On the Optimum Range Resolution of Radar Signals in Noise*

N. J. NILSSON†, MEMBER, IRE

Summary—Optimum radar resolution is recognized to be a problem in distinguishing between different possible target configurations. Radar reception systems which perform optimum range resolution are then designed using the principles of statistical decision theory. In particular, the design of the optimum resolution system is carried out for a squared-error loss function, modified to provide extra penalties for wrong guesses about the number of targets present. Such a system is capable of simultaneously deciding the number of targets present, their spatial positions (ranges) and their relative amplitudes. The analysis also includes a discussion of an optimum device for the resolution of distributed (clutter-like) targets.

INTRODUCTION

THE ability to resolve multiple-echo signals in time determines the range resolution of radar systems. The resolution rule-of-thumb for pulse radars is that echo pulses separated by a pulse length can be resolved, but echo pulses which overlap to any significant extent appear as only one target. Woodward[1] has developed a generalization of this rule which is applicable even for radar (sounding) signals whose time-bandwidth products are larger than unity. After defining the so-called Radar Ambiguity Function, Woodward infers that the time resolution cell for *any* signal is equal to the reciprocal of the (sounding) signal bandwidth.

However, this classical definition of resolution takes only qualitative account of the fact that the signals are embedded in noise. Even very narrow-band signals should be resolvable for arbitrarily small time separations in the complete absence of noise. We intuitively expect, then, that our ability to resolve two or more known signals in noise should depend, not only on the signal bandwidth, but on the echo-to-noise power ratio. In this paper, we shall treat the resolution problem as a problem of combined signal detection and signal estimation. Systems which achieve optimum range resolution will be derived, and their characteristics compared with those systems which are optimum only in the single-target detection sense.

Various authors have recognized the need for treating the general signal resolution problem in a more precise fashion than can be done by appeals to Woodward's ambiguity function alone. Swerling[2] has analyzed the problem of resolving two radar targets at the same range, but at (slightly) different angles within the antenna beamwidth. Helstrom[3,4] discusses the problem of distinguishing between two noise-corrupted signals whose form and location are known exactly. In this paper, we shall attempt first to formulate a suitable definition of radar resolution and then to apply this definition to the design of systems which perform optimum radar range resolution.

What is meant by "radar resolution?" A radar system achieving "good" resolution should be able to provide continuous and reliable answers to the following four questions which specify the target configuration:

a) How many (point) targets are there?
b) What are their relative (spatial) positions?
c) What are their relative velocities?
d) What are their relative amplitudes (cross sections)?

We define resolution in the following way. *An ensemble of target configurations (specified by the set of all situations which are allowable answers to the above four questions) is radar resolvable with average loss $\mathfrak{L}$, relative to a certain level of additive noise, a certain sounding signal, and a certain loss function L, if their respective composite echo returns can be distinguished by the Bayes decision device with minimum average loss $\mathfrak{L}$.*

Distinguishing among the many different possible target configurations is a problem in statistical decision theory. After defining the loss function L incurred for wrong guesses about the target configuration, we may calculate the average loss for any decision system. Decision systems with the lowest average loss are called Bayes decision systems and, therefore, according to the above definition, achieve the best resolution.

This definition of resolution prompts us to ask three more questions:

1) What *is* the optimum or Bayes radar resolving system relative to the noise, the sounding signal, and the loss function?

* Received by the PGIT, December 7, 1960. The research reported in this paper was conducted during the author's recently-completed term of active duty at Rome Air Dev. Ctr., Griffiss AFB, Rome, N. Y.

† Stanford Res. Inst., Menlo Park, Calif.

[1] P. M. Woodward, "Probability and Information Theory with Applications to Radar," McGraw-Hill Book Co., Inc., New York, N. Y.; 1953.

[2] P. Swerling, "The resolvability of point sources," in "Proceedings of Symposium on Decision Theory and Applications to Electronic Equipment Development, Vol. I," Rome Air Dev. Ctr. Griffiss AFB, Rome, N. Y., RADC-TR-60-70A; April, 1960.

[3] C. W. Helstrom, "The resolution of signals in white Gaussian noise," Proc. IRE, vol. 43, pp. 1111–1118; September, 1955.

[4] C. W. Helstrom, "Statistical Theory of Signal Detection," Pergamon Press, New York, N. Y., ch. X; 1960.

2) How well does it resolve (what is the minimum average loss) relative to the noise and the sounding signal for the given loss function?
3) For what (which) particular sounding signal(s) is the resolution best?

The present paper is devoted to finding the optimum system asked for in question 1). The optimum system should provide answers to questions a)–d) about the target configuration and distinguish optimally between different target configurations. We shall simplify the analysis, however, by assuming that all targets have the same angle and have zero velocity; that is, we inquire only about the number of targets, their relative ranges and their relative amplitudes. Questions 2) and 3), which have to do with quality of resolution and optimum sounding signal selection, are beyond the scope of this paper. Future treatment of these interesting questions should provide answers which will replace the rule-of-thumb about the reciprocal of the signal bandwidth with much more precise statements about resolution quality.[5]

Bayes Decision Procedures

The Target Density Function

When a radar (sounding) signal of form $s(t)$ is transmitted into a target environment, the received echo will be a linear superposition of like signals. In this paper, it will be assumed that the targets are stationary so that the individual echo returns comprising the total received echo will differ only in time of arrival and amplitude, corresponding to different target ranges and cross sections. The total received echo signal $S(t)$ can then be written as

$$S(t) = \int_0^T A(\tau)s(t-\tau)\,d\tau \qquad 0 \le t \le T \tag{1}$$

where

$A(\tau)$ = target density function $0 \le \tau \le T$
$s(t)$ = transmitted (sounding) signal
T = maximum possible target range (in seconds), assumed to be much greater than the reciprocal of the sounding signal bandwidth.

$s(t)$ is normalized such that

$$\int_0^T s^2(t)\,dt = 1. \tag{2}$$

$A(\tau)$ describes everything that is relevant about a stationary target configuration. Over that range of τ where a distributed target may exist, $A(\tau)$ is a continuous function of τ. On the other hand, point targets are represented by Dirac delta functions. For example, if there is a point target of amplitude A_1 at range τ_1, and another of amplitude A_2 at range τ_2, then

$$A(\tau) = A_1\,\delta(\tau-\tau_1) + A_2\,\delta(\tau-\tau_2)$$

and from (1)

$$S(t) = A_1 s(t-\tau_1) + A_2 s(t-\tau_2).$$

In what follows, we may at times restrict the possible target configurations to some given set. Let us call the set of allowed target configurations $\mathcal{A}$ and let A stand for any member $A(\tau)$ of the set $\mathcal{A}$. Suppose further that we know (or may ascribe) some *a priori* probability measure $p[A]$ to each member A of $\mathcal{A}$. Such a probability measure is necessary in the Bayes decision procedure, and it is well to state at the outset our assumptions about $p[A]$, even though they be of questionable merit. Sometimes we shall also define a probability measure $p[S]$ on the members S in the set $\mathcal{S}$ of all possible received echos. S, of course, stands for $S(t)$. When there is a one-to-one correspondence between the members of sets $\mathcal{A}$ and $\mathcal{S}$, then $p[A]$ will be identical with $p[S]$.

The Received Datum

We shall assume that the composite echo $S(t)$ is accompanied by additive, stationary, Gaussian noise denoted by $N(t)$. For simplicity, let $N(t)$ have zero mean value. The total received waveform is then

$$X(t) = S(t) + N(t) \qquad 0 \le t \le T. \tag{3}$$

We shall use the notation X as representing an arbitrary received waveform $X(t)$ belonging to some set of waveforms x.

Loss Functions and Bayes Decisions

The problem posed in this paper is: after reception of X, we must decide in an "optimum" manner which A in $\mathcal{A}$ represents the actual target configuration. We shall denote the result of this decision as $\hat{A}$, our estimate of A. In this paper, we shall equate the set $\hat{\mathcal{A}}$ of all possible estimates with the set $\mathcal{A}$ of all possible target density functions. That is, we shall never make an estimate $\hat{A}$ corresponding to an impossible target configuration.

We have defined an "optimum" decision as a Bayes decision. Let us denote the Bayes decision for A as $\hat{A}_b$. To make a Bayes decision, we must define a loss function L which fixes the loss incurred for erroneous decisions. That is, if A represents the actual target situation, but we decide $\hat{A}$, then we lose an amount $L[\hat{A}, A]$. The Bayes estimate $\hat{A}_b$ is that $\hat{A}$ which minimizes the average value of L.[6] Since choice of any $\hat{A}$ also implies by (1) a composite echo $S(\hat{A})$, we may choose to define our loss function in terms of $S(\hat{A})$ and S, that is, $L = L[S(\hat{A}), S]$. [When it

[5] For an interesting study of pairwise resolution and a technique for providing some answers to the problem of pairwise resolution quality, see G. W. Preston, "The Advanced Theory of Radar Measurements," Final Rept. to the Rome Air Dev. Ctr. on Contract AF 30(602)-2120, General Atronics Corp. Rept. No. 799-207-12; August 20, 1960.

[6] For a general discussion of Decision Theory, see: D. Blackwell and M. A. Girshick, "Theory of Games and Statistical Decisions," John Wiley and Sons, Inc., New York, N. Y.; 1954. For applications of Decision Theory to signal detection, see: D. Middleton, "Random processes, signals, and noise—an introduction to statistical communication theory," in "Pure and Applied Physics, Introductory Series," McGraw-Hill Book Co., Inc., New York, N. Y., ch. 21; 1960.

is convenient to write $S(\hat{A})$ explicitly as a function of time, we shall denote it by $\hat{S}(t)$.]

It is obvious that one way to minimize the average value of L is to choose $\hat{A}$ as some function of X in such a way that for *every* X, L averaged over the *a posteriori* probability measure for A, denoted by $p[A \mid X]$, is minimized.[7] That is, minimize the conditional expectation of L given X, denoted by $E_X[L]$. If L is defined in terms of $\hat{A}$ and A:

$$E_X[L(\hat{A}, A)] = \sum_{\mathcal{A}} p[A \mid X] L[\hat{A}, A] \tag{4}$$

where $p[A \mid X]$, an *a posteriori* probability measure, is called the *a posteriori* likelihood of A being the target density function given X. The sum over the set $\mathcal{A}$ represents an average over this set. $p[A \mid X]$ is obtained from the defined *a priori* probability measure $p[A]$ using Bayes' rule and the noise statistics. In case L is defined in terms of $S(\hat{A})$ and S, and, if there is a one-to-one correspondence between members of the sets $\mathcal{A}$ and $\mathcal{S}$, then we may write

$$E_X[L\{S(\hat{A}), S\}] = \sum_{\mathcal{S}} p[S \mid X] L[S(\hat{A}), S] \tag{5}$$

where $p[S \mid X]$, an *a posteriori* probability measure, is called the *a posteriori* likelihood of the composite echo S given X. $p[S \mid X]$ can be calculated from the probability measure $p[S]$ using Bayes' rule and the noise statistics. In both (4) and (5), $E_X[L]$ is a function of $\hat{A}$. The Bayes estimate A_b minimizes $E_X[L]$ over all other estimates $\hat{A}$ in the set of possible target configurations $\mathcal{A}$. In order to proceed further to see what sort of decision procedures arise, we must assume some particular loss functions.

Estimation of the Parameters of n Point Targets

The Loss Function

Before trying to design a system to guess how many targets exist, let us assume that a known number n exists and that we must estimate their parameters (ranges and amplitudes). Let us define the loss function L as being the integrated squared error in terms of the echo signal.

$$L = \int_0^T \{\hat{S}(t) - S(t)\}^2 \, dt \tag{6}$$

where $S(t)$ is the actual echo signal (a random process), and $\hat{S}(t)$ is the composite echo signal implied in (1) by the choice of the estimate $\hat{A}(\tau)$. $\hat{S}(t)$ is therefore a function of the estimate $\hat{A}$. Let us also restrict the set $\mathcal{A}$ to those target density functions which represent a collection of n point targets, that is

$$A(\tau) = \sum_{i=1}^{n} A_i \, \delta(\tau - \tau_i) \tag{7}$$

where

A_i = amplitude of ith target.
τ_i = range (in seconds) of ith target, $\tau_i \leq T$ for all i.

[7] Middleton, *op. cit.*, p. 1028.

Estimation of a particular target density function $\hat{A}$ is now achieved by estimating the components of the n-vectors

$$\hat{\boldsymbol{\tau}} = [\hat{\tau}_1, \hat{\tau}_2, \cdots \hat{\tau}_n] \quad \text{and} \quad \hat{\mathbf{A}} = [\hat{A}_1, \hat{A}_2, \cdots \hat{A}_n].$$

Making use of (6) and (7), and assuming a one-to-one correspondence between the members of sets $\mathcal{A}$ and $\mathcal{S}$, the conditional expectation of the loss is

$$E_X\left[\int_0^T \left\{\sum_{i=1}^{n} \hat{A}_i s(t - \hat{\tau}_i) - S(t)\right\}^2 dt\right]$$

where the $\hat{A}_i$ and $\hat{\tau}_i$ are estimates of A_i and τ_i, respectively, and the expectation is taken over the *a posteriori* likelihood for the waveform S. Expansion of this expression yields

$$\begin{aligned} E_X[L] = &\int_0^T \left[\sum_{i=1}^{n} \hat{A}_i s(t - \hat{\tau}_i)\right]^2 dt \\ &- 2\int_0^T \left[\sum_{i=1}^{n} \hat{A}_i s(t - \hat{\tau}_i)\right] E_X[S(t)] \, dt \\ &+ \int_0^T E_X[S^2(t)] \, dt. \end{aligned} \tag{8}$$

Since $E_X[L]$ is an average taken over the *a posteriori* likelihood for $S(t)$, the term $E_X[S(t)]$ in the above equation will be a waveform or function of time.

Minimization of the Average Loss

Since we are trying to minimize $E_X[L]$ by choice of $\hat{A}_i$ and $\hat{\tau}_i (i = 1, 2, \cdots n)$, we need only maximize the expression

$$\begin{aligned} J_n = &2 \sum_{i=1}^{n} \hat{A}_i \int_0^T E_X[S(t)] s(t - \hat{\tau}_i) \, dt \\ &- \sum_{i=1}^{n} \sum_{j=1}^{n} \hat{A}_i \hat{A}_j \int_0^T s(t - \hat{\tau}_i) s(t - \hat{\tau}_j) \, dt. \end{aligned} \tag{9}$$

The above equation can be written as

$$J_n = 2 \sum_{i=1}^{n} \hat{A}_i \hat{\phi}_i - \sum_{i=1}^{n} \sum_{j=1}^{n} \hat{A}_i \hat{A}_j \hat{\lambda}_{ij} \tag{10}$$

where

$$\hat{\phi}_i = \phi(\hat{\tau}_i) = \int_0^T E_X[S(t)] s(t - \hat{\tau}_i) \, dt$$

and

$$\begin{aligned} \hat{\lambda}_{ij} = \lambda(\hat{\tau}_i - \hat{\tau}_j) &= \int_0^T s(t - \hat{\tau}_i) s(t - \hat{\tau}_j) \, dt \\ &= \text{the ``autocorrelation function'' of } s(t). \end{aligned}$$

$\hat{\phi}_i$ is obtained as the output of a filter, at time $\hat{\tau}_i$, which is matched to the transmitted signal $s(t)$ when $E_X[S(t)]$ is the filter input waveform.

We shall first maximize J_n by proper choice of the $\hat{A}_i$ and then obtain final maximization by choice of the $\hat{\tau}_i$. We can differentiate (10) with respect to $\hat{A}_k$ and set the result equal to zero to derive the following relationship that must hold for the maximizing amplitude estimates $\hat{A}_{io}$:

$$\sum_{i=1}^{n} \hat{A}_{io}\lambda_{ki} = \hat{\phi}_k. \tag{11}$$

If both sides of the above equation are multiplied by $\hat{A}_{ko}$ and summed over k, we obtain

$$\sum_{i,k=1}^{n} \hat{A}_{io}\hat{A}_{ko}\lambda_{ki} = \sum_{k=1}^{n} \hat{A}_{ko}\hat{\phi}_k. \tag{12}$$

Since (12) must also be satisfied by the maximizing estimates $\hat{A}_{io}$, then J_n, maximized over the $\hat{A}_i$, can be written from (10) as

$$J_{no} = \sum_{i=1}^{n} \hat{A}_{io}\hat{\phi}_i \tag{13}$$

where the $\hat{A}_{io}$ must satisfy (11).

It is perhaps more convenient to express the above relationships in matrix notation. Let us represent all the $\hat{A}_{io}$ by the the n-dimensional column vector

$$\hat{\mathbf{A}}_0 = \begin{bmatrix} \hat{A}_{1o} \\ \hat{A}_{2o} \\ \vdots \\ \hat{A}_{io} \\ \vdots \\ \hat{A}_{no} \end{bmatrix} \tag{14}$$

and group all the λ_{ij} into an $n \times n$ matrix

$$\hat{\boldsymbol{\lambda}} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1j} & \cdots & \lambda_{1n} \\ \lambda_{21} & & & & & \\ \vdots & & & & & \\ \lambda_{i1} & & \cdots & \lambda_{ij} & \cdots & \lambda_{in} \\ \vdots & & & & & \vdots \\ \lambda_{n1} & & & \cdots & & \lambda_{nn} \end{bmatrix}. \tag{15}$$

Then the set of (11) may be written as the matrix equation

$$\hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\lambda}}\hat{\mathbf{A}}_o \tag{16}$$

where $\hat{\boldsymbol{\phi}}$ is the n-dimensional column vector

$$\hat{\boldsymbol{\phi}} = \begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_i \\ \vdots \\ \hat{\phi}_n \end{bmatrix}. \tag{17}$$

Both $\hat{\boldsymbol{\phi}}$ and $\hat{\boldsymbol{\lambda}}$ are functions of the vector $\hat{\boldsymbol{\tau}} = [\hat{\tau}_1, \hat{\tau}_2, \cdots \hat{\tau}_n]$. The maximizing estimate $\hat{\mathbf{A}}_o$ must then satisfy the matrix equation

$$\hat{\mathbf{A}}_o = \hat{\boldsymbol{\lambda}}^{-1}\hat{\boldsymbol{\phi}} \tag{18}$$

where $\hat{\boldsymbol{\lambda}}^{-1}$ is the inverse of $\hat{\boldsymbol{\lambda}}$.

J_{no} can also be written from (13) in vector notation as the dot product

$$J_{no} = \hat{\boldsymbol{\phi}} \cdot \hat{\mathbf{A}}_o \tag{19}$$

or, using the relationship given in (18),

$$J_{no} = \hat{\boldsymbol{\phi}} \cdot (\hat{\boldsymbol{\lambda}}^{-1}\hat{\boldsymbol{\phi}}). \tag{20}$$

Eq. (20) is an expression for J_n maximized over the vector $\hat{\mathbf{A}}$. J_{no} is still a function of the vector $\hat{\boldsymbol{\tau}}$. J_{no} must then be further maximized by choice of a $\hat{\boldsymbol{\tau}}_b$ which fixes $\hat{\boldsymbol{\phi}}_b$ and $\hat{\boldsymbol{\lambda}}_b^{-1}$.

The final maximum for J_n can then be written as

$$J_{nb} = \hat{\boldsymbol{\phi}}_b \cdot (\hat{\boldsymbol{\lambda}}_b^{-1}\hat{\boldsymbol{\phi}}_b). \tag{21}$$

Upon discovering the maximizing $\hat{\boldsymbol{\tau}}_b$, we may write an expression for the Bayes estimate $\hat{\mathbf{A}}_b$ from (18) in the following way

$$\hat{\mathbf{A}}_b = \hat{\boldsymbol{\lambda}}_b^{-1}\hat{\boldsymbol{\phi}}_b. \tag{22}$$

Let us summarize the above expressions by stating the rule for finding the Bayes estimates $\hat{A}_{1b} \cdots \hat{A}_{nb}, \hat{\tau}_{1b}, \cdots \hat{\tau}_{nb}$ for the positions and amplitudes of a known number n of targets:

1) Form the quantity $\phi(\tau)$ by passing $E_X[S(t)]$ through a filter matched to $s(t)$.

2) Form an n-vector $\hat{\boldsymbol{\phi}} = [\phi(\hat{\tau}_1), \cdots \phi(\hat{\tau}_n)]$ and an $n \times n$ matrix $\hat{\boldsymbol{\lambda}}$ by selecting estimates $\hat{\tau}_1, \cdots \hat{\tau}_n$ in such a way that the quadratic form $J_{no} = \hat{\boldsymbol{\phi}} \cdot (\hat{\boldsymbol{\lambda}}^{-1}\hat{\boldsymbol{\phi}})$ is maximized. Let us say that maximization occurs for the vector $\hat{\boldsymbol{\tau}}_b = [\hat{\tau}_{1b}, \cdots \hat{\tau}_{nb}]$ which determines a $\hat{\boldsymbol{\phi}}_b$ and a $\hat{\boldsymbol{\lambda}}_b$.

3) Using the maximizing $\hat{\boldsymbol{\tau}}_b$, calculate the Bayes amplitude estimates $\hat{A}_{1b}, \cdots \hat{A}_{nb}$ by

$$\hat{\mathbf{A}}_b = \hat{\boldsymbol{\lambda}}_b^{-1}\hat{\boldsymbol{\phi}}_b.$$

In short, the receiver has only to calculate the conditional mean $E_X[S(t)]$ and pass this waveform through a matched filter to obtain $\phi(\tau)$. Then, using its knowledge of $\lambda(\tau)$, it performs certain maximizing operations to obtain $\hat{\mathbf{A}}_b$ and $\hat{\boldsymbol{\tau}}_b$. If $S(t)$ is a T-second sample from a stationary Gaussian process,[8] $E_X[S(t)]$ is equal to that $S_o(t)$ which maximizes the *a posteriori* likelihood $p[S \mid X]$. (In a Gaussian distribution, the mode equals the mean.)

[8] The requirement that $S(t)$ must be a sample from a stationary process involves ignoring the radar "range-to-the-fourth-power" law for reasonable ensembles of target configurations. The assumption that $S(t)$ is Gaussian seems reasonable if the proper rationalizations about $s(t)$ and $p[A]$ are made. In particular, for any $s(t)$, S will be Gaussian if the A_i are Gaussian. However, the relationship between the statistics of $S(t)$ and $A(t)$ for various sounding signals should be investigated.

Youla[9] has shown that $S_o(t) = E_X[S(t)]$ is given by the following set of integral equations:

$$S_o(t) = \int_0^T X(\tau)h(t-\tau)\,d\tau, \qquad 0 \le t \le T$$

$$\int_0^T [R_S(\xi-\tau) + R_N(\xi-\tau)]h(t-\tau)\,d\tau = R_S(\xi - t) \qquad (23)$$

where

$R_S(\tau)$ = autocovariance function of the random process of which $S(t)$ is a T-second sample

and

$R_N(\tau) = \langle N(t)N(t+\tau)\rangle$, the autocovariance function of the noise.

These two equations state that $S_o(t)$ can be obtained from a linear filter with input $X(t)$ and impulse response $h(\tau)$ where $h(\tau)$ is the solution to a modified Wiener-Hopf integral equation. As the equations stand, $h(\tau)$ is not physically realizable, but this situation can be corrected if we are willing to tolerate, at most, a delay of T seconds in obtaining $S_o(t)$. In the limit as T becomes very large, the filter specified in (23) approaches the unrealizable case of Wiener's least-mean-square-error filter. Finally, to obtain $\phi(\tau)$, $E_X[S(t)]$ is passed through a filter matched to $s(t)$ as shown in Fig. 1.

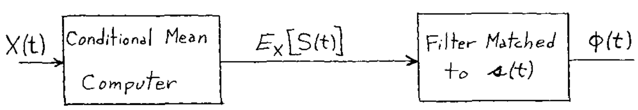

Fig. 1—Processing for $\phi(\tau)$.

Special Case: $n = 1$

When it is known that there is only one target present, the Bayes estimates for A_1 and τ_1 are obtained in a straightforward manner. $\hat{\lambda} = \hat{\lambda}^{-1} = (1)$ in the one-dimensional case so that, from (20),

$$J_{1,o} = \hat{\phi}_1^2. \qquad (24)$$

That is, we must choose an estimate $\hat{\tau}_1$, which maximizes $\phi(\tau)$, and thus $\phi^2(\tau)$. Then, from (22)

$$\hat{A}_{1b} = \hat{\phi}_{1b}. \qquad (25)$$

Referring to Fig. 2, we locate the maximum of $\phi(\tau)$ and equate the target strength with this maximum. Such a procedure using a matched filter operating on $X(t)$, instead of on $E_X[S(t)]$, has long been the accepted procedure.

[9] D. C. Youla, "The use of the method of maximum likelihood in estimating continuous-modulated intelligence which has been corrupted by noise," IRE Trans. on Information Theory, no. IT-3, pp. 90–105; March 1954.

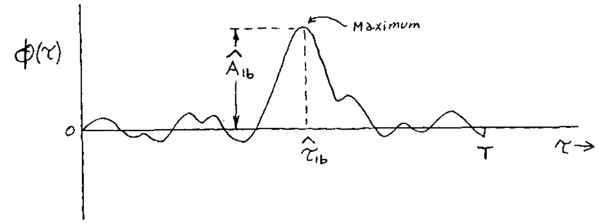

Fig. 2—Bayes estimate for a single target.

Special Case: $n = 2$

For two targets, we must make estimates $\hat{A}_{1b}$, $\hat{A}_{2b}$, $\hat{\tau}_{1b}$, and $\hat{\tau}_{2b}$. We first calculate $\hat{\lambda}^{-1}$ from $\hat{\lambda}$.

$$\hat{\lambda} = \begin{bmatrix} 1 & \lambda_{12} \\ \lambda_{12} & 1 \end{bmatrix} \qquad (26)$$

where $\lambda_{12} = \lambda_{21}$ because $\lambda(\tau)$ is an even function. Inverting the above matrix yields

$$\hat{\lambda}^{-1} = \begin{bmatrix} \dfrac{1}{1-\lambda_{12}^2} & \dfrac{-\lambda_{12}}{1-\lambda_{12}^2} \\ \dfrac{-\lambda_{12}}{1-\lambda_{12}^2} & \dfrac{1}{1-\lambda_{12}^2} \end{bmatrix} \qquad (27)$$

so that from (20)

$$J_{2,o} = \frac{\hat{\phi}_1^2 - 2\lambda_{12}\hat{\phi}_1\hat{\phi}_2 + \hat{\phi}_2^2}{1-\lambda_{12}^2}. \qquad (28)$$

We must choose a $\hat{\tau}_{1b}$ and a $\hat{\tau}_{2b}$ such that the resulting $\hat{\phi}_{1b}$, $\hat{\phi}_{2b}$, and λ_{12b} maximize $J_{2,o}$. $\hat{\tau}_{1b}$ and $\hat{\tau}_{2b}$ are the Bayes estimates of the target positions. Then, using $\hat{\boldsymbol{\phi}}_b$ and $\hat{\lambda}_b^{-1}$ in (22), we calculate

$$\hat{A}_{1b} = \frac{\hat{\phi}_{1b} - \lambda_{12b}\hat{\phi}_{2b}}{1-\lambda_{12b}^2} \qquad (29)$$

and

$$\hat{A}_{2b} = \frac{\hat{\phi}_{2b} - \lambda_{12b}\hat{\phi}_{1b}}{1-\lambda_{12b}^2}.$$

Eq. (29) is identical with the result of Helstrom[4] who discusses a similar resolution problem, except that Helstrom uses $X(t)$ as the input to the matched filter instead of $E_X[S(t)]$.

Specialized Two-Target Configurations

It is seen that a generalization of the simple matched-filter is optimum when two targets of unknown positions and amplitudes are present. Maximization of (28) in two dimensions is somewhat more complex, though, than looking for the maximum output of a simple filter. $\hat{\phi}_{1b}$ and $\hat{\phi}_{2b}$ will not in general occur at relative maxima of $\phi(\tau)$ unless these relative maxima are very high indeed and are also separated by more than a correlation interval. (It is beyond the scope of this paper to discuss possible means of implementing a device for automatically computing $\hat{\tau}_{1b}$ and $\hat{\tau}_{2b}$; we shall limit ourselves to the task of specifying mathematically the optimum operations.) Let us study some limiting-case two-target configurations to gain some insight into properties of the optimum procedure.

1) *Two Targets Known to be Separated by More Than the Correlation Interval*: If $|\tau_2 - \tau_1|$ is always greater than the reciprocal of the sounding signal bandwidth, λ_{12} will approach 0, and, from (28)

$$J_{2,o} = \hat{\phi}_1^2 + \hat{\phi}_2^2. \tag{30}$$

It is obvious that (30) can be maximized by choosing for $\hat{\phi}_{1b}$ and $\hat{\phi}_{2b}$ the two highest peaks in $\phi(\tau)$ which are separated by more than the correlation interval. Eq. (29), with $\lambda_{12b} = 0$, is then used to compute $\hat{A}_{1b}$ and $\hat{A}_{2b}$. This technique is close to the way in which multitarget situations are handled by present-day radars.

2) *Two Targets Indistinguishably Close*: If it is known that $\tau_1 \cong \tau_2$, and it is only needed to find the mean range, the two-target estimation problem reduces to the one-target problem. As $\hat{\tau}_1 \to \hat{\tau}_2$, and thus $\hat{\lambda}_{12} \to 1$, and $\hat{\phi}_2 \to \hat{\phi}_1$, (28) becomes

$$J_{2,o} = \hat{\phi}_1^2 = \hat{\phi}_2^2 \tag{31}$$

which can be maximized by finding the maximum of $\phi(\tau)$.

3) *A Small Target Near a Large One*: If $A_1 \gg A_2$, we may be justified in setting $\hat{\phi}_{1b} = \max \phi(\tau)$, ignoring what we may later decide as estimates of τ_2 and the consequent $\hat{\lambda}_{12}$. Thus, $\hat{\tau}_{1b}$ is approximately the τ which maximizes $\phi(\tau)$. For any received waveform, this maximum will be some constant $\hat{\phi}_{1b}$.

Regarding $\hat{\phi}_1$ as a constant in (28), we may write

$$J_{2,o} = \hat{\phi}_{1b}^2 + \frac{[\hat{\phi}_2 - \lambda(\hat{\tau}_2 - \hat{\tau}_{1b})\hat{\phi}_{1b}]^2}{1 - \lambda^2(\hat{\tau}_2 - \hat{\tau}_{1b})}. \tag{32}$$

$J_{2,o}$ can be maximized by choice of $\hat{\tau}_2$ by selecting the τ which maximizes the quantity

$$\frac{[\phi(\tau) - \hat{\phi}_{1b}\lambda(\tau - \hat{\tau}_{1b})]^2}{1 - \lambda^2(\tau - \hat{\tau}_{1b})}.$$

The above expression describes approximately what the receiver must do to locate a small target in the presence of a large one. First, the maximum of $\phi(\tau)$ is found and its time of occurrence noted. The large target is guessed to be located at this point. Its effect is *subtracted* from $\phi(\tau)$, and the result is squared and divided by $[1 - \lambda^2(\tau - \hat{\tau}_{1b})]$. This new waveform is then scanned in τ for a maximum which occurs, say, at $\hat{\tau}_{2b}$. Now that $\hat{\tau}_{1b}$ and $\hat{\tau}_{2b}$ are known, (29) allows us to calculate estimates for the target amplitudes.

A Multitarget Detection and Estimation Problem

The Hybrid Loss Function and Its Minimization

One important assumption that we have made in our development must now be extended. So far, we have assumed a known number n of targets, making the problem one of estimation rather than one of detection. If n is random, we are faced with a combined detection and estimation problem, the solution of which will lead to optimum resolution systems as they were defined in the Introduction. We will also have to modify somewhat the loss function given by (6) so that extra penalties can result when the wrong *number* of targets is guessed.

Let us consider the following loss function, written with explicit reference to the number of targets guessed to be present and actually present:

$$L[S(\hat{A}), S; i, j] = \int_0^T [S(\hat{A}) - S(t)]^2\, dt + \alpha_{ij}. \tag{33}$$

In the above equation, the notation is the same as that in (6). The added term α_{ij} is the *extra* loss incurred for guessing i targets present when really there are j. Let us compose a matrix from the components α_{ij}:

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_{00} & \alpha_{01} & \alpha_{02} & \cdots & \\ \alpha_{10} & & & & \\ \vdots & & & & \\ \cdot & & \alpha_{ij} & \cdots & \\ \vdots & & \vdots & & \end{bmatrix}. \tag{34}$$

$\boldsymbol{\alpha}$ is an infinite matrix with elements α_{ij} defined for all $i = 0, 1, 2, \cdots$, and $j = 0, 1, 2, \cdots$.

Proceeding exactly as in the previous section, we want to minimize the conditional expectation of the loss, given $X(t)$. It is easily shown that this minimization is equivalent to maximizing the expression

$$K_n = J_n - E_X[\alpha_{nj}] \tag{35}$$

by simultaneous choice of n, $\hat{\boldsymbol{\tau}}$ and $\hat{\mathbf{A}}$; where J_n is given, for each n, by (10), and $E_X[\alpha_{nj}]$ = the expectation given $X(t)$ of α_{nj} over all possible numbers of targets j.

The following procedure is used to maximize K_n. First, for each n, $J_{nb} = \hat{\boldsymbol{\phi}}_b \cdot (\hat{\boldsymbol{\lambda}}_b^{-1}\hat{\boldsymbol{\phi}}_b)$ is found by selecting an n-dimensional vector $\hat{\boldsymbol{\tau}}_b$ such that the resulting $\hat{\boldsymbol{\phi}}_b$ and $\hat{\boldsymbol{\lambda}}_b^{-1}$ maximize the n-quadratic form $\hat{\boldsymbol{\phi}} \cdot (\hat{\boldsymbol{\lambda}}^{-1}\hat{\boldsymbol{\phi}})$. Then we calculate

$$\begin{aligned} E_X[\alpha_{nj}] &= \alpha_{n0}p_X(0) + \alpha_{n1}p_X(1) \\ &\quad + \cdots \alpha_{nj}p_X(j) + \cdots \\ &= \boldsymbol{\alpha}_n \cdot \mathbf{p}_X \end{aligned} \tag{36}$$

where

$\boldsymbol{\alpha}_n$ = a Hilbert space vector with components α_{n0}, α_{n1}, α_{n2}, $\cdots$

$\mathbf{p}_X$ = a Hilbert space vector with components $p_X(0)$, $p_X(1)$, $p_X(2)$, $\cdots$

and

$p_X(i)$ = the *a posteriori* probability that i targets are present given the received waveform $X(t)$.

By using J_{nb} instead of J_n in (35), we have the maximized expression

$$K_{no} = J_{nb} - \boldsymbol{\alpha}_n \cdot \mathbf{p}_X \tag{37}$$

which must be further maximized by choice of $\hat{n}$. If it were not for the term $\boldsymbol{\alpha}_n \cdot \mathbf{p}_X$, K_{no} would have no maximum over n because it can be shown that $J_{n+1,\,b} \geq J_{nb}$ for all n.[10] In practice, the highest value of n to be tested will be some finite number, probably two or three, so that all the K_{no} could, in principle, be calculated and the largest found. Once $\hat{n}$ is determined, the $\hat{n}$-vector $\hat{\tau}_b$ gives the Bayes range estimates, and the corresponding $\hat{\boldsymbol{\phi}}_b$ and $\hat{\boldsymbol{\lambda}}_b^{-1}$ are used to calculate $\hat{\mathbf{A}}_b$.

Let us now review the procedure for optimum estimation and detection in a multitarget environment:

1) The receiver calculates $\phi(\tau)$ and $\mathbf{p}_X$ from the received waveform $X(t)$. $\phi(\tau)$ is again the matched-filtered version of $E_X[S(t)]$. The calculation of $E_X[S(t)]$ for the case of random n will be discussed later in greater detail. $\mathbf{p}_X$ is a vector whose ith component is just the *a posteriori* probability that i targets are present.

2) From $\phi(\tau)$ and $\lambda(\tau)$ the receiver finds, for each n, that $\hat{\boldsymbol{\phi}}_b$ and $\hat{\boldsymbol{\lambda}}_b^{-1}$ which maximize the n-quadratic form $\hat{\boldsymbol{\phi}} \cdot (\hat{\boldsymbol{\lambda}}^{-1}\hat{\boldsymbol{\phi}})$. The maxima of the n-quadratic forms are called J_{nb}.

3) The quantity $K_{no} = J_{nb} - \boldsymbol{\alpha}_n \cdot \mathbf{p}_X$ is the then calculated and maximized over n. That is, if $K_{\hat{n}o} > K_{no}$ for all integers $n \neq \hat{n}$, then $\hat{n}$ is an optimum or Bayes estimate for the number of targets present.

4) The range estimates of the $\hat{n}$ targets are then the components of the $\hat{n}$-vector $\hat{\boldsymbol{\tau}}_b$ which determined the maximizing $\hat{\boldsymbol{\phi}}_b$ and $\hat{\boldsymbol{\lambda}}_b^{-1}$ for the $\hat{n}$th dimension.

5) The $\hat{n}$-dimensional vector $\hat{\boldsymbol{\phi}}_b$ and the $\hat{n} \times \hat{n}$ matrix $\hat{\boldsymbol{\lambda}}_b$ are then used to calculate the Bayes amplitude estimates $\hat{A}_{1b}, \cdots \hat{A}_{\hat{n}b}$ of the $\hat{n}$ targets by the expression $\hat{\mathbf{A}}_b = \hat{\boldsymbol{\lambda}}_b^{-1}\hat{\boldsymbol{\phi}}_b$. The above procedure completes the combined Bayes detection and estimation technique. For the loss function assumed, it provides optimum resolution. Special examples of this procedures will be considered below.

Calculation of $E_X[S(t)]$

When the number of targets n is a random variable, the method of Youla[9] cannot be directly applied to the calculation of $E_X[S(t)]$. When n is random, we may not assume that the *a priori* probability density function for the waveform $S(t)$ is a multidimensional Gaussian distribution if there is a finite possibility that $S(t) = 0$, *i.e.*, no targets. If the *a posteriori* probability for zero targets, given $X(t)$, is $p_X(0)$, and if *otherwise* $p[S]$ and $p[S \mid X]$ can be considered multidimensional Gaussian likelihood functions, then

$$E_X[S(t)] = [1 - p_X(0)]S_0(t) \tag{38}$$

where $S_o(t)$ is that S which maximizes the "continuous" or Gaussian portion of $p[S \mid X]$. $S_o(t)$ can be obtained by the same linear filter described in (23).

[10] For any n, a J_{n+1} which equals J_{nb} can always be obtained by choosing two of the $\hat{\tau}_i$ to be equal in the $(n+1)$-vector $\hat{\tau}$ to make it in reality an n-vector. Thus, $J_{n+1,\,b} \geq J_{nb}$.

Some Special Cases

Let us now consider some examples to illuminate the combined detection and estimation theory that has been developed. First, we must select an appropriate matrix $\boldsymbol{\alpha}$. Following somewhat the philosophy of Bennion,[11] who uses a loss function which penalizes a given amount for false alarms but penalizes only with the squared error for false rest, we write

$$\boldsymbol{\alpha} = \alpha \begin{bmatrix} 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 1 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 2 & 1 & 0 & 0 & \cdot & \cdot & \cdot \\ 3 & 2 & 1 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}. \tag{39}$$

That is, we lose an extra amount 3α if we say 3 targets present when there are really none, etc.

1) *Detection of a Single Target*: Suppose we need only choose the larger of $K_{0,o}$ or $K_{1,o}$ and, if $K_{1,o}$ is our choice, make estimates of the single target's range and amplitude. Such a situation is a combined single-target detection and estimation problem. From (10), guessing no target present implies $J_{o,b} = 0$, so that

$$K_{0,o} = -\boldsymbol{\alpha}_o \cdot \mathbf{p}_X$$

and

$$K_{1,o} = \max_{\tau} \phi^2(\tau) - \boldsymbol{\alpha}_1 \cdot \mathbf{p}_X. \tag{40}$$

We must choose the maximum, which leads to the rule: say a target is present if

$$K_{1,o} - K_{0,o} = \{\max_{\tau} \phi^2(\tau) - [\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0] \cdot p_X\} > 0 \tag{41}$$

or if

$$\max_{\tau} \phi^2(\tau) > \alpha p_X(0)$$

otherwise say no target is present. If a target is declared present, $\hat{\tau}_{1b}$ and $\hat{A}_{1b}$ are calculated as before. Notice in this example that the combined detection and estimation problem involves essentially a threshold detection scheme even though the value of the threshold is a function of the received waveform [through $p_X(0)$].

2) *Pairwise Resolution*: To answer the question: "Are there two targets present or just one?" merely involves choosing between $K_{1,o}$ and $K_{2,o}$. Now the $J_{2,b}$ and $J_{1,b}$ are calculated using $\phi(\tau)$ and $\lambda(\tau)$. If

$$J_{2,b} - J_{1,b} > \alpha[p_X(1)] \tag{42}$$

then say two targets are present, but if not, then say only one target is present. Estimates for A_i and τ_i can then be made. If a small target is to be detected in the presence of a large one, the procedure under (32) may be

[11] D. Bennion, "Some Results in the Estimation of Signal Parameters," Stanford Electronics Lab., Stanford University, Stanford, Calif., Tech Rept. No. 10; September 10, 1956.

used to obtain the rule: say two targets are present if

$$\max_{\tau} \frac{[\phi(\tau) - \hat{\phi}_{1b}\lambda(\tau - \hat{\tau}_{1b})]^2}{1 - \lambda^2(\tau - \hat{\tau}_{1b})} > \alpha p_X(1) \tag{43}$$

otherwise say only one target present. The estimates for A_i and τ_i are obtained as before.

3) *Widely Separated Targets*: If it is known *a priori* that any targets will be widely separated, $\hat{\lambda}$ is an identity matrix as is $\hat{\lambda}^{-1}$. In this case, $J_{nb} = \sum_{i=1}^{n} \hat{\phi}_{ib}^2$ and

$$K_{no} = \sum_{i=1}^{n} \hat{\phi}_{ib}^2 - \boldsymbol{\alpha}_n \cdot \mathbf{p}_X. \tag{44}$$

The following procedure may be used to maximize $K_{n,o}$ by choice of n:

a) The receiver computes $\phi(\tau)$ and finds its maximum, ϕ_{1b}. If $\hat{\phi}_{1b}^2 > \alpha p_X(0)$, a target is announced and its parameters are computed.

b) The next highest peak in $\phi(\tau)$ which is more distant from τ_{1b} then a correlation interval is selected as $\hat{\phi}_{2b}$. If $\hat{\phi}_{2b}^2$ is greater than $\alpha[p_X(0) + p_X(1)]$, then a second target is announced and its parameters are computed.

c) This process is repeated until the threshold condition is not satisfied, giving a decision of, say, $\hat{n}$ targets and their parameters.

Discussion

We have proposed an optimum technique for simultaneously performing the generalized detection (how many targets) and the estimation (where are they and what are their amplitudes) problems. The detection problem involves the maximization of a set of quadratic forms ($n = 0, 1, 2, \cdots$). From each maximized quadratic form, a constant is subtracted which is a function of the *a posteriori* probability of a certain number of targets being present. Then, the resulting terms are compared to find the largest. We have assumed a modified squared-error loss function which inflicts extra penalties for false guesses about the number of targets present.

The receiver must compute the conditional expectation of the composite echo signal as well as the *a posteriori* probabilities for various numbers of targets being present. The conditional expectation of the echo signal can sometimes be obtained by a modified type of Wiener filtering operation on the received, noise-corrupted, waveform. If no assumptions about *a priori* echo statistics are made, the received datum (signal plus noise) is usually used in place of the conditional expectation of the composite echo signal. Such practice is a consequence of using a maximum likelihood procedure instead of a squared-error Bayes procedure.

So far, we have concerned ourselves only with point targets which cause $A(\tau)$ to take the form of delta functions. If optimum estimates of $A(\tau)$ for distributed targets, such as clutter, are to be found, a somewhat different tack must be taken. Now we shall discuss some results for this important problem.

Distributed Targets

The Turin Filter

Suppose $A(\tau)$ does not consist of point targets, but, instead, represents distributed targets such as clutter. Then a loss function involving the integrated squared error between $A(\tau)$ and the estimate $\hat{A}(\tau)$ might be appropriate. Such a problem was considered by Turin,[12] although he restricted his estimator to a linear filter.[13] Under the assumption that $A(\tau)$ is a T-second sample from a stationary Gaussian random process with autocovariance function $R_A(\tau)$, the transfer function of the optimum linear estimator, as derived by Turin is

$$T(\omega) = \frac{F^*(\omega)}{F(\omega)F^*(\omega) + \dfrac{G_N(\omega)}{G_A(\omega)}} \tag{45}$$

where

$F(\omega)$ = Fourier transform of $s(t)$
$F^*(\omega)$ = conjugate of $F(\omega)$
$G_N(\omega)$ = power spectral density of $N(t)$
= Fourier transform of $R_N(\tau)$
$G_A(\omega)$ = power spectral density function of the random process of which $A(t)$ is a T-second sample
= Fourier transform of $R_A(\tau)$.

Special Properties of the Turin Filter

The filter shown in Fig. 3 is a "crispening filter."

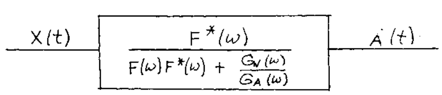

Fig. 3—Turin filter. The optimum linear estimator for the target density function.

(We disregard the extra time delay needed to insure realizability.) The $F^*(\omega)$ in the numerator "compresses" the received waveform by causing phase reinforcement, while the denominator accents the high-frequency components allowing for fast risetimes. For example, if the average noise power is much less than the average echo power, the Turin filter approaches an inverse filter. An inverse filter has a delta function output every time $s(t)$ occurs at the input. On the other hand, if the average noise power is much greater than the average echo power, then the Turin filter becomes a simple matched filter. No rise-time crispening can be done in the face of so much noise. In any case, if $|F(\omega)|$ is rectangular and if the noise and $A(\tau)$ are white, then the filter also becomes a matched filter.

[12] G. L. Turin, "On the estimation in the presence of noise of the impulse response of a random, linear filter," IRE Trans. on Information Theory, vol. IT-3, pp. 5–10; March, 1957.

[13] Which is no restriction under the assumption that $A(\tau)$ is Gaussian.

The resolution enhancing properties of this filter come about due to its crispening action. The lower the noise-to-echo power ratio, the greater is the crispening action, and hence, the finer is the potential range (time) resolution capability. It is assumed that $G_N(\omega)$ is known, and $G_A(\omega)$ can be calculated from the statistics of the ensemble of target density functions.

Conclusion

The problem considered in this paper, briefly stated, concerned optimum methods of signal detection. The signal form, however, was not known exactly except for the fact that is was a composite of known signal forms echoing from a random configuration of targets. We stated in the Introduction that the solution of such a detection problem also solves the optimum resolution problem if resolution is defined to mean distinguishing between the different possible composite echos. Various procedures were derived for making estimates of the target density function (which may either be discrete or continuous) under special assumptions regarding *a priori* probabilities and loss functions.

No attempt was made to evaluate different radar systems, including optimum ones, to determine just how well they are able to distinguish between different target configurations in the face of random noise. This problem was posed as question 2) in the Introduction and must be answered before quantitative statements can be made about resolution ability. Also unanswered is the problem of deciding what sounding signal to transmit so that resolution ability can be further enhanced. It is hoped that the present study might provide stimulation for a complete set of answers to these questions. In addition to the above two important problems, the following suggestions are offered as topics for further research:

1) Investigation of the *a priori* and *a posteriori* statistics of $S(t)$ for various assumptions about the statistics of the target configuration, the statistics of the noise, and the known sounding signal.
2) Generalization of the results of this paper to include the other spatial dimensions (angle) and velocity.
3) Investigation of other loss functions.
4) Instrumentation of a two-target resolver using (28) and measurement of its performance in noise.
5) Analysis of methods to calculate approximately the *a posteriori* probabilities for the number of targets present.

Appendix

List of Symbols

A, $A(\tau)$ = Target density function
A_i = Amplitude of ith target
$\mathbf{A}$ = Target amplitude vector whose ith component is A_i
$\mathcal{A}$ = Set of all possible $A(\tau)$
$E_X[\ ]$ = Conditional expectation of [ ]
$F(\omega)$ = Fourier transform of $s(t)$
$G_A(\omega)$, $G_N(\omega)$ = Power spectral density functions of $A(\tau)$ and $N(t)$, respectively
J_{no}, J_{nb} = The quadratic form $\hat{\boldsymbol{\phi}} \cdot (\hat{\boldsymbol{\lambda}}^{-1}\hat{\boldsymbol{\phi}})$ and its maximum, respectively
K_{no} = An expression derived from J_{nb} by subtraction of $\boldsymbol{\alpha}_n \cdot \mathbf{p}_X$
L = The loss function defined for each combination of $[\hat{A}(\tau), A(\tau)]$ or $[S(\hat{A}), S(t)]$
n = Number of point targets present
N, $N(t)$ = Zero-mean, stationary, Gaussian noise
$p[\ ]$, $p[\ |\ X]$ = Likelihood and conditional likelihood, respectively, of the waveform [ ]
$p_X(i)$ = Conditional probability of i targets present given X
$\mathbf{p}_X$ = A Hilbert space vector whose ith component is $p_X(i)$
$R_A(\tau)$, $R_N(\tau)$, $R_S(\tau)$ = Autocovariance functions of the stationary random processes $A(\tau)$, $N(t)$, and $S(t)$, respectively
s, $s(t)$ = Sounding (transmitted) signal form
S, $S(t)$ = Composite echo signal consisting of a linear superposition of weighted and delayed $s(t)$'s
$\mathcal{S}$ = Set of all possible $S(t)$ as determined by $\mathcal{A}$
T = Maximum possible target range (delay time)
X, $X(t)$ = Received datum = $S + N$
$\boldsymbol{\alpha}$ = A loss matrix with terms α_{ij} being the extra loss incurred for guessing i targets when there are really j targets
$\boldsymbol{\alpha}_i$ = A vector whose jth component is α_{ij}
$\delta(\tau)$ = Dirac delta function
$\lambda(\tau) = \int_0^T s(t)s(t+\tau)\,dt$, "autocorrelation function" of $s(t)$
$\hat{\boldsymbol{\lambda}}$ = A matrix consisting of terms $\lambda_{ij} = \lambda(\hat{\tau}_i - \hat{\tau}_j)$
$\hat{\boldsymbol{\lambda}}^{-1}$ = The inverse of $\hat{\boldsymbol{\lambda}}$
ϕ, $\phi(\tau)$ = Matched-filtered version of $E_X[S(t)]$
$\hat{\boldsymbol{\phi}}$ = A vector whose ith component is $\hat{\phi}_i = \phi(\hat{\tau}_i)$
τ_i = Range (delay) of the ith target

The symbol ˆ is used over a parameter to indicate an estimate of that parameter; for example, $\hat{A}_i$ is an estimate of A_i. The subscript b is used with an estimate to indicate a Bayes estimate; for example, $\hat{A}_{ib}$ is the Bayes estimate of A_i. These symbols are also used in conjunction with terms like ϕ and λ when they are evaluated using Bayes estimates. For example, $\hat{\phi}_{ib} = \phi(\hat{\tau}_{ib})$. The symbols $\langle\ \rangle$ are used to indicate the (unconditional) expectation of the enclosed quantity.

Acknowledgment

The work reported in this paper is the result of the author's studies on radar resolution problems at the Rome Air Development Center. The author gratefully acknowledges the support and encouragement of his associates at the Center. Specific thanks are due to Lt. J. K. Wolf of the Rome Air Development Center for many illuminating suggestions. Also appreciated were the many stimulating discussions with Prof. N. Abramson of the Electrical Engineering Department of Stanford University.