Epoch Detection — A Method For Resolving Overlapping Signals

By TZAY Y. YOUNG

(Manuscript received October 20, 1964)

The purpose of this paper is to discuss an epoch detection procedure which is very useful for the resolution and detection of signals overlapping in time. An epoch is the beginning instant of a signal. The epoch detection procedure is based on the following hypotheses: On the null hypothesis H_0 that a certain instant t is not an epoch, analytical continuation exists at t, and one may predict the signal in the future based on past experience or vice versa. On the hypothesis H_1 that t is an epoch, the analytic continuation is disrupted at t.

Based on this idea and the assumption of a Gaussian noise, a test statistic is derived from the maximum likelihood principle. The test statistic may be obtained at the output terminal of a linear filter. The performance of such a system is considered. Also discussed briefly are the cases of overlapping stochastic signals and overlapping radar signals. Some experimental results obtained from a digital computer are shown.

I. INTRODUCTION

Consider a signal composed of a train of overlapping wavelets.* The wavelets may, for one reason or another, arrive at the receiver (or measuring apparatus) delayed by different amounts of time. The time delays of the individual wavelets are unknown, but their differences may be relatively small so that the wavelets overlap. The beginning instant of each wavelet is called an epoch. These signals are corrupted with Gaussian noise. Our problem is to detect the overlapping in time. In other words, we wish to design a practical system which enables us to resolve the received signal train into overlapping wavelets and to describe them individually.

The theory of statistical detection of signals buried in noise has been well established.[1-4] In the field of resolving overlapping wavelets, Hel-

* We use the word "wavelets" for the individual overlapping wavelets, and reserve the word "signal" for the over-all signal train.

strom[5] discussed the optimum detection of two overlapping wavelets. With his assumption that one wavelet is separated from the other by known amount of time, the problem is considerably simplified and relatively easy to handle. Nilsson[6] discussed the problem of resolving N overlapping wavelets by deriving an equation to be maximized in an N-dimensional parameter space. Even in the case $N = 2$, the maximization of this equation is very complex and practically unsolved. Root[7] considered the general resolvability of radar signals, but gave no decision rule. Other studies related to signal resolution place most emphasis on the study of ambiguity functions[8,9] and on the design of a radar waveform which is inherently suitable for signal resolution.[10]

Generally speaking, for N overlapping wavelets, an optimum detection procedure would always involve searching for the maximum value of a likelihood function in an N-dimensional parameter space.[6] For N large this is hardly practical, and furthermore, if the number N is unknown, the problem becomes even more complicated. In a recent memorandum,[11] the author suggested an epoch detection procedure based on the properties of the signal at the epochs. The basic idea was to use a portion of the received signal in the past to predict the signal in the future, and to announce the arrival of a new wavelet if the prediction failed sufficiently badly. The present paper originated from that work. We intend to formalize and to develop the principle of epoch detection.

Consider a signal $f(t)$ consisting of two overlapping wavelets as shown in Fig. 1. The function $f(t)$ is analytic everywhere except at the two epochs t_1 and t_2. For any instant t which is not an epoch, it is possible to use the signal immediately prior to t to predict the signal immediately after. This is indeed the property of analytic continuation. However, at the two epochs, the statement is no longer true. Indeed we may define an epoch as an *instant* at which *analytic continuation is disrupted.*

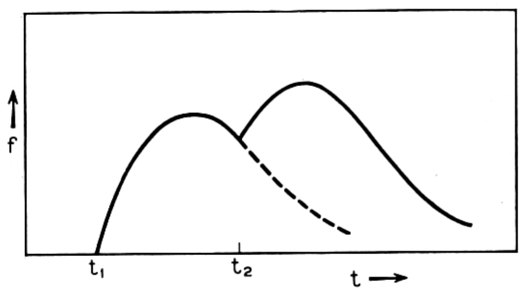

Fig. 1 — Overlapping signals.

It is precisely this disruption of analytic continuation that enables us to detect the epochs. In practice, we shall use the signal representation technique to describe such disruptions.

We make the assumption that any wavelets, though close enough to cause overlapping, are separated by at least T seconds, i.e.,

$$|t_j - t_k| \geqq T, \tag{1}$$

where t_j, t_k are any arbitrary epochs and T is a predetermined quantity. This assumption is necessary for our formulation, since for any instant t we shall utilize the information in the time interval $(t - T, t + T)$ to determine whether t is likely to be an epoch. We further assume that *any* $2T$-second segment of the individual *wavelets* is representable by a set of known component functions. Then the disruption of analytical continuation simply means that if an epoch exists in a certain $2T$ interval the *signal* in that interval is no longer representable by the set of component functions. Likelihood functions may be formulated in accordance with these criteria. The instant $\hat{t}$ which corresponds to a maximum value of a likelihood ratio is then the estimate of the epoch. This is of course the well-known maximum likelihood method of signal extraction, which has some theoretical advantages.[12,13] Other parameters of the wavelet may be estimated simultaneously.

Using the epoch detection scheme, we have in fact reduced an N-dimensional problem to N one-dimensional problems. Undoubtedly, in a process such as this, some information is lost, and one cannot expect optimum signal resolution except for some extreme cases. However, the simplicity and the practicality of the process justify our investigation. The process should be especially useful in the case of strong signals for which the advantage of a simple system outweighs that of optimality. In addition, the concept of epoch detection deserves to be studied and developed on its own right.

II. STATISTICAL EPOCH DETECTION

Let us denote the deterministic signal by $f_s(t)$, the random Gaussian noise by $f_n(t)$, and the noisy signal by $f_{s+n}(t)$. In this section, we shall consider the case that the deterministic signal consists of N overlapping wavelets with each wavelet being of the same waveform. Then we may write

$$\begin{aligned} f_{s+n}(t) &= f_s(t) + f_n(t) \\ &= \sum_{k=1}^{N} A_k f_w(t - t_k) + f_n(t), \end{aligned} \tag{2}$$

where A_k and t_k are the amplitude and the true epoch of the kth wavelet respectively. The function $f_w(\tau)$ represents the waveform of the individual wavelets.

To begin with, let us assume that each wavelet is representable by a set of known component functions. This assumption presents no theoretical difficulty since, by using a set of component functions that constitute a complete set, one may represent any continuous signals to any degree of accuracy.[14] However, practical considerations limit us to use a set of a finite number of component functions. We are particularly interested in the classes of component functions known as *generalized exponentials*, which include real and complex exponentials, sinusoids, polynomials and possible sums of products of such functions. The generalized exponentials have the following important property. A finite and properly-chosen set of generalized exponentials, as a set, goes into itself under the translation of time.[15] As a result, if a wavelet $f_w(\tau)$ is exactly representable by a properly chosen set of m generalized exponentials $\varphi^{(i)}(\tau)$, $i = 1, 2 \cdots m$, i.e.,

$$f_w(\tau) = \begin{cases} 0, & -\infty \leqq \tau < 0, \\ \sum_{i=1}^{m} c_i(0)\varphi^{(i)}(\tau), & 0 \leqq \tau \leqq \infty, \end{cases} \tag{3}$$

then the tail of $f_w(\tau)$ is also exactly representable by the same set of generalized exponentials,

$$f_w(t + \tau) = \sum_{i=1}^{m} c_i(t)\varphi^{(i)}(\tau), \qquad 0 \leqq t \leqq \infty, \quad 0 \leqq \tau \leqq \infty, \tag{4}$$

where $c_i(t)$ is the ith coefficient for a time translation of t seconds. Obviously, under this condition our earlier assumption that every $2T$ segment of the individual wavelets is representable by the set of component functions is fulfilled. The full significance of this property will be appreciated later, when we derive the test statistic for epoch detection.

The assumption of generalized exponentials is not as restrictive as it first appears. For one thing, most physical wavelets may be represented by a few terms of these functions. Furthermore, almost all commonly used functions for signal representation or curve fitting belong to the classes of generalized exponentials, and if we are willing to tolerate some inaccuracies by an approximate representation, practically all waveshapes may be represented by them. It is interesting to note that for the generalized exponentials $\varphi^{(i)}(\tau)$ analytic continuation exists everywhere except at $\tau = 0$. Consequently, the epoch of a wavelet $f_w(\tau)$ described by (3) satisfies our earlier definition of disruption of analytic continuation.

Next, consider the Gaussian noise $f_n(t)$ having a covariance function $R(\tau)$.[16] The covariance function may be taken as the kernel of an integral equation,

$$\int_0^{2T} R(t-\tau)\psi^{(j)}(\tau)d\tau = \lambda_j\psi^{(j)}(\tau). \tag{5}$$

For our problem, $R(\tau)$ is real and symmetric, the eigenvalues, λ_j, are positive, and the eigenfunctions, $\psi^{(j)}(\tau)$, are orthonormal real functions. Both deterministic and random signals may be expressed in terms of these eigenfunctions.[16,17] Thus, we may write

$$\begin{aligned} f_{s+n}(t+\tau) &= \sum_j v_j(t)\psi^{(j)}(\tau), \\ f_s(t+\tau) &= \sum_j s_j(t)\psi^{(j)}(\tau), \\ f_n(t+\tau) &= \sum_j n_j(t)\psi^{(j)}(\tau), \\ & \qquad 0 \leqq \tau \leqq 2T, \end{aligned} \tag{6}$$

and

$$\varphi^{(i)}(\tau) = \sum_j u_{ij}\psi^{(j)}(\tau), \qquad 0 \leqq \tau \leqq 2T, \tag{7}$$

with

$$\begin{aligned} v_j(t) &= \int_0^{2T} f_{s+n}(t+\tau)\psi^{(j)}(\tau)d\tau, \\ s_j(t) &= \int_0^{2T} f_s(t+\tau)\psi^{(j)}(\tau)d\tau, \\ n_j(t) &= \int_0^{2T} f_n(t+\tau)\psi^{(j)}(\tau)d\tau, \end{aligned} \tag{8}$$

and

$$u_{ij} = \int_0^{2T} \varphi^{(i)}(\tau)\psi^{(j)}(\tau)d\tau. \tag{9}$$

It is essential to note that by this expansion, the random variables n_j (and also v_j) are *independent variables* with variances λ_j. Since we are not interested in the singular case,[16] we assume that

$$\sum_{j=1}^{\infty} \frac{| s_j(t) |^2}{\lambda_j} < \infty, \qquad \sum_{j=1}^{\infty} \frac{| u_{ij} |^2}{\lambda_j} < \infty. \tag{10}$$

We are now in a position to derive the test statistic for epoch detection. Let us start with the simplest case.

2.1 *Known Wavelet at Known Epoch*

In this case, we assume that there are reasons to believe that a wavelet in the form of $A_k f_w(t - t_k)$ may arrive. Both A_k and t_k are assumed known. In assuming known t_k, it is also implied that no epoch other than t_k may appear in the time interval $(t_k - T, t_k + T)$. Let us define a function

$$f_h(\tau) = \begin{cases} f_w(\tau - T), & T \leqq \tau \leqq 2T, \\ 0, & \text{elsewhere.} \end{cases} \tag{11}$$

We wish to test the hypothesis H_1 that the wavelet arrives against the null hypothesis H_0 that it does not. Thus, we write

$$H_0 : f_s(t_k - T + \tau) = \sum_i b_i \varphi^{(i)}(\tau), \qquad 0 \leqq \tau \leqq 2T, \tag{12}$$

and

$$H_1 : f_s(t_k - T + \tau) = f_g(\tau) + \sum_i a_i \varphi^{(i)}(\tau), \qquad 0 \leqq \tau \leqq 2T, \tag{13}$$

with $f_g(\tau)$ defined as

$$f_g(\tau) \equiv A_k f_h(\tau) - A_k \sum_i r_i \varphi^{(i)}(\tau), \qquad 0 \leqq \tau \leqq 2T. \tag{14}$$

The constants r_i will be defined later. Let us explain these two hypotheses. In the first place, we notice that in using generalized exponentials as component functions, it is implied that $\varphi^{(i)}(\tau)$ and consequently $f_w(\tau)$ extends from $\tau = 0$ to $\tau = \infty$, as clearly indicated in (3). (The case of overlapping pulses will be treated later.) Therefore, on the null hypothesis H_0, although the new wavelet does not arrive, there will be tails of previously arrived wavelets appearing in the time interval $(t_k - T, t_k + T)$. Since every $2T$-second segment of these previously arrived wavelets is representable by the component functions $\varphi^{(i)}(\tau)$ with $0 \leqq \tau \leqq 2T$, we obtain (12) with the coefficients b_i to be estimated.

On the hypothesis H_1, the wavelet arrives at $t = t_k$. The term $A_k f_h(\tau)$

in (14) simply reflects this fact, since $f_h(\tau) = 0$ for $\tau < T$, as shown in (11). It is also this term that causes the disruption of analytic continuation at t_k. In addition to the kth wavelet, there are also tails of previously arrived wavelets, and we might have written for the hypothesis H_1, $f_s = A_k f_h + \sum q_i \varphi^{(i)}$. For reasons that will be pointed out later, we simply split q_i into two terms, $q_i = a_i - A_k r_i$, and obtain (13).

The random variables are independent when expressed in terms of eigenfunctions, and consequently we expand, similar to (6), $f_h(\tau)$ and $f_g(\tau)$ into

$$f_h(\tau) = \sum_j h_j \psi^{(j)}(\tau), \quad 0 \leqq \tau \leqq 2T,$$
$$f_g(\tau) = \sum_j g_j \psi^{(j)}(\tau), \quad 0 \leqq \tau \leqq 2T, \tag{15}$$

with

$$h_j = \int_0^{2T} f_h(\tau)\psi^{(j)}(\tau)\,d\tau,$$
$$g_j = A_k h_j - A_k \sum_i r_i u_{ij}. \tag{16}$$

The joint probability density for the null hypothesis may then be written as

$$P_0(v;b_i) = \frac{1}{\prod_j (2\pi\lambda_j)^{\frac{1}{2}}} \exp\left[-\sum_j \frac{(v_j - \sum_i b_i u_{ij})^2}{2\lambda_j}\right] \tag{17}$$

according to (6), (7), and (12). Similarly, we write for hypothesis H_1 the joint probability density

$$P_1(v;a_i) = \frac{1}{\prod_j (2\pi\lambda_j)^{\frac{1}{2}}} \exp\left[-\sum_j \frac{(v_j - \sum_i a_i u_{ij} - g_j)^2}{2\lambda_j}\right]. \tag{18}$$

In the absence of a priori information on the tails of previously arrived wavelets, a reasonable test is the maximum likelihood test which is given by

$$L = \frac{\max_a P_1(v;a)}{\max_b P_0(v;b)} \geqq \exp(\eta) \tag{19}$$

with the threshold η to be determined either by the Bayes criterion or by the Neyman-Pearson criterion. Equation (19) is equivalent to

$$\log L = \max_a \left[-\sum_j \frac{(v_j - \sum_i a_i u_{ij})^2 - 2(v_j - \sum_i a_i u_{ij}) g_j + g_j^2}{2\lambda_j} \right] - \max_b \left[-\sum_j \frac{(v_j - \sum_i b_i u_{ij})^2}{2\lambda_j} \right] \geqq \eta. \tag{20}$$

In order to simplify (20) somewhat, let us write

$$u_{ij}^* = \frac{u_{ij}}{\lambda_j}, \qquad \varphi^{(i)*}(\tau) = \sum_j u_{ij}^* \psi^{(j)}(\tau). \tag{21}$$

We assume that it is possible to write

$$\sum_j \frac{u_{lj} u_{ij}}{\lambda_j} = \delta_{li} \tag{22}$$

where δ_{li} is the Kronecker delta. Remembering the orthonormality of eigenfunctions, (22) may be written as

$$\begin{aligned} \sum_j \frac{u_{lj} u_{ij}}{\lambda_j} &= \sum_j u_{lj} u_{ij}^* \int_0^{2T} \psi^{(j)}(\tau) \psi^{(j)}(\tau)\, d\tau \\ &= \sum_{j,k} \int_0^{2T} u_{lk} \psi^{(k)}(\tau) u_{ij}^* \psi^{(j)}(\tau)\, d\tau \\ &= \int_0^{2T} \varphi^{(l)}(\tau) \varphi^{(i)*}(\tau)\, d\tau \\ &= \delta_{li}. \end{aligned} \tag{23}$$

Thus (22) is simply a consequence of the fact that $\varphi^{(l)}(\tau)$ and $\varphi^{(i)*}(\tau)$ form a biorthonormal system.[18] Furthermore,

$$\begin{aligned} \varphi^{(i)}(\mu) &= \sum_j u_{ij} \psi^{(j)}(\mu) \\ &= \sum_j \lambda_j u_{ij}^* \psi^{(j)}(\mu) \\ &= \sum_j u_{ij}^* \int_0^{2T} R(\mu - \tau) \psi^{(j)}(\tau)\, d\tau \\ &= \int_0^{2T} R(\mu - \tau) \varphi^{(i)*}(\tau)\, d\tau. \end{aligned} \tag{24}$$

Consequently, $\varphi^{(i)*}(\tau)$ is indeed the solution of an integral equation and may be obtained for given $\varphi^{(i)}(\mu)$ and $R(\mu - \tau)$.[19] Thus it is always possible to achieve the biorthonormalization of a set of known component functions by means of a process similar to the Gram-Schmidt process of orthonormalization. It is appropriate to point out here that the assumption of a biorthonormal system is solely for the purpose of mathematical simplicity.

Now let us define the constant r_i as

$$r_i \equiv \sum_j \frac{u_{ij}}{\lambda_j} h_j . \tag{25}$$

Then, according to (16), we obtain

$$\begin{aligned} \sum_j \frac{u_{ij} g_j}{\lambda_j} &= A_k r_i - A_k \sum_{j,l} \frac{u_{ij} r_l u_{lj}}{\lambda_j} \\ &= A_k r_i - A_k \sum_l r_l \delta_{il} \\ &= 0. \end{aligned} \tag{26}$$

In other words, $f_g(\tau)$ is orthogonal to $\varphi^{(i)*}(\tau)$. Returning to (20), we notice that because of (26), log L may be simplified into the form of

$$\begin{aligned} \log L = \sum_j \frac{v_j g_j}{\lambda_j} - \sum_j \frac{g_j^2}{2\lambda_j} + \max_a \left[- \sum_j \frac{(v_j - \sum_i a_i u_{ij})^2}{2\lambda_j} \right] \\ - \max_b \left[- \sum_j \frac{(v_j - \sum_i b_i u_{ij})^2}{2\lambda_j} \right]. \end{aligned} \tag{27}$$

However, the last two terms are indeed identical. Thus,

$$\log L = \sum_j \frac{v_j g_j}{\lambda_j} - \sum_j \frac{g_j^2}{2\lambda_j} . \tag{28}$$

The last term in (28) is only a constant, and we may use the statistic

$$G = \sum_j \frac{v_j g_j}{\lambda_j} \geqq \xi \tag{29}$$

for testing the arrival of the wavelet at the instant $t = t_k$. Here ξ is the threshold for testing G.

The results may also be expressed in the form of integral equations. Using a procedure similar to that used in (23) and (24), the statistic shown in (29) may be expressed as

$$G = \int_0^{2T} f_{s+n}(t_k - T + \tau) f_g^*(\tau)\, d\tau \geqq \xi \tag{30}$$

with $f_g^*(\tau)$ being the solution of the integral equation

$$f_g(\mu) = \int_0^{2T} R(\mu - \tau) f_g^*(\tau)\, d\tau. \tag{31}$$

The function $f_g(\mu)$ has been defined in (14) and is rewritten here.

$$f_g(\mu) = A_k f_h(\mu) - A_k \sum_i r_i \varphi^{(i)}(\mu). \tag{14}$$

The constants r_i, written in integral form, become

$$r_i = \int_0^{2T} f_h(\tau) \varphi^{(i)*}(\tau)\, d\tau \tag{32}$$

according to (25). For a white noise with a covariance function $\delta(\tau)$, the results are considerably simpler since in this case,

$$\begin{aligned} \varphi^{(i)*}(\tau) &= \varphi^{(i)}(\tau), \\ f_g^*(\tau) &= f_g(\tau). \end{aligned} \tag{33}$$

The test statistic G shown in (30) may be obtained by a linear filter. If we use a linear filter whose weighting function is characterized by $f_g^*(\tau)$ — or, in other words, if the impulse response of the filter is $f_g^*(-\tau)$ — then with $f_{s+n}(t)$ as input, the output of the filter gives us the desired statistic G with a time delay of T seconds.[14] As examples, we show in Fig. 2 some wavelets and the weighting functions of their corresponding "matched" filters for epoch detection in white noise. (See Appendix.)

The weighting functions shown in the figure are calculated according to (14). It is essential to note the difference between our "matched" filter and the standard matched filter for the detection of non-overlapping signals. Without interfering signals, the matched filter would be $f_w(\tau)$, while in our case, a term in the form of $\sum r_i \varphi^{(i)}(\tau)$ is to be subtracted from the original waveform, as clearly shown in (14). It is indeed the subtraction of this term that enables us to suppress the effect of previously arrived wavelets. It is also this subtraction that represents the price we pay.

We wish to compute the false alarm and detection probabilities for the epoch detection system which is based on the statistic G. Since G is obtained from a linear operation on a Gaussian-distributed variable, G is also Gaussian-distributed.[17] Under the hypothesis H_0, its mean value is

$$E[G \mid H_0] = \int_0^{2T} f_g^*(\tau) \sum_i b_i \varphi^{(i)}(\tau)\, d\tau = 0 \tag{34}$$

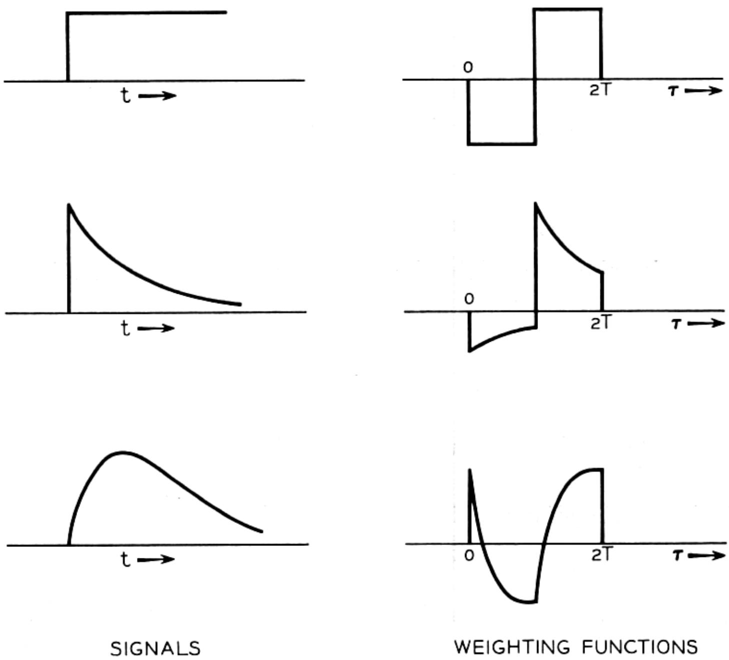

Fig. 2 — "Matched" filters for epoch detection in white noise.

since $f_g^*(\tau)$ and $\varphi^{(i)}(\tau)$ are orthogonal. Under the hypothesis H_1, the mean value becomes

$$\begin{aligned} E[G \mid H_1] &= \int_0^{2T} f_g^*(\tau)[f_g(\tau) + \sum_i a_i\varphi^{(i)}(\tau)]\,d\tau \\ &= \int_0^{2T} f_g^*(\tau)f_g(\tau)\,d\tau. \end{aligned} \tag{35}$$

The variance of G under either hypothesis is

$$\begin{aligned} \operatorname{Var} G &= \int_0^{2T}\int_0^{2T} f_g^*(\tau)f_g^*(\mu)\overline{f_n(\tau)f_n(\mu)}\,d\mu d\tau \\ &= \int_0^{2T}\int_0^{2T} f_g^*(\tau)f_g^*(\mu)R(\tau - \mu)\,d\mu d\tau \\ &= \int_0^{2T} f_g^*(\tau)f_g(\tau)\,d\tau, \end{aligned} \tag{36}$$

where we have used (31). Thus,

$$d^2 = \int_0^{2T} f_g^*(\tau) f_g(\tau) d\tau, \tag{37}$$

a *dimensionless constant*, plays the role of signal-to-noise ratio (SNR).[3,4] The probability density functions of G are

$$p_0(G) = (2\pi d^2)^{-\frac{1}{2}} \exp(-G^2/2d^2), \tag{38}$$
$$p_1(G) = (2\pi d^2)^{-\frac{1}{2}} \exp[-(G - d^2)^2/2d^2].$$

and the false alarm and detection probabilities are, respectively[3]

$$\begin{aligned} Q_0 &= \operatorname{erfc}(\xi/d), \\ Q_d &= \operatorname{erfc}\left(\frac{\xi}{d} - d\right), \end{aligned} \tag{39}$$

where erfc (x) is the error-function integral.

2.2 *Unknown Amplitude and Unknown Epoch*

In this case, the two hypotheses become

$$H_0 : f_s(t - T + \tau) = \sum_i b_i \varphi^{(i)}(\tau), \qquad 0 \leqq \tau \leqq 2T, \tag{40}$$

$$H_1 : f_s(t - T + \tau) = A(t) f_g(\tau) + \sum_i a_i \varphi^{(i)}(\tau), \qquad 0 \leqq \tau \leqq 2T, \tag{41}$$

where

$$f_g(\tau) \equiv f_h(\tau) - \sum_i r_i \varphi^{(i)}(\tau), \qquad 0 \leqq \tau \leqq 2T. \tag{42}$$

Notice the slight difference between the definition of $f_g(\tau)$ shown in (42) and that of (14). The joint probability densities are, similar to the previous case,

$$\begin{aligned} P_0(v;b_i) &= \frac{1}{\prod_j (2\pi\lambda_j)^{\frac{1}{2}}} \exp\left[- \sum_j \frac{(v_j - \sum_i b_i u_{ij})^2}{2\lambda_j} \right] \\ P_1(v;a_i, A) &= \frac{1}{\prod_j (2\pi\lambda_j)^{\frac{1}{2}}} \exp\left[- \sum_j \frac{(v_j - \sum_i a_i u_{ij} - A g_j)^2}{2\lambda_j} \right]. \end{aligned} \tag{43}$$

Using the principle of maximum likelihood estimation, we first make for each instant t an estimate of the amplitude, $\hat{A}(t)$, and then make an estimate of the epoch, $\hat{t}$, which corresponds to the maximum value of the likelihood ratio of the hypothesis H_1 against the hypothesis H_0. Thus

$$L(\hat{A},\hat{t}) = \max_t L(\hat{A},t) = \max_t \left[\frac{\max_{a,A} P_1(v;a_i\,,A)}{\max_b P_0(v;b_i)} \right]. \tag{44}$$

By the same argument leading to (28), we obtain

$$\begin{aligned} \log L(\hat{A},\hat{t}) &= \max_t \log L(\hat{A},t) = \max_t \,[\max_A \log L(A,t)] \\ &= \max_{t,A} \left[\sum_j \frac{2v_j A(t) g_j - A^2(t) g_j^2}{2\lambda_j} \right]. \end{aligned} \tag{45}$$

Taking the partial derivative with respect to A,

$$\frac{\partial \log L(A,t)}{\partial A} = 0, \tag{46}$$

we get $\hat{A}(t)$, the maximum likelihood estimate of $A(t)$,

$$\hat{A}(t) = \left[\sum_j \frac{v_j g_j}{\lambda_j} \right] \Big/ \left[\sum_j \frac{g_j^2}{\lambda_j} \right]. \tag{47}$$

It should be noted that the random variable v_j is also a function of t, as shown in (8). Substituting (47) into (45) gives us

$$\log L(\hat{A},\hat{t}) = \max_t \left[\frac{1}{2} \hat{A}^2(t) \sum_j \frac{g_j^2}{\lambda_j} \right]. \tag{48}$$

If we normalize function $f_g(\tau)$ such that

$$\begin{aligned} \sum_j \frac{g_j^2}{\lambda_j} &= \int_0^{2T} f_g^*(\tau) f_g(\tau)\, d\tau \\ &= \int_0^{2T} \int_0^{2T} R(\tau - \mu) f_g(\mu) f_g(\tau)\, d\mu\, d\tau \\ &= 1, \end{aligned} \tag{49}$$

then

$$\hat{A}(t) = \sum_j \frac{v_j g_j}{\lambda_j} = \int_0^{2T} f_{s+n}(t - T + \tau) f_g^*(\tau)\, d\tau, \tag{50}$$

and

$$\log L(\hat{A},\hat{t}) = \max_t \log L(\hat{A}, t) = \max_t \frac{1}{2} \,|\, \hat{A}(t) \,|^2. \tag{51}$$

Thus the instant $\hat{t}$ that corresponds to the maximum value of log $L(\hat{A},t)$ will be our estimate of the epoch. Indeed we may base our estimate on the maximum value of $|\hat{A}(t)|$. Equation (50) may of course be generated by means of a linear filter. If $f_h(\tau)$ and consequently $f_w(\tau)$ are properly "normalized" in the sense of (49) and (42), the signal-to-noise ratio (SNR) for the kth wavelet is A_k^2. Following the procedure used by Woodward and Davis,[20] it can be shown that, for the strong-signal case, the variance of the epoch estimate $\hat{t}$ is inversely proportional to SNR and to the square of the filter bandwidth.

We notice that the performance of an epoch detection system is related to the component functions only indirectly. It is the signal waveform itself that is important. As a rule of thumb, the smaller the absolute values of the constants r_i are, the more effective the system will be. In fact, we may define a useful figure of merit,

$$\rho \equiv \frac{\text{SNR for epoch detection}}{\text{SNR for the detection of } f_h(\tau)} = \frac{\int_0^{2T} f_g^*(\tau) f_g(\tau) d\tau}{\int_0^{2T} f_h^*(\tau) f_h(\tau) d\tau} \tag{52}$$

as the efficiency of the epoch detection system, where $f_h^*(\tau)$ is defined in the same way as we did for $f_g^*(\tau)$. Using (42) and (49) and the fact that $f_g^*(\tau)$ is orthogonal to $\varphi^{(i)}(\tau)$, we have

$$\rho = \frac{1}{1 + \sum_i r_i^2}. \tag{53}$$

As a result, $\rho < 1$. In the limit as every r_i approaches zero, $\rho \to 1$ and the epoch detection system approaches the optimum detection system for non-overlapping pulses of duration T seconds.

2.3 *Overlapping Pulses*

A pulse of duration T_0 seconds may be regarded as two overlapping wavelets with epochs separated by T_0 seconds. For instance, an exponential pulse, exp $(-\tau)$ for $0 \leqq \tau \leqq T_0$, may be regarded as the sum of two exponential functions, exp $(-\tau)$ with $0 \leqq \tau \leqq \infty$ and $-$exp $(-\tau)$ with $T_0 \leqq \tau \leqq \infty$. We assume that several pulses may overlap. Thus, arrival of a pulse is characterized by the simultaneous existence of a wavelet $f_w^b(\tau)$ at the beginning epoch t_k^b and a wavelet

$f_w^e(\tau)$ at the ending epoch t_k^e where $t_k^e = t_k^b + T_0$. For mathematical simplicity, we assume that the Gaussian noise in the time interval $(t_k^b - T, t_k^b + T)$ and the noise in the interval $(t_k^e - T, t_k^e + T)$ are uncorrelated. In other words, we assume

$$R(\tau) = 0 \quad \text{for} \quad \tau \geqq T_0 - 2T, \tag{54}$$

thus enabling us to treat independently the random variables in the two time intervals.

Again we formulate two hypotheses.

$$\begin{aligned} H_0 : f_s(t - T + \tau) &= \sum_i b_i^b \varphi^{(i)}(\tau), \\ f_s(t + T_0 - T + \tau) &= \sum_i b_i^e \varphi^{(i)}(\tau), \qquad 0 \leqq \tau \leqq 2T, \end{aligned} \tag{55}$$

and

$$\begin{aligned} H_1 : f_s(t - T + \tau) &= A(t) f_g^b(\tau) + \sum_i a_i^b \varphi^{(i)}(\tau), \\ f_s(t + T_0 - T + \tau) &= A(t) f_g^e(\tau) + \sum_i a_i^e \varphi^{(i)}(\tau), \\ & \qquad 0 \leqq \tau \leqq 2T, \end{aligned} \tag{56}$$

where

$$\begin{aligned} f_g^b(\tau) &\equiv f_h^b(\tau) - \sum_i r_i^b \varphi^{(i)}(\tau), \\ f_g^e(\tau) &\equiv f_h^e(\tau) - \sum_i r_i^e \varphi^{(i)}(\tau), \\ & \qquad 0 \leqq \tau \leqq 2T, \end{aligned} \tag{57}$$

with $f_h^b(\tau)$ and $f_h^e(\tau)$ defined in the same way as $f_h(\tau)$, and the constants r_i^b and r_i^e defined in the same way as r_i. Let us define

$$f_g(\tau) \equiv f_g^b(\tau) + f_g^e(\tau - T_0). \tag{58}$$

Notice that $f_g^e(\tau - T_0) = 0$ for $\tau < T_0$. The function $f_g(\tau)$ is normalized in the sense that

$$\int_0^{T_0+2T} f_g^*(\tau) f_g(\tau) d\tau = 1, \tag{59}$$

with

$$f_g^*(\tau) = f_g^{b*}(\tau) + f_g^{e*}(\tau - T_0) \tag{60}$$

and

$$f_g^b(\mu) = \int_0^{2T} R(\mu - \tau) f_g^{b*}(\tau) d\tau,$$
$$f_g^e(\mu) = \int_0^{2T} R(\mu - \tau) f_g^{e*}(\tau) d\tau. \tag{61}$$

Equations (58) through (61) can be justified only on the assumption of (54).

Using the maximum likelihood principle, we write

$$L(\hat{A},\hat{t}) = \max_t \left[\frac{\max\limits_{a,A} P_1(v; a_i^b, a_i^e, A)}{\max\limits_b \ P_0(v; b_i^b, b_i^e)} \right]. \tag{62}$$

With a derivation parallel to that of 2.2, we obtain the final results

$$\hat{A}(t) = \sum_j \frac{v_j g_j}{\lambda_j} = \int_0^{T_0+2T} f_{s+n}(t - T + \tau) f_g^*(\tau) d\tau, \tag{63}$$

and

$$\log L(\hat{A},\hat{t}) = \max_t \tfrac{1}{2} \mid \hat{A}(t) \mid^2. \tag{64}$$

Thus, based on the value of $| \hat{A}(t) |$, we may obtain the estimate of the epoch $\hat{t}$. Again a simple linear filter with a weighting function $f_g^*(\tau)$ defined in (60) will suffice to generate $\hat{A}(t)$.

III. OVERLAPPING STOCHASTIC SIGNALS

Again we consider a train of overlapping wavelets corrupted with a Gaussian noise. Each wavelet is assumed to be representable by a set of m known generalized exponential functions. However, the wavelets are stochastic in the sense that their exact waveforms are unknown and that each wavelet may differ from the other. As a result, the two hypotheses become

$$H_0 : f_s(t - T + \tau) = \sum_i b_i \varphi^{(i)}(\tau), \qquad 0 \leqq \tau \leqq 2T, \tag{65}$$

$$H_1 : f_s(t - T + \tau) = \sum_i c_i(t) \chi^{(i)}(\tau) + \sum_i a_i \varphi^{(i)}(\tau), \qquad 0 \leqq \tau \leqq 2T, \tag{66}$$

where

$$\chi^{(i)}(\tau) = \sum_l \beta_{il} \varphi^{(l)}(\tau - T) - \sum_l \gamma_{il} \varphi^{(l)}(\tau) \tag{67}$$

with

$$\varphi^{(l)}(\tau - T) = 0 \quad \text{for} \quad \tau < T. \tag{68}$$

The constants β_{il} and γ_{il} are constrained by the biorthonormality relationships. Let $\psi^{(j)}(\tau)$ be the eigenfunctions of the covariance function $R(\tau)$. For the sake of the independence of random variables, we expand the component functions, the noise, etc., in terms of $\psi^{(j)}(\tau)$. For $\chi^{(i)}(\tau)$, we then write

$$\chi^{(i)}(\tau) = \sum_j x_{ij}\psi^{(j)}(\tau), \qquad 0 \leqq \tau \leqq 2T. \tag{69}$$

The constraints of biorthonormality are

$$\sum_j \frac{u_{ij}u_{lj}}{\lambda_j} = \delta_{il}\,, \tag{70}$$

$$\sum_j \frac{x_{ij}u_{lj}}{\lambda_j} = 0, \tag{71}$$

and

$$\sum_j \frac{x_{ij}x_{lj}}{\lambda_j} = \delta_{il}\,. \tag{72}$$

A direct result of (70) and (71) is, similar to (25) and (32),

$$\begin{aligned} \gamma_{il} &= \sum_{j,k} \frac{u_{lj}}{\lambda_j}\beta_{ik}\int_0^{2T} \varphi^{(k)}(\tau - T)\psi^{(j)}(\tau)d\tau \\ &= \sum_k \beta_{ik}\int_0^{2T} \varphi^{(k)}(\tau - T)\varphi^{(l)*}(\tau)d\tau. \end{aligned} \tag{73}$$

To illustrate the procedure for formulating $\chi^{(i)}(\tau)$, let us consider the simple case of white noise for which the biorthonormality reduces to orthonormality. The first step is of course to orthonormalize with respect to the time interval $(0,2T)$ the m component functions by means of the Gram-Schmidt procedure. Next we may choose any value of β_{il} for (67) as long as the m functions $\sum \beta_{il}\varphi^{(l)}(\tau - T)$, $i = 1, 2 \cdots m$ are linearly independent. Using (73) for the calculation of γ_{il} guarantees that $\chi^{(i)}(\tau)$ is orthogonal to $\varphi^{(l)}(\tau)$. Finally, by means of the Gram-Schmidt process, we may combine the functions $\chi^{(i)}(\tau)$ linearly to make them orthonormal. In this way, all three conditions, (70), (71) and (72), are satisfied. For colored noise, the procedure is similar.

Under these assumptions of biorthonormality, an application of maximum likelihood principle then gives us

$$\hat{c}_i(t) = \sum_j \frac{v_j x_{ij}}{\lambda_j} = \int_0^{2T} f_{s+n}(t - T + \tau)\chi^{(i)}*(\tau)d\tau, \qquad (74)$$

where $\chi^{(i)*}(\tau)$ is the solution of the equation

$$\chi^{(i)}(\mu) = \int_0^{2T} R(\mu - \tau)\chi^{(i)}*(\tau)d\tau, \qquad (75)$$

and

$$\log L(\hat{c}_i, \hat{t}) = \max_t \log (\hat{c}_i, t) = \max_t \left[\frac{1}{2}\sum_i |\hat{c}_i(t)|^2\right]. \qquad (76)$$

The estimated epoch $\hat{t}$ is the instant that corresponds to the maximum value of log $L(\hat{c}_i, t)$. The epoch detection system which generates log $L(\hat{c}_i, t)$ will then consist of a summing amplifier, m squarers and m linear filters characterized by the m weighting functions $\chi^{(i)*}(\tau)$.

For stochastic signals, a proper definition of SNR for the kth wavelet is

$$\begin{aligned} d_k^2 &\equiv \frac{1}{m}\sum_{i=1}^m c_i^2(t_k) \int_0^{2T} \chi^{(i)}*(\tau)\chi^{(i)}(\tau)d\tau \\ &= \frac{1}{m}\sum_{i=1}^m c_i^2(t_k), \end{aligned} \qquad (77)$$

where we have used (72). The coefficient, $c_i(t_k)$, is defined by (66) with t_k, the kth epoch, substituted for t.

Let us now show some experimental results obtained from a digital computer. Fig. 3 illustrates the detection of overlapping wavelets, each consisting of two exponentials, $e^{-\tau}$ and $e^{-2\tau}$. Although in our experiment the three overlapping wavelets have the same waveshape, they are regarded as *stochastic* since we do not assume the a priori knowledge of the proportion of the two exponentials that constitute the wavelets. The signals are additively corrupted with white noise as shown in the second row. With the definition of (77), the signal-to-noise ratios for our examples are, in decibels, ∞, 15 and 8, respectively. Since we know the component functions, $e^{-\tau}$ and $e^{-2\tau}$, what we need to do would be simply to estimate the coefficients $\hat{c}_i(t)$ by means of linear filters prescribed by (67) and then calculate log $L(\hat{c}_i, \hat{t})$ according to (76). What we actually did is based on a more primitive model;[11] nevertheless, the basic philosophy is the same. Using this primitive model, the logarithms of the likelihood ratios, log $L_f(t)$, are calculated over the noisy signals, and shown in the third row. It is clear from the figure that the estimated epochs $\hat{t}$ which correspond to the maximum value of log $L_f(t)$, almost coincide with the true epochs. However, for the 8-db

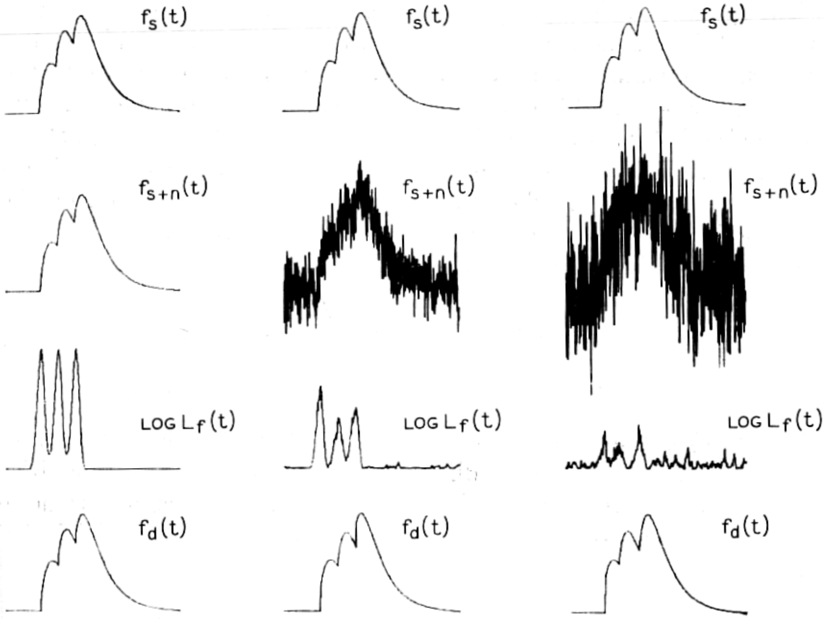

Fig. 3 — Detection of overlapping wavelets.

case, the detection probability is low, and the peak corresponding to the second epoch is almost not distinguishable from the peaks which are due to the random noise alone. With the epochs estimated, one may estimate in a piecewise manner the signal between the epochs, and the signals $f_d(t)$ thus detected are shown in the last row.

IV. OVERLAPPING RADAR SIGNALS

The most important application of statistical detection theory is in radar signal detection. We shall consider typical radar pulses which are sinusoidal signals modulated by square waves. Each pulse, as discussed in Section II, can be characterized by a beginning epoch and an ending epoch. For simplicity, we treat them separately as two epochs.

For overlapping radar signals, we may write

$$f_s(t) \cos (\omega_c t + \theta) = \sum_{k=1}^{N} A_k f_w(t - t_k) \cos (\omega_c t + \theta_k), \quad (78)$$

where we have regarded $f_s(t)$ and $f_w(\tau)$ as envelopes of the sinusoidal signals and θ_k are phase angles. The pulse envelope $f_w(\tau)$ is a square wave. Similarly we consider $f_{s+n}(t)$ as the envelope of the noisy signal.

The epoch detection system developed previously is difficult to implement in this case because it is sensitive to radio-frequency phase. For this reason, we assume the use of a perfect envelope detector to take the signal envelope first.[17] The noise under this condition becomes narrow-band noise. Let us designate, for a certain instant t, the envelope of the sinusoidal signal by α and the envelope of the noisy signal by v. With a variance λ, the probability density for the envelope at time t on the hypothesis that the sampled waveform is sine-wave plus noise is, according to the classic work of Rice,[21]

$$p(v,\alpha) = \begin{cases} \dfrac{v}{\lambda} \exp\left(-\dfrac{v^2+\alpha^2}{2\lambda}\right) I_0\left(\dfrac{\alpha v}{\lambda}\right), & v \geqq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{79}$$

where $I_0(x)$ is the modified Bessel function of the first kind and order zero. It is known as the modified Rayleigh distribution or the Rice distribution. On the hypothesis that the sampled waveform is noise alone, $\alpha = 0$ and $I_0(0) = 1$, and (79) is reduced to the Rayleigh distribution.

$$p_0(v) = \begin{cases} \dfrac{v}{\lambda} \exp\left(-\dfrac{v^2}{2\lambda}\right) & v \geqq 0, \\ 0, & \text{otherwise,} \end{cases} \tag{80}$$

We again formulate a null hypothesis H_0 and a hypothesis H_1 that an epoch has arrived. Thus,

$$H_0 : f_s(t - T + \tau) = \alpha_0, \qquad 0 \leqq \tau \leqq 2T. \tag{81}$$

$$H_1 : f_s(t - T + \tau) = \begin{cases} \alpha_1, & 0 \leqq \tau \leqq T, \\ \alpha_2, & T \leqq \tau \leqq 2T. \end{cases} \tag{82}$$

We may look for a coordinate system such that the random variables on these coordinates are statistically independent. However, unlike the Gaussian distribution, it is very difficult to find such a coordinate system. The usual procedure, which we shall follow here, is to use for coordinates samples of the envelope waveform taken at regular intervals and far enough apart so that it is a reasonable approximation to suppose them statistically independent. We take in the time interval $(t - T, t + T)$ $2M$ measurements at $2M$ uniformly spaced instants separated by $\Delta\tau$ seconds apart. Let us write for the instant t

$$v_j(t) = f_{s+n}(t - T + j\Delta\tau). \tag{83}$$

Then on the null hypothesis H_0, the joint probability density is

$$P_0(v;\alpha_0) = \prod_{j=1}^{2M} \frac{v_j}{\lambda} \exp\left(-\frac{v_j^2 + \alpha_0^2}{2\lambda}\right) I_0\left(\frac{\alpha_0 v_j}{\lambda}\right), \tag{84}$$

while on the hypothesis H_1, we have

$$\begin{aligned} P_1(v;\alpha_1, \alpha_2) &= \prod_{j=1}^{M} \frac{v_j}{\lambda} \exp\left(-\frac{v_j^2 + \alpha_1^2}{2\lambda}\right) I_0\left(\frac{\alpha_1 v_j}{\lambda}\right) \\ &+ \prod_{j=M+1}^{2M} \frac{v_j}{\lambda} \exp\left(-\frac{v_j^2 + \alpha_2^2}{2\lambda}\right) I_0\left(\frac{\alpha_2 v_j}{\lambda}\right). \end{aligned} \tag{85}$$

The maximum likelihood principle then requires

$$L(\hat{\alpha}, \hat{t}) = \max_t \left[\frac{\max_\alpha P_1(v;\alpha_1, \alpha_2)}{\max_\alpha P_0(v;\alpha_0)}\right]. \tag{86}$$

By considering the logarithm of the likelihood functions and taking partial derivatives, we can easily show that

$$\begin{aligned} \sum_{j=1}^{M} \left[-\frac{\hat{\alpha}_1}{\lambda} + \frac{v_j}{\lambda} \frac{I_0'\left(\frac{\hat{\alpha}_1 v_j}{\lambda}\right)}{I_0\left(\frac{\hat{\alpha}_1 v_j}{\lambda}\right)}\right] &= 0, \\ \sum_{j=M+1}^{2M} \left[-\frac{\hat{\alpha}_2}{\lambda} + \frac{v_j}{\lambda} \frac{I_0'\left(\frac{\hat{\alpha}_2 v_j}{\lambda}\right)}{I_0\left(\frac{\hat{\alpha}_2 v_j}{\lambda}\right)}\right] &= 0, \end{aligned} \tag{87}$$

and

$$\sum_{j=1}^{2M} \left[-\frac{\hat{\alpha}_0}{\lambda} + \frac{v_j}{\lambda} \frac{I_0'\left(\frac{\hat{\alpha}_0 v_j}{\lambda}\right)}{I_0\left(\frac{\hat{\alpha}_0 v_j}{\lambda}\right)}\right] = 0, \tag{88}$$

where $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\alpha}_0$ are the maximum likelihood estimates of α_1, α_2, α_0 respectively, and I_0' is the derivative of I_0. The estimated epoch then corresponds to

$$\begin{aligned} \log L(\hat{\alpha}, \hat{t}) = \max_t \Bigg\{ &\sum_{j=1}^{M} \left[-\frac{\hat{\alpha}_1^2}{2\lambda} + \log I_0\left(\frac{\hat{\alpha}_1 v_j}{\lambda}\right)\right] \\ &+ \sum_{j=M+1}^{2M} \left[-\frac{\hat{\alpha}_2^2}{2\lambda} + \log I_0\left(\frac{\hat{\alpha}_2 v_j}{\lambda}\right)\right] \\ &- \sum_{j=1}^{2M} \left[-\frac{\hat{\alpha}_0^2}{2\lambda} + \log I_0\left(\frac{\hat{\alpha}_0 v_j}{\lambda}\right)\right]\Bigg\}. \end{aligned} \tag{89}$$

It should be noted that $\hat{\alpha}$ and v_j are functions of t. Our problem would have been solved if we had been able to solve (87) and (88) for $\hat{\alpha}$, substitute them into (89) and then search for $\hat{t}$ that corresponds to the maximum value of $\log L(\hat{\alpha},t)$. An explicit solution in this case is, to say the least, very difficult. Certain approximations are needed. We shall discuss the case of strong signals and the case of weak signals separately.

In the first place, we notice that if the signal-to-noise ratio is sufficiently high — i.e., $\alpha_1^2/2\lambda \gg 1$ and $\alpha_2^2/2\lambda \gg 1$ [$(\alpha_1^2 - \alpha_2^2)/2\lambda$ may be small] — the Rice distribution approaches the Gaussian distribution and the discussion in Section II is directly applicable. A linear filter may thus be used. On the other hand, if $\alpha_1^2/2\lambda \gg 1$ and $\alpha_2^2/2\lambda$ small or vice versa, we encounter the epoch of a large pulse. Therefore a sub-optimal epoch detection scheme may be used, and again a linear filter may be chosen for its simplicity.

Finally, consider the case that $\alpha_1^2/2\lambda \ll 1$ and $\alpha_2^2/2\lambda \ll 1$. It is well-known that for small x (see Ref. 22),

$$
\begin{aligned}
I_0(x) &= 1 + \left(\frac{1}{2}x\right)^2 + \frac{1}{1^2\cdot 2^2}\left(\frac{1}{2}x\right)^4 + \cdots \\
\log I_0(x) &= \left(\frac{1}{2}x\right)^2 - \frac{1}{4}\left(\frac{1}{2}x\right)^4 + \cdots .
\end{aligned}
\tag{90}
$$

If we substitute (90) into (87) and (88) and retain only those terms that involve $\hat{\alpha}/\sqrt{\lambda}$ and $(\hat{\alpha}/\sqrt{\lambda})^3$, we obtain as approximations

$$
\begin{aligned}
\frac{\hat{\alpha}_1^{\,2}}{2\lambda} &\approx \sum_{j=1}^{M}\left[\frac{v_j^{\,2}}{2\lambda} - 1\right] \Big/ \sum_{j=1}^{M}\frac{v_j^{\,4}}{8\lambda^2}, \\
\frac{\hat{\alpha}_2^{\,2}}{2\lambda} &\approx \sum_{j=M+1}^{2M}\left[\frac{v_j^{\,2}}{2\lambda} - 1\right] \Big/ \sum_{j=M+1}^{2M}\frac{v_j^{\,4}}{8\lambda^2},
\end{aligned}
\tag{91}
$$

and

$$
\frac{\hat{\alpha}_0^{\,2}}{2\lambda} \approx \sum_{j=1}^{2M}\left[\frac{v_j^{\,2}}{2\lambda} - 1\right] \Big/ \sum_{j=1}^{2M}\frac{v_j^{\,4}}{8\lambda^2}. \tag{92}
$$

Similarly, substituting (90) into (89) gives us

$$
\begin{aligned}
\log L(\hat{\alpha},\hat{t}) = \max_t \Bigg\{ &\frac{\hat{\alpha}_1^{\,2}}{2\lambda} \sum_{j=1}^{M}\left[\frac{v_j^{\,2}}{2\lambda} - 1 - \frac{1}{2}\frac{\hat{\alpha}_1^{\,2}}{2\lambda}\frac{v_j^{\,4}}{8\lambda^2}\right] \\
&+ \frac{\hat{\alpha}_2^{\,2}}{2\lambda} \sum_{j=M+1}^{2M}\left[\frac{v_j^{\,2}}{2\lambda} - 1 - \frac{1}{2}\frac{\hat{\alpha}_2^{\,2}}{2\lambda}\frac{v_j^{\,4}}{8\lambda^2}\right] \\
&- \frac{\hat{\alpha}_0^{\,2}}{2\lambda} \sum_{j=1}^{2M}\left[\frac{v_j^{\,2}}{2\lambda} - 1 - \frac{1}{2}\frac{\hat{\alpha}_0^{\,2}}{2\lambda}\frac{v_j^{\,4}}{8\lambda^2}\right]\Bigg\}.
\end{aligned}
\tag{93}
$$

An epoch detection system based on (91), (92), and (93) is very difficult to implement. It needs further simplification. We notice that, for M sufficiently large, the random variables take on nearly all possible values of the probability functions. The summations then represent an ensemble average. For example, with $\alpha_1^2/2\lambda \ll 1$,

$$\begin{aligned}\sum_{j=1}^{M} \frac{v_j^2}{2\lambda} &\approx M \overline{\left(\frac{v_j^2}{2\lambda}\right)} = M \int_0^\infty \frac{x^2}{2\lambda}\, p(x,\alpha_1)dx \approx M\left(1 + \frac{\alpha_1^2}{2\lambda}\right), \\ \sum_{j=1}^{M} \frac{v_j^4}{8\lambda^2} &\approx M \overline{\left(\frac{v_j^4}{8\lambda^2}\right)} = M \int_0^\infty \frac{x^4}{8\lambda^2}\, p(x,\alpha_1)dx \approx M\left(1 + \frac{\alpha_1^2}{\lambda}\right),\end{aligned} \tag{94}$$

where $p(x,\alpha_1)$ is the Rice distribution shown in (79). We may of course do the same for $\alpha_2^2/2\lambda$ and $\alpha_0^2/2\lambda$. It is indeed from this consideration that we include the term

$$\frac{1}{2}\frac{\hat{\alpha}^2}{2\lambda}\frac{v_j^4}{8\lambda^2}$$

in (93), since it is of the same order as the difference of the remaining two terms.

From the consideration of order of magnitude, it should be obvious that for (91) and (92) the denominators may be replaced by M and $2M$ respectively. As a result, we may write

$$\frac{\hat{\alpha}_0^2}{2\lambda} = \frac{1}{2}\left(\frac{\hat{\alpha}_1^2}{2\lambda} + \frac{\hat{\alpha}_2^2}{2\lambda}\right). \tag{95}$$

Using (91), (92), (94), and (95) for (93) and simplifying, we finally obtain

$$\begin{aligned}\log L(\hat{\alpha},\hat{t}) &= \max_t \left[\frac{M}{2}\left(\frac{\hat{\alpha}_1^2}{2\lambda}\right)^2 + \frac{M}{2}\left(\frac{\hat{\alpha}_2^2}{2\lambda}\right)^2 - M\left(\frac{\hat{\alpha}_0^2}{2\lambda}\right)^2\right] \\ &= \max_t \frac{M}{4}\left|\frac{\hat{\alpha}_2^2}{2\lambda} - \frac{\hat{\alpha}_1^2}{2\lambda}\right|^2 \\ &= \max_t \frac{1}{4M}\left|\sum_{j=M+1}^{2M} \frac{v_j^2(t)}{2\lambda} - \sum_{j=1}^{M} \frac{v_j^2(t)}{2\lambda}\right|^2.\end{aligned} \tag{96}$$

A test may naturally be based on the quantity inside the absolute sign. A large positive value for the quantity indicates the arrival of a beginning epoch, and a large negative value corresponds to an ending epoch. A pulse is of course marked by the arrival of both epochs. The result is consistent with the conventional square-law detector for small, nonoverlapping signals.

V. CONCLUSION

We have investigated the problem of epoch detection. A test statistic, which may be obtained from a simple, linear filter, has been derived for Gaussian noise. In the derivation, we have assumed that each wavelet is representable by a set of known generalized exponentials. This is not as restrictive as it appears, considering the fact that any continuous signal may be represented with a least-square error as small as we wish by using a sufficient number of component functions.

The epoch detection scheme is particularly useful for the resolution and detection of overlapping signals. For N overlapping wavelets, the procedure reduces the resolution problem from an N-dimensional problem to N one-dimensional problems. Some information is lost in this reduction, and consequently it is not a scheme for optimal resolution. However, it has the essential advantage of simplicity and practicality.

The performance of the epoch detection system has been considered briefly. The discussions of overlapping stochastic signals and overlapping radar signals show that the method is applicable to these cases, and the experimental results enhance our confidence in the detection procedure.

VI. ACKNOWLEDGMENTS

The author wishes to express his appreciation to W. H. Huggins of the Johns Hopkins University for many helpful suggestions. He is also grateful to J. F. Kaiser of Bell Telephone Laboratories for his comments on the manuscript. Part of the preliminary work was done while the author was with the Carlyle Barton Laboratory, the Johns Hopkins University, Baltimore, Maryland, under a contract supported by the U. S. Air Force.

APPENDIX

The weighting functions of the "matched" filters are calculated according to the equation

$$f_g(\tau) = f_h(\tau) - \sum_i r_i \varphi^{(i)}(\tau),$$

where r_i's are chosen in such a way that for white noise $f_g(\tau)$ is orthogonal to every $\varphi^{(i)}(\tau)$. This orthogonality (or biorthogonality for Gaussian noise) is the central idea of epoch detection, and has been discussed in the paper.

As an example, for the last signal in Fig. 2, the signal (or wavelet) is of the form

$$f_w(\tau) = e^{-\tau} - e^{-2\tau},$$

and therefore from (11)

$$f_h(\tau) = \begin{cases} e^{-(\tau-T)} - e^{-2(\tau-T)}, & T \leqq \tau \leqq 2T, \\ 0, & \text{elsewhere.} \end{cases}$$

Thus we have

$$f_g(\tau) = \begin{cases} -r_1e^{-\tau} - r_2e^{-2\tau}, & 0 \leqq \tau < T, \\ (e^T - r_1)e^{-\tau} - (e^{2T} + r_2)e^{-2\tau}, & T \leqq \tau \leqq 2T, \end{cases}$$

with r_1 and r_2 to be determined by the orthogonality relationships

$$\int_0^{2T} f_g(\tau)e^{-\tau}d\tau = 0,$$

$$\int_0^{2T} f_g(\tau)e^{-2\tau}d\tau = 0,$$

In our example, $T = 0.7$, and then the solution of the above two equations is

$$r_1 = 0.62 \quad \text{and} \quad r_2 = -0.76.$$

A substitution of these values into the equation of $f_g(\tau)$ results in

$$f_g(\tau) = \begin{cases} -0.62e^{-\tau} + 0.76e^{-2\tau}, & 0 \leqq \tau < 0.7, \\ 1.39e^{-\tau} - 3.30e^{-2\tau}, & 0.7 \leqq \tau \leqq 1.4, \end{cases}$$

which, except for a scale factor, is the weighting function shown in Fig. 2.

REFERENCES

1. Woodward, P. M., *Probability and Information Theory with Applications to Radar*, Pergamon Press, New York, 1953.
2. Middleton, D., *An Introduction to Statistical Communication Theory*, McGraw-Hill, New York, 1960.
3. Helstrom, C. W., *Statistical Theory of Signal Detection*, Pergamon Press, New York, 1960.
4. Wainstein, L. A., and Zubakov, V. D., *Extraction of Signals from Noise*, Prentice-Hall, Englewood Cliffs, N. J., 1962.
5. Helstrom, C. W., The Resolution of Signals in White Gaussian Noise, Proc. IRE, *43*, Sept., 1955, p. 1111.
6. Nilsson, N. J., On the Optimum Range Resolution of Radar Signals in Noise, IRE Trans. Information Theory, *IT-7*, Oct., 1961, p. 245.
7. Root, W. L., Radar resolution of closely spaced targets, IRE Trans. Military Electronics, *MIL-6*, April, 1962, p. 197.

8. Sussman, S. M., Least-Square Synthesis of Radar Ambiquity Functions, IRE Trans. Information Theory, *IT-8*, April, 1962, p. 246.
9. Rihaczak, A. W., Radar Resolution Properties of Pulse Trains, Proc. IEEE, *52*, Feb., 1964, p. 153.
10. Klander, J. R., Price, A. C., Darlington, S., and Albersheim, W. L., The Theory and Design of Chirp Radars, B.S.T.J., *39*, July, 1960, p. 745.
11. Young, T. Y., Statistical Epoch Detection of Overlapping Signals, unpublished memorandum, Carlyle Barton Laboratory, Johns Hopkins University, 1963.
12. Slepian, D., Estimation of Signal Parameters in the Presence of Noise, IRE Trans. Information Theory, *PGIT-3*, Mar., 1954, p. 68.
13. Middleton, D., and Van Meter, D., Detection and Extraction of Signals in Noise from the Point of View of Statistical Decision Theory, J. Soc. Indust. Appl. Math., *3*, Dec., 1955, p. 192; *4*, June, 1956, p. 86.
14. Huggins, W. H., Signal Theory, IRE Trans. Circuit Theory, *CT-3*, Dec., 1956, p. 210.
15. Ule, L. A., Weighted Least-Square Smoothing Filters, IRE Trans. Circuit Theory, *CT-2*, June, 1955, p. 197.
16. Kelly, E. J., Reed, I. S., and Root, W. L., The Detection of Radar Echoes in Noise, J. Soc. Indust. Appl. Math., *8*, June, 1960, p. 309; Sept., 1960, p. 481.
17. Davenport, W. B., Jr., and Root, W. L., *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York, 1958.
18. Akhiezer, N. I., *Theory of Approximation*, Frederick-Ungar, New York, 1956.
19. Zadeh, L. A., and Ragazzini, J. R., Optimum Filters for the Detection of Signals in Noise, Proc. IRE, *40*, Oct., 1952, p. 1123.
20. Woodward, P. M., and Davis, I. L., A Theory of Radar Information, Phil. Mag., *41*, 1950, p. 1001.
21. Rice, S. O., Mathematical Analysis of Random Noise, B.S.T.J., *23*, 1944, p. 282; *24*, 1945, p. 46.
22. Dwight, H. B., *Tables of Integrals and Other Mathematical Data*, Macmillan, New York, 1957.