Information Theory and the Design of Radar Receivers*

PHILIP M. WOODWARD†

Summary—**The paper deals with the problem, frequently encountered in radar, of extracting simple numerical information from a noisy waveform. It is suggested that the only ideal way of doing this is to use the principle of inverse probability and convert the waveform into a probability distribution for the quantity sought. The method is applied to the problem of determining the time delay of a periodically modulated rf waveform in the presence of white Gaussian noise when the undelayed waveform without noise is exactly known. As a result, the matched predetection filter of Van Vleck and Middleton is automatically specified, and the theory of ideal detection is briefly indicated.**

I. Introduction

THE OBJECT of this paper is to outline, in terms of a somewhat idealized example, a mathematical method by which a theoretically ideal radar receiver may always be specified in principle. It was for some time customary to regard signal-to-noise ratio as an all-important quantity in receiver design. Efforts were made to ensure that as high a ratio as ideally possible was obtained at the output. This seems now to be a mistaken philosophy, since signal-to-noise ratio does not measure information, and is something which can often be artificially enhanced by passing the waveform through a nonlinear device which does not alter the information content at all. The present method deals, not with signal-to-noise ratio, nor even with quantities of information, but with the information itself.

In radar, we have to answer such questions as whether a target is present or absent, what its range is, whether it is moving, and so on. If we attempt to design a receiver which would answer any or all of such questions exactly, we are attempting the impossible, because of the noise which must inevitably introduce false indications. But if we demand, on every occasion, an automatic assessment of the relative probabilities of all possible answers, we are being completely realistic and no receiving device can possibly do better. The present paper shows how this idea works out in one rather familiar problem—that of determining the time delay of a periodic waveform of known shape and amplitude. This amounts, in radar, to determining the range of a stationary target known to be present and giving an echo of known strength, and is obviously an artificial problem. But it suffices to illustrate the method, and is not altogether without practical interest. The quantity of range information latent in such a radar waveform has been evaluated elsewhere[1] without reference to any actual method of extracting it. Here the emphasis is on designing the ideal receiver rather than on evaluating its actual performance.

The whole of the present approach is based on the principle of inverse probability briefly summarized in the following section. There is nothing new in this principle, but in spite of the growing application of probability theory to such problems as arise in radar, as for instance by Kaplan and McFall,[2] there does not seem to have been a systematic attempt to apply inverse probability. The necessity for doing so becomes quite apparent, once the foundations of modern communication theory[3] have been studied. The present paper may perhaps help to "set the ball rolling," for the method can be applied to a very wide variety of radar problems.

II. The Principle of Inverse Probability and the Information Function

A "direct" probability describes the chance of an event happening on a given hypothesis, but if the event has actually happened and there are various hypotheses which would explain it, one is faced with a problem of inverse probability. Prior to the event, the various hypotheses may not have seemed equally probable, and such previous knowledge is expressed in terms of the *a priori* probabilities of the hypotheses. After the event, which will usually be an experimental observation specifically performed to test the hypotheses, their relative probabilities may become changed. The principle of inverse probability expresses the *a posteriori* probabilities in terms of the corresponding *a priori* ones, by utilizing the probabilities that the actual observation would have been obtained if each hypothesis in turn had been true. It is valid only when the hypotheses are mutually exclusive and exhaustive.

Let $H_1 H_2$, and so on, denote the hypotheses,

$P(H_n)$	the *a priori* probability of H_n
$P(Ob\|H_n)$	the probability of the observation if H_n were true
$P(H_n\|Ob)$	the *a posteriori* probability of H_n after the observation is known.

The theorem[4] may then be written briefly in the form

* Decimal classification: R537.13. Original manuscript received by the Institute, June 23, 1950; revised manuscript received, March 9, 1951. The present paper is a revised version of the paper originally communicated to the Institute by G. G. Macfarlane, under the title "The decoding of radar information."

† Telecommunications Research Establishment, Ministry of Supply, T.R.E., Great Malvern, England.

[1] P. M. Woodward and I. L. Davies, "A theory of radar information," *Phil. Mag.*, ser. 7, vol. 41, p. 1001; October, 1950.

[2] S. M. Kaplan and R. W. McFall, "The statistical properties of noise applied to radar range performance," PROC. I.R.E., vol. 39, pp. 56–60; January, 1951.

[3] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. Jour.*, vol. 27, pp. 379, 623; July and October, 1948.

[4] Harold Jeffreys, "Theory of Probability," chap. I; Oxford University Press; 1939.

$$P(H_n \mid Ob) \propto P(H_n)P(Ob \mid H_n). \qquad (1)$$

In the present application, the hypotheses are all the possible delay times of a given periodic waveform, and the "observation" is simply the given waveform delayed an unknown amount, and with white Gaussian noise added to it. We shall call this the "received waveform," but it is not the output from a receiver because this would prejudge the whole question. It is the waveform as it enters the receiving system, including any noise which the receiver itself may subsequently introduce. The *a priori* probabilities of the hypotheses will form a continuous probability density distribution for the unknown delay time τ, and for simplicity this distribution will be taken uniform over an interval equal to one period of the waveform. In other words, we take all inherently unambiguous delay times to be equally likely *a priori*, though any other prior knowledge can subsequently be inserted in the theory.

The best that any receiver can do is to form the *a posteriori* distribution of probability for τ, from the received waveform. This distribution is the actual information sought, and is most conveniently handled in logarithmic form. We shall call its logarithm the "information function" and denote it by $Q(\tau)$. Equation (1) may now be written

$$Q(\tau) = \log P(Ob \mid \tau) + \text{constant}, \qquad (2)$$

where $P(Ob|\tau)$ denotes the probability density for the received waveform on the hypothesis τ. The constant term is merely the logarithm of the normalizing factor for the *a posteriori* distribution. It serves no useful purpose and is in the future omitted.

III. Evaluation of the Information Function

Let us consider a particular occasion when the true value of the delay time τ happens to be τ_0, and write the received waveform in terms of real functions as

$$Y(t) = G(t - \tau_0) + I(t). \qquad (3)$$

Here $G(t)$ is the rf waveform which would have been received in the absence of noise or of any time delay, and is assumed known *a priori*. It is also assumed that $G(t)$ is periodically modulated, and although it is convenient to use the language of pulses, the theory is, in fact, valid for any periodic modulation whatever, including frequency modulation. The function $I(t)$ represents added noise. The probability distribution for the magnitude of $I(t)$ at any particular time t is assumed Gaussian, but it is necessary here to generalize this concept. It can be shown, either by resorting to sampling-point analysis[5] or by means of a statistical mechanical argument, that the probability density for the whole waveform $I(t)$, in an appropriate number of dimensions, is proportional to

$$\exp\left[-\frac{1}{N_0}\int I(t)^2 dt\right], \qquad (4)$$

where N_0 is the mean noise power per unit bandwidth.[1] The observer has access to $Y(t)$ but not to τ_0 directly, and must therefore try out all possible values of τ in turn. On the hypothesis τ, he can argue that the noise waveform alone would have to be $Y(t)-G(t-\tau)$, for which the probability density is proportional to

$$\exp\left\{-\frac{1}{N_0}\int [Y(t) - G(t-\tau)]^2 dt\right\}. \qquad (5)$$

Consequently, by (2), we have

$$Q(\tau) = -\frac{1}{N_0}\int [Y(t) - G(t-\tau)]^2 dt. \qquad (6)$$

The information function is thus proportional to the integrated square of the departure of the received waveform from a hypothetical noise-free waveform of lag τ. As the hypothetical τ is varied, the value which gives least-mean-square departure from the received waveform produces a maximum in the information function, and this, from the observer's point of view, is the most probable value of τ.

The limits of integration in (6) are chosen to correspond with whatever portion of the received waveform is being examined, and it is necessary to take this to be a whole number of repetition periods of the modulation. If the integrand be expanded into three terms, it will be found that the G^2 integral is independent of τ because of periodicity, and the Y^2 integral depends on τ_0 but not on τ, from (3). These two terms can consequently be omitted from $Q(\tau)$, since anything independent of τ may be absorbed into the normalization, which has already been omitted. We are left with

$$Q(\tau) = \frac{2}{N_0}\int Y(t)G(t-\tau)dt. \qquad (7)$$

The integrand may be said to exist for all values of t, but τ is confined within certain fixed *a priori* limits, say between 0 and the repetition period R of $G(t)$. The domain of the integrand may be represented diagrammatically (Fig. 1) as a strip of indefinite length in the t direction, and of width R in the τ direction. If it is required to

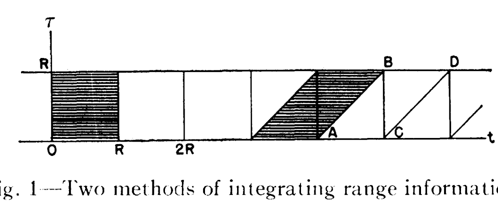

Fig. 1—Two methods of integrating range information.

form the information function $Q(\tau)$ to represent all the information contained in $Y(t)$ in the interval (O, R), then $Y(t)G(t-\tau)$ must be integrated with respect to t between the limits 0 and R for each value of τ as indicated by the shaded square. Further observation of the received waveform will entail further integrations in the obvious manner, and the limits of integration will always be represented by vertical lines on the diagram. The trouble with this process, from a practical point of view, is that all the accumulated information of any one in-

[5] C. E. Shannon, "Communication in the presence of noise," Proc. I.R.E., vol. 37, p. 10; January, 1949.

terval becomes available instantaneously at the end of the interval, and is followed by a lull while fresh integration is being performed.

Fortunately, there is an alternative method which is more natural, though mathematically less straightforward. By carrying out the integration, not in successive squares, but in parallelograms such as the one indicated on the diagram, the information function becomes available from $\tau=0$ to R (A to B in the diagram) progressively in time. The moment B has been reached, a fresh trace can be developed from C to D. In this way, we have set up a saw-tooth relationship, a time base in fact, connecting τ and t.

This process of "progressive" integration does not, of course, correspond to fixed limits on the integral (7), but to a pair of limits advancing together in time. In fact, the value of the information trace at time $t=nR+\tau$, where n specifies the nth trace and τ is confined between 0 and R, is given by

$$Q_n(\tau) = \frac{2}{N_0}\int_{t-R}^{t} Y(t)G(t-\tau)dt. \tag{8}$$

It is mathematically inconvenient that the information function Q is the logarithm of an *a posteriori* probability distribution for τ only when the limits of integration are constant. In other words, each hypothesis τ should, strictly, be tested out on the same piece of the received waveform. Space does not permit a full discussion of this point, and it must suffice to remark that the progressive probability distribution

$$P_n(\tau) = e^{Q_n(\tau)}, \tag{9}$$

normalization omitted, behaves for all practical purposes as though it were a strict *a posteriori* distribution. In particular, when the information from successive periods of the received waveform is combined, either by summing the Q_n over n or by multiplying together the P_n, the resulting distribution differs from a true *a posteriori* distribution only because of end-effects, which become progressively less and less important.

IV. The Ideal Predetection Filter

The progressive information function given by (8) happens to have a very simple electronic interpretation. The form of the expression, being a linear superposition, will be recognized as that of the output from a linear filter. In fact, it is the output at time $t=nR+\tau$ from a filter whose input is the received waveform $Y(t)$, and whose impulsive response is given by

$$\delta(t) \to \begin{cases} \dfrac{2}{N_0}G(-t), & 0 < t < R \\ 0, & t < 0 \text{ and } t > R. \end{cases} \tag{10}$$

Such a filter, apart from the special scaling factor $2/N_0$, has been discussed by Van Vleck and Middleton,[6] who show that it is the unique linear filter which gives maximum peak signal-to-noise performance. (This property is, however, irrelevant to the present theory.) The filter has a frequency response which, in amplitude, has the same shape as the amplitude spectrum of one period of the input signal G, but which in phase is equal and opposite to that of G.

The output from the filter is, of course, a modulated radio-frequency waveform, and when passed through an exponential rectifying device in accordance with (9), it becomes the progressive *a posteriori* probability distribution for τ. It has already been pointed out that several traces of Q may be added together, before such rectification, if the information from several periods of the received waveform is to be combined. This is simply phase-coherent pulse-to-pulse summation, and must not be performed unless τ is completely independent of time, as has so far been assumed.

The underlying effect of this ideal filter is to cause the output signal peak to look exactly like a particularly large noise peak; all the pattern information, which originally distinguished signal from noise, has been extracted from the waveform and converted into amplitude discrimination. This may seem surprising in view of the fact that the signal and noise outputs have different power spectra. Or again, if the input signal is a square pulse, the noise output can be regarded as a multitude of overlapping square pulses, while the signal output will be a triangular pulse. However, the fact remains that a multitude of square pulses overlapping at microscopic intervals to form Gaussian noise provides a background against which the pattern of a single triangular pulse cannot be distinguished. If, indeed, pattern information could still be utilized in the filter output, one is led to a *reductio ad absurdum*, since it has been shown that the most probable value of τ is given by selecting, regardless of pattern, the largest amplitude in the filter output.

V. Discussion

The *a posteriori* probability distribution for the delay time τ has been shown to take the form of a modulated rf waveform, obtained at the output from a linear filter, distorted in amplitude or "rectified," by means of a device having an exponential characteristic. This naturally results in a function of τ whose envelope, if the signal is large enough, is peaked near the true value τ_0, but which contains under its envelope a multitude of fine peaks produced by the carrier. This fine structure represents a succession of probable and improbable values of τ resulting from comparison of the carrier phase in $Y(t)$ and $G(t)$. When this highly ambiguous knowledge of range is of no interest, it may be removed by smoothing or "detecting" the *a posteriori* distribution in such a way that areas over intervals of an rf cycle are preserved. In fact, when τ changes with time sufficiently rapidly to render the rf information out of date from one trace to the next, but not rapidly enough to affect the modulation appreciably, successive

[6] J. H. Van Vleck and D. Middleton, "A theoretical comparison of the visual, aural and meter reception of pulsed signals in the presence of noise," *Jour. Appl. Phys.*, vol. 17, p. 940; November, 1946.

traces may be combined only after removal of the rf. The detected *a posteriori* distributions must then be multiplied together, or alternatively their logarithms must be added. Mathematical readers will see immediately that, in this way, an ideal detection characteristic (of the form log I_0, where I_0 is the modified Bessel function) is uniquely specified by the theory when post-detection pulse-to-pulse summation is to be performed.

Without, for the present, developing any further the theory of removing any of the idealizations, it should be clear that any problem of extracting all the information from a noisy waveform, can, in principle, be solved uniquely and ideally by one universal method. One simply has to state the question, write down the *a posteriori* probability distribution for all possible answers to it, and interpret the resulting formula in terms of a physical device, on the principle that anything which can be computed mathematically, can also be computed electronically. No problem of waveform decoding then remains, for the *a posteriori* distribution *is* the required information; anything further is pure guesswork.

ACKNOWLEDGMENT

Acknowledgment is made to the Chief Scientist, British Ministry of Supply, for permission to publish this paper (British Crown copyright reserved). The paper is reproduced with the permission of the Controller, H.B.M. Stationery Office, England.

Notes on an Automatic Radio-Frequency Repeater System*

JAMES A. CRAIG†

This paper is published with the approval of the IRE Professional Group on Vehicular Communications, and has been secured through the co-operation of that Group.—*The Editor.*

***Summary*—The paper describes the basic principles involved in an operating system of rf repeaters, and discusses the planning and installation of this system in Cuba. It mentions the types of antennas used and briefly describes the physical layout of the equipment involved. Also included is a resumé of the difficulties that arose and how they were overcome.**

ONE OF THE principal radio-broadcasting chains indicated about 2 years ago their need for some form of network transmission system that would enable them to overcome the inadequacies of the available Cuban telephone lines. The service provided by these lines was undependable and, when available, was noisy and wholly unsuitable from a program-quality standpoint.

The plan involved the use of radio repeaters to relay programs originating in Santiago studios near the Eastern end of Cuba and to service a network of AM broadcasting stations extending some 500 miles west to Havana. At an average 50-mile distance between repeaters this would necessitate at least ten repetitions of the program material.

To use the conventional form of repeater would have been practically impossible since demodulation and remodulation at each repeater point would have introduced a prohibitive amount of noise and distortion due to nonlinearity in the detection and modulation systems. This factor alone usually dictates a maximum of 5 or 6 repeaters even for voice-communication circuits, and the requirement in this case was for program quality. The common form of repeater, when stripped of embellishments, remains essentially a radio receiver and a radio transmitter operating back-to-back.

To overcome the foregoing limitations a radio-frequency repeater system was developed. In this repeater, as the name implies, rf is used exclusively. Frequencies between 150 and 180 mc can be accommodated by this particular system although the principles are being readily applied to the other commonly used vhf and uhf bands. The actual frequency range used in the Cuban system is from 163 to 170.2 mc.

From the functional block diagram in Fig. 1 it is seen that a typical operating frequency of 165.0 mc has been chosen as the incoming carrier for the purpose of illustration. This frequency is fed from an antenna to a single stage rf amplifier. From this amplifier it is heterodyned with a locally generated signal originating in a "channel" oscillator. This oscillator is controlled by a quartz crystal operating as a harmonic oscillator at a frequency of approximately 45 mc. The crystal frequency is tripled and then amplified at 137.5 mc, as chosen in the block diagram, and is then fed to the first mixer stage. This crystal is the only one that requires changing when a different carrier frequency is to be used. The beat frequency between the incoming signal and the local oscillator is always 27.5 mc for any incoming carrier, and this frequency enters the first intermediate frequency amplifier consisting of two high-gain stages.

As the signal leaves the first IF amplifier, the user can elect to retransmit on a frequency which is either

* Decimal classification: R480. Original manuscript received by the Institute, March 22, 1951; revised manuscript received May 22, 1951. Presented, National Conference of the IRE Professional Group on Vehicular Communications, Detroit, Michigan, November 3, 1950.

† Link Radio Corp., 125 W. 17th St., New York 11, N. Y.