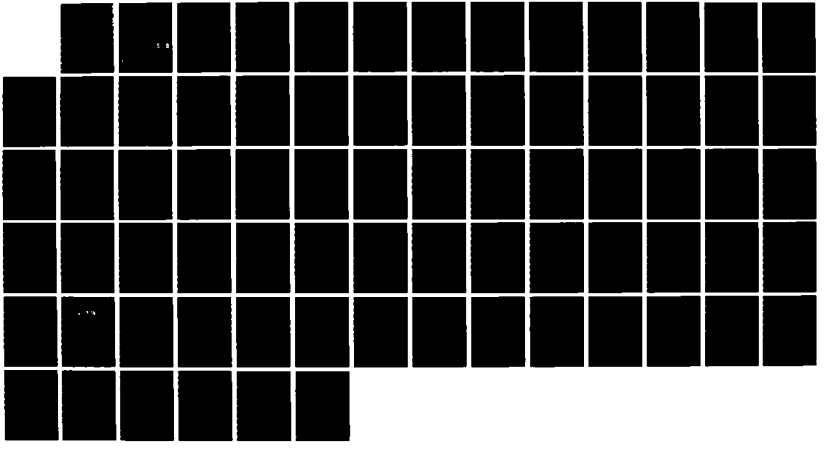
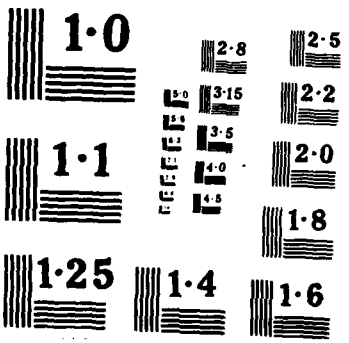
AD-A169 937 ZERO-CROSSINGS ANALYSIS(U) MARYLAND UNIV COLLEGE PARK 1/1
DEPT OF MATHEMATICS B KEDEM JAN 86 AFOSR-TR-86-0413
AFOSR-82-0187

UNCLASSIFIED F/G 12/1 NL




1·0
2·8
2·5
3·15
2·2
1·1
3·5
4·0
2·0
1·8
1·25
1·4
1·6

(2)

AD-A169 937

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UC	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 86-0413
6a. NAME OF PERFORMING ORGANIZATION: University of Maryland; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR
6c. ADDRESS (City, State and ZIP Code): College Park, MD 20742	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Air Force Office of Scientific Research; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-82-0187
8c. ADDRESS (City, State and ZIP Code): Bolling AFB, Washington, DC 20332	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
~~AFOSR 82-0187~~ 61102F		2304/A5	

11. TITLE (Include Security Classification): Zero-Crossings Analysis	
12. PERSONAL AUTHOR(S): Benjamin Kedem	
13a. TYPE OF REPORT: Technical Report; 13b. TIME COVERED: FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day): January 1986; 15. PAGE COUNT: 70
16. SUPPLEMENTARY NOTATION	

DTIC ELECTE JUL 24 1986 S D

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	HIGHER ORDER CROSSINGS

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

We advance a coherent development of zero-crossing based methods and theory appropriate for fast signal analysis. Quite a few ideas pertaining to zero-crossing counts found in the literature can be expressed and interpreted with the help of this more general setup. A central issue addressed in some detail is the fruitful connection which exists between zero-crossing counts and linear filtering. This connection is explored and interpreted with the help of a certain zero-crossing spectral representation, and is then applied in spectral analysis, detection and discrimination. Zero-crossing counts in filtered time series are called higher order crossings. The theme of this work is that higher order crossings analysis provides a useful descriptive as well as analytical tool that can in many respects match spectral analysis. To a great extent these two types of analysis are in fact equivalent, but each emphasizes a different point of view. Advantages offered by higher order crossings are great simplicity and a drastic data reduction.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS ☐	21 ABSTRACT SECURITY CLASSIFICATION: UC
22a. NAME OF RESPONSIBLE INDIVIDUAL: B. KEIL [illegible]	22b TELEPHONE NUMBER (Include Area Code): [illegible] 22c OFFICE SYMBOL: NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE

ZERO-CROSSINGS ANALYSIS

by

Benjamin Kedem*

Department of Mathematics
University of Maryland
College Park, Maryland
20742

MD85-44
TR85-38

November 1985

ZERO-CROSSINGS ANALYSIS

by

Benjamin Kedem

Department of Mathematics, University of Maryland, College Park

ABSTRACT

We advance a coherent development of zero-crossing based methods and theory appropriate for fast signal analysis. Quite a few ideas pertaining to zero-crossing counts found in the literature can be expressed and interpreted with the help of this more general setup. A central issue addressed in some detail is the fruitful connection which exists between zero-crossing counts and linear filtering. This connection is explored and interpreted with the help of a certain zero-crossing spectral representation, and is then applied in spectral analysis, detection and discrimination. Zero-crossing counts in filtered time series are called higher order crossings. The theme of this work is that higher order crossings analysis provides a useful descriptive as well as analytical tool that can in many respects match spectral analysis. To a great extent these two types of analysis are in fact equivalent, but each emphasizes a different point of view. Advantages offered by higher order crossings are great simplicity and a drastic data reduction.

The support of Grant AFOSR 82-0187 is gratefully acknowledged.

ZERO-CROSSINGS ANALYSIS

by

Benjamin Kedem

1. INTRODUCTION

We shall be concerned in this paper with developing a certain systematic approach to the analysis of random signals based on zero-crossing counts. The need for such a development stems from the interest in zero-crossing based methods and techniques in fields such as signal processing [1], [2], [3], fluid mechanics [4], [5], speech processing [6], [7], biomedical engineering [8], optics [9], [10], neurophysiology [11], structural dynamics [12], [13], communications [14], [15], and image processing [16], [17]. It is felt that a more general setup can provide a better understanding of these techniques and of their outcomes, and furthermore, suggests interesting and useful new methods of analysis.

The first serious attempt to study properties of zero-crossings is the pioneering work of Kac [18] and Rice [19] who were mainly interested in moment and distribution problems pertaining to zero-crossing counts. This work and related level crossing problems are discussed and reviewed in [2], [20], [21], [22], [23], [24], where numerous additional references can be found. None of this voluminous work, though, will be reproduced here.

There is a great deal more to zero-crossings other than the classical distribution and moment problems, and we shall be concerned with somewhat more general properties of zero-crossing counts and zero-crossing based methods of analysis. In particular,

the present work explores the fruitful connection which exists between zero-crossing counts and time invariant linear filtering. As we shall wee, this connection sheds light on interesting and surprisingly useful properties of zero-crossings appropriate for the fast analysis of random signals. Furthermore, it will become evident that zero-crossing counts in random signals and their filtered versions essentially constitute a *domain* by itself which in many respects is equivalent to the spectral domain. This will be demonstrated and illustrated by a fair number of examples, using real and artificial data.

To motivate some key ideas presented in this paper, it is instructive and helpful at this early stage to consider an example [25], [26], [27] which is most illuminating. Figure 1 presents the graphs of two signals in discrete time, $t = 1,2,\ldots,100$, each made of a superposition of two finite sinusoids with angular frequencies 0.1 and 3. As a general rule we adhere to discrete time which is more convenient for machine calculations. Consider the first graph (a) and suppose it is desired to estimate the highest frequency which in this case is 3. Let D_1 and D_2 denote the number of zero-crossings and the number of local minima and maxima, respectively, when the discrete time points are connected by line segments. Since the low frequency component in (a) is obviously dominant, $D_1 = 7$ is too small for serving as a reasonable estimate. On the other hand, $D_2 = 92$ gives a good deal more information about the periodicity of the high frequency component as it better captures the component's oscillation characteristics. Indeed, by using proper normalization, our estimate for the highest frequency

is $\pi \cdot 92/99 \simeq 2.9$ which is quite reasonable. Now D_2, ignoring end effects, is the number of zero-crossings in the differenced signal and a difference is a high-pass filter which amplifies high frequency components. Thus, the crucial point is that in this case we must render the high frequency component dominant by a filter in order to be able to estimate its frequency from zero-crossings. On the other hand, in the second case (b) the high frequency component is already dominant and $D_1 = 95$ yields $\pi \cdot 95/99 \simeq 3$ as expected. Note that in this second case $D_2 \simeq 94$ is very close to D_1 which is an indication [28], [29] of the presence of a strong periodic component. It is seen that for the cases considered the pair (D_1, D_2) provides useful spectral information.

Another observation pertains to the discrimination potential of (D_1, D_2). The two graphs appear to be obviously different and this difference in appearance results in different pairs (D_1, D_2). This shows that differences in signals can be expressed very economically by zero-crossings observed in the signal and in functions thereof.

In what follows, these ideas will be developed within a convenient framework which is very effective in describing the domain of zero-crossings.

The paper is divided into four parts. The first, consisting of Section 2, introduces some basic notions and terminology of stationary random processes. The second (3.1 - 3.4) discusses general properties of zero-crossings. As a matter of convenience important points and facts are summarized at times in the form of theorems. We do not prove these theorems but give intuitive

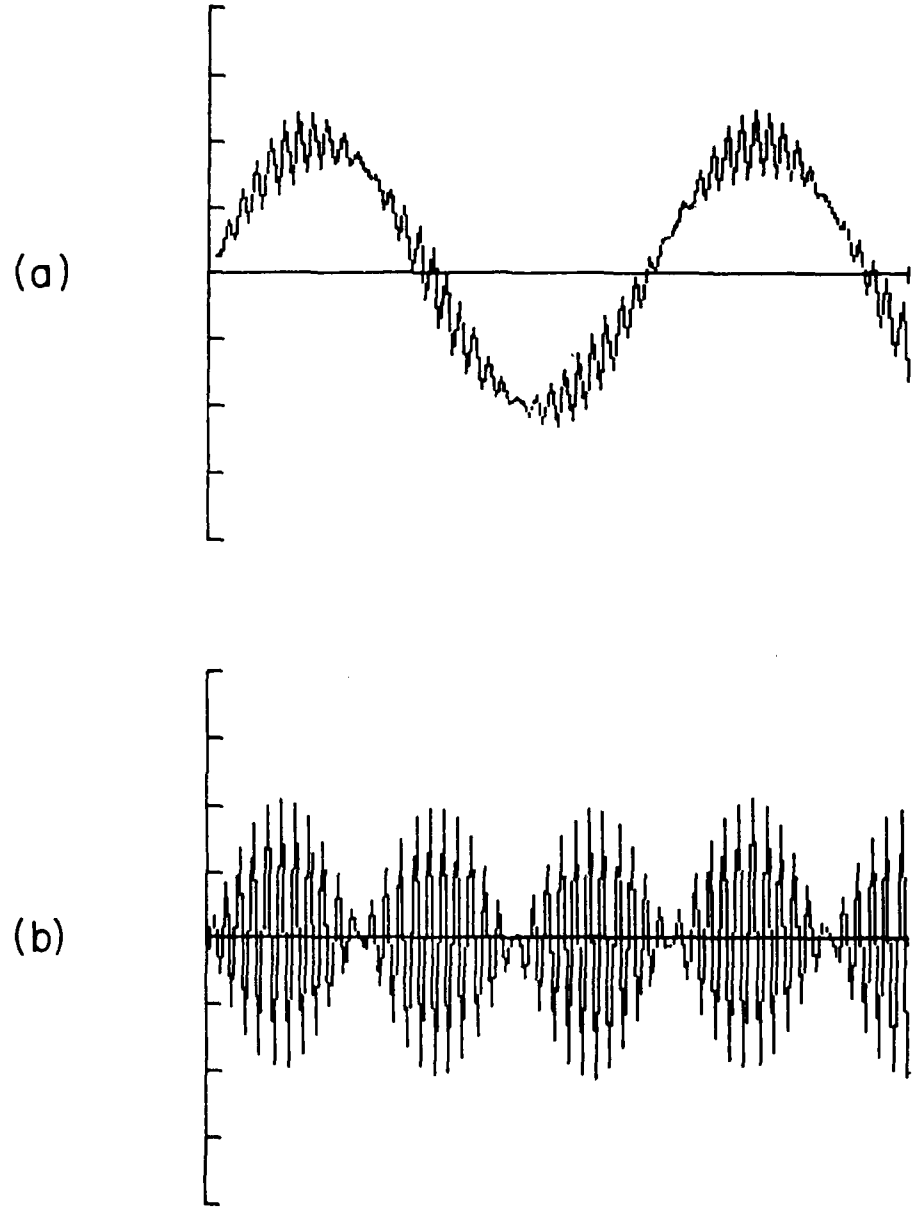


Figure 1. Superpositions of two sinusoids with angular frequencies 0.1 and 3.

arguments instead of formal proofs. The third part (4.1 - 4.5) discusses spectral analysis by zero-crossings, and the fourth (5.1 - 5.5) is devoted to discrimination between time series. The paper contains quite a few examples and illustrations which enhance its tutorial nature.

2. SOME PRELIMINARIES AND NOTATION

A random signal in discrete time is commonly referred to as a stochastic or random process. Without loss of generality we let "discrete time" be the set of integers and denote the random process by $\{Z_t\}$, $t = \ldots,-1,0,1,\ldots$, where for each fixed t, Z_t is a random quantity. Examples of such random processes are any electronic noise sampled at integral time points and random digit streams. E will denote mathematical expectation or mean. In the context of random processes E stands for ensemble average. Many of the ideas which we would like to bring up are best explained if we make the simplifying assumption that $\{Z_t\}$ is a real stationary Gaussian random process with mean zero. By this we mean that "statistics" do not change in time and that the joint probability distribution of any finite vector $(Z_{t_1},\ldots,Z_{t_k})$ has a multivariate normal distribution. In this case for each fixed t, Z_t has the usual Gaussian distribution. There are a number of very good references [21], [30], [31], [32], [33], [34] which cover the theory and applications of stationary processes and to which the reader is referred for complete treatment. Here, however, we summarize very briefly some basic notions needed in the sequel.

A stationary random process is characterized by the requirement that the mean $EZ_t \equiv \mu$ be a constant and that the covariance function

$$\gamma_k \equiv E(Z_t - \mu)(Z_{t+k} - \mu)$$

is a function of the lag k only. Then the celebrated Wiener-Khintchine Theorem [33], p. 75, states that there exists a monotone

increasing function $F(\omega)$ such that

$$\gamma_k = \int_{-\pi}^{\pi} \cos(k\omega)\, dF(\omega)$$

where π is the highest possible angular frequency due to the discreteness of t. F is called the spectral distribution function of the random process. $F(\pi)$ is equal to the variance (total power) of Z_t since

$$\gamma_0 = \int_{-\pi}^{\pi} dF(\omega).$$

Except for a constant, F behaves like a cumulative probability distribution function and therefore can be decomposed, for all practical purposes, into two components

$$F(\omega) = F_1(\omega) + F_2(\omega).$$

F_1 is a nondecreasing continuous function determined by a nonnegative symmetric spectral density, f, by the relation

$$F_1(\omega) = \int_{-\pi}^{\omega} f(\lambda)\, d\lambda.$$

F_2 is a nondecreasing step function determined by a symmetric spectral function, q, by the relation

$$F_2(\omega) = \sum_{\lambda_j \leq \omega} q(\lambda_j)$$

where λ_j $(= \lambda_{-j})$ is a discrete set of frequencies. The spectral mass or power associated with a set of frequencies Λ is determined from the continuous and discrete components by the expression [33]

$$\int_{\Lambda} dF(\omega) = \sum_{\lambda_j \in \Lambda} q(\lambda_j) + \int_{\Lambda} f(\lambda)\, d\lambda.$$

That is, both continuous and discrete components contribute to the power. The spectrum is said to be continuous, discrete or of mixed type according to the three cases ($F_1 \not\equiv 0$, $F_2 \equiv 0$), ($F_1 \equiv 0$, $F_2 \not\equiv 0$) or ($F_1 \not\equiv 0$, $F_2 \not\equiv 0$), the last one being the most general case. An important example of a process with a mixed spectrum is furnished by a trigonometric polynomial with random amplitudes plus noise. This type of process is useful for our purpose.

By a time invariant linear filter [33], [35] we mean a linear operation L on $\{Z_t\}$, of the form

$$Y_t \equiv L(Z_t) = h_0 Z_t + h_1 Z_{t-1} + h_2 Z_{t-2} + \ldots$$

which produces a new stationary process $\{Y_t\}$. The transfer function $H(\omega)$ is defined by the sum

$$H(\omega) = \sum_r h_r e^{-i\omega r}.$$

$|H(\omega)|$ is called the gain associated with the filter

When a time invariant linear filter with transfer function H is applied to a stationary process $\{Z_t\}$ the resulting process is also stationary provided the matching condition

$$\int_{-\pi}^{\pi} |H(\lambda)|^2 dF(\lambda) < \infty$$

is satisfied. In this case the output spectral density and spectral function are given by $|H(\lambda)|^2 f(\lambda)$ and $|H(\lambda)|^2 q(\lambda)$, $-\pi < \lambda \leq \pi$, respectively.

In many cases it is convenient to normalize γ_k by γ_0 and define a new quantity ρ_k by

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \qquad k = 0, \pm 1, \ldots .$$

ρ_k is called the correlation or autocorrelation function and satisfies $|\rho_k| \leq 1$.

One last point, any finite realization of a random process, $Z_1, Z_2, \ldots, Z_N$, is called a time series. Throughout most of the paper, we deal with time series of length N from stationary Gaussian random processes with zero mean. The only exceptions are bounded processes discussed very briefly at the end.

3. SOME GENERAL PROPERTIES OF ZERO-CROSSINGS

In this section zero crossing counts in discrete time are defined and are shown to possess a certain spectral representation which shows their intriguing connection with the spectrum and linear filtering. The representation also illustrates the fundamental principle that the zero-crossing rate tends to admit values in a neighborhood of a dominant frequency. This principle is the key factor in many of our results.

3.1. A Zero-crossing Spectral Representation

Let $Z_1,\ldots,Z_N$ be a zero mean stationary Gaussian time series. To define what is meant by zero-crossings in discrete time we consider the associated clipped binary series X_t defined by

$$X_t \equiv \begin{cases} 1, & Z_t \geq 0 \\ 0, & Z_t < 0 \end{cases}, \qquad t = 1,\ldots,N,$$

and let d_t be the indicator function at time t,

$$d_t \equiv X_t + X_{t-1} - 2X_tX_{t-1}.$$

Then d_t is 0 or 1. When $d_t = 1$ we say that a zero-crossing occurs at time t. The number of zero-crossings in $Z_1,\ldots,Z_N$ is denoted by D_1 and is defined by the sum

$$D_1 \equiv d_2 + \ldots + d_N.$$

As an illustration consider the time series Z_t, $t = 1,\ldots,12$, and its associated two binary series X_t, d_t, $t = 1,\ldots,12$ in Figure 2. We have $D_1 = d_2 + \ldots + d_{12} = 6$, and it is seen that the number

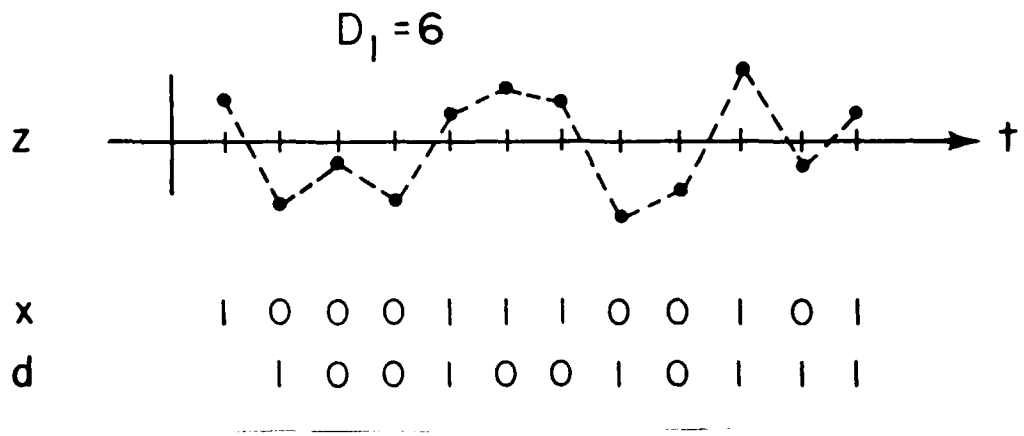


Figure 2. A time series of size 12 with 6 zero-crossings.

of zero-crossings is precisely the number of symbol changes in the X series. As a digression we remark that the number of symbol changes is a so-called sufficient statistic for stationary binary Markov chains and as such this number carries information about the chains' parameters [36].

In the classical theory of zero-crossings in continuous time the expected number of zero-crossings per unit time is a useful quantity [19], [2], and it is interesting to see where ED_1, the expected number of zero-crossings, leads us in the present case of discrete time.

Since $\{Z_t\}$ is Gaussian with zero mean, $EX_t = 1/2$ and [37]

$$EX_t X_{t-1} = \frac{1}{4} + \frac{1}{2\pi} \sin^{-1}\rho_1$$

which gives $ED_t = \frac{1}{2} - \frac{1}{\pi} \sin^{-1}\rho_1$ and therefore

$$ED_1 = (N-1)(\frac{1}{2} - \frac{1}{\pi} \sin^{-1}\rho_1).$$

By rearranging terms we obtain the basic formula [36]

$$\rho_1 = \cos\left(\frac{\pi ED_1}{N-1}\right). \tag{1}$$

But from the definition of ρ_1 and by the Wiener-Khintchine Theorem follows the relation

$$\cos\left(\frac{\pi ED_1}{N-1}\right) = \frac{\int_{-\pi}^{\pi} \cos(\omega)\,dF(\omega)}{\int_{-\pi}^{\pi} dF(\omega)} \tag{2}$$

to which we refer as the *zero-crossing spectral representation* [27], and one which plays an important role in our development. In particular, suppose $\{Z_t\}$ is operated on by a linear filter L with transfer function H. Then the output is again Gaussian with mean zero and its spectrum is given by $|H(\omega)|^2 dF(\omega)$. Let D_{H1} denote the number of zero-crossings in the filtered series $L(Z_1),\ldots,L(Z_N)$. Then (2) implies the representation

$$\cos\left(\frac{\pi ED_{H1}}{N-1}\right) = \frac{\int_{-\pi}^{\pi} \cos(\omega)\,|H(\omega)|^2 dF(\omega)}{\int_{-\pi}^{\pi} |H(\omega)|^2 dF(\omega)}. \tag{3}$$

(3) connects zero-crossings and linear filtering. Clearly, when L is the identity (or do nothing) filter then $H \equiv 1$ and (3) reduces to (2) as it should. We can now extend (3) to any sequential filtering. Thus the sequential filter $L_1 L_2 \ldots L_k$ with transfer function $H_1 \cdot H_2 \cdots H_k$ yields the generalization

$$\cos\left(\frac{\pi E D_{H_1 \ldots H_k 1}}{N-1}\right) = \frac{\int_{-\pi}^{\pi} \cos(\omega)\,|H_1(\omega)|^2 \ldots |H_k(\omega)|^2 dF(\omega)}{\int_{-\pi}^{\pi} |H_1(\omega)|^2 \ldots |H_k(\omega)|^2 dF(\omega)} \tag{4}$$

where $D_{H_1 \ldots H_k 1}$ is the zero-crossing count in the sequentially filtered time series. In subsequent sections some interesting properties and corollaries of (3), (4) are studied and discussed in some detail. Of central importance to us is the degree to which zero-crossings before and after filtering determine F.

The idea that zeros of filtered signals are of relevance in applications is not entirely new. For example, in the continuous time case in order to determine the expected number of extrema per unit time we simply find the expected number of zeros per unit time in the derivative of the random signal [19]. Upcrossings of differenced time series are discussed in [38] while in [39] a filter is applied prior to some counting procedures. The fact that an application of a filter affects the zero-crossing count is also recognized in [19]. (3) was introduced in [27], [40] for the purpose of defection of periodicities in time series. The continuous time analog of (2), (3) is known as Rice's formula [19]. Another analog for a non-Gaussian case is given in [41].

3.2. The Dominant Frequency Principle

When $F(\omega)$ is continuous at the origin, an assumption we shall adopt throughout the paper, symmetry implies that (2) can be expressed in terms of positive frequencies as in the expression

$$\cos\left(\frac{\pi E D_1}{N-1}\right) = \frac{\int_0^{\pi} \cos(\omega)\,dF(\omega)}{\int_0^{\pi} dF(\omega)}. \tag{5}$$

This shows clearly that the normalized expected zero-crossing rate $\pi ED_1/(N-1)$ is a *weighted average* of the spectral mass. Therefore, when a certain frequency band becomes dominant, i.e., carries more power than other bands, it attracts the normalized expected zero-crossings and $\pi ED_1/(N-1)$ admits values in this band. Likewise when a certain frequency ω_0, say, becomes significantly dominant then (5) implies that $\pi ED_1/(N-1) \simeq \omega_0$. In the extreme case when only ω_0 is present, that is, for $\omega \in [0,\pi]$

$$F(\omega+) - F(\omega-) > 0, \qquad \omega = \omega_0$$
$$= 0, \qquad \omega \neq \omega_0,$$

we have the equality

$$\frac{\pi ED_1}{N-1} = \omega_0 .$$

Thus, replacing ED_1 simply by D_1, we can see that when a certain frequency becomes dominant the quantity $\pi D_1/(N-1)$ will land at or near this frequency. In other words, a dominant frequency, when it exists in the spectrum, can be quickly detected by zero-crossings only. Now, when a time invariant linear filter is applied to a stationary time series it modifies the spectral weight given to the frequencies in the range $[0,\pi]$, emphasizing some bands while attenuating others. Consequently, when a discrete frequency exists in the spectrum it can be enhanced by a filter and then estimated by $\pi D_{H1}/(N-1)$. This tendency of zero-crossings (after proper scaling) to admit values in a neighborhood of a significantly dominant frequency will be called the *dominant*

frequency principle. This principle can be recognized as the basis of many of our theoretical results. See also [41], [42].

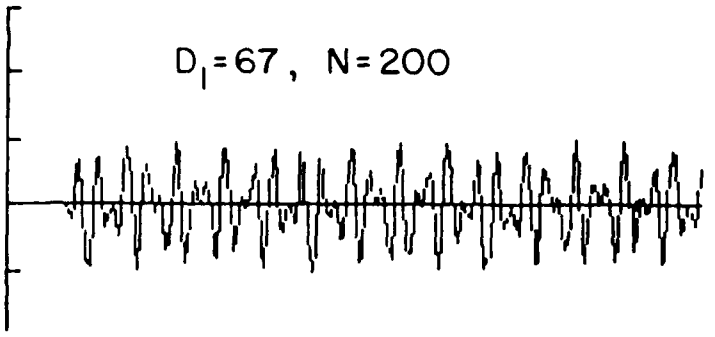
To illustrate this principle and its application in fast frequency detection consider the superposition

$$Z_t = A\cos(0.8t) + B\cos(1.25t), \qquad t = 1,\ldots,200. \tag{6}$$

Realizations of this series before and after filtering are plotted in Figure 3. The graphs were scaled but this has no effect on zero-crossings! From (2) or equivalently (5) we know that D_1 is a function of the weights A, B. When $A = B = 1$, no frequency is dominant. We obtain from Figure 3 (a) $D_1 = 67$ and $\pi D_1/199 = 1.057$ which is between 0.8 and 1.25 as expected. When $A = 0.8$ and $B = 1$, the frequency 1.25 is dominant. From Figure 3 (b) we obtain $D_1 = 79$ and $\pi D_1/199 = 1.247$ is very close to the dominant frequency in agreement with the dominant frequency principle. By operating on this second series with a low-pass filter we can render the frequency 0.8 dominant. Figure 3 (c) gives the scaled graph of the series in (b) after it was operated on with a low-pass linear filter with transfer function $(1 + e^{-i\lambda})^6$. Now the frequency 0.8 is dominant, $D_{H1} = 51$ and $\pi D_{H1}/199 = 0.805$ as expected.

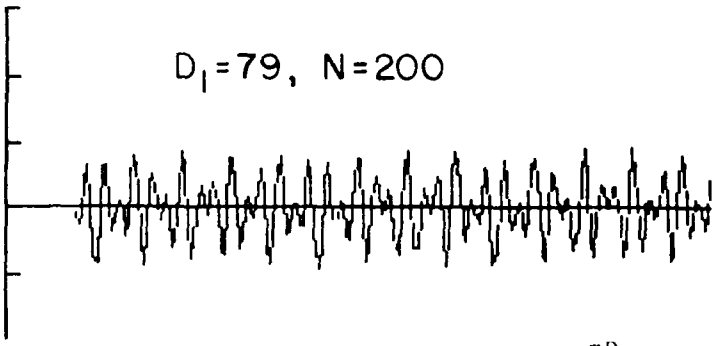
At this point the reader may be puzzled by the fact that (6) does not at all look random while our theory pertains to random time series. This ambiguity, should it arise, will be resolved in Section 4.1.

(a)



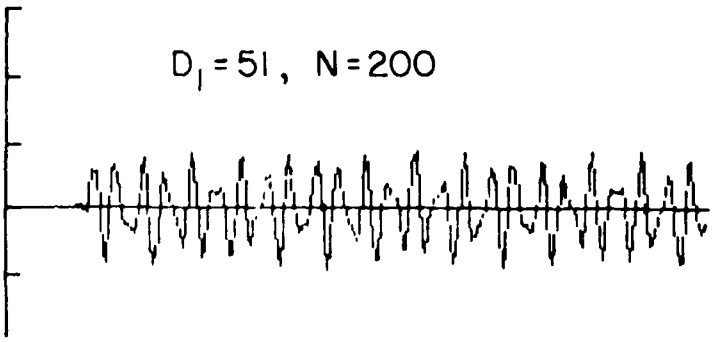
$Z_t = \cos(0.8t) + \cos(1.25t)$. Neither frequency is dominant. $\frac{\pi D_1}{N-1} = 1.057$.

(b)



$Z_t = 0.8\cos(0.8t) + \cos(1.25t)$. 1.25 is dominant. $\frac{\pi D_1}{N-1} = 1.247$.

(c)



Low-pass filtering of (b). 0.8 is dominant. $\frac{\pi D_{H1}}{N-1} = 0.805$.

Figure 3. Demonstration of the dominant frequency principle. Zero-crossing counts detect dominant frequencies. All graphs are scaled.

3.3. Zero-crossings of Repeatedly Differenced Series

A useful filter associated with zero-crossings is the difference operator. To introduce this operator it is convenient to define the shift operator B as

$$BZ_t = Z_{t-1}$$

with transfer function $e^{-i\omega}$. The difference operator ∇ is defined by $\nabla \equiv 1 - B$ with transfer function $1 - e^{-i\omega}$. Then the first difference of $\{Z_t\}$ is

$$\nabla Z_t \equiv (1 - B)Z_t = Z_t - Z_{t-1}$$

and the second difference is

$$\nabla^2 Z_t = \nabla(\nabla Z_t) \equiv (1 - B)^2 Z_t = Z_t - 2Z_{t-1} + Z_{t-2}.$$

In general the k'th difference of $\{Z_t\}$ is given by

$$\nabla^k Z_t \equiv (1 - B)^k Z_t = \sum_{j=0}^{k} \binom{k}{j} (-1)^j Z_{t-j}.$$

Define

$$D_k = \#\text{ of zero-crossings by } \nabla^{k-1} Z_t, \quad t = 1, \ldots, N.$$

Then D_1 is as before the number of zero-crossings in $Z_1, \ldots, Z_N$ while D_2 is the number of zero-crossings in $\nabla Z_1, \ldots, \nabla Z_N$ and D_3 is the number of zero-crossings in $\nabla^2 Z_1, \ldots, \nabla^2 Z_N$ and so on. The D_k are called higher order crossings [43], [44]. When $\{Z_t\}$ is first operated on by a filter L with transfer function H and is then repeatedly differenced, we use the notation D_{Hk} to denote higher order crossings in the filtered series $L(Z_t)$.

Unlike the process $\{Z_t\}$ which is defined over all the integers, in practice we only have finite time series and lose an observation with each difference. To avoid end-effects we must discard the beginning of the series, and start indexing the observations or samples by moving to the right. Thus if it is desired to evaluate k higher order crossings the time index $t = 1$ is given to the k'th or to a later observation. The D_j must be evaluated from differenced records *of the same length*. For example, suppose the given record is

$$1 \quad 6 \quad 1 \quad 7 \quad 8 \quad 9 \quad 2 \quad 3 \quad 0 \quad 7.$$

Subtracting the average we have the centered series

$$-3.4 \quad 1.6 \quad -3.4 \quad 2.6 \quad 3.6 \quad 4.6 \quad -2.4 \quad -1.4 \quad -4.4 \quad 2.6 \ .$$

In order to evaluate, say, D_1, D_2, D_3 we record the series starting with $Z_1 = -3.4$ and $Z_2 = 2.6$ while reserving the first -3.4 and 1.6 for Z_{-1} and Z_0 respectively. We can now evaluate Z_t, ∇Z_t, $\nabla^2 Z_t$ for $t = 1,\ldots,8$ as

Z:	-3.4	2.6	3.6	4.6	-2.4	-1.4	-4.4	2.6
∇Z:	-5.0	6.0	1.0	1.0	-7.0	1.0	-3.0	7.0
$\nabla^2 Z$:	-10.0	11.0	-5.0	0.0	-8.0	8.0	-4.0	10.00

Then $D_1 = 3$, $D_2 = 5$, $D_3 = 7$ and are all evaluated from records of length 8. Observe that by our definition a shift from a negative value to 0 is counted as a crossing. Similar remarks apply to other filters as well.

Since ∇ is a high-pass filter it pushes the spectral mass to the right so that $\{\nabla Z_t\}$ becomes more oscillatory than $\{Z_t\}$. We therefore expect that $D_2 > D_1$. ∇^2 amplifies high frequencies even more than ∇ so that $\{\nabla^2 Z_t\}$ is more oscillatory than $\{\nabla Z_t\}$ and we expect that $D_3 > D_2$, and so on. Also, if ω^*, say, is the highest frequency present in the spectrum, we see that by applying to $\{Z_t\}$ the filter ∇^k and letting $k \to \infty$, the power is pushed all the way to ω^* and renders it dominant! Therefore by the dominant frequency principle $\pi D_j/(N-1)$ should converge in some sense to ω^*. This heuristic argument is made precise in terms of the sequence of expected higher order crossings $\{ED_j\}$ [27], [28].

Theorem 1. Suppose $\{Z_t\}$, $t = \ldots,-1,0,1,\ldots$, is stationary and Gaussian with mean 0 and suppose ω^* is the highest frequency in the spectrum, $\omega^* \leq \pi$. Then

(i) $0 \leq ED_1 \leq ED_2 \leq \ldots \leq (N-1)$

(ii) $\dfrac{\pi ED_j}{N-1} \to \omega^*, \quad j \to \infty.$

The theorem says that the sequence $\{\pi ED_j/(N-1)\}$ is monotone and bounded and therefore it converges to its least upper bound which is ω^*. Under fairly general conditions we can also show [27] that

$$\lim_{N\to\infty} \lim_{j\to\infty} \mathrm{Var}(D_j/N) = 0.$$

It follows that $\pi D_j/(N-1)$ (without "E") approximates ω^* for sufficiently large j and N. As we shall see in the next section this estimator converges as j increases remarkably fast provided ω^* is

sufficiently removed away from π. This discussion explains why D_2 is in general a better choice than D_1 in the estimation of the highest frequency [25], [26], [27].

The monotone property of the sequence $\{ED_j\}$ in (i) follows from (1) and the relation

$$\rho_1 \geq \rho_\nabla(1) \tag{7}$$

where $\rho_\nabla(1)$ is the first correlation in $\{\nabla Z_t\}$ [28]. In general the inequality

$$D_{j+1} \geq D_j - 1 \tag{8}$$

holds for $j \geq 1$. Experience shows however that we really have $D_{j+1} \geq D_j$ for sufficiently long series; e.g. $N = 500$ [43], [44].

Figure 4 gives the graphs of two time series of length $N = 200$ from a stochastic difference equation

$$Z_t = \phi Z_{t-1} + u_t \tag{9}$$

where $\{u_t\}$ are uncorrelated normal random variables with mean zero and $\phi = \pm 0.5$. It is seen that even for these relatively short series the first few D_j tend not to decrease. (9) satisfies the requirement for being a stationary Gaussian process when $|\phi| < 1$ [45].

When some of the D_j "touch", it is an indication of a strong periodic component in the data. Consider the extreme case of a sinusoid with a certain frequency. Then the differences are again sinudoids with the same frequency but different amplitudes, and by the dominant frequency principle *all* the ED_j are the same provided

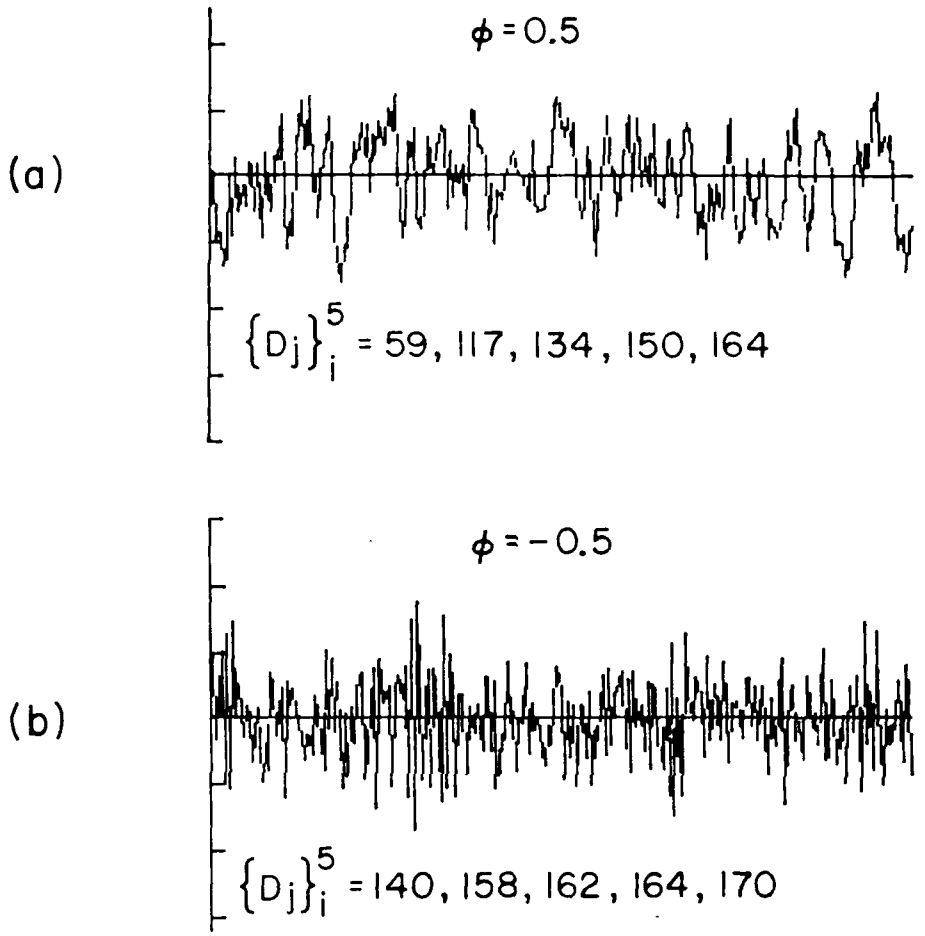


Figure 4. Higher order crossings from (9), N = 200.

they are obtained from differenced series of the same length. For a Gaussian series, we can show that the converse holds as well [28], [29]. This is phrased in the following theorem which is called the Sinusoidal Limit Theorem (SLT).

Theorem 2 (SLT). Let Z_t be a zero-mean Gaussian stationary process. Assume $ED_1 > 0$, and suppose

$$\frac{ED_1}{N-1} = \frac{ED_2}{N-1}. \tag{10}$$

Then $\{Z_t\}$ is a pure sinusoid with period $2(N-1)/ED_1$.

The surprising fact is that when (10) holds *all* the ED_j are equal. A more general Sinusoidal Limit Theorem of which Theorem 2 is a special case is due to Slutsky [46].

Perhaps the most general statement that can be made about higher order crossings is that they determine uniquely the spectral distribution function F up to a constant. To see that, first note that the higher order crossing spectral representation is given by

$$\cos\left(\frac{\pi ED_{k+1}}{N-1}\right) = \frac{\int_{-\pi}^{\pi} \cos(\omega)(\sin \omega/2)^{2k} dF(\omega)}{\int_{-\pi}^{\pi} (\sin \omega/2)^{2k} dF(\omega)} \tag{11}$$

from which we obtain after some algebra the "long formula" [46],

$$\cos\left(\frac{\pi ED_{k+1}}{N-1}\right) = \frac{-\binom{2k}{k-1} + \rho_1\left[\binom{2k}{k} + \binom{2k}{k-2}\right] - \ldots + (-1)^k \rho_{k+1}}{\binom{2k}{k} - 2\rho_1\binom{2k}{k-1} + \ldots + (-1)^k 2\rho_k}. \tag{12}$$

$$= -\frac{\nabla^{2k}\rho_{k-1}}{\nabla^{2k}\rho_k}$$

where ∇ now operates on ρ_k and ρ_{k-1}. (12) provides a recursion for obtaining $\rho_1, \rho_2, \ldots$ from $ED_1, ED_2, \ldots$. For example, for $k = 0$ (12) reduces to (1) so that ρ_1 is determined from ED_1. For $k = 1$ (12) gives

$$\rho_2 = 1 - 2(1 - \cos(\frac{\pi ED_1}{N-1}))(1 + \cos(\frac{\pi ED_2}{N-1})),$$

while for $k = 2$ ρ_3 is determined similarly from ED_1, ED_2, ED_3. In general ρ_k is determined from $(ED_1,\ldots,ED_k)$. Now recall that ρ_k is the k'th Fourier coefficient of $\bar{F}(\omega) \equiv F(\omega)/\gamma_0$, and so the ρ_k uniquely determine $\bar{F}$. But from (12) the sequence $\{\rho_k\}$ is uniquely determined by the sequence of higher order crossings $\{ED_j\}$. It follows that F is completely determined by $\{ED_j\}$. We have

Theorem 3. For a zero mean stationary Gaussian process, the sequence of expected higher order crossings $\{ED_j\}$ uniquely determines the normalized spectral distribution function $\bar{F}$.

The above discussion can be summarized by the symbolism

$$\{ED_j\} \iff \{\rho_k\} \iff \bar{F} \tag{13}$$

which is another indication that zero-crossings of filtered series contain useful information about the spectral properties of the process. (13) may be viewed as a ramification of the Wiener-Khintchine relation, and is an evidence of the existence of a domain to which we refer rather loosely as the D-domain. The D_j however are by no means the only features in this domain as is evident from the next section.

3.4. Zero-crossings of Repeatedly Summed and Differenced Series

The counterparts of the D_j are zero-crossings of repeated summation. Consider the summation filter $(1+B)$ which gives

$$(1+B)Z_t = Z_t + Z_{t-1}.$$

Define

$$_jD \equiv \#\text{ of zero-crossings by } (1+B)^{j-1}Z_t, \quad t = 1,\ldots,N.$$

Then $_1D \equiv D_1$, and again we refer to the $_jD$ as higher order crossings. Whereas the D_j tend to increase, the $_jD$ tend to decrease, but in many respects these two types of higher order crossings are quite analogous. Thus, the counterpart of Theorem 1 is the following fact.

Theorem 4. Suppose $\{Z_t\}$, $t = \ldots,-1,0,1,\ldots$, is stationary and Gaussian with mean 0, and suppose $_*\omega > 0$ is the lowest frequency in the spectrum. Then,

(i) $E\,_1D \geq E\,_2D \geq \ldots > 0$

(ii) $\frac{\pi E\,_jD}{N-1} \to {}_*\omega, \quad j \to \infty.$

Evidently Theorem 4 is another manifestation of the dominant frequency principle, and provides a way for estimating the lowest frequency. Again we can obtain a "long formula" as in (12) in terms of $\{E\,_jD\}$, by replacing in (12) all the minus signs by plus signs, which establishes the equivalence of this sequence of zero-crossings and $\bar{F}$.

There is no reason why we cannot put the D_j and $_jD$ together to produce more versatile features. A more general definition of higher order crossings is the following. Consider the filter defined by m differences which follow n sums,

$$(1-B)^m(1+B)^n \tag{14}$$

whose squared gain is given by

$$|G(\omega)|^2 = 2^{m+n}(1 - \cos\omega)^m(1 + \cos\omega)^n. \qquad (15)$$

Define the sequence $\{{}_jD_k\}$ by

$${}_jD_k \equiv \#\text{ of zero-crossings in } (1-B)^{k-1}(1+B)^{j-1}Z_t, \quad t = 1,\ldots,N.$$

Again we refer to the ${}_jD_k$ as higher order crossings. Evidently, by definition

$$D_1 \equiv {}_1D \equiv {}_1D_1$$

because in these cases no filtering is applied to $\{Z_t\}$. When m,n are large and $m/n = c$ then $\pi\,{}_1D_1/(N-1)$ will tend to admit a value in the neighborhood of λ_c where

$$\lambda_c = \cos^{-1}\left(\frac{1-c}{1+c}\right),$$

provided λ_c is a point in the spectrum. This is so since $|G(\omega)|^2$ in (15) is unimodal with a peak occurring at λ_c which makes this frequency dominant.

When prior to the application of (14) the series is operated on by a linear filter with transfer function H we shall use the notation

$${}_jD_{Hk}$$

to denote the higher order crossings in the filtered series. If the filter is just the summation filter $1+B$, then clearly ${}_1D_{Hk} \equiv {}_2D_k$.

4. SPECTRAL ANALYSIS BY ZERO-CROSSINGS

Our discourse leads us naturally to the important problem of spectrum analysis. The relevance of higher order crossings to this problem will become self-evident. A highlight of this section is the design of a filter used in the delicate problem of detecting a very weak signal from higher order crossings.

4.1. Analysis of Discrete Spectra by Zero-crossings

We are now finally ready to apply the above results in discrete spectrum estimation. For this purpose consider the process

$$Z_t = \sum_{j=1}^{p} (A_j \cos \omega_j t + B_j \sin \omega_j t), \tag{16}$$

$t = \ldots,-1,0,1,\ldots,$ where the ω_j and p are constants and $\{A_j\}$, $\{B_j\}$ are taken as uncorrelated normal random variables such that

$$EA_j = EB_j = 0 \qquad \text{for all } j$$

$$EA_iA_j = EB_iB_j = 0, \qquad i \neq j$$

$$= \sigma_j^2, \qquad i = j$$

and

$$EA_iB_j = 0 \qquad \text{for all } i,j.$$

Further, without loss of generality we assume that $0 < \omega_1 < \omega_2 < \ldots < \omega_p \leq \pi$. It follows [32] that $\{Z_t\}$ is a stationary Gaussian process with mean zero. Its spectral distribution function is a step function with jumps of magnitude of size $\frac{1}{2}\sigma_j^2$ at $\pm\omega_j$, $j = 1,\ldots,p$. In order to realize Z_t, $t = 1,\ldots,N$, it is useful to think of the $\{A_j\}$, $\{B_j\}$ as first being determined by some random

mechanism so that they become just constants throughout each realization. This resolves the ambiguity mentioned earlier, for the constant values there may be thought of as being the constant *values* of random variables. The point being made is that what we *actually observe* is a trigonometric polynomial of the form (16) with some unknown coefficients. Processes of the form (16) are ideal for an illustration of all the foregoing discussion.

The spectral representation of the ED_j (11) reduces to

$$\cos\left(\frac{\pi ED_{j+1}}{N-1}\right) = \sum_{r=1}^{p} W_r(j,\underset{\sim}{\omega})\cos(\omega_r), \tag{17}$$

where

$$W_r(j,\underset{\sim}{\omega}) \equiv \frac{1}{1 + \sum\limits_{\substack{i=1 \\ i\neq r}}^{p} \left(\frac{\sigma_i}{\sigma_r}\right)^2 \left[\frac{\sin\,\omega_i/2}{\sin\,\omega_r/2}\right]^{2j}}.$$

An interesting representation analogous to (17) has been obtained in [41]. It is seen that $W_r(j,\underset{\sim}{\omega}) \geq 0$ and $\sum\limits_{r=1}^{p} W_r(j,\underset{\sim}{\omega}) = 1$. Also, since the ω_i are ordered in $(0,\pi]$ and since $\sin(x/2)$ is monotone there we have

$$W_r(j,\underset{\sim}{\omega}) \to \begin{cases} 0, & r \neq p \\ \\ 1, & r = p \end{cases}, \qquad j \to \infty.$$

Therefore as $j \to \infty$

$$\cos\left(\frac{\pi ED_{j+1}}{N-1}\right) \to \cos(\omega_p),$$

or, since $\cos(x)$ is monotone in $[0,\pi]$,

$$\frac{\pi ED_j}{N-1} \to \omega_p, \qquad j \to \infty, \tag{18}$$

and we have demonstrated (ii) of Theorem 1 in the purely discrete spectrum case. That is, the normalized ED_j converge to the highest angular frequency ω_p. Similarly, directly or from (ii) of Theorem 4

$$\frac{\pi E_j D}{N-1} \to \omega_1, \qquad j \to \infty. \tag{19}$$

Now (2) reduces by (17) to

$$\cos\left(\frac{\pi ED_1}{N-1}\right) = \frac{\sigma_1^2 \cos \omega_1 + \ldots + \sigma_p^2 \cos \omega_p}{\sigma_1^2 + \ldots + \sigma_p^2}. \tag{20}$$

Suppose $\sigma_r \to \infty$. That is, suppose ω_r becomes dominant as expressed by associating more power with the amplitudes A_r, B_r. Then (20) obviously converges to $\cos(\omega_r)$ and we have by monotonicity

$$\frac{\pi ED_1}{N-1} \to \omega_r, \qquad \sigma_r \to \infty, \tag{21}$$

in accordance with the dominant frequency principle. Another useful fact due to (17) is that

$$\omega_1 \le \frac{\pi ED_1}{N-1} \le \frac{\pi ED_2}{N-1} \le \ldots \le \omega_p. \tag{22}$$

Observe that $\omega_1 = \omega_p$ if and only if $\pi ED_1/(N-1) = \pi ED_2/(N-1)$ by the SLT.

(18), or equivalently (19), provides a readily available method for obtaining all the ω's *and* p as shown in the following algorithm. We assume that N is sufficiently large so that $ED_j/(N-1)$ in (18) can be replaced by $D_j/(N-1)$ for all practical purposes.

Algorithm 1. Computation of $\omega_1,\ldots,\omega_p$ and p given $Z_1,\ldots,Z_N$ from the harmonic process (16). Assume $0 < \omega_1 < \omega_2 < \ldots < \omega_p \leq \pi$.

Step 0: Let $p^* = 0$.

Step 1: Determine ω_p from $\lim_{j\to\infty} \pi D_j/(N-1)$.

If $\omega_p = 0$, go to Step 4.

If $\omega_p \neq 0$, let $p^* = p^* + 1$. Print ω_p.

Step 2: Filter out ω_p with an ideal low-pass filter with cutoff frequency ω_p^-.

Step 3: Replace p by $p - 1$. Go to Step 1 with the filtered series obtained in Step 2.

Step 4: Let $p = p^*$. Print p. End computation.

When p is known apriori the algorithm can be easily amended to stop after the detection of p frequencies. Also, in practice we do not have ideal filters and must be content with approximations [33], [35].

To demonstrate the essence of the algorithm consider the simpler case when p is known. Suppose (16) is of the form

$$Z_t = \cos(0.8t) - 1.3\cos(1.25t) + \sin(2t), \quad t = 1,\ldots,200, \tag{23}$$

where $p = 3$. This is a fairly short series and it is interesting to see what the algorithm produces in this case. Instead of a close to ideal low-pass filter we use the much cruder low-pass filter $(1+B)^k$ which attenuates high frequency and which suffices for this demonstration. The results with only eight higher order crossings are given in Table 1 and are quite satisfactory given the length of the data. The column corresponding to $_{31}D_j$ starts too

j	D_j	$_7D_j$	$_{31}D_j$
1	77	50	47
2	81	61	48
3	89	79	48
4	108	79	48
5	126	80	49
6	126	80	49
7	127	80	49
8	127	79	50
Estimated highest frequency	$\hat{\omega}_3 = 2.005$	$\hat{\omega}_2 = 1.247$	$\hat{\omega}_1 = 0.789$

Table 1. Demonstration of Algorithm 1 with $p = 3$, $N = 200$, and $\omega_1 = 0.8$, $\omega_2 = 1.25$, $\omega_3 = 2$. The ω_j estimates are obtained by multiplying the last D in each column by $\pi/199$.

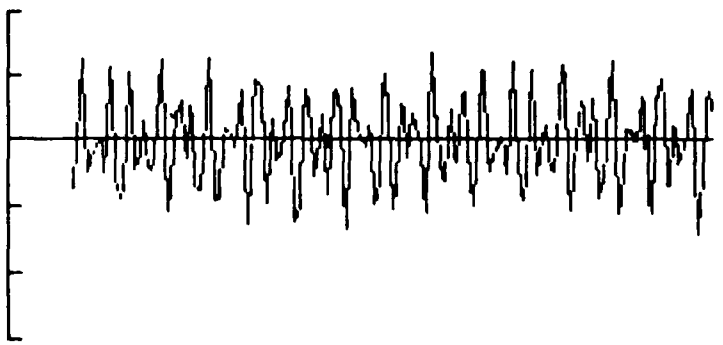


Figure 5. The series in (23). N = 200.

low due to an end effect produced by the shortness of the original series which was 220. For another example using the sine-Butterworth low-pass filter see [40].

4.2. Mixed Spectrum Analysis by Zero-crossings

Consider again the harmonic process (16), but suppose we now add to it a noise component to form a signal plus noise process,

$$Z_t = \sum_{j=1}^{p} (A_j \cos \omega_j t + B_j \sin \omega_j t) + \varepsilon_t, \qquad (24)$$

where the A_j, B_j, ω_j are as in (16) and $\{\varepsilon_t\}$ is white noise made of uncorrelated normal random variables with mean 0 and variance σ_ε^2. It is assumed that $\{\varepsilon_t\}$ is independent of the $\{A_j\}$, $\{B_j\}$. Then $\{Z_t\}$ is again a stationary Gaussian process with zero mean since it is the sum of two independent Gaussian components, the signal and the noise. As before the problem is to determine the ω_j and p. This time however, the problem is more involved since $\{\varepsilon_t\}$ adds to the process a continuous spectrum whose highest frequency is π. Therefore by Theorem 1, $\pi D_j/(N-1)$ will converge to π and not to ω_p. It is true that ω_p is still the highest discrete frequency, but this is of no great help now. It follows that the technique of the previous section is not in general suitable for the mixed spectrum case.

Luckily enough, the dominant frequency principle prevails in this case too but in a rather subtle way which requires an explanation.

The zero-crossing spectral representation (2) for the signal plus noise process (24) reduces to

$$\cos\left(\frac{\pi E D_1}{N-1}\right) = \frac{\sigma_1^2 \cos\omega_1 + \ldots + \sigma_p^2 \cos\omega_p}{\sigma_1^2 + \ldots + \sigma_p^2 + \sigma_\varepsilon^2}, \qquad (25)$$

from which we can see directly that when ω_r becomes dominant (by rendering σ_r^2 larger than σ_ε^2 and σ_j^2, $j \neq r$) it can be estimated by $\pi D_1/(N-1)$. Evidently, our problem is solved if we can render the ω's dominant in succession starting with ω_1.

Let us envision an ideal situation where ω_1 becomes dominant due to a low-pass filter which pushes the spectral mass to the left. Then $\pi D_{H1}/(N-1)$ will detect it. By applying the high-pass difference filter ∇ the power shifts to the right and suppose ω_2 now becomes dominant instead. Then $\pi D_{H2}/(N-1)$ will detect it. Ideally, we apply ∇ again, push the power even further to the right, render ω_3 dominant and detect it by $\pi D_{H3}/(N-1)$. This iterative process can be repeated until all the ω's are detected by the D_{Hj}. This idealization cannot of course be expected to hold exactly, but the degree to which it does hold is surprising. The fact observed in numerous cases is that the normalized higher order crossings $\pi D_{Hj}/(N-1)$ tend to "visit" the ω's on their way to the highest frequency as j increases. The same holds true for $\pi_k D_H/(N-1)$ by symmetry, except that now the sequence tends to "visit" the ω's on its way to the lowest frequency as k increases. This fidelity of higher order crossings has been first observed in [27].

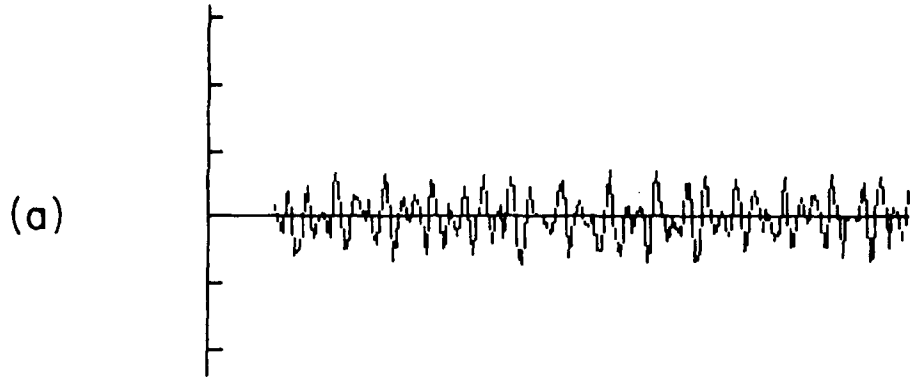
To demonstrate the curious tendency of normalized D_{Hj} to either land on or pass very near the ω_j, consider again the process in (6) with additive random noise,

$$Z_t = 0.8\cos(0.8t) + \cos(1.25t) + \varepsilon_t, \quad t = 1,\ldots,200, \tag{26}$$

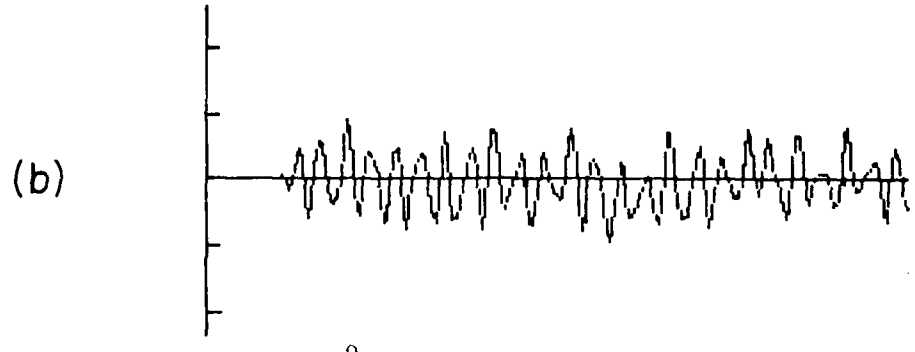
$\mathrm{Var}(\varepsilon_t) = 1/12$. For a low-pass filter we use $(1+B)^9$ so that $D_{Hj} \equiv {}_{10}D_j$. From Table 2 we see that without prefiltering the sequence $\{\pi D_j/199\}$ starts off near the dominant frequency 1.25 and then increases toward π. Consequently no information is available about the smaller frequency 0.8. But by applying the low-pass filter, the sequence $\{\pi_{10}D_j/199\}$ starts off near 0.8 and then as it increases it also visits the second 1.25. Thus both frequencies were visited. For more examples see [27].

	No filtering		Application of $(1+B)^9$	
j	D_j	$\pi D_j/199$	${}_{10}D_j$	$\pi_{10}D_j/199$
1	78	1.231	51	0.805 (*)
2	79	1.247 (*)	51	0.805 (*)
3	95	1.500	66	1.042
4	110	1.737	78	1.231
5	130	2.052	79	1.247 (*)
6	144	2.273	79	1.247 (*)
7	154	2.431	80	1.263
8	160	2.526	80	1.263
9	164	2.589	79	1.247 (*)
10	166	2.621	82	1.295

Table 2. Higher order crossings from (26) visit ω_1 and ω_2. A visit is indicated by (*).



Original series (26). N = 200.



(26) filtered by $(1+B)^9$.

Figure 6. The scaled series (26) before and after filtering.

4.3. Detection of a Weak Signal

The fact that the sequence $\{\pi_k D_{Hj}/(N-1)\}$ tends to contain values which are attracted to discrete frequencies in the spectrum, provides a powerful tool in detection. This will now be illustrated by applying the foregoing ideas in the estimation of a single frequency of a sinusoid buried in random noise. In this connection we make use of the periodogram evaluated at higher order crossings,

$$I\left(\frac{\pi_k D_{Hj}}{N-1}\right) = \frac{2}{N}\left|\sum_{t=1}^{N} Z_t \exp\left(-\frac{i\pi_k D_{Hj}}{N-1}t\right)\right|^2, \quad j = 1,2,\ldots, \tag{27}$$

and $i = \sqrt{-1}$. When a certain $\pi_k D_{Hj}/(N-1)$ is close to a discrete frequency, the periodogram will in general be inflated. In this way we can tell which of the normalized higher order crossings landed at or near a discrete frequency.

Periodogram analysis is an important tool in time series analysis [47], [48]. Usually the periodogram is evaluated at points of the form $2\pi k/N$, $k = 0,\ldots,[N/2]$, and the idea is to observe significantly large periodogram ordinates. A point $2\pi k/N$ which gives rise to an inflated periodogram ordinate indicates that the data contain a periodic component with period N/k. In our use of (27), the periodogram is evaluated at only a few points which are not necessarily of the form $2\pi k/N$.

Consider the process (24) with $p = 1$,

$$Z_t = A_1 \cos \omega_1 t + B_1 \sin \omega_1 t + \varepsilon_t, \tag{28}$$

$t = 1,\ldots,N$, and again suppose we want to estimate ω_1 from zero-crossings. This problem was considered in [1], [3], [39], [41], [42], [49], [50]. It is actually possible [49] to solve for ω_1 in terms of D_1 and D_2 but it was shown in [50] that the solution may be extremely sensitive to small changes in these higher order crossings. The incorporation of certain filters gives more robust estimates.

The first order low-pass filter is defined by the recursion

$$L(Z_t) = (1-\alpha)Z_t + \alpha L(Z_{t-1}), \tag{29}$$

$0 < \alpha < 1$. Its squared gain is given by [33]

$$|G(\omega)|^2 = \frac{(1-\alpha)^2}{1 - 2\alpha\cos\omega + \alpha^2}.$$

The combination of this filter and repeated summation was found very useful in higher order crossing analysis owing much to the following fact [50].

Theorem 5. Let $\{Z_t\}$ be given by (28) and without loss of generality assume that $0 < \omega_1 < \pi/2$. Then regardless of the signal-to-noise ratio, $\sigma_1/\sigma_\varepsilon$, we have

(a) For α sufficiently close to 0

$$\omega_1 \leq \frac{\pi ED_{G1}}{N-1} \leq \cos^{-1}(\alpha)$$

with equality holding when $\alpha = \cos(\omega_1)$.

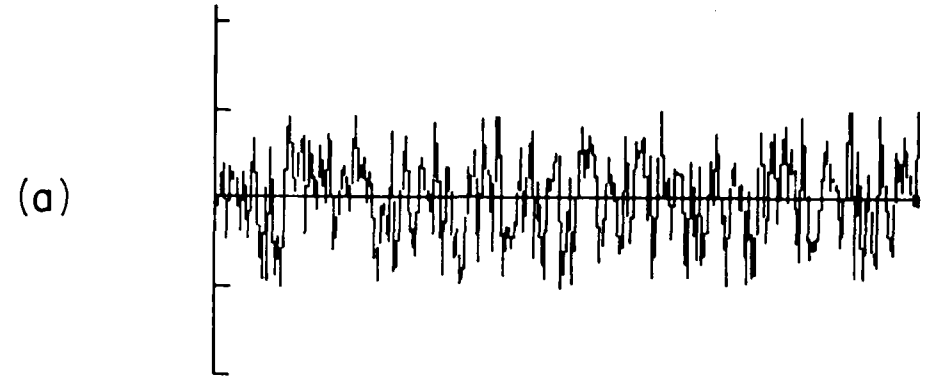
(b) For α sufficiently close to 1 the inequalities are reversed.

Theorem 5 tells us that as α increases away from 0 toward 1, $\pi D_{G1}/(N-1)$ has a good chance for "visiting" ω_1. Our strategy in light of the foregoing discussion is to increase α in small steps and then evaluate (27) at each step. When a visit occurs, the periodogram will be inflated. Since the theorem holds regardless of the signal-to-noise ratio, $\sigma_1/\sigma_\varepsilon$, this procedure gives hope in detecting even very weak signals. For small $\sigma_1/\sigma_\varepsilon$ it is perhaps better to prefilter with $(1+B)^k$ for some k, e.g. $k = 2$.

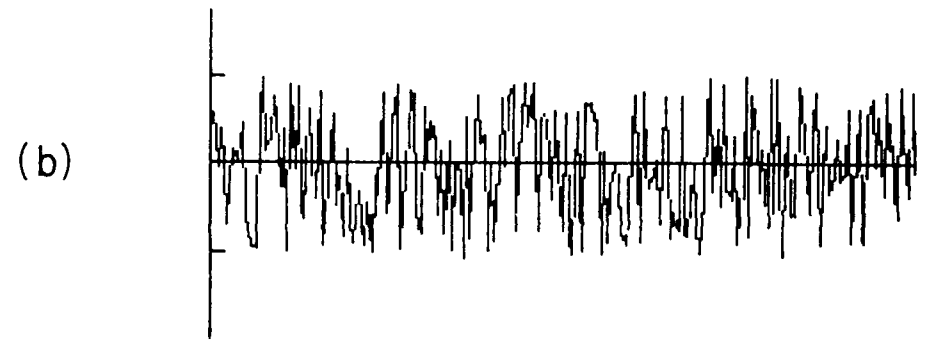
As an illustration consider a computer simulation of (28) with $\omega_1 = 0.75$, $\sigma_1/\sigma_\varepsilon = -20$dB, $N = 2000$. For each α-value, $\alpha = 1/30,\ldots,10/30$, we obtained the zero-crossings ${}_3D_{G1}$. The normalized ${}_3D_{G1}$ and the corresponding *relative* (out of 10 cases) periodogram ordinates are given in Table 3. The highest relative periodogram ordinate corresponds of course to 0.75 as it should. The more startling fact though is that the normalized ${}_3D_{G1}$ corresponding to $\alpha = 0.1$ landed exactly on ω_1.

α	$\pi\ {}_3D_{G1}/1999$	Relative periodogram ordinate
1/30	0.791	0.121
2/30	0.765	0.049
3/30	0.750 (*)	0.583
4/30	0.724	0.110
5/30	0.709	0.008
6/30	0.693	0.025
7/30	0.684	0.026
8/30	0.655	0.030
9/30	0.640	0.002
10/30	0.627	0.046
		1.000

Table 3. Relative periodogram ordinates evaluated at ten $\pi\ {}_3D_{G1}/1999$ from (28). $\sigma_1/\sigma_\varepsilon = -20$dB, $\omega_1 = 0.75$. (*) indicates a visit.



Completely random series.



The series (26) with $a_1 = 0.75$ and $\sigma_1/\sigma_\varepsilon = -20$ dB.

Figure 7. Random noise versus weak signal plus random noise.

The problem of detecting a weak periodic signal in noise by the so-called harmogram is considered also in [51]. An example suggested there is treated successfully by higher order crossings in detecting the fundamental frequency [27].

4.4. Search for Periodicities

As a final example of mixed spectrum analysis we consider briefly a geophysical time series whose graph is given in Figure 8. We do not attempt complete spectrum analysis but only further illustrate the interesting interplay between zero-crossings and linear filtering in a case of real data. In this example the series consists of 600 observations sampled every 10 minutes. The search for periodicities is conducted mostly by low-pass filtering followed by extraction of higher order crossings [40]. We confine ourselves to low-pass filters which are combinations of repeated summation, the first order low-pass filter (29) and repeated application of the sine Butterworth filter with squared gain [35]

$$|Q(\omega)|^2 = \left[\frac{1}{1 + \left(\frac{\sin \frac{1}{2}\omega}{\sin \frac{1}{2}\omega_b}\right)^{12}}\right]^L \tag{30}$$

where ω_b is an ideal cutoff point and L is the number of repetitions of the filter. For each combination of filters we extracted 30 higher order crossings and the periodogram (27) was evaluated for each of the 30 cases.

Starting with the filter $(1-B)(1+B)^{12}$ followed by (29), with $\alpha = 0.25$, the relative periodogram ordinate (27) gave about 90% weight to $\pi_{13}D_{G2}/(N-1) = 0.1049$. This corresponds to a period of $2\pi/(6 \times 0.1049) = 9.983$ hours.

For the same combination of filters but with $\alpha = 0.6$, the relative periodogram ordinate gave nearly 100% weight to $\pi_{13}D_{G2}/(N-1) = 0.0839$. This corresponds to a period of 12.481 hours.

To detect even lower significant frequencies (or higher periods) we used (30) with $\omega_b = 0.084$, $L = 2$ together with $(1-B)(1+B)^{10}$ and (29) with $\alpha = 0.1$. The relative periodogram ordinate gave about 50% weight to $\pi_{11} D_{GQ3}/(N-1) = 0.0427$ which corresponds to a period of 24.525 hours.

Therefore the data contain periodic components with periods of about 10, 12.5 and 24.5 hours. For a discussion of significance tests using periodogram ordinates in conjunction with higher order crossings, see [27].

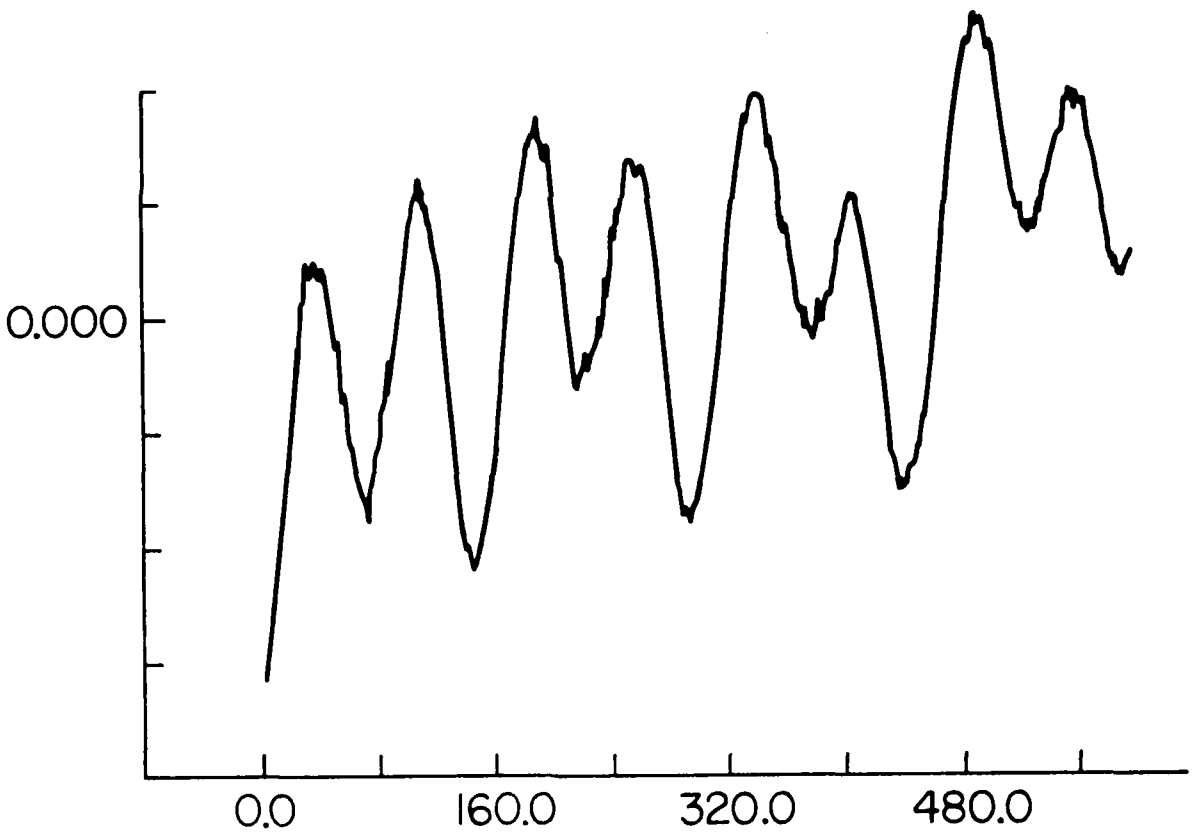


Figure 8. A geophysical time series.

4.5. A Continuous Spectrum Case

An important class of stationary stochastic processes is the one characterized by stochastic difference equations of the form

$$Z_t = \phi_1 Z_{t-1} + \ldots + \phi_p Z_{t-p} + \varepsilon_t. \tag{31}$$

When $\{\varepsilon_t\}$ is white Gaussian noise and the roots of $1 - \phi_1 x - \ldots - \phi_p x^p = 0$ lie outside the unit circle, $\{Z_t\}$ is stationary and Gaussian with mean zero [45]. (31) provides a useful model for many applications [45], [52], and is referred to as an autoregressive process of order p. This process has continuous spectrum (only!) with spectral density [32]

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \cdot \frac{1}{|1 - \phi_1 e^{-i\omega} - \ldots - \phi_p e^{-i\omega p}|^2}, \quad -\pi < \omega \leq \pi. \tag{32}$$

Except for the constant term $\sigma_\varepsilon^2/2\pi$, $f(\omega)$ is determined by higher order crossings in agreement with Theorem 3. To see that, define

$$\underset{\sim}{\phi} = \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_p \end{pmatrix}, \quad \underset{\sim}{\rho} = \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_p \end{pmatrix}, \quad \underset{\sim}{R} = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & & & \vdots \\ \rho_{p-1} & \rho_{p-2} & & \cdots & 1 \end{pmatrix}.$$

Then, the so-called Yule-Walker equations imply [45] that

$$\underset{\sim}{\phi} = \underset{\sim}{R}^{-1} \underset{\sim}{\rho}. \tag{33}$$

From (12) we know that $\rho_1, \ldots, \rho_p$ are uniquely determined by $ED_1, \ldots, ED_p$, and so $\phi_1, \ldots, \phi_p$ by (33) are completely determined from expected higher order crossings. This and the fact that

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \rho_1\phi_1 - \ldots - \rho_p\phi_p}$$

imply that the normalized spectral density ($\bar{f}(\omega) \equiv f(\omega)/\gamma_0$)

$$\bar{f}(\omega) = \frac{1 - \rho_1\phi_1 - \ldots - \rho_p\phi_p}{2\pi|1 - \phi_1 e^{-i\omega} - \ldots - \phi_p e^{-i\omega p}|^2} \tag{34}$$

is completely determined by $ED_1,\ldots,ED_p$. It follows that we can obtain fast estimates of (34) by solving for $\underset{\sim}{\rho}$ and $\underset{\sim}{\phi}$ in terms of the observed $D_1,\ldots,D_p$.

As an example consider the first order case of $p = 1$. Then (34) becomes

$$\bar{f}(\omega) = \frac{1 - \phi_1^2}{2\pi(1 - 2\phi_1\cos\omega + \phi_1^2)} . \tag{35}$$

But ϕ_1 can be estimated by [36]

$$\hat{\phi}_1 = \cos\left(\frac{\pi D_1}{N-1}\right) . \tag{36}$$

Thus $\bar{f}(\omega)$ can be estimated very fast from D_1 only.

Similarly, in the second order case a fast spectral estimate is obtained from D_1, D_2 by substituting these two higher crossings for ED_1, ED_2 in (12). From this we obtain in succession estimates for ρ_1, ρ_2 and then estimates for ϕ_1, ϕ_2 from (33), and finally an estimate for $\bar{f}(\omega)$ by substitution in (34).

This type of spectral estimate is reminiscent of autoregressive and maximum entropy spectral estimates [52], [53], but has not

previously been studied in depth. Experience suggests however that this procedure can be of value for small p.

5. SIGNAL DISCRIMINATION BY ZERO-CROSSINGS

In addition to the fact that higher order crossings contain useful spectral information, they also serve as a simple and effective means for achieving drastic data reduction important for discrimination and classification purposes. In what follows, we discuss various ways for using the higher order crossings $\{D_j\}$ in discrimination between time series by appealing to their monotone property and the rate at which they increase. It will become apparent that in general very few D_j are needed as their discrimination potency decreases rather rapidly as j increases.

Terms such as discrimination, classification, goodness of fit and signature analysis are sometimes used somewhat loosely to mean basically the same thing. By discrimination between time series we shall mean the process of determining from the D_j the degree of similarity between two time series, or between a given time series and an hypothesized one. For example, given the D_j of an engine signature, the discrimination problem is to determine how close these D_j are to the higher order crossings of a reliable engine.

Throughout this section it is assumed that the highest frequency is π as, for example, is always the case in the presence of white noise.

5.1. The Higher Order Crossings Theorem

To quantity in some sense the amount of information in the D_j pertaining to discrimination, it is convenient to introduce a binary process $\{X_t^{(j)}\}$ which we have been using implicitly all along. Define $\{X_t^{(j)}\}$ by clipping $\{\nabla^{j-1} Z_t\}$,

$$X_t^{(j)} \equiv \begin{cases} 1, & \nabla^{j-1} Z_t \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

Clearly $X_t^{(1)} \equiv X_t$, and D_j is precisely the number of symbol changes in $X_1^{(j)}, \ldots, X_N^{(j)}$. Of interest is the behavior of $\{X_t^{(j)}\}$ large j.

Because it is assumed that π is included in the spectral support, (ii) of Theorem 1 tells us that

$$\frac{ED_j}{N-1} \to 1, \qquad j \to \infty. \tag{37}$$

This and the SLT imply that in (i) of Theorem 1 the inequalities are strict,

$$ED_1 < ED_2 < \ldots, \tag{38}$$

since otherwise if $ED_j = ED_{j+1}$ for some j we immediately have a sinusoid and $ED_j/(N-1)$ will converge to the frequency of this sinusoid divided by π, which contradicts (37). Now, replacing ED_j by D_j, we expect from (37), (38) that as j increases $\{X_t^{(j)}\}$ approaches a *limiting state* of alternating symbols:

$$\ldots 010101 \ldots . \tag{39}$$

This degenerate state is independent of the initial process $\{Z_t\}$ as long as π is in the spectrum. This is stated formally in the Higher Order Crossings Theorem (HOCT) [44].

Theorem 6 (HOCT). Let $\{Z_t\}$, $t = \ldots,-1,0,1,\ldots$, be a stationary Gaussian process with mean zero, and assume that π is included in the spectral support. Then,

(i) $X_t^{(j)} \Longrightarrow \begin{cases} \ldots010101\ldots, & \text{with probability } 1/2 \\ \ldots101010\ldots, & \text{with probability } 1/2, \end{cases}$

as $j \to \infty$.

(ii) $\lim_{j\to\infty} \lim_{N\to\infty} D_j/(N-1) = 1$ with probability 1.

The theorem is demonstrated in two special cases in Figure 9 where we see that the degeneracy (39) is achieved rather fast. The binary arrays should be thought of as sections from infinite binary arrays while the rightmost column gives $D_1,\ldots,D_{16}$ from series of length $N = 1000$. The process from which these data were taken is (9) with $\phi = 0.8, 0.5$.

From the above discussion we see that discrimination based on D_j for large j is a fruitless effort since different processes yield almost the same D_j. Moreover, experience shows that even for moderate j the D_j are not very helpful. In Figure 9, for example, later D_j are relatively more similar than earlier ones for two different processes. In other words, the initial rate of convergence toward the degenerate state (39) is quite fast. This initial rate however provides a very useful discrimination measure and consequently only very few D_j are needed for effective discrimination [43], [44], [54]. Figure 10 gives the D_j, $j = 1,\ldots,10$, again from a first order autoregressive process (9) corresponding to different values of the parameter ϕ, when $N = 1000$. The rate of increase in the D_j

j	$\{X_t^{(j)}\}$	D_j	$\{X_t^{(j)}\}$	D_j
1	1001000001111111	207	0111000011111111	331
2	0011001101111111	513	0110001110100 10	567
3	0011001101010101	659	0100011110100 10	690
4	0110011001010101	715	0100011010101010	738
5	0110010011010001	745	0101010010101010	774
6	0100110011010001	773	0101010110101000	798
7	0100110111010111	807	0101010110101000	810
8	0101100110010111	823	0101010100101010	829
9	0101100100010100	829	0101010100101010	837
10	0101001101010100	849	0101010101101010	848
11	0101001101010100	855	0101010101101010	854
12	0101011101010100	865	0101010101001010	858
13	0101010001010100	875	0101010101001010	858
14	0101010101010100	883	0101010101011010	862
15	0101010101010100	885	0101010101011010	868
16	0101010101010100	893	0101010101010010	870
	$\phi = 0.8$		$\phi = 0.5$	
	(a)		(b)	

Figure 9. Demonstration of the Higher Order Crossings Theorem using a first order autoregressive process. N = 1000.

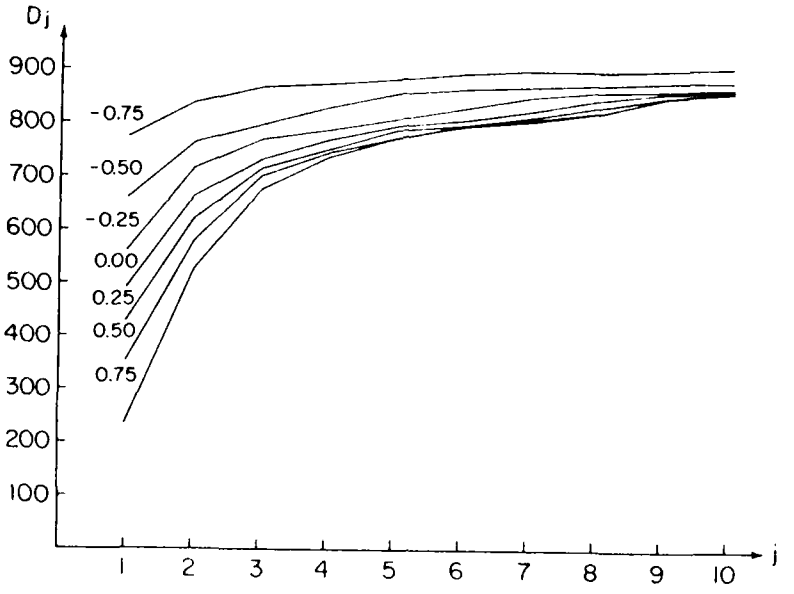


Figure 10. Plots of D_j, $j = 1,\ldots,10$, from a first order autoregressive process with parameter ϕ, N = 1000.

(compare with (38)) differs initially from process to process, but as j continues to increase the differences taper off.

Another example is provided by speech data obtained from the utterances of "zero", "one", and "two". See Figure 11 for a typical appearance of such data. The sampling rate was 8000 samples per second and we used 2000 samples in obtaining $D_1,\ldots,D_{20}$ in each of the three cases. The D_j in each of the three cases are plotted in Figure 12 and are seen to exhibit different initial rates of increase corresponding to the various cases. The figure also suggests that successful discrimination can be achieved by fewer than twenty higher order crossings since from D_{10} on the rate of increase in the three cases is very similar. In many cases in practice eight or even as few as six D_j suffice for effective discrimination.

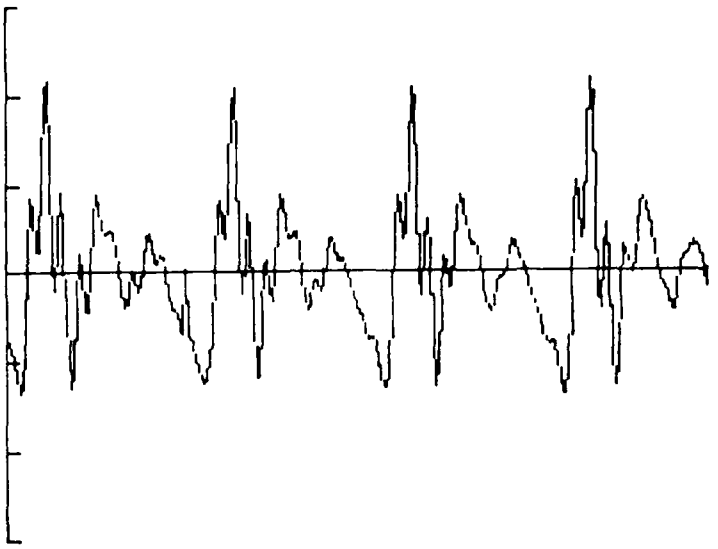


Figure 11. 250 samples from "zero" corresponding to 1/32 second.

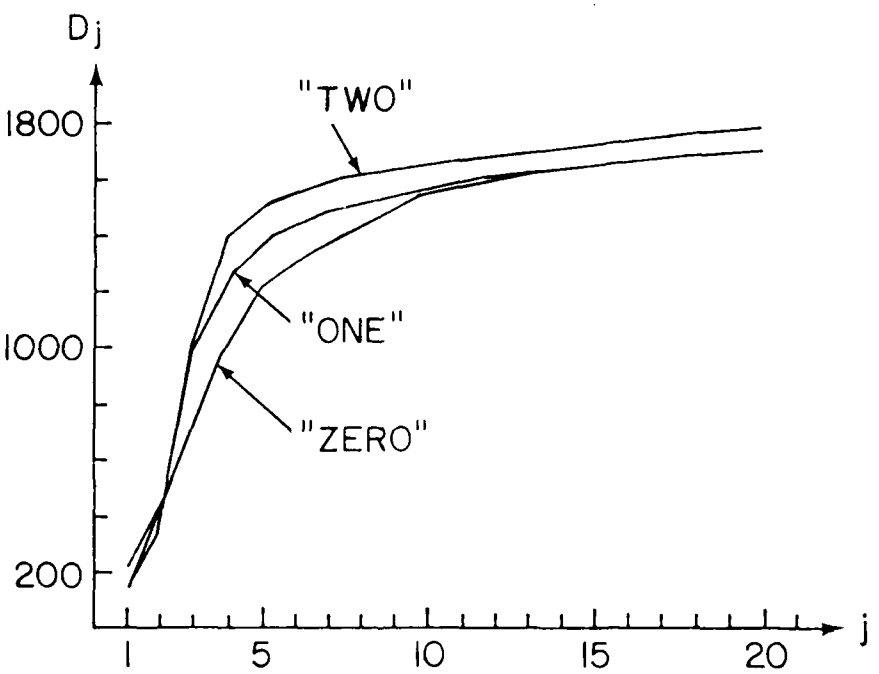


Figure 12. Plots of D_j, $j = 1,\ldots,20$, from speech data.

It is interesting to note that the apparent diminishing utility of higher order crossings for discrimination purposes is essentially a graphical fact. When observing the graph of a time series, the most conspicuous features are zero-crossings, peaks and troughs, and inflection points. But these three features correspond to D_1, D_2, D_3 and the inability of humans to discern much beyond these features parallels the fact that the D_j for large j have virtually no discrimination potency.

5.2. An Application to White Noise Testing

To measure the degree of similarity between time series, we can use plots such as those in Figures 10 and 12. By themselves, though, the plots are not sufficient since they do not provide a clear-cut decision rule. To circumvent this problem we can obtain probability limits which contain a given D_j plot with known probability. In a wide range of applications this goal can be achieved if the variance of D_j is known. Although simple closed form expressions for $\mathrm{Var}(D_j)$ do not exist, various approximations are available [54], [55]. A very useful and simple approximation can be obtained if $\rho_k \to 0$, as $k \to \infty$, sufficiently fast. Define

$$\lambda_1^{(k)} \equiv \Pr(X_t^{(k)} = 1 \mid X_{t-1}^{(k)} = 1).$$

Then it was shown in [55] that when $\rho_k = 0$, $k > m$, for some positive integer m, we have

$$\lim_{k \to \infty} \frac{\mathrm{Var}\, D_k}{(N-1)\lambda_1^{(k)}(1-\lambda_1^{(k)})} = 1. \tag{40}$$

Thus, for sufficiently large k

$$\mathrm{Var}(D_k) \sim (N-1)\lambda_1^{(k)}(1-\lambda_1^{(k)}). \tag{41}$$

It has been observed however that (41) provides a very good approximation for $\mathrm{Var}(D_k)$ also for small k in the white noise case.

Under the hypothesis of white noise D_k has an asymptotic normal distribution and also we know ED_k exactly. Thus we can use the approximation (41) in forming probability limits for D_k. Approximate 95% probability limits for D_k are given by [55]

$$(N-1)[\frac{1}{2} + \frac{1}{\pi}\sin^{-1}(\frac{k-1}{k})] \pm 1.96\{(N-1)[\frac{1}{4} - (\frac{1}{\pi}\sin^{-1}(\frac{k-1}{k}))^2]\}^{\frac{1}{2}}. \tag{42}$$

The hypothesis of white noise is rejected if at least one D_j, $j = 1,\ldots,K$, falls outside the limits (41). When all the D_j, $j = 1,\ldots,K$, fall inside the limits (42), the initial rate of increase in the D_j resembles that of white noise and the hypothesis of white noise is accepted. For statistical considerations regarding this test see [54], [55].

As an illustration of this graphical test, consider the geophysical series in Figure 8. The first ten D_j together with the limits (42) are plotted in Figure 13. Since the first two D_j lie outside the limits, the series is not white noise as is well expected.

As another example consider the series is Figure 14. For this case, the path of higher order crossings is completely inside the limits (42) and so the hypothesis of white noise is accepted.

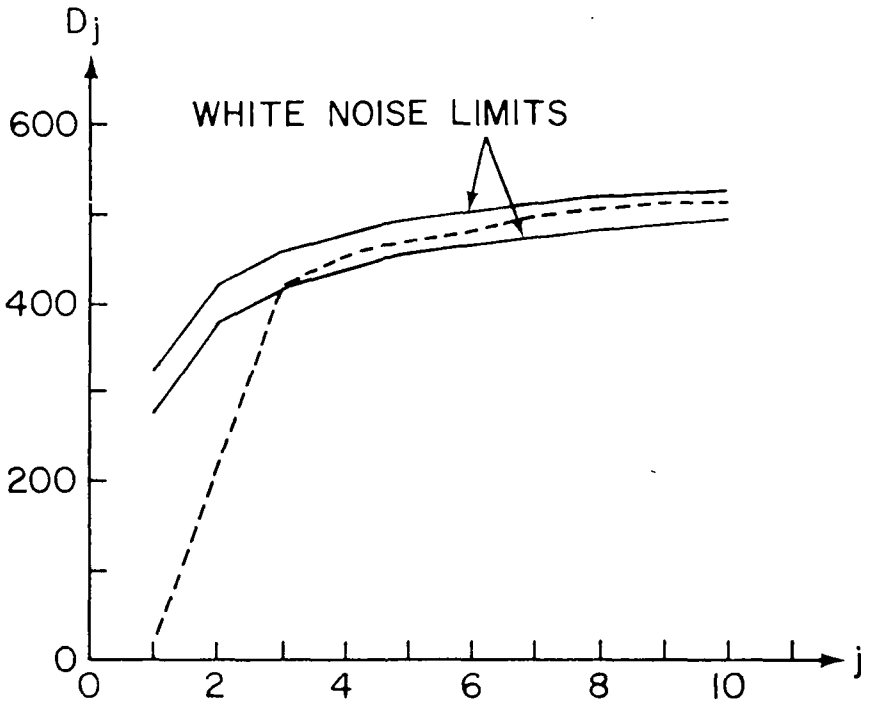


Figure 13. Higher order crossings plot (broken line) from the geophysical data. The hypothesis of white noise is rejected.

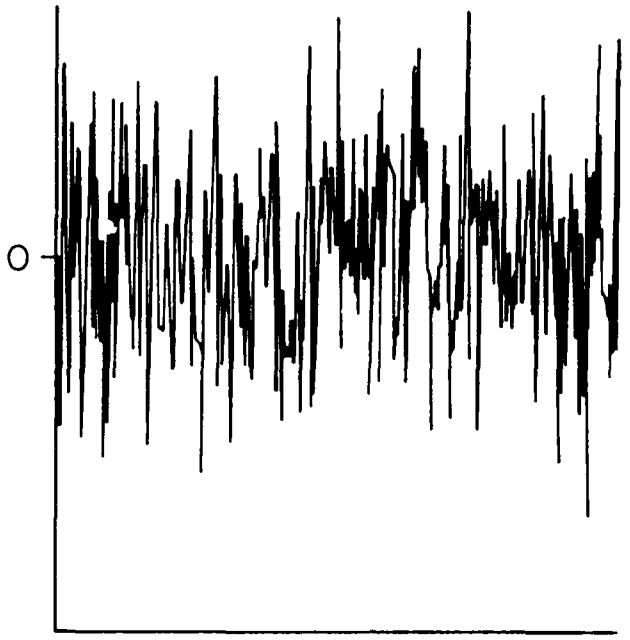


Figure 14. An oscillating time series.

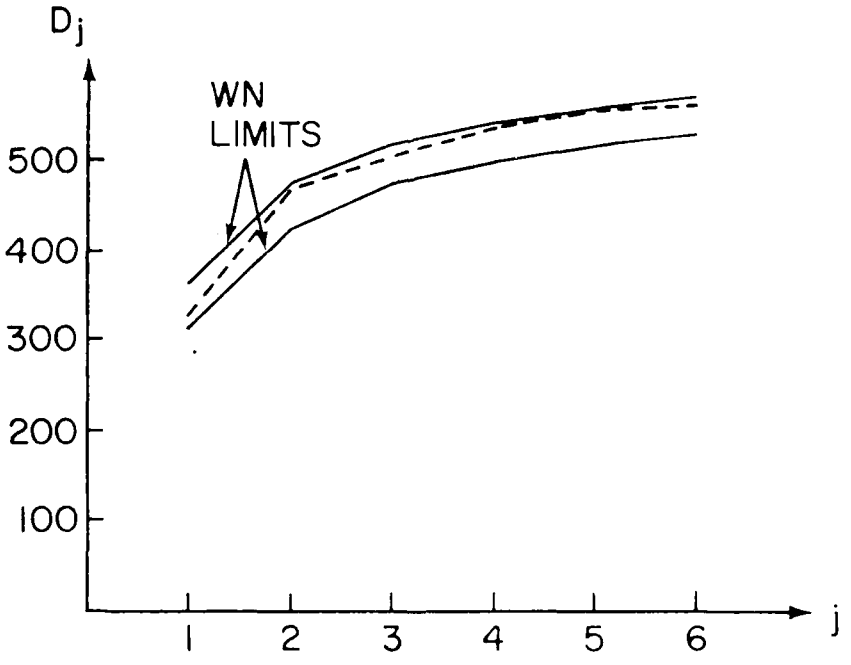


Figure 15. Higher order corssings plot from the series in Figure 14 (broken line). The series oscillates as white noise.

5.3. A Graphical Similarity Measure

We have used above the rate of increase in the first few higher order crossings as a basis for a white noise test. The same procedure can be used in testing the hypothesis that a given time series follows a specific model; e.g. (31) with given parameters. When the parameters are specified, under the hypothesis, ED_k and $\lambda_1^{(k)}$ can be evaluated exactly from (1) as follows. Define $\rho_1^{(k)}$ to be the first autocorrelation of $\{\nabla^k Z_t\}$,

$$\rho_1^{(k)} \equiv \text{corr}(\nabla^k Z_t, \nabla^k Z_{t-1}),$$

where now $\rho_1^{(0)} \equiv \rho_1$. Observe that $\rho_1^{(k)}$ is given by (12). (1) immediately yields

$$ED_k = (N-1)\left(\frac{1}{2} - \frac{1}{\pi} \sin^{-1}\rho_1^{(k-1)}\right) \tag{43}$$

and

$$\lambda_1^{(k)} = \frac{1}{2} + \frac{1}{\pi} \sin^{-1}\rho_1^{(k-1)}. \tag{44}$$

Experience shows that substitution of (44) in (41) still provides a reasonable fast approximation to $\text{Var}(D_k)$, and so approximate 95% probability limits for D_k are given by

$$ED_k \pm 1.96\{\text{Var } D_k\}^{\frac{1}{2}}. \tag{45}$$

To obtain (45) we follow the following steps.

1. Specify a model for the time series by a hypothesis.
2. Obtain the first few autocorrelations for the hypothesized model.
3. Obtain $\rho_1^{(k)}$ from (12).

4. Evaluate (43), (44).

5. Obtain the probability limits (45).

As an example, consider the D_j is Figure 9 (b). Figure 16 gives the probability limits under various null hypotheses H_0. The higher order crossings plot lies completely within the limits (45) for a first order autoregressive process with parameter 0.5, as it should, and the hypothesis is accepted. In all the other cases the hypothesis is rejected. We conclude that the series from which the D_j were extracted *oscillates* as a first order autoregression with the indicated parameter.

For a power study of this test, as well as a refinement of (45) see [54].

5.4. The Psi-square Statistic

Another device which exploits the monotonicity in the D_j is a very simple quadratic form defined as follows. Define the increments in the D_j by

$$\Delta_k \equiv \begin{cases} D_1 & k = 1 \\ D_k - D_{k-1} & k = 2,\ldots,K-1 \\ (N-1) - D_{k-1} & k = K. \end{cases}$$

When N is sufficiently large, $0 \le \Delta_k \le (N-1)$ and

$$\sum_{k=1}^{K} \Delta_k = N-1.$$

Let $m_k \equiv E\Delta_k$. Viewing the Δ_k as "frequencies" in a multinomial experiment, we define a general similarity measure by [43], [44]

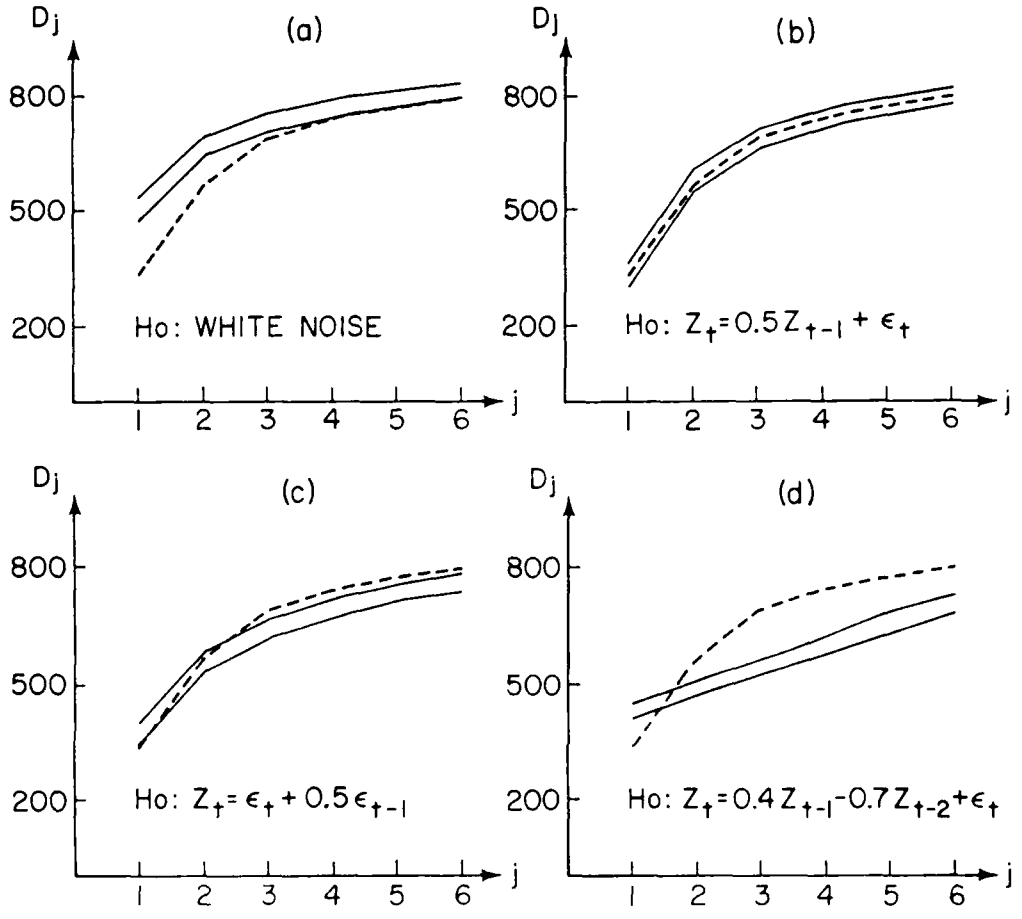


Figure 16. The rate of increase in D_j in (b) resembles that of a first order autoregressive process with parameter $\phi = 0.5$. $\{\epsilon_t\}$ are uncorrelated $N(0,\sigma^2)$ random variables. N = 1000. The solid lines are the probability limits obtained under H_0.

$$\psi^2 \equiv \sum_{k=1}^{K} \frac{(\Delta_k - m_k)^2}{m_k} . \tag{46}$$

The ψ^2 statistic can be used in testing the hypothesis that $m_k = m_k^{(0)}$, $k = 1,\ldots,K$, or simply as a distance measure, in measuring the degree of similarity between time series.

The probability distribution of the psi-square statistic (42) has been studied in [56], but no closed forms expressions are available at present. It has been observed however that when the Δ_k come from linear processes [45] of the type presented in Figure 16, the tail of the probability density of ψ^2 varies very little from process to process [43], [44]. In fact, computer simulations show that in most cases $Pr(\psi^2 > 30) < 0.05$, for $K = 9$ and sufficiently large N as can also be seen from Figure 17. We can use the ψ^2-value 30 as a critical value which corresponds to a significance level of less than 0.05. This means that the probability of rejecting a true hypothesis is less than 0.05.

As an example, consider again the D_j from Figure 9 (b), and the four models in Figure 16. We obtain, by using (12), (43), the m_j under each of the four hypotheses. The Δ_j are computed from the D_j. In case (a), the hypothesis is that of white noise and with $K = 9$, $N = 1000$, we have

$$\psi^2 = 57.12 + 26.24 + 49.23 + \ldots + 0.50 > 30$$

and the hypothesis of white noise is strongly rejected. This of course is not a surprise since the D_j were generated by a nonwhite noise process, and we expect a large ψ^2 value. In case (b) both

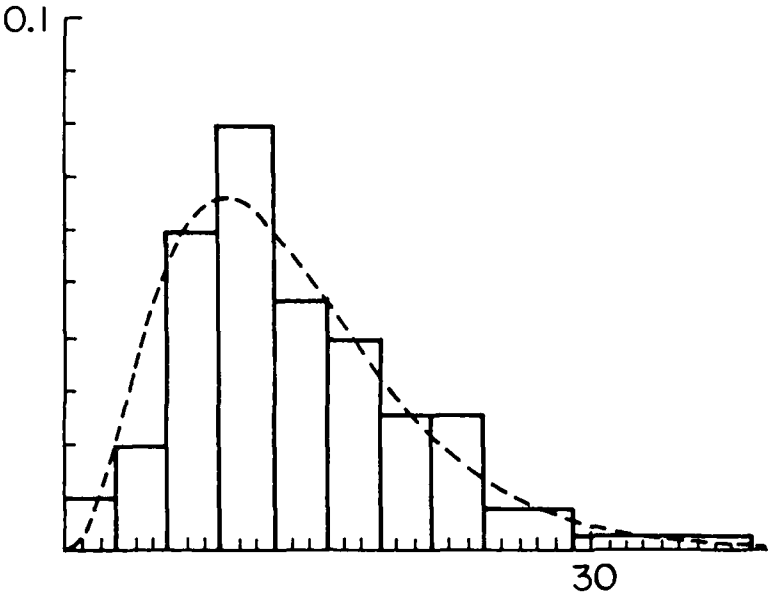


Figure 17. A density of $1.847\chi^2_{(7)}$ versus an empirical distribution of ψ^2 (K = 9, N = 1000) obtained from 100 ψ^2 values. The underlying process is a first order autoregression with parameter 0.5. In 100 ψ^2 values only two exceeded 30, three exceeded 29, four exceeded 28.

m_j and Λ_j come from the same process and we expect a small ψ^2. Indeed, we obtain $\psi^2 \simeq 7.5 < 30$ and the hypothesis of a first order autoregressive process with parameter 0.5 is accepted. For the two other cases the ψ^2 values are over 40 and over 600, respectively, and the two corresponding hypotheses are rejected. We see that evaluating ψ^2 quantifies the results of D_j plots such as those in Figure 16.

When the m_j in (46) are replaced by observed Λ_j from another process, ψ^2 becomes a measure of distance. Suppose in the speech

data the Δ_j from "zero" are taken as "expected" and Δ_j in the other two cases as observed quantities. Then the ψ^2 value from "zero" and "one" is smaller than the ψ^2 value from "zero" and "two". That is, "one" is closer to "zero" than "two", but this is of course well known.

5.5. Extension to Two Dimensions

The Higher Order Crossings Theorem can be extended to two dimensional processes $Z(t_1,t_2)$, $t_j = \ldots,-1,0,1,\ldots$, $j=1,2$ [57], so that the above methodology can be extended in a natural way to discrimination between pictures and images. To describe the degenerate state corresponding to (39) introduce the notation ∇_s to denote a difference with respect to the index s,

$$\nabla_s Z(s,t) = Z(s,t) - Z(s-1,t).$$

Let $X^{(k)}(t_1,t_2) = 0$ unless $\nabla_{t_1}^{k-1}\nabla_{t_2}^{k-1} Z(t_1,t_2) \geq 0$, in which case $X^{(k)}(t_1,t_2) = 1$. As $k \to \infty$, $\{X^{(k)}\}$ converges to

$$\begin{array}{c}
\vdots \\
\vdots \\
\vdots \\
\ldots 0\ 1\ 0\ 1\ 0\ 1 \ldots \\
\ldots 1\ 0\ 1\ 0\ 1\ 0 \ldots \\
\ldots 0\ 1\ 0\ 1\ 0\ 1 \ldots \\
\ldots 1\ 0\ 1\ 0\ 1\ 0 \ldots \\
\ldots 0\ 1\ 0\ 1\ 0\ 1 \ldots \\
\ldots 1\ 0\ 1\ 0\ 1\ 0 \ldots \\
\vdots \\
\vdots \\
\vdots
\end{array} \qquad (47)$$

with probability 1/2 or to its shift, again with probability 1/2. (47) is the degenerate state analogous to (39) when (π,π) is in the spectral support. The higher order crossings are now defined as the total number of horizontal and vertical symbol changes in the binary arrays of fixed size $\{X^{(k)}(t_1,t_2)\}$, $1 \leq t_1, t_2 \leq N$, and enjoy similar properties as in the one dimensional case. For an application of these higher crossings in texture discrimination see [16].

6. CONCLUDING REMARKS

We have described in some detail the domain of zero-crossings to which we refer as the D-domain on account of the higher order crossings ${}_{k}D_{Hj}$. It is a graphical-combinatorial domain. In retrospect, several points need further clasification. First, although many of our results depend on the Gaussian assumption , our experience shows that reasonable deviations from Gaussianity have almost no bearing on our previous conclusions. This is due chiefly to the dominant frequency principle which is independent of any Gaussian assumption. Still, it is of interest to know whether a parallel theory of zero-crossings exists which disposes of the Gaussian assumption. In our opinion this is possible provided there exists a relation such as (2), but by no means the same, which connects the zero-crossing count or a similar quantity with the spectral distribution function. The case we have in mind is the class of *bounded* processes.

It is possible to obtain a representation analogous to (2) but in terms of the crossings or a random curve. Suppose $\{Z_t\}$ is a zero mean stationary bounded process,

$$|Z_t| < A \quad \text{for all } t.$$

Obviously $\{Z_t\}$ cannot be a Gaussian process. In the range $(-A,A)$ define a uniform, completely random, process $\{U_t\}$ which consists of independent random variables uniformly distributed in $(-A,A)$. For each t, U_t has the probability density $1/2A$ defined (i.e., positive there) over $(-A,A)$. Corresponding to $Z_1,\ldots,Z_N$ we now

have the "random curve" $U_1,\ldots,U_N$, so that the time series $Z_1,\ldots,Z_N$ and the random curve $U_1,\ldots,U_N$ cross each other as in Figure 18.

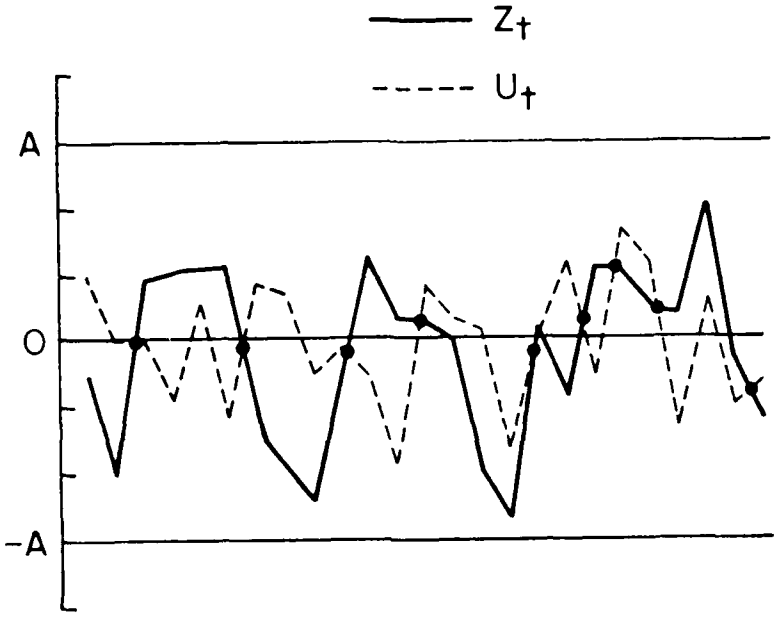


Figure 18. Random curve crossings.

Let C_1 be the number of such mutual crossings. Then it can be shown that [58] the analog of (2) is given by

$$A^2\left(1 - \frac{2E(C_1)}{N-1}\right) = \int_{-\pi}^{\pi} \cos(\omega)\,dF(\omega). \tag{48}$$

In principle we can now work out a new theory which parallels the above development, but the new results do not assume simple forms.

The second point which needs a clarification is the choice of a level (not random) at which we clip to obtain crossings. In the Gaussian case, there is a certain advantage in clipping at times at levels other than 0 [59], [60] but then the strong intuitive appeal of zero-crossings is lost.

Our last point concerns a curious similarity between Rice's formula and the zero-crossing spectral representation (2). The two are not too far apart after all. To see that, apply to the cosines in (2) a second order Taylor's series expansion to obtain after some cancellation the approximation

$$\frac{ED_1}{N-1} \simeq \frac{1}{\pi}\left[\frac{\int_{-\pi}^{\pi} \omega^2 dF(\omega)}{\int_{-\pi}^{\pi} dF(\omega)}\right]^{\frac{1}{2}} . \tag{49}$$

The left hand side of (49) is our definition of the expected zero-crossing rate, while the right hand side has the form of Rice's well-known formula [19, eq. 3.3-11] for the zero-crossing rate when expressed in terms of the 2nd spectral moment, but modified to suit the discrete time case.

ACKNOWLEDGMENT

The author is grateful to the Naval Research Laboratory for the speech data.

REFERENCES

[1] S. O. Rice, "Statistical properties of a sine wave plus random noise," Bell System Tech. J., vol. 27, pp. 109-157, Jan. 1948.

[2] J. S. Bendat, Principles and Applications of Random Noise Theory, New York: Wiley & Sons, 1958, ch. 10, pp. 370-414.

[3] A. J. Rainal, "Another zero-crossing principle for detecting narrowband signals," IEEE Trans. Instrum. and Measure., IM-16, pp. 135-138, June 1967.

[4] M. A. Badri Narayanan, S. Rajagopalan and R. Narasimha, "Experiments on the structure of turbulence," J. Fluid Mech., vol. 80, part 2, pp. 237-257, Apr. 1977.

[5] K. R. Sreenivasan, A. Prabhu and R. Narasimha, "Zero-crossings in turbulent signals," J. Fluid Mech., vol. 137, pp. 251-272, Dec. 1983.

[6] J. C. Anderson, "Improved zero-crossing method enhances digital speech," Everything Designers Need, vol. 27, pp. 171-174, Oct. 1982.

[7] R. J. Niederjohn and P. F. Castelaz, "Zero-crossing analysis methods and their use for automatic speech recognition," 21st Midwest Symp. on Circ. and Systems, Iowa State Univ., Aug. 1978, pp. 274-281.

[8] T. Masuda, H. Miyano and F. Sadoyama, "The measurement of muscle fiber conduction velocity using a gradient threshold zero-crossing method," IEEE Trans. on Biomedical Engin., vol. BME-29, pp. 673-678, Oct. 1982.

[9] N. Takai, T. Iwai, T. Ishizaka and T. Asakura, "Zero-crossing study on dynamic properties of speckles," J. Optics (Paris), vol. 11, pp. 93-101, no. 2, 1980.

[10] J. Ohtsubo, "Exact solution of zero-crossing rate of a differentiated speckle pattern," Optics Communications, vol. 42, pp. 13-18, June 1982.

[11] A. Fuglsang-Frederiksen, M. L. Monaco and K. Dahl, "Integrated electrical activity and number of zero-crossings during a gradual increase in muscle force in patients with neuromuscular diseases," Electroencephalog. Clinical Neurophys. (Netherlands), vol. 58, pp. 211-219, no. 3, 1984.

[12] Y. K. Lin, Probabilistic Theory of Structural Dynamics, New York: McGraw-Hill, 1967, ch. 9, pp. 293-337.

[13] J. R. Rice and F. P. Beer, "On the distribution of rises and falls in a continuous random process," J. of Basic Engineering, Ser. D, vol. 87, pp. 398-404, June 1965.

[14] H. B. Voelcker, "Zero-crossing properties of angle-modulated signals," IEEE Trans on Communic., vol. Com.-20, pp. 307-315, June 1972.

[15] R. G. Wiley, H. Schwarzlander and D. D. Weiner, "Demodulation procedure for very wide-band FM," IEEE Trans. on Communic., vol. Com.-25, pp. 318-327, Mar. 1977.

[16] B. Kedem, "Qualitative texture discrimination," First IPA Conference on Image Proc. Comput. Graph. and Pat. Recog., Beersheva, June 1983, pp. 15-21.

[17] R. M. Haralick, "Digital step edges from zero-crossing of second directional derivatives," IEEE Trans. on Pat. Analysis and Machine Intel., vol. PAMI-6, pp. 58-68, Jan. 1984.

[18] M. Kac, "On the average number of real roots of a random algebraic equation," Bull. Amer. Math. Soc., vol. 49, pp. 314-320, Apr. 1943.

[19] S. O. Rice, "Mathematical analysis of random noise," Bell System Tech. J., vol. 24, pp. 46-156, Jan. 1945.

[20] V. N. Tikhonov, "Characteristics of overshoots in random processes (review)," Radio Engin. and Electronic Physics, vol. 9, pp. 295-320, Mar. 1964.

[21] H. Cramér and M. R. Leadbetter, Stationary and Related Stochastic Processes, New York: Wiley, 1967.

[22] I. F. Blake and W. C. Lindsey, "Level-crossing problems for random processes," IEEE Trans. of Infor. Theory, vol. IT-19, pp. 295-315, May 1973.

[23] S. M. Berman, "Sojourns and extremes of stationary processes," Annals of Probab., vol. 10, pp. 1-46, Feb. 1982.

[24] M. R. Leadbetter, G. Lindgren, and H. Rootzen, Extremes and Related Properties of Random Sequences and Processes, New York: Springer-Verlag, 1983.

[25] A. Becker and R. S. Mackay, "Maximum frequency indication: minima counting versus zero-crossing counting," IEEE Trans. on Biomed. Engin., vol. BME-29, pp. 213-215, Mar. 1982.

[26] F. E. Offner, "Comments on "Maximum frequency indication: minima counting versus zero-crossing counting"," IEEE Trans. on Biomed. Engin., vol. BME-30, p. 78, Jan. 1983.

[27] B. Kedem, "Detection of hidden periodicities by means of higher order crossings I, II," Dept. of Math., Univ. of Md. Reports TR84-55 and TR84-56, 1984.

[28] B. Kedem, "On the sinusoidal limit of stationary time series," Annals of Stat., vol. 12, pp. 665-674, June 1984.

[29] B. Kedem, "A stochastic characterization of the sine function," to appear in the American Math. Month.

[30] A. Papoulis, Probability, Random Variables and Stochastic Processes, New York: McGraw-Hill, 1965.

[31] N. M. Blachman, Noise and its Effect on Communication, New York: McGraw-Hill, 1966.

[32] T. W. Anderson, The Statistical Analysis of Time Series, New York: Wiley, 1971, ch. 7, pp. 371-437.

[33] L. K. Koopmans, The Spectral Analysis of Time Series, New York: Academic Press, 1974.

[34] S. Karlin and H. M. Taylor, A First Course in Stochastic Processes, 2nd ed., New York: Academic Press, 1975, ch. 9, pp. 443-535.

[35] R. K. Otnes and L. Enochson, Applied Time Series Analysis, New York: Wiley, 1978, chs. 4,5, pp. 106-218.

[36] B. Kedem, Binary Time Series, New York: Dekker, 1980.

[37] T. J. Stieltjes, "Extrait d'une lettre adressée a M. Hermite," Bull. Sci. Math. (2:E Série), vol. 13, pp. 170-172, July 1889.

[38] Z. A. Lomnicki and S. K. Zaremba, "Some applications of zero-one processes," J. Royal Statist. Soc. B, vol. 17, pp. 243-255, no. 2, 1955.

[39] F. Durst, A. Melling and J. H. Whitelaw, Principles and Practice of Laser-Dopler Anemometry, London: Academic Press, 1976, ch. 9, pp. 239-271.

[40] B. Kedem, "Search for periodicities by axis-crossings of filtered time series," to appear in Signal Processing.

[41] N. M. Blachman, "Zero-crossing rate for the sum of two sinusoids or a signal plus noise," IEEE Trans. on Infor. Theory, vol. IT-21, pp. 671-675, Nov. 1975.

[42] R. C. Higgins, "The utilization of zero-crossing statistics for signal detection," J. Acoust. Soc. America, vol. 67, pp. 1818-1820, May 1980.

[43] B. Kedem and E. Slud, "On goodness of fit of time series models: an application of higher order crossings," Biometrika, vol. 68, pp. 551-556, Aug. 1981.

[44] B. Kedem and E. Slud, "Higher order crossings in the discrimination of time series," Annals of Statist., vol. 10, pp. 786-794, Sept. 1982.

[45] G. E. P. Box and G. M. Jenkins, Time Series Analysis Forecasting and Control, 2nd ed., San Francisco: Holden Day, 1976, ch. 3, pp. 46-84; appendix A7.5, pp. 274-284.

[46] E. E. Slutsky, "The summation of random causes as the source of cyclic processes," English translation in Econometrica, vol. 5, pp. 105-146, Apr. 1937.

[47] A. Schuster, "On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena," Terr. Mag. Atmos. Elect., vol. 3, pp. 13-41, no. 1, 1898.

[48] M. B. Priestley, Spectral Analysis and Time Series, vol. 1. London: Academic Press, 1981, ch. 6, pp. 389-501.

[49] M. Dechambre and J. Lavergnat, "Frequency determination of a noisy signal by zero-crossings counting," Signal Processing, vol. 8, pp. 93-105, Feb. 1985.

[50] B. Kedem, "On frequency detection by zero-crossings,", to appear in Signal Processing.

[51] M. J. Hinich, "Detecting a hidden periodic signal when its period is unknown," IEEE Trans. on Acoust. Speech and Sign. Process., vol. ASSP-30, pp. 747-750, Oct. 1982.

[52] T. J. Ulrych and T. N. Bishop, "Maximum entropy spectral analysis and autoregressive decomposition," Rev. of Geophys. Space Physics, vol. 13, pp. 183-200, Feb. 1975.

[53] H. Akaike, "Power spectrum estimation through autoregressive model fitting," Ann. Inst. Statist. Math., vol. 21, pp. 407-419, no. 3, 1969.

[54] B. Kedem, "A graphical similarity measure for time series models," Dept. of Math., Univ. of Md. Report TR85-10, pp. 1-21, 1985.

[55] B. Kedem and G. Reed, "On the asymptotic variance of higher order crossings with special reference to a fast white noise test,", to appear in Biometrika.

[56] G. Reed, "Some properties and applications of higher order crossings," Ph.D. thesis, Dept. of Math., Univ. of Md., pp. 1-128, 1983.

[57] B. Kedem, "A two dimensional higher order crossing theorem," IEEE Trans. on Infor. Theory, vol. IT-29, pp. 159-161, Jan. 1983.

[58] B. Kedem, "Some graphical considerations in time series analysis," IEEE Trans. on Pattern Anslys. Mach. Intel., vol. PAMI-4, pp. 493-499, Sept. 1982.

[59] G. Lindgren, "Spectral moment estimation by means of level crossings," Biometrika, vol. 61, pp. 401-418, Dec. 1974.

[60] M. B. Masanori, B. Okamoto and K. Iwase, "A statistical index representing a gross feature of several climatic time series by hard limiting," J. Meteoro. Soc. of Japan, vol. 60, pp. 726-738, Apr. 1982.

END

DTIC

8-86