raddis, 10-30-89

THEORY OF RADAR TARGET DISCRIMINATION

E. T. Jaynes
Wayman Crow Professor of Physics
Washington University, St. Louis MO 63130

Abstract: We give a direct application of probability theory to the problem of deciding which of a set of possible targets is present. The reliability of discrimination depends on the noise level, the background hash, the variation of echo with target aspect angles, the energy and shape of the transmitted pulse, and the number of pulses. The effect of each of these variables is calculated and discussed, leading to some new conclusions about optimal radar design and optimal data processing. We think that the tactics which might succeed are quite different from those that have been tried in the past, and give elementary intuitive explanations of why this is so.

1. INTRODUCTION

If radar systems could distinguish different targets from each other, there would be big advantages in air safety. Airport traffic controllers have made serious errors from their inability to determine which echo on their screen represents which flight. In the recent Persian Gulf incident, it appears that a passenger plane was shot down because a shipboard radar could not distinguish its echo from that of a fighter plane. In the near future it will become important to identify different space vehicles. Presumably, good target discrimination would be helpful also in radar weather forecasting; and the same principles will apply as well in ultrasound imaging for medical diagnosis. But although the technical problem of target discrimination has been well recognized and studied for many years, no good solutions have been forthcoming.

With recent renewed emphasis on the importance of the problem, it appeared that better understanding of the theoretical problem is a prerequisite for any practical hardware improvements. Past efforts have tended to consider the problem as one of physics (electromagnetic/acoustic scattering theory, etc.). But although the physics is well understood, this alone has not led to progress. More fundamentally, it is a problem of *information processing*, calling for a full application of probability theory. There have been few past efforts to use probability theory, and they have been based on "sampling theory" methods which are unable to deal with nuisance parameters such as aspect angle, or to make use of all the supplementary information available to a radar operator or system.

In the present work we go back to fundamentals and consider the problem from the start as one of probabilistic inference, in which the knowledge from physics is taken for granted and used to tell us how to formulate the problem. Most important, we use full Bayesian probability theory, which overcomes the limitations of sampling theory.

A transmitted pulse $f(t)$ gives rise to an echo from a target, of the form

$$y(t) = \int r(t-t')f(t')dt' \qquad (1-1)$$

where $r(t)$ is the "impulse response function", or as we shall call it, the *reflection function* , of the target, which we consider defined for all time. Presumably,

$$r(t) = 0 \quad \text{when } t < 2d/c \qquad (1-2)$$

where d is the distance to the nearest part of the target, c the velocity of light. In the theory, however, we do not assume this; the final formulas turn out to have the same general form whether or not (2) is satisfied. Thus our results would hold also in discrimination problems where the variable t is not a time, and the physical causality condition (1-2) need not hold.

More important are the meanings of $f(t)$ and $y(t)$. One could take these to be the forms of the actual electromagnetic fields in space; if so, practically all of the following theory would remain valid but for the addition of position variables as parameters: $f(x,t)$, $y(x,t)$. However, these results would then need to be convolved with the properties of antennas and matching circuits before they would be expressed in terms of the easily measurable quantities, the voltages and currents at the actual transmitter and receiver terminals.

It is much more convenient to take $f(t)$ to be the transmitted pulse as measured at the transmitter terminals (presumably a certain reference plane in a coaxial line or waveguide); and $y(t)$ to be the echo part of the received signal as measured similarly at the receiver terminals. With this interpretation the following theory is exact as given, and all the functions needed to apply it are directly measurable with standard laboratory equipment.

Our reflection functions are then convolutions:

$$r(t) = (\text{transmitter function}) * (\text{echo in space}) * (\text{receiver function}). \qquad (1-3)$$

But in the frequency domain this reduces to a simple product

$$R(\omega) = A_T(\omega)\ E(\omega)\ A_R(\omega), \qquad (1-4)$$

and the effects that depend only on the target are separated automatically from the radar design parameters. In any event, the properties of our targets that are relevant for discrimination with a given (*i.e.,* already built) antenna system are the $R(\omega)$ functions, not the $E(\omega)$ functions.

Note also that the physics of the problem (both electromagnetic scattering theory and antenna theory) is contained entirely in the $r(t)$ or $R(\omega)$ functions. Whether these are expressed in a modal expansion, singularity expansion, creeping wave analysis, or just measured experimentally, makes no difference. What is relevant to the problem before us (decide between a set of possible targets) is simply the numerical values of the $r(t)$ functions themselves, because they carry all the information about the target that is in the echoes.

We stress this point because of a widespread belief that determining the poles of the singularity expansion is essential to target identification, because they are aspect independent. Indeed, if a few poles could be determined from the received echo, that would lead to the desired identification. However, separate identification of the poles does not appear feasible in practice because of receiver noise and the rapid decay of the echo. But it seems obvious that separate pole identification, while sufficient if it could be accomplished, cannot be necessary.

The reason for this is that probability theory will give us its final verdict on any particular target in the form of a single number, the probability that it is the one present. In calculating it, probability theory will automatically take into account all the information in the data that is relevant to this question, and whatever prior information is available. The result will be, presumably, some average over the joint probability distribution for all the pole positions. To channel the analysis through a phase of estimating the separate pole positions is not only a larger calculation, but a less informative one, for this ignores not only correlations in that joint probability distribution, but also other relevant information that may be in the data.

Indeed, if the returned echoes depend on aspect, it follows that any prior information about aspect that we have, will help us to make target identification. But once the course of a target is known, we know a great deal about its aspect. It would be self-defeating to concentrate our attention on the poles because they are aspect-independent, so strongly that we ignore this highly cogent information about aspect.

When the physics has been done, in whatever way, and we have the reflection functions $r(t)$ for our targets, then the real problem (probability analysis of incomplete information) is ready to begin. If the poles are indeed the essential factor in target identification, this analysis will tell us so automatically; and it will tell us also the quantitative way in which they enter into the problem. Until the results of this analysis are at hand, we are not in a position to judge what role the poles may play in the problem, beyond intuitive guesswork.

The total data set $D \equiv \{d(t)\}$ available for processing is not just the target echo $y(t)$. It has in general two other unavoidable components:

$$d(t) = y(t) + h(t) + n(t) \qquad (1-5)$$

where $h(t)$ is "hash", representing ground clutter and echoes from any other objects in the antenna beam or side lobes:

$$h(t) = \int r_H(t-t')f(t')dt' \qquad (1-6)$$

and $n(t)$ is noise. This always includes at least the universal noise from thermal radiation falling on the antenna. At the frequencies and temperatures of concern to us, $(hf \ll kT)$, thermal radiation follows the Rayleigh-Jeans equipartition law for the normal modes of space, leading to the Nyquist thermal noise law corresponding to the antenna radiation resistance [mean-square open circuit voltage in a bandwidth Δf Hz of $\delta V^2 = 4R_{rad}kT\Delta f$, where k is Boltzmann's constant, 1.36 E-23 joules/degree Kelvin]. In addition, $n(t)$ may have contributions from the internally generated noise of an imperfect receiver, as discussed in Sec. 2 below, as well as atmospheric disturbances and jamming signals.

For our purposes, the functional distinction between hash and noise is not that they have different physical origins, but that they have different effects on target discrimination, because $h(t)$ is systematic (*i.e.*, the same on successive pulses) while $n(t)$ varies from one pulse to the next in a way that we can neither predict nor control.

Of course, any data function $d(t)$ which can be recorded for computer processing will be digitized and sampled only at discrete times; but we expect this digitizing to be so good that the continuum approximation used here is accurate enough for all practical purposes. In any event, the final results are such that the effects of coarse digitizing are evident.

Consider now the simplest imaginable problem of discrimination; to decide between two possible fixed targets, without the complications of aspect angle and hash; and we analyze only the data from a single pulse. Almost all the conceptual subtleties that have been troublesome in the past are present already in this simple "baby" version of the problem. After we have worked out its full solution and understood it thoroughly, we shall find it relatively easy to deal with the complications of the real world, which are matters of technical detail rather than basic understanding.

2. DISCRIMINATION BETWEEN TWO TARGETS

If target A is present, the echo function is

$$y_A(t) = \int r_A(t-t')f(t')dt', \qquad (2-1a)$$

while if target B is present it is

$$y_B(t) = \int r_B(t-t')f(t')dt'. \qquad (2-1b)$$

Then our data from a single pulse will be

$$d(t) = y_A(t) + n(t), \text{ if the target is A}, \qquad (2-2a)$$

$$d(t) = y_B(t) + n(t), \text{ if the target is B}. \qquad (2-2b)$$

We shall take $n(t)$ to be white Gaussian noise, with expected square σ^2; *i.e.*, we take the probability of a given noise function $n(t)$ to be proportional to

$$p(n(t)|\sigma) \propto \exp\left\{-\frac{1}{2\sigma^2}\int n^2(t)dt\right\}. \qquad (2-3)$$

We indicate this as a conditional probability, conditional on knowledge of σ. If σ is unknown, it must be estimated from the data and probability theory tells us the proper way of doing this, as shown by Bretthorst (1988). But in the present problem, σ will be known in advance because it is essentially the noise temperature of the receiver [see Eq. (2-54) below], and one will surely have determined this before trying to test the system at all. This simplifies our calculation.

Now if we knew that A is in fact the true target present, then the probability of getting a given data function $d(t)$ would be just the probability that the noise would make up the difference in (2-3):

$$p(D|A\sigma) \propto \exp\left\{-\frac{1}{2\sigma^2}\int [d(t) - y_A(t)]^2\, dt\right\} \qquad (2-4a)$$

while if B is true, this probability is

$$p(D|B\sigma) \propto \exp\left\{-\frac{1}{2\sigma^2}\int [d(t) - y_B(t)]^2\, dt\right\} \qquad (2-4b)$$

where we are using D as an abbreviation for the entire run of data $d(t)$. These are the "sampling probabilities" for our problem.

But the probabilities we need are the other way around: what is the probability, given the data, that A is the true target? These are

$$p(A|D\sigma), \qquad p(B|D\sigma). \qquad (2-5)$$

Probability theory tells us how to obtain them from the sampling probabilities (2-4). By the product rule, the probability that both A and D are true is

$$p(AD|\sigma) = p(A|\sigma)\, p(D|A\sigma) = p(D|\sigma)\, p(A|D\sigma) \qquad (2-6)$$

since the proposition 'AD' on the left-hand side is the same as 'DA' (*i.e.*, Boolean logic is commutative). Therefore,

$$p(A|D\sigma) = p(A|\sigma)\, \frac{p(D|A\sigma)}{p(D|\sigma)}. \qquad (2-7a)$$

The first factor on the right is the "prior probability" $p(A|\sigma)$, which is clearly necessary in all inference from data. That is, to ask "What do you know about A after getting the data D?" does not make sense — it is not a well posed question — unless we take into account, "What did you know about A before getting D?". The second factor is the "likelihood", which shows how the prior probability is updated as a result of getting the evidence of the data D. Likewise, the probability that B is the true target, is

$$p(B|D\sigma) = p(B|\sigma)\, \frac{p(D|B\sigma)}{p(D|\sigma)}. \qquad (2-7b)$$

Now in the simple problem being considered, we are given at the outset that there are only two possible targets, A and B. Therefore

$$p(A|\sigma) + p(B|\sigma) = 1, \qquad (2-8)$$

and this is still true after getting the data, so

$$p(A|D\sigma) + p(B|D\sigma) = 1. \qquad (2-9)$$

Then (2-7) and (2-9) give us

$$p(D|\sigma) = p(D|A\sigma)\, p(A|\sigma) + p(D|B\sigma)\, p(B|\sigma), \qquad (2-10)$$

which is a special case of a more general probability rule: given a set of any mutually exclusive and exhaustive propositions $\{A_1, \ldots, A_n\}$ and any propositions X, Y,

$$p(X|Y) = \sum_{i=1}^{n} p(XA_i|Y) = \sum_i p(X|A_iY)\, p(A_i|Y), \qquad (2-11)$$

which we shall need later in dealing with multiple targets.

For many purposes we can eliminate $p(D|\sigma)$ by considering probability ratios, or odds, instead of probabilities. In the present binary problem these are the same thing; the ratio of probabilities of A and B is

$$\frac{p(A|D\sigma)}{p(B|D\sigma)} = \frac{p(A|\sigma)}{p(B|\sigma)}\, \frac{p(D|A\sigma)}{p(D|B\sigma)}, \qquad (2-12)$$

while the odds on any proposition X with probability $p(X)$ are $o(X) = p/(1-p)$. But because of (2-8), (2-9)

$$\frac{p(A|D\sigma)}{p(B|D\sigma)} = \frac{p(A|D\sigma)}{1 - p(A|D\sigma)} = o(A|D\sigma) \qquad (2-13)$$

so it does not matter which terminology we use. With multiple targets, odds and probability ratios are no longer the same.

Using (2-7), the normalization constants that we left out of (2-3) cancel out anyway and the odds on target A reduce to

$$o(A|D\sigma) = o(A|\sigma) \exp\left\{ \frac{1}{\sigma^2}[d\cdot(y_A - y_B) + \frac{1}{2}(y_B\cdot y_B - y_A\cdot y_A)] \right\} \qquad (2-14)$$

where we have used the abbreviations

$$d\cdot y_A \equiv \int d(t)\, y_A(t)\, dt, \qquad (2-15)$$

$$y_A\cdot y_A \equiv \int y_A(t)\, y_A(t)\, dt,$$

etc. A term $(d\cdot d)$ has cancelled out. If we have any prior information about which target is likely to be present, this should be expressed in the prior odds term $o(A|\sigma)$. If, as usual, we have no such information, this term is equal to unity. In either case, $\sigma^2 \log[o(A|D\sigma)/o(A|\sigma)]$ is the fundamental quadratic form, on which all depends.

One reason for past confusion is that different workers have appealed only to their differing intuitions about how the data should be analyzed, without making any attempt to see what probability theory has to tell us about the problem. Intuition can give us bits and pieces of the truth; but it almost never gives us the whole truth.

Now we see from (2-12) that, since the data appear nowhere else, the import of the data for this problem resides entirely in the "likelihood ratio" $L = p(D|A\sigma)/p(D|B\sigma)$. All other aspects of the data are irrelevant for the problem of deciding between A and B; few people have perceived this intuitively. Probability theory tells us, in (2-14), how the data should be processed for optimal discrimination between targets. With Gaussian noise, a simple linear operation on the data is the optimal computation which generates the posterior log-odds in favor of one target over the other.

Now if A is in fact the true target present, then $d(t) = y_A(t) + n(t)$, and the result of this data processing will be

$$\sigma^2 \log o(A|D\sigma) = n \cdot (y_A - y_B) + y_A \cdot (y_A - y_B) + \frac{1}{2}(y_B \cdot y_B - y_A \cdot y_A) \qquad (2-16)$$

or,

$$\sigma^2 \log o(A|D\sigma) = n \cdot (y_A - y_B) + \frac{1}{2}(y_A - y_B) \cdot (y_A - y_B) \qquad (2-17a)$$

a "random noise" part and a systematic part. If B is the true target, our computer will find instead the log-odds in favor of A of

$$\sigma^2 \log o(A|D\sigma) = n \cdot (y_A - y_B) - \frac{1}{2}(y_A - y_B) \cdot (y_A - y_B) \qquad (2-17b)$$

in which the systematic term has a reversed sign.

The term $n \cdot (y_A - y_B)$ represents an unavoidable confusion due to noise. Eqs. (2-17) tell us that if $n(t)$ happens to resemble $(y_A - y_B)$, this term will be positive and it will incline us in the direction of favoring A. If $n(t)$ happens to have the opposite sign, so it resembles $(y_B - y_A)$, it will make us tend to favor B. We shall estimate the magnitude of this term presently [Eq. (2-57)]; but from symmetry it is as likely to be positive as negative, so the expected log-odds in favor of A comes from the systematic term alone:

$$\sigma^2 \, \langle \log o(A|D\sigma) \rangle = \pm\frac{1}{2} \, (y_A - y_B) \cdot (y_A - y_B) \qquad (2-18)$$

with the plus sign if A is true.

Evidently, for best discrimination between A and B we want to make the magnitude of $(2-18)$ as large as possible. To see how this depends on the reflection functions and the transmitted pulse, write the difference in reflection functions (2-1) as

$$r(t-t') \equiv r_A(t-t') - r_B(t-t'). \qquad (2-19)$$

We have from (2-1),

$$(y_A - y_B) \cdot (y_A - y_B) = \int dt \left[\int dt' \, r(t-t') \, f(t')\right] \left[\int dt'' \, r(t-t'') \, f(t'')\right]$$

$$= \int \int f(t') g(t', t'') f(t'') dt' dt'' \qquad (2-20)$$

where

$$g(t', t'') \equiv \int dt \;\, r(t-t') \, r(t-t''). \qquad (2-21)$$

Abbreviating the integral in (2-20) by '$\int\int fgf$ ', this is

$$(y_A - y_B) \cdot (y_A - y_B) = \int \int fgf \qquad (2-20')$$

We discuss the maximization problem first in the time domain, then in the frequency domain.

Time Domain

The condition that (2-20) be a maximum for a given total amount of energy radiated, ' $\int f^2$ ' = $\int [f(t)]^2 dt$, is found by Lagrange multipliers: in our shorthand notation,

$$0 = \delta \left[\int \int fgf - \lambda \int f^2 \right] = \int 2\delta f \cdot \left[\int gf - \lambda f \right] \qquad (2-22)$$

or, the condition for stationarity of $\int \int fgf$ is the integral equation

$$\int g(t-t')\, f(t')\, dt' = \lambda f(t). \qquad (2-23)$$

To understand the condition for a maximum, note that if this integral equation had a discrete set of eigenvalues and normalized eigenfunctions:

$$\int g(t-t')\, \phi_i(t')\, dt' = \lambda_i\, \phi_i(t), \qquad i = 1, 2, \cdots \qquad (2-24)$$

then we could view it in a very simple way. Given any function $f(t)$, expand it in the eigenfunctions:

$$f(t) = \sum_i a_i\, \phi_i(t) \qquad (2-25)$$

Then we find that

$$\frac{\int \int fgf}{\int f^2} = \frac{\sum_i |a_i|^2\ \lambda_i}{\sum_i |a_i|^2} \qquad (2-26)$$

is a weighted average of the eigenvalues. This makes it obvious that the absolute maximum is achieved when $f(t)$ is proportional to that eigenfunction belonging to the greatest eigenvalue, and (2-26) shows how much the performance will deteriorate when $f(t)$ is not optimal. This would give us essentially complete understanding of the problem.

However, our $g(t, t')$ is not of this type; it has continuous eigenvalues and non-normalizable eigenfunctions. To see this, note from (2-21) that it is translationally invariant:

$$g(t', t'') = g(t' - t'') \qquad (2-27)$$

and so, if $f(t)$ was an eigenfunction of (2-23), then $f(t-s)$ would be one also for all real s. There are two possibilities: (1) there is an infinite degeneracy; (2) $f(t)$ is an exponential function. This is symptomatic that things will be simpler in the frequency domain.

Frequency Domain

Taking note of (2-27), define the fourier transforms

$$G(\omega) \equiv \int g(t) e^{i\omega t} dt \qquad (2-28)$$

$$F(\omega) \equiv \int f(t) e^{i\omega t} dt \qquad (2-29)$$

Then we need some Parseval-type formulas:

$$\int g(t'-t'') f(t'') dt'' = \int dt'' f(t'') \int \frac{d\omega}{2\pi} G(\omega) e^{-i\omega(t'-t'')} = \int \frac{d\omega}{2\pi} G(\omega) F(\omega) e^{-i\omega t'} \qquad (2-30)$$

and

$$\int\int fgf = \int dt' f(t') \int \frac{d\omega}{2\pi} G(\omega) F(\omega) e^{-i\omega t'} = \int \frac{d\omega}{2\pi} G(\omega) |F(\omega)|^2 \qquad (2-31)$$

and the conventional Parseval theorem:

$$\int f^2(t) dt = \int \frac{d\omega}{2\pi} |F(\omega)|^2 . \qquad (2-32)$$

The ratio to be maximized is now

$$\frac{\int\int fgf}{\int f^2} = \frac{\int d\omega \; |F(\omega)|^2 \; G(\omega)}{\int d\omega \; |F(\omega)|^2} \qquad (2-33)$$

which is, analogous to (2-26), a weighted average of the values of $G(\omega)$, weighted according to the power density of the transmitted pulse at frequency ω. This makes it, again, obvious how the quality of discrimination for a given transmitted energy depends on the properties of the targets as described by $G(\omega)$, and on the spectrum of the transmitted pulse as described by $F(\omega)$.

Now let us relate $G(\omega)$ more directly to the target reflection functions. Referring to Equations $(2-19)-(2-21)$, we can make another Parseval-type relation:

$$g(t', t'') = \int dt \, r(t-t') \int \frac{d\omega}{2\pi} R(\omega) \, e^{-i\omega(t'-t'')} = \int \frac{d\omega}{2\pi} \; |R(\omega)|^2 \, e^{-i\omega(t'-t'')} \qquad (2-34)$$

In other words, we have simply

$$G(\omega) = |R_A(\omega) - R_B(\omega)|^2 \qquad (2-35)$$

which makes (2-33) appear very cogent and sensible. This is the usual outcome of a Bayesian probability analysis; a final result that intuition would never have found for us, but which seems intuitively right after a little meditation.

The transmitted pulse that is optimal for purposes of target discrimination will then have its spectrum concentrated near the frequency where $|R_A(\omega) - R_B(\omega)|$ reaches its absolute maximum. In fact, however, in existing radar systems the transmitted pulse will have a spectrum concentrated rather sharply near some carrier frequency ω_o which was not chosen with this problem in mind at all. Then the combined result of the above equations is that, if A is the true target, a single pulse will give us an expected log-odds in favor of A of approximately

$$\langle \log o(A|D\sigma) \rangle \simeq \frac{\int f^2(t) dt}{2\sigma^2} \; |R_A(\omega_o) - R_B(\omega_o)|^2 \; , \qquad (2-36)$$

provided that $G(\omega)$ is not rapidly varying in the neighborhood of ω_o. The first factor on the right is a kind of signal/noise ratio; *i.e.,* it is something vaguely like

$$\frac{\text{(energy radiated in a pulse)}}{\text{(noise energy incident on the receiver)}} \qquad (2-37)$$

But to make this precise we must now examine the noise term $n(t)$, its probability distribution, and some of the facts of life concerning receiver operation, a little more closely. The term σ^2 in (2-36) is essentially the receiver noise temperature T_N, but with a conversion factor that requires some effort to derive. Previously we defined σ only by the probability distribution (2-3).

To find this conversion factor exactly, we need first a short digression on the meaning of our transmitter and receiver signals $f(t)$, $y(t)$, $n(t)$. We decided before to define these as the values measured at certain reference planes in the coaxial cables or waveguides connecting transmitter and receiver to their antennas; but until now we did not need to decide whether they are voltages, currents, travelling wave amplitudes, etc.

Transmission Lines and Receivers

In a transmission line of characteristic impedance Z (which might be the wave impedance of a waveguide mode), there is a voltage and current $v(t)$, $i(t)$ at this reference plane (which in a waveguide represent the amplitudes of the transverse electric and magnetic fields of the mode being used). The forward and backward traveling wave amplitudes are

$$f_\pm(t) = \frac{1}{2}\left[\frac{v(t)}{\sqrt{Z}} \pm i(t)\sqrt{Z}\right], \qquad (2-38)$$

with the meaning that f_+^2 and f_-^2 are the instantaneous powers in watts, carried by the forward and backward waves. We verify that, indeed, the difference

$$f_+^2(t) - f_-^2(t) = v(t)i(t) \qquad (2-39)$$

is the net instantaneous power flow.

Now we define the transmitter pulse $f(t)$ as the forward wave amplitude, at the transmitter reference plane, travelling from transmitter to transmitting antenna. Likewise, by $y(t)$ and $n(t)$ we mean the components of the travelling wave amplitudes at the receiver reference plane, travelling from receiving antenna to receiver.

We should be aware that there is a difference in the circuit conditions for these two waves. In the transmitter, one will take pains to match the transmission system to the antenna so that all the energy in the forward wave $f(t)$ is radiated out into space instead of being wasted setting up standing waves in the transmission line. It will be desirable also to match the receiver transmission line to the receiving antenna, and we assume henceforth that this has been done.

One might then think naïvely that we should take equal care to match the receiver to its transmission line so that all the energy captured by the receiving antenna is actually delivered to the receiver. However, this is not the case for a good receiver. In order to detect radiation it is not necessary to absorb it; for a magnetic field can deflect a charged particle in an observable way without delivering any energy to it. An electric field can deflect a charged particle in an observable way while actually removing energy from it. In fact, an ideal receiver does not run on energy at all, but reflects back all the energy incident on it!

The point here is that the receiver is designed not for maximum energy, but for maximum signal/noise ratio, at its output. Matching the receiver to its transmission line would indeed give maximum output energy for a given gain; but that is not what we want. How much of the noise at the receiver output is amplified noise presented to its input, how much is generated internally by imperfections (Nyquist thermal noise, shot noise, etc.) in the receiver?

An ideal receiver is one that generates no internal noise, but delivers at the output a signal/noise ratio equal to that at the input. Suppose, then, that there is a desired signal $y(t)$ and unwanted noise $n(t)$, which are wave amplitudes travelling toward the receiver, giving an incident signal/noise ratio $(S/N)_{inc} = y^2/n^2$. If the receiver presents an infinite impedance at this reference plane $[\,i(t) = 0]$, then from (2-38) there is a signal voltage $v_{sig}(t) = 2\sqrt{Z}\ y(t)$ and a noise voltage $v_{noise}(t) = 2\sqrt{Z}\ n(t)$, leading to a signal/noise ratio $v^2_{sig}/v^2_{noise} = (S/N)_{inc}$, which the ideal receiver amplifies and delivers to its output. If the incident noise $n(t)$ is Nyquist noise, carrying average power $P = \langle n^2 \rangle = kT\Delta f$ in a bandwidth Δf, then the average v^2_{noise} is $4Z\ kT\Delta f$.

Now if we match the receiver to the input transmission line, the signal voltage is cut in half but the noise voltage is not because we must reckon with a new source of thermal noise, that generated by the receiver input impedance Z. The impedance which determines the total noise voltage at the reference plane is now Z/2, the parallel combination of the impedances looking toward receiver and antenna, and the RMS noise voltage at the reference plane will be reduced only by a factor $\sqrt{2}$ rather than 2. Even if the receiver is ideal from this point on, its output signal/noise ratio cannot be better than that at the input reference plane, which is now 3 db lower than $(S/N)_{inc}$.

So, if the receiver generates no internal noise, we would lose 3 db in output signal/noise ratio by matching it to its transmission line. If the receiver input impedance at the reference plane is zero rather than infinite, interchange voltage and current in the above arguments and the 3 db loss conclusion still holds. If the receiver input impedance is purely reactive, then it will appear infinite or zero at some other reference plane, at which these arguments will apply. So quite generally, in order to deliver the maximum signal/noise ratio at its output, an ideal receiver must reflect all the energy incident upon it.

It is only in the limit of an "infinitely bad" receiver, in which all the output noise is generated internally, that matched input impedance becomes the condition for maximum signal/noise ratio at the output. Actual receivers are somewhere between ideal and infinitely bad, and so they perform best when partially matched, so that a part of the incident energy is reflected and radiated back out the receiver antenna.

This fact surprises many people on first hearing; but we note that it is so general that it remains true in quantum theory, at optical frequencies where $hf \gg kT$ and the Nyquist noise formula no longer holds. For initiation of a photochemical reaction it is not necessary that the light energy be absorbed. For example, it might be thought that the eyes of animals adapted to seeing in the dark would have pupils that act as perfect black bodies, absorbing all the incident light energy. On the contrary, it is a familiar fact that the animals with best night vision have eyes that reflect the incident light strongly, looking like search-lights in the dark.

The result of this little digression is that while the transmitted signal $f(t)$ is looking into a matched transmission line, the received signal $y(t) + n(t)$ will not be in general, and the noise which interferes with target discrimination does not come entirely down the transmission line from the receiving antenna.

Receiver Noise Considerations

The noise performance of receivers must be specified in a way that includes both the noise actually incident on the receiver terminals, and the internally generated noise. In effect, we note the output signal/noise ratio and then imagine an ideal receiver, which would have the same S/N ratio at its input. The input noise of this ideal receiver is greater than the Nyquist value for the ambient temperature; but of course it can be written in the Nyquist form with some higher temperature.

Thus we take for the effective average of $n^2(t)$

$$\langle n^2(t) \rangle = kT_N \ \Delta f \tag{2-40}$$

where Δf is the bandwidth amplified by the receiver, and T_N its "noise temperature". One also speaks of the "noise figure" of a receiver, being the ratio of its noise temperature to the ambient temperature. Thus a receiver with a "6 db noise figure" is one whose noise temperature is four times room temperature: $4 \times (20 + 273) = 1172^o K$.

The fact that we are concerned with a finite bandwidth greatly simplifies the probability description of the noise, because it means that the sampling theorem representation is available. Given a fourier transform pair

$$F(\omega) = \int f(t) e^{i\omega t} dt, \tag{2-41}$$

$$f(t) = \int \frac{d\omega}{2\pi} F(\omega) e^{-i\omega t} \tag{2-42}$$

if it is band limited to frequencies less than Ω:

$$F(\omega) = 0, \qquad |\omega| \le \Omega \tag{2-43}$$

then define the Nyquist sampling times and sampling functions:

$$t_k \equiv \frac{\pi k}{\Omega}, \quad k = 0, \ \pm 1, \ \pm 2, \dots \tag{2-44}$$

$$s_k(t) \equiv \frac{\sin \Omega (t - t_k)}{\Omega (t - t_k)} \tag{2-45}$$

Then the theorem is that a band-limited function is a sum of $(\sin x / x)$ functions:

$$f(t) = \sum_{k=-\infty}^{\infty} f(t_k) s_k(t) \tag{2-46}$$

Furthermore, this is an expansion in orthogonal functions, for

$$\int_{-\infty}^{\infty} s_j(t) s_k(t) dt = \frac{\pi}{\Omega} s_j(t_k) = \frac{\pi}{\Omega} \delta_{jk}. \tag{2-47}$$

Then the integral of a product of band-limited functions is

$$\int n(t) g(t) dr = \int dt \sum_{jk} n_j g_k s_j(t) s_k(t) = \frac{\pi}{\Omega} \sum_j n_j g_j \tag{2-48}$$

and, in the special case $g(t) = n(t)$,

$$\int n^2(t)dt = \frac{\pi}{\Omega} \sum_j n_j^2. \qquad (2-49)$$

Note that (2-48) and (2-49) are not merely discrete sum approximations to the integrals; for band-limited functions they are exact.

Now we defined the quantity σ in (2-3) by saying that the noise is supposed white, and the probability of a noise function $n(t)$ shall be

$$p(n(t)|\sigma) \propto \exp\left\{-\frac{1}{2\sigma^2}\int n^2(t)dt\right\}. \qquad (2-50)$$

This is now, from (2-49),

$$\exp\left\{-\frac{\pi}{2\sigma^2\Omega}\sum_j n_j^2\right\}. \qquad (2-51)$$

But this states that the variables $n_j = n(t_j)$ are assigned independent Gaussian distributions with means $\langle n_j \rangle = 0$ and second moments

$$\langle n_j n_k \rangle = \frac{\Omega}{\pi}\,\sigma^2 \delta_{jk.} \qquad (2-52)$$

In other words, our definition (2-50) plus the band-limited condition implies white noise in the sense that values of $n(t)$ separated by Nyquist intervals are independent. The noise is as "white" as it can be in view of the band limiting.

This enables us to find the missing conversion factor between σ and the noise temperature T_N. From our definition of $n(t)$ as the amplitude of a travelling wave, the expectation of energy carried by it in the frequency bandwidth $\Delta f = \Omega/2\pi$ in some long time interval τ is from (2-49), (2-52),

$$\frac{\pi}{\Omega}\sum_{t_j=0}^{\tau} n_j^2 = kT_N \cdot \frac{\Omega\tau}{2\pi} \qquad (2-53)$$

The number of terms in the sum is $\tau/\delta t = \Omega\tau/\pi$, where $\delta t = \pi/\Omega$ is the Nyquist sampling interval. By (2-52) these terms are all equal. Therefore Ω and τ cancel out, and (2-53) becomes simply

$$\sigma^2 = \frac{1}{2}kT_N \qquad (2-54)$$

just the average thermal energy per degree of freedom according to the Rayleigh-Jeans law, at temperature T_N. Although the argument leading to this result has been long, we are rewarded in the end with a pleasant surprise: a beautifully simple formula.

Another equally nice result is the estimated value of the integral (2-48). As we noted before [Eqs. (2-17), (2-18)] its expectation is, trivially $\langle (n \cdot g) \rangle = 0$; but now we can calculate its expected square. Using (2-48) we have

$$\langle (n \cdot g)^2 \rangle = \frac{\pi^2}{\Omega^2}\sum_{jk}\langle n_j n_k \rangle g_j g_k = \frac{\pi}{\Omega}\,\sigma^2 \sum_j g_j^2$$

or, in view of (2-49),

$$\langle (n \cdot g)^2 \rangle = \sigma^2 \int g^2(t)dt. \qquad (2-55)$$

That this turns out so simple and neat is another pleasant surprise. Now we can return to the log-odds calculation (2-17), (2-36) with all factors known.

Final Results:
The approximate expected log-odds (2-36) in favor of target A is now simply

$$\langle \log o(A|D\sigma)\rangle \simeq \frac{(\text{Energy radiated per pulse})}{kT_N}\, |R_A(\omega_o) - R_B(\omega_o)|^2, \qquad (2-56)$$

the product of two dimensionless factors, one enormously large and one enormously small; we estimate them separately below. But how much can the calculated log-odds (2-17) vary due to noise? For reliable discrimination the systematic part (2-56) of the log-odds must be large compared to its random variability. In (2-17) we saw that the noise contributes a random confusion term to the log-odds of $\sigma^{-2} n \cdot (y_A - y_B)$, and from (2-55) we can estimate this as

$$\pm \left[\frac{\int [y_A(t) - y_B(t)]^2 dt}{\sigma^2}\right]^{1/2} \qquad (2-57)$$

But this integral is just the ‘ $\int\int fgf$ ’ that we have evaluated in $(2-31)$ and $(2-35)$:

$$\int [y_A(t) - y_B(t)]^2 \; dt = \int d\omega \; |F(\omega)|^2 \; G(\omega) \simeq |R_A(\omega_o) - R_B(\omega_o)|^2 \int f^2(t) dt \qquad (2-58)$$

and we have yet another pleasant surprise: the square of (2-57) is just twice the expected log-odds (2-36).

Therefore our final conclusion for this "baby" version of the problem can be stated very simply: given the echo functions $y_A(t)$ and $y_B(t)$ for the two possible targets and the data $d(t)$ obtained by the receiver from a pulse echo, calculate the dimensionless number

$$L_A \equiv \frac{[d \cdot (y_A - y_B) + \frac{1}{2}(y_B \cdot y_B - y_A \cdot y_A)]}{kT_N} \qquad (2-59)$$

This is the log odds in favor of target A given by a single pulse. The mean value, or "expected value" of L_A is given by (2-18), (2-54) as

$$M = \langle L_A \rangle = \frac{(y_A - y_B) \cdot (y_A - y_B)}{kT_N} \; . \qquad (2-60)$$

Different pulses, with randomly varying samples of noise, will yield varying conclusions given approximately by

$$\log o(A|D\sigma) \simeq M \pm \sqrt{2M}. \qquad (2-61)$$

Thus if $M > 10$ the targets can be distinguished quite reliably. We could hardly have hoped for an easier prescription. Note that (2-59) and (2-60) are exact; they do not have the approximation made in (2-36) and (2-56) which supposed that the transmitted pulse spectrum is sharply peaked at a frequency where $G(\omega)$ is not rapidly varying.

Numerical Estimates:
It remains to estimate the numerical values that we might hope for in a real situation. For example, if the transmitter radiates one Megw for one microsecond and the receiver has a noise temperature of 1000K, the energy ratio in (2-56) is about

$$\frac{1 \text{ joule}}{1.36 \cdot 10^{-23} \cdot 1000 \text{ joules}} = 0.75 \times 10^{22}. \qquad (2-62)$$

Then to achieve reliable discrimination between any two targets A and B, the reflection function factor in (2-56) must be large compared to 10^{-22}.

To estimate magnitudes for this small factor, we need reasonable guesses for our antenna gains and the scattering cross-section of a target. Suppose our transmitter radiates the total power P_{rad} watts. The antenna concentrates the energy as much as possible in the direction of the target, so the power density incident on the target at distance d is

$$P_{inc} = G\,\frac{P_{rad}}{4\pi d^2} \text{ watts/m}^2. \qquad (2-63)$$

where G is the antenna gain, relative to an isotropic radiator. It can be estimated two ways, from the beam width or, using the reciprocity theorem, from its absorption cross-section. We illustrate both methods.

Suppose our antenna is a parabolic dish of diameter 2a, operating at a wavelength λ. Its beam width is, crudely, $\delta\theta \simeq \lambda/2a$, so its main beam fills in space a solid angle of about $\Omega \simeq \pi(\delta\theta/2)^2$. Thus we estimate its gain as

$$G = \frac{4\pi}{\Omega} \simeq \left(\frac{4a}{\lambda}\right)^2. \qquad (2-64)$$

On the other hand, consider its absorption properties. An infinitesimal dipole has an absorption cross section of $3\lambda^2/8\pi$; i.e. the maximum power that it can extract from a passing plane wave is the power incident on this area. But this has a gain of 3/2 because of the slight concentration of fields in the dipole's equatorial plane (the average of $\sin^2\theta$ over a sphere is 2/3). Therefore the hypothetical but nonexistent isotropic radiator would have an absorption cross-section of $\lambda^2/4\pi$. Now the absorption cross-section of our dish antenna is about equal to its area, πa^2 (actually, slightly less because the dish is not uniformly illuminated by the feeder), and so we estimate the gain as

$$G \simeq \frac{\pi a^2}{\lambda^2/4\pi} = \left(\frac{2\pi a}{\lambda}\right)^2 \qquad (2-65)$$

in approximate agreement with (2-64). For example, for an 18 inch dish at X band ($\lambda = 3$ cm) we estimate a gain of about $G \simeq 2000$.

To get crude estimates of scattering cross-sections, suppose that our target is a perfectly conducting sphere of radius r, large compared to λ so that we can use geometrical optics. Consider the radiation of density P_{inc} incident on a small area A of the spherical surface. This area fills a solid angle, as seen from the center of the sphere, of $\Omega = A/r^2$. But it is reflected back at twice the angle of incidence, thus going into a solid angle 4Ω. Thus the reflected energy appears at a distance d from the sphere with a density

$$P_{refl} = \frac{P_{inc}A}{4\Omega d^2} = P_{inc}\,\frac{\pi r^2}{4\pi d^2}, \qquad (2-66)$$

confirming our intuitive feeling that in the geometrical optics limit the back scattering cross-section Σ of a perfectly conducting sphere should be just its projected shadow area: $\Sigma = \pi r^2$. Indeed, in this limit the back scattering cross-section of a perfectly conducting object of any shape, integrated over all angles, should be its shadow area, because that is intuitively the amount of energy it intercepts [However, this intuition fails in the exact forward direction, because of some subtleties about creeping waves, the Arago bright spot, etc. which do not concern us here].

Then in practice, we expect that the strongest echoes from a metallic target will come from that part of its surface which presents a perpendicular aspect to the radar system, and has the greatest radius of curvature. If that flattest perpendicular surface has principal radii of curvature r_1, r_2, then we estimate the back scattering cross-section from it to be

$$\Sigma \simeq \pi r_1 r_2. \qquad (2-67)$$

If there is more than one such surface, their echoes will interfere, varying the net backward cross-section in a way critically dependent on aspect angle.

For a small airplane the single-surface cross-section (2-67) might be, conceivably, less than one square meter; perhaps 2 or 3 square meters is a reasonable average guess. Of course, at much lower frequencies, where the geometric optics approximation does not hold and the wing dipole resonance appears, the back scattering cross-section can be much greater than this, of the order of the aforementioned $3\lambda^2/8\pi$. If the wing dipole resonance of a large airplane is at 6 MHz, this would lead to $\Sigma \simeq 300\ m^2$.

Now combining Equations (2-63) - (2-67), we estimate the reflected energy density back at the radar system to be

$$P_{refl} \simeq P_{rad} \cdot \frac{G_t}{4\pi d^2} \cdot \frac{\Sigma}{4\pi d^2} \quad \text{watts/m}^2 \qquad (2-68)$$

where G_t is the gain of the transmitting antenna. The power intercepted by the receiver antenna will be P_{refl} times its absorption cross-section, which is by (2-65), $A_r = G_r \lambda^2/4\pi$. Finally, the power delivered to the receiver is, in terms of antenna absorption cross-sections,

$$P_{rec} \simeq P_{trans} \cdot \frac{\Sigma}{4\pi\lambda^2} \cdot \frac{A_t A_r}{d^4}, \qquad (2-69)$$

which is separated into two dimensionless factors, one depending on the target, the other on the radar antenna design. We compare this with our previous theoretical results. From the definition (1-1) of our reflection functions, we have

$$\frac{\text{(Energy received)}}{\text{(Energy transmitted)}} = \frac{\int d\omega\ |F(\omega)|^2\ |R(\omega)|^2}{\int d\omega\ |F(\omega)|^2} \qquad (2-70)$$

Therefore, if the transmitted energy spectrum is concentrated near ω_o, we have the estimate

$$|R(\omega_o)|^2 \simeq \frac{\Sigma}{4\pi\lambda^2} \cdot \frac{A_t A_r}{d^4}. \qquad (2-71)$$

in which $\lambda = 2\pi c/\omega_o$.

For example, if $\Sigma = 3\ m^2$, $\lambda = 10$ cm, $A_t = A_r = 1\ m^2$, d = 10 km, then (2-71) is about

$$\frac{3 \times 10^4}{4\pi \times 100} \cdot \frac{1}{10^{16}} = 2 \times 10^{-15}. \qquad (2-72)$$

The aforementioned receiver with noise temperature of 1000 K has in a bandwidth 1 MHz an effective input noise power of

$$kT\Delta f = 1.36 \times 10^{-23} \times 1000 \times 10^{6} = 1.36 \times 10^{-14} \text{ watts}, \qquad (2-73)$$

so if the transmitter radiates 1 Megw, we estimate that the echo can be detected with a signal/noise ratio

$$S/N = \frac{10^{6} \times 2 \times 10^{-15}}{1.36 \times 10^{-14}} = 1.5 \times 10^{5} \qquad (2-74)$$

or about 52 db, about as good as an audio cassette tape recording. This means that small differences in the echo from different targets should be easily detectable, as far as noise is concerned. The problem is with the information aspect; we need to know in advance what difference to look for.

3. GENERALIZATION TO ASPECT ANGLE

Eq. (2–59) represents the solution to the data processing problem which takes full account of the noise, but applies only to the special case where the echo from a target is always the same function $y(t)$ and there is no background hash interference $h(t)$. Before we have a useful solution for real problems, we need to make allowance for three complicating features. The echo function always depends on at least two parameters, the target range and aspect angle; and our signal will always be contaminated with hash (ground returns from fixed nearby objects).

We shall consider the hash problem relatively trivial, because we can always see some returns, which we know are pure hash, when no target is in the beam. Therefore the hash, for a given orientation direction of the antenna, can be known very accurately, and it is rather clear how to make allowance for it; just subtract the hash $h(t)$ from the data $d(t)$.

Indeed, when any complicating feature is known very accurately, then probability theory will tell us simply to adjust the data by subtracting off its effect (or dividing it out, etc.) so as to take it into account; and then to proceed as if the complication were not present. This has seemed intuitively obvious to most people without any theoretical analysis (such as when economists do detrending or seasonal adjustment on their data before analyzing for other effects), although we do not think that anyone has been able to see intuitively the exact conditions under which this "data fudging" rule is valid, much less what to do when it is not.

It is when a complication is not known accurately that new difficulties of principle arise, and we need to re-examine from the start what probability theory has to say about the problem; what is the optimal way to make allowance for its possible disturbing effects, while still extracting from the data all the information possible bearing on the question of interest?

First let us look at the "new complications" problem in a very general way, to see how probability theory supports the above statements. Suppose that the reflection function $r(t)$ from target A depends on some additional parameter α, so that the received echo function $y(t)$ depends on it. Thus when A is the true target present,in place of (1–1) we have

$$y_A(t,\alpha) = \int r_A(t-s;\alpha)\, f(s)\, ds \qquad (3\text{–}1a)$$

and the probability of getting a data set $D = \{d(t)\}$ becomes, in place of (2–4a)

$$p(D|A\alpha\sigma) \propto \exp\Big\{ -\frac{1}{2\sigma^2} \int [d(t) - y_A(t,\alpha)]^2\, dt \Big\} \qquad (3\text{–}2a)$$

Likewise, target B has another parameter β, and we have

$$y_B(t,\beta) = \int r_B(t-s;\beta)\, f(s)\, ds \tag{3–1b}$$

$$p(D|B\beta\sigma) \propto \exp\Big\{-\frac{1}{2\sigma^2}\int [d(t) - y_B(t,\beta)]^2\, dt\Big\} \tag{3–2b}$$

But how do we deal with the fact that α and β are unknown?

There are two different ways of organizing the probability calculation to answer this. First, note that the basic rule (2–7) is still valid without change:

$$p(A|D\sigma) = p(A|\sigma)\,\frac{p(D|A\sigma)}{p(D|\sigma)} \tag{3–3}$$

But now the sampling probability that we have, $p(D|A\alpha\sigma)$ contains α and the sampling probability that we want, $p(D|A\sigma)$, does not; and similarly for target B. To get from one to the other, apply the sum rule and then the product rule:

$$p(D|A\sigma) = \int p(D\alpha|A\sigma)\, d\alpha = \int p(D|A\alpha\sigma)\, p(\alpha|A\sigma)\, d\alpha \tag{3–4}$$

This is a weighted average of all possible values of $p(D|A\alpha\sigma)$, weighted according to the prior probability $p(\alpha|A\sigma)$.

Therefore the odds ratio for comparing target A with target B still takes the form (2–12):

$$\frac{p(A|D\sigma)}{p(B|D\sigma)} = \frac{p(A|\sigma)}{p(B|\sigma)}\,\frac{\int p(D|A\alpha\sigma)\, p(\alpha|A\sigma)\, d\alpha}{\int p(A|B\beta\sigma)\, p(\beta|B\beta)\, d\beta} \tag{3–5}$$

Thus probability theory tells us, very sensibly, that if α is unknown, then the best we can do is to "hedge our bets" by making allowance for all possible values that it might have, taking into account any information about how likely the different possible values are.

The calculation could be organized differently by applying the sum rule and product rule directly to the final probability $f(A|D\sigma)$:

$$p(A|D\sigma) = \int p(A\alpha|D\sigma)\, d\alpha = \int p(A|D\alpha\sigma)\, p(\alpha|D\sigma) \tag{3–6}$$

which is a weighted average, now using probabilities of α conditional on the data. Then we apply the rule (3–3) with a different choice of propositions:

$$p(A|D\alpha\sigma) = p(A|\alpha\sigma)\,\frac{p(D|A\alpha\sigma)}{p(D|\alpha\sigma)} \tag{3–7}$$

Of course, the calculation via (3–3) – (3–5) is entirely equivalent to the one using (3–6), (3–7), and we are free to choose whichever one is more convenient computationally. But let us view this another way. Suppose our aim were to estimate α from returns known to originate from target A. Then probability theory would tell us to do the calculation

$$p(\alpha|DA\sigma) = p(\alpha|A\sigma)\,\frac{p(D|A\alpha\sigma)}{p(D|A\sigma)} \tag{3–8}$$

Now in the right-hand side of (3–8) we recognize the integrand of (3–4). That integrand is just proportional to the probability density for α, given the data D. Therefore we recognize three cases:

I. The prior information alone (for example, information obtained from the returns of previous pulses) is enough to determine α quite accurately. Then in (3–4) the prior probability $p(\alpha|A\sigma)$ is not far from a delta function peaking at the indicated value $\hat{\alpha}$, and we should act as if we knew α. This is rather accurately the situation when α represents some property of the hash (in which case α and β are the same parameter).

II. The data D contain enough information to determine α accurately, even though the prior information does not. Then (3–8) is sharply peaked at the indicated value $\hat{\alpha}$, and most of the contribution from the integral (3–4) comes from the immediate neighborhood of this peak. This is the case if α is the target range R; which is very accurately known from the echo time even when we have no prior information about it.

III. The data and prior information are not sufficient to determine α very well. Then the integrand of (3–5) remains broad, and we have no choice but to use the full integral formula. In this case, failure to know α is almost sure to cause a deterioration in our ability to resolve targets. Therefore *it becomes crucially important that we make use of every bit of prior information about α that we can acquire*. This may be the case if α is the aspect angle of the target.

Of course, everything we have said about α applies equally well to β.

What Happened to the Poles? Note that the effect of poles in the singularity expansion of the scattering, although not explicitly visible in the above, has been taken into account automatically by probability theory – but in much greater generality than just poles. For if there is *any* feature of the likelihood $p(D|A\alpha\sigma)$ that does not depend on α, then that feature will come through the averaging over α in (3–4) unchanged. Then if this feature is different for target A and target B, it will be part of the information in the odds ratio (3–5) and in the final log–likelihood for A over B.

Indeed, if this α–independent feature is the *only* significant difference between target A and target B, then it will become automatically the only thing that is contributing to that log–likelihood; in that case the target identification will arise just from the difference in the poles; and from nothing else.

Thus our very different basic approach to the target identification problem has not in any way disregarded the perfectly valid argument that poles, being independent of aspect, may provide an important clue to identification. Rather, our analysis will complete that argument by showing in exactly what way pole information is to be used optimally in analyzing the data (i.e., what specific function of the pole positions is the one relevant to the identification), and by recognizing that in general other information might also be cogent for target identification, and it should of course be taken into account.

But we stress again that, if returned echoes depend on aspect, this does not mean that we should look only for aspect–independent features. On the contrary, prior information about aspect may become *necessary* for target identification.

CONCLUSION

The above analysis has indicated in a very general way the calculations that should be performed by a computer analyzing radar data, in order to achieve the maximum possible discrimination between different targets. Still to be done is to find more explicitly: (1) What are the actual reflection functions $r(t;\alpha)$ for various targets and aspect angles? In what detail must this information be stored in the computer in order to achieve near-optimal performance? (2) What prior information is available about aspect in real situations? Then one would be in a position to write the explicit computer programs which draw on the stored information and carry out the calculations indicated above.

It is not possible to predict, at present, exactly how well the resulting systems will perform, because this depends on information about details of the reflection functions (how much do they differ for different targets) that we do not have. However, from the way this theory has been derived directly from fundamentals, we can say confidently that the data processing indicated here will yield the best performance that it is possible to obtain from the information assumed. Therefore major efforts to obtain the reflection function information for the targets anticipated and wavelengths available are justified. Once that information is at hand, we would be in a position to predict the discrimination performance from the theory given here.

It is the writer's belief that, since the signal/noise considerations turned out to be quite favorable, very reliable target discrimination is possible in principle, using existing memory capacities and computing power. For its realization in practice, the present top priority job is to obtain the aforementioned reflection function information.