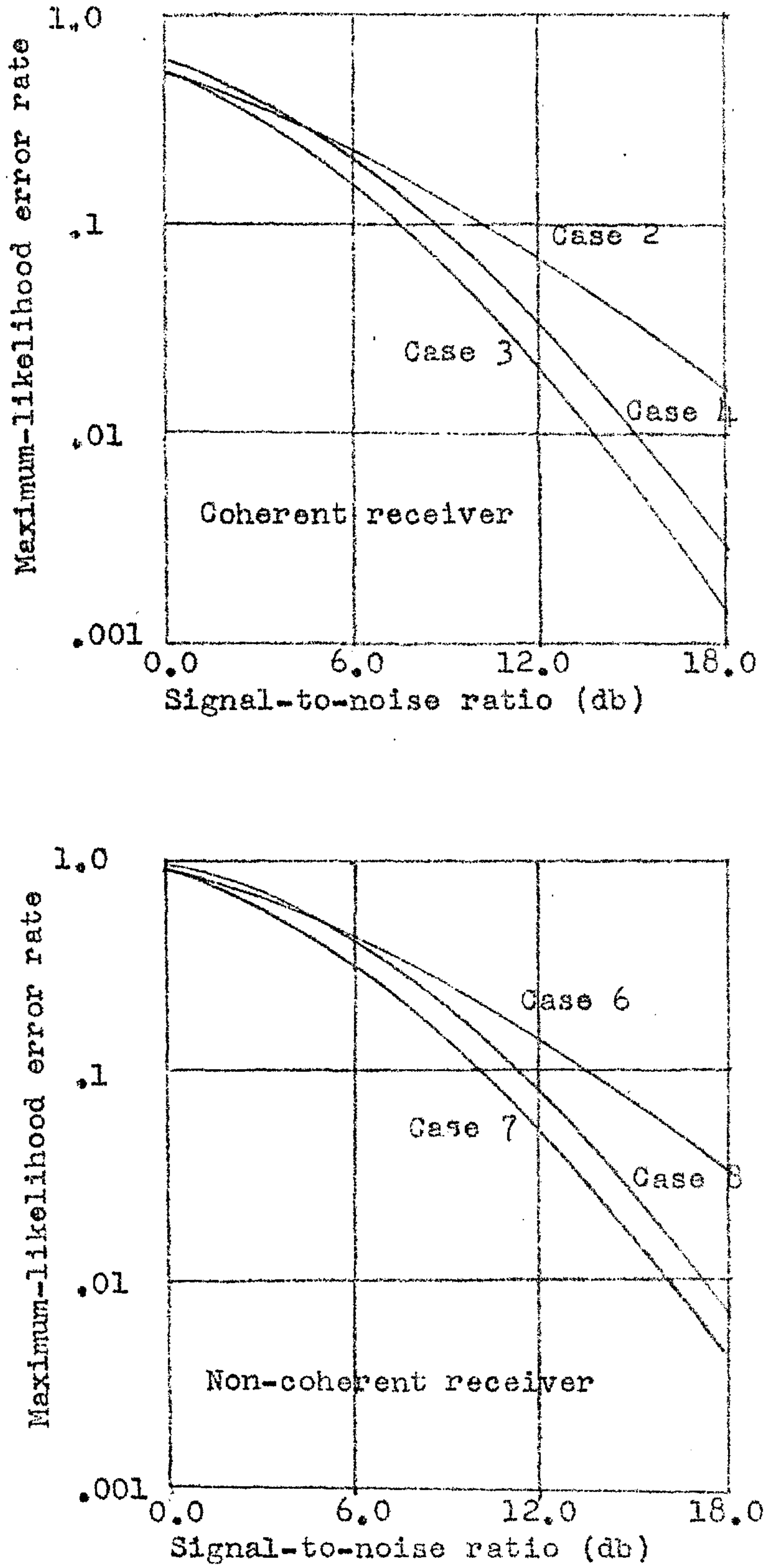


FIGURE 6.12 Comparion of fading conditions

Figure 6.12 shows that the advantage of within-element selectivity relative to between-element selectivity may be expressed as an economy of 1 to 2 db in the signal power necessary to achieve a specified error rate. This is a low number compared with the advantage of between-element selectivity relative to completely non-selective fading which is an increasing function of signal-to-noise ratio. Since 1 to 2 db is not a significant power difference in practical telegraph systems it may be said that both types of selective fading result in approximately the same maximum-likelihood performance and that this performance is significantly better than it would be if there were no selectivity at all.*

In terms of multiplex design, this evaluation implies that code elements should be assigned to time-frequency regions with the objective of insuring that each message is affected by at least one of the two types of selective fading. This objective is best met if the 14 regions associated with each message are distributed as widely as possible over the rectangle which represents the multiplex frame. Such a distribution guarantees that each symbol will be received with a diversified set of element amplitudes if the

---

* These observations pertain to the detectors considered in this thesis which have no access to measurements of the two amplitudes, A_0 and A_1, of signals characterized by within-element selectivity. If these measures were available, the best results would be obtained under the condition of fading which exhibits both types of selectivity.

fading is in any way selective over the time slots or frequencies in the multiplex frame. With this type of assignment, therefore, some selectivity advantage is maintained over a wide variety of propagation conditions. Since within-element selecticity is particularly desirable, the two regions associated with each element should be well-spaced in both time and frequency.

Figure 6.13 demonstrates the assignments of the code elements of two messages to time-frequency regions in a 24-message multiplex package which has been designed on the basis of these criteria. The 14 x 24 = 336 time-frequency regions are formed with 12 characteristic frequencies and 28 time intervals. With respect to every message in the package, the two regions assigned to a single element are separated by at least 5 frequency intervals and 12 time intervals. The minimum spacing between two successive elements of a symbol is 3 frequency units and 6 time units. This arrangement contrasts with a conventional multiplex scheme which contains six frequency-multiplexed two-tone signals, each conveying the information of four time-multiplexed telegraph messages. In this case there is no time diversity within elements and no frequency diversity between elements. In conventional systems, therefore, only a restricted class of fading patterns results in selectivity with respect to all 24

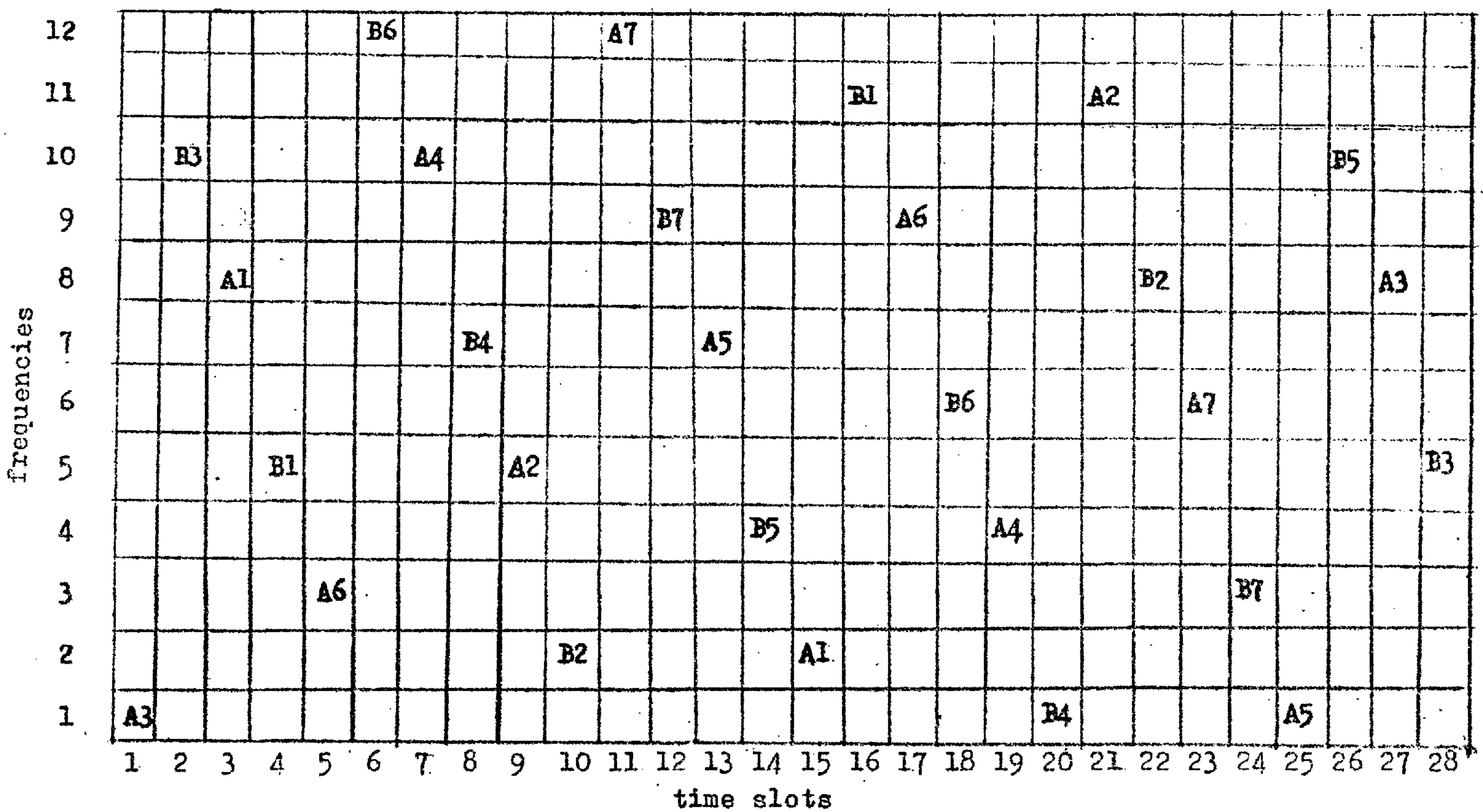


FIGURE 6.13 Multiplex assignment for achieving fading selectivity over a wide range of physical conditions. The two rectangles labelled Ak are assigned to the kth element of symbols of message A.

messages.

For binary detection, non-selectivity may be as acceptable as, or more desirable than, selective fading. In systems with a maximum-likelihood detector selective fading is clearly preferable.

6.6.3. Comparison of detectors

The last two sentences of the preceding section suggest that the value of maximum-likelihood detection relative to binary detection is lowest when the fading is non-selective. This is confirmed by Figures 6.14 and 6.15 which contain curves of the function $D(\rho)$ for the various types of amplitude fading at coherent and non-coherent receivers respectively. For either type of phase information, the proportion of binary-detector erasures correctly interpreted by a maximum-likelihood detector is lowest for a given signal-to-noise ratio when the fading is completely non-selective (Cases 2 and 6). Over the range of signal-to-noise ratios plotted, $D(\rho)$ is less than .55 at a coherent receiver and .45 at a non-coherent receiver under the condition of non-selective fading. With $D(\rho)$ at these levels, the messages produced by a maximum-likelihood detector are likely to be no more than marginally better than those available from a binary

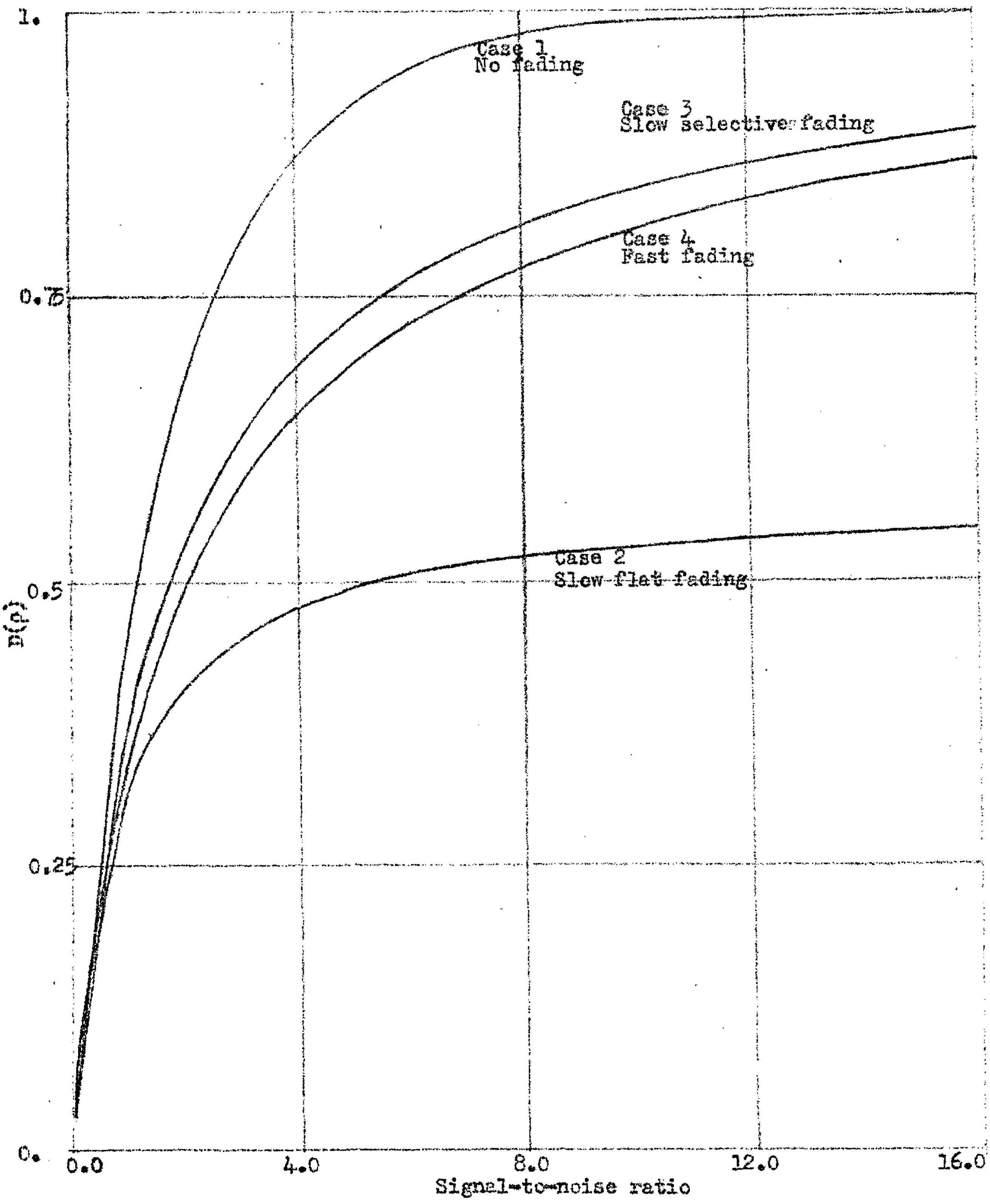


FIGURE 6.14 Maximum-likelihood figure of merit--coherent reception

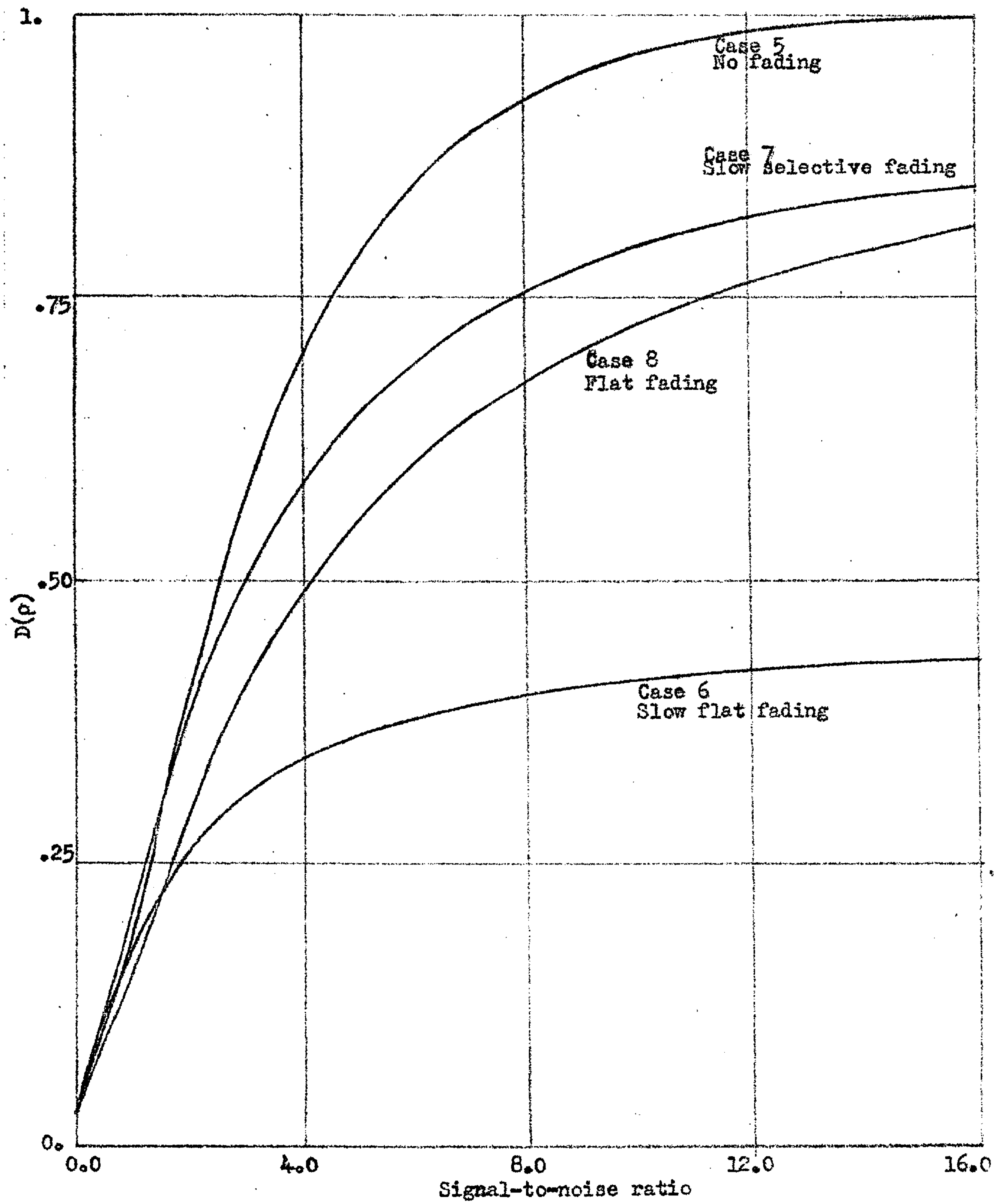


FIGURE 6.15 Maximum-likelihood figure of merit--non-coherent reception

detector. If a quality-assessment mechanism of the type described in Section 6.4 could be introduced, maximum-likelihood detection would be preferable even though approximately half of the relatively uncertain symbols would be incorrect.

In the selective fading case, $D(\rho)$ rises more steeply with average signal-to-noise ratio and it exceeds .75 for ρ greater than 7.0 at a coherent receiver and 11.0 at a non-coherent receiver. When $D(\rho) = .75$, three erasures are correctly translated for each incorrect translation. Assessment of simulated transmissions has confirmed that this level of success results in maximum-likelihood messages which are considerably more useful than those produced by a binary detector, even when there is no quality-assessment mechanism built into the maximum-likelihood detection process.

In commercial telegraph practice there are many one-way and broadcast systems which represent information by means of the five-unit start/stop code over the entire route from teleprinter to teleprinter. Because the five-unit code is non-redundant, the output of a binary detector contains no erasures and the error rates of the received messages may be compared directly with the error rates of messages produced by a maximum-likelihood detector of symbols represented by the seven-unit code. The inefficiency of the five-unit code is mentioned in Section 3.3 and it is now demonstrated in Figures 6.16 and 6.17 which show for each propagation condition the error rates associated with binary detection of the five-unit code and maximum-likelihood detection of the seven-unit code. In computing the five-unit-code error rate it has been assumed that the stop element has the same length as an information element and that this is identical to the element length in the seven-unit-code system. Both codes, therefore, convey telegraph information at the same speed.

Although its advantage is least pronounced when the fading is non-selective, the maximum-likelihood detector offers a lower error rate than the binary detector for all of the propagation conditions. Figures 6.16 and 6.17 thus demonstrate that the performance of a telegraph system can, under any of the conditions, be improved by a code conversion from five-unit to seven-unit and adoption of a maximum-likelihood detector.

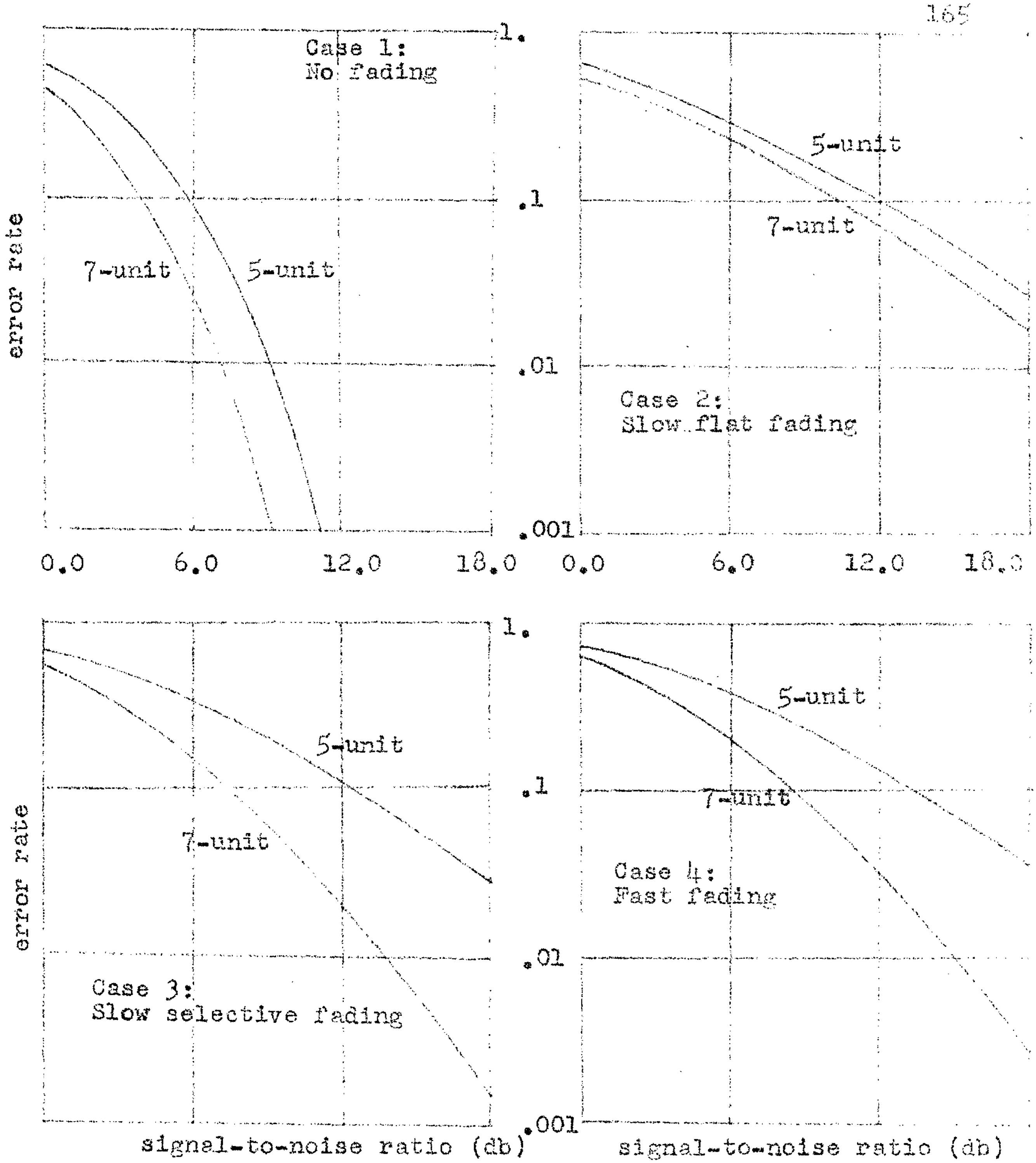


FIGURE 6.16 Comparison of error rates: binary detection of five-unit code, maximum-likelihood detection of seven-unit code (coherent receiver)

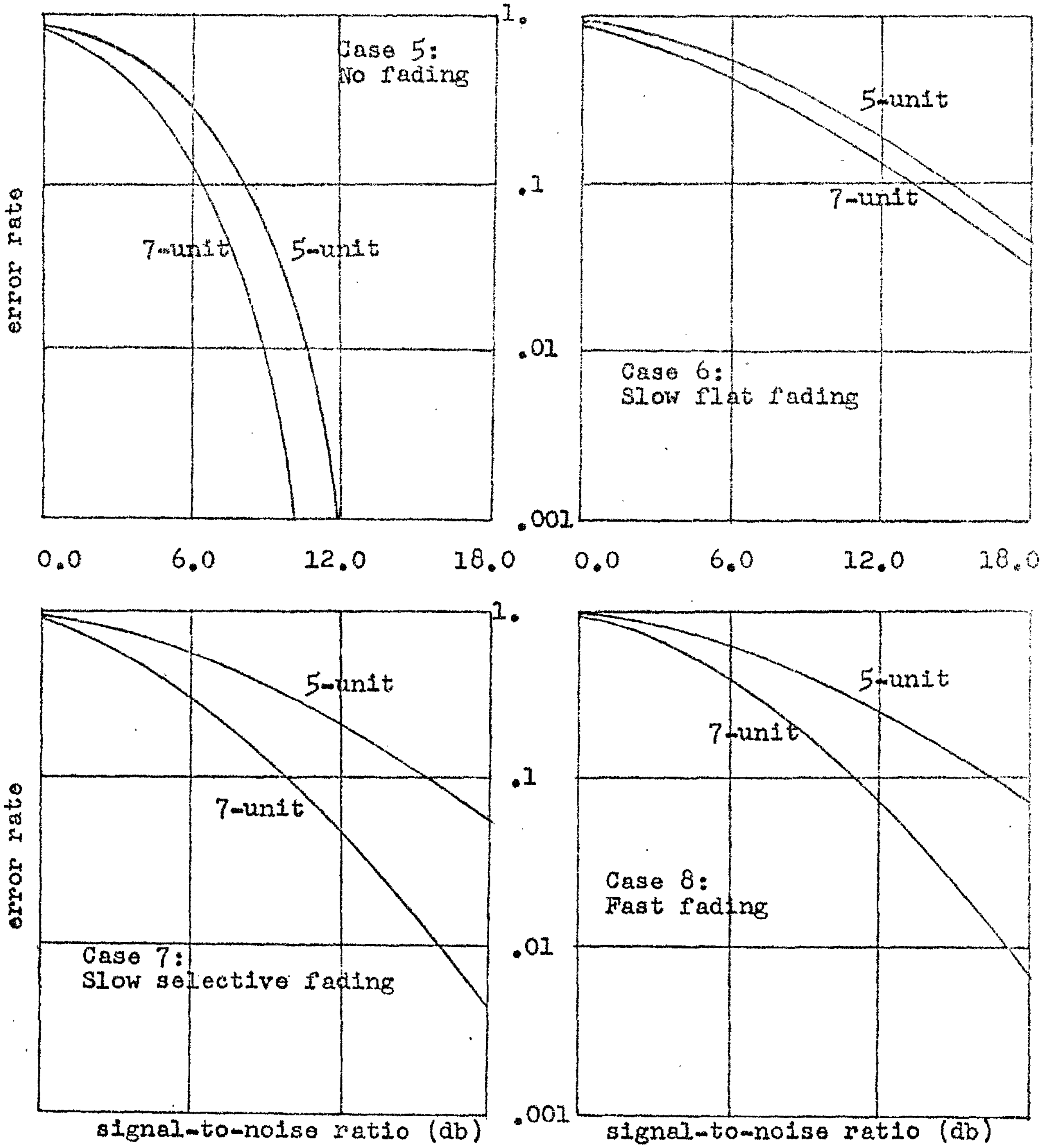


FIGURE 6.17 Comparison of error rates: binary detection of five-unit code, maximum-likelihood detection of seven-unit code (non-coherent receiver)

CHAPTER 7: FEEDBACK SYSTEMS

In this chapter the performance characteristics presented in Chapter 6 are applied to a comparison of two types of feedback communication system, using auto-sending. A feedback system contains facilities for transmitting signals in both directions between a pair of terminal stations. In addition to the message information passed through the system, the two stations exchange control information which serves to maintain the quality of communication by adopting system operation to changing channel conditions. Figure 7.1 illustrates the flow of information in a one-way system with feedback. Messages are transmitted only in the A-B direction and control information relevant to the signals received at B is fedback to the transmitter at terminal A. Feedback facilities may also be built into two-way systems such as the one shown in Figure 7.2. In this case, each of the two one-way channels conveys both message information and control information.

The technique of ARQ signalling (15,16,57), described in Section 4.4, is an example of a feedback mechanism used in a two-way telegraph system. In the ARQ technique, as well as in many other feedback methods, the control mechanism involves an adjustment

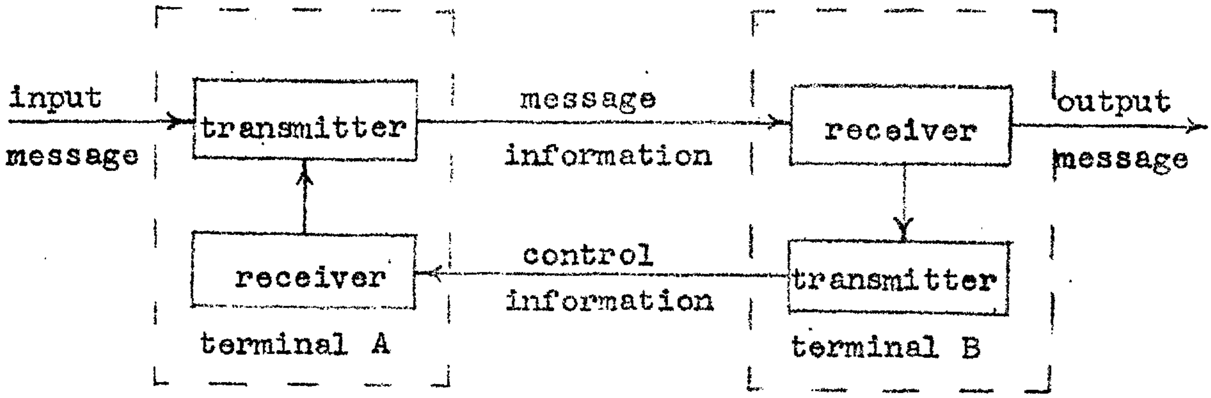


FIGURE 7.1 One-way system with feedback.

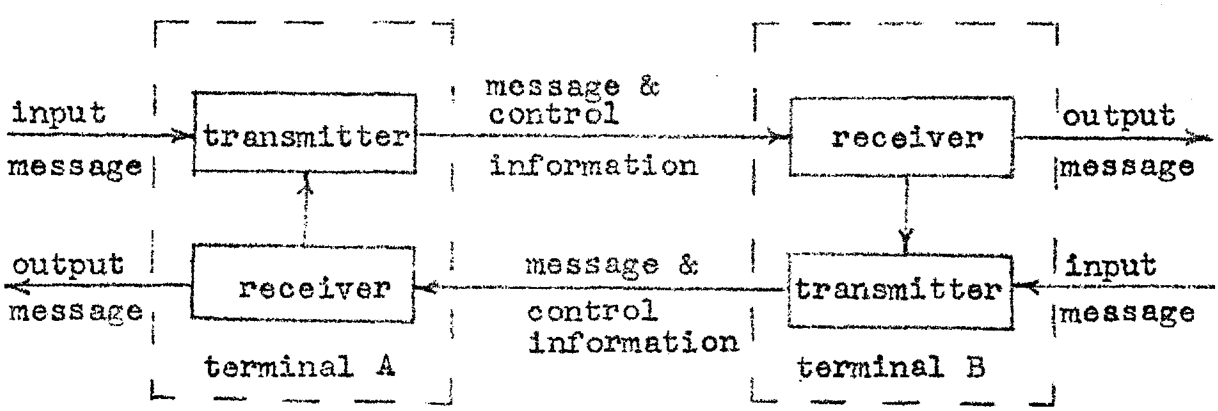


FIGURE 7.2 Two-way system with feedback.

of the effective transmission speed of message symbols*. The adjustment is achieved by means of repeated transmissions of portions of the signal sequence, a procedure which conforms to the model of a sequential decision problem described in Section 2.1 (62). In the terms of this model, the detector may be viewed as making two types of decision with respect to each transmitted symbol. One type is a terminal decision which results in the output of a teleprinter symbol and the other is a decision to request the transmission of additional data relevant to the identity of the transmitted symbol. This request results in the retransmission of the symbol's signal elements. The decision to request a retransmission corresponds, in the ARQ system, to the detection of an erasure.

One weakness of the ARQ technique is the requirement that retransmissions be organized in repetition cycles consisting of several message symbols. This requirement derives from the long path delay of most hf telegraph systems and it results in an appreciable decrease in transmission speed relative to the case of a one-symbol repetition cycle (see Figure 7.6). If

---

* In principle it is possible to adjust other system parameters such as bandwidth and power but in practice the constraints on these parameters severely limit the ranges over which they may be varied.

each erasure were to cause only one retransmission, the speed of the ARQ system would increase but a considerable symbol-storage requirement would be imposed on the receiver (52). Another weakness of the ARQ method is its incapacity to utilize the information in the originally-received signals associated with symbols which are retransmitted. The set of received waveforms corresponding to the several transmissions of a single symbol may be viewed as branch signals in a time-diversity system. The ARQ system combines the branch signals in a manner somewhat similar to the optimal selection method defined in Section 4.2. The introduction of a better combination method would result in an improvement in performance but it would require the installation at each receiver of complicated information-storage facilities.

In addition to techniques which may be described as sequential decision procedures, there are other means of using a feedback facility to vary the transmission speed of a telegraph system in order to control the error rate (21). A method proposed in the present chapter involves adjustment of the time duration of transmitted signal elements in response to changing channel conditions. A maximum-likelihood detector is employed so that no erasures are detected and there are no requests for retransmission of message symbols. The independent variable of the maximum-likelihood performance characteristics in Chapter 6 is the element signal-to-noise ratio which is directly proportional to the element time duration. If the signal component of the received waveform has non-zero amplitude, the element duration may be a adjusted to a value which causes the system to operate at any desired point on a $C(\rho)$ characteristic.

In order to perform this adjustment, the transmitter must obtain information related to a measurement of the channel parameters (43). This type of information contrasts with the detector-decision information which controls the occurrence of retransmissions in the ARQ system. The latter type of information is known as postdecision feedback and the channel-parameter information required by the variable-duration system is known as predecision feedback. A practical example of a predecision feedback scheme is the "Janet" system of vhf tropospheric propagation. In this system, the variable-speed mechanism is binary. When the measurement of channel quality exceeds a threshold, message information is transmitted at a fixed rate. When the measurement falls below the threshold, message transmission ceases and the speed is zero. The predecision feedback system proposed in this chapter calls for a much finer adjustment of transmission speed.

In Section 7.2, the performance of the proposed system is compared with that of an ARQ system operating in the same noise and fading environment. For the purpose of this comparison, it is assumed that the ideal condition of no noise in the feedback channel exists. This assumption permits the specification of a transmission-speed characteristic for the predecision feedback system which is identical to that of the ARQ system. With both systems operating at the same speed, their relative merits may be assessed on the basis of their error rates alone. The error rates of the two systems and their common speed are illustrated as functions of signal-to-noise ratio in Section 7.2. Under each of the propagation conditions, the two error-rate curves cross at a value of ρ between 7 and 17 db. At lower signal-to-noise ratios, the variable-duration system offers better performance while at higher values of ρ the ARQ system is superior.

The variable-duration system thus offers performance improvements over ARQ operation under poor channel conditions. It also admits more flexibility in the specification of speed and error-rate characteristics. For these reasons, it appears that the development of a practical variable-duration system would be desirable. Although a detailed practical design is beyond the scope of this thesis, preliminary design considerations are discussed in Section 7.3.

7.1 Relative speed and effective signal-to-noise ratio

At each receiver of an ARQ system, the number of symbols generated per unit time is a function of the channel signal-to-noise ratio. The detection of erasures at either end of the system causes extra symbols to be transmitted in both directions. The relative speed (sometimes called the "throughput" (7,8)) of a one-way transmission may be defined as the average number of output symbols per transmitted symbol. If the relative speed in one direction of an ARQ system is $s(\rho)$, the average signal energy per output symbol is greater by a factor of $1/s(\rho)$ than the energy per transmitted symbol.

If the performance of an ARQ system is to be compared with that of a system which has no retransmission mechanism, the speed of the latter system should be reduced until symbols are generated at the average ARQ output rate. This reduction may be accomplished by an increase in the time duration of each signal element by a factor of $1/s(\rho)$. The signal-to-noise ratio is thereby increased from ρ to

$$\tilde{\rho} = \frac{\rho}{s(\rho)} \tag{7.1}$$

which may be defined as the "effective signal-to-noise ratio" of the ARQ system.

The relative speed may be calculated from the binary-detector performance characteristics of Chapter 6 according to the formula given by Van Duuren (57):

$$s(\rho) = \frac{a_f(a_f a_r + a_f B_r + B_f)}{\nu - (\nu - a_f)(a_f a_r + a_f B_r + B_f)} \tag{7.2}$$

The subscripts f and r refer respectively to the "forward" one-way channel, whose relative speed is $s(\rho)$, and the "return" channel. ν is the number of symbols per repetition cycle and a and B are binary-detector performance characteristics. $a(\rho)$ is the probability of correct detection -- this is $1-A(\rho)$ where $A(\rho)$ is the curve plotted in Figures 6.3 to 6.16 -- and $B(\rho)$ is the probability of incorrect detection given in the same curves.

Figures 7.3 and 7.4 are the result of using the data of Figure 6.7 to calculate relative speeds and effective signal-to-noise ratios at a non-coherent receiver of non-fading signals. The ARQ system has a four-symbol repetition cycle and the different curves pertain to three situations which span a

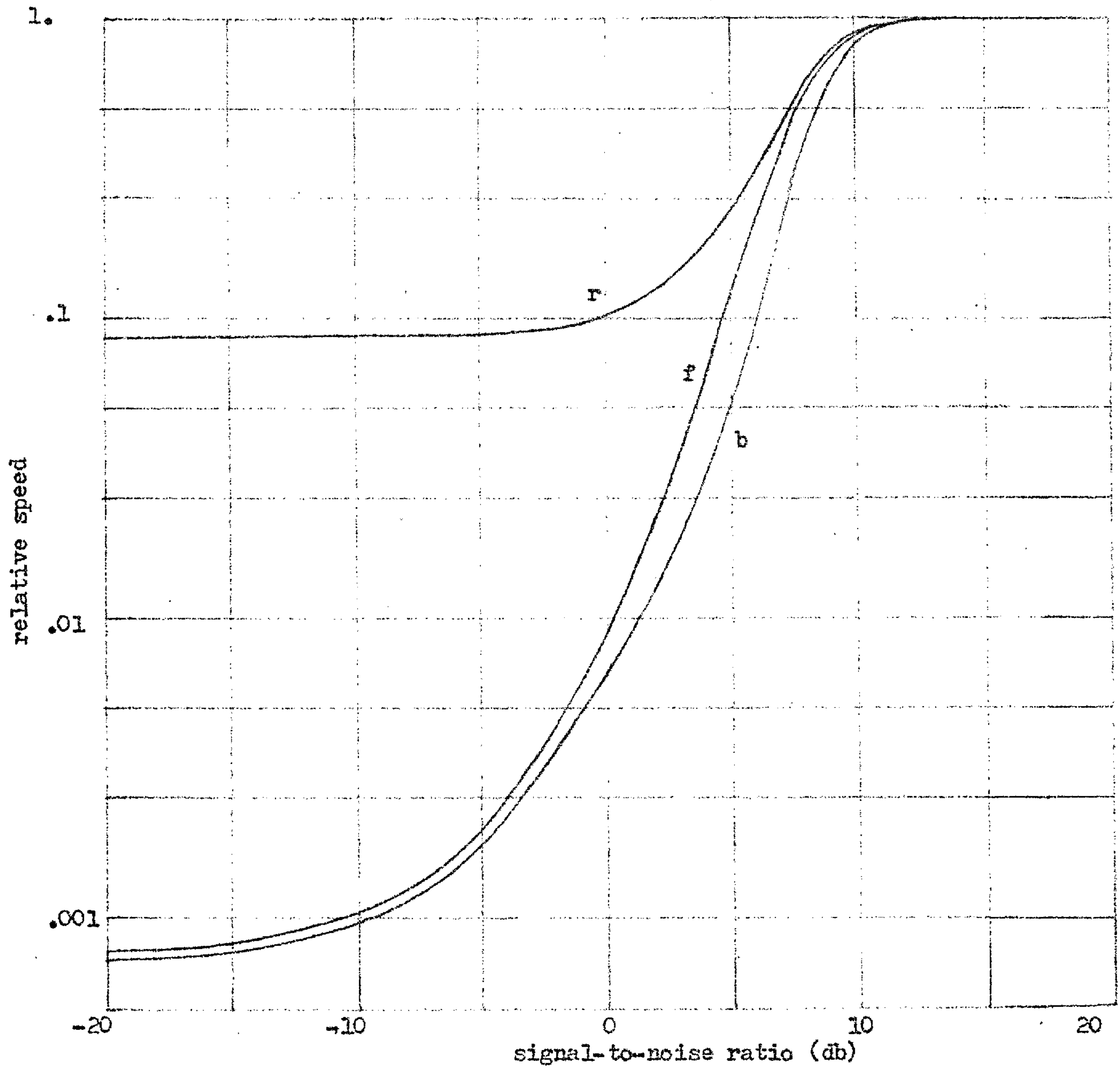


FIGURE 7.3 ARQ-system relative-speed characteristics (4-symbol repetition cycle, Case 5 propagation)

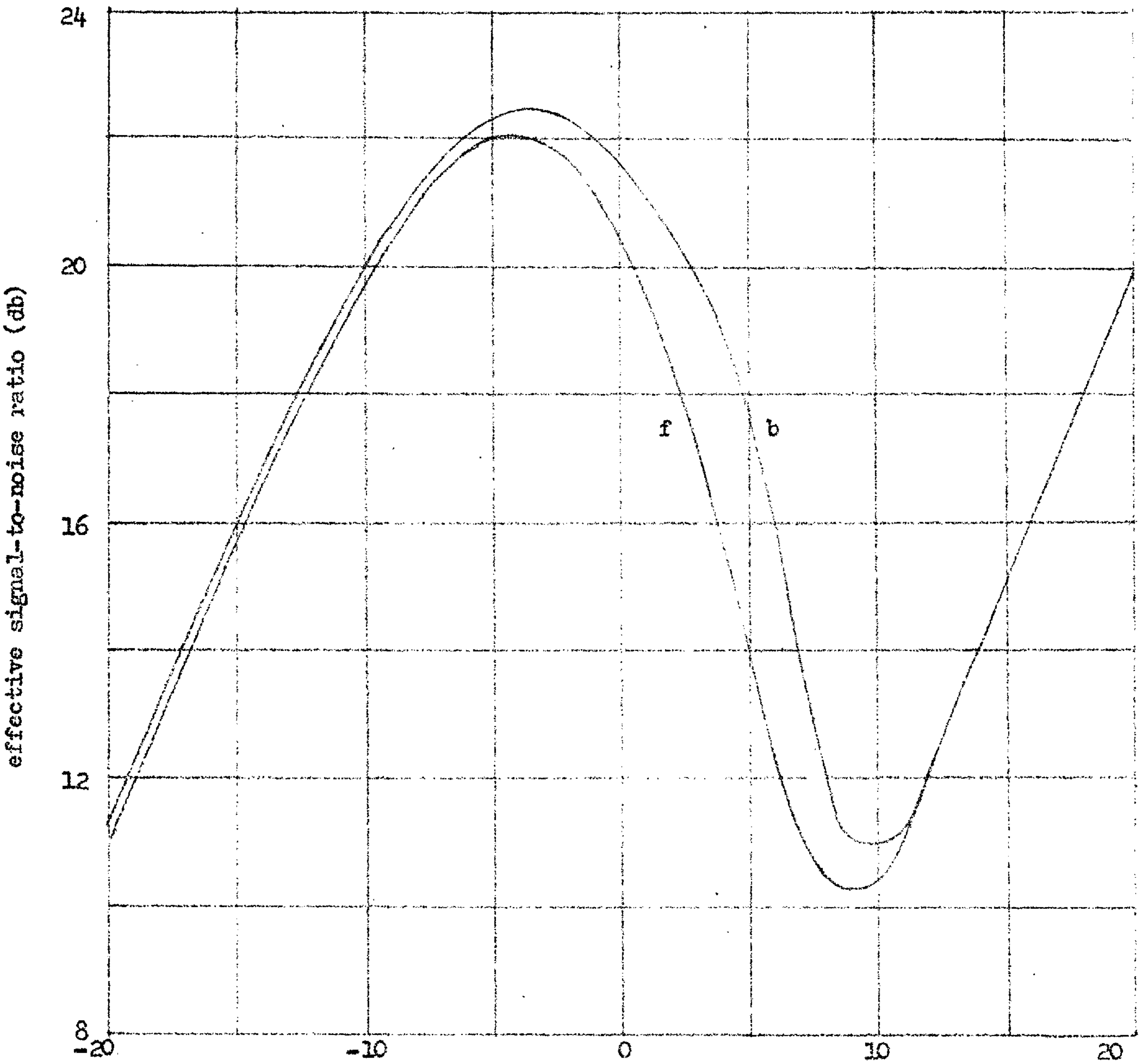


FIGURE 7.4 Effective-signal-to-noise-ratio characteristics
(4-symbol repetition cycle, Case 5 propagation)

wide range of physical conditions. In the computation of the curves labeled "f" it has been assumed that the return channel is noise-free so that all of the retransmissions are caused by erasures detected at the output of the forward channel. Curve r illustrates the speed of the forward channel when erasures occur only in the return direction. The forward-channel signal-to-noise ratio is assumed infinite and the independent variable of Curve r is the signal-to-noise ratio of the return channel. The curves labeled "b" apply to the condition of equal signal-to-noise ratios in both channels. Retransmissions, under this condition are initiated in response to erasures detected at both ends of the system.

The relative-speed curves are S-shaped -- nearly horizontal for low and high signal-to-noise ratios and monotonic increasing for intermediate values of ρ. Under poor channel conditions, $\tilde{\rho}(\rho)$ rises very steeply [with a slope of approximately $1/s(\rho)$] but it gradually levels off and begins to fall as the slope of $s(\rho)$ becomes significant. Eventually $\tilde{\rho}(\rho)$ attains a local minimum and for high signal-to-noise ratios it is nearly equal to ρ because the number of erasures is insignificant. In situations where both channels have finite but unequal signal-to-noise ratios, the curves of Figure 7.3 represent bounds on the relative speeds. If the forward-channel signal-to-noise ratio is ρ, the independent variable of Figure 7.3, the relative speed is bounded by Curves f and b. The effective signal-to-noise ratio is likewise bounded by the two curves of Figure 7.4.

7.2 System comparison

Neither the messages generated by a maximum-likelihood detector nor the output messages of an ARQ system contain erasures. Each type of message may therefore be characterized by a single error probability. In order to compare the two types of operation, error-rate curves are presented in this section which have been derived under the assumption that both systems operate in the same noise and fading environment and that both generate output symbols at the same average rate. For a given signal-to-noise ratio, this rate is determined by the relative speed of the ARQ system. The speed of the maximum-likelihood system is altered by an extension of its element-duration time from T seconds to

$$\tilde{T} = \frac{T}{s(\rho)} \qquad (7.3)$$

Its signal-to-noise ratio is thus increased to $\tilde{\rho}(\rho)$ the effective signal-to-noise ratio of the ARQ system.

In general, in this chapter, the variable $\tilde{\rho}$, when used in connection with the variable-duration maximum-likelihood system, will denote the system signal-to-noise ratio which depends on the manner in which the element duration varies. The symbol, ρ, as in other chapters of this thesis, denotes an objective measure of channel quality. It is the signal-to-noise ratio corresponding to a nominal duration T, a system constant. If $\tilde{T}$ is the actual duration,

$$\tilde{\rho} = \frac{\tilde{T}}{T}\rho \tag{7.4}$$

and the speed relative to T is

$$s(\rho) = \frac{T}{\tilde{T}} . \tag{7.5}$$

Thus in both the ARQ system and the variable-duration system, $s(\rho)$ is the average number of output symbols generated in 7T seconds and 7.1 applies. As a function of ρ, the error rate of the maximum-likelihood system may be found on the $C(\rho)$ characteristic in Chapter 6 as

$$C_M(\rho) = C(\tilde{\rho}) . \tag{7.6}$$

To facilitate the comparisons of this chapter it has been assumed that a perfect return channel exists so that $a_r = 1$ and $B_r = 0$ and

$$s(\rho) = \frac{a_f(a_f+B_f)}{\nu-(\nu-a_f)(a_f+B_f)} . \tag{7.7}$$

All of the retransmissions are due to erasures detected at the output of the forward channel. In this condition the ARQ error rate may be defined as

$$C_A(\rho) = \frac{B_f}{a_f + B_f} , \qquad (7.8)$$

the proportion of incorrect symbols in the ARQ output message. This function represents the message quality more explicitly than the error rate plotted by Van Duuren (57), who calculates the ratio of incorrect outputs to the total number of transmitted symbols, a figure which is lower than $C_A(\rho)$ by a factor of $s(\rho)$.

Relative to the performance obtained in practical systems, the error rates defined in Equations 7.6 and 7.8 are unrealistic in two respects. They depend upon a hypothetical means of adjusting the duration of the signal elements in the maximum-likelihood system and they fail to account for the effects of noise in the return channel of the ARQ system. If there were noise in the return channel, the relative speed of the forward channel would be lower than it is under the assumed conditions and the effective signal-to-noise ratio would be higher. The maximum-likelihood error rate would, therefore, be lower than the value obtained when the return channel is perfect. The ARQ error rate would, on the other hand, be higher* because some of the correctly detected symbols would be diverted to supervise retransmissions over the return channel. The assumption of a perfect return channel, therefore, makes the ARQ system appear more favourable relative to a maximum-likelihood system than it is in a practical situation. It is the maximum-likelihood system which is overrated, however, by the assumption that it adjusts its element time duration to match the maximum-likelihood transmission rate to the ARQ output rate. An important attraction of the ARQ system is its easily implemented means of adjusting the relative speed. In this chapter, the ARQ error-rate curves represent the characteristics of a practical system while the maximum-likelihood error rates indicate what might be achieved if an adaptive mechanism were available. The possibility of developing one is discussed in Section 7.3.

From the point of view of an external observer, there are four variables involved in a comparison of the two systems. For a given type of fading and phase

---

* The general formula for $C_A(\rho)$ is:

$$C_A(\rho) = \frac{B_f}{a_f(a_r+B_r)+B_f}$$

information, the quality of the channel is measured by ρ, the independent variable of the comparison. The dependent variables are $C_M(\rho)$ and $C_A(\rho)$, the two error rates and $s(\rho)$ which indicates the speed at which messages are produced. Another variable is ν, the number of symbols in the ARQ-system repetition cycle. It is not directly observable outside of the system but it may be set independently of ρ and it has an important effect on the relative performance of the two systems. For the most part, the material of this chapter pertains to ARQ systems with a four-symbol repetition cycle because this is the length most common in practical applications. Figure 7.3 illustrates the relative speed of such a system at a non-coherent receiver of non-fading signals. Figure 7.5 compares the error rate of this system with that of a maximum-likelihood system. The $C_M(\rho)$ characteristic has been obtained, according to Equation 7.6, from Curve f in Figure 7.4, which gives $\tilde{\rho}$ and from the $C(\rho)$ curve of Figure 6.7.

Two important aspects of the data in Figure 7.5 are the difference in shape of the two curves and the fact that they cross at a signal-to-noise ratio of approximately 9. For lower values of ρ, the maximum-likelihood system offers better performance than the ARQ system. For higher signal-to-noise

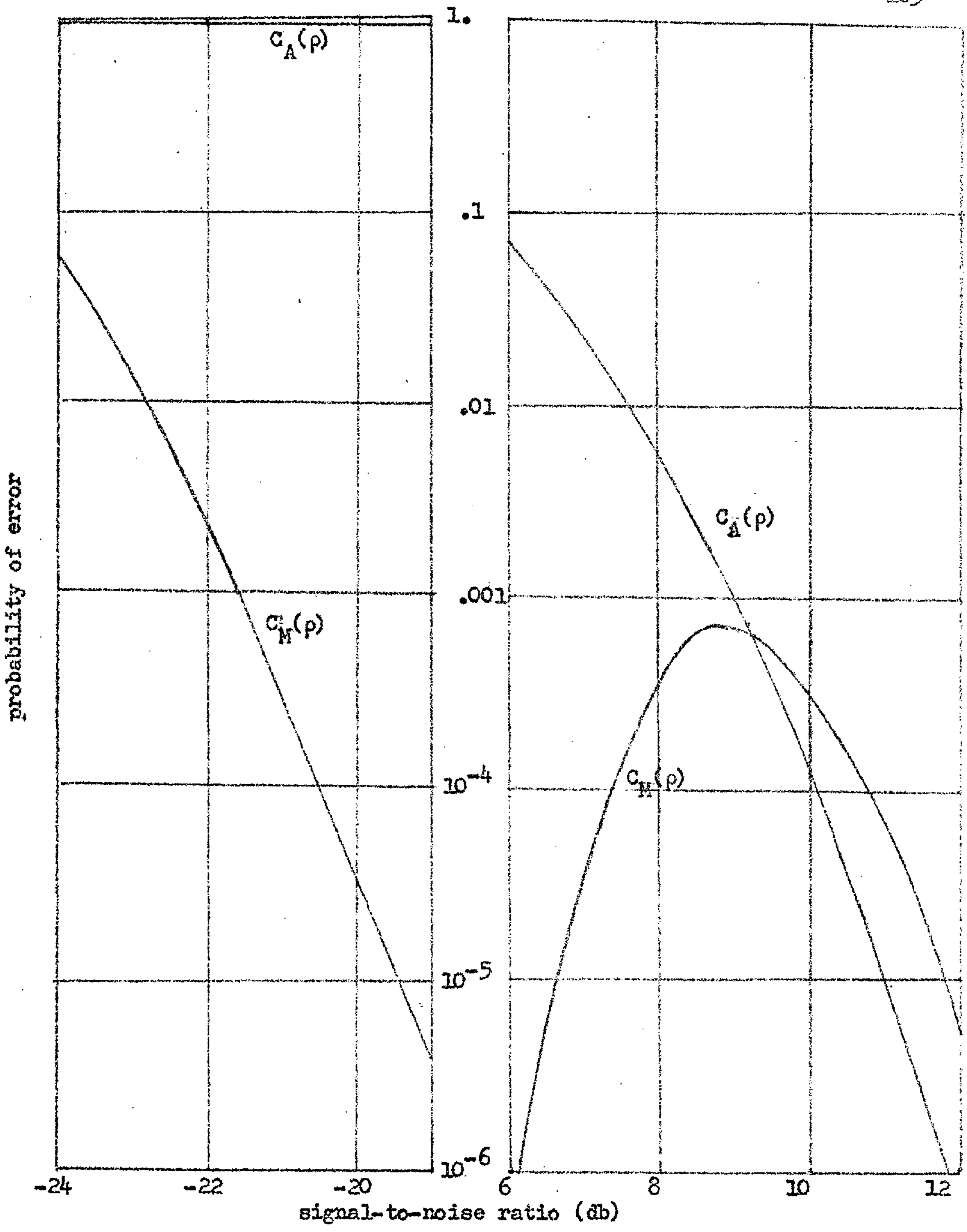


FIGURE 7.5 Comparison of error rates--ARQ system and variable-duration maximum-likelihood system

ratios, the ARQ system is preferable. Under very good channel conditions, the noise is only occasionally strong enough to prevent the detection of a correct symbol. It is only when one of the infrequent erasures appears that the ARQ system initiates a retransmission. In this way it uses additional energy only when the channel demands it and this efficient operation at high values of ρ leads to an error rate which asymptotically approaches $B(\rho)$. In the maximum-likelihood system, the energy is distributed uniformly over all of the message symbols, but the incremental energy per symbol due to repetitions in the ARQ system becomes negligible as ρ increases. Under very good channel conditions, $\tilde{\rho} \approx \rho$ and $C_M(\rho)$ asymptotically approaches $C(\rho)$.

At lower signal-to-noise ratios, on the other hand, effectively strong noise samples occur more consistently. A high proportion of the symbols transmitted by the ARQ system are detected as erasures or are ignored by the detector because they are part of a block of four symbols beginning with an erasure. Although the received signal contains useful information about these symbols, the detector makes no use of this information when it requests their retransmission. The energy used to transmit them is wasted. In the maximum-likelihood system all of the energy associated with a symbol is utilized by the detector. As the signal-to-noise ratio falls and the slope of the relative-speed curve becomes significant, the energy per symbol in the maximum-likelihood system rises steeply. The result is a very low maximum-likelihood error rate over a wide range of signal-to-noise ratios.

Except for the condition of virtually no signal, the maximum-likelihood error rate in Figure 7.5 does not exceed the respectable value of .0008. The error-rate curve of the ARQ system on the other hand rises monotonically with decreasing ρ to a value at $\rho = 0$ of 34/35. This is also the maximum-likelihood error rate when there is no signal but due to the very low relative speed in this condition, the effective signal-to-noise ratio is approximately 1,300 times as great as ρ. Thus, for ρ as low as .01 (-20db) the maximum-likelihood error rate falls to .00006. For ρ=.01, the ARQ system rejects almost all of the transmitted symbols. In its output message 97% are incorrect. If the ARQ system has an element transmission rate of 100 bauds the output rate of the two systems is one symbol every 90 seconds, a speed which is unacceptable in most telegraph applications. In situations where there is a use for messages produced at this rate, very reliable communication may be obtained with a maximum-likelihood system. An ARQ system is worthless in the low signal-to-noise ratio environment, however, due to its very high error rate.

Practical systems may be judged on their ability to meet minimum speed and maximum error rate specifications. If the lowest tolerable speed is equivalent to a relative speed less than 0.9 [the value of $s(\rho)$ corresponding to the cross-over signal-to-noise ratio of Figure 7.5], and the highest allowable error rate exceeds 0.0008, the maximum-likelihood system is able to meet the performance specifications over a wider range of channel conditions than the ARQ system. For high-quality channels ($\rho \geq 9$), the ARQ system offers the better results but both systems meet the speed and error requirements. The greatest advantage of the maximum-likelihood system is realized when the signal-to-noise ratio is between -20db and +6db. Over this range, the relative speed rises from .00075 to .219 while the error rate is neglible (less than 10^{-7}, one error in 10 million symbols). The ARQ error rate is never less than .06 at these signal-to-noise ratios. Section 7.3 contains a more general discussion of attainable maximum-likelihood characteristics.

In order to demonstrate the effects of the repetition-cycle length, Figures 7.6 and 7.7 present the speeds and error rates for several values of ν under the propagation conditions (Case 5 in Table 3.7) pertaining to Figures 7.3, 7.4 and 7.5. In Figure 7.6, the relative-speed characteristic for $\nu = 4$ is identical to Curve f in 7.3 and the characteristics corresponding to other values of ν have similar S-shapes. For low signal-to-noise ratios, $s(\rho)$ is approximately proportional to $1/\nu$. For high values of ρ, almost every error initiates one and only one repetition cycle so that the relative speed differs from unity by approximately ν times the probability of erasure. Thus for ρ high,

$$s(\rho) \approx 1-\nu[A(\rho) - B(\rho)] .$$

The error rate of the ARQ system is independent of ν and depends only on $A(\rho)$ and $B(\rho)$ in Figure 6.7. The speed of an ARQ system varies with the length of its repetition cycle; the error rate is a function only of the channel signal-to-noise ratio. The quality of maximum-likelihood output messages varies inversely with speed and thus improves as ν increases. For low values of $\rho, \tilde{\rho}$ is approximately proportional to $\nu\rho$ and the error-rate curves are roughly parallel and separated from the $\nu=1$ curve by $10 \log \nu$ db in the horizontal direction. At high signal-to-noise ratios, $\tilde{\rho} \approx \rho$ regardless of ν so that

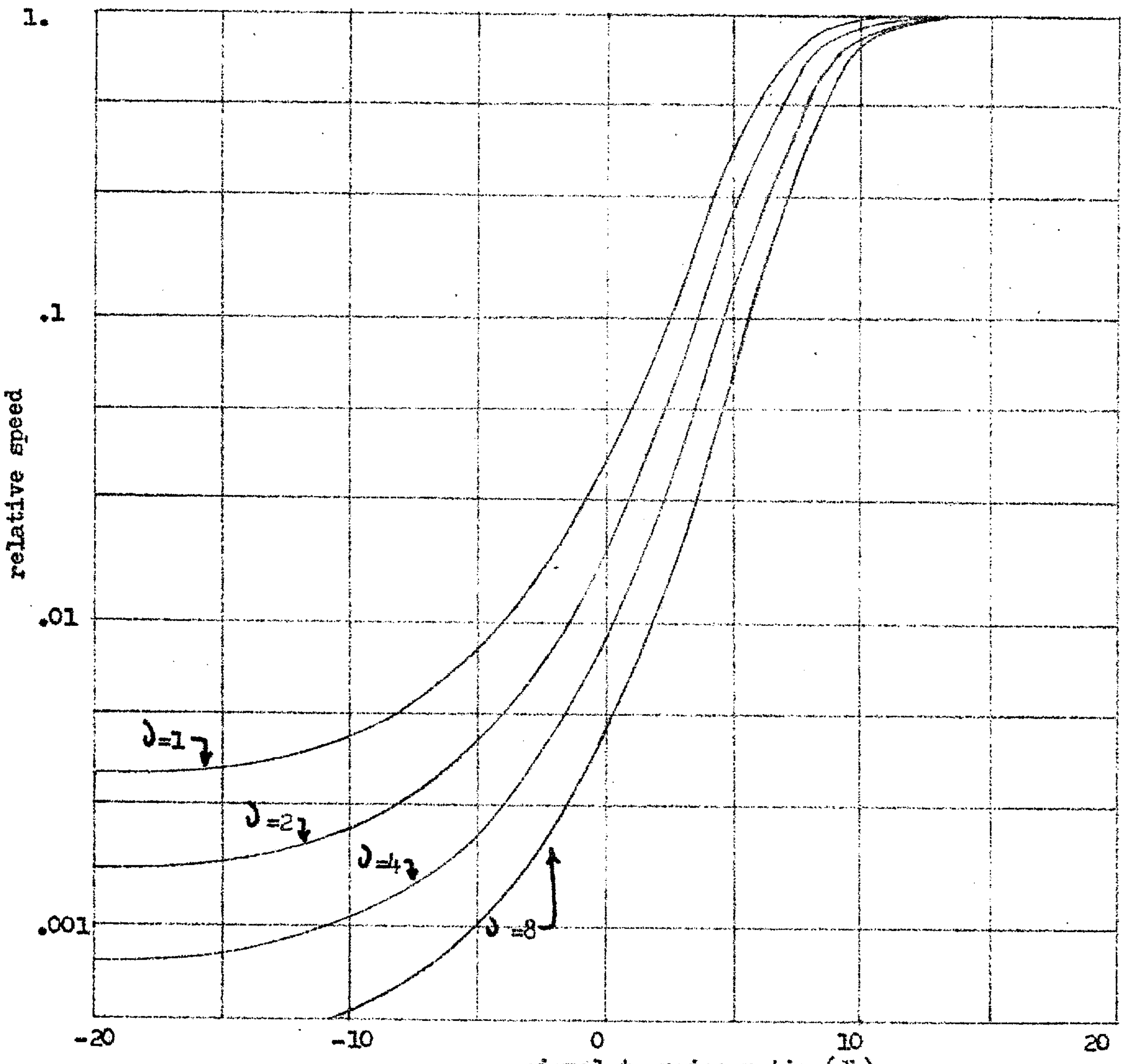


FIGURE 7.6 Relative-speed characteristics--effects of variations in repetition-cycle length (Case 5 propagation)

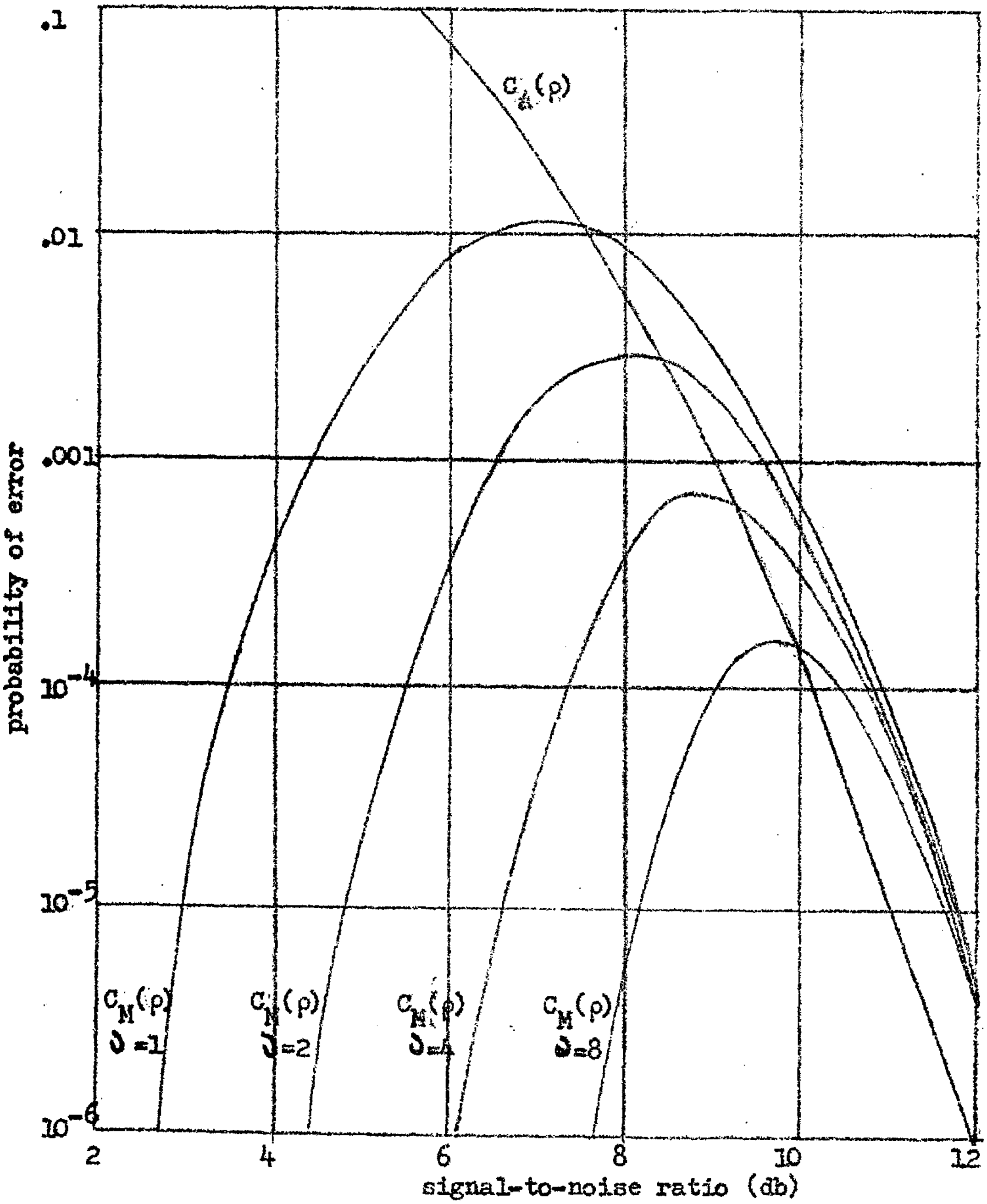


FIGURE 7.7 Comparison of error rates--effects of variations in repetition-cycle length

all of the curves approach $C(\rho)$. The value of ρ for which both systems offer equal performance is an increasing function of ν.

Figures 7.8 to 7.14 contain performance data for the seven other propagation conditions. In each figure, the relative speed and the two error rates are shown as functions of signal-to-noise ratio for $\nu=4$. The curves have been derived from the characteristics of Chapter 6 and the maximum-likelihood error rates corresponding to fading channels with low average signal-to-noise ratios are presented only as approximate guides to be accepted with caution. As the relative speed decreases with decreasing ρ, the element-duration increases and the assumptions underlying the maximum-likelihood characteristics become vulnerable in one of two respects. For moderately low values of $s(\rho)$ (between 0.5 and 0.1 approximately), the fading becomes increasingly selective between elements. The assumption of between-element correlation, a basis of the curves of Cases 2,3, 6 and 7, thus becomes increasingly untenable and the maximum-likelihood characteristic of Figure 7.11 or 7.14 (depending on whether reception is coherent or non-coherent) becomes increasingly applicable regardless of the physical nature of the fading.

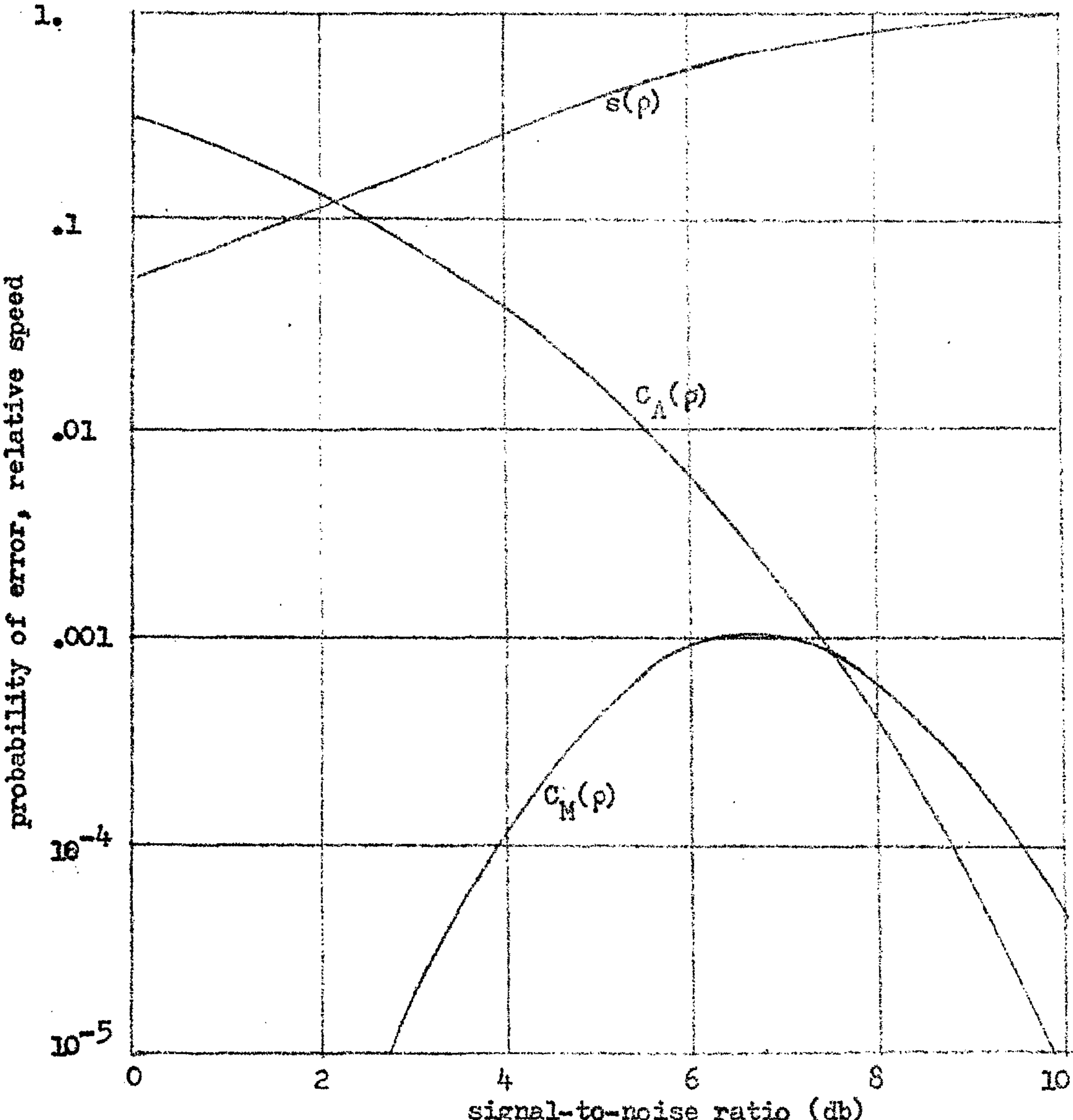


FIGURE 7.8 ARQ and variable-duration characterisitics
Case 1: Coherent receiver, no fading

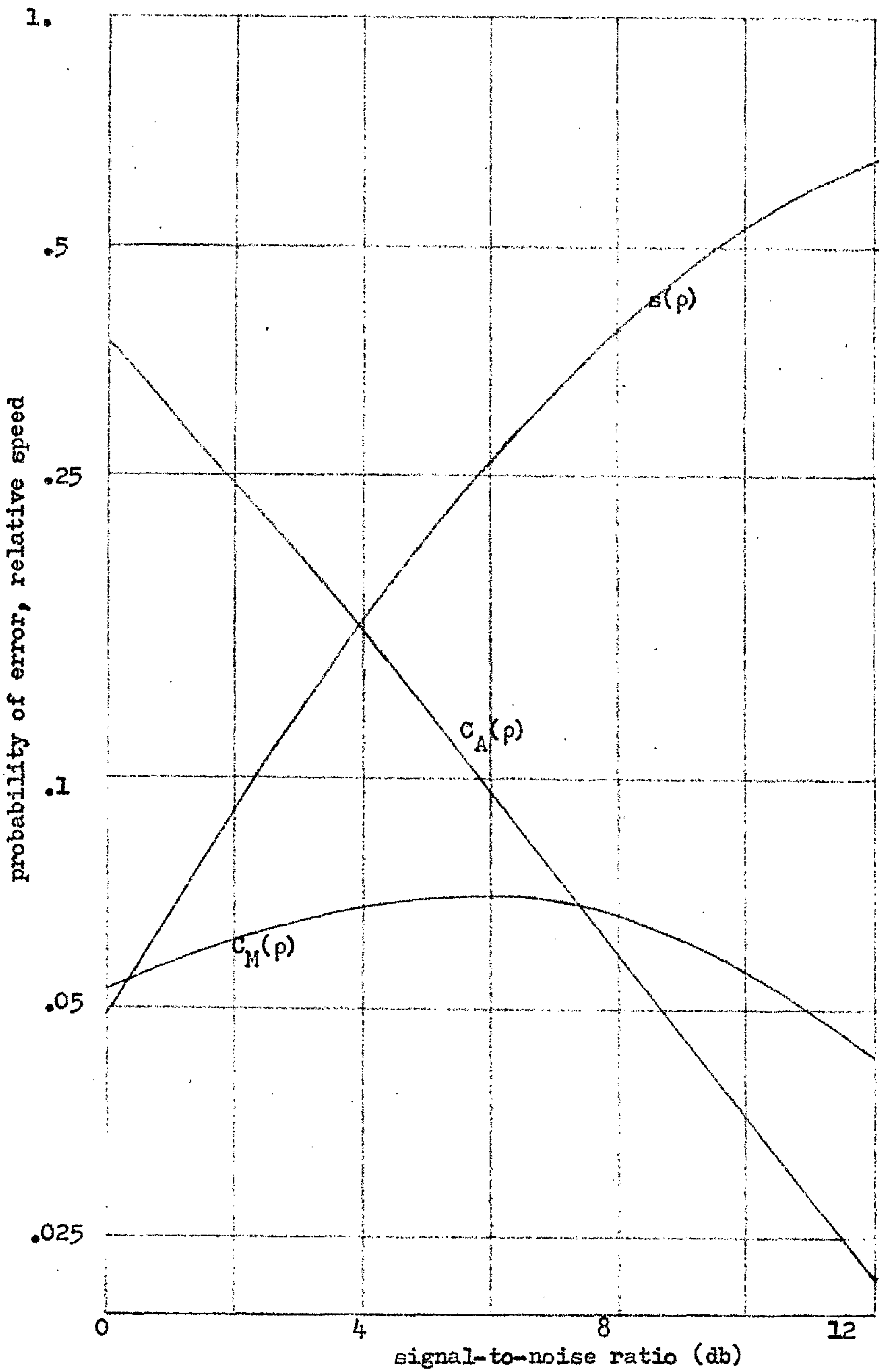


FIGURE 7.9 ARQ and variable-duration characteristics
Case 2: coherent receiver, slow flat fading

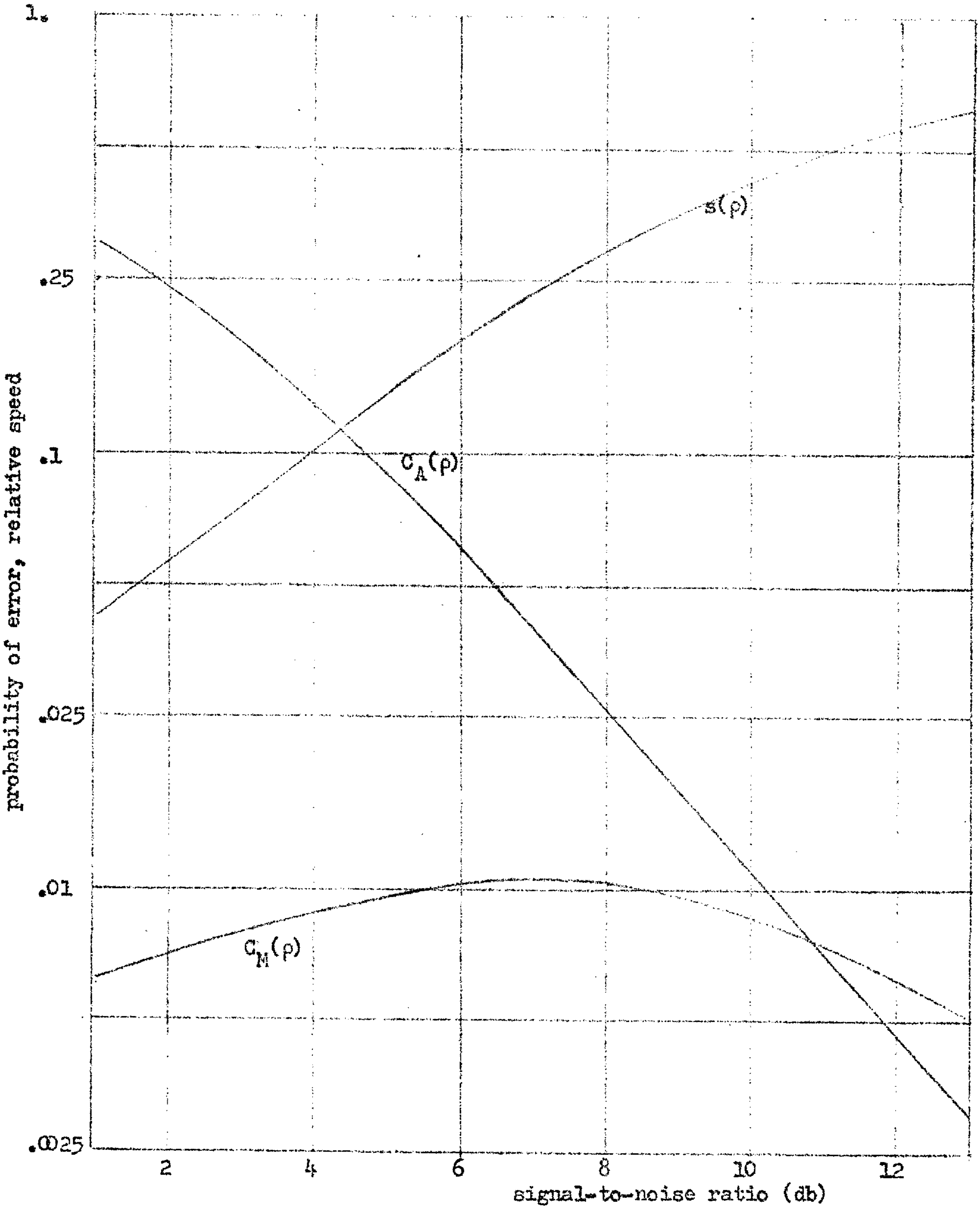


FIGURE 7.10 ARQ and variable-duration characteristics
Case 3: Coherent receiver, slow selective fading

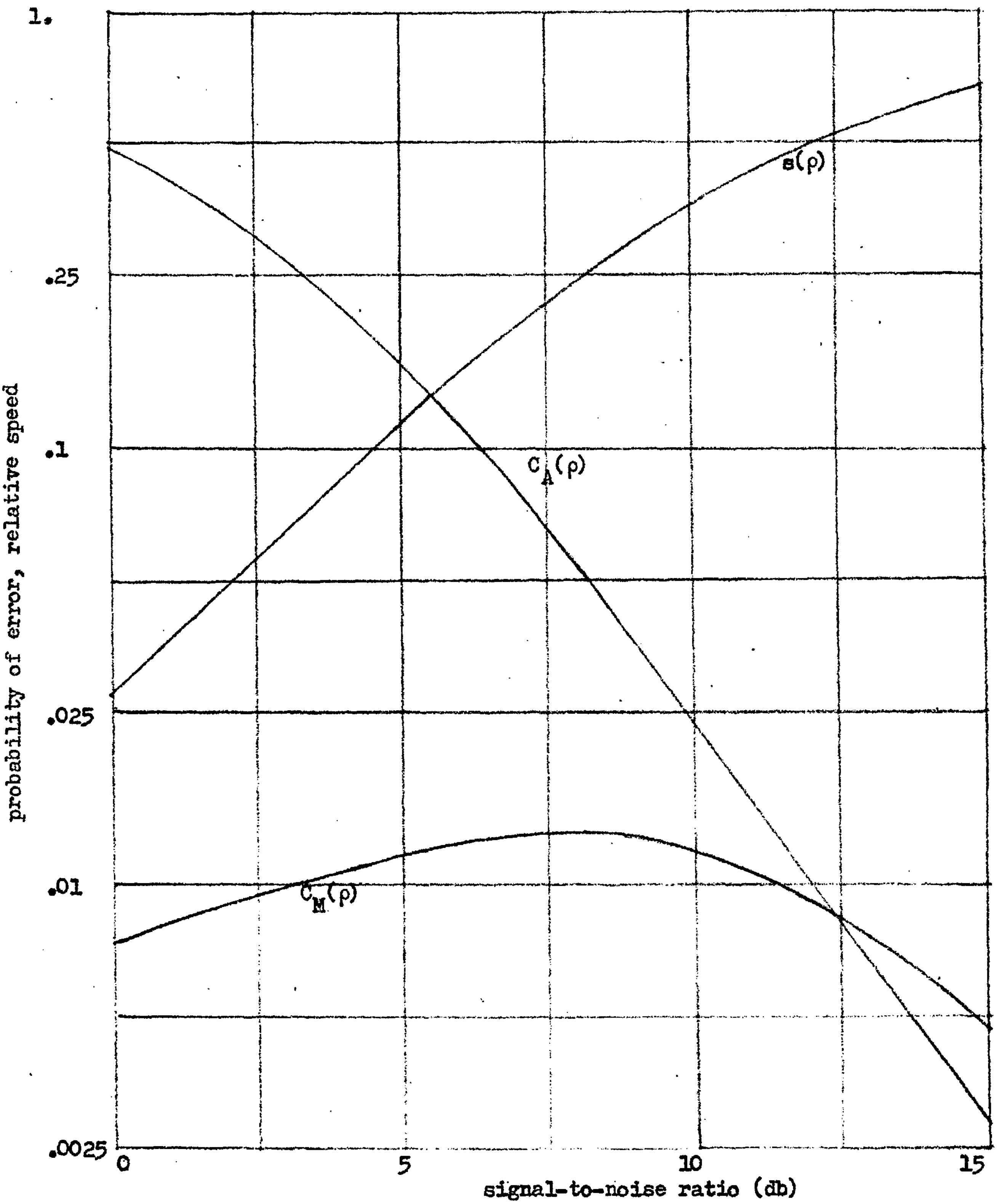


FIGURE 7.11 ARQ and variable-duration characteristics
Case 4: Coherent receiver, fast fading

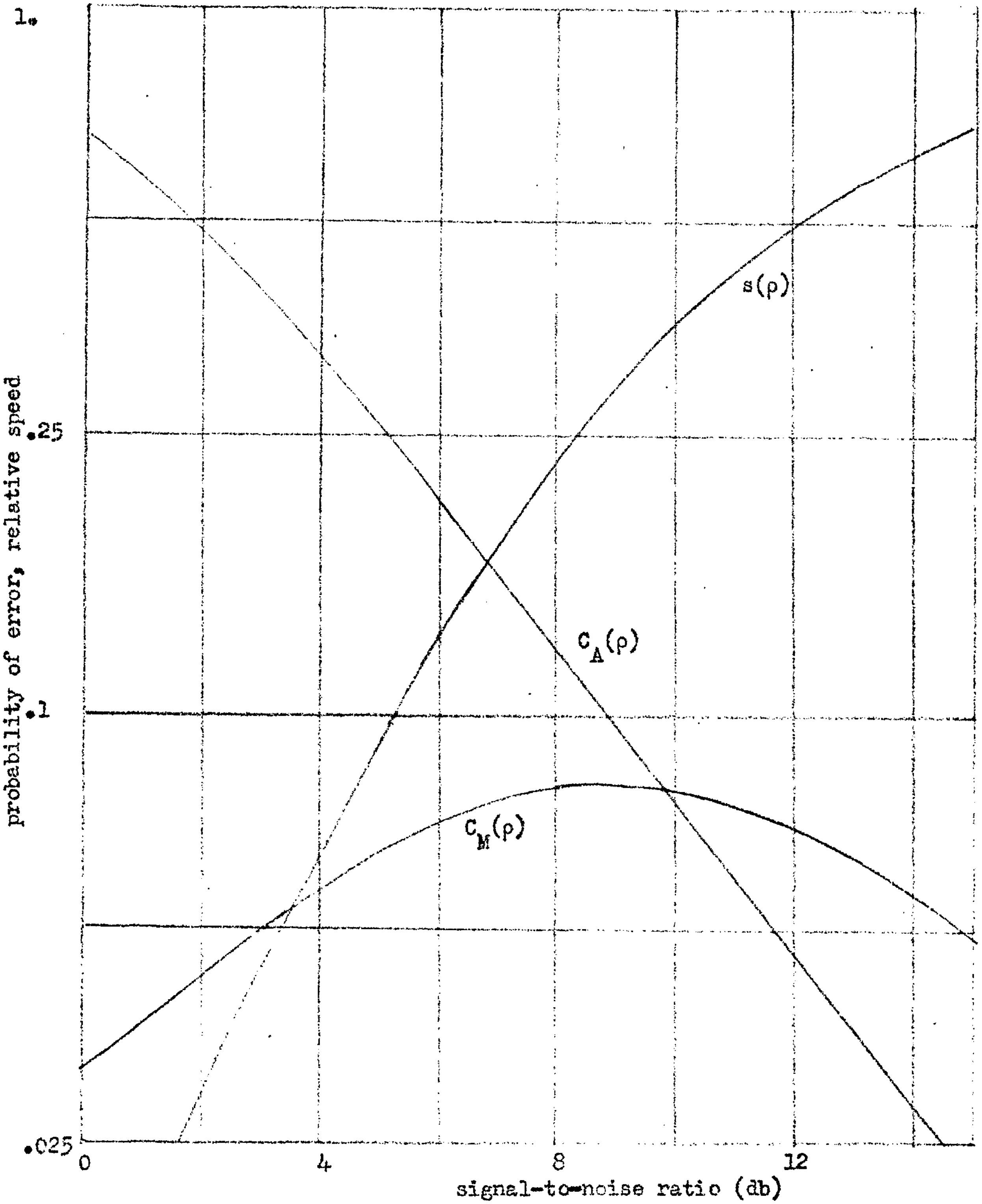


FIGURE 7.12 ARQ and variable-duration characteristics
Case 6: non-coherent receiver, slow flat fading

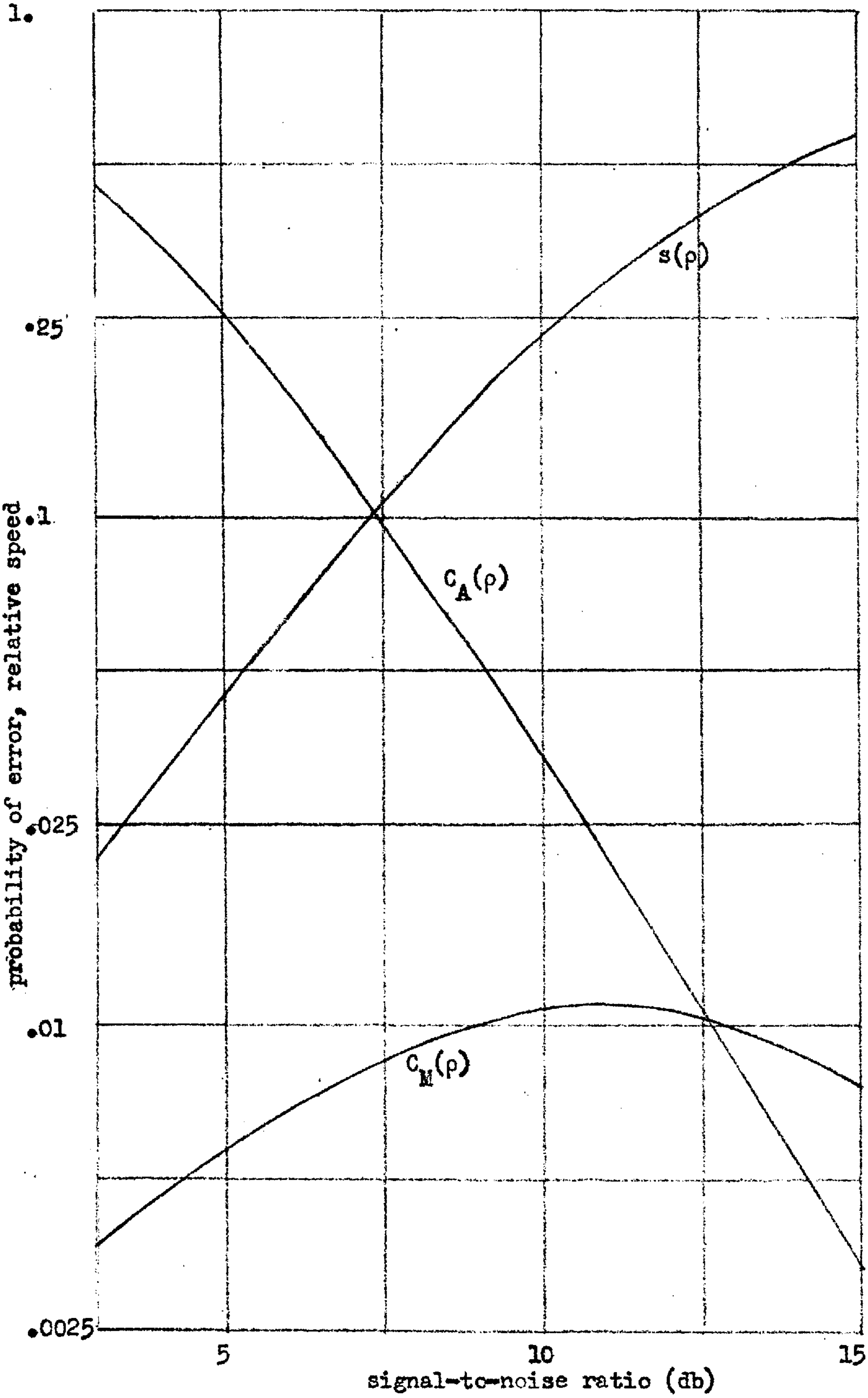


FIGURE 7.13 ARQ and variable-duration characteristics.
Case 7: non-coherent receiver, slow selective fading

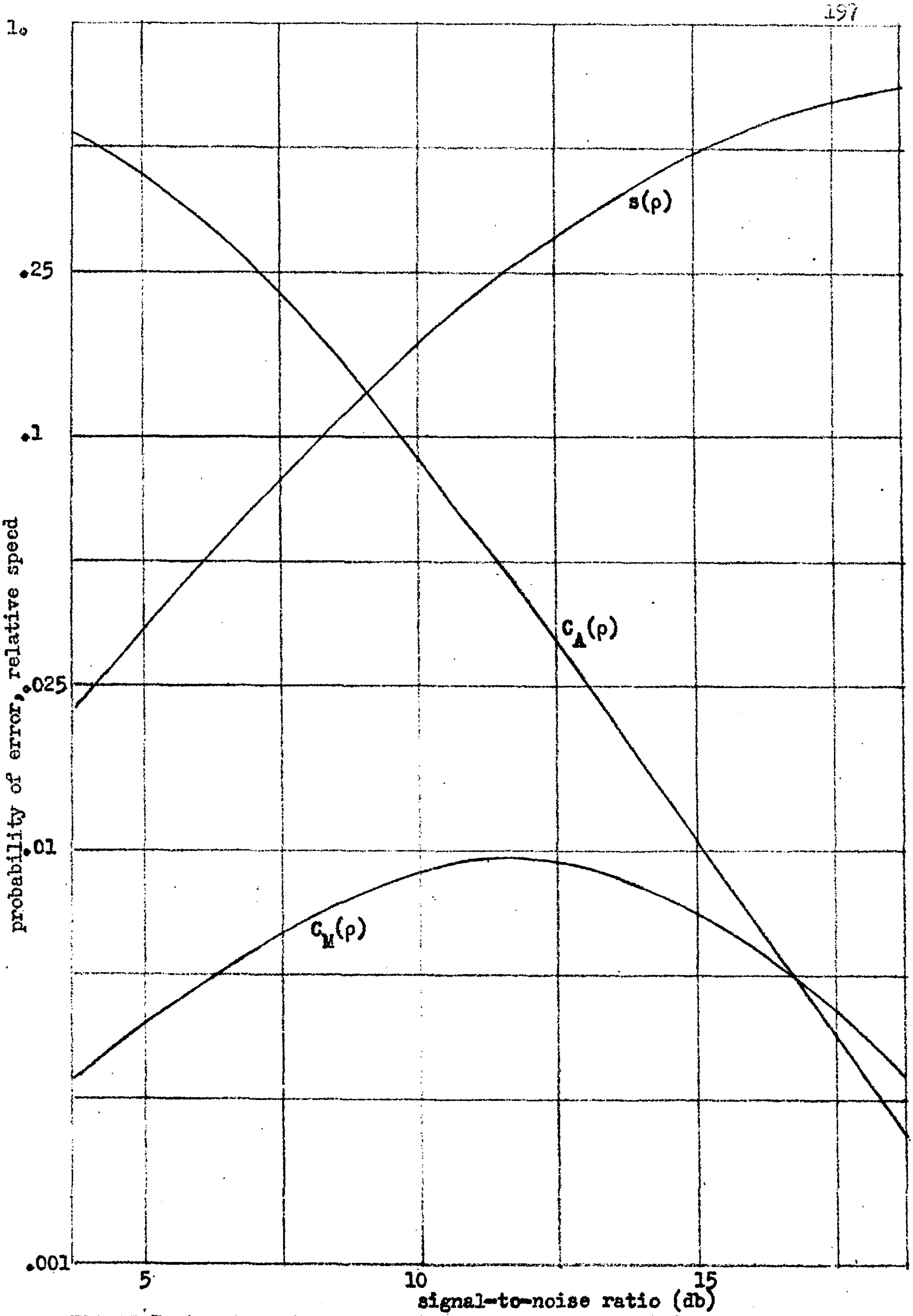


FIGURE 7.14 ARQ and variable-duration characterisitics
Case 8: non-coherent receiver, fast fading

For lower relative speeds, a second assumption -- that of constant amplitude over the duration of each element -- is likely to be violated. Since the validity of the measurement procedure of Chapter 4 is based on this assumption, the entire detection problem must be re-analyzed in order to provide accurate performance data for a variable-duration system operating at low speed in a fading environment. An analysis of this situation, which conforms to a model of signal detection in non-stationary noise, is beyond the scope of this thesis.

7.3 Practical application.

The performance characteristics of the preceding section offer a comparison of two means of varying the symbol output rate of a telegraph system in order to control the error rate. The ARQ method is based on repetition of the symbols detected as erasures by a binary detector. At high signal-to-noise ratios, its error-control capability is better than that of the other technique; its feasibility has been demonstrated and is verified by its widespread use in commercial systems. For a given operating speed, the variable-duration maximum-likelihood method has the lower error rate at low and intermediate signal-to-noise ratios. Its advantage is especially pronounced in extremely adverse conditions. As the signal-to-noise ratio falls, the quantity of useful information generated at the output of the ARQ system deminishes rapidly. A maximum-likelihood system operating at the same speed as the ARQ system produces high quality messages at signal-to-noise ratios as low as -20db.

These properties of the two systems suggest that the ARQ system is the one more suited to cable telegraphy because in most practical situations a high signal-to-noise ratio can be consistently maintained over this transmission medium. In the case of hf radio channels, on the other hand, it is not unusual for the signal-to-noise ratio to lapse to very low values. This property of the hf channel implies that the quality of a two-way radiotelegraph system could be enhanced by the introduction of maximum-likelihood detectors and a mechanism which uses information about the condition of each channel to control the time duration of its signal elements. Although the practical implementation of a maximum-likelihood detector is a straightforward matter (44), the configuration of a practical variable-speed mechanism remains to be established. Further study is required to determine the means of extracting at a receiver information relevant to the state of the channel (43), a means of reliably communicating this information over the noisy return path from receiver to transmitter (60), and a procedure for using the information fedback from the receiver to control the duration of transmitted elements. Until these techniques are developed and their feasibility demonstrated, ARQ signalling, in spite of its inferior theoretical characteristics, will remain more attractive to system designers because of its established performance.

Although a detailed design investigation of a variable-duration mechanism is outside the scope of this thesis, some of the desired properties of such a mechanism may be discussed in general terms. The basic system objective would be the maintenance of a message error rate which does not exceed a specified limit. Given this limit, ε, and the propagation conditions under which the system must operate, the designer may use the appropriate $C(\rho)$ characteristic in Chapter 6 to determine the system signal-to-noise ratio, $\tilde{\rho}_\varepsilon$, necessary to achieve the specified error rate. Thus,

$$\varepsilon = C(\tilde{\rho}_\varepsilon) \quad . \tag{7.9}$$

The quality of the channel is reflected in the value of ρ, the average signal-to-noise ratio corresponding to a nominal duration, T. For the purpose of this discussion, T may be taken as the shortest time duration which satisfies the requirements of orthogonality in time (no appreciable interelement interference due to multipath propagation) and orthogonality in frequency (Equation 4.4)

When the channel signal-to-noise ratio, ρ, is higher than $\tilde{\rho}_\varepsilon$, the system signal-to-noise ratio necessary to meet the error-rate specification, the system may operate at maximum speed (with elements of T seconds duration) so that the system signal-to-noise ratio, $\tilde{\rho}$, is equal to ρ. The system error rate, $C_M(\rho)$, equals, in this situation, $C(\rho)$, a value lower than ε. Under poorer channel conditions, when $\rho < \tilde{\rho}_\varepsilon$ the element duration must be extended to

$$\tilde{T} = \frac{\tilde{\rho}_\varepsilon}{\rho} T \qquad (7.10)$$

so that the system signal-to-noise ratio, given in Equation 7.4, is

$$\tilde{\rho} = \tilde{\rho}_\varepsilon \qquad (7.11)$$

and the error rate is ε. The error-rate characteristic of a system operating according to this plan may be expressed as

$$C_M(\rho) = \varepsilon \quad \text{for } \rho \leq \tilde{\rho}_\varepsilon$$

$$C_M(\rho) = C(\rho) \quad \text{for } \rho \geq \tilde{\rho}_\varepsilon \quad . \qquad (7.12)$$

The transmission speed relative to T (Equation 7.5) becomes

$$s(\rho) = \frac{\rho}{\tilde{\rho}_\varepsilon} \quad \text{for } \rho \leq \tilde{\rho}_\varepsilon$$

$$s(\rho) = 1 \quad \text{for } \rho \geq \tilde{\rho}_\varepsilon \quad . \qquad (7.13)$$

Speed and error-rate characteristics corresponding to three possible values of ε are illustrated in Figure 7.15 for the case of non-coherent propagation of non-fading signals.

The performance objectives proposed in Equations 7.12 and 7.13 may be met exactly only by an ideal system in which precise information of the state of each one-way channel is available at its transmitting terminal. A practical feedback system must perform three control functions: channel measurement, communication of control information and adaptation of signal characteristics. The first two are inherently liable to error due to the randomness of the channels and the design of all three operations must anticipate and minimize the harmful effects of these errors on system performance. In addition to meeting this reliability requirement, a practical feedback system must be economical in terms of the proportion of total signalling capacity devoted to the communication of control information.

One means of realizing the variable-duration

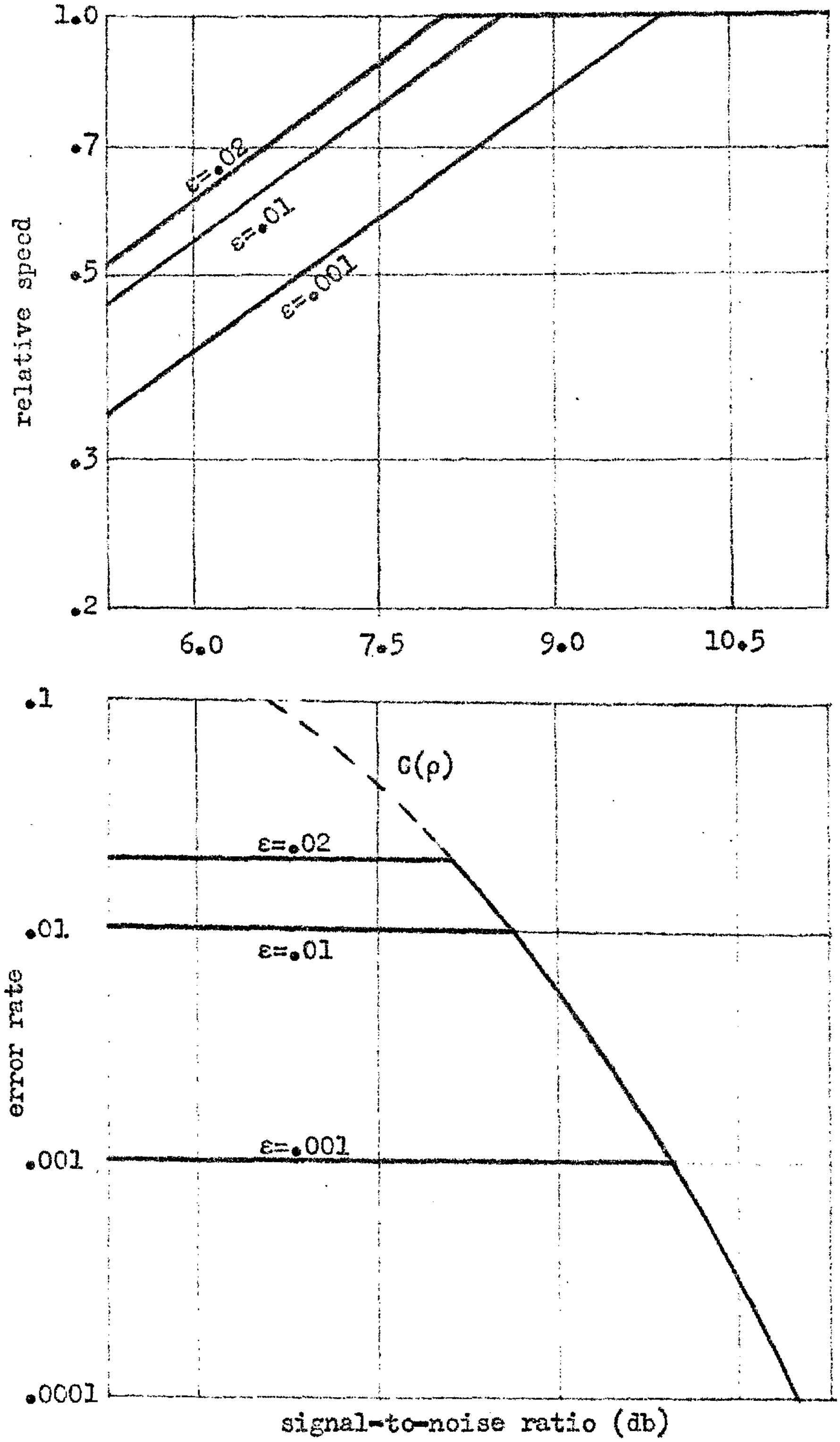


FIGURE 7.15 Proposed characteristics of variable-duration maximum-likelihood system

mechanism would be to organize the measurement and control functions on the basis of multiplex frames. If one message in a multiplex group of 24 were reserved for the transmission of control information, the capacity of the system to convey telegraph messages would be reduced by less than 4.2%*, a reasonable price for the error-control advantages of variable-duration operation. The variable which controls the element duration of each multiplex frame is the average signal-to-noise ratio, a statistical parameter of the signal measurements obtained from the matched filters. Techniques of parameter estimation may be applied, therefore, in order to obtain a measure of the signal-to-noise ratio by means of computations involving the 336 signal measurements associated with each multiplex frame. A new estimate may be computed after the reception of each multiplex frame and the estimate may contain information relating to previous frames if it can be assumed that the average signal-to-noise ratio varies slowly relative to the frame duration time.

In order to achieve reliable transmission of control information, the number of possible control

---

*The effect of this reduction on the performance characteristics may be accounted for by multiplying the relative speed by .958.

signals should be restricted. Such a restriction would permit highly redundant coding of the control information and insure a very low error probability in its reception, especially because the speed of both channels is controlled in order to maintain a high effective signal-to-noise ratio, $\tilde{\rho}$. The effects of the few errors which would still occur in the reception of control information may be mitigated by a restriction on the maximum allowable change in element duration from frame to frame.

The adoption of this type of plan would preclude the attainment of the performance specifications of Equations 7.12 and 7.13. These idealized objectives have the status of a guide to the shape of the operating characteristics which must be specified within the constraints of a practical design. In any event, the foregoing plan is only a tentative outline; the formal design of a variable-duration system requires extensive study. The performance comparisons of Section 7.2 indicate that the benefits over ARQ operation to be derived from a variable-duration maximum-likelihood system justify this further effort.

CHAPTER 8 : MAXIMUM LIKELIHOOD DETECTION OF FIXED-RATIO CODES

Maximum-likelihood detection of symbols represented by a fixed-ratio code may be viewed as an extension of the detection procedure described by Silverman and Balser in their paper on Wagner codes (51). The fixed-ratio code constraint implies that the identity of each maximum-likelihood output symbol may be determined from the relative magnitudes of the binary likelihood ratios; the detector is not required to compute and compare the symbol likelihoods.

The present chapter is a mathematical presentation of the properties of the maximum-likelihood detector applicable to a general fixed-ratio code. The decision process presented in Section 6.1 (with reference to the four-out-of-seven code) is formally shown in Section 8.2 to conform to the maximum-likelihood decision rule. In Section 8.3, the general formula (Equation 8.10) for the detector error rate is derived and it is applied in Section 8.4 to systems operating under the four fading conditions formulated in Chapter 3. Sections 8.5 and 8.6 provide the basis of the derivation of the error-rate formulas given in 8.7 as integral functions of signal-to-noise ratio. It is these formulas (Table 8.1) that have been integrated numerically to produce the performance characteristics of Chapter 6.

The detection method formulated in this chapter is described by Barrow (6) who presents a formula equivalent to Equation 8.10 and performance characteristics applicable to detectors operating at coherent receivers. Rosie(44) has developed a practical detector of symbols represented by the two-out-of-five code. The principal contribution of this thesis to the subject is an extension of the detection principles to the case of non-coherent reception. For this situation a square-law approximation (13) to the theoretically precise likelihood statistic is adopted (in Section 8.5).

8.1 Notation and assumptions

Message symbols are represented by an n-element binary code which is constrained such that the code sequence for each element has m "ones" and s = n-m "zeros". The alphabet of symbols

$$X = X_1, X_2, \ldots, X_M$$

has therefore,

$$M = \frac{n!}{s!m!} \tag{8.1}$$

members. The sequence of elements in the binary code word for X_i is

$$b_{i1}\ b_{i2}\ \dots\ b_{in}$$

such that

$$\sum_{k=1}^{n} b_{ik} = m \ . \tag{8.2}$$

Each code element causes a two-tone modulator to generate a signal element of T seconds duration. The binary signal elements have equal energy and are uncorrelated. They are transmitted over a high-frequency radio path characterized by amplitude fading and by additive white Gaussian noise whose power spectral density is N. The signal-to-noise ratios at the receiver are

$$\rho_o = \frac{A_o^2 T}{4N} \quad \text{and} \quad \rho_1 = \frac{A_1^2 T}{4N} \tag{8.3}$$

where A_o and A_1 are the amplitudes of the signal components of the received waveform. At a coherent receiver, the phase of the received signal element is a known constant. At a non-coherent receiver it is a random variable, uniformly distributed over the range 0 to 2π radians. At the receiver, the waveform associated with each transmitted element is processed so that two data -- a_o and a_1 at a coherent receiver and c_o and c_1 at a non-coherent receiver -- are measured. The set of 2n measurements associated with a telegraph symbol is denoted $\overline{\alpha}$. The symbol likelihood function is $f_i(\overline{\alpha})$, the probability density of $\overline{\alpha}$ when X_i is transmitted. The likelihood may be expressed in terms of the 2n element likelihoods $\lambda_o^{(k)}$ and $\lambda_1^{(k)}$ (k = 1, 2, ..., n), which may be computed from the element measurements -- according to Equation 4.8 at a coherent receiver and 4.9 at a non-coherent receiver. The symbol and element likelihoods are related through the code configuration by the formula:

$$f_i(\overline{\alpha}) = \prod_{k=1}^{n} \left[\lambda_o^{(k)}(1-b_{ik}) + \lambda_1^{(k)} b_{ik}\right] . \tag{8.4}$$

8.2 The Maximum-Likelihood Decision Rule

The decision Y_j that X_j is the transmitted symbol conforms by definition to a maximum-likelihood decision rule if $f_j(\overline{\alpha})$ is a maximal member of

$$f_1(\overline{\alpha}),\ f_2(\overline{\alpha}),\ \ldots,\ f_M(\overline{\alpha}) \quad .$$

If a unit-cost (minimum-symbol-error-rate) fidelity criterion is adopted, this is a Bayes decision rule relative to the uniform prior probability function.

If Y_j is a maximum-likelihood symbol, then

$$\frac{f_j(\overline{\alpha})}{f_i(\overline{\alpha})} = \prod_{k=1}^{n} \frac{\lambda_o^{(k)}(1-b_{jk}) + \lambda_1^{(k)}\ b_{jk}}{\lambda_o^{(k)}(1-b_{ik}) + \lambda_1^{(k)}\ b_{ik}} \geq 1 \tag{8.5}$$

for i = 1, 2, ..., M. Each of the n factors in this product may be expressed in terms of the likelihood ratio

$$L_k = \frac{\lambda_1^{(k)}}{\lambda_o^{(k)}} ,$$

in the following manner:

$$\frac{\lambda_o^{(k)}(1-b_{jk}) + \lambda_1^{(k)} b_{jk}}{\lambda_o^{(k)}(1-b_{ik}) + \lambda_1^{(k)} b_{ik}} \quad \begin{aligned} &= L_k \text{ if } b_{jk}=1 \text{ and } b_{ik}=0 \\ &= 1 \text{ if } b_{jk} = b_{ik} \\ &= \frac{1}{L_k} \text{ if } b_{jk} = 0 \text{ and } b_{ik}=1 \end{aligned} \tag{8.6}$$

Because an equal number of "ones" and "zeros" must be inverted if the code sequence of X_i is to be transformed into that of X_j, Equation 8.5 is the product for each i of an equal number of L_k and $^1/L_k$ factors. Wherever L_k appears, $b_{jk}=1$ and wherever $^1/L_k$ appears $b_{jk} = 0$. The inequality in 8.5 will always be satisfied, therefore, if Y_j is chosen such that $b_{jk}=1$ for the m values of k corresponding to the m highest values of L_k.

The output of a maximum-likelihood detector depends, therefore, on the ordering of the L_k rather than their actual values. If $\hat{L}_k$ is a monotonic function of L_k, the ordering of the n values of $\hat{L}_k$ will be identical to that of the likelihood ratios themselves and the detector may thus operate on the set of $\hat{L}_k$ rather than on the likelihood ratios explicitly. This argument confirms, therefore, the specification in Sections 6.1 and 6.2 of a maximum-likelihood detector as a device which calculates from the n pairs of measured data, the n numbers $\hat{L}_k$ which have the same order as the set of binary likelihood ratios. It decides that the transmitted symbol is X_j such that $b_{jk}=1$ for the m highest values of $\hat{L}_k$.

8.3 Probability of Error

In this section, the general formula for the probability of error of a maximum-likelihood detector of fixed-ratio codes is derived. Equation 8.10 is a function of the two element signal-to-noise ratios, ρ_o and ρ_1, which are parameters of the conditional probability densities $F_o(u)$ and $F_1(u)$. For each element, the likelihood statistic $\hat{L}_k$ is a random variable which depends on the specific noise waveform encountered. The statistical properties of $\hat{L}_k$ are reflected in the functions $F_o(u)$, the conditional probability density of the event $\hat{L}_k=u$ when "zero" is transmitted and $F_1(u)$, the probability density that $\hat{L}_k=u$ when "one" is transmitted.

Correct detection of a telegraph symbol depends on the order of the n values of $\hat{L}_k$ associated with the symbol. If the m "ones" in the code word of X_i, the transmitted symbol, are associated with higher values of $\hat{L}_k$ than of any of the s "zeros", the correct decision, Y_i will result. The lowest $\hat{L}_k$ associated with a "one" must, therefore, be greater than all of the $\hat{L}_k$ associated with "zeros". If the value of the lowest $\hat{L}_k$ associated with a "one" is denoted τ, the probability that at least one of the m "ones" will result in τ is

$$mF_1(\tau)d\tau \quad . \tag{8.7}$$

For each "zero" the probability that $\hat{L}_k < \tau$ is the cumulative distribution function

$$\int_{-\infty}^{\tau} F_o(u)\, du$$

and the joint probability that all s "zeros" generate values of $\hat{L}_k < \tau$ is

$$\left[\int_{-\infty}^{\tau} F_o(u)du\right]^s \quad . \tag{8.8}$$

The joint probability that the other m-1 "ones" have $\hat{L}_k \geq \tau$ is

$$\left[\int_{\tau}^{\infty} F_1(u)\, du\right]^{m-1} . \tag{8.9}$$

A correct result requires the coincidence of the three events: a) $\hat{L}_k = \tau$ for one of the "ones" (probability given by 8.7), b) $\hat{L}_k < \tau$ for all of the "zeros" (probability given by 8.8) and c) $\hat{L}_k \geq \tau$ for the other m-1 "ones" (probability given by 8.9). The joint probability of these events is the product of their individual probabilities. Since the correct symbol is detected for any real number τ which is the lowest $\hat{L}_k$ associated with a transmitted "one", the probability of correct detection is the integral of this product over all real numbers and the probability of error is

$$C(\rho_o,\rho_1) = 1-\int_{-\infty}^{\infty} mF_1(\tau)\left[\int_{-\infty}^{\tau} F_o(u)du\right]^s \times \left[\int_{\tau}^{\infty} F_1(u)du\right]^{m-1} d\tau. \quad (8.10)$$

The signal-to-noise ratios appear in the right side of 8.10 as parameters of F_o and F_1 and this formula may be applied to any type of fading and any additive noise which is independent from element to element.

8.4 The Influence of Fading

When the received signal is subject to fading, $C(\rho_o,\rho_1)$ in Equation 8.10 is a random variable whose nature depends on the statistical properties of ρ_o and ρ_1. The average error rate is the expectation of 8.10. When there is no fading, the signal-to-noise ratios are equal constants, $\rho_o = \rho_1 = \rho$, and Equation

8.10 may be applied directly. Thus

$$C_{steady}(\rho) = C(\rho,\rho) \quad . \tag{8.11}$$

Under the three conditions of amplitude fading, ρ is the mean value of the random variables ρ_o and ρ_1. It is the statistical average of 8.3 when A_o and A_1 are Rayleigh-distributed random variables with mean-square value S. Therefore,

$$\rho = \frac{ST}{4N} \quad . \tag{8.12}$$

In the derivations of the error-rate formulas, the two instantaneous signal-to-noise ratios are best expressed in terms of the normalized measurements v_o and v_1 as

$$\rho_o = \tfrac{1}{2}\rho v_o^2 \quad \text{and} \quad \rho_1 = \tfrac{1}{2}\rho v_1^2 \quad . \tag{8.13}$$

The probability densities of v_o and v_1 are

$$p(v_o) = v_o \exp(-\tfrac{1}{2}v_o^2)$$

$$\text{and} \quad p(v_1) = v_o \exp(-\tfrac{1}{2}v_1^2) \quad . \tag{8.14}$$

8.4.1 Slow Fading

When the fading is slow, ρ_o and ρ_1, parameters of 8.7, 8.8 and 8.9, are constant over each symbol. For flat fading $\rho_o = \rho_1$ and the average error rate is

$$C_{flat}(\rho) = \int_o^\infty v \exp(-\tfrac{1}{2}v^2)C(\tfrac{1}{2}\rho v^2, \tfrac{1}{2}\rho v^2)dv$$

$$= \int_0^\infty v \exp(-\tfrac{1}{2}v^2) C_{steady}(\tfrac{1}{2}\rho v^2) dv \quad . \qquad (8.15)$$

In the event of selective fading, A_o and A_1 are statistically independent and the joint probability density of v_o and v_1 is the product of the two formulas of 8.13,

$$p(v_o,v_1) = v_o v_1 \exp[-\tfrac{1}{2}(v_o^2+v_1^2)] \quad , \qquad (8.16)$$

so that 8.10 must be averaged with respect to each of the two random variables. Thus,

$$C_{sel}(\rho) = \int_0^\infty\int_0^\infty v_o v_1 \exp[-\tfrac{1}{2}(v_o^2+v_1^2)] \times$$

$$C(\tfrac{1}{2}\rho v_o^2, \tfrac{1}{2}\rho v_1^2) dv_o dv_1 \quad . \qquad (8.17)$$

8.4.2. Fast Fading

Under this condition, the signal-to-noise ratios assume independent values for each of the n signal elements. Because the integral of 8.10 may be viewed as the product of n factors, the signal-to-noise ratio parameter of each factor is a random variable independent of the other signal-to-noise ratios. The average value of 8.10 may be derived, therefore in terms of the average values of its n factors. For this purpose, an average conditional probability density function,

$$\overline{F}_1(\tau) = \int_0^\infty v \exp(-\tfrac{1}{2}v^2) F_1(\tau) d\tau \ , \qquad (8.18)$$

may be defined. Its parameter is ρ, the average signal-to-noise ratio when the parameter of $F_1(\tau)$ in the integrand of 8.18 is given the value $\frac{1}{2}\rho v^2$. $\overline{F}_o(u)$ may be similarly defined and the average value of the integrand in 8.8 may be derived and expressed in terms of $\overline{F}_o(u)$ in the following manner:

$$\int_0^\infty v\ \exp(-\tfrac{1}{2}v^2)\left[\int_{-\infty}^{\tau} F_o(u)du\right]dv =$$

$$= \int_{-\infty}^{\tau}\left[\int_0^\infty v\ \exp(-\tfrac{1}{2}v^2)F_o(u)dv\right]du$$

$$= \int_{-\infty}^{\tau} \overline{F}_o(u)\ du\ .$$

In the same manner, the average of the integral (in 8.9) of $F_1(u)$ may be expressed as:

$$\int_{\tau}^{\infty} \overline{F}_1(u)du.$$

Thus for fast fading, the symbol error rate is:

$$C_{fast}(\rho) = 1 - \int_{-\infty}^{\infty} m\overline{F}_1(\tau)\left[\int_{-\infty}^{\tau}\overline{F}_o(u)du\right]^s \times$$

$$\left[\int_{\tau}^{\infty}\overline{F}_1(u)du\right]^{m-1} d\tau\ . \qquad (8.19)$$

8.5 The Likelihood Statistic $\hat{L}_k$

The error-rate formulas derived in this chapter are applicable to detectors which perform the data processing operations specified in Section 6.2. At a coherent receiver,

$$\hat{L}_k = a_1 - a_o \tag{8.20}$$

is computed although it is an optimum likelihood statistic only when the attenuations of the two signal elements are equal. When there is selective fading $\hat{L}_k$ is best computed as

$$\frac{(A_1 a_1 - \rho_1)}{N} - \frac{(A_o a_o - \rho_o)}{N} \tag{8.21}$$

which requires knowledge of the values of the random variables A_1, A_o, ρ_1 and ρ_o which are measures of the state of the channel. If these measures were assumed to be available and 8.21 computed, the error rates under the conditions of both fast and slow selective fading would be lower than those derived in Section 8.7.

At a non-coherent receiver,

$$\log L_k = \left[\log I_o\left(\frac{A_1 c_1}{N}\right) - \rho_1\right] - \left[\log I_o\left(\frac{A_o c_o}{N}\right) - \rho_o\right] \tag{8.22}$$

is an optimum likelihood statistic and the adoption of

$$\hat{L}_k = c_1^2 - c_o^2$$

is based upon the assumption of equal A_o and A_1 and the approximation of $\log I_o(\chi)$ by $\frac{\chi^2}{4}$ (13). The validity of this approximation may be evaluated from the series expansion for $I_o(\chi)$:

$$I_o(\chi) = 1 + \frac{\chi^2}{4} + \frac{1}{(2!)^2}\left(\frac{\chi^2}{4}\right)^2 + \frac{1}{(3!)^2}\left(\frac{\chi^2}{4}\right)^3 + \cdots$$

$$= 1 + \sum_{i=1}^{\infty} \frac{1}{(i!)^2}\left(\frac{\chi^2}{4}\right)^i .$$

This may be substituted in the series

$$\log(1+y) = y - \frac{y^2}{2} + \frac{y^3}{3} - \frac{y^4}{4} + \cdots$$

which converges for $|y| < 1$. The result is

$$\log I_o(\chi) = \frac{1}{4}\left(\chi^2 - \frac{\chi^4}{16} + \frac{\chi^6}{144} - \frac{11\chi^8}{12288} + \cdots\right), \quad (8.23)$$

a series which converges for $|\chi| < 1.8$.

Equation 8.23 may be used to demonstate that

$$\log I_o\left(\frac{\blacksquare Ac_1}{N}\right) - \log I_o\left(\frac{\blacksquare Ac_o}{N}\right)$$

is very nearly proportional to $\hat{L}_k$ when $\frac{\blacksquare Ac_1}{N}$ and $\frac{\blacksquare Ac_o}{N}$ are low. Both of these statistics are consistently close to zero when the signal strength is weak and it is in this condition that the system is most susceptible to error. The choice of $c_1^2 - c_o^2$ as a likelihood statistic insures, therefore, that the detector's performance approaches the optimum when the propagation conditions are poorest.

8.6 The Conditional Probability Densities

In order to express Equation 8.10 as an explicit function of ρ_o and ρ_1, it is necessary to derive the conditional probability density functions of the likelihood statistics. Although $a_1 - a_o$ and $c_1^2 - c_o^2$ are the data-processing operations specified for a practical detector, it is more convenient mathematically to work with the normalized measurements defined in Equation 5.4. If all of the physical measurements are weighted by the normalizing constant $(NT)^{-\frac{1}{2}}$, the error performance of the detector is unaffected.

8.6.1 Coherent Receiver

The normalized likelihood statistic,

$$\alpha = (NT)^{-\frac{1}{2}} (a_1 - a_o) \tag{8.24}$$

is a normal random variable with unit variance. For the condition "zero" transmitted, the mean value of α is $-\sqrt{\rho_o}$ and when "one" is transmitted, the mean value is $\sqrt{\rho_1}$. The conditional probability densities at a coherent receiver are therefore

$$G_o(\alpha) = \frac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}(\alpha + \sqrt{\rho_o})^2\right]$$

and

$$G_1(\alpha) = \frac{1}{\sqrt{2\pi}} \exp\left[-\tfrac{1}{2}(\alpha - \sqrt{\rho_1})^2\right] . \tag{8.25}$$

The integrals in 8.8 and 8.9 may be expressed in terms of the function $\varphi(x)$ defined as:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-\tfrac{1}{2}t^2)dt \tag{8.26}$$

which is related to erfc(x), the complementary error function (1) by

$$\varphi(x) = \tfrac{1}{2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) . \tag{8.27}$$

The integrals of G_o and G_1 are

$$\int_{-\infty}^{\tau} G_o(u)\, du = \varphi(-\tau-\sqrt{\rho_o}\,)$$

and

$$\int_{\tau}^{\infty} G_1(u)\, du = \varphi(\tau-\sqrt{\rho_1}) \tag{8.28}$$

so that the general error-rate formula for a coherent receiver (except when there is fast fading) is

$$C(\rho_o,\rho_1) = 1- \frac{m}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp[-\tfrac{1}{2}(\tau-\sqrt{\rho_1})^2] \times$$

$$\varphi^s(-\tau-\sqrt{\rho_o})\, \varphi^{m-1}(\tau-\sqrt{\rho_1})d\tau \tag{8.29}$$

in which $\varphi^i(x)$ denotes $[\varphi(x)]^i$.

8.6.2 <u>Non-coherent Reception</u>

In this case, the normalized likelihood statistic is

$$\gamma = \frac{1}{NT}(c_1^2 - c_o^2). \tag{8.30}$$

The conditional probability densities of the normalized measurements γ_o and γ_1 are the Rice probability densities listed in Table 5.2. The probability density functions of $\mu_o = \gamma_o^2$ and $\mu_1 = \gamma_1^2$ may be derived from the formulas of Table 5.2 by means of the general rule that $q(y)$, the probability density of $y = x^2$, is related to $p(x)$, the probability density of x, through

$$q(y) = \frac{1}{2\sqrt{y}} \, p(\sqrt{y}) \quad . \tag{8.31}$$

The probability density of μ_1, when "one" is transmitted is, therefore,

$$\exp(-\mu_1-\rho_1) I_o(2\sqrt{\mu_1\rho_1}) \quad \text{for } \mu_1 \geq 0 \tag{8.32}$$

and that of μ_o is

$$\exp(-\mu_o) \quad \text{for } \mu_o \geq 0 \tag{8.33}$$

so that the joint conditional probability density function of these independent random variables is

$$h_1(\mu_o,\mu_1) = \exp(-\mu_o-\mu_1-\rho_1) I_o(2\sqrt{\mu_1\rho_1})$$

$$\text{for both } \mu_o \text{ and } \mu_1 \geq 0$$

$$h_1(\mu_o,\mu_1) = 0 \quad \text{for either } \mu_o \text{ or } \mu_1 < 0 \tag{8.34}$$

The conditional probability density of $\gamma = \mu_1-\mu_o$ is the integral of $h_1(\mu_o,\mu_o+\gamma)$ over all possible values of μ_o. The function h_1 has value zero except when both its arguments are non-negative. Thus its integral is

$$H_1(\gamma) = \int_{\Lambda}^{\infty} \exp(-2\mu_o-\gamma-\rho_1) I_o[2\sqrt{\rho_1(\mu_o+\gamma)}]\, d\mu_o \tag{8.35}$$

in which Λ is the lowest value of μ_o for which μ_o and $\gamma+\mu_o$ are both non-negative. For $\gamma \geq 0$, $\Lambda = 0$ and for $\gamma < 0$, $\Lambda = -\gamma$. Equation 8.35 may be developed by a change of the variable of integration to $t = 2\sqrt{\mu_o+\gamma}$ and the expression of the integral in the form:

$$H_1(\gamma) = \tfrac{1}{2}\exp(\gamma-\tfrac{1}{2}\rho_1)\int_{2\sqrt{\Lambda+\gamma}}^{\infty} t \exp[-\tfrac{1}{2}(t^2+\rho_1)] I_o(t\sqrt{\rho_1})\, dt$$

$$= \tfrac{1}{2}\exp(\gamma-\tfrac{1}{2}\rho_1)\; Q(\sqrt{\rho_1},\, 2\sqrt{\Lambda+\gamma}) \; . \tag{8.36}$$

The Q function is related to the cumulative probability function of the Rice distribution and is defined by

$$Q(x,y) = \int_{y}^{\infty} t \exp[-\tfrac{1}{2}(t^2+x^2)]\; I_o(tx)dt \; . \tag{8.37}$$

It has been published in tabular form (31). Substitution of the appropriate values of Λ in 8.36 and application of the identity $Q(x,0) = 1$ yields

$$H_1(\gamma) = \tfrac{1}{2} \exp(-\tfrac{1}{2}\rho_1) \exp(\gamma)\; Q(\sqrt{\rho_1},\, 2\sqrt{\gamma}) \quad \text{for } \gamma \geq 0$$

$$H_1(\gamma) = \tfrac{1}{2} \exp(-\tfrac{1}{2}\rho_1) \exp(\gamma) \quad \text{for } \gamma < 0. \tag{8.38}$$

A similar derivation results in the probability density of γ under the condition "zero" transmitted,

$$H_o(\gamma) = \tfrac{1}{2}\exp(-\tfrac{1}{2}\rho_o)\exp(-\gamma) \qquad \text{for } \gamma \geq 0$$

$$H_o(\gamma) = \tfrac{1}{2}\exp(-\tfrac{1}{2}\rho_o)\exp(-\gamma)Q(\sqrt{\rho_o},\ 2\sqrt{-\gamma}) \qquad \text{for } \gamma \leq 0. \tag{8.39}$$

Figure 8.1 shows $H_1(\gamma)$ as a function of γ for three values of the parameter ρ_1. When $\rho_o = \rho_1$, $H_o(\gamma)$ is symmetrical about $\gamma = 0$ to $H_1(\gamma)$.

In order to consolidate the expressions to follow, the function $\theta(x,y)$ may be defined as

$$\theta(x,y) = Q(\sqrt{2x},\ \sqrt{2y}) - \tfrac{1}{2}\exp(-\tfrac{1}{2}x+y)Q(\sqrt{x},\ 2\sqrt{y}) \tag{8.40}$$

The integrals of $H_o(\gamma)$ and $H_1(\gamma)$ are

$$\int_\tau^\infty H_1(\gamma)d\gamma = \theta(\rho_1,\tau) \qquad \text{for } \tau \geq 0$$

$$= 1-\tfrac{1}{2}\exp(-\tfrac{1}{2}\rho_1+\tau) \qquad \text{for } \tau < 0$$

and

$$\int_{-\infty}^\tau H_o(\gamma)d\gamma = 1-\exp(-\tfrac{1}{2}\rho_o-\tau) \qquad \text{for } \tau \geq 0$$

$$= \theta(\rho_o,\ -\tau) \qquad \text{for } \tau < 0. \tag{8.41}$$

Substitution of expressions 8.38 and 8.41 into Equation 8.10 yields the general probability-of-error formula for non-coherent reception

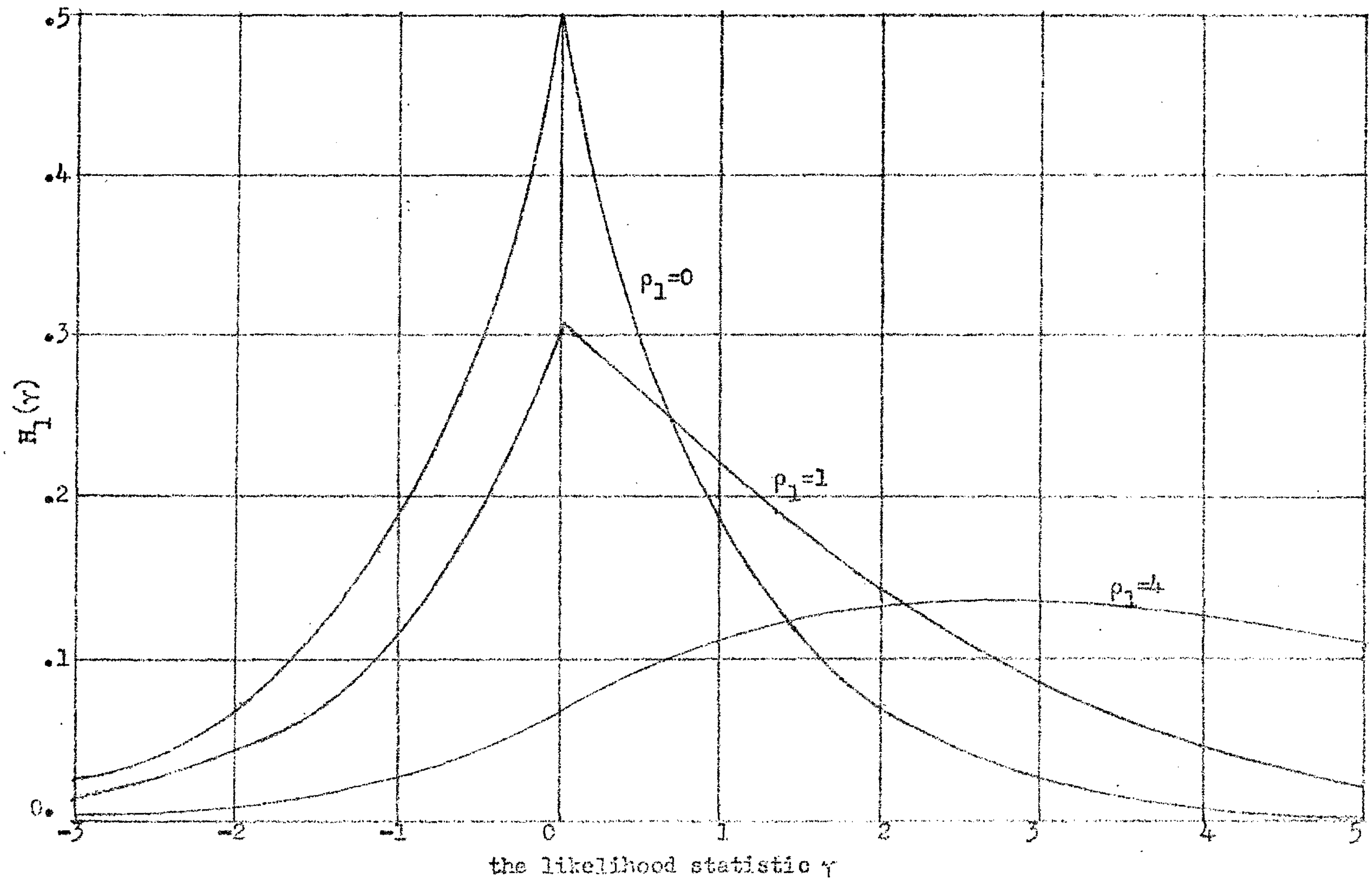


FIGURE 8.1 Conditional probability density of the likelihood statistic at a non-coherent receiver

$$C(\rho_o,\rho_1) = 1-m\int_{-\infty}^{0} \tfrac{1}{2}\exp(-\tfrac{1}{2}\rho_1+\tau)[1-\tfrac{1}{2}\exp(-\tfrac{1}{2}\rho_1+\tau)]^{m-1}\theta^{s}(\rho_o,-\tau)d\tau$$

$$-m\int_{0}^{\infty} \tfrac{1}{2}\exp(-\tfrac{1}{2}\rho_1+\tau)Q(\sqrt{\rho_1},2\sqrt{\tau})[1-\tfrac{1}{2}\exp(-\tfrac{1}{2}\rho_o-\tau)]^{s}\times$$

$$\times\ \theta^{m-1}(\rho_1,\tau)d\tau \qquad (8.42)$$

Equation 8.42 may be written as the single integral:

$$C(\rho_o,\rho_1) = 1-\tfrac{1}{2}m\ \exp(-\tfrac{1}{2}\rho_1)\int_{0}^{\infty} \{\exp(-\tau)[1-\tfrac{1}{2}\exp(-\tfrac{1}{2}\rho_1-\tau)]^{m-1}\ \times$$

$$\times\ \theta^{s}(\rho_o,\tau)\ +\ \exp(\tau)\ Q(\sqrt{\rho_1},2\sqrt{\tau})\ \times$$

$$\times\ [1-\exp(-\tfrac{1}{2}\rho_o-\tau)]^{s}\ \theta^{m-1}(\rho_1,\tau)\}\ d\tau\ . \qquad (8.43)$$

8.6.3 Fast Fading

Equation 8.19 is the basic error-rate formula for the fast-fading condition. The terms of the formula are integrals of the probability density functions of the likelihood statistics. At a coherent receiver the terms are

$$\overline{G}_1(\tau) = (1-\beta^2)\left\{\frac{1}{\sqrt{2\pi}}\exp(-\frac{\tau^2}{2})+\tau\beta\varphi(-\tau\beta)\exp[-\tfrac{1}{2}\tau^2(1-\beta^2)]\right\},$$

$$\int_{-\infty}^{\tau}\overline{G}_o(\alpha)d\alpha = \varphi(-\tau)+\beta\varphi(\tau\beta)\exp[-\tfrac{1}{2}\tau^2(1-\beta^2)] =$$

$$= q_1(-\tau)$$

and

$$\int_{\tau}^{\infty}\overline{G_1}(\alpha)d\alpha = \varphi(\tau)+\beta\varphi(-\tau\beta)\exp[-\tfrac{1}{2}\tau^2(1-\beta^2)] =$$

$$= q_1(\tau) \qquad (8.44)$$

in which the notation $q_1(\tau)$ is introduced for the purpose of abbreviation and the average signal-to-noise ratio ρ is expressed in terms of the parameter

$$\beta = \frac{\sqrt{\rho}}{\sqrt{2+\rho}} \tag{8.45}$$

In the case of a non-coherent receiver all of the terms of 8.19 are exponential functions:

$$\begin{aligned}
\overline{H}_1(\tau) &= \frac{1}{2+\rho} \exp\left[-\frac{\tau}{1+\rho}\right] && \text{for } \gamma \geq 0 \\
&= \frac{1}{2+\rho} \exp(\tau) && \text{for } \gamma < 0 \\
\int_{-\infty}^{\tau} \overline{H}_o(\gamma)d\gamma &= 1 - \frac{1}{2+\rho} \exp(-\tau) && \text{for } \tau \geq 0 \\
&= \frac{1+\rho}{2+\rho} \exp\left[\frac{\tau}{1+\rho}\right] && \text{for } \tau < 0 \\
\int_{\tau}^{\infty} \overline{H}_1(\gamma)d\gamma &= \frac{1+\rho}{2+\rho} \exp\left[-\frac{\tau}{1+\rho}\right] && \text{for } \tau \geq 0 \\
&= 1 - \frac{1}{2+\rho} \exp(\tau) && \text{for } \tau < 0
\end{aligned} \tag{8.46}$$

8.7 The Error-Rate Formulas

The probability of error associated with a specific propagation condition is derived by substitution of the probability density functions of Section 8.6 in the appropriate formula in Section 8.4.

8.7.1 Coherent receiver, no fading (Case 1)

$$C_1(\rho) = 1 - \frac{m}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\tfrac{1}{2}(\tau-\sqrt{\rho})^2\right]\varphi^s(-\tau-\sqrt{\rho})\varphi^{m-1}(\tau-\sqrt{\rho})d\tau \tag{8.47}$$

8.7.2 Coherent receiver, slow flat fading (Case 2)

$$C_2(\rho) = 1 - \int_{-\infty}^{\infty} v \exp(-\tfrac{1}{2}v^2)C_1(\tfrac{1}{2}\rho v^2)dv \tag{8.48}$$

8.7.3 Coherent receiver, slow selective fading (Case 3)

To derive the error probability for this condition, the order of integration has been changed from that indicated in 8.17, which implies integration(of Equation 8.29) with respect to τ to find $C(\frac{1}{2}\rho v_o^2, \frac{1}{2}\rho v_1^2)$ followed by a double integration with respect to v_o and v_1. On the other hand, the factors of 8.29 containing ρ_o and ρ_1 may be integrated separately:

$$q_2(\tau) = \int_o^{\infty} v \exp(-\tfrac{1}{2}v^2)\varphi^s(-\tau-v\sqrt{\tfrac{1}{2}\rho})dv$$

$$q_3(\tau) = \int_o^{\infty} v \exp[-\tfrac{1}{2}v^2-\tfrac{1}{2}(\tau-v\sqrt{\tfrac{1}{2}\rho})^2]\varphi^{m-1}(\tau-v\sqrt{\tfrac{1}{2}\rho})dv \tag{8.49}$$

The product of these functions is integrated to produce the probability of error

$$C_3(\rho) = 1 - \frac{m}{\sqrt{2\pi}} \int_{-\infty}^{\infty} q_2(\tau)q_3(\tau)d\tau \quad : \tag{8.50}$$

which is, effectively, a double-integral expression.

8.7.4 Coherent receiver, fast fading (Case 4)

$$C_4(\rho) = 1-4(1-\beta^2)\int_{-\infty}^{\infty} \left\{ \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}\tau^2) + \tau\beta\varphi(-\beta)\exp\left[-\tfrac{1}{2}\tau^2(1-\beta)^2)\right]\right\} q_1{}^{s}(-\tau)q_1{}^{m-1}(\tau)d\tau \tag{8.51}$$

in which $q_1(\tau)$ and β are defined in Equations 8.44 and 8.45, respectively.

8.7.5 Non-coherent receiver, no fading (Case 5)

$$C_5(\rho) = 1-\tfrac{1}{2}m \exp(-\tfrac{1}{2}\rho)\int_0^{\infty} \{\exp(-\tau)[1-\tfrac{1}{2}\exp(-\tfrac{1}{2}\rho-\tau)]^{m-1} \times \theta^{s}(\rho,\tau)+\exp(\tau)Q(\sqrt{\rho}, 2\sqrt{\tau})[1-\exp(-\tfrac{1}{2}\rho-\tau)]^{s} \times \theta^{m-1}(\rho,\tau)\}\, d\tau \tag{8.52}$$

8.7.6 Non-coherent receiver, slow flat fading (Case 6)

$$C_6(\rho) = 1 - \int_0^{\infty} v \exp(-\tfrac{1}{2}v^2)C_5(\tfrac{1}{2}\rho v^2)dv \tag{8.53}$$

8.7.7 Non-coherent receiver, slow selective fading (Case 7)

As in Case 3, the triple integral implied by Equation 8.17 may be reduced to a double integral by means of a change in the order of integration. In this case, it is the factors of 8.43 which are integrated separately. The expressions

$$q_4(\tau) = \exp(-\tau)\int_0^\infty v \; \exp[-\tfrac{1}{2}v^2(1+\tfrac{1}{2}\rho)][1-\tfrac{1}{2}\exp(-\tau-\frac{\rho v^2}{4})]^{m-1}dv$$

and

$$q_5(\tau) = \exp(\tau) \int_0^\infty v \; \exp(-\tfrac{1}{2}v^2)[1-\tfrac{1}{2}\exp(-\tau-\frac{\rho v^2}{4})]^{s} \; dv \qquad (8.54)$$

are exponential functions contained in the double-integral formula:

$$C_7(\rho) = 1-\tfrac{1}{2}m\int_0^\infty\int_0^\infty v \; \exp(-\tfrac{1}{2}v^2)\{q_4(\tau)\theta^s(\tfrac{1}{2}\rho v^2,\tau)$$

$$+q_5(\tau)\exp(-\frac{\rho v^2}{4})Q(v\sqrt{\tfrac{1}{2}\rho},2\sqrt{\tau})\,\theta^{m-1}(\tfrac{1}{2}\rho v^2,\tau)\} \; dvd\tau \; . \qquad (8.55)$$

8.7.8 Non-coherent receiver, fast fading (Case 8)

For any m and s, the error rate may be expressed as an algebraic function of ρ which is derived from the following integral of weighted exponential functions:

$$C_8(\rho) = \frac{m}{2+\rho}\int_0^\infty \left\{\left[\frac{1+\rho}{2+\rho}\right]^s \exp\left[-\tau-\frac{s\tau}{1+\rho}\right]\left[1-\frac{1}{2+\rho}\exp(-\tau)\right]^{m-1}\right.$$

$$\left. + \left[\frac{1+\rho}{2+\rho}\right]^{m-1} \exp\left[-\frac{m\tau}{1+\rho}\right]\left[1-\frac{1}{2+\rho}\exp(-\tau)\right]^s\right\}d\tau \qquad (8.56)$$

TABLE 8.1

ERROR PROBABILITIES OF THE FOUR-OUT-OF-SEVEN CODE

Case 1 : $$C_1(\rho) = 1 - \frac{4}{\sqrt{2\pi}} \int_0^\infty \{\exp[-\tfrac{1}{2}(\tau-\sqrt{\rho})^2] + \exp[-\tfrac{1}{2}(\tau+\sqrt{\rho})^2]\}[\varphi(-\tau-\sqrt{\rho})]^3 d\tau$$

Case 2 : $$C_2(\rho) = \int_0^\infty v \exp(-\tfrac{1}{2}v^2) C_1(\tfrac{1}{2}\rho v^2) dv$$

Case 3 : $$C_3(\rho) = 1 - \frac{4}{\sqrt{2\pi}} \int_0^\infty \{[q_2(\tau)q_3(\tau)]^3 + [q_2(-\tau)q_3(-\tau)]^3\} d\tau$$

Case 4 : $$C_4(\rho) = 1 - (1-\beta^2)\int_0 \left\{ \frac{2}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}\tau^2) + \tau\beta[1-2\varphi(\tau\beta)]\exp[-\tfrac{1}{2}\tau^2(1-\beta^2)]\right\} \times$$
$$\times [q_1(\tau)q_1(-\tau)]^3 \, d\tau$$

Case 5 : $$C_5(\rho) = 1 - 2\exp(-\tfrac{1}{2}\rho)\int_0^\infty [\exp(-\tau) + \exp(\tau)Q(\sqrt{\rho}, 2\sqrt{\tau})][1-\exp(-\tfrac{1}{2}\rho-\tau)]^3 \times$$
$$\times \theta^3(\rho,\tau) d\tau$$

continued ...

TABLE 8.1 (continued)

<u>Case 6</u> : $$C_6(\rho) = \int_0^\infty v \exp(-\tfrac{1}{2}v^2)\, C_5(\tfrac{1}{2}\rho v^2)\, dv$$

<u>Case 7</u> : $$C_7(\rho) = 1 - 2\int_0^\infty\int_0^\infty v \exp\left(-\frac{v^2}{2}\right)\theta^3(\tfrac{1}{2}\rho v^2, \tau)\left\{ \frac{\exp(-\tau)}{1+\frac{1}{2}\rho} - \frac{3\exp(-2\tau)}{2+2\rho} + \frac{3\exp(-3\tau)}{4+6\rho} - \frac{\exp(-4\tau)}{8+16\rho} + \left[1 - \frac{3\exp(-\tau)}{2+\rho} + \frac{3\exp(-2\tau)}{4+4\rho} - \frac{\exp(-3\tau)}{8+12\rho}\right] \times \exp\left[\tau - \frac{\rho v^2}{4}\right] Q(v\sqrt{\tfrac{1}{2}\rho}, 2\sqrt{\tau})\right\} dv\,d\tau$$

<u>Case 8</u> : $$C_8(\rho) = 1 - 4\left[\frac{1+\rho}{2+\rho}\right]^4 \left\{ \frac{1}{4} + \frac{1}{4+\rho} - \frac{3}{2+\rho}\left[\frac{1}{5+\rho} + \frac{1}{5+2\rho}\right] + \frac{3}{(2+\rho)^2}\left[\frac{1}{6+2\rho} + \frac{1}{6+3\rho}\right] - \frac{1}{(2+\rho)^3}\left[\frac{1}{7+3\rho} + \frac{1}{7+4\rho}\right]\right\}$$

CHAPTER 9: CONCLUSIONS

Three approaches to the telegraph-detection problem are reported in the earlier chapters of this thesis: 1) a computer simulation of four detectors, 2) a presentation and assessment of binary-detector and maximum-likelihood-detector performance characteristics and 3) a study of the feasibility of developing a variable-speed feedback mechanism. Each of these investigations has resulted in preliminary conclusions and each requires further attention if its practical significance is to be fully understood.

The simulated telegraph system has been devised in order to demonstrate the configuration and performance of the minimum-risk detector and to compare it with other telegraph detectors. The cost matrix of the minimum-risk detector takes into account the serious effects of control-symbol errors on the legibility of detected messages. Although the minimum-risk output messages of Appendix B appear equally legible as or more legible than the corresponding outputs of other detectors, the initial conclusion of the simulation is that the computational complexity of a minimum-risk detector precludes its adoption in most practical situations.

A formal verification of this conclusion requires a consideration of specific practical circumstances and the performance of a series of subjective tests designed to evaluate quantitatively the relative merits of the detection methods. Such tests could possibly establish the desirability of a minimum-risk detector in certain applications. They could also lead to modifications of the [redacted] cost matrix of Section 5.1. The simulation method of Chapter 5 is thus presented as an instrument to be used in further studies of the results obtainable from the formulated detectors and in other investigations where an examination of possible detection results is sought prior to the physical construction of an entire system.

The most practical proposals of the thesis stem from the maximum-likelihood error-rate formulas derived in Chapter 8. The performance characteristics of Chapter 6 have been calculated from these formulas and they indicate the advantages offered by maximum-likelihood detectors of symbols represented by the seven-unit code. Where the five-unit code is used, conversion to the seven-unit format and the introduction of maximum-likelihood detectors would enhance the quality of detected messages and extend the range of signal-to-noise ratios which admit reliable communication. The effects of signal fading on detectability are influenced by the organization of message-multiplex schemes. When maximum-likelihood detection is adopted, the most harmful effects are caused by completely non-selective fading. The greatest immunity to this type of fading is achieved when the 14 time-frequency regions associated with each message in a multiplex system are as widely dispersed as possible in both time and frequency.

Because the propagation conditions encountered by physical systems often differ significantly from any which have been formulated mathematically, the performance improvements indicated in Chapter 6 must be verified by experiments with equipment operating in realistic conditions. In an experimental program currently in progress, a pair of matched filters and a data-logging system have been developed to perform the measurement role of a telegraph detector. The computational role is performed by a digital computer so that any one of several decision rules may be readily implemented. The aim of the experiments is to compare the performances of binary and maximum-likelihood detectors operating on physical signals and thereby to test the validity of the theoretical characteristics of Chapter 6.

The investigation in Chapter 7 of two-way feedback systems points to the desirability of developing a system in which the signal-element time duration is adapted to changing channel conditions. This system would incorporate maximum-likelihood detectors and maintain reliable communication over a considerably wider range of signal-to-noise ratios than ARQ systems. Preliminary design considerations for a variable-duration system are presented in Chapter 7 but further study is required to overcome practical obstacles and to demonstrate the validity of a detailed design proposal.

"Finally. It was stated at the outset, that this system would not be here, and at once, perfected. You cannot but plainly see that I have kept my word. But now I leave my ... system standing thus unfinished, even as the great Cathedral of Cologne was left, with the crane still standing upon the top of the uncompleted tower. For small erections may be finished by their first architects; grand ones, true ones, ever leave the copestone to posterity. Heaven keep me from ever completing anything. This whole book is but draught - nay, but the draught of a draught. Oh, Time, Strength, Cash, and Patience!"

from "Moby Dick" by Melville

Bibliography

1. Abramowitz, M. and I.A. Stegun, ed., "Handbook of Mathematical Functions", National Bureau of Standards, Washington, D.C., 1964.

2. Alekseyev, A.I., "Optimum Noise Immunity of Noise Like Signals", Telecommunications and Radio Engineering, Part 2, vol.20, pp. 79-83, Aug., 1965.

3. Allnatt, J.W., et al., "Frequency Diversity in Reception of Selectively Fading Binary Frequency-Modulated Signals", Proc. IEE, vol.104B, pp 98-110, March, 1957.

4. Baghdady, E.J., ed., "Lectures on Communication System Theory", McGraw Hill, New York, 1961.

5. Baghdady, E.J., "Diversity Techniques", in ref.4, pp. 125-175.

6. Barrow, B.B., "Error Probabilities for Data Transmission Over Fading Radio Paths", SHAPE Air Defense Centre, The Hague, Feb., 1962.

7. Benice, R.J. and A.H. Frey, "An Analysis of Retransmission Systems", IEEE Transactions on Communication Technology, vol. COM-12, pp. 135-145, Dec., 1964.

8. Benice, R.J. and A.H. Frey, "Comparisons of Error Control Techniques", IEEE Transactions on Communication Technology, vol. COM-12, pp. 146-154, Dec., 1964.

9. Blackwell, D. and M.A. Girshick, "Theory of Games and Statistical Decisions", Wiley, New York, 1954.

10. Bennett, W.R. and S.O. Rice, "Spectral Density and Autocorrelation Functions Associated with Binary Frequency-Shift Keying", BSTJ, vol.XLII, pp. 2355-2385, Sept., 1963.

11. Brennan, D.G. and M.L. Phillips, "Phase and Amplitude Variability in Medium Frequency Transmission", Lincoln Laboratory Technical Report No.93, Sept., 1957.

12. Brennan, D.G., "Linear Diversity Combining Techniques", Proc. IRE, vol. 47, pp. 1075-1102, June, 1959.

13. Bussgang, J.J. and W.L. Mudgett, "A Note of Caution on the Square-Law Approximation to an Optimum Detector", IRE Transactions on Information Theory, vol. IT-6, pp. 504-505, Sept., 1960.

14. Cherry, C., "On Human Communication", Chapter 5, MIT Press, Cambridge Mass., 1957.

15. Croisdale, A.C., "Automatic Error Correcting Systems Used on H.F. Radio Links", ref.24, pp. 187-210.

16. Croisdale, A.C., "Teleprinting Over Long Distance Radio Links", Post Office Electrical Engineers' Journal, vol.51, pp. 88-93, July, 1958 and pp. 219-225, Oct., 1958.

17. Doeltz, M.L., "Predicted Wave Radio Teleprinter", Electronics, vol.27, pp. 166-169, Dec., 1954.

18. Egan, R.D. and D.S. Pratt, "Oblique Sounding and HF Communications", ref. 24, pp. 47-51.

19. Freebody, J.W., "Telegraphy", Pitman, London, 1959.

20. Frith, K. and P.N. Clarkson, "A Transistorized Regenerative Repeater", ref. 24, pp. 346-349.

21. Green, P.E., "Feedback Communication Systems", in ref. 4, pp. 345-368.

22. Grisdale, G.L., et al., "Fading of Long-Distance Radio Signals and Comparison of Space- and Polarization-Diversity Reception in the 6-18 Mc/s Range", Proc. IEE, vol. 104B, pp. 39-51, Jan., 1957.

23. Helstrom, C.W., "The Resolution of Signals in White Gaussian Noise", Proc. IRE, vol.43, pp. 1111-1118, Sept., 1955.

24. Institution of Electrical Engineers, "Convention on HF Communication", London, March 1963.

25. Joint Technical Advisory Committee, "Radio Spectrum Utilization", IEEE, New York, 1957.

26. Kahn, L.R., "Ratio Squarer", Proc. IRE, vol.42, pg. 1704, Nov., 1954.

27. Kotel'nikov, V.A., "The Theory of Optimum Noise Immunity", McGraw Hill, New York, 1959.

28. Law, H.B., "The Detectability of Fading Radio-telegraph Signals in Noise", Proc. IEE, vol.104B, pp. 130-140, March, 1957.

29. Machol, R.E., ed., "Information and Decision Processes", McGraw Hill, New York, 1960.

30. Machol, R.E. and P. Gray, ed., "Recent Developments in Information and Decision Processes", Macmillan, 1962.

31. Marcum, J.I., "Table of Q functions", Rand Corporation Report RM 339, Jan., 1950

31a. Mason, W.C. et al., "Tables of the Distribution Functions of Sums of Rayleigh Variables", pre-publication copies referred to in ref. 6 and 12. No record of the publication of this paper has been found by the present author.

32. McNicol, R.W.E., "The Fading of Radio Waves of Medium and High Frequencies", Proc. IEE, vol.96, part III, pp. 517-524, Dec., 1949.

33. Middleton, D., and D. Van Meter, "Detection and Extraction of Signals in Noise From the Point of View of Statistical Decision Theory", Journal of the Society of Industrial Applied Mathematics, vol.3, pp. 192-253, Dec., 1955 and vol.4, pp. 86-119, June, 1956.

34. Middleton, D. and D. Van Meter, "On Optimum Multiple-Alternative Detection of Signals in Noise", IRE Transactions on Information Theory, vol. IT-1, pp. 1-9, Sept., 1955.

35. Mitra, S.K., "The Upper Atmosphere", (2nd edition), The Asiatic Society, Calcutta, 1952.

36. Parry, C.A., "The Present Status and Future Utilization of Advanced Technology In High Frequency Systems Engineering", (Abstract only), IEEE International Convention Record, vol.14, part 2, pg. 50.

37. Raiffa, H. and R. Schlafer, "Applied Statistical Decision Theory", Division of Research, Graduate School of Business Administration, Harvard U., Boston, Mass., 1961

38. Ratcliffe, J.A., "Diffraction From the Ionosphere and the Fading of Radio Signals", Nature, vol.162, pp. 9-11, 3 July, 1948.

39. Ratcliffe, J.A., "Physics of the Upper Atmosphere", Academic Press, London, 1960.

40. Rice, S.O., "Mathematical Analysis of Random Noise", BSTJ, vol.23, pp. 282-332, July, 1944 and vol. 24, pp. 46-156, Jan. 1945.

41. Ridout, P.M., and L.K. Wheeler, "The Choice of Multichannel Telegraph Systems For Use on HF Radio Links", Proc. IEE, vol. 110, pp. 1402-1410, Aug., 1963.

42. Robin, H.K., et al., "Multitone Signalling System Employing Quenched Resonators For Use on Noisy Radio-Teleprinter Circuits", Proc. IEE, vol. 110, pp. 1554-1568, Sept., 1963.

43. Root, W.L., "On the Measurement and Use of Time-Varying Communication Channels", Information and Control, vol. 8, pp. 390-422, Aug., 1965.

44. Rosie, A.M., "Correlation Recognition of a Redundant 5-Unit Code", Proc. IEE, vol.110, pp.493-500, March, 1963.

45. Saitskii, Y.I. and V.M. Timofeev, "Analysis of Five-Element Codes for Teleprinter Equipments", Electrosvyaz', no.7, pp. 57-62, 1957.

46. Schlaiffer, R., "Probability and Statistics for Business Decisions", McGraw Hill, New York, 1959.

47. Schwartz, L.S., "Feedback For Error Control and Two-Way Communication", IEEE Transactions on Communications Systems, vol. CS-11, pp. 49-56, March 1963.

48. Shannon, C.E. and W. Weaver, "The Mathematical Theory of Communication", University of Illinois Press, Urbana, Illinois, 1949.

49. Shannon, C.E., "Communication In the Presence of Noise", Proc. IRE, vol.37, pp. 10-21, Jan., 1949.

50. Shannon, C.E., "Coding Theorems for a Discrete Source with a Fedility Criterion", in ref.29, pp. 93-126.

51. Silverman, R.A. and M. Balser, "Coding for Constant Data Rate Systems", Proc. IRE, vol. 42, pp. 1428-1435, Sept., 1954 and vol.43, pp. 728-733, June, 1955.

52. Stuart, R.D., "An Insert System for Use with Feedback Communication Links", IEEE Translactions on Communications Systems, vol. CS-11, pp. 142-143, March, 1963.

53. Swets, J.A., ed., "Signal Detection and Recognition By Human Operators", Wiley, New York, 1964.

54. Tribus, M., "The Use of the Maximum Entropy Estimate in the Estimation of Reliability", in ref. 30, pp. 102-140.

55. Turin, G.L., "Error Probabilities for Binary Symmetric Ideal Reception Through Non-Selective Slow Fading and Noise", Proc. IRE, vol.46, pp. 1603-1618, Sept., 1958.

56. Turin, G.L., "An Introduction to Matched Filters", IRE Transactions on Information Theory, vol.IT-6, pp. 311-329, June 1960.

57. Van Duuren, H.C.A., "Error Probability and Transmission Speed on Circuits Using Error Detection and Automatic Repetition of Signals", IRE Transactions on Communications Systems", vol. CS-9, pp. 38-50, March, 1961.

58. Van Meter, D. and D. Middleton, "Modern Statistical Approaches to Reception In Communication Theory", Transactions of the Professional Group on Information Theory, vol. PGIT-4, pp. 119-145, Sept., 1954.

59. Van Vleck, J.H. and D. Middleton, "A Theoretical Comparison of the Visual Aural and Motor Reception of Pulsed Signals in the Presence of Noise", Journal of Applied Physics, vol.17, pp. 940-971, Nov., 1946.

60. Vasil'ev, P.V., "Optimum Length of Code Combination For the Feedback Signals in a Data Transmission System Using a Reverse Channel", Telecommunications and Radio Engineering, Part 1, vol.19, pp. 36-41, 1965.

61. Viterbi, A.J., "On Coded Phase Coherent Communications", IEEE Transactions on Space Electronics and Telemetry, vol. SET-7, pp. 3-14, March, 1961.

62. Viterbi, A.J., "The Effect of Sequential Decision Feedback on Communication Over the Gaussian Channel", Information and Control, vol.8, pp.80-92, Feb., 1965.

63. Viterbi, A.J., "Optimum Detection and Signal Selection for Partially Coherent Binary Communication", IEEE Transactions on Information Theory, vol.IT-11, pp. 239-246, April, 1965.

64. Voelcker, H.B., "Simple Codes for Fading Circuits", IRE Transactions on Communications Systems, vol. CS-6, pp. 47-52, Dec. 1958.

65. Voelcker, B.H., "Phase Shift Keying in Fading Channels", Proc. IEE, vol.107B, pp. 31-38, Jan., 1960.

66. Wald, A., "Contributions to the Theory of Statistical Estimation and Testing Hypotheses", Annals of Mathematical Statistics, vol.10, pp. 299-326, Dece., 1939.

67. Wald, A., "On the Principles of Statistical Inference", Notre Dame Mathematical Lectures, no.1, 1942.

68. Wald, A., "Sequential Analysis", Wiley, New York, 1947.

69. Wald, A., "Statistical Decision Functions", McGraw Hill, New York, 1950.

70. Weiss, L., "Statistical Decision Theory", McGraw Hill, New York, 1961.

71. Wilks, S.S., "Mathematical Statistics", Wiley, New York, 1962.

72. Woodward, P.M. and I.L. Davies, "Information Theory and Inverse Probability in Telecommunication", Proc. IEE, vol.99, part III, pp. 37-44, March, 1952.

73. Woodward, P.M., "Probability and Information Theory with Applications to Radar", Pergamon, London, 1953.

74. Wozencraft, J.M. and I.M. Jacobs, "Principles of Communication Engineering", Wiley, New York, 1965.

75. Zelen, M., and N.C. Severo, "Probability Functions", in ref.1, pp. 925-995.

Appendix A

The Cost Matrix of Section 5.1

1. X_i (transmitted symbol) and Y_j (received symbol) are identical.

$$c_{ii} = 0 .$$

2. X_i and Y_j are both printed characters ($i \neq j$).

$$c_{ij} = 1 .$$

3. X_i is printed character.

   a. Y_j is carriage return.

The beginning of a line is overprinted with the symbols following the inserted carriage return. If the insertion may occur at all positions on the line with equal probability, half of the line will, on the average, be illegible. Thus

$$c_{ij} = \tfrac{1}{2}k_1 .$$

   b. Y_j is line feed.

Although printing continues on a new line only the transmitted character is lost so that

$$c_{ij} = 1 .$$

c. Y_j is letter shift.

If teleprinter is in the letters mode when this error occurs, only the transmitted symbol is lost. If teleprinter is in figures mode, half (on the average) of the symbols in the figure shift block are lost. The probability that the teleprinter is in the letters mode is

$$\frac{k_2}{k_2+k_3} .$$

The probability it is in the figures mode is

$$\frac{k_3}{k_2+k_3}$$

and therefore,

$$c_{ij} = \frac{k_2}{k_2+k_3} + \tfrac{1}{2}k_3 \frac{k_3}{k_2+k_3}$$

d. Y_j is figure shift.

$$c_{ij} = \frac{k_3}{k_2+k_3} + \tfrac{1}{2}k_2 \frac{k_2}{k_2+k_3}$$

The reasoning is the same as that of 3c.

4. X_i is carriage return.

a. Y_j is printed character.

Printing continues in the line's margin until the margin is exhausted, after which characters overprint the last position on the line. The cost is

$$c_{ij} = k_1 - k_4 \ .$$

b. Y_j is line feed.

The effect is the same as that described in 4a and

$$c_{ij} = k_1 - k_4 .$$

c. Y_j is letter shift.

Added to the cost described in 4a is the cost given in 3c of printing figure-shift symbols in the wrong mode. Thus

$$c_{ij} = k_1 - k_4 + \tfrac{1}{2}k_3 \frac{k_3}{k_2 + k_3}$$

d. Y_j is figure shift.

As in 4c,

$$c_{ij} = k_1 - k_4 + \tfrac{1}{2}k_2 \frac{k_2}{k_2 + k_3} \ .$$

5. X_i is line feed.

a. Y_j is printed character.

It is likely that the carriage is at the left edge of the page and that the preceding line will be overprinted by the line of print following the line feed. In this event both lines are lost to a reader and

$$c_{ij} = 2k_1.$$

b. Y_j is carriage return.

The overprinting described in 5a occurs and

$$c_{ij} = 2k_1.$$

c. Y_j is letter shift.

The overprinting described in 5a occurs. If a figure-shift block is printed in the wrong mode, the loss described in 3c may be greater than $2k_1$ characters so that

$$c_{ij} = \mathrm{Max}(2k_1,\ \tfrac{1}{2}k_3\ \frac{k_3}{k_2+k_3}).$$

d. Y_j is figure shift.

The effect is similar to that of 5c. Thus,

$$c_{ij} = \mathrm{Max}(2k_1,\ \tfrac{1}{2}k_2\ \frac{k_2}{k_2+k_3}).$$

6. X_i is letter shift.

a. Y_j is printed character.

The letter-shift block will be printed in the wrong mode so that

$$c_{ij} = k_2$$

b. Y_j is carriage return.

The effects described in 5a and 3a both occur and

$$c_{ij} = \text{Max}(k_2, \tfrac{1}{2}k_1).$$

c. Y_j is line feed.

The letter-shift block is lost. Thus,

$$c_{ij} = k_2$$

d. Y_j is figure shift.

Here too, the letter-shift block is lost and

$$c_{ij} = k_2$$

7. X_i is figure shift. The effects are similar to those of 6.

a. Y_j is printed character.

$$c_{ij} = k_3$$

b. Y_j is carriage return.

$$c_{ij} = \text{Max.}(k_3, \tfrac{1}{2}k_1)$$

c. Y_j is line feed.

$$c_{ij} = k_3$$

d. Y_j is letter shift.

$$c_{ij} = k_3 .$$

APPENDIX B: EXAMPLE OF SIMULATION RESULTS

The following pages demonstrate the results of the simulated transmission of an English text through the telegraph system formulated in Chapter 5. The original text, consisting of 1069 teleprinter symbols, is presented on this page and is followed by the outputs of the four detectors operating at signal-to-noise ratios of 2.0, 3.0, 4.0, 5.0 and 6.0. The assumed propagation conditions are constant channel attenuation and coherent reception.

TEXT OF THE TRANSMITTED MESSAGE

TELEGRAPHIC COMMUNICATION CAN BE ESTABLISHED USING VARIOUS CODES. BECAUSE IT IS MOST ECONOMICAL, THE MOST WIDESPREAD CODE IN TELEPRINTER USE IS THE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DEVELOPMENT OF TELEPRINTER TECHNIQUE VARIOUS TOPICS WERE TAKEN INTO CONSIDERATION IN WORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WAS CONSTRUCTED SO THAT VOWELS, BEING THE MOST FREQUENTLY ENCOUNTERED, COULD BE SENT BY THE RIGHT HAND ONLY. THE BASIC CRITERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS OF THE MATTER TO BE SENT, SO THAT THE MOST FREQUENTLY OCCURRING LETTERS WERE REPRESENTED BY COMBINATIONS WITH THE LEAST NUMBER OF CHANGES FROM ONE POLARITY OF CURRENT TO THE OTHER. THIS DEVICE ENABLED A REDUCTION IN THE EFFECTS OF CHARACTERISTIC DISTORTIONS. CODED COMBINATIONS MOST SUBJECT TO THIS DISTORTION WERE ALLOCATED TO THE LEAST FREQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CITED THERE ARE OTHER CONSIDERATIONS IN DEVELOPING A CODE.

BINARY DETECTOR ρ=2.0

6DOODODOSHODVODOSDOOODOOSOSOODODODOOOYOSBOODODSOODSOOODOOSOODOSROOOO
•66'863465θ9,θ6θθ294θ6θ 9θθ θθ3θθKθ3-ELθθθNTθθOθE.θθ
θθθθ ?θθθ95 θ9•θ
θθθ2θθθCθ6θθθ10θθHθθE•WθCθθθθEθθFSθθθEθθEθθθθTθθθθθANθθθθLθXM θθE θ
θTHθ SθθMEθθ
θθθθθθAθTθθ θO θEθθθNθθNθCSO θHAθ TθθθMθST θRθQUθNTLY OCCUθRθθ
θCθθNGθθθθθθθMRθθθθθθθθYθθOθθθNθθIONSθWIθθθHE LEAθθ NUMθERθOθθ
TY OF θURθθθTθTOθθHE θTHERθ ITθIθθDθVθCEθθθNθθ
θ

θ?θ'θ•θ'θθθ3 θθ85θ48θθθθ5θθθθθθθ3θ-4θ θ5θθθθ:θ,θ8θ3θ3θ89,'θθ,
•θθ3)θθθθ• - θ9•3.

MAXIMUM-LIKELIHOOD DETECTOR ρ=2.0

TθEEθθθθNθθ θθRθθθθCθθθNθθθNLWθTθθYKθST WIDZSHREAD
CODE IN
TELEPRIYTER USE HS YHE FLV -13) .3,5 ?8,•4/. AT SEVEGAL STAGES
BEYEEθPWEθθNθθUTELS N THE
KARIOUS TOVICU WERESTAKECGINTOC
DONSIDERBTION INRWORKING OUT THERFIME-ELEMENT NOBE.-
5θ3 ?-7,95
9•3 2θ
θ
θ43173,5)6 3,:97,574-•, COUUD BE SSNT BB THE CIGHTCHAND ONLYXM THE
BASRC CRNTERION HSRE WAS EAKE OF REMEMAERING THE CODEθ(

THE SIEMENS Cθ
F5θBZ.θABTθ459772536,9::7]θ7;θ-
'35-
T GS KUI
SREPRESENTEDGAY COMBINATIONS WITO YHE LEAST NUMBER OF
CHANGZS MRPM ONU POLAN
TY OF CURR NT TO THE OTHER. ITHIS DEVICEGBENBEθ
CODDDSCOMBINAQIONS MOST SUEWAPT VO TUIS DISTORWION WEREPALLO ATED TO
THE LEKSTGFLZKUENTLY-USED LETTERS.

9?34853480 :
53• 5θ343 -43 :5θ3 :9,38•343θ89,'θ8,
•3(3)908.• - 99•3.

MINIMUM-ERROR DETECTOR ρ=2.0

TθEEθθθθHθθ θθRθθθθCθθθθθθBN,BθHESMθST WIDISHREADICODE IN
DθθθθθθθθθθθOθθEθθθθRRθNEEθUθEθθθθθθθθθθθθθθθθTθθθθθθθθθθESTAKEθGINTθHE
OTHE BAUNOT NOθ
θ
FREQUENTLY ENCOUNTERAD, COUUD BE SSNT BB THE CIGHTCHAND ONLYAM THE
BASRC CRNTERION HSRE WAS EASE OF REMEMAERING THE CODE80

THE SIEMENS Cθ
F5θθE.θABTθ459772536,9::7]θ75θ-
T GS BUIESREPRESENTEDGAY COMBINATIONS WITO THE LEAST NUMBER OF
CHANGIS MRPM ONE POLAN
ENBBLED A REDUCTIHN INTYHEFEEθθθTθTOθOCθθBAOTθθθSTθθHθSSθθθTθθNS.
TODDDSCOMBINAQIONS MOST SUEWAPT VO TUIS DISTORWION WEREPALLO ATED TO
THE LEASTGFLEKUENTLY-USED LETTERS.

9?34853480 :353• 5θ343 -43 55θ3 :9,38•343589,'θ8,
•3(3)9-8.θ - 99•3.

MINIMUM-RISK DETECTOR ρ=2.0

TBLEQFOBBUC CORZBUSCOCDOSCUN BE ISYB
BECAUE IT
S
MOST EONTICAL,3THE MOST WIDISHREAD
CODE IN
TEOEPRIYTER USE HS NHE FLV -EL IE
T BINERY. AT SEVERL STAGES
N THE
BENEEBPMENT OFLTES
FNIGU
KARIOUS TO
ICU WERESTAKEGINTO
NSIDERBTION INRWORKING OUT THE FIME-ELEMENT
OBEM
THE BAUNOT
ODE WAS
CONSRRUCTED SO
WNAT VOTELS, B NG THE MOST
FREQUENTLY ENCOUNTERAD, COUUD BE SSNT BB THE CIGHT
HAND ONLYM THE
ASRC CRNTERION HSRE WAS EASE OF REMEM
ERING THE CODE80
THE SIEME
S CODE
F THE MARTER TO
E
5θ3 .9'5 θ43177,586T9::7943θ-
'35-
T GS BUI
SREPRESENTEDGAY COMBINATIONS WITO THE LEAST NUMBER OF
CHANGIS MRM ONE POLAN
TY OF CURR
T
TO THE OTHER. ITHIS
EBBBLED A REDUCTIHN IN THE EFFECTS OF CMARACTERISTN DEVICE
QISTORTIONSWM
OD
DSCOMBINAQIO
S MOST SU
WAPT VO TUIS DISTORWION WEREPALLO ATED TO
THE LEASTGFLEKUENTLYSAUSED LETTERS.
BEGBDEBRBH C
TED THERE ARE TTHER CONEIDERETIONSGIN
DEKELOAIMG A OODE.

BINARY DETECTOR ρ=3.0

[illegible] THE
DE[illegible]PMENT[illegible]OF[illegible]TE[illegible]SPRI[illegible]TER [illegible]EC[illegible]N[illegible] [illegible]ARIOU[illegible]TO[illegible]ICU [illegible]ER[illegible]TAKE[illegible]INTO[illegible]
[illegible]O[illegible]SIDER[illegible]T[illegible]ON I[illegible]W[illegible]R[illegible]ING OU[illegible] THE[illegible]FI[illegible]E-EL[illegible]MENT [illegible]O[illegible]E.[illegible]
[illegible]HE B[illegible]U[illegible]OT [illegible]O[illegible]E[illegible]
[illegible]FREQ[illegible]N[illegible]LY ENCOUN[illegible]R[illegible]D[illegible]COU[illegible]D BE S[illegible]N[illegible] BY TH[illegible] [illegible]IGHT[illegible]HAND ONL[illegible]M THE [illegible]
BASRC C[illegible]ERION HERE WAS EA[illegible]E OF REME[illegible]ERING[illegible]THE[illegible]CO[illegible]E[illegible]

TH[illegible] SIEMENS C[illegible]
F TH[illegible] MART[illegible] TO [illegible]N[illegible], [illegible]'9 5[illegible] 5[illegible]5 .9'5 043177,5)6 9::7[illegible]4[illegible],[illegible] [illegible]S-
T GS [illegible]PRESE[illegible]TE[illegible]Y [illegible]OM[illegible]INATIONS[illegible]IT[illegible] [illegible]HE LEAS[illegible] NUM[illegible]ER C[illegible]
[illegible]TO[illegible]
[illegible]

[illegible]73'8+[illegible]'[illegible]3 [illegible]53[illegible] :[illegible]5[illegible]+ 5[illegible]3[illegible]3 -43 [illegible]5[illegible]:9,[illegible]343589,'[illegible],
+3[illegible]3)9[illegible] - [illegible]9+3.

MAXIMUM-LIKELIHOOD DETECTOR ρ=3.0

TELEGRAPHIC COMMUNICTION CUN BE ESYKBLISIED USUNG VARIOUS CODES.
BECAUSE IT IS MOST EGONMICAL, THE MOST WIDZSHREAD CODE IN
TELEPRIYTER USE IS YHE FLVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE
DE ELOPMENT OFLTELSPRINTER TECHNIQUE VARIOUS TOPICU WERE TAKENGINTO
DONSIDERATION INRWORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WASVCONSTRUCTED SO
THAT VOWELS, BE NG THE MOST
FREQUENTLY ENCOUNTERED, COUUD BE SSNT BY THE RIGHT HAND ONLY. THE
BASRC CRNTERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS
F5U8E.[illegible]ABT[illegible]3T[illegible]7[illegible])5E[illegible]T:7[illegible]4[illegible],5[illegible]-
T GS JEIESREPRESENTEDGBY COMBINATIONS WITH THE LEAST NUMBER OF
CHANGZS FROM ONE POLAN
EGABLED A REDUCTION INT[illegible]DFE[illegible]TO[illegible]OC[illegible]AOTE[illegible]STI[illegible]H[illegible]S[illegible]RT[illegible]ONS.
CODEDSCOMBINAQIONS MOST SUBJACT VO THIS DISTORWION WEREPALLOCATED TO
THE LEASTGFLEQUENTLY-USED LETTERS.

BESIDES THE CRITERIH C
TED THERE ARE CTHE :9,'8+343589,'8B,
+3*3)90B,+ - 99+3.

MINIMUM-ERROR DETECTOR ρ=3.0

TELEGRAPHIC COMMUNICTION CUN BE ESYABLISIED USUNG VARIOUS CODES. BECAUSE IT IS MOST EONOMICAL, THE MOST WIDESHREAD CODE IN TELEPRIYTER USE IS NHE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DE ELOPMENT OF TELSPRINTER TECHNIQUE VARIOUS TOPICU WERE TAKENGINTO DONSIDERATION INRWORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT NODE WAS CONSTRUCTED SO
THAT VOWELS, BE NG THE MOST
FREQUENTLY ENCOUNTERED, COUUD BE SENT BY THE RIGHT HAND ONLY. THE BASRC CRNTERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS F3$HE MARTER TO BE EENT, :'9 5@-5 5@3 .9'5 @43177,5)6 9::7447, T GS WEIESREPRESENTEDGBY COMBINATIONS WITH THE LEAST NUMBER OF CHANGES FROM ONE POLAN
ENABLED A REDUCTION INTTHOFEEBEOESTOFOCMHBACTEBRSTICHOSSDEUTCENS. CODED COMBINATIONS MOST SUEJACT VO THIS DISTORTION WERE ALLOCATED TO THE LEASTGFLEQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CETED THERE ARE TTHE :9,'8+343589,'@8, +3*3)908,e - 99+3.

MINIMUM-RISK DETECTOR ρ=3.0

TELEGRAPHIC COMMUNICTION CUN BE ESYABLISIED USUNG VARIOUS CODES. BECAUSE IT IS MOST EONMICAL, THE MOST WIDESHREAD
CODE IN
TELEPRIYTER USE IS NHE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DE ELOPMENT OF TELSPRINTER TECHNIQUE VARIOUS TOPICU WERE TAKENGINTO DONSIDERATION INRWORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT
ODE WAS
CONSTRUCTED SO
THAT VOWELS, BE NG THE MOST
FREQUENTLY ENCOUNTERED, COUUD BE SENT BY THE RIGHT HAND ONLY. THE BASRC CRNTERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS BETHE MARTER TO BE EENT, SO THAT THE MOST FREQUUNTLY OCCURRUN T GS WEIESREPRESENTEDGBY COMBINATIONS WITH THE LEAST NUMBER OF CHANGES FROM ONE POLAN
EBABLED A REDUCTION INTYHOFEEBBOESTOFOCMHBAOTEBRSTICHOSSDORTCONS.
COD
D COMBINATIONS MOST SU
JACT VO THIS DISTORTION WERE ALLOCATED TO
THE LEASTGFLEQUENTLY-USED LETTERS.

BESIDES THE CRITERIA C
TED THERE ARE TTHER CONSIDERETIONSGIN
DEVELOPING A CODE.

BINARY DETECTOR ρ=4.0

[illegible]ICAN, [illegible] CODES.
TELEPRIeTER USE IS eHE FIVE-ELEMENT BINoRe. AT SeVERAL STAGES IN THE
DEeeOPMENT OFeTEeEPRIeTER eECHNeeE VARIOUSeTOPICe WEREeTAKEeeINTO
CONSIDEReTeON INeWOReING OUT THE FIVE-ELeMENT CODE.ee
eFREQeeNTLY ENCOUNTEReD, eCOUeD BE SeNT BY THE eIGHTeHeND [illegible]
BASRC CeeeERION HeRE WAS EASE OF REMEeBERING THE CODEe

THE SIEMENS CODEeWAe DESIGNED TO TeKE INTO ACeOeNT THE SeATISTICS
F THe MAeTER TO eE eeNT, [illegible] eeS-
T eS eEeeeREPRESEeTEDeeY COMeINeTIONSeWITe eHE LEAST NUMeER Oe
[illegible]
CODeD COMBINAeIOeS MOST SUeJEeT VO THIS DISTOReIeN WEREeALLOeATED TO
THEeLEeSTeeeEeUENTLY-USED LETTEeS.

BESIDeS eeE CRITERIe CeTED THEeE ARE eTHee CONSIeERETIONSeIN
DEeELOPeNe A eODE.

MAXIMUM-LIKELIHOOD DETECTOR ρ=4.0

TELEGRAPHIC COMMUNICTION CAN BE ESTABLISHED USUNG VARIOUS CODES.
BECAUSE IT IS MOST EGONOMICAL, THE MOST WIDESPREAD CODE IN
TELEPRIYTER USE IS YHE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE
DEVELOPMENT OF TELEPRINTER TECHNIQUE VARIOUS TOPICU WERE TAKENGINTO
CONSIDERATION INRWORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WAS CONSTRUCTED SO THAT VOWELS, BEING THE MOST
FREQUENTLY ENCOUNTERED, COULD BE SENT BY THE RIGHT HAND ONLY. THE
BASRC CRNTERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS
F THE MARTER TO BE SENT, SO THAT THE MOST FREQUUNTLY OCCURRING SET-
T GS WERE REPRESENTEDGBY COMBINATIONS WITH THE LEAST NUMBER OF
[illegible]
CODED COMBINATIONS MOST SUBJECT VO THIS DISTORTION WERE ALLOCATED TO
THE LEASTGFREQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CITED THERE ARE OTHER CONSIDERETIONSGIN
DEVELOPING A CODE.

MINIMUM-ERROR DETECTOR $\rho=4.0$

TELEGRAPHIC COMMUNICTION CAN BE ESTABLISHED USUNG VARIOUS CODES. BECAUSE IT IS MOST EONOMICAL, THE MOST WIDESHREAD CODE IN TELEPRIYTER USE IS NHE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DE ELOPMENT OF TELEPPINTER TECHNIQÜE VARIOUS TOPICS WERE TAKENGINTO CONSIDERATION INRWORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WAS CONSTRUCTED SO THAT VOWELS, EEING THE MOST FREQUENTLY ENCOUNTERED, COULD BE SENT BY THE RIGHT HAND ONLY. THE BASRC CRNTERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS F THE MARTER TO BE SENT, SO THAT THE MOST FREQUENTLY OCCURRING SET- T GS WERE REPRESENTEDGBY COMBINATIONS WITH THE LEAST NUMBER OF CHABGES ARBEDOGKIDGLARITHEOEFEUGBENOFTCHARECDERESTICTHISTDRVIONS. CODED COMBINATIONS MOST SUBJECT TO THIS DISTORTION WERE ALLOCATED TO THE LEAST FREQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CITED THERE ARE OTHER CONSIDERETIONSGIN DEVELOPING A CODE.

MINIMUM-RISK DETECTOR $\rho=4.0$

TELEGRAPHIC COMMUNICTION CAN BE ESTABLISHED USUNG VARIOUS CODES. BECAUSE IT IS MOST EONOMICAL, THE MOST WIDESHREAD CODE IN TELEPRIYTER USE IS NHE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DE ELOPMENT OF TELEPRINTER TECHNIQUE VARIOUS TOPICS WERE TAKENGINTO CONSIDERATION INRWORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WAS CONSTRUCTED SO THAT VOWELS, BEING THE MOST FREQUENTLY ENCOUNTERED, COULD BE SENT BY THE RIGHT HAND ONLY. THE BASRC CRNTERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS F THE MARTER TO BE SENT, SO THAT THE MOST FREQUENTLY OCCURRING SET- T GS WERE REPRESENTEDGBY COMBINATIONS WITH THE LEAST NUMBER OF CHANGES FROM ONE POLAR TY OF CURRENT TO THE OTHER. ITHIS DEVICE ENABLED A REDUCTION IN THE EFFECTS OF CHARACTERISTIC DISTORTIONS. CODED COMBINATIONS MOST SUBJECT TO THIS DISTORTION WERE ALLOCATED TO THE LEAST FREQUENTLY-USED LETTERS.

BESIDES THE CRITERIA C TED THERE ARE OTHER CONSIDERETIONSGIN DEVELOPING A CODE.

BINARY DETECTOR $\rho=5.0$

BECBOSEP@TCIGOMBSTIC@CUORICAN.@THESMOBEIWCODSPBCADEGBREOUB CODES. TELEPRI@TER USE IS @HE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DE@ELOPMENT OF@TELEPRI@TER TECHNIQUE VARIOUS TOPICS WERE@TAKEN@INTO CONSIDER@TION IN@WORKING OUT THE FIVE-EL@MENT CODE.@

@FREQUENTLY ENCOUNTERED,@COU@D BE S@NT @Y THE @IGHT@HADDRONDOT @ODE @ BAS@C CR@TERION H@RE WAS EASE OF REMEMBERING THE CODE@

THE SIEMENS CODE@WA@ DESIGNED TO T@KE INTO AC@OUNT THE S@ATISTICS OF TH@ MA@TER TO BE @ENT, SO THA@ THE MOST FREQU@NTLY OCCURR@NG @ET-T @S @E@E@REPRESENTED@BY COMBIN@TIONS@WIT@ THE LEAST NUMBER O@ @EANB@ED@R@REDNCTRONA@OTYHOFECBRBEBTOTOCMMBACTBE@STOCHOSSDDOTCONS. COD@D COMBINA@IONS MOST SUBJ@@T TO THIS DISTORTI@N WERE@ALLO@ATED TO THE LE@ST @@EQUENTLY-USED LETTERS.

BESIDES THE CRITERIA C@TED THERE ARE OTHE@ CONSI@ERETIONS@IN DE@ELOPIN@ A @ODE.

MAXIMUM-LIKELIHOOD DETECTOR $\rho=5.0$

TELEGRAPHIC COMMUNICATION CAN BE ESTABLISHED USUNG VARIOUS CODES. BECAUSE IT IS MOST ECONOMICAL, THE MOST WIDESPREAD CODE IN TELEPRIYTER USE IS THE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DEVELOPMENT OF TELEPRINTER TECHNIQUE VARIOUS TOPICS WERE TAKENGINTO CONSIDERATION INRWORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WAS CONSTRUCTED SO THAT VOWELS, BEING THE MOST FREQUENTLY ENCOUNTERED, COULD BE SENT BY THE RIGHT HAND ONLY. THE BASRC CRITERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS OF THE MARTER TO BE SENT, SO THAT THE MOST FREQUENTLY OCCURRING LET-T RS WERE REPRESENTED BY COMBINATIONS WITH THE LEAST NUMBER OF ENANGED PRBEDOOEIBOLANITMEOEFEORSENOFTOHPRECBERESTICTBISTOBVIOBS. CODED COMBINATIONS MOST SUBJECT TO THIS DISTORTION WERE ALLOCATED TO THE LEAST FREQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CITED THERE ARE OTHER CONSIDERETIONS IN DEVELOPING A CODE.

MINIMUM-ERROR DETECTOR ρ=5.0

TELEGRAPHIC COMMUNICATION CAN BE ESTABLISHED USUNG VARIOUS CODES. BECAUSE IT IS MOST ECONOMICAL, THE MOST WIDESPREAD CODE IN TELEPRINTER USE IS THE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DEVELOPMENT OF TELEPRINTER TECHNIQUE VARIOUS TOPICS WERE TAKENGINTO CONSIDERATION IN WORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WAS CONSTRUCTED SO THAT VOWELS, BEING THE MOST FREQUENTLY ENCOUNTERED, COULD BE SENT BY THE RIGHT HAND ONLY. THE BASRC CRITERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS OF THE MATTER TO BE SENT, SO THAT THE MOST FREQUENTLY OCCURRING LET-T RS WERE REPRESENTED BY COMBINATIONS WITH THE LEAST NUMBER OF ENARGEB ARBBDOGEIBOLARITHEOEFEBGEENOFTOHARECOERIBTICTBISTOBVIOES. CODED COMBINATIONS MOST SUBJECT TO THIS DISTORTION WERE ALLOCATED TO THE LEAST FREQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CITED THERE ARE OTHER CONSIDERETIONS IN DEVELOPING A CODE.

MINIMUM-RISK DETECTOR ρ=5.0

TELEGRAPHIC COMMUNICTION CAN BE ESTABLISHED USUNG VARIOUS CODES. BECAUSE IT IS MOST EONOMICAL, THE MOST WIDESPREAD CODE IN TELEPRINTER USE IS THE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DEVELOPMENT OF TELEPRINTER TECHNIQUE VARIOUS TOPICS WERE TAKENGINTO CONSIDERATION IN WORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WAS CONSTRUCTED SO THAT VOWELS, BEING THE MOST FREQUENTLY ENCOUNTERED, COULD BE SENT BY THE RIGHT HAND ONLY. THE BASRC CRITERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS OF THE MATTER TO BE SENT, SO THAT THE MOST FREQUENTLY OCCURRING LET-T RS WERE REPRESENTED BY COMBINATIONS WITH THE LEAST NUMBER OF CHANGES FROM ONE POLARITY OF CURRENT TO THE OTHER. THIS DEVICE ENABLED A REDUCTION IN THE EFFECTS OF CHARACTERISTIC DISTORTIONS. CODED COMBINATIONS MOST SUBJECT TO THIS DISTORTION WERE ALLOCATED TO THE LEAST FREQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CITED THERE ARE OTHER CONSIDERETIONS IN DEVELOPING A CODE.

BINARY DETECTOR ρ=6.0

ΣEGΣUBΣPΣΣCISOMΘSΣIECΣΣOΘICAΣ,ΣΣHESΣOΣΣIΣΣΣDSΣΣΣΣDΘΘΘΣΣOΣΣ CODES. TELEPRIΘTER USE IS ΘHE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DEΘELOPMENT OFΘTELEPRIΘTER TECHNIQUE VARIOUS TOPICS WEREΘTAKENΘINTO CONSIDERATION INΘWORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT ΘODE WAS CONSTRUCTED SO ΘHAT VOWELS, BEING THE MOST FREQUENTLY ENCOUNTERED,ΘCOULD BE SΘNT BY THE RIGHTΘHAND ONLY. THE BASΘC CRΘTERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODEΘWAS DESIGNED TO TΘKE INTO ACΘOUNT THE STATISTICS OF THΘ MAΘTER TO BE ΘENT, SO THAΘ THE MOST FREQUENTLY OCCURRΘNG ΘET-T RS ΘEREΘREPRESENTEDΘBY COMBINΘTIONS WITH THE LEAST NUMBER OF ΘΣΣΣΣΣΣDΘΣΟΣΣΣΣΣΣΣΣΣΣΣΣΣΣYΣΣFΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣΣONS. CODΘD COMBINATIONS MOST SUBJEΘT TO THIS DISTORTION WERE ALLOCATED TO THE LEΘST ΘΘEQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CΘTED THERE ARE OTHEΘ CONSIDERETIONSΘIN DEΘELOPING A ΘODE.

MAXIMUM-LIKELIHOOD DETECTOR ρ=6.0

TELEGRAPHIC COMMUNICATION CAN BE ESTABLISHED USUNG VARIOUS CODES. BECAUSE IT IS MOST ECONOMICAL, THE MOST WIDESPREAD CODE IN TELEPRINTER USE IS THE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DEVELOPMENT OF TELEPRINTER TECHNIQUE VARIOUS TOPICS WERE TAKENGINTO CONSIDERATION IN WORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WAS CONSTRUCTED SO THAT VOWELS, BEING THE MOST FREQUENTLY ENCOUNTERED, COULD BE SENT BY THE RIGHT HAND ONLY. THE BASIC CRITERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS OF THE MATTER TO BE SENT, SO THAT THE MOST FREQUENTLY OCCURRING LET-T RS WERE REPRESENTED BY COMBINATIONS WITH THE LEAST NUMBER OF CHANGES FROM ONE POLARITY OF CURRENT TO THE OTHER. THIS DEVICE ENABLED A REDUCTION IN THE EFFECTS OF CHARACTERISTIC DISTORTIONS. CODED COMBINATIONS MOST SUBJECT TO THIS DISTORTION WERE ALLOCATED TO THE LEAST FREQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CITED THERE ARE OTHER CONSIDERETIONS IN DEVELOPING A CODE.

MINIMUM-ERROR AND MINIMUM-RISK DETECTORS p=6.0

TELEGRAPHIC COMMUNICATION CAN BE ESTABLISHED USUNG VARIOUS CODES. BECAUSE IT IS MOST ECONOMICAL, THE MOST WIDESPREAD CODE IN TELEPRINTER USE IS THE FIVE-ELEMENT BINARY. AT SEVERAL STAGES IN THE DEVELOPMENT OF TELEPRINTER TECHNIQUE VARIOUS TOPICS WERE TAKEN INTO CONSIDERATION IN WORKING OUT THE FIVE-ELEMENT CODE.

THE BAUDOT CODE WAS CONSTRUCTED SO THAT VOWELS, BEING THE MOST FREQUENTLY ENCOUNTERED, COULD BE SENT BY THE RIGHT HAND ONLY. THE BASIC CRITERION HERE WAS EASE OF REMEMBERING THE CODE.

THE SIEMENS CODE WAS DESIGNED TO TAKE INTO ACCOUNT THE STATISTICS OF THE MATTER TO BE SENT, SO THAT THE MOST FREQUENTLY OCCURRING LET- T RS WERE REPRESENTED BY COMBINATIONS WITH THE LEAST NUMBER OF CHANGES FROM ONE POLARITY OF CURRENT TO THE OTHER. THIS DEVICE ENABLED A REDUCTION IN THE EFFECTS OF CHARACTERISTIC DISTORTIONS. CODED COMBINATIONS MOST SUBJECT TO THIS DISTORTION WERE ALLOCATED TO THE LEAST FREQUENTLY-USED LETTERS.

BESIDES THE CRITERIA CITED THERE ARE OTHER CONSIDERETIONS IN DEVELOPING A CODE.

TABLE B.1

SUMMARY OF SIMULATION RESULTS

Detector	Signal-to-noise ratio 2.0	3.0	4.0	5.0	6.0
	Message cost				
Binary	2284	1216	674	445	126
Maximum-likelihood	1239	338	169	126	4
Minimum-error	1207	296	166	123	3
Minimum-risk	540	227	51	7	3
	Number of errors				
Binary (erasures)	425	251	147	83	54
Binary (incorrect symbols)	41	18	7	3	2
Maximum-likelihood	173	65	24	9	4
Minimum-error	147	51	21	6	3
Minimum-risk	157	53	23	7	3
	Error rate				
Binary (erasures)	.398	.235	.137	.077	.050
Binary (incorrect symbols)	.038	.017	.007	.003	.002
Maximum-likelihood	.162	.061	.022	.008	.004
Minimum-error	.138	.048	.020	.006	.003
Minimum-risk	.147	.050	.022	.007	.003

Appendix C

Binary-Detector Performance Characteristics

The detection probabilities, $A(\rho)$ and $B(\rho)$, illustrated in Figures 6.3 to 6.10 have been calculated from the formulas presented in this appendix.

C.1 Cases 1, 4, 5, 8: No fading and fast fading

$$A(\rho) = (1-q)^7$$

$$B(\rho) = 12q^2-60q^3+138q^4-174q^5+118q^6-34q^7$$

where $q = q(\rho)$ is the binary error rate given by:

	Coherent receiver	Non-coherent receiver
No fading	Case 1: $q(\rho) = \varphi(\sqrt{\rho})$	Case 5: $q(\rho) = \frac{1}{2}\exp(-\frac{1}{2}\rho)$
Fast fading	Case 4: $q(\rho) = \frac{1}{2}-\frac{1}{2}(\frac{\rho}{2+\rho})^{\frac{1}{2}}$	Case 8: $q(\rho) = \frac{1}{2+\rho}$

C.2 Cases 2, 3, 6, 7: Slow fading

The detection probabilities are expressed in terms of the functions: $R_1(\rho)$, $R_2(\rho)$, ..., $R_7(\rho)$ defined as

$$R_i(\rho) = \int_0^\infty v \exp(-\frac{v^2}{2}) \varphi^i(v\sqrt{\tfrac{1}{2}\rho})dv$$

at a coherent receiver and

$$R_i(\rho) = (\tfrac{1}{2})^{i-1} (\frac{1}{2+i\rho})$$

at a non-coherent receiver. For flat fading (Cases 2 and 6) the characteristics are:

$$A(\rho) = 7R_1-21R_2+35R_3-35R_4+21R_5-7R_6+R_7$$

$$B(\rho) = 12R_2-60R_3+138R_4-174R_5+118R_6-34R_7.$$

In the case of selective fading (Cases 3 and 7), the formulas are:

$$A(\rho) = 7R_1-9R_2-12R_1^2+5R_3+30R_1R_2-R_4-18R_2^2-16R_1R_3 + +18R_2R_3+3R_1R_4-4R_3^2-3R_2R_4+R_3R_4$$

$$B(\rho) = 12R_1^2-60R_1R_2+90R_2^2+48R_1R_3-12R_1R_4-162R_2R_3+ +76R_3^2+42R_2R_4-34R_3R_4.$$