Technical Report No. 161

03674-5-T

# THEORY OF SIGNAL DETECTABILITY: COMPOSITE DEFERRED DECISION THEORY

by

Richard A. Roberts

Approved by: 

B. F. Barton

for

COOLEY ELECTRONICS LABORATORY

Department of Electrical Engineering
The University of Michigan
Ann Arbor, Michigan

Contract No. Nonr-1224(36)
Office of Naval Research
Department of the Navy
Washington 25, D. C.

March 1965

## ERRATA

| Page | Correction |
|---|---|
| 29 | Add "\|" to $f(L_{n-1}, \Theta_{n-1} \mid L_n, \Theta_n) = 0$ in second paragraph. |
| 53 | Add a "sub zero" to the A in Fig. 3.11. |
| 62 | Multiply right-hand side of Eq. 3.53 by $\exp(-D/2)$. |
| | Add to right-hand side of Eq. 3.54 $D/2$. |
| | Add to right-hand side of Eq. 3.55 $k\,D/2$. |
| 77 | Left-hand side of Eq. 4.41 should be P("A"\|SN). |
| 83 | Change second word on page 83 from "these" to "this." |
| 86 | Add $\Phi$ to integrand of right-hand side of Eq. 4.50. |
| | Change lower limit of second integral from 0 to c. |
| 133-136 | Switch the figure on page 133 with that on page 136. |
| 175 | Change Fig. 4.40 to Fig. 4.38 (3 times). |
| 183 | Add "$\sqrt{}$" to the x and $2\pi$ in the denominator of Eq. B4. |
| 192 | Add "$\{\ \}$" to Eq. E.6 enclosing the argument of $\ell n$. |
| 193 | Add "\|" to left-hand side of Eq. E.12. |
| 194 | Add "$\sqrt{}$" to the g in the denominator of Eq. E.19. |

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

## ACKNOLWEDGMENTS

The author wishes to thank the members of his committee, Professor W. M. Brown, Professor A. B. Macnee, Professor E. L. McMahon, Professor M. O. Reade, and the chairman, Professor H. W. Farris, for their discussions and suggestions in the writing of this dissertation. The author is indebted to Mr. T. G. Birdsall, principal investigator of the contract on which this work was performed, for his constant interest and direction. We would also like to thank the Acoustic Branch of the Office of Naval Research for their support in the research and writing of this thesis.

The rough drafts and final copy were typed by Miss Carol Kruse and Mrs. Lillian Thurston, respectively. Thanks are due to both for a fine job.

The computer programs were written by C. V. Kimball. Mr. Kimball incorporated several special techniques in the computations due to the complicated functions that arose in the analysis. The programs represent a great amount of very good work. Appendix C is due to Mr. Kimball.

## LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

## LIST OF SYMBOLS

| | |
|---|---|
| $\alpha$ | parameter of the signal amplitude distribution |
| a | signal amplitude |
| a(t) | received signal envelope function |
| $A_i$ | parameter of the phase distribution |
| "A" | the response, "signal-and-noise present" |
| ANO | average number of observations |
| $B_i$ | parameter of the phase distribution |
| "B" | the response, "noise alone present" |
| $\beta_c$ | threshold or cut value |
| C | cost of a single observation |
| cd | cost of an observation of quality d |
| d | quality of a single observation |
| D | total available quality |
| E | received signal energy, $\int_0^T [a(t) \cos(\omega t - \phi)]^2 \, dt$ |
| f( ) | probability density function of ( ) |
| f(a\|b, c) | probability of a under conditions b and c |
| $F_k( )$ | optimum expected loss function at stage k |
| g | parameter of the signal amplitude distribution |
| $G_k( )$ | expected look-ahead loss function at stage k |
| h | parameter of the signal amplitude distribution |
| $I_0( )$ | Bessel function of zero order and pure imaginary argument |
| L | log-odds ratio, $\ell n[P(SN)/P(N)]$ |
| $\ell( )$ | likelihood ratio of ( ), $\ell(z) = f(z\|SN)/f(z\|N)$ |
| M | normalizing constant for probability density functions |
| N(a, b) | normal density function with mean a and variance b |
| N | noise power at receiver; also the condition noise alone |
| $N_0$ | noise power density |

## LIST OF SYMBOLS (Cont.)

| | |
|---|---|
| ONP | the optimum nonsequential procedure |
| P( ) | probability of ( ) |
| $Q(\alpha, b)$ | the Q function $\int_{\beta}^{\infty} x \exp\left[-\frac{x^2}{2}\right] \exp\left[-\frac{\alpha^2}{2}\right] I_0(\alpha x)\, dx$ |
| Q | the separable subspaces |
| $R(\alpha, \beta, \gamma)$ | the R function, $\frac{1}{I_0(\alpha\gamma)} \int_{\beta}^{\infty} x \exp\left[-\frac{x^2}{2}\right] I_0(\alpha x) \exp\left[-\frac{\alpha^2}{2} - \frac{\gamma^2}{2}\right] I_0(\gamma x)\, dx$ |
| ROC | receiver operating characteristic |
| s(t) | received signal envelope |
| SKEA | signal known except for amplitude |
| SKEP | signal known except for phase |
| SN | the condition signal-and-noise |
| T(L) | optimum terminal loss function |
| $u_k(L)$ | the average number of observations in SN at stage k |
| $v_k(L)$ | the average number of observations in N at stage k |
| $(W, \Delta_0)$ | the loss due to terminal decision errors |
| x(T) | $(2/EN_0)^{1/2} \int_0^T z(t)\, a(t) \cos \omega t\, dt = r \cos \theta$ |
| y(T) | $(2/EN_0)^{1/2} \int_0^T z(t)\, a(t) \sin \omega t\, dt = r \sin \theta$ |
| $x_k(L)$, P("A"\| N) | the probability of false alarm (at stage k) |
| $y_k(L)$, P("A"\|SN) | the probability of detection (at stage k) |
| r | $(x^2 + y^2)^{1/2}$ |
| $\theta$ | received signal phase, signal parameter variable |
| z(t) | receiver input |
| $\phi$ | true signal phase |
| $\omega$ | true signal frequency |
| $\Theta$ | signal distribution parameter vector |

| | |
|---|---|
| $(\Gamma_k, \Delta_k)$ | decision boundaries in log-odds ratio, L |
| $\Gamma_{h_k}, \Delta_{h_k}$ | decision boundaries in terms of h |
| $\Phi(x)$ | $= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{t^2}{2}\right] dt$ |
| $\phi(x)$ | $= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right]$ |
| $\omega(x)$ | $= \Phi(x)/\phi(x)$ |
| $\tilde{x}_k(L)$ | conditional probability of false alarm (at stage k) |
| $\tilde{y}_k(L)$ | conditional probability of detection (at stage k) |
| $\tilde{u}_k(L)$ | conditional ANO in N at stage k |
| $\tilde{v}_k(L)$ | conditional ANO in SN at stage k |

## LIST OF APPENDICES

## ABSTRACT

### THE THEORY OF SIGNAL DETECTABILITY: COMPOSITE DEFERRED DECISION THEORY

By Richard A. Roberts

The theory of sequential detection problems is extended to procedures of a composite signal hypothesis, such as signals whose amplitude or phase is unknown. Observations of signal-and-noise are assumed not to be independent, so that information can be extracted about the signal parameters for use in the detection process. The yes-no terminal decision must be made within a finite period.

A general theory is presented for the solution of practical problems in the sequential detection of composite signal hypotheses. The theory specifies the optimum stopping rule needed for the sequential detector and the dependence between the observer's opinion of the cause of the reception and the distribution of the unknown signal parameters. The general theory implies the information that must be extracted from the observation is the likelihood ratio and the a posteriori signal parameter distribution. The form of the optimum detector, derived from the general theory, must include an adaptive capability, so that the signal parameter distribution can be sequentially updated.

Applications of the general theory to a signal of unknown phase and a signal of unknown amplitude are presented. The numerical results include (1) optimum receiver designs for both sequential and nonsequential observation procedures, (2) evaluations of the optimum receivers in terms of error performance and average observation time, and (3) comparisons of optimum sequential and nonsequential detectors.

The unknown amplitude problem results in several new conclusions. Some of the more important are: (1) as signal uncertainty increases, the average observation time and range of a priori opinions for which it is profitable to observe, decreases, (2) the savings of

sequential procedures over nonsequential procedures is primarily in error performance rather than observation time, (3) due to the large signal uncertainty of unknown amplitude, the optimum detector must observe for long periods of time to obtain acceptable detection performance, and (4) the graph of the optimum decision boundaries and the mean-motion trajectories of different signal amplitudes can be used as a good approximate method for describing the operation of the optimum sequential detector in the unknown amplitude problem.

# CHAPTER I

## INTRODUCTION

The development and study of methods for extracting information from a background of random interference is called signal detection theory. This thesis is an extension of the present theory of signal detectability in the area of sequential observation procedures and is based on three areas of study: classical detection theory of Grenander, Middleton and Van Meter, and Peterson, Birdsall and Fox (Refs. (1-3); sequential analysis of Wald (Ref. 4); and deferred decision theory of Goode, Birdsall and Roberts (Refs. 5, 6).

Signal detection theory is statistical in nature because of the presence of random interference; thus, a basic mathematical operation of the theory is to make decisions in the face of uncertainty. The method of statistical inference employed throughout this study is the Bayesian philosophy of subjective or personal probabilities, perhaps, best expounded by Savage (Ref. 7). Bayesian statistics are based on a definition of probability as a measure of the opinions of ideally consistent people. The modification of these opinions upon receipt of information, through observation or otherwise, is achieved according to Bayes' Rule.

To place the present work in proper perspective, it is necessary to review certain aspects of classical detection theory, Wald's sequential analysis, and deferred decision theory.

### 1.1 Background and Related Work

Classical detection theory, as developed in the early 1950's (Refs. 1-3), has emphasized a relatively simple detection problem; that of deciding the presence or absence of a signal in noise after a fixed observation length. This highly idealized problem is an appropriate model for the study of many basic sensory and decision processes. However, if a cost is associated with the observation process, classical detection theory must be extended to detection problems in which the observation length is based on observations already taken, i. e., sequential-observation procedures.

The major theory of sequential observation and decision has been based primarily on Wald's theory of sequential analysis (Ref. 4) developed in the 1940's. In this theory Wald assumes that the available observation time is unbounded, i. e., the observer has the option of deferring his terminal decision endlessly if he so chooses. Wald's standard theory is, therefore, often not applicable to detection problems, because approximations are made which are unreliable for small numbers of observations typical in many detection situations. Moreover, standard sequential analysis does not consider as explicit variables the important parameters of observation cost, terminal decision error loss, and a priori opinions, i. e., probabilities of the occurrences of signal-and-noise or noise alone. Standard sequential analysis cannot consider the effects of time varying costs, nonstationary observation statistics, and nonindependent observations in signal-and-noise.

The general formulation of observation-decision procedures has been considered by Wald in a very general and abstract form (Ref. 8). This formulation, because of its generality, contributes little to an understanding of any specific problem. Only in specific applications is it possible to ascertain the effect of various parameters on observation-decision procedures. [1]

Deferred decision theory is a theory of sequential observation and decision characterized by a finite allowable observation time. That is, the observer must make a terminal decision within a specified time. Thus, by choosing a small allowable observation time, it is possible to obtain accurate quantitative results for small numbers of observations. In addition, deferred decision theory considers as explicit variables, observation costs, terminal decision error loss, and a priori probabilities of the occurrence of signal-and-noise or noise alone. It is also possible to investigate the effects of time varying costs, nonstationary observation statistics, dependent signal-and-noise observations, and other complications.

1. 1. 1 Optimization of Sequential-Observation Decision Procedures. A basic consideration of this work is the design of optimum detectors. Therefore, it is important to define "optimum."

---

[1] "Observation-decision procedures" is used, in the context of a detection problem, to denote an idealization of a detection device or receiver used to ascertain signal presence.

In fixed-observation procedures, i. e., the classical detection problem, optimum can be defined in terms of the error performance[1] of the procedure. Fixed procedures are optimized by the correct balance of terminal decision errors. The optimum balance between the miss and false alarm probabilities is obtained by using the likelihood ratio[2] of the observation as the basis for the terminal decision. Birdsall (Ref. 9) has shown, in great generality, that optimum performance based on the likelihood ratio includes any definition of optimum such that detections are not bad and false alarms are not good.

From the above discussion, we can conclude that optimum terminal decisions are based on the likelihood ratio of the total observation. Therefore, in optimum sequential procedures, the terminal decision must be based on the total likelihood ratio (or any monotone function of the likelihood ratio). Classical detection theory supplies the answer of how to respond yes or no after the decision has been made to stop observing. The optimization of sequential procedures, therefore, reduces to a study of how and when to terminate the observation process. The method or rule used to determine how to stop the observation process is termed the "stopping rule" or "observation rule."

In observation-decision procedures for which the observation is sequential in nature, optimum cannot be defined in terms of the error performance alone. We must also consider the average length of observation necessary to reach a terminal decision. The optimum procedure is a balance between the error performance and the observation time.

In deferred decision theory losses are assigned to terminal decision errors and a cost is associated with the observation process. The correct balance between error performance and observation time is achieved by minimizing the expected or average loss for a terminal decision. In other words, the optimum deferred decision procedure is the procedure which minimizes the expected loss for a terminal decision, and this is accomplished by the optimum stopping rule.

The familiar method of optimization of sequential procedures is that used by Wald in sequential analysis. In this optimization the error performance is chosen at some

[1] Error performance refers to the probability of the two errors that can result from a terminal decision. One can response "signal present" when noise is the true cause of the receiver input, or one can respond "signal absent" when signal-and-noise is the true cause. These errors are termed a false alarm and miss, respectively.

[2] The likelihood ratio of the input z is denoted by $\ell(z)$ and is equal to the probability of z in signal-and-noise over the probability of z in noise alone, i. e., $\ell(z) = f(z|SN)/f(z|N)$.

preselected level; the procedure is then optimized by minimizing the average time necessary to reach the prescribed error performance. The often quoted rule-of-thumb is that the sequential procedures save 50 percent in the average time over comparable fixed procedures. This method of optimization places a primary concern on error performance. Deferred decision, in contrast, takes account of the error performance and observation time according to their relative importance as quantitatively described by the losses assigned to errors and the cost assigned to observing.

1.1.2 Background in Deferred Decision Theory. Deferred decision theory was formulated abstractly by Wald (Ref. 8) and Blackwell and Girshick (Ref. 10).[1] As mentioned, the mathematics implied by the above authors becomes useful in signal detection theory only as it is applied to specific problems.

The first application of the mathematical theory of deferred decision theory was published by the late H. Goode (Ref. 5). Goode studied the problem of a signal known exactly in added white Gaussian noise and developed a computing algorithm from the basic mathematics of Blackwell and Girshick. The work of Goode has recently been expanded by Birdsall and Roberts (Ref. 6). They extended the work of Goode and evaluated deferred decision receivers. They also compared the optimum nonsequential procedure and deferred decision and showed that is is not possible to predict where the savings of the sequential procedure occurs, i.e., in error performance or observation time, independently of the available observation time.

Deferred decision theory as studied by Goode, Birdsall, and Roberts assumes: (1) all costs are time invariant, (2) all costs are stationary, and (3) observations are independent in both noise and signal-plus-noise. These assumptions are sufficient to show that (1) the likelihood ratio contains all the observation information needed to determine the optimum stopping rule, and (2) deferred decision processes reduce to Wald's sequential tests

---

[1] The mathematical formulation of the deferred decision problem is given in Section 10.2 of Blackwell and Girshick. The following is quoted from Section 10.2,

> "It is to be observed that, while the averaging process is laborious from a computational point of view, the fact that the determination of the stopping regions and the Bayes risk involves nothing more complicated than taking expectations is of theoretical interest. Also this method can be considered as an iterative procedure for obtaining the Bayes risk and the stopping regions of the nontruncated sequential procedure."

for large available observation times. The implementation of the stopping rule, in the latter case, results in two thresholds, constant in time.

This thesis is an extension of the previous work in deferred decision discussed above, hereafter termed simple deferred decision theory. The theory of simple deferred decision can be generalized in several aspects. For example, the effects of nonstationary observation statistics or time varying costs might be studied. The research reported here generalizes simple deferred decision theory to include uncertainties in the signal hypothesis. In other words, the signal is no longer known exactly, but instead possesses one or more distributed parameters. The problem is termed a composite signal hypothesis. The presence of signal uncertainty in the detection problem implies that observations in signal-and-noise are no longer independent, because there is the possibility of "learning" about the signal parameters as additional observations are taken. In simple deferred decision theory, the concept of learning or adaptation[1] is not applicable, because observations are assumed independent in noise and signal-and-noise. In other words, if signal uncertainties exist, they remain unknown throughout the entire detection process. A practical aspect of a composite signal hypothesis is that a more realistic modeling of many physical problems can be achieved than by a simple signal hypothesis.

An understanding of what is implied by a composite and a simple hypothesis is important for the remainder of this work. To define a composite and simple hypothesis, let $\theta_1, \ldots, \theta_k$ be unknown parameters of the distribution of a random variable z. A statement of the values of $\theta_1, \ldots, \theta_k$ is called a simple hypothesis if it determines all k parameters uniquely. The statement is called a composite hypothesis if it is consistent with more than one value for some parameters. For example, if $\theta_1$ and $\theta_2$ are the mean and variance, respectively, of the normally distributed random variable z, then the statement that z is distributed normally with mean equal zero ($\theta_1 = 0$) and variance equal one ($\theta_2 = 1$) is a simple hypothesis. The statement that z is distributed normally with mean equal zero ($\theta_1 = 0$) and unknown variance ($\theta_2$) is a composite hypothesis.

---

[1] Adaptation is defined as the process of updating probability density functions upon the receipt of information (obtained usually through observations). The updating procedure is given by Bayes' Rule.

Classical detection theory, sequential analysis, and deferred decision theory all implicitly assume that the optimum detector is an unlimited memory device. In other words, the optimum receiver can remember everything needed to solve any specific problem. As a practical consideration, optimum detectors should possess a finite soft memory.[1] A detector with an infinite soft memory is probably not of engineering interest and its performance may be impossible to evaluate. Unlimited memory detectors possessing a finite soft memory are termed adequate-memory devices. The reduction of receiver memory to less than adequate can be justified only on the basis of receiver cost, since the performance of such a receiver is, clearly, suboptimum.

## 1.2 Problem Formulation and Notation

With the preceding as background, let us formulate the detection problem we wish to consider. The basic problem is to design and evaluate detection receivers for the detection of signals in noise. A cost is associated with the observation process. Three types of observation-decision problems will be considered: the fixed-time or classical detection problem, the optimum fixed procedure, and deferred decision.

The optimum fixed procedure allows the observer the freedom of choosing his observation time before the start of the observation. The procedure is called optimum because the observation length is based on all of the observer's a priori knowledge and results in the minimum average loss for a terminal decision. This knowledge includes the cost per unit time, the losses due to a terminal decision error, the "quality of observation,"[2] and the a priori probability of signal-and-noise or noise alone.

As an aside and in order to give the reader a more intuitive understanding of optimum detectors, consider the implementation of the optimum detection receiver for classical detection theory, Wald's sequential analysis, and simple deferred decision theory.

The detection receiver of classical detection theory is a likelihood ratio processor followed by a simple threshold device. In the study of a detection device, we usually

---

[1] Soft memory is a function of the observations and is usually erasable, e.g., magnetic storage in a digital computer. Hard memory is not a function of the observations, e.g., the wiring configuration of an electronic circuit.

[2] "Quality of observation" is a phrase used here to denote such factors as output signal-to-noise ratio, noise figure, and other similar quantities. In each specific problem it is quantified by an appropriate definition.

partition the device into two cascaded sections. The first section processes the physical waveforms, and in the optimum device, transforms them into the likelihood ratio. This first section is a transformer of the multidimensional input waveform into the one dimensional likelihood ratio. Classical detection theory places primary interest on the first section of the receiver. The second section operates on the output of the first section and has as an output the actual decision "A", "signal present," or "B", "signal absent." The second section has been previously defined as the stopping or observation rule. Figure 1.1 presents a schematic diagram of a typical fixed detection device.

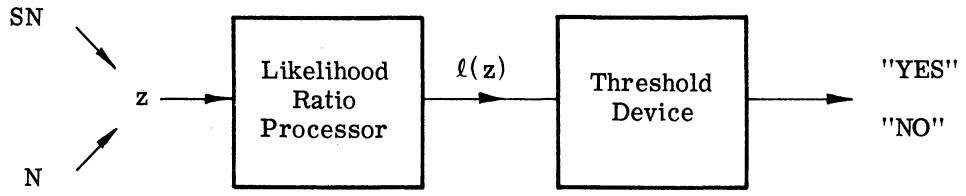


Fig. 1.1. The optimum detection receiver for a fixed observation-decision procedure.

To implement a detection device for a Wald sequential procedure, the first section would again be a likelihood ratio processor, since the optimum terminal decision must be based on the total likelihood ratio of the input. The second section, however, is more complicated than the simple threshold device of a classical detector. The stopping rule for a Wald sequential procedure consists of two thresholds, $(\Gamma, \Delta)$, constant in time. The values of thresholds or decision boundaries depend on the desired error performance. The observation continues as long as the logarithm of the total likelihood ratio remains between the two thresholds. Figure 1.2 depicts the detection receiver for a Wald procedure.

The detection receiver for a simple deferred decision theory is similar to the receiver of Fig. 1.2. The stopping rule for simple deferred decision theory is implemented by a pair of thresholds which converge as the time remaining for observation decreases. The decision boundaries on thresholds meet at a point at the end of the allowable observation time. Figure 1.3 is a schematic representation of a simple deferred decision detector.

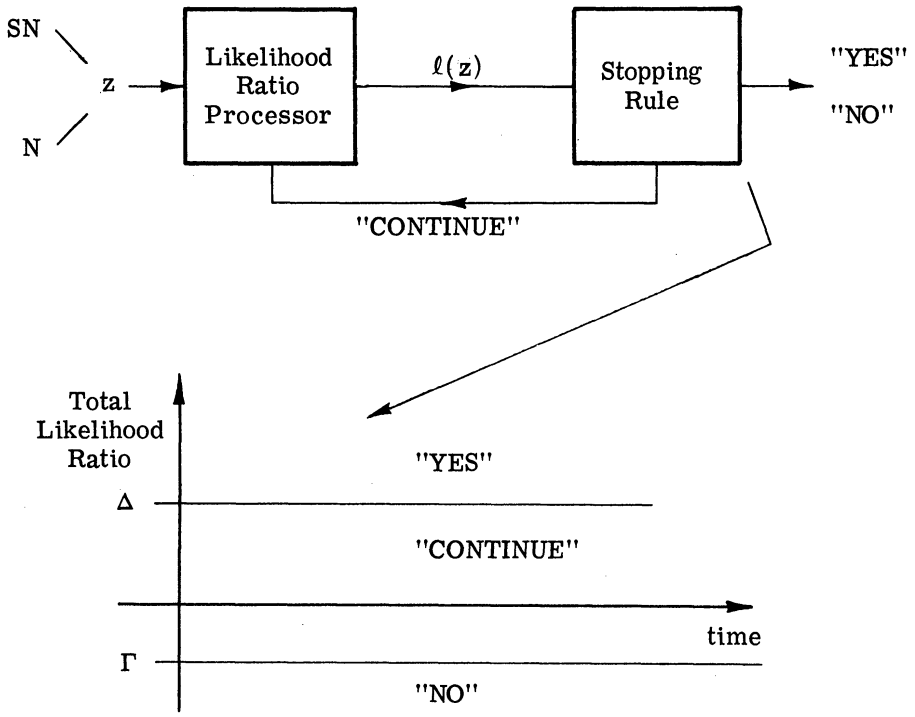


Fig. 1.2. The optimum detection receiver for a Wald sequential procedure.

To state the abstract deferred decision problem, consider a detection situation with possible alternatives signal-plus-noise (SN) or noise alone (N). The observer responds "A" if he decides SN is the cause of the input, otherwise he responds "B". The possible alternative causes, SN and N, have known a priori probabilities P(SN) and P(N), respectively. The receiver input, z, is a random variable because of the added noise. In noise, $z(t) = n(t)$, where n(t) is the noise. In signal-plus-noise, $z(t) = n(t) + S_\Theta(t)$, where $S_\Theta(t)$ is the received signal and, in general, has one or more parameters, $\Theta = (\theta_1, \theta_2, \ldots, \theta_k)$, known statistically, i.e., the probability distribution of $\Theta$ is known. More concisely, the signal hypothesis is composite. The noise hypothesis is assumed to be simple. The a priori density function of the signal parameters, $f(\Theta)$, is assumed known. The conditional probability density functions of the input, z, are also assumed known and are denoted by $f(z|SN)$ and $f(z|N)$. The probability density function of the input under the condition signal-plus-noise is given by

$$f(z|SN) = \int_\Theta f(z|SN, \Theta)\, f(\Theta)\, d\Theta \qquad (1.1)$$

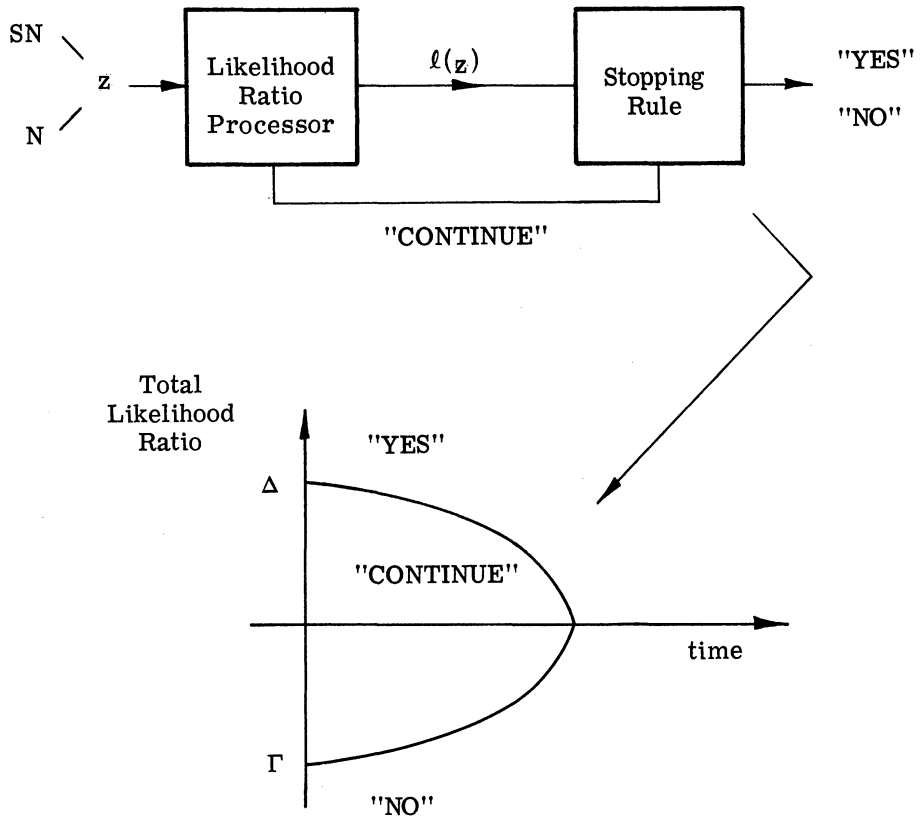


Fig. 1.3. The optimum detection receiver for a deferred decision procedure.

Due to the added noise two types of errors may occur when a terminal decision is made. The observer may respond "A" when the actual cause of the input is N. Similarly, the response "B" may be given when the actual cause is SN. The first error is termed a false alarm, the second error a miss. A loss $W_{FA}$ is associated with a false alarm and $W_M$ is associated with a miss. The cost of a single observation is C. The maximum allowable number of observations or deferrals, $n_{max}$, is finite and known. In general, the density functions $f(z|SN)$ and $f(z|N)$, the losses $W_{FA}$ and $W_M$, the observation quality, and observation cost can all vary with the allowable number of deferrals, n.

The deferred deferred decision problem is to find the procedure which minimizes the total expected loss for a terminal decision, given the parameters of the problem discussed above. The solution to this minimization problem determines the optimum receiver design for the deferred decision problem.

The remainder of this chapter will be devoted to several loosley connected concepts which are common to the specific problems of Chapters 3 and 4. We mention these concepts here to present the reader the necessary notation and ideas needed for the sequel.

The present mathematical formulation assumes the observation process is discrete. That is, the intermediate decisions of whether to terminate or continue observing are made at discrete times rather than continuously. This does not imply that the observation itself is necessarily discrete. It means that the times when this decision is permitted are discrete, usually equally spaced throughout the complete observation.

The abstract deferred decision problem is linked to the physical world, primarily, through the conditional probability density functions of the input, $f(z|SN)$ and $f(z|N)$. The specification of these quantities in a physical problem is, perhaps, the most difficult problem of the receiver designer. The density function of the input under the cause of noise, $f(z|N)$, is usually the easier of the two to specify. This is because in most physical situations the cause of the input is usually noise; there are long periods of time when no signal is transmitted. Thus, the statistics of the noise can be observed for long periods of time allowing $f(z|N)$ to be readily evaluated.

The application of the general theory assumes the noise is added white Gaussian noise (WGN). This type of noise is, in many cases, an appropriate model for physical problems. For example, vacuum tube noise, noise due to thermal motion of electrons, and certain types of atmospheric noise (Ref. 11) are all modeled very closely by WGN. The assumption of added WGN also is mathematically tractable, and this property is often reason enough for using WGN as the noise hypothesis.

Deferred decision detectors are evaluated in terms of the expected loss for a terminal decision. The expected loss is generally expressed as a function of the log-odds ratio, L, defined as

$$L = \ell n\left[\frac{P(SN)}{P(N)}\right] \tag{1.2}$$

The log-odds ratio is a quantitative measure of the observer's opinion of receiver input cause. The log-odds ratios before and after an observation, z, are related by Bayes' Rule. If P(SN|z) is the a posteriori probability of SN, then Bayes' Rule states

$$P(SN|z) = \frac{P(SN)\, f(z|SN)}{P(SN)\, f(z|SN) + P(N)\, f(z|N)} \tag{1.3}$$

Expressed in terms of the log-odds ratio, Eq. 1.3 is

$$L \text{ (based on z)} = L + \ell n\left[\frac{f(z|SN)}{f(z|N)}\right] \tag{1.4}$$

The last term in Eq. 1.4 is the likelihood ratio of the observation. Equation 1.4 states that the log-odds ratio after an observation is the log-odds ratio before the observation plus the logarithm of the likelihood ratio of the input observation.

If the value of the log-odds ratio at the time of a terminal decision is L, then the corresponding probabilities of SN and N are

$$P(SN) = \frac{e^{L}}{1+e^{L}}, \qquad P(N) = \frac{1}{1+e^{L}} \tag{1.5}$$

If the response "B" is made, the probability of error is P(SN). If the response "A" is made, the probability of error is P(N). The minimum expected loss for a terminal decision is, therefore, the minimum of the expected loss for an "A" response, $\frac{W_{FA}}{1+e^{L}}$, and the expected loss for a "B" decision, $\frac{W_{M}}{1+e^{-L}}$. This minimum average loss is denoted by T(L).

A more symmetric form for T(L) is obtained if W and $\Delta_o$ are defined as follows.

$$\frac{2}{W} = \frac{1}{W_M} + \frac{1}{W_{FA}}, \qquad \Delta_o = \ell n\left[\frac{W_{FA}}{W_M}\right] \tag{1.6}$$

Using the above equations, the average terminal loss for a log-odds ratio of L is

$$T(L) = \frac{1}{1+e^{-L}}\frac{W}{2}\left(1+e^{-\Delta_o}\right), \qquad L < \Delta_o$$

$$= \frac{1}{1+e^{L}}\frac{W}{2}\left(1+e^{\Delta_o}\right), \qquad L > \Delta_o \tag{1.7}$$

To obtain this minimum average loss, the response "B" must be made whenever $L < \Delta_0$ and "A" whenever $L > \Delta_0$. For $L = \Delta_0$ either response or a random mixture of the two is permitted. A plot of a typical terminal loss function is shown in Fig. 1.4.

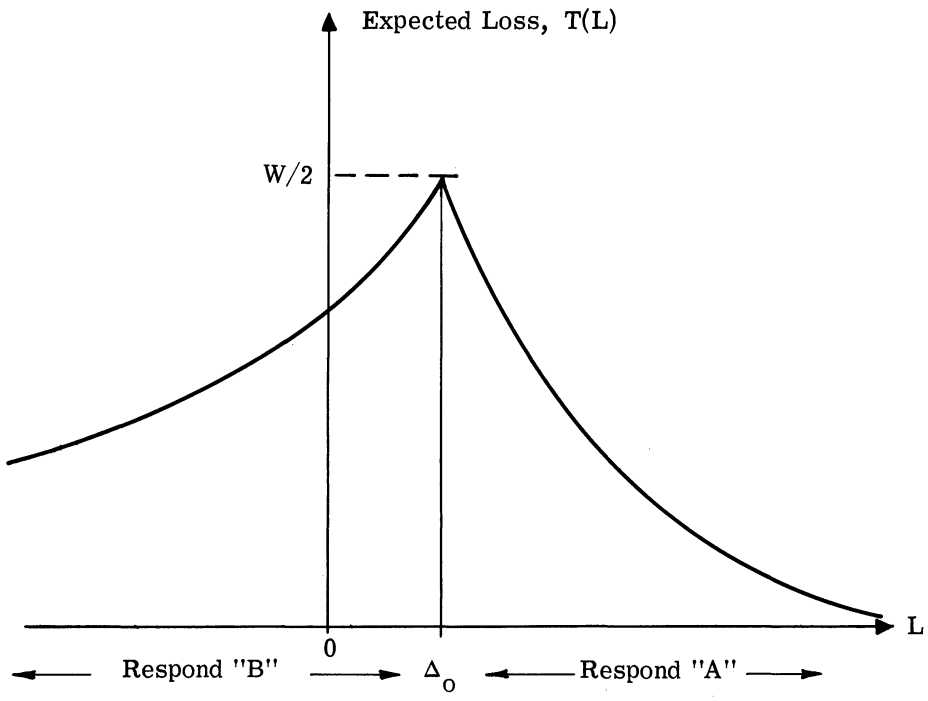


Fig. 1.4. The optimum average terminal loss function, T(L), as a function of the log-odds ratio.

The evaluation of deferred decision procedures is made on the basis of the expected risk for a terminal decision. The evaluation is made more meaningful if one decomposes the total average risk into the average risk due to errors and the average cost due to observation. Since the losses of terminal decision errors and the cost associated with observing are assumed known, this decomposition consists of specifying the error performance and average time necessary to reach a terminal decision.

Error performance, in this study, is given in terms of the receiver operating characteristic (ROC). The ROC is a plot of the probability of detection, P("A"|SN), versus the probability of false alarm, P("A"|N). The parameter along the curve is the log-odds ratio.

A ROC is called normal if the curve can be parameterized by the normal probability distribution function. This means that the probability of detection, P("A"|SN), and the probability of false alarm, P("A"|N), can be written as

$$P(\text{"A"}|SN) = \Phi(\lambda + \sqrt{d}), \quad \text{when} \quad P(\text{"A"}|N) = \Phi(\lambda) \tag{1.8}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt \tag{1.9}$$

Normal ROC curves arise whenever the logarithm of the likelihood ratio is normally distributed under N and SN with equal variances and means separated by the variance, say d. From Eqs. of 1.8 we see that normal ROC curves can be parameterized by the quantity, d, the quality of an observation.

The utility of normal ROC curves is that a physical significant can be attributed to the parameter d. In general, d is the output signal-to-noise ratio of the likelihood ratio processor. For example, a d of one represents a zero db output signal-to-noise ratio. In the special case of a signal known exactly in added white Gaussian noise, a greater physical significance can be given to d. In this case d equals $2E/N_o$ where E is the received signal energy and $N_o$ is the noise power density at the receiver.

The concept of normal ROC curves provides some basis for comparison between various ROC curves. The ROC curves obtained in this work are not generally normal, however, the curves are often approximately normal over the L range of interest. The quality of observation, d, serves as a convenient quantitative measure for both normal and approximately normal ROC curves. For nonnormal ROC curves d is usually measured on the negative diagonal. Using special paper,[1] normal ROC curves plot as a straight line with a slope of one. Figure 1.5 is a graph of several normal ROC's with parameter d. Figure 1.6 shows two nonnormal ROC curves with the implied parameter d.

The average length of observation is measured by the average number of observations (ANO) which is a function of the log-odds ratio. A separate ANO curve exists for

[1]No. 42, 453 of Codex Book Co., Norwood, Massachusetts.

each total available observation time. In other words, the average observation time depends on the available observation. The ROC curves are also a function of the observation time.

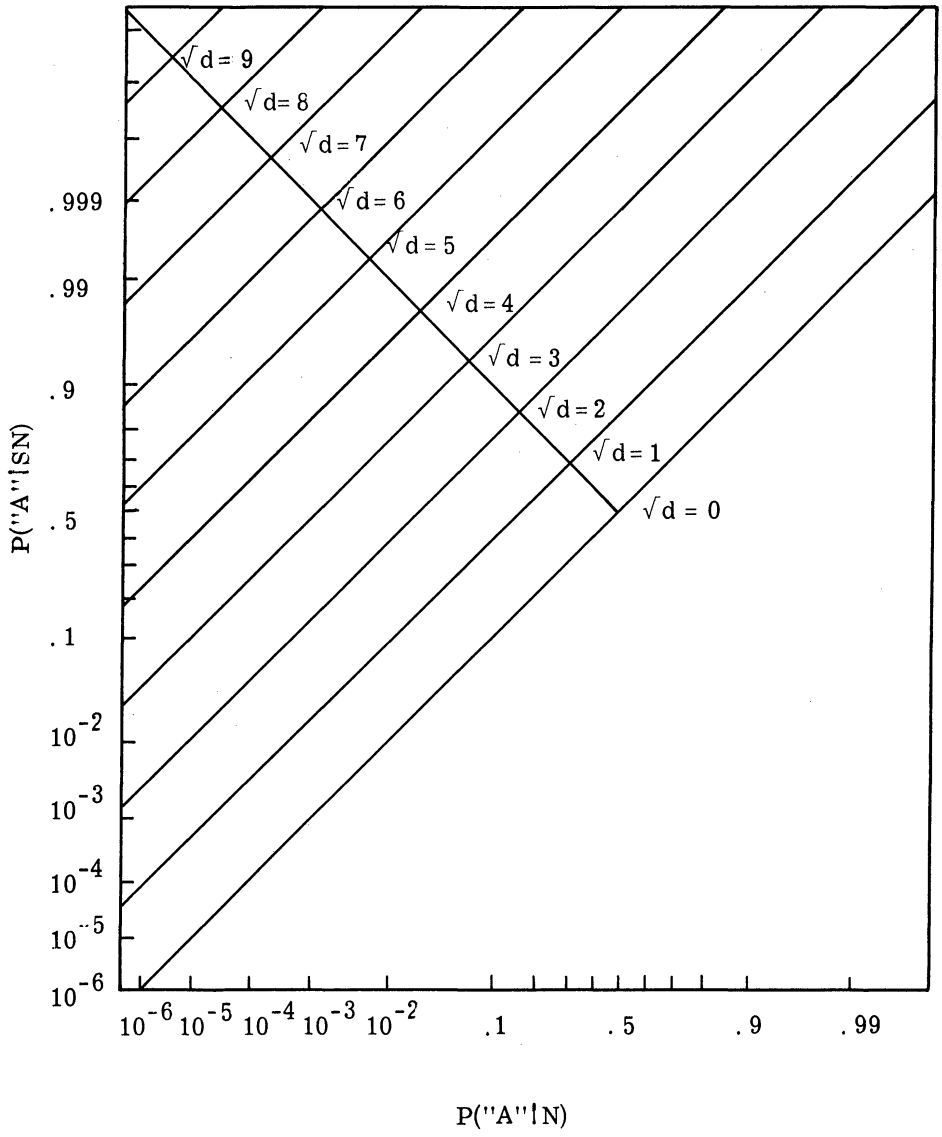


Fig. 1.5. Normal ROC curves with parameter d plotted on normal coordinate paper.

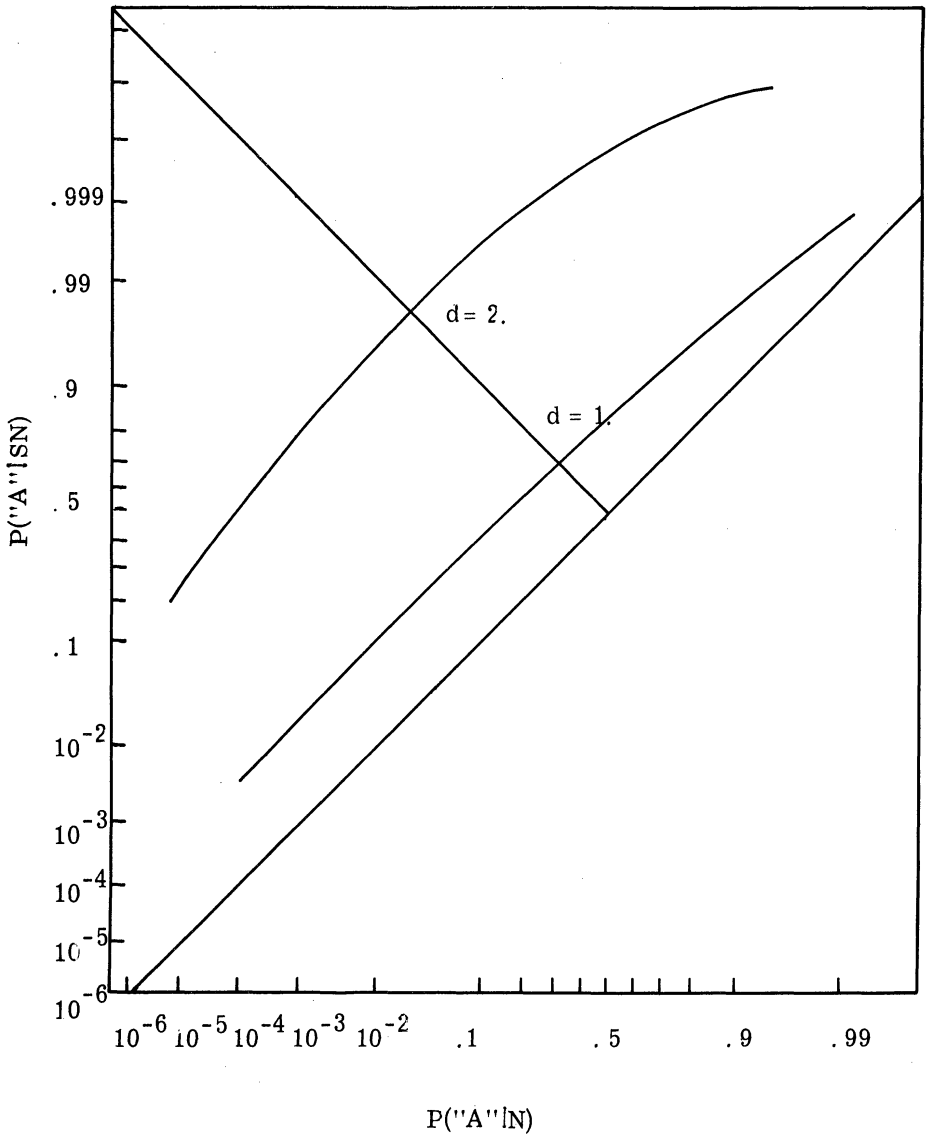


Fig. 1.6. Two nonnormal curves with implied parameter d plotted on normal coordinate paper.

## 1.3 Contributions of the Present Work

Previous applications of deferred decision theory have assumed a simple signal hypothesis, (Refs. 5,6). This assumption implies that observations in N and SN are independent. Thus, the concept of learning about the signal during the observation process is not applicable.

The extension and application of deferred decision theory to include composite signal hypotheses is our basic contribution. In addition to the general formulation and method of solution, given in Chapter 2, we study two specific applications of the general theory. In Chapter 3, we examine the signal uncertainty of unknown phase. Chapter 4 examines a greater amount of signal uncertainty in the form of unknown signal amplitude.

One of the interesting aspects of our work concerns the use of adaptation in optimum detectors. There seems to be a great deal of confusion in the literature on adaptive receivers. The literature (Ref. 12) leads one to believe that adaptive receivers will increase detectability beyond present day optimum receivers. This is not the case. Adaptation does not buy better performance than already optimum detectors. It is possible for adaptation to decrease receiver memory and thereby decrease receiver cost (Ref. 13). In certain problems adaptation is necessary because of receiver memory limitations (Ref. Ibid). Our study of sequential detection of a composite signal hypothesis is another example where adaptation must be used to obtain an optimum solution.

Applications of the general theory to the problems of unknown signal phase and unknown signal amplitude result in receiver designs, receiver evaluations, and conclusions concerning the effect of signal uncertainty in sequential detectors. For example, from the evaluation of the optimum detector for a signal of unknown phase, we can conclude that a phase dispersion of less than $65^{o}$ is, for detection, a signal known exactly. From the numerical evaluations of the optimum detector of a signal of unknown amplitude, we find that increasing signal uncertainty decreases the average number of observations, and results in a high error rate (40%) as compared to the detection of a signal known exactly. Other conclusions conerning the applications of the general theory are given as they occur in the text.

The concept of receiver design which is employed here for composite signal hypotheses is basically different from work in the published literature. The usual method is to design a signal-known-exactly receiver based on some estimate of the distributed signal parameter. We have chosen to include the signal parameter distribution in the overall design without resorting to the concept of estimation and design. Our method is not without its difficulties, as will be seen, however, it is an optimum detector design rather than a suboptimum design obtained by estimating the signal parameter.

# CHAPTER II

## THE GENERAL SOLUTION AND PROPERTIES OF THE COMPOSITE-SIGNAL-HYPOTHESIS, DEFERRED DECISION PROBLEM

### 2.1 The General Solution

The general solution to a composite-signal-hypothesis, deferred decision problem can be given in terms of an iterative algorithm similar to the simple deferred decision problem of Refs. 5 and 6. The iterative solution is possible because at any stage of the observation one can minimize the expected loss for a terminal decision without regard to any loss previously incurred. In the language of state variables,[1] we say that the state of the observation-decision procedure can be specified at any stage of the observation, and the minimum expected loss can be expressed as a function of the state and future observations. Readers familiar with the work of Bellman (Ref. 14) will recognize this as a statement of the principle of optimality.

Consider a sequential, composite-signal-hypothesis problem in which the observer's state of knowledge before the observations begin includes:

(1) $f(\Theta)$, the probability density function of the signal parameters, $\Theta$, prior to the start of observation,

(2) $f(z|SN)$, $f(z|N)$, the distributions of the input, z, in N and SN,

(3) W and $\Delta_0$, the losses due to terminal decision errors,

---

[1] The state of a physical system is the specification of a minimum set of variables, called state variables, from which it is possoble to predict the future behavior of the system if the future inputs to the system are known. Similarly, the state of an observation-decision procedure is the minimum set of variables needed to predict the future behavior of the process given the future observations.

(4) C, the cost of an observation, and

(5) $n_{max}$, the maximum allowable number of observations or deferrals.

We have two possible alternatives, N and SN, which we know from prior knowledge occur with probability $1/(1 + e^{L})$ and $1/(1 + e^{-L})$, respectively. To make decisions, we observe the random variable z with the probability distributions of (2) above. The penalty for taking observations is the cost of an observation. Our problem is to choose N or SN on the basis of the observations and our prior knowledge in such a fashion that the expected loss for our terminal decision is a minimum. Whether we should delay the terminal decision will depend on whether the expected loss for additional observations is less than the loss for making an immediate terminal decision. The expected loss of taking another observation will depend on the expected loss of taking another observation beyond that, etc., until we have exhausted all possible observations. In other words, the expected loss of taking additional observations depends on the allowable number of observations, the actual value being the result of a complicated series of expectations.

1.2.1 General Theory. To begin the process of solution, suppose no observations may be taken, i.e., an immediate terminal decision must be made. The optimum expected loss function as a function of the log-odds ratio for no observations has been derived in Section 1.3, i.e., T(L). This minimum loss is obtained, as is shown in Section 1.3, by responding "A" whenever $L > \Delta_0$ and "B" whenever $L < \Delta_0$. As a matter of notation, we denote the optimum expected loss for no observations by another symbol, $F_0(L) \equiv T(L)$.

Consider next the possiblity of one allowable deferral, i.e., $n_{max} = 1$. For a given log-odds ratio, L, the "look-ahead loss," the loss incurred if an observation is taken, is compared with the loss if no observation is taken, the optimum terminal loss function. The observation is taken only if the average look-ahead loss is smaller than terminal loss function for the log-odds ratio in question.

If "z" is observed and $L_1$ is the a priori log-odds for n = 1, then the log-odds after the observation is $L_0$ and is given by

$$L_0 = L_1 + \ell n[\ell(z)] \tag{2.1}$$

where $\ell(z)$ is the likelihood ratio of the input observation. The loss for this observation z is $T(L_0) = T\{L_1 + \ell n[\ell(z)]\}$ . The average loss if the observation is taken is found by averaging $T\{L + \ell n[\ell(z)]\}$ over all possible observations, z. To this must be added the cost of the observation. The sum of these two quantities is the average look-ahead loss at stage n = 1, $G_1(1)$.

$$G_1(L) = \int T\{L + \ell n[\ell(z)]\} f(z)\, dz + C \tag{2.2}$$

where f(z) is the probability density function of z. The optimum expected loss function at stage n = 1 is $F_1(L)$.

$$F_1(L) = \min [T(L), G_1(L)] \tag{2.3}$$

The intersection of $G_1(L)$ and T(L) define two points, $(\Gamma_1, \Delta_1)$, which constitute the optimum stopping rule for n = 1. If, with one more possible observation, the log-odds has value between $\Gamma_1$ and $\Delta_1$, then the next observation is profitable; otherwise, the observation should not be paid for, and instead, an immediate terminal decision should be made. Figure 2.1 depicts the alternatives that may be taken for the n = 1 state.

The composite nature of the signal hypothesis has thus far been suppressed. In order to calculate the likelihood ratio, $\ell(z)$, however, we must use our knowledge of the composite-signal-hypothesis. The likelihood ratio of the input is an average likelihood ratio obtained by a weighted average with respect to the signal parameter density function, $f_1(\Theta)$.

$$\ell(z) = \int_{\Theta} \ell(z|\Theta)\, f_1(\Theta)\, d\Theta \tag{2.4}$$

The conditional likelihood ratio $\ell(z|\Theta)$ can be written as

$$\ell(z|\Theta) = \frac{f(z|SN, \Theta)}{f(z|N)} \tag{2.5}$$

Thus, Eq. 2.4 can be written as

$$\ell(z) = \int_{\Theta} \frac{f(z|SN, \Theta)}{f(z|N)}\, f_1(\Theta)\, d\Theta \tag{2.6}$$

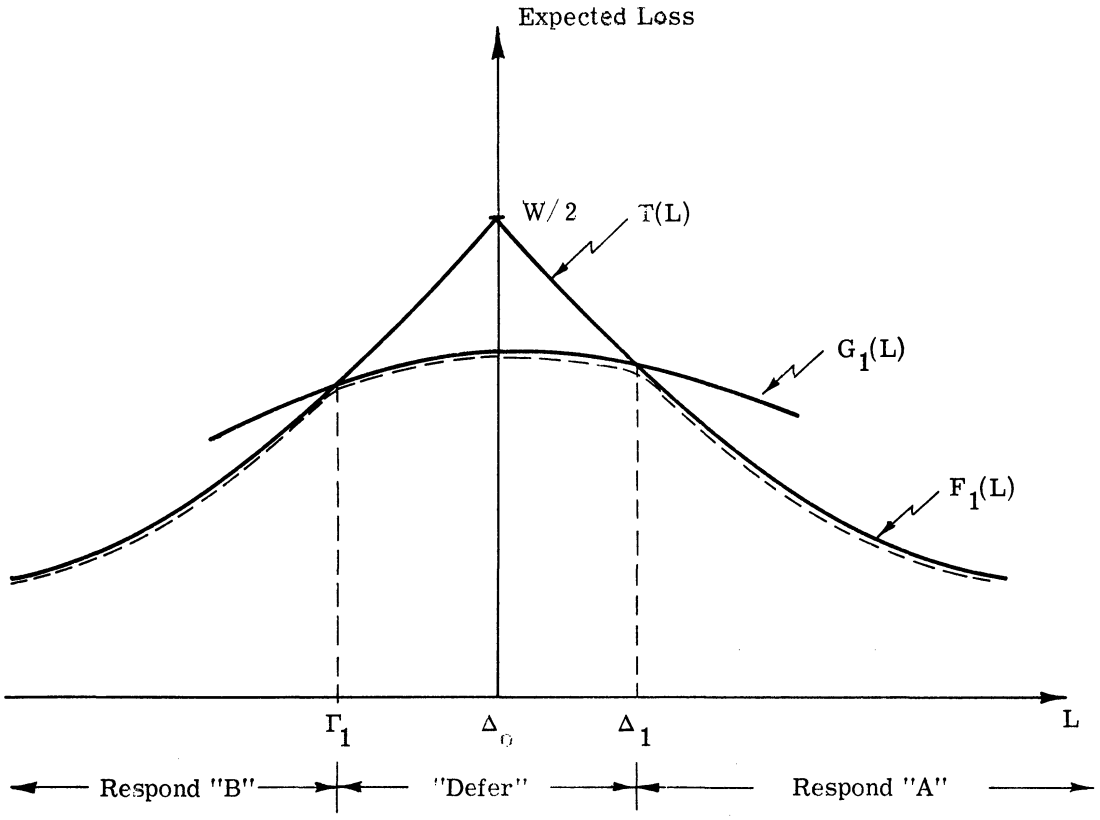


Fig. 2.1. The look-ahead average loss function, the optimum average loss function, and the possible alternatives for one allowable observation.

Equation 2.4 states that the likelihood ratio of the observation, $\ell(z)$, is given by averaging the likelihood ratio for known signal parameters with respect to the probability of occurrence of the signal parameters. The probability density function of the input can be written in terms of the conditional probability density function as

$$f(z) = P(SN) \cdot f(z|SN) + P(N) \cdot f(z|N) \tag{2.7}$$

$$= \frac{1}{1 + e^{-L}} f(z|SN) + \frac{1}{1 + e^{L}} f(z|N) \tag{2.8}$$

The conditional probability density function $f(z|SN)$ can be expressed by

$$f(z|SN) = \int_{\Theta} f(z|SN, \Theta)\, f_1(\Theta)\, d\Theta \tag{2.9}$$

If z is observed, the probability density function of the signal parameters after the observation can be found by Bayes' Rule as shown in Eq. 2.10

$$f_0(\Theta) = f_1(\Theta|z) = \frac{f(z|\Theta)}{f(z)} f_1(\Theta) \tag{2.10}$$

Equation 2.10 can be rewritten as

$$f_0(\Theta) = \frac{P(SN) \cdot f(z|SN, \Theta) + P(N) f(z|N)}{P(SN) \cdot f(z|SN) + P(N) f(z|N)} \cdot f_1(\Theta) \tag{2.11}$$

Since it is possible to update the signal parameters only under the condition SN, Eq. 2.11 can be written as

$$f_0(\Theta) = \frac{\ell(z|\Theta)}{\ell(z)} f_1(\Theta) \tag{2.12}$$

The one-allowable-observation case does not make use of the a posteriori distribution of Eq. 2.12 in the solution. However, for n greater than one, the updated probability distributions of the signal parameters must be used in the solution for the calculation of the average likelihood ratio. This continual modification of the observer's opinion of the signal parameters is part of the adaptation that is used in the decision process. The observer's opinion of the presence or absence of a signal is also modified through the observation in a similar manner.

The calculation of the updated probability distribution of the signal parameters as observations are taken can also be used as a classification output. Classification is the next hierarchy of decision beyond the simple yes-no decision on signal existence. It is the specification of the values of the signal parameters after one has decided a signal is present. For a continuous distribution on the signal parameters, classification is the specification of the probability distribution of the signal parameters.

To illustrate the use of the a posteriori distribution of the signal parameters, consider next the case of $n_{max}$ equal to two, i.e., there is a maximum of two allowable deferrals. The solution procedure is to compare the average look-ahead loss at stage n = 2 with the terminal loss function, T(L). If the average look-ahead loss function, $G_2(L)$ is smaller that T(L), then it is profitable to take the observation. The observation information modifies both the a priori signal parameters distribution and the a priori log-odds ratio.

After one observation, these updated quantities are used to determine the optimum procedure for the remaining observation. The method of solution must be such that the $n_{max} = 2$, "one-observation-taken" state is identical to the $n_{max} = 1$, "no-observation-taken" state to effect an economical computer solution (see Section 2.3).

At stage n = 2, let $L_2$ be the a priori log-odds ratio and $f_2(\Theta)$ the a priori distribution on the signal parameters. The terminal loss if no observations are taken is again T(L). The average look-ahead loss is $F_1\{L_2 + \ell n[\ell(z)]\}$ averaged with respect to the distribution of z. This average look-ahead loss plus the cost of the observation at stage n = 2 is $G_2(L)$.

$$G_2(L) = \int F_1\{L + \ell n[\ell(z)]\}\, f(z)\, dz + C \tag{2.13}$$

The optimum expected loss at stage n = 2 is $F_2(L)$

$$F_2(L) = \min\,[T(L),\ G_2(L)] \tag{2.14}$$

The intersections of T(L) and $G_2(L)$ determine the optimum decision points $(\Gamma_2, \Delta_2)$, in log-odds ratio. For the log-odds ratio greater than the upper decision point, $\Delta_2$, the response is "A", for L less than the lower decision point, $\Gamma_2$, the response is "B", otherwise the response is "C", continue. As before the likelihood ratio is calculated by averaging the conditional likelihood ratio, $\ell(z|\Theta)$, with respect to the a priori distribution of the signal parameters at n = 2, i.e., $f_2(\Theta)$. This is Eq. 2.15.

$$\ell(z) = \int_{\Theta} \ell(z|\Theta)\, f_2(\Theta)\, d\Theta \tag{2.15}$$

The a posteriori signal parameters density function is $f_1(\Theta)$ and is given by

$$f_1(\Theta) = \frac{\ell(z|\Theta)}{\ell(z)}\, f_2(\Theta) \tag{2.16}$$

Using Eq. 2.4, Eq. 2.16 can be written

$$f_1(\Theta) = \frac{\ell(z|\Theta)}{\int_{\Theta} \ell(z|\Theta) \cdot f_2(\Theta)\, d\Theta}\, f_2(\Theta) \tag{2.17}$$

The observer is now in the n = 1 state with a new log-odds ratio, $L_1$, determined from the a priori log-odds, $L_2$, by Bayes' Rule

$$L_1 = L_2 + \ell n[\ell(z)] \tag{2.18}$$

It has been demonstrated above that it is necessary to calculate the "current opinion" of the signal parameter distribution at each stage of the observation. It is up to the receiver designer whether to make explicit use of this function or not. In the case of detection, there is no need to design the equipment so that this density function can be recovered. However, in the case that one wanted a classification output, it is probable that one could design the receiver to obtain the density function. This adaptive feature of the problem solution is a natural consequence of the manner in which the solution is obtained. The adaptation is a by-product which one may choose to use or ignore depending on the engineering application.

## 2.2 The Closure Property

The general formulation of the deferred decision problem presented in Section 2.1.1 is rather straightforward. It is similar to the formulation of simple deferred decision theory. Again, as in the case of simple deferred decision theory, it is the application of the general theory to specific problems that is of interest in the design and evaluation of detection receivers. The application of the general theory, of course, implies computations of numbers and in the computations certain difficulties arise. The remainder of this chapter is devoted to a discussion of two computational problems.

Consider the stage-by-stage calculation of the distribution of the signal parameters as given by Eq. 2.12. Equation 2.12 is a simple application of Bayes' Rule and given any distribution on the parameters, the stage-by-stage calculation of the resulting distributions is, conceptually, simple. However, if the functional form of the distributions changes with succeeding observations, then there is a practical problem concerned with receiver memory. The solution procedure we have presented is predicated on known the stage-by-stage a posteriori distribution of the signal parameters. Therefore, a problem in which the signal parameter distirubtion changes functional form with each observation means the receiver must possess an infinite soft memory, or at least a very large amount of soft memory. Because we are interested in (1) engineering applications of the theory presented here, and

(2) full memory receivers, the specific problems we consider must be restricted to problems in which full memory receivers possess a finite soft memory, i. e., adequate memory receivers.

A necessary condition for a receiver design to result in an adequate memory receiver is the closure property. The closure property is the invariant character of the functional form of the a priori signal parameters distribution after an observation. In other words, the functional form of the distribution must remain the same after it is modified by observation; the only changes that may occur are changes in the value of the signal parameters.

In addition to the engineering design considerations, the closure property is important in the computational aspects of deferred decision problems. Because of the complexity of the functions involved, analytical solutions to deferred decision problems are not possible, even for the very simplest problems. To obtain quantitative results, computations must be programmed on high-speed digital computers. If the parameter distribution is not closed, computations become prohibitive even for large computers. Thus, all parameter distributions considered in this thesis will possess the closure property.

Some simple consequences of the closure property are evident from Eq. 2. 19 below.

$$f(\Theta \mid \text{based on } z) = \frac{\ell(z \mid \Theta)}{\ell(z)} \cdot f(\Theta) \tag{2.19}$$

Consider the problem of choosing the functional form of $f(\Theta)$ such that $f(\Theta)$ is closed under observation. The specific form of $f(\Theta)$, of course, depends on the particular problem in question. A suitable choice for $f(\Theta)$ can be found without resorting to a trial and error method by noting the functional form $\ell(z \mid \Theta)$. Since $\ell(z)$ is a normalizing constant independent of $\Theta$, $\ell(z \mid \Theta)$ is the function which determines the closure property. For specific examples the reader is referred to Chapters 3 and 4.

Another property evident from Eq. 2. 19 is that independent of the form of $f(\Theta)$, the functional form of $f(\Theta)$ after several observations is, in many cases, a closed form times a term proportional to the original distribution. This can be seen more clearly by examining the form of $f(\Theta)$ after two independent observations, $z_1$ and $z_2$. After two such observations

the distribution of $f(\Theta)$ is

$$f(\Theta|z_1, z_2) = \frac{\ell(z_2|\Theta)}{\ell(z_2)} \cdot \frac{\ell(z_1|\Theta)}{\ell(z_1)} \cdot f(\Theta) \tag{2.20}$$

(A similar expression results for more than two observations.) The functional form of the conditional likelihood ratio, $\ell(z|\Theta)$, in a specific problem does not change with observation. For example, if $\ell(z|\Theta)$ is an exponential in z and $\Theta$, then $\ell(z_2|\Theta)\,\ell(z_1|\Theta)$ is again an exponential in z and $\Theta$. For many problems $\ell(z_k|\Theta)\,\ell(z_{k-1}|\Theta)\cdots$ is a closed form. Since $\ell(z_k)\,\ell(z_{k-1})\cdots$ is independent of $\Theta$, the functional form of $f(\Theta)$ after several observations is, therefore, a closed form times a term proportional to the original distribution.

The concept of closed distributions can be characterized using the theory of sufficient statistics. A sufficient statistic of an observation is information equivalent to the observation for the purposes of the problem. For example, in the yes-no detection problem the likelihood ratio of the input (or any monotone function of the likelihood ratio) is a sufficient statistic of the observation.

In symbols, the concept of a sufficient statistic can be characterized as follows. If $f(\Theta|z)$ is the a posteriori probability of $\Theta$ based on the observation of y and

$$f(\Theta|z) \sim f[\Theta|S_1(z), S_2(z), \ldots, S_k(z)] \tag{2.21}$$

then $\{S_j(z)\}_{j=1}^{k}$ is a sufficient statistic for the observation z. (The symbol "$\sim$" means "is equivalent to.") From the preceding discussion, we can infer that a closed distribution exists for a particular problem if there exists a sufficient statistic of the observation with a finite number of parameters and the number of these parameters is independent of the number of observations taken. Equation 2.21 is a statement that the signal parameter distribution has a finite number of parameters dependent only on the observations and independent of the number of observations taken.

An independent study of closed distributions has been carried out by Spragins (Ref. 15). Spragins investigates closed distributions abstractly without specific application to a particular problem and claims to exhibit a method for constructing closed distributions in any given problem. The method is basically the same as discussed in the paragraph below Eq. 2.19. The work by Spragins gives a tabulation of ten types of closed density functions but does not include the closed distributions derived in Chapters 3 and 4.

One further point concerning closed distributions should be mentioned. There are two points of view of how to carry out the updating procedure defined by Eq. 2.19. We can update the signal parameter distribution under the possiblity that either N or SN is present, or we can investigate the signal parameter distribution only under the SN hypothesis. Although, at first glance, the two methods may seem very similar they result in different conclusions. The present study assumes that the updating procedure occurs only under the condition SN, because only when the signal is being transmitted is it possible to obtain information on the signal parameter distribution. Referring to Fig. 2.2, the switch must be connected to the signal generator before we can apply Eq. 2.19.

These two methods of considering closed distributions can be stated in a more formal manner. If z is an observation and $\{z\}$ a sequence of observations, then to update under N or SN is to investigate closed distributions for which the a posteriori distribution is of the form $f(\Theta|\{z\})$ . The present study assumes the input is due to signal-plus-noise, i.e., the a posteriori distribution is of the form $f(\Theta|\{z\}, SN)$ .

The results of the latter study are altogether distinct from the results of updating $f(\Theta|\{z\})$ . For example, in the detection of a signal of unknown amplitude, developed in Chapter 4, if one updates $f(\Theta|\{z\})$ , it is not possible to obtain a closed signal amplitude distribution. Using the formulation of the present work, however, a very general set of closed distributions on signal amplitudes are obtained. (See Chapter 4.)

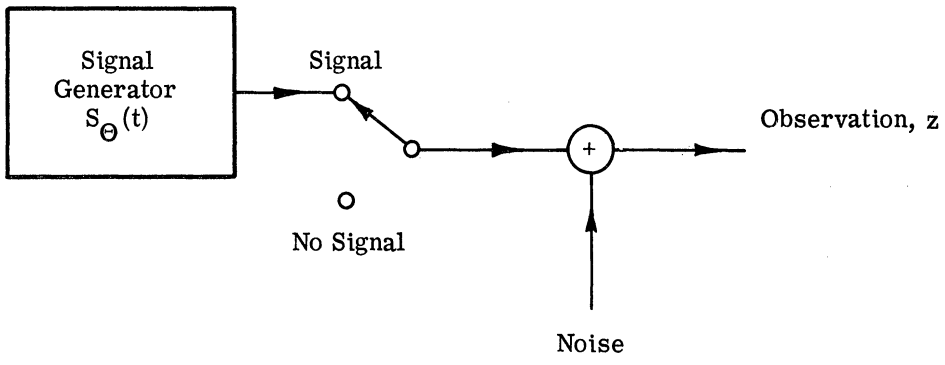


Fig. 2.2. A schematic diagram depicting two situations for updating a signal parameter distribution.

## 2.3 Subspace Separability

In the development of the general method of solution for sequential, composite-signal-hypothesis problems presented in Section 2.1, the basic computation is the calculation of expected values. The expected values are used in determining the average look-ahead risk functions which, in turn, are compared with the terminal loss function to ascertain the optimum stopping rule.

The computation of an expected value is a weighted average, and in our problem, is an average with respect to the distribution of the observation. The distribution of the observation is a function of the values of the signal parameters. The observation distribution, therefore, changes after each observation because of the updating procedure on the signal parameter distribution.

For purposes of discussion consider a distribution of signal parameters given by

$$f(\Theta) = g(S, \theta_1, \theta_2, \ldots, \theta_k) = g(S, \Theta) \tag{2.23}$$

where $\Theta$ is a vector with component values equal to the value of the parameters in the signal $S_\Theta(t)$. At the nth stage of observation the average look-ahead risk, $G_n$, is a function of the updated log-odds ratio, $L_n$, and the updated values of the signal parameters, $\Theta_n$. (The distribution of signal parameters is assumed closed.) Thus

$$G_n = G_n(L_n; \theta_{1,n}, \theta_{2,n}, \ldots, \theta_{k,n}) \tag{2.24}$$

The risk at the nth stage is found by averaging the optimum expected risk function of the $(n-1)^{\underline{st}}$ stage.

$$G_n(L_n; \Theta_n) = E\left[F_{n-1}(L_{n-1}; \Theta_{n-1})\right] \tag{2.25}$$

The notation of Eq. 2.25 is compressed in that $L_{n-1}$ and $\Theta_{n-1}$ are functions of $L_n$ and $\Theta_n$ and the observation, z.

$$L_{n-1} = L_{n-1}(L_n, z) \tag{2.26}$$

$$\Theta_{n-1} = \Theta_{n-1}(\Theta_n, z) \tag{2.27}$$

The averaging process implied by Eq. 2.25 can be calculated in many equivalent ways. For example, in the case of simple signal hypothesis (Ref. 5), this averaging process is accomplished by averaging with respect to the distribution of $\ell n[\ell(z)]$. This is a "one-dimensional" average in that the average risk functions are calculated as a function of a single variable, L.

Let us consider some possible methods of calculating the expectation of Eq. 2.25. For example, it is theoretically possible to take the expected value with respect to the observation space. This means that for each possible observation the transformation of $L_n \rightarrow L_{n-1}$ and $\Theta_n \rightarrow \Theta_{n-1}$ is calculated. Each observation is then multiplied by an appropriate weight in accordance with its frequency of occurrence. These results are summed to obtain the expected value. Typically, the dimensionality of the space of observations is the dimensionality of the input observation in 2WT space. In practice, this might be easily $10^6$ or more. Thus, this manner of computing the expectation of Eq. 2.25 involves far too many computations.

If the probability distribution of the signal parameters is closed, then it is possible to reduce the dimensionality from that of the observation space to k + 1 dimensions, where k is the number of signal parameters, $\Theta = (\theta_1, \theta_2, \ldots, \theta_k)$. This can be seen by examining the state of the observation process at any stage of the observation. The state of the observation-decision process is given by L, $\Theta$, n, $n_{max}$. Suppressing the overt specification of n and $n_{max}$, to calculate the average look-ahead risk function at stage n, $G_n$, it is sufficient to know the optimum risk function at stage n - 1 for a k + 1 dimensional space, i.e., the space $L_{n-1} \times \Theta_{n-1}$. To calculate $G_n$ for one point in the k + 1 dimensional space of $L_n \times \Theta_n$, the optimum risk function at stage n - 1 must be known. Thus, the averaging process necessary to calculate the average look-ahead loss function is an average calculated over k + 1 random variables. To calculate one point on the $G_{n+1}$ look-ahead loss function, the solution over the entire L x $\Theta$ space for all previous n stages must be known.

Is this a satisfactory situation for computations to proceed economically? To answer this question, consider a problem in which the parameter space is one dimensional, i.e., $\Theta = \theta_1$. The L x $\Theta$ space is thus two dimensional. At the nth stage of the observation in order to calculate one point on the average look-ahead loss function, $G_n$, the expected value of the optimum loss function at the n - 1 stage, $F_{n-1}$, must be known. This expectation

must be taken over the complete L x $\theta_1$ space at the n - 1 stage. Assume that the L x $\theta_1$ can be represented approximately by 100 points along the L axis and 100 points along the $\theta_1$ axis. For this hypothetical case, to obtain one point on the $G_n$ look-ahead loss function, the averaging process must be computed over $10^2 \times 10^2 = 10^4$ points. To continue the iterative process to the n + 1 stage, the risk function for the complete L x $\theta_1$ space at the nth stage must be known. If the L x $\theta_1$ is represented by $10^4$ points and each point is obtained by averaging between $10^4$ points, then $10^8$ averages must be computed to continue the process one deferral. Even with today's high-speed computers, the two-dimensional average discussed above represents a large amount of computation time. For a larger dimensional parameter space the number of computations is, of course, very much larger.

From the calculations presented in the preceding paragraph it would seem that the detailed solution to a specific composite-signal-hypothesis problem in a sequential observation procedure is not economically possible. However, it may be possible to reduce the dimensionality of a deferred decision problem from the dimension of the L x $\Theta$ space to a lower dimension; in particular, to one dimension. In certain problems the probability density function of the signal parameter(s) at the $(n-1)^{st}$ stage is zero except on some proper subspace of L x $\Theta$ at the nth stage. In symbols $f(L_{n-1}, \Theta_{n-1}\ L_n, \Theta_n) = 0$ except on some proper subspace of $L_n \times \Theta_n$ where f is the probability of obtaining $(L_{n-1}, \Theta_{n-1})$. If this is the case, then instead of having to solve for all the points in a k + 1 dimensional subspace of the $L_{n-1} \times \Theta_{n-1}$ space to calculate one point on the average look-ahead loss, $G_n$, one need carry out the average over a smaller subspace, namely, the dimension of the proper subspace of $L_n \times \Theta_n$.

More precisely, let Q be the proper subspace of L x $\Theta$. If there exists a sequence of proper subspaces $\{Q_j\}_{j=1}^{n}$ such that the average look-ahead loss function at the jth stage, $G_j(Q_j)$, depends only on $F_{j-1}(Q_{j-1})$, then the problem is said to be "subspace separable." The solution to the deferred decision problem at the jth stage in a subspace-separable problem is the solution on the union of the subspaces from j = 1 to the jth stage. If $Q_j'$ and $Q_j$ are subspaces of L x $\Theta$ at the jth stage, then either

$$Q_j' \cap Q_j \equiv Q_j \quad \text{or} \quad Q_j' \cap Q_j \equiv \phi \tag{2.27}$$

That is, either the subspaces at the jth stage are identical or they are disjoint. The subspaces are layers like sheets of mica, in the $L \times \Theta$ space, and one subspace, $Q_j$, does not intersect another subspace $Q_j'$, for all j. As discussed previously, for an economical solution Q must be a one-dimensional subspace of the $L \times \Theta$ space. This is true for the two problems considered in detail in Chapters 3 and 4.

In summarizing the results of this chapter, we emphasize again that the general formulation is much too abstract to give any insight into sequential observation procedures as they pertain to detection problems. It is only in the applications of the general theory that one is able to draw conclusions about sequential detection procedures. From the discussion presented in this chapter, it is clear that it may be computationally impractical to apply the general theory unless two simplifying assumptions are met: (1) closed signal parameter distributions, and (2) one-dimensional subspace separability.

# CHAPTER III

## DETECTION OF A SIGNAL KNOWN EXCEPT FOR PHASE

The first problem we consider in detail is the detection of a signal known except for phase (SKEP) in added white Gaussian noise (WGN). The solution to this problem is well known (Refs. 3, 16-18) for an a priori uniform phase distribution and a fixed observation-decision procedure, i. e., a detection procedure in which the terminal decision is made after a fixed observation time, or more generally, fixed observation quality. Under these assumptions the optimum (likelihood ratio) receiver can be realized by a matched filter (Refs. 3, 16-18).

### 3.1 Problem Statement and Notation

The detection problem of a SKEP in added WGN is generalized here to include: (1) optimum receiver design and evaluation for fixed observation-decision procedures for a class of a priori phase distributions, (2) optimum receiver design and evaluation of the optimum nonsequential procedure for the same class of a priori phase distributions considered in (1), and (3) the optimum likelihood processor for the receiver of a deferred decision procedure.

The functional form of the a priori phase distribution contains two parameters, and by proper choice of these parameters, we can model unimodal phase distributions which include the uniform distribution and phase known exactly as special cases. The form of the phase distribution was chosen (1) to include the uniform distribution, and (2) so that the functional form of the distribution did not change with observations of a stable SKEP in added WGN, i. e., the phase distribution is closed under normal observations.

The only distribution closed under normal observations and including the uniform distribution on the interval $[0, 2\pi]$ is that given by Eq. 3.1. This can be seen by calculating the a posteriori distribution that results from updating a uniform a priori distribution. This

distribution is the phase distribution obtained by an ideally consistent person taking measurements of a stable SKEP in added WGN assuming, a priori, all phases are equally likely.

$$f(\phi) = \frac{\exp\left[A_0 \cos (B_0 - \phi)\right]}{2\pi\left[I_0(A_0)\right]} , \quad 0 \leq \phi \leq 2\pi$$

$$= 0 , \quad \text{otherwise} \qquad (3.1)$$

The parameter $A_0$ is a measure of the dispersion of the distribution about the parameter $B_0$. For large $A_0$, the distribution approaches a Gaussian distribution, the variance of the distribution, in this case, being $A_0^{-1}$. As $A_0$ varies from zero to infinity, the distribution varies from a uniform distribution on $[0, 2\pi]$ to a phase known exactly. If we plot the half angle spread at one-half the most probable value for $\phi$ as a function of $A_0$, we obtain the graph of Fig. 3.1. Figure 3.1 can be used to measure the dispersion of $\phi$ about $B_0$ as a function of $A_0$ (if one wishes to characterize the phase distribution by a single number). Of course, it is better to present the entire phase distribution rather than a single number. In Fig. 3.2 two distributions are plotted for $p(\phi)$ for values of $A_0$ equal to 2 and 4. Note also that the distribution of Eq. 3.1 offers the advantage of being naturally defined on $[0, 2\pi]$. If a Gaussian phase distribution is assumed, for example, one must make the further approximation of disregarding the tails of the distribution.

Before continuing let us introduce some additional notation to describe the problem. Let $s(t) = a(t) \cos (\omega t - \phi)$ be the transmitted signal and $E = \int_0^T [a(t) \cos (\omega t - \phi)]^2 \, dt$ be the received signal energy. Further, let the receiver input be denoted $z(t)$ and defined for all times t, $0 \leq t \leq T$. The noise is assumed limited to a bandwidth of W cycles per second with a noise power density of $N_o$. The Nyquist sampling theorem (Ref. 16) permits the representation of each receiver input as a point in a 2WT dimensional space; the coordinates of each point are the values of the function at the sampling points $t_i = i/2W$, $1 \leq i \leq 2WT$. The true signal phase is $\phi$ and the true signal frequency is $\omega$.

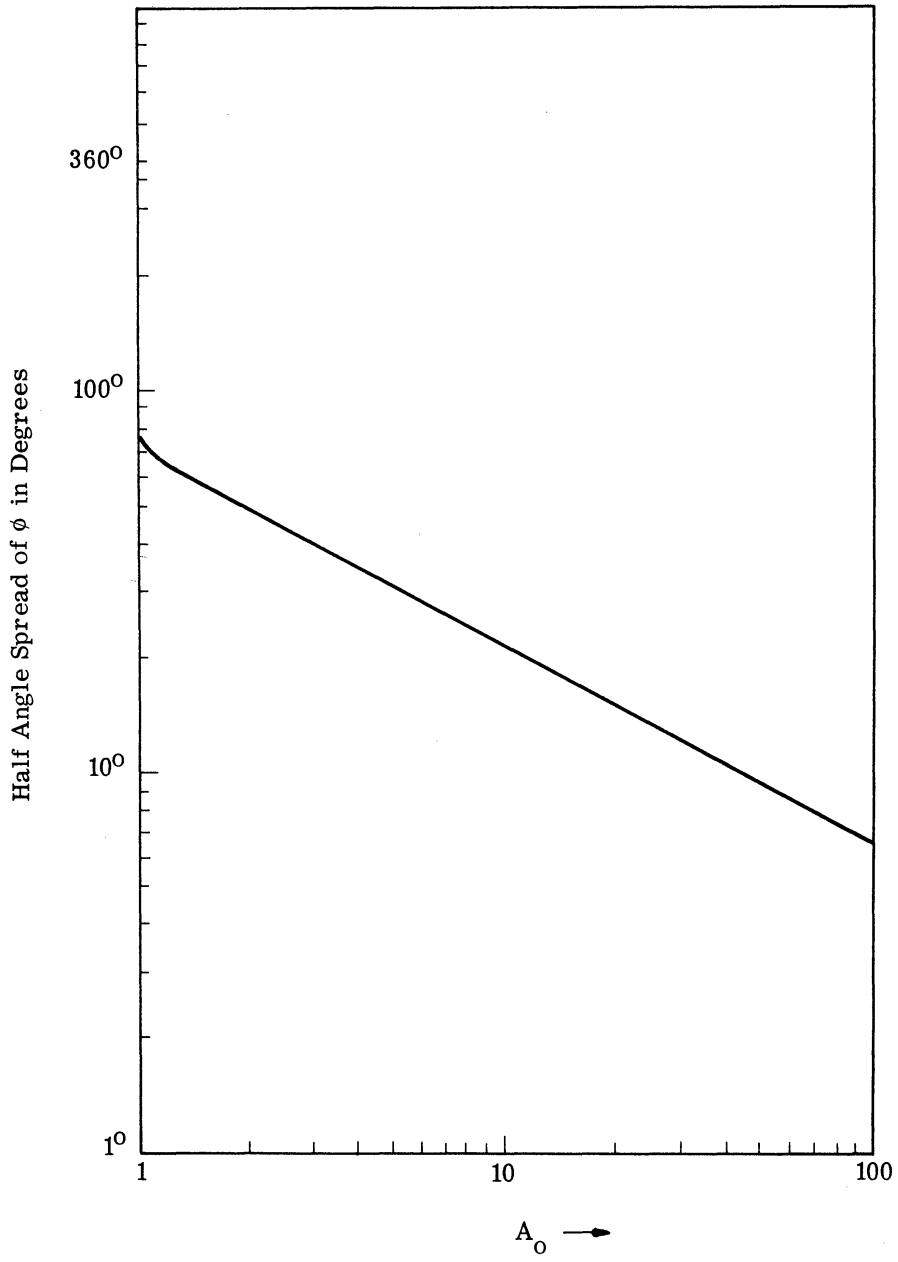


Fig. 3.1. A graph of the half-angle separation for the phase distribution of Eq. 3.1 as a function of the parameter $A_0$.

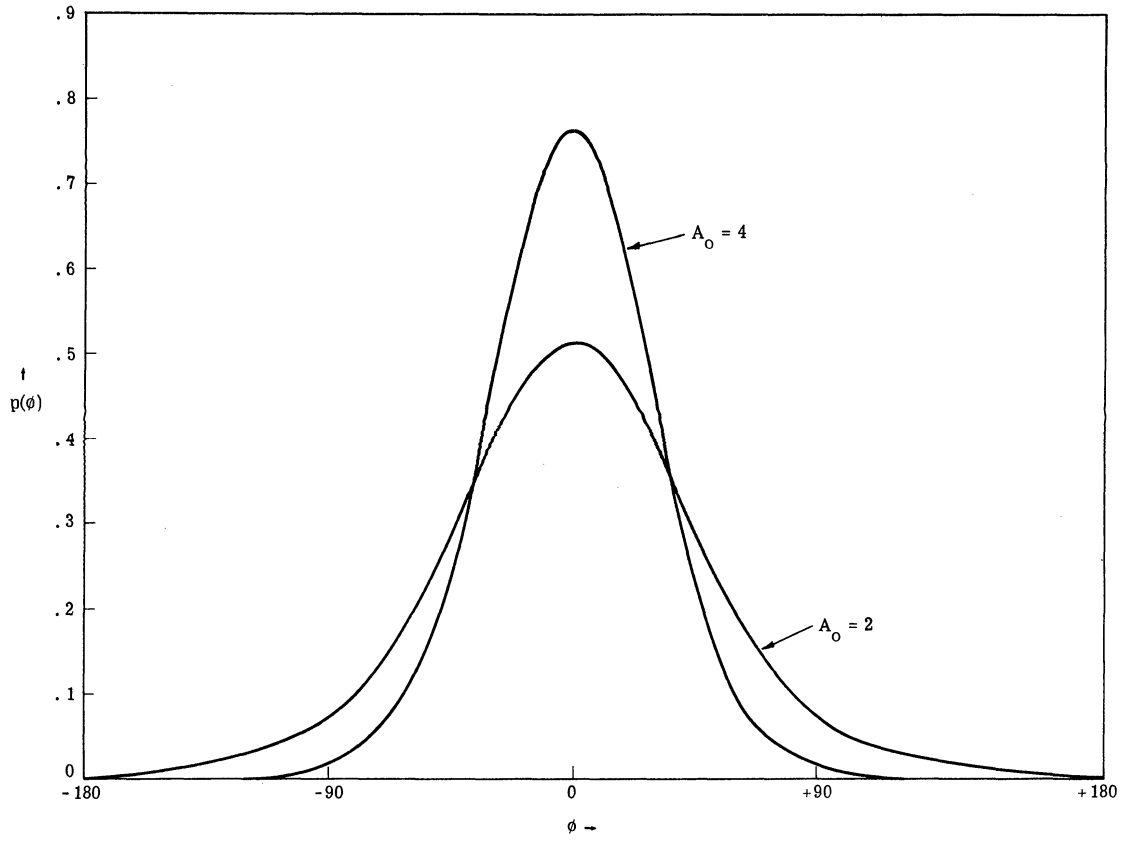


Fig. 3.2. A graph of two phase distributions for two values of the parameter $A_0$.

## 3.2 The Average Likelihood Ratio for a Signal Known Except for Phase with A Priori Phase Information

Optimum receiver design is based on the likelihood ratio of the received waveform z(t). To derive the likelihood ratio of z(t), we use the distributions of z(t) in noise alone and signal-plus-noise. The derivation is made simpler by defining the two quantities, x(t) and y(t).

$$x(T) = \sqrt{\frac{2}{EN_0}} \int_0^T z(t)\, a(t) \cos(wt)\, dt = \frac{1}{W\sqrt{2EN_0}} \sum_{i=1}^{2WT} z(t_i)\, a(t_i) \sin(wt_i) \qquad (3.2)$$

$$y(T) = \sqrt{\frac{2}{EN_0}} \int_0^T z(t)\, a(t) \sin(wt)\, dt = \frac{1}{W\sqrt{2EN_0}} \sum_{i=1}^{2WT} z(t_i)\, a(t_i) \sin(wt_i) \qquad (3.3)$$

The right-hand expressions in Eqs. 3.2 and 3.3 are true by virtue of the sampling theorem. Using the right-hand expressions for x(T) and y(T) the distributions of x(T) and y(T) in N and SN can be found.

Consider first, the distribution of x(T) in noise alone. [The same logic applies to y(T).] Since in noise alone each $z(t_i)$ is due to noise only, each $z(t_i)$ has a normal distribution with zero mean and variance equal to $N_0W = N$ where N is the noise power. The $z(t_i)$ are independent; therefore, the sum $\left(1/W\sqrt{2EN_0}\right)\sum_{i=1}^{2WT} z(t_i)\, a(t_i) \cos(wt_i)$ has a normal distribution with the mean equal to the sum of the means, i.e., zero, and variance equal to the sum of the variances, i.e., $\left(1/W\sqrt{2EN_0}\right)\sum_{i=1}^{2WT} a^2(t_i) \cos^2(wt_i) = 1$. The distributions of x(T) and y(T) are independent because they are orthogonal.

In signal-plus-noise each $z(t_i)$ is due to the noise and the signal. The signal is known except for the carrier phase and the phase is, in general, not uniformly distributed. If the observer knew the transmitted phase, then the sum $\left(1/W\sqrt{2EN_0}\right)\sum_{i=1}^{2WT} z(t_i)\, a(t_i) \cos(wt_i)$ would be normally distributed with mean $(2E/N_0)^{1/2} \cos\phi$ and variance equal to unity. For this problem we define the quality of observation as the square of the separation between the mean value of z(t) in noise alone and in signal-plus-noise. The quality of observation is denoted by D and is equal to $2E/N_0$.

From the above discussion it is possible to determine the joint probability density function of x and y in noise alone and in signal-plus-noise. Because of the independence of x and y, we can write $f(x, y|N)$ as the product of the marginal distributions.

$$f(x, y|N) = f(x|N) \cdot f(y|N) \tag{3.4}$$

$$= \frac{1}{2\pi} \exp\left[-\frac{x^2 + y^2}{2}\right] \tag{3.5}$$

In polar coordinates Eq. 3.5 is

$$f(r, \theta|N) = \frac{1}{2\pi}\, r \exp\left[-\frac{r^2}{2}\right] \tag{3.6}$$

The r multiplying the exponential term in Eq. 3.6 is the Jacobian of the transformation. In signal-plus-noise, if the observation quality is D and the transmitted phase is known to be $\phi$,

$$f(x,\ y|SN,\ \phi) = \frac{1}{2\pi} \exp\left[-\frac{(x - \sqrt{D}\cos\phi)^2 + (y - \sqrt{D}\sin\phi)^2}{2}\right] \tag{3.7}$$

Rewriting Eq. 3.7 in polar coordinates results in Eq. 3.8 below.

$$f(r,\ \theta|SN,\ \phi) = \frac{1}{2\pi}\, r\, \exp\left[-\frac{r^2}{2} - \frac{D}{2} + \sqrt{D}\, r\cos(\theta - \phi)\right] \tag{3.8}$$

Using Eqs. 3.6 and 3.8, the conditional likelihood ratio of the input, $\ell(r,\ \theta|\phi)$, is

$$\ell(r,\theta|\phi) = \frac{f(r,\ \theta|SN,\ \phi)}{f(r,\ \theta|N)} = \exp\left[-\frac{D}{2} + \sqrt{D}\, r\cos(\theta - \phi)\right] \tag{3.9}$$

The average likelihood ratio is found by averaging Eq. 3.9 with respect to a specific distribution for phase. If $f(\phi)$ is given by Eq. 3.1, the average likelihood ratio is

$$\begin{aligned}\ell(r,\ \theta) &= \int_0^{2\pi} \ell(r,\ \theta|\phi)\, f(\phi)\, d\phi \\ &= \int_0^{2\pi} \exp\left[-\frac{D}{2} + \sqrt{D}\, r\cos(\theta - \phi)\right] \frac{\exp[A_0\cos(B_0 - \phi)]}{2\pi\, I_0(A_0)}\, d\phi\end{aligned} \tag{3.10}$$

where $I_0$ is the Bessel function of the zero order and pure imaginary argument. Expanding the integrand of Eq. 3.10 and collecting terms, we obtain

$$\ell(r,\ \theta) = \frac{\exp\left[-\frac{D}{2}\right]}{2\pi\, I_0(A_0)} \int_0^{2\pi} \exp\left[A_1\cos(B_1 - \phi)\right] d\phi = \frac{\exp\left[-\frac{D}{2}\right] I_0(A_1)}{I_0(A_0)} \tag{3.11}$$

The quantities $A_1$ and $B_1$ in Eq. 3.11 are given by

$$A_1^2 = Dr^2 + A_0^2 + 2A_0\sqrt{D}\, r\cos(B_0 - \theta) \tag{3.12}$$

$$B_1 = \tan^{-1}\left[\frac{\sqrt{D}\, r\sin\theta + A_0\sin B_0}{\sqrt{D}\, r\cos\theta + A_0\cos B_0}\right] \tag{3.13}$$

The reduction of the integrand of Eq. 3.10 to the integrand of Eq. 3.11 demonstrates the closure of $f(\phi)$ as given by Eq. 3.1. Recall that the updating equation for $f(\phi)$ is given by Eq. 2.5.

$$f(\phi|r,\ \theta) = \frac{\ell(r,\ \theta|\phi)}{\ell(r,\ \theta)} \cdot f(\phi) \tag{2.5}$$

Substituting in Eq. 3.4, we obtain Eq. 3.14,

$$f(\phi|r, \theta) = \frac{\exp\left[-\frac{D}{2} + \sqrt{D}\, r \cos(\theta - \phi)\right]}{\frac{\exp\left[-\frac{D}{2}\right] I_0(A_1)}{I_0(A_0)}} \cdot \frac{\exp[A_0 \cos(\theta - B_0)]}{2\pi I_0(A_0)} \tag{3.14}$$

Expanding the numerator of Eq. 3.14, we have

$$\exp\left[-\frac{D}{2} + \sqrt{D}\, r \cos(\theta - \phi) + A_0 \cos(B_0 - \phi)\right] = \exp\left[-\frac{D}{2}\right] \exp\left\{\cos\left[\phi(\sqrt{D}\, r \cos\theta + A_0 \cos B_0)\right] + \sin\left[\phi(\sqrt{D}\, r \sin\theta + A_0 \sin B_0)\right]\right\} = \exp\left[-\frac{D}{2}\right] \exp[A_1 \cos(B_1 - \phi)] \tag{3.15}$$

where $A_1$ and $B_1$ are given by Eqs. 3.12 and 3.13. Equation 3.15 is the numerator of Eq. 3.14. Substituting Eq. 3.15 into Eq. 3.14 and reducing it to simplest terms, we obtain

$$f(\phi|r, \theta) = \frac{\exp[A_1 \cos(B_1 - \phi)]}{2\pi I_0(A_1)} \tag{3.16}$$

thereby demonstrating the closure of $f(\phi)$ under normal observations.

## 3.3 Optimum Receiver Design and Evaluation for Fixed-Observation Procedures

3.3.1 Receiver Design, Nonsequential Realization. Receiver design, in a fixed procedure, is the realization of the mathematical expression for the likelihood ratio in block diagram form. Equivalent receiver performance is obtained by basing receiver design on any monotone function of the likelihood ratio. The art in receiver design is the method of obtaining different receiver designs from the same mathematical expression by mathematical manipulations and transformations. These manipulations and transformations serve to indicate many different physical devices capable of the same receiver performance (since they all result from the same mathematical expression). The choice of a particular physical device depends on the designer's experience and ingenuity in realizing block diagrams from a mathematical expression. This is an art best learned by experience.

For example, experience has shown that certain monotone functions of the likelihood ratio point to simple receiver designs, and one such expression useful in optimum receiver design is the logarithm of the likelihood ratio. Receiver design based on the logarithm of the likelihood ratio is, of course, optimum, since the logarithmic function is a monotone

function of its argument. For the problem of a SKEP, the logarithm of the likelihood ratio, Eq. 3.17, can be used as the starting point of the receiver design presented in this section.

$$\ell n\left[\ell(r, \theta)\right] = -\frac{D}{2} + [\ell n\, I_0(A_1)] - \ell n[I_0(A_0)] \qquad (3.17)$$

Equation 3.17 is the mathematical expression whose realization is the optimum detection receiver for a fixed observation-decision procedure which incorporates the possibility of a priori phase information of varying degrees. Since the $I_0$ function is a monotone function of its argument, Eq. 3.17 indicates that $A_1$ is monotone with the likelihood ratio of the input. Therefore, the optimum detector can be realized by calculating the quantity $A_1$. In Fig. 3.3 we schematically represent the quantities in the defining equation for $A_1$, Eq. 3.12.

The operation of the optimum nonsequential detection receiver is quickly obtained with reference to Fig. 3.3. As mentioned above, the optimum receiver may calculate $A_1$ or any monotone function of $A_1$. The calculation of $A_1$ can be accomplished by adding to the x and y components of $A_0$, the x and y components of the reception, respectively. This process is repeated for each additional observation. The pair $(A_1, B_1)$ is modified by the reception in the same manner $(A_0, B_0)$ is modified. The above manner of realizing Eq. 3.17 is achieved by the receiver shown in Fig. 3.4. The reception, z(t), is first multiplied, integrated, and sampled as shown, to obtain x(T) and y(T); to these orthogonal components are added the x and y components of $A_0$; the resulting sum is $A_1$.

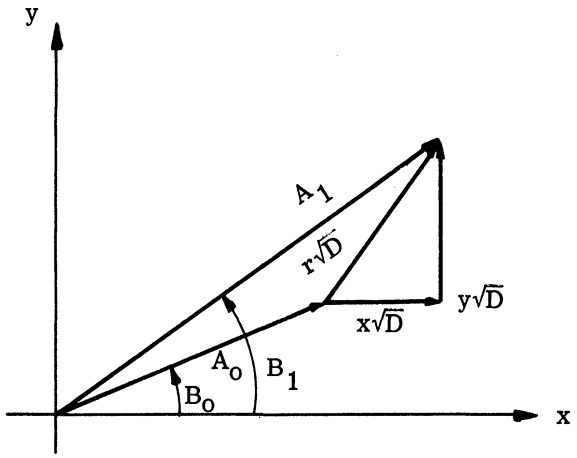


Fig. 3.3. A schematic representation of the terms of Eq. 3.12.

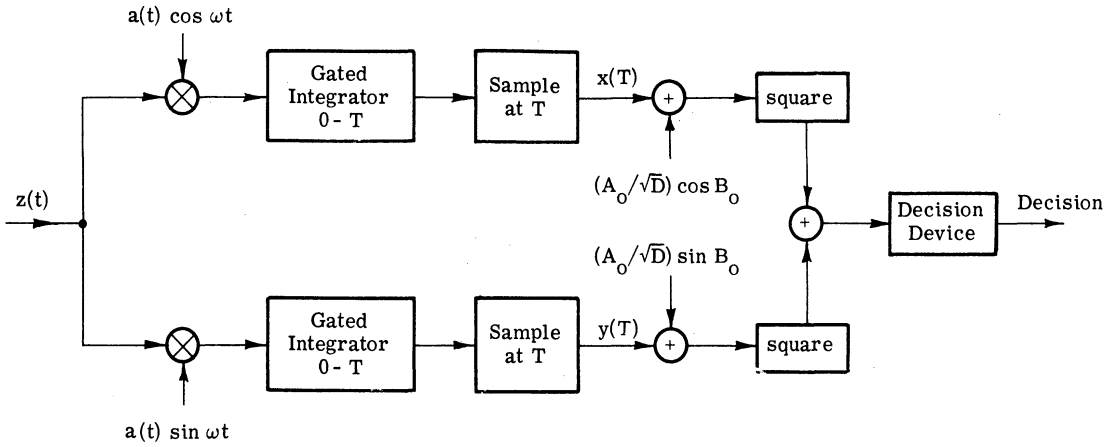


Fig. 3.4. A block diagram of a fixed-observation detector incorporating the a priori phase distribution of Eq. 3.1.

3.3.2 Receiver Design, Sequential Realization. Equivalent realization of the likelihood ratio can be obtained for fixed procedures by breaking the total observation into several smaller sequential observations. The likelihood ratio of the total observation is obtained by updating the likelihood ratio of each smaller observation taken in sequence. This is in contrast to the usual method of considering the total observation as a single reception as was done in the realization presented in Fig. 3.4. The equivalence of the likelihood ratios obtained by these two derivations has been proved in generality by Nolte (Ref. 13).

The receiver for the optimum sequential-observation procedure consists of the sequential realization presented in this section in cascade with a decision device consisting of a set of dual thresholds. If the total observation quality is D and the total quality can be divided into n observations of quality d, such that d = D/n, then for n observations the performance of the sequential realizations are identical to the fixed realization of Fig. 3.4.

Generalizing the results of Eqs. 3.11 and 3.16, the likelihood ratio and phase distribution after k observations of quality d can be written as

$$\ell(r, \theta | \text{k obs.}) = \frac{\exp\left[-\frac{d}{2}\right] I_0(A_k)}{I_0(A_k)} \tag{3.19}$$

$$p(\phi|\,k \text{ obs.}) = \frac{\exp\left[A_k \cos (B_{k-\phi})\right]}{2\pi I_0(A_k)} \tag{3.20}$$

where

$$A_k^2 = A_{k-1}^2 + r^2 d + 2A_{k-1} r\sqrt{d} \cos (B_{k-\phi}) \tag{3.21}$$

$$B_k = \tan^{-1}\left[\frac{\sum_{}^{k} y_i + A_0 \sin B_0}{\sum_{}^{k} x_i + A_0 \cos B_0}\right] \tag{3.22}$$

Schematically, we can represent the quantities in Eqs. 3.21 and 3.22 as shown in Fig. 3.5.

Equation 3.19 is the basis of the sequential realizations of the likelihood ratio presented in Figs. 3.6 and 3.7. The sequential receiver design of Fig. 3.6 is similar to the nonsequential version of Fig. 3.4. Sequential operation and a priori phase information have added an integrator and summer, as shown in Fig. 3.6. The adaptive version of Fig. 3.7 is designed so that it appears to "learn" the signal phase as observations are taken by updating the values of $A_k$ and $B_k$ according to Bayes' Rule. This design offers the observer a classification output, i.e., the receiver not only decides on the presence or absence of a signal but also presents the observer a distribution of the transmitted phase angle.

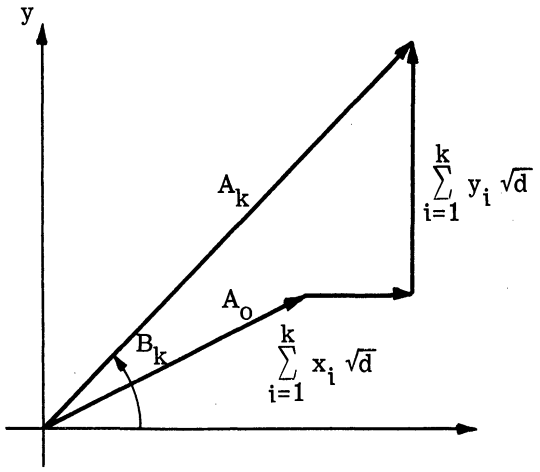


Fig. 3.5. A schematic representation of the terms of Eqs. 3.21 and 3.22.

Although this design appears to learn the signal phase, its performance is identical to the sequential realization of Fig. 3.6 (and the receiver of Fig. 3.4, if the total observed quality is the same in each case). The sequential design of Fig. 3.7 also offers the observer a method by which he can determine phase dispersion as a function of the observation time

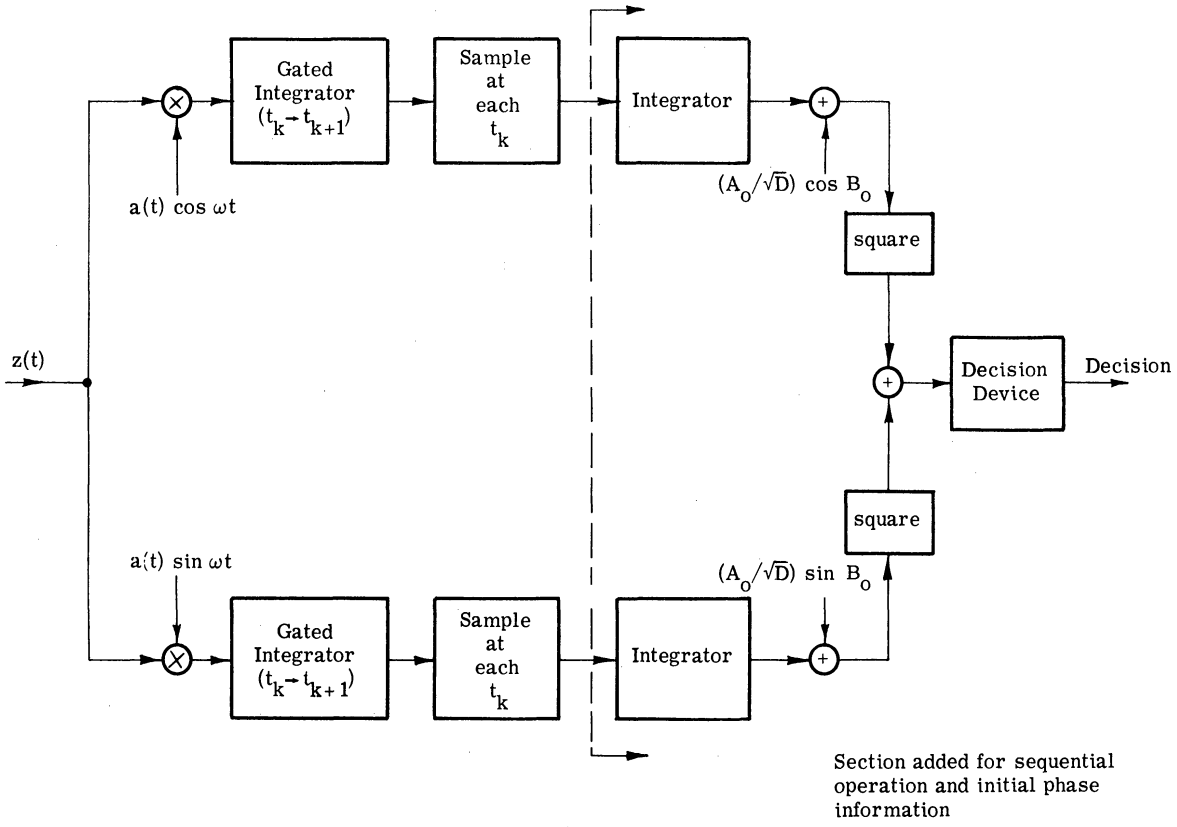


Fig. 3.6. A block diagram of a sequential realization of a fixed-observation detector for a SKEP.

and input signal-to-noise ratio. This is achieved by calculating $A_k$ from the $x_i$ and $y_i$, $i = 1, 2, \ldots, k$, and using the graph of Fig. 3.1 to determine the phase angle dispersion.

The above discussion points out the art of receiver design. There is no optimum receive realization, since many equivalent receivers, all different from the standpoint of equipment, result in exactly the same performance. The particular receiver realization one chooses depends on the engineering application.

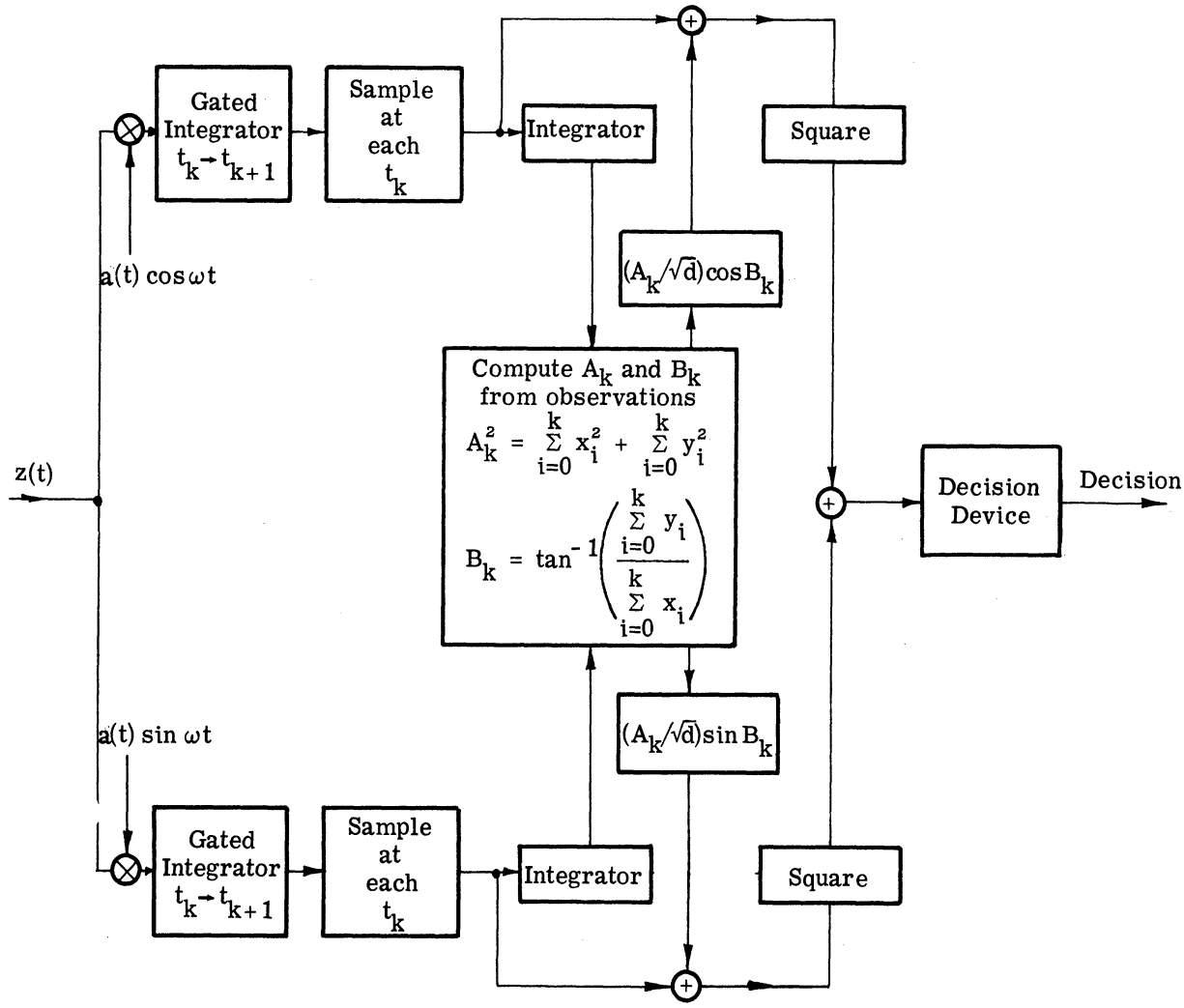


Fig. 3.7. A block diagram of an adaptive realization of a fixed-observation detector for a SKEP.

3.3.3 Evaluation of the Optimum Detection Receiver in a Fixed-Observation Procedure. The evaluation of detection receivers is, usually, considerably more difficult than their design. For fixed-observation procedures, the evaluation can be completely described by the error performance of the receiver (see Section 1.1). Error performance is conveniently displayed by means of the ROC curves presented in Fig. 3.8. There are two parameters, D and $A_0$, associated with each ROC curve. For different amounts of initial phase information, i.e., different values of $A_0$, the quality of observation, D, is varied. From the ROC in Fig. 3.8, we see that the performance of the optimum fixed receiver is within one db of the performance of the optimum receiver detecting a signal known exactly

for $A_0$ equal to four or more. $A_0$ equal to four corresponds to, roughly, a half-angle spread of $33^o$. In other words, we need know only the quadrant the phase angle lies in to approximate the detection performance of a phase known exactly.

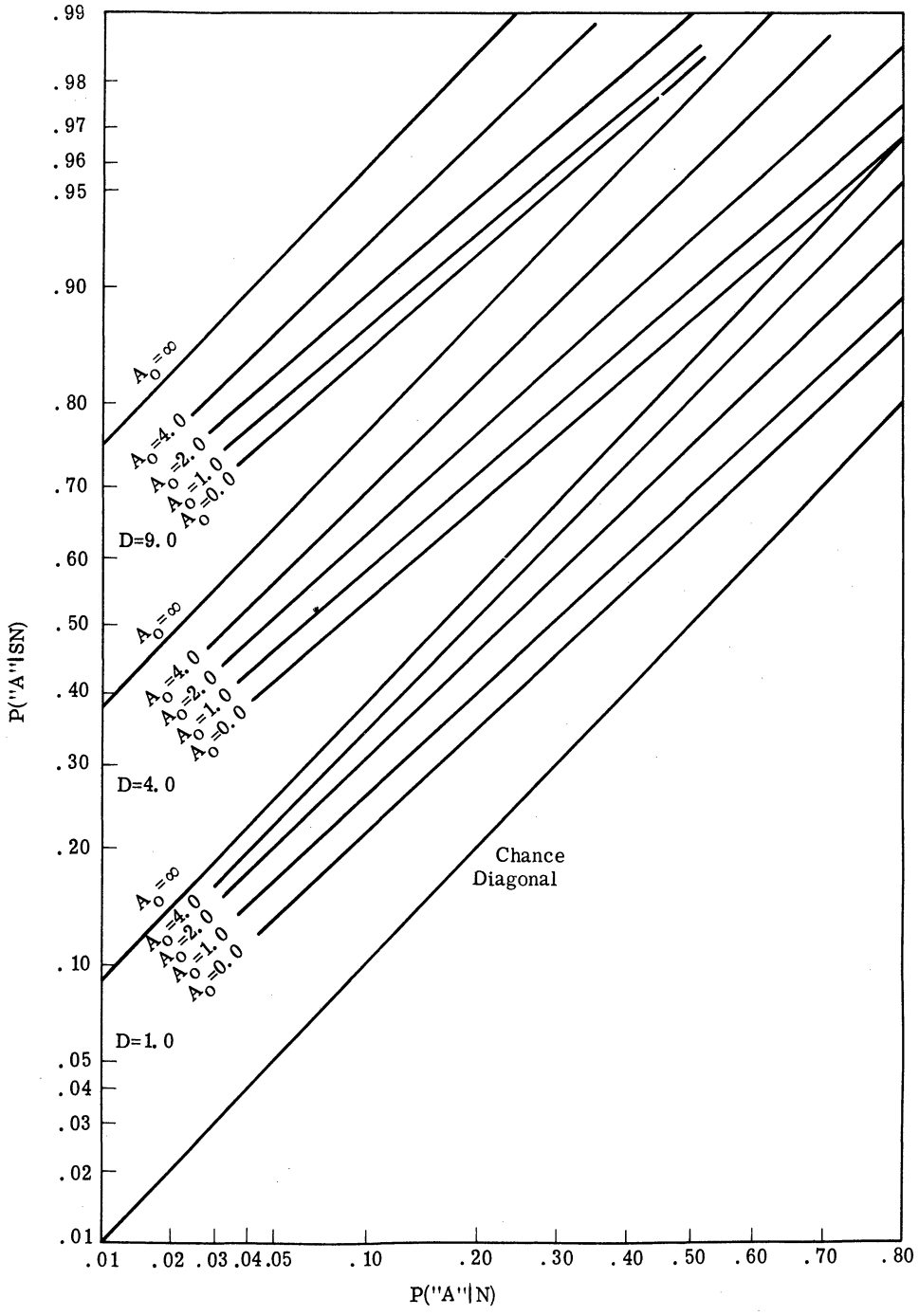


Fig. 3.8. The ROC curves for the optimum fixed-observation detector as a function of a priori phase dispersion and observation quality.

The design method we have used in incorporating a distributed signal parameter in the design of the optimum detector is, to the author's knowledge, unique. The usual design method for considering distributed signal parameters is to estimate the signal parameter and design a fixed receiver with the estimated value assumed to be "true." This method of receiver design is suboptimum, although it is a reasonable engineering solution. The design method for including a distributed signal parameter presented here is optimum for fixed procedures. It represents a significant departure from the philosophy of estimation and fixed receiver design.

The remainder of this section is devoted to the derivations of the probability of detection and false alarm for use in plotting the ROC curves of a SKEP, such as those of Fig. 3.8.

Consider first the calculation of the false alarm probability. The probability of false alarm is the probability of responding "A" when noise alone is the actual cause of the input. It is denoted by P("A"|N). The response "A" occurs whenever the likelihood ratio of the input exceeds a critical value, denoted $\beta_0$. In symbols, we can write

$$P(\text{"A"}|N) = P[\ell(r, \theta) > \beta_c|N] \tag{3.23}$$

To calculate the false alarm probability, therefore, we need to determine the distribution of the likelihood ratio in noise alone, or equivalently, the distribution in noise of any quantity monotone in the likelihood ratio.

$A_1$ has been shown to be monotone with the likelihood ratio. Therefore, Eq. 3.23 can be written as

$$P(\text{"A"}|N) = P(R > \beta_c|N) \tag{3.24}$$

where

$$R^2 = \frac{A_1^2}{D} = r^2 + \frac{A_0^2}{D} + \frac{2A_0 r}{\sqrt{D}} \cos(B_0 - \theta) \tag{3.25}$$

The problem of determining P("A"|N) is reduced to finding the distribution of R in noise alone. If we knew the distribution of R in noise alone, then P("A"|N) is given by calculating the amount of the R distribution outside of an arbitrary circle centered at the origin of the x-y

plane with radius equal to $\beta_c$, as pictured in Fig. 3.9. We have previously determined that the joint probability density function of x and y in noise alone is bivariate Gaussian with mean centered at the tip of the $A_0/\sqrt{D}$ vector and variance unity. (The $A_0$ vector represents a priori knowledge of the phase.) Thus, the probability of false alarm is reduced to finding the "volume" under a bivariate Gaussian distribution outside the radius of a circle as shown in Fig. 3.9. The center of the circle is not coincident with the center of the distribution.

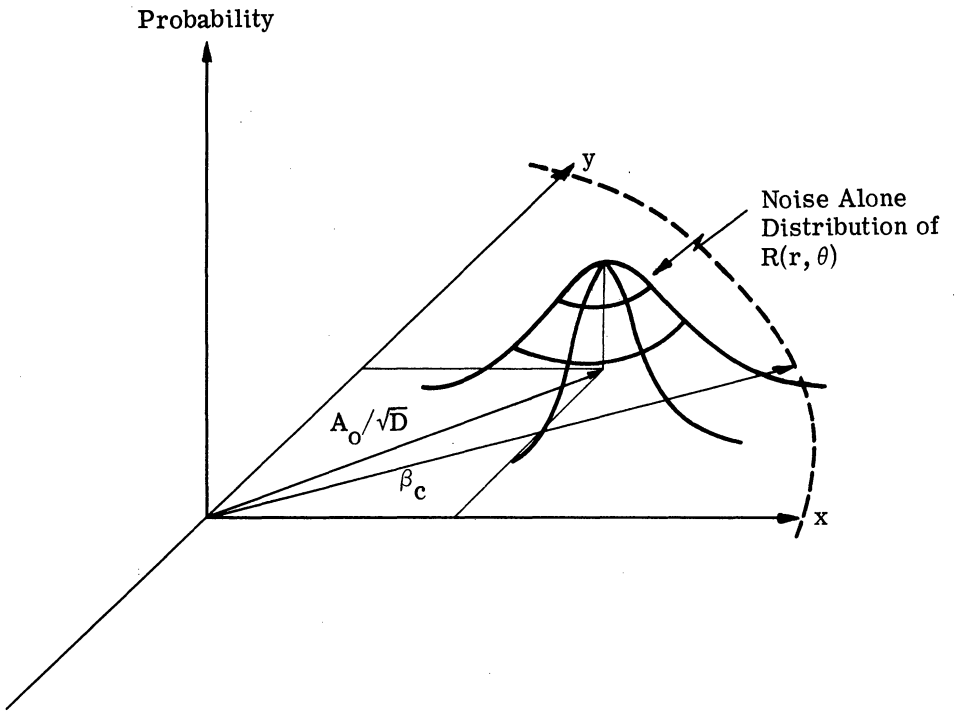


Fig. 3.9. A schematic diagram of the noise alone distribution in the x-y plane.

This is a well-known problem in statistics; the answer is given by the cumulative distribution function of the noncentral chi-squared distribution with two degrees of freedom and noncentrality parameter equal to the distance between the center of the circle and the center of the distribution squared (Ref. 19). For purposes of illustration and completeness, let us derive the false alarm probability using Eq. 3.24.

The distribution of r and $\theta$ in noise alone is given by Eq. 3.6. Thus, the probability that $R(r, \theta)$ is greater than some critical threshold, $\beta_c$, is

$$P[R > \beta_c \,|N] = \int_{\beta_c}^{\infty} \int_{0}^{2\pi} \frac{r}{2\pi} \exp\left[-\frac{r^2}{2}\right] d\Delta dr \tag{3.26}$$

Equation 3.26 represents the "volume" under the bivariate Gaussian distribution outside a circle of radius $\beta_c$; the center of the circle and the distribution are separated by $A_0/\sqrt{D}$ as shown in Fig. 3.9.

The integration of Eq. 3.26 is facilitated by a change of coordinates from (r, $\Delta$) coordinates to (R, $\psi$) coordinates, as shown in Fig. 3.10. Under the transformation, the differential area $d\Delta dr$ becomes $J\left(\frac{r\Delta}{R\psi}\right) d\psi dR$ where $J\left(\frac{r\Delta}{R\psi}\right)$ is the Jacobian of the transformation and is equal to R/r. Rewriting Eq. 3.26 in the (R, $\psi$) coordinate system, we obtain

$$P[R > \beta_c \,|N] = \int_{\beta_c}^{\infty} \int_{0}^{2\pi} \frac{r}{2\pi} \exp\left[-\frac{R^2}{2} - \frac{A_0^2}{2D} + \frac{A_0 R \cos\psi}{\sqrt{D}}\right] \cdot \frac{R}{r}\, d\psi dR$$

$$= \exp\left[-\frac{A_0^2}{2D}\right] \int_{\beta_c}^{\infty} R \exp\left[-\frac{R^2}{2}\right] I_0\left[\frac{A_0 R}{D}\right] dR \tag{3.27}$$

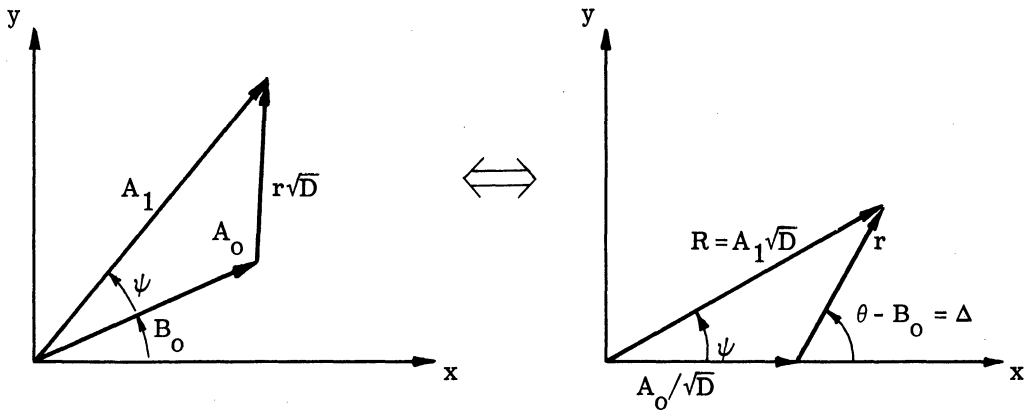


Fig. 3.10. A schematic diagram depicting the change of coordinates carried out in Eq. 3.26.

The integral of Eq. 3.27 arises in the evaluation of the detection of a signal of unknown phase and was defined by Rice (Ref. 20), and tabulated by Marcum in an unpublished report of the Rand Corporation (Ref. 21). It is called the Q function and is a function of two

parameters: (1) the radius of the circle, $\beta_c$, and (2) the separation between the center of the circle and the center of the noise distribution, $A_0/\sqrt{D}$. Equation 3.27 written in terms of the Q function is

$$P["A"|N] = P[R > \beta_c|N] = Q(A_0/\sqrt{D}, \beta_c) \tag{3.28}$$

The derivation of the corresponding detection probability, P("A"|SN), is derived in a similar manner. The probability of detection is the probability of responding "A" when the true cause of the input is due to signal-plus-noise. The response "A" occurs whenever the likelihood ratio of the input exceeds some critical value, $\beta_c$. In symbols

$$P["A"|SN] = P[\ell(r, \theta) > \beta_c|SN] \tag{3.29}$$

Equation 3.29 can be written, using the same logic as we used in the determination of P["A"|N], as

$$P["A"|SN] + P[R > \beta_c|SN] \tag{3.30}$$

where R is defined by Eq. 3.25. As before, the basic problem is the determination of the distribution of R in SN. Knowing the distribution of R in signal-plus-noise, we find the probability of detection by calculating the volume under the R distribution (not bivariate Gaussian) outside a circle of radius $\beta_c$.

To calculate the detection probability, we proceed as we did in the calculation of the false alarm probability. R is a function of r and $\theta$ whose distribution in signal-plus-noise given the transmitted phase, is given by Eq. 3.8. Thus, the probability that $R(r, \theta)$ is greater than some critical threshold, $\beta_c$, is

$$P[R > \beta_c|SN] = \int_{\beta_c}^{\infty} \int_0^{2\pi} \int_0^{2\pi} f(r, \theta|SN, \phi)\, d\phi\, d\Delta dr \tag{3.31}$$

Substituting the expressions for $f(r, \theta|SN, \phi)$ and $f(\phi)$ and carrying out the $\phi$ integration, we obtain

$$P[R > \beta_c|SN] = \int_{\beta_c}^{\infty} \int_0^{2\pi} \frac{r}{2\pi} \exp\left[-\frac{r^2}{2}\right] \exp\left[-\frac{D}{2}\right] \frac{I_0(R\sqrt{D})}{I_0(A_0)}\, d\Delta dr \tag{3.32}$$

If we again write Eq. 3.32 in the (R, $\psi$) coordinate system, then we obtain

$$P[R > \beta_c | SN] = \int_{\beta_c}^{\infty} \int_{0}^{2\pi} \frac{r}{2\pi} \exp\left[-\frac{1}{2}R^2 + \frac{A_0^2}{D} - \frac{2A_0 R\cos\psi}{\sqrt{D}}\right] \exp\left[-\frac{D}{2}\right] \frac{I_0(R\sqrt{D})}{I_0(A_0)} \frac{R}{r}\, d\psi dR \tag{3.33}$$

$$= \frac{\exp\left[-\frac{A_0^2}{2D}\right] \exp\left[-\frac{D}{2}\right]}{I_0(A_0)} \int_{\beta_c}^{\infty} R \exp\left[-\frac{R^2}{2}\right] I_0\left(\frac{A_0 R}{\sqrt{D}}\right) I_0(R\sqrt{D})\, dR \tag{3.34}$$

Equation 3.34 defines a new function of three variables we designate by $R(\alpha, \gamma, \beta_c)$. This function is examined in detail in Appendix A.

$$R(\alpha, \gamma, \beta_c) = \int_{\beta_c}^{\infty} \frac{\exp\left[-\frac{\alpha^2}{2}\right] I_0(\alpha t) \exp\left[-\frac{\gamma^2}{2}\right] I_0(\gamma t)}{I_0(\alpha\gamma)}\, t \exp\left[-\frac{t^2}{2}\right] dt \tag{3.35}$$

The probability of detection corresponding to a false alarm probability of $Q(A_0/\sqrt{D}, \beta_c)$ is $R(A_0/\sqrt{D}, \sqrt{D}, \beta_c)$ where $A_0/\sqrt{D}$ is the separation between the noise alone distribution and the center of the "threshold circle" of radius $\beta_c$ and D is the quality of an observation. As shown in Appendix A, the quantitative evaluation of $R(\alpha, \gamma, \beta_c)$ is possible only by use of numerical integration. In the calculation of both $Q(\alpha, \beta_c)$ and $R(\alpha, \gamma, \beta_c)$ various approximations are made for large and small arguments to facilitate calculations. These approximations are given in Appendix B. The average computing time on an IBM 7090 is approximately 20 seconds per ROC curve as plotted in Fig. 3.8.

## 3.4 The Optimum Nonsequential Observation-Decision Procedure

3.4.1 Analytic Derivation. In the preceding sections of this chapter, we have considered fixed-observation procedures in which the observation time (or quality) is fixed independently of the observer. This is, as we have mentioned, the framework of classical signal detection theory. We have generalized previous studies of a SKEP in a fixed-observation-decision procedure by: (1) assuming the phase distribution is taken from a class of distributions which include the uniform distribution and phase known exactly as special cases, and (2) considering receiver realizations which are sequential in nature. Evaluations of the

optimum receiver for a SKEP in a distributed phase environment have also been presented. In this section we study another form of observation and decision: the optimum nonsequential-observation-decision procedure for a SKEP.

The optimum nonsequential-observation-decision procedure, hereafter called the optimum nonsequential procedure (ONP) is a fixed-observation process optimized, in the sense of minimum expected loss for a terminal decision, by the correct choice of the observation length. The observation length is given in terms of the more general quantity of quality rather than time for the SKEP problem. The choice of the optimum observation quality is based on all information relevant to the detection problem known before the start of the actual receiver reception. The relevant information consists of: (1) the a priori distribution of signal phase, (2) the a priori log-odds ratio, (3) the cost of observation per unit quality, (4) the losses due to terminal decision errors, (5) the observation statistics, and (6) the total available quality of the observation.

Receivers for the ONP are similar to receivers used in classical fixed procedures. The fixed procedure is optimized by correct operation of the receiver, and optimum operation is accomplished by: (1) correct choice of the observation length, and (2) observing only for a specified set of a priori log-odds ratios dependent on the ratio of error loss to observation cost.

The method for determining the optimum observation length is straightforward. We express the value of an observation in terms of the parameters $\{A_0, B_0; W, \Delta_0, L, c, D\}$ where $(A_0, B_0)$ are the parameters of the a priori phase distribution, $(W, \Delta_0)$ represent the losses due to a false alarm and a miss, L is the a priori log-odds ratio, c is the cost of an observation of quality one, and D is the total available observation quality. From the expression for the value of an observation, the maximum value due to an observation in terms of the observed quality is found by differentiating the value expression with respect to D (treated as a continuous variable), setting the resultant expression equal to zero, and solving the differentiated expression for D.

The value of an observation is defined as the amount to be gained observation. It is the optimum terminal loss function, T(L), minus the expected look-ahead loss function, $G(L, D_0, A_0)$. (We have chosen to suppress the dependence of G on the error losses and observation cost.) $D_0$ is the actual observed quality, L is a priori log-odds, and $A_0$ is the

dispersion parameter of the a priori phase distribution. In symbols, the value associated with an observation of quality $D_0$ is

$$V(L, D_0, A_0) = T(L) - G(L, D_0, A_0) \tag{3.36}$$

The average loss associated with an observation of quality $D_0$ is

$$G(L, D_0, A_0) = P(SN) \cdot W \cdot P("B"|SN) + P(N) \cdot W \cdot P("A"|N) + c \cdot D_0 \tag{3.37}$$

The functional form of T(L) changes depending on whether L is less or greater than $\Delta_0$. This implies there are two functional forms for the value function, V. Assuming the false alarm loss equals the miss loss, i.e., $\Delta_0 = 0$, we have for $L \leq \Delta_0$

$$\begin{aligned} V(L, D_0, A_0) &= P(SN) \cdot W - G(L, D_0, A_0) \\ &= P(SN) \cdot W \cdot P("A"|SN) - P(N) \cdot W \cdot P("A"|N) - cD_0 \\ &= \frac{e^L}{1+e^L} \cdot W \cdot P("A"|SN) - \frac{1}{1+e^L} \cdot W \cdot P("A"|N) - cD_0 \end{aligned} \tag{3.38}$$

In a similar manner for $L \geq \Delta_0$, we obtain

$$V(L, D_0, A_0) = \frac{e^L}{1+e^L} \cdot W \cdot P("B"|SN) + \frac{1}{1+e^L} \cdot W \cdot P("B"|N) - cD_0 \tag{3.39}$$

The probabilities of detection and false alarm, correct rejection and miss used in Eqs. 3.38 and 3.39, are related as shown in Eqs. 3.40 and 3.41.

$$P("A"|SN) = 1 - P("B"|SN) \tag{3.40}$$

$$P("B"|N) = 1 - P("S"|N) \tag{3.41}$$

From Eqs. 3.40 and 3.41, we see that it is sufficient to consider only the probability of detection and false alarm to calculate the value function in terms of L.

From Section 3.3, we have that when the probability of false alarm is $Q(A_0/\sqrt{D}, \beta_c)$, the corresponding detection probability is $R(A_0/\sqrt{D}, \sqrt{D}, \beta_c)$. Thus, the value associated with observation of a stable SKEP can be written as

$$V(L, D_0, A_0) = \frac{W}{1+e^{-L}} R(A_0/\sqrt{D_0}, \sqrt{D_0}, \beta_c) - \frac{W}{1+e^{L}} Q(A_0/\sqrt{D_0}, \beta_c) - cD_0 \qquad (3.42)$$

To find the optimum observation quality from Eq. 3.42, we solve for $D_0$ after setting the differentiated equation (differentiated with respect to $D_0$) equal to zero. Before we can do the differentiation of Eq. 3.42 with respect to $D_0$, we must express the critical threshold value, $\beta_c$, in terms of the parameters of the problem, i.e., $\{A_0, B_0; L, W, D, c\}$.

The expression for $\beta_c$ in terms of the above parameters can be found by calculating the probability of detection or false alarm and expressing the cut value in terms of the parameters of the problem. Consider the calculation of the probability of detection. The probability of detection is the probability that the a posteriori log-odds ratio is greater than $\Delta_0$ in signal-and-noise, in symbols

$$\begin{aligned} P(\text{"A"}|SN) &= P[L + \ell n\, \ell(r, \theta) \geq \Delta_0 | SN] \\ &= P[\ell n\, \ell(r, \theta) \geq \Delta_0 - L | SN] \end{aligned} \qquad (3.43)$$

Using the equation for likelihood ratio, $\ell(r, \theta)$, we can write Eq. 3.43 as

$$P(\text{"A"}|SN) = P\left\{ \ell n \left[ \frac{\exp\left[-\frac{D}{2}\right] I_0(A_1)}{I_0(A_0)} \right] \geq \Delta_0 - L \Big| SN \right\} \qquad (3.44)$$

where $A_1$ is defined by Eq. 3.12. Solving the expression in Eq. 3.44 for $A_1$, we obtain

$$\begin{aligned} P(\text{"A"}|SN) &= P\left\{ -\frac{D_0}{2} - \ell n\, I_0(A_0) + \ell n\, I_0(A_1) \geq \Delta_0 - L | SN \right\} \\ &= P\left\{ A_1 \geq I_0^{-1} \left\{ \exp\left[ \Delta_0 - L + \frac{D_0}{2} + \ell n\, I_0(A_0) \right] \right\} \Big| SN \right\} \end{aligned} \qquad (3.45)$$

However, we have determined that the probability $A_1/\sqrt{D_0}$ is greater than some quantity is given by $R(A_0/\sqrt{D_0}, \sqrt{D_0}, \beta_c)$ where $\beta_c$ is

$$\beta_c = 1/\sqrt{D_0} \cdot I_0^{-1} \left\{ \exp\left[ \Delta_0 - L + \frac{D_0}{2} + \ell n\, I_0(A_0) \right] \right\} \qquad (3.46)$$

In a similar manner the probability of false alarm, P("A"|N) is

$$P(\text{"A"}|N) = P\left\{A_1 \geq I_0^{-1}\left\{\exp\left[\Delta_0 - L + \frac{D_0}{2} + \ell n\, I_0(A_0)\right]\right\}\middle| N\right\}$$

$$= Q(A_0/\sqrt{D_0},\ \beta_c) \qquad (3.47)$$

where $\beta_c$ is again given by Eq. 3.46.

We are now in a position to differentiate the value function, Eq. 3.42, with respect to $D_0$. This is a very complicated function of $D_0$; the functional dependence of $D_0$ occurs in the integrand of the integrals of the Q and R functions and also in the limits of these integrals ($\beta_c$). Because of the complicated functions involved we used an IBM 7090 to obtain quantiative results.

A convenient graph for displaying the value function and the optimum observation quality is the "contour graph of the value of observation." This is a graph of the quality of observation versus the log-odds ratio with contour curves of constant value. For each W/c and $A_0$ there corresponds a set of value contours. We are primarily interested in the contour of zero value and the optimum observation quality line. The latter is the optimum observation quality as a function of L, for fixed W/c and $A_0$, and results in the minimum expected loss for a terminal decision. This is the observation quality that is obtained by numerical differentiation of Eq. 3.42 with respect to $D_0$. The computer programs used in obtaining the contour graphs are presented in Appendix C.

3.4.2 Numerical Results. For any specific set of parameters, the computer programs given in Appendix C can be used to determine a contour graph of the value of observation. Figures 3.11 and 3.12 depict two contour graphs for W/c = 30 and W/c = 100 The contours shown are the zero value contours for several values of $A_0$, the dispersion parameter of the phase distribution. The approximately horizontal dotted lines represent the optimum observation quality for various $A_0$ values. Any point that falls inside the zero value contour, for a specific set of parameters, represents a profitable observation, and conversely, any point that falls outside the zero value contour represents an unprofitable observation.

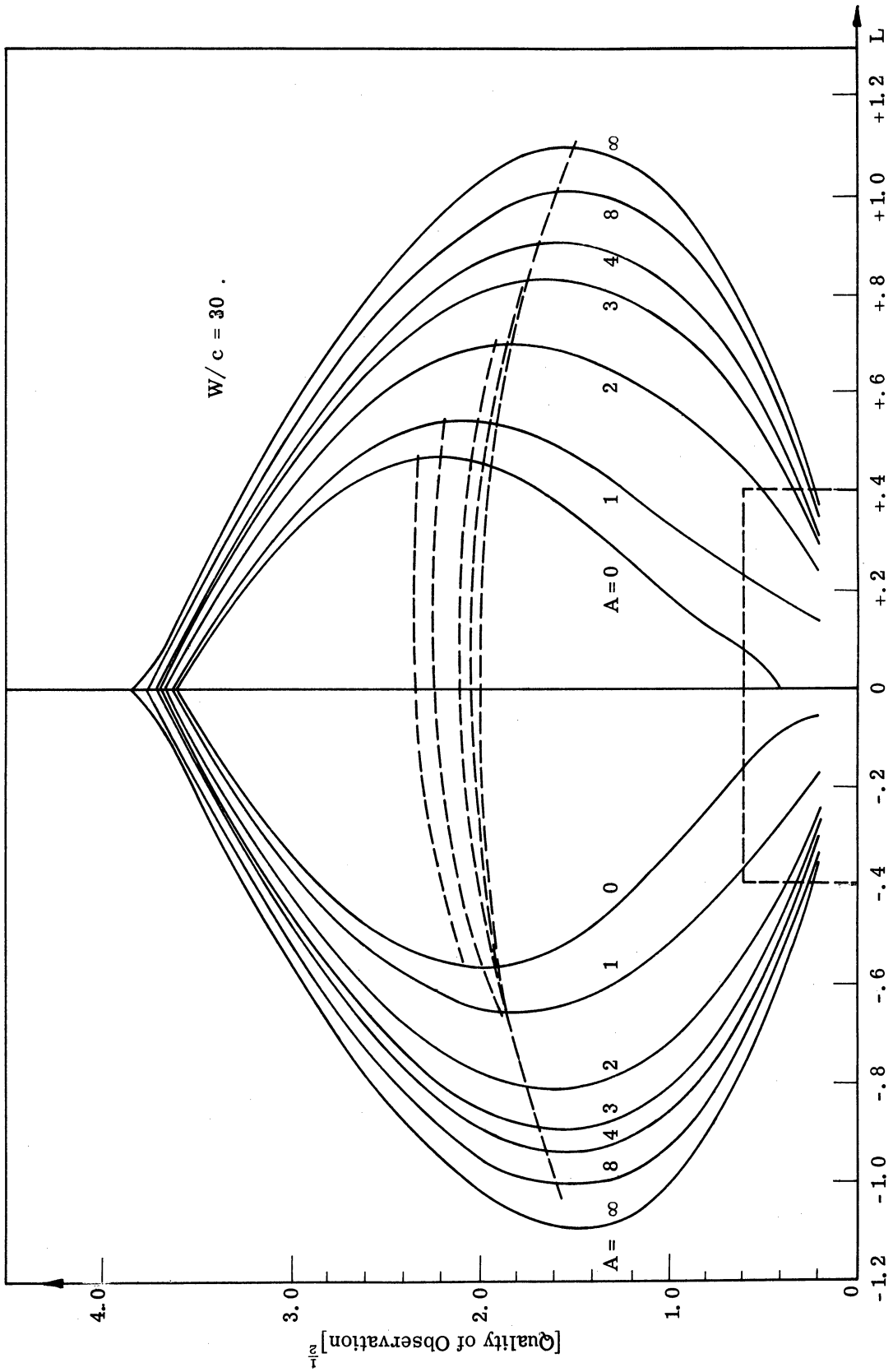


Fig. 3.11. The value contour graph for the ONP for W/c = 30, $\Delta_0 = 0$, and several values of $A_0$.

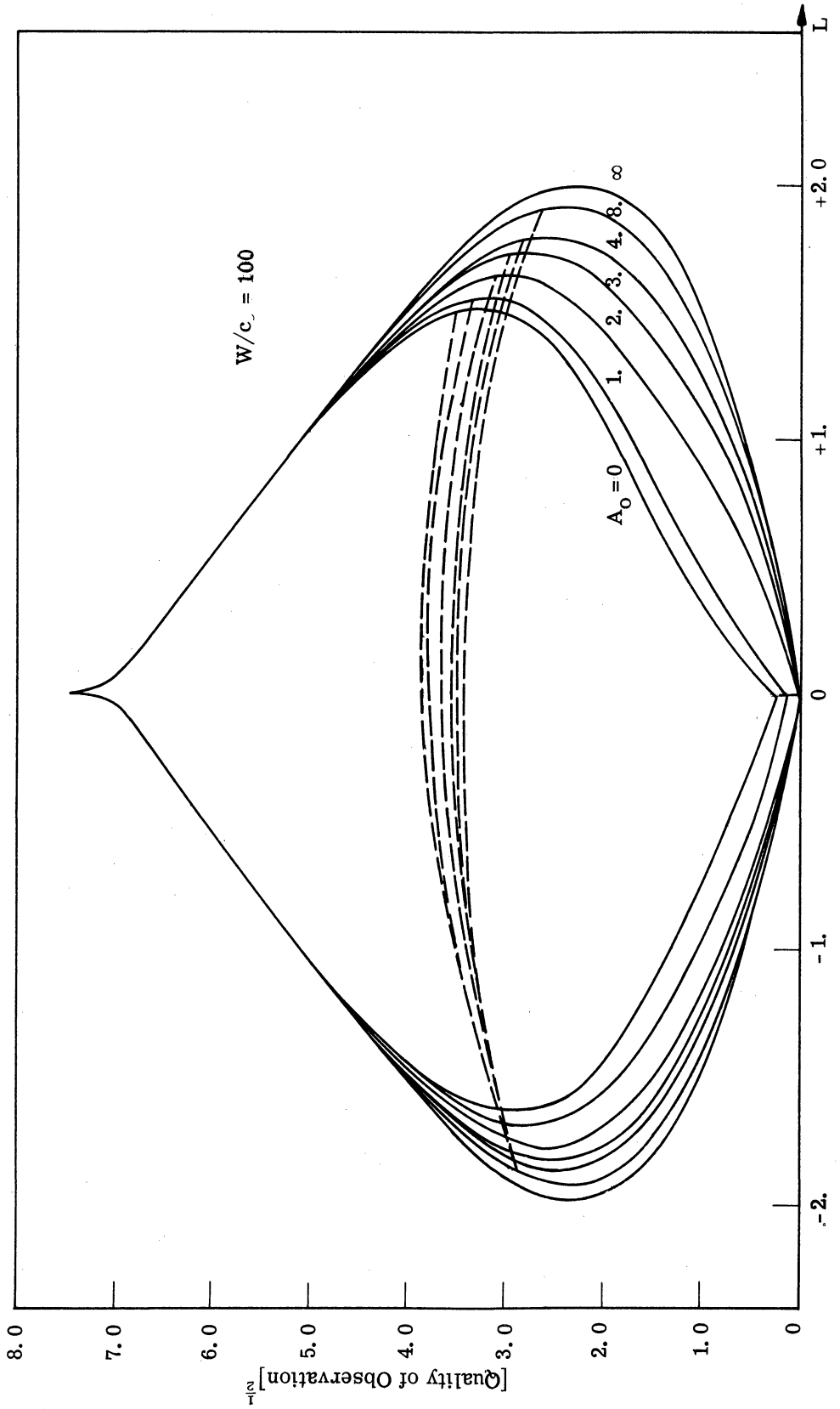


Fig. 3.12. The value contour graph for the ONP for W/c = 100, $\Delta_0 = 0$, and several values of $A_0$.

The computer programs used in the determination of Figs. 3. 11 and 3. 12 are not accurate for large and small values of D simultaneously. Figure 3. 13 is an expansion of the contour graph for small values of observation quality for W/c equal to 30. We are primarily interested in observation quality near the optimum observation quality.

Consider the characteristics of the ONP which can be inferred from the value contour graph for a SKEP. Referring to Fig. 3. 11, one immediate conclusion is the existence of a finite range of a priori log-odds ratios for which an observation is profitable. This range of L values is a function of $A_0$ and the W/c ratio. For example, if W/c = 30 and $A_0 = 4$, then observations are profitable only for $-0.95 < L < 0.90$. For log-odds ratios outside of this range, it is more profitable to make an immediate terminal decision. In Fig. 3. 14 one zero-value contour graph for W/c = 30 and $A_0 = 4$ has been repeated from Fig. 3. 11 to depict various aspects of the ONP with less confusion.

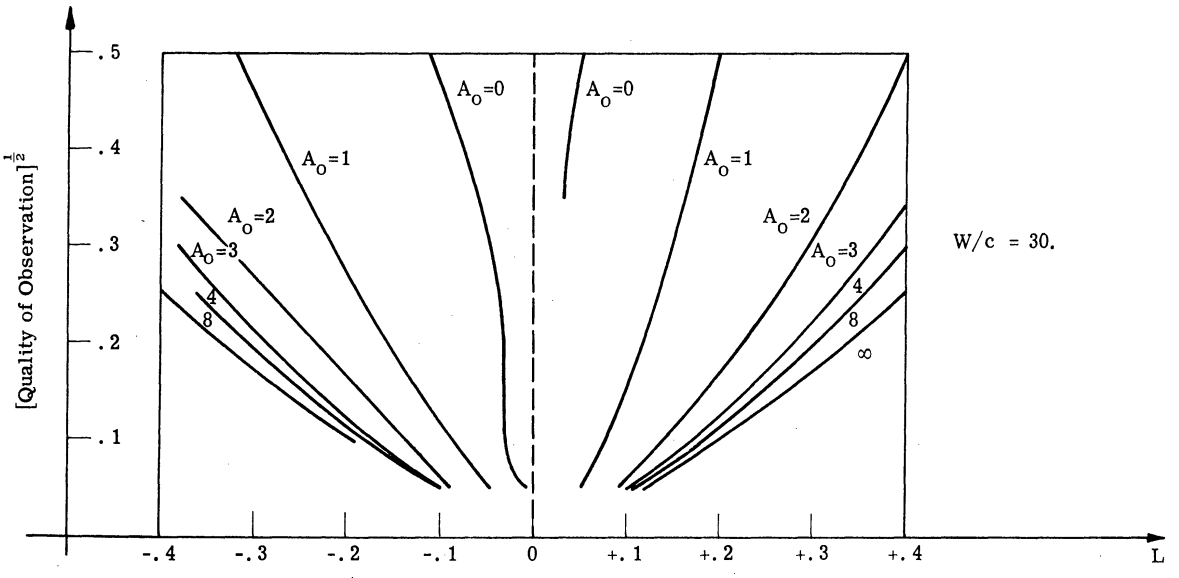


Fig. 3. 13. An expanded section of the value contour graph of Fig. 3. 11.

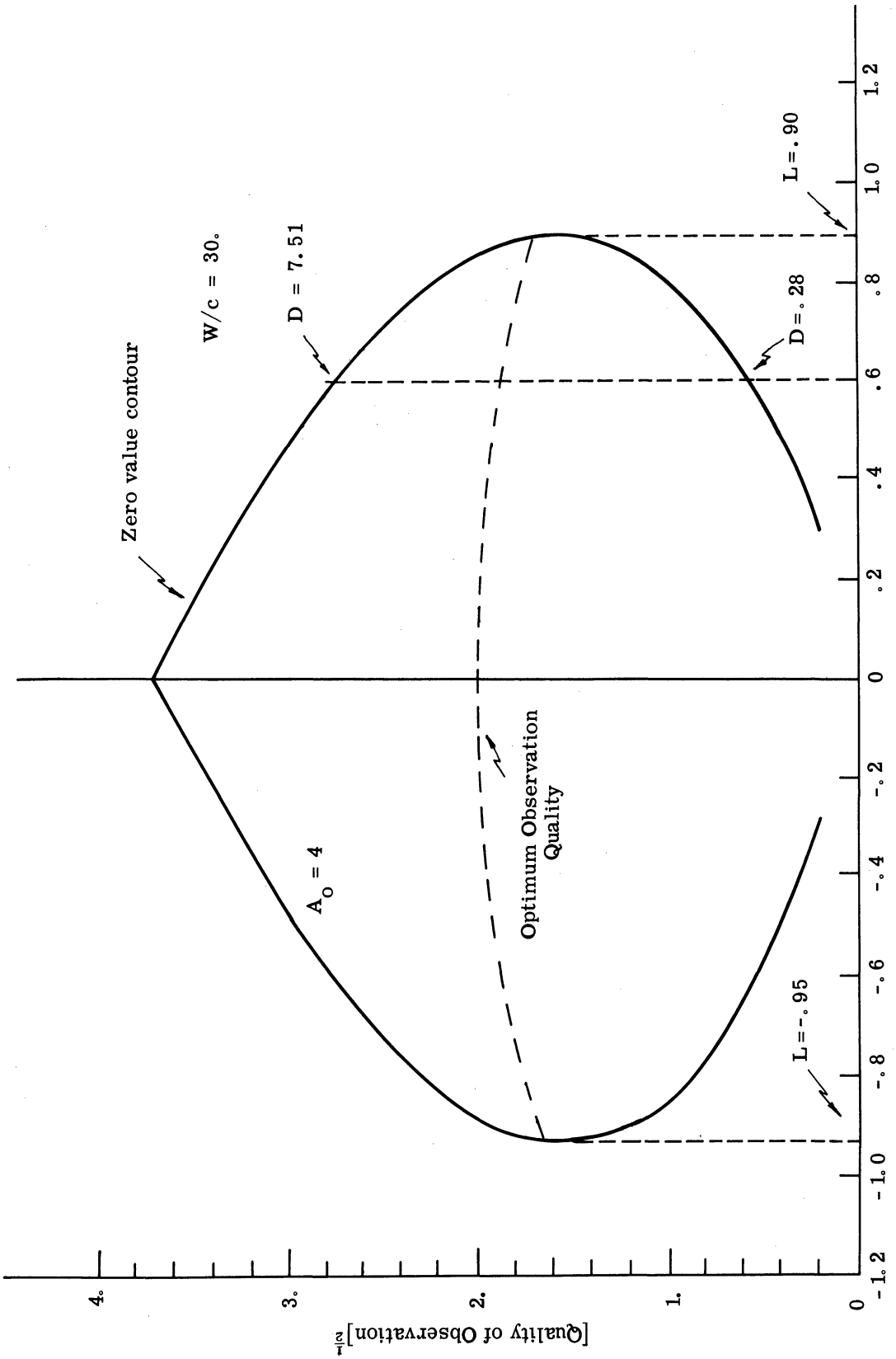


Fig. 3.14. The zero value contour graph for W/c = 30, $\Delta_0 = 0$, and $A_0 = 4$ depicting various aspects of the ONP.

Referring to Fig. 3.14, if the log-odds ratio is equal to 0.6, then we see that it is profitable to observe and the optimum observation quality (the dotted line) is 3.56. An observation quality of 3.56 results in the minimum average loss for a terminal decision; although, any observation quality between 0.28 and 7.51 is profitable, i.e., an observation quality between 0.28 and 7.51 results in a value greater than zero. For the same set of parameters, an observation quality above 7.51 or below 0.28 costs more than it is worth.

Other characteristics of the optimum nonsequential procedure can also be inferred from the value contour graph. The zero value contour for the same $A_0$ value but different W/c ratios expands as the W/c ratio increases. A large W/c ratio means it costs less for the same amount of observation quality; or equivalently, the loss due to a terminal decision is large. Thus, it is profitable to buy larger amounts observation quality than for smaller W/c ratios.

With reference to Figs. 3.11 and 3.12, we also see that as $A_0$ increases, the zero value contour "grows" in both L and D, i.e., as initial phase information becomes more certain, it is more profitable to observe. The outermost zero value contour, labeled "$\infty$", is the phase known exactly zero value contour. This latter curve was not obtained by use of the program in Appendix C but rather from programs used in Ref. 6 for the case of a signal known exactly in added WGN. Since the signal known exactly contour is the most certain case, it should fall outside any zero value contour in which $A_0 < \infty$. The agreement in the gross characteristics of the zero value contour for a signal known exactly compared with the results of SKEP zero value contours serves as a cross-check on the calculations and numerical results presented here.

Again referring to Figs. 3.11 and 3.12, we notice that for $A_0 < \infty$ the zero value contours are all skewed toward negative log-odds ratios. The skewness of the contour graph is also present in Chapter 4 in the study of unknown amplitude signal uncertainty; thus, it is appropriate to explain the skewness of the value contour graph.

The discussion that follows is heuristic in nature; however, because it is consistant for different forms of signal uncertainty it seems most probable. The basic concept of this discussion is that an observation is profitable only for those situations in which the observer's a priori log-odds ratio is likely to be changed sufficiently by the observation so that the appropriate terminal decision is not the terminal decision dictated by the original L.

For example, in the case of a SKEP the value contour is skewed toward negative L's and away from positive L's. From this we infer that for the same absolute value of log-odds ratio, all other parameters being equal, we are more likely to obtain a positive indication of signal-and-noise than noise alone. In other words, we are more likely to change our opinion of the cause of input if we think, a priori, that the cause of the input is noise alone rather than signal-and-noise. For $A_0 < \infty$ is there a reasonable explanation of why there is a better indication of signal-and-noise? (Notice for $A_0 = \infty$, phase known exactly, the value contour graph is symmetric.)

Yes, there is, for in signal-and-noise there exists the possibility of learning the signal phase, thereby increasing the detectability of the signal in noise. For the same conditions in noise alone no such possibility exists. Thus, for comparable a priori log-odds, the observer is more likely to change his opinion for negative a priori log-odds than for positive a priori log-odds. Therefore, the value contour graph is skewed toward negative L's for SKEP in added WGN.

From Figs. 3.11 and 3.12, we see that the optimum observation quality in nearly independent of L, in the continue region,[1] for all values of $A_0$. It is also evident that as signal uncertainty increases, i.e., as $A_0$ decreases, the optimum observation length also increases, all other parameters being equal. Intuitively, we might expect this to occur; to extract the same amount of "information" from an observation, for greater signal uncertainty, we have to observe for a longer period.

The error performance of the ONP is given by the ROC curves of Figs. 3.15 and 3.16 for W/c = 30. and W/c = 100., respectively. These ROC curves are identical to the ROC curves of the fixed-observation procedure except they are truncated for L's not in the continue region. The curves are labeled as in Section 3.2 with $A_0$, the dispersion parameter of the initial distribution, and D, the quality of observation.

---

[1] "Continue region" is a phrase used to denote the L region for which it is profitable to take an observation.

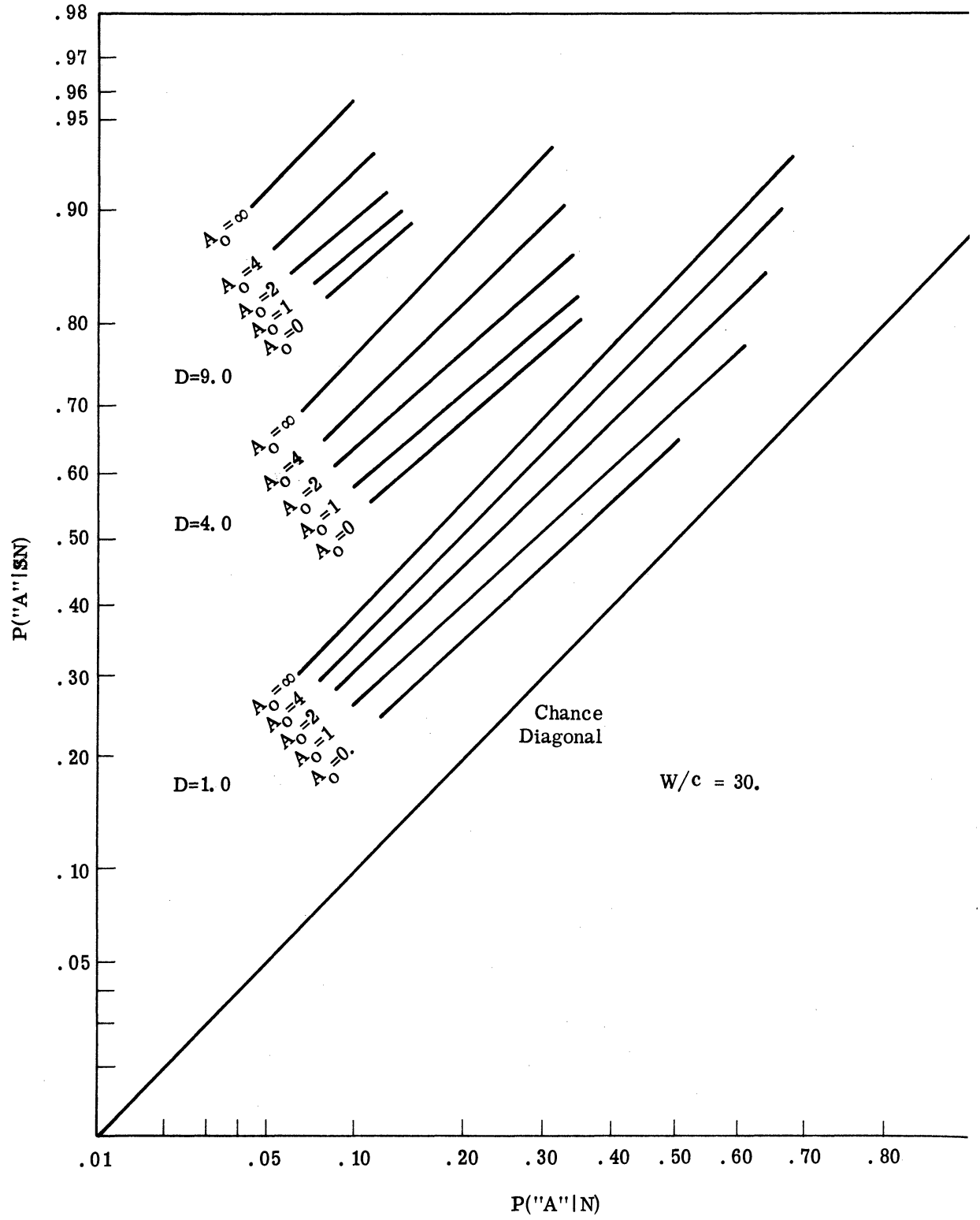


Fig. 3.15. The ROC curves for the ONP for W/c = 30, $\Delta_0 = 0$, and several values of $A_0$.

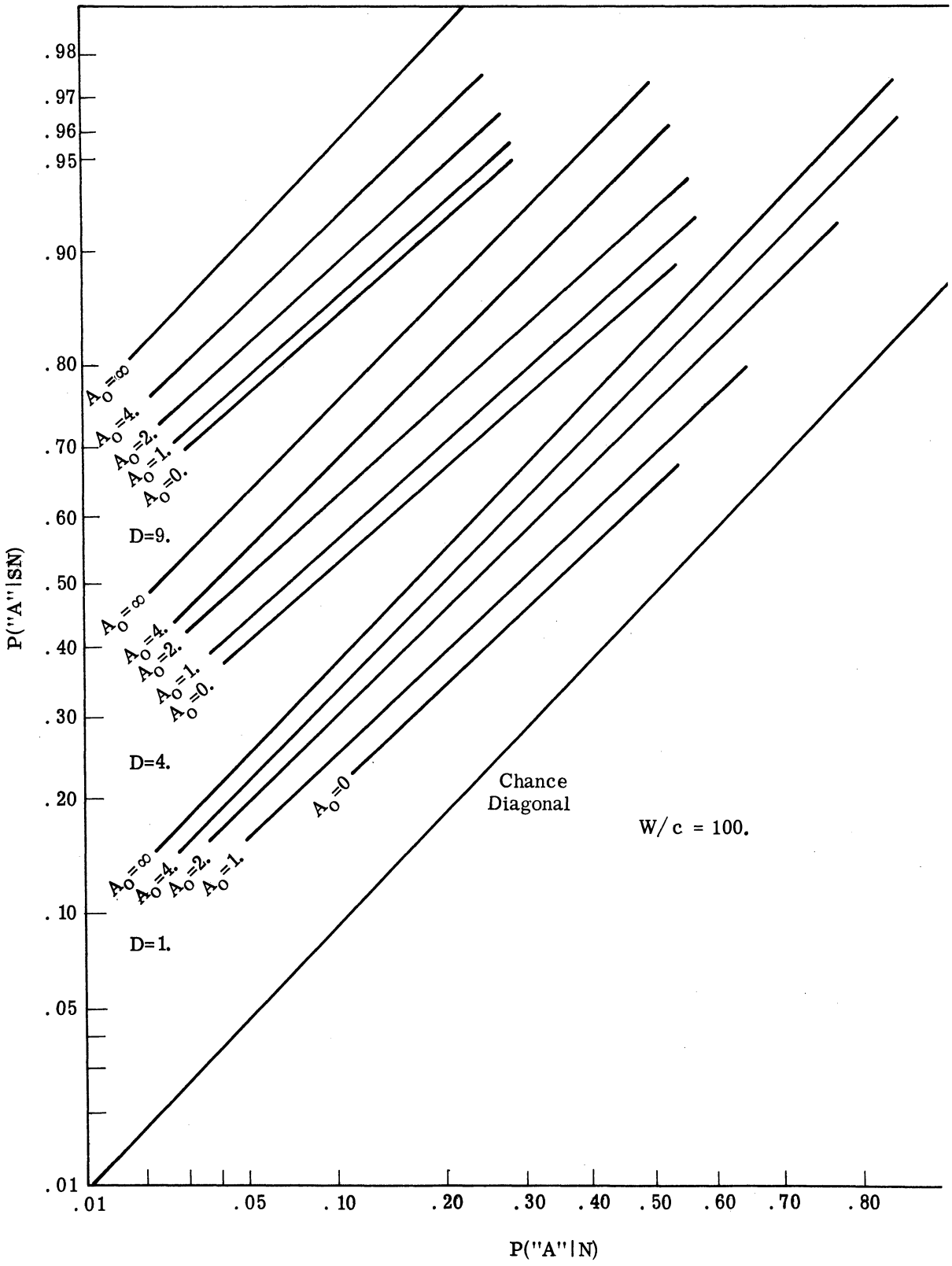


Fig. 3.16. The ROC curves for the ONP for W/c = 100, $\Delta_0 = 0$, and several values of $A_0$.

## 3.5 Derivation of the Deferred Decision Procedure for a Signal Known Except for Phase

In this section we present the analytic derivation of the equations necessary to solve a deferred decision procedure. Quantitative results are not presented because, from an engineering standpoint, the use of deferred decision in the SKEP problem does not appear to be justified. The signal uncertainty that is due to unknown phase is a small uncertainty; thus, the increase in performance of a deferred decision procedure over the ONP for the SKEP problem is approximately the same as the increase in performance of a deferred decision procedure, over the ONP for signal known exactly problem. The amount of computer time necessary to obtain quantitative results in a deferred decision problem is considerable. Since the results of the deferred decision SKEP problem would be approximately those of the signal known exactly problem, which has been solved in great detail in Ref. 6, the numerical calculations of the deferred decision SKEP problem have not been carried out. The equations are derived here for completeness and to demonstrate the subspace separable property as applied to the SKEP problem.

Let us consider first the determination of the separable subspaces and the relationship between the signal parameters in the phase distribution, $\{A_i, B_i\}$ and the log-odds ratio, $L_i$, on the separable subspaces.

Let n be the available number of observations in a deferred decision procedure. By Bayes' Rule, we can write

$$L_{n-1} = L_n + \ell n\, \ell(r, \theta) \tag{3.48}$$

Again using Bayes' Rule, the updating equation for the phase distribution is

$$f_{n-1}(\phi) = \frac{\ell(r, \theta | \phi)}{\ell(r, \theta)}\, f_n(\phi) \tag{3.49}$$

Substituting the expression for likelihood ratio from Eq. 3.49 into Eq. 3.48 and rearranging terms, we obtain

$$L_{n-1} - L_n = \ell n \left[ \frac{\ell(r, \theta | \phi)\, f_n(\phi)}{f_{n-1}(\phi)} \right] \tag{3.50}$$

Solving for $f_{n-1}(\phi)$, we obtain the following general expression used to determine the separable subspaces

$$f_{n-1}(\phi) = \exp[L_1 - L_{n-1}]\, \ell(r, \theta|\phi) \cdot f_n(\phi) \tag{3.51}$$

To find the separable subspaces in the SKEP problem, substitute the expressions for the probability density functions $f_{n-1}(\phi)$ and $f_n(\phi)$ in Eq. 3.51.

$$\frac{\exp[A_{n-1}\cos(B_{n-1} - \phi)]}{2\pi I_0(A_{n-1})} = \exp[L_n - L_{n-1}] \exp\left[-\frac{D}{2} + \sqrt{D}\, r \cos(\theta - \phi)\right] \frac{\exp[A_n \cos(B_n - \phi)]}{2\pi I_0(A_n)} \tag{3.52}$$

The last two numerator terms on the right-hand side of Eq. 3.52 combine to form $[A_{n-1}\cos(B_{n-1} - \phi)]$, cancelling the same term on the left side of the equation. Cancelling the 2 $\pi$'s in the denominators on each side of Eq. 3.52, we can write

$$\frac{1}{I_0(A_{n-1})} = \exp[L_n - L_{n-1}] \frac{1}{I_0(A_n)} \tag{3.53}$$

Taking logarithms of both sides of Eq. 3.53 and rearranging terms, we have

$$\ell n\, I_0(A_n) + L_n = \ell n\, I_0(A_{n-1}) + L_{n-1} \tag{3.54}$$

Equation 3.54 implies that the relationship $\ell n\, I_0(A_k) + L_k$ is independent of the value of k; therefore, we can set $\ell n\, I_0(A_k) + L_k$ equal to a constant Q, thereby defining a fixed relation between $A_k$ and $L_k$ at any stage of the observation in a sequential observation procedure.

$$Q = \ell n\, I_0(A_k) + L_k \tag{3.55}$$

The value of Q, in Eq. 3.55, defines the separable subspaces of the SKEP problem assuming the phase distribution is of the form given by Eq. 3.1. As we vary the values of Q, we vary the relationship between the observer's log-odds ratio (his opinion of the presence of a signal) and the observer's opinion of the phase distribution (represented by the value of $A_0$). Consider two values of Q, one negative and the other positve, for a given phase distribution. A negative value of Q represents a situation in which the observer's opinion of the presence of signal is less than even odds for the given phase distribution. The converse is true for positive values of Q. The latter case represents a situation in which the observer's opinion of signal presence is better than even odds for the given phase distribution.

As explained in Chapter 2, the relationship between $A_k$ and $L_k$ throughout a series of observations in a deferred decision procedure is locked together through the value of the subspace, i.e., in such a manner that Eq. 3.55 is always satisfied. This property is the reason that it is possible to obtain a computing algorithm that is economically feasible to use.

To illustrate the subspace separability property, consider the solution of the SKEP deferred decision problem from two initial states. First, by specifying the initial phase distribution and log-odds ratio and second, by specifying the state by the value of Q and the log-odds ratio. The first method of specifying the state of the decision process is perhaps the more "natural" method. The idea of using a separable subspace to describe the state of a detection problem is a new concept necessitated by the dependent SN observations.

Consider first a deferred decision procedure in which we specify our initial state of knowledge by the log-odds ratio and the phase distribution, i.e., L and A. The solution procedure is as described in Chapter 2. The solution process begins at stage $n_{max} = 0$ for which the optimum expected loss is $F_0(L) = T(L)$.

The solution process continues by solving the $n_{max} = 1$ stage. The a priori log-odds ratio is $L_1$ and the a priori phase $f_1(\phi)$. The questions we must answer are: (1) for a log-odds ratio of $L_1$, should the observation be taken? and (2) what is the average or expected loss if the observation is taken? The average loss if the observation is taken is $G_1(L)$, given by Eq. 3.56.

$$G_1(L) = E\{F_0[L + \ell n\, \ell(r, \theta)]\} + C = \int F_0[L + \ell n\, \ell(r, \theta)]\, f_1(r, \theta)\, dr\, d\theta + C \qquad (3.56)$$

The observation is taken only if $G_1(L)$ is less than $T(L)$. The intersections of $G_1(L)$ and $T(L)$ determine the decision points in L, $(\Gamma_1, \Delta_1)$. The set of decision points, $(\Gamma_i, \Delta_i)_{i=1}^{n}$ constitute the optimum stopping rule for the deferred decision problem. The minimum of $G_1(L)$ and $T(L)$ is denoted by $F_1(L)$ and represents the minimum expected loss for one observation possible. We note that no use is made of the updated density function of phase angle for $n_{max} = 1$ problems.

Consider next the $n_{max} = 2$ problem, i.e., there are at most, two possible observations before a terminal decision must be made. We have an a priori log-odds ratio

and phase distribution, $L_2$ and $f_2(\phi)$, respectively. As before we determine the average loss if an observation is taken, denoted by $G_2(L)$.

$$G_2(L) = E\{F_1[L + \ell n\, \ell(r, \theta)]\} + C = \int F_1[L + \ell n\, \ell(r, \theta)]\, f_2(r, \theta)\, dr\, d\theta + C \qquad (3.57)$$

The observation is taken only if $G_2(L)$ is less than $T(L)$: The optimum average loss function, $F_2(L)$, is

$$F_2(L) = \min[G_2(L),\ T(L)] \qquad (3.58)$$

After one observation, we are in the $n_{max} = 1$ state. The log-odds ratio is now $L_1$ and the phase distribution is $f_1(\phi)$. These quantities are obtained from $L_2$ and $f_2(\phi)$ by Bayes' Rule.

$$L_1 = L_2 + \ell n\, \ell(r, \theta) \qquad (3.59)$$

$$f_1(\phi) = \frac{\ell(r, \theta|\phi)}{\ell(r, \theta)}\, f_2(\phi) \qquad (3.60)$$

If the problem solution is to iterate, after one observation from $n_{max} = 2$ problem, the state of the problem must be mapped into the $n_{max} = 1$, "no-observation-taken" state previously solved. At stage n = 1 we solved the optimum expected loss function given any $L_1$ and a specific $f_1(\phi)$. After an observation from the $n_{max} = 2$ state, the a priori log-odds, $L_2$, is transformed into $L_1$, a random variable dependent on the observation $(r, \theta)$. The a priori phase distribution, $f_2(\phi)$ is transformed into $f_1(\phi)$, or, equivalently, $A_2$ is transformed into $A_1(r, \theta)$, a random variable dependent on the observation $r, \theta$. Because we have solved the $n_{max} = 1$ problem for only one phase distribution, the solution process does not iterate. In order to iterate the solution, the entire $L \times A_1$ space must be known. As discussed in Chapter 2, this is computationally impractical.

Suppose that instead of specifying the initial state in the "natural manner" i. e., by specifying L and the phase distribution, we specify the initial state in terms of the separable subspace, Q and L. (We can obtain the phase distribution that corresponds to this state.)

For $n_{max} = 0$ the optimum terminal loss function is again $F_0(L) = T(L)$.

Consider the solution for the $n_{max} = 1$ problem. The initial state is Q and $L_1$ from which

we can determine the initial phase distribution if we desire. (We need also the initial value of B to express the actual phase distribution.) As before, the average look-ahead loss is $G_1(L)$ and the optimum expected loss is $F_1(L)$, the minimum of $G_1(L)$ and $T(L)$.

Let us proceed now to the inductive step. For the $n_{max} = 2$ problem, the initial state is $L_2$ and Q. If an observation is taken, the average loss is $G_2(L)$.

$$G_2(L) = E\{F_1[L + \ell n\, \ell(r, \theta)]\} + C \tag{3.61}$$

The optimum expected loss is $F_2(L) = \min[G_2(L), T(L)]$. We are now in the n = 1 state (after the observation has been taken). The log-odds ratio is now $L_2 + \ell n\, \ell(r, \theta)$; however, Q is the same since the subspace remains invariant under observation. The new value of $A_2$, i.e., $A_1$, is determined by the subspace specification Q through Eq. 3.55. Notice that $n_{max} = 1$, "no-observation-taken" problem state is identical to the $n_{max} = 2$, "one-observation-taken" state. Thus, the solution process iterates in a one-dimensional subspace, namely, in the L space.

It should be mentioned that the value of B in the phase distribution is never used in the calculations. As we carry through the mathematics, the mode value, B, of the phase distribution is integrated out of every expression in which it appears. This is a reasonable circumstance when we consider that our problem is one of detection rather than identification (or classification). It is the probability of the mode value, represented by A, rather than a particular mode value that is important. This, of course, assumes that the initial value of B assumed is consistent with the observer's opinion of the phase distribution. In order to obtain the phase distribution at each stage of the observation, the value of B must be recorded from observation to observation.

## 3.6 Summary of Chapter III

The problem of detecting a signal of unknown phase has been solved here for the first time for: (1) fixed procedures in which the a priori phase distribution is other than uniform, (2) the optimum nonsequential procedure, including the evaluation of the optimum nonsequential detector, and (3) the deferred-decision procedure. Numerical calculations have been presented for added WGN for (1) and (2) above in addition to several receiver designs. The detailed conclusions of the evaluations and receiver designs are too lengthy to be repeated here; they can be found at the conclusion of the various sections.

# CHAPTER IV

## DETECTION OF A SIGNAL KNOWN EXCEPT FOR AMPLITUDE

We consider in this chapter the detection of a signal known except for amplitude (SKEA) in added WGN. The solution to this problem has been published only for the classical detection situation, i.e., a fixed-observation procedure, assuming the amplitude is distributed according to a Gaussian or Rayleigh distribution, (Ref. 16).

### 4.1 Problem Statement and Notation

4.1.1 Preliminary Remarks. The amount of signal uncertainty can be varied over a much larger range in the detection problem of a SKEA than is possible in the detection of a SKEP discussed in Chapter III. The possibility of using more sophisticated detection receivers for detecting a SKEA may, therefore, result in large increases in detection performance when compared to the simple fixed-observation receiver of classical detection theory. Large increases in detection performance over previous detection receivers do not occur in the SKEP problem, because unknown phase represents a relatively small amount of signal uncertainty.

Previous studies of the detection of a SKEA in noise are generalized here, for the first time, to include (1) optimum receiver design and evaluation for the classical detection problem for three classes of a priori amplitude distributions, (2) sequential realizations of the optimum receiver for fixed-observation procedures, (3) optimum receiver design and evaluation of the optimum nonsequential procedure (ONP), and (4) optimum receiver design and evaluation for the optimum sequential procedure, deferred decision. Comparisons of the ONP detector and the deferred decision detector are also presented. The three classes of signal amplitude distributions considered are discussed below.

4.1.2 The Signal Amplitude Distributions. The functional form of the amplitude distribution was chosen (1) to meet the closure condition under observations in normal added noise, and (2) to model as many different types of signal amplitude uncertainty as possible.

The functional form of the amplitude distribution is given by Eq. 4. 1, and is called the Halsted distribution (Ref. 22).

$$f(a) = M\, p(a) \exp\left[ah - \frac{a^2 g}{2}\right], \quad a \geq 0$$
$$= 0, \quad \text{otherwise} \tag{4.1}$$

The quantity $p(a) = A_0 a^{\alpha} + A_1 a^{\alpha-1} + \ldots + A_{n-1} a + A_n$ is, in general, a polynomial in a. We will confine our study to $p(a) = a^{\alpha}$. M is the normalizing constant for the amplitude density function.

The number of distributions included in the Halsted distribution is large. For example, we have (1) for $\alpha = 0$ and $h = 0$, the Gaussian distribution, (2) for $\alpha = 0$, the truncated Gaussian distribution, (3) for $\alpha = 0$ and $g = 0$, the exponential distribution, and (4) for $a = 1$ and $g = 0$, the Rayleigh distribution.

In applying the general theory to obtain numerical results, we will limit the number of signal amplitude distributions to three classes corresponding to $\alpha = 0$, $\alpha = 1$, and $\alpha = h = 0$ in Eq. 4. 1. The latter distribution, $\alpha = h = 0$, is of interest in itself and as an approximation for signal amplitude distributions for $\alpha \geq 2$. Each class of signal amplitude distribution must be solved as a separate problem. Although the basic concepts in detecting an unknown amplitude signal do not depend on the particular form of the amplitude distribution, the detailed mathematics of the solution does depend on amplitude distribution. Thus, we will present the mathematics for three problems but will concentrate on the truncated Gaussian distribution, i. e., $\alpha = 0$, in the numerical results.

Before continuing, let us derive the normalizing constants for the three classes of distributions and discuss some of the properties of each distribution.

For $\alpha = 0$, the amplitude distribution is

$$f(a) = M \exp\left[ah - \frac{a^2 g}{2}\right], \quad a \geq 0$$
$$= 0, \quad \text{otherwise} \tag{4.2}$$

Since f(a) is a probability density function, we have

$$\int_0^{\infty} f(a)\, da = M \int_0^{\infty} \exp\left[ah - \frac{a^2 g}{2}\right] da = 1 \tag{4.3}$$

Solving Eq. 4.3 for M, we obtain

$$M = \frac{1}{\int_0^\infty \exp\left[ah - \frac{a^2 g}{2}\right] da} = \frac{\sqrt{g}\,\phi\left(\frac{h}{\sqrt{g}}\right)}{\Phi(h/\sqrt{g})} \tag{4.4}$$

where $\phi(\alpha) = (1/\sqrt{2\pi})\exp[-x^2/2]$ and $\Phi(x)$ is as defined in Eq. 1.9. The quantity $\Phi(x)/\phi(x)$ occurs so often in the analysis of the SKEA problem that it is given a new symbol, $\omega(x)$.

$$\omega(x) = \frac{\Phi(x)}{\phi(x)} \tag{4.5}$$

Thus, for $\alpha = 0$, the amplitude distribution is given by

$$f(a) = \frac{\sqrt{g}}{\omega(h/\sqrt{g})} \exp\left[ah - \frac{a^2 g}{2}\right], \quad a \geq 0$$

$$= 0 \quad , \quad \text{otherwise} \tag{4.6}$$

Equation 4.6 can also be written as

$$f(a) = \frac{\sqrt{g}}{\Phi(h/\sqrt{g})}\, \phi\left(\frac{a - h/g}{1/\sqrt{g}}\right), \quad a \geq 0$$

$$= 0 \quad , \quad \text{otherwise} \tag{4.7}$$

Equation 4.7 is a convenient form since it points out that h/g and $1/\sqrt{g}$ are approximately the mean and variance of the distribution, respectively. The actual mean and variance of the distribution, given below in Eqs. 4.8 and 4.9, respectively, are complicated functions of h and g.

$$E(a) = \int_0^\infty a\, f(a)\, da = \frac{1}{\sqrt{g}\ \omega(h/\sqrt{g})} + \frac{h}{g} \tag{4.8}$$

$$E(a^2) = \int_0^\infty a^2 f(a)\, da = \frac{h}{g}\left[\frac{1}{g} + \frac{h\,\omega(h/\sqrt{g})}{g^{3/2}}\right] + \frac{\omega(h/\sqrt{g})}{g^{3/2}} \tag{4.9}$$

We can find the normalizing constant for the $\alpha = 1$ distribution in a similar manner. This results in the probability density function of Eq. 4.10.

$$f(a) = \frac{g}{1 + \frac{h}{\sqrt{g}}\,\omega(h/\sqrt{g})}\; a \exp\left[ah - \frac{a^2 g}{2}\right], \qquad a \geq 0$$

$$= 0 \qquad , \qquad \text{otherwise} \qquad (4.10)$$

The density function for $\alpha = 1$ possesses a first-order zero which means that the probability of zero amplitude signal is zero. For the $\alpha = 0$ density function, the probability of a zero amplitude signal is nonzero. The latter distribution models a situation for a signal amplitude close to zero. For $\alpha \geq 2$, the Halsted distribution possesses higher-order contacts at the origin. This distributions can be effectively modeled by the Gaussian distribution given in Eq. 4.11.

$$f(a) = \sqrt{g}\,\phi\left(\frac{a - h/g}{1/\sqrt{g}}\right), \qquad -\infty < a < \infty \qquad (4.11)$$

Graphs of the signal amplitude density functions for different values of $\alpha$, h, and g are included throughout this chapter as they occur.

4.1.3 Notation. The notation used in this chapter is consistent with notation thus far introduced. The receiver input is denoted z(t) and the cause of z(t) is due to N or SN. In symbols, we write z(t) = n(t) in N and z(t) = a u(t) + n(t) in SN. The quantity u(t) is a unit signal. More generally, u(t) is a normalized autocorrelation, i.e., if a · s(t) is the SKEA and s(t) is a known waveform, then

$$a\,u(t) = \frac{a \int s(t) \cdot s(t)\,dt}{\left|\int s(t) \cdot s(t)\,dt\right|} \qquad (4.12)$$

The quantity a in Eq. 4.12 is distributed according to the Halsted distribution. The receiver input is assumed to be defined for all time t, $0 \leq t \leq T$. The noise is assumed limited to the receiver bandwidth, W cycles per second, with a noise power density equal to $N_0$. The observation, $z_i$, in noise alone are normalized such that each $z_i$ is N(0, 1).

## 4.2 The Average Likelihood Ratio of a Signal Known Except for Amplitude

Optimum receiver design in fixed-observation procedures is based entirely on

the average likelihood ratio[1] of the input observation. In this section, we derive the average likelihood ratio for the three classes of signal amplitude distributions previously discussed. These average likelihood ratios are, in turn, used to obtain optimum receiver designs for fixed-observation problems. The results are also used in the design of optimum sequential detectors.

In general, the average likelihood ratio of a distributed parameter variate is

$$\ell(z) = \int_{\Theta} \ell(z|\Theta)\, f(\Theta)\, d\Theta \tag{4.13}$$

For a SKEA,

$$\ell(z|\Theta) = \ell(z|a) = \exp\left[az - \frac{a^2}{2}\right] \tag{4.14}$$

The right-hand equality of Eq. 4.14 is evident when we note that the likelihood ratio, given the amplitude, is merely the likelihood ratio of a signal known exactly. Thus, the average likelihood ratio for the three classes of amplitude distributions can be found by substituting the appropriate density function for $f(\Theta)$ in Eq. 4.13.

(1) For $\alpha = 0$, we have that the average likelihood ratio is

$$\ell(z) = \int_0^{\infty} \exp\left[az - \frac{a^2}{2}\right] \cdot \frac{\sqrt{g}}{\omega(h/\sqrt{g})} \cdot \exp\left[ah - \frac{a^2 g}{2}\right] da$$

$$= \sqrt{\frac{g}{g+1}} \cdot \frac{1}{\omega(h/\sqrt{g})} \cdot \omega\left(\frac{h+z}{\sqrt{g+1}}\right) \tag{4.15}$$

(2) For $\alpha = 1$, we obtain

$$\ell(z) = \int_0^{\infty} \exp\left[az - \frac{a^2}{2}\right] \cdot \frac{g}{1 + \frac{h}{\sqrt{g}}\, \omega(h/\sqrt{g})} \cdot a \exp\left[ah - \frac{a^2 g}{2}\right] da$$

$$= \frac{1}{1 + \frac{h}{\sqrt{g}}\, \omega(h/\sqrt{g})} \cdot \frac{g}{g+1} \cdot \left[1 + \frac{h+z}{\sqrt{g+1}}\, \omega\left(\frac{h+z}{\sqrt{g+1}}\right)\right] \tag{4.16}$$

---

[1] Any monotone function of the likelihood ratio is equivalent to the likelihood ratio for detection.

(3) For the Gaussian case, we have

$$\ell(z) = \int_{-\infty}^{+\infty} \exp\left[az - \frac{a^2}{2}\right] \cdot \frac{\sqrt{g}}{\sqrt{2\pi}} \cdot \exp\left[-\frac{g}{2}\left(a - \frac{h}{g}\right)^2\right] da$$

$$= \sqrt{\frac{g}{g+1}} \exp\left[-\frac{h^2}{2g}\right] \exp\left[\frac{(h+z)^2}{2(g+1)}\right] \qquad (4.17)$$

## 4.3 Optimum Receiver Design and Evaluation for Fixed-Observation Procedures

4.3.1 Receiver Design, Nonsequential Realization. The basis of optimum receiver design for a SKEA for signal amplitude distributions corresponding to $\alpha = 0$, $\alpha = 1$, and $\alpha = h = 0$ is given by Eqs. 4.15 through 4.17, respectively. Because the functions $\omega(\ )$ and exp [ ] are monotone in their arguments, the average likelihood ratio, for all three signal amplitude distributions considered, is monotone with the quantity h + z. Thus, the optimum detection receiver can be realized by calculating the quantity h + z for the three amplitude distributions considered. One possible nonsequential realization is shown in Fig. 4.1. It should not be surprising that the optimum receiver is independent of the particular amplitude distribution chosen, since the physical problem in all three cases is basically the same.

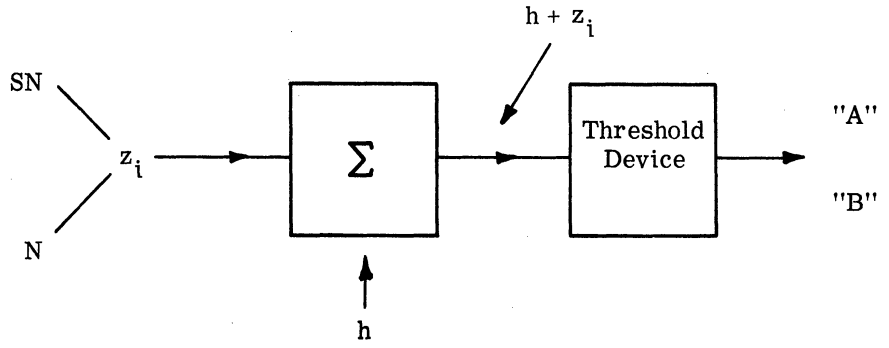


Fig. 4.1. An optimum fixed-time detector for a signal of unknown amplitude.

Although we have determined the optimum detector for only three classes of amplitude distributions, the results imply that the receiver of Fig. 4.1 is the optimum detector for any amplitude distribution. This optimum receiver is an integrator of the input followed by a threshold device.

4.3.2 Receiver Design, Sequential Realization. In order to effect a practical, in terms of receiver memory, sequential realization of the average likelihood ratio, we must demonstrate closure of the amplitude density function under normal observations. If the amplitude density function is closed, then the total observation can be broken into several smaller cascaded observations, and the average likelihood ratio of the total observation can be calculated by updating the likelihood ratio of each smaller observation taken in sequence. The closure property permits the updating process to be carried out by an adequate memory receiver.

Let us demonstrate the closure of the amplitude distribution for $\alpha = 0$. The updating process is achieved according to Bayes' Rule as given by Eq. 4.18.

$$f(a|z) = \frac{\ell(z|a)}{\int_0^\infty \ell(z|a)\, f(a)\, da} \cdot f(a) \tag{4.18}$$

The quantities on the right-hand side of Eq. 4.18 have all been previously calculated. Substituting these quantities into Eq. 4.18, we obtain

$$f(a|z) = \frac{\exp[az - a^2/2]}{\sqrt{\frac{g}{g+1}} \cdot \frac{1}{\omega(h/\sqrt{g})} \cdot \omega\left(\frac{h+z}{\sqrt{g+1}}\right)} \cdot \frac{\sqrt{g}}{\omega(h/\sqrt{g})} \exp\left[ah - \frac{a^2 g}{2}\right] \tag{4.19}$$

Combining terms in Eq. 4.19, we have

$$f(a|z) = \frac{\sqrt{g+1}}{\omega\left(\frac{h+z}{\sqrt{g+1}}\right)} \exp\left[a(h+z) - \frac{a^2}{2}(g+1)\right], \quad a \geq 0$$

$$= 0 \qquad , \text{ otherwise} \tag{4.20}$$

thereby demonstrating the closure of f(a) under normal observations. The observation maps $h \to h + z$, $g \to 1$, and $\alpha \to \alpha$. Thus, there is one constant parameter, $\alpha$, one deterministic parameter, g, and one random variable, h. Generalizing the results of Eq. 4.20 for k observations, we obtain

$$f(a|\,k \text{ obs.}) = \frac{\sqrt{g_k}}{\omega\left(\frac{h_k}{\sqrt{g_k}}\right)} \exp\left[ah_k - \frac{a^2}{2} g_k\right], \quad a \geq 0$$

$$= 0 \qquad , \quad \text{otherwise} \tag{4.21}$$

where

$$h_k = h_0 + \sum_{i=1}^{k} z_i \tag{4.22}$$

$$g_k = g_0 + k \tag{4.23}$$

and $g_0$, $h_0$ are the parameters of the initial amplitude distribution.

The average likelihood ratio after two observations can be easily obtained from the results of Eqs. 4.21 through 4.23, using Eq. 4.24

$$\ell(z_1, z_2) = \ell(z_2|z_1) \cdot \ell(z_1) \tag{4.24}$$

The conditional likelihood ratio $\ell(z_2|z_1)$ is the likelihood ratio of $z_2$ given the observation $z_1$, and can be written as

$$\ell(z_2|z_1) = \sqrt{\frac{g_1}{g_2}} \frac{\omega\left(\frac{h_2}{\sqrt{g_2}}\right)}{\omega\left(\frac{h_1}{\sqrt{g_1}}\right)} \tag{4.25}$$

Thus

$$\ell(z_1, z_2) = \sqrt{\frac{g_1}{g_2}} \frac{\omega\left(\frac{h_2}{\sqrt{g_2}}\right)}{\omega\left(\frac{h_1}{\sqrt{g_1}}\right)} \cdot \sqrt{\frac{g_0}{g_1}} \frac{\omega\left(\frac{h_1}{\sqrt{g_1}}\right)}{\omega\left(\frac{h_0}{\sqrt{g_0}}\right)}$$

$$= \sqrt{\frac{g_0}{g_2}} \frac{\omega\left(\frac{h_2}{\sqrt{g_2}}\right)}{\omega\left(\frac{h_0}{\sqrt{g_0}}\right)} \tag{4.26}$$

Generalizing, the likelihood ratio after k observations is

$$\ell(z \mid k \text{ obs}) = \sqrt{\frac{g_0}{g_k}} \cdot \frac{\omega\left(\frac{h_k}{\sqrt{g_k}}\right)}{\omega\left(\frac{h_0}{\sqrt{g_0}}\right)} \tag{4.27}$$

A sequential realization of the optimum fixed-observation detection receiver can be obtained by using Eq. 4.27 as the basis of the design. Since $\omega(\ )$ is a monotone function of its argument and $g_k$ is deterministic, $h_k$ is a monotone function of the likelihood ratio of the input observations. Thus, the optimum receiver can be realized as shown in Fig. 4.2. The performance of the receiver in Fig. 4.2 is identical to the performance of the nonsequential realization presented in Fig. 4.1.

The sequential realizations of the optimum receiver for the $\alpha = 1$ and Gaussian amplitude distribution are the same as shown in Fig. 4.2. The average likelihood ratios after k observations for the $\alpha = 1$ and the Gaussian amplitude distributions are given by Eqs. 4.28 and 4.29, respectively.

$$\ell(z \mid k \text{ obs}) = \frac{g_0}{g_k} \cdot \frac{1 + \frac{h_k}{\sqrt{g_k}} \cdot \omega\left(\frac{h_k}{\sqrt{g_k}}\right)}{1 + \frac{h_0}{\sqrt{g_0}} \cdot \omega\left(\frac{h_0}{\sqrt{g_0}}\right)} \tag{4.28}$$

$$\ell(z \mid k \text{ obs}) = \sqrt{\frac{g_0}{g_k}} \cdot \exp\left[\frac{h_k^2}{2g_k} - \frac{h_0^2}{2g_0}\right] \tag{4.29}$$

the optimum receiver for $\alpha = 1$ and the Gaussian distributions (amplitude) again calculates $h_k$ which is a monotone function of the likelihood ratio.

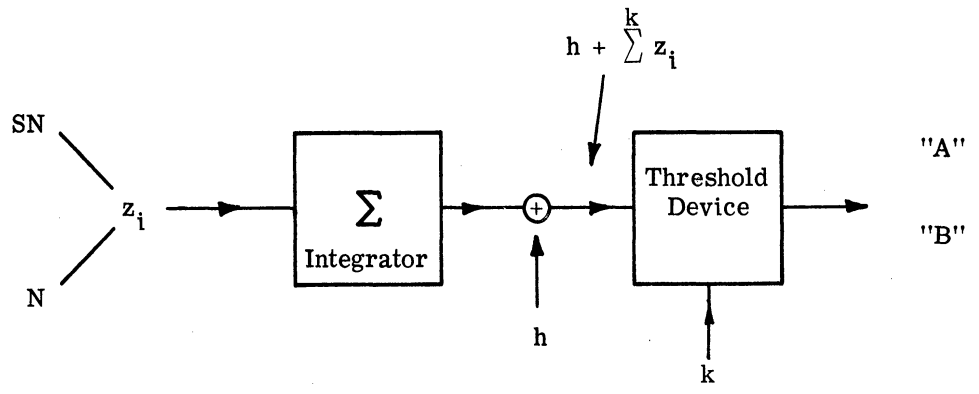


Fig. 4.2. A sequential realization of an optimum fixed-time detector for a signal of unknown amplitude.

4.3.3 Evaluation of the Optimum Fixed-Observation Receiver. Evaluation of fixed-observation procedures can be given in terms of the error performance of the procedure. We present in this section the error performance for the $\alpha = 0$ amplitude distribution in terms of ROC curves. The length of observation is measured in the number of observations taken rather than by the more general concept of observation quality. The reason for this is discussed later in this section along with the units of the time scale used. The ROC curves presented are parameterized according to the length of the observation, i.e., the number of observations taken.

Consider the calculation of the false alarm probability P("A"|N). A false alarm occurs whenever the log-odds ratio, at the time of termination, is greater than the threshold, $\Delta_0$, in noise alone. In symbols

$$P(\text{"A"}|N) = P(L_k > \Delta_0|N) \tag{4.30}$$

where

$$L_k = L_0 + \ell n\, \ell(z_1, z_2, \ldots, z_k) \tag{4.31}$$

Using Eq. 4.27, the above equation can be written as

$$L_k = L_0 + \ell n \left[ \frac{\sqrt{g_0}}{\omega\left(\frac{h_0}{\sqrt{g_0}}\right)} \right] + \ell n \left[ \frac{\omega\left(\frac{h_k}{\sqrt{g_k}}\right)}{\sqrt{g_k}} \right] \tag{4.32}$$

The zero subscript again corresponds to the a priori state of the detection problem. Substituting Eq. 4.32 into Eq. 4.30, the probability of false alarm can be written as

$$P("A" \,|\, N) = P\left\{ \ell n \left[ \frac{\omega\left(\frac{h_k}{\sqrt{g_k}}\right)}{\sqrt{g_k}} \right] > \Delta_0 - L_0 + \ell n \left[ \frac{\omega\left(\frac{h_0}{\sqrt{g_0}}\right)}{\sqrt{g_0}} \right] \,\Big|\, N \right\} \tag{4.33}$$

Because the logarithmic and $\omega(\ )$ functions are monotone in their arguments, Eq. 4.33 is equivalent to

$$P("A" \,|\, N) = P(h_k > h^* \,|\, N) \tag{4.34}$$

where $h^*$ is the solution to the equation

$$\ell n \left[ \frac{\omega\left(\frac{h^*}{\sqrt{g_k}}\right)}{\sqrt{g_k}} \right] = \Delta_0 - L_0 + \ell n \left[ \frac{\omega\left(\frac{h_0}{\sqrt{g_0}}\right)}{\sqrt{g_0}} \right] \tag{4.35}$$

Equation 4.35 is a transcedental equation and is solvable only by approximate methods. Using the solution for $h^*$ from Eq. 4.35, the false alarm probability is easily obtained. The quantity $h_k$ is defined by Eq. 4.22 as

$$h_k = h_0 + \sum_{i=1}^{k} z_i \tag{4.22}$$

Thus

$$h_k - h_0 = \sum_{i=1}^{k} z_i \tag{4.36}$$

In noise alone $\sum_{i=1}^{k} z_i$ is distributed according to a Gaussian distribution with zero mean and variance k. Thus, we have

$$P("A"|N) = P\left(\frac{h_k - h_0}{\sqrt{k}} > \frac{h^* - h_0}{\sqrt{k}} \,\middle|\, N\right) = \Phi\left(\frac{h_0 - h^*}{\sqrt{k}}\right) \tag{4.37}$$

The corresponding probability of detection, $P("A"|SN)$, is found in a similar manner. The detection probability that the log-odds ratio is greater than the threshold, $\Delta_0$, in SN. This can be written as

$$P("A"|SN) = P(L_k > \Delta_0|SN) \tag{4.38}$$

The total log-odds ratio, $L_k$, is again given by Eq. 4.22 and, as in the derivation of $P("A"|N)$, the probability that $L_k > \Delta_0$ can be written

$$P("A"|SN) = P(h_k > h^*|SN) \tag{4.39}$$

where $h^*$ is the solution to Eq. 4.35. The quantity $h_k - h_0$ is defined in Eq. 4.36 as $\sum_{i=1}^{k} z_i$. The distribution of $h_k - h_0$ in SN, if the signal amplitude is known is, therefore, Gaussian with mean ka and variance k. Thus, the probability of detection, given the signal amplitude, can be written as

$$P("A"|SN,a) = P\left[\frac{h_k - h_0}{\sqrt{k}} > \frac{h^* - h_0}{\sqrt{k}} \,\middle|\, SN, a\right] \tag{4.40}$$

The probability of detection for a distributed signal amplitude is found by a weighted average of Eq. 4.40 with the signal amplitude density function. That is,

$$P("A"|SN, a) = \int_0^\infty P("A"|SN, a)\, f(a)\, da \tag{4.41}$$

Since $\sum_{i=1}^{k} z_i$ is distributed N(ka, k), then the random variable, $\xi$, defined below is distributed N(0, 1)

$$\xi = \frac{\sum_{i=1}^{k} z_i}{k} - a\sqrt{k} \tag{4.42}$$

Thus, we can rewrite Eq. 4.40 in terms of $\xi$ as

$$P("A"|SN, a) = P\left(\xi > \frac{h^* - h_0}{\sqrt{k}} - a\sqrt{k} \,\middle|\, SN, a\right) \tag{4.43}$$

From the expression for the conditional detection probability, Eq. 4.43, we find P("A"|SN) by averaging P("A"|SN, a) with respect to the amplitude distribution. An equivalent calculation of P("A"|SN) is obtained by averaging P("A"|SN, $\xi$) with respect to the $\xi$ distribution, i.e., the Gaussian distribution of zero mean and unity variance. In symbols, we have

$$P(\text{"A"}|SN, \xi) = P\left(a > \frac{h^* - h_0}{\sqrt{k}} - \frac{\xi}{\sqrt{k}} \,\middle|\, SN, \xi\right) \tag{4.44}$$

The reason this alternate function for P("A"|SN) is given is that the averaging process for $\xi$ assumed known is easier and faster to perform on the IBM 7090. The proof that the two formulations for obtaining P("A"|SN) are equivalent is given in Appendix D.

Continuing from Eq. 4.44, we have that the probability that "a" is greater than any threshold, in this case $(h^* - h_0)/k - \xi/\sqrt{k}$ is the integral of f(a) from zero to the threshold value. The amplitude is assumed to be always nonnegative. Therefore, if the right-hand side of the inequality contained in Eq. 4.44 is negative, then "a" is greater than threshold with probability one. That is, if $(h^* - h_0)/k < \xi/\sqrt{k}$, then

$$P\left(a > \frac{h^* - h_0}{k} - \frac{\xi}{\sqrt{k}} \,\middle|\, SN\right) = 1.0 \tag{4.45}$$

However, if $\frac{h^* - h_0}{k} > \frac{\xi}{k}$, then

$$P\left(a > \frac{h^* - h_0}{k} - \frac{\xi}{\sqrt{k}} \,\middle|\, SN\right) = \int_{-\infty}^{+\infty} \frac{\sqrt{g_0}}{\Phi\left(\frac{h_0}{\sqrt{g_0}}\right)} \Phi\left(\frac{\frac{h_0}{g_0} + \frac{h_0 - h^*}{k} + \frac{\xi}{\sqrt{k}}}{\frac{1}{\sqrt{g_0}}}\right) d\Phi(\xi) \tag{4.46}$$

Summarizing, P("A"|SN) is given by Eq. 4.45 for $\xi > \frac{h^* - h_0}{\sqrt{k}}$ and by Eq. 4.46 for $\xi < \frac{h^* - h_0}{\sqrt{k}}$.

The preceding derivation pertains only to the $\alpha = 0$ signal amplitude distribution. A similar derivation must be performed for the $\alpha = 1$ and Gaussian signal amplitude distributions. These derivations of P("A"|N) and P("A"|SN) contribute little to a basic understanding of the evaluation problem, and so we have placed these derivations in Appendix E.

4.3.4 Numerical Results of the Fixed-Observation Procedure. The association of a cost with the observation process implies that the optimum receiver must balance the observation cost and terminal decision error performance. The length of observation is, therefore, an important explicit variable in discussing optimum detectors.

In the detection of a SKEP, we were able to define the length of observation in terms of the observation quality, a measure of the separation between the noise alone and signal-plus-noise distributions. In the problem of a SKEA, there is no simple quantity that can be defined as the quality of observation. We define the length of an observation in terms of the number of observations taken.

Let T be the observation time and W the receiver noise bandwidth. By the Nyquist sampling theorem, it is necessary to sample every 1/2W seconds in order to completely describe the receiver waveform. Thus,

$$\text{sampling time} = \frac{\text{observation time}}{\text{number of samples}}$$

$$= \frac{1}{2W} = \frac{T}{k} \tag{4.47}$$

Since the noise power is equal to $N_0W$, Eq. 4.47 implies that the length of an observation is

$$T = \frac{2Nk}{N_0} \tag{4.48}$$

where k is the number of observations.

The numerical evaluation of the probability of false alarm and probability of detection as derived in Section 4.3.3 is accomplished by means of an IBM 7090 programmed according to the algorithm presented in Appendix F. The results that follow pertain to the $\alpha = 0$ signal amplitude distribution. Results for the $\alpha = 1$ and Gaussian signal amplitude distributions are similar to those obtained for the $\alpha = 0$ distributions.

Figure 4.3 depicts the ROC curves for none truncated Gaussian signal amplitude distributions. Several of the amplitude distributions are shown in Figs. 4.4 and 4.5. The ROC curves represent performance of a single observation, i.e., k = 1.

The mean and variance of the $\alpha = 0$ amplitude distributions are approximately h/g and $1/\sqrt{g}$, respectively. It would be very useful if we could order the ROC curves, in terms of distance from the (0, 1) point, by some simple combination of the parameters h and g.

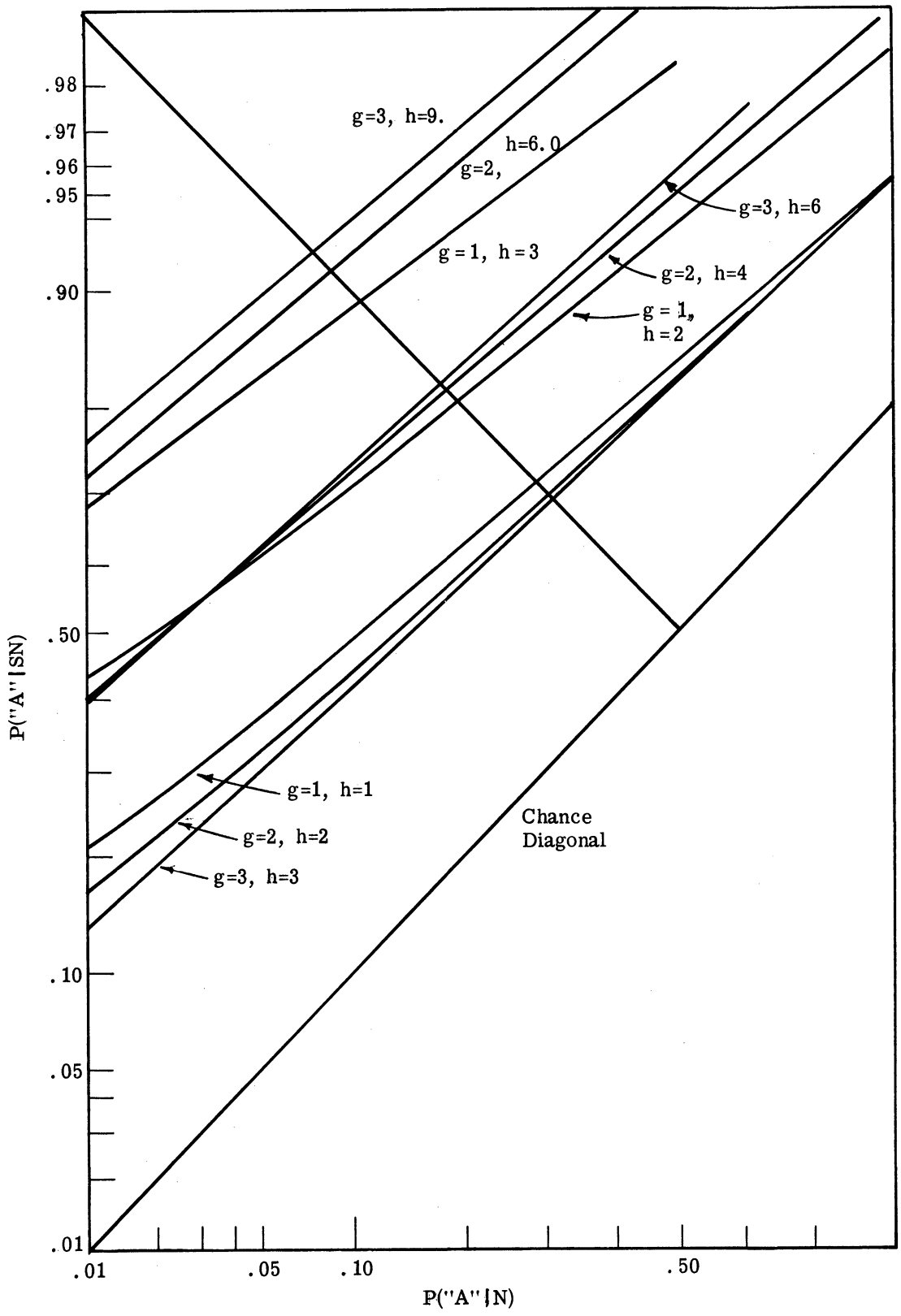


Fig. 4.3. ROC curves for the optimum fixed-time detector for several truncated Gaussian amplitude distributions.

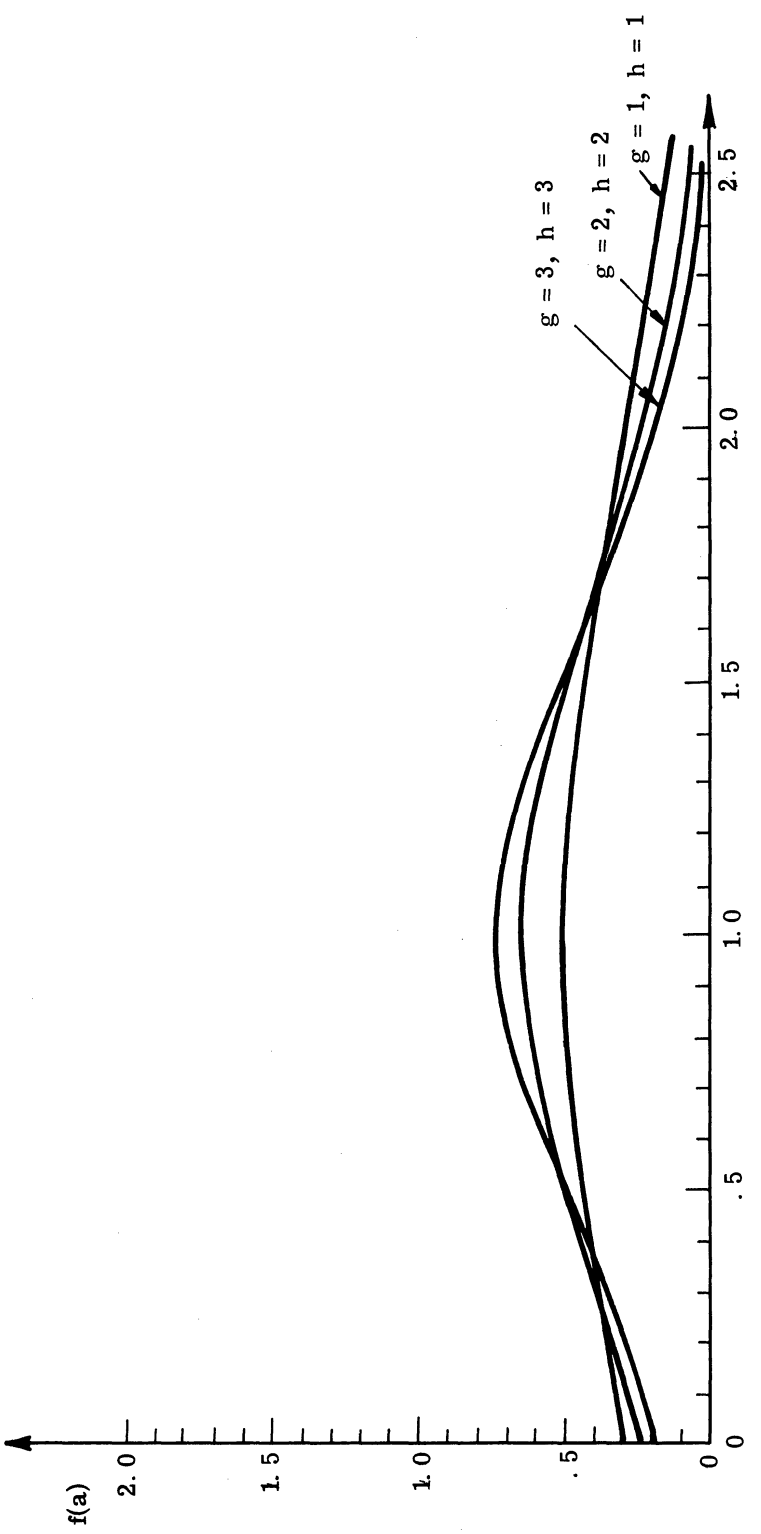


Fig. 4.4. Graphs of truncated Gaussian probability density functions with parameters h = g = 1, h = g = 2, and h = g = 3.

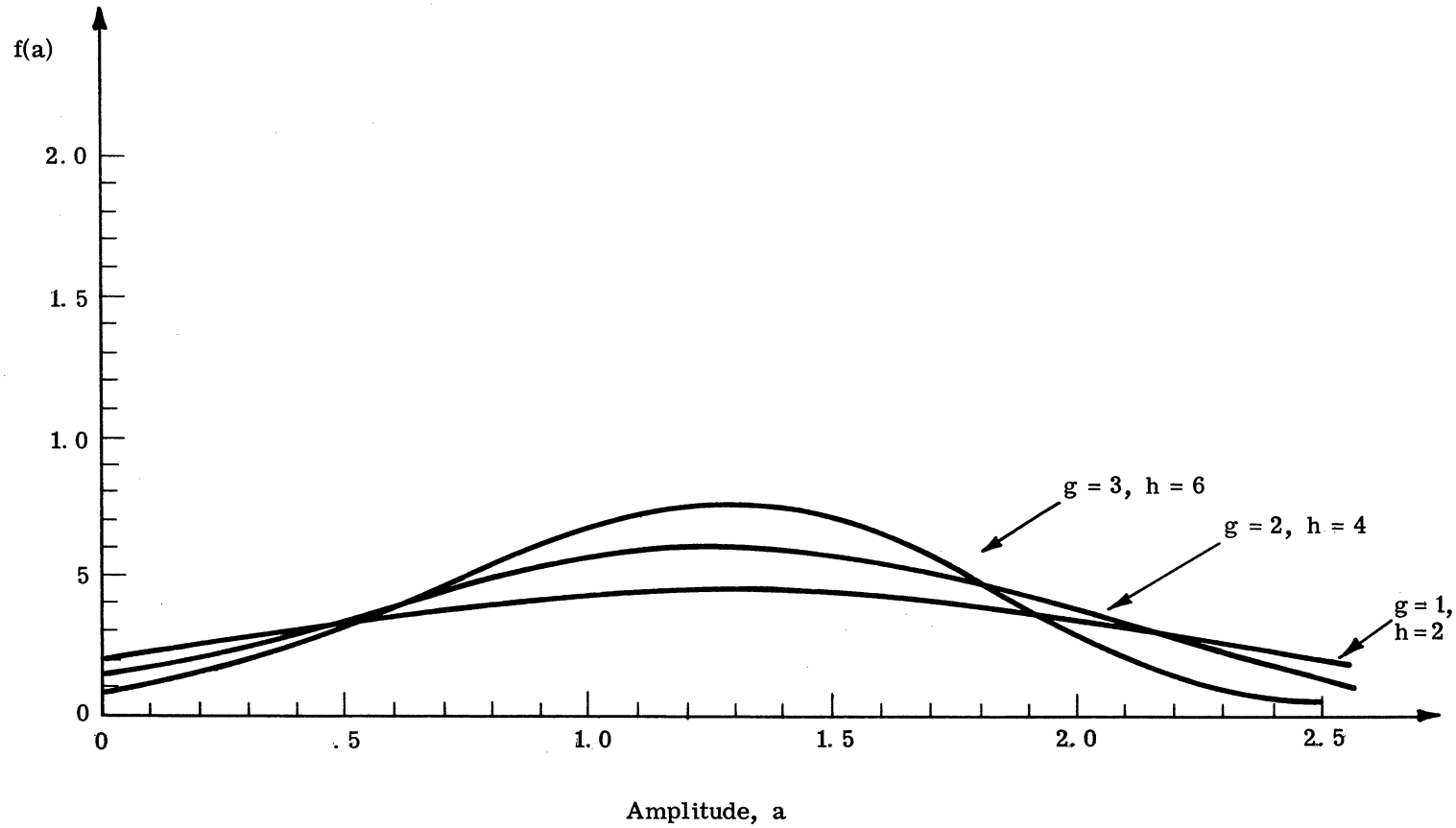


Fig. 4.5. Graphs of truncated Gaussian probability density functions with parameters y = 1 and h = 2, g = 2 and h = 4, and g = 3 and h = 6.

If these were possible, we could predict the performance order of a set of $\alpha = 0$ distributions and also predict the performance (approximate) of $\alpha = 0$ distributions from the knowledge of a small number of amplitude distributions. Unfortunately, the author has not been able to find a simple combination of h and g which can be used to order the ROC curves over the complete range from (0, 0) to (1, 1). A few generalities can be made.

Referring to Fig. 4.3, we see that for the same approximate mean, h/g, the ROC curves with the smaller variances, $1/\sqrt{g}$ result in fewer detections for low false alarms and greater detection for high false alarms. As the mean value of the amplitude distributions increase, the crossover point, where the last statement is true, shifts toward lower false alarm rates.

If observations of increasing longer length are taken of the same truncated amplitude distribution, we obtain a set of ROC curves similar to those shown in Figs. 4.6 and 4.7. The parameter across the set of ROC curves is the length of observation, i.e., k = 1, 2, ... . The striking feature of the ROC curves shown in Figs. 4.6 and 4.7 is the manner in which they converge as longer observations are taken. A comparable set of ROC curves for a signal known exactly results in ROC curves that do not converge. The convergence of the ROC curves implies that ratio of the mean to the variance of the distribution of the likelihood ratio does not continue to increase at the same rate as the receiver input is observed for increasing longer times.

The ROC curves of Figs. 4.6 and 4.7 will continue to be monotonely better as the observation time increases. However, the rate of increase in performance is small. Let us show that as longer observations are taken the ROC curves continue to improve monotonely.

If the amplitude of the signal is known, the ROC curves are normal. Consider the rate of increase in detections for a given false alarm probability, say, $x = \Phi(\lambda)$. The corresponding detection probability for a given amplitude and length of observation is $y(x|a, k) = \Phi(\lambda + \sqrt{k}a)$. The detection probability for an amplitude distributed according to f(a) is

$$y(x) = \int_0^\infty \Phi(\lambda + \sqrt{k}a)\, f(a)\, da \qquad (4.49)$$

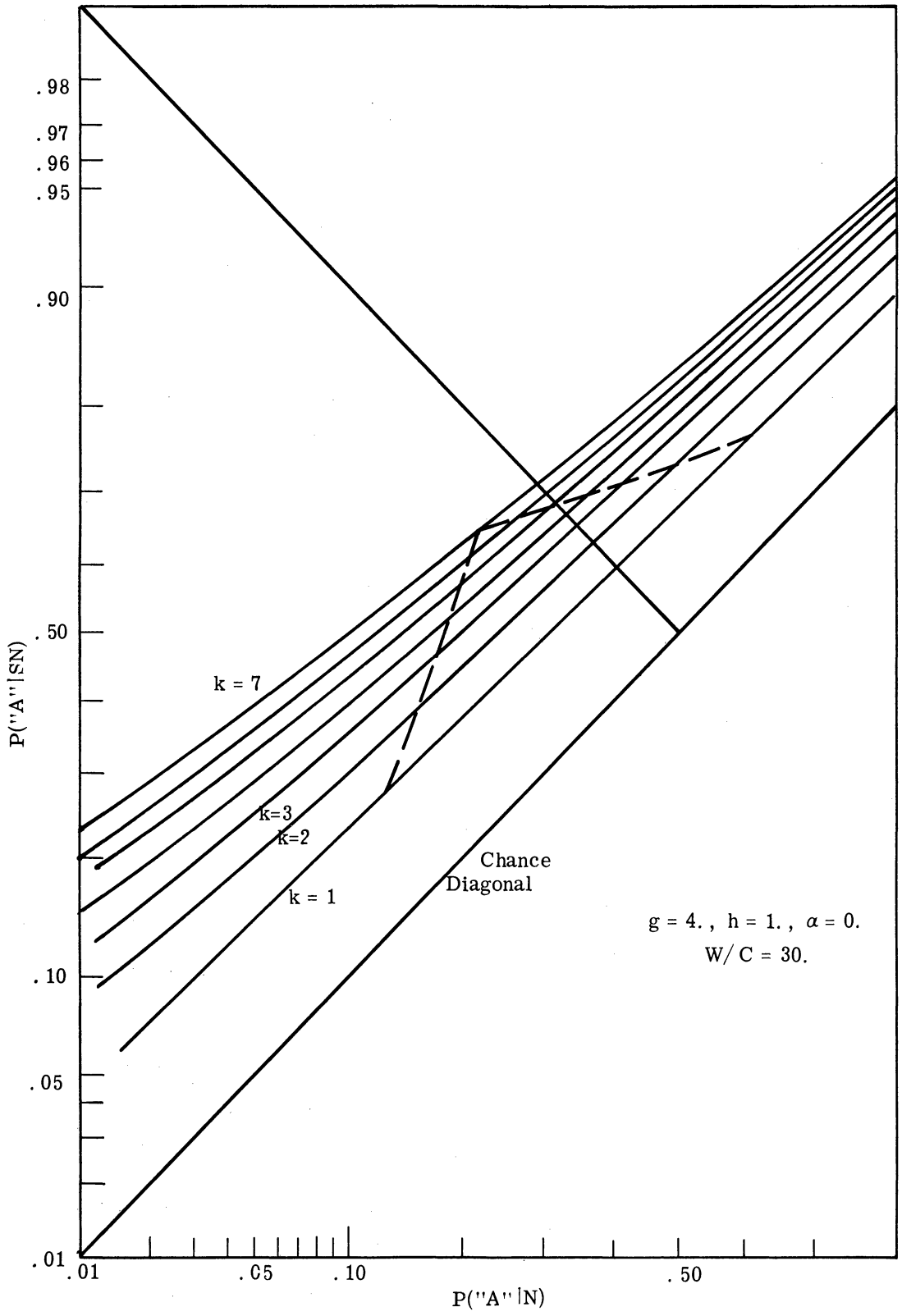


Fig. 4.6. ROC curves for the optimum fixed-time detector for the truncated Gaussian amplitude distribution with parameters h = g = 1.

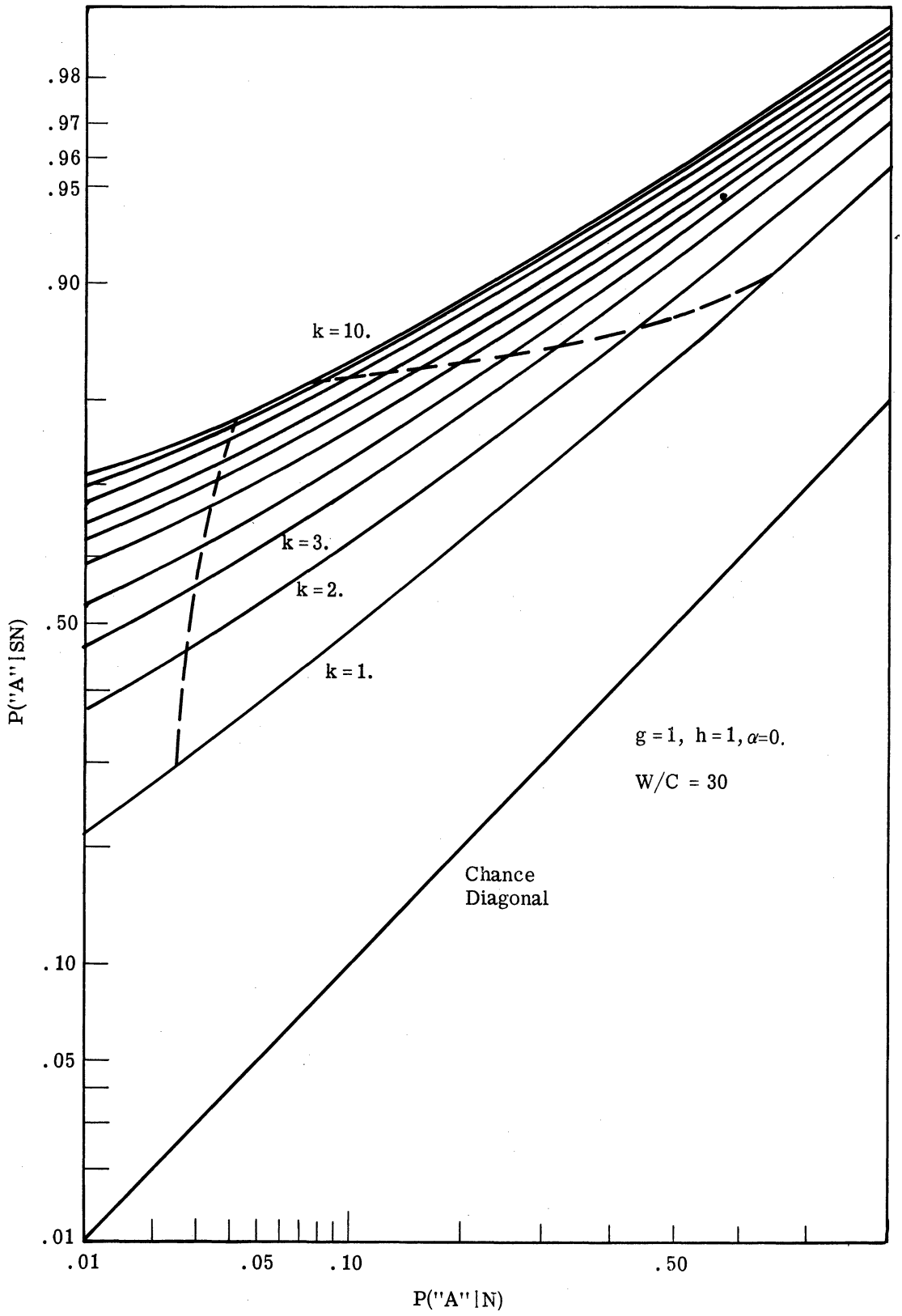


Fig. 4.7. ROC curves for the optimum fixed-time detector for the truncated Gaussian amplitude distribution with parameters h = 1 and g = 4.

We will show that for any false alarm probability, $\Phi(\lambda)$, the detection probability, $\Phi(\lambda + \sqrt{k}a)$, can be made as close to one as we wish as we let k increase, and thus the integral $\int_0^\infty \Phi(\lambda + \sqrt{k}a)\, f(a)\, da$ can be made as close to one as we wish. Formally we wish to show that $\lim_{k\to\infty} \int_0^\infty \Phi(\lambda + \sqrt{k}a)\, f(a)\, da = 1$. That is, given $\lambda_0$ and $\epsilon > 0$, there exists a number $k_0$ such that for $k > k_0$ the integral $\int_0^\infty \Phi(\lambda_0 + \sqrt{k}a)\, f(a)\, da > 1 - \epsilon$. We can write

$$\begin{aligned}
\int_0^\infty (\lambda_0 + \sqrt{k}a)\, f(a)\, da &= \int_0^c \Phi(\lambda_0 + \sqrt{k}a)\, f(a)\, da + \int_0^\infty \Phi(\lambda_0 + \sqrt{k}a)\, f(a)\, da \\
&\geq \int_0^c \Phi(\lambda_0)\, f(a)\, da + \int_c^\infty \Phi(\lambda_0 + \sqrt{k}c)\, f(a)\, da \\
&\geq \Phi(\lambda_0) \int_0^c f(a)\, da + \Phi(\lambda_0 + \sqrt{k}c) \int_c^\infty f(a)\, da \\
&\geq \Phi(\lambda_0)\, F(c) + \Phi(\lambda_0 + \sqrt{k}c)\, [1 - F(c)] \qquad (4.50)
\end{aligned}$$

where F(c) is the distribution function of f(a). Rewriting Eq. 4.50, we have

$$\int_0^\infty \Phi(\lambda_0 + \sqrt{k}a)\, f(a)\, da \geq \Phi(\lambda_0 + \sqrt{k}c) - F(c)\, [\Phi(\lambda_0 + \sqrt{k}c) - \Phi(\lambda_0)] \qquad (4.51)$$

The quantity in brackets is less than or equal to one, and thus if we replace the brackets by one we insure the inequality. Thus we have

$$\int_0^\infty \Phi(\lambda_0 + \sqrt{k}a)\, f(a)\, da \geq \Phi(\lambda_0 + \sqrt{k}c) - F(c) \qquad (4.52)$$

Choose $c > 0$ such that $F(c) < \epsilon/2$; then choose $k = k(\lambda, c, \epsilon)$ such that $\Phi(\lambda_0 + \sqrt{k}c) > 1 - \epsilon/2$. This demonstrates the monotone increase of the ROC curves with increasing k.

The most significant consequence of the convergence of the ROC curves is that it is profitable to observe for only a relatively short time in the SKEA problem. This conclsuion will become more apparent when the results of the optimum nonsequential procedure are presented.

## 4.4 The Optimum Nonsequential Observation-Decision Procedure

4.4.1 Analytic Derivation. We consider in this section the derivation of the optimum nonsequential procedure (ONP) for the case of a SKEA. This is the same observation-decision procedure as described in Section 3.4.1. In the interest of continuity and

clarity, we will derive the ONP in detail for the unknown amplitude problem. There are three mathematical problems to solve, one for each amplitude distribution. A detailed derivation of the ONP is given for the $\alpha = 0$ amplitude distribution and only the results are summarized for the $\alpha = 1$ and Gaussian amplitude distribution.

The ONP is a fixed-length observation procedure optimized by the correct choice of the observation length. The observation length is based upon the relevant parameters of the detection problem prior to the actual start of the observations. This information includes (1) the a priori amplitude distribution, (2) the a priori log-odds ratio, (3) the cost of a single observation, (4) the losses due to terminal decision errors, (5) the observation statistics, and (6) the total available observation time.

The optimum observation length is obtained by determining the maximum of the value for an additional observation in terms on the parameters described in the previous paragraph. The maximum of the value function, as a function of observation time, is determined, as in the SKEP problem, by differentiating the value function with respect to observation time. The differentiated expression is set equal to zero, and the resultant equation is solved for the optimum observation time.

The value of an observation is defined as the amount to be gained by observation and is the optimum terminal loss function, T(L), minus the average look-ahead loss, $G(L; k; g, h, \alpha)$. L is the a priori log-odds ratio, k is the available number of observations, and g, h, and $\alpha$ are the signal amplitude distribution parameters. Referring to Eqs. 3.38 and 3.39, the value of an observation can be written in terms of probability of detection and the probability of false alarm as

$$V(L, k) = \frac{W}{1 + e^{-L}} P(\text{"A"}|SN) - \frac{W}{1 + e^{L}} P(\text{"A"}|N) - Ck, \quad L < \Delta_0$$

$$= \frac{W}{1 + e^{-L}} P(\text{"B"}|SN) + \frac{W}{1 + e^{L}} P(\text{"B"}|N) - Ck, \quad L > \Delta_0 \qquad (4.53)$$

Consider the determination of the optimum observation length for $L < \Delta_0$. (The same logic applies for $L > \Delta_0$.) From Eqs. 4.37, 4.45, and 4.46, we obtain the expressions for P("A"|SN) and P("A"|N) in terms of the parameters of the observation procedure.

Because of the complicated functions, the differentiation of Eq. 4. 53 with respect to observation time was carried out numerically by use of an IBM 7090. The program used is given in Appendix F.

Before presenting the numerical results of the ONP, consider the manner in which that state of ONP is given. In order to compare the results of the ONP with deferred decision, a description of the state of the decision problem must be given in terms of the separable subspaces. Thus, the numerical results of the ONP will be presented for a state specified in two ways: (1) by specification of the separable subspaces and the log-odds ratio, and (2) in the more "natural" manner, by specifying the amplitude distribution and the log-odds ratio. Rather than derive the definitions of the separable spaces here, we present the definitions below and leave the derivations for Section 4. 5.

The separable subspaces for the $\alpha = 0$ signal amplitude distribution are given by the value of Q as defined in Eq. 4. 54.

$$Q = L + \ell n \left( \frac{\sqrt{g}}{\omega(h/\sqrt{g})} \right) \tag{4.54}$$

For the $\alpha = 1$, signal amplitude distribution the value of Q is defined as

$$Q = L + \ell n \left( \frac{g}{1 + \frac{h}{\sqrt{g}}\, \omega(h/\sqrt{g})} \right) \tag{4.55}$$

And finally, for the Gaussian amplitude distribution, the separable subspaces are defined by

$$Q = L + \ell n \sqrt{g} - \frac{h^2}{2g} \tag{4.56}$$

4. 4. 2 Numerical Results of the Optimum Nonsequential Procedure. The numerical results of the ONP are best presented in the form of the contour graph for the value of observation as in the SKEP problem. Each contour graph is parameterized by the error loss-to-observation cost ratio, W/C, and amplitude distribution. The amplitude distributions are defined in two ways: (1) by specification of the actual distribution through the set of parameters ($\alpha$, g, h), or (2) by the specification of the separable subspace Q, the log-odds ratio, and g. As the former method of description is more natural and intuitive, let us consider the value contour graph from this point of view first.

Each class of amplitude distribution, i. e., $\alpha = 0$, $\alpha = 1$, and the Gaussian case, presents its individual problems in the calculations of quantitative results. Results are grossly the same for the three classes of amplitude distributions studied. Thus, extensive computations are presented only for the $\alpha = 0$ signal amplitude distribution. The general discussions and conclusions concerning the numerical results of the $\alpha = 0$ distribution carry over to the $\alpha = 1$ and Gaussian distributions.

If the state of the decision problem is given in terms of the $\alpha = 0$ amplitude distribution, i. e., g and h, L, and W/C, we obtain the value contour graphs shown in Figs. 4. 9 through 4. 14. In Figs. 4. 9 through 4. 14, the zero value contour is plotted as a function of the number of observations and L for W/C = 30. The approximately horizontal dotted line in each graph is the optimum observation length, as a function of L, for the minimum average loss for a terminal decision. Each value contour graph corresponds to a different $\alpha = 0$ amplitude distribution. Several of the amplitude distributions are shown in Figs. 4. 4, 4. 5 and 4. 8.

Several conclusions can be drawn from the set of value contour curves shown in Figs. 4. 9 through 4. 14. For example, the optimum observation length is relatively small for all the amplitude distributions of Figs. 4. 9 through 4. 14. The corresponding ROC curves shown in Figs. 4. 15 through 4. 20 exhibit a convergence phenomena as observations of increasing longer length are taken. This implies that it does not pay in decreased error performance to take proportionately longer observations, since the rate of increase in error performance is small.

We note that as the mean of the amplitude distribution increases (for the same variance), the zero value contour graph expands. Thus, the zero value contour for h = 3 and g = 1 is larger than for h = 2 and g = 1. This can be explained intuitively by the fact that as the mean becomes larger, the uncertainty of presence or absence of the signal decreases, and whenever "cause uncertainty" decreases, the zero value contour increases in size. (This can also be seen in SKEP problems. See Fig. 3. 11. )

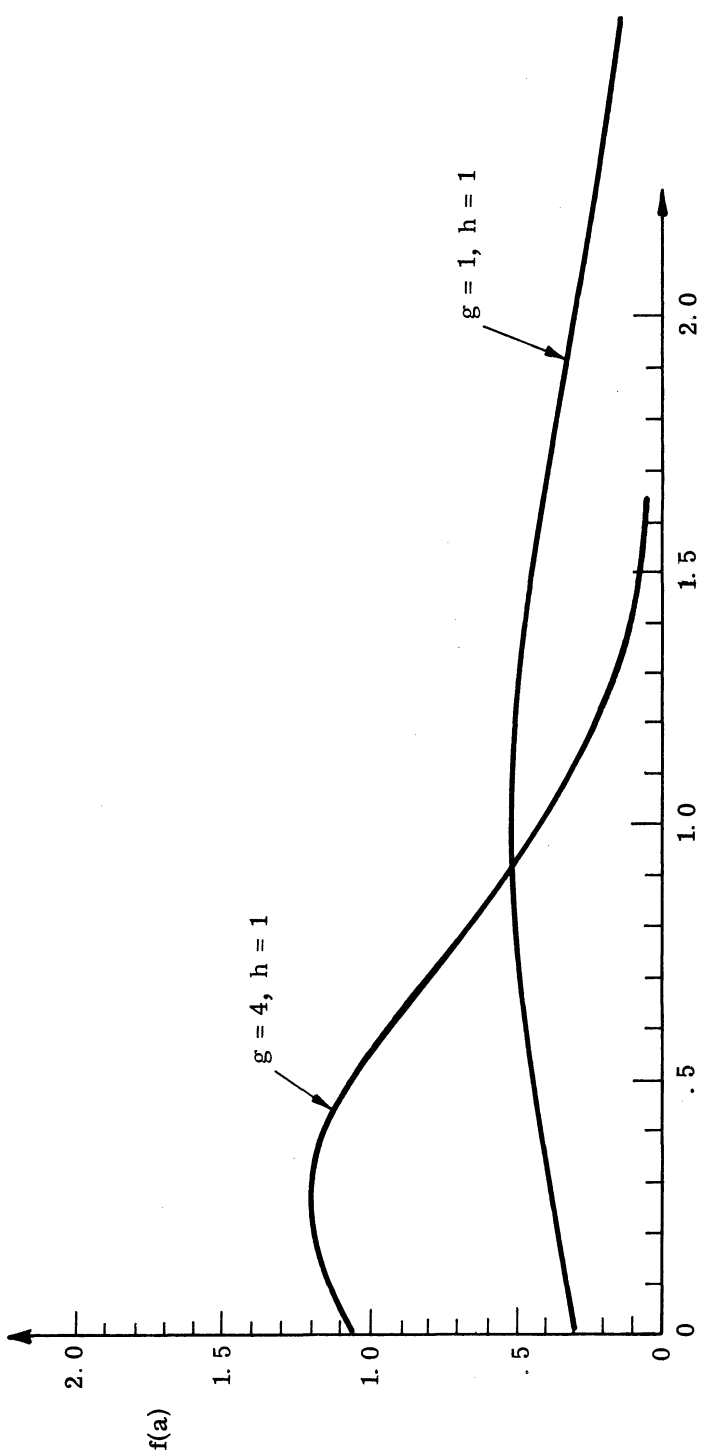


Fig. 4.8. Graphs of truncated Gaussian probability density functions with parameters h = g = 1 and h = 1 and g = 4.

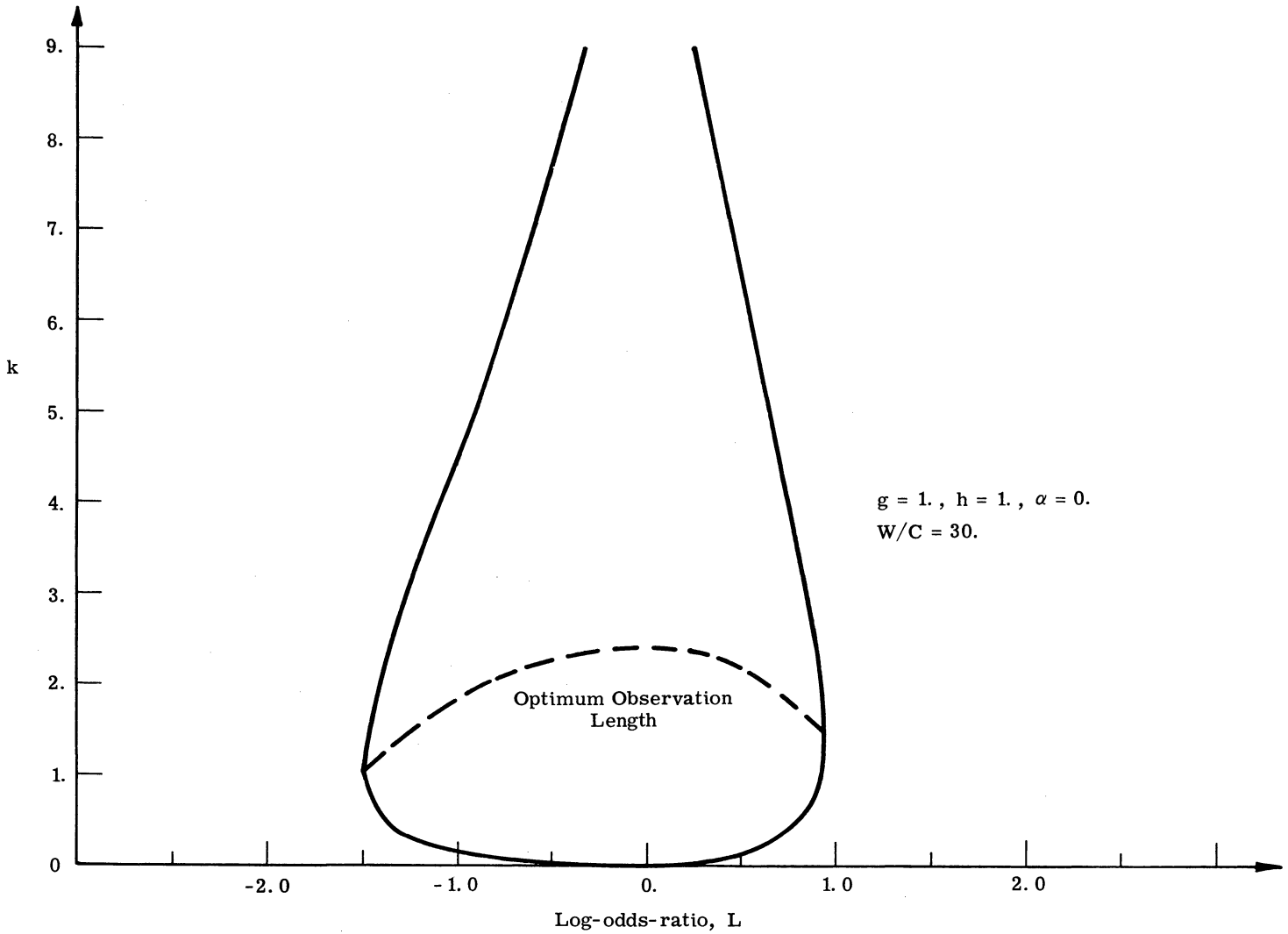


Fig. 4.9. The value contour graph for the ONP for a SKEA with parameters W/C = 30, h = g = 1, and $\alpha = 0$.

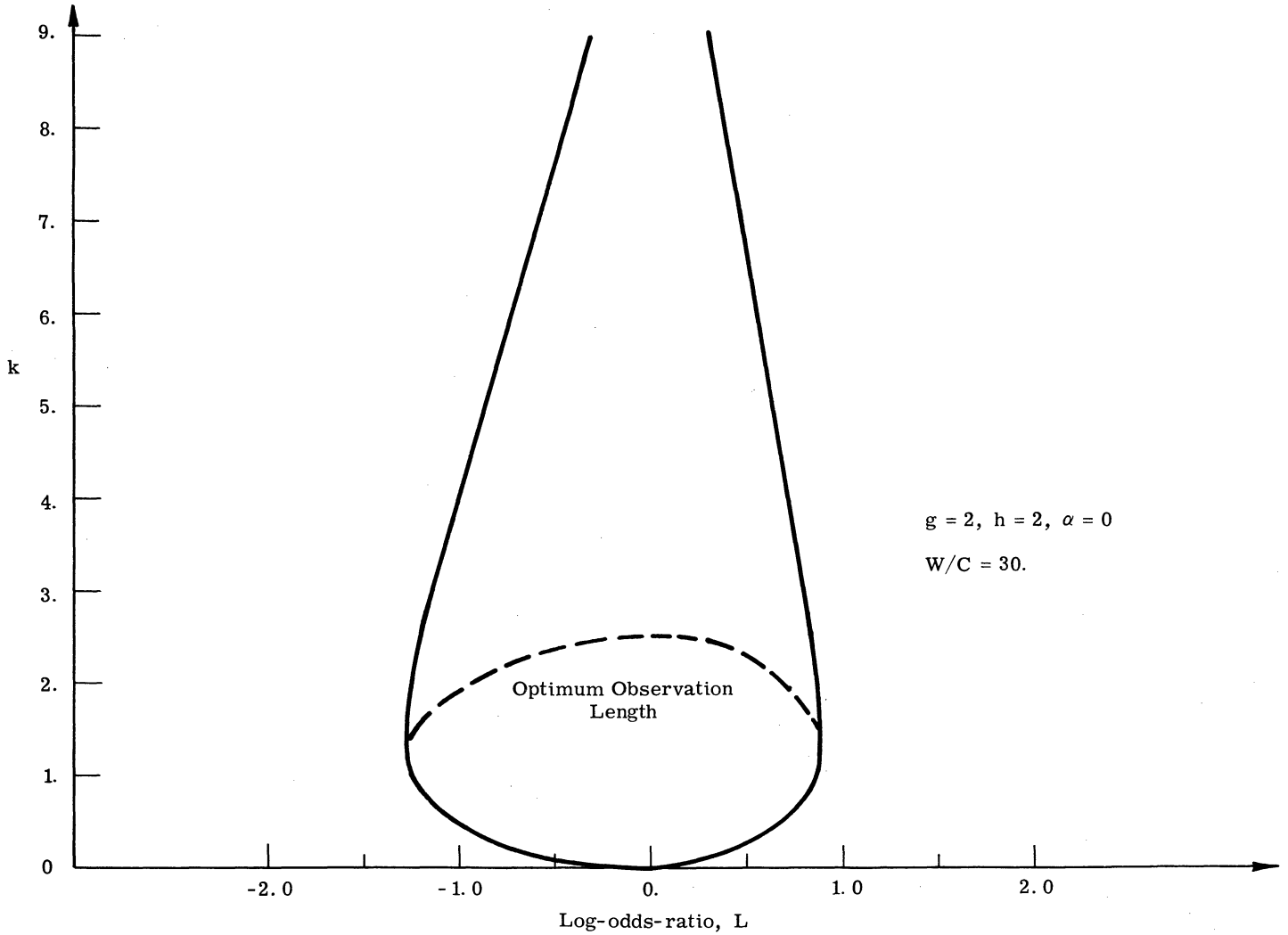


Fig. 4.10. The value contour graph for the ONP for a SKEA with parameters W/C = 30, h = g = 2, and $\alpha = 0$.

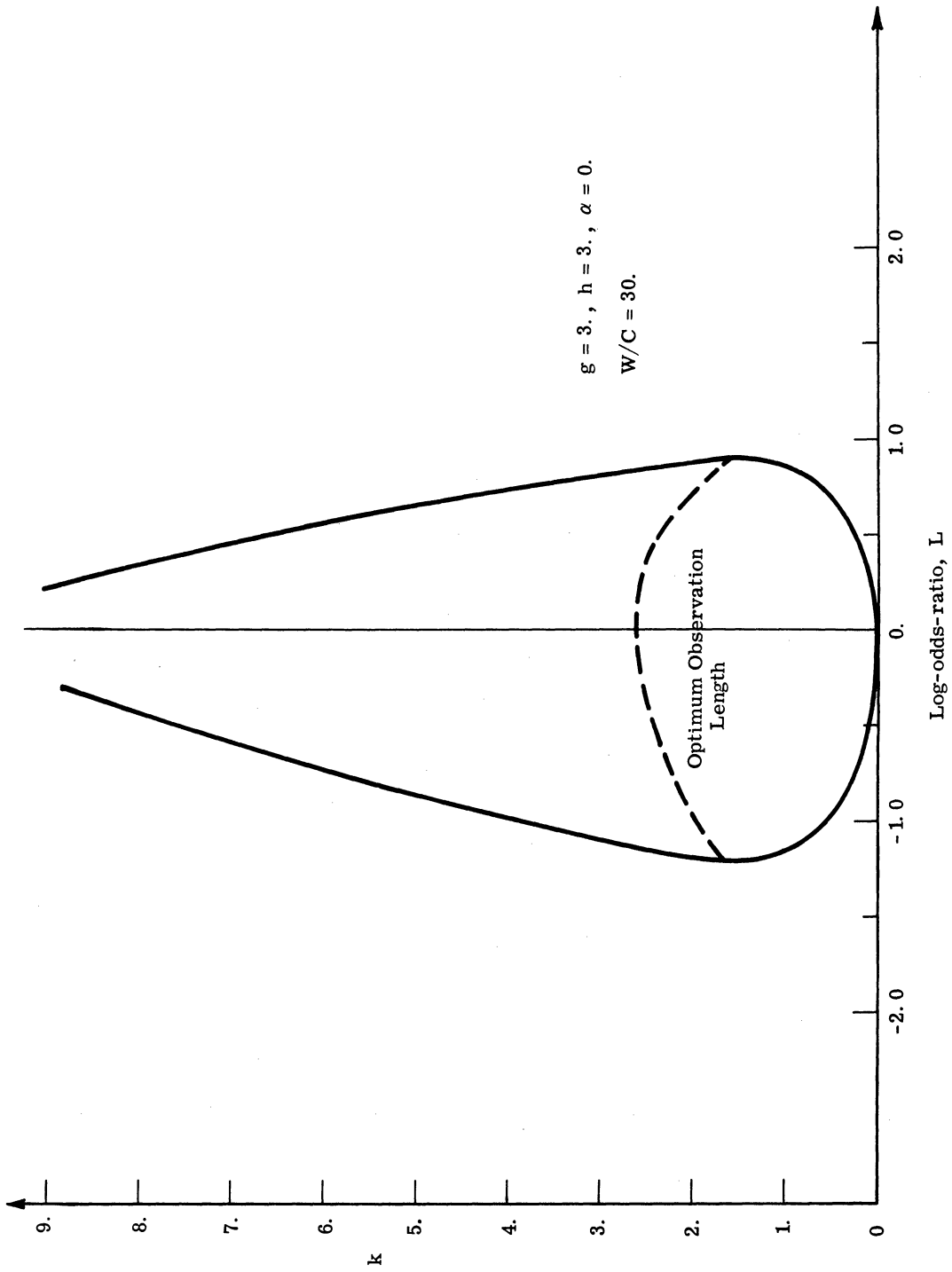


Fig. 4.11. The value contour graph for the ONP for a SKEA with parameters W/C = 30, h = g = 3, and $\alpha = 0$.

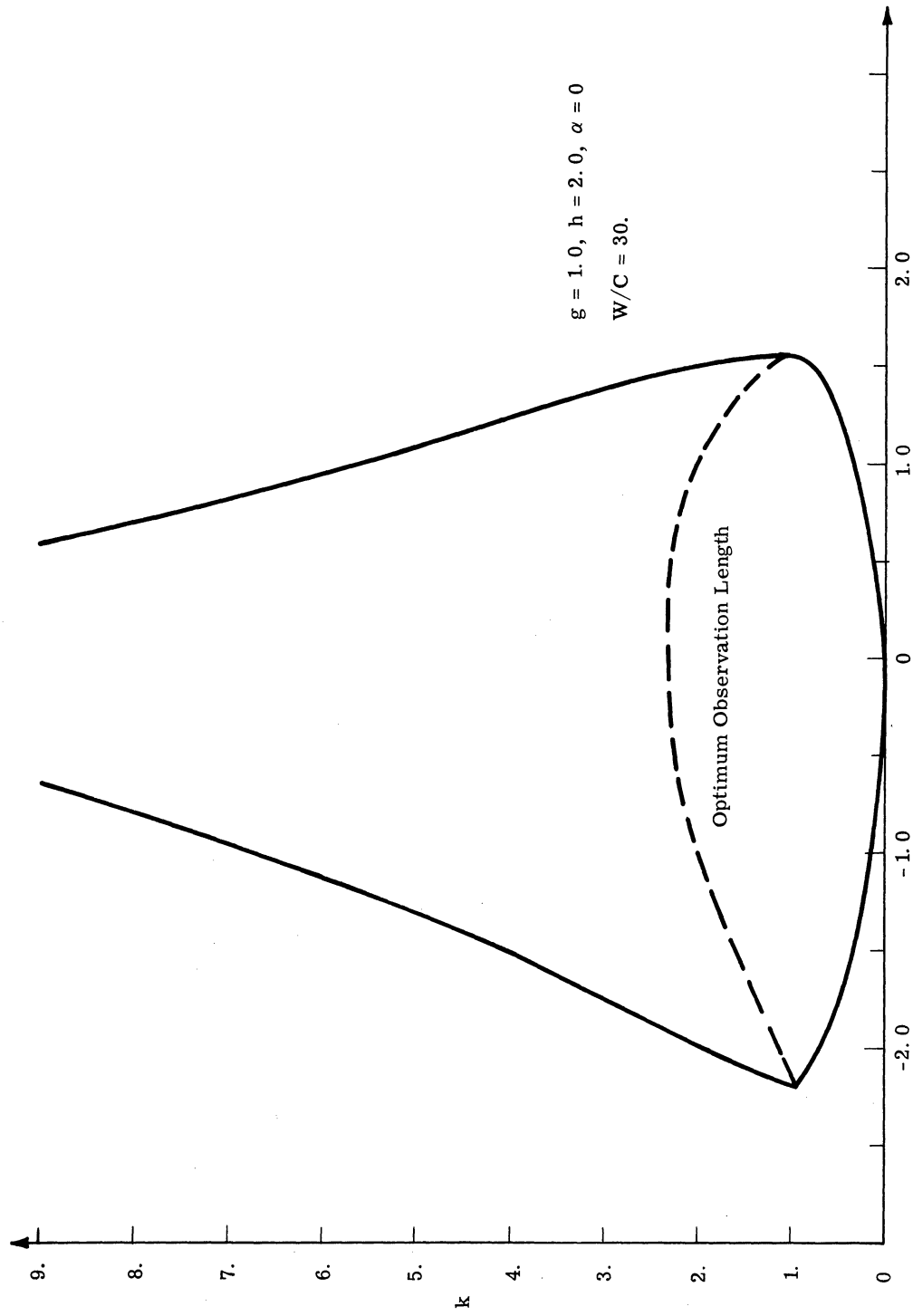


Fig. 4.12. The value contour graph for the ONP for a SKEA with parameters W/C = 30, h = 2 and g = 1, and $\alpha = 0$.

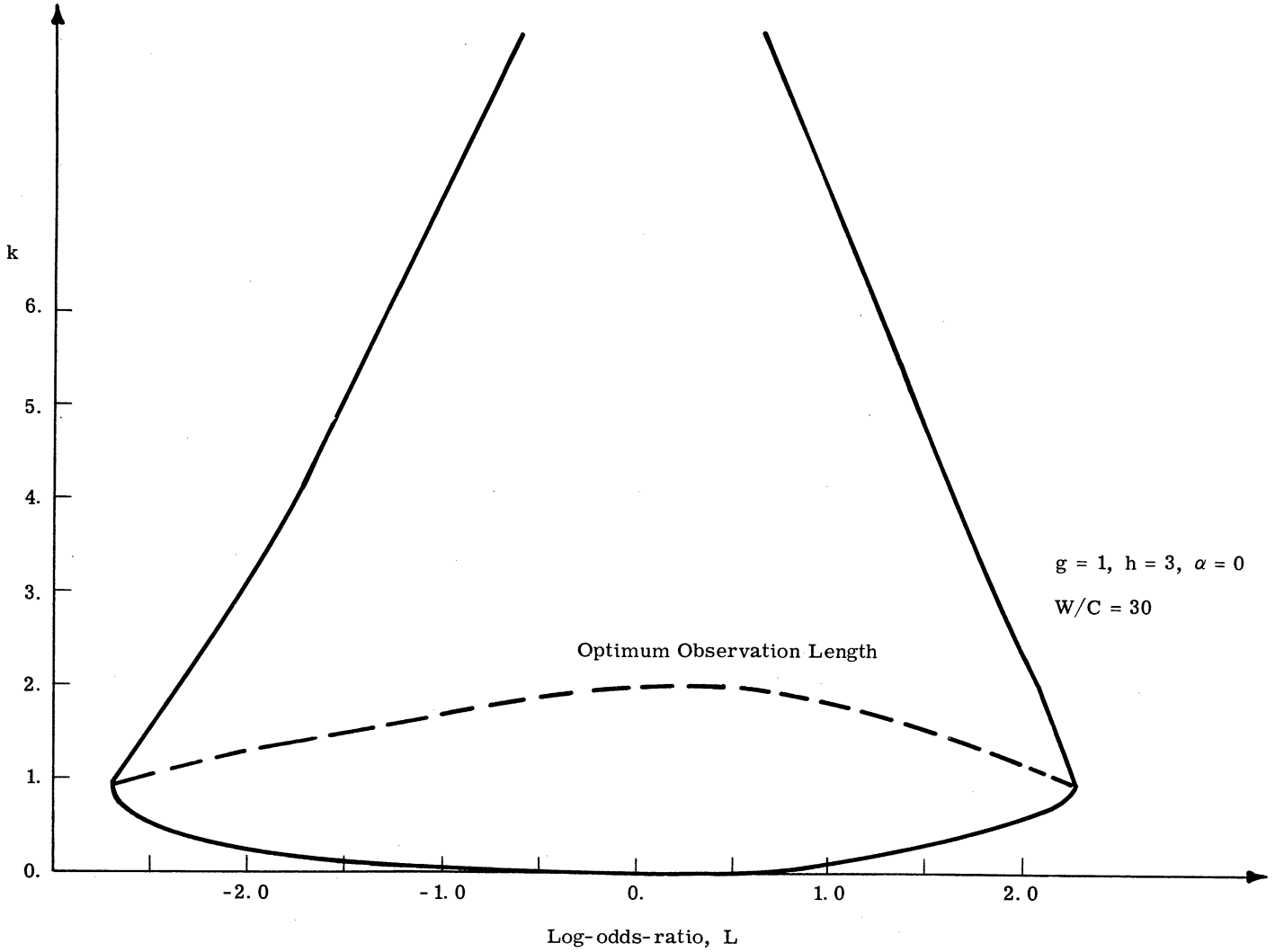


Fig. 4.13. The value contour graph for the ONP for a SKEA with parameters W/C = 30, h = 3 and g = 1, and $\alpha = 0$.

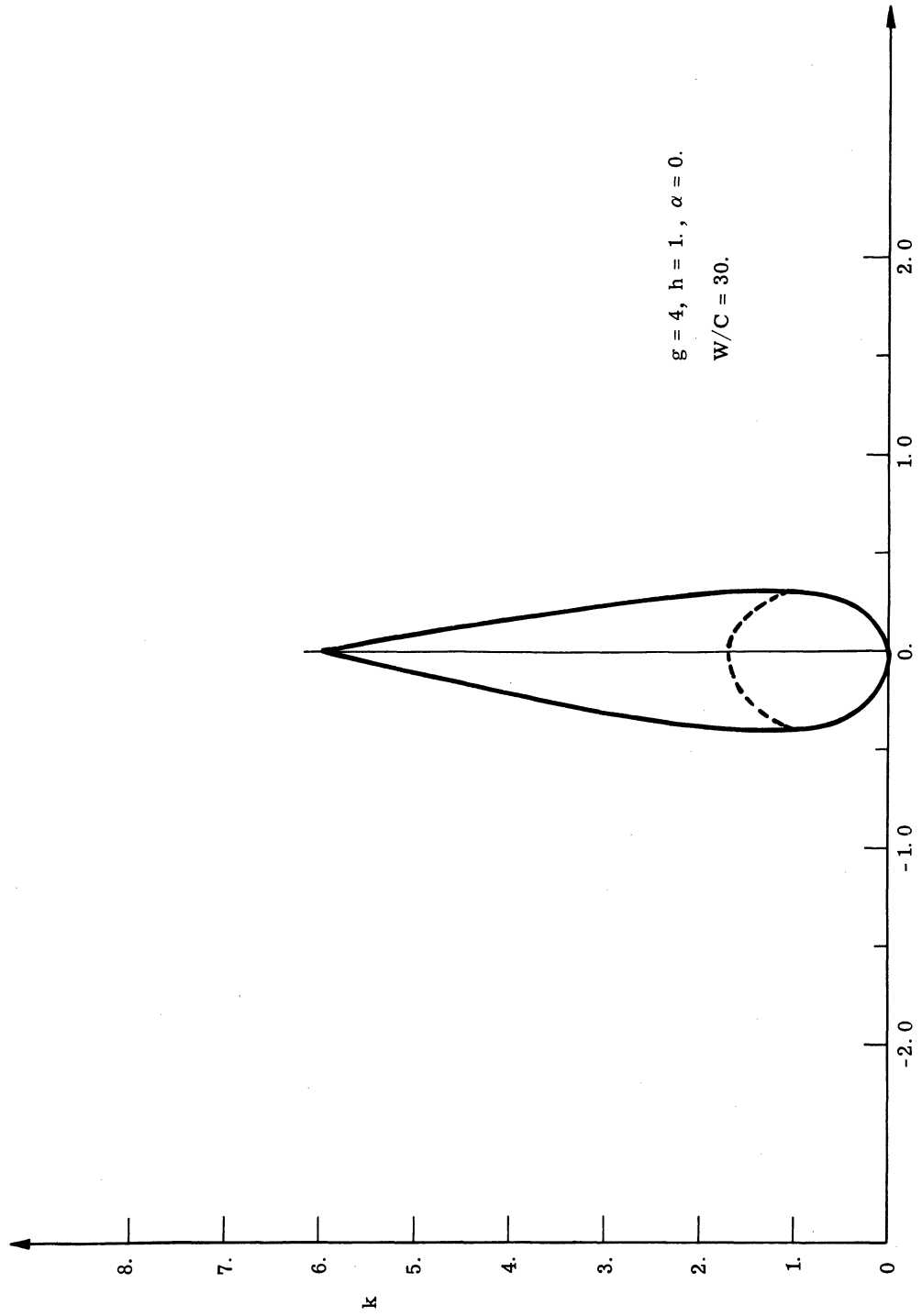


Fig. 4.14. The value contour graph for the ONP for a SKEA with parameters W/C = 30, h = 1 and g = 4, and $\alpha = 0$.

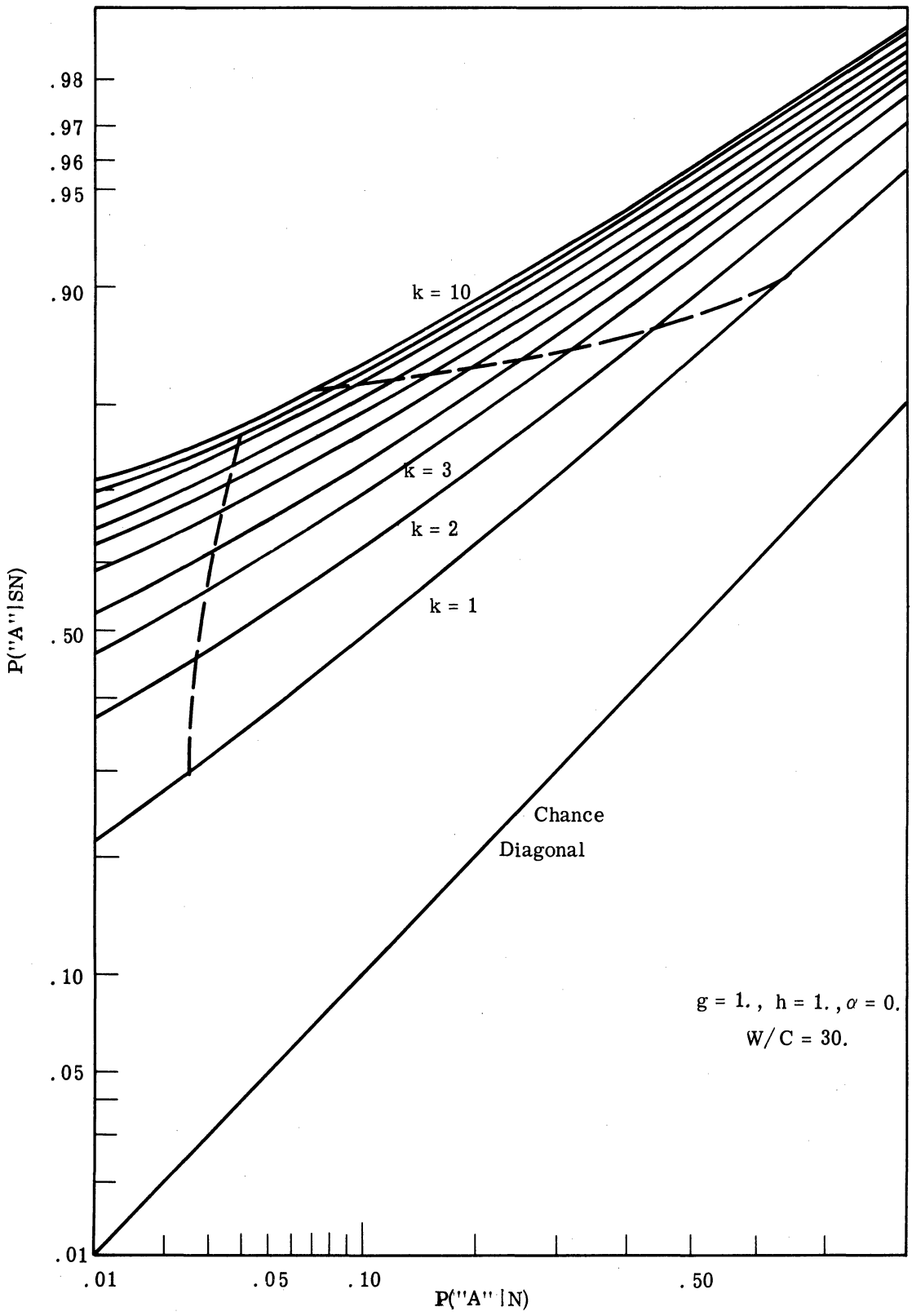


Fig. 4.15. The ROC curves for the ONP for a SKEA with parameters W/C = 30, h = g = 1, and $\alpha$ = 0.

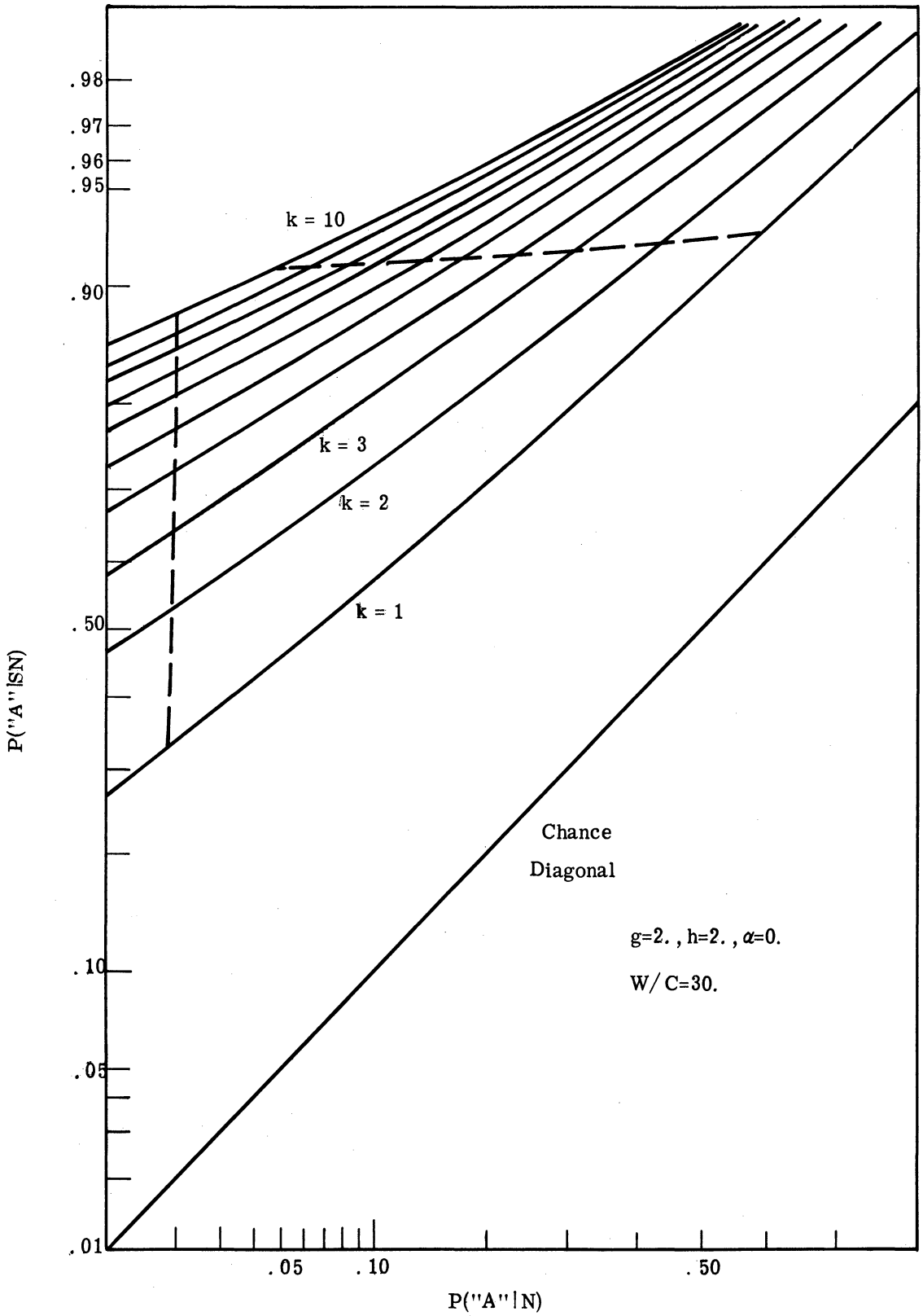


Fig. 4. 16. The ROC curves for the ONP for a SKEA with parameters W/C = 30, h = g = 2, and $\alpha$ = 0.

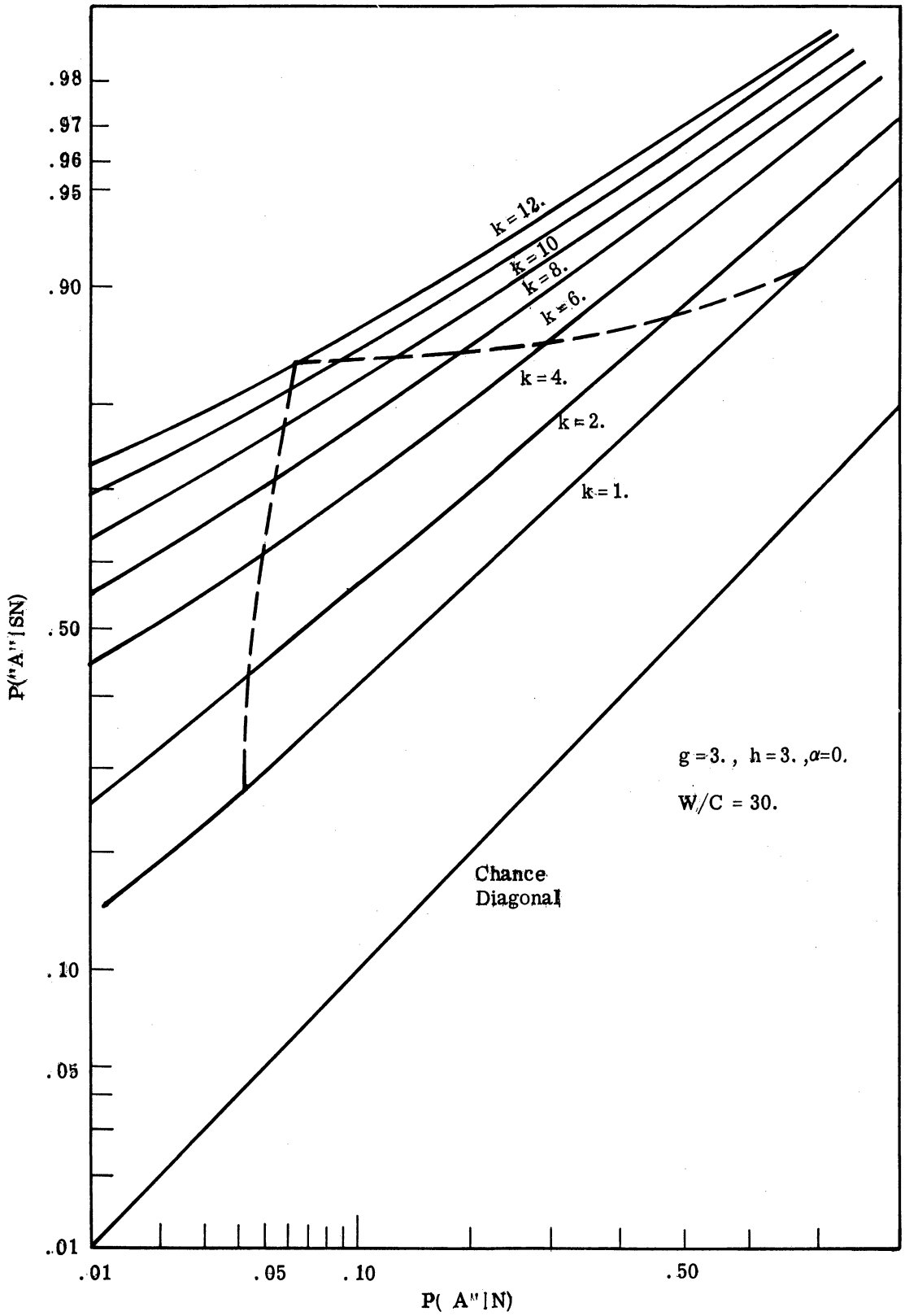


Fig. 4. 17. The ROC curves for the ONP for a SKEA with parameters W/C = 30, h = g = 3, and $\alpha = 0$.

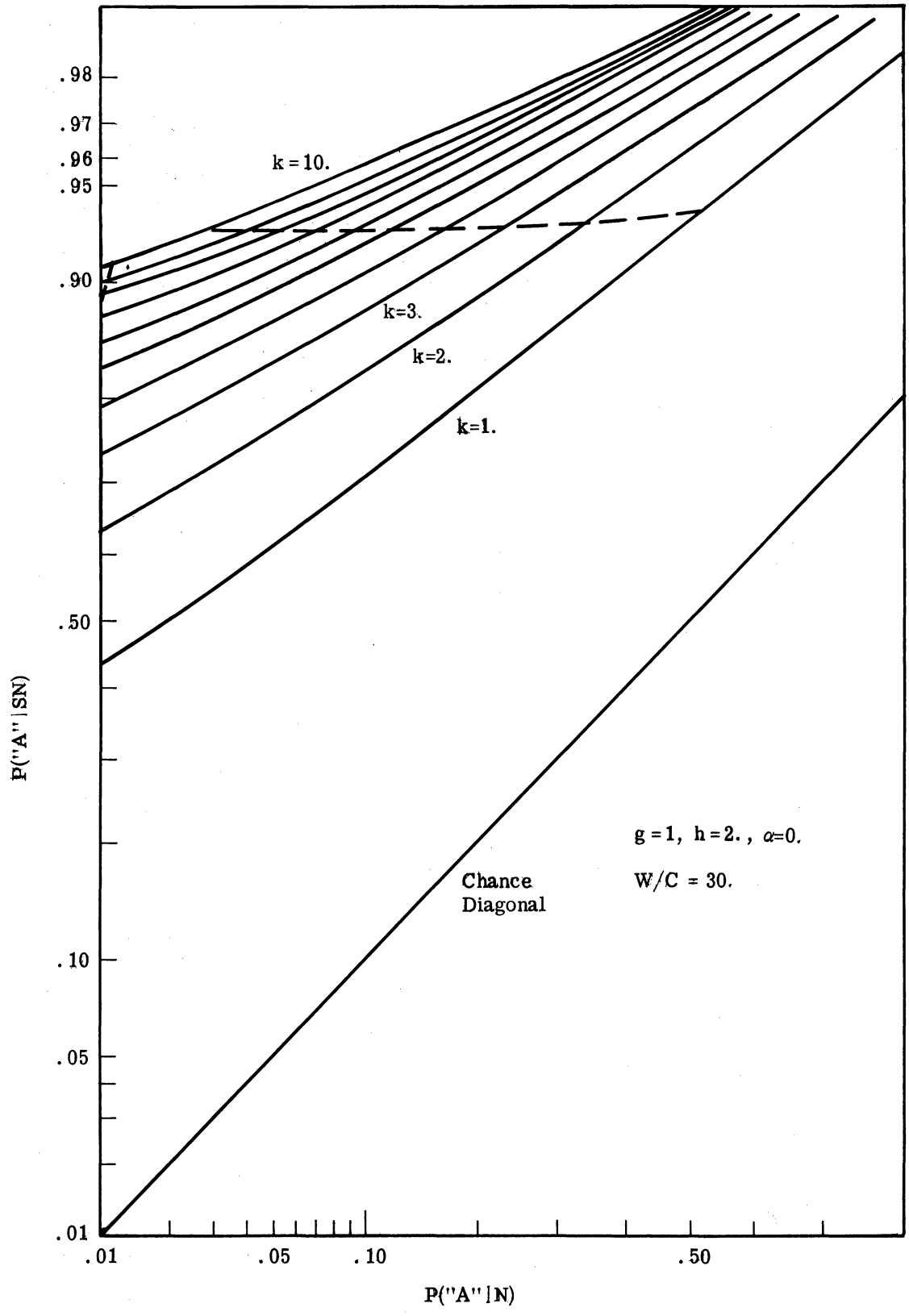


Fig. 4. 18. The ROC curves for the ONP for a SKEA with parameters W/C = 30, h = 2 and g = 1, and $\alpha$ = 0.

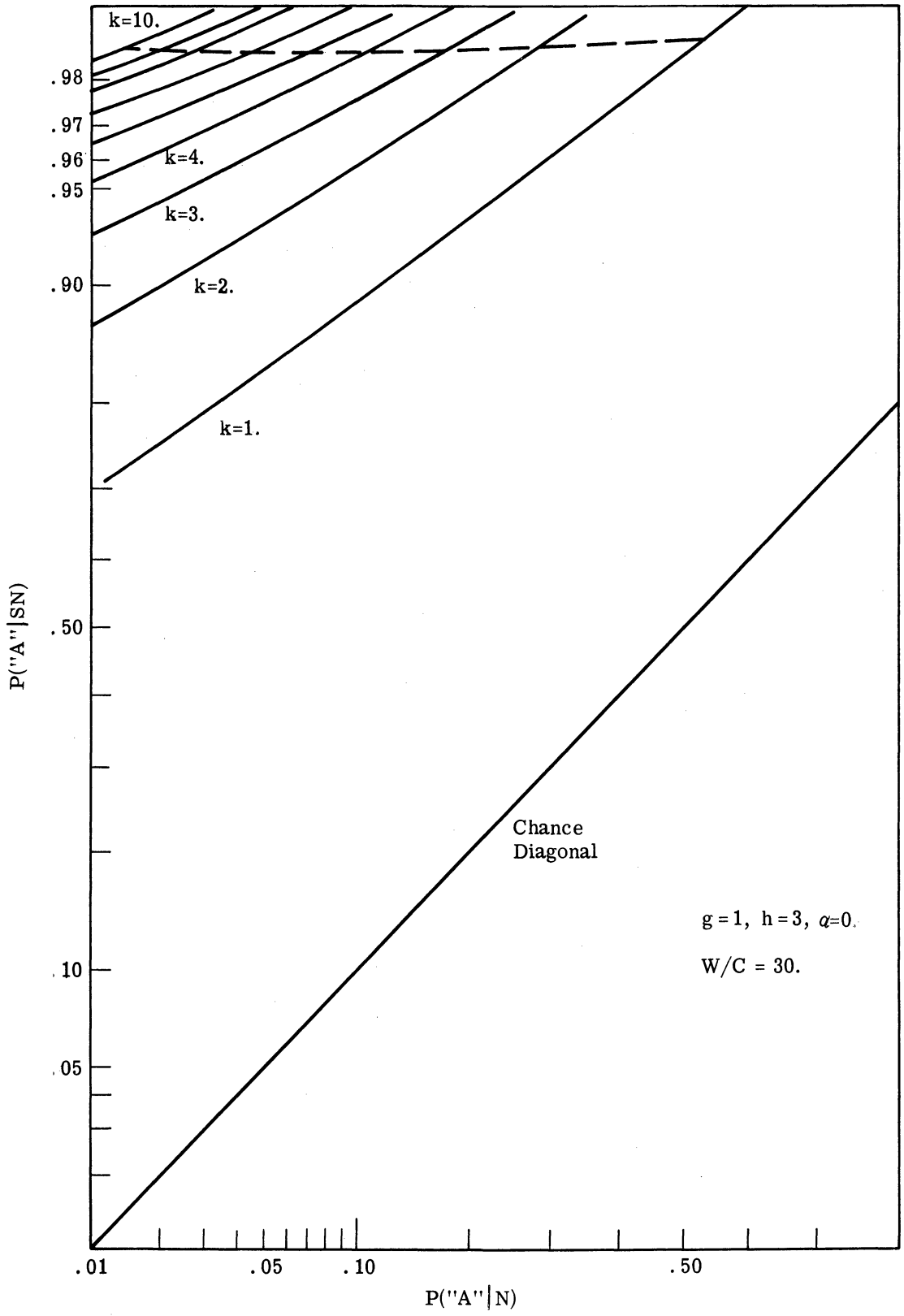


Fig. 4.19. The ROC curves for the ONP for a SKEA with parameters W/C = 30, h = 3 and g = 1, and $\alpha$ = 0.

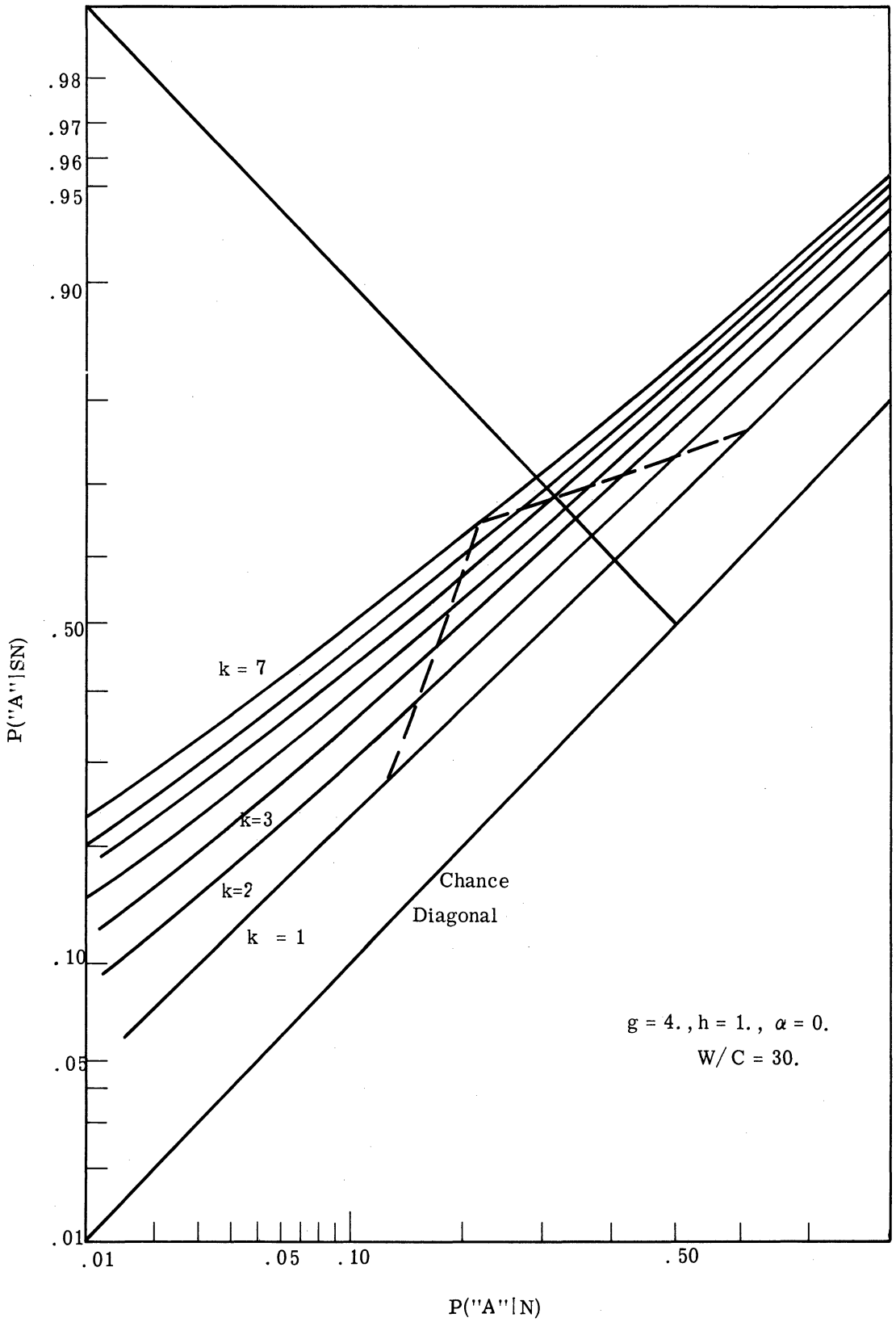


Fig. 4.20. The ROC curves for the ONP for a SKEA with parameters W/C = 30, h = 1 and g = 4, and $\alpha = 0$.

Referring to Figs. 4.9 through 4.14, we note that the zero value contour is skewed toward negative log-odds ratio. As explained in Section 3.4.2, an observation is profitable only for those situations in which the appropriate terminal decision is not the terminal decision dictated by the original log-odds ratio. In other words, observations are profitable only when it is possible to change the observer's opinion of the cause of the input. The skewness toward negative log-odds ratio in Figs. 4.9 through 4.14 results because there is a greater indication of signal-and-noise for negative L's than there is of noise alone. Thus, if we compare amplitude distributions with the same mean but different variances, the distributions with the larger variances should be skewed farther toward negative L's. This last statement is verified by comparing the value contour graphs for $h = g = 1$, $h = g = 2$, and $h = g = 3$.

The ROC curves of Figs. 4.15 through 4.20 show the trace of the zero value contour. (In two of the figures only one side of the trace falls within the scale of the graph.) The trace of the zero value contour shows the region of the ROC for profitable operation. This operating region is relatively small and is confined approximately to the region centered about the negative diagonal.

Specifying the state of the decision process in terms of the separable subspaces, Q, we obtain value contour graphs shown in Figs. 4.21 through 4.26. The contour graphs are plotted as functions of the number of observations and the log-odds ratio. For each log-odds ratio and value of g, the corresponding amplitude distribution can be found by solving for the value of h through the definiing equation for the subspace. The numerical results displayed in the graphs of this section were obtained by means of an IBM 7090 programmed according to the programs of Appendix C.

The value contour graphs obtained by specifying the initial state in terms of Q and g skew toward positive L's. Recall that when the state is given in terms of h and g, the value contour graphs skew toward negative L's. However, the same explanation can be given for both phenomena. A few of the amplitude distributions for the ROC curves of Figs. 4.21 through 4.26 are shown in Figs. 4.27 through 4.31. Because of the subspace specification each log-odds ratio, for fixed g, results in a new amplitude distribution. We see that as g

increases, for a given L, the mean of the signal amplitude distribution shifts toward zero and the variance decreases. As the log-odds increases, for fixed g, the mean of the distribution shifts away from zero. The latter statement implies that a positive value of L is more likely to be changed by observation toward negative L's than vice versa, because noise alone is a more definite observation than is signal-plus-noise. This is a result of the Q surface curvature. On a Q surface the distribution of amplitude and the opinion of presence of a signal are not independent. If the log-odds ratio is negative, then the amplitude distribution shifts toward very small amplitude distributions. On the other hand, if the log-odds ratio is positive, the amplitude distribution tends to be displaced from the origin.

Referring to the value contour graphs of Figs. 4.21 through 4.26, we see that as g increases, for Q fixed, the value contour graphs shrink. This is caused by a shift of the mean of amplitude distribution toward zero and a decrease in the variance as g increases. The observer's opinion of the signal amplitude is, that if the signal is present, it is small. Small signal amplitudes mean that there is less possibility of the observer changing his a priori log-odds ratio. Thus, it becomes less profitable for one to observe for large values of L, positive or negative, as g becomes larger.

The concept of Q surfaces is new, and thus it is helpful to examine many combinations of the parameters Q, g, and L to obtain an intuitive understanding of Q surfaces. For example, if the value of g is held fixed and the value of the Q surface is decreased, we obtain results similar to those depicted in Figs. 4.30 and 4.31. We see that as the value of Q decreases, for a given g value, the amplitude signal distribution becomes less difuse. It follows that it is more profitable to observe as the value of Q decreases, if all other parameters remain the same. Therefore, the value contour graphs should increase in size as Q decreases in value. This is illustrated in Figs. 4.25 and 4.26.

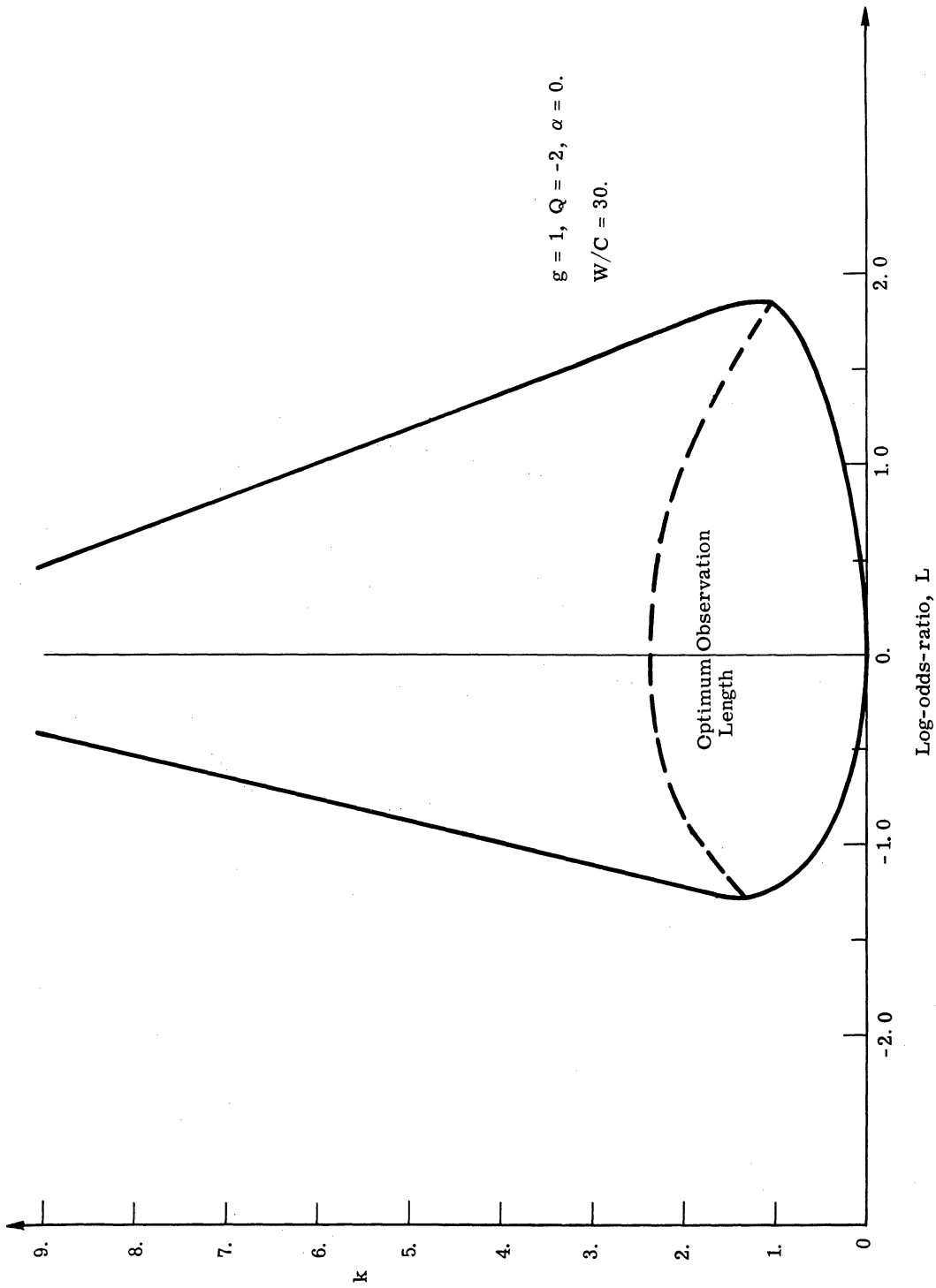


Fig. 4.21. The value contour graph for the ONP for a SKEA with parameters W/C = 30, g = 1, and Q = -2.

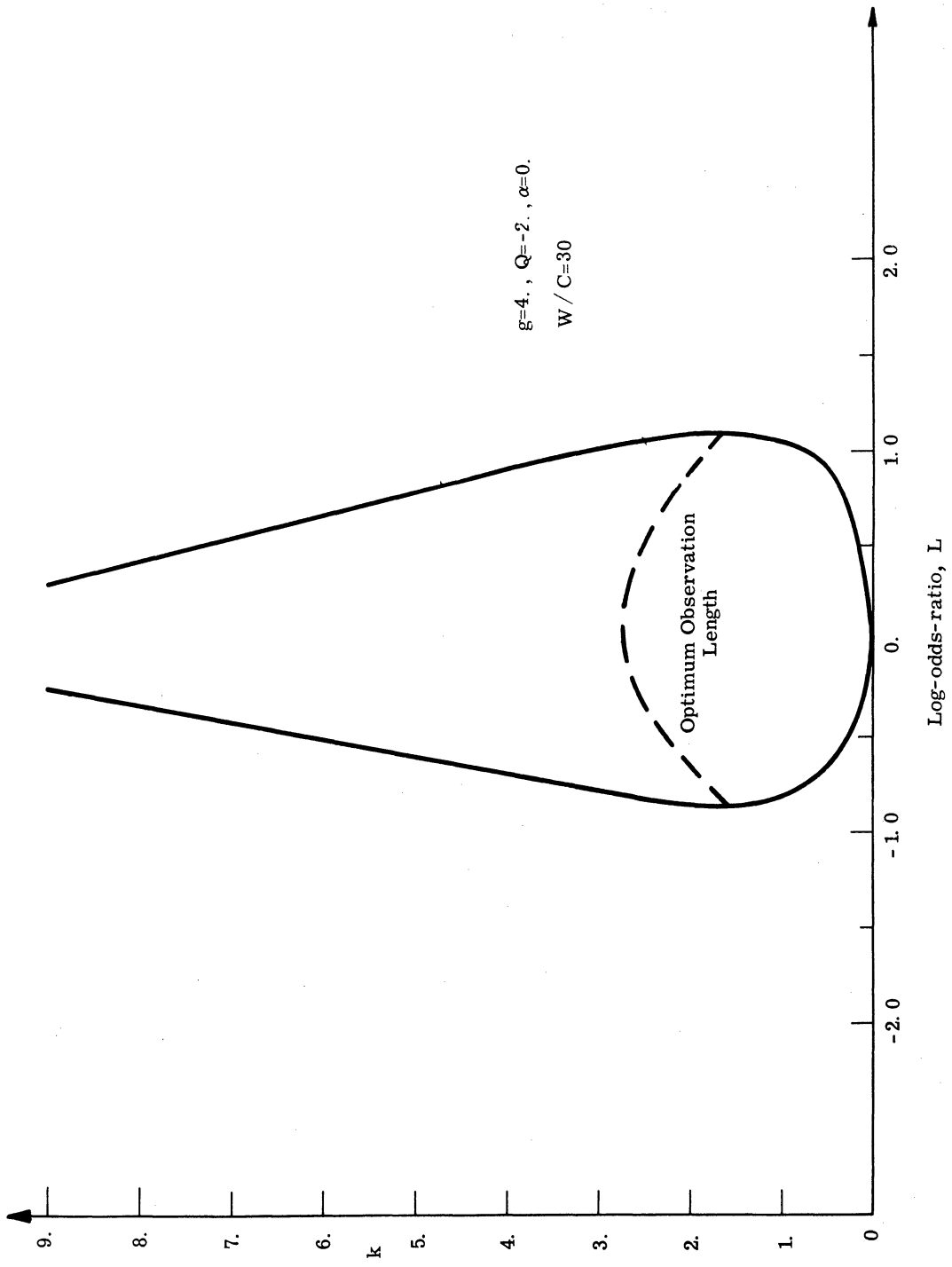


Fig. 4. 22. The value contour graph for the ONP for a SKEA with parameters W/C = 30, g = 4, and Q = -2.

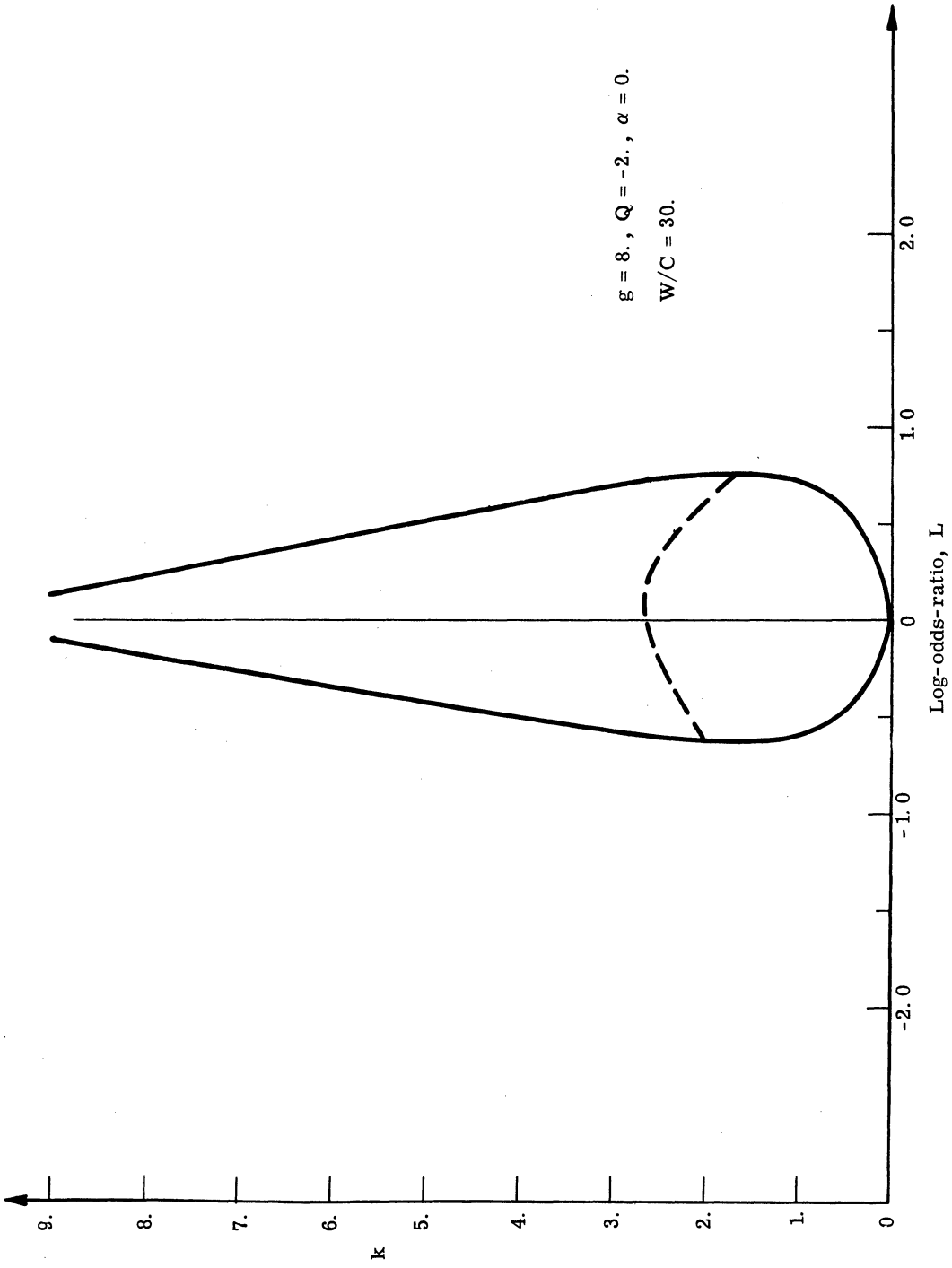


Fig. 4.23. The value contour graph for the ONP for a SKEA with parameters W/C = 30, g = 8, and Q = -2.

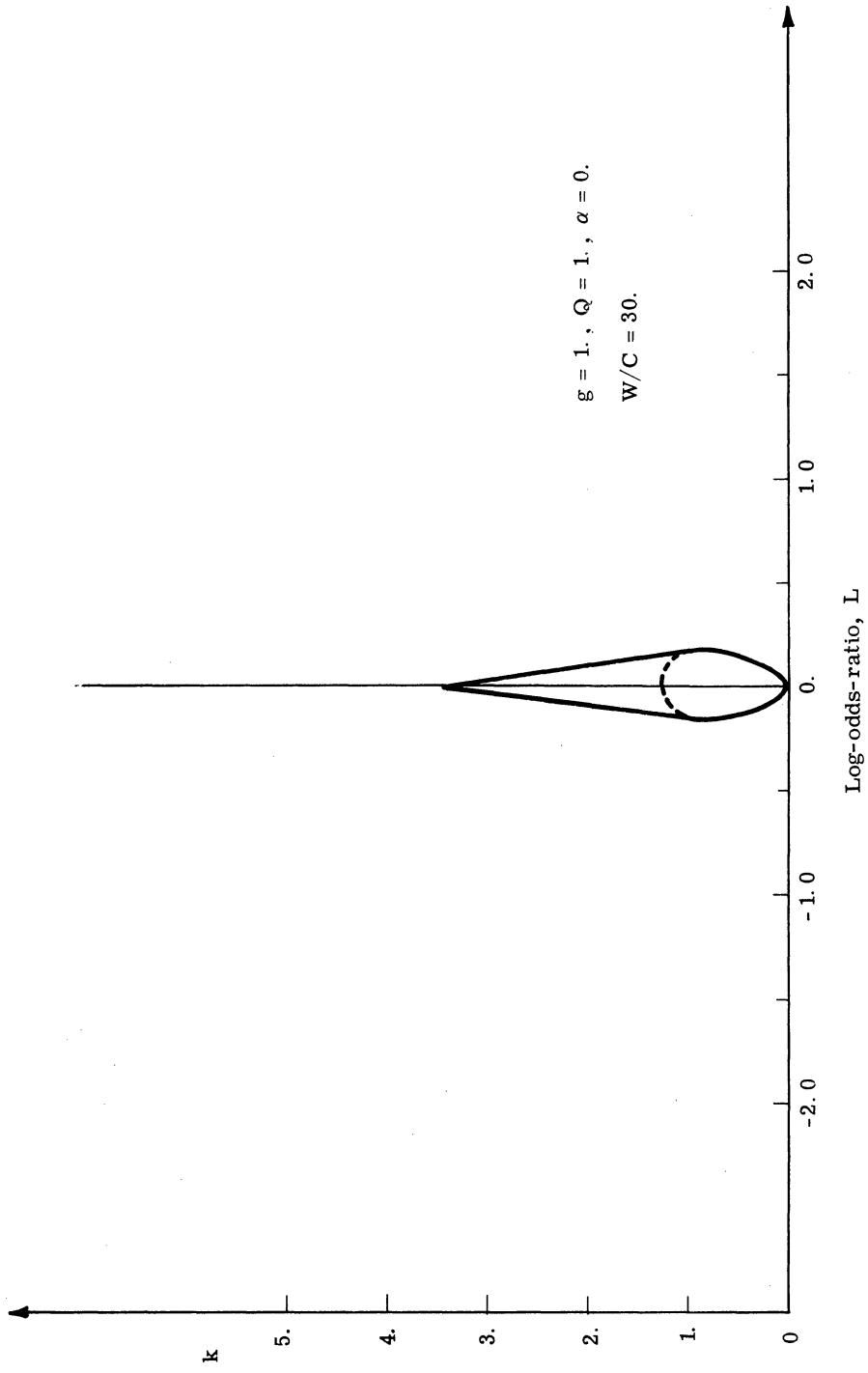


Fig. 4.24. The value contour graph for the ONP for a SKEA with parameters W/C = 30, g = 1, and Q = 1.

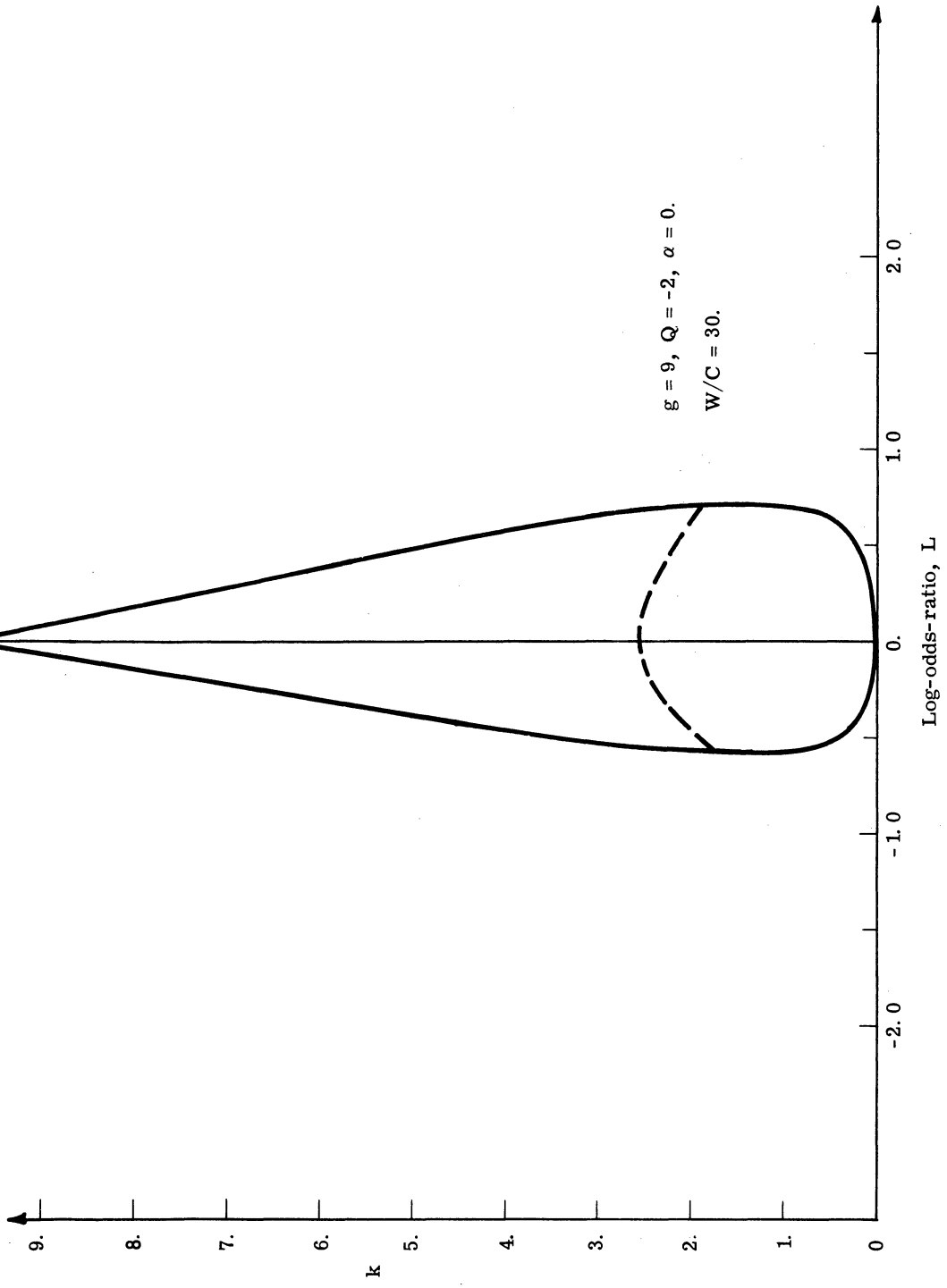


Fig. 4.25. The value contour graph for the ONP for a SKEA with parameters W/C = 30, g = 9, and Q = -2.

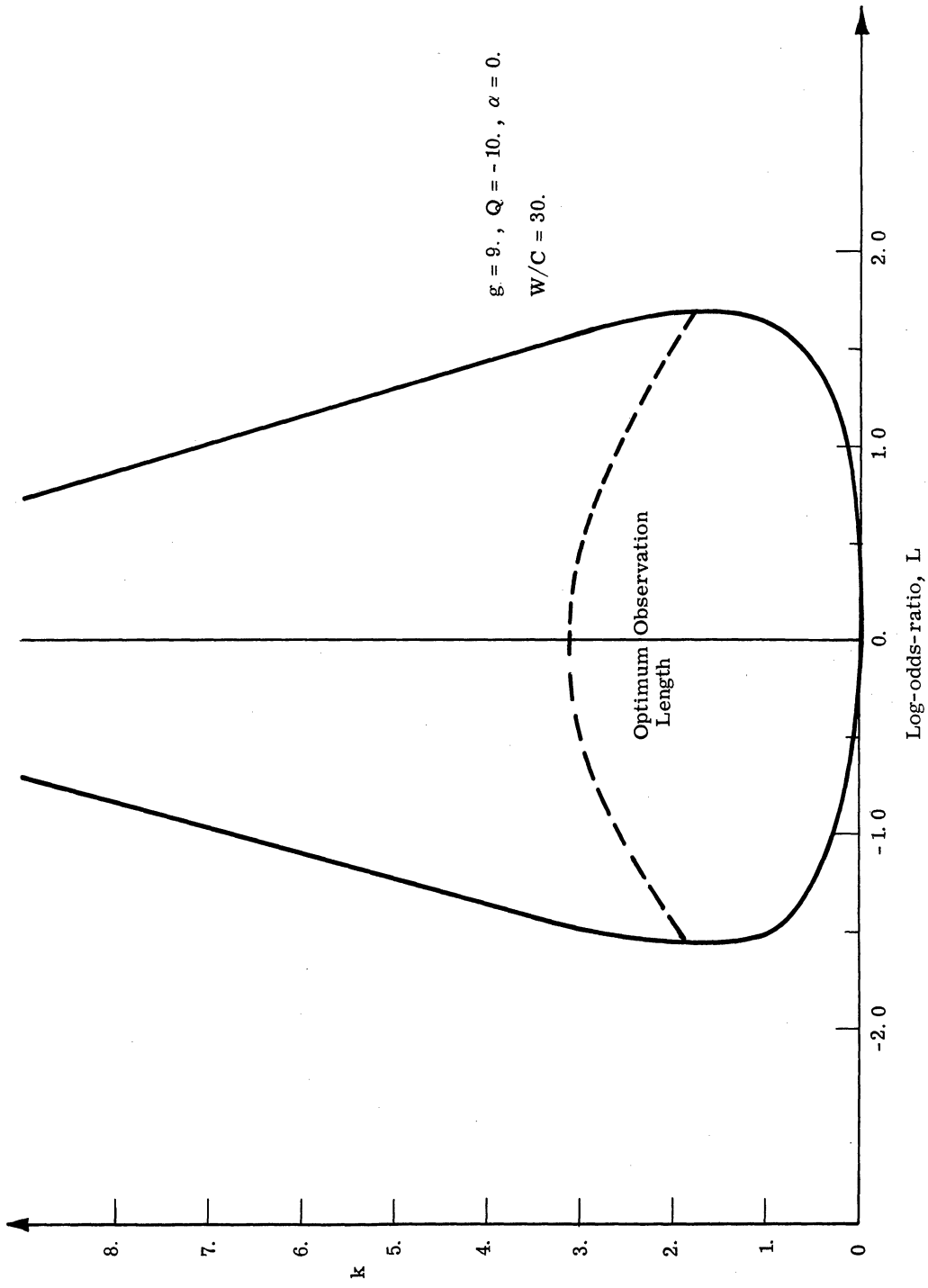


Fig. 4.26. The value contour graph for the ONP for a SKEA with parameters W/C = 30, g = 9, and Q = -10.

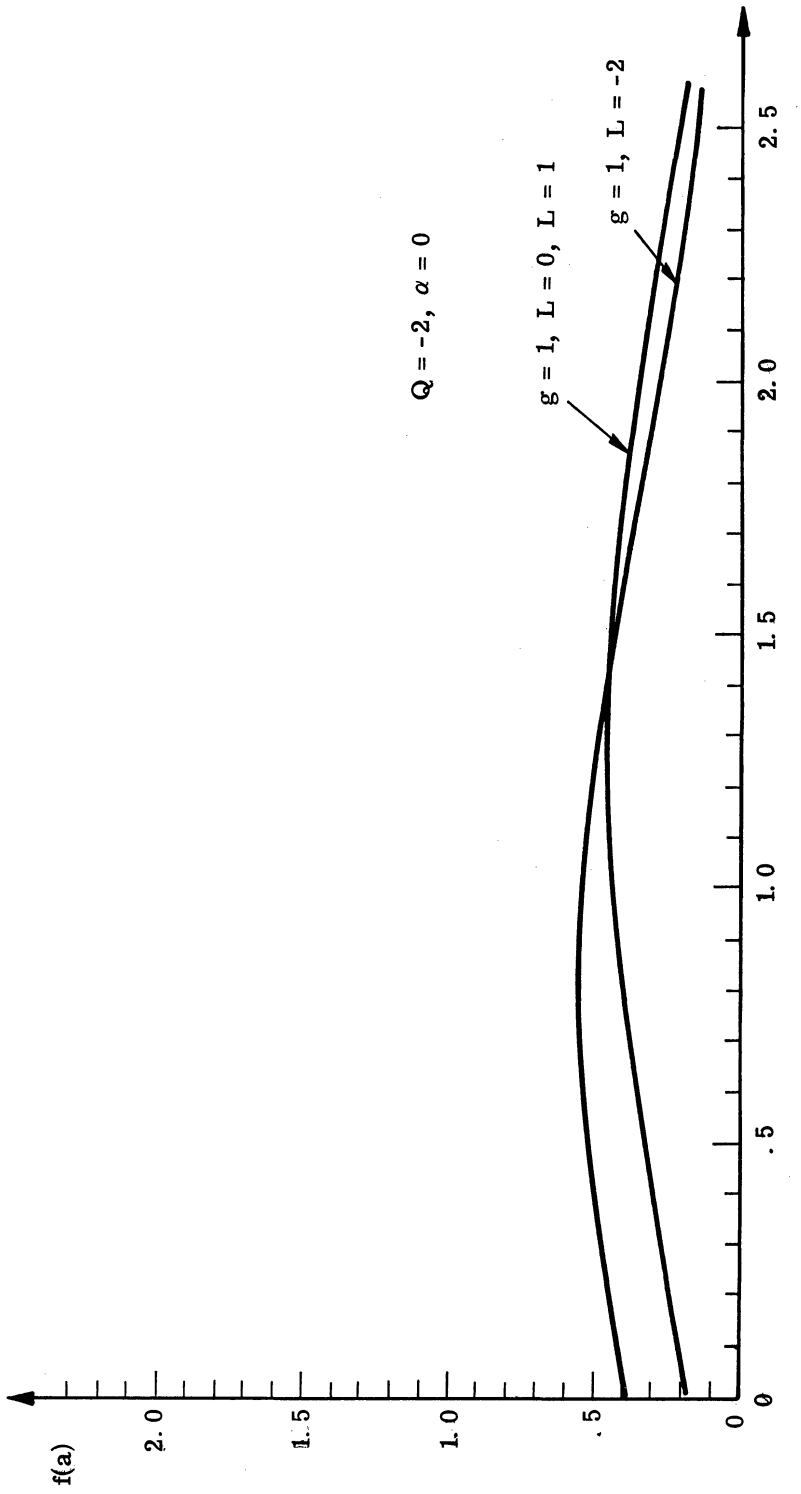


Fig. 4.27. Graphs of truncated Gaussian probability density functions with parameters g = 1, Q = -2, and L = -1, 0, and 1.

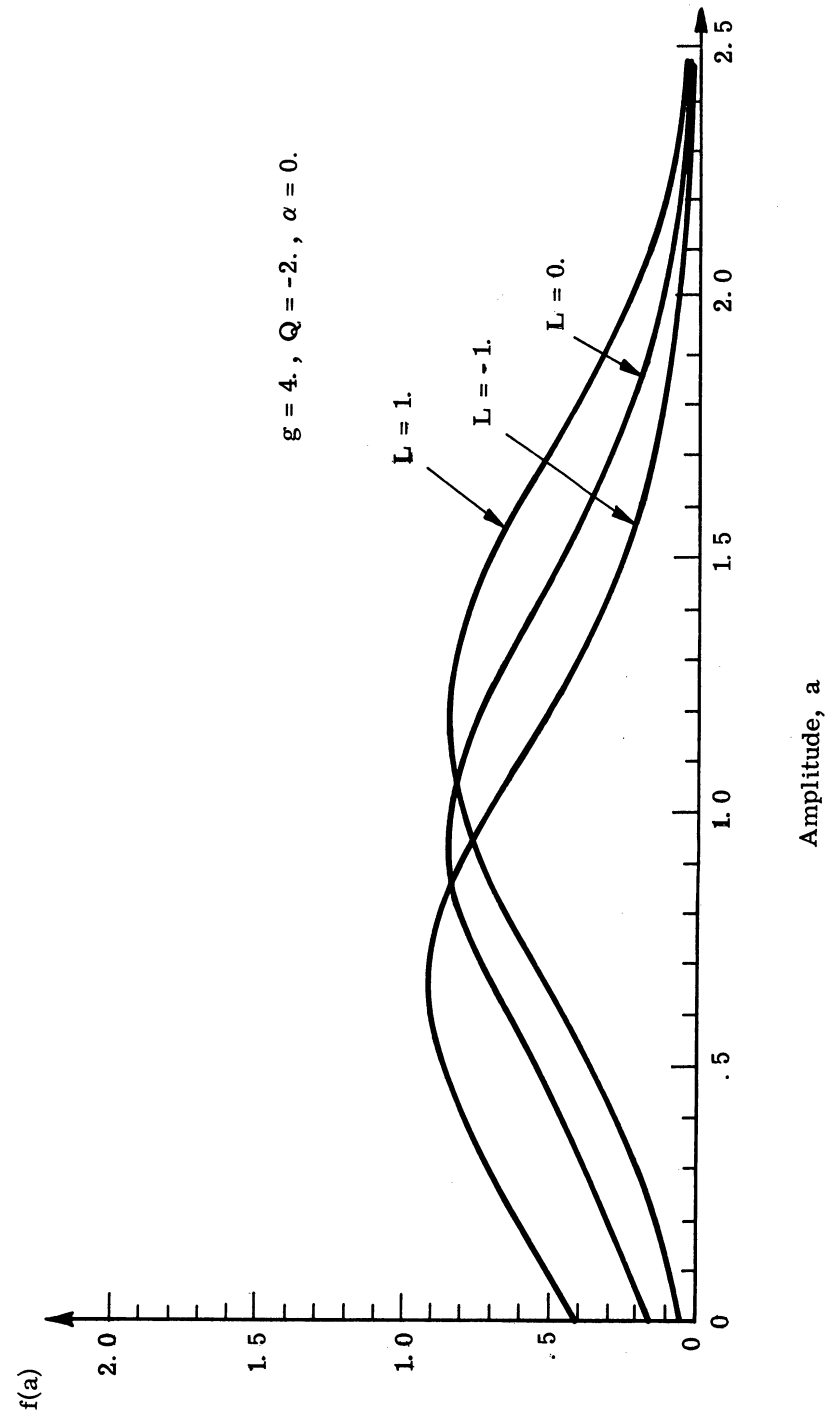


Fig. 4.28. Graphs of truncated Gaussian probability density functions with parameters g = 4, Q = -2, and L = -1, 0, and 1.

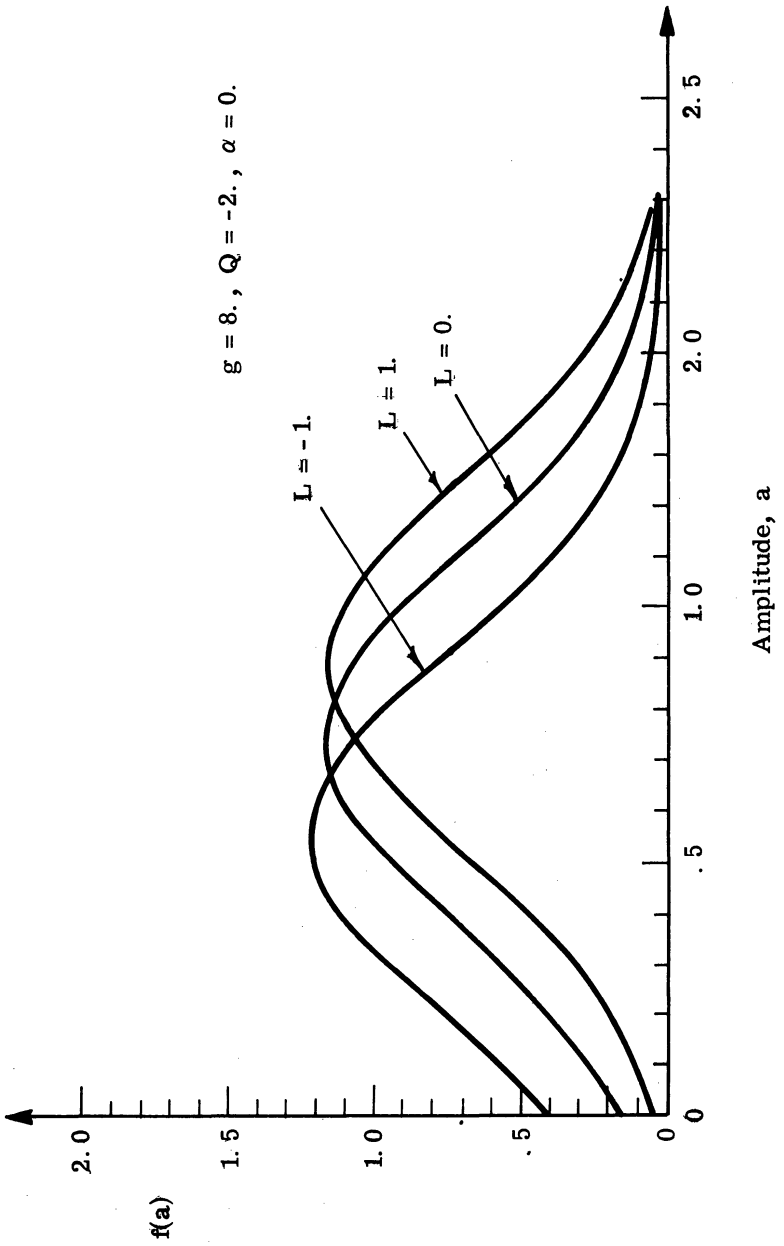


Fig. 4.29. Graphs of truncated Gaussian probability density functions with parameters g = 8, Q = -2, and L = -1, 0, and 1.

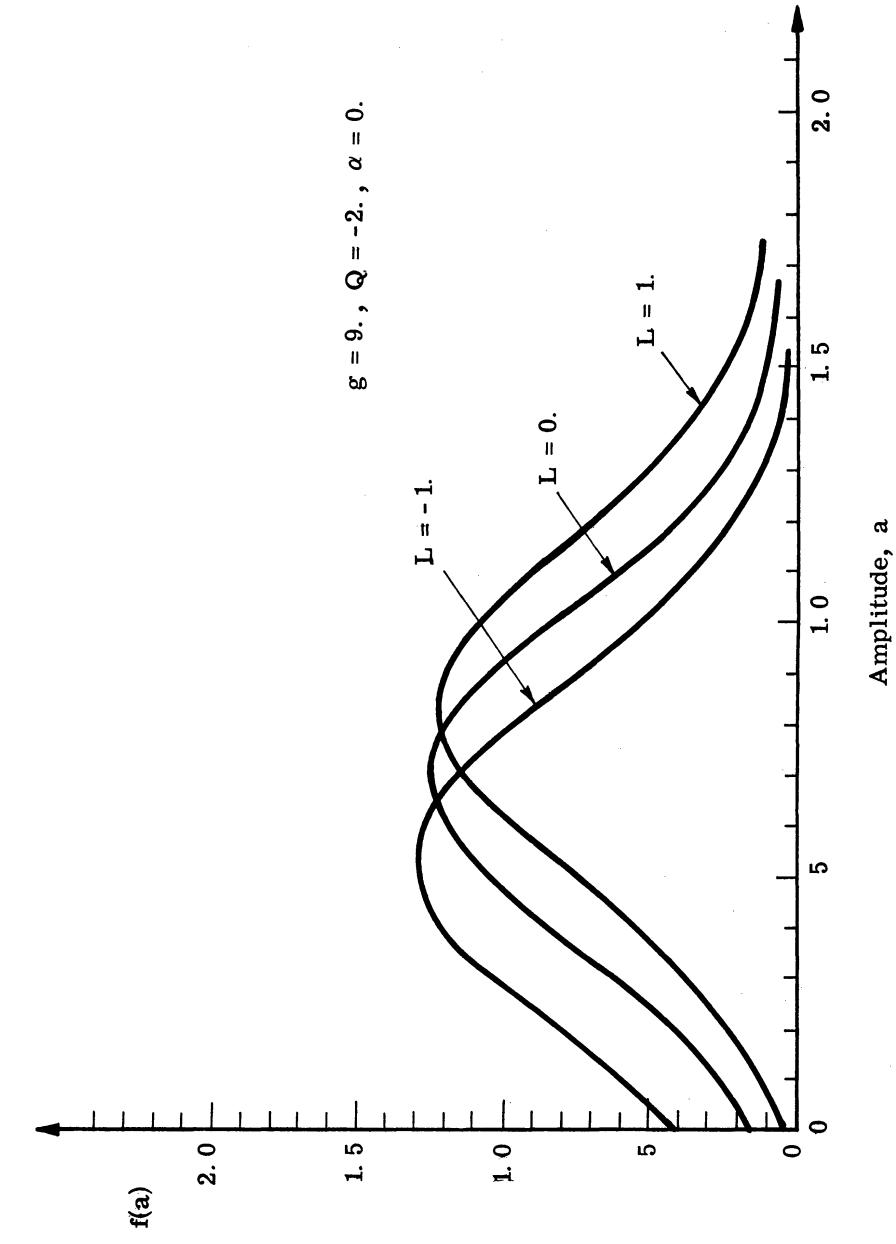


Fig. 4.30. Graphs of truncated Gaussian probability density functions with parameters g = 9, Q = -2, and L = -1, 0, and 1.

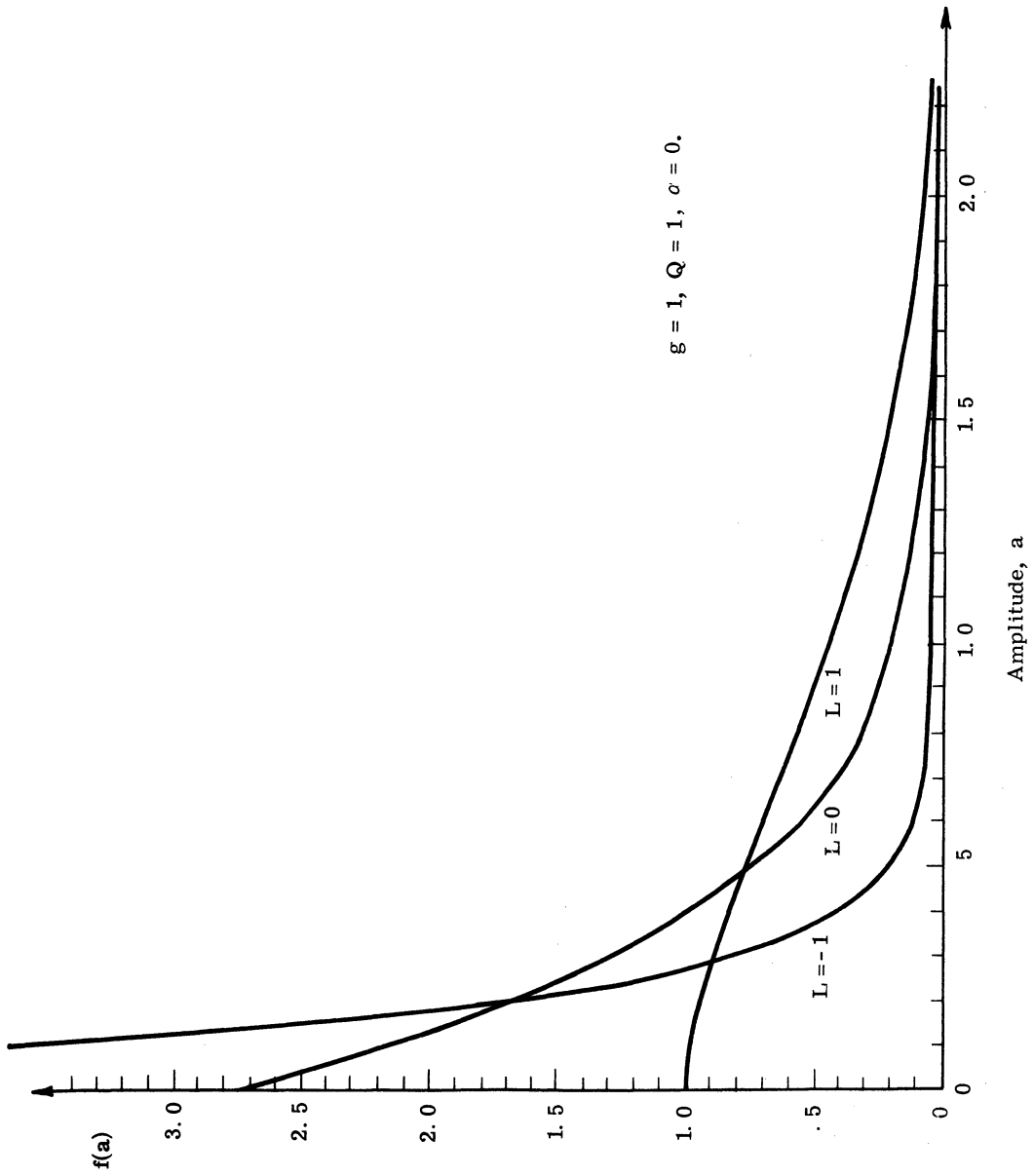


Fig. 4.31. Graphs of truncated Gaussian probability density functions with parameters g = 1, Q = 1, and L = -1, 0, and 1.

The corresponding ROC curves for the value contour graphs of Figs. 4.21 through 4.26 are shown in Figs. 4.32 through 4.37. We again note the convergence phenomena of the ROC curves as longer observations are taken. The dotted lines on Figs. 4.32 through 4.37 are the trace of zero value contour and they are a function of the W/C ratio. If the W/C ratio increases, there is a corresponding widening of the zero value trace. In Fig. 4.32 the optimum ROC curve results from observing the optimum amount of time as determined from the value contour graph of Fig. 4.22. Notice the optimum ROC differs greatly from a normal ROC. (A normal ROC plots as a straight line with a slope of one.)

## 4.5 Deferred Decision Theory for a Signal Known Except for Amplitude

### 4.5.1 Derivation of the Deferred Decision Procedure.

The derivation of the optimum sequential procedure for an unknown amplitude signal parallels our previous derivation in Section 3.5 concerning the SKEP. Numerical results are obtained by means of an IBM 7090 as hand calculations are too tedious and long to be of any practical use. Each class of amplitude distribution creates its own special problems both with the mathematics and with the numerical computations. We have completed derivations for the $\alpha = 0$, $\alpha = 1$, and Gaussian signal amplitude distributions. However, in order to facilitate easier reading of this material, we will present only the detailed mathematics and computations for the $\alpha = 0$ signal amplitude distribution. The corresponding derivations for the $\alpha = 1$ and Gaussian distribution are relegated to the appendices. Gross characteristics of the decision problem do not depend on whether the analytical form of the amplitude distribution is the $\alpha = 0$, $\alpha = 1$, or Gaussian distribution.

As we have discussed previously (see Sections 2.4 and 3.5), to effect an economical computer solution, the state of a sequential procedure must be specified in terms of the separable subspaces. We have already defined the separable subspaces for the three amplitude distributions under consideration in Section 4.3.3. The derivation of the Q surface for the $\alpha = 0$ amplitude distribution follows. The derivations of the separable subspaces for $\alpha = 1$ and the Gaussian amplitude distribution are given in Appendix G.

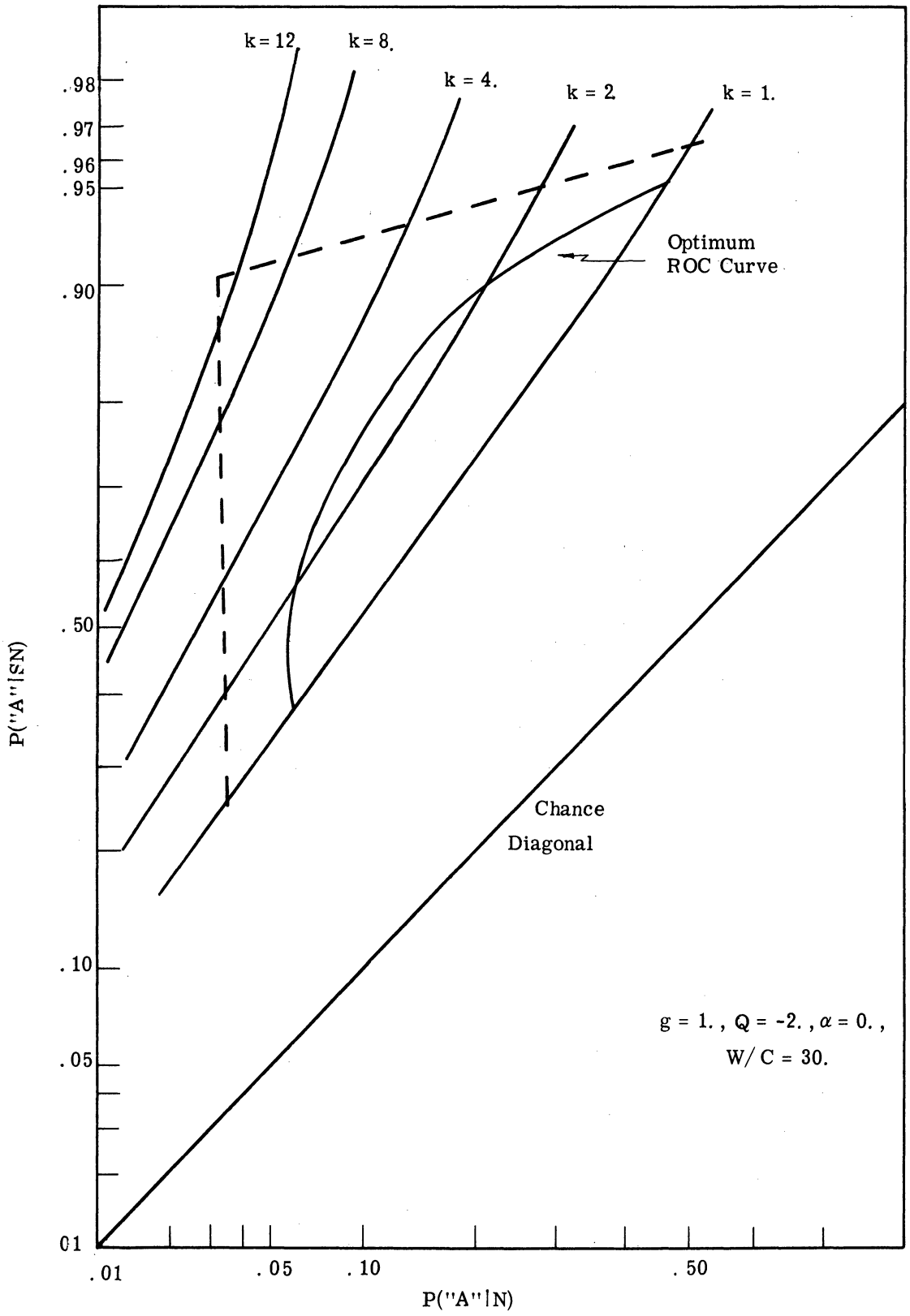


Fig. 4.32. The ROC curves for the ONP for a SKEA with parameters W/C = 30, g = 1, and Q = -2.

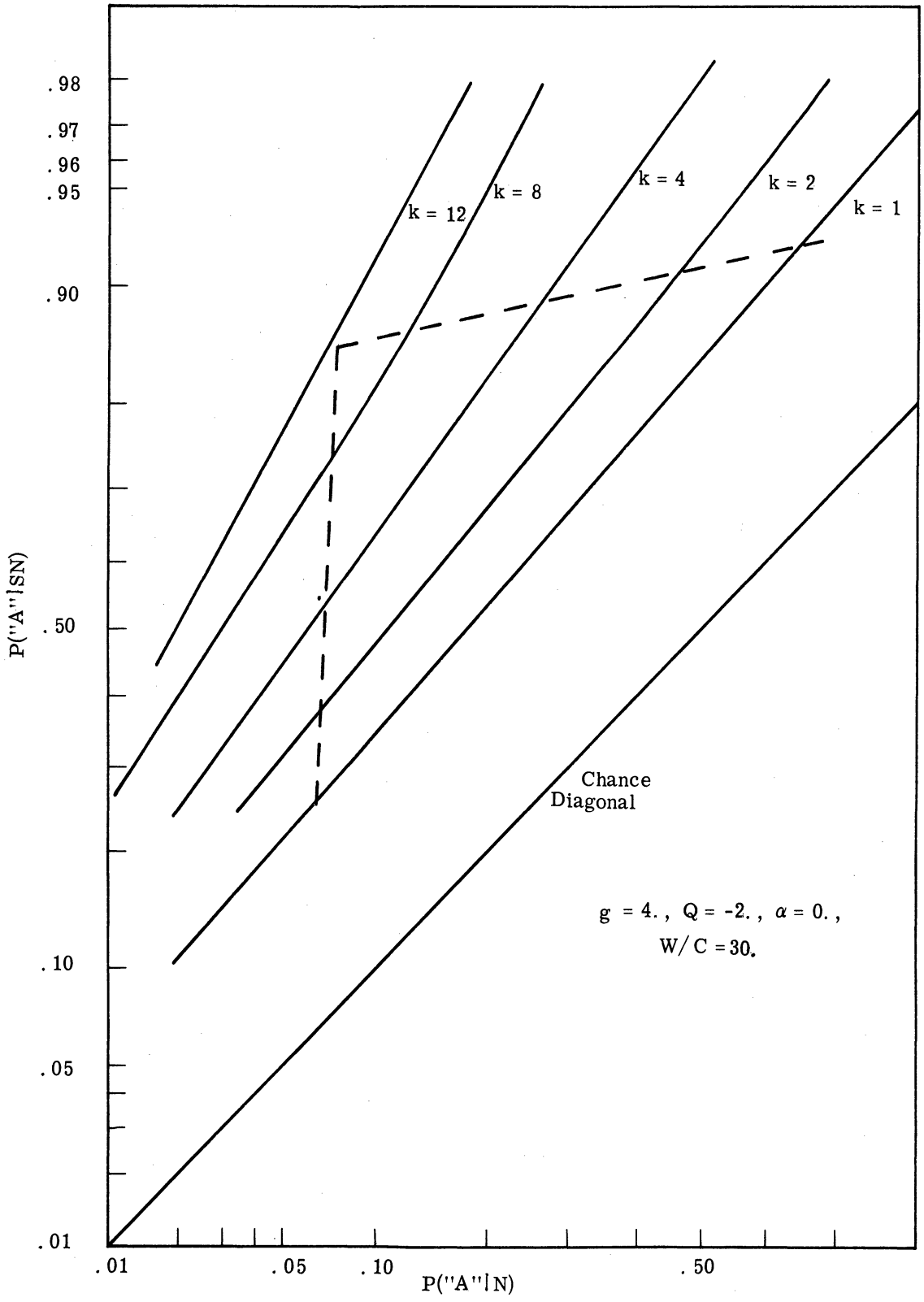


Fig. 4.33. The ROC curves for the ONP for a SKEA with parameters W/C = 30, g = 4, and Q = -2.

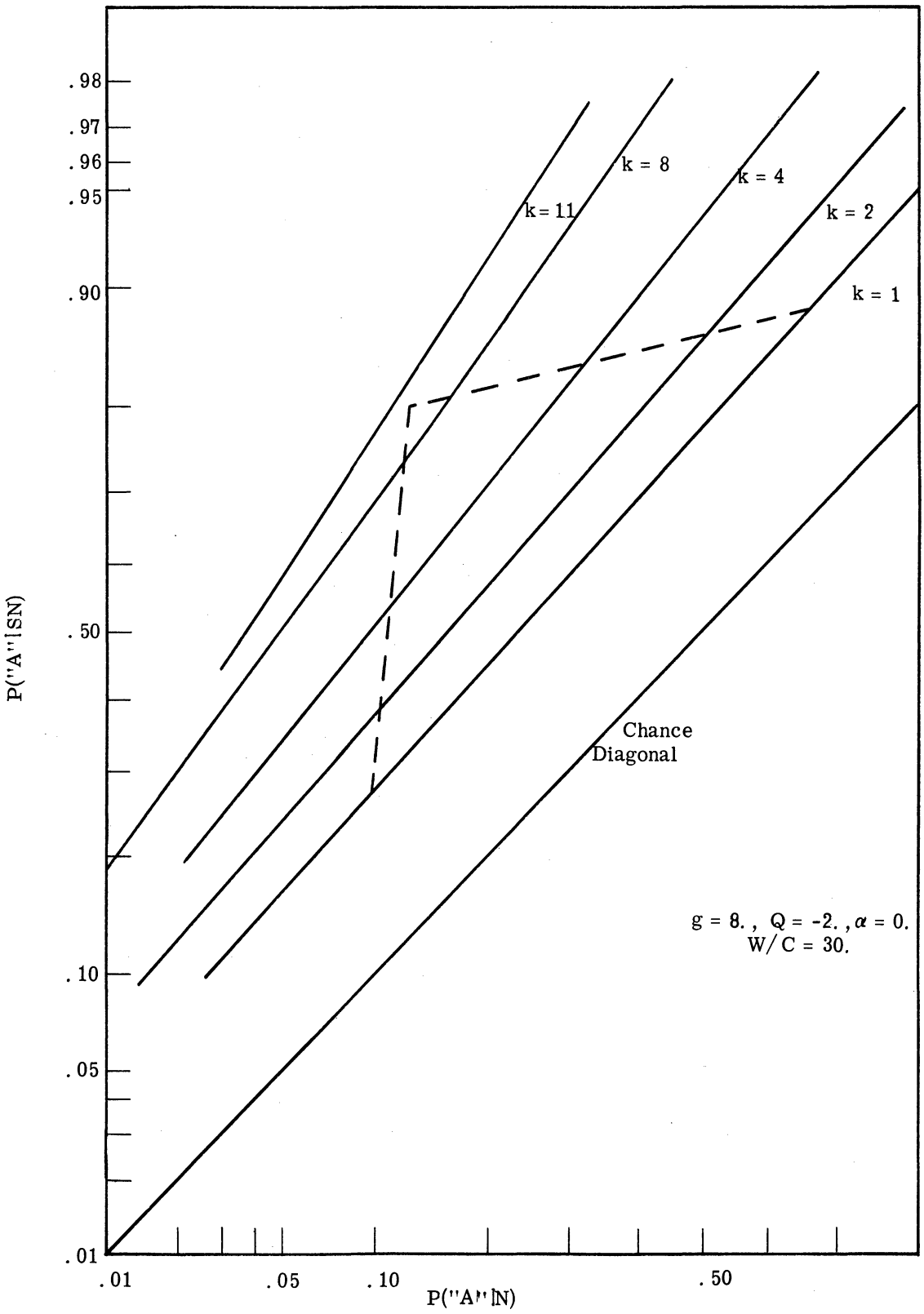


Fig. 4.34. The ROC curves for the ONP for a SKEA with parameters W/C = 30, g = 8, and Q = -2.

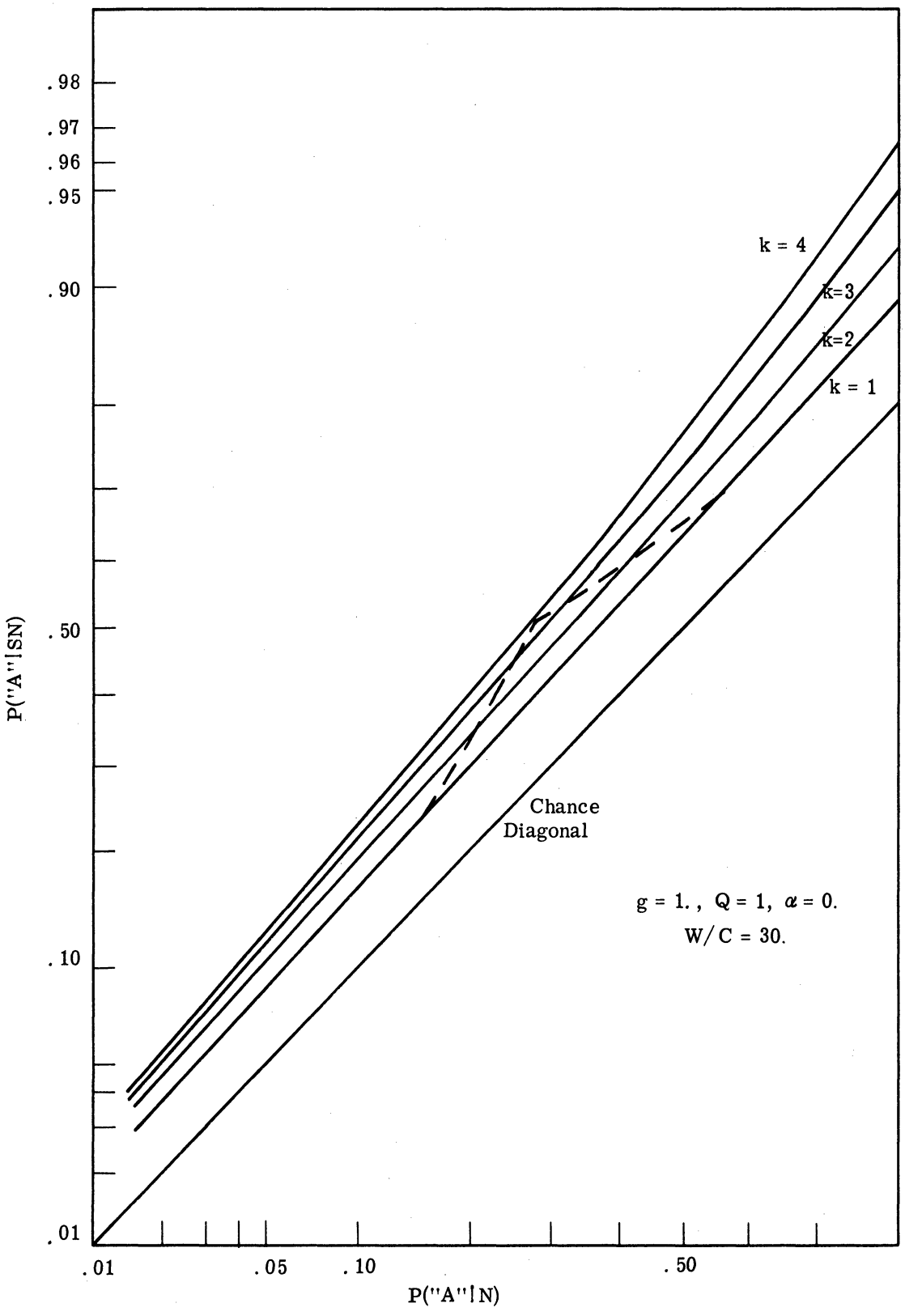


Fig. 4.35. The ROC curves for the ONP for a SKEA with parameters W/C = 30, g = 1, and Q = 1.

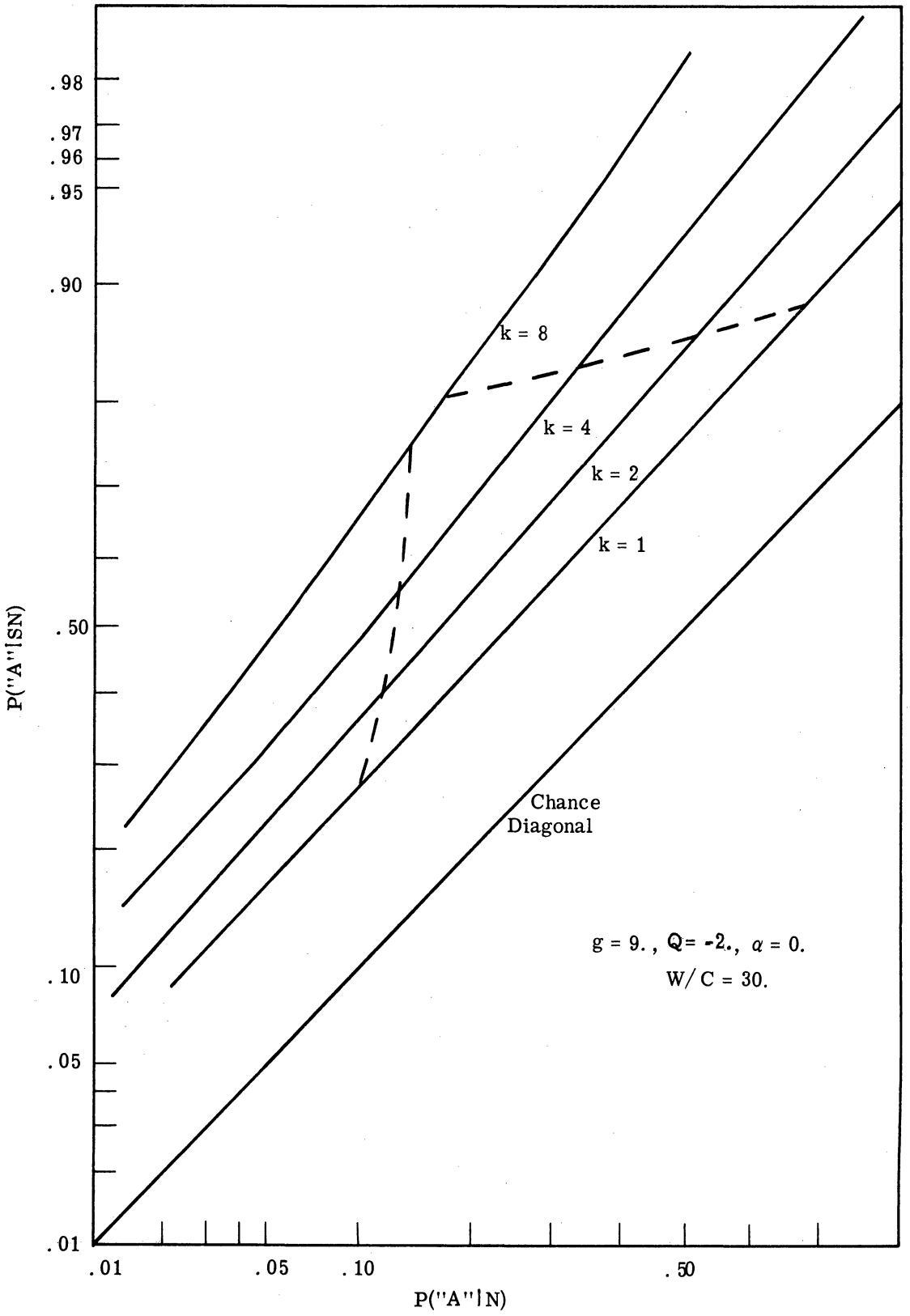


Fig. 4.36. The ROC curves for the ONP for a SKEA with parameters W/C = 30, g = 9, and Q = -2.

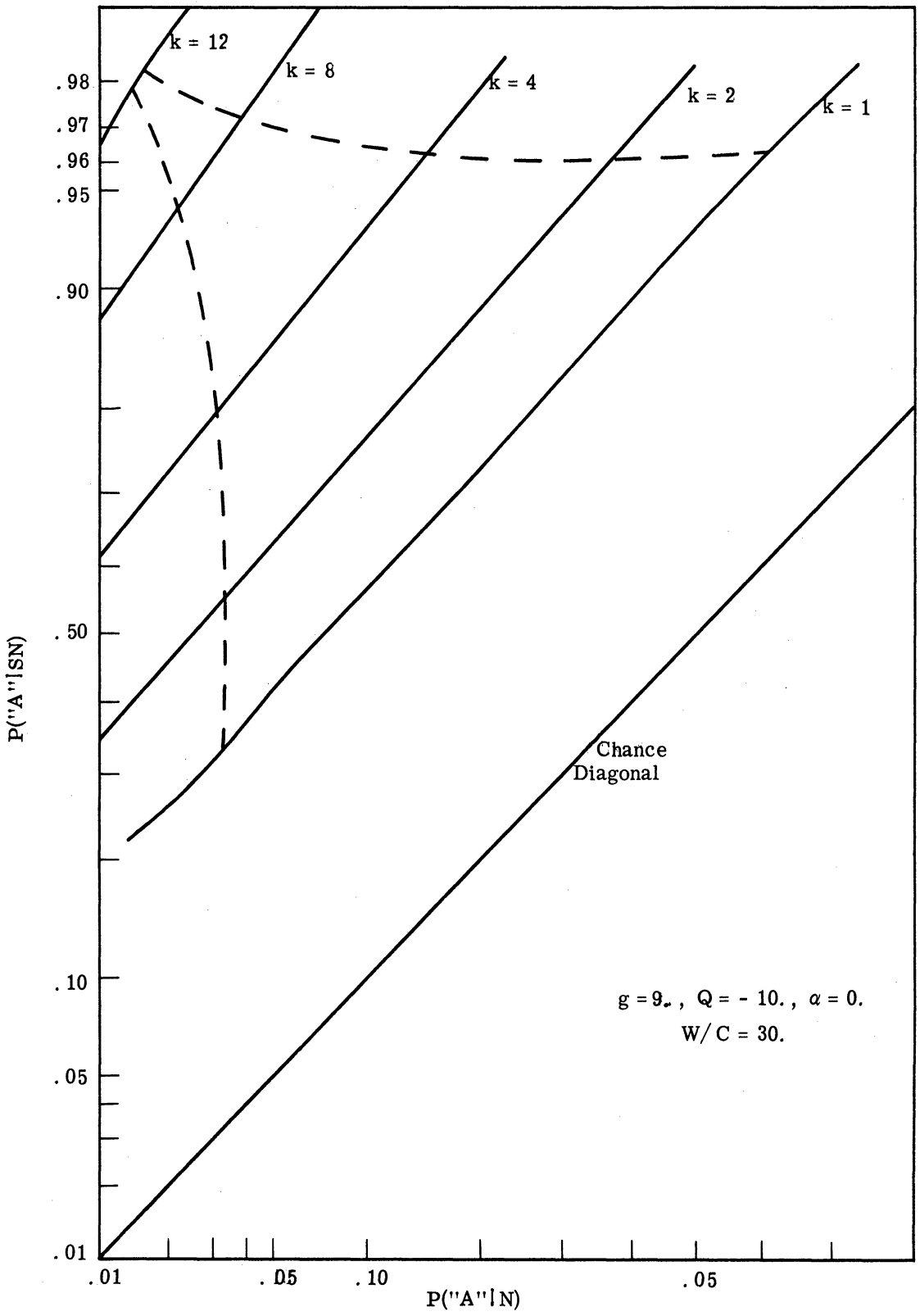


Fig. 4.37. The ROC curves for the ONP for a SKEA with parameters W/C = 30, g = 9, and Q = -10.

The Q surfaces determine the relationship betweeen the log-odds ratio and the amplitude distribution at any stage of observation. The functional form of this relationship is invariant for a given amplitude distribution. If n is the available number of observations in a deferred-decision procedure, then by Bayes' Rule, the log-odds after observation, $L_{n-1}$, is

$$L_{n-1} = L_n + \ell n\, \ell(z) \tag{4.57}$$

Applying the Bayes' Rule again, the updated amplitude distribution is

$$f_{n-1}(a) = f(a|z) = \frac{\ell(z|a)}{\ell(z)} \cdot f_n(a) \tag{4.58}$$

Substituting the expression for the likelihood ratio from Eq. 4.57 into 4.58 and rearranging the terms, we obtain

$$L_{n-1} - L_n = \ell n \left[ \frac{\ell(z|a) \cdot f_n(a)}{f_{n-1}(a)} \right] \tag{4.59}$$

Solving for the a posteriori amplitude distribution, $f_{n-1}(a)$, we obtain the following general expression for determining the separable subspaces.

$$f_{n-1}(a) = \exp\left[L_n - L_{n-1}\right] \ell(z|a)\, f_n(a) \tag{4.60}$$

If the expression for the probability density functions and the conditional likelihood ratio are substituted into Eq. 4.60, we obtain

$$M_{n-1} \exp\left[ah_{n-1} - \frac{a^2 g_{n-1}}{2}\right] = \exp\left[L_n - L_{n-1}\right] \exp\left[az - \frac{a^2}{2}\right] M_n \exp\left[ah_n - \frac{a^2 g_n}{2}\right] \tag{4.61}$$

where $M_{n-1}$ and $M_n$ are the normalizing constant for the density functions. Equation 4.61 can be reduced to

$$\ell n\, M_{n-1} + L_{n-1} = \ell n\, M_n + L_n \tag{4.62}$$

Equation 4.62 implies that $\ell n\, M + L$ is independent of the stage of observation, and thus, defines the invariant relationship between L and amplitude distribution parameters,

g, h, and $\alpha$. Recalling that the normalizing constant M is equal to g $\omega$(h g), we have

$$Q = L + \ell n \left[ \frac{\sqrt{g}}{\omega(h/\sqrt{g})} \right] \tag{4.63}$$

The value of Q as defined above specifies the separable subspaces for the $\alpha = 0$ distribution. The value of Q determines the relationship between L and the amplitude parameters, h and g. By appropriate choice of the Q surface, we can model various relationships between L and the amplitude distribution. For example, if we wish to model a problem in which the signal amplitude is fairly well known, (a large mean value and small variance), then we would choose a large negative value for Q and a large positive value for g.

It is of interest to determine the value of Q corresponding to a signal known exactly so that past work in simple deferred decision theory can be related to the present work. To determine the Q value of a signal known exactly, consider the functional form of the signal amplitude distribution and the role of the parameter g. The quantity g is the reciprocal variance of the amplitude distribution, and thus large values of g correspond to well-defined amplitude distributions. Thus, in order to determine the Q value for a signal known exactly, let the value of g approach infinity while the mean of distribution remains finite. The Q value resulting from the above procedure is a good approximation to the Q surface for a signal known exactly. Using the approximation that $\omega(h/\sqrt{g}) \cong \frac{1}{\sqrt{2\pi}} \exp\left[\frac{h^2}{2g}\right]$, we obtain, by substituting into Eq. 4.63, the following approximate formula for Q.

$$\begin{aligned} Q &\cong L + \frac{1}{2} \ell ng - \frac{h^2}{2g} + \ell n \sqrt{2\pi} \\ &\cong L + \frac{1}{2} \ell ng - \frac{1}{2} (\text{mean})^2 g + \ell n \sqrt{2\pi} \end{aligned} \tag{4.64}$$

As g increases, the value of Q decreases, because the logarithm of g is dominated by the term $\frac{1}{2}$ (mean)$^2$ g. We conclude from this heuristic discussion that the signal known exactly surface is a Q of minus infinity. The above discussion is not rigorous, however, it does point out a useful rule of thumb: as Q becomes more negative the amplitude distributions become less diffuse. This fact was also noted in the discussion of the numerical results of the ONP in Section 4.4.2.

In Section 3. 5 we showed that the state of the decision procedure must be given in terms of Q, the separable subspaces, to obtain an iterative procedure in terms of a one-dimensional expectation. It is easy to prove this same result for the unknown amplitude problem. Rather than repeat the proof for the unknown amplitude problem, we refer the reader to the similar proof for the unknown phase problem given in Section 3. 5. The derivation of the deferred-decision procedure for a SKEA with the state specified in terms of Q follows.

We begin the solution for a state with no available observations. The optimum average loss for a terminal decision is, as derived previously, the terminal loss function $T(L) = F_0(L)$.

Consider next the optimum procedure for one available observation. The state of the decision process includes n = 1, W/C, $\Delta_0$, L, g, and Q. Our problem is to decide whether to take an observation or to make an immediate terminal decision such that average loss for terminal decision is a minimum. If we take the observation, the average loss is equal to (1) the cost of the observation, and (2) the average of the optimum loss function of stage n - 1 weighted with respect to the probability density function of the observations. This average look-ahead loss is denoted $G_1(L)$ for stage n = 1.

$$G_1(L) = E\{F_0[L + \ell n\, \ell(z)]\} + C \tag{4.65}$$

The observation is taken only if $G_1(L)$ is less than average loss for an immediate decision (terminal), T(L). The intersection of $G_1(L)$ and T(L) define two decision points in L, namely, $\Gamma_1$ and $\Delta_1$. The set of decision points, $(\Gamma_i, \Delta_i)_{i=1}^{n}$, consitutes the optimum stopping rule. The minimum of $G_1(L)$ and T(L) is denoted by $F_1(L)$ and is the optimum average loss for one available observation. The signal amplitude distribution can be found from the initial state $\{L, g, \ell, g_t - g\}$ by using Eq. 4. 63 to determine the corresponding value of h. The quantity $g_t - g$ is another equivalent method of specifying the available number of observations. For $g_t - g$ equal to one, i. e., one available observation, no use is made of the updated amplitude distribution.

Suppose there are two available observations and the initial state is $L_2$, $g_2$, Q, $g_t - g_2 = 2$. If an observation is taken, the average loss, $G_2(L)$, is

$$G_2(L) = E\{F_1[L + \ell n\, \ell(z)]\} + C \tag{4.66}$$

The observation is taken only if $G_2(L)$ is less than $T(L)$. The intersection of $G_2(L)$ and $T(L)$ determine the decision points, $\Gamma_2$ and $\Delta_2$. The optimum average loss function is $F_2(L)$ and is given by

$$F_2(L) = \min[G_2(L),\ T(L)] \tag{4.67}$$

After one observation we have the possibility of one more observation. The log-odds ratio after one observation is $L_1 = L_2 + \ell n\, \ell(z)$. The value of Q, however, does not change since the decision processes can never change the initial Q state variate. The value of $g_2$ becomes $g_1$ as given by $g_1 = g_2 + 1$. From the above discussion, it is clear that the $g_t - g_2 = 2$ and "one-observation-taken" state is identical to the $g_t - g_1 = 1$ and "no-observation-taken" state. Therefore, the solution procedure iterates in the one-dimensional L space. The general set of iterative equations necessary to solve the deferred-decision problem is

$$L_{n-1} = L_n + \ell n\, \ell(z) \tag{4.68}$$

$$g_{n-1} = g_n + 1 \tag{4.69}$$

$$G_n(L) = E\{F_{n-1}[L + \ell n\, \ell(z)]\} + C \tag{4.70}$$

$$F_n(L) = \min[G_n(L),\ T(L)] \tag{4.71}$$

$$f_{n-1}(a) = \frac{\ell(z|a)}{\int \ell(z|a)\, f_n(a)\, da}\, f_n(a) \tag{4.72}$$

Consider the application of the general Eqs. 4.68 through 4.72 to the specific problem of a SKEA for a signal amplitude distribution corresponding to $\alpha = 0$. The initial state of the decision process for one available observation is $\{L_1,\ g_1,\ Q,\ g_t - g_1 = 1;\ W/C,\ \Delta_0,\ \alpha\}$. The amplitude distribution is specified completely by calculating h from the defining equation for Q, Eq. 4.63. Thus

$$h = \sqrt{g_1}\, \omega^{-1}[\sqrt{g_1} \exp (L - Q] \tag{4.73}$$

The log-odds ratio after one observation is $L_0$ and is given by

$$L_0 = L_1 + \ell n\, \ell(z)$$

$$= L_1 + \ell n \left[ \frac{\sqrt{g_1}}{\omega\left(\frac{h_1}{\sqrt{g_1}}\right)} \right] + \ell n \left[ \frac{\omega\left(\frac{h_1 + z}{\sqrt{g_1 + 1}}\right)}{\sqrt{g_1 + 1}} \right]$$

$$= Q + \ell n \left[ \frac{\omega\left(\frac{h_1 + z}{\sqrt{g_1 + 1}}\right)}{\sqrt{g_1 + 1}} \right] \tag{4.74}$$

Using Eq. 4.74, we can determine the average loss of taking one observation, i.e., $G_1(L)$, by

$$G_1(L) = \int_{-\infty}^{\infty} T\left\{ Q + \ell n \left[ \frac{\omega\left(\frac{h_1 + z}{\sqrt{g_1 + 1}}\right)}{\sqrt{g_1 + 1}} \right] \right\} f_1(z)\, dz + C \tag{4.75}$$

where $f_1(z)$ is the probability density function of the observations. This density function can be written as

$$f_1(z) = \frac{1}{1 + e^{-L}} \int_0^{\infty} f(z \,|\, SN, a)\, f(a)\, da + \frac{1}{1 + e^{L}} f(z \,|\, N) \tag{4.76}$$

From the assumptions of the problem, the distribution of z in noise alone is normal with zero mean and unity variance.[1] The distribution of z in signal-and-noise, given the amplitude, is normal with mean equal to the given amplitude and unity variance. Thus, Eq. 4.76 can be written as

$$f_1(z) = \frac{1}{1 + e^{-L}} \int_0^{\infty} \phi(z - a)\, \frac{\sqrt{g_1}}{\Phi\left(\frac{h_1}{\sqrt{g_1}}\right)}\, \phi\left(a\sqrt{g_1} - \frac{h_1}{g_1}\right) da + \frac{1}{1 + e^{L}}\, \phi(z)$$

[1] The variance of the noise may be any number without complicating the problem. We have normalized the variance for convenience.

$$= \frac{1}{1+e^{-L}} \left[ \sqrt{\frac{g_1}{g_1+1}} \; \frac{\phi(z)}{\omega\left(\frac{h_1}{\sqrt{g_1}}\right)} \; \omega\left(\frac{h_1+z}{\sqrt{g_1+1}}\right) \right] + \frac{\phi(z)}{1+e^{L}} \tag{4.77}$$

The average look-ahead loss is obtained by evaluating Eq. 4.75 where $f_1(z)$ is given by Eq. 4.77. The optimum average loss is then obtained by choosing the minimum of $G_1(L)$ and $T(L)$, i.e., $F_1(L) = \min[G_1(L), T(L)]$.

Generalizing the results of stage $g_t - g_1 = 1$, we have for k available observations and a state $\{Q, L_k, g_k = g_{k-1}-1, g_t - g_k = k\}$, the following iterative equations.

$$L_{k-1} = Q + \ell n \left[ \frac{\omega\left(\frac{h_k+z}{\sqrt{g_k+1}}\right)}{\sqrt{g_k+1}} \right] \tag{4.78}$$

$$G_k(L) = \int_{-\infty}^{\infty} F_{k-1}\left\{ Q + \ell n \left[ \frac{\omega\left(\frac{h_k+z}{\sqrt{g_k+1}}\right)}{\sqrt{g_k+1}} \right] \right\} f_k(z)\, dz + C \tag{4.79}$$

$$f_k(z) = \frac{1}{1+e^{-L}} \left[ \sqrt{\frac{g_k}{g_k+1}} \; \frac{\phi(z)}{\omega\left(\frac{h_k}{\sqrt{g_k}}\right)} \; \omega\left(\frac{h_k+z}{\sqrt{g_k+1}}\right) \right] + \frac{\phi(z)}{1+e^{L}} \tag{4.80}$$

$$F_k(L) = \min[G_k(L), T(L)] \tag{4.81}$$

The intersection of $G_k(L)$ and $T(L)$ defines the decision boundary points which are used in determining the correct response of "A", "B", or "continue." The set of decision boundary points for k = 1, 2, ..., n constitutes the optimum stopping rule. Note that the decision boundary is intimately involved with the initial state. In simple deferred-decision theory the stopping rule is a function of the available observation time and the costs involved in the detection problem. However, in composite deferred-decision theory, the stopping rule is not only a function of the available observation time and the costs, but it is

also a function of the initial probability distribution (as specified as Q, g, and L) and the observation time already used. These latter two dependencies are a direct result of the learning or adaptation that occurs in the observation process.

The reader may question where use is made of the updated amplitude distributions. Overt use of these distributions is not necessary because of the closure property, and thus it may appear, at first glance, that there is no updating process. Equations 4.78 and 4.80, however, are a direct result of updating the signal amplitude distributions.

The numerical results presented in Section 4.5.3 were obtained by a computing algorithm programmed for an IBM 7090. The program, given in Appendix H, computes the average risk functions and decision points in h space rather than L space because computation in h space requires a direct interpolation of the $\omega(\ )$ function rather than an inverse interpolation. The corresponding L values are found in the program by use of Eq. 4.63.

For a detailed derivation of the deferred-decision procedure for the $\alpha = 1$ and the Gaussian signal amplitude distribution, the reader is referred to Appendix I.

4.5.2 Evaluation of the Deferred-Decision Procedure for a Signal Known Except for Amplitude. The evaluation of a deferred-decision procedure is given by the average loss for a terminal decision as a function of L. It is helpful for an intuitive understanding of the evaluation to decompose the average loss into the average time spent in observing and the average error performance obtained in making terminal decisions. The latter quantity is displayed by means of ROC curves and the former is given by the average number of observations (ANO). The complete evaluation requires a separate ROC and ANO curve for each stage of the observation procedure. The ROC curves are calculated as a function of L and presented as an implicit function of L. The ANO curves are presented as an explicit function of L.

We can solve for the ROC and ANO curves in a manner similar to that used for obtaining the optimum average loss functions. The solution is again an iterative type of solution solved, first, for one deferral, then two deferrals, etc. The derivations of the ROC and ANO curves follow.

The specific functions of L which describe the ROC at stage n of the observation procedure are denoted by $y_n(L)$ and $x_n(L)$. The function $y_n(L)$ is the probability of responding "A" given that SN is the cause of the receiver input and $L_n = L$. Similarly, $x_n(L)$

is the probability of responding "A" given that N is the cause of the receiver input and $L_n = L$.

The basic iterative equations for the ROC curves, with n possible deferrals, are

$$\begin{aligned} y_n(L) &= 0 && , \quad L \leq \Gamma_n \\ &= \int y_{n-1}[L + \ell n\, \ell(z)]\, f_n(z|SN)\, dz && , \quad \Gamma_n < L < \Delta_n \\ &= 1 && , \quad L \geq \Delta_n \end{aligned} \tag{4.82}$$

$$\begin{aligned} x_n(L) &= 0 && , \quad L \leq \Gamma_n \\ &= \int x_{n-1}[L + \ell n\, \ell(z)]\, f_n(z|N)\, dz && , \quad \Gamma_n < L < \Delta_n \\ &= 1 && , \quad L \geq \Delta_n \end{aligned} \tag{4.83}$$

The quantities $(\Gamma_k, \Delta_k)_{k=1}^{n}$ are the optimum decision boundaries obtained from the optimum average loss functions. The probability density functions $f_n(z|N)$ and $f_n(z|SN)$ are defined as in Section 4.5.1. These equations are repeated below.

$$f_n(z|N) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \tag{4.84}$$

$$f_n(z|SN) = \sqrt{\frac{g_n}{g_n + 1}} \cdot \frac{\phi(z)}{\omega\left(\frac{h_n}{\sqrt{g_n}}\right)} \cdot \omega\left(\frac{h_n + z}{\sqrt{g_n + 1}}\right) \tag{4.85}$$

The specific quantities used to define the average time spent in observing, i.e., the ANO curves, are denoted by $u_n(L)$ and $v_n(L)$. The function $u_n(L)$ is the average number of observations under the conditions: SN is the cause of the receiver input, there are n deferrals, and $L_n = L$. The function $v_n(L)$ is the average number of observations under the conditions: N is the cause of the input, there are n deferrals, and $L_n = L$.

The basic iterative equations for the ANO functions, with n deferrals, are

$$\begin{aligned} u_n(L) &= 0 \\ &= 1 + \int u_{n-1}[L + \ell n\, \ell(z)]\, f_n(z|SN)\, dz \end{aligned} \tag{4.86}$$

and

$$v_n(L) = 0$$

$$= 1 + \int v_{n-1}[L + \ell n\, \ell(z)]\, f_n(z|N)\, dz \qquad (4.87)$$

where $f_n(z|N)$ and $f_n(z|SN)$ are defined by Eqs. 4.84 and 4.85, respectively.

Equations 4.82, 4.83, 4.86 and 4.87 can be used to find the ROC and ANO curves at any stage of observation if the ROC and ANO curves for the n = 1 stage are known. For n = 1, the ROC curves are identical to the ROC curves derived in Section 4.3.3. For n = 1, the ANO functions are equal to zero for $L \leq \Gamma_1$ and $L \geq \Delta_1$ and are equal to one for L in the continue region. Equations 4.82, 4.83, 4.86 and 4.87 were solved iteratively in the same program that determined the optimum average loss curves given in Appendix H. The average loss curves were found first to determine the decision points. These were then used in the determination of the ROC and ANO curves. The computational program is explained further in Appendix H.

The derivations of the optimum average loss functions and the ROC and ANO curves for the $\alpha = 1$ and Gaussian signal amplitude distributions are similar to the derivations presented for the $\alpha = 0$ distribution. The derivations for these amplitude distributions are given in Appendix I, since they do not serve to increase the reader's understanding of the basic detection problem.

## 4.6 Numerical Results of the Deferred-Decision Procedure for a Signal Known Except for Amplitude

The numerical results of deferred-decision theory for a signal known except for amplitude are divided into two parts. The first section is concerned with the optimum decision boundaries and the related performance of the optimum receiver. The second section is concerned with comparisons of the optimum deferred-decision receiver and other suboptimum receivers.

### 4.6.1 Decision Boundaries of the Optimum Detection Receiver for a Signal Known Except for Amplitude.

The optimum decision boundaries, for a composite signal hypothesis, are a function of the initial state of the detection problem which, for a SKEA, is

given by $\{Q, L, g, g_t - g\}$. The last statement implies that the decision boundaries are a function of the a priori signal amplitude distribution (because of the dependence of the state on Q, L, and g). Thus, for each a priori signal amplitude distribution, there is a set of decision boundaries. Therefore, to completely describe the optimum stopping rule for a given W/C, g, and $g_t - g$, we must solve the continum of all Q surfaces in the range of interest. In the specific case of a distributed amplitude problem, the decision boundaries can be represented as a three-dimensional solid with axes L, g, and h. Different values of Q cut the solid into two dimentional surfaces. The extent of the surface, in terms of L, g, and h, constitutes the optimum stopping rule for a given Q surface. The value of W/C determines how large the three-dimensional solid is in the L and h directions.

Using what we know of the optimum stopping rule, the optimum detection receiver for a composite signal hypothesis takes the form shown in Fig. 4.38. In previous optimum receiver designs, as shown in Figs. 1.1 through 1.3, the optimum receiver consists of a likelihood ratio processor and a stopping rule. In the case of a composite signal hypothesis, a further complication of the receiver is necessary. The input, z, is used in determining not only the likelihood ratio, but also in determining the a posteriori amplitude distribution. The latter process is necessary for calculating the average likelihood ratio and for determining the decision boundaries.

The modification of the signal parameter distribution with each succeeding observation is achieved in accordance with Bayes' Rule. The continual updating process is what many authors term learning or adaptation. We see that this process is a natural consequence of the design of the optimum deferred-decision procedure rather than a preconceived concept. In other words, the fact that the learning or adaptation process is necessary is a natural outgrowth of our optimum receiver and not a planned result that the optimum receiver must be adaptive.

The updating process of the signal amplitude distribution also furnishes the observer with a classification output. The classification output is another natural consequence of the solution procedure and not an overt design objective. In the remainder of this section, we are primarily interested in the optimum stopping rule.

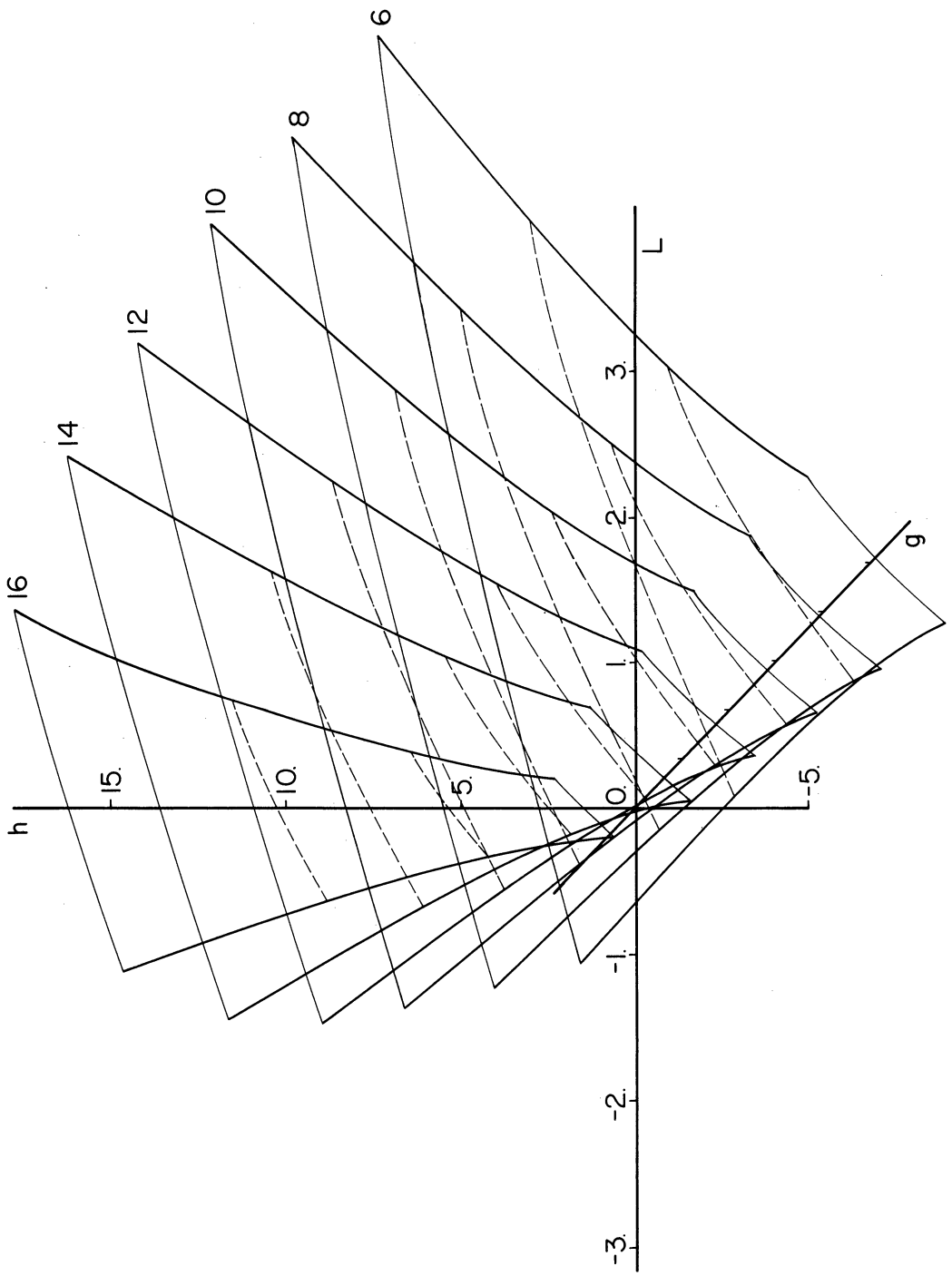


Fig. 4.38. The optimum deferred-decision detector for a composite-signal-hypothesis.

The optimum stopping rule for a SKEA can be represented as a three-dimensional solid in L x h x g space, as we have previously mentioned. Because we describe the state of the problem in terms of $\{Q, L, g, g_t - g\}$, the value of h (and hence, the amplitude distribution) must be found from the defining equation for Q. The more natural method of describing the state is in terms of the amplitude distribution and the log-odds ratio, i. e., in terms of $\{L, g, h, g_t - g\}$. If one wishes to specify the state by $\{L, g, h, g_t - g\}$, then it is necessary to solve the problem in terms of $\{Q, L, g, g_t - g\}$ and use the correspondence between Q and h as defined by Eq. 4. 63.

A degree of understanding Q surfaces can be obtained by plotting the surface Q-L as a function of g and h. From the defining equation for Q, we have

$$Q\text{-}L = \ell n\left(\frac{\sqrt{g}}{\omega(h/\sqrt{g})}\right) \tag{4.88}$$

Plotting Q-L versus h and g, we obtain the graph shown in Fig. 4. 39. Suppose, for purposes of discussion, the log-odds ratio is zero. We see, from Fig. 4. 39, that as Q becomes more negative, the Q surface approaches a plane, and thus becomes independent of the parameter h. This collaborates an earlier discussion given in Section 4. 5. 1 showing that the signal known exactly surface is Q equal to minus infinity. All Q surfaces for the $\alpha = 0$ amplitude distribution are a translate of the surface shown in Fig. 4. 39 along the Q-L axis.

The optimum stopping rule for a specific case of the $\alpha = 0$ amplitude distribution is shown in Fig. 4. 40. The parameters of the problem are W/C = 100, $\Delta_0$ and $g_t = 21$. The axes are L, h, and g, and physically represent the observer's opinion of the cause of the input, the value of the integrated input, and the time observed, respectively. The stopping rule shown in Fig. 4. 40 was obtained by means of an IBM 7090 programmed according to the program given in Appendix H. The Q surfaces shown correspond to Q = +1, 0, -2, and -6.

The Q surface corresponding to a value of -6 is the upper surface and exhibits the least amount of curvature. As the value of Q becomes more positive, the signal amplitude distribution becomes more diffuse, and the L range for which it is profitable to observe decreases. The spacing of the approximately vertical sides of the stopping rule is dependent on the W/C ratio. As the W/C increases, the extent of the stopping rule increases in the L and h directions.

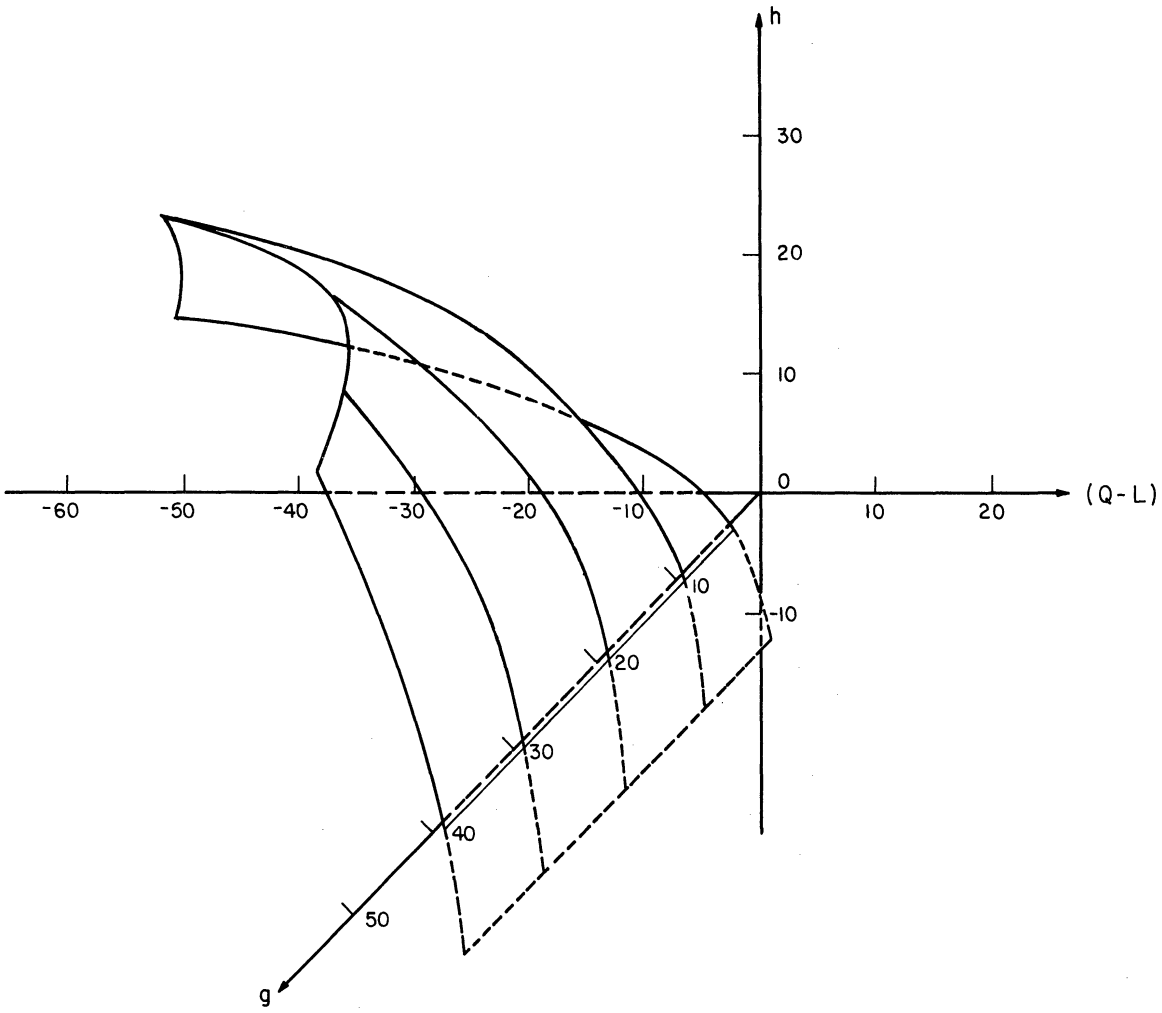


Fig. 4.39. A graph of the Q-L surface as a function of h and g.

From Fig. 4.40, we can obtain the relationship between the observer's log-odds ratio and the amplitude distribution on a Q surface as a function of observed time, g. Referring to Fig. 4.40, we see that Q surface tips downward in the negative L direction. This means that if the observer thinks the cause of the input is noise alone, then he also thinks that the signal amplitude distribution is near zero, and visa versa.

The concept of a Q surface is useful not only in computations, but also in understanding the operation of the detection receiver. In operation the receiver always operates in such a manner that it's output remains on the initial Q surface. Thus, by examining the

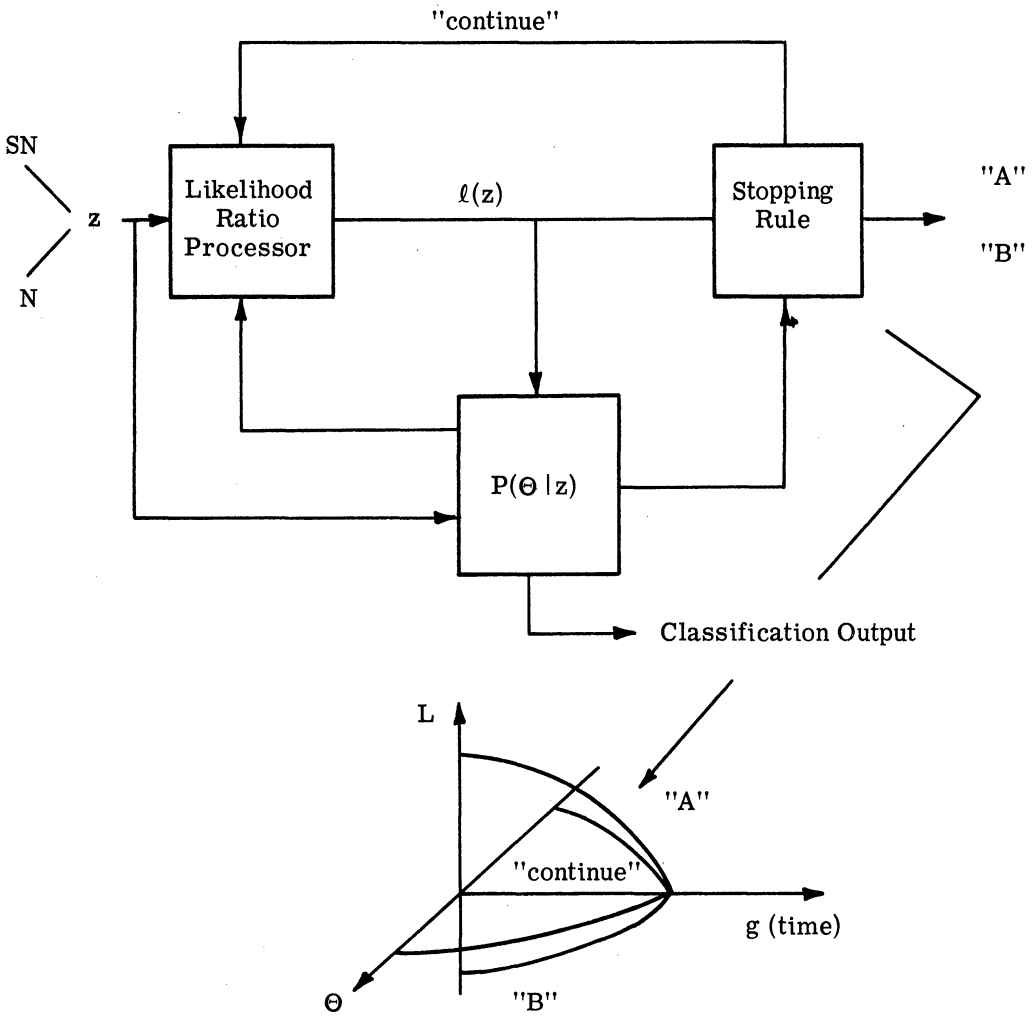


Fig. 4.40. The optimum deferred-decision boundaries for a SKEA with parameters W/C = 100, $g_t$ = 21, and $\alpha$ = 0.

operation of the receiver on a set of representative Q surfaces, we obtain a reasonably accurate picture of the receiver's functions.

Notice that as the time spent in observing increases (g increasing), the Q surface tips upward in the positive h direction. Physically, h is the integrated receiver input. Thus, we expect, when signal-and-noise is the cause of the receiver input, the value of the integrated input to increase. For noise alone the integrator output will wander about some value near the initial condition of the integrator. Therefore, we would expect, from the physical operation of the receiver, the decision boundaries in terms of h to increase as g increases.

The complete optimum stopping rule for a SKEA consists of a specification of the three-dimensional stopping rule similar to that depicted in Fig. 4.38. However, the complete stopping rule involves a great deal of computation and is valid for only a given W/C and $g_t$. In the remainder of this section, we discuss various aspects of the optimum stopping rule directed toward approximate methods for finding the decision boundaries as various parameters of the detection problem are changed.

Suppose, for example, we fix the W/C ratio and decrease the value of $g_t$. Can any conclusions be drawn concerning the resultant optimum stopping rule? We can, for example, show that the stopping rule for the smaller $g_t$ is contained in the stopping rule with the large $g_t$. The proof follows below.

If $F_k(L, g|g_t, Q, W/C)$ is the optimum average loss function for a fixed set $g_t$, Q, and W/C with k possible deferrals, and $G_k(L, g)$ is the average look-ahead loss function with k possible deferrals, then $G_k(L, g) \leq G_{k-1}(L, g)$. It is clear that, given the same state in a decision process, the possibility of an additional observation results in an average loss less than or equal to the average loss for an immediate terminal decision. In symbols, this can be expressed as

$$F_1(L, g|g_t = g^*, Q, W/C) \leq F_0(L, g|g_t = g^* - 1, Q, W/C) = T(L). \tag{4.89}$$

Now

$$G_2(L, g) = E\{F_1(L, g|g^*)\} + C \tag{4.90}$$

and

$$G_1(L, g) = E\{F_0(L, g|g^* - 1)\} + C \tag{4.91}$$

implying that $G_2(L, g) \leq G_1(L, g)$. By induction, we can establish that $G_k(L, g) \leq G_{k-1}(L, g)$. From the last statement it follows that $G_k(L, g)$ intersects T(L) at a larger L value than $G_{k-1}(L, g)$ implying that the decision boundary for a given $g_t^*$ contains the decision boundary for all $g_t < g_t^*$.

A numerical verification of the preceding discussion is depicted in Fig. 4.41. The optimum decision boundaries are shown for the same Q and W/C and two values of $g_t$. The boundaries for $g_t = 10$ are everywhere inside the boundaries for $g_t = 21$. Although we

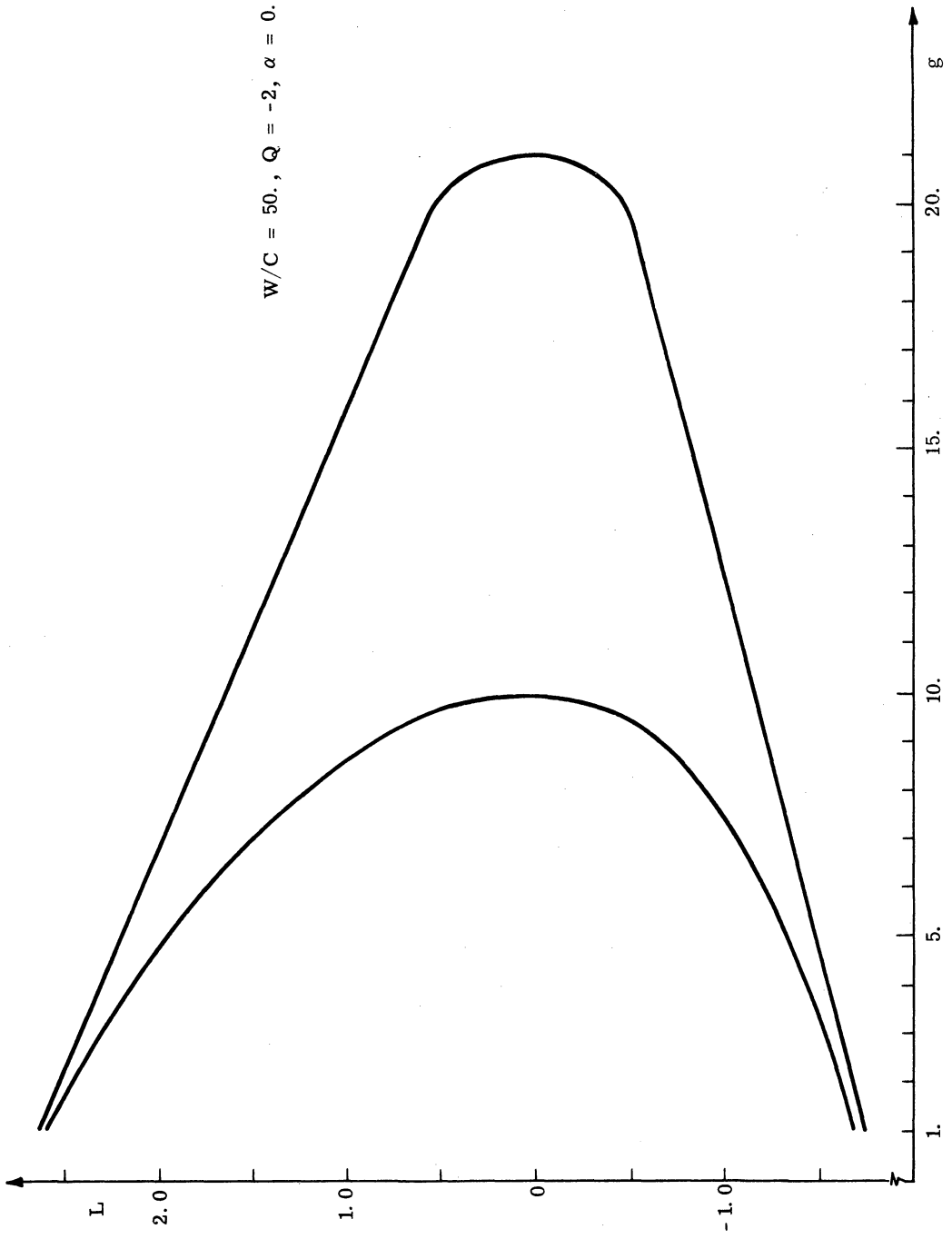


Fig. 4.41. The optimum deferred-decision boundaries for a SKEA for $\alpha$ = a, Q = -2, W/C = 50, and $g_t$ = 10 and 21.

have shown the boundaries projected only onto the L x g plane, the same result holds true for the boundaries projected onto the h x g plane.

The three-dimensional graph of the stopping rule presented in Fig. 4. 38 does not give as accurate a description of the stopping rule as does the projection of the decision boundary, for a given Q surface, onto the L x g or h x g plane. The projections of the optimum stopping rule are shown in Figs. 4. 42 through 4. 44 for a detection problem with parameters W/C = 50 and 100 and $g_t$ = 21.

From Figs. 4. 42 and 4. 43, we see that the general character of the projection of the optimum stopping rule onto the L x g plane is a convergence of the boundaries as the available observation time decreases coupled with a definite expansion of the decision boundaries at the time when observations begin, i. e , g = 1, 2, . . . , . The latter phenomenon is due to the desire to make good decisions on large amplitude signals which, if they are present, cause early terminations.

Figure 4. 44 depicts the projection onto the h x g plane of the same optimum stopping rule of Figs. 4. 42 and 4. 43. Again, we note the expansion of the decision boundaries for small g values and the convergence of the decision boundaries for small available observation times, i. e., $g_t$ - g small. Notice also that as the value of Q decreases, the slope of the h versus g decision boundaries increases. This implies that the integrator output in SN increases faster per unit time as the value of Q decreases. This again verifies our conclusion that signal amplitude becomes better known as the value of Q decreases.

The evaluation of the optimum receiver for a simple deferred-decision problem (Ref. 6) pointed out that the performance of the optimum receiver could be approached very closely by a suboptimum receiver whose decision boundaries are near the optimum decision boundaries. In simple deferred-decision theory, the decision boundaries for large available observation time, i. e., the asymptotic boundaries are convenient to use in operating the receiver, because the boundaries are independent of time. The asymptotic boundaries also provide the connecting link between simple deferred-decision theory and Walds' theory of of sequential analysis. Thus, the asymptotic solution of simple deferred-decision theory is an important aspect of the over-all solution. From the above discussion it would appear that the asymptotic solution to composite deferred-decision theory might also be quite useful.

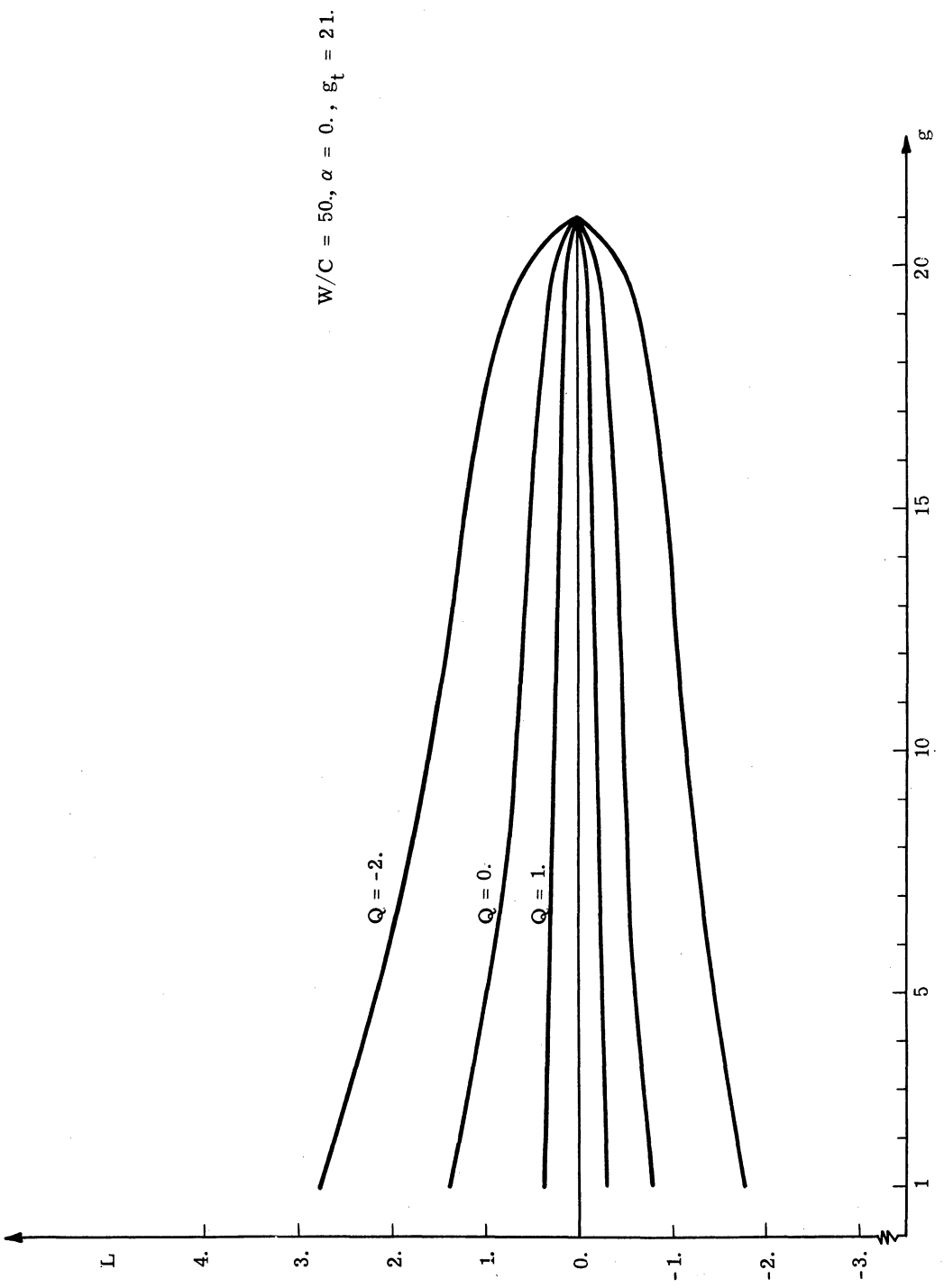


Fig. 4.42. The optimum deferred-decision boundaries for a SKEA projected on the L x g plane for W/C = 50, $g_t$ = 21, and $\alpha$ = 0.

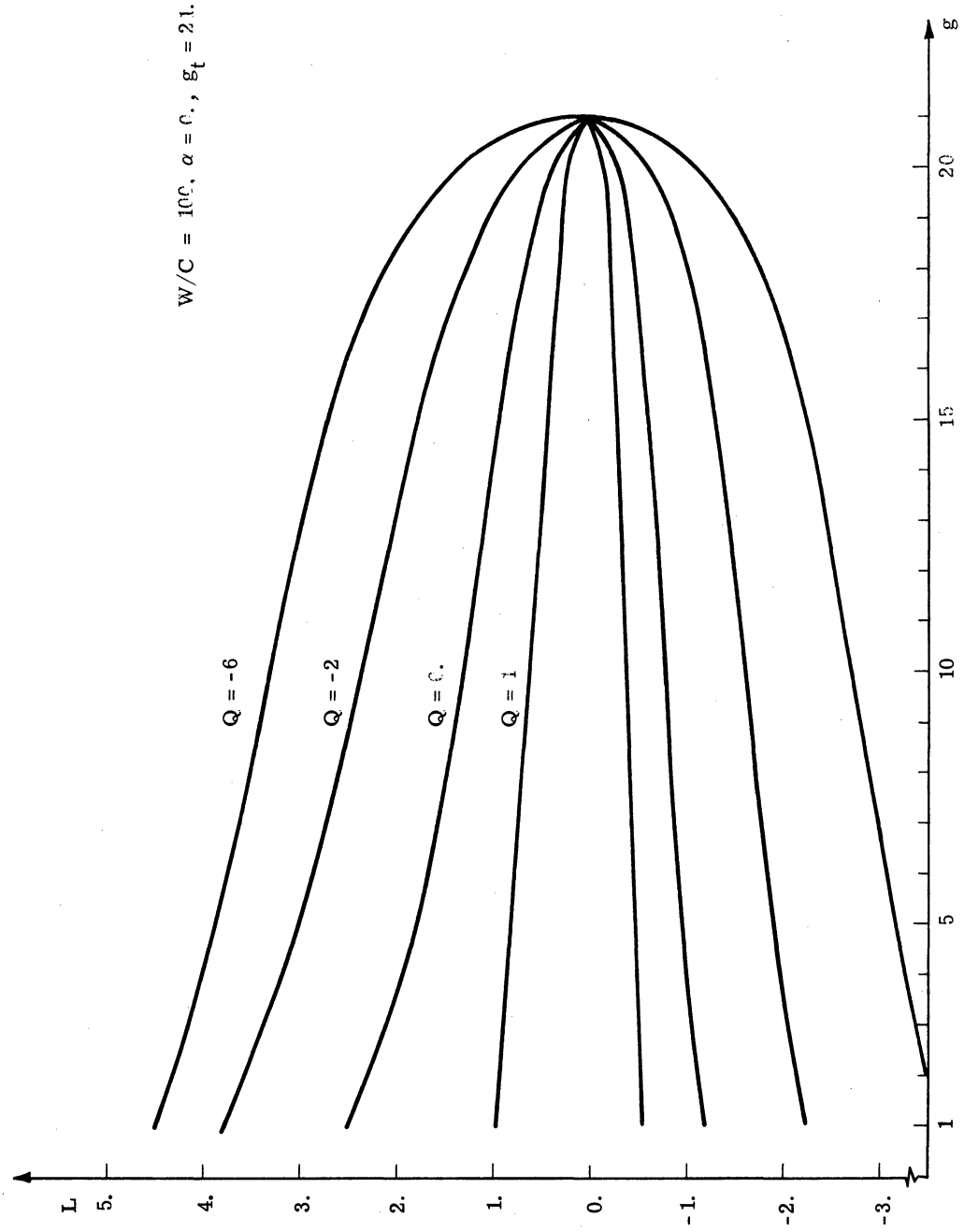


Fig. 4.43. The optimum deferred-decision boundaries for a SKEA projected on the L x g plane for W/C = 100, $g_t$ = 21, and $\alpha$ = 0.

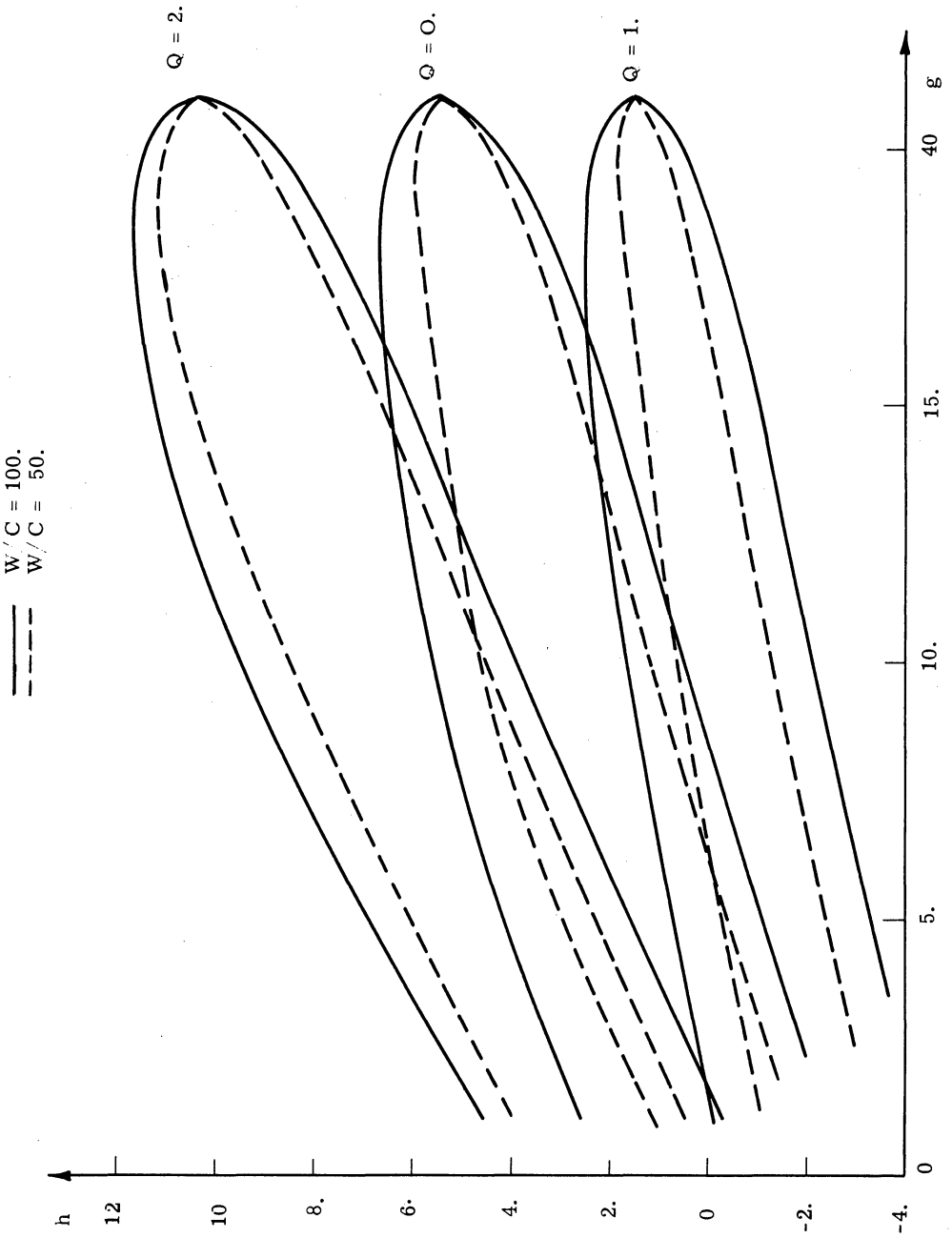


Fig. 4.44. The optimum deferred-decision boundaries for a SKEA projected on the h x g plane for W/C = 50 and 100, $g_t = 21$, and $\alpha = 0$.

Unfortunately, there is no asymptotic theory for every composite deferred-decision problem, and the present SKEA problem is an example.

This, perhaps, startling statement is quite evident when we examine the state of the detection problem as observations are taken. The state is given by $\{Q, L, g, g_t - g\}$ at any stage of the observation process. The absence of an asymptotic solution for the SKEA problem is directly attributed to the dependence of the state on $g_t - g$ and g simultaneously. In the simple deferred-decision problem the state depends on the available number of observation, $g_t - g$, but not on observations already taken, g. It is because of the dependence of the state on the available number of observations and the observations already taken that the SKEA problem possesses no asymptotic solution. This dependence can be attributed to the learning or adaptation process. The possibility of learning from observation to observation implies that we can never reach a situation in which the decision boundaries are independent of observations already taken and available observations left.

In simple deferred-decision theory, we can solve a nested set of problems because the solution for n available deferrals contains the solution for k available deferrals, $k \leq n$. Thus, we may extend any simple deferred-decision solution to $n + 1, n + 2, \ldots,$ available observations. This type of extension is not possible for the SKEA problem, because the state of the problem depends not only on the available observation time, $g_t - g$, but also on the time spent in observing, g.

The relation of the projection of the optimum decision boundary onto the L x g plane to the W/C ratio was examined over a range of W/C ratios as shown in Figs. 4.45 and 4.46. The computations pictured in Figs. 4.45 and 4.46 indicate that (1) the optimum decision boundary for fixed g as a function of $\ell n(W/C)$ is approximately linear over a wide range of W/C, and (2) as Q decreases and g increases, the approximation becomes better. Figure 4.45 is for the $\alpha = 0$ distribution, and Fig. 4.46 is for the Gaussian amplitude distribution.

Additional insight into the general character of the optimum decision boundary can be obtained by studying the average trajectory of a known amplitude signal as a function of L, g, and h. Since we use L as an explicit variable, it is convenient to examine the average trajectories in the L x g or (L-Q) x g plane.

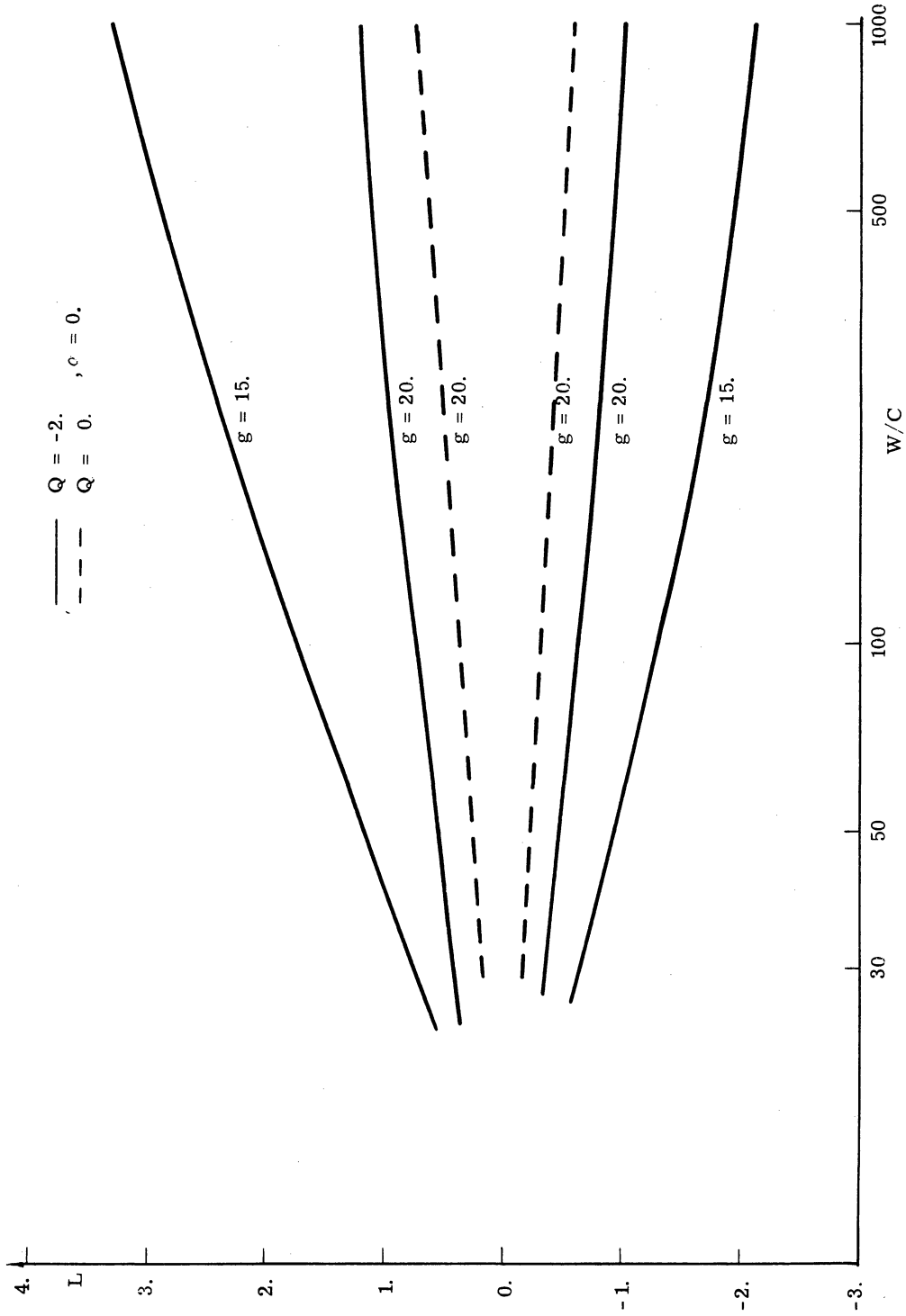


Fig. 4.45. The optimum deferred-decision boundaries for fixed g as a function of W/C for Q = 0 and -2, $g_t$ = 21, and $\alpha$ = 0.

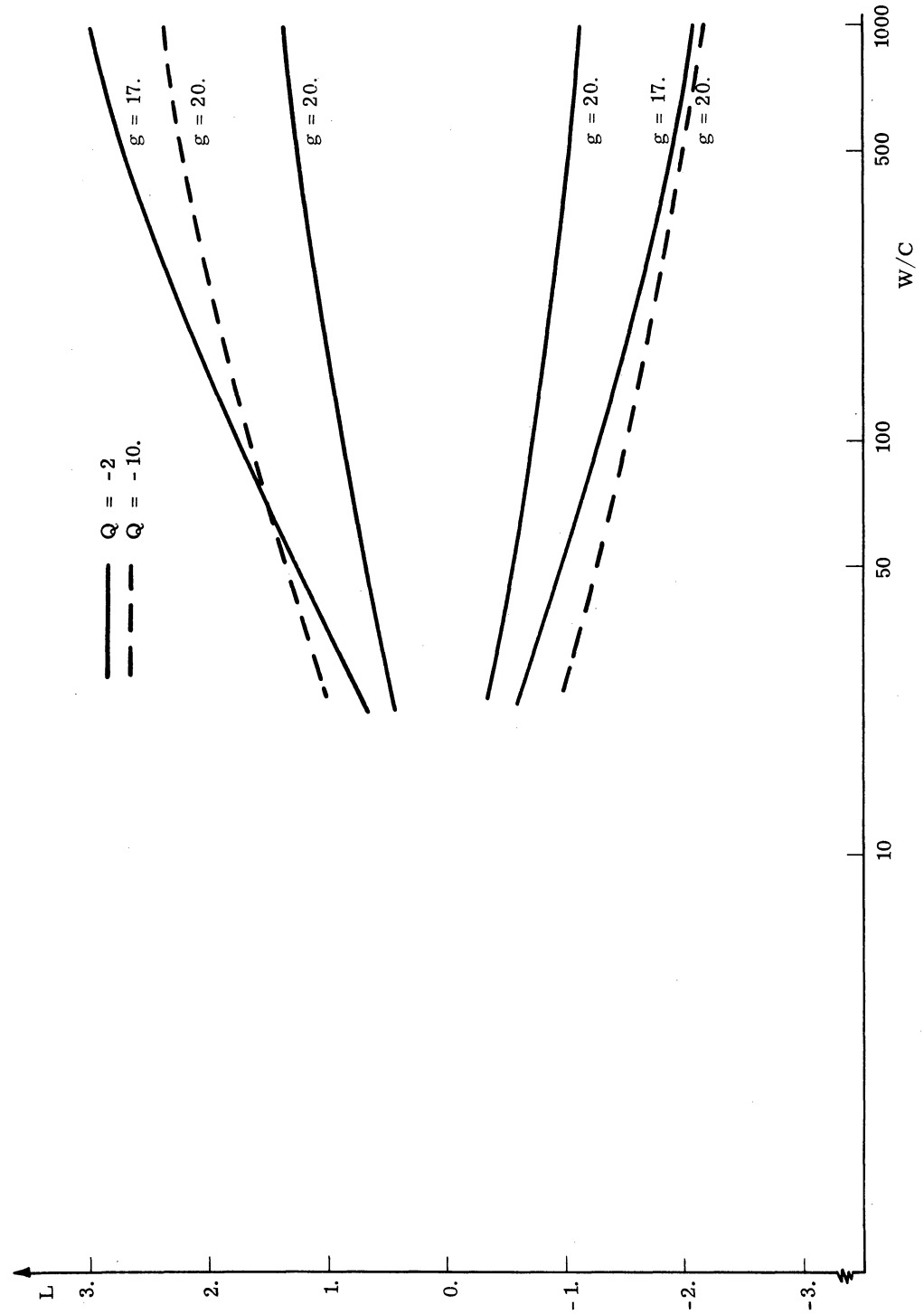


Fig. 4.46. The optimum deferred-decision boundaries for fixed g as a function of W/C for Q = -2 and -10, $g_t$ = 21, and a Gaussian amplitude distribution.

A knowledge of the average trajectories for various size signal amplitudes can be used in an approximate placement of the lower decision boundary. For example, the average trajectories of a small amplitude signal and noise alone have the form shown in Fig. 4.47. The lower decision boundary should discriminate against the noise alone trajectory, but still permit small amplitude signals to terminate in an "A" response. Thus, the lower decision boundary for moderate g values would have a value approximately equal to the dotted line shown in Fig. 4.47.

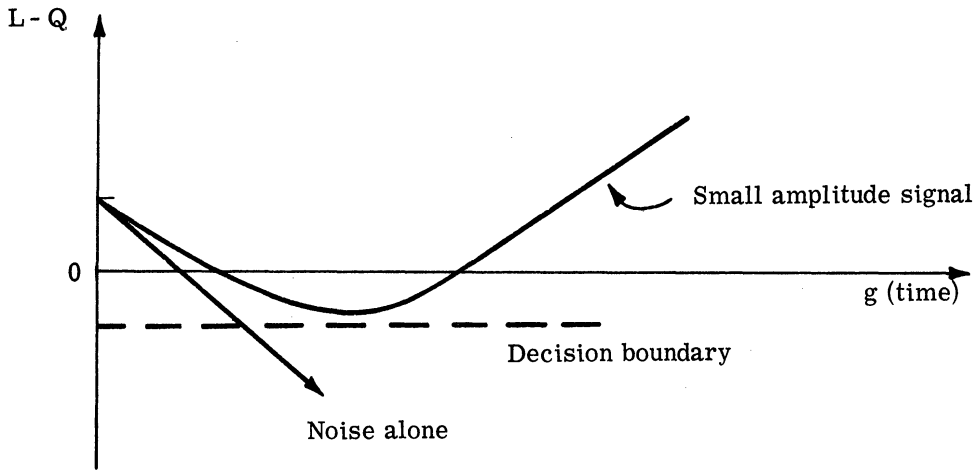


Fig. 4.47. The average trajectories of noise alone and small amplitude signals in the (L - Q) x g plane.

A numerical example of several amplitude trajectories are shown in Fig. 4.48 for the amplitude density function given by

$$f(a) = \sqrt{\frac{2}{\pi}} \exp\left(\frac{-a^2}{2}\sqrt{\frac{\pi}{2}}\right), \quad a \geq 0$$

$$= 0 \quad , \quad \text{otherwise} \qquad (4.92)$$

The above density function represents the "maximum diffuseness" or uncertainty one can obtain with an $\alpha = 0$ distribution.

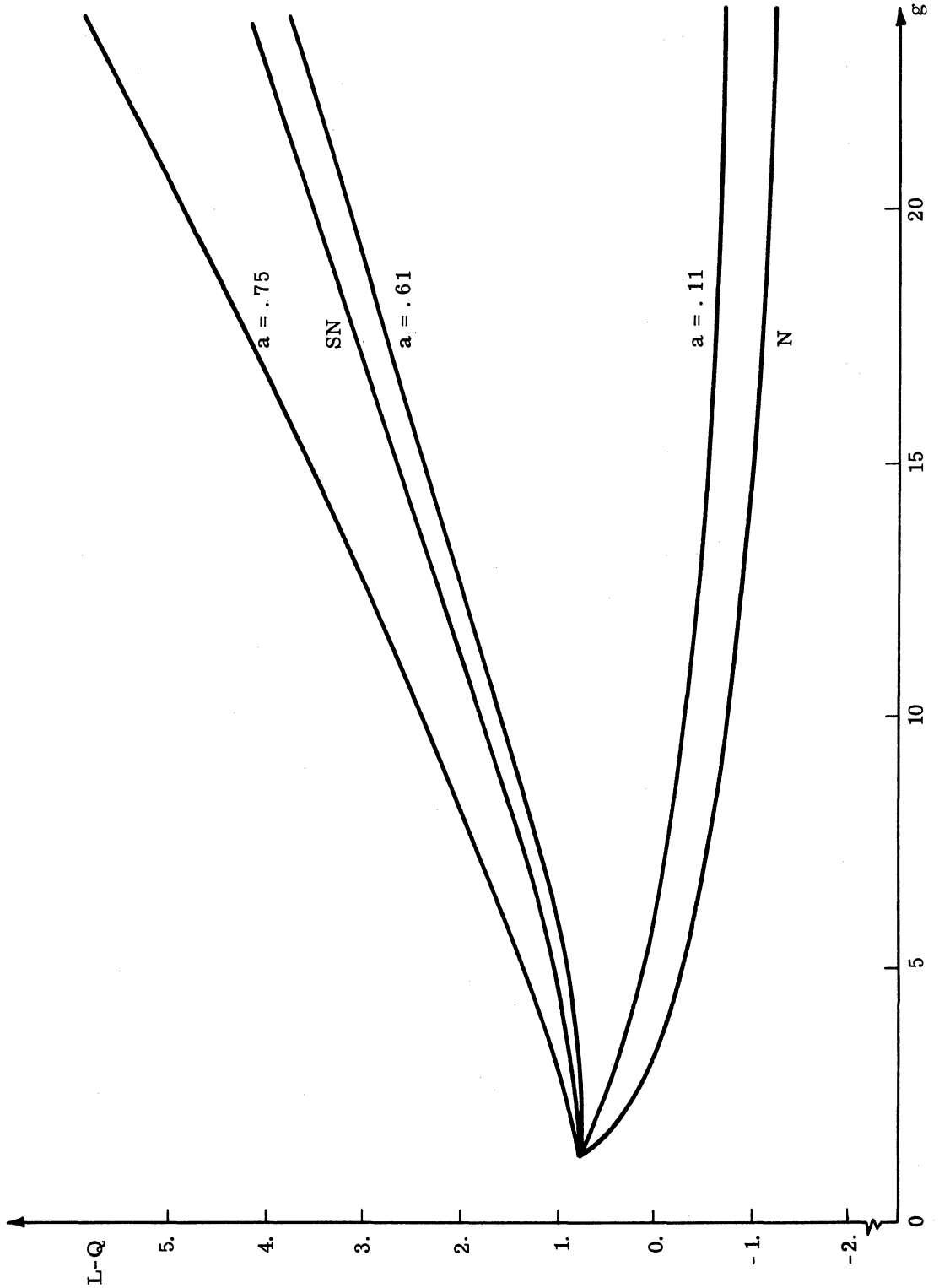


Fig. 4. 48. The average trajectories of four signals in the (L - Q) x g plane for a specific truncated amplitude distribution.

Plotted in Fig. 4.48 are trajectories in the (L-Q) x g plane for signal amplitudes representing 0% probability (noise alone), 10% probability (a = 0.11), 50% probability (a = 0.61), and 60% probability (a = 0.75). The 10% probability amplitude means 10% of the time the signal amplitude is less than or equal to a = 0.11. Also shown in Fig. 4.48 is the trajectory of the average motion in signal-and-noise. From Fig. 4.48, we conclude that (1) the noise alone trajectory is quite slow downward as compared to the signal-and-noise trajectory upward, and (2) the spread in amplitude trajectories is fairly broad. The latter conclusion seems to indicate that amplitude information is important. That is, it does not appear that we can use an estimated signal amplitude to predict average performance.

Signals with amplitudes less than or equal to 0.11 take a very long time to integrate up to values such that they would result in an "A" response. For example, at g = 400 the value of the a = 0.11 trajectory is only 0.46. Thus, to find small amplitude signals the receiver has to operate for a long period of time.

This completes our discussion of the optimum stopping rule for the SKEA problem. The next section presents evaluations of optimum receivers in terms of ROC and ANO curves.

4.6.2 Performance of the Optimum Deferred-Decision Receiver. The performance of the optimum deferred-decision receiver is given in terms of the average loss functions. As before, we use the ROC and ANO functions for an equivalent evaluation. The mathematical derivation of the above two functions is given in Section 4.5.2. The necessary computations are performed by an IBM 7090 programmed according to the algorithm given in Appendix H. In this section only the numerical results of the $\alpha = 0$ signal amplitude distribution are discussed in detail. The evaluations for the $\alpha = 1$ and Gaussian amplitude distributions result in similar conclusions.

Typical evaluation functions for the optimum deferred-decision detector are shown in Figs. 4.49 through 4.52. Figures 4.49 and 4.50 are the ANO and ROC curves, respectively, corresponding to an observation state of $\{Q = -2, g_t = 21, W/C = 100, g = 20, 10, \ldots,\}$. Figures 4.51 and 4.52 are the ANO and ROC curves, respectively, for the same observation state as above expect Q = 0 instead of Q = -2. From previous discussions, Q = 0 represents an observation state of greater signal uncertainty than Q = -2.

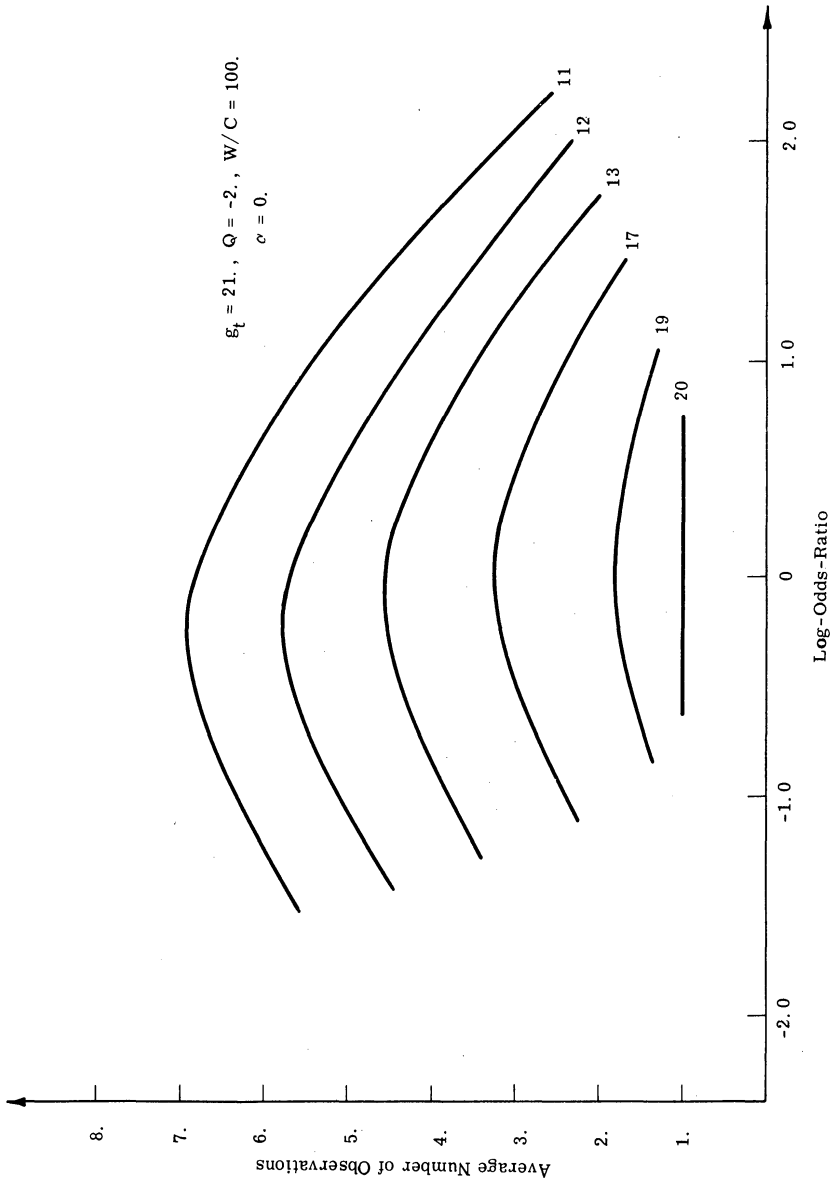


Fig. 4.49. The average number of observations for a deferred-decision receiver as a function of the log-odds ratio with parameters Q = -2, $g_t$ = 21, W/C = 100, and $\alpha$ = 0.

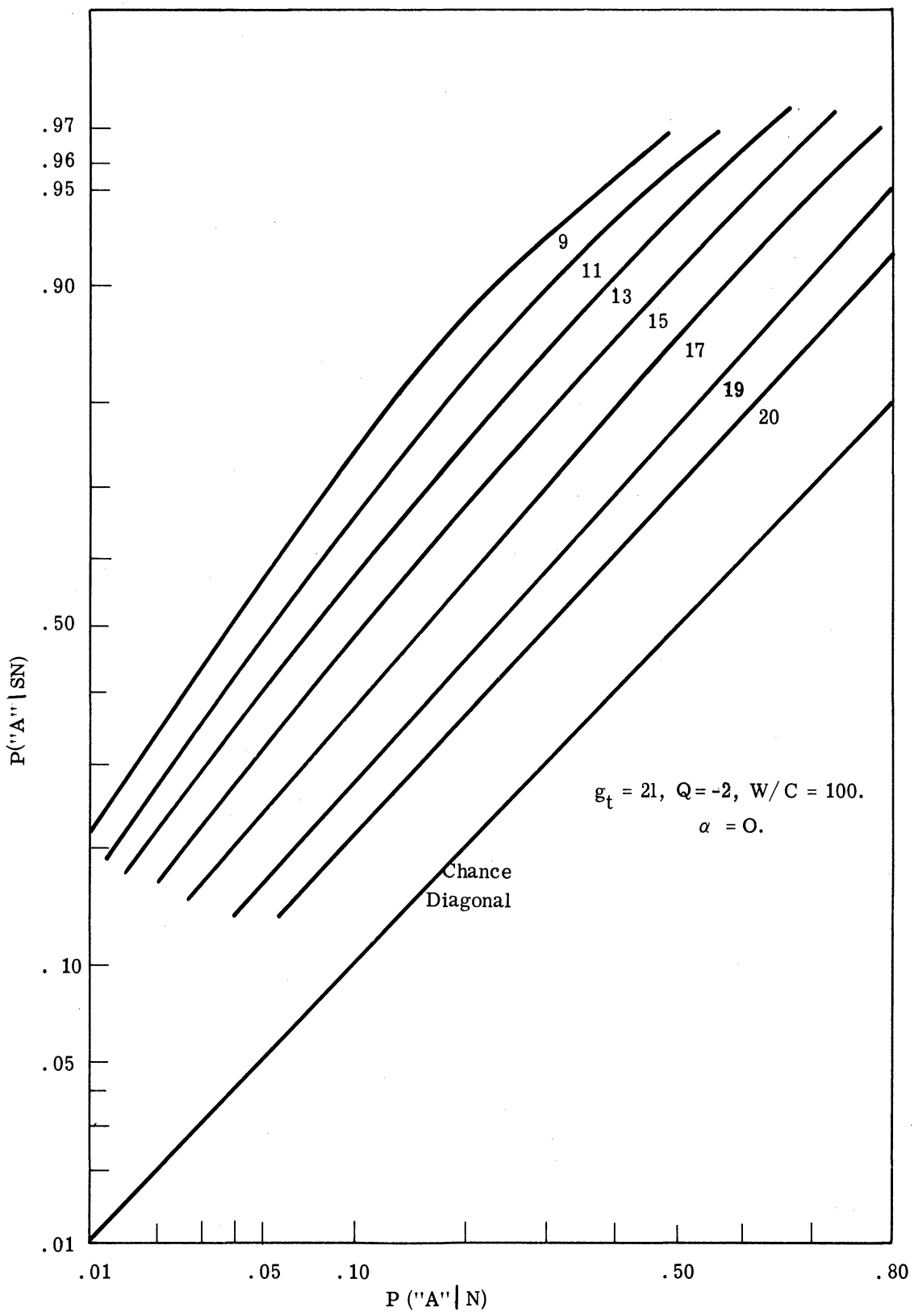


Fig. 4.50. The ROC curves of a deferred-decision receiver with parameters Q = -2, $g_t$ = 21, W/C = 100, and $\alpha$ = 0.

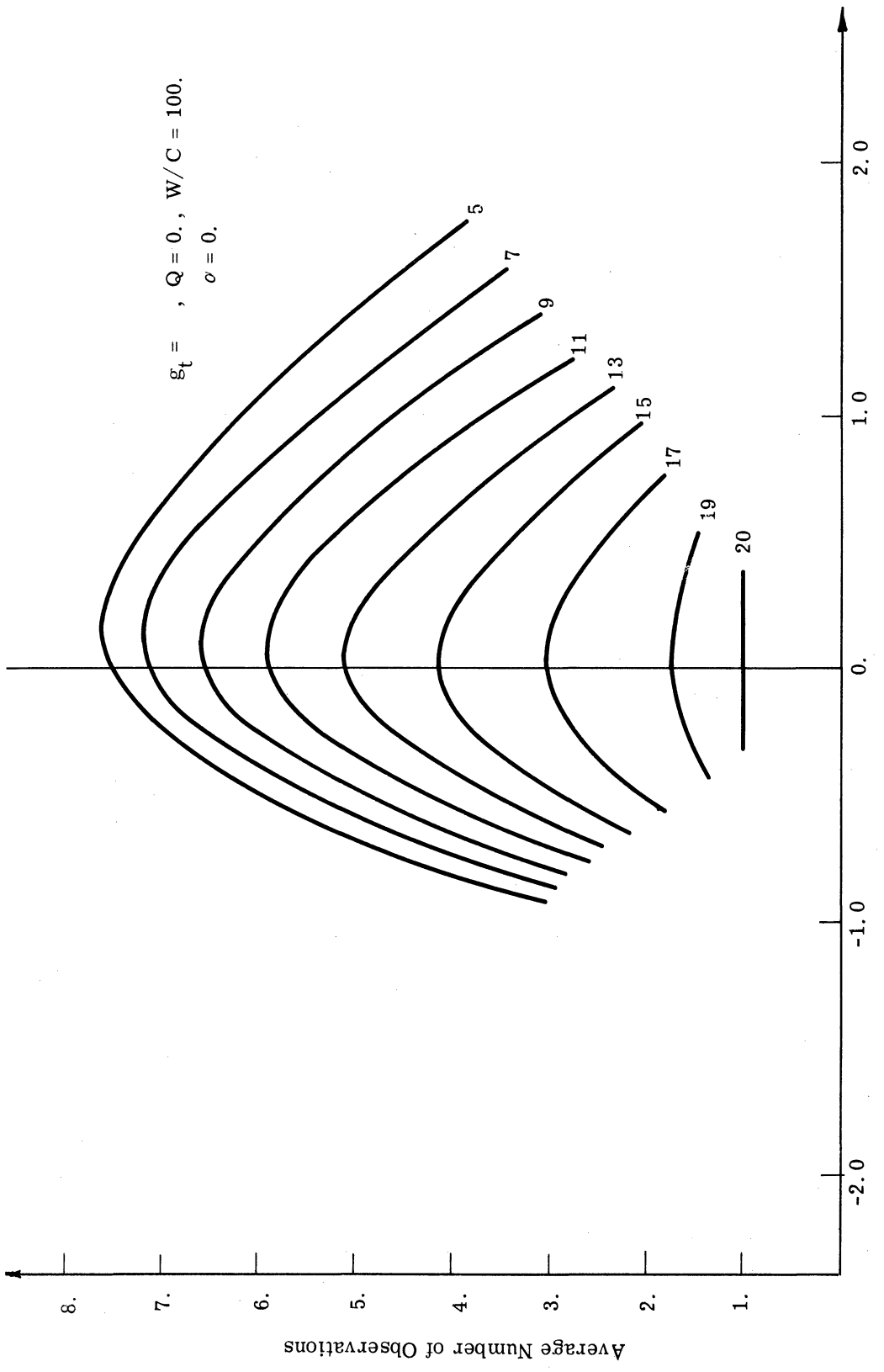


Fig. 4.51. The average number of observations for a deferred-decision receiver as a function of the log-odds ratio with parameters $Q = 0$, $g_t = 21$, $W/C = 100$, and $\alpha = 0$.

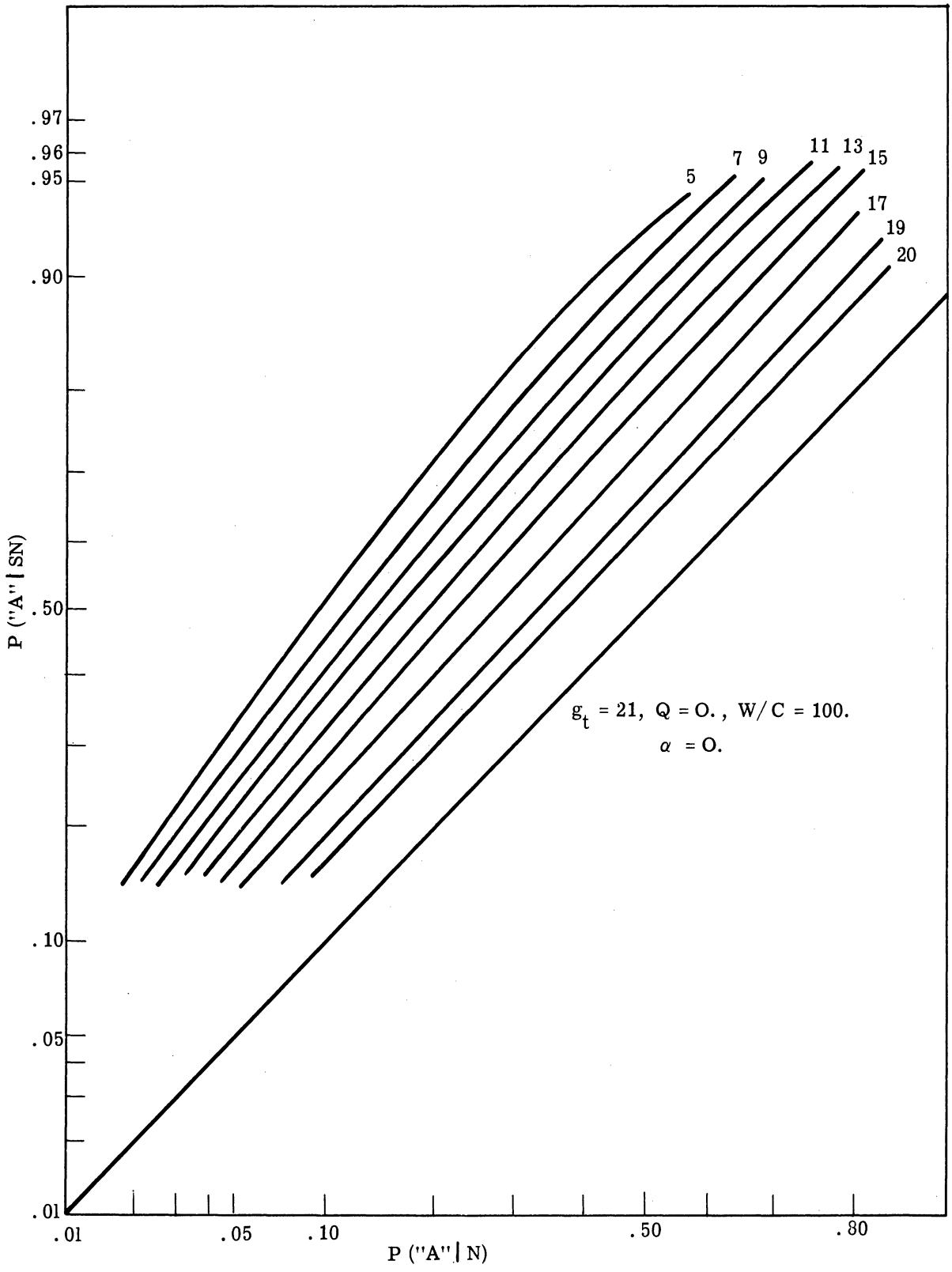


Fig. 4.52. The ROC curves of a deferred-decision receiver with parameters Q = 0, $g_t$ = 21, W/C = 100, and $\alpha$ = 0.

Consider the ROC curves of Figs. 4.50 and 4.52. The curves are plotted with g as a parameter. (If the value of g is subtracted from $g_t$, the resulting number is the available number of observations.) The shape of the ROC curves closely approximates normal ROC curves. (A normal ROC plots as a straight line with unity slope.) Notice that the increase in error performance as more observations are taken is approximately linear. This is in contrast to simple deferred-decision theory in which the ROC curves converge as longer observations are taken.

The comparison between the ROC curves for Q = 0 and Q = -2 shows, as we expect, that the error performance for greater amplitude uncertainty is poorer, i.e., the ROC curves in Fig. 4.52 fall below the comparable ROC curves in Fig. 4.50.

The ANO curves of Figs. 4.49 and 4.51 are plotted as a function of L with g as a parameter. The most outstanding feature of the curves shown in Figs. 4.49 and 4.51 is the large amount curvature they possess. The average observation time becomes small quickly as the log-odds ratio differs from 50-50 odds (L = 0). This phenomenon is more pronounced as the signal amplitude becomes less certain and, again, points to the interdependence between the observer's opinion of the cause of the receiver input and his opinion of the signal amplitude. If the observer's opinion of the cause of the receiver input is strongly in favor of SN, then his opinion of the signal amplitude is biased toward large amplitudes. Conversely, if the oberver's L is negative, he is biased toward small amplitude signals. In either case, his observation time is relatively small. The interdependence of L and f(a) is governed by the value of Q; as Q becomes larger the interdependence becomes greater.

A comparison of Figs. 4.49 and 4.51 shows that greater signal amplitude uncertainty causes a decrease in the average observation time. In other words, as signal uncertainty becomes greater, it becomes less profitable to observe. This same phenomenon has been demonstrated in the optimum nonsequential procedure discussed in Sections 3.4.2 and 4.4.2.

The preceding discussion of the evaluation curves for the SKEA deferred-decision problem is for the $\alpha$ = 0 signal amplitude distribution. The same general conclusions are true for the $\alpha$ = 1 and Gaussian distributions. The next Section, 4.6.3, considers another aspect of the deferred-decision evaluation, comparisons with suboptimum receivers.

4.6.3 Comparisons of the Deferred-Decision Receiver with Suboptimum Receivers. We have presented in this chapter two optimum procedures, one sequential and the other nonsequential. In order to determine the advantage of sequential procedures as compared to nonsequential procedures, we will compare the deferred-decision procedure with the ONP of Section 4.4. All other nonsequential procedures fall below the performance of the ONP, the "upper bound" on performance for fixed-observation procedures.

The standard comparison between sequential and nonsequential procedures, in the past, has been to compare a Wald sequential procedure and a fixed-time test. The comparison is made by assuming the error performance of the Wald procedure and fixed-time test are equal. One then compares the average time necessary to reach the desired error performance. The usual result quoted is that the sequential procedure saves approximately 50% in the average number of observations taken (Ref. 4).

The comparison of a simple deferred-decision procedure and the optimum fixed procedure (Ref. 6) results in quite a different conclusion. The optimization of the two procedures is based on minimizing the average loss for a terminal decision. Under this method of optimization, the fixed observer no longer operates with the error probabilities of sequential observer and, as a result, it is not possible to predict where the savings of the sequential procedure occurs, i.e., in better error performance of small observation times. If the available output signal-to-noise ratio is small (on the order of +4 db in $2E/N_0$), simple deferred-decision procedures consume about 60% of the time of fixed procedures and the resultant error probabilities are approximately equal. However, if the available output signal-to-noise ratio is large (on the order of +10 db in $2E/N_0$), then simple deferred-decision procedures and ONP consume about the same time and the savings occurs in improved error performance.

The comparison of composite deferred-decision procedures and ONP is usually considerably more difficult. For example, in the problem of a SKEA a difficulty arises because the state of the observation depends not only on the available number of observations, $g_t - g$, but also on the number of observations already taken, g. For a meaningful comparison, the two procedures should be compared for the same initial state or a priori information.

The state of the ONP is given by $\{g, Q, L\}$ and $g_t$, if $(g_t - g)$ is less than the optimum observation time. That is, the state of the ONP is dependent on $g_t$ only if the choice of the optimum observation time is restricted by the value of $g_t$. The state of a deferred-decision procedure is $\{g, Q, L, g_t - g\}$ and must always include $g_t$. Thus, there are two methods of comparing deferred decision and ONP for a SKEA. For Q, W/C, and $\Delta_0$ fixed, we can compare, as a function of L, deferred decision and the ONP by varying g with $g_t$ fixed, or by varying $g_t$ with g fixed.

The first method of comparison involves the solution of one deferred-decision procedure and many solutions of the ONP. One interpretation of this type of comparison is, for $(g_t - g) >$ optimum observation time, the effect of a priori amplitude information on detection performance. The second method of comparison involves the solution of many deferred-decision procedures and one ONP solution. The interpretation we give to this comparison is the effect of termination and the usefulness of available observation time, again for $(g_t - g) >$ optimum observation time.

In order to exhibit the characteristics of a comparison between deferred decision and the ONP, consider the following numerical example. The amplitude is distributed according to the $\alpha = 0$ distribution and the parameters of the procedure are $W/C = 50$, $g_t = 21$, $\Delta_0 = 0$. We have chosen to compare the two procedures using two functions, the average loss due to terminal decision errors and the average observation time. The results of the comparisons are shown in Figs. 4.53 and 4.54.

Figure 4.53 depicts the probability of a terminal decision error as a function of the available observation time for a log-odds ratio of zero. The graph for the ONP was obtained by evaluating twenty ONP's with an initial value of g equal to 1, 2, ..., 20. For $(g_t - g)$ greater than the optimum observation, the optimum observation time was used in the comparison; otherwise, $(g_t - g)$ was used for the observation time. Thus, Fig. 4.53 represent data from many ONP's and one deferred-decision procedure.

Figure 4.54 displays the average observation time as a function of the available observation time for a log-odds ratio of zero. From the two figures, we conclude that (1) deferred decision saves primarily through better quality decisions and, except for very small available observation times, takes longer to reach terminal decision than the ONP, and (2) the quality of decision is relatively poor, i.e., the number of terminal decision errors is

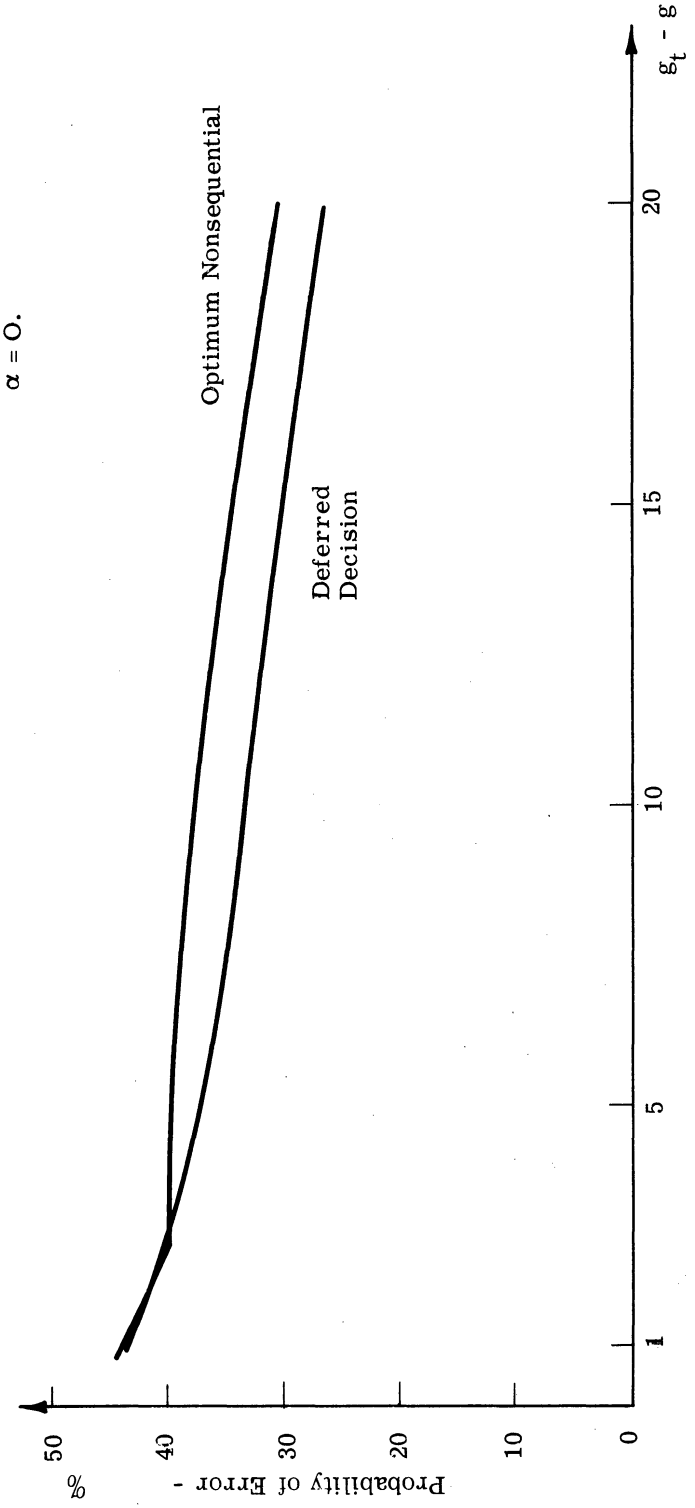


Fig. 4.53. A comparison of probability of terminal decision error as a function of the available observation time for a deferred decision and optimum non-sequential procedure with parameters $g_t = 21$, $Q = 0$, $W/C = 50$, $\alpha = 0$, and $L = 0$.

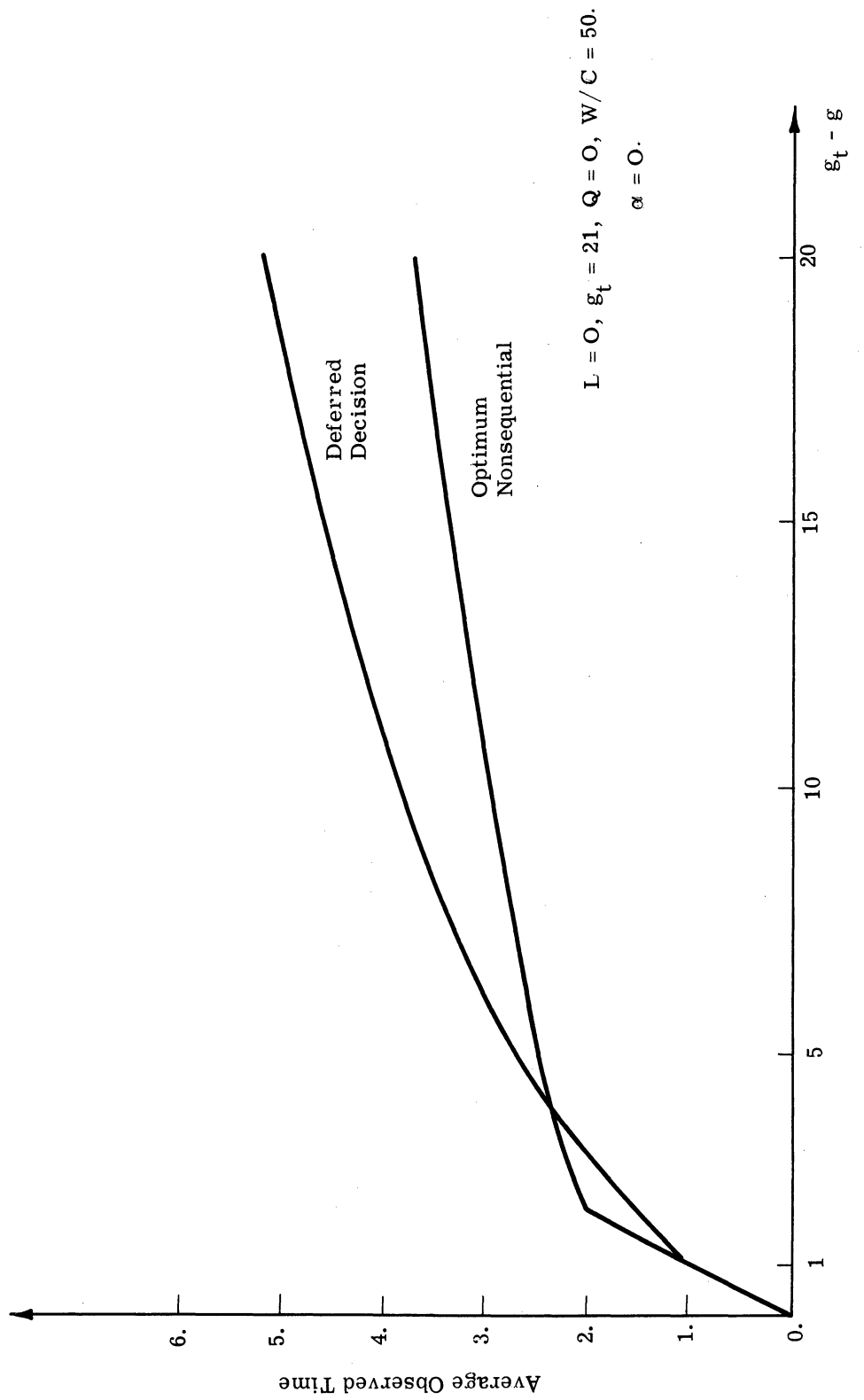


Fig. 4.54. A comparison of average observation time as a function of the available observation time for deferred decision and optimum nonsequential procedure with parameters $g_t = 21$, $Q = 0$, $W/C = 50$, $\alpha = 0$, and $L = 0$.

quite large for the parameters evaluated. In fact, the high rate of terminal decision error, for even the optimum procedure, might be prohibitive in many practical problems.

Notice also that the ONP has a relatively short optimum observation time. This indicates that the nonsequential observer does not bother with small amplitude signals, but instead, detects only large amplitude signals and terminates. Interpreting the graph in Figs. 4.53 and 4.54 as the effect of a priori amplitude information on detection performance, we see the reasonable result that as the a priori amplitude distribution becomes less diffuse (g increasing), the performance of the optimum receiver improves.

The reader may be puzzled as to why the average observed time for the ONP increases as the available observation time increases. This does not occur for simple-signal-hypothesis problems. However, in composite-signal-hypothesis problems the average observed time increases in the ONP because the receiver is constrained to operate on a Q surface.

The comparison of deferred decision and the ONP made by fixing the value of g and varying the value of $g_t$ results in evaluation curves as shown in Figs. 4.55 and 4.56. The numerical example is the same as discussed in Figs. 4.53 and 4.54, namely, Q = 0, W/C = 50, $\Delta_0 = 0$, and L = 0. The initial g value is three.

From the comparison shown in Figs. 4.55 and 4.56, we see, again, that the savings of the sequential procedure occurs primarily in error performance rather than observation time. If we interpret this comparison as exhibiting the effect of termination, we conclude that termination effects are present only in the ONP, a fact we knew beforehand. The deferred-decision procedure does not exhibit any unusual effects as we increase $g_t$. The average observed time increases, with increasing $g_t$, while the probability of terminal decision error decreases, pointing out the balance that occurs between error performance and observation time. As in the previous comparison, the second comparison shows that the optimum procedure incurs a large number of terminal decision errors.

The two most striking conclusions of the comparisons presented between deferred decision and the ONP are: (1) sequential procedures save primarily in better error performance, and (2) the optimum sequential procedure, for the parameters evaluated, incur a large number of terminal decision errors.

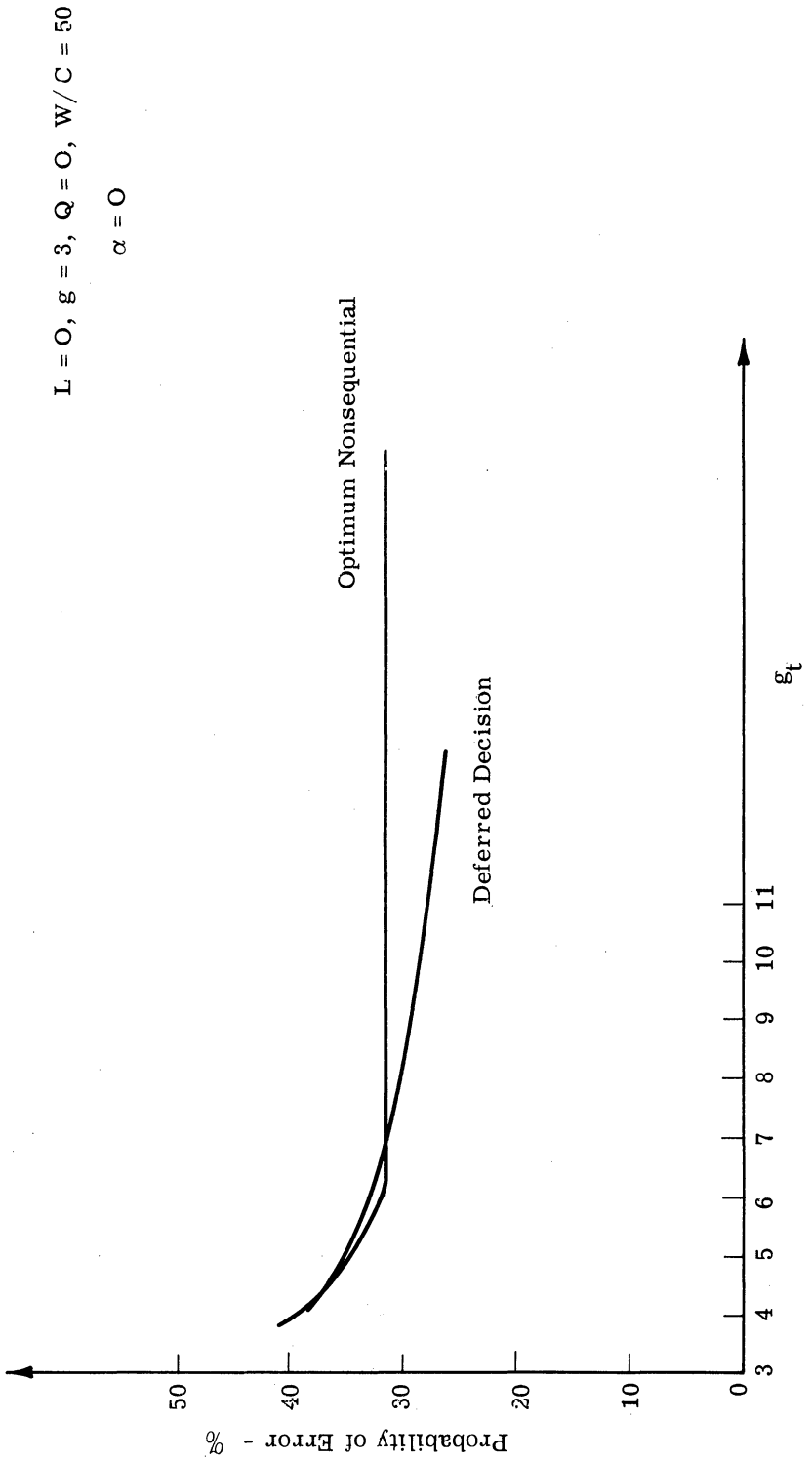


Fig. 4.55. A comparison of probability of terminal decision error as a function of the available observation time for a deferred decision and optimum nonsequential procedure with parameters g = 3, Q = 0, W/C = 50, $\alpha$ = 0, and L = 0.

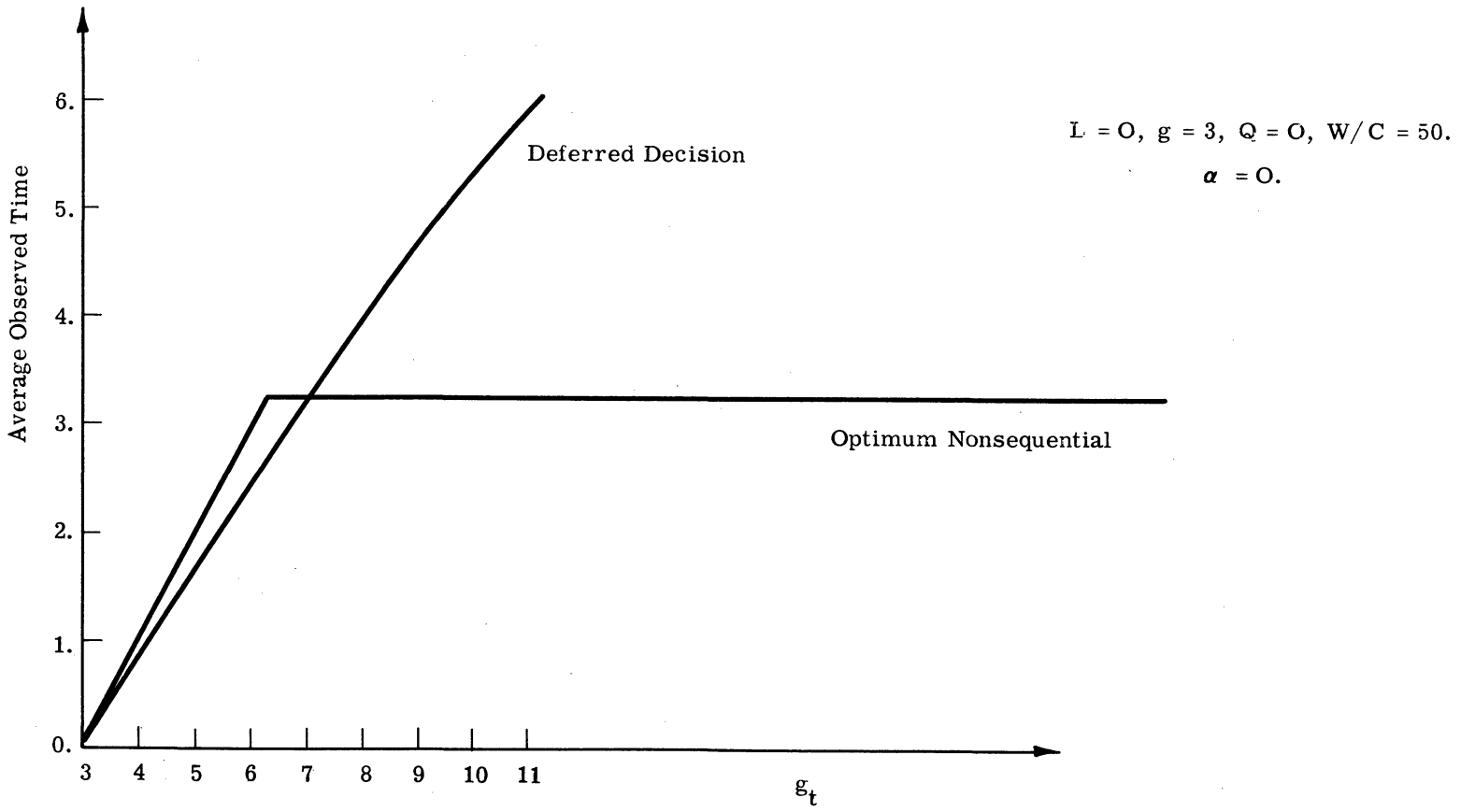


Fig. 4.56. A comparison of average observation time as a function of the available observation time for deferred decision and optimum nonsequential procedure with parameters g = 3, W/C = 50, $\alpha$ = 0, and L = 0.

Nonsequential procedures, if given the option of choosing the observation time, will choose a relatively small observation time. The small observation time implies that the receiver detects only the large amplitude signals. For a nonsequential receiver it does not pay to spend time in order to detect small amplitude signals.

Deferred-decision receivers, on the other hand, attempt to take advantage of "small" and "middle" amplitude signals by continuing to observe. In this way, although observation times are fairly long, better error performance is obtained. However, even the optimum procedures make a large amount of errors.

The question of what is a "middle" amplitude signal is of concern to the sequential observer because these signals take the most time to detect. Small and large amplitude signals should terminate in small observation times because of the Q surface interdependence. The "middle" size signal tends to wander down the center of the stopping rule, resulting in long observation times.

To determine middle amplitude signals, we have evaluated the performance of the optimum receiver for any given amplitude signal. The conditional evaluation is given in terms of the ROC and ANO curves. The derivation of these conditional evaluation curves follows.

If $z_i$ is a receiver input in noise alone, then $z_i$ is N(0, 1); in signal-and-noise $z_i$ is N(a, 1) for a given signal amplitude, a.

The ROC curves can be derived in an iterative manner as follows. For one available observation, the probability of false alarm is

$$\tilde{x}_1(h) = P(\text{"A"} \mid N) = P(L_0 > 0 \mid N)$$

$$= P\left\{ Q + \ell n \left[ \frac{\omega\left(\frac{h_0}{\sqrt{g_0}}\right)}{\sqrt{g_0}} \right] > 0 \middle| N \right\}$$

$$= P\left\{ \omega\left(\frac{h_0}{\sqrt{g_0}}\right) > \sqrt{g_0}\ e^{-Q} \middle| N \right\} = P(h_0 > h^* \mid N) \qquad (4.93)$$

where

$$h^* = \sqrt{g_0}\ \omega^{-1}\left(\sqrt{g}\ e^{-Q}\right) \qquad (4.94)$$

and

$$g_0 = g_1 + 1 \tag{4.95}$$

$$h_0 = h_1 + z \tag{4.96}$$

The distribution $(h_0 - h_1)$ is N(0, 1). Therefore, the probability of false alarm is

$$\begin{aligned} \tilde{x}_1(h) &= \Phi(h_1 - h^*) , \qquad \Gamma_{h_1} < h_1 < \Delta_{h_1} \\ &= 0 , \qquad \Gamma_{h_1} \geq h_1 \\ &= 1 , \qquad \Delta_{h_1} \leq h_1 \end{aligned} \tag{4.97}$$

For two available observations, we average the previous stage's probability of false alarm. Thus

$$\begin{aligned} \tilde{x}_2(h) &= \int \tilde{x}_1(h + z)\, f(z\ N)\, dz \\ &= \int \tilde{x}_1(h + z)\, d\Phi(z) , \qquad \Gamma_{h_2} < h_2 < \Delta_{h_2} \\ &= 0 , \qquad \Gamma_{h_2} \geq h_2 \\ &= 1 , \qquad \Delta_{h_2} \leq h_2 \end{aligned} \tag{4.98}$$

For more than two available observations, the probability of false alarm is found in an iterative manner in the manner shown above.

The probability of detection for a given signal amplitude is also a straightforward application of the iterative procedure, we have presented several times previously. For one available observation, we have

$$\tilde{y}_1(h) = P(\text{"A"} \,|\, SN, a) = P\left\{ Q + \ell n \left[ \frac{\omega\left(\frac{h_0}{\sqrt{g_0}}\right)}{\sqrt{g_0}} \right] > 0 \,\middle|\, SN, a \right\} \tag{4.99}$$

$$= P(h_0 > h^* \,|\, SN, a)$$

where $h^*$, $h_0$, and $g_0$ are defined in Eqs. 4.94 through 4.96. The distribution of $h_0 - h_1$, for a given amplitude signal is N(a, 1). Therefore, we have

$$\begin{aligned}
\tilde{y}_1(h) &= (a + h_1 - h^*) , && \Gamma_{h_1} < h_1 < \Delta_{h_1} \\
&= 0 , && h_1 \geq \Gamma_{h_1} \\
&= 1 , && \Delta_{h_1} \leq h_1
\end{aligned} \tag{4.100}$$

For two available observations, we obtain

$$\begin{aligned}
\tilde{y}_2(h) &= \int \tilde{y}_1(h + z)\, f(z|SN, a)\, dz \\
&= \int \tilde{y}_1(h + z)\, d\Phi(z - a) , && \Gamma_{h_2} < h_2 < \Delta_{h_2} \\
&= 0 , && \Gamma_{h_2} \geq h_2 \\
&= 1 , && \Delta_{h_2} \leq h_2
\end{aligned} \tag{4.101}$$

For more than two available observations, the probability of detection is found by extending the iterative shown above.

The conditional average number of observations is given by a similar iterative solution. For one available observation, we have

$$\begin{aligned}
\tilde{u}_1(h) &= 1 , && \Gamma_{h_1} < h_1 < \Delta_{h_1} \\
&= 0 , && \text{otherwise}
\end{aligned} \tag{4.102}$$

$$\begin{aligned}
\tilde{u}_1(h) &= 1 , && \Gamma_{h_1} < h_1 < \Delta_{h_1} \\
&= 0 , && \text{otherwise}
\end{aligned} \tag{4.103}$$

where $\tilde{u}_k(h)$ and $\tilde{v}_k(h)$ are the ANO curves for k available observations, given the signal amplitude, in SN and N, respectively. For two available observations, we obtain

$$\begin{aligned}
\tilde{v}_2(h) &= \int \tilde{u}_1(h + z)\, d\Phi(z - a) , && \Gamma_{h_2} < h_2 < \Delta_{h_2} \\
&= 0 && \text{otherwise}
\end{aligned} \tag{4.104}$$

$$\tilde{v}_2(h) = \int \tilde{v}_1(h+z)\, d\Phi(z) \quad , \quad \Gamma_{h_2} < h_2 < \Delta_{h_2}$$
$$= 0 \quad , \quad \text{otherwise} \qquad (4.105)$$

For more than two available observations, the ANO curves are found by extending the iterative procedure shown above.

Using the iterative equations derived for the conditional ROC and ANO curves, we can obtain the conditional evaluation curves for the optimum SKEA receiver. The numerical results that follow were obtained by use of an IBM 7090 programmed according to the algorithm given in Appendix K.

Conditional performance curves can be presented in many forms to exhibit various aspects of optimum receiver operation. For example, to determine the middle amplitude signals, we can obtain the performance for various amplitude signals. These conditional performance curves are the performance of the optimum receiver in signal-and-noise for a given amplitude signal. Numerical results of conditional evaluation curves for a deferred-decision procedure with parameters $Q = 0$, $W/C = 50$, $g_t = 21$, $L = 0$, and $\alpha = 0$ are shown in Figs. 4.57 and 4.58.

Figure 4.57 is a plot of the probability of error versus the available observation time for various amplitude signals under the condition signal-and-noise. Because the condition at the input is signal-and-noise, the probability of error is the probability of a miss. Figure 4.58 is a plot of the average observation time versus the available observation time for various amplitude signals in signal-and-noise.

From Fig. 4.58, we see that an amplitude signal of approximately 0.2 results in the longest average observation time. Thus, we conclude that an amplitude of 0.2 is the "middle" amplitude signal. This same result can be obtained by referring to Fig. 4.59 depicting the optimum decision boundaries in terms of h and g, and shows the $L = 0$ contour (dotted line) and the "average motions" of various amplitude signals. The "average motions" are the average trajectories of the various amplitude signals in h x g space. In other words, the various arrows represent the average operational path of a given amplitude signal on the h x g plane. Notice, from Fig. 4.59, that there exists a middle amplitude signal of approximately 0.25 resulting in an average motion down the center of the decision boundaries.

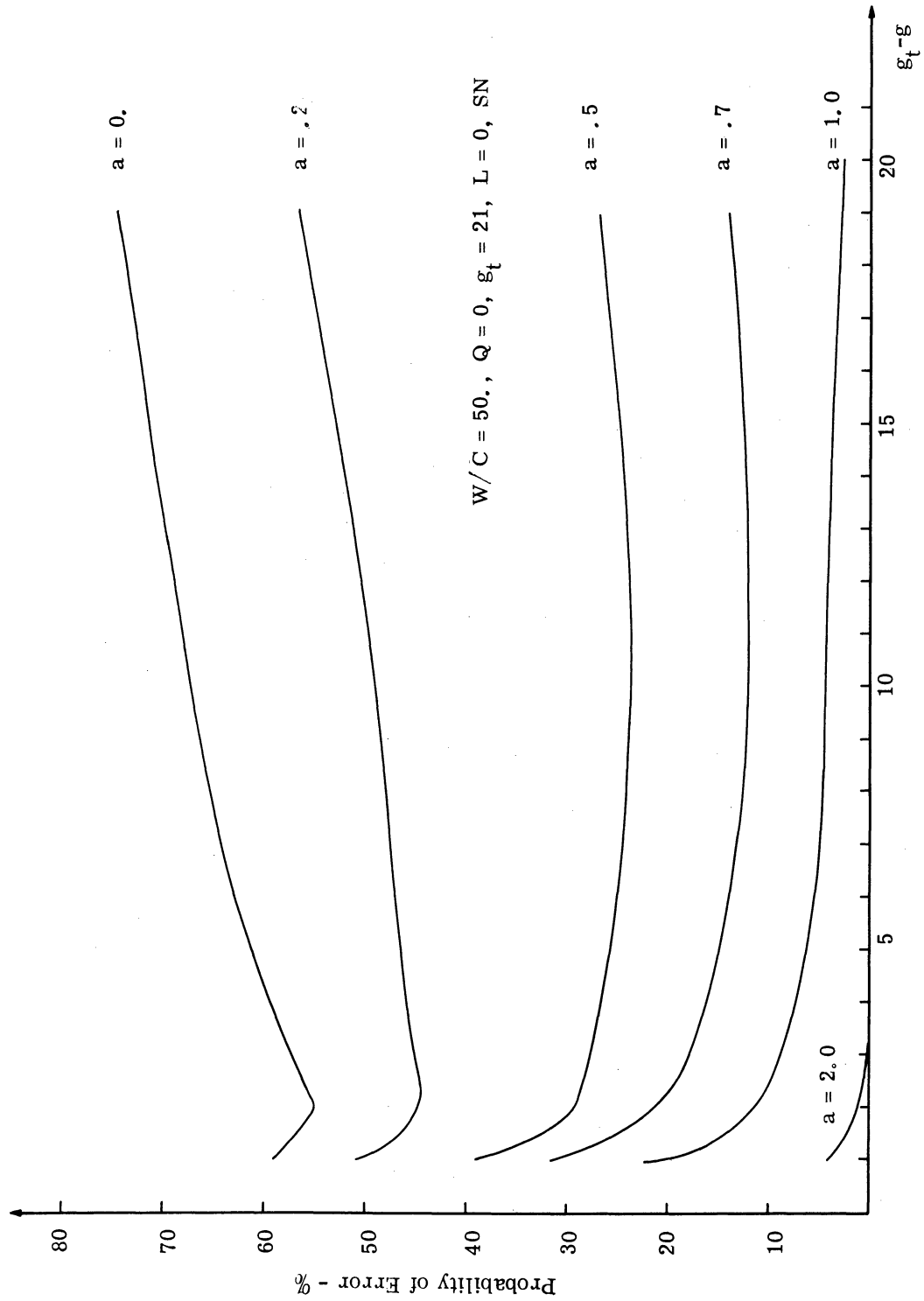


Fig. 4.57. The conditional probability of error (miss) curves as a function of the available observation time for several amplitude signals in signal-and-noise.

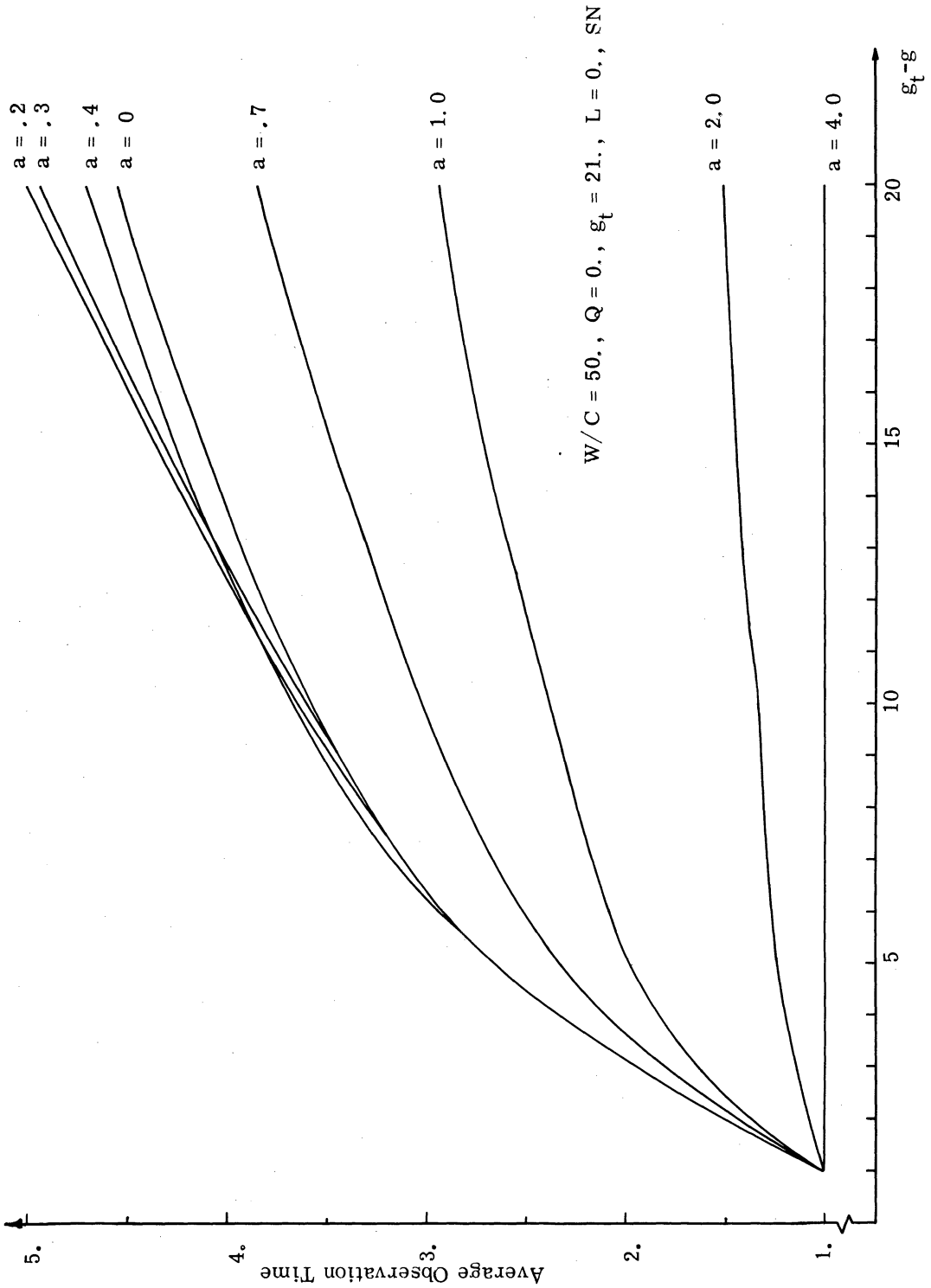


Fig. 4.58. The conditional number of observation curves as a function of the available observation time for several amplitude signals in signal-and-noise.

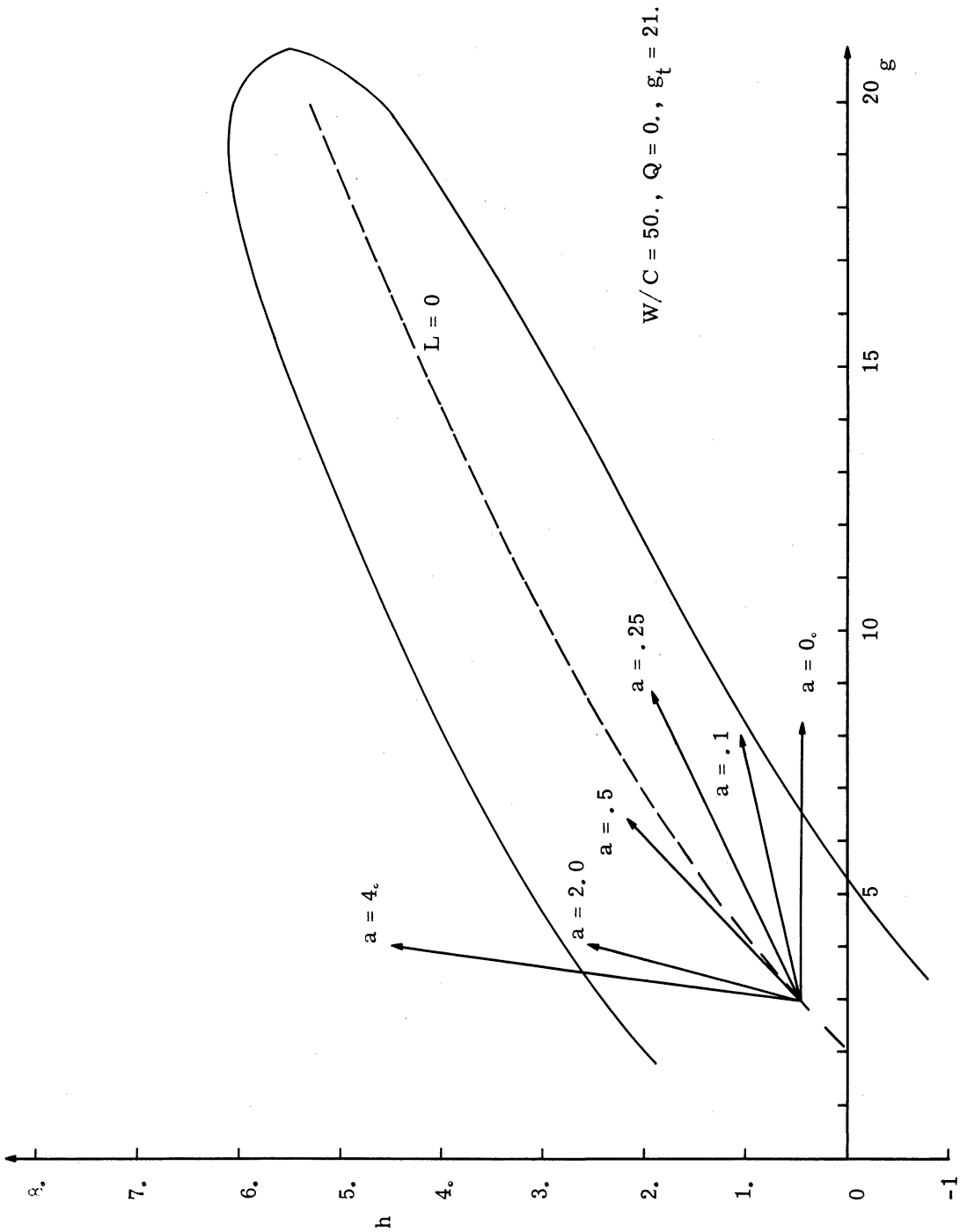


Fig. 4.59. The optimum deferred-decision boundaries and average trajectories of several amplitude signals for W/C = 50, Q = 0, $g_t$ = 21, and $\alpha$ = 0.

This, of course, implies that the longest observation time occurs for signals of value approximately 0.25, agreeing very well with computer computations depicted in Fig. 4.58.

It is most convenient and practical to obtain results from Fig. 4.59, since this figure can be obtained from the computer program for the optimum boundaries. If it is possible to predict the operation of various amplitude signals from a knowledge of the optimum decision boundaries and the average motion, then we can eliminate a large amount of computation in the evaluation programs. We have demonstrated the ability to predict the value of the middle amplitude signal. Let us consider some further properties of the conditional performance curves of Figs. 4.57 and 4.58, and our ability to predict certain of these properties by means of the decision boundaries and the amplitude mean motions.

From Fig. 4.58, the average observed times for each of the amplitudes examined are presented. It is possible, using Fig. 4.59, to predict the approximate average observation time for those amplitudes which terminate relatively quickly. Thus, the conditional average time for an amplitude of zero is approximately four observations from Fig. 4.59 which agrees reasonably well with the computer computations. It is difficult to predict the average observation time for the middle amplitude signals from Fig. 4.59, because the variance of the tip of the mean motion vector is large in comparison to the width of the boundaries.

Notice, in Figs. 4.57, the increase in the miss probability as the available observation time increases for small signal amplitudes. This can also be explained from the decision boundary--mean motion graph. The increase in miss probability as $(g_t - g)$ increases is due to the curvature of the boundaries. A miss occurs whenever the integrated input reception crosses the lower decision boundary in signal-and-noise. As we increase the available observation time (decrease g), the curvature of the boundaries causes small amplitude signals to have a greater probability of crossing the lower boundary. Therefore, the error probability increases for small amplitude signals. For large amplitude signals, the mean motion is so strongly upward that it overcomes any boundary curvature.

From the preceding discussion, it is clear that a great deal of both qualitative and quantitative information of optimum receiver operation can be obtained from a decision boundary--mean motion graph. This graph gives one a good intuitive feel of the performance of various size signals, and can be used as a method of predicting the middle amplitude signal.

The middle amplitude signal may be an important practical design parameter. Although the mathematical problem assumes an amplitude distribution defined on $[0, \infty)$, a practical receiver must work over a limited dynamic range. The middle amplitude signal would appear to be an appropriate signal to center the receiver's dynamic range. Thus, it might be of practical interest to know the value of middle amplitude signals even in the design of suboptimum receivers.

From Figs. 4.57 through 4.59, we might classify amplitudes below 0.7 -1, as relatively small amplitudes and amplitudes above 1.5 -2.0 as large amplitude signals. It is interesting to compare the probability of the large amplitudes against the small amplitude signals, since the optimum SKEA receiver is, in some sense, a balanced design taking into account both the large and small amplitude signals. A plot of a four-representative amplitude distributions, shown in Fig. 4.60, indicates that the probability of the large amplitudes is less than 5% of the total amplitude distribution function.

The final numerical computations we present below are the performance curves of the optimum SKEA receiver for various size signals. These curves represent the performance we would expect if the noise and signal were present an equal amount of time, i.e., $L = 0$. The results are similar to the conditional signal-and-noise only performance of Figs. 4.57 and 4.58. The addition of noise alone for half the time has tended to smooth out large differences between the different signal amplitudes. Conclusions we obtain from Figs. 4.61 and 4.62 do not add to previously obtained concepts.

The evaluation of suboptimum receivers is a fascinating but never-ending job. We have written a computer program that can be used to evaluate various suboptimum receivers. Many suboptimum receivers were evaluated, but the results have not been included, because they are, in retrospect, largely self evident. If a suboptimum receiver is a receiver with decision boundaries approximating the optimum decision boundaries, the receiver's performance is close to optimum performance. A poor approximation to the optimum decision boundaries results in poor average performance. Until such time as we can obtain approximate methods for obtaining stopping rules based on the a priori amplitude distribution, the evaluation of suboptimum receivers does not appear to yield any further insight into the design of optimum detectors.

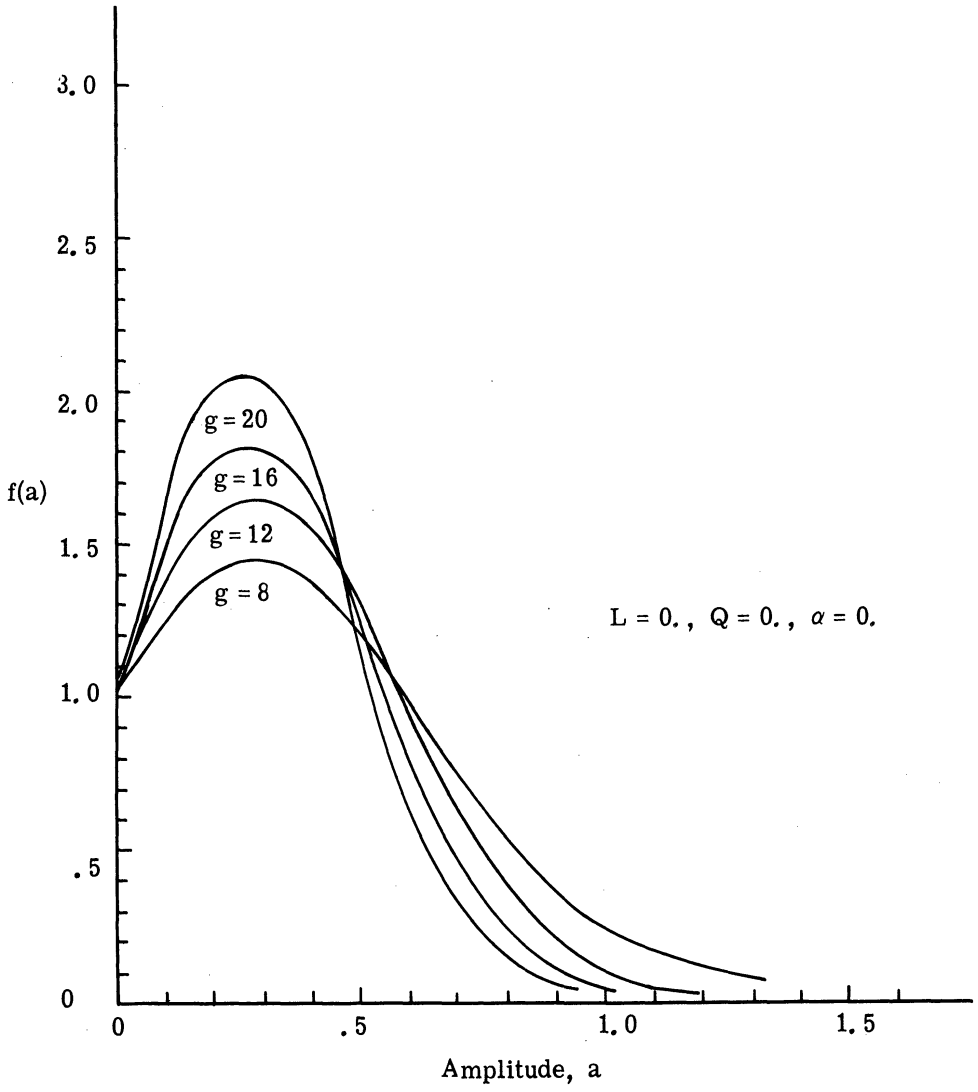


Fig. 4.60. Probability density functions for the truncated Gaussian distribution with parameters Q = 0, L = 0, and g = 8, 12, 16, and 20.

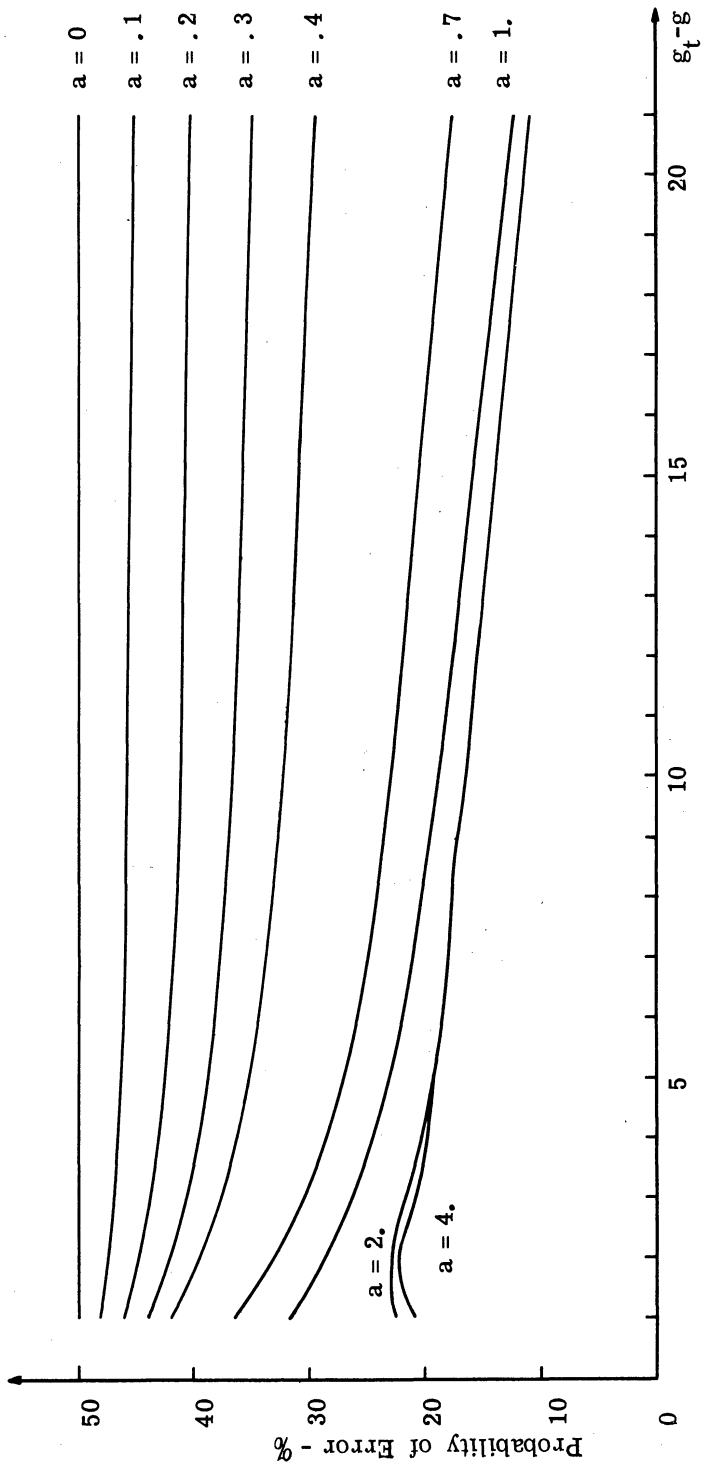


Fig. 4.61. The probability of error curves for a terminal decision as a function of the available observation time for several amplitude signals.

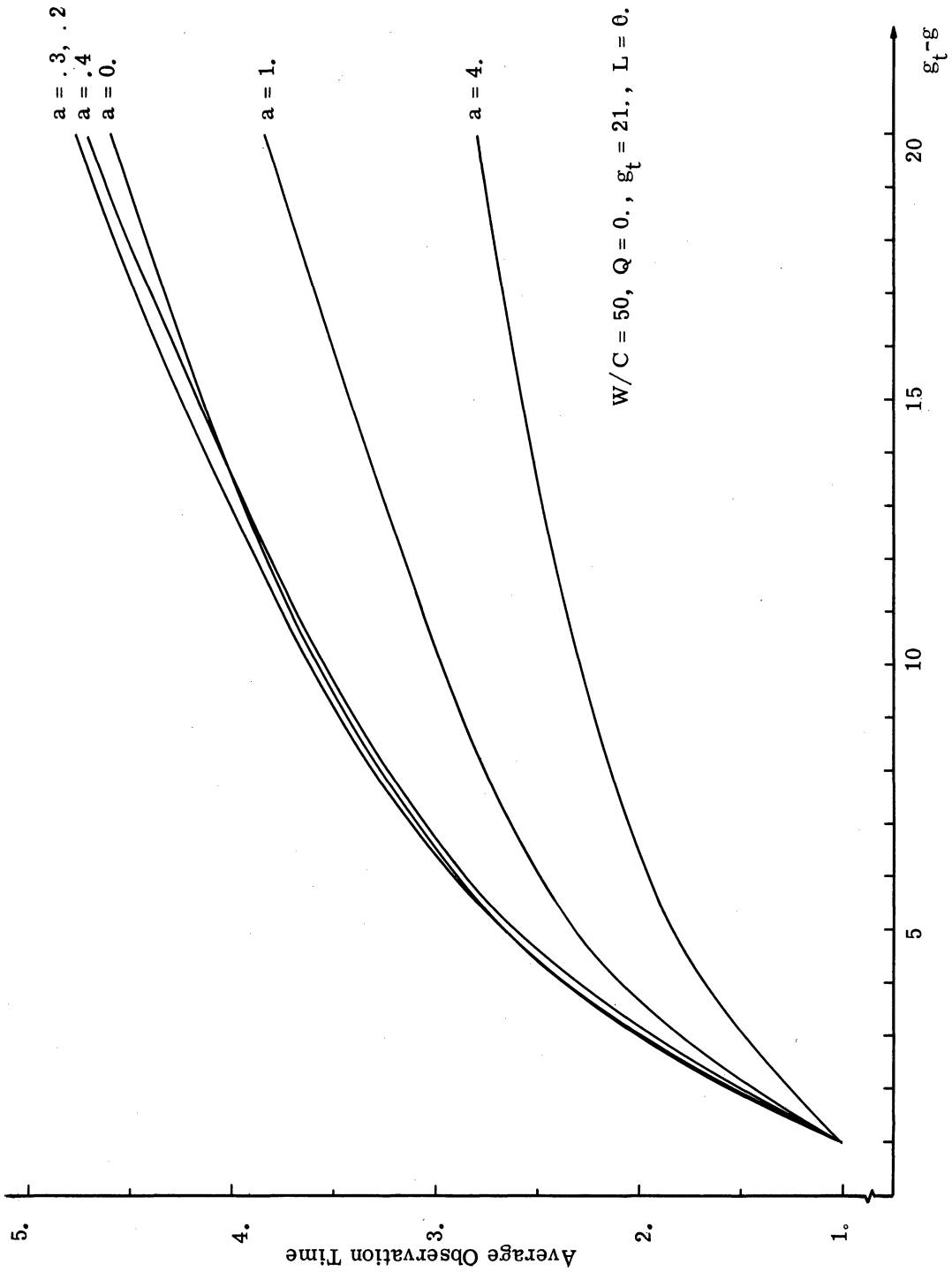


Fig. 4.62. The average number of observation curves as a function of the available observation time for several amplitude signals.

## 4.7 Summary of Chapter IV

The study of detecting a distributed signal in noise presented here has, for the first time, included: (1) fixed-detection procedures for a priori amplitude distributions other than the Rayleigh and Gaussian, (2) the optimum nonsequential procedure and evaluation, (3) the optimum sequential procedure, deferred decision, and evaluation, and (4) comparisons between (2) and (3) above. The form of the optimum sequential detector for a SKEA indicates that the receiver must include an adaptive capability and must extract from the observation, not only the likelihood ratio, but also information to update the signal parameter distribution. Numerical results of the detection of a SKEA presented here are, perhaps, the most important contributions of this chapter. These results are summarized in the sections in which they appear but are not repeated here because of their length.

# CHAPTER V

We have derived the optimum, i. e., minimum expected loss for a terminal decision, sequential and nonsequential detection procedure for a composite signal hypothesis. In Chapter II the sequential procedure is presented in generality. Chapters III and IV examine two specific problems, a signal known except for phase and a signal known except for amplitude, respectively. The numerical results of these two problems include evaluations of the optimum procedure and comparisons with suboptimum procedures.

The conclusions and results of this work can be separated into two parts: (1) those that pertain to the general derivation and the general form of the optimum detector, and (2) those that stem from the numerical calculations.

## 5. 1 Conclusions from the General Solution

The general formulation and solution of composite deferred-decision problems are given in Chapter II. The general derivation leads to some striking conclusions. These results are, perhaps, best summarized by the block diagram of the optimum receiver shown in Fig. 4. 38.

Recall that in simple deferred decision theory (Ref. 6), the likelihood ratio of the input observation is sufficient to ascertain both the optimum terminal decision and the optimum stopping rule. To be sure, other parameters of the problem must be known, but from the observation itself only the likelihood ratio is necessary. In other words, in simple deferred decision theory the observation information needed for optimum detection is completely contained in the likelihood ratio of the observation.

The results of Chapter II, however, show that the likelihood ratio of the observation does not give the necessary observation information needed for optimum detections in composite-hypothesis problems. The optimum detector must, from the observations, obtain information to update the signal parameter distribution. Without this information the receiver cannot determine the average likelihood ratio needed in terminal decisions and the correct set of decision boundaries needed for the optimum stopping rule. The receiver must

process the observation to calculate the updated signal parameter distribution in addition to the likelihood ratio. This updating (learning or adaptation) is performed after each observation or whenever new information pertaining to the detection problem is obtained.

The general form of a composite deferred-decision detector is shown in Fig. 4.40, The adaptive portion of the receiver is the block that modifies the signal parameter density function after each observation. From Fig. 4.40, we conclude that the detection receiver for a composite-signal-hypothesis problem is a device where an adaptive form must be used to obtain optimum performance. The adaptive nature of the optimum receiver is a result of optimizing detection performance rather than ad hoc postulation that adaptation is necessary for optimum performance. This last point is made because in present day literature, adaptation is sometimes postulated as a necessary part of the receiver. It is important to note that the adaptive portion of the composite deferred-decision receiver may not be overtly present in every receiver realization. That is, it may be difficult to separate the receiver into blocks as shown in Fig. 4.40.

A second conclusion drawn from the general block diagram of Fig. 4.40, that we will only mention, is the classification output that is available. The probability density function of the signal parameteris the observer's opinion of signal parameter. Thus, it is precisely the function needed for classification. Again, it may be difficult to design a receiver in which the classification is easily obtained. Classification is a "by-product" of the solution of composite deferred decision rather than a primary objective of the solution just as the adaptive form of the receiver is a "by-product."

The final summary remark concerning the general derivation of a composite deferred-decision problem is not obtained from any result per se, but rather pertains to our over-all receiver design philosophy. We have presented the solution for a distributed signal parameter detection problems by incorporating the complete signal parameter distribution in the optimum receiver design. This has not been done previously in sequential procedures: The more generally used design philosophy is to estimate the signal parameter in some manner, and design a fixed receiver for the estimated value assumed to be the "true" value. This eliminates any possibility of adaptation or learning during the observation process, and, for sequential procedures, reduces the problem to that of simple deferred-decision theory. Designing receivers on the basis of estimates is a reasonable engineering

procedure. However, it is suboptimum to incorporating the complete signal parameter distribution in the over-all design. There are, to be sure, problems connected with our design philosophy. For example, we must use closed-signal parameter distributions. However, since the solution method presented is optimum, it may prove valuable in, at least, establishing an upper limit on performance.

## 5.2 Conclusions From Numerical Computations

The results and conclusions of the numerical computations have been given throughout this work as they occur. They are too numerous to repeat in detail here. As a summary of the numerical results, we present general trends and concepts indicated by our calculations of the effect of signal uncertainty in sequential and nonsequential detection problems.

Consider the results obtained from ONP calculations for a SKEP and a SKEA. The computations are best summarized by the value contour graph of an observation as a function of observation time (or quality) and the log-odds ratio. The value contour graphs of the SKEP and SKEA problem indicate that: (1) as signal uncertainty increases, the L range over which it is profitable to observe decreases, (2) it is profitable to observe only if it is possible for the observer to change his a priori terminal decision upon receipt of observation information, and (3) large amounts of signal uncertainty imply relatively small optimum observation times. From the ROC curves of the fixed procedures, we conclude that the rate of improvement in performance decreases with longer observation times. This last conclusion can be seen from the convergence of the ROC curves as longer observations are taken.

The results of the computations for the deferred-decision procedure for a SKEA can be divided into two parts; those pertaining to the decision boundary, and the results concerned with performance curves of the optimum receiver.

Consider the numerical computations of the optimum stopping rule. The complete description of the stopping rules is a time-consuming computational job because of its three-dimensional character. Computations indicate, however, that reasonable approximations to the projected decision boundary in, say, the L x g space result in performance close to optimum. For example, if the optimum boundary is approximated by two constant

boundaries (in time) which are spaced to fit the optimum boundary, then there is very little difference in the performance of a receiver using the optimum and suboptimum boundaries.

One of the most interesting results of the SKEA problem is that the deferred-decision process does not permit an asymptotic solution. This is the only sequential detection problem, known to the author, that does not permit unbounded available observation times. This is a result of the signal uncertainty of an unknown amplitude signal. The corresponding problem for a signal of unknown phase does possess an asymptotic solution. The absence of an asymptotic solution creates a problem, because it eliminates the possibility of useful approximations. One of the most useful approximations in simple deferred-decision theory is the approximation of the optimum decision boundary by the very simple asymptotic solution to the decision boundary. Asymptotic solutions are useful as they can often be solved analytically without resorting to computer computations.

The conclusions taken from the performance curves for a SKEA receiver are the most appropriate for an intuitive understanding of the basic detection problem. The performance curves that result from a comparison of the ONP and deferred decision indicate that deferred decision saves primarily in error performance. Quantitatively, the average performance of even the optimum procedure is not good enough for many practical problems, i. e., the average error rate (40% errors) is too high. The result of a high average number of errors is because the optimum decision boundaries are spaced very close together. Comparable decision boundaries for a signal known exactly are five or six times more widely spaced. The close spacing of the SKEA decision boundaries is a result of the amplitude distribution squeezing down near zero for negative L values, i. e., the L and f(a) dependence on a Q surface. This dependence creates a situation in which it does not pay to "work on" small signals, thereby causing the boundaries to be spaced close together. For comparable W/C ratios, the error performance for SKEA is about an order of magnitude worse than for a signal known exactly.

Conditional performance curves of the optimum receiver in signal-and-noise indicate the existence of a middle amplitude signal which has an average observation time longer than very small or very large signals. The middle size signal is the signal that causes the integrated input to drift down the center of the decision boundaries. This signal

may be an important design parameter in practical receivers for use in centering the receiver's dynamic range. The middle size signal can be found by means of a most useful graph, the decision boundary-mean motion graph. This graph depicts the decision boundaries, for a given Q surface, and the average trajectory of a given amplitude signal. The graph is very helpful in gaining an intuitive feeling of the operation of the optimum receiver.

## 5.3 Additional Results of the SKEP and SKEA Detection Problems

In addition to the conslusions and results presented in Sections 5.1 and 5.2, mention should be made of the following new results. These results are not logically placed in either Section 5.1 or 5.2.

The design and evaluation of the optimum SKEP receiver for fixed procedures produced the following new results. From the ROC curves, we conclude that phase information to within $65^{o}$ is, for detection, a signal known exactly. The optimum receiver design of Fig. 3.7 presents an adaptive receiver which yields a classification output. Using this receiver and Fig. 3.1, one can obtain phase angle dispersion as a function of the input signal-to-noise ratio. This method of obtaining phase angle dispersion as a function of the input signal-to-noise ratio and integration time is a new and, perhaps, practical result. The optimum sequential receiver designs of Figs. 3.6 and 3.7 are practical receiver designs for present-day technology, and could be implemented as they are presented in block diagram form.

The conclusions from the design and evaluation of the optimum SKEA receiver have been discussed in Sections 5.1 and 5.2.

## 5.4 Future Work

The results of this work can be used as the basis of considering further types of signal uncertainty in sequential observation-decision procedures. For example, one might consider the detection of a signal of unknown phase and amplitude as a logical extension of the present work. This problem may be simpler than the unknown amplitude problem, and is, basically, a problem of determining an unknown position in two space.

As we have already mentioned, the average error rate obtained by the optimum SKEA receiver is too high for the W/C ratios investigated. This implies that we should increase the W/C ratio, thereby making it profitable to detect small signals. However, increasing the W/C ratio by an order of magnitude results in drastic computational problems.

The main cause of the high rate of error can be traced to the effect of negative L's on the signal amplitude distribution, as we have previously explained (Section 5.2). In order to limit the high error rate without resorting to large W/C ratios, it may be possible to work with truncated signal amplitude distributions so that negative L's would shift the amplitude distribution not to zero, but to some preset truncation value greater than zero. This, it appears, may permit one to investigate the more practical detection range of 99% detections rather than the present 60% detections.

The problem of considering time uncertainties in the signal ensemble is another problem area which has not been investigated. Thus far we have assumed complete time knowledge in the signal ensemble. In our present framework, we would have to consider problems in which the time uncertainty is learnable. That is, the unknown time pattern would have to exhibit some type of dependence from observation to observation.

Another problem that has not been considered is the problem of embedding the present detection problem into a series of detection problems. In this type of problem the optimization of a single terminal decision may not be as crucial as optimizing the maximum number of correct terminal decisions or some other quantity. This problem might first be considered for a signal known exactly and later extended to include signals with distributed parameters.

## APPENDIX A

## INVESTIGATION OF THE INTEGRAL

$$\int_{\beta_c}^{\infty} \exp\left[\frac{-\alpha^2}{2}\right] I_0(\alpha\sqrt{2}\,t) \exp\left[\frac{-\gamma^2}{2}\right] I_0(\gamma\sqrt{2}\,t)\, t \exp\left[\frac{-t^2}{2}\right] dt$$

The following integral arises in the calculation of the detection probability of a signal known except for phase. Two questions are of interest: (1) the value of the definite integral from zero to infinity, and (2) a closed-form expression for the integral in terms of well-known (tabulated) functions. The answer to the first question is the normalizing constant for P("A"|SN). The answer to the second question is the value of P("A"|SN). The integral is

$$I = \int_{\beta_c}^{\infty} \exp\left[\frac{-\alpha^2}{2}\right] I_0(\alpha\sqrt{2}\,t) \exp\left[\frac{-\gamma^2}{2}\right] I_0(\gamma\sqrt{2}\,t)\, t \exp\left[\frac{-t^2}{2}\right] dt \qquad (A.1)$$

Now

$$I_0(\alpha\sqrt{2}\,t)\, I_0(\gamma\sqrt{2}\,t) = \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} \frac{\alpha^{2k}\,\gamma^{2n}\left(\frac{\sqrt{2}t}{2}\right)^{2k}\left(\frac{\sqrt{2}t}{2}\right)^{2n}}{n!\,k!\,\Gamma(n+1)\,\Gamma(k+1)} \qquad (A.2)$$

Using the relationships $\Gamma(x+1) = x!$ and $\sum_{n=0}^{\infty}\sum_{k=0}^{n} A(n,k) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} A(n-k,k)$, we can rewrite Eq. A.2 as

$$I_0(\alpha\sqrt{2}\,t)\, I_0(\gamma\sqrt{2}\,t) = \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{\alpha^{2k}\,\gamma^{2n-2k}\,(t/2)^{n}}{[(n-k)!]^2\,(k!)^2} \qquad (A.3)$$

The definition of the ordinary hypergeometric function, ${}_2F_1$, can be used to rewrite Eq. A.3. The ordinary hypergeometric function is defined by Eq. A.4.

$${}_2F_1\left[\begin{matrix} a, b; \\ c; \end{matrix} z\right] = \sum_{n=0}^{\infty} \frac{(a)_n\,(b)_n\,z^n}{(c)_n\,n!} \qquad (A.4)$$

The notation $(a)_n$ means

$$(a)_n = a(a+1)(a+2), \ldots, (a+n-1) \qquad (A.5)$$

From Eq. A. 4, we have the following

$$2F1\left[\begin{matrix}-n, -n; \\ 1;\end{matrix}\ \left(\frac{\alpha}{\gamma}\right)^2\right] = \sum_{k=0}^{\infty} \frac{(-1)^k n!\,(-1)^k n!\left(\frac{\alpha}{\gamma}\right)^{2k}}{(n-k)!\,(n-k)!\,(1)_k\,k!} \tag{A.6}$$

$$= \sum_{k=0}^{n} \frac{(n!)^2\left(\frac{\alpha}{\gamma}\right)^{2k}}{[(n-k)!]^2\,(k!)^2} \tag{A.7}$$

Therefore, using Eq. A. 7, we can write Eq. A. 3 as

$$I_0(\alpha\sqrt{2}\,t)\,I_0(\gamma\sqrt{2}\,t) = \sum_{n=0}^{\infty} 2F1\left[\begin{matrix}-n, -n; \\ 1;\end{matrix}\ \left(\frac{\alpha}{\gamma}\right)^2\right] \frac{\gamma^{2n}\,(t/2)^n}{(n!)^2} \tag{A.8}$$

There is a relationship between the ordinary hypergeometric function and the Legendre polynomials (Ref. 23), as shown in Eq. A. 9.

$$P_n(x) = \left(\frac{x+1}{2}\right)^n 2F1\left[\begin{matrix}-n, -n; \\ 1;\end{matrix}\ \frac{x-1}{x+1}\right] \tag{A.9}$$

Using the results of Eq. A. 9, we can write the integral in Eq. A. 1 as

$$I = \int_{\beta_c}^{\infty} \exp\left[\frac{-\alpha^2}{2}\right] \exp\left[\frac{-\gamma^2}{2}\right] \sum_{n=0}^{\infty} P_n\left[\frac{\gamma^2+\alpha^2}{\gamma^2-\alpha^2}\right] \frac{\gamma^2-\alpha^2}{(n!)^2}\,(t/2)^n\,t\,\exp\left[\frac{-t^2}{2}\right] dt \tag{A.10}$$

For $\beta_c$ equal to zero, we have

$$I = \exp\left[\frac{-\alpha^2}{2}\,\frac{-\gamma^2}{2}\right] \sum_{n=0}^{\infty} P_n\left[\frac{\gamma^2+\alpha^2}{\gamma^2-\alpha^2}\right] (\gamma^2-\alpha^2)^n \int_0^{\infty} \frac{t^n \exp(-t)}{2^n (n!)^2}\,dt \tag{A.11}$$

However,

$$\int_0^{\infty} t^n \exp[-t]\,dt = n! \tag{A.12}$$

Therefore, Eq. A. 11 can be written as

$$I = \sum_{n=0}^{\infty} P_n\left[\frac{\gamma^2+\alpha^2}{\gamma^2-\alpha^2}\right] \left(\frac{\gamma^2-\alpha^2}{2}\right)^n \frac{1}{n!} \exp\left[\frac{-\alpha^2}{2}\,\frac{-\gamma^2}{2}\right] \tag{A.13}$$

Equation A. 13 can be simplified in terms of Bessels function by means of Eq. A. 14.

$$\exp[xt]\ J_0(t\sqrt{1-x^2}) = \sum_{n=0}^{\infty} \frac{P_n(x)\, t^n}{n!} \qquad \text{(A. 14)}$$

By use of the identity given in Eq. A. 14, we can write the integral of interest as

$$\begin{aligned} I &= \exp\left[\frac{\gamma^2+\alpha^2}{\gamma^2-\alpha^2}\right] \cdot \frac{\gamma^2-\alpha^2}{2} \quad J_0\left[\frac{\gamma^2-\alpha^2}{2}\sqrt{1-\left(\frac{\gamma^2+\alpha^2}{\gamma^2-\alpha^2}\right)^2}\right] \exp\left[\frac{-\gamma^2}{2}\,\frac{-\alpha^2}{2}\right] \\ &= J_0(i\alpha\gamma) \\ &= I_0(\alpha\gamma) \end{aligned} \qquad \text{(A. 15)}$$

Equation A. 15 is the normalizing constant for the integrand of Eq. A. 1, and is the answer to the first question. The answer to the second question posed is evident from Eq. A. 11. For $\beta_c \neq 0$, Eq. A. 11 can be written

$$I = \exp\left[\frac{-\alpha^2}{2}\,\frac{-\gamma^2}{2}\right] \sum_{n=0}^{\infty} P_n\left[\frac{\gamma^2+\alpha^2}{\gamma^2-\alpha^2}\right] (\gamma^2-\alpha^2)^n \int_{\beta_c}^{\infty} \frac{t^n \exp[-t]}{2^n (n!)^2}\, dt \qquad \text{(A. 16)}$$

The difficulty in obtaining a closed-form answer lies in the evaluation of the integral $\int_{\beta_c}^{\infty} \frac{t^n \exp[-t]}{2^n (n!)^2}\, dt$. This integral is the incomplete gamma function. The author was unable to find a closed-form solution in terms of tabulated functions. The numerical evaluation of Eq. A. 1 is obtained by means of numerical integration. Using the results of Eq. A. 15, we define a new function of three variables, $R(\alpha, \gamma, \beta_c)$ as

$$R(\alpha, \gamma, \beta_c) = \frac{1}{I_0(\alpha\gamma)} \int_{\beta_c}^{\infty} \exp\left[\frac{-\alpha^2}{2}\right] I_0(\alpha t) \exp\left[\frac{-\gamma^2}{2}\right] I_0(\gamma t)\, t \exp\left[\frac{-t^2}{2}\right] dt \qquad \text{(A. 17)}$$

The approximations used in the evaluation of $R(\alpha, \gamma, \beta_c)$ are given in Appendix B.

## APPENDIX B

## APPROXIMATIONS TO $R(\alpha, \gamma, \beta_c)$ AND $Q(\alpha, \beta_c)$

Numerical evaluation of $Q(\alpha, \beta_c)$ and $R(\alpha, \gamma, \beta_c)$ is facilitated by various approximations depending on the values of the arguments $\alpha$, $\gamma$, and $\beta_c$. Consider first the evaluation of $R(\alpha, \gamma, \beta_c)$.

$$R(\alpha, \gamma, \beta_c) = \int_{-\infty}^{\beta_c} dR(\alpha, \gamma, t)\, dt \tag{B. 1}$$

where

$$dR(\alpha, \gamma, t) = \frac{1}{I_0(\alpha, \gamma)} \exp\left[\frac{-\alpha^2}{2}\,\frac{-\gamma^2}{2}\right] t \exp\left[\frac{-t^2}{2}\right] I_0(\alpha t)\, I_0(\gamma t) \tag{B. 2}$$

Thus

$$dR(\alpha, \gamma, t) = \exp\left\{-\ell n[I_0(\alpha\gamma)] + \ell n\, t + \ell n[I_0(\alpha t)] + \ell n[I_0(\gamma t)] - \frac{(\alpha^2 + \gamma^2 + t^2)}{2}\right\} \tag{B. 3}$$

The modified Bessel function, $I_0(x)$, is approximated to within 1 part in $10^7$ (Ref. 24) by

$$I_0(x) = \frac{\exp(x)}{x}\left\{\frac{1}{2\pi} + (0.013285917)\,\frac{3.75}{x} + (0.00225387)\,\frac{3.75}{x}^2 + \ldots\right\}$$

$$= \frac{\exp(x)}{2\pi x}\left\{1 + \frac{0.12488}{x} + \frac{0.079425}{x} + \ldots\right\} \tag{B. 4}$$

Using the approximations in Eq. B. 4 for the $I_0$ function in Eq. B. 3, we obtain the following

$$dR(\alpha, \gamma, t) = \exp\left\{-\alpha\gamma + \frac{1}{2}\,\ell n(\alpha\gamma) + \ell n\sqrt{2\pi} + \alpha t - \frac{1}{2}\,\ell n(\alpha t) - \ell n\sqrt{2\pi}\right.$$

$$\left. = \ell n\, t + \gamma t - \frac{1}{2}\,\ell n(\gamma t) - \ell n\sqrt{2\pi} - \frac{(\alpha^2 + \gamma^2 + t^2)}{2} + \ldots\right\} \tag{B. 5}$$

If we assume $\alpha$, $\gamma$, and t are large, then the terms in the expansion not evaluated are approximately zero. Simplifying Eq. B. 5, under the above assumption, we obtain

$$dR(\alpha, \gamma, t) \cong \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\,(t - \alpha + \gamma)^2\right] \tag{B. 6}$$

Thus, the approximation for $\alpha$, $\gamma$, and t large is

$$dR(\alpha, \gamma, t) \cong \phi[t - (\alpha + \gamma)] \tag{B.7}$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right]$$

Thus

$$R(\alpha, \gamma, \beta_c) \cong \Phi[\beta_c - (\alpha + \gamma)] \tag{B.8}$$

Equation B. 9 is a good approximation whenever $\alpha\beta_c$, $\gamma\beta_c$, and $\alpha\gamma$ are simultaneously greater than 5.

If $\alpha\beta_c$ and $\gamma\beta_c$ are simultaneously greater than 5, then by using the approximation to the $I_0$ function given in Eq. B. 4 for the $\alpha\beta_c$ and $\gamma\beta_c$ arguments, we obtain the approximation

$$R(\alpha, \gamma, \beta_c) \cong \Phi[\beta_c - (\alpha + \gamma)] \frac{\exp(\alpha\gamma)}{\sqrt{2\pi\,\alpha\gamma}\, I_0(\alpha\gamma)} \tag{B.9}$$

As an example of the approximations given above, consider the evaluation of $R(1, 3, \beta_c)$ where $\beta_c$ is allowed to vary from zero to infinity. The approximation is

$$R(1, 3, \beta_c) \cong \int_{\beta_c}^{5} \exp\left[-\frac{1}{2} - \frac{a}{2}\right] I_0(t)\, I_0(3t)\, t \exp\left[\frac{-t^2}{2}\right] + \frac{\Phi(1)\, e^3}{\sqrt{6\pi}\, I_0(3)}, \quad \beta_c < 5 \tag{B.10}$$

$$\cong \frac{\Phi(\beta_c - 4)\, e^3}{\sqrt{6\pi}\, I_0(3)}, \quad \beta_c > 5 \tag{B.11}$$

The approximations to the Q function are similar to the same as the approximations used in the R functions, since we can write the Q function as

$$Q(\alpha, \beta_c) = R(\alpha, 0, \beta_c) \tag{B.12}$$

The Q and R functions were evaluated numerically by use of an IBM 7090. The computer programs used in all numerical calculations in this work were written in an algebraic source language developed at The University of Michigan called the Michigan Algorithm Decoder (MAD). We present flow diagrams of the important programs and omit the complete MAD listings because of their great bulk and questionable worth. Interested readers can obtain the binary decks for all programs used in this work by contacting the author. The Q and R programs are merely an evaluation of the formulas given above, and because the programs are straightforward, we have omitted them.

## APPENDIX C

## GENERAL OPTIMUM NONSEQUENTIAL SUBROUTINE

To optimize the fixed-observation procedure for the signal known except for phase (SKEP) problem, a generalized computer subroutine was developed. This system otpimizes any fixed time procedure in which the detection and false alarm probabilities are known as a function of the initial log-odds ratio, L, and the quality of observation, d. The operation and use of this system is described below.

Optimization of a fixed-observation procedure requires an analysis of the value of the observation over the L - $\sqrt{d}$ plane. A contour plot of this surface yields two useful boundaries for the decision process. When the value for a given L and $\sqrt{d}$ is greater than zero, observation yields smaller expected losses than an immediate terminal decision. Thus, the value equal zero contour bounds the region where an observation should be taken. The optimum $\sqrt{d}$ for a given initial L corresponds to a peak in the value surface on a cross section along that L. A plot of these optimum $\sqrt{d}$'s as a function of the initial L describes a ridge line on the value surface. The subroutines GONS1., GONS2. and GONS3. provide for a systematic search of the value surface and locate both boundaries.

Figure C. 1 shows a flow diagram for a typical main program using the GONS routines. For discussion purposes, it will be assumed that the main program, like the subroutines themselves, is written in the MAD language and that the reader is familiar with that language. Intermediate print outs of the probabilities, risks, etc., have been omitted from the flow diagram, but are highly recommended in practice. The three subroutines of the system are described in the order of their use in the following paragraphs.

The GONS1. routine sets up the GONS system by establishing a grid in the L - $\sqrt{d}$ plane and by reading in the W/C ratio. The value of the observation of each point of this grid will be computed by the second routine. Decision boundaries are then determined by linear interpolation from these points by the same routine. Current storage assignments in the routines limit the size of this grid to a maximum of 100 x 100 points. The following call for GONS1. produces the grid shown in Fig. C. 2: GONS1. (DP, L, DPSTRT, DPSTEP, DPSTOP, LMIN, LSTEP, LMAX, W2C).

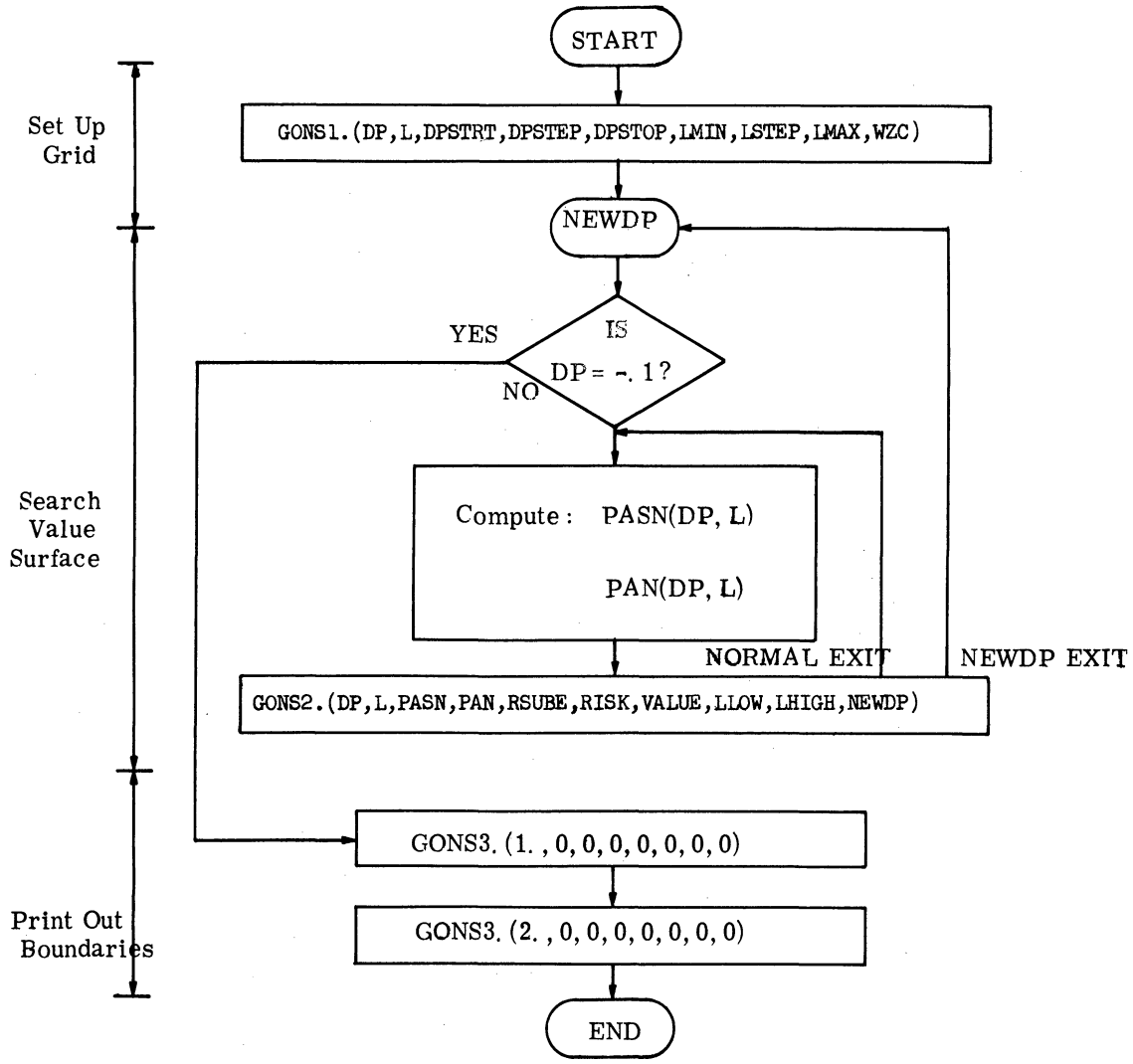


Fig. C. 1. Flow diagram for typical main program.
(Intermediate print outs have been deleted. )

Once the grid on the L - $\sqrt{d}$ plane has been defined, the main program begins searching the value surface under the control of the GONS2. subroutine. The main program computes the probabilities P(A|SN) and P(A|N) for the L - $\sqrt{d}$ coordinates given it by either GONS1. or GONS2. and gives these to GONS2. From the probabilities, GONS2. computes the value, risk, and risk due to error and returns them to the main program. GONS2. also returns the coordinates of a new point in the L - $\sqrt{d}$ plane for which the main program will computer the next pair of probabilities. The search continues on a line of constant $\sqrt{d}$, as indicated in Fig. C. 2, until the entire line is transversed, then the routine transfers to the

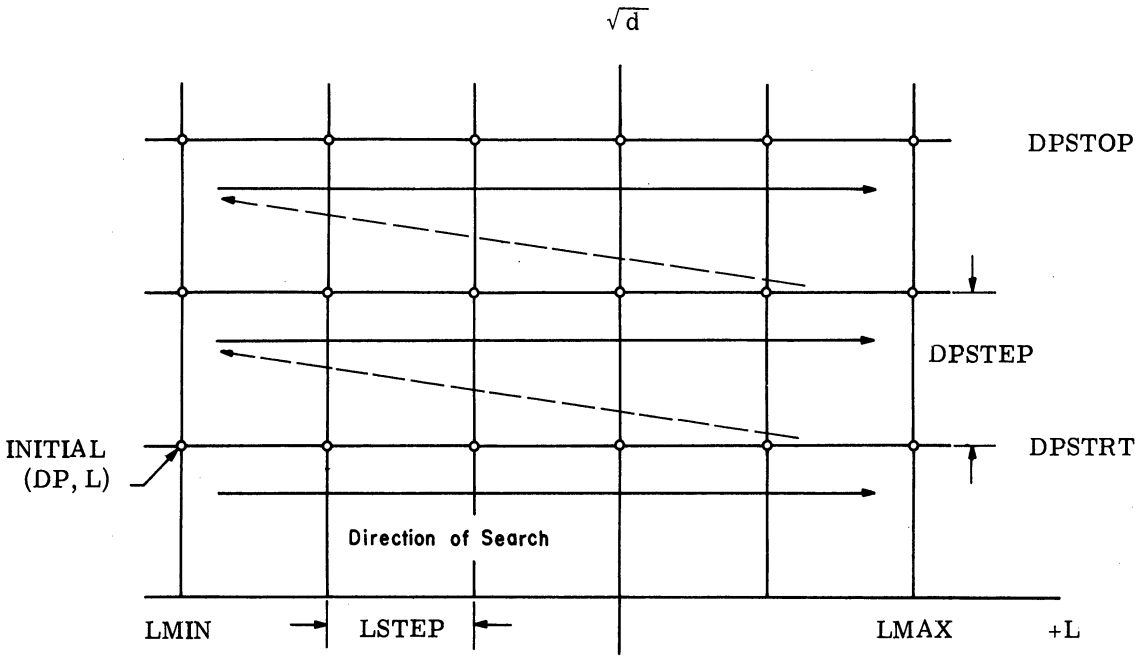


Fig. C.2. Grid produced by call of GONS1. routine.

location specified in its last argument. At this location the main program may print out the observation region boundaries, then it continues to provide GONS2. with the probabilities as before. A typical call for GONS21 might be:

GONS2. (DP, L, PASN, PAN, RERROR, RISK, VALUE, LLOW, LHIGH, NEWDP)

where:

DP, L — are the coordinates of the next point for which the probabilities are to be found. These are returned by GONS2.

PASN, PAN — are the detection and false alarm probabilities provided to GONS2. by the main program.

RSUBE, RISK

VALUE — are the risk due to error, risk and value, respectively computed by GONS2.

LLOW, LHIGH — are the boundaries of the observation region provided by GONS2. after the value surface has been searched along a line of constant $\sqrt{d}$.

NEWDP is the location to transfer to after completing a search across a line of constant $\sqrt{d}$.

When the search of the value surface is complete, the routine GONS2. sets the value of $\sqrt{d}$ to -1. This is recognized by the main program as a flag to stop the search process and print out the boundaries. At this time the GONS routines have both boundaries stored within them and the GONS3. routine is used to print them out.

The call to print out the observation region boundaries as a function of $\sqrt{d}$ would be:

GONS3. (1. 0. , 0. , 0. , 0. , 0. , 0)

To obtain the optimum $\sqrt{d}$ as a function of L, the call would be:

GONS3. (2. , 0. , 0. , 0. , 0. , 0. , 0. )

Both calls print the variables out on $S20,2F30.6*$ format.

## APPENDIX D

## A PROOF ON THE INTERCHANGE OF AN AVERAGING PROCESS

In this appendix we present a proof which allows one to interchange the manner in which the averaging process used in the calculation of the probability of detection is carried out. (See Section 4.3.3.)

Theorem: If p(a) and g(x) are continuous probability density functions, then

$$\int_{-\infty}^{\infty} P(x > a|a)\, p(a)\, da = \int_{-\infty}^{\infty} P(a < x|x)\, g(x)\, dx \qquad (D.1)$$

Proof:

$$P(x > a|a) = 1. - \int_{-\infty}^{a} g(x)\, dx \quad \text{by definition}$$

Also

$$P(a < x|x) = \int_{-\infty}^{a} p(a)\, da$$

Thus, it is equivalent to show

$$\int_{-\infty}^{\infty} \left(1. - \int_{-\infty}^{a} g(x)\, dx\right) p(a)\, da \stackrel{?}{=} \int_{-a}^{\infty} \int_{-\infty}^{x} p(a)\, da\, dx \qquad (D.2)$$

The above can be written

$$\int_{-\infty}^{\infty} g(x) \left( \int_{-\infty}^{x} p(a)\, da \right) dx \stackrel{?}{=} 1. - \int_{-a}^{\infty} p(a) \left( \int_{-\infty}^{a} g(x)\, dx \right) da \qquad (D.3)$$

Rewriting the above, we obtain

$$\int_{-\infty}^{\infty} \int_{-\infty}^{x} h(x, a)\, da\, dx \stackrel{?}{=} 1. - \int_{-\infty}^{\infty} \int_{-\infty}^{a} h(x, a)\, dx\, da \qquad (D.4)$$

where h(x, a) = g(x) p(a) and can be considered as the joint probability density function of x and a. If we consider the mass density in the (x, a) plane, the term on the left-hand side of Eq. D.4 is the total mass below the line x = a in the (x, a) plane. The double integral on the right-hand side is the mass above the line x = a in the (x, a) plane. Thus, since the total mass in the (x, a) plane is one, Eq. D.4 is an equality.

Q.E.D.

APPENDIX E

## DERIVATIONS OF THE PROBABILITY OF FALSE ALARM AND DETECTION FOR THE GAUSSIAN AMPLITUDE AND $\alpha = 1$ AMPLITUDE DISTRIBUTIONS

This appendix includes the derivations of the probability of detection and false alarm for the Gaussian amplitude distribution and the amplitude distribution corresponding to $\alpha = 1$. The derivations are similar to the derivation presented in Section 4.3.3 for the truncated Gaussian amplitude distribution.

### Gaussian Amplitude Distribution

Consider the calculation of the probability of detection and the probability of false alarm for the amplitude distributed according to a Gaussian distribution as defined by Eq. E.1.

$$f(a) = \sqrt{\frac{g}{2\pi}} \exp\left[-\frac{g}{2}(a - h/g)^2\right] = \phi(a\sqrt{g} - h/g), \quad -\infty < a < \infty \qquad \text{(E.1)}$$

After a single observation the distribution is given by Eq. E.2.

$$f(a|z) = \frac{\ell(z|a)}{\ell(z)} f(a) \qquad \text{(E.2)}$$

Subsituting in Eq. E.2, we obtain

$$f(a\ z) = \frac{\exp\left[za - a^2/2\right]\frac{\sqrt{g}}{\sqrt{2\pi}} \exp\left[-\frac{g}{2}(a - h/g)^2\right]}{\int_{-\infty}^{\infty} \exp\left[za - a^2/2\right]\frac{\sqrt{g}}{\sqrt{2\pi}} \exp\left[-\frac{g}{2}(a - h/g)^2\right] da} \qquad \text{(E.3)}$$

The integral in the denominator of Eq. E.3 can be evaluated by comgining the exponential terms and completing the square. This results in Eq. E.4 below.

$$f(a|z) = \sqrt{\frac{g+1}{2\pi}} \exp\left[-\frac{g+1}{2}\left(a - \frac{h+z}{g+1}\right)^2\right]$$
$$= \phi\left(a\sqrt{g+1} - \frac{h+z}{\sqrt{g+1}}\right) \qquad \text{(E.4)}$$

Thus, from Eq. E.4, we see that under normal observations h is transformed into h + z, a random variable, and g is transformed into g + 1, a determinsitic variable.

We can now obtain the total log-odds ratio after one or more observations using the results of the updating process in the amplitude distribution. The total log-odds ratio after an observation is given by

$$L \text{ (based on z)} = L_0 + \ell n\, \ell(z) \tag{E.5}$$

If the amplitude distribution is given by Eq. E. 1, then the total log-odds ratio is

$$\begin{aligned} L_1 &= L_0 + \ell n \sqrt{\frac{g_0}{2\pi}} \int_{-\infty}^{\infty} \exp\left[az - a^2/2\right] \exp\left[-\frac{g_0}{2}(a - h_0/g_0)^2\right] da \\ &= L_0 + \ell n \sqrt{\frac{g_0}{2\pi}} + \ell n \left\{ \int_{-\infty}^{\infty} \exp\left[az - a^2/2\right] \exp\left[-\frac{g_0}{2}(a - h_0/g_0)^2\right] da \right\} \\ &= L_0 + \ell n \sqrt{g_0} - \ell n \sqrt{g_0 + 1} - \frac{h_0^2}{2g_0} + \frac{(h_0 + z)^2}{2(g_0 + 1)} \\ &= L_0 + \ell n \sqrt{\frac{g_0}{g_1}} - \frac{h_0^2}{2g_0} + \frac{h_1^2}{2g_1} \end{aligned} \tag{E.6}$$

where $h_1 = h_0 + z$ and $g_1 = g_0 + 1$.

Generalizing the results of Eq. E. 6, the log-odds ratio after k observations is given by

$$L_k = L_0 + \ell n \sqrt{\frac{g_0}{g_k}} - \frac{h_0^2}{2g_0} + \frac{h_k^2}{2g_k} \tag{E.7}$$

where

$$h_k = h_0 + \sum_{i=1}^{k} z_i \quad \text{and} \quad g_k = g_0 + k \tag{E.8}$$

The probability of false alarm is the probability of responding "A" in noise alone, and this response is appropriate whenever the log-odds ratio, at the time of a terminal decision, is greater than the threshold, $\Delta_0$. In symbols,

$$\begin{aligned} P(\text{"A"}|N) &= P(L > \Delta_0|N) \\ &= P\left\{ L_0 + \ell n \sqrt{\frac{g_0}{g_k}} - \frac{h_0^2}{2g_0} + \frac{h_k^2}{2g_k} > \Delta_0 \middle| N \right\} \\ &= P\left\{ h_k > \left[ 2g_k \left( \Delta_0 - L_0 + \frac{h_0^2}{2g_k} + \ell n \frac{g_k}{g_0} \right) \right]^{1/2} \middle| N \right\} \end{aligned} \tag{E.9}$$

Equation E. 9 can be written as

$$P(\text{"A"}|N) = P\left(h_k > h^*|N\right) \tag{E. 10}$$

where h* is the solution to the equation

$$h^* = \left[2g_k\left(\Delta_0 - L_0 + \frac{h_0^2}{2g_0} + \ell n\sqrt{\frac{g_k}{g_0}}\right)\right]^{1/2} \tag{E. 11}$$

From Eq. E. 8, we have that $(h_k - h_0) = \sum_{i=1}^{k} z_i$ and so $(h_k - h_0)$ is distributed N(0, k). The probability of false alarm is, therefore, given by

$$\begin{aligned} P(\text{"A" } N) &= P(h_k - h_0 > h^* - h_0|N) \\ &= \Phi\left(\frac{h_0 - h^*}{\sqrt{k}}\right) \end{aligned} \tag{E. 12}$$

The corresponding probability of detection is the probability that the total log-odds ratio is greater than the threshold, $\Delta_0$, i. e. ,

$$\begin{aligned} P(\text{"A"}|SN) &= P(L_k > \Delta_0|SN) \\ &= P(h_k > h^*|SN) \end{aligned} \tag{E. 13}$$

where h* is the solution to Eq. E. 11. Equation E. 13 can be written

$$P(\text{"A"}|SN) = P(h_k - h_0 > h^* - h_0|SN) \tag{E. 14}$$

The quantity $(h_k - h_0) = \sum_{i=1}^{k} z_i$ is distributed, if the amplitude is known, according to a Gaussian distribution with mean equal to ka and variance equal to k. Thus, given the amplitude, the probability of detection is

$$P(\text{"A"}|SN, a) = P\left(\frac{h_k - h_0}{\sqrt{k}} > \frac{h^* - h_0}{\sqrt{k}}\Big|SN, a\right) \tag{E. 15}$$

We define a new variable $\xi = \frac{h_k - h_0}{\sqrt{k}} - a\sqrt{k}$ which is distributed N(0, 1). Equations E. 15 can be written in terms of $\xi$ as

$$P(\text{"A"}|SN, a) = P\left(a > \frac{h^* - h_0}{\sqrt{k}} - \frac{\xi}{\sqrt{k}}\Big|SN, \xi\right) \tag{E. 16}$$

Thus, the probability of detection is given by

$$P(\text{"A"}|SN) = \int_{-\infty}^{\infty} P(\text{"A"}|SN, \xi)\, d\Phi(\xi)$$

$$= \int_{-\infty}^{\infty} \Phi\left( \frac{\frac{h_0}{g_0} + \frac{\xi}{\sqrt{k}} + \frac{h_0 - h^*}{k}}{1/\sqrt{g_0}} \right) d\Phi(\xi) \qquad \text{(E. 17)}$$

The evaluation of Eq. E. 17 was accomplished by means of numerical integration using the general program of Appendix C with the appropriate subroutines. The specific evaluation of the integral in Eq. E. 17 was accomplished using Stieltjes integration, by approximating the normal distribution, $d\Phi(\xi)$, with 50 values chosen to represent two percent probability. Numerically, Eq. E. 17 is evaluated according to

$$\int_{-\infty}^{\infty} \Phi\left( \frac{\frac{h_0}{g_0} + \frac{\xi}{\sqrt{k}} + \frac{h_0 - h^*}{k}}{1/\sqrt{g_0}} \right) d\Phi(\xi) \cong 0.02 \sum_{i=1}^{50} \Phi\left( \frac{\frac{h_0}{g_0} + \frac{A_i}{\sqrt{k}} + \frac{h_0 - h^*}{k}}{1/\sqrt{g_0}} \right) \qquad \text{(E. 18)}$$

Amplitude Distribution for $\alpha = 1$

The calculation of the probability of detection and false alarm for the $\alpha = 1$ amplitude distribution parallels the derivation of the Gaussian amplitude distribution. We will not present the detail of the derivation given above.

The amplitude distribution for the $\alpha = 1$ case is as defined in Eq. E. 19.

$$f(a) = \frac{g}{1 + \frac{h}{g}\,\omega\left(\frac{h}{g}\right)}\, a \exp[ah - ag^2/2], \qquad a \geq 0$$

$$= 0, \qquad \text{otherwise} \qquad \text{(E. 19)}$$

After a single observation, the updated amplitude distribution is

$$f(a|z) = \frac{g + 1}{1 + \frac{h + z}{\sqrt{g + 1}}\,\omega\left(\frac{h + z}{\sqrt{g + 1}}\right)}\, a \exp[a(h + z) - a^2(g + 1)/2], \qquad a \geq 0$$

$$= 0, \qquad \text{otherwise} \qquad \text{(E. 20)}$$

A normal observation transforms the parameter h into the random variable h + z and transforms g into g + 1. For k normal observations, h is transformed into $h + \sum_{i=1}^{k} z_i$ and g is transformed into g + k.

The total log-odds ratio after k observations is obtained from Eq. E.21.

$$L_k = L_0 + \ell n\, \ell(z_1, z_2, \ldots, z_k) \tag{E.21}$$

In a manner analogous to the derivation of the Gaussian amplitude distribution, we obtain

$$L_k = L_0 + \ell n \left[ \frac{g_0}{1 + \frac{h_0}{\sqrt{g_0}}\, \omega\left(\frac{h_0}{\sqrt{g_0}}\right)} \right] + \ell n \left[ \frac{1 + \frac{h_k}{\sqrt{g_k}}\, \omega\left(\frac{h_k}{\sqrt{g_k}}\right)}{g_k} \right] \tag{E.22}$$

where

$$h_k = h_0 + \sum_{i=1}^{k} z_i \quad \text{and} \quad g_k = g_0 + k \tag{E.23}$$

The probability of false alarm is equal to the probability that $L_k$ is greater than $\Delta_0$ in noise alone. In terms of the random variable $h_k$, the probability of false alarm is the probability that $h_k$ is greater than h*, where h* is the solution to the equation

$$\frac{h^*}{\sqrt{g_k}}\, \omega\left(\frac{h^*}{\sqrt{g_k}}\right) = g_k \left\{ \exp\,(\Delta_0 - L_0) + 1 + \frac{h_0}{\sqrt{g_0}}\, \omega\left(\frac{h_0}{\sqrt{g_0}}\right) \right\} + 1 \tag{E.24}$$

The quantity $(h_k - h_0) = \sum_{i=1}^{k} z_i$ in noise alone is N(0, k). Therefore, we can write the probability of false alarm as

$$\begin{aligned} P(\text{"A"} \,|\, N) &= P(h_k - h_0 > h^* - h_0 \,|\, N) \\ &= \Phi\left(\frac{h_0 - h^*}{\sqrt{k}}\right) \end{aligned} \tag{E.25}$$

The corresponding probability of detection is obtained in terms of the random variable h in a manner analogous to the derivations used in the Gaussian amplitude distribution and the truncated Gaussian distribution. The probability of detection, if the amplitude is known, can be written

$$P(\text{"A"}|\,SN,\ a) = P\left(a > \frac{h^* - h_0}{k} - \frac{\xi}{\sqrt{k}}\,\middle|\,SN,\ \xi\right) \tag{E.26}$$

where $\xi = (h_k - h_0)/\sqrt{k} - a\sqrt{k}$ is distributed N(0, 1) and h* is given by Eq. E.24. Thus, the probability of detection is given by averaging Eq. E.26 with respect to the normal distribution with mean zero and unit variance. The probability that "a" is greater than some threshold value, in this case the threshold is equal to $(h^* - h_0)/k - \xi\sqrt{k}$, is given by the cumulative probability distribution function of "a".

The cumulative distribution function of "a" is defined as probability that "a" is less than some cut value. (We are interested in one minus the distribution function.) In symbols, the cumulative distribution function is

$$\begin{aligned} P(a > x|SN,\ \xi) &= 1 - \int_0^x f(a)\,da \\ &= 1 - \int_0^x \frac{ag}{\phi(h/\sqrt{g}) + \frac{h}{\sqrt{g}}\,\Phi(h/\sqrt{g})}\,\phi\left(\frac{a - h/g}{1/\sqrt{g}}\right)da \end{aligned} \tag{E.27}$$

The evaluation of the integral in Eq. E.27 can be obtained in terms of the normal distribution function and the normal density function as

$$\begin{aligned} P(a > x|SN,\ \xi) = 1 - \frac{g}{\phi(h/\sqrt{g}) + \frac{h}{\sqrt{g}}\,\Phi(h/\sqrt{g})} \Bigg\{ &- \frac{1}{g}\left[\phi\left(\frac{x - h/g}{1/\sqrt{g}}\right) - \phi\left(\frac{h}{\sqrt{g}}\right)\right] \\ &+ \frac{h}{g}\left[\Phi\left(\frac{x - h/g}{1/\sqrt{g}}\right) - 0.5\right]\Bigg\} \end{aligned} \tag{E.28}$$

Using Eq. E.28, the probability of detection can be found from

$$P(\text{"A"}|\,SN) = \int_{-\infty}^{\infty} \Phi\left(a > \frac{h^* - h_0}{k} - \frac{\xi}{\sqrt{k}}\,\middle|\,SN,\ \xi\right) f(a)\,da \tag{E.29}$$

The numerical evaluation of Eq. E.29 was obtained by use of an IBM 7090. The computer program used is given in Appendix C.

APPENDIX F

## THE COMPUTER FLOW DIAGRAM OF THE SUBROUTINE FOR THE PROBABILITY OF DETECTION AND FALSE ALARM FOR A SKEA

This appendix presents the computer flow diagram of the subroutines for the probability of detection and false alarm for a single observation of a signal of unknown amplitude. The program shown in Fig. F. 1 is an external function used in conjunction with the general calling program presented in Appendix C. The external function contains two entry points; one for the detection probability, and the other for the false alarm probability. The amplitude distribution is assumed to be truncated Gaussian, i. e., $\alpha = 0$.

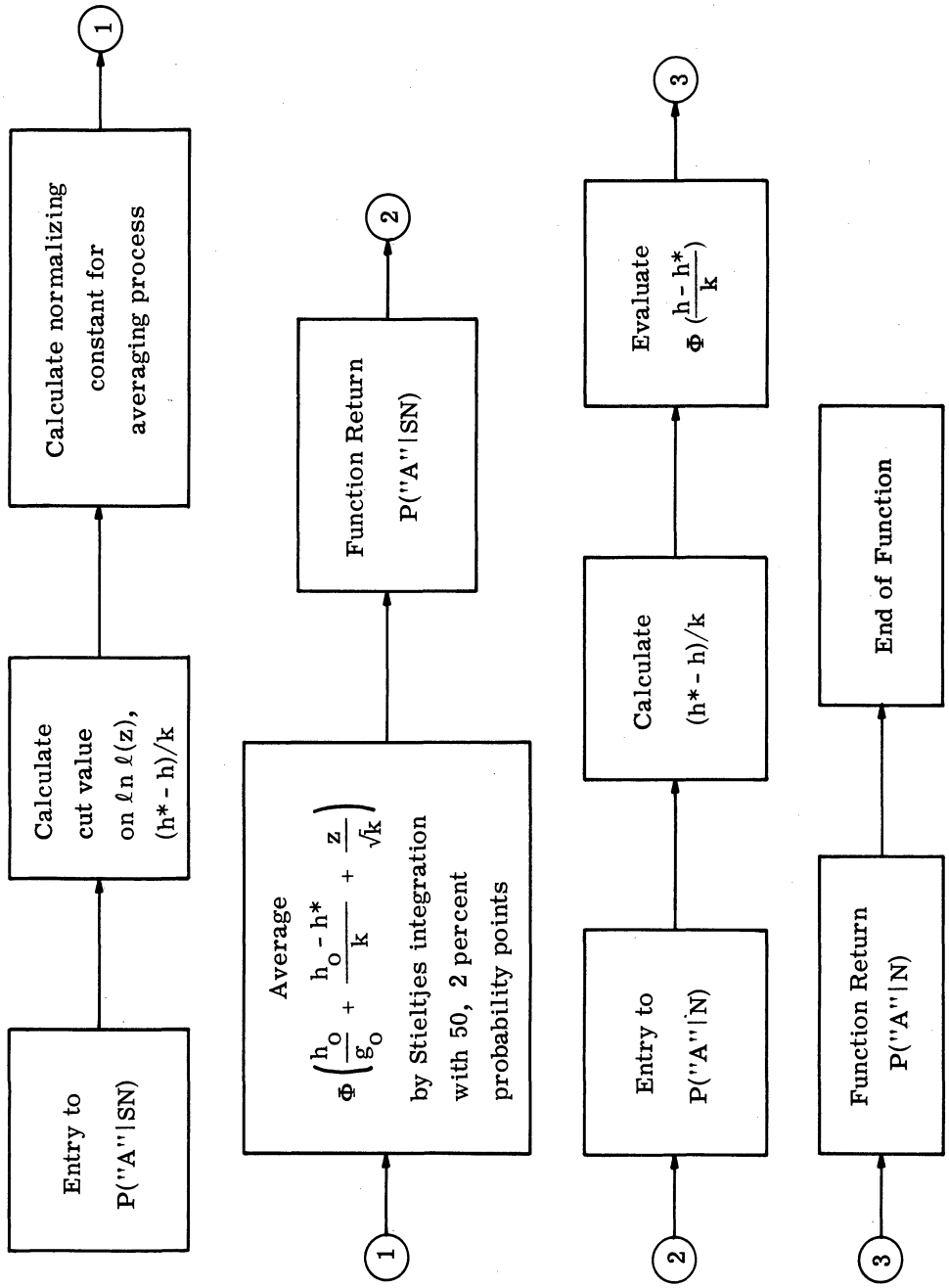


Fig. F. 1. Subroutine for P("A"|SN) and P("A"|N) for SKEA.

APPENDIX G

## DERIVATIONS OF THE SEPARABLE SUBSPACES FOR THE GAUSSIAN AND $\alpha = 1$ AMPLITUDE DISTRIBUTIONS

The determination of separable subspaces for the Gaussian amplitude and the $\alpha = 1$ amplitude distribution are obtained easily by using the general equality of Eq. 4.67 repeated below.

$$f_{n-1}(a) = \exp[L_n - L_{n-1}]\ \ell(z|a)\, f_n(a) \tag{4.67}$$

The subscripts n and n-1 refer to post and prior observation, respectively.

### The Gaussian Amplitude Distribution

For the Gaussian amplitude distribution, the substitution of the expression for the a posteriori and a priori density function in Eq. 4.67 results in

$$\sqrt{\frac{g_{n-1}}{2\pi}}\ \exp\left[-\frac{g_{n-1}}{2}\left(a - \frac{h_{n-1}}{g_{n-1}}\right)^2\right] = \exp[L_n - L_{n-1}]\ \exp[az - a^2/2]\sqrt{\frac{g_n}{2\pi}}\ \exp\left[-\frac{g_n}{2\pi}\left(a - \frac{h_n}{g_n}\right)^2\right] \tag{G.1}$$

Expanding Eq. G.1 and rearranging terms, we obtain

$$\sqrt{\frac{g_{n-1}}{2\pi}}\ \exp\left[-a^2 g_{n-1}/2 + a\, h_{n-1} - \frac{h^2_{n-1}}{2g_{n-1}}\right] = \frac{g_n}{2\pi}\ \exp[L_n - L_{n-1}]\ \exp\left[-a^2(g_n+1)/2 + a(h_n + z) - \frac{h_n^2}{2g_n}\right] \tag{G.2}$$

Using the relations $h_{n-1} = h_{n+z}$ and $g_{n-1} = g_n + 1$, we can write Eq. G.2 as

$$\ell n\ \sqrt{g_{n-1}}\ - \frac{h^2_{n-1}}{2g_{n-1}} + L_{n-1} = \ell n\ \sqrt{g_n}\ - \frac{h^2_n}{2g_n} + Ln \tag{G.3}$$

Equation G.3 implies that the quantity of $\ell n\ \sqrt{g} - h^2/2g$ is invariant with observation, and this is the specification of the separable subspace, Q, for the Gaussian amplitude distribution

$$Q = L + \ell n\ \sqrt{g} - h^2/2g \tag{G.4}$$

The Amplitude Distribution for $\alpha = 1$

The separable subspaces for the amplitude distribution corresponding to $\alpha = 1$ are also found by using Eq. 4.67 as a starting point. Substituting the appropriate functions into Eq. 4.67, we obtain the following

$$M_{n-1}\, a \exp[a\, h_{n-1} - a^2 g_{n-1}/2] = \exp[L_n - L_{n-1}] \exp[az - a^2/2]\, M_n\, a \exp[a\, h_n - a^2 g_n/2] \tag{G.5}$$

where $M_{n-1}$ and $M_n$ are the normalizing constants for the amplitude density functions. Expanding Eq. G.5, collecting terms, and taking logarithms of both sides of the equation, we obtain

$$\ell n\, M_n + L_n = \ell n\, M_{n-1} + L_{n-1} \tag{G.6}$$

From Eq. G.6, we conclude that $\ell n\, M + L$ is independent of the stage of observation and this defines the separable subspaces for the $\alpha = 1$ amplitude distribution. Recalling that the normalizing constant M is equal to $g \Big/ \left[1 + \frac{h}{\sqrt{g}}\, \omega(h/\sqrt{g})\right]$, we have for the separable subspaces

$$Q = L + \ell n \left[ \frac{g}{1 + \frac{h}{\sqrt{g}}\, \omega(h/\sqrt{g})} \right] \tag{G.7}$$

# APPENDIX H

## THE COMPUTER FLOW DIAGRAMS FOR THE DESIGN AND EVALUATION OF THE OPTIMUM SKEA DETECTOR FOR THE $\alpha = 0$ AND $\alpha = 1$ AMPLITUDE DISTRIBUTIONS

This appendix presents computer flow block diagrams for calculation of the optimum risk functions, the probability of detection, the probability of false alarm, and the ANO functions for the $\alpha = 0$ and $\alpha = 1$ amplitude distributions. Figure H. 1 presents the flow diagram for calculation of the decision boundary from the optimum risk functions. Figure H. 2 presents the flow diagram for calculation of the ROC and ANO curves for sub-optimum decision boundaries that are read in as data.

The programs in Figs. H. 1 and H. 2 both use a set of external tables occupying approximately a third of the IBM 7090 core memory. The tables are of the $\omega(\ )$, $\Phi(\ )$ and $\phi(\ )$ functions, and their use in the programs permits a more rapid computation time than calculation of these functions from power series (as is usually done).

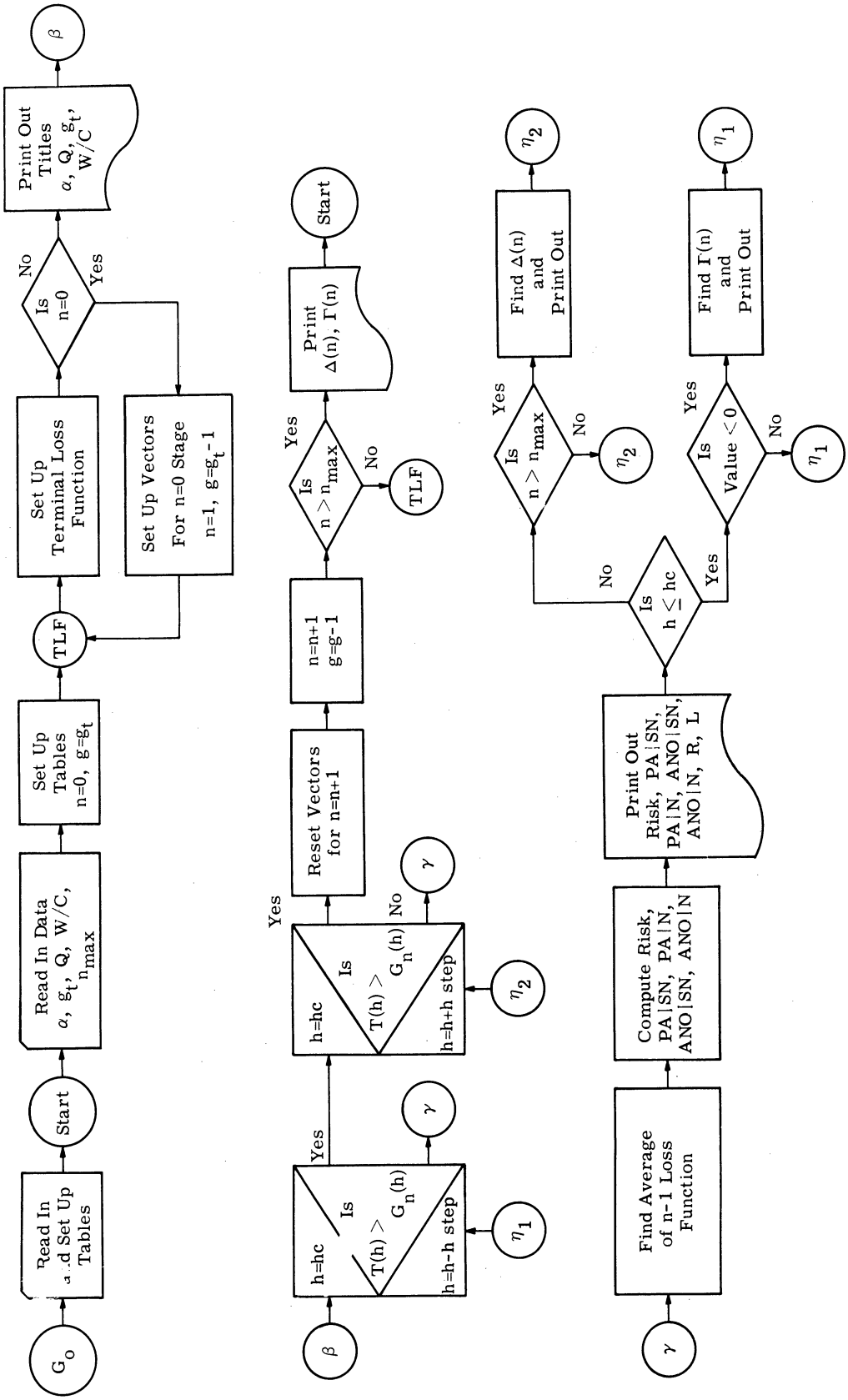


Fig. H. 1. Computer flow diagram for the computation of the decision boundaries for the $\alpha = 0$ and $\alpha = 1$ amplitude distributions for a SKEA.

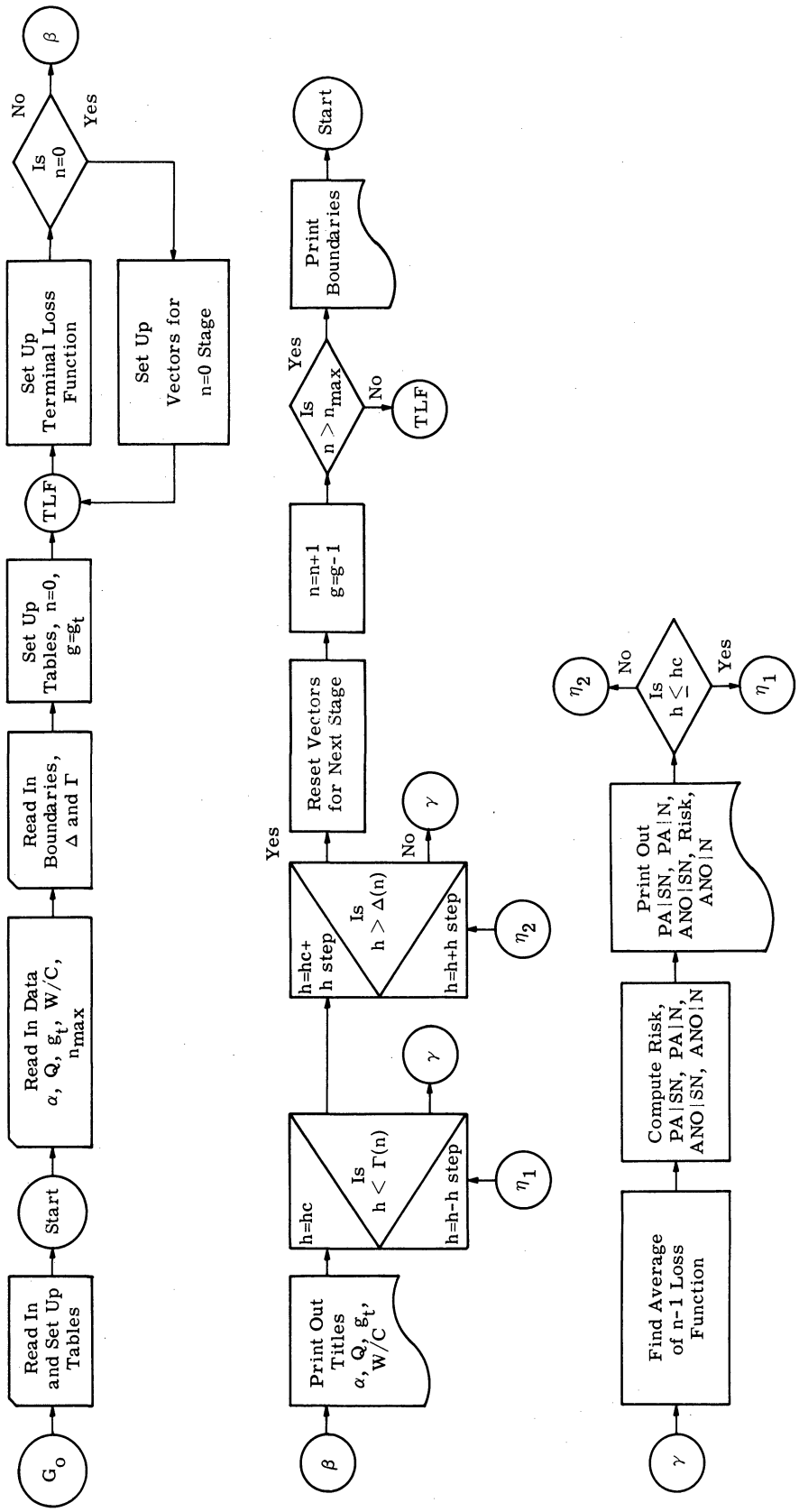


Fig. H. 2. Computer flow diagram for the computation of the ROC and ANO functions for the $\alpha = 0$ and $\alpha = 1$ amplitude distributions for a SKEA.

APPENDIX I

# DERIVATIONS OF THE OPTIMUM RISK FUNCTIONS AND EVALUATION FUNCTIONS FOR THE GAUSSIAN AND $\alpha = 1$ AMPLITUDE DISTRIBUTIONS

This appendix presents the derivations of the optimum average loss functions and the ROC and ANO functions for the Gaussian and $\alpha = 1$ amplitude distributions. The basic method of solution is identical to that presented in Sections 4.5.1 and 4.5.2. The mathematics, however, are distinct from the truncated Gaussian amplitude distribution problem.

## Gaussian Amplitude Distribution

Consider first the derivation of the optimum average loss function. The amplitude distribution is given in Eq. I.1.

$$f(a) = \sqrt{\frac{g}{2\pi}} \exp\left[-\frac{g}{2}(a - h/g)^2\right] = \phi(a\sqrt{g} - h/g), \quad -\infty < a < \infty \tag{I.1}$$

As in the case of truncated amplitude distribution, the state of the decision process must be expressed in terms $\{Q, L, g, g_t - g\}$ in order to effect a one-dimensional averaging process in obtaining the optimum average loss functions. The separable subspaces, Q, are derived in Appendix G. They are given by

$$Q = L + \ell n \sqrt{g} - \frac{h^2}{2g} \tag{I.2}$$

Rather than carry through the iteration steps in detail, as presented in Section 4.5.1, the general iteration step is stated. The same logic used in Section 4.5.1 applies both to this problem and the $\alpha = 1$ amplitude distribution problem.

To obtain the general iteration step for the determination of the optimum average loss function, assume the state of the decision process is $\{Q, L_k, g_k, g_t - g_k = n\}$. Since $g_t - g_k = n$, this imples the available number of deferrals is equal to n. Our problem is to decide whether to take an observation or make an immediate terminal decision. The immediate terminal decision results in an average loss of T(L) while a deferral results in an average

loss of $G_k(L)$, for a log-odds ratio $L = L_k$. The average look-ahead loss, $G_k(L)$, is given by

$$G_k(L) = \int_{-\infty}^{\infty} F_{k-1}[L + \ell n\, \ell(z)]\ f_k(z)\, dz + C \tag{I. 3}$$

where $F_{k-1}(L)$ is the optimum average loss function at stage k - 1, i. e., the stage for the available number of observations one less than stage k. The optimum average loss function for stage k is $F_k(L)$.

$$F_k(L) = \min[G_k(L),\ T(L)] \tag{I. 4}$$

The intersections of $G_k(L)$ and T(L) are, again, the optimum decision boundary points.

The quantities $f_k(z)$ and $\ell n\, \ell(z)$ of Eq. I. 3 are the probability density function and the log likelihood ratio of the observations. For the Gaussian amplitude distribution, we have

$$\begin{aligned} f_k(z) &= \frac{1}{1 + e^{-L}}\ f_k(z|SN) + \frac{1}{1 + e^{L}}\ f_k(z|N) \\ &= \frac{1}{1 + e^{-L}} \int_{-\infty}^{\infty} f(z|SN,\ a)\ f_k(a)\, da + \frac{1}{1 + e^{L}}\ f_k(z|N) \end{aligned} \tag{I. 5}$$

Now the distribution of the observation, z, in noise alone is assumed to be N(0, 1). The distribution of z in signal-and-noise, if we know the amplitude, is N(a, 1). Thus, Eq. I. 5 can be written

$$\begin{aligned} f_k(z) &= \frac{1}{1 + e^{-L}} \int_{-\infty}^{\infty} \phi(z - a)\ \phi\left(a\sqrt{g_k} - \frac{h_k}{g_k}\right) da + \frac{\phi(z)}{1 + e^{L}} \\ &= \frac{1}{1 + e^{-L}}\ \phi\left(a\sqrt{g_k + 1} - \frac{h_k + z}{g_k + 1}\right) + \frac{1}{1 + e^{L}}\ \phi(z) \end{aligned} \tag{I. 6}$$

And finally, we have the updating equation for the log-odds ratio

$$\begin{aligned} L_{k-1} &= L_k + \ell n\, \ell(z) \\ &= Q + \frac{(h_k + z)^2}{2(g_k + 1)} - \frac{1}{2}\, \ell n(g_k + 1) \end{aligned} \tag{I. 7}$$

The updating equation for the probability amplitude density function is given in Appendix E.

The evaluation of the optimum procedure is, again, given in terms of the average loss functions as a function of L. The decomposition of these functions into loss due to terminal decision errors and the loss due observation time is accomplished in terms of the ROC and ANO functions, respectively. The derivations of these functions exactly parallels the derivations presented in Section 4.5.2. The initial ROC curves for one allowable deferral are derived in Appendix E. The solution to these equations were obtained by means of an IBM 7090. The various computer programs are presented in Appendix J.

Amplitude Distribution for $\alpha = 1$

The optimum average loss functions, the ROC and the ANO functions are derived in the same manner as for the Gaussian amplitude distribution. The equations that correspond to Eqs. I.1 through I.7 are presented below. The words that explain the equations are the same as used in the Gaussian amplitude distribution problem given above.

$$f(a) = \frac{g}{1 + \frac{h}{\sqrt{g}}\,\omega\left(\frac{h}{\sqrt{g}}\right)}\; a \exp\left(ah - \frac{a^2 g}{2}\right), \qquad a \geq 0$$

$$= 0 \qquad , \qquad \text{otherwise} \tag{I.8}$$

$$Q = L + \ell n \left[\frac{g}{1 + \frac{h}{\sqrt{g}}\,\omega\left(\frac{h}{\sqrt{g}}\right)}\right] \tag{I.9}$$

$$G_k(L) = \int_{-\infty}^{\infty} F_{k-1}[L + \ell n\, \ell(z)]\, f_k(z)\, dz + C \tag{I.10}$$

$$F_k(L) = \min[G_k(L),\ T(L)] \tag{I.11}$$

$$f_k(z) = \frac{1}{1 + e^{-L}}\, f_k(z|SN) + \frac{1}{1 + e^{L}}\, f_k(z|N)$$

$$= \frac{1}{1 + e^{-L}} \int_{-\infty}^{\infty} f(z|SN,\ a)\, f_k(a)\, da + \frac{1}{1 + e^{L}}\, f_k(z|N)$$

$$= \frac{1}{1 + e^{-L}} \int_{-\infty}^{\infty} \phi(z - a)\, \frac{g_k}{1 + \frac{h_k}{\sqrt{g_k}}\,\omega\left(\frac{h_k}{\sqrt{g_k}}\right)}\, a \cdot \exp\left[ah_k - \frac{a^2}{2} g_k\right] da$$

$$+ \frac{1}{1 + e^{-L}}\, \phi(z)$$

$$= \frac{1}{1+e^{-L}} \cdot \frac{g_k}{g_k+1} \cdot \frac{1+\frac{h_k+z}{\sqrt{g_k+1}}\,\omega\left(\frac{h_k+z}{\sqrt{g_k+1}}\right)}{1+\frac{h_k}{\sqrt{g_k}}\,\omega\left(\frac{h_k}{\sqrt{g_k}}\right)}\,\phi(z)$$

$$+ \frac{1}{1+e^{L}}\,\phi(z) \qquad \text{(I. 12)}$$

$$L_{k-1} = L_k + \ell n\, \ell(z)$$

$$= Q - \ell n(g_k+1) + \ell n\left[1+\frac{h_k+z}{\sqrt{g_k+1}}\,\omega\left(\frac{h_k+z}{\sqrt{g_k+1}}\right)\right] \qquad \text{(I. 13)}$$

The updating equation for the probability density function on the amplitude is given in Appendix E.

The evaluation of the optimum procedure in terms on the ROC curves and ANO functions, again, parallels the derivations presented in Section 4.5.2. The initial ROC curves for one allowable deferral are derived in Appendix E. The numerical solution of the equations comprising the solution to the $\alpha = 1$ amplitude distribution problem was obtained by means of an IBM 7090. The program is the same program used in the truncated amplitude distribution prlblem and is given in Appendix H. This program also gives the ROC and ANO functions for the $\alpha = 1$ amplitude distribution.

APPENDIX J

## THE COMPUTER FLOW DIAGRAMS FOR THE DESIGN AND EVALUATION OF THE OPTIMUM SKEA DETECTOR FOR THE GAUSSIAN AMPLITUDE DISTRIBUTION

This appendix presents the computer flow block diagrams for the design and evaluation of the optimum SKEA receiver for a signal amplitude distributed according to Gaussian distribution. The gross characteristics of the operation and evaluation of this receiver are the same as for the receivers designed for the $\alpha = 0$ and $\alpha = 1$ distributions. The Gaussian distributions differ from the $\alpha = 0$ and $\alpha = 1$ distributions in that negative as well as positive amplitudes may occur.

Figure J. 1 is the computer flow graph for the determination of the optimum decision boundaries. The evaluation of the Gaussian amplitude distribution is accomplished by means of a second program as shown in Fig. J. 2.

The terms "old" and "new" functions correspond to the optimum risk function, $F_n$, before and after the averaging process, respectively. The initial optimum risk function is $T(L) = F_0(L)$, the terminal loss function. The computer program stores vectors, one filled with $F_n$-old and the other filled with $F_n$-new.

A more detailed flow diagram is available from the author upon request along with the binary cards used in the IBM 7090.

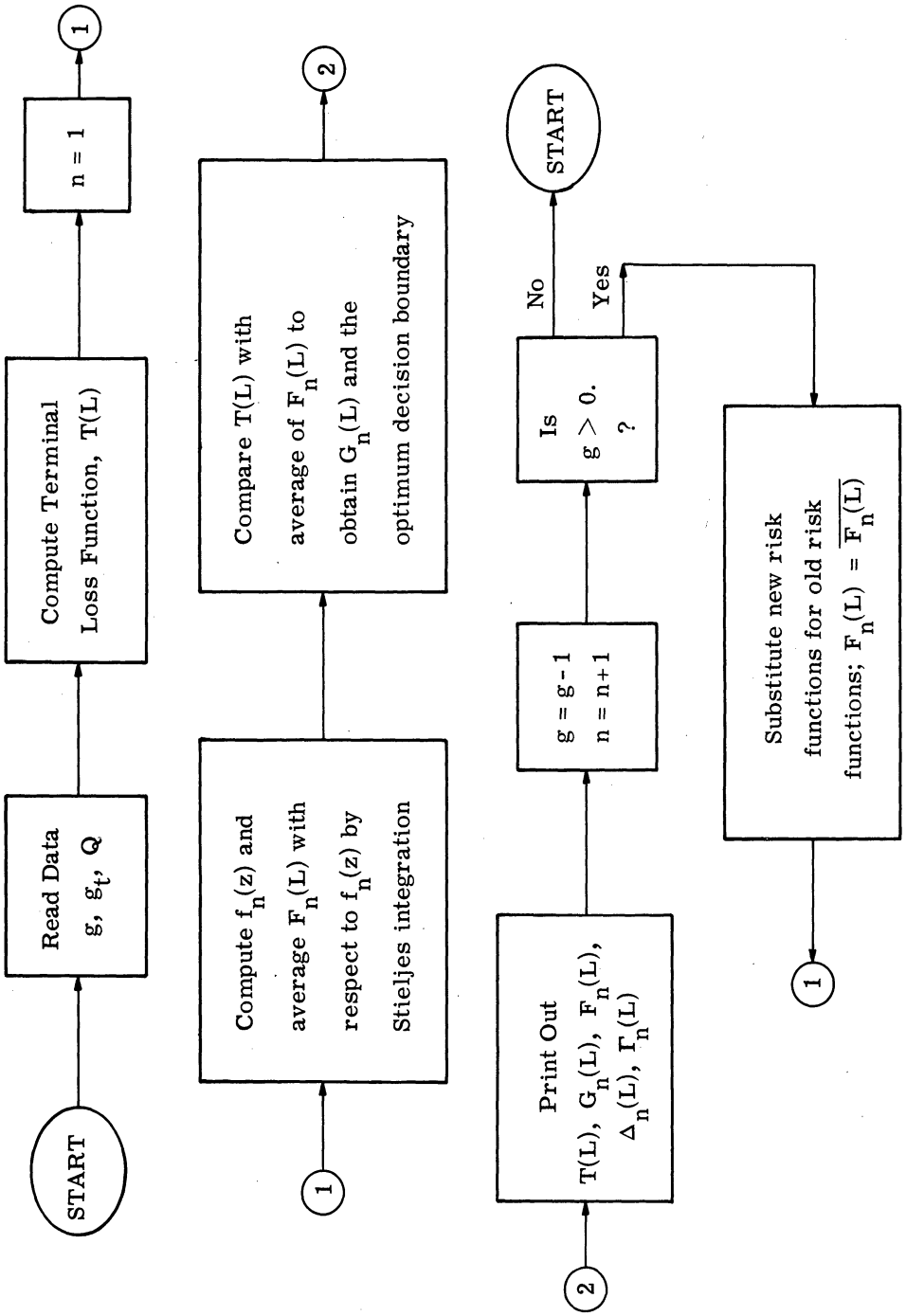


Fig. J. 1. The optimum boundary program for a Gaussian distributed amplitude signal.

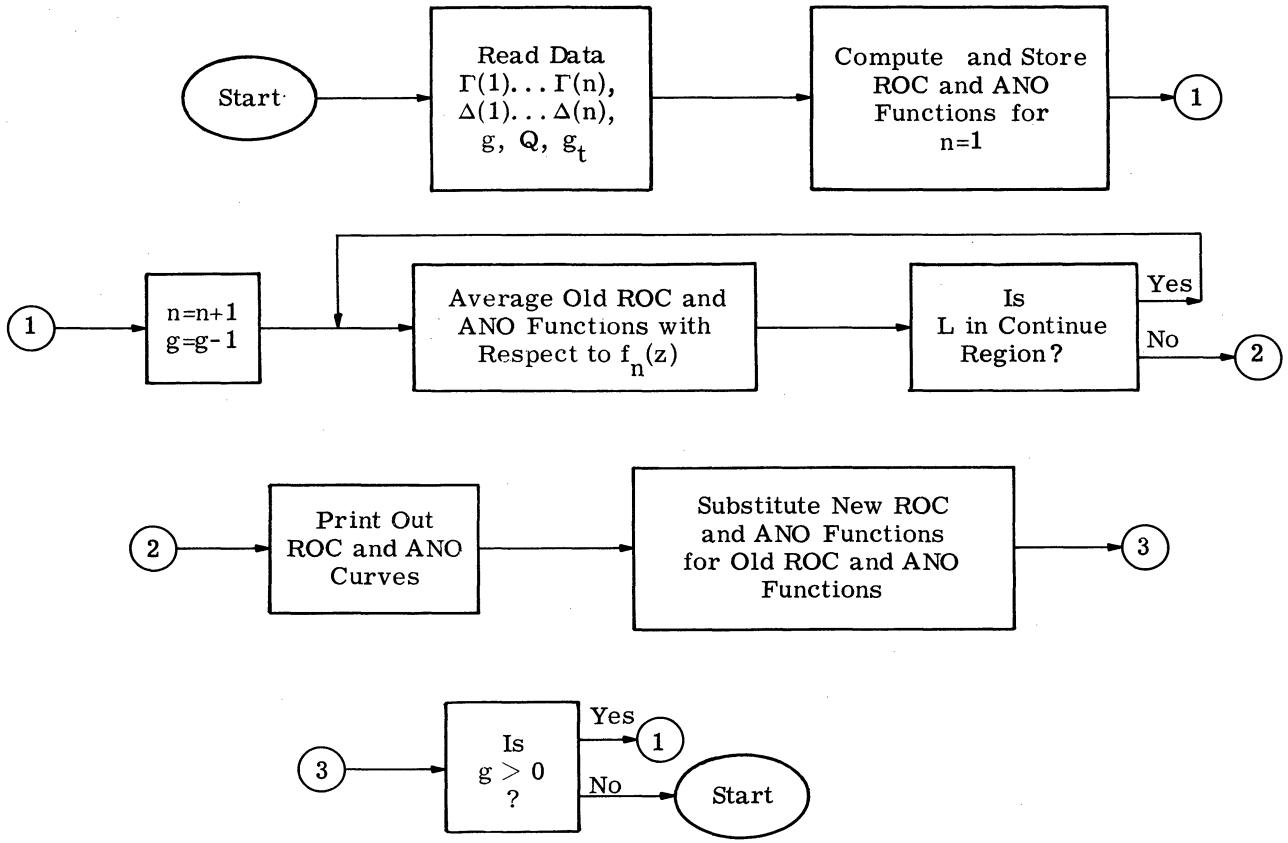


Fig. J. 2. The evaluation program for a Gaussian distributed amplitude signal.

APPENDIX K

## THE COMPUTER FLOW DIAGRAM FOR CALCULATION OF THE CONDITIONAL PERFORMANCE EVALUATORS FOR A SKEA

This appendix presents the computing algorithm, in block diagram form, for the conditional performance curves presented in Section 4.6.3. The program calculates conditional performance in signal-and-noise, in noise alone, and in any mixture of SN and N. The program is written in terms of the parameter h rather than the log-odds ratio, L. The input data needed include $g_t$, g, Q, and the signal amplitude. The program can be thought of as a simulation of the optimum detection receiver for a SKEA in which specific amplitude signals are received at the input. The notation used in the block diagrams is consistent with notation introduced in Chapter IV. The computer block diagram is shown in Fig. K.1.

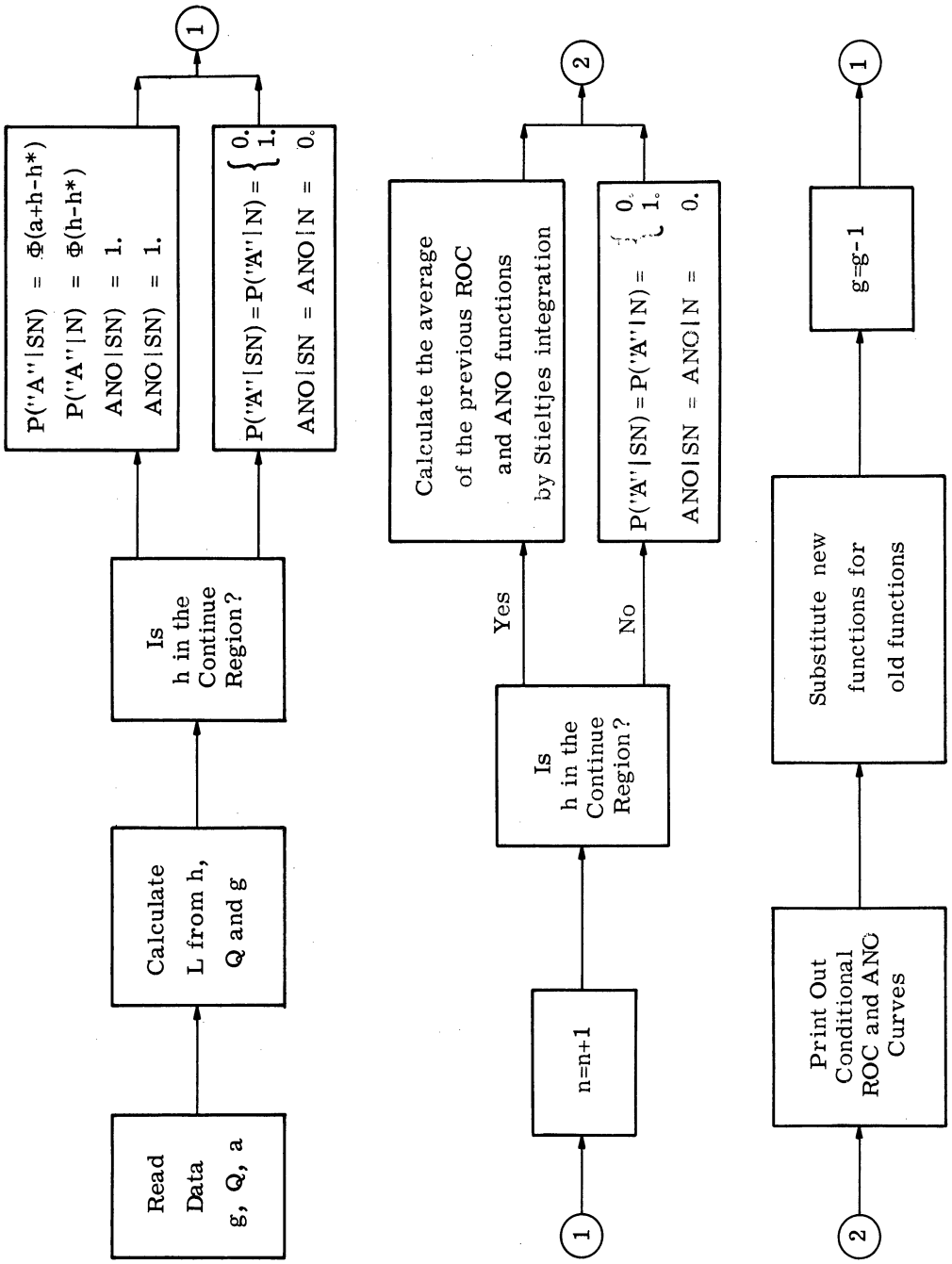


Fig. K. 1. The computer flow diagram for conditional performance curves, ROC and ANO, and for a SKEA.

## REFERENCES


1. U. Grenander, "Stochastic Processes and Statistical Inference," Arkiv det Mat., 1, 1950, pp. 195-277.

2. D. Middleton and D. Van Meter, "Modern Statistical Approaches to Reception in Communication Theory," IRE Trans. on Information Theory, IT-4, 1954.

3. W. W. Peterson, T. G. Birdsall and W. C. Fox, "The Theory of Signal Detectability," IRE Trans. on Information Theory, IT-4, 1954.

4. A. Wald, Sequential Analysis, John Wiley and Sons, Inc., New York, 1947.

5. H. H. Goode, Deferred Decision Theory, Cooley Electronics Laboratory Technical Report No. 123, The University of Michigan, Ann Arbor, Michigan, July 1961.

   This material may also be found in: R. E. Machol and R. Gray, Recent Developments in Information and Decision Processes, New York, The McMillan Company, 1962.

6. T. G. Birdsall and R. A. Roberts, The Theory of Signal Detectability: Observation-Decision Procedures, Cooley Electronics Laboratory Technical Report No. 136, The University of Michigan, Ann Arbor, Michigan, January 1964.

7. L. J. Savage, The Foundations of Statistics, John Wiley and Sons, Inc., New York, 1954.

8. A. Wald, Statistical Decision Functions, John Wiley and Sons, Inc., New York, 1950.

9. T. G. Birdsall, Random Processes, Engineering Summer Conference, The University of Michigan, 1963.

10. D. Blackwell and M. A. Girshick, Theory of Games and Statistical Decisions, John Wiley and Sons, Inc., New York, 1954.

11. V. A. Kotel'nikov, The Theory of Optimum Noise Immunity, Translated from the Russian by R. A. Silverman, McGraw-Hill Book Company, New York, 1959.

12. E. M. Glaser, "Signal Detection by Adaptive Filters," IRE Trans. on Information Theory, April 1961.

13. L. Nolte, "Adaptive Detection and Classification of Signals in Noise," J. of Underwater Acoustics, October 1964.

14. R. Bellman, Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

15. J. D. Spragins, Reproducing Distributions for Machine Learning, Systems Theory Laboratory, Stanford Electronics Laboratory Technical Report No. 6103-7, Stanford University, Stanford, California, November 1963.

16. L. A. Wainstein and V. D. Zubakov, Extraction of Signals from Noise, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1962.

17. P. M. Woodward, Probability and Information Theory, with Applications to Radar, Pergamon Press, London, 1953.

18. C. W. Helstrom, Statistical Theory of Signal Detection, Pergamon Press, New York, 1960.

REFERENCES (Cont.)

19. P. B. Patnaik, "The Noncentral $\chi^2$ and F Distributions and Their Applications," Biometrika, 36, 1949, pp. 202-232.

20. S. O. Rice, "Mathematical Analysis of Random Noise," Bell System Technical Journal, Vols. 23 and 24, 1944-5.

21. J. I. Marcum, Tables of Q Functions, Project Rand Report RM-399 (unpublished report of the Rand Corporation).

22. L. Halstead, T. G. Birdsall and L. N. Nolte, On the Detection of a Randomly-Distorted Signal in Gaussian Noise, Cooley Electronics Laboratory Technical Report No. 129, The University of Michigan, Ann Arbor, Michigan, October 1962.

23. E. D. Rainville, Special Functions, The McMillan Company, New York, 1960.

24. Y. L. Luke, Integrals of Bessel Functions, McGraw-Hill Book Company, New York, 1962.

## DISTRIBUTION LIST

| No. of Copies | |
|---|---|
| 4 | Office of Naval Research (Code 468), Department of the Navy, Washington 25, D. C. |
| 1 | Office of Naval Research (Code 436), Department of the Navy, Washington 25, D. C. |
| 1 | Office of Naval Research (Code 437), Department of the Navy, Washington 25, D. C. |
| 6 | Director, U. S. Naval Research Laboratory, Technical Information Division, Washington 25, D. C. |
| 1 | Director, U. S. Naval Research Laboratory, Sound Division, Washington 25, D. C. |
| 1 | Commanding Officer, Office of Naval Research Branch Office, 230 N. Michigan Avenue, Chicago 1, Illinois |
| 10 | Commanding Officer, Office of Naval Research Branch Office, Box 39, Navy No. 100, FPO, New York, New York |
| 20 | Defense Documentation Center, Cameron Station, Building No. 5, 5010 Duke St., Alexandria 4, Virginia |
| 2 | Commander, U. S. Naval Ordnance Laboratory, Acoustics Division, White Oak, Silver Spring, Maryland |
| 1 | Commanding Officer and Director, U. S. Navy Electronics Laboratory, San Diego 52, California |
| 1 | Director, U. S. Navy Underwater Sound Reference Laboratory, Office of Naval Research, P. O. Box 8337, Orlando, Florida |
| 2 | Commanding Officer and Director, U. S. Navy Underwater Sound Laboratory, Fort Trumbull, New London, Connecticut (Attn: Mr. W. R. Schumacher, Mr. L. T. Einstein) |
| 1 | Commander, U. S. Naval Air Development Center, Johnsville, Pennsylvania |
| 1 | Commanding Officer and Director, David Taylor Model Basin, Washington 7, D. C. |
| 1 | Office of Chief Signal Officer, Department of the Army, Pentagon, Washington 25, D. C. |
| 2 | Superintendent, U. S. Navy Postgraduate School, Monterey, California (Attn: Prof. L. E. Kinsler, Prof. H. Medwin) |

DISTRIBUTION LIST (Cont.)

| No. of Copies | |
|---|---|
| 1 | U. S. Naval Academy, Annapolis, Maryland (Attn: Library) |
| 1 | Harvard University, Acoustics Laboratory, Division of Applied Science, Cambridge 38, Massachusetts |
| 1 | Brown University, Department of Physics, Providence 12, R. I. |
| 1 | Western Reserve University, Department of Chemistry, Cleveland, Ohio (Attn: Dr. E. Yeager) |
| 1 | University of California, Department of Physics, Los Angeles, California |
| 2 | University of California, Marine Physical Laboratory of the Scripps Institution of Oceanography, San Diego 52, California (Attn: Dr. V. C. Anderson, Dr. Philip Rudnick) |
| 1 | Dr. M. J. Jacobson, Department of Mathematics, Rensselaer Polytechnic Institute, Troy, New York |
| 1 | Director, Columbia University, Hudson Laboratories, 145 Palisade Street, Dobbs Ferry, N. Y. |
| 1 | Woods Hole Oceanographic Institution, Woods Hole, Massachusetts |
| 1 | Johns Hopkins University, Department of Electrical Engineering, Johns Hopkins University, Baltimore 18, Maryland (Attn: Dr. W. H. Huggins) |
| 1 | Director, University of Miami, The Marine Laboratory, # 1 Rickenbacker Causeway, Miami 49, Florida (Attn: Dr. J. C. Steinberg) |
| 1 | Litton Industries, Advanced Development Laboratories, 221 Crescent St., Waltham, Massachusetts (Attn: Dr. Albert H. Nuttall) |
| 1 | Institute for Defense Analysis, Communications Research Division, von Neumann Hall, Princeton, New Jersey |
| 1 | Commander, U. S. Naval Ordnance Test Station, Pasadena Annex, 3202 E. Foothill Blvd., Pasadena 8, California |
| 1 | Chief, Bureau of Ships (Code 688), Department of the Navy, Washington 25, D. C. |
| 1 | Chief, Bureau of Naval Weapons (Code RU-222), Department of the Navy, Washington 25, D. C. |
| 1 | Cornell Aeronautical Laboratory, Inc., P. O. Box 235, Buffalo 21, New York (Attn: Dr. J. G. Lawton) |
| 1 | Autonetics, A Division of North American Aviation, Inc., 3370 East Anaheim Road, Anaheim, California (Attn: Dr. N. Schalk) |
| 1 | Mr. Charles J. Loda, Institute for Defense Analyses, 400 Army-Navy Drive, Arlington, Virginia 22202 |

DISTRIBUTION LIST (Cont.)

| No. of Copies | |
|---|---|
| 1 | Catholic University of America, Department of Physics, Washington, D. C. 20017, Attn: Dr. Frank Andrews |
| 1 | Dr. B. F. Barton, Director, Cooley Electronics Laboratory, The University of Michigan, Ann Arbor, Michigan |
| 108 | Cooley Electronics Laboratory, The University of Michigan, Ann Arbor, Michigan |

Unclassified

Security Classification

**DOCUMENT CONTROL DATA - R&D**

*(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)*

| | |
|---|---|
| 1. ORIGINATING ACTIVITY *(Corporate author)*<br>Cooley Electronics Laboratory<br>The University of Michigan<br>Ann Arbor, Michigan | 2a. REPORT SECURITY CLASSIFICATION<br>Unclassified<br>2b. GROUP |

3. REPORT TITLE

Theory of Signal Detectability: Composite Deferred Decision Theory

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Summary, 1963 - 1965

5. AUTHOR(S) *(Last name, first name, initial)*

Roberts, Richard A.

| 6. REPORT DATE | 7a. TOTAL NO. OF PAGES | 7b. NO. OF REFS |
|---|---|---|
| March, 1965 | 217 | 24 |

| | |
|---|---|
| 8a. CONTRACT OR GRANT NO.<br>Nonr 1224(36)<br>b. PROJECT NO.<br>Task 187-200<br>c.<br>d. | 9a. ORIGINATOR'S REPORT NUMBER(S)<br>9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*<br>Tech. Report 161, Cooley Elec. Lab. |

10. AVAILABILITY/LIMITATION NOTICES

| | |
|---|---|
| 11. SUPPLEMENTARY NOTES | 12. SPONSORING MILITARY ACTIVITY<br>Office of Naval Research<br>Acoustic Branch, Code 468<br>Washington 25, D. C. |

13. ABSTRACT

The theory of sequential detection problems is extended to procedures of a composite signal hypothesis, such as signals whose amplitude or phase is unknown. Observations of signal-and-noise are assumed not to be independent so that information can be extracted about the signal parameters for use in the detection process. The yes-no terminal decision must be made within a finite period.

A general theory is presented for the solution of practical problems in the sequential detection of composite signal hypotheses. The theory specifies the optimum stopping rule needed for the sequential detector and the dependence between the observer's opinion of the cause of the reception and the distribution of the unknown signal parameters. The general theory implies the information that must be extracted from the observation is the likelihood ratio and the a posteriori signal parameter distribution. The form of the optimum detector, derived from the general theory, must include an adaptive capability, so that the signal parameter distribution can be sequentially updated.

Applications of the general theory to a signal of unknown phase and a signal of unknown amplitude are presented.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

| 14. KEY WORDS | LINK A | | LINK B | | LINK C | |
|---|---|---|---|---|---|---|
| | ROLE | WT | ROLE | WT | ROLE | WT |
| optimum receivers<br>signal processing<br>adaptive detector<br>adaptive signal detection<br>signal detectability theory<br>deferred decision theory<br>sequential decision procedures<br>unknown phase detection<br>unknown amplitude detection<br>sequential signal detection | | | | | | |

## INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.