NONLINEAR SPECTRAL ANALYSIS OF RADAR CLUTTER

TO JELISAVETA,

BRANKO, AND OLIVERA

KESLER

NONLINEAR SPECTRAL ANALYSIS
OF RADAR CLUTTER

By

Stanislav B. Kesler, M.Eng.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University
December, 1977

DOCTOR OF PHILOSOPHY (1977)
(Electrical Engineering)

McMaster University
Hamilton, Ontario

TITLE: Nonlinear Spectral Analysis of Radar Clutter

AUTHOR: Stanislav B. Kesler
(Dipl. Eng., University of Belgrade, 1965)
(M. Eng., University of Belgrade, 1973)

SUPERVISOR: Dr. S. S. Haykin

NUMBER OF PAGES: xiii, 202

ABSTRACT

This thesis is concerned primarily with the spectral analysis of clutter in an air traffic control (ATC) radar environment, which is produced by such objects as weather disturbances and migrating flocks of birds. The aim is to provide a means for the on-line classification of the different forms of clutter in such an environment. With the clutter identified and displayed for the operator, aircraft can be vectored in a way to avoid hazardous areas. This separation and identification of radar echoes may also be useful in studies relating to meteorology and ornithology.

Since the mechanisms which give rise to weather clutter and bird echoes are different, it is reasonable to expect that they have different statistical parameters. One of these parameters, spectral spread, provides a convenient measure for comparison. To perform the separation of clutter based on spectral spread, a high resolution method is necessary. In this thesis, the maximum entropy method (MEM) is applied to clutter classification. It is shown that the resolution and stability properties of the MEM are well-suited for the separation of different types of clutter; particularly weather from birds. The analysis is performed on both computer-simulated and actual clutter signals. In each case, the results are compared with those obtained by applying the Welch method of averaged modified periodograms. This comparison has revealed that the Welch method does not have the necessary resolution for the classification of clutter.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the help and guidance to him by his supervisor, Dr. S. S. Haykin througnout the preparation of the thesis. He would also thank Dr. J. B. Anderson, Dr. C. R. Carter, and Dr. T. Timusk for serving on his supervisory committee.

The author is grateful for financial support from the National Research Council of Canada. The author would like to thank Dr. H. C. Chan and Mr. B. W. Currie for their permission to use some of their computer programs and their many enjoyable and helpful discussions. Thanks are also due to the research personnel of the Communications Research Laboratory at McMaster University for participating in many lively discussions.

Finally, I would like to express my appreciation for the loving manner in which my wife and colleague, Jelisaveta, and my children tolerated the time spent on this research and thesis preparation; without their understanding and constant support this effort would not have been possible.

TABLE OF CONTENTS

Table of Contents Continued

LIST OF ILLUSTRATIONS

List of Illustrations Continued

List of Illustrations Continued

List of Illustrations Continued

List of Illustrations Continued

List of Illustrations Continued

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Purpose of the Research

This dissertation developes a method for radar clutter analysis and classification. Based on this method, an adaptive procedure for Moving Target Indication (MTI) is proposed.

The classical approach to the problem of clutter spectra analysis is to use linear spectral estimation methods [1] - [4]. It has been demonstrated, however, that nonlinear methods, in particular, the maximum entropy method (MEM) [5], provide superior tools for the analysis of clutter and for discrimination amongst different types of clutter. The capability of these methods to adapt to the signals being processed enables one to construct a simple adaptive procedure for clutter elimination.

To accommodate the use of the MEM for a coherent video signal we first develop a complex form of the method. Both the MEM and the Welch methods are then applied to the analysis of synthetic clutter signals. The obtained results confirmed the advantages of MEM for short data record lengths.

The MEM has then been applied to the analysis of video signals recorded at a coherent radar site. The availability of different types of clutter enabled us to make the spectral estimates of ground clutter, weather clutter, and bird flocks. It is shown that the differences in spectra are clearly recognizable.

Using the concept of prediction error filtering (PEF), which is the basis for MEM, a simple adaptive MTI processing scheme is proposed. The resulting filter is absolutely stable, minimum phase, and fast-converging. It is also suitable for real-time implementation.

1.2 Motivation for the Research

Radar was first developed during the second World War as a detection device to extract information about the presence and position of aircraft. It was soon realized that much more information from the target than merely range could be extracted by transmitting different waveforms, and by applying sophisticated signal processing techniques. Moreover, information could also be obtained about the operating environment of the radar. Examples of these applications are weather radar [6], and sideway-looking airborne radar [7]. The use of radar is now extended to such broad areas as detection and tracking satellites [8], speed measuring in highway traffic control [9], radar astronomy [10], exploration and prediction of metereological conditions [11], and examination of bird migrations [12].

The last two applications are typical examples of clutter signal processing. Weather radars require means for discrimination amongst different environmental conditions. Metereological radars try to determine the type and intensity of weather patterns, while in ornithology radar is used to explore migration habits of various bird species. The method developed in this research is well-suitable for both applications.

When clutter is removed from the plan-position indicator (PPI), there is a possibility of vectoring an aircraft into a hazardous area or dense bird formations. It is, therefore, of interest not only to simply reject the clutter echoes, but also to analyze them, so that the safety of the aircraft can be improved. This requires on-line classification of the different forms of clutter. The work reported in this thesis describes a procedure for achieving this requirement. The procedure is based on a spectral density estimate of the clutter. Before describing the estimation procedure in detail, it is instructive to begin by reviewing the clutter spectral characteristics and the different methods that are available for spectral estimation.

1.3 Statistical Properties of Clutter

The characterization of clutter has been being looked at from two viewpoints:

(1) The amplitude statistics of the signal.

(2) The statistics of the time variations of the signal.

Knowledge of the amplitude statistics of the target and clutter returns enables a detection procedure utilizing hypothesis testing. As far as probability density functions of the target models are concerned, the original works of Marcum [13] and Swerling [14], as well as the extensions of Swerling [15], and generalizations of Kanter [16], have been found to bracket a wide range of practical cases. The most relevant work on radar detection is collected in [17]. However, there does not seem to be any distribution that can be claimed as truly representative of clutter, but two that are often quoted are Rayleigh and lognormal [18], [19].

Since the essence of most of the clutter eliminating procedures is filtering, the time-frequency characteristics of a clutter signal are of great importance. The spectral characteristics of the clutter signal have been extensively studied and many attempts have been made to derive models that fit the observed spectra. Since there are a great variety of clutter scatterers, no one model can cover all cases.

The major contribution to spectral spread in clutter is radial movement of the scatterers. This is related to the power spectral density of the clutter return via the Doppler relation

$$f_d = \frac{2v_r}{c} \cdot f_c = \frac{2v_r}{\lambda_c} \qquad (1.3.1)$$

where $c = 3 \times 10^8$ meters per second is the propagation velocity, f_c is the carrier frequency in hertz, λ_c is the wavelength in meters, and v_r is the radial velocity component of the scatterers in meters per second.

There has been much conjecture on the shape of the power spectral density of clutter and it is often approximated as [20]-[24]

$$S(f) = e^{-f^2/2\sigma_d^2} \qquad (1.3.2)$$

where the rms spectral spread, σ_d, is related to the rms velocity spread of the scattering elements, σ_v in meters per second, by the Doppler relation

$$\sigma_d = \frac{2\sigma_v}{\lambda_c} \, . \qquad (1.3.3)$$

Since Eq. (1.3.2) is of the same form as a Gaussian probability density function with zero mean, σ_d and σ_v are frequently referred to as the standard deviations of the spectral and velocity distributions, respectively.

More recently it has been found that, in some cases, a better fit to measured data is obtained when the shape of the spectral density is given by an all-pole rational expression of order three [25]

$$S(f) = \frac{1}{1+(f/f_3)^3}, \qquad (1.3.4)$$

or of order four [26,p.9]

$$S(f) = \frac{1}{1+(f/f_4)^4}, \qquad (1.3.5)$$

where f_3 and f_4 are the corresponding cutoff frequencies. All mentioned results indicate that there is really no unique representation for the power spectral density of clutter.

Scatterer scintillation has been taken into account in an attempt to get better agreement between measured data and theoretical models [27], [28]. It turns out that modelling the scatterers as rotating dipoles [27] yields a better representation of certain types of clutter.

Antenna rotation [29] and instabilities in the local oscillator [30] also contribute to an overall clutter spectral spread.

The usual way of combatting clutter in coherent pulse radars is an MTI filtering [23,24,29,31,32]. The use of digital hardware for canceller circuitry has resulted in outstanding performance [33]-[35],hence, modern radars are now using digital MTI filters almost exclusively.

1.4 Linear Methods of Spectral Analysis

We say that a discrete complex random process

$$x(n\Delta t) = x^r(n\Delta t) + jx^i(n\Delta t), \; n = \ldots,-2,-1,0,1,2,\ldots \quad (1.4.1)$$

is stationary if the real part $x^r(n\Delta t)$ and the imaginary part $x^i(n\Delta t)$ are jointly stationary processes [36]. Prameter Δt in Eq. (1.4.1) is the time spacing between samples. Unless otherwise stated, we shall assume $\Delta t = 1$ in the reminder of the work. The expected value of $x(n)$ is

$$\mu_x = E[x(n)] = E[x^r(n)] + jE[x^i(n)]$$
$$= \sum_{n=-\infty}^{\infty} x^r(n) \cdot p_{x_n^r}[x^r(n)] + j \sum_{n=-\infty}^{\infty} x^i(n) \cdot p_{x_n^i}[x^i(n)] \quad , (1.4.2)$$

where $p_{x_n^r}[x^r(n)]$ and $p_{x_n^i}[x^i(n)]$ are probability distribution functions of the real and the imaginary parts, respectively.

For the stationary process, given by Eq. (1.4.1), the autocorrelation function (ACF) at lag $\tau = m\Delta t = m$ is given by

$$R_x(m) = E[x(n) \cdot x^*(n+m)]. \quad (1.4.3)$$

The asterisk signifies a complex conjugate. An important property of $R_x(m)$ is

$$R_x(-m) = R_x^*(m) \; . \quad (1.4.4)$$

Experimental investigations have confirmed that video signals representing clutter returns, can be assumed to be ergodic random processes. We may therefore use time averages instead of ensemble averages to estimate the autocorrelation function. The time autocorrelation function is defined by

$$R_x(m) = \lim_{M\to\infty} \frac{1}{2M+1} \sum_{n=-M}^{M} x(n) \cdot x^*(n+m) . \qquad (1.4.5)$$

Both the ensemble and the time autocorrelation functions in Eqs.(1.4.3) and (1.4.5) are denoted with the same symbols, since we shall be using the ergodicity assumption in the remainder of the work.

Another useful second-order characteristic of a random process is the power spectral density (PSD). It may be defined as the mean-square value of the portion, $x(n,f,\Delta f)$, of the signal $x(n)$ in the frequency range from f to $f + \Delta f$ [37]:

$$E\left[|x(n,f,\Delta f)|^2\right] = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} |x(n,f,\Delta f)|^2 . \qquad (1.4.6)$$

When Δf tends to zero, the ratio of $E\left[|x(n,f,\Delta f)|^2\right]$ and Δf approaches the power spectral density at a frequency f

$$S_x(f) = \lim_{\Delta f\to 0} \frac{E\left[|x(n,f,\Delta f)|^2\right]}{\Delta f} = \lim_{\Delta f\to 0} \frac{1}{\Delta f} \left[\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} |x(n,f,\Delta f)|^2\right] . \qquad (1.4.7)$$

It is clear from (1.4.7) that $S_x(f)$ is a real-valued, non-negative function of frequency.

The Wiener-Khintchine theorem states that the autocorrelation function, $R_x(m)$, and the power spectral density, $S_x(f)$ form a Fourier transform pair as follows:

$$S_x(f) = \sum_{m=-\infty}^{\infty} R_x(m)\cdot e^{-j2\pi m\Delta tf} \tag{1.4.8}$$

and

$$R_x(m) = \int_{-1/2\Delta t}^{1/2\Delta t} S_x(f)\cdot e^{j2\pi m\Delta tf}\,df, \tag{1.4.9}$$

where $1/\Delta t$ is the Nyquist frequency. In the case of normalized Δt, this frequency is equal to unity, and the limits of integration in Eq.(1.4.9) are -0.5 and 0.5. It follows from Eqs.(1.4.4) and (1.4.5) that $S_x(f)$ is not a symmetrical function of frequency as in the case of real processes and, hence, cannot be expressed as one-sided density function.

In the analysis of physical processes, we are given only a limited amount of data, so that we can only estimate the quantities defined by Eqs.(1.4.5) to (1.4.9). Estimation of statistical moments of the random process has become an extremely important branch of time series analysis due to the fact that it has applications in virtually every field of science and engineering.

Spectral analysis methods can be divided into two classes: linear and nonlinear. In the first case, the estimates of the ACF and PSD are related by Fourier transformation, while in the second case these estimates are connected in some other way.

Assume we are given a sample function of length N of a second-order stationary band-limited discrete complex random process with zero-mean:

$$\{x(n)\} = \{x(1), x(2), \ldots, x(N)\}, \tag{1.4.10}$$

from which we wish to estimate the second order statistics. The zero-mean restriction does not, in fact, introduce a loss of generality, since we can always extract a sample mean:

$$\mu_x = \frac{1}{N}\sum_{n=1}^{N} x'(n) \tag{1.4.11}$$

from the original process $\{x'(n)\}$ and subtract it from each term to get $\{x(n)\} = \{x'(n) - \mu_x\}$. For the estimation of the autocorrelation function, we may use a biased estimate:

$$\hat{R}(m) = \frac{1}{N}\sum_{n=0}^{N-|m|} x(n)\cdot x^*(n+m)\,, \qquad -M \le m \le M, \tag{1.4.12}$$

or an unbiased estimate:

$$\hat{R}'(m) = \frac{1}{N-|m|}\sum_{n=0}^{N-|m|} x(n)\cdot x^*(n+m)\,, \qquad -M \le m \le M, \tag{1.4.13}$$

where M (<N) is a maximum lag [38].

For the estimation of the power spectral density, we may use one of the following procedures:

(1) The Blackman-Tukey procedure [1]: This consists of multiplying the unbiased ACF estimate, Eq.(1.4.13) with an appropriate lag window and then Fourier transforming the product to get the spectral estimate:

$$\hat{S}_{BT}(f) = \hat{S}'(f) \circledast W(f) \tag{1.4.14}$$

where $\hat{S}'(f)$ is the Fourier transform of $\hat{R}'(m)$, the spectral window, W(f), is the Fourier transform of the lag window, w(m), and $\circledast$ denotes convolution.

(2) The Periodogram Method: The periodogram is defined as the squared magnitude of the finite Fourier transform of the sequence {x(n)}, n = 0,1,...,N-1, divided by N:

$$I_N(f) = \frac{1}{N}\left|X(f)\right|^2 = \frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)\cdot e^{-j2\pi fn}\right|^2. \qquad (1.4.15)$$

It follows from Eqs.(1.4.15) and (1.4.12) that the periodogram is the Fourier transform of the biased autocorrelation estimate, $\hat{R}(m)$, that is

$$I_N(f) = \sum_{m=-N+1}^{N-1} \hat{R}(m)\cdot e^{-j2\pi fm}. \qquad (1.4.16)$$

The periodogram is a biased and nonconsistent estimator of the power spectral density [39], [40]. These limitations of the periodogram may be overcome by using a spectral window. For example, the use of a Bartlett window [2] is equivalent to dividing a data sequence x(n), n = 0,1,...,N-1 into K segments of L samples each, so that N = KL as shown in Fig. 1.4-1. That is, we form the segments

$$x^{(k)}(n) = x[n+(k-1)L], \quad 0 \le n \le L-1, \quad 1 \le k \le K, \qquad (1.4.17)$$

and compute the K periodograms

$$I_L^{(k)}(f) = \frac{1}{L}\left|\sum_{n=0}^{L-1} x^{(k)}(n)\cdot e^{-j2\pi fn}\right|^2, \quad 1 \le k \le K. \qquad (1.4.18)$$

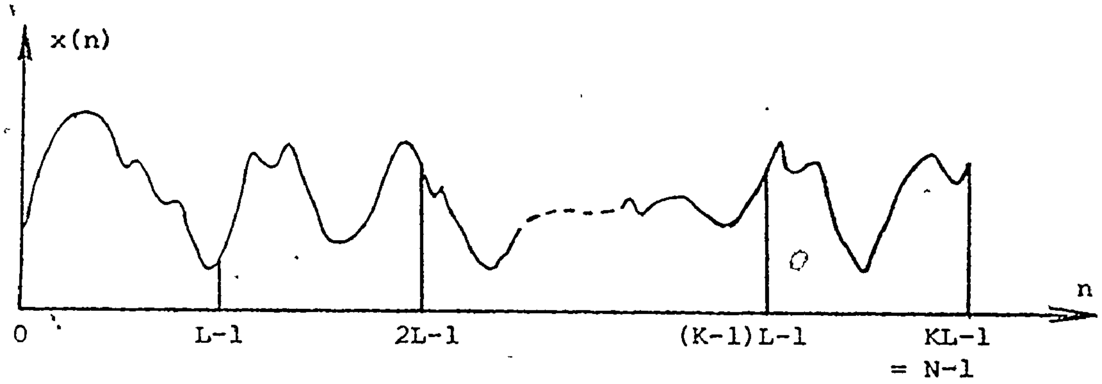


Fig. 1.4 - 1 A data sequence x(n) of length N divided into K segments of L samples each.

The smoothed spectral estimator, $\hat{S}_B(f)$, is then defined by the sample mean

$$\hat{S}_B(f) = \frac{1}{K}\sum_{k=1}^{K} I_L^{(k)}(f) . \qquad (1.4.19)$$

Averaging over periodograms increases the bias of the estimate, but also decreases the variance. We can compromise between bias and variance by changing M and L for a fixed N.

(3) <u>The Modified Periodogram Method</u>: Another smoothing procedure, introduced by Welch [4], involves a modification of the Bartlett method which is particularly well-suited to direct computation of a power spectral density estimate using the fast Fourier transform (FFT) algorithm [41]. The data record is again sectioned into K = N/L segments of M samples each as defined by Eq.(1.4.17). In this case, however, the lag window, w(n), is applied directly to the data segments before computation of the periodogram. Thus, we define the K modified periodograms:

$$J_L^{(k)}(f) = \frac{1}{LU}\left|\sum_{n=0}^{L-1} x^{(k)}(n)\cdot w(n)\cdot e^{-j2\pi fn}\right|^2, \quad 1 \le k \le K, \qquad (1.4.20)$$

where

$$U = \frac{1}{L}\sum_{n=0}^{L-1} w^2(n) \qquad (1.4.21)$$

is the energy of a spectral window. The spectral density estimate is defined by

$$\hat{S}_W(f) = \frac{1}{K}\sum_{k=1}^{K} J_L^{(k)}(f) \quad . \qquad (1.4.22)$$

Welch has shown [4] that the estimator (1.4.22) is asymptotically unbiased. Moreover, for nonoverlapping segments, the variance of the estimate is

$$\mathrm{var}\,[\hat{S}_W(f)] \approx \frac{1}{K}\,S^2(f) \qquad (1.4.23)$$

where $S(f)$ is a true spectral density. Further reduction in the variance can be obtained by overlapping the segments of $x(n)$, so that the modified periodograms are not independent. Thus, by windowing the data segments before computing the periodograms, we achieve the variance reduction of the original Bartlett procedure and at the same time achieve smoothing of the spectrum. In this case the spectral window is proportional to the squared magnitude of the Fourier transform of the lag window rather than simply the Fourier transform itself. This means that, no matter which lag window is used, the spectral window will always be non-negative and thus the estimate $\hat{S}_W(f)$ will always be non-negative as well.

It is considered that the Welch method based on averaged modified periodograms is the best procedure for spectral estimation in the sense that it provides compromise between statistical stability (consistency) and spectral resolution (bias) Thus, the Welch method is chosen in this work for comparison with the MEM ~~for~~ the analysis of simulated and actual clutter signals.

1.5 Nonlinear Methods of Spectral Analysis

Several nonlinear methods with increased resolution have been recently proposed for spectral estimation. These include the maximum likelihood method (MLM) proposed by Capon [42], [43], and the maximum entropy method (MEM) proposed by Burg [5], [44] and Parzen, who refers to it as an autoregressive (AR), or an all-pole spectral estimator [45], [46]. These methods do not have a fixed window function associated with them. Instead, when power is being estimated at one frequency, the methods adjust themselves to be least disturbed by power at other frequencies. As pointed out by Lacoss [47], one can think of a window function adapting itself to the spectrum of the process under analysis. This is why the MLM and the MEM are referred to as data adaptive methods.

We present here a brief analysis of the MLM and MEM, and also mention the other methods. Empirical as well as theoretical investigations indicate that the resolution properties of the MEM are somewhat better than those of the MLM [47]-[49]. Consequently, less data is needed for MEM than for MLM to obtain comparable resolution.

The MLM [42] has been used by Capon for frequency-wavenumber spectral analysis with arrays [43]. The method can be adopted to single time series analysis. It can be represented in terms of minimum variance unbiased estimator of the spectral components. At a particular frequency, f, it turns out that the ML estimator, $\hat{S}_L(f)$ is the output power of a filter designed to pass that frequency undistorted and to reject all other frequency components in an optimal manner. If we denote the vector impulse response of this filter as $\underline{H} = [h_1, h_2, \ldots, h_M]^T$, and introduce a vector of complex exponentials, $\underline{E} = \left[1, e^{j2\pi f}, \ldots, e^{j2\pi(M-1)f}\right]^T$, it can be shown [47] that the optimum filter is given by

$$\underline{H} = \frac{\underline{R}^{-1}\cdot\underline{E}^*}{\underline{E}^T\cdot\underline{R}^{-1}\cdot\underline{E}^*} \tag{1.5.1}$$

where $\underline{R}$ is an MxM autocorrelation matrix of the input noise process. The superscript T signifies a transpose. "Optimum" here means that the output of the filter gives an unbiased estimate of the particular frequency component of the input signal, which has minimal variance. This minimal variance is, in fact, a maximum likelihood estimate of this particular spectral component, and is given by

$$\hat{S}_L(f) = \mathrm{var}_{(\min)}[y_{out}] = \frac{1}{\underline{E}^T\cdot\underline{R}^{-1}\cdot\underline{E}^*}\,. \tag{1.5.2}$$

It has been demonstrated by Capon that the ML estimator is asymptotically unbiased and consistent [42]. Also, Capon and Goodman [50] have shown that the estimators are multiples of the chi-square random variables with $N/M - M - 1$ degrees of freedom if an autocorrelation function is estimated up to the Mth lag from N values of a signal sample. Since this result holds asymptotically for large N and M, the number of degrees of freedom is approximately equal to that of the Blackman-Tukey estimator with the appropriate window function.

The MEM is based, as stated by Burg [51], on choosing the spectrum which corresponds to the most random or the most unpredictable time series whose ACF agrees with the known values. This assumption corresponds to the concept of maximum entropy. The problem can be mathematically expressed as finding the real positive value of the spectral estimate, $\hat{S}_E(f)$ which maximizes the entropy of the corresponding random process

$$H = \int_{-0.5}^{0.5} \ln \hat{S}(f)\,df, \tag{1.5.3}$$

under the set of constraints

$$R(m) = \int_{-0.5}^{0.5} \hat{S}(f) e^{j2\pi fm} df, \qquad -M \leq m \leq M, \tag{1.5.4}$$

where R(m)'s are the ACF values which are known (or estimated) up to the value M. Set of equations (1.5.4) actually contains only M+1 equations by recalling that $R(-m) = R^*(m)$. The solution of this constrained maximization problem has been shown to be [51], [52]

$$\hat{S}_E(f) = \frac{P_{M+1}}{\left|\sum_{m=0}^{M} a_{M,m} \cdot e^{-j2\pi fm}\right|^2}, \tag{1.5.5}$$

where the coefficients P_{M+1} and $a_{M,m}$, $m = 0,1,\ldots,M$ are given by the set of linear equations

$$\sum_{k=0}^{M} R(m-k) \cdot a_{M,m} = P_{M+1} \cdot \delta_{m,o}, \qquad m = 0,1,\ldots,M, \tag{1.5.6}$$

and $\delta_{m,o}$ is a Kronecker delta function. As will be shown in the next chapter, the coefficients $\{a_{M,m}\}$, $m = 0,1,\ldots,M$ represent the impulse response sequence of a prediction error filter (PEF) with the prediction distance equal to unity, and P_{M+1} is its output power.

Since the ME estimator is nonlinear, it is not possible to obtain general analytical expressions for its statistical properties. However, asymptotic results obtained by Kromer [53] and Berk [54] indicate that the ME estimator is asymptotically unbiased and consistent with variance

$$\mathrm{var}[\hat{S}_E(f)] = \frac{2M}{N} S^2(f). \tag{1.5.7}$$

This result holds for large N and M. It has been found recently [49] that the transfer function of the PEF has a probability density which may be interpreted as a generalization of the student's t distribution. Again, the asymptotic properties of the ME estimator may be considered approximately equal to those of the Blackman-Tukey estimator with the appropriate window function.

The main advantage of the nonlinear methods described above is the absence of a fixed window function. Thus, there are no unrealistic assumptions about the unknown portion of the data. Zero extension or periodic extension of the autocorrelation estimate, inherent in the linear methods, are quite arbitrary and, consequently, may be used satisfactorily only when a large amount of data is available. In the cases where the record length is relatively short, either because the phenomena can be considered stationary only over limited intervals of time, or because data can be collected only over the finite time interval, or both, the nonlinear methods are generally superior.

The MEM has been successfully applied to noisy sinusoidal waves as test signals [55]-[57], to geophysical data [58]-[62], and radar data

[63], [64]. The application of the method in the spatial processing of array data has also been indicated [65]. It has been observed, however, that in some test cases the MEM does not adequatly represent the true spectra [56], [57], [66]. In order to solve this problem, Ulrych and Clayton [67] proposed a variation of the original ME procedure, which they called least-square (LS) procedure. Since this modification does not necessarily produce a true PEF, Fougere suggested another variation [68], [69] which does result in a minimum phase PEF, and the disagreement with actual spectra has been completely resolved. The only shortcoming of this method is somewhat increased computation time.

The method proposed by Pisarenko [70] deals with retrieval of harmonics from a covariance function. As pointed out, the ME spectrum for harmonic processes in additive noise is a smoothed version of this method. Finally, Treitel et.al. [71], [72] state that the applicability of a particular nonlinear method depends somewhat on the type of signal under analysis. Experimental analysis indicate that linear methods are, in general, best suited for moving-average (all-zero) processes [73], while the MEM is best-suited for the AR processes. A method for spectral estimation of mixed autoregressive-moving-average (ARMA) processes has also been proposed in [71], [72].

CHAPTER 2

MAXIMUM ENTROPY (AUTOREGRESSION) METHOD OF SPECTRAL ANALYSIS

2.1 Introductory Remarks

As mentioned earlier, because of the finite length of the available data record, a spectral estimator represents the true PSD only approximately. The degree of approximation depends on the statistical properties of the data sample and on the assumptions imposed on the chosen estimation procedure. If we require the spectrum to be the most random (i.e., to have the maximum entropy) of all the spectra with known ACF up to the Mth lag, we have the maximization problem expressed by Eqs.(1.5.3) and (1.5.4). The reason why the solution of these equations is a good spectral estimator can be qualitatively described as follows: If we are given a finite data record, we can calculate the ACF values up to the lag M' which is about 40% of the record length, by using MEM with a PEF having the impulse response of length M'. If we now calculate the same function with a filter of length M, which is about 25% of the original data record, and use the ME extension to find the values for $m = M+1, M+2, \ldots, M'$, we find that the extended values will be very close to the previously calculated values. We shall say more about ACF extension later on in this chapter.

The maximum entropy concept is equivalent to fitting an AR (all-pole) model of finite order to the given data [45],[46]. This equivalence was first noted by van den Bos [74] to determine certain statistical properties of the ME spectral estimator. In the sequel we will use both terms interchangably.

2.2 Time-Domain Analysis of the ME Spectral Estimator

Under the assumptions that we have made concerning the process $x(n)$, defined by Eq. (1.4.1), it is possible to represent it by an all-pole process of order M:

$$x(n) = -a_1 x(n-1) - a_2 x(n-2) - \dots -a_M x(n-M) + w(n) \qquad (2.2.1)$$

where $a_1, a_2, \dots, a_M$ are autoregressive coefficients, and $w(n)$ is an uncorrelated process. Since $x(n-i)$; $i = 1,2,\dots,M$ are zero-mean Gaussian random variables, so is $w(n)$ as their linear combination. Thus, $w(n)$ is a Gaussian random process, with a sample variance P_{M+1}. The process $w(n)$ is frequently called the innovation or prediction error process. Rewriting Eq. (2.2.1) in a compact form, we have

$$\sum_{m=0}^{M} a_m \cdot x(n-m) = w(n) \qquad (2.2.2)$$

where $a_0 \equiv 1$.

To get the relationship between the ACF values, noise power, and AR coefficients, we evaluate the variance of both sides of Eq. (2.2.2)

$$E\left[\left(\sum_{m=0}^{M} a_m \cdot x(n-m)\right)\left(\sum_{k=0}^{M} a_k \cdot x(n-k)\right)^*\right] = E[w(n) \cdot w^*(n)] = P_{M+1} . \qquad (2.2.3)$$

Since the operations in Eq. (2.2.3) are linear and the summations are finite, we can interchange the order of summation and expectation, obtaining

$$\sum_{m=0}^{M} a_m \cdot \sum_{k=0}^{M} a_k^* \cdot E[x(n-m)\cdot x^*(n-k)] = \sum_{m=0}^{M}\sum_{k=0}^{M} a_m a_k^* R(m-k) = P_{M+1}, \quad (2.2.4)$$

where R(m-k) is the ACF value at the lag m-k. The left-hand side of Eq. (2.2.4) represents a Hermitian form in M+1 filter coefficients a_m with the associated Hermitian matrix of M+1 ACF values:

$$\underline{R}(m) = \begin{bmatrix} R(0) & R^*(1) & R^*(2) & \cdots & R^*(M) \\ R(1) & R(0) & R^*(1) & \cdots & R^*(M-1) \\ R(2) & R(1) & R(0) & \cdots & R^*(M-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R(M) & R(M-1) & R(M-2) & \cdots & R(0) \end{bmatrix}. \quad (2.2.5)$$

If we define an (M+1)x1 column vector of filter coefficients

$$\underline{a} = \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}, \quad (2.2.6)$$

we can write Eq. (2.2.4) in matrix form as follows

$$\underline{a}^H \cdot \underline{R}(m) \cdot \underline{a} = P_{M+1}. \quad (2.2.7)$$

where the superscript H signifies complex conjugate transpose (Hermitian transpose). If we now define an (M+1)x1 column vector

$$\underline{P}_{M+1} = \begin{bmatrix} P_{M+1} \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \tag{2.2.8}$$

and note that

$$\underline{a}^H \cdot \underline{P}_{M+1} = P_{M+1}, \tag{2.2.9}$$

we can write Eq. (2.2.7) in the form

$$\underline{R}(m) \cdot \underline{a} = \underline{P}_{M+1}, \tag{2.2.10}$$

or, in the expanded form

$$\begin{bmatrix} R(0) & R^*(1) & \cdots & R^*(M) \\ R(1) & R(0) & \cdots & R^*(M-1) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ R(M) & R(M-1) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_M \end{bmatrix} = \begin{bmatrix} P_{M+1} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \tag{2.2.11}$$

If we write the above system in the form of the M+1 simultaneous linear equations we get a set of normal equations for the Wiener prediction error filter (PEF) with prediction distance equal to unity [75].

2.3 The Relation Between the Wiener Prediction Filtering and Prediction Error Filtering

We can derive a set of equations (2.2.10) by starting with the corresponding Wiener prediction filter. To do this, we define the input sequence as in Eq. (1.4.10)

$$\{x(n)\} = \{x(1), x(2), \ldots , x(N)\} . \tag{2.3.1}$$

We wish to design a filter of order M which will give at its output the best approximation, in a least-square sense, to the time advanced version of the input. Thus, the desired output is $d(n) = x(n+1)$. If we denote the filter impulse response by the column vector

$$\underline{b} = [b_1, b_2, \ldots , b_M]^T, \tag{2.3.2}$$

then, the optimum values for the filter coefficients are obtained from the solution of the normal equations [75], [76]

$$\begin{bmatrix} R(0) & R^*(1) & \ldots & R^*(M-1) \\ R(1) & R(0) & \ldots & R^*(M-2) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ R(M-1) & R(M-2) & \ldots & R(0) \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_M \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \cdot \\ \cdot \\ \cdot \\ R(M) \end{bmatrix} . \tag{2.3.3}$$

In the general case for any desired output, the column vector on the right-hand side of Eq. (2.3.3) contains positive lag crosscorrelation values between the input and the desired output

$$R_{dx}(m) = \frac{1}{M}\sum_{n=0}^{M-m} x(n)\cdot d^*(n+m), \qquad m = 0,1,\ldots,M-1. \qquad (2.3.4)$$

In a special case of prediction filtering with prediction distance equal to unity, $d(n) = x(n+1)$, and Eq. (2.3.4) becomes

$$R_{dx}(m) = \frac{1}{M}\sum_{n=0}^{M-m} x(n)\cdot x^*(n+m+1) = R(m+1), \quad m = 0,1,\ldots,M-1. \quad (2.3.5)$$

These values appear on the right-hand side of Eq. (2.3.3).

To get the system of equations (2.2.11) we first subtract the vector $[R(1), R(2), \ldots, R(M)]^T$ from both sides of Eq. (2.3.3) and augment the system by adding the equation

$$-R(0) + b_1R(1) + b_2R(2) + \ldots + b_MR(M) = -P_{M+1}\,. \qquad (2.3.6)$$

to the top of the modified system. We thus obtain the system (2.2.11) where

$$a_0 \equiv 1 \qquad (2.3.7(a))$$

$$a_i = -b_i\,, \qquad i = 1,\ldots,M \qquad (2.3.7(b))$$

$$\hat{P}_{M+1} = \sum_{i=0}^{M} a_iR(i)\,. \qquad (2.3.7(c))$$

We now see that the filter of Eq. (2.2.11) can be identified as a unit PEF associated with the prediction filter of Eq. (2.3.3). Equations (2.3.7) give the relations between the corresponding parameters.

Figure 2.3 - 1 illustrates a prediction filtering model described in the present section. The calculation of the filter coefficients

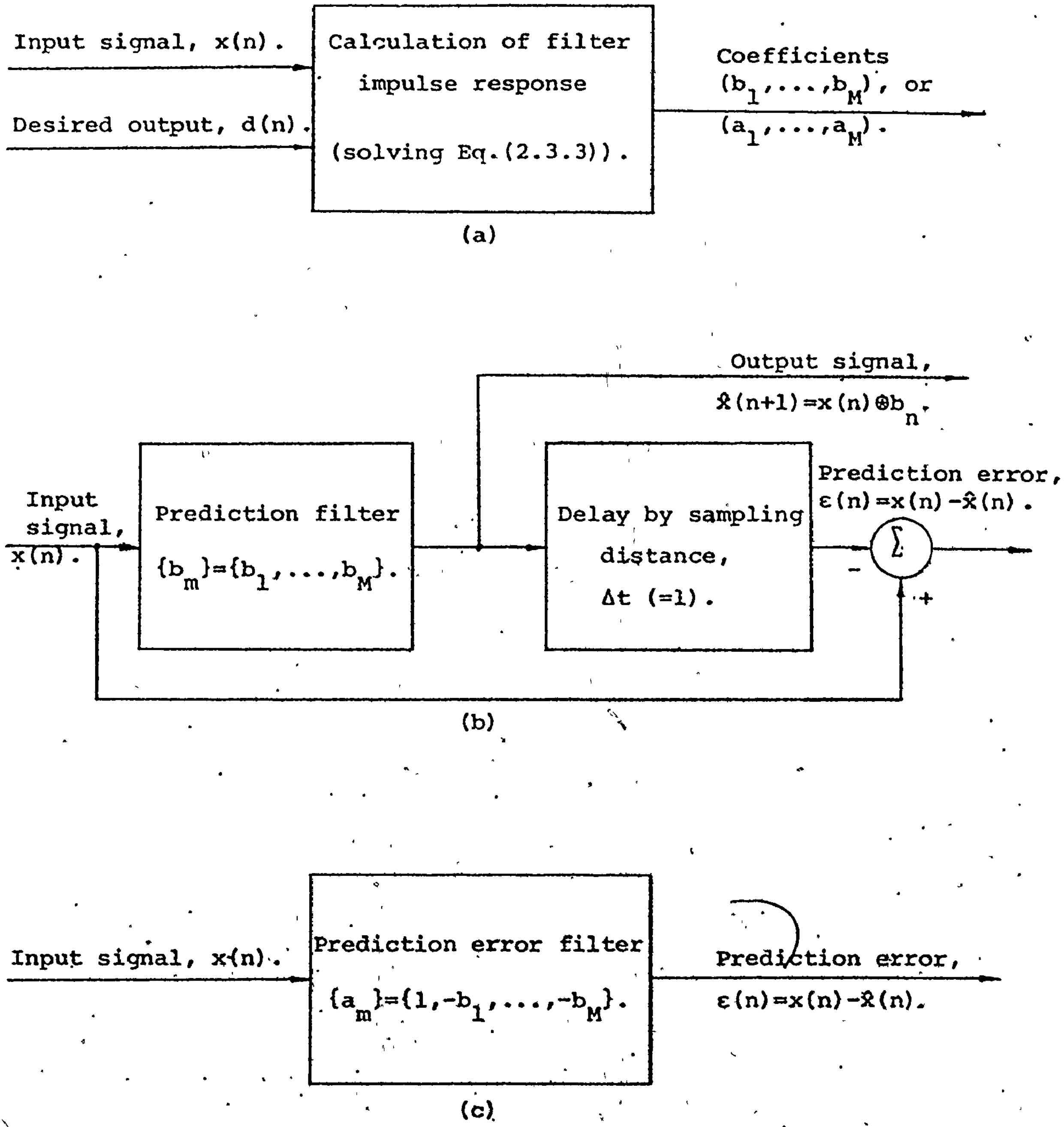


Fig. 2.3 - 1 The predictive filter model: (a) design of filter coefficients, (b) application of prediction filter, (c) alternative form with prediction error filter.

is shown in Fig. 2.3 - 1(a), and its application in prediction is given in Fig. 2.3 - 1(b). The error sequence in the prediction of a future value of the input signal is obtained by delaying the filter output and subtracting the current estimate from the true value. The error sequence, $\varepsilon(n) = x(n) - \hat{x}(n)$, is the output of the alternate form, the PEF, shown in Fig. 2.3 - 1(c).

2.4 Revised Levinson Algorithm

An efficient algorithm for solving Eq. (2.2.11) recursively is given by Levinson in a paper which appears as an Appendix in Wiener's classic work on least-squares filtering [77]. In this procedure use is made of the Toeplitz structure of the autocorrelation matrix (Eq. (2.2.2)) to reduce the number of computations from N^3 for a general matrix inversion case, to N^2.

The Levinson algorithm is based on the minimization of the mean-square error with respect to all filter coefficients. This procedure may sometimes lead to an autocorrelation matrix which is not positive-definite. The modified algorithm, suggested by Burg [44] avoids this problem. It starts from the known matrix of the first M+1 estimates of the ACF values given by Eq. (2.2.5). We now introduce superscripted notation for this matrix, $\underline{R}^{M+1}(m)$, to emphasize its dependence of the number of lags; M+1 is the dimension of the (square) matrix. The MEM gives an estimate of the ACF value, $R(M+1)$, such that the extended matrix $\underline{R}^{M+2}(m)$ is non-negative definite. Hence, the determinant $\det\underline{R}^{M+2}(m)$ must be non-negative. Being a quadratic function of $R(M+1)$, $\det\underline{R}^{M+2}(m)$ becomes zero for two values of the unknown R(M+1). Since the second derivative of $\det\underline{R}^{M+2}(m)$ with respect to R(M+1) is

$$\frac{d^2[\det\underline{R}^{M+2}(m)]}{d[R(M+1)]^2} = -2\cdot\det\underline{R}^{M}(m) \quad , \tag{2.4.1}$$

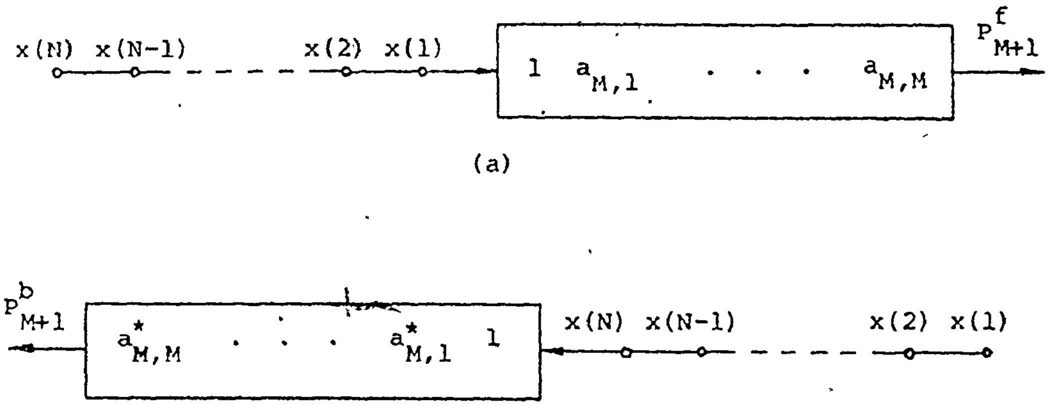
which is always negative, it follows that the permissible values for R(M+1) are those between the solutions of the quadratic equation $\det\underline{R}^{M+2}(m) = 0$, including the solutions themselves. Selecting the value for the estimate of R(M+1) to be exactly in the middle of the permissible range, $\det\underline{R}^{M+2}(m)$ will be a maximum. The entropy of the (M+2)-dimensional Gaussian probability density function with AC matrix $\underline{R}^{M+2}(m)$ is equal to [74],[78]

$$H(M+1) = \ell n\{(2\pi e)^{(M+2)/2}\cdot[\det\underline{R}^{M+2}(m)]^{1/2}\} \quad . \tag{2.4.2}$$

Therefore, maximizing the determinant $\det\underline{R}^{M+2}(m)$ is equivalent to maximizing the entropy.

We can now solve Eq. (2.2.11) iteratively by increasing the matrix dimension from M+1 to M+2. From now on we introduce a double subscript for filter coefficients, to distinguish between the filter order (first subscript) and the order of the coefficient of the particular filter (second subscript). Thus, $a_{M,m}$ is the mth coefficient of the Mth order filter and $a_{M,M}$ is its last coefficient. Note that $a_{M,0} \equiv 1.0$.

Following a procedure similar to that given by Andersen [79] for real valued data, we start with M=0, the corresponding prediction error filter being trivial with a single coefficient equal to unity. The power P_1 is the value of the ACF for zero lag, given by



(a)

(b)

Fig. 2.4 - 1 The illustration of (a) forward and (b) backward prediction error filtering.

$$P_1 \triangleq R(0) = \frac{1}{N}\sum_{n=1}^{N} x(n)\cdot x^*(n) = \frac{1}{N}\sum_{n=1}^{N} |x(n)|^2 . \qquad (2.4.3)$$

For M=1, we have that the first-order filter with output error power P_2 defined by

$$\begin{bmatrix} R(0) & R^*(1) \\ R(1) & R(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{1,1} \end{bmatrix} = \begin{bmatrix} P_2 \\ 0 \end{bmatrix} . \qquad (2.4.4)$$

In order to solve Eq. (2.4.4) for the three unknowns R(1), $a_{1,1}$, and P_2, an additional condition is required, which is obtained by minimizing the average output error power P_2 with respect to $a_{1,1}$. This power is obtained by averaging the power, P^f_{M+1}, from forward filtering and that from backward filtering, P^b_{M+1}, i.e., when the filter is conjugated, reversed, and run backward in time. The forward and backward filtering are illustrated in Figs. 2.4 - 1(a), and (b), respectively, for the general case of the Mth order filter [80].

We see that the power resulting from forward filtering is

$$P_2^f = \frac{1}{N-1} \sum_{n=1}^{N-1} |x(n+1) + a_{1,1} \cdot x(n)|^2 \tag{2.4.5}$$

while the power resulting from backward filtering is

$$P_2^b = \frac{1}{N-1} \sum_{n=1}^{N-1} |x(n) + a_{1,1}^* \cdot x(n+1)|^2 \ . \tag{2.4.6}$$

The additional condition required is obtained by minimization of the average of Eqs. (2.4.5) and (2.4.6)

$$P_2 = \frac{1}{2} (P_2^f + P_2^b) \tag{2.4.7}$$

with respect to $a_{1,1}$. Setting $\partial P_2/\partial a_{1,1}$ to zero, we find that the output power is minimum for

$$a_{1,1} = \frac{-2 \cdot \sum_{n=1}^{N-1} x^*(n) \cdot x(n+1)}{\sum_{n=1}^{N-1} (|x(n)|^2 + |x(n+1)|^2)} \ . \tag{2.4.8}$$

From Eq. (2.4.4), P_2 is obtained in terms of $a_{1,1}$ and P_1 as

$$P_2 = P_1(1 - a_{1,1} \cdot a_{1,1}^*) = P_1(1 - |a_{1,1}|^2) \tag{2.4.9}$$

which, of course, is a real quantity.

Extending this procedure to higher order filters we find from Fig. 2.4 - 1 that the average output power from the Mth order filter is

$$P_{M+1} = \frac{P^f_{M+1}+P^b_{M+1}}{2} = \frac{1}{2(N-M)}\sum_{n=1}^{N-M}\left[\left|\sum_{m=0}^{M} a_{M,m}x(n+M-m)\right|^2+\left|\sum_{m=0}^{M} a^*_{M,m}x(n+m)\right|^2\right] . \tag{2.4.10}$$

Differentiating P_{M+1} with respect to the last coefficient, $a_{M,M}$, and setting the derivative to zero, we obtain the optimum value of $a_{M,M}$

$$a_{M,M} = \frac{-2\sum\limits_{n=1}^{N-M}\left[\left(\sum\limits_{m=0}^{M} a_{M-1,m}x(n+M-m)\right)\left(\sum\limits_{m=0}^{M} a^*_{M-1,m}x(n+m)\right)^*\right]}{\sum\limits_{n=1}^{N-M}\left[\left|\sum\limits_{m=0}^{M} a_{M-1,m}x(n+M-m)\right|^2 + \left|\sum\limits_{m=0}^{M} a^*_{M-1,m}x(n+m)\right|^2\right]} , \tag{2.4.11}$$

where $a_{M,m} \equiv 0$ for $m>M$.

The introduction of the average error power by filtering in both directions ensures that the corresponding autocorrelation matrix is positive-definite in each recursion step [51]. Note also from Fig. 2.4 - 1 that the filter is not run off the ends of the data sample. Thus no assumptions about the signal before and after the data sample are required. This is in contrast in either the zero or periodic extension of the estimated ACF which is inherent in linear methods.

The remaining coefficients $a_{M,m}$ $m = 1,2,\ldots,M-1$ are obtained from the lower M equations in (2.2.11) (Appendix A) in terms of $a_{M,M}$ and the coefficients $a_{M-1,m}$, $m = 1,2,\ldots,M-1$ (Eq. (A.8))

$$a_{M,m} = a_{M-1,m} + a_{M,M}\cdot a^*_{M-1,M-m} , \quad m = 1,2,\ldots,M-1. \tag{2.4.12}$$

Note that the set of equations (2.4.12) hold formally for $m=0,1,\ldots,M$, if we recall the definitions for $a_{M,0}(\equiv 1.0)$, and $a_{M,m}(\equiv 0.0$, for $m > M)$.

The minimum output power P_{M+1} can, of course, be obtained by substituting (2.4.11) into (2.4.10). However, as shown in Appendix A, P_{M+1} is given by a simple expression (A.10)

$$P_{M+1} = P_M(1 - a_{M,M}\cdot a^*_{M,M}) = P_M(1 - |a_{M,M}|^2) . \quad (2.4.13)$$

We have now calculated all the quantities necessary for evaluating the maximum entropy power spectral estimate given by Eq. (1.5.5). After the calculation of the prediction error of the desired length $M = M_{max}$, by using Eqs. (2.4.11) to (2.4.13) in a recursive manner, we introduce those quantities into Eq. (1.5.5), choose the desired frequency resolution and calculate the spectrum. Although the algorithm does not require knowledge of the ACF values, we can calculate them simultaneously at any step M in the recursion, by using the bottom equation in (A.2), which yields

$$R(M) = - \sum_{m=0}^{M-1} a^*_{M,M-m}\cdot R(m) = - \sum_{m=1}^{M} a^*_{M,m}\cdot R(M-m) . \quad (2.4.14)$$

The chioce of M_{max} depends on the statistical properties of the input signals and on the desired resolution. We shall say more about it in section 2.6. Here we note that once $M = M_{max}$ is chosen we can calculate the ACF values up to M_{max}. However, we may wish to extrapolate R(m) up to and including the lag value L > M. This may often happen when the chosen value of M in the iteration process is very small. The criterion for the ACF estimates beyond the lag M is that these values cannot reflect any additional information about the signal, that is, the change in entropy after the Mth iteration should be equal to zero. It will be shown later that this condition is equivalent to [81]

$$a_{M+1,M+1} = a_{M+2,M+2} = \dots = a_{L,L} = 0. \qquad (2.4.15)$$

Thus, using the second form of the R(M) from Eq. (2.4.14) in Eq. (2.4.15), we have the extrapolated values for R(m)

$$R(M+\ell) = -\sum_{m=1}^{M} a^{*}_{M,m} R(M-m+\ell) \quad , \quad \ell = 1,2,\dots,L-M. \qquad (2.4.16)$$

The values obtained by Eq. (2.4.16) have the property that they tend to zero as L increases, thus equation (2.4.16) is sometimes usefull for reducing truncation errors without introducing spurious information.

2.5 Procedure for Computer Calculation of the Complex Filter Coefficients

A numerical procedure for the iterative solution of the PEF parameters follows. Substituting Eq. (2.4.12) into (2.4.10) we have for the average error power

$$P_{M+1} = \frac{1}{2(N-M)} \sum_{n=1}^{N-m} \left[\left| \sum_{m=0}^{M} a_{M-1,m} x(n+M-m) + a_{M,M} \sum_{m=0}^{M} a^{*}_{M-1,M-m} x(n+M-m) \right|^{2} \right.$$
$$\left. + \left| \sum_{m=0}^{M} a^{*}_{M-1,m} x(n+m) + a^{*}_{M,M} \sum_{m=0}^{M} a_{M-1,M-m} x(n+m) \right|^{2} \right] . \qquad (2.5.1)$$

Summations with the respect to m in Eq. (2.5.1) do not depend on $a_{M,M}$. Noting that the first sum in the square brackets equals the last one (only the order of summation is reversed), and the same for the second and the third sum, we denote them as follows

$$p_M(n) = \sum_{m=0}^{M} a^*_{M-1,M-m} x(n+M-m) = \sum_{m=0}^{M} a^*_{M-1,m} x(n+m) \qquad (2.5.2)$$

$$q_M(n) = \sum_{m=0}^{M} a_{M-1,m} x(n+M-m) = \sum_{m=0}^{M} a_{M-1,M-m} x(n+m) \qquad (2.5.3)$$

where $n = 1,2,\ldots,N-M$. Introducing (2.5.2) and (2.5.3) into (2.5.1) or directly into (2.4.11), the condition $\partial P_{M+1}/\partial a_{M,M} = 0$ yields [80]

$$a_{M,M} = \frac{-2\sum_{n=1}^{N-M} p^*_M(n)\cdot q_M(n)}{\sum_{n=1}^{N-M}\left(|p_M(n)|^2 + |q_M(n)|^2\right)} \ . \qquad (2.5.4)$$

Recursive formulae for $p_M(n)$ and $q_M(n)$ may be derived by using the relationships between the filter coefficients, Eq. (2.4.12). The derivation is given in Appendix B. Equations (B.3) and (B.6) in Appendix B give

$$p_M(n) = p_{M-1}(n) + a^*_{M-1,M-1}\cdot q_{M-1}(n) \qquad (2.5.5)$$

$$q_M(n) = q_{M-1}(n+1) + a_{M-1,M-1}\cdot p_{M-1}(n+1) \ . \qquad (2.5.6)$$

For the "zeroth" order filter, $M = 0$, we have simply

$$p_o(n) = q_o(n) = x(n), \quad n = 1,2,\ldots,N. \qquad (2.5.7)$$

Since the recursion starts from $M = 1$, the actual starting values are

$$p_1(n) = x(n), \qquad n = 1,2,\ldots,N-1, \qquad (2.5.8)$$

$$q_1(n) = x(n+1), \qquad n = 1,2,\ldots,N-1. \qquad (2.5.9)$$

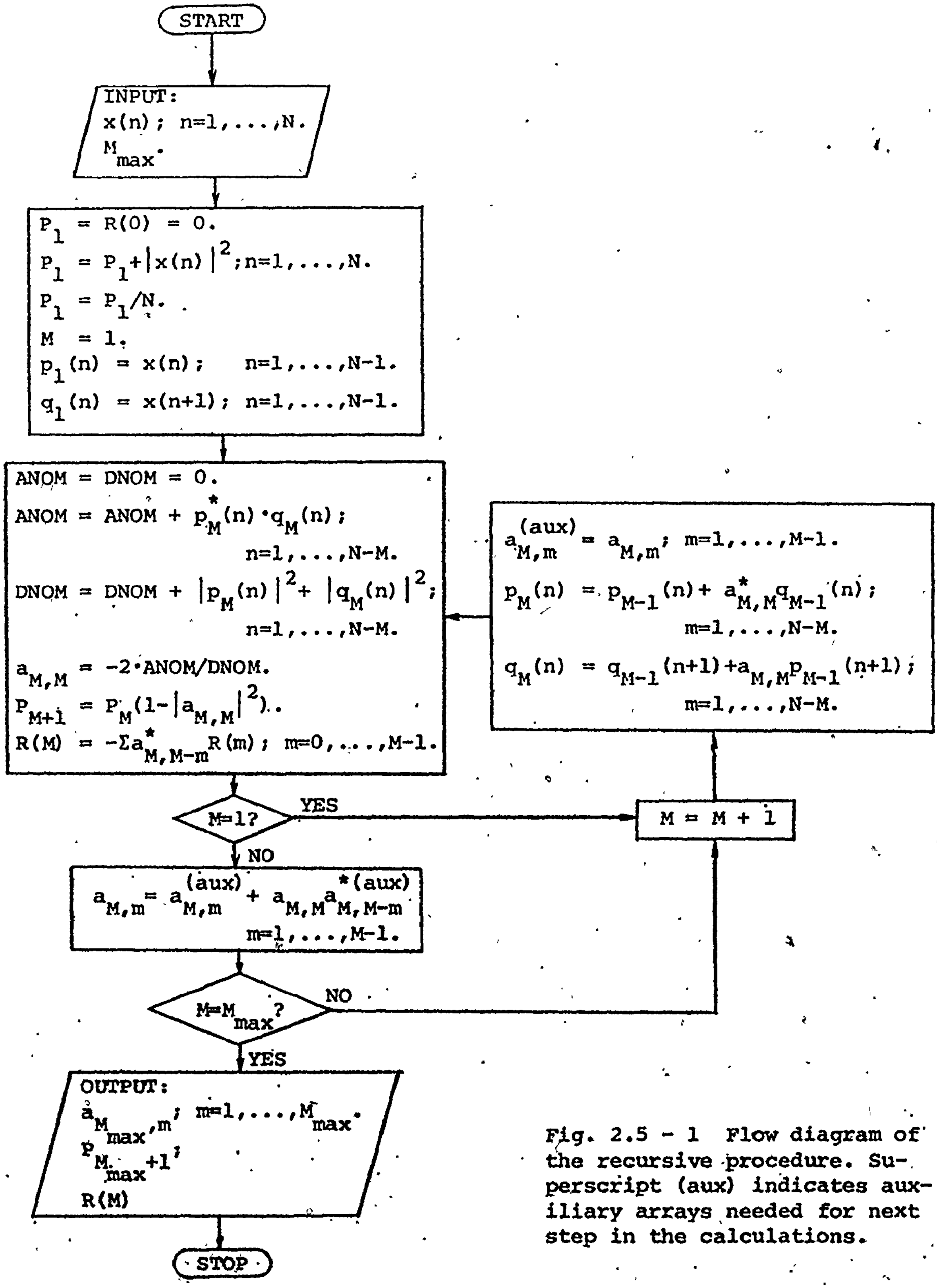


Fig. 2.5 - 1 Flow diagram of the recursive procedure. Superscript (aux) indicates auxiliary arrays needed for next step in the calculations.

We can now outline the complete calculation procedure for the filter coefficients. Note that the calculation of the ACF values in each step is independent of the remaining calculations and may thus be optional.

The steps in the iterative procedure are as follows:

(1) From a given signal sample $x(n)$, $n = 1,2,\ldots,N$ calculate the power P_1 $(= R(0))$ using Eq. (2.4.3), and the parameters $p_1(n)$ and $q_1(n)$ by means of Eqs. (2.5.8) and (2.5.9).

(2) Using $M = 1$, calculate $a_{1,1}$ from (2.5.4), P_2 from (2.4.13), and R(1) from (2.4.14).

(3) Increment M by 1, and calculate the next values for $p_M(n)$ and $q_M(n)$ from Eqs. (2.5.5) and (2.5.6).

(4) For the current value of M, calculate $a_{M,M}$, P_{M+1}, and R(M) using Eqs. (2.5.4),(2.4.13), and (2.4.14), respectively.

(5) Calculate the remaining coefficients $a_{M,m}$, $m = 1,2,\ldots,$ M-1 through (2.4.12) and store them.

(6) Repeat steps (3) - (5) for each value of M, up to the preselected value M_{max}.

We now have all the quantities necessary for calculating the spectral density estimate given by Equation (1.5.5). The recursive procedure is summarized in the flow diagram shown in Fig. 2.5 - 1. Superscript (aux) indicates auxiliary arrays needed for the next step.

2.6 Entropy Considerations and the Choice of the PEF Order

Let us consider the process x(n) with the joint probability density function of the first M samples $p[x(1),\ldots,x(M)]$, In general, the entropy of x(n) is defined to be

$$H(x) = -\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} p[x(1),\ldots,x(M)]\cdot\log\{c_M p[x(1),\ldots,x(M)]\}dx(1)\ldots dx(M) \tag{2.6.1}$$

where the base of the logarithm is any positive number and c_M is an arbitrary scale factor. If x(n) is a Gaussian process, Eq. (2.6.1) reduces to

$$H(x) = \log\frac{(2\pi e)^{M/2}}{c_M} + \frac{1}{2}\log\{\det\underline{R}^{M+1}(m)\}\ . \tag{2.6.2}$$

Choosing the scale factor to be $c_M = (2\pi e)^{M/2}$, and using natural logarithms, we have

$$H(M) = \frac{1}{2}\ln\{\det\underline{R}^{M+1}(m)\}\ . \tag{2.6.3}$$

In Equation (2.6.3) we have emphasized the dependence of the entropy on the order of the known autocorrelation lag, M. Thus, in each iteration, M, H(M) is the measure of entropy exhibited by the corresponding spectral density. For example, if the iteration process is terminated after the "zeroth" recursion, the spectral density (Eq. (1.5.5) with M = 0) is constant for all frequencies

$$\hat{S}_E(f)_{(M=0)} = \frac{P_1}{|a_{0,0}\cdot e^{0}|^2} = P_1 = R(0)\ . \tag{2.6.4}$$

The autocorrelation matrix is reduced to a single term R(0), so that the measure of entropy exhibited by this spectrum is

$$H(0) = \frac{1}{2} \ln R(0) \quad . \tag{2.6.5}$$

The measure of entropy thus obtained would be realistic for the case of white noise only. If there is the some amount of information in the relationships which exist between pairs of adjacent points, this information would be indicated by

$$H(1) = \frac{1}{2} \ln R(0) + \frac{1}{2} \ln (1 - |a_{1,1}|^2) \quad , \tag{2.6.6}$$

which can be calculated either by finding the corresponding all-pole rational spectral density for M = 1 or, by simply finding the value of $\det \underline{R}^2(m)$ from Eq. (2.4.4). Extending this procedure to the higher order filter, we find that the measure of entropy after the Mth recursion is

$$H(M) = \frac{1}{2} \ln R(0) + \frac{1}{2} \sum_{m=1}^{M} \ln (1 - |a_{m,m}|^2) \quad . \tag{2.6.7}$$

This may be considered to be a measure of information remaining "in" the data after the Mth recursion. Thus, the amount of information "extracted " by the Mth recursion is the difference

$$\Delta H(M) = H(M-1) - H(M) = - \frac{1}{2} \ln (1 - |a_{M,M}|^2) \quad . \tag{2.6.8}$$

Since $|a_{M,M}| < 1$, the logarithm in Eq. (2.6.8) is negative, and therefore $\Delta H(M) > 0$ for any M. Moreover, one can expect that in each iteration, additional information has been extracted, so that $\Delta H(M)$ decreases with M. Of course, with M approaching the data length, N, statistical accuracy decreases. Thus, we may use $\Delta H(M)$ as a subjective criterion for terminating the recursion. Depending on the nature of the data, we can decide to choose for M_{max} that value of M for which $\Delta H(M)$ falls below a certain value or below a certain percentage of H(0). Even with more objective criteria, given below, we can use the concept of entropy increment to get some kind of generalization.

Three different objective criteria have been proposed for choosing the filter order. The first one, called final prediction error (FPE) has been suggested by Akaike [82] - [84]. The FPE is defined as an estimate of the mean-square error in prediction (i.e., the prediction variance in a case of zero-mean process) expected when a predictor, calculated from one observation of the process, is applied to another independent observation of the same process. For a filter of order M, the FPE is defined by

$$\mathrm{FPE}(M) = P_{M+1} \frac{N + M + 1}{N - M - 1} . \qquad (2.6.9)$$

Since P_{M+1} decreases with M, while the second term in Eq. (2.6.9) increases, the FPE(M) will have a minimum at some value $M = M_{opt}$. This value is used for the PEF order in an iterative procedure for coefficient calculation.

Another criterion, also suggested by Akaike [85], is the minimization of the log-likelihood of the prediction error variance as a function of M. This criterion, called the information theoretic criterion (AIC) is given by

$$AIC(M) = \ln(P_{M+1}) + \frac{2M}{N} . \qquad (2.6.10)$$

Again, AIC(M) may be calculated in each recursion and M_{opt} is a value for which the AIC(M) is minimum.

A third criterion has been proposed by Parzen [86] and is known as the autoregressive transfer function criterion (CAT). The order M_{opt} is obtained where the estimate of the difference in the mean-square errors between the true filter, which exactly gives the prediction error, and the estimated filter, is minimum. Parzen [86] has shown that this difference can be calculated, without explicitly knowing the exact infinite filter, by using

$$CAT(M) = \frac{1}{N}\sum_{m=1}^{M} \frac{N-m}{N} P_m^{-1} - \frac{N-M}{N} P_M^{-1} . \qquad (2.6.11)$$

The application of any of the above criteria yields a value for M_{opt} which is, in most cases, the best compromise between the variance and the bias of spectral estimate. It has been demonstrated that for one period of the complex exponential signal, for example, all three criteria give the same value for the M_{opt} [62].

Based on the experience gained by using the described criteria for the clutter return analysis, and on the experimental results presented in [62], we have chosen the FPE criterion for the analysis in the following chapters, since it gives, in general, slightly better results than the other two.

2.7 Summary

Chapter 2 deals with a detailed description of the MEM which is used in this thesis for spectral analysis and classification of clutter signals. The connection between the estimates of the PSD and the ACF is given. It was shown that the corresponding estimates are related by the coefficients of the Wiener PEF. The relation between the prediction filtering and the prediction error filtering is given, because it is closely related to the method. Burg's revision of the Levinson algorithm is reviewed, which ensures positive-definiteness of the autocorrelation matrix and minimum phase property for the PEF. We have developed a complex version of the MEM algorithm, since it is needed for processing coherent video signals.

The method has also been discussed from the entropy standpoint. Three different criteria for selecting the optimum filter order have been reviewed. Akaike's FPE criterion exhibits slightly better results in the analysis of clutter data than the other two.

CHAPTER 3

CLUTTER CONSIDERATIONS AND ADEQUACY OF AN ALL-POLE MODEL

Before proceeding with the analysis of simulated and actual clutter data, it is necessary to explore the applicability of the methods used. In this chapter we first give examples of estimating the spectra of a narrow-band second-and fourth-order AR processes. After discussing the clutter models and the criterion for choosing the appropriate one for the analysis, we give a rationale for an all-pole representation of clutter spectra.

3.1 Spectral Estimation of All-Pole Processes

As mentioned in Chapter 1 the asymptotic behavior of spectral analysis methods have been investigated. It has been found that, in general, the linear ond nonlinear methods give spectral estimates with similar statistical behavior when the record length N, and the maximum estimated autocorrelation lag M, become very large. Even though such analyses give some insight into the behavior of the spectral estimator, they tend to be misleading in comparing the merits of the linear techniques, and methods such as the maximum entropy estimation which is based on modelling the observed data.

For example, if a set of data can be modelled by a narrow-band second-order all-pole process, it will take a prediction error filter of low order to generate an ME estimate of its spectrum with minimum bias.

On the other hand, the number of lags of the autocorrelation estimate required to obtain a spectral window with satisfactorily resolution might be large enough to make the corresponding linear estimate unacceptable due to its instability.

To illustrate the applicability of the maximum entropy method to the analysis of of data records that can be approximated by the autoregressive model of a finite order, we will consider the three examples of the all-pole process. The usefulness of the minimum final prediction error, discussed in Chapter 2 will also be demonstrated in these examples.

Recalling Eq. (2.2.1) the second-order process can be represented by

$$x(n) = -a_1 x(n-1) - a_2 x(n-2) + w(n) \tag{3.1.1}$$

where $\{w(n)\}$ is a white noise sequence with zero mean and variance σ_w^2. The theoretical spectral density of $x(n)$ is given by [38]

$$S_x^{(2)}(f) = \frac{\sigma_w^2}{|1 + a_1 e^{-j2\pi f} + a_2 e^{-j4\pi f}|^2} . \tag{3.1.2}$$

We wish to have a process whose autocorrelation function is a damped sine wave, with spectral peak at some frequency f_o, and with most of its power located at the lower frequency range. To meet these requirements the parameters a_1 and a_2 should satisfy the following relations [38]

$$-(1 + a_2) < a_1 < 0 \tag{3.1.3}$$

$$|a_1(1 + a_2)| \leq |4a_2| \tag{3.1.4}$$

$$a_1^2 - 4a_2 < 0. \tag{3.1.5}$$

The location of a spectral peak, f_o, is related to the parameters a_1 and a_2 by

$$\cos(2\pi f_o) = - \frac{a_1(1 + a_2)}{4a_2} . \qquad (3.1.6)$$

Parameters a_1 and a_2 have been chosen to satisfy the inequalities (3.1.3) to (3.1.5) and to give the peak frequency $f_o = 0.155$. With $a_1=-1$, the second parameter is found from Eq. (3.1.6) to be $a_2= 0.8$, and σ_w^2 is set to the value 2.56. With these parameters the plot of Eq. (3.1.2) is as shown in Figure 3.1 - 1. This curve is labelled "theoretical". The peak value of the plot is 90. and occurs at $f_o = 0.155$. The negative portion of the curve is symmetric with respect to the ordinate axis and thus, it was not plotted. The same is true for the other curves in Fig. 3.1 - 1 since the generated signal is real. The remaining curves represent the ME spectral density estimates obtained from 128 samples of x(n), with w(n) obtained from a Gaussian-distributed random number generator. The first 100 samples generated are ignored in order to reduce transient effects. The plot for M = 9 corresponds to the optimum prediction error filter, for which the FPE is a minimum. In Fig. 3.1 - 1 we see that for both M = 5 and M = 14, the ME spectral density estimator exhibits minor peaks which do not appear in the plot for M_{opt}. The number of these peaks and their amplitudes increase with increasing filter order. Also, the location of the mean peak shifts to the right and its amplitude increases with increasing M. We may thus consider the curve corresponding to M_{opt} as an optimum estimate in the sense that it is a compromise between the

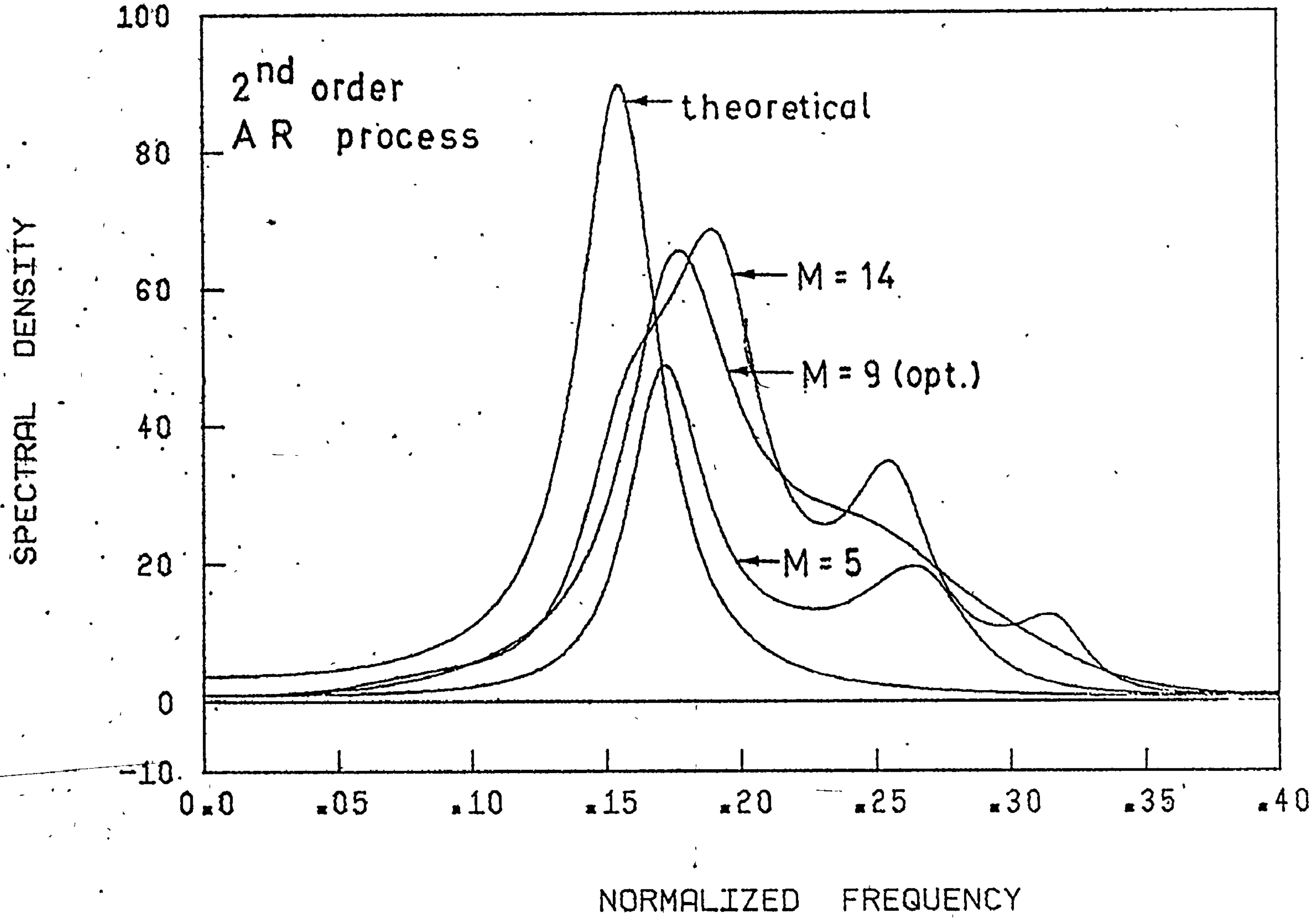


Fig. 3.1 - 1 Theoretical power spectral density for the all-pole process $x(n) = x(n-1) - 0.8x(n-2) + z(n)$; $\sigma_w^2 = 2.56$, and its ME estimates for $M = 6$, $M = M_{opt} = 9$, and $M = 14$.

fidelity of the estimate and the computation time for the filter coefficients. Note that since the curve for FPE as a function of M has a relatively broad minimum, the resulting estimate is somewhat insensitive to the choice of M in the immediate vicinity of M_{opt}.

For a second example, consider the case of a complex second-order all-pole process involving data generated in a manner similar to that in the previous example. In this case, however, two different sets of random numbers are generated according to Eq. (3.1.1), which form the real and imaginary parts of the process. The parameters of the process are $a_1 = -1$ and $a_2 = 0.8 - j0.2$, while σ_w^2 is chosen to be

$$\sigma_w^2 = |1 + a_1 + a_2|^2 = 0.68, \qquad (3.1.7)$$

so that the normalized spectral density at zero frequency is equal to unity. The theoretical spectral density of such a process, and its ME estimates for three different values are shown in Figures 3.1 - 2(a), (b), and (c). The theoretical spectrum exhibits assymetry in that the peak in the negative frequency region is broader, further away from the origin, and has a significantly smaller magnitude than the peak in the positive frequency region.

The spectral density estimates in Fig. 3.1 - 2(a) are obtained with 64 data samples of the process, with the first 50 sample ignored in order to reduce transient effects. It can be seen from Fig 3.1 - 2(a) that the spectral estimates give a good representation of the theoretical spectrum in the positive frequency region, again with closest approximation obtained for the optimal value of M; in the negative frequency region, however, the estimated peaks are shifted away from their actual

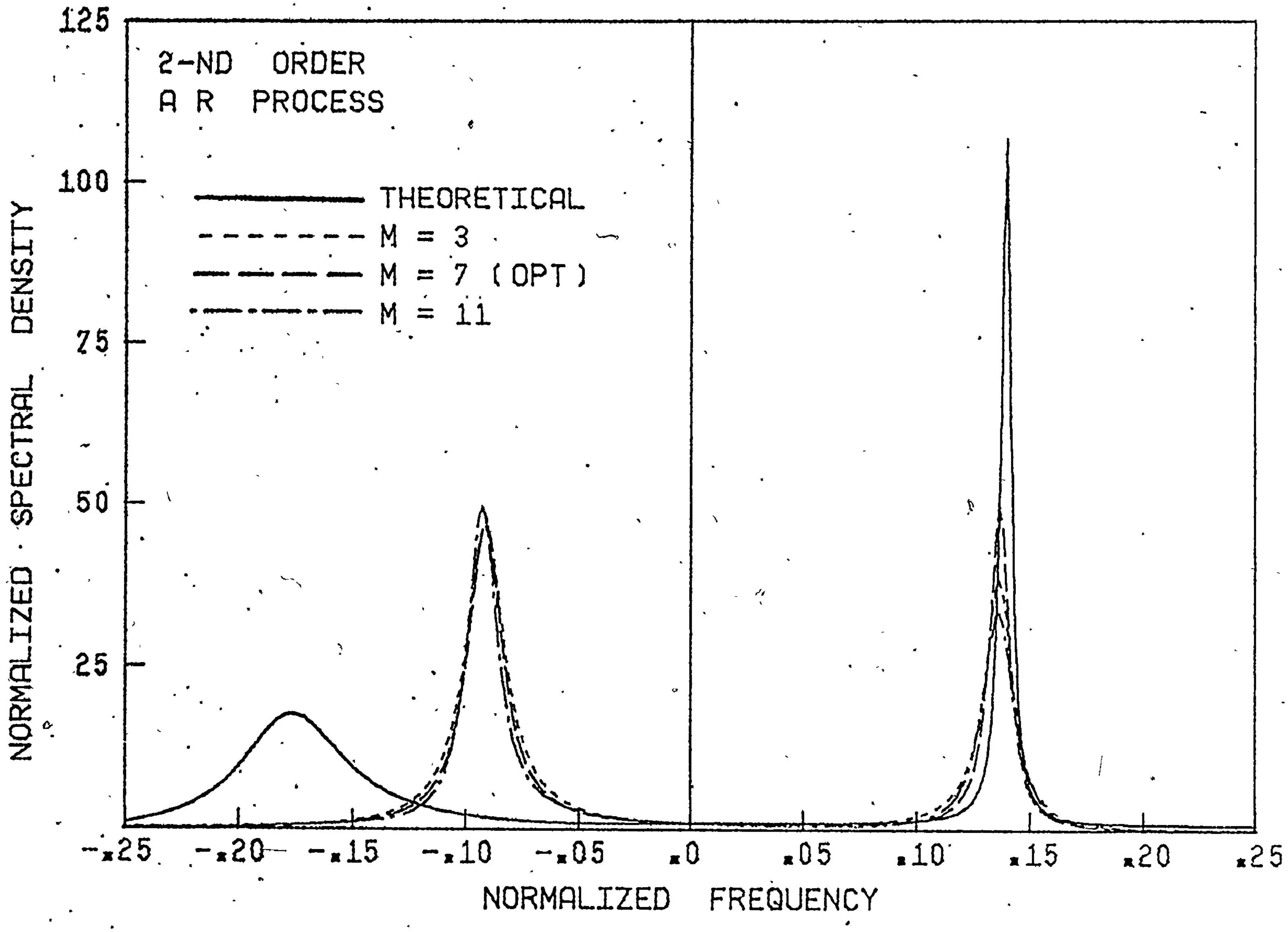


Fig. 3.1 - 2 (a). The theoretical power spectral density for the complex second-order all-pole process with $a_1 = 1.$, $a_2 = 0.8-j0.2$, and $\sigma_w^2 = 0.68$, and its ME estimates with the first 50 samples ignored and N = 64.

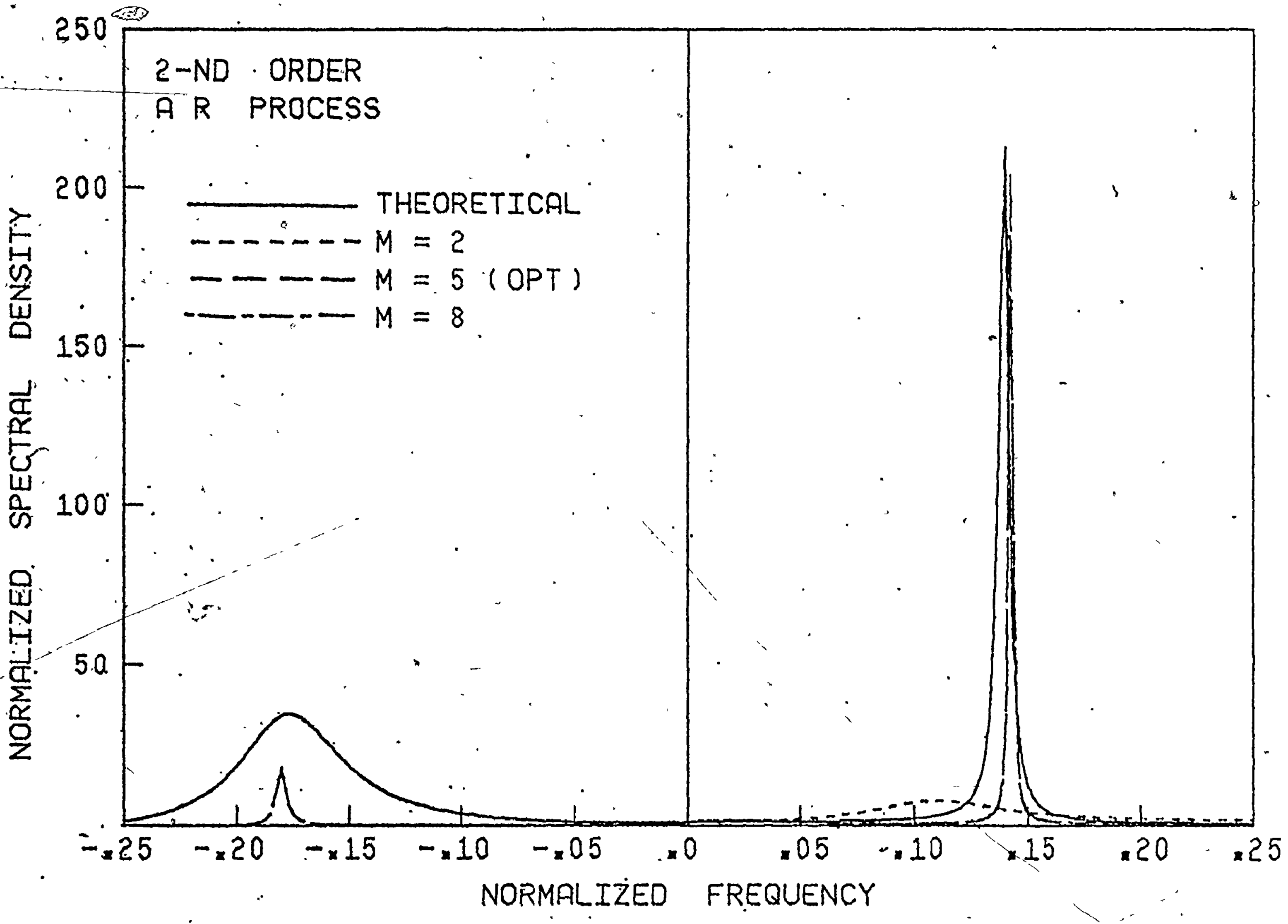


Fig. 3.1 - 2 (b) The same spectra as in Fig. 3.1 - 2 (a) except that the first 114 samples are ignored and N = 64.

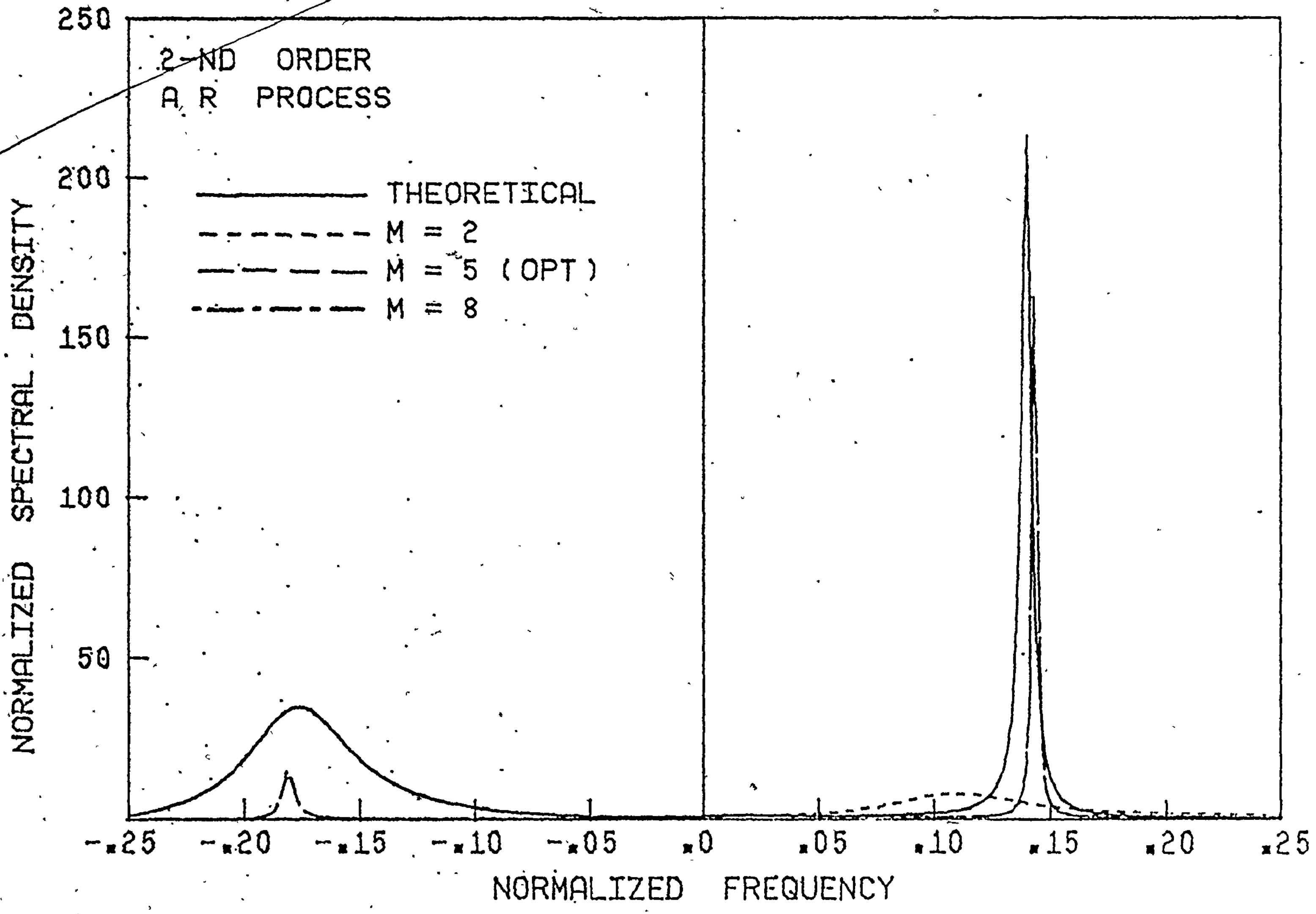


Fig. 3.1 - 2(c) The same spectra with the first 114 samples ignored and N = 192.

position. The reason for this phenomenon is that the number of front-end data samples of the process that were ignored were not adequate to ensure that all transient effects have died away completely. This observation is confirmed in Fig. 3.1 - 2(b) where 114 data samples were ignored as compared to 50 in the case of Fig. 3.1 - 2(a). Note, however, that in Fig. 3.1 - 2(b) the optimum length of the filter is equal to 5. We thus find that in Fig. 3.1 - 2(b) the ME estimates provide a closer approximation to the actual position of the peak for negative frequencies, for both the optimal value of M and for M = 8, than the previous estimates in Fig. 3.1 - 2(a). Note also that the filter of order two gives a poor estimate of the peak in the positive frequency region, and that the peak in the negative frequency region is not resolved at all.

Provided that an adequate number of front-end data samples are ignored so as to make transients effects negligible, the ME estimates are relatively insensitive to the number of data samples used for driving the estimators. This is confirmed in Fig. 3.1 - 2(c) where the number of data samples ignored is the same as that in Fig. 3.1 - 2(b), but the number of data samples used in the spectral analysis is three times as many.

A third example is a fourth-oder all-pole process described by

$$x(n) = -a_1x(n-1) - a_2x(n-2) - a_3x(n-3) - a_4x(n-4) + w(n) \qquad (3.1.8)$$

with a corresponding spectral density

$$S_x^{(4)}(f) = \frac{\sigma_w^2}{\left|1+a_1e^{-j2\pi f}+a_2e^{-j4\pi f}+a_3e^{-j6\pi f}+a_4e^{-j8\pi f}\right|^2} . \qquad (3.1.9)$$

The parameters a_i, $i = 1,2,3,4$ were chosen in such a way to get the two close spectral peaks at frequencies $f_1 = 0.3$ and $f_2 = 0.33$ with respective amplitudes of 8.3 and 10.0. The standard deviation of the innovation (white noise) process w(n) was chosen to be $\sigma_w = |1 + a_1 + a_2 + a_3 + a_4|$ so that the value for $S_x^{(4)}(0)$ is normalized to unity. The parameter values are given in Table 3.1 - 1.

parameter	a_1	a_2	a_3	a_4	σ_w	σ_w^2
value	1.6	2.4	1.44	0.81	7.25	52.56

TABLE 3.1 - 1.

Parameter values for the 4th order all-pole process.

The curve labelled "theoretical" in Fig. 3.1 - 3 is a plot of Eq. (3.1.9), while the remaining three curves represent the ME estimates obtained from 128 samples of x(n). The random number generator used for generating w(n) was the same as in the first example. Again, the plots are shown for positive frequencies only. In Fig. 3.1 - 3 we see that for M = 3 (the value which is less than the theoretical order of an all-pole process), the existing peaks have not been resolved. Besides, the single peak is shifted to the left from the larger peak by 10.4% of the full frequency range. With larger values of M both peaks can be resolved.

Table 3.1 - 2 summarizes the locations and magnitudes of both peaks, as well as the per cent shift from the true peak location. We see that, by increasing M, the positions of the maxima in the corresponding spectral estimates become closer to the exact positions. In this sense the estimate corresponding to M = 11 is better than the one for the op-

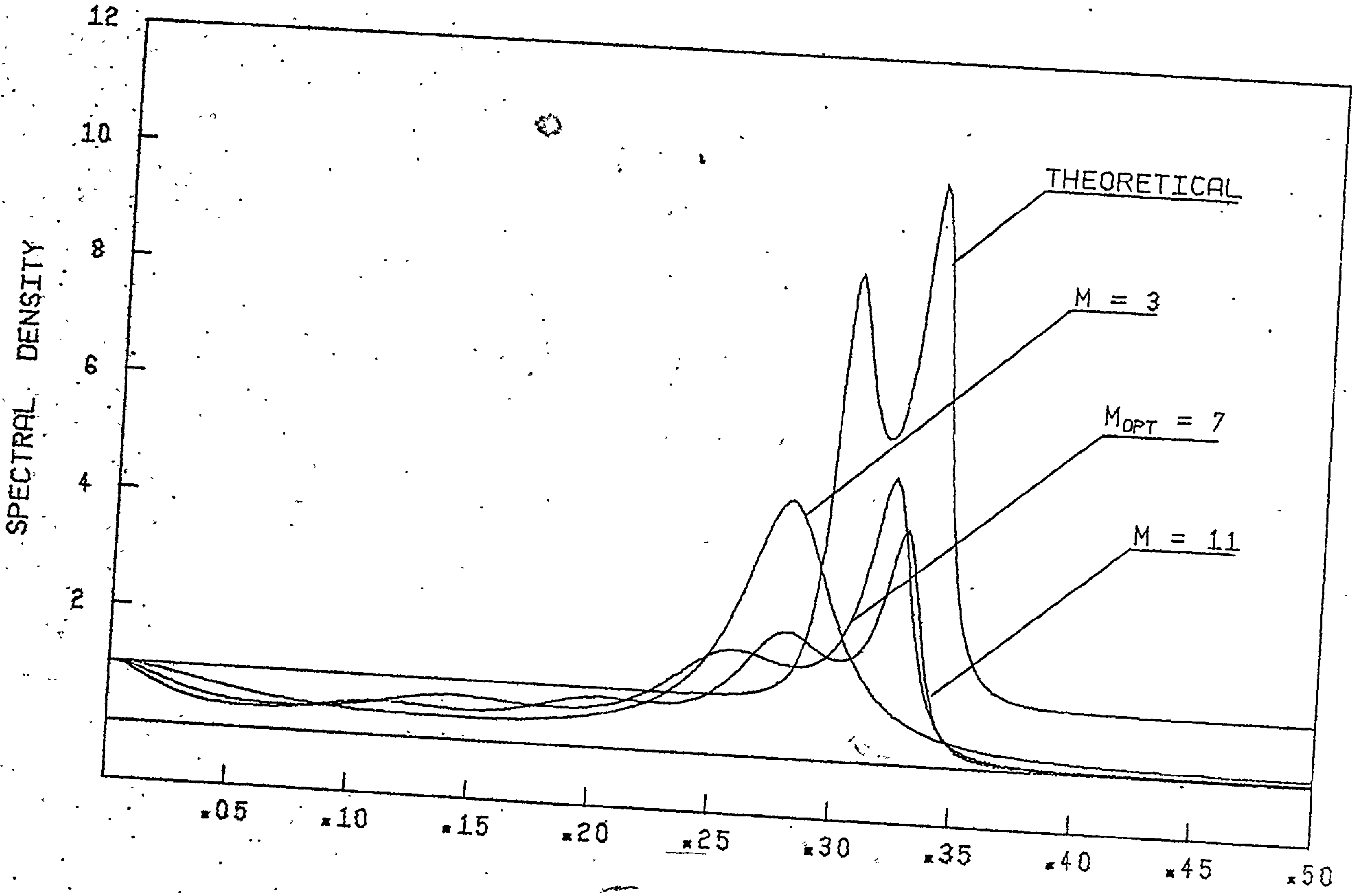


Fig. 3.1 - 3 Theoretical power spectral density for the all-pole model $x(n) = -1.6x(n-1) -2.4x(n-2) - 1.44x(n-3) - 0.81x(n-4) + z(n)$; $\sigma_w^2 = 52.56$, and its ME estimates for $M = 3$, $M = M_{opt} = 7$, and $M = 11$.

Parameter values	I maximum			II maximum		
Spectra	f_1	$\|f_1-f_{1t}\|$ (%)	$S_x^{(4)}(f_1)$	f_2	$\|f_2-f_{2t}\|$ (%)	$S_x^{(4)}(f_2)$
Theoretical	0.300	0.0	8.3	0.330	0.0	10.0
$M = 3$	-	-	-	0.278	10.4	4.4
$M_{opt} = 7$	0.257	8.4	1.8	0.321	1.8	4.9
$M = 11$	0.280	4.0	2.2	0.327	0.6	4.0

TABLE 3.1 - 2.

Locations and magnitudes of the spectral peaks for the theoretical 4th order all-pole spectrum and its ME estimates with $M = 3$, 7(opt), and 11.

timum value of M. It should be noted, however, that by increasing M, the number of minor peaks in the left-hand portion of the diagram also increases.

The results of the above analysis show that the proposed objective criteria for the prediction error filter order can only be used as a guide. A priori knowledge of the signal statistics should also be considered. We see that when the signal sample can be modelled by an all-pole process, the value of M_{opt} will be equal or slightly greater than the order of the process. If this is not the case (for example, when a high level noise is added to the signal), the estimate for M_{opt} is generally not good and the larger values of M should be chosen. In these cases, the curves for the FPE(M) or for the other criteria, usually have a broader minimum. This is in agreement with the resuls of section 2.6.

3.2 Description of the Selected Clutter Model

Amongst the various models for clutter available in the literature, we have chosen the one proposed by Hawkes and Haykin [87], since it gives a good representation of real clutter signal for a large number of cases and, yet, it is computationally simple and efficient. The model is characterized by four parameters which directly correspond to the parameters of scatterers that reflect the energy transmitted by radar. We can vary these parameters at will, thus we can simulate the signals with different statistical properties. The description of the clutter model follows.

When an RF pulse is transmitted from a radar toward a cloud of random scatterers, part of the energy is reflected back to the radar by the scatterers independently of one another. As the transmitted wave reaches the scatterers, echoes are received at a rate that depends on the local density of the cloud. Wong et. al. [27],[28] have developed a clutter model by treating the scatterers as dipoles, instead of points with variable cross-sections. This assumption permits the determination of the effects of rotation of the scatterers, which were previously neglected.

For the case of linear transmit-linear receive, the radar cross-section developed in [28] reduces to one that varies sinusoidally with time. In particular, the functional dependence of the radar echo from a dipole scatterer located at azimuth θ_i and range k may be expressed as follows.

$$s_{i,k}(t,f_r) = s'_{i,k}\{C + Ue^{j2(2\pi f_r t+\alpha)} + Le^{-j2(2\pi f_r t+\alpha)}\} \quad (3.2.1)$$

where

$s'_{i,k}$ = broadside echo voltage of the dipole with the electric field vector parallel to its axis,

f_r = rotational frequency of the dipole,

α = initial angle between the dipole axis and the plane perpendicular to the direction of propagation,

C,U,L = random variables which depend on polarization and orientation of the rotation axis of the dipole.

For the case of linear transmit-linear receive, we have

$$\frac{1}{4} E[C^2] = E[U^2] = E[L^2] \quad . \tag{3.2.2}$$

These quantities represent the relative magnitudes of the mean backscattered power contained in the three spectral components which make up the received echo from a single dipole scatterer.

Consider Fig. 3.2 - 1 where it is assumed that the antenna is scanning around at ω_s radian per second; at time t = 0, θ = 0. The model assumes uniformly distributed scatterers in a small range-azimuth cell both along the radial lines and along the azimuth increment. Each radial line is identified by azimuth angle, θ_i and along it there are scatterers distributed at uniform intervals, indexed by the letter "k".

We shall characterize the received clutter signal in terms of its autocorrelation function. For this purpose, we first compute the return from one elemental dipole scatterer, sum all the individual contributions to determine the clutter signal, and then form the autocorrelation function of the signal. The calculation is detailed in Appendix C.

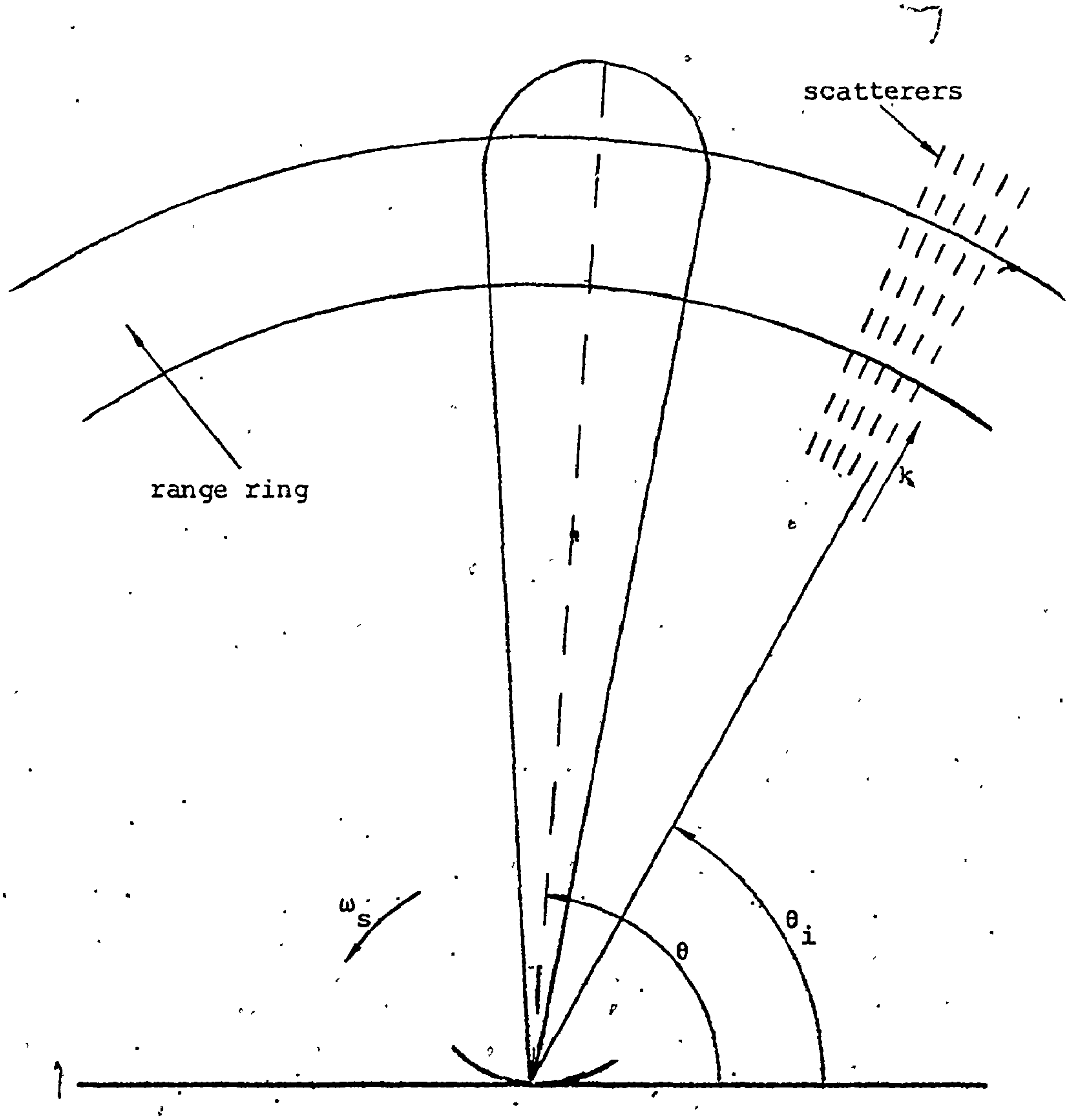


Fig. 3.2 - 1 The assumed clutter model.

For the transmitted signal, we shall assume a coherent pulse train defined by

$$\psi_T(t) = \mu_T(t) \cdot e^{j2\pi f_c t} \qquad (3.2.3)$$

where f_c is the radar carrier frequency. The envelope of this signal is

$$\mu_T(t) = \sum_{p=0}^{P-1} a_p \mu_s(t - pT_R) \qquad (3.2.4)$$

where

$\mu_s(t)$ = envelope of a single pulse,

a_p = weighting given to each pulse by the antenna,

P = number of pulses transmitted,

T_R = interpulse period.

For the assumed Gaussian distributions for the dipole rotation frequency, f_r, and Doppler shift, f_d, the normalized autocorrelation function of the clutter is given by Eq. (C.7) which is reproduced here for convenience

$$R_I(q) = \frac{2}{3}\left[1 + \frac{1}{2}\cos(4\pi q \mu_r) e^{-8\pi^2 q^2 \sigma_r^2}\right] \cdot e^{-j2\pi q \mu_d - 2\pi^2 q^2 \sigma_d^2 - \frac{q^2}{2T_o^2}} . \qquad (3.2.5)$$

In Eq. (2.3.5) T_o and the model parameters μ_r, μ_d, σ_r, and σ_d have been normalized with respect to the interpulse period T_R.

The model used for the spectral analysis of clutter signal is based on Eq. (C.3). As mentioned in Appendix C the dynamic range for the main lobe of the antenna, and the size of the scatterer array have been chosen to satisfy the trade-off between the adequate representation of the scatterer field and the computational time.

3.3 Adequacy of an All-Pole Representation of Clutter Spectra

As mentioned in section 1.4, the applicability of a particular method of spectral analysis depends on the statistical properties of the signal being analyzed. For linear processes, Treitel et. al. have shown [71],[72] that for the theoretical models corresponding to MA, AR, and ARMA processes the best procedures are, respectively, the lag-window, the maximum entropy, and the iterative least-square procedures. In the case of some physical signals encountered in practice it is difficult, if not impossible, to determine which model best fits to the observed data. There is no simple and reliable test which can identify a real life situation with some particular model. A number of theoretical tests have been devised for this purpose (e.g. see [88,ch.5]), but none appear to be satisfactory from a practical viewpoint.

In the case of clutter signals which are obtained by reflection of transmitted waves from various objects, there exists no unique way of determining the spectral properties of the signal. However, a considerable amount of effort has been expended in order to find a model which best fits the particular situation. In the classic works by Barlow [20] and Goldstein [21] the clutter spectrum is assumed to be Gaussian for the majority of fluctuacting clutter targets. This spectral shape has since been confirmed by other authors [22],[23]. A somewhat convenient form of the expression for clutter spectrum developed by Barlow [20] is

$$S(f) = e^{-f^2/2\sigma_f^2} \tag{3.3.1}$$

where σ_f is the spectral spread. However, other investigators have found that a better fit to experimental data is obtained when the power spectral density is represented by an all-pole spectrum of order three [25] or of order four [26,p.9]. It is therefore apparent that there is no unique representation of the clutter spectrum which will be valid for all cases. Both the transcedental (Gaussian shaped) and rational (all-pole) spectra have been used as a reasonable approximation for the particular representation.

The clutter may be modelled as an autoregressive (all-pole) process of a finite order. The representations in [25], and [26,p.9] give the model orders of three and four, respectively. On the other hand, the Gaussian shaped model, assumed by the other workers, may be approximated accurately by an all-pole rational model. For example, the absolute errors in approximating the function

$$S(f) = \frac{1}{\sqrt{2\pi}} e^{-f^2/2} \qquad (3.3.2)$$

by an autoregressive model of orders 6 and 10 are less than 2.7×10^{-3} and 2.3×10^{-4}, respectively [89]. Thus, the application of the MEM to the spectral analysis of clutter data is theoreticaly justified. The experimental analysis given in Chapters 4 and 5 confirmes the practical usefullness of the method. Indeed, virtually every stohastic process encountered in practice (after appropriate preprocessing) can be approximated to any desired accuracy by a finite autoregressive model of a sufficiently high degree [90,p.327]. Since the clutter spectrum is reasonably smooth, the degree need not be too large.

3.4 Summary

In this chapter we discussed the applicability of the MEM to spectral analysis of clutter. For the analysis to be justified, the clutter has to be represented by an all-pole model. The order of the model have been chosen to be two and four with both signals having relatively narrow spectral peaks. The obtained results have shown that the positions of spectral peaks in the estimates become closer to the true peak locations, as number of filter coefficients, M, increase. However, as M becomes a more significant fraction of the number of data samples, N, statistical instabilities occur; the problem well recognized in spectral analysis [1]. There exists some optimal value, M_{opt}, for which the best compromise between resolution, stability, and computation time is achieved. We have used the minimum FPE criterion to determine M_{opt}, since it seems to be the best criterion for the type of signals we are interested in. Resolution capabilities of the MEM, as a function of M, have been considered on the example of fourth order all-pole process which contained two, very close, spectral peaks.

Consideration of clutter models has led us to choose a model developed by Hawkes and Haykin [87] since it is general and yet computationally fairly simple.

Finally, we discussed the adequacy of the MEM in clutter spectral analysis, which is equivalent to the problem of representing the clutter signal by a finite order AR model. We have demonstrated that, in virtually all cases, such a representation is valid. The reasons are summarized below:

(1) The spectral densities of clutter signal given in [25] and [26,p.9] are given in terms of rational all-pole functions of order three and four, respectively. These models allow direct applications of the MEM. The suggested orders of a process are rather low, indicating a computationally efficient procedure.

(2) Non-rational spectra suggested in [20],[27] and some other authors are all based on Gaussian shaped function which, obviously, do not have finite order expansion. However, a reasonably accurate approximation can be obtained by an all-pole representation with a relatively low order.

(3) Theoretically, virtually every random process can be represented by an all-pole process [67],[90]. The efficiency of the MEM as applied to spectral estimation depends on how high is the order with which the particular process can be accurately described. The analysis performed in this work as well as the results of previous work indicate that all-pole model fitting to various clutter signals is perfectly feasible.

CHAPTER 4

MAXIMUM ENTROPY ESTIMATION OF COMPUTER-SIMULATED CLUTTER MODEL AND COMPARISON WITH WELCH'S METHOD

We present here the spectral analysis of computer-generated clutter based on the model described in the previous chapter. This analysis is carried out using the MEM and Welch's method.

4.1 Short-Term Spectral Estimation of Computer-Generated Clutter

Spectral analysis of clutter returns in a scanning radar may be helpful in forming a "clutter map" of the surrounding environment. For example, the spectral characteristics of the clutter can be used as a means of controlling the coefficients of an adjustable filter in an adaptive moving target indicator (MTI) filtering scheme [91],[92]. The speed of adaptation depends on how fast the power spectral density can be computed from the observed clutter data. For a typical Air Traffic Control (ATC) radar, the number of data samples collected from a range-azimuth cell in one scan is between 10 and 20. To get a reliable estimate of the spectral density of the clutter in this cell with conventional methods, we need to use data gathered from several scans to perform the analysis. In this case, the measured spectral density represents the time average over those scans. Adjustment of the adaptive filter coefficients takes place after this analysis is completed.

As an example, assume that we need 500 data samples from a particular range-azimuth cell to perform the analysis by using an FFT-based method. For a typical rotating antenna with 10 revolutions per minute (RPM) and 20 hits per beamwidth, we need data collected from 25 scans, corresponding to 2.5 minutes. Thus, the obtained spectral density estimates in every scan represent average spectra for the previous 25 revolutions. They give reliable representation of clutter signal if the changes in the statistical properties of the surrounding environment are not too fast.

If, however, the statistical properties of the environment change continuously or somewhat rapidly, because of prevailing stormy weather conditions or bird migration, it may be necessary to use a short-term estimate of the pertinent clutter signal. In this regard, we may apply the MEM, since the estimate obtained in this manner, may be calculated from the data collected from two or three consecutive scans or even from a single scan only. To illustrate this possibility, the analyses were run with 64, 32, and 16 data points. Figure 4.1 - 1 shows the spectral density estimates (normalized with respect to the values at zero frequency) obtained from 64, 32, and 16 data samples. The model generates both in-phase and quadrature components of the video signal for the coherent pulse radar. Thus, the complex form of the MEM algorithm [80] is employed in the analysis. The PEF length is automatically set to the value corresponding to the minimum value of the final prediction error. For simplicity of presentation, we consider only the radial motion of scatterers and assume zero-mean radial velocity with variance σ_d^2.

We observe in Fig. 4.1 - 1 that the spectra exhibit three peaks. Two of these peaks are in the positive frequency domain at the locations

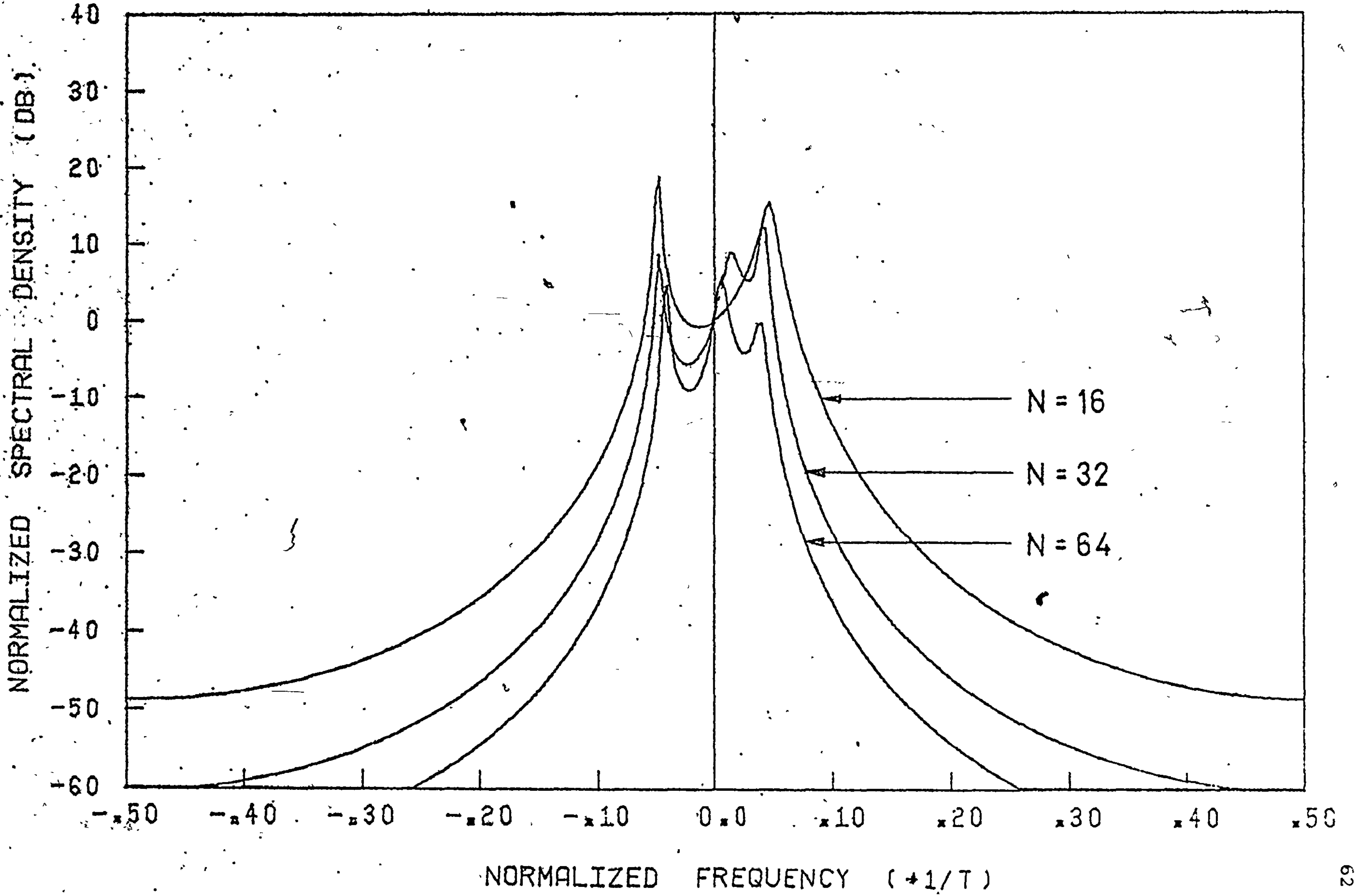


Fig. 4.1 - 1 The ME estimates of a clutter signal using first 16, first 32, and first 64 samples.

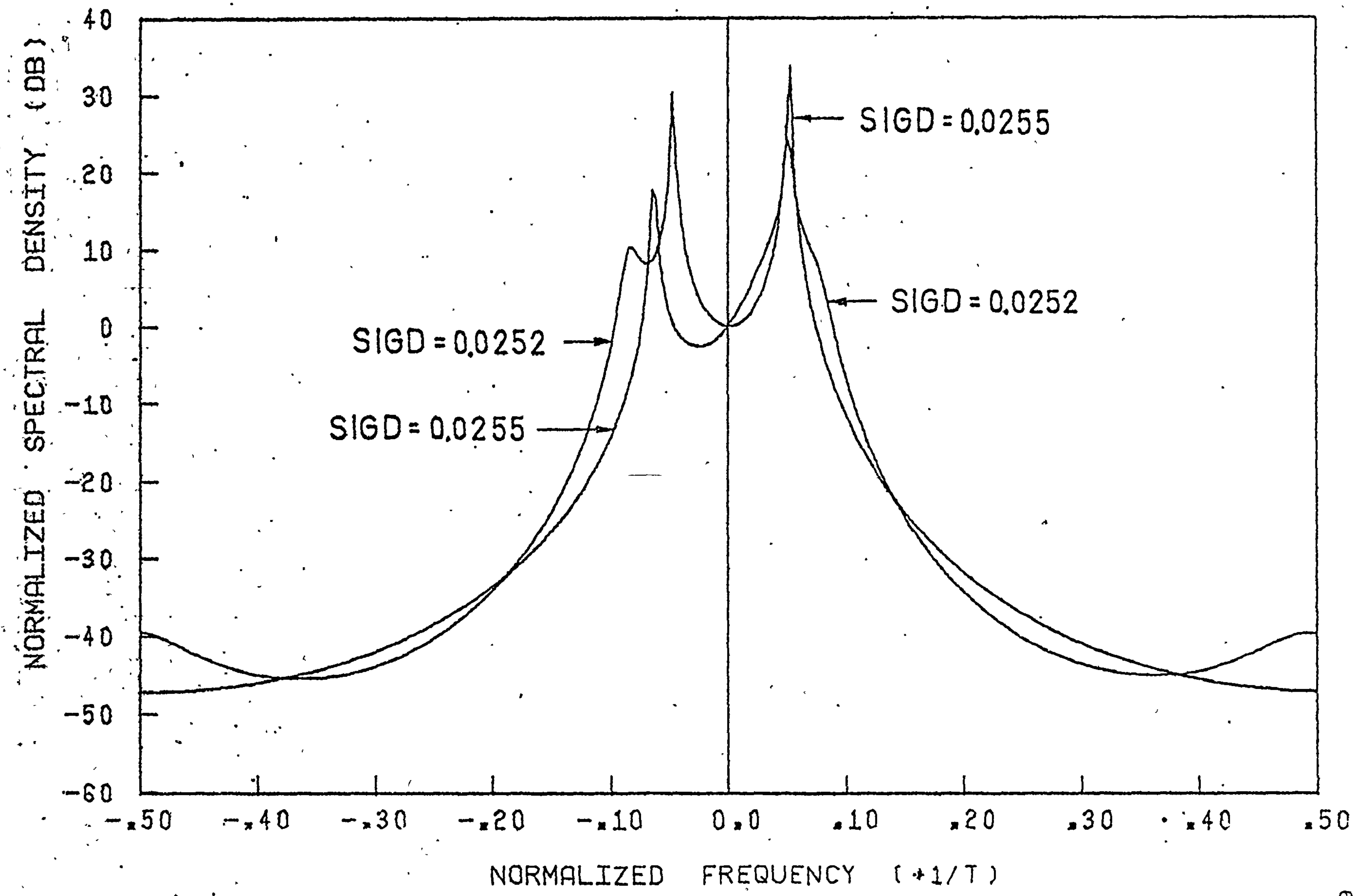


Fig. 4.1 - 2 The ME estimates of clutter signals having two different values of σ_d, with 16 samples.

designated as f_{p1} and f_{p2}. The third one, f_n, is in the negative frequency domain. The locations of these peaks, for different number of data samples used, are summarized in Table 4.1 - 1. Frequencies f_n and f_{p2} are located nearly symmetrically away from zero frequency and the corresponding amplitudes slightly increase with decreasing the number of data samples. Another peak appears at very low frequencies. It is shifted to the right for the case of 32 data points, while for the case of 16 data points it vanishes. This is due to the reduced resolution of the estimate for a very small number of data samples [93].

N \ f	f_n	f_{p1}	f_{p2}
64	-0.040	0.0067	0.0402
32	-0.047	0.0134	0.0410
16	-0.047	-	0.0470

TABLE 4.1 - 1.

Locations of the spectral peaks for different number of data samples.

It should be noted that the clutter data for the cuves in Figure 4.1 - 1 were generated using three independent sets of the random number generator. The assumption about stationarity cannot be reliably confirmed with the number of data samples that are used in the procedure. Nevertheless, the results show that the curves in Fig. 4.1 - 1 do not differ significantly. The shift of the peak location at f_n or f_{p2}, for example, does not exceed 0.8% of the pulse repetition frequency (PRF). If we take a difference between two -10 dB points as a measure of spectral spread, we see from Fig. 4.1 - 1 that the variation of these points is less than 7% of the PRF, when the length of record is reduced from 64 to 16 data samples.

The possibility of resolving clutter signals with slightly different spectral spreads is illustrated in Figure 4.1 - 2. The diagrams represent the ME estiamtes for two different realizations of the clutter signal with the Doppler standard deviation σ_d equal to 0.0252 and 0.0255, with the number of data samples $N = 16$. Both curves are generated with identical output sets of random number generator. For a radar operating at carrier frequency $f_c = 1$ GHz and PRF = 600 Hz, the absolute difference in Doppler spread is

$$\Delta\sigma_{d(abs)} = (\sigma_{d2} - \sigma_{d1})\cdot \text{PRF} = 0.18 \text{ m/sec}, \tag{4.1.1}$$

which corresponds to a spectral spread of

$$\Delta\sigma_f = 2\cdot\frac{f_c}{c}\cdot\Delta\sigma_{d(abs)} = 1.2 \text{ Hz} \tag{4.1.2}$$

where $c = 3\times10^8$ meters per second is the velocity of light. Although this difference is small, and the record is only 16 data samples long, nevertheless, the difference is detectable in the ME spectral density estimates.

4.2 Comparison of the MEM and Welch's Method

In this section we present the results of a comparative analysis of the MEM and Welch's method for power spectrum estimation. The analysis is based on the computer model developed in [87] which was described in Chapter 3. To simplify the analysis, we vary only one model parameter, σ_d, while the others are set equal to zero (the influence of μ_d, μ_r, and σ_r on the shape of the spectral density is shown in detail in [26]). With this simplification we are able to examine the influence of other parameters intrinsic in the problem of spectral estimation using a finite number of data samples.

As shown in Chapter 2 the prediction error filter length M is not too critical when the underlying spectra are reasonably smooth, which is the case in almost every clutter situation. This is, of course, valid if M lies in the vicinity of the value M_{opt} for which one of the three previously mentioned criteria is minimum. We have chosen the final prediction error (FPE) criterion, so that all the results obtained correspond to the value of M_{opt} for which FPE is minimum. This value is automatically determined by the computer program which compares, at each recursion, a new value for FPE with the previous one, and terminates the recursion when the condition

$$FPE(M+1) > FPE(M) \qquad (4.2.1)$$

occurs. This procedure is experimentally justified by checking the monotonicity of FPE(M). In all the cases examined here, FPE has a single minimum when plotted as a function of M. We have checked all the cases by by computing FPE for the values of M which are at least three times greater than M_{opt}. All the diagrams exhibit a broad minimum at M_{opt}.

For the purpose of comparison, Welch's method has been chosen as representative of the linear spectral estimation procedures. The reasons for such a choice are given in Section 1.4 along with the description of the method itself. The window functions have been chosen in accordance with the requirements for resolution and statistical stability. The Bartlett window has the narrowest main lobe (apart from the rectangular window which has large sidelobes), the Parzen window exhibits the lowest sidelobe level, while the Hamming window is a good compromise between the Bartlett and Parzen windows. The analytical expressions for lag windows are given below.

Bartlett window:

$$w_B(m) = 1 - \left| \frac{m-M'}{M''} \right| , \qquad 0 < m < M-1 , \tag{4.2.2}$$

Hamming widow:

$$w_H(m) = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi(m+1/2)}{M} \right) , \qquad 0 \le m < M-1 , \tag{4.2.3}$$

and Parzen window:

$$w_P(m) = \begin{cases} 1-6\left(\frac{m-M'}{M''} \right)^2 + 6 \left(\frac{|m-M'|}{M''} \right)^3 , & \frac{M'}{2} < m < \frac{3}{2} M' \\ 2\left(1- \frac{|m-M'|}{M''} \right)^3 , & 0 < m < \frac{M'}{2} \text{ and } \frac{3}{2} M' < m < M-1. \end{cases} \tag{4.2.4}$$

where $M' = (M-1)/2$ and $M'' = (M+1)/2$.

In contrast with the analysis performed in the previous section, the minimum number of data samples was chosen to be 32, rather than 16, because the spectral densities by Welch's procedure were too broad for a small number of samples. The maximum number was chosen to be 256, large enough to give a qualitative indication of a bound where both methods give approximately tha same results. When Welch's method was applied all the data were divided, for convenience, into 4 nonoverlapping segments.

Figure 4.2 - 1 shows the spectral density estimates obtained with the MEM and Welch's method (based on the three window functions) using 256 data samples. Three spectral peaks can be observed with both methods. Specifically, we have a peak at zero frequency and two at nearly symmetrically located frequencies. The appearance of the latter two peaks is atributed to the fact that a limited number of data samples were taken and that the spectra were not averaged. There are also assymetries in the computed spectra. This is so, because the number of scatterers moving in any one direction is not balanced exactly by an equal number moving in the opposite direction. However, these assymetries are not excessive. In all cases where the methods were compared, the generated clutter samples were identical [94],[95].

With the exception of the Parzen window applied to the data segments in Welch's procedure, all three peaks are easily resolvable with both methods. Also, the spectral spread of the main lobe is greater in the case of the Parzen window; on the other hand, the power density level at higher frequencies is about 15 dB below that of the estimate with the Hamming window and roughly 5 dB below the Bartlett window case. Note

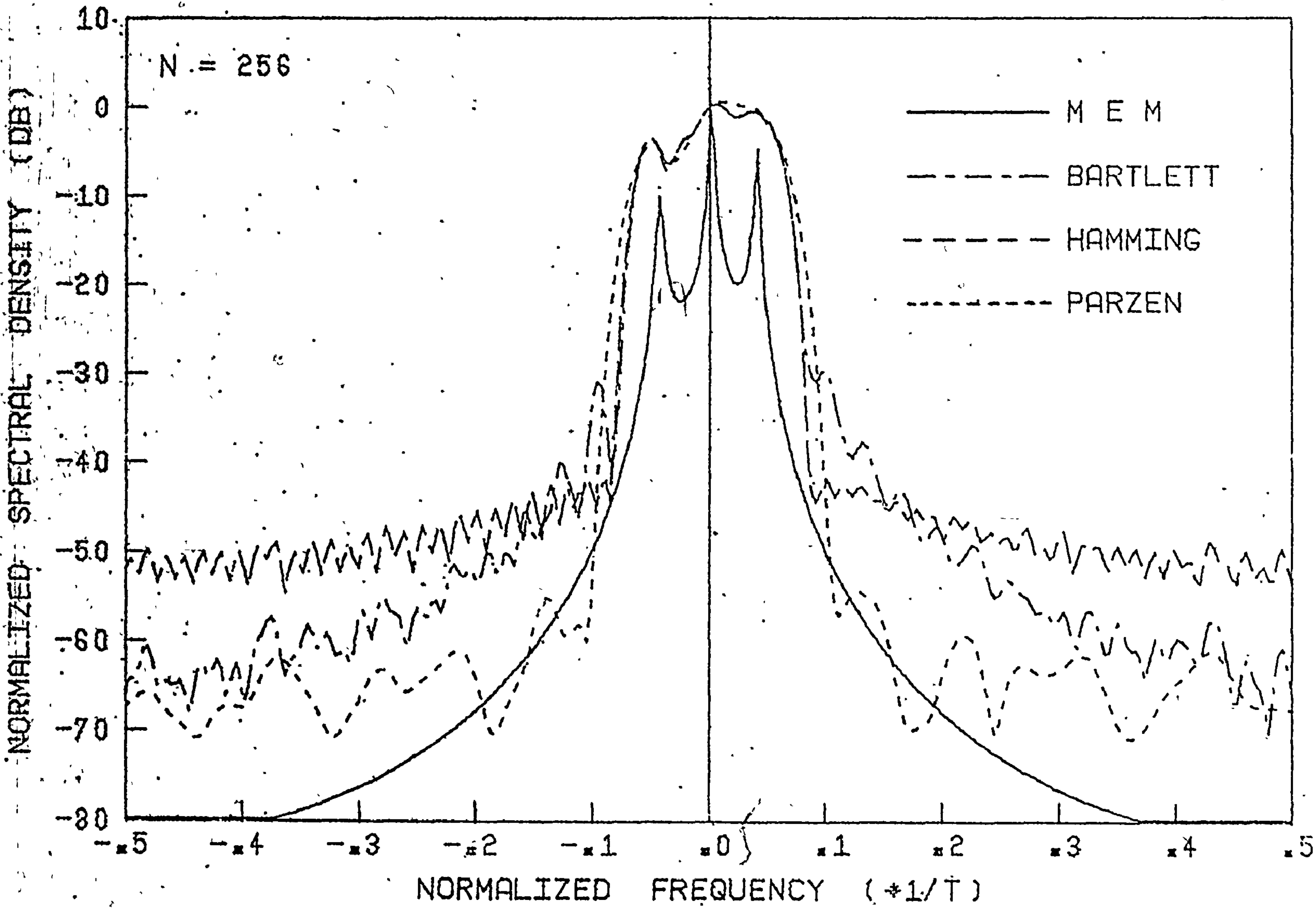


Fig. 4.2 − 1 The power spectral density estimates of 256 data points using the MEM and Welch's method (based on the Bartlett, Hamming, and Parzen windows).

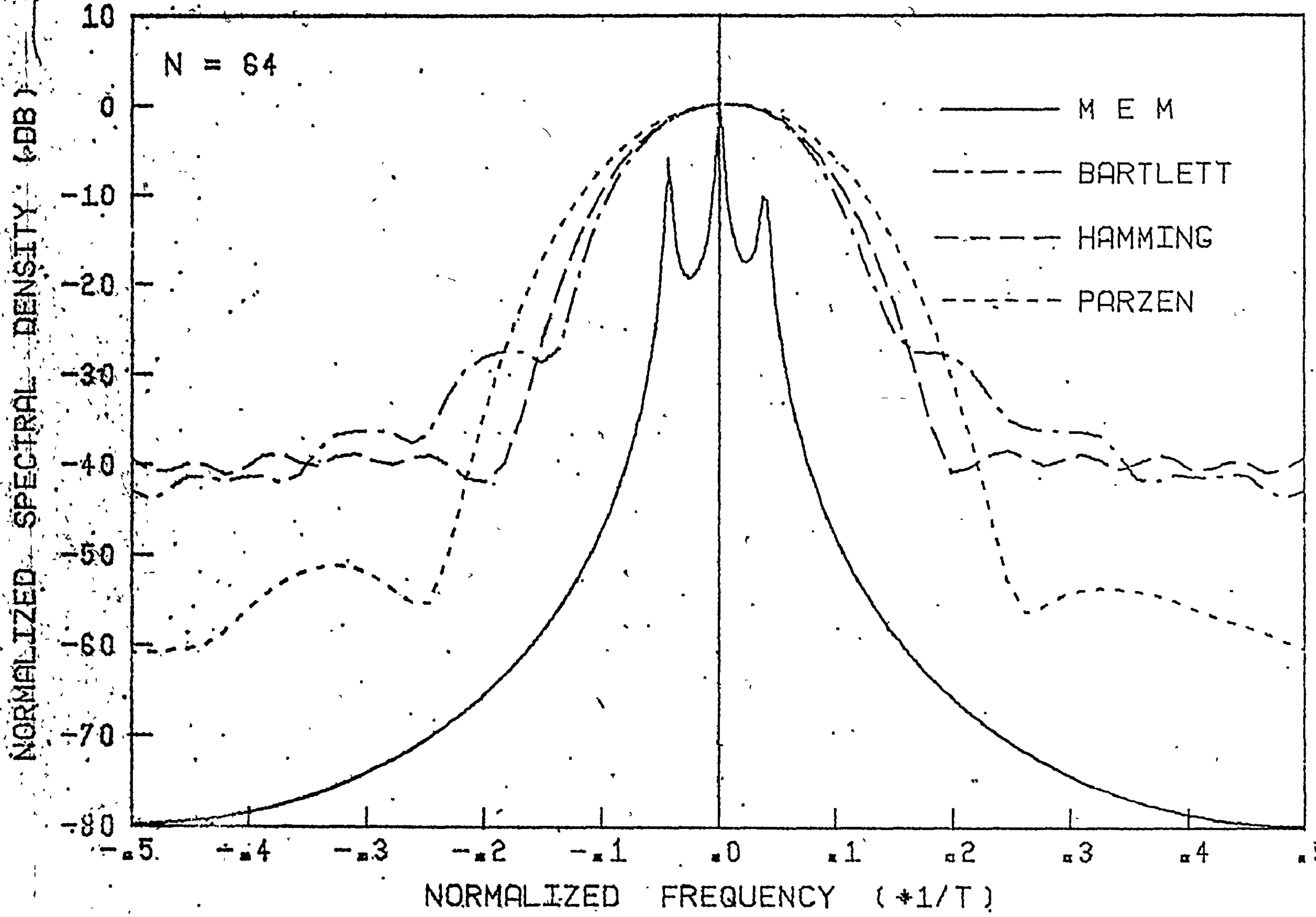


Fig. 4.2 - 2 The power spectral density estimates of 64 data points using the MEM and Welch's method (based on the Bartlett, Hamming, and Parzen windows).

that the "sidelobe structure" inherent in all lag-winow methods does not exist in the maximum entropy spectrum.

If we reduce the number of data samples to the first 64, keeping all other parameters unchanged, we obtain the results shown in Figure 4.2 - 2. While the maximum entropy estimate still reveals three peaks with slight changes in their amplitudes, the resolution capability of Welch's method has completely vanished. We also observe that both the spectral spread and the sidelobe levels have increased considerably in the case of Welch's method, while the corresponding increase in the spectral spread in the case of MEM is only about one per cent at the -50 dB points.

The effect of varying the number of data samples on the spectral density estimate is illustrated in Fig. 4.2 - 3 for the MEM and in Fig. 4.2 - 4 for Welch's method. The Parzen window was chosen in the latter case because it has the lowest spectral sidelobes amongst the three windows used in the analysis. We see that by halving the record length, the spectral spread obtained by Welch's method is approximately doubled, while the MEM results in only a relatively small change in the spectral spread. In Figures 4.2 - 3 and 4.2 - 4 we have also included the theoretical spectral density obtained by taking the discrete Fourier transform of the Eq. (3.2.5), using 64 samples padded with an equal number of zeros.

The resolution capabilities of the two methods are compared in Figures 4.2 - 5 and 4.2 - 6. The curves in Fig. 4.2 - 5 show Welch's spectral density estimates of the clutter data consisting of 256 samples, with three different values for the parameter σ_d, which differ by no more than 3%. The selected values for normalized σ, as well as their absolute

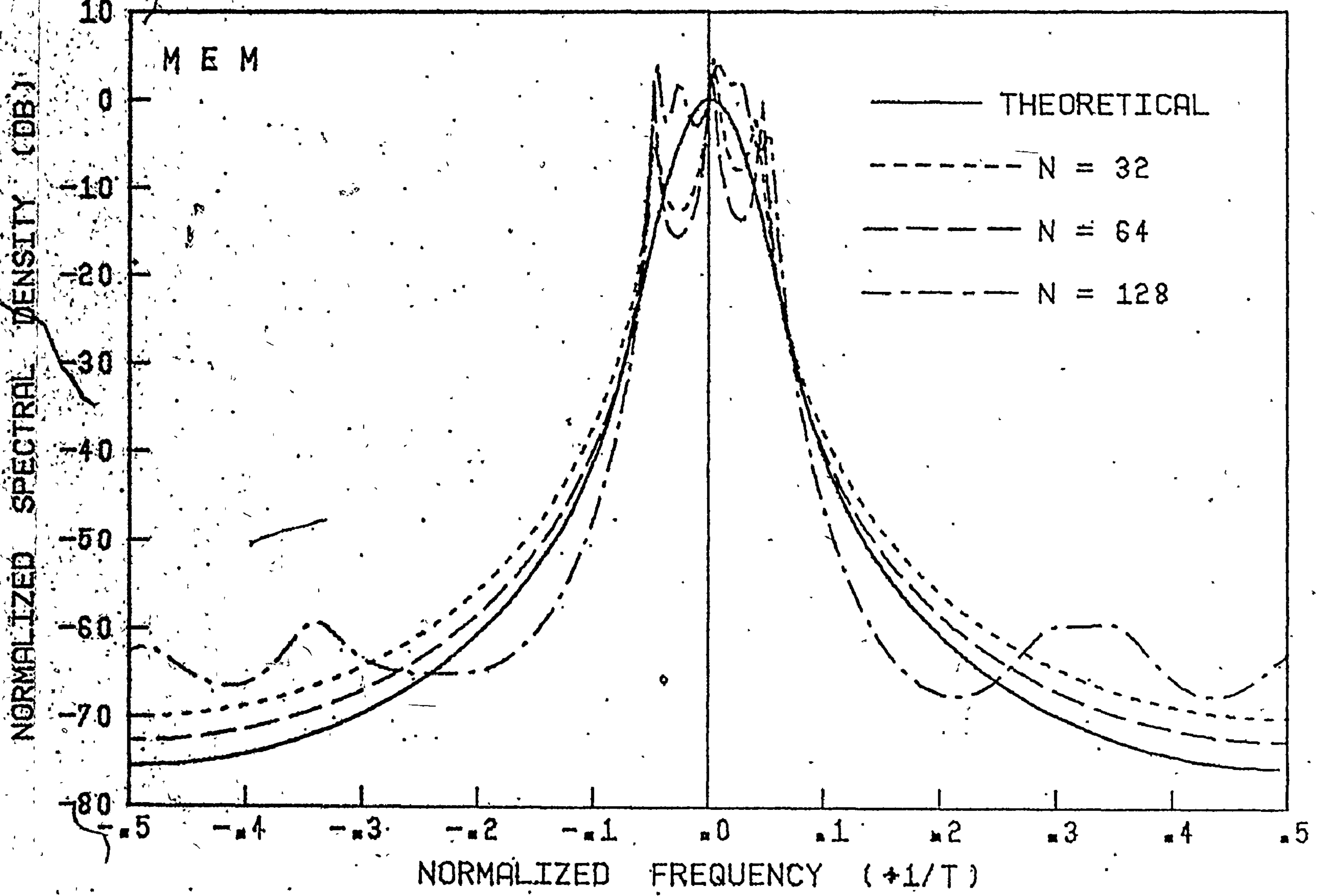


Fig. 4.2 - 3 The ME estimates of a clutter signal using the first 32, first 64, and first 128 samples. The theoretical spectrum is obtained by Fourier transformation of Eq.(3.2.5).

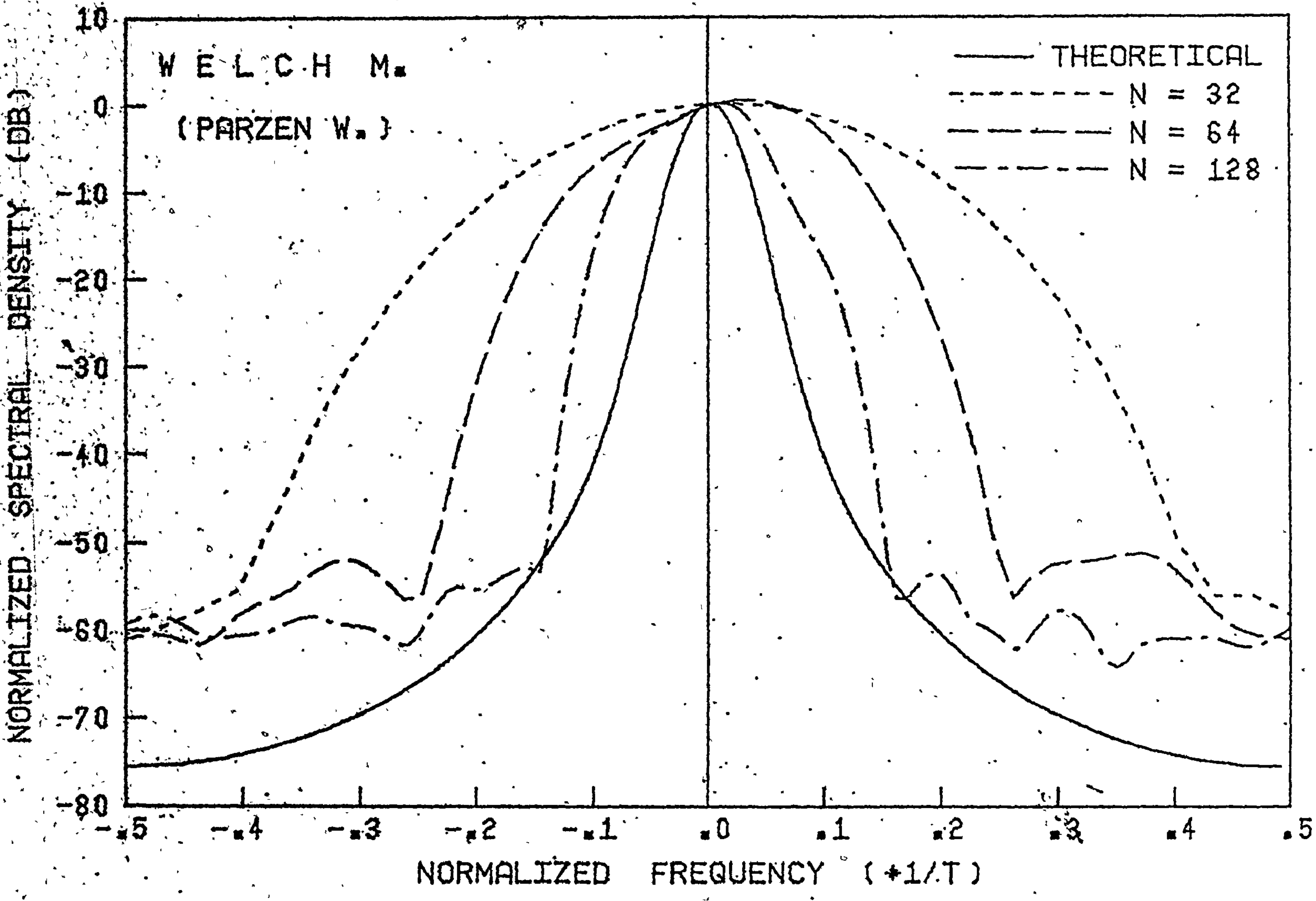


Fig. 4.2 - 4 Estimates obtained by using Welch's method (based on the Parzen window) applied to the first 32, first 64, and first 128 samples of the clutter signal. The theoretical spectrum is obtained by Fourier transformation of Eq.(3.2.5).

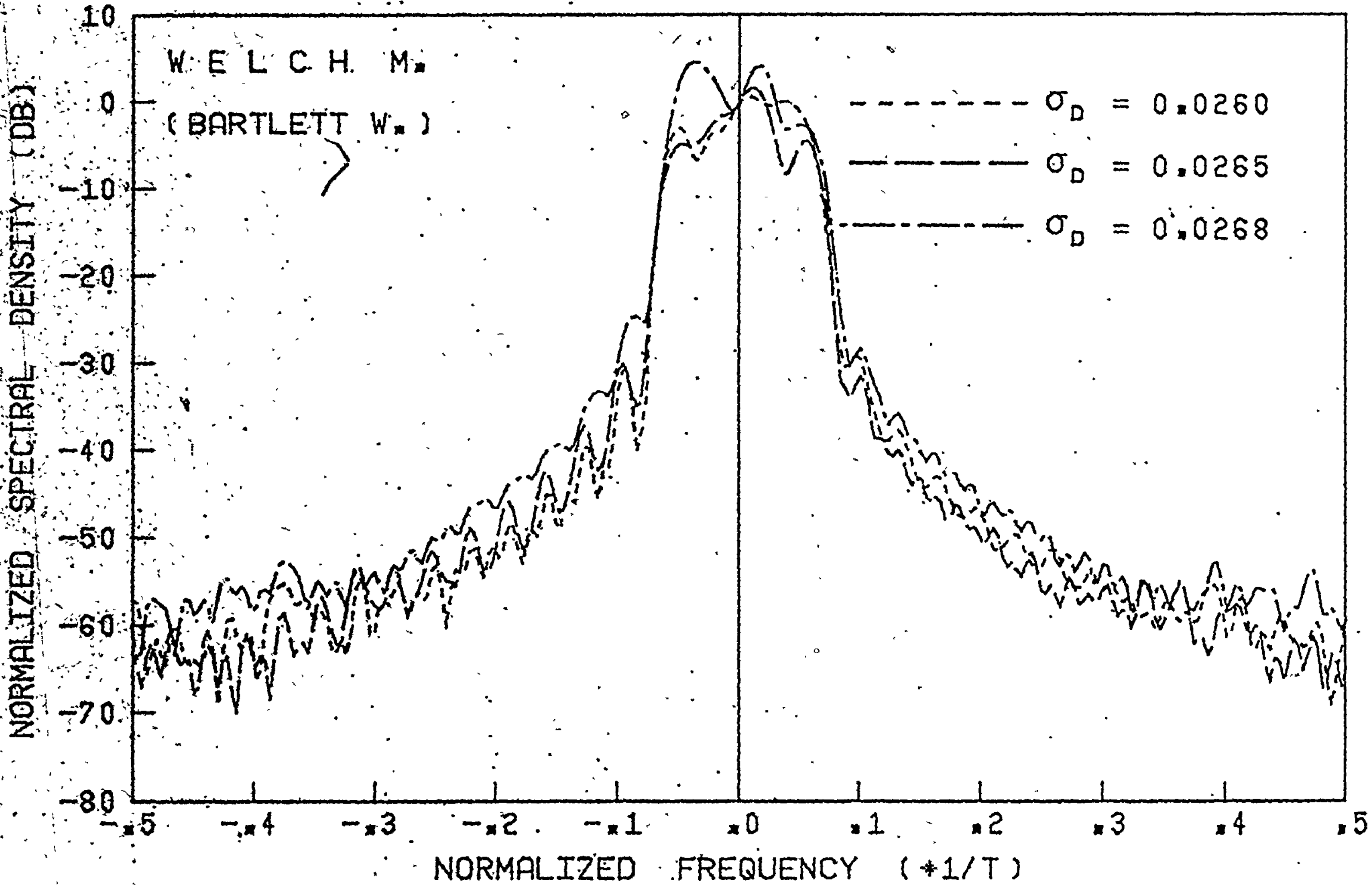


Fig. 4.2 - 5 Estimates obtained by using Welch's method (based on Bartlett window) applied to clutter signals with three different values of σ_d, ~~with~~ 256 data samples.

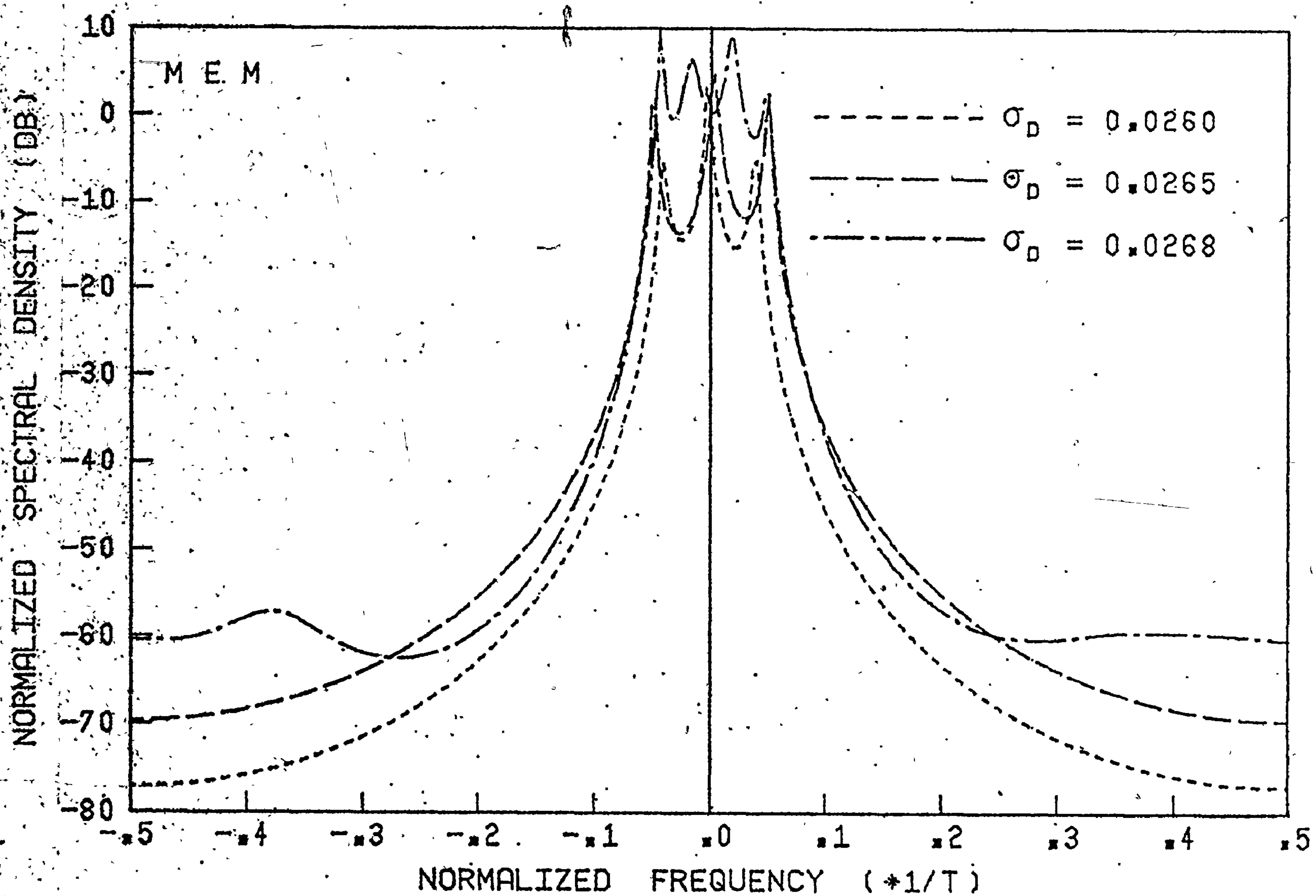


Fig. 4.2 - 6 The ME estimates of clutter signals having three different values of σ_d, with 128 data samples.

values corresponding to PRF = 600 Hz are summarized in Table 4.2 - 1. The Bartlett window was chosen in this case, since it has the narrowest main lobe compared to the Hamming and Parzen windows, and thus the best resolution. We see from Fig. 4.2 - 5 that the effect of this small change in σ_d on the spectral density characteristics cannot be detected by applying Welch's method, in spite of using a relatively long data record (i.e., 256 data samples). On the other hand, by applying the MEM as in Fig.4.2 - 6, the different values of σ_d are readily resolved, even though the analysis is based on a reduced record length, namely 128 data samples only.

σ_d (normalized)	0.0260	0.0265	0.0268
$\sigma_d \cdot$PRF (Hz)	15.6	15.9	16.08

TABLE 4.2 - 1.

Normalized (to PRF = 600 Hz) and absolute values of σ_d, selected for diagrams in Figs. 4.2 - 5 and 4.2 - 6.

4.3 Computation Time

The number of operations required to perform a spectral analysis with both methods is given below. By operation we mean a complex multiplication followed by a complex addition.

To compute a finite Fourier transform of a sequence of length L by means of the FFT algorithm we require a number of operation proportional to $L \cdot \log_2 L$. Suppose, we have N_W data samples and wish to divide them into K segments of length L overlapped by an amount L/2 to perform Welch's method. Then we need approximately

$$K \cdot L \cdot \log_2 L = 2N_W \cdot \log_2 L \tag{4.3.1}$$

operations for performing the Fourier transformation. If, for example, there is no overlapping between segments, the number of operations is halved, since $K = N_W/L$ in that case. Additional computations for calculating a window function, multiplying the data by a window, and averaging, require an order of LK operations. Thus, in the most favourable case with no overlapping, we need roughly $N_W \cdot \log_2 L$ operations for performing Welch's method.

The algorithm for calculation of the PEF coefficients in the MEM requires $2(N_E - m)$ operations in the mth recursion step to calculate the quantities $p_m(n)$ and $q_m(n)$ by using Eqs. (2.5.5) and (2.5.6), where N_E is the data sample length used in the analysis. Four additional operations are required to calculate $a_{m,m}$. For a PEF of order M we thus need

$$M(4 + 2[N_E - (M+1)]) = M(2N_E - M) \tag{4.3.2}$$

operations. Next, computation of each spectral point requires about M operations, so that the calculation of the spectral density estimate requires about $M \cdot N_f$ operations. The number of points on the frequency axis, N_f, is not related to either M or N_E and is chosen in such a way as to satisfy the resolvability requirements. We thus need approximately

$$M(2N_E + N_f - M) \approx M(2N_E + N_f) \tag{4.3.3}$$

operations for the maximum entropy spectral calculation.

To get an idea of the order of magnitude of the number of operations in a typical example, assume a signal sample of length $N = 2^{10} = 1024$ and divide it into $K = 4$ nonoverlapping segments of $L = 256$ samples. The number of operations in Welch's method is approximately 8000. For MEM, assuming the same number of samples and $M = 10$, the total number of computational operations, required for, say, 256 spectral points, is approximately 20,000. We see that the number of operations in this particular example is greater for MEM than for Welch's method by a factor of 2.5. However, as mentioned previously, for a given number of samples, the spectral resolution achieved by using the MEM is superior to that achieved by using Welch's method. Also, for the spectra usually encountered in clutter analysis, the minimum FPE criterion gives a relatively small value for the optimal order of the PEF.

4.4 Summary

To test the applicability of the MEM to clutter spectra analysis, we have made use of computer-simulated clutter. This approach is convenient, because we can predetermine the values of the parameters of the signal and vary them in a way that we can quantitatively determine the corresponding variations in the estimates. The results show that reliable spectra can be obtained with a small number of signal samples. Even with such a small number we are able to distinguish between the spectral estimates of two signals with slightly different predetermined values for the Doppler standard deviation.

In the second part of the Chapter we performed a comparative analysis of two spectral estimation methods: linear method based on modified periodograms which was proposed by Welch, and nonlinear, MEM, proposed by Burg. The analysis was based on a realistic, computationaly tractable computer model of clutter signal. From the results obtained we can draw the following conclusions:

(1) For a short record length, the MEM is superior to Welch's method in both resolution and stability.

(2) When the record length is sufficiently long both methods give approximately the same results.

(3) As indicated in Fig. 4.2 - 4, the MEM seems very promising in distinguishing between slight changes in spectral parameters. This capability may be very usefull in making decisions about the source of the clutter.

(4) For reasonably smooth spectra, the computation times are of the same order of magnitude for both methods. For high values of the optimum PEF lenghts the MEM requires more calculations, while Welch's method is independent of the type of signal being analyzed.

CHAPTER 5

ANALYSIS AND CLASSIFICATION OF VIDEO RECORDED CLUTTER

In this chapter we present the results of spectral analysis performed on actual clutter data obtained by recording the video signal of an air traffic control radar under different environmental conditions. We begin the chapter by giving a brief description of the radar and recording facilities used in the experiment.

5.1 Radar Characteristics, Video Recording, and Interface

Video tape recordings of four radars have been made: an AASR-1 (Airport and Airways Surveillance Radar) in Ottawa, an ASR-803 in Vancouver (copied), an AASR-1 in Toronto, and an ASR-8 in Bagotville. The first three were manufactured by Raytheon, and the last one by Texas Instruments. The ASR-8 was chosen for the spectral analysis since only this radar was coherent and provided I and Q bipolar signals, as well as unipolar noncoherent signals [96].

Table D - 1 in Appendix D lists the characteristics of the ASR-8 radar. The received RF signal is heterodyned against the STALO to produce a 30 MHz IF signal. The IF signal is envelope-detected by means of crystal diode to produce a normal video. The IF signal is also mixed with the 30 MHz output of COHO, as well as with a 90° shifted output of COHO to produce the I and Q components, which are then phase-detected. Both components are converted into digital form and passed through a digital dual-MTI canceller system. The outputs of the phase detectors were recorded simultaneously on a video tape.

The antenna system consists of a normal (low) beam and a passive (high) beam for avoiding ground clutter. The calculated free-space elevation pattern is shown in Fig. D - 1, while Fig. D - 2 shows the azimuth pattern. The traces under the main lobe are cross-polarized measurements, i.e. responses of vertically polarized receiving pattern to harizontally polarized transmitting one, or vice versa [96].

The radar video signals analyzed in this research were recorded using an RCA Airborne Wideband Video Recorder/Reproducer, Adviser, herein referred to as the video recorder. The following is a summary of its basic characteristics. A detailed description of operation of the recorder is given in [97].

The recorder records the video signals onto two Wideband Video Channels, and the low frequency signals onto two audio channels called the auxiliary channels. The frequency range for the video channels extends from 15 Hz to 5 MHz, with 3 dB attenuation at 10 Hz and 6 MHz. The upper cutoff frequency for the auxiliary channels is 15 kHz, while the lower cutoff frequency is 100 Hz for the Aux 1 and 350 Hz for Aux 2. Two types of video signals can be recorded simultaneously on the two video channels A and B. For the analysis of coherently detected video signals we used the recordings of I and Q components of video signals, which are available at the outputs of the phase detectors.

In addition to the video signals, timing signals are required to allow the determination of range and azimuth of a particular video segment. Range of a target is measured as the time delay between transmission of a pulse and reception of the echo. Hence, the time of pulse transmission must be known. This timing is given by the system trigger. Azimuthal information is given by the position of the antenna. This position is given by the synchro signals S1, S2, S3 which are sinusoidal waves, of frequency equal to the antenna scan rate, 120° apart in phase, which modulate a 60 Hz carrier. These three signals can be resolved to provide the azimuth of the antenna. The azimuth digitizer is driven by these synchro signals to obtain a series of azimuth change pulses (ACP), usually 4096 per revolution, as well as an azimuth reference pulse (ARP), usually at north reference.

Therefore there are just three signals needed for positional information: system triggers, ACP's, and ARP's. It was decided to record the ACP's directly on Aux 1, and the ARP's directly onto Aux 2. The pulses are rounded because of the limited bandwidth, but remain sufficiently well defined to allow their recovery upon replay. Since the system trigger is intimately tied to the video, and the video at time $t = 0$ (at pulse transmission) is of no use, it was decided to mix the system trigger with the video (I component) prior to recording. It was mixed in as a negative-going pulse with the amplitude sufficient for distinguishing it from the video.

The interface between the recorder and computer is provided by Radar Video Display and Sampler/Digitizer (RVDSD) [98].The RVDSD is logically divided into two parts: display circuitry and sampling/interface circuitry. The basic purpose of the display circuitry is to take the video and timing signals from the recorded video tape and configure them to produce a pseudo PPI display suitable for use on any oscilloscope with

X, Y, Z input capabilities. The sampling/interface circuitry has two basic purposes. The first is to sample at a proper rate and analog-to-digital convert the selected video within the sampling area. The samples are stored in random access memory (RAM). The second purpose is to interface the RVDSD to a CDC 1700 computer to allow the computer to access the data in the RAM and to control the other registers, such as those defining the sampling area, if desired. The following is the brief description of the system parameters.

As mentioned above, the display circuitry produces a PPI display on a scope with X, Y, and Z capabilities. The display is circular and shows range increasing from zero at the centre to a maximum at the perimeter and azimuth by the position of the range sweep. Echo intensity is shown by modulating the Z-axis with the corresponding video; the greater the echo intensity, the brighter the display. With the storage oscilloscope one can get a visual representation of the surrounding area. The sampling/interface circuitry samples the video within the predetermined sampling area. The number of resolution cells per sweep (range cells) is typically 2500. With 4096 azimuth positions per scan, there are approximately $2500 \cdot 4096 \approx 10$ million samples per scan. In order to accomodate the slow speed of the digital computer relative to the recorder, only a small portion of the PPI display can be processed at any one time. In any case, a large portion of PPI display contains no apparent echoes, and often the target or clutter of interest only extends over a small area. Therefore, this restriction to a relatively small area is of no major consequence. If a larger area is required, several sampling passes through the same section of tape could be made.

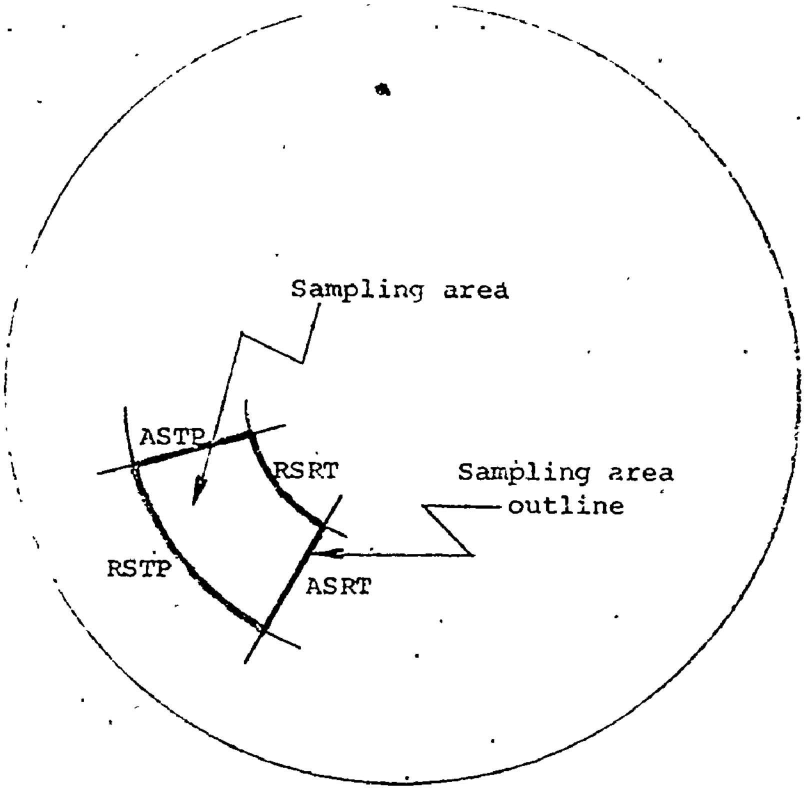


Fig. 5.1 - 1 Sampling area definition.

The area to be sampled is called the sampling area and is defined on the PPI by the values contained in four boundary registers: azimuth start (ASRT), azimuth stop (ASTP), range start (RSRT), and range stop (RSTP), as shown in Figure 5.1 - 1. Each digital sample is addressed by its range and azimuth, i.e. the corresponding values of 12 bit registers, called present range (PR) and present azimuth (PA), respectively. The value of PR register is reset to zero by system trigger since this indicates the start of a new sweep. Then it simply counts clock pulses, since the number of samples per sweep equals the number of resolution cells per sweep. Similarly, the PA register counts ACP's from 0 to 4095 and is reset to zero by ARP's which indicate north reference. The values PR and PA are continually compared to the values of the four boundary registers. A digital signal SA (sampling area) is high when PR and PA are within their respective limits, including that the corresponding point on the PPI is within the sampling area and a sample should be taken. SA is low otherwise. The signal SA is used to gate the sampling clock, hence control the sampling. Also, when mixed with the basic video, it produces an outline of the sampling area on the PPI display, as indicated in Figure 5.1 - 1. Another option causes display of only the video within the sampling area, and masks or grounds the video signal outside the sampling area.

The values for the boundary registers can be loaded either from the front panel switch register or from the computer. The choice is made mannually using the CONTROL switch. This possibility is very convenient since the interesting area may be chosen mannually by looking at the PPI display and, after the choice has been made, control may be transfered to the computer, and the samples from that area collected.

A total of eight registers can be displayed on the front panel. The value of PR is not displayed because it changes so rapidly that the lights would be simply a blur. In addition to the PA and the four boundary registers, three others are displayed. These are present north mark (PNM), north mark stop (NMSTP), and number of samples (NSAMP). PNM counts the number of scans, while NMSTP can be loaded with the number of scans desired. NSAMP counts the number of samples taken per scan. This number must not exceed a limit predetermined by the RAM storage capacity which, in case of sampling I and Q components of a video signal, is 2048 samples.

The maximum number of consecutive scans, which can be processed in one run is 100, but can be easily increased with a slight change in software.

With the presently available circuitry, the total conversion cycle for sample and hold (SH) and A/D permits sampling at a rate of 1 MHz for 6 bit conversion. Since the signal-to-noise ratio of the video recorder is 36 dB, which corresponds to 63.1 or about 64, there is no need for resolution higher than 6 bits.

In order to save storage space, the RVDSD have been designed to store both I and Q component into one computer word. This was accomplished with 16 bits word length, still leaving four bits per computer word unused.

5.2 Data Collection and Analysis

The recorded radar signal is two-dimensional by nature in that it is a function of the range and azimuth coordinates. On the other hand, the processing is done in one dimension. A choice must therefore be made about the number of samples in range and those in azimuth, and the order in which the samples are processed. Suppose that we wish to analyze an area that contains N_S samples so that we have

$$N_R \cdot N_A = N_S = \text{Const.} \tag{5.2.1}$$

where N_R and N_A are the numbers of samples in range and azimuth, respectively. Now, the azimuth bin is defined by 3 dB beamwidth. The number of azimuth change pulses (ACP's) within the specified bin determines the number of hits per beamwidth. For the radar used in the analysis the azimuth 3 dB beamwidth is 1.35° so that the number of hits per beamwidth is

$$N_H = \frac{4096}{360^\circ} \cdot 1.35^\circ = 15.36 \,, \qquad (5.2.2)$$

which, for the purpose of analysis, can be rounded to 16. On the other hand, the range bin is determined by the transmitted pulse width. The procedure used for arrangement of the signal samples is to take 16 samples from the first range-azimuth cell, then, the next 16 samples from the neighbouring cell in azimuth, and so on, up to the last azimuth bin in the same range ring. Then we switch to the next range ring and repeat the same procedure, until we reach the N_R-th ring. For example, if we choose two range rings ($N_R = 2$) then, from Eq. 5.2.1, $N_S = 2N_A$, and the arrangement is as shown in Fig. 5.2 - 1(a).

Instead of going along the azimuth we may, of course, choose the successive samples along range in one azimuth bin, then shift to the next bin, etc. This order of data arrangement is exactly the one which exists in computer memory, since the incoming data appear in that order. The arrangements with 2 and 8 successive range rings are shown in Figures 5.2 - 1(b) and (c) respectively. In order to compare power density spectra for various sample arrangement, each selected area of clutter has been defined in two ways: one corresponding to Fig. 5.2 - 1(b), and another corresponding to Fig. 5.2 - 1(c). In both ways the initial sample (number 1 in both Figures) was the same. In either case computer program provides an option which enables reordering data in a manner shown in Figure 5.2 - 1(a).

We may now briefly outline the steps of the measurement procedure:

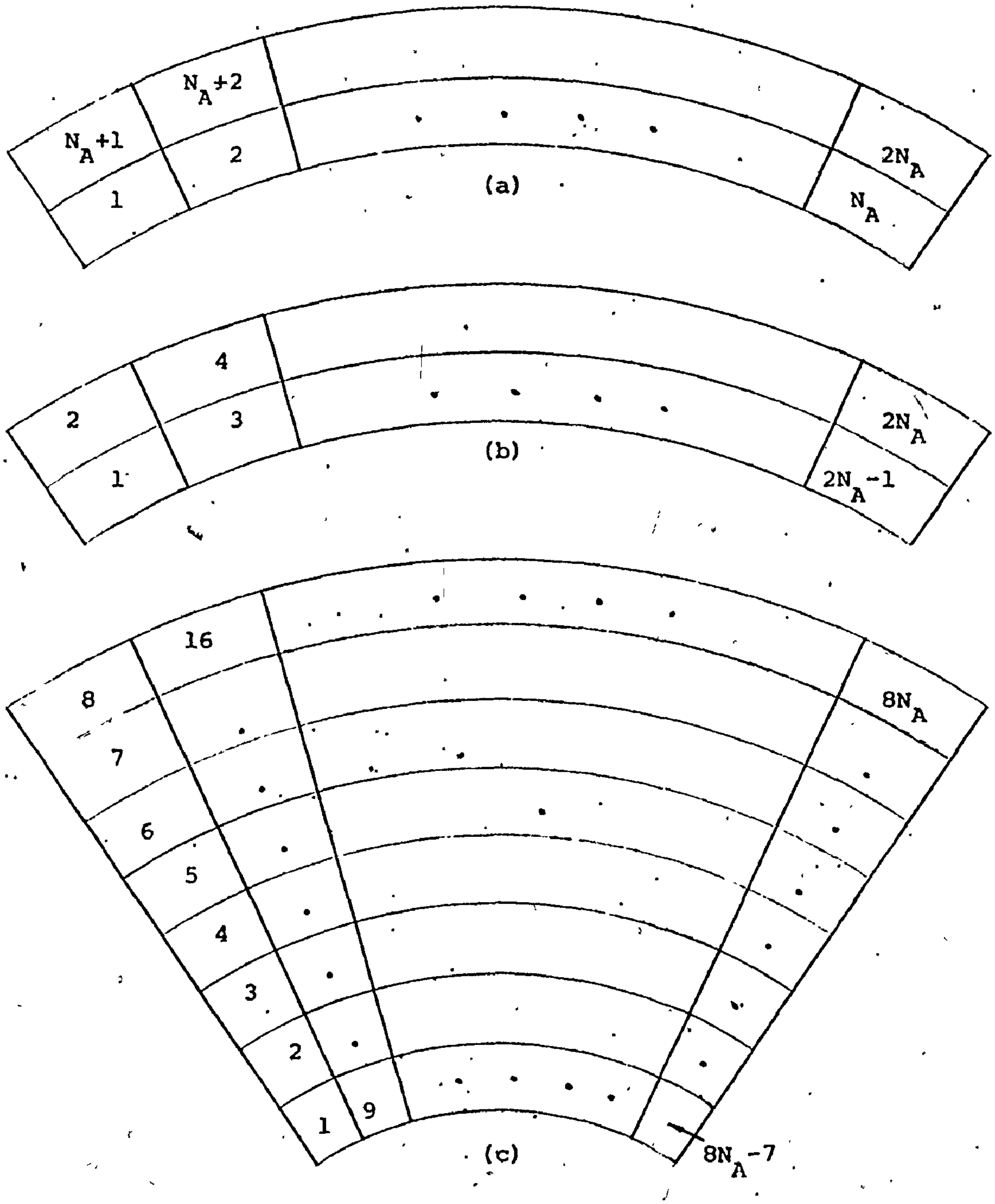


Fig. 5.2 - 1 Different ordering of sample bins for processing:
(a) one range segment at a time;
(b) two range segments in succession;
(c) eight range segments in succession.

(1) With the aid of visual inspection of the PPI display, the desired sampling area is chosen, and the values of the four boundary registers are set mannually.

(2) Control is then transferred to the computer, and the chosen boundary values, as well as the desired number of scans, are entered via RVDSD software into the computer. The starting and final sector addresses on disk are also chosen and the data are written onto the disk.

(3) After the transfer of data is completed, samples are picked up from disk, separated into I and Q samples, arranged in one of the three ways, explained above, and then written onto the pseudotape unit in a format required by the programs which perform spectral analysis. All the operations in this step are performed by program SAMARR and supporting routines.

(4) The program which calculates the power spectral density function is then called up and the results obtained are plotted on a cathode ray tube (CRT), and also printed out. Program MEMFP1 calculates the maximum entropy spectra, while PER64 evaluates the spectra by means of Welch's method.

Documented listings of the computer programs and supporting subroutines are given in Appendix E.

5.3 Analysis of the Reflections from Ground, Storm Clouds and Birds

As mentioned in Chapter 4, the minimum final prediction error (FPE) criterion gives the optimun value of a prediction error filter (PEF) order, in the sense that it represents the best compromise between bias and variance of the spectral estimate. Therefore, this criterion is applied in all the following measurements. It has been shown that the choice of M in the vicinity of M_{opt} yields almost identical results. It turns out that for all measurements the value of M_{opt} is in range of 0.05N to 0.2N.

Selection of ground and weather clutter was based on observations of the PPI display for a relatively long period. These observations were combined with the weather report available at the radar site during the time of recording. Ground clutter patterns are easily recognizable on PPI since they are fixed and exist on each recording. Clouds were prominent in that part of the display where they were reported. Also, their movement was easily recognized in a few scans only. Figure 5.3 - 1(a) shows the PPI display of normal video where the cloud formations between north and east are easily recognizable. Reflections from those clouds break through MTI filter and appear on PPI display of MTI video, as shown in Fig 5.3 - 1(b). Recognition of bird flocks was not possible on the storage oscilloscope PPI display because of its lower sensitivity, smaller display area, and lesser persistence than the original PPI. However, at night when migration was expected, the long exposure photograph (over 15 scans) of an original PPI display was made, which revealed the presence of scattered strikes which had indicated a large number of small point targets in motion. This photograph is shown in Figure 5.3 - 2. With this evidence, one can say with high probability that the bird flocks were displayed on PPI. As will be seen later, this can also be deduced from the spectral analysis results for the echoes reflected from this area.

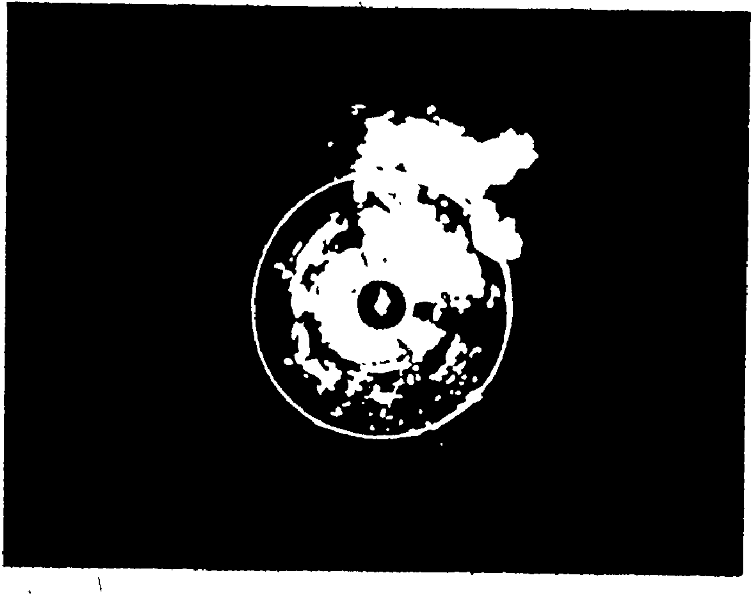


Fig. 5.3 - 1(a) PPI display of the area of stormy clouds, normal video.

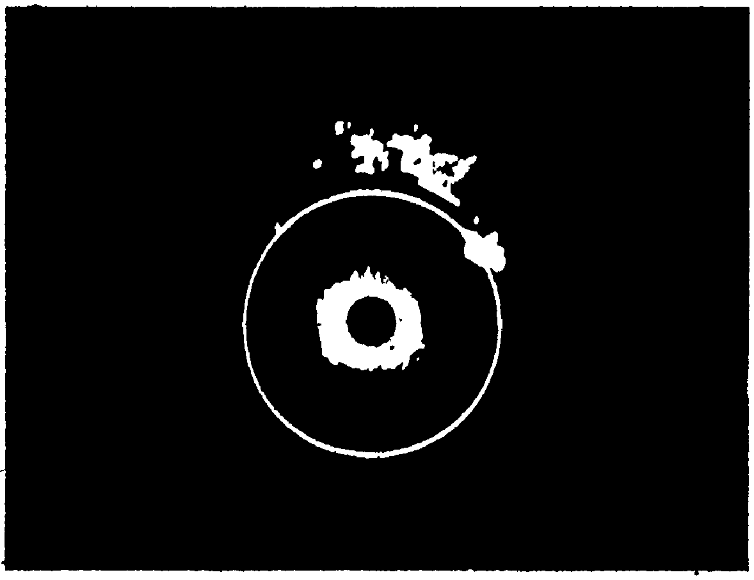


Fig. 5.3 - 1(b) PPI display of the area of stormy clouds, MTI video.

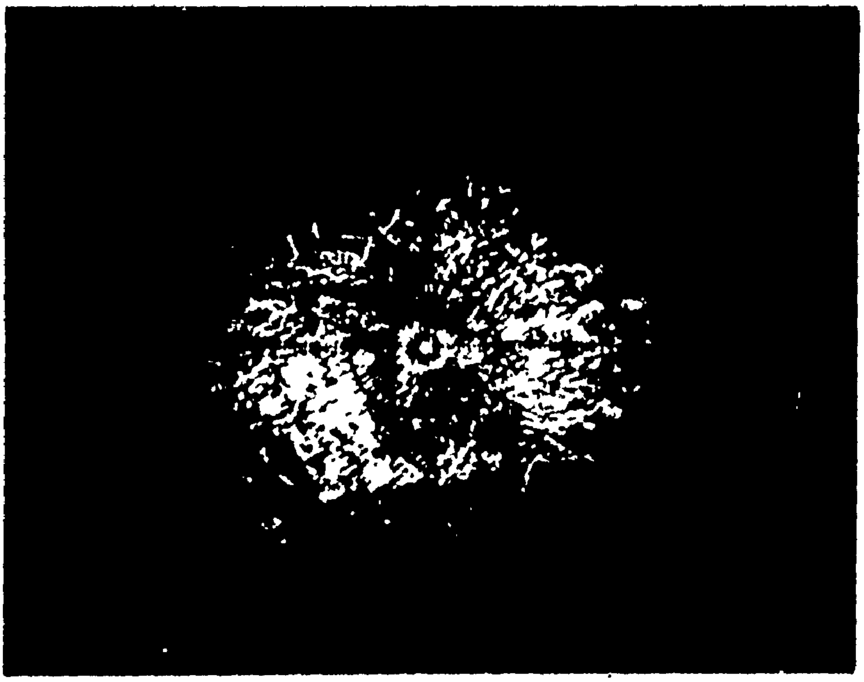


Fig. 5.3 - 2 Original PPI display of birds.

Ground clutter has been analyzed by using two range cells with different numbers of data samples. Figure 5.3 - 3(a) shows the spectral density estimates obtained with MEM using minimum FPE criterion, for different numbers of data samples. Frequency values are normalized with respect to the average PRF which is 1040 Hz, and the diagrams are drawn in the frequency range symmetrical with respect to zero frequency. Spectral density values are also normalized with respect to the maximum value, so that they extend from 0 dB to the limiting value which is selected in advance. These normalizations are done in all diagrams that follow. Figure 5.3 -3(b) shows the parts of the curves in Fig. 5.3 -3(a) for the frequency range -0.25 to 0.25. A feature that is immediately obvious from these diagrams is the presence of a very narrow spectral peak at zero frequency, and very low sidelobes (below -40 dB, except for N = 8). Furthermore, there is not much difference between the spectral curves for N = 256, 100, and 14. It is only when the number of data samples is low (N = 8) that we find the spread of the curve is relatively wide and the sidelobes relatively high. Figures 5.3 - 4(a) and (b) correspond to the set of data collected from the neighbouring sampling area. Fig. 5.3 - 5 depicts the arrangements of these areas. Segments are separated by 96 words because this is the number of words stored on one sector of the disk. We see that the curves corresponding to N = 8, N = 16, and N = 100 are obtained by processing completely independent sets of data (except for N = 100 where only 4 samples overlap), while in the case of 256 signal samples, 160 samples are common to both runs. We see that the curves for N = 256 in Figs. 5.3 - 3(a) and 5.3 - 4(a) (or equivalently, in Figs. 5.3 - 3(b) and 5.3 - 4(b)) are similar. Although somewhat less pronounced,

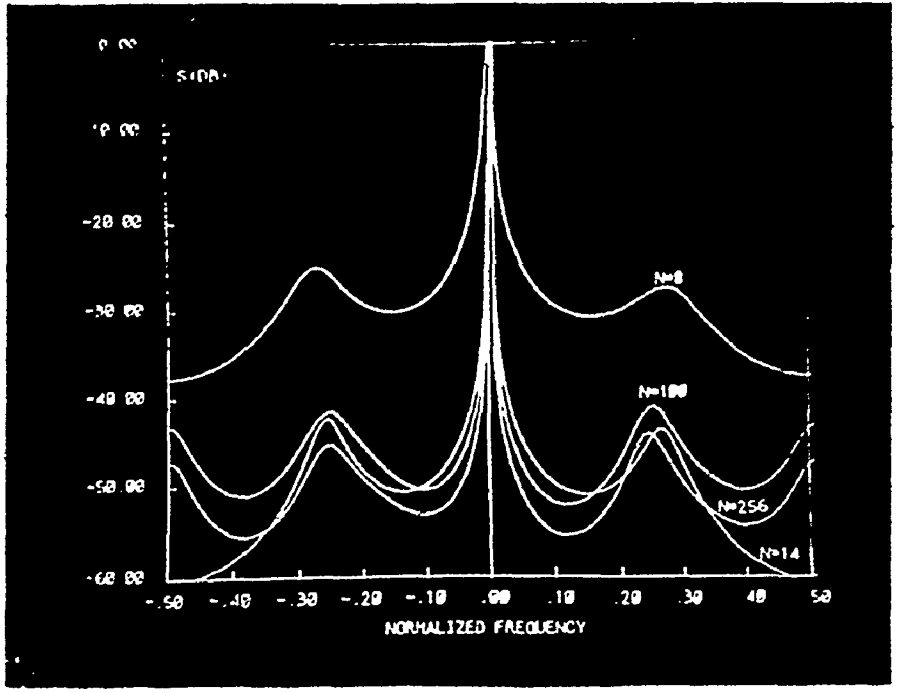


Fig. 5.3 - 3a) ME spectral estimates of ground clutter, with N = 256, 100, 14, and 8; frequency range from -0.5 to 0.5 .

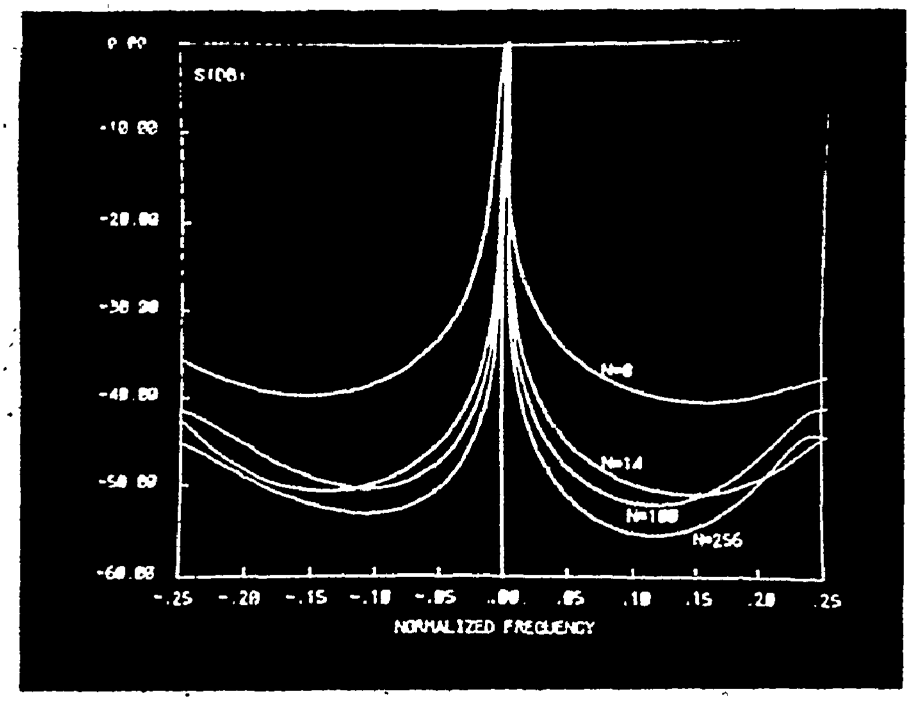


Fig. 5.3 - 3b) ME spectral estimates of ground clutter, with N = 256, 100, 14, and 8; frequency range from -0.25 to 0.25 .

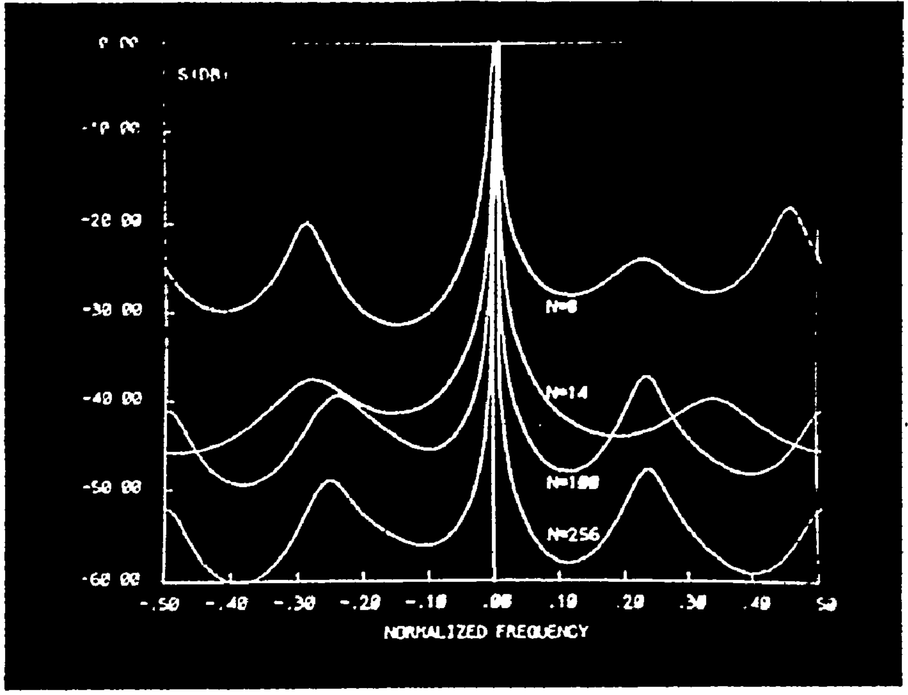


Fig. 5.3 - 4a) Estimates of the reflections from the cell adjacent to that in Fig. 5.3 - 3; frequency range from -0.5 to 0.5 .

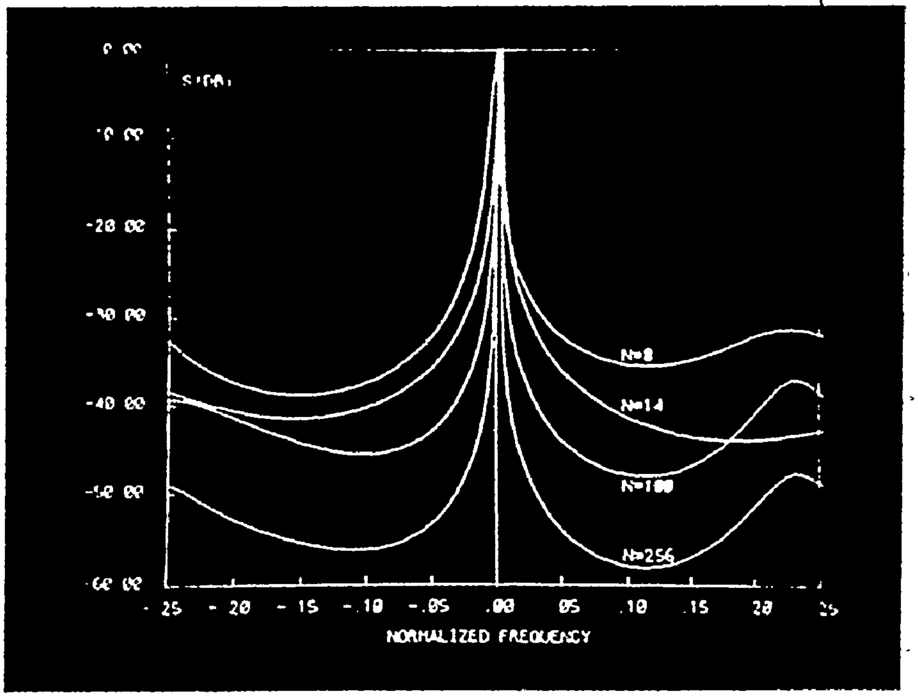


Fig. 5.3 - 4b) Estimates of the reflections from the cell adjacent to that in Fig. 5.3 - 3; frequency range from -0.25 to 0.25 .

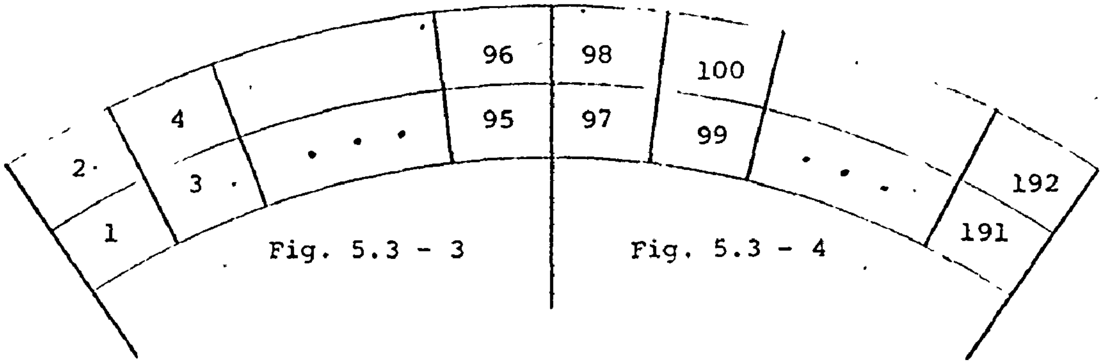


Fig. 5.3 - 5 Arrangements of the sampling areas represented by the spectra in Figures 5.3. - 3 and 5.3 - 4 .

this similarity is also apparent in the other curves too. This result is to be expected because the probability of variation in ground reflectivity characteristics within the contiguous areas is small. It is interesting to note that the analysis of the same data ordered as indicated in Fig. 5.2 - 1(a) (taking all the samples from one range ring and switching then to another) produced practically identical results with those ordered as in Fig. 5.3 - 5, suggesting that the spatial correlations of samples in azimuth and range are of the same order of magnitude. These results were also obtained when the data were ordered as in Fig. 5.2 - 1(c). The characteristics of narrow spectral peaks and low sidelobe level, show that the ground clutter spectra are estimated with high fidelity even with as low as 14 signal samples. The level, position, and shape of the sidelobes depend on the number of data samples, N; in any event their level is below approximately -40 dB.

Similar observations may be made concerning the influence of the number of signal samples on the spectral analysis of storm clouds. Fig. 5.3 - 6(a) shows that the variation in spectra for different N is mini-

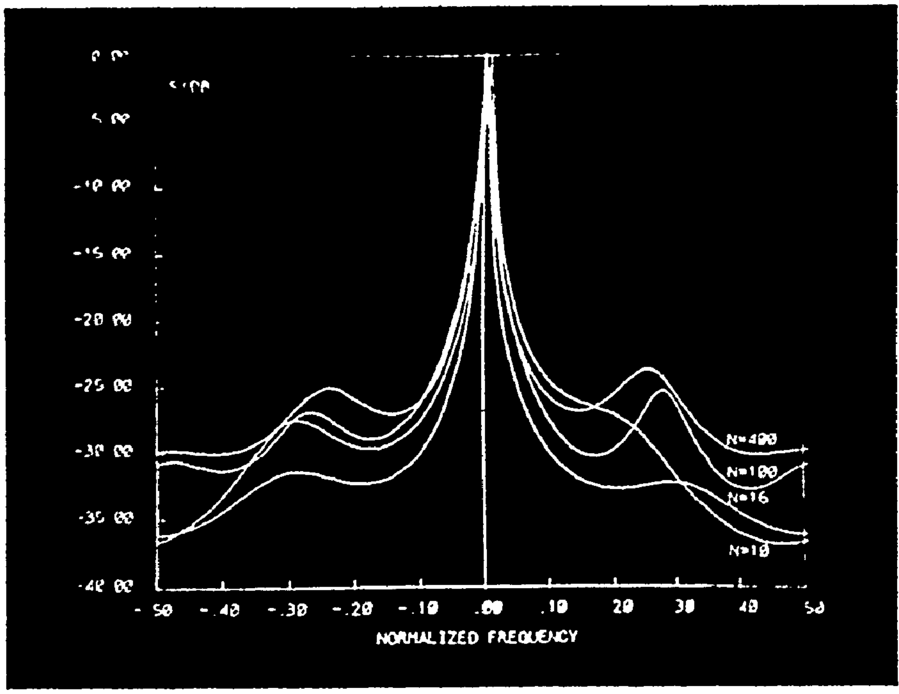


Fig. 5.3 - 6(a) ME spectral estimates of weather clutter, with N = 400, 100, 16, and 10; frequency range from -0.5 to 0.5.

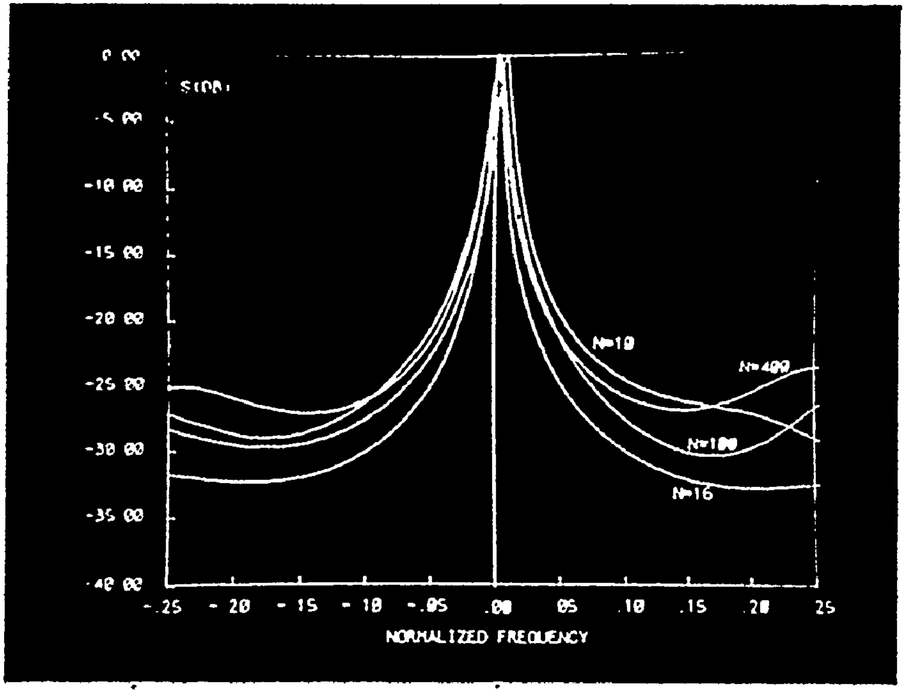


Fig. 5.3 - 6(b) ME spectral estimates of weather clutter, with N = 400, 100, 16, and 10; frequency range from -0.25 to 0.25.

mal. It is also of interest to note that the spectral spread for weather is larger than for the case of ground clutter. This is particularly revealing in Fig.5.3 - 6(b) where the frequency scale has been expanded. The sidelobe levels are about (10 - 15) dB higher, compared with the case of ground clutter.

Figures 5.3 - 7(a) and (b) show the spectra from two adjacent areas with two different frequency scales. Data have been ordered in eight range bin sequences, so that the separation between the adjacent areas was 12 azimuth pulses (ACP's), that is, only 3/4 of a 3 dB azimuth bin. The change in spectral shape is quite noticable.

It is apparent from the results reported above that there is no need for analyzing large numbers of data samples. The spectra are almost the same for all values of N above some minimum value. This value was found to be in the neighbourhood of 10 to 30 samples for all the measurements, regardless of the type of clutter. Therefore, subsequent analyses were based on a relatively short data record.

The power density spectra in the case of birds were found to differ much more from one area to the next, compared with clutter due to clouds and ground. This is in agreement with birds' variable spatial and temporal distribution within flocks. A typical set of normalized spectral densities is shown in Figures 5.3 - 8(a) and (b). The difference in shape between these curves and the previous ones is immediately recognizable. First, the spectral spread is significantly wider; second, the sidelobe levels are higher (note the change in ordinate scale); finally, some of the main peaks are shifted away from zero frequency, thereby resulting in a pronounced asymmetry. These variation will be explored in more detail in section 5.5.

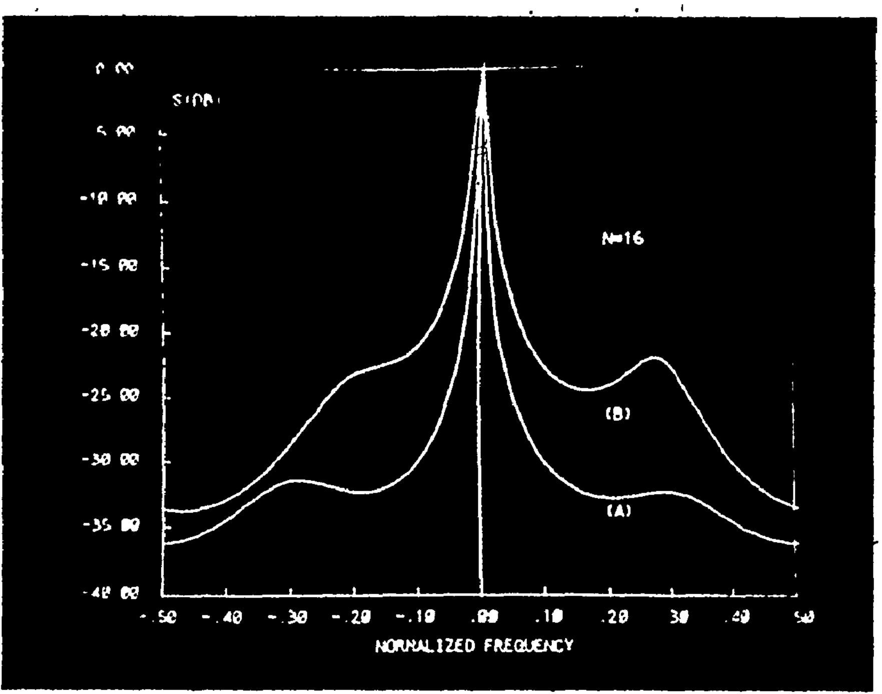


Fig. 5.3 - 7(a) ME spectral estimates of weather clutter from two adjacent cells, N = 16; frequency range from -0.5 to 0.5.

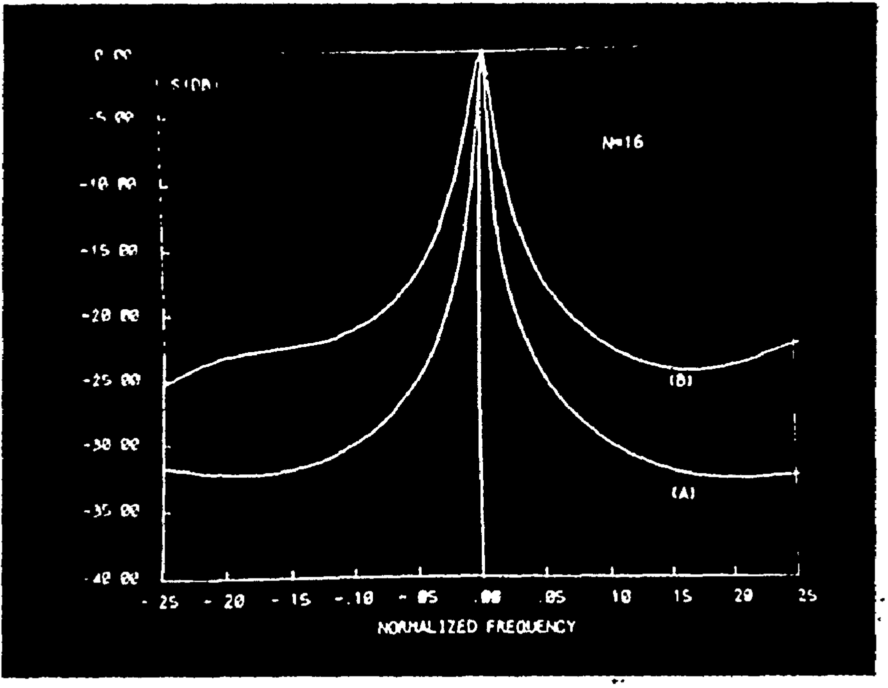


Fig. 5.3 - 7(b) ME spectral estimates of weather clutter from two adjacent cells, N = 16; frequency range from -0.25 to 0.25.

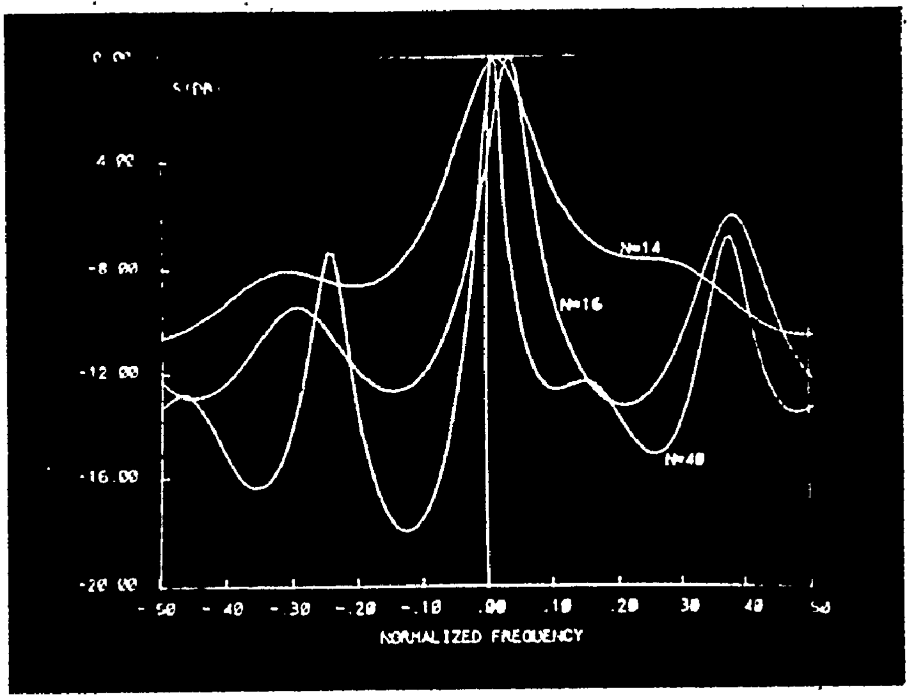


Fig. 5.3 - 8(a) ME spectral estimates of birds clutter, with N = 40, 16, and 14; frequency range from -0.5 to 0.5 .

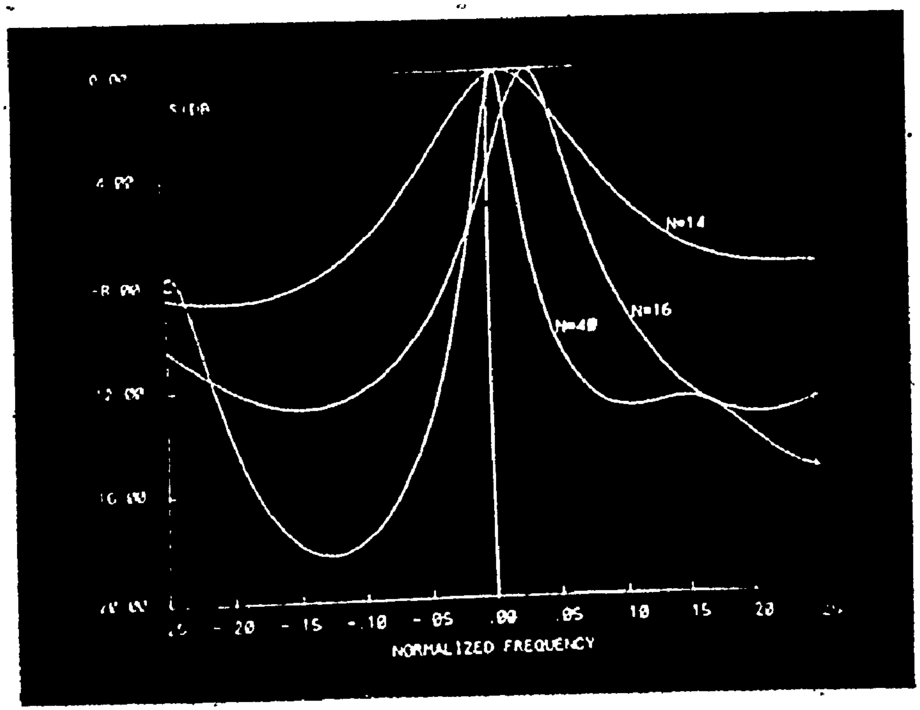


Fig. 5.3 - 8(b) ME spectral estimates of birds clutter, with N = 40, 16, and 14; frequency range from -0.25 to 0.25 .

5.4 Comparison of the MEM With Welch's Method of Spectral Analysis

In this section we present results of the spectral analysis of actual clutter data, which illustrate the comparison of the MEM with the periodogram method based on use of the FFT. In the latter case, the particular procedure chosen is based on Welch's method.

When apply the method of averaged modified periodograms to perform spectral analysis, we have to specify several parameters. When the number of samples is selected, we decide on the number of zeros with which the series will be padded. This number determines the degree of overlap with adjacent periods of spectra and also the density of points on the frequency scale. There is also the number of periodograms that have to be averaged, and the choice of a suitable window function. Since it is not possible to present the influence of all these parameters, it was decided to check experimentally over some of them, and to choose the most appropriate ones. Thus, for example, the Hamming (raised cosine) window was selected as a good compromise between several window functions tested. It was used in all the measurements. For the spectral analysis of 16, 32, and 64 data samples, a 128-point DFT was used, while for the analysis of 128 and 256 data samples, a 512-point DFT was used. Thus, the data were padded with appropriate numbers of zeros. Since the number of zeros added is at least equal to the number of samples analyzed, the effect of overlap is minimized.

The computer programs written for $N = 16$, $N = 32$, and $N = 64$ have an option for the averaging of a maximum of 4 segments of samples. That is, the analysis of 32 data samples, for example, involving the averaging of four modified periodograms, required, in fact, 4x32 = 128 samples. For N = 128 and N = 256 no averaging was done, because we wanted to compare Welch's method with MEM for relatively short data sequences. In any case, the core storage requirements would have been too large for the computer.

The results of the spectral analysis of ground clutter using the Welch method are given in Figures 5.4 - 1(a) and (b) for the case of 16, 32, and 64 data samples. The analysis for N = 256 is shown separately in Fig. 5.4 - 2 for normalized frequency range from -0.25 to 0.25 only. Comparison of Figures 5.4 - 1(a), 5.4 - 1(b), and 5.4 - 2 with Figures 5.3 - 3(a) and 5.3 - 3(b) of the previous section clearly shows that the frequency spread in the estimates obtained using the Welch method is wider than the one in the ME estimates. It is also highly dependent on N, because of the fixed window function applied to the data. The sidelobe level also depends on N, and is about (10 - 20)dB higher in Figure 5.4 - 1(a) than in Figure 5.3 - 3(a). These observations confirm the results obtained for computer-simulated clutter given in the previous chapter.

The spectra in Figs. 5.4 - 1(a) and (b) are obtained by averaging over 4 data segments. For comparison, the results, when no averaging was performed, are given in Fig. 5.4 - 3. As expected, the sidelobe region in the latter Figure is much less smoothed than in the former. Thus, it was decided to average over 4 data segments in the remainder of the analysis.

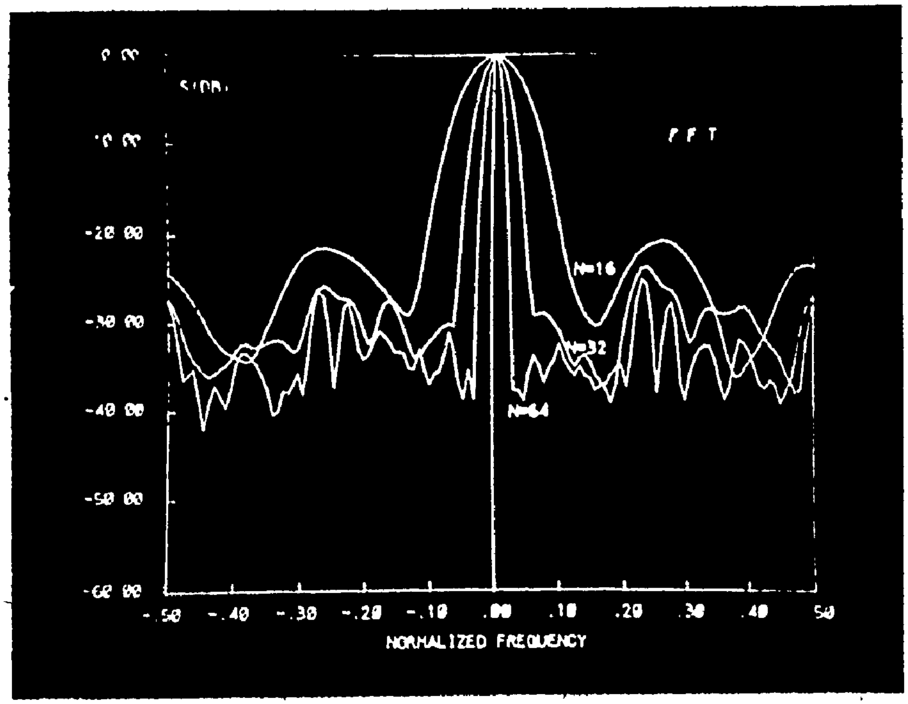


Fig. 5.4 - 1(a) Welch's spectral estimates of ground clutter, with N = 64, 32, and 16, averaged over 4 data blocks; frequency range from -0.5 to 0.5 .

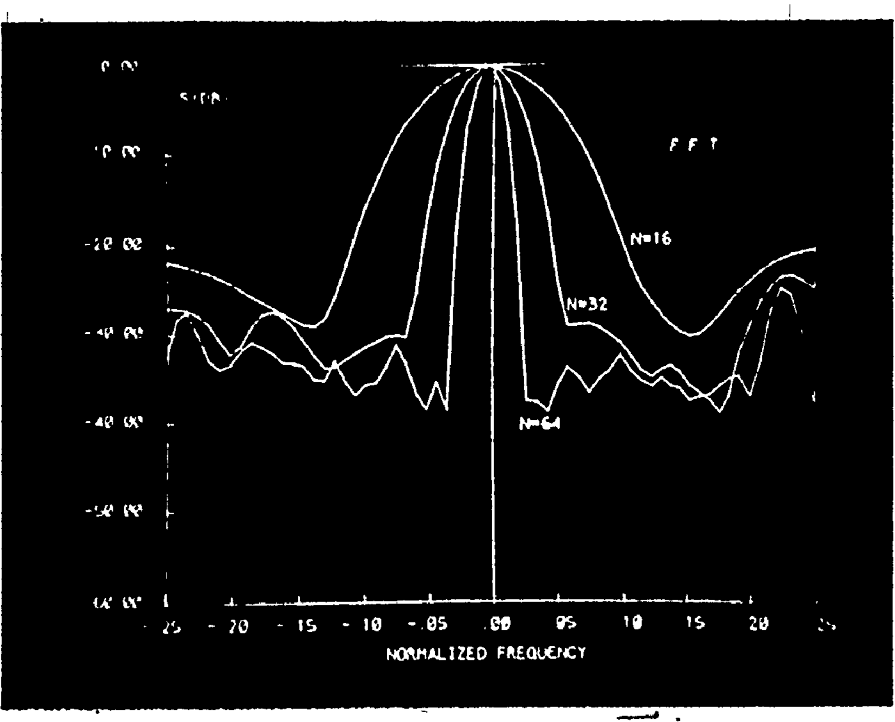


Fig. 5.4 - 1(b) Welch's spectral estimates of ground clutter, with N = 64, 32, and 16, averaged over 4 data blocks; frequency range from -0.25 to 0.25 .

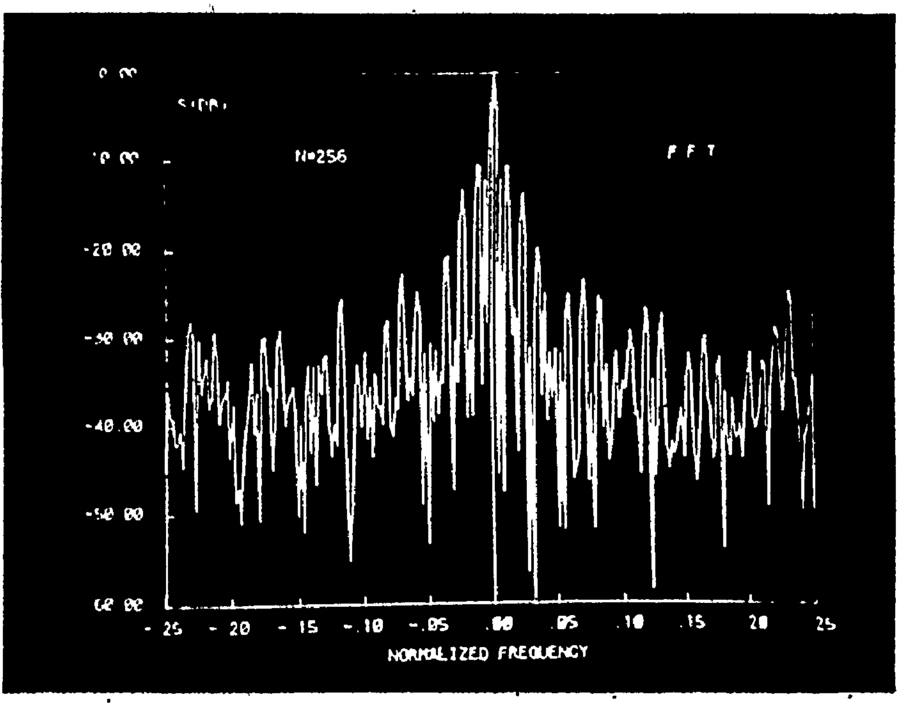


Fig. 5.4 - 2 Welch's spectral estimates of ground clutter, with N = 256; frequency range from -0.25 to 0.25 .

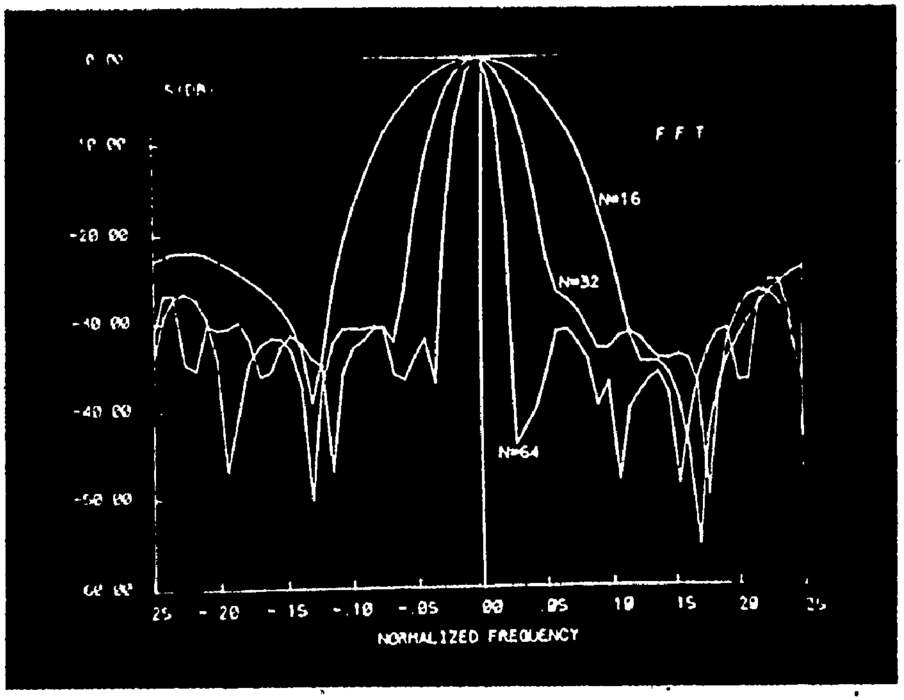


Fig. 5.4 - 3 Welch's spectral estimates as in Fig. 5.4 - 1(b) without averaging.

Figures 5.4 - 4 and 5.4 - 5 show the spectra obtained using the Welch method for weather clutter. Comparing these results with those in Fig. 5.3 - 6(b) (with the same coordinate scales), we may make the same observations as for the case of ground clutter. We may also add that the spectral spreads in Figures 5.4 - 1(b) (for groung clutter), and 5.4 - 4 (for weather) do not differ significantly. The only visible difference is a somewhat smoother transition from the mainlobe to the sidelobe region in Fig. 5.4 - 4.

A direct comparison between the Welch method and the MEM is shown in Figures 5.4 - 6 (for N = 16), and 5.4 - 7 (for N = 100). The differences, mentioned in the previous paragraphs, are clearly displayed in these two Figures.

The last issue in this section is the comparative analysis of all three types of clutter with both methods. Based on previous results, typical sampling areas were selected for the measurements and the results are shown for N = 128 and N = 16. Figures 5.4 - 8 and 5.4 - 9 show the ME and the Welch spectral estimates, respectively, with 128 samples of data used in the analysis. Different spectral spreads for different clutter signals are easily recognizable in the ME estimates, while this is not the case for the Welch ones. Note also the minor peak which appears at the frequency f ≈ 0.13 in both estimates for bird reflection. That peak is not discernible in the ME estimate with 16 data samples (Figure 5.4 - 10) whereas it is still visible in Figure 5.4 - 11 which shows the Welch estimate for the same area as in Figures 5.4 - 8 and 5.4 - 9. However, the spectral spreads are much wider and the sidelobe regions much higher than the corresponding parameters in Fig. 5.4 - 10.

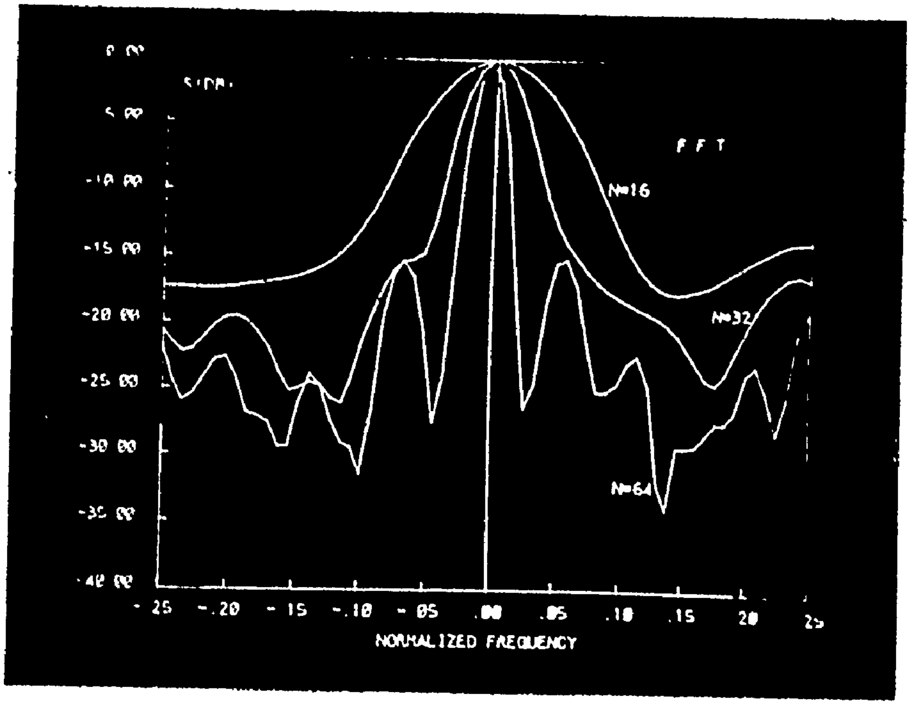


Fig. 5.4 - 4 Welch's spectral estimates of weather clutter, with N = 64, 32, and 16, averaged over 4 data blocks.

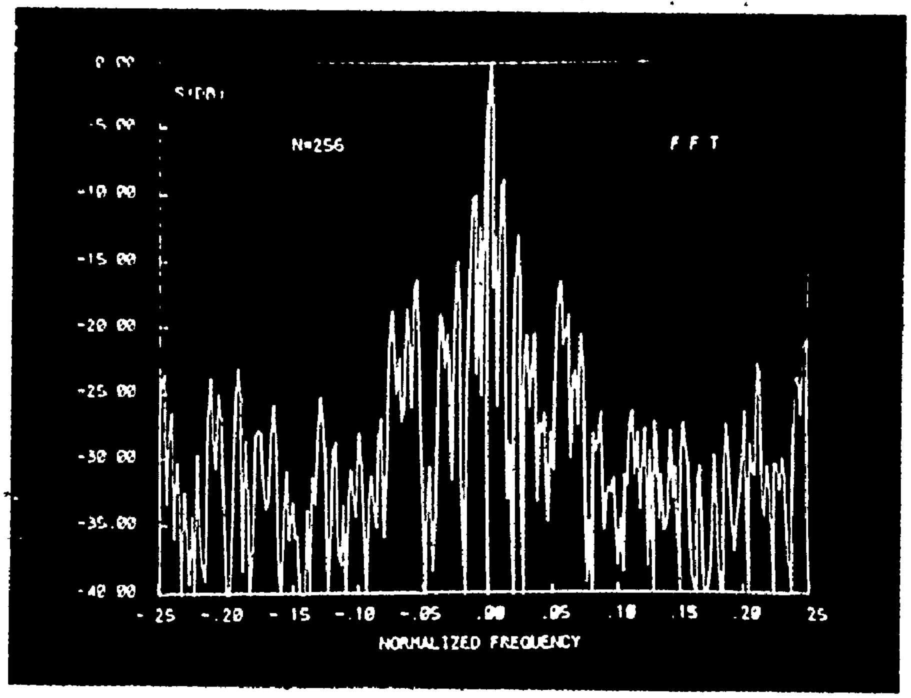


Fig. 5.4 - 5 Welch's spectral estimates of weather clutter, with N = 256.

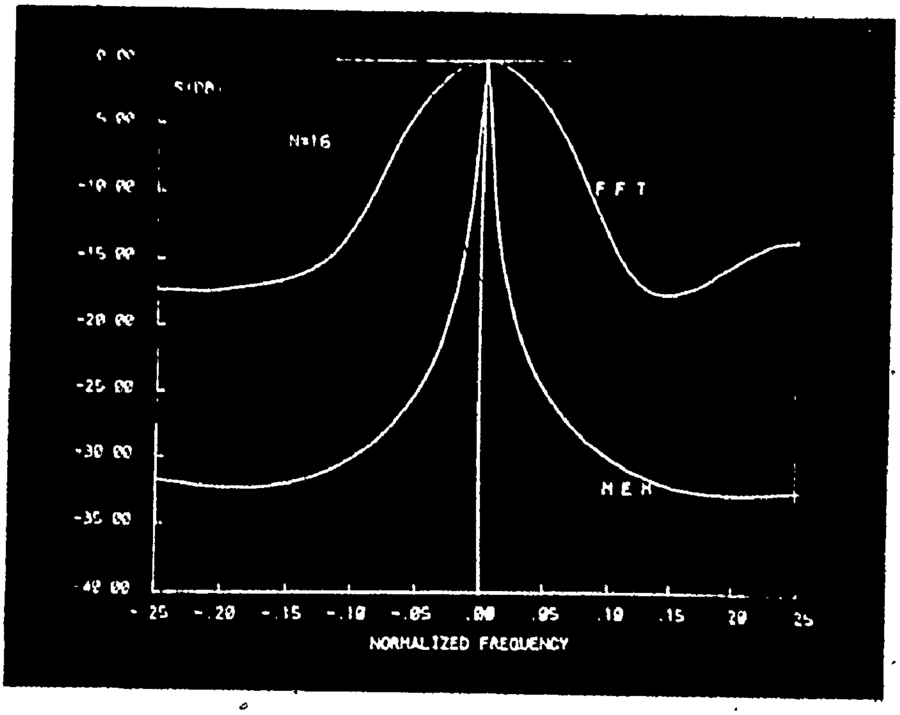


Fig. 5.4 - 6 ME and Welch's estimates of weather clutter, with N = 16.

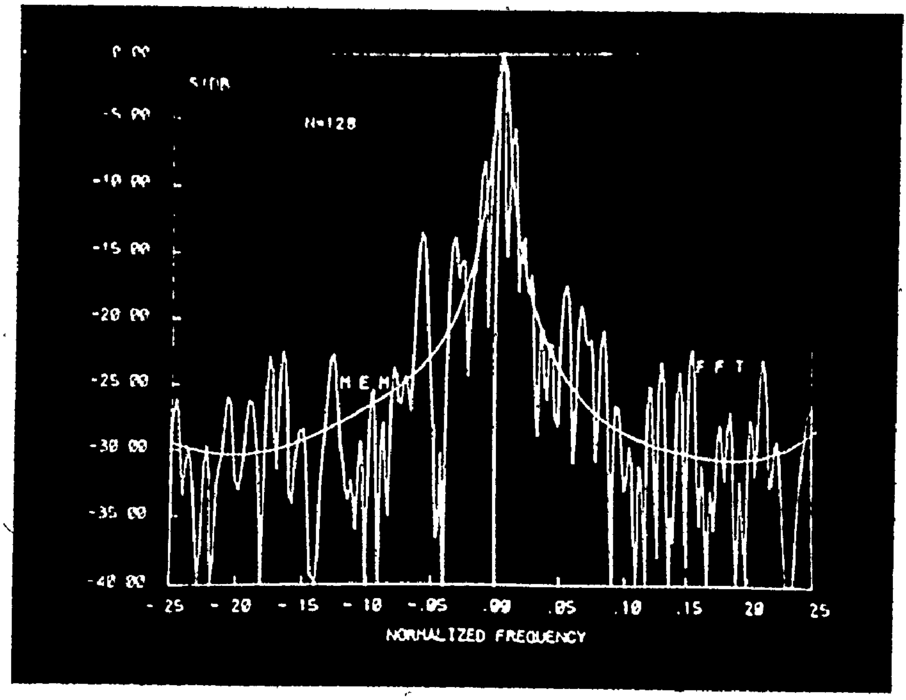


Fig. 5.4 - 7 ME and Welch's estimates of weather clutter, with N = 128.

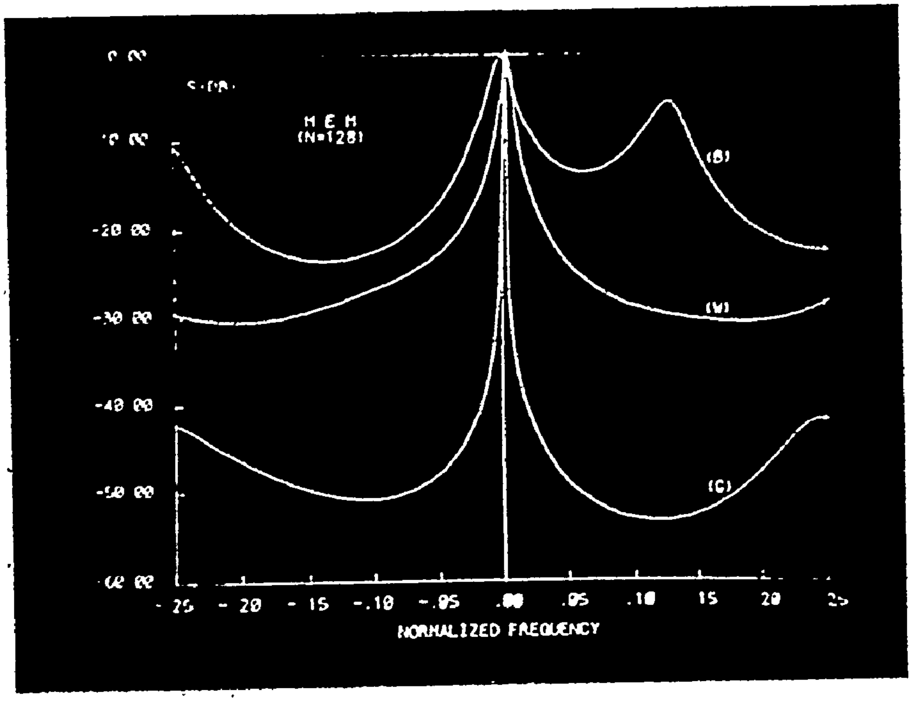


Fig. 5.4 - 8 ME estimates of typical ground, weather, and bird clutter signals, with N = 128.

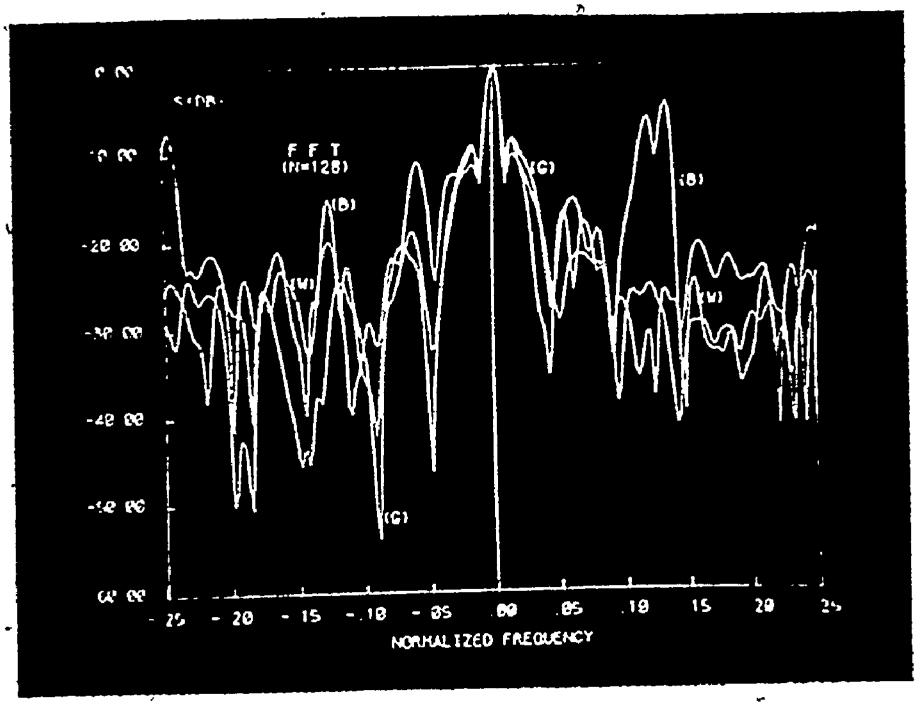


Fig. 5.4 - 9 Welch's estimates of typical ground, weather, and bird clutter signals, with N = 128.

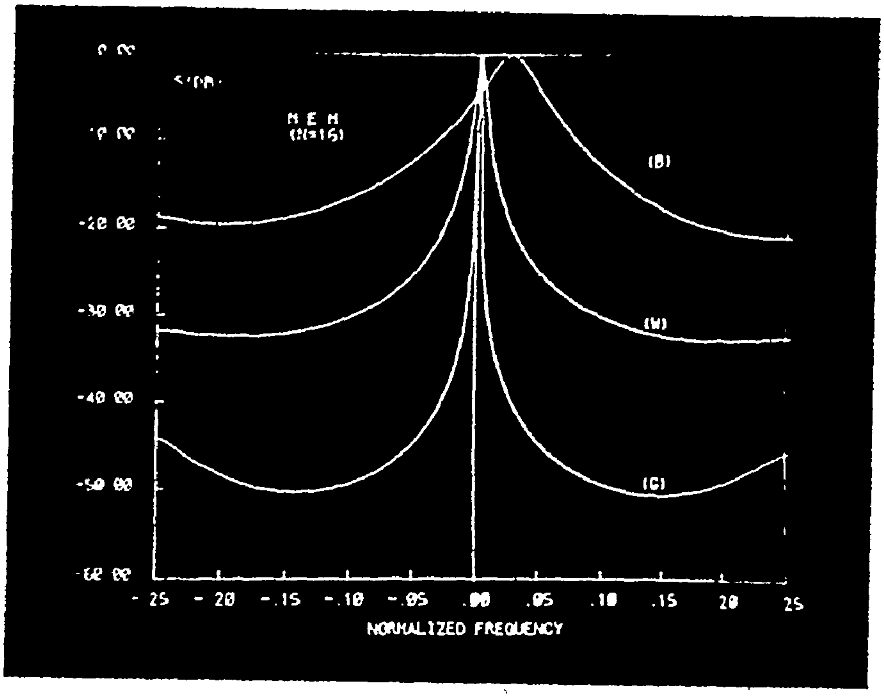


Fig. 5.4 - 10 ME estimates of typical ground, weather, and bird clutter signals, with N = 16 .

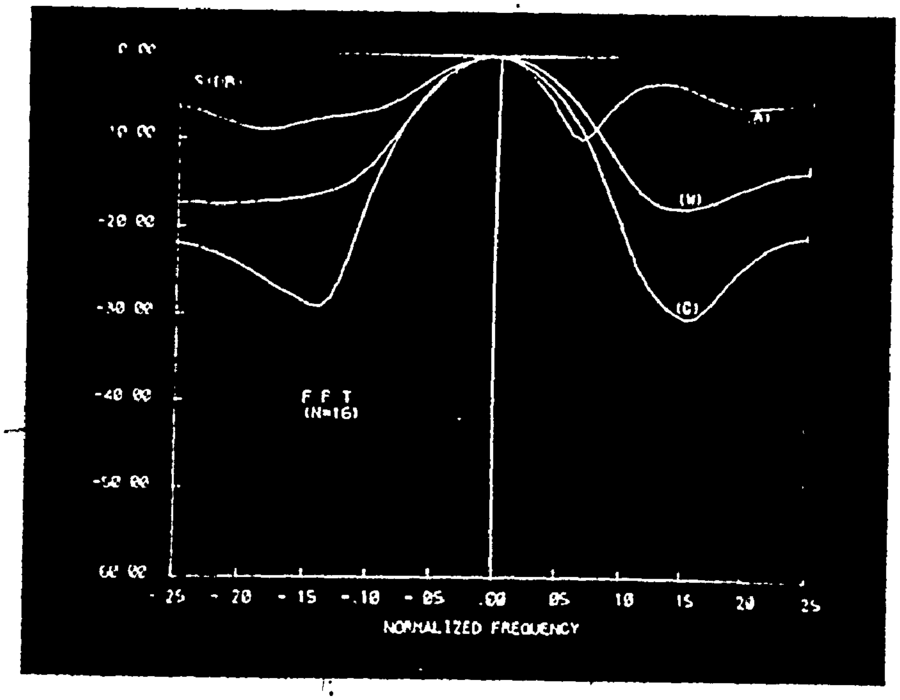


Fig. 5.4 - 11 Welch's estimates of typical ground, weather, and bird clutter signals, with N = 16 .

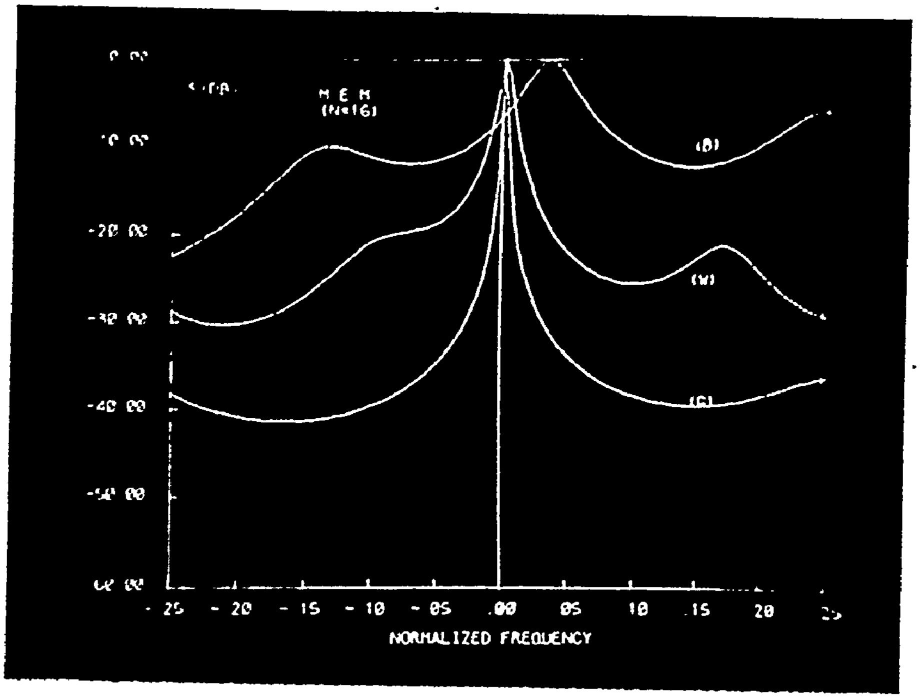


Fig. 5.4 - 12 ME estimates as in Fig. 5.4 - 10 from the adjacent cells.

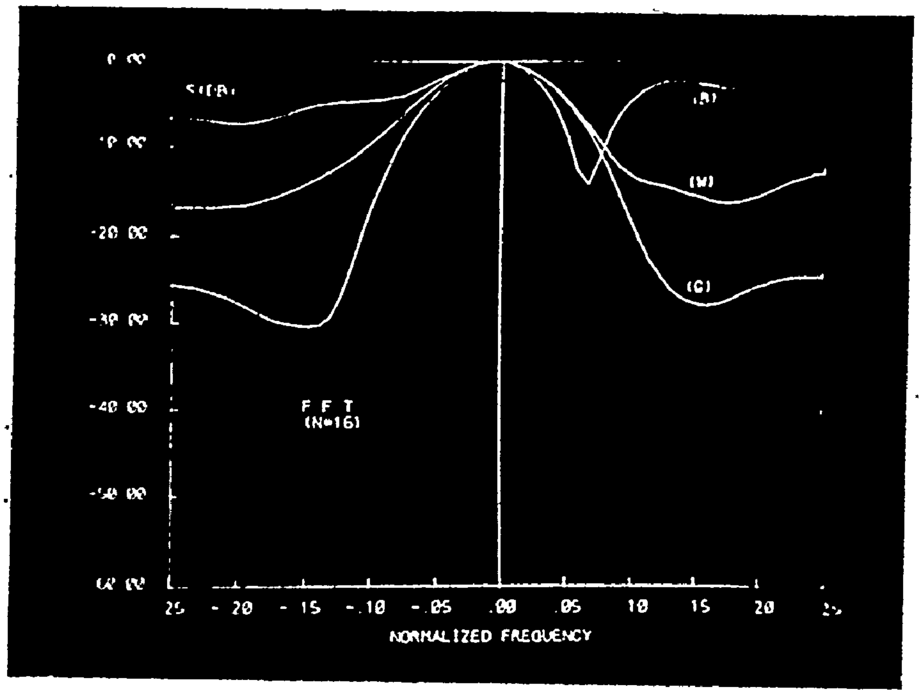


Fig. 5.4 - 13 Welch's estimates as in Fig. 5.4 - 11 from the adjacent cells.

The important point to note is that the spectral spreads are still recognizable in the ML estimates, whereas the main lobes of the curves in Fig. 5.4 - 11 almost overlap.

Finally, Figures 5.4 - 12 and 5.4 - 13 show the same results as Figures 5.4 - 10 and 5.4 - 11, for the adjacent sampling areas. Some minor changes appear in both Figures, but the resolving capabilities are about the same as in Figures 5.4 - 10 and 5.4 - 11.

The results reported above suggest two conclusions:

(1) In the spectral analysis of recorded radar clutter, the MEM has the same advantages over the Welch method, which were predicted from theoretical considerations in Chapter 1, and confirmed by means of computer-simulated clutter in Chapter 4. That is, more reliable estimates for small number of data samples and higher spectral resolution.

(2) The ME estimation provides a possible method for distinguishing amongst the different types of clutter on the basis of different spectral spreads. This possibility is explored more thoroughly in the next section.

5.5 Classification of Different Clutter Returns on the Basis of Statistical Analysis Using the MEM

Figures 5.4 - 8, 5.4 - 9, and 5.4 - 12 indicate that typical spectra for clutter due to ground, weather, and birds can be separated on the basis of producing different spectral spreads. However, because the problem is statistical in nature, we need to analyze a large number of data.

Data have been collected during 4 scans from the adjacent areas in a manner similar to that described in Figure 5.4 - 10 and 5.4 - 12. Memory units (pseudotapes), which served as an input device to the programs for spectral analysis, were filled with the data, and data were analyzed in their proper order. For the comparative analysis to be useful, we need to select a convenient parameter (or parameters). It is apparent from the preceding analysis that the most suitable parameter is the spectral spread. We can define this spread as the frequency band for which the magnitude of the spectral estimate exceeds some predefined value. This value depends on the nature of the problem, and is usually given as the percentage of the maximum value of the magnitude. For example, we may define the spectral spread as the separation between 2 frequencies at which the spectral density drops to 1% (-20 dB) of the maximum value. While this particular value is well-defined in the case of ground and weather spectra, the spectral density of bird echoes, in the majority of cases studied, did not fall below the -20 dB line. Hence, the -10 dB limits were chosen as the basis for measuring spectral spread.

Figures 5.5 - 1 and 5.5 - 2 show the set of spectral estimates of ground clutter for N = 14 and N = 8, respectively. Recognizing that the abscissa in Fig. 5.5 - 1 is extended 5 times with respect to that in Fig. 5.5 - 2 we see that the spectral estimates for various signals are grouped more densly for N = 14 than for N = 8, which is, of course, to be expected. For comparison, sets of curves for N = 100 and N = 256 are presented in Figures 5.5 - 3 and 5.5. - 4, respectively, on the same frequency scale as in Fig. 5.5 - 1. It is apparent that the curves are grouped more closely as the number of analyzed data samples increases. One

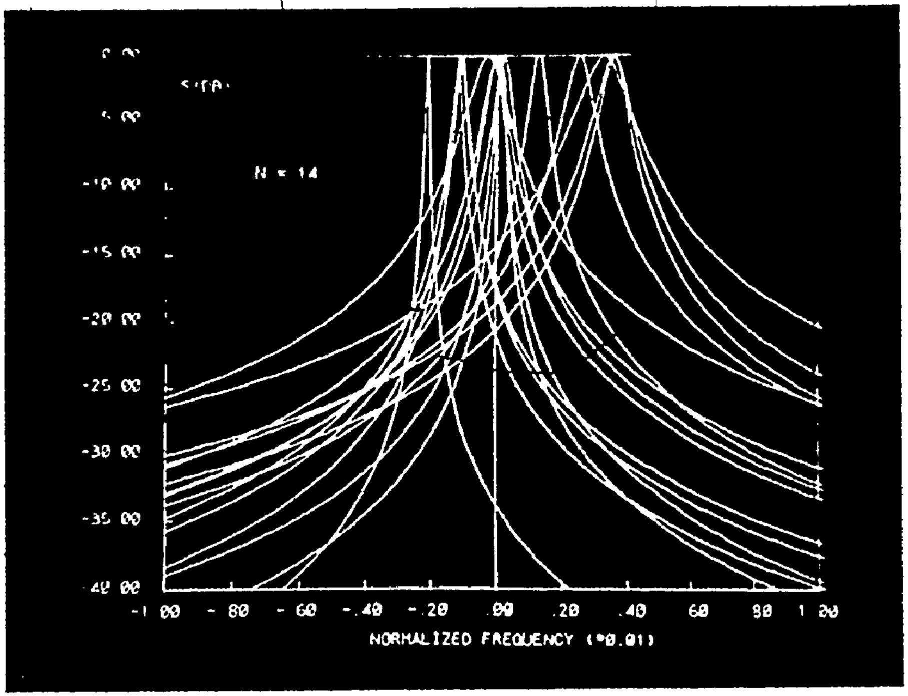


Fig. 5.5 - 1 Typical ME estimates of ground clutter, with N = 14, from several cells and several scans; frequency range from -0.01 to 0.01 .

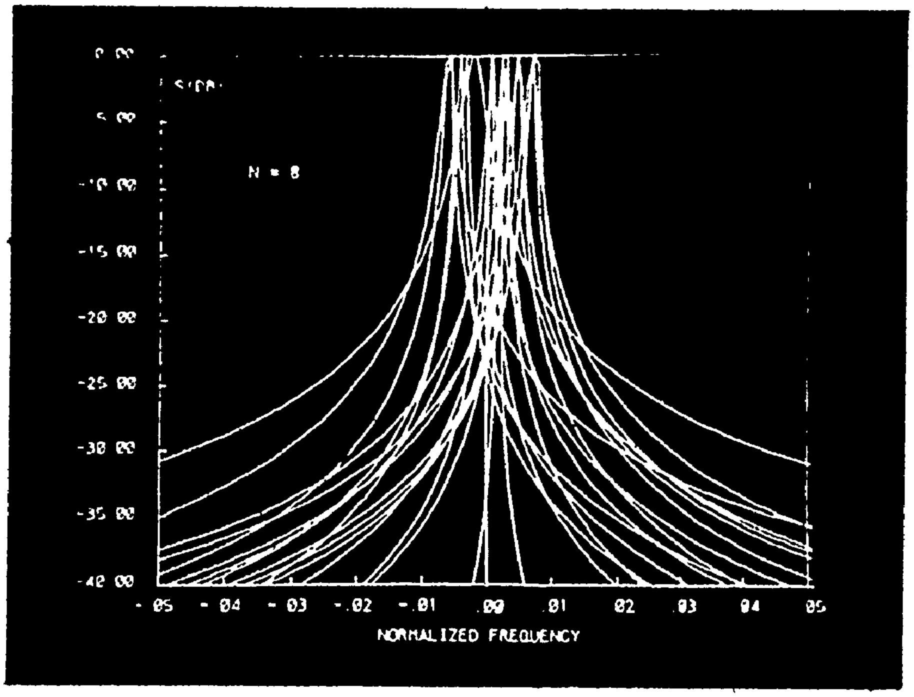


Fig. 5.5 - 2 Typical ME estimates of ground clutter, with N = 8, from several cells and several scans; frequency range from -0.05 to 0.05 .

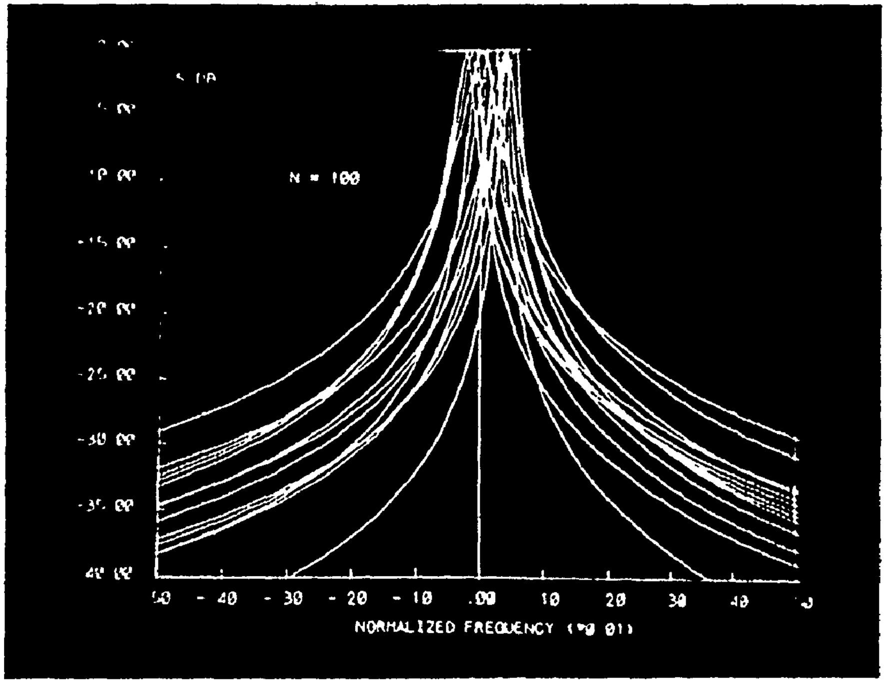


Fig. 5.5 - 3 Typical ME estimates of ground clutter, with N = 100, from several cells and several scans; frequency range from -0.005 to 0.005 .

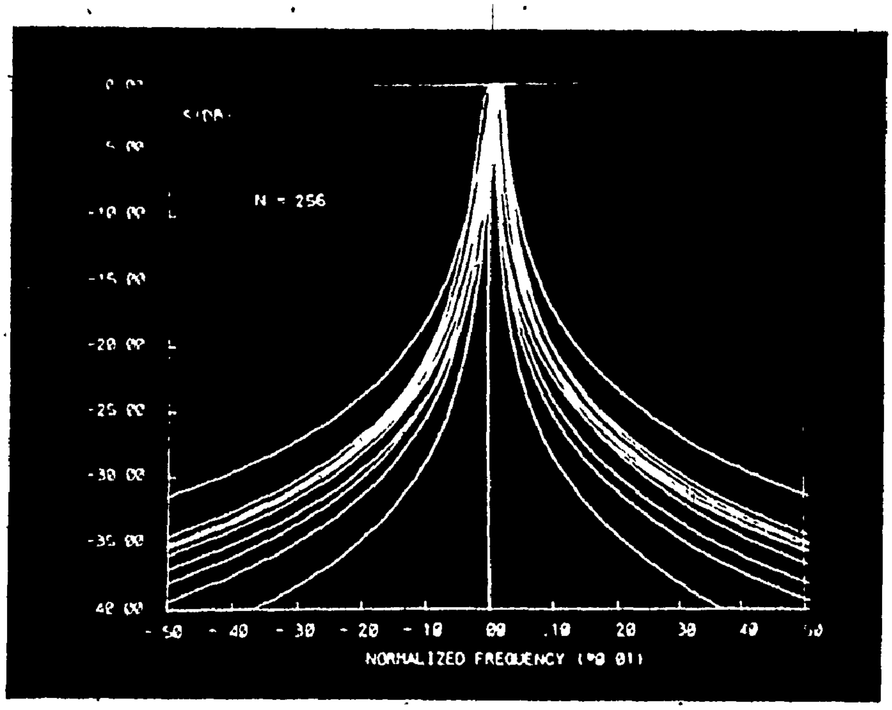


Fig. 5.5 - 4 Typical ME estimates of ground clutter, with N = 256, from several cells and several scans; frequency range from -0.005 to 0.005 .

should have in mind, however, that there is a higher degree of correlation between curves for a larger number of samples, since more samples are common to the adjacent program runs, as mentioned in section 5.3.

These observations also hold for the case of weather clutter presented in Figures 5.5. - 5 and 5.5 - 6. It should be noted that the differencies in the curves are significantly larger than in the case of ground clutter. This is apparent even though the frequency scales differ by an order of magnitude. We also see that there is a noticable shift in the peak frequencies away from the origin.

The differencies in the various spectral estimates are even more pronounced in the case of birds. This is shown in Figures 5.5 - 7 (for N = 40) and 5.5 - 8 (for N = 16). Moreover, most of the curves differ significantly in shape. In fact, the estimates of the signals from the same areas, collected in contiguous scans, somehow resemble each other. To emphasize this resemblance, we have separated the diagrams corresponding to the particular cells within the sampling area. The curves corresponding to the first area, collected over 4 contiguous scans, with N = 40 are shown in Fig. 5.5 - 9(a). The collection from the next three areas are shown in Figs. 5.5 - 9(b), (c), and (d). These diagrams illustrate a small spatial correlation of the echoes reflected from bird flocks, and a relatively slow variation of their spatial distribution from scan to scan. These characteristics of bird echoes have been confirmed by several investigators [93]-[100].

These observations, made for N = 40 samples, are also valid for N = 16. Figures 5.5 - 10(a), (c), (d), and (e) correspond to Figures 5.5 - 9(a), (b), (c), and (d), respectively. For comparison, we have

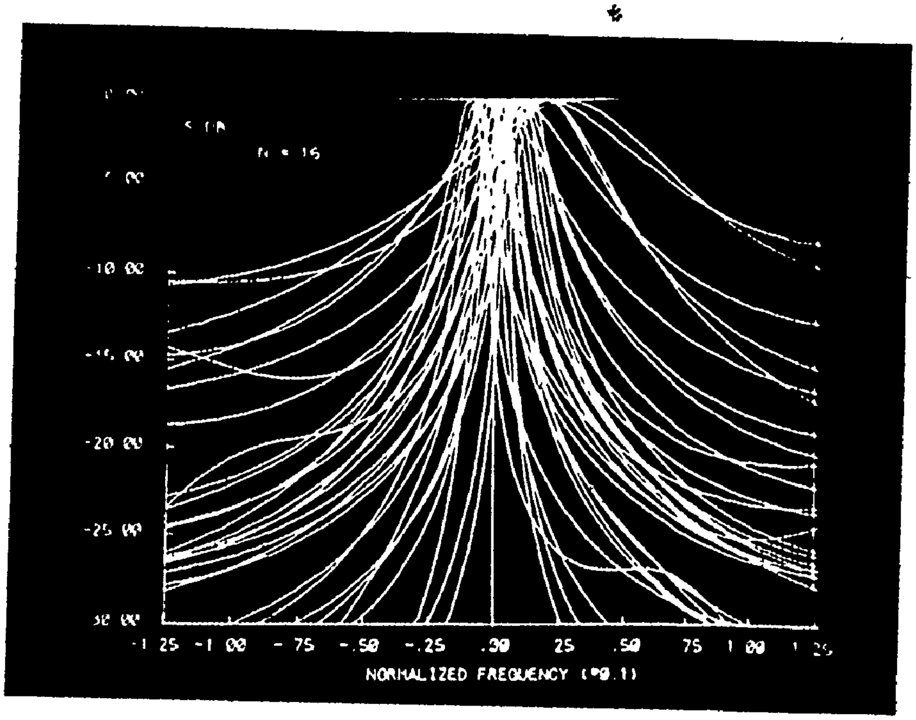


Fig. 5.5 - 5 Typical ME estimates of weather clutter, with N = 16, from several cells and several scans; frequency range from -0.125 to 0.125 .

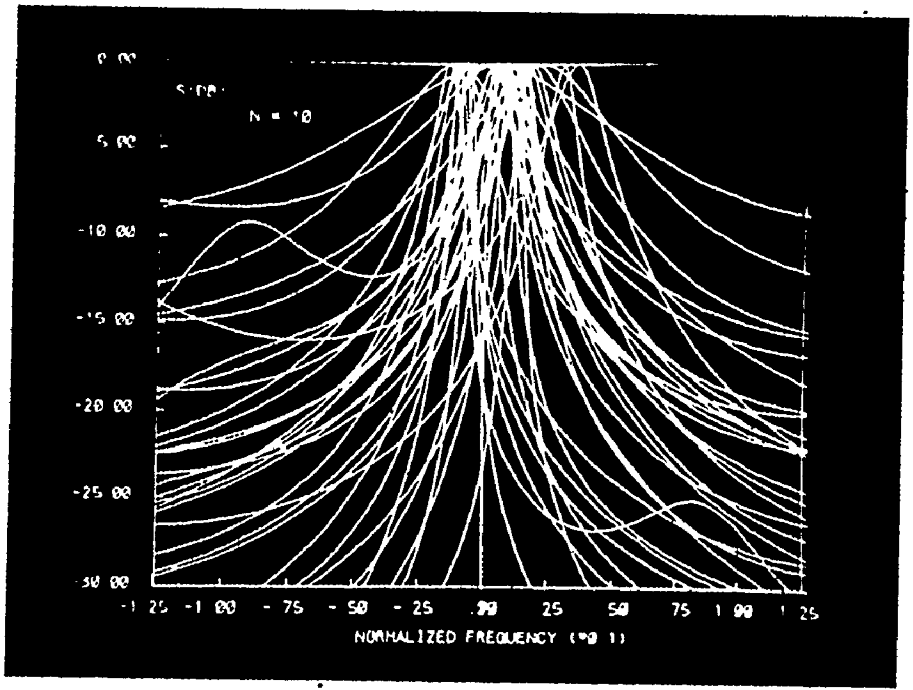


Fig. 5.5 - 6 Typical ME estimates of weather clutter, with N = 10, from several cells and several scans; frequency range from -0.125 to 0.125 .

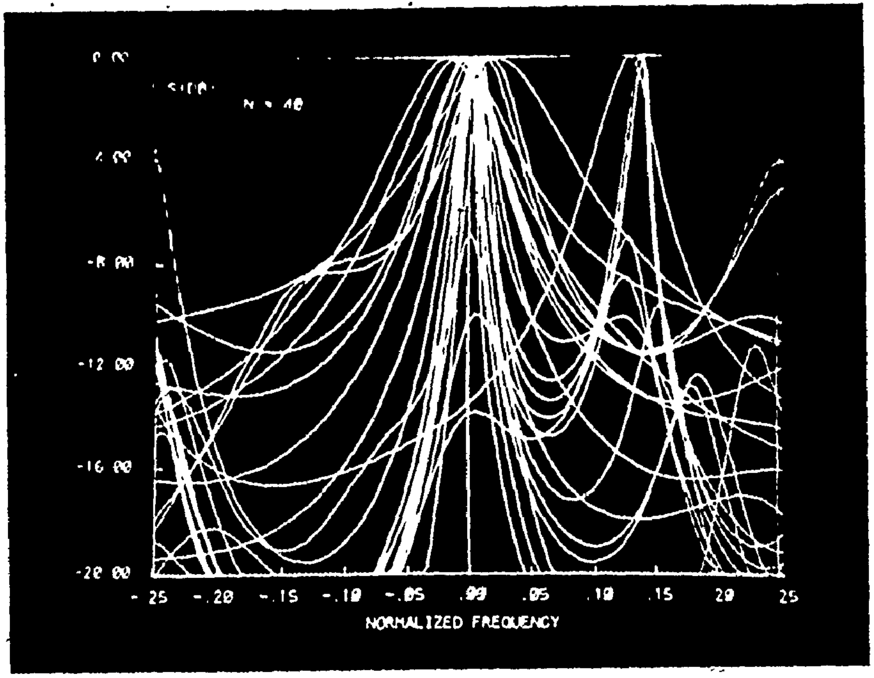


Fig. 5.5 - 7 Typical ME estimates of birds clutter, with N = 40, from several cells and several scans; frequency range from -0.25 to 0.25 .

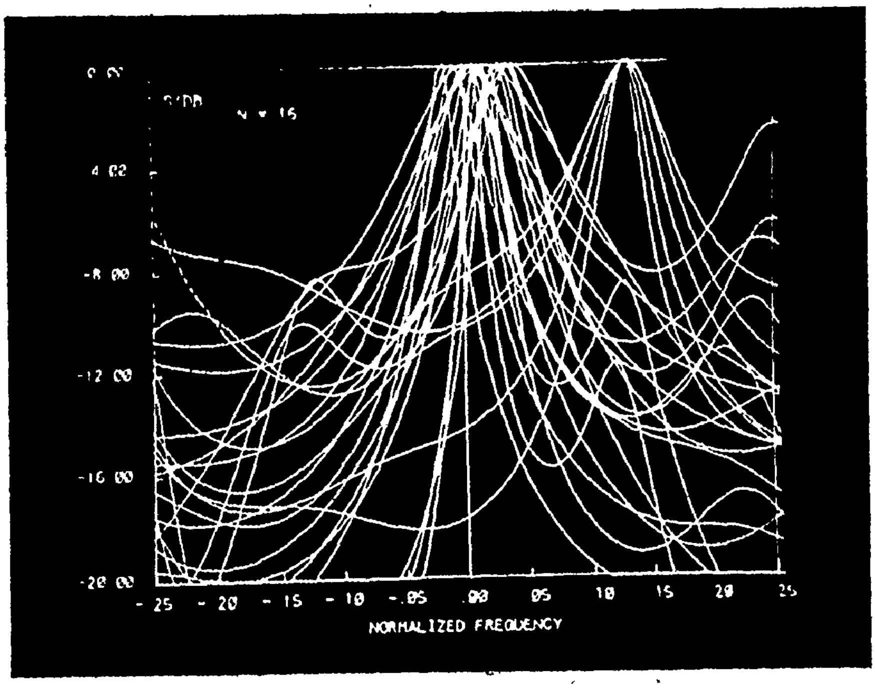


Fig. 5.5 - 8 Typical ME estimates of birds clutter, with N = 16, from several cells and several scans; frequency range from -0.25 to 0.25 .

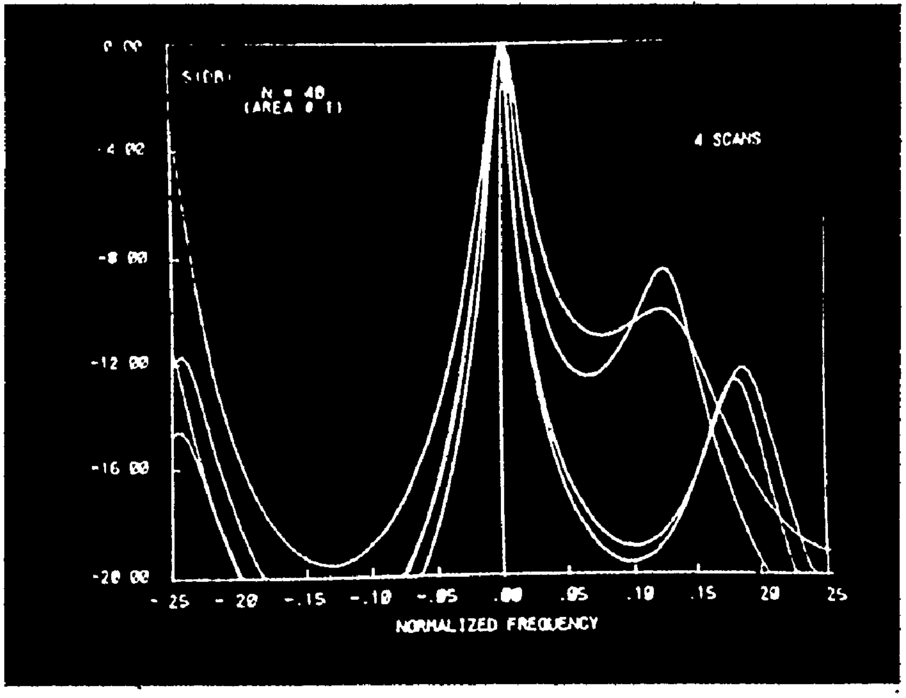


Fig. 5.5 - 9(a) ME estimates of birds from one area, with N = 40, collected over 4 consecutive scans.

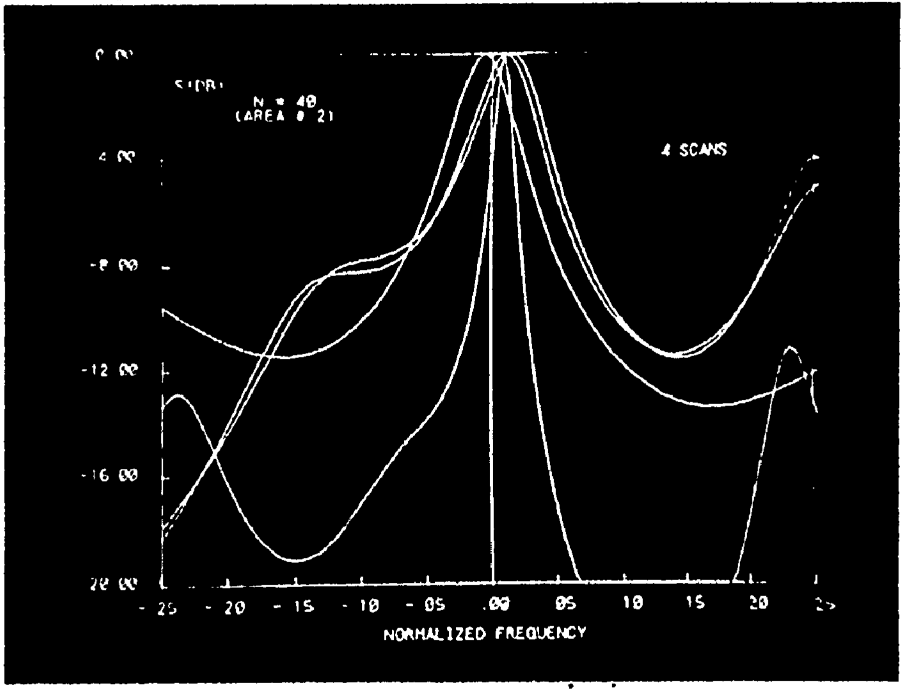


Fig. 5.5 - 9(b) ME estimates of birds from the next adjacent area, with N = 40, collected over 4 consecutive scans.

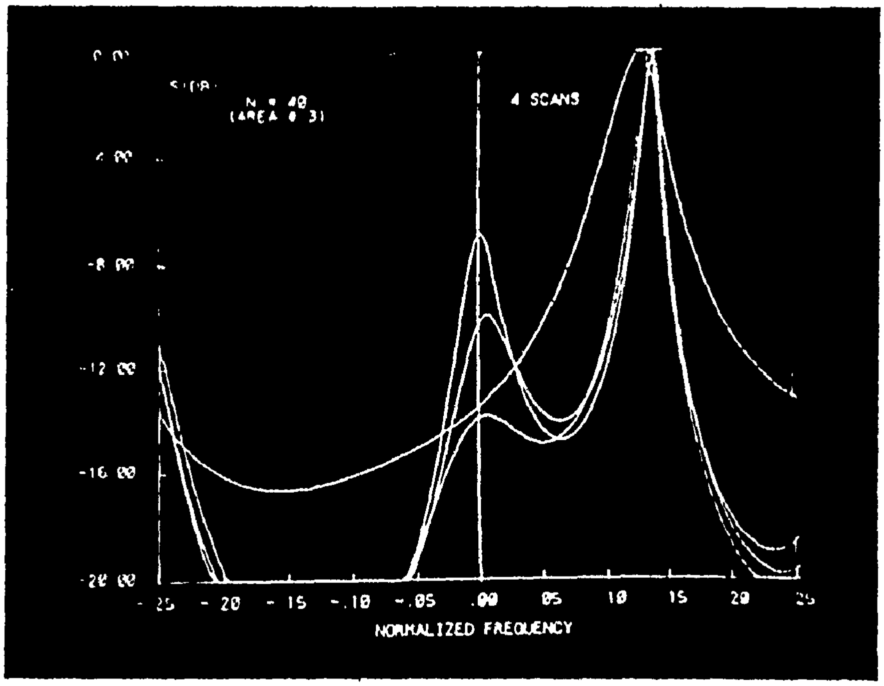


Fig. 5.5 - 9(c) ME estimates of birds from the third adjacent area, with N = 40, collected over 4 consecutive scans.

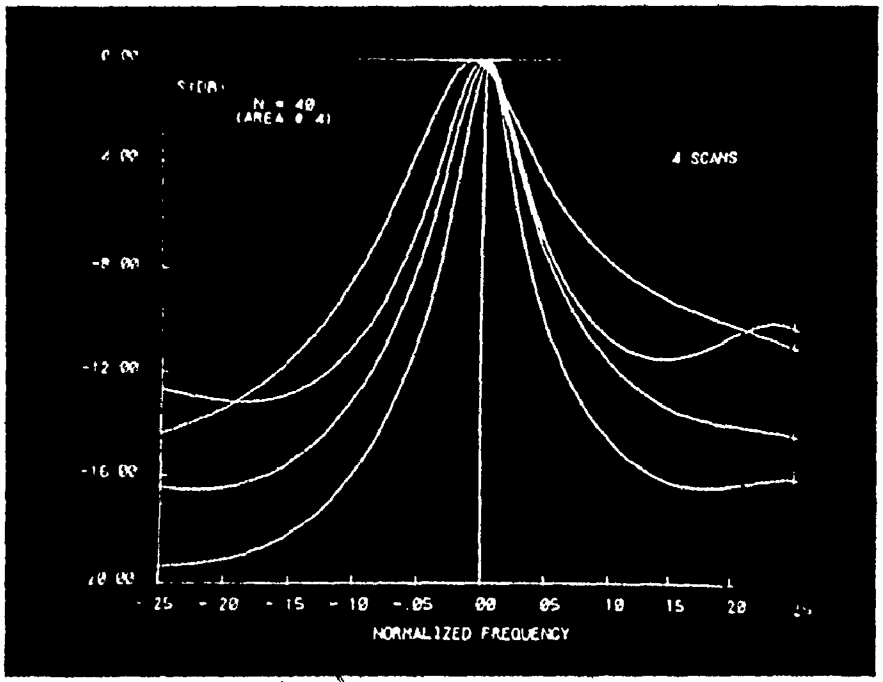


Fig. 5.5 - 9(d) ME estimates of birds from the fourth adjacent area, with N = 40, collected over 4 consecutive scans.

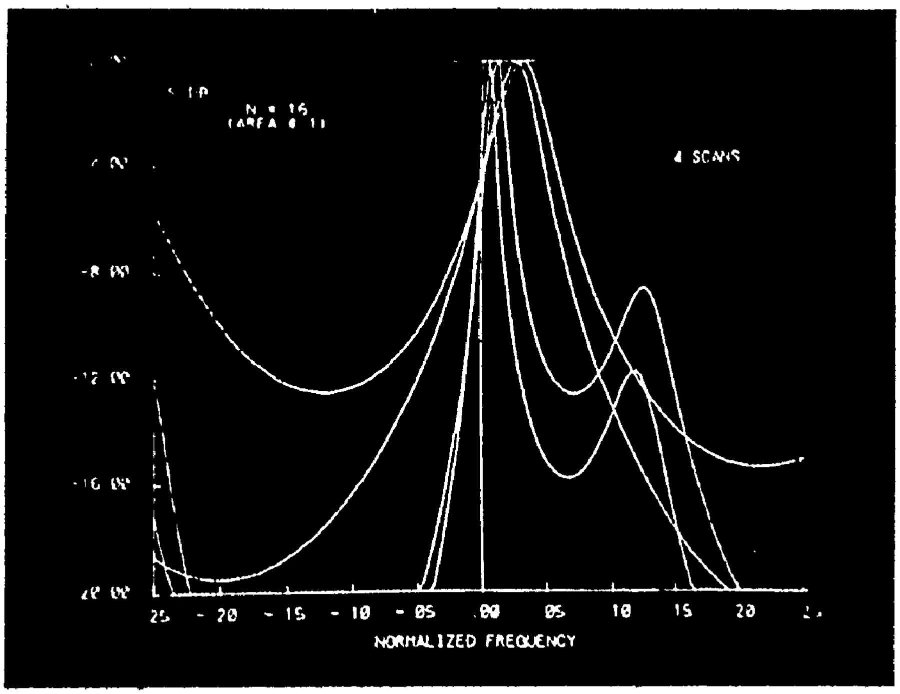


Fig. 5.5 - 10(a) ME estimates of birds from the first area, with N = 16, collected over 4 consecutive scans.

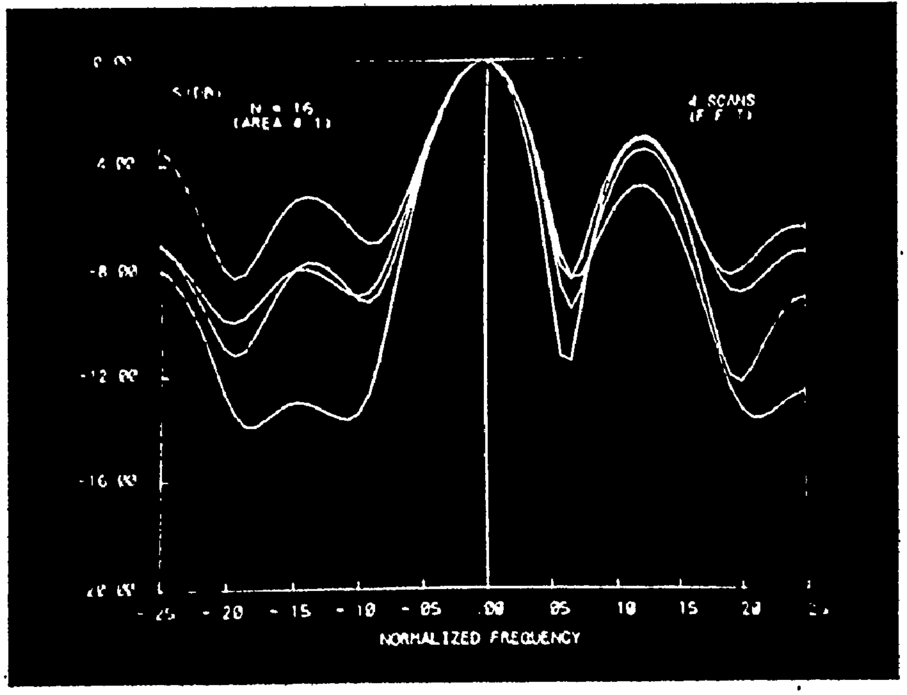


Fig. 5.5 - 10(b) Welch's estimates of birds from the same area as in Fig. 5.5 - 10(a) with N = 16, collected over 4 consecutive scans.

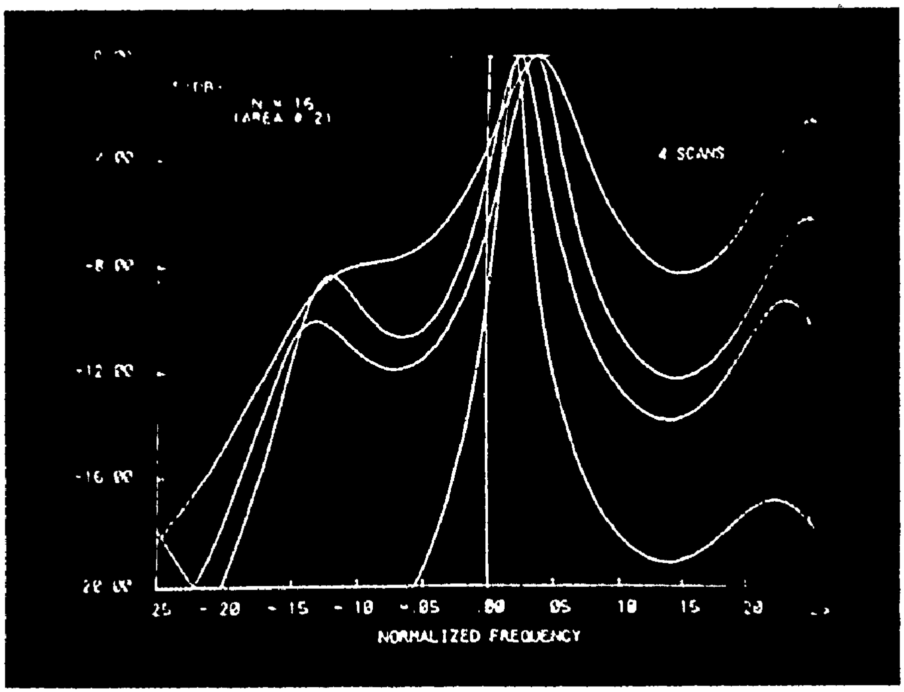


Fig. 5.5 - 10(c) ME estimates of birds from the second area, with N = 16, collected over 4 consecutive scans.

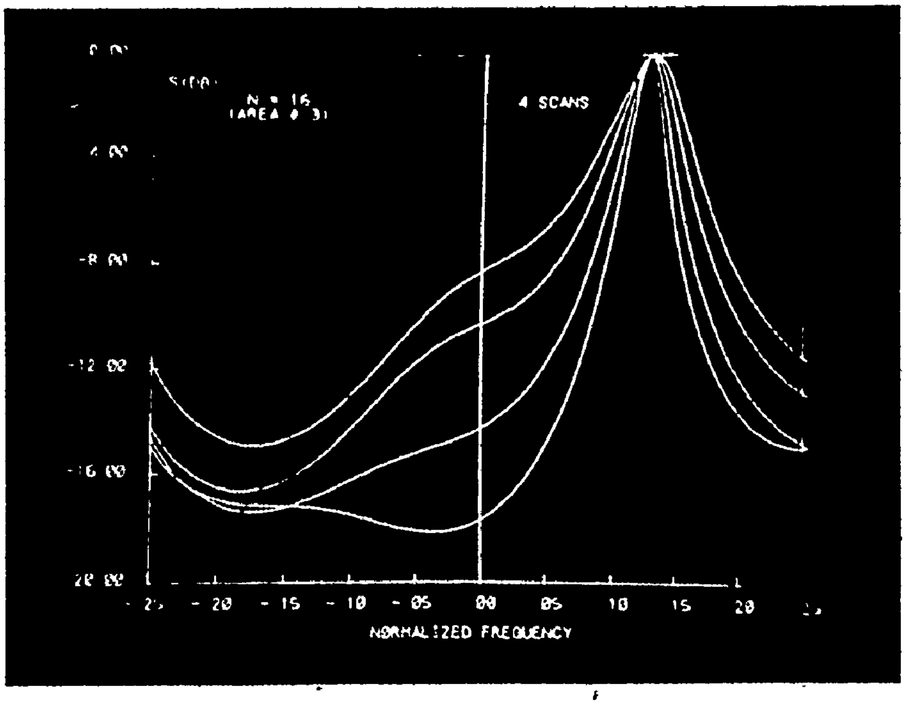


Fig. 5.5 - 10(d) ME estimates of birds from the third area, with N = 16, collected over 4 consecutive scans.

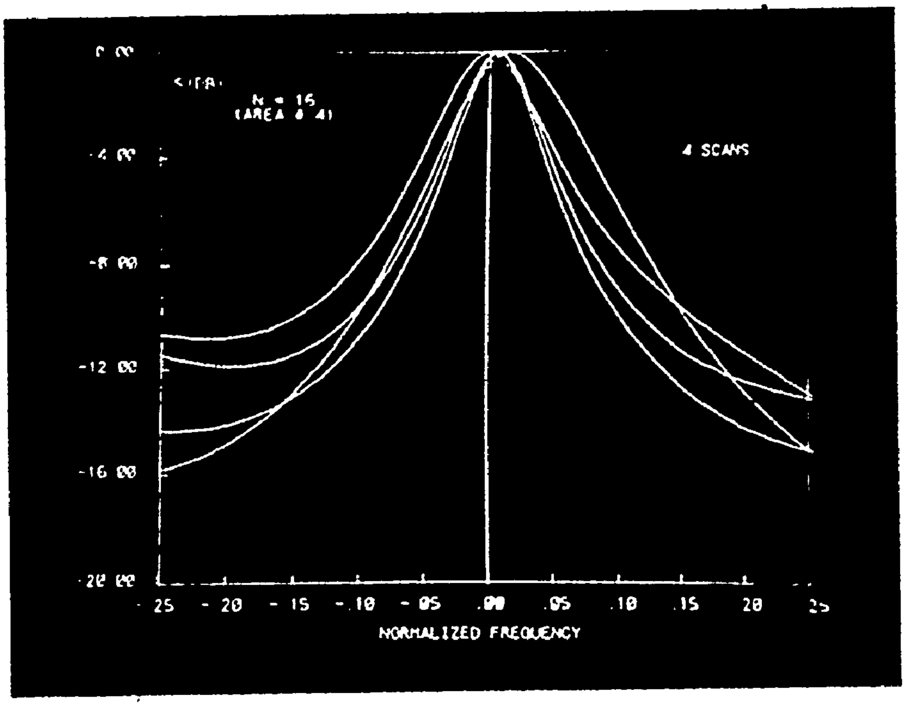


Fig. 5.5 - 10(e) ME estimates of birds from the fourth area, with N = 16, collected over 4 consecutive scans.

inserted 16 point estimates using the Welch method for area # 1 (Figure 5.5 - 10(b)). The spectra in Figures 5.5. - 10(a) and (b) reveal similar patterns, but the sidelobe levels in the Welch estimate of only (3 - 6)dB below the central peak, illustrate, once again, drawback of the method with such short signal records.

The qualitative results, described above, indicate that it is possible to provide a practical means for the discrimination of the various types of clutter, based on the maximum entropy analysis of power spectral densities. In order to provide feeling for the statistical variability of spectral spread for different clutter conditions and, also, for different record lengths, a large number of spectra were calculated, and the spectral spread, corresponding to the -10 dB points, were extracted. Next, the sample mean, the sample variance, and the sample standard deviation of the spread were calculated by using the following formulas:

$$\mu_{\Delta F} = \frac{1}{N}\sum_{n=1}^{N} \Delta F_n \quad . \tag{5.5.1}$$

$$\sigma^2_{\Delta F} = \frac{1}{N-1}\cdot\sum_{n=1}^{N} (\Delta F_n - \mu_{\Delta F})^2 \tag{5.5.2}$$

$$\sigma_{\Delta F} = \sqrt{\sigma^2_{\Delta F}} \tag{5.5.3}$$

Table 5.5 - 1 shows the results of using Eqs.(5.5.1) to (5.5.3) for all three types of clutter, with both $N = 16$ and $N = 40$. In each case, the spectra were calculated using 48 runs. The values of μ, σ^2, and σ are given for both the spectral spread ΔF and the spectral peak location F_o. The latter results were included because it was found that

No. of samples			N = 16			N = 40		
Spec par.	Sam. aver	Multp. factor	Ground	Weather	Birds	Ground	Weather	Birds
ΔF	μ	10^{-3}	0.847	25.83	151.3	1.785	32.54	133.7
	σ	10^{-3}	0.451	22.30	106.0	0.664	25.60	115.0
	σ^2	10^{-6}	0.204	496.00	11140.0	0.440	656.00	13300.0
F_o	μ	10^{-3}	0.013	1.30	27.5	0.740	1.38	25.6
	σ	10^{-3}	0.612	12.50	43.6	0.737	5.94	52.5
	σ^2	10^{-6}	0.380	156.00	1900.0	0.540	35.24	2760.0

TABLE 5.5 - 1.

Sample averages for spread, ΔF, and peak frequency, F_o, with 16 and 40 data samples. Note that, since the normalizing frequency, PRF, is 1.04x 10^3 Hz, entries for μ and σ are very close to absolute values in hertz.

the location of the spectral peak depends on the type of clutter. Hence, it may provide additional information for the classification of clutter. We see that the spectral spreads for the three different cases are easily discernible.

Based on the results described above, we conclude that the resolution capability of the MEM provides a practical means for the discrimination of the three different types of clutter. As mentioned in the introductory chapter, this possibility can prove useful in ATC radars for vectoring aircrafts in a way to avoid hazardous areas of migrating flocks of birds or heavy storms. Also, it may find applications in meteorological radars used for high resolution studies of weather conditions.

To illustrate the differencies in the statistical averages of spectral spread for the three types of clutter, we present the results of Table 5.5 - 1 in Figures 5.5 - 11 (for N = 16) and 5.5 - 12 (for N = 40). The solid vertical lines in these Figures represent the sample means, while the dotted lines represent the boundary values determined by standard deviations.

Figure 5.5- 13 shows the histogram for the values of ground clutter with N = 16. Normalized values of ΔF are given on the horizontal axes. For convenience, the frequency of occurrence is also normalized, so that the value of the vertical axis is dimensionless quantity, as shown by

$$n_c = \frac{n_{\Delta F}}{n \cdot \delta F} \qquad (5.5.4)$$

where

n = total number of observations,

δF = length of the selected interval on the ΔF-axis, and

$n_{\Delta F}$ = number of observations that fall inside δF.

Thus, the total area under the histogram is equal to unity.

Similar histograms for weather clutter and bird echoes are given in Figures 5.5 - 14 and 5.5 - 15, respectively. The horisontal scales in the last three Figures are different because of different spreads. As a consequence of Eq. (5.5.4), the vertical scales differ, too. Comparison of the histograms was limited to the cases of weather and birds only for two reasons. First, the results of previous analysis suggest

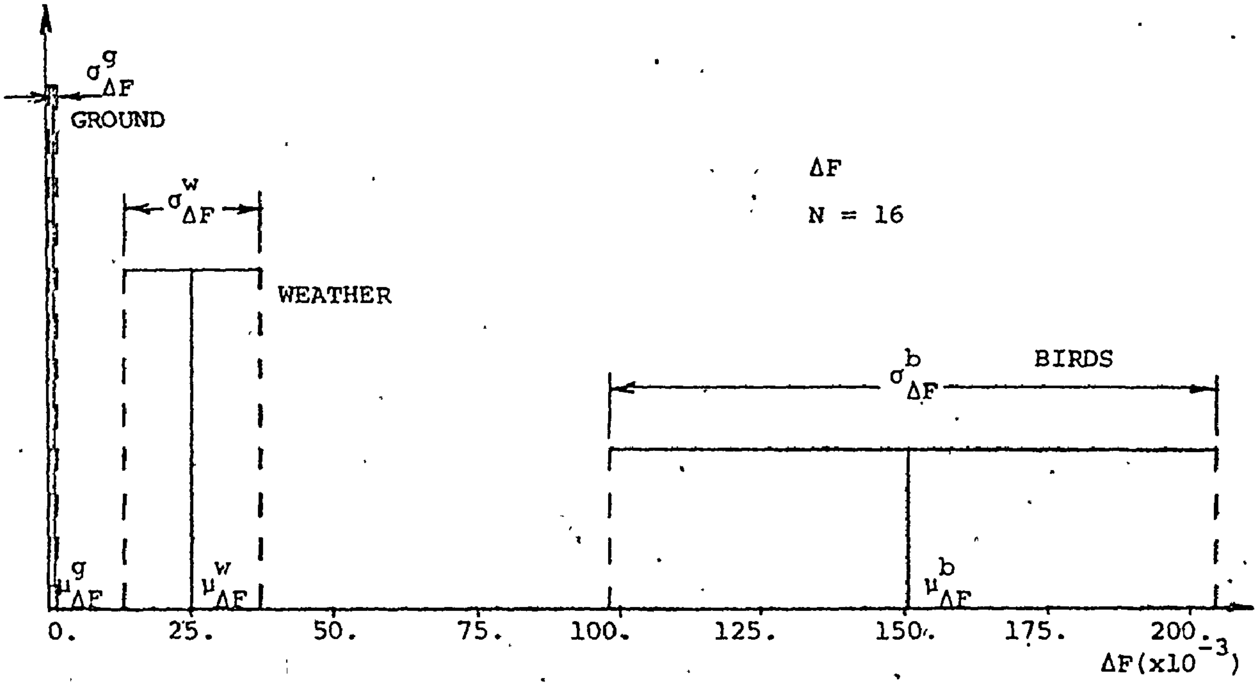


Fig. 5.5 - 11 Comparison of the statistical averages of ΔF for ground, weather, and birds; N = 16 .

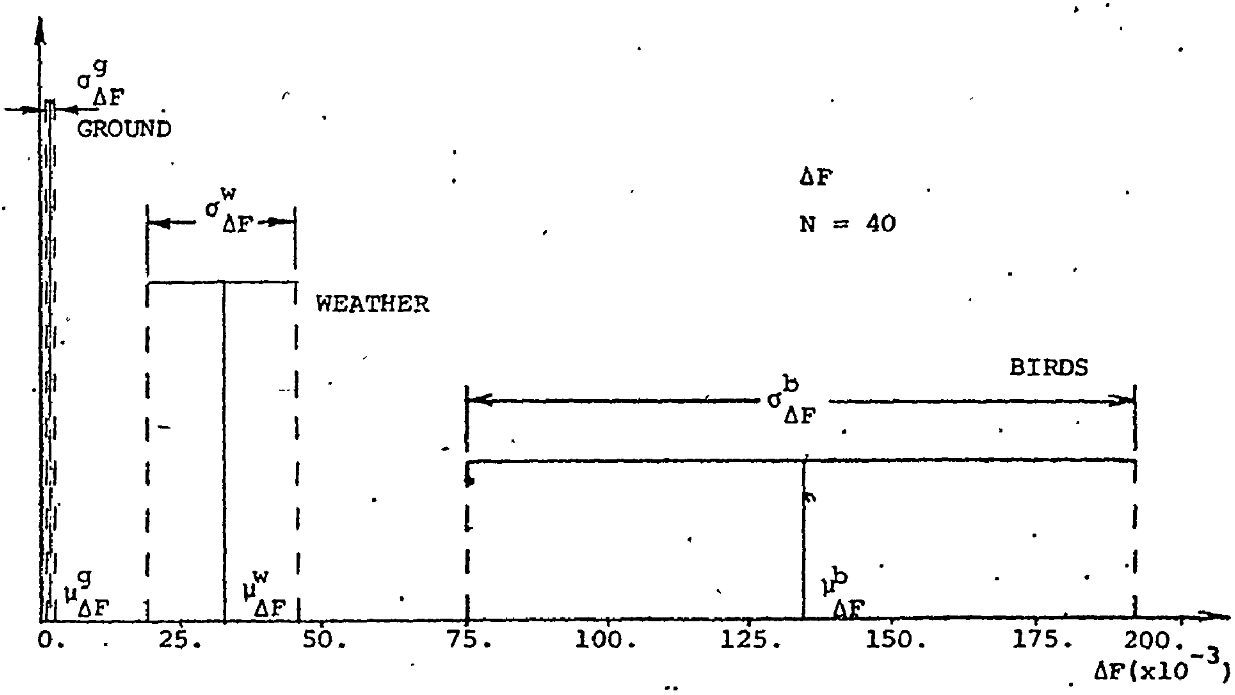


Fig. 5.5 - 12 Comparison of the statistical averages of ΔF for ground, weather, and birds; N = 40 .

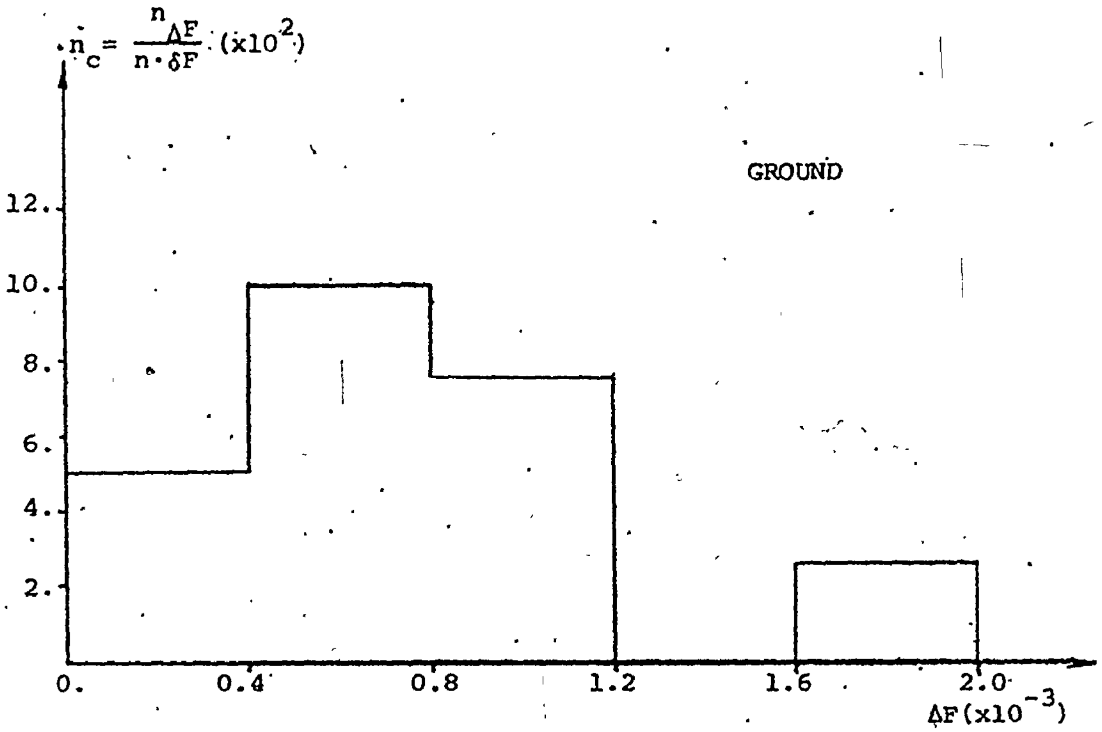


Fig. 5.5 - 13 Histogram for ground clutter spectral spread; N = 16.

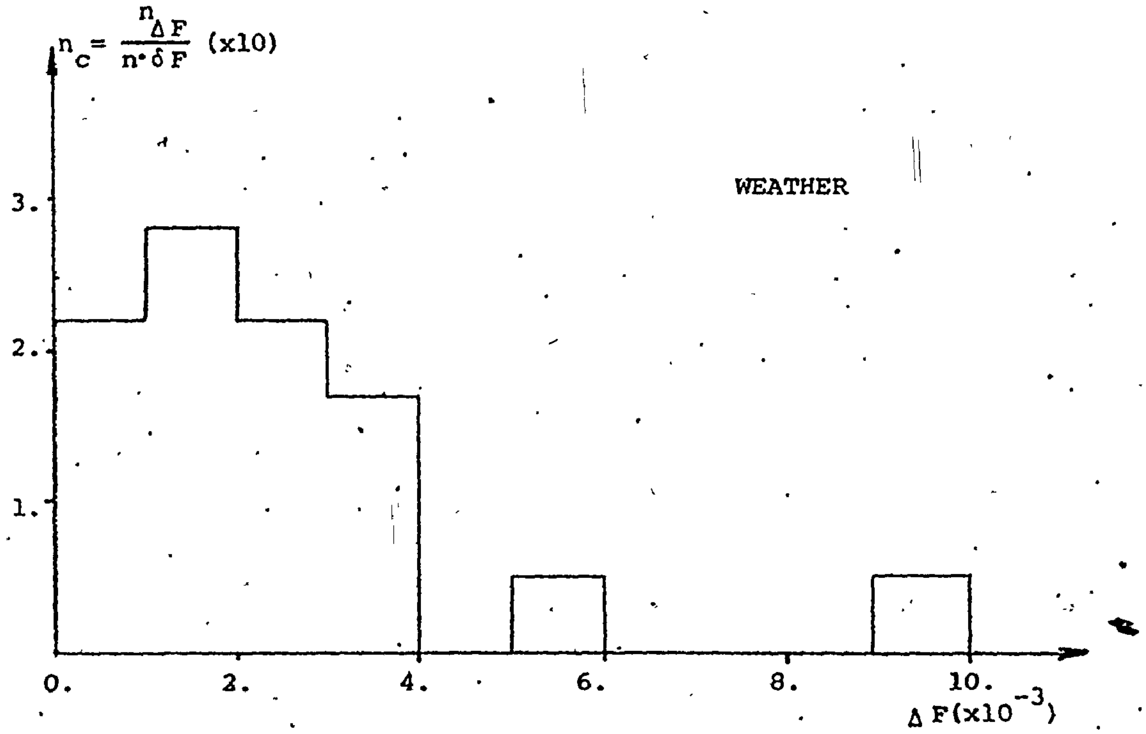


Fig. 5.5 - 14 Histogram for weather clutter spectral spread; N = 16.

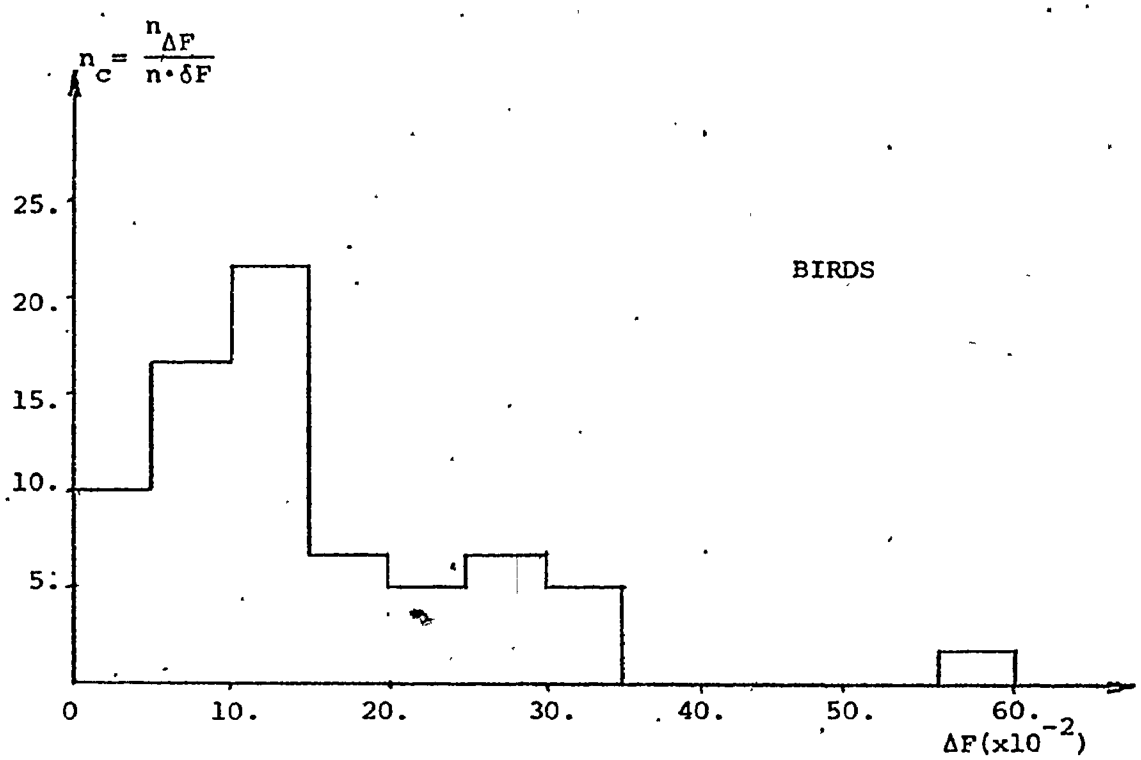


Fig. 5.5 - 15 Histogram for birds spectral spread; N = 16.

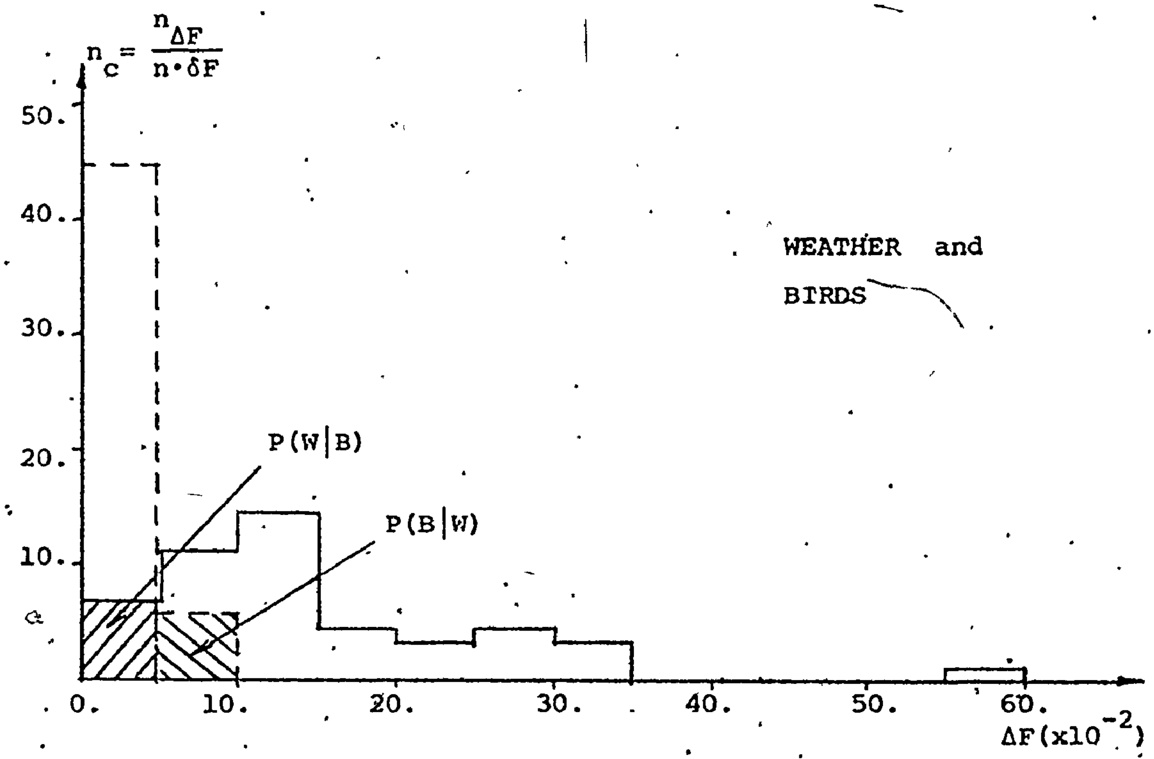


Fig. 5.5 - 16 Histograms for weather and birds spectral spreads; N = 16.

that there is practically no difficulty in separating ground clutter from the other two types. Second, the values of spread for ground clutter are an order of magnitude smaller than the corresponding values for weather and birds (Figure 5.5 - 11), and cannot therefore be represented meaningfully on the same scale. Referring to Figures 5.5 - 14, 5.5 - 15 and 5.5 - 16, we see that there is some overlap between the two curves corresponding to weather and birds. We may thus treat the problem as a simple binary hypothesis testing problem with a prescibed threshold. If the measured spectral spread is below the threshold we declare its origin as being from weather backscatter. If it is above the threshold, we assign it to birds. For the example in Figure 5.5 - 16 the obvious choice for threshold is $\Delta F_t = 5\mathrm{x}10^{-2}$. In this case, the probabilities of mistaking one signal for another are given by the hatched areas. The average probability of error is given by

$$\bar{P}_e = \frac{1}{2}\left(P(W|B) + P(B|W)\right) = 12.4\% \qquad (5.5.5)$$

where $P(W|B)$ is a probability of declaring weather when the actual signal comes from birds, and $P(B|W)$ is the probability of declaring birds when weather backscatter is present. Even in this simple example, the probability of error is rather small.

There are two more issues that should be mentioned. The first is the amount of information which can be extracted from the peak location F_o. To explore this issue we have drawn a number of curves based on the Table 5.5 - 1, similar to those in Figures 5.5 - 11 and 5.5 - 12. The results are given in Figures 5.5 - 17 (for $N = 16$) and 5.5 - 18 (for $N = 40$). We see that the means and standard deviations of F_o are, roughly,

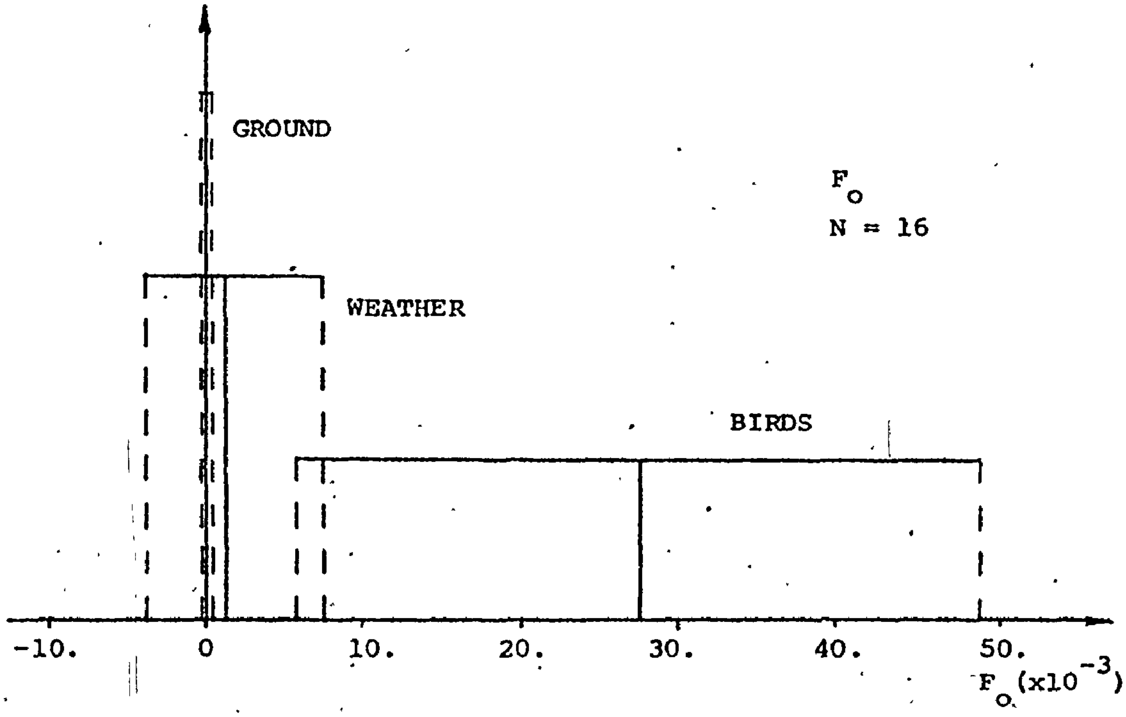


Fig. 5.5 - 17 Comparison of the statistical averages of F_o for ground, weather, and birds; N = 16 .

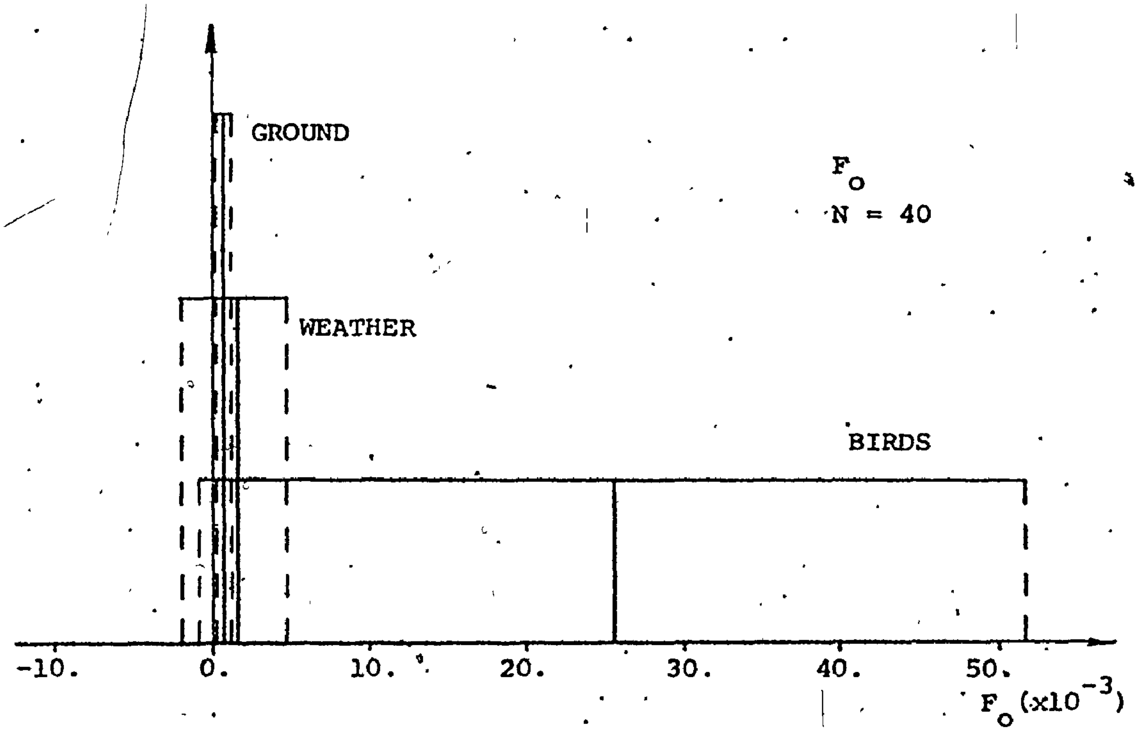


Fig. 5.5 - 18 Comparison of the statistical averages of F_o for ground, weather, and birds; N = 40 .

of the same order of magnitude as the corresponding values for ΔF. Also, the areas for different types of clutter overlap significantly. Thus, we see that clutter classification cannot be made on the basis of statistical analysis of the peak location. This agrees with the fact that the spectral peak location indicates the magnitude and direction of the radial component of scatterer velocities.

The second important issue is to examine whether there is any adventage to be gained by performing the spectral analysis of the video signal after MTI filtering. Since the original MTI filtered video was not available, a simple subroutine was written for a two-pulse canceller. The data were arranged in the order depicted in Fig. 5.2 - 1(a), i.e., all samples from one range ring, than those from the next one, etc. and then MTI filtering was applied. Examples of filtering ground clutter and birds are given in Figures 5.5 - 19 and 5.5 - 20, respectively. The spectra of both the original and filtered signals are shown. The effects of this filtering can be easily seen in both Figures. In particular, we see that while the low frequency components have been removed, the high frequency patterns have remained similar to those of the original signals. Since the spectral shape of the filtered signal depends more heavily on sidelobe patterns, the distinguishing of the spectra on the basis of the filtered signals is more difficult than on the basis of the original spectra.

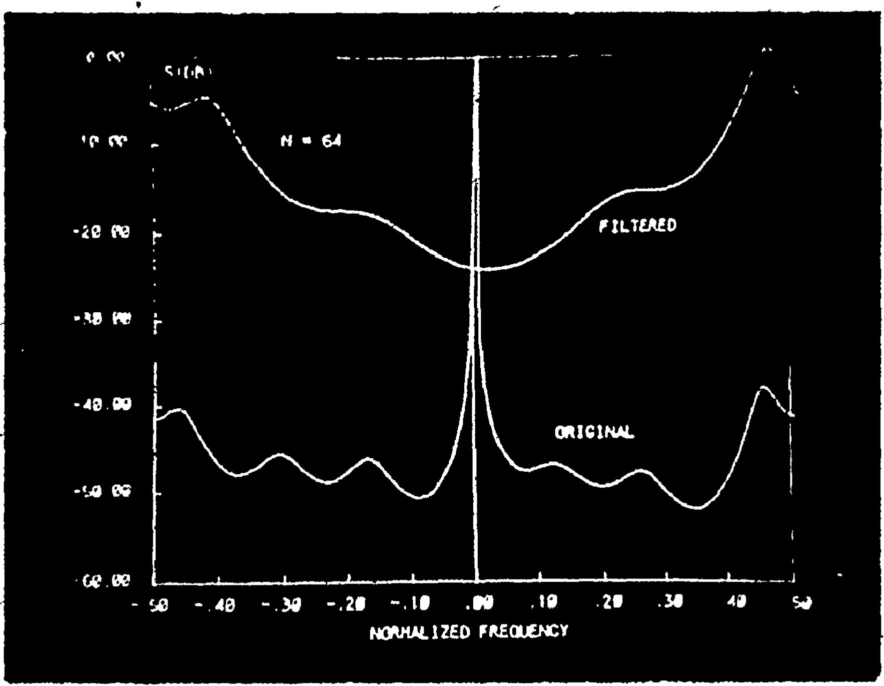


Fig. 5.5 - 19 ME spectral estimates of normal video and MTI filtered video, ground clutter, N = 64.

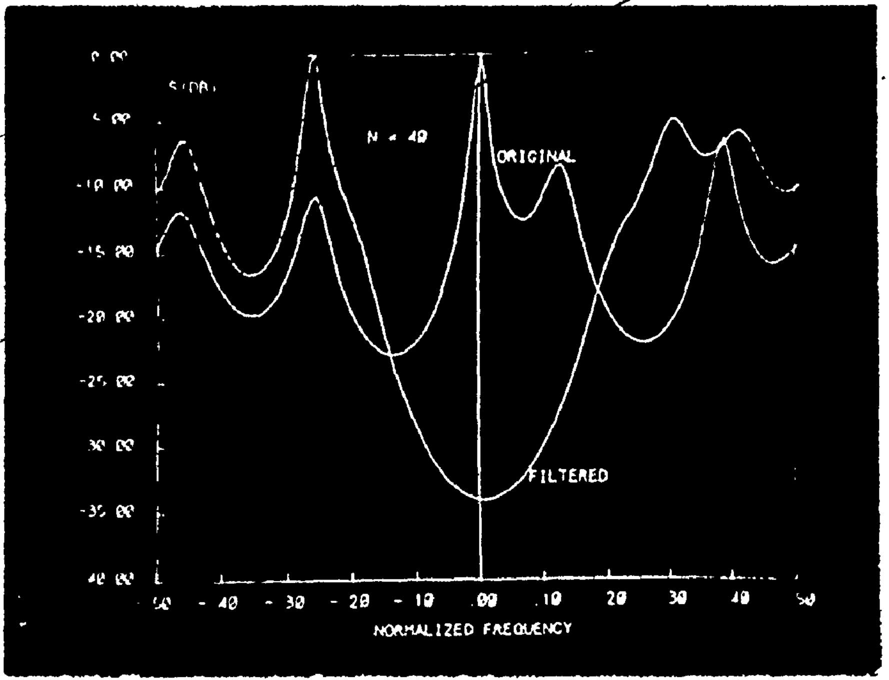


Fig. 5.5 - 20 ME spectral estimates of normal video and MTI filtered video, birds, N = 40.

5.6 Summary

In this chapter we have applied the maximum entropy method to the analysis of recorded radar clutter data. The results obtained confirm the resolution properties of the MEM, which were established in the previous chapter, using computer-simulated clutter.

The last part of the chapter dealt with a statistical analysis of the results. It was demonstrated that use of the MEM can significantly help in clussifying the different types of clutter by means of different spectral spreads.

CHAPTER 6

ADAPTIVE DIGITAL PREDICTIVE FILTERING BASED ON THE MEM

6.1 Introduction

The results of spectral analysis of clutter signal, reported in the previous chapters suggest two possible applications for prediction error filtering in coherent radar systems. The first one, mentioned in Chapter 5, is the classification of the different forms of clutter based on on-line spectral estimation. The second application is in the area of adaptive MTI filtering, based on the similar filter structure. In this chapter, after a brief review of adaptive MTI filtering, we shall present the basic idea of adaptive predictive filtering and its proposed usage in clutter cancellation. This approach to adaptive filtering has been studied by several investigators [101]-[108].

6.2 Review of Adaptive MTI Filtering

As mentioned in Chapter 1, modern radar systems use digital MTI filters almost exclusively. The advantages of digital filter over its analogue counterpart are greater hardware simplicity, and a high degree of accuracy and stability. The block diagram af a fully coherent single-delay digital MTI canceller is shown in Fig. 6.2 - 1. The local oscillator is divided into two paths such that there is a 90° differential phase shift at the two mixers whose video outputs form the I and Q channels. These signals are sampled and converted to digital form. The digital word

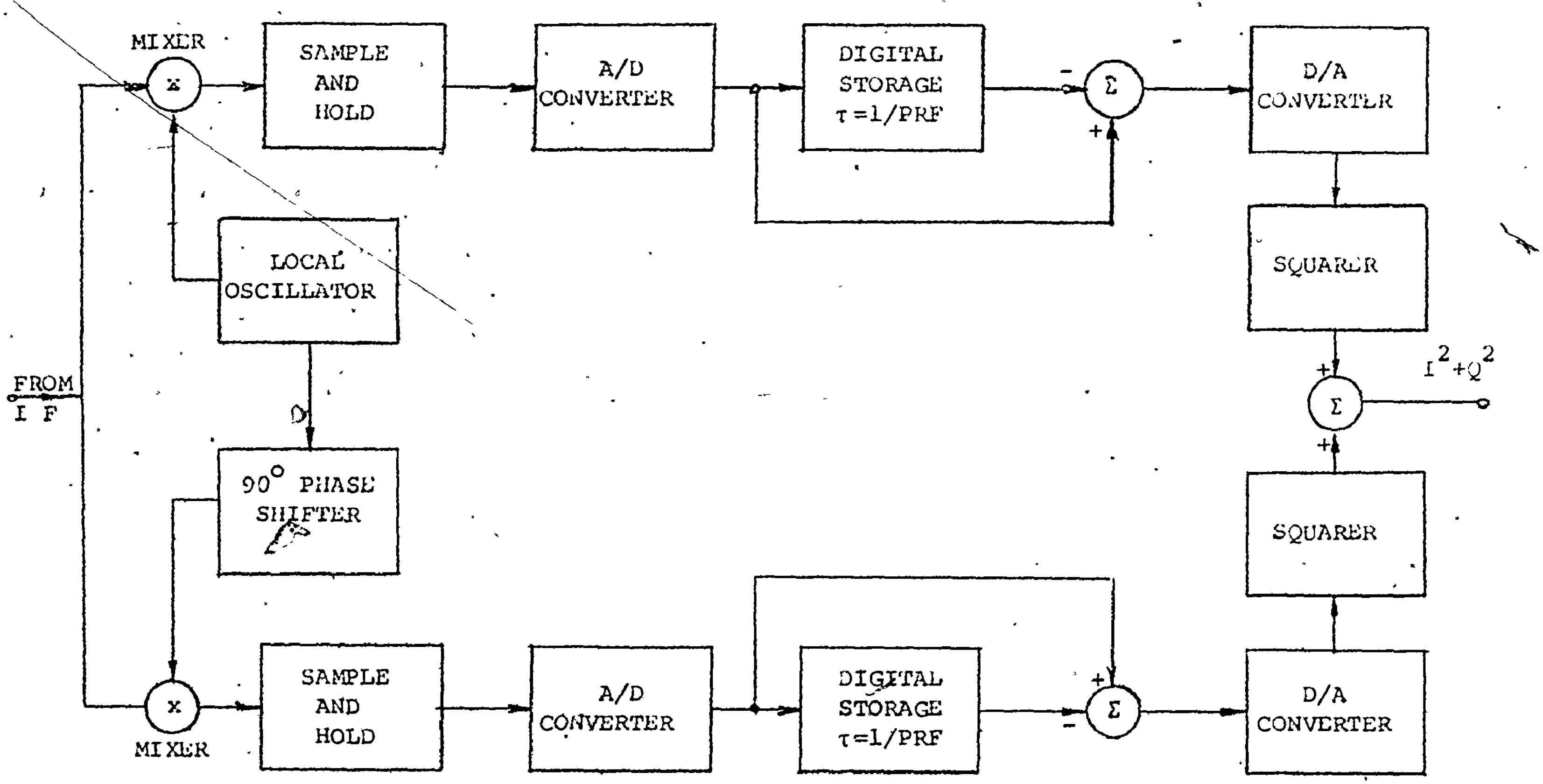


Fig. 6.2 - 1 Block diagram of a fully coherent single-delay MTI canceller.

for each range bin is held in storage until the echo from the second transmitted pulse (from the same range bin) is converted into digital form. The difference between the two signals is then converted back into analogue form and combined with the quadrature channel. Summation of the magnitudes of the I and Q residues is a good approximation to a square-law detector [22,Ch.14].

In a radar environment in which the clutter signal undergoes significant changes, for example, due to variations in environmental conditions, it is desirable to use a clutter suppression scheme which is adaptive in such a way that it is self-optimizing according to an appropriate criterion. The block diagram of an adaptive digital MTI filtering scheme is shown in Fig. 6.2 - 2. For convenience, the processing of only one channel is shown. The system consists of five basic components: (1) clutter parameter estimator, (2) coefficient calculator, which computes the optimum filter coefficients, (3) digital filter, (4) phase-error algorithm that adjusts the local oscillator and phase-shifter circuitry, and (5) automatic detector. Such a system is often called a closed-loop system because there exists a feedback path from the output of the digital filter to the input. Roy and Lowenschuss [109], and Hansen et. al. [110] have described an adaptive filter of this kind.

If the feedback is removed from the block diagram in Figure 6.2 -2, the system will be of an open loop type. In an MTI problems this type is usually preferred, because there is no way of ever knowing whether the filter selected is the optimum one, since there is no a priori knowledge about the existence of a target echo in the received radar signal. The open loop adaptive digital MTI filter is described in [111],

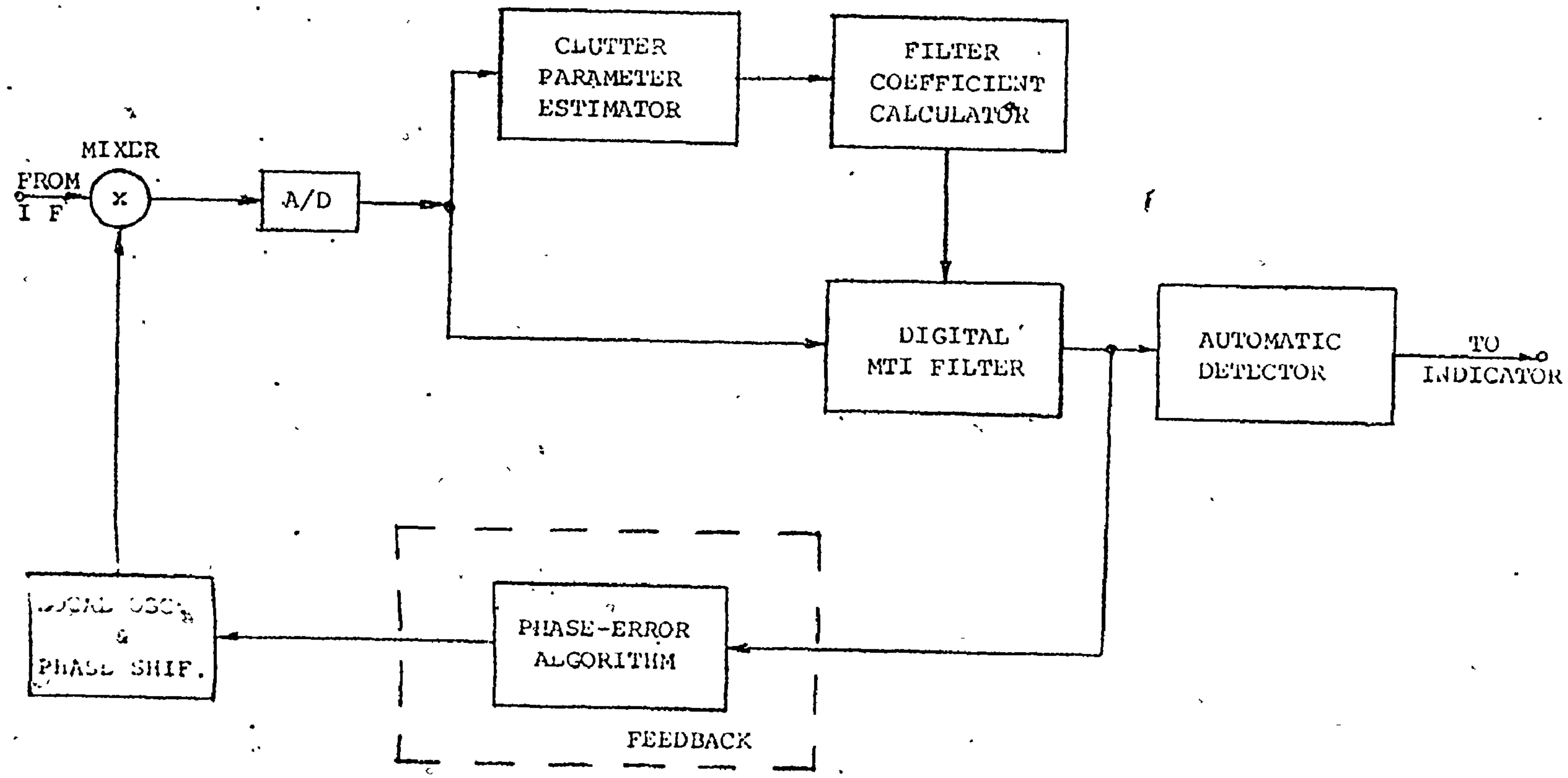


Fig. 6.2 - 2 A general block diagram of an adaptive MTI filtering.

where the difference phase shift is used as a criterion for determining a filter frequency response. A more practical performance measure, the false alarm probability, is used by Haykin and Hawkes [91],[92] as the criterion for adjusting the filter. In this scheme, an autocorrelator is used as the clutter-parameter estimator, while the filter coefficient calculator block consists of an arithmetic logic unit and read only memory (ROM) unit for selecting the set of optimum filter coefficients.

The ability of the adaptive MTI filter to suppress clutter is directly dependent on the statistical accuracy of the spectral density estimate obtained. Thus, the quality of the estimator in Fig. 6.2 - 2 has a direct influence on the system performance. On the other hand, when there are rapid changes in environmental conditions, the filter should adapt quickly to them. In this case, the clutter estimation should be performed with a relatively short clutter sample record obtained from a few scans. In extreme cases, e.g. thunderstorms with high wind, it may be necessary to have a short-term spectral estimate from one scan only. In such cases, adaptive predictive MTI filtering (to be described in the next 3 sections) seems to be a suitable processing scheme.

6.3 Clutter Whitening Filter

Consider the coherent video signal at the output of the phase detector in an MTI radar system. In general, this signal consists of three components: target echoes, clutter, and receiver noise. Let the power spectra of these components be denoted as T(f), C(f), and W(f) (= Const. = $N_o/2$), respectively. It is assumed that T(f) and C(f) occupy

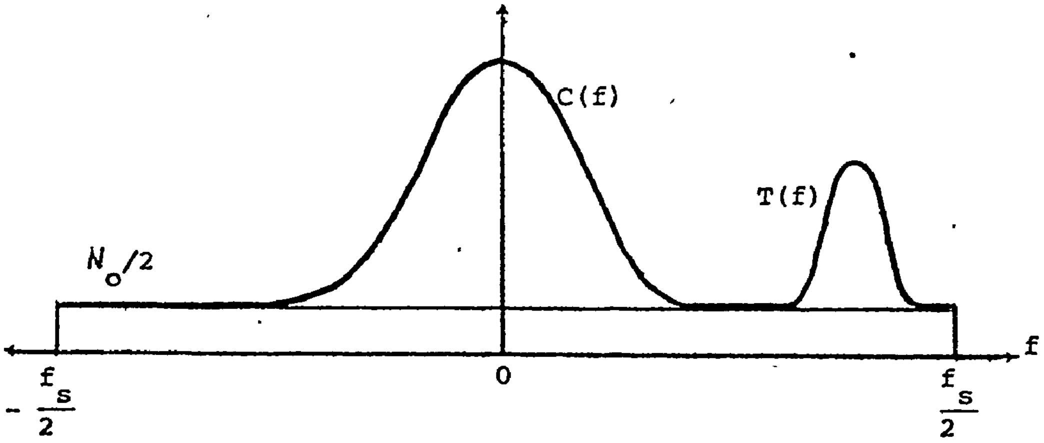


Fig. 6.3 - 1 Typical spectral pattern (within one sampling interval) of a radar video.

non-overlapping frequency intervals. A typical spectral pattern is sketched in Figure 6.3 - 1. If we pass the video signal through a finite impulse response (FIR) digital filter with amplitude response H(f) (Fig. 6.3 - 2), the power spectrum of the resulting output signal is given by

$$Y(f) = |H(f)|^2 [T(f) + C(f) + N_o/2] . \qquad (6.3.1)$$

Now, suppose that, during the process of filtering, we can somehow automatically adjust the filter coefficients in a way that the squared magnitude of the amplitude response H(f) becomes

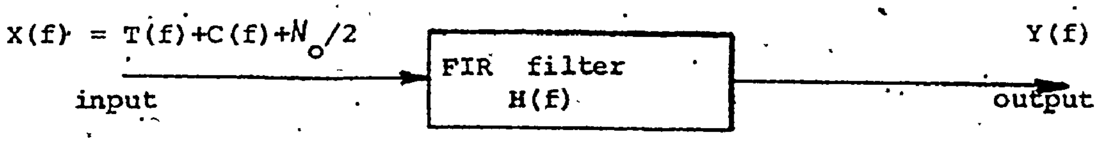


Fig. 6.3 - 2 FIR filtering of a coherent video signal.

$$|H(f)|^2 = \frac{K}{N_o/2 + C(f)} , \qquad (6.3.2)$$

where K is a constant. We shall describe the process of filter adjustment later. Combining Eqs. (6.3.1) and (6.3.2) we now have for the output power spectrum:

$$Y(f) = \frac{K}{N_o/2 + C(f)} \cdot [T(f) + C(f) + N_o/2]$$

$$= K \cdot \left[1 + \frac{T(f)}{N_o/2 + C(f)} \right] . \qquad (6.3.3)$$

In the low frequency region where T(f) = 0, the output spectrum is constant and is approximately equal to K. In the region corresponding to target frequencies, C(f) = 0 and the output spectrum is

$$Y(f) = K \cdot \left[1 + \frac{2T(f)}{N_o} \right] . \qquad (6.3.4)$$

Except for the additive constant, the output spectrum is a scaled version of the target spectrum. When there is no target, the output spectrum will be flat, that is, Y(f) = K. The value of K is determined by normalizing the value $|H(0)|^2$ to unity. From (6.3.2) we have

$$K = N_o/2 + C(0) . \qquad (6.3.5)$$

From the above considerations, it is clear that the filter H(f) is an inverse filter with respect to the clutter component of the video signal, i.e., a clutter whitening filter (since the spectral density

$N_o/2$ is a constant, it only determines a pedestal for the clutter spectrum). This means that the clutter signal is constantly being whitened, while the target is being enhanced. In the absence of a target the output signal has a white spectrum; such a situation is used for "training" the adaptive algorithm, i.e., for adjusting the filter coefficients. There is usually enough time for adaptation, since the time on target is in most cases only a fraction of the time for antenna revolution.

By using the white noise spectral level as an additive component in defining the inverse filter amplitude response, as in Eq. (6.3.2), we avoid numerical instabilities in performing calculations based on the equations (6.3.3) and (6.3.4), when the level of the slutter spectrum is extremly low. This may be a common situation in the region of target frequencies.

In the following section we describe the adaptive procedure for adjusting filter coefficients.

6.4 Adaptive Procedure

As shown in Chapter 3, we can, in most cases, model clutter as being in the class of all-pole (autoregressive) signals. That is, the power spectrum may be represented in the form

$$\hat{C}(f) = \frac{K_c}{\left|a_o + a_1 e^{-j2\pi f} + \ldots + a_M e^{-j2\pi Mf}\right|^2}\ , \quad a_o \equiv 1.0. \qquad (6.4.1)$$

The error between the all-pole representation, $\hat{C}(f)$, and the true value $C(f)$ can be made arbitrarily small by making the order of the process, M, large enough. It has been demonstrated, however, that in most practical situations, M does not exceed 10. Assuming that $N_o/2 << K_c$ we can see from Eqs. (6.3.2) and (6.4.1), that the inverse filter contains zeros only, which is an FIR filter.

Referring to the block diagrams of the forward-and backward-prediction error filters in Figures 2.4 - 1(a) and (b), we find that their actual FIR structures are as shown in Figures 6.4 - 1(a) and (b). These two configurations can be combined in the form of a ladder filter structure as shown in Figure 6.4 - 2. The output of the forward filter is also the actual filter output. Note that the filter coefficients in Figure 6.4 - 2 represent the coefficients $\tilde{a}_{m,m}$, m = 1, 2, ... , M, given by Eq. (2.4.11) rather than the coefficients $a_{M,m}$, m = 1, 2, ... , M, for the Mth order filter, given by Eqs. (2.4.12). This means that all lower order filtered outputs are at all time available as well as the Mth order prediction error series. By displaying these coefficients simultaneuously we obtain valuable information for determining the length or number of ladder stages to be used. Equivalence between the two structures given in Figures 6.4 - 1 and 6.4 - 2 is proved in Appendix F.

The performance measure of the adaptive filter is an average forward and backward output power at each stage of the ladder. As shown in Chapter 2, the power for the mth stage is given by

$$P_{m+1} = \frac{1}{2(N-m)} \sum_{n=1}^{N-m} \{ |y_m^f(n) + a_{m,m} y_m^b(n)|^2 + |y_m^b(n) + a^*_{m,m} y_m^f(n)|^2 \} , \qquad (6.4.2)$$

where $y_m^f(n)$ and $y_m^b(n)$ are, respectively, the forward and the backward

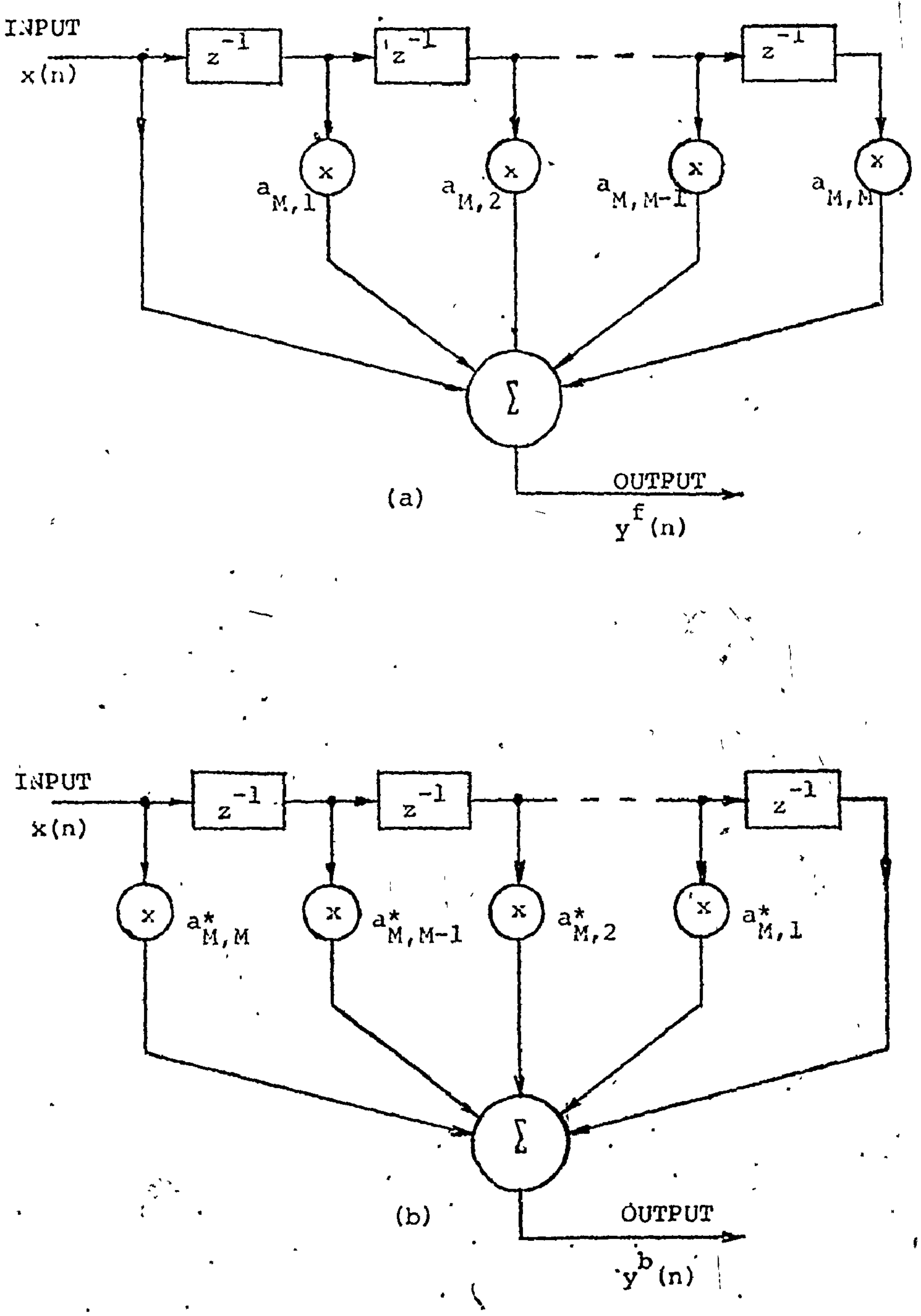


Fig. 6.4 - 1 Forward (a), and backward (b) FIR prediction error filter.

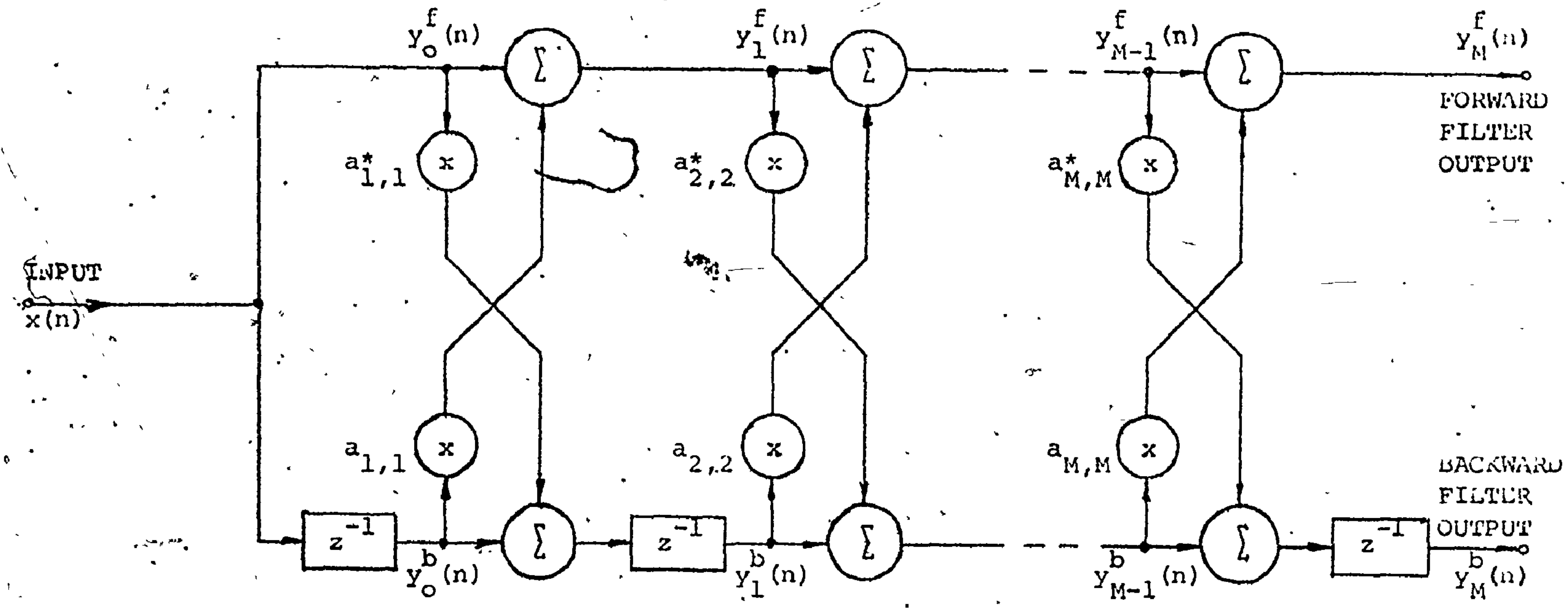


Fig. 6.4 - 2 Prediction error filter that combines forward and backward filtering.

error series at the output of mth stage, and $m = 0,1,2,\ldots,M$. The criterion of adaptation is based on minimization of the average output power, (6.4.2), at each stage with respect to the filter coefficients. If the average value of the squared-output of each stage is minimized, then the total average output power of the system as a whole will also be minimized. The value of $a_{m,m}$ which minimizes P_{m+1} is given by

$$a_{m,m} = \frac{-2 \sum_{n=1}^{N-m} y_m^{f*}(n) \cdot y_m^{b}(n)}{\sum_{n=1}^{N-m} \{ |y_m^{f}(n)|^2 + |y_m^{b}(n)|^2 \}} . \qquad (6.4.3)$$

From Eq. (6.4.3) it follows that $|a_{m,m}| < 1$. The numerator on the right hand side of Eq. (6.4.3) is the negative cross-power and the denominator is the averaged auto-powers of the m point forward and backward prediction errors. The summation will be taken over past values of the numerator and denominator. In effect, we are training the system to respond to what happened in the past. Since we may expect changes to occur in the input signal, the statistics close to the point we are trying to predict would be weighted more heavily than those further away. Then, as the predictable portion of the input signal changes the change will be quickly noted and incorporated into the filter adjustments.

A reasonable way of weighting the coefficients is to multiply each term with the exponential function of the form

$$w(n) = e^{-(n-1)/\tau} , \quad n = 1,2,\ldots,N-m , \qquad (6.4.4)$$

where the relaxation time of the adaptation, τ, is defined as 1/e times the decay time along the time axis. Thus, the weighted optimal coefficients are given by

$$a^w_{m,m} = \frac{-2\sum_{n=1}^{N-m} w^2(n)\cdot y^{f*}_m(n)\cdot y^b_m(n)}{\sum_{n=1}^{N-m}\{|y^f_m(n)|^2 + |y^b_m(n)|^2\}} \quad . \qquad (6.4.5)$$

The close-in statistics used to compute the filter are more heavily weighted against those further away which are gradually "forgotten". The $a^w_{m,m}$'s using the form given by Eq. (6.4.5) are conveniently computed from the current and immediate past values of the numerator and denominator by simple feedback. As the filtering proceeds, the time window moves along with the filter. The data are continuously processed and the filter is continuously updated. The result is that the PEF adapts to continually minimize the weighted average value of the output power.

The choice of the relaxation time is governed by a trade-off between the speed of adaptation and a reliable estimation of the filter coefficients. The range of values for τ should be

$$1/f_{t\ell} < \tau < 1/f_{ch} \quad . \qquad (6.4.6)$$

where f_{ch} is the cutoff frequency of the clutter spectrum and $f_{t\ell}$ is the lower cutoff frequency of target spectrum.

6.5 Configuration of an Adaptive MTI Filter

In this section we present a complete block diagram of the adaptive filter based on the theory presented above. Figure 6.5 - 1 shows the implementation of the least mean-square prediction error algorithm (LMSPEA) given by Eq. (6.4.5). The blocks indicated by an asterisk denote complex conjugation operations. Division of the numerator, N, by the denominator, D, is indicated by the block N/D, while the circuit in the feedback loop serves as an accumulator of the successive values. The attenuation, K, is inserted in the loop to avoid instability. The output of LMSPEA is the optimal value of $a_{m,m}$ and its conjugate. The block diagram of the complete adaptive filter in complex form is shown in Figure 6.5 - 2. Each block labeled LMSPEA is of identical structure, shown in Fig. 6.5 - 1. The number in parenthesis indicates the order of a stage.

With such an adaptive filter we obtain the minimum possible value for the average error power at the output of the overall filter. Availability of control at each stage gives valuable information for selecting the filter order. This is achieved, however, at the expense of relatively complicated circuitry, since we require M adjusting circuits for a filter of order M. This does not seem to be a problem, first, because all the LMSPEA circuits are identical, and second, we need a significantly lower filter order than other adaptive circuits based on FIR filtering [26].

In an operating environment involving a coherent MTI radar, we are given the I and Q components of the video signal at the phase detector output. We shall denote these components as well as the real and

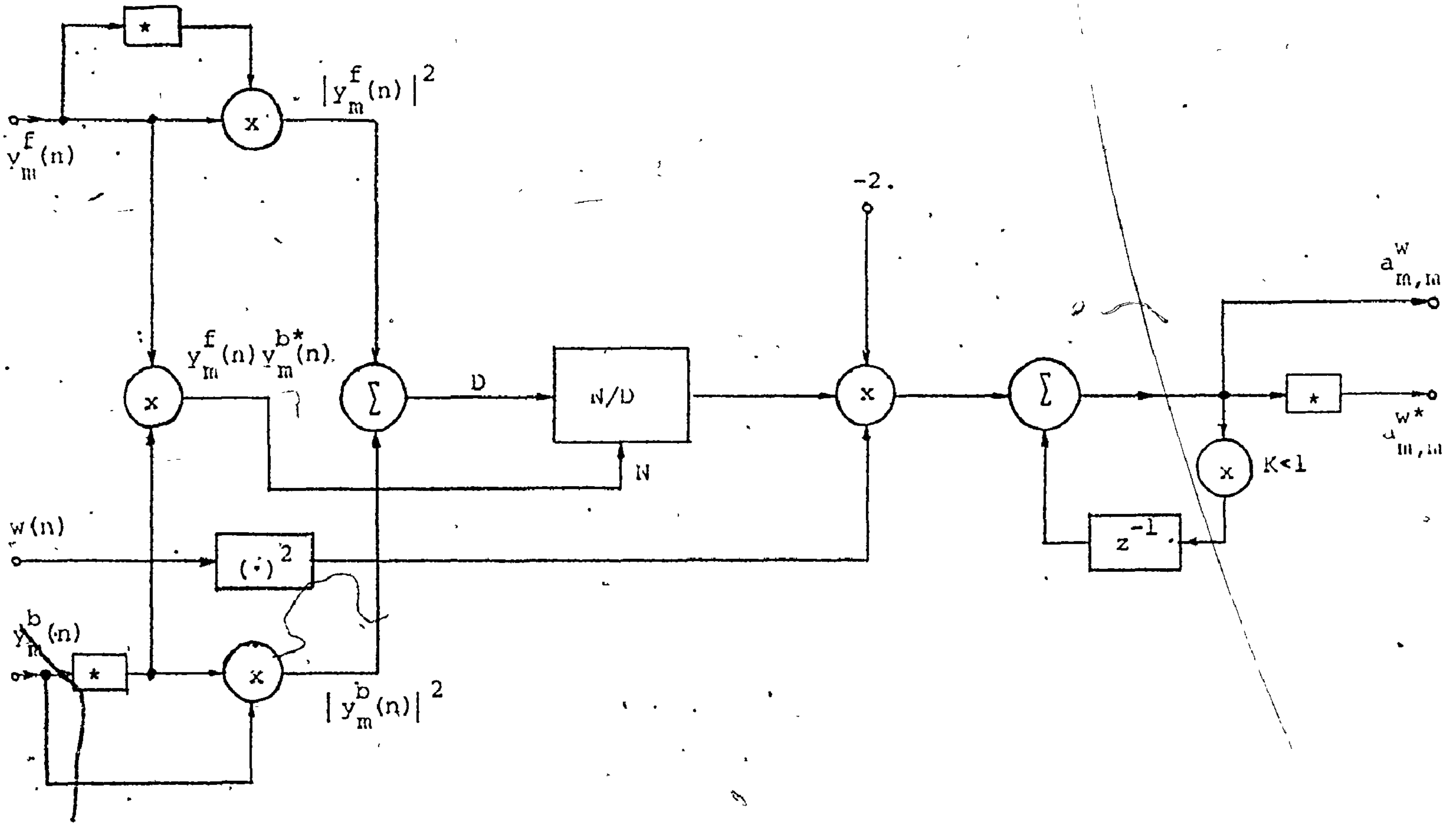


Fig. 6.5 - 1 Implementation of the LMSPEA, Eq. (6.4.5).

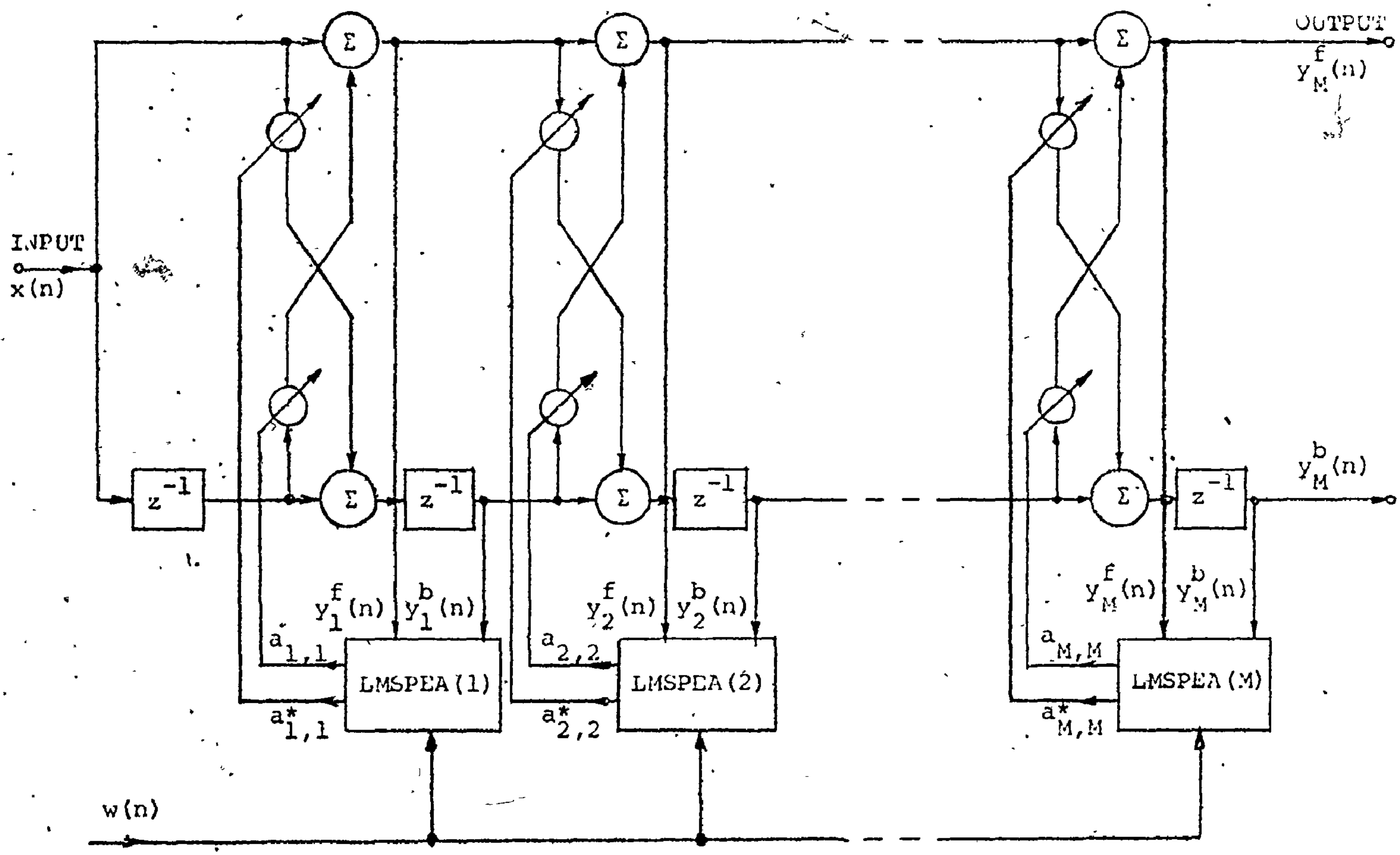


Fig. 6.5 - 2 Block diagram realization of an adaptive MTI filter in a complex form.

imaginary components of $a_{m,m}$'s with additional subscripts (r) and (i), respectively. The same holds for the forward-and backward-output signal components. Based on this notation, the real and imaginary parts of Eq. (6.4.5) are as follows, respectively:

$$a^w_{m,m,(r)} = -2 \sum_{n=1}^{N-m} \frac{w^2(n)\,[y^f_{m,(r)}(n)\,y^b_{m,(r)}(n)+y^f_{m,(i)}(n)\,y^b_{m,(i)}(n)]}{|D_m(n)|^2} \qquad (6.5.1(a))$$

$$a^w_{m,m,(i)} = 2 \sum_{n=1}^{N-m} \frac{w^2(n)\,[y^f_{m,(r)}(n)\,y^b_{m,(i)}(n)-y^f_{m,(i)}(n)\,y^b_{m,(r)}(n)]}{|D_m(n)|^2} \qquad (6.5.1(b))$$

where

$$|D_m(n)|^2 \triangleq \left(y^f_{m,(r)}(n)\right)^2+\left(y^f_{n,(i)}(n)\right)^2+\left(y^b_{m,(r)}(n)\right)^2+\left(y^b_{m,(i)}(n)\right)^2 \qquad (6.5.2)$$

When the filter in Fig. 6.5 - 2 is represented in terms of the real and imaginary parts of the input signal and those of the filter coefficients it has two inputs and two pairs of outputs, namely, the real and imaginary parts of the forward and backward prediction error series. For the sake of simplicity, only the first stage is shown in Fig. 6.3 - 3. The following stages are simply cascaded at the corresponding output points. Also, to simplify the notation, the indices (1,1) have been omitted in the subscripts for real and imaginary components of the coefficient $a_{1,1}$. Although the structures shown in Figures 6.4 - 3 and 6.4 - 4 may appear complicated, their realization is quite straightforward.

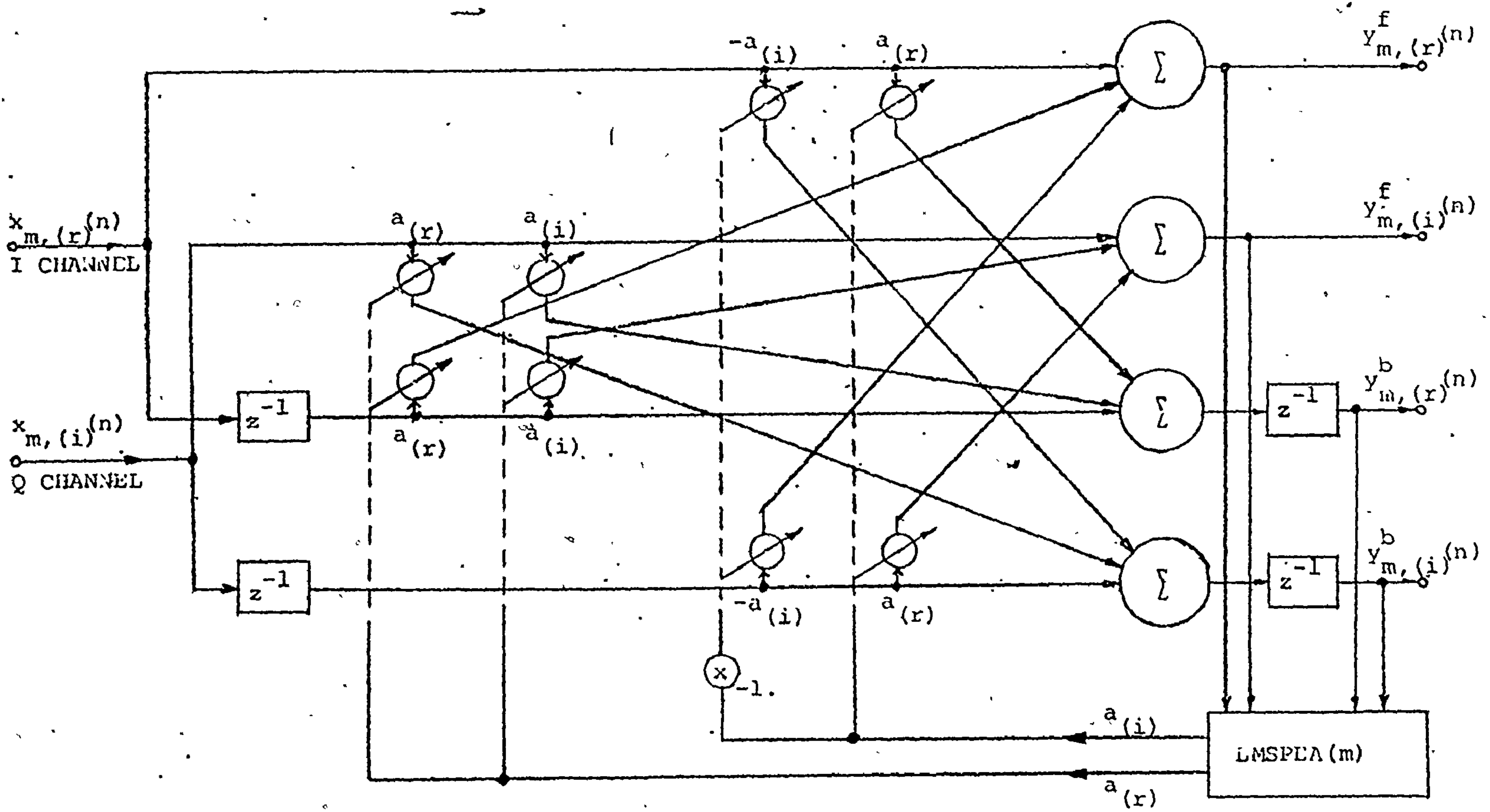


Fig. 6.5 - 3 Filter from Fig. 6.5 - 2 split into "real" and "imaginary" channels (one stage only).

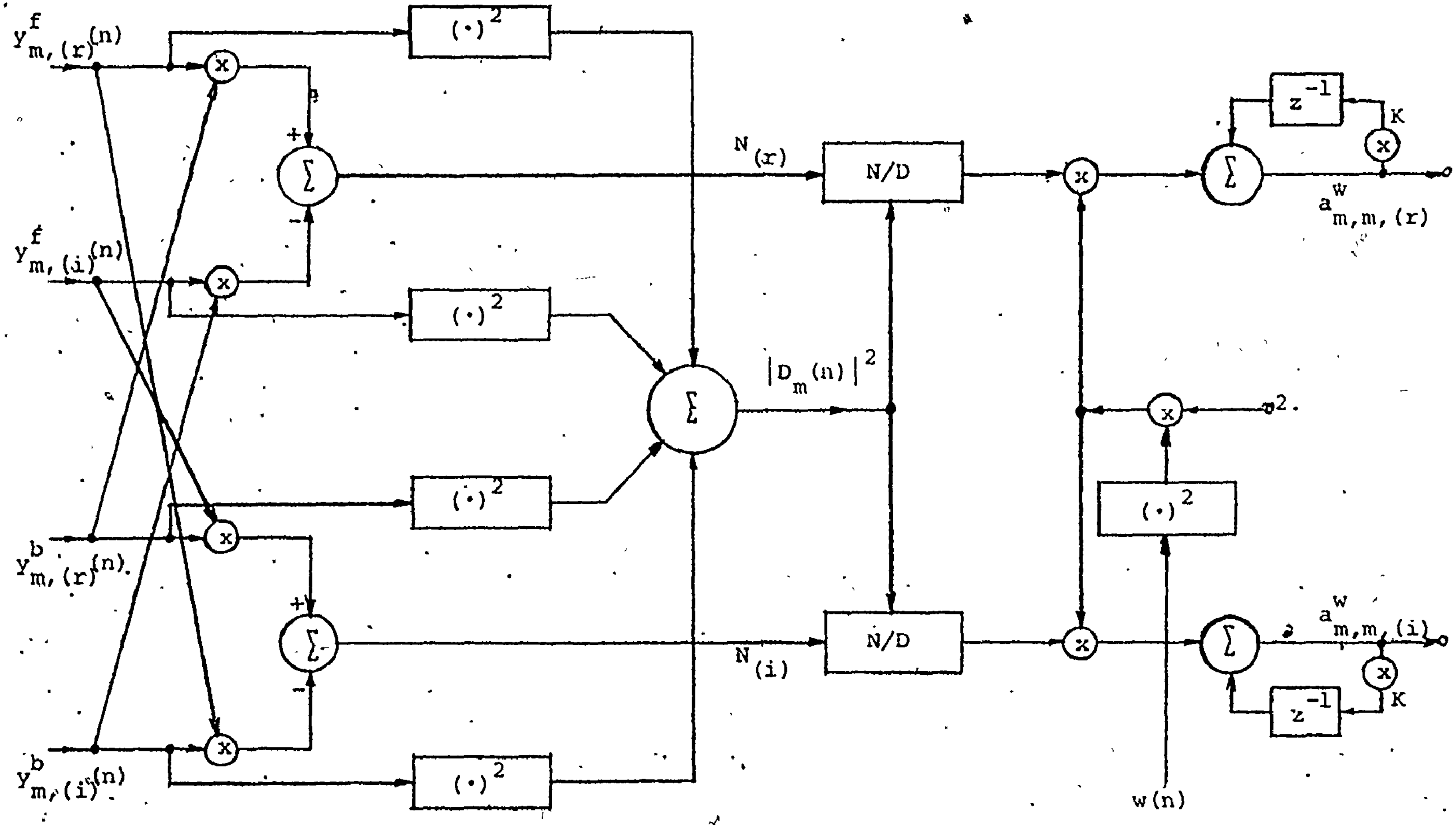


Fig. 6.5 - 4 Implementation of the LMSPEA in terms of real components, Eq. (6.5.1).

We may summarize the important features of the adaptive MTI filtering procedure described above as follows:

(1) The filtering process is continuous, with the filter coefficients continuously updated while the data is being filtered.

(2) Since the filter coefficients estimated are always less than unity, the system is always stable regardless of the rate of adaptation. Also, the filter is always minimum -phase.

(3) The system has a capability for real-time operation. The stages in the ladder could easily be built as discrete hardware elements and cascaded to form a filter of any desired length.

(4) All lower-order filtered outputs are at all time available as well as the Mth order error signals. Displaying these simultaneuously provides valuable information with regard to filter order and number of stages to be used.

(5) For a prescribed filter order, the input signal spectrum can be calculated using the values of the average output power and the estimated coefficients. This provides a means for on-line classification of the input clutter component, as discussed previously.

The performance of the basic structure of the filter ladder, only with real coefficients, is illustrated in Figure 6.5 - 5 taken from [101]. Autoregressive process of fourth order was generated with the known parameters. The first half of the time series was produced with one set of values for the $a_{m,m}$'s and the other half with another set of values. The spectra of both halves of the input signal are given in Fig. 6.5 - 5(a). The preselected values for the $a_{m,m}$'s are shown by the dashed lines in

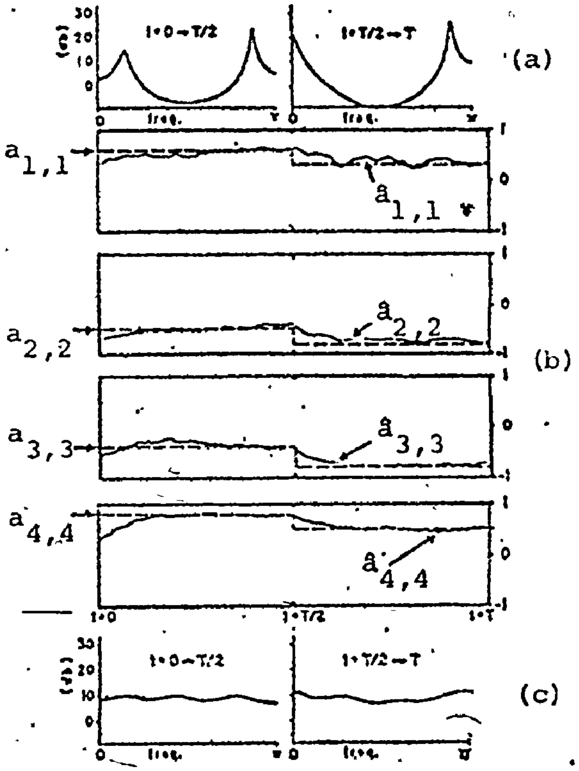


Fig. 6.5 - 5 Adaptation of fourth order filter coefficients to a step discontinuity in signal parameters: (a) spectra of the first and the second half of the input signal, (b) filter coefficients adaptation, (c) spectra of the output prediction error series for both parts of the signal.

in Figure 6.5 - 5(b), for m = 1,2,3, and 4. The solid lines represent the estimated values, $\hat{a}_{m,m}$'s. As the filtering operation proceeds, we see that the estimated values soon track the true values. Then, the predetermined signal parameters abruptly change. The adaptive filter soon adjusts itself to the new environment and the estimated coefficients begin to move in the direction of the true values. Note that the exponential behavior of the estimated coefficients following the step generally obeys a decay rate corresponding to the relaxation time set for adaptation. The diagrams in Figure 6.5 - 5(c) show the spectra of the output prediction error series for both parts of the input signal. It is evident that the spectra have been significantly whitened by the filter.

6.6 Summary

This chapter has dealt with a proposal for adaptive MTI filtering based on the PEF. It has been shown that a simple open loop adaptive procedure can be designed for clutter rejection for use in a coherent pulse radar. An important feature of the filter is the possibility of continual adaptation of the filter and, at the same time, on-line spectral estimation, which provides the means for discrimination between the different types of the clutter.

CHAPTER 7

SUMMARY AND RECOMMENDATIONS

7.1 Summary

In this thesis we have developed system for on-line classification of radar clutter, based on the maximum entropy method. The separation between weather and bird echoes is made possible by a prediction error filtering technique which exhibits excellent resolution capabilities. A statistical analysis of computer-simulated and recorded coherent radar signals have confirmed the method's ability to resolve the spectra of ground clutter, weather clutter, and bird echoes even with very short data record lengths.

Based on the prediction error filtering or, equivalently, the maximum entropy estimation of power spectral density, an adaptive filtering scheme for suppression of nonstationary clutter signals, has been developed. Thus, it is possible to classify the signal while removing it from the PPI display.

The important contributions of the thesis are summarized below:

(1) The complex form of the algorithm for calculating the PEF coefficients, which is used for processing the I and Q components of video signal in coherent pulse radar, has been derived.

(2) The adequacy of an all-pole representation of clutter has been demonstrated.

(3). Using computer-simulated clutter data, it has been demonstrated that small variations in the clutter-model parameters are detectable by using the maximum entropy method of spectral analysis.

(4) Detailed comparison with the best known linear method (the Welch method) of spectral analysis has revealed the superiority of the MEM in respect to record length and resolution. The computation times for both methods are approximately of the same order of magnitude.

(5) Using recorded clutter data, it has been verified that the separation of different forms of clutter is possible by using the ME analysis of the spectral spread. Statistical analysis of different clutters' spreads has shown that one is able to distinguish between ground, weather, and birds returns with a relatively low probability of error. Again, comparison with Welch's method has shown that the latter lacks the necessary resolution for classification.

(6) A simple open-loop adaptive MTI filtering procedure, based on prediction error power minimization, has been derived. The adaptive scheme is fast, absolutely stable, and relatively simple to implement.

7.2 Recommendations for Further Research

(1) In order to provide a more complete statistical analysis of the clutter-classification procedure, it is necessary to collect the data from various radar sites. In particular, data from the high-density bird migration corridors of Canada should be recorded. The analysis of both coherent and noncoherent video returns would also be of interest.

(2) A more thorough statistical analysis of the different clutter returns should be performed for different environmental conditions.

(3) The statistical theory of adaptivity should be evaluated for both simulated and actual radar data.

(4) Since the adaptive MTI filtering scheme, together with the filter coefficients adjusting circuits, is relatively simple, it would be of interest to implement it in hardware form and test it.

REFERENCES

[1] Blackman, R.B., and Tukey, J.W., "The Measurements of Power Spectra from the Point of View of Communications Engineering", Dover Publications, Inc., New York, N.Y., 1959.

[2] Bartlett, M.S., "An Introduction to Stohastic Processes with Special Reference to Methods and Applications", Cambridge University Press, Cambridge, 1953.

[3] Jones, R.H., "A Reappraisal of the Periodogram in Spectral Analysis", Technometrics, vol.7, no.4, pp.531-542, Nov. 1965.

[4] Welch, P.D., "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on the Time Averaging Over Short, Modified Periodograms", IEEE Trans.Audio.Electroacoust., vol.AU-15, no.3, pp.70-73, June 1967.

[5] Burg, J.P., "Maximum Entropy Spectral Analysis", paper presented at the 37th Annual International Meeting, Soc. of Explor. Geophys., Oklahoma City, Okla., Oct. 31, 1967.

[6] Battan, L.J., "Radar Observation of the Atmosphere", The University of Chicago Press, 1973.

[7] Harger, R.O., "Synthetic Aperture Radar Systems; Theory and Design", Academic Press, New York, N.Y., 1970.

[8] Learch, C.S., Jr., "Phased Array Radar for Satellite Tracking", 1962 IRE Natl. Conv. vol.10, pp.50-57, 1962.

[9] Ward, K.J., "Techniques for Radar Speed Detection", Traffic Dig. Rev., Sept., Oct., Nov., 1966; reprint by Muniquip Corp., Elk Grove Village, Ill.

[10] Evans, J.V., and Hagfors, T., (eds.) "Radar Astronomy", McGraw-Hill Book Company, New York, N.Y., 1968.

[11] Atlas, D., "Advances in Radar Metereology", in H.E. Landsberg and J. van Mieghen (eds.): "Advances in Geophysics", vol.10, pp.318-468, Academic Press, Inc., New York, N.Y., 1964.

[12] Eastwood, E., "Radar Ornithology", Methuen and Co., 1967.

[13] Marcum, J.I., "A Statistical Theory of Target Detection by Pulsed Radar", IRE Trans.Information Thery, vol.IT-6, no.2, pp.59-267, April 1960.

References Continued

[14] Swerling, P., "Probability of Detection for Fluctuacting Targets", IRE Trans.Information Theory, vol.IT-6, no.2, pp.271-308, April 1960.

[15] Swerling, P., "Recent Development in Target Models for Radar Detection Analysis", AGARD Conf. Proc., no.66, pp.7-1:7-12, May, 1970.

[16] Kanter, I., "A Generalization of the Detection Theory by Swerling", EASCON 74, pp.198-205, Oct. 1974.

[17] Haykin, S.S., (ed.) "Detection and Estimation, Application to Radar", Benchmark Papers in Electrical Engineering and Computer Science, vol.13, Dowden, Hutchinson & Ross, Inc., 1976.

[18] Dodsworth, E.J., and Hathaway, J., "Computer Simulation of MTI", Royal Radar Establishment Technical Note no.759, 1970.

[19] Siegert, J.F., "On the Fluctuation Signals Returned by Many Independent Scatterers", MIT Lincoln Laboratory Report no.465, Nov. 12, 1943.

[20] Barlow, E.J., "Doppler Radar", Proc. IRE, vol.37, pp.340-355, April 1949.

[21] Goldstein, H., "Sea Echo, the Origins of Echo Fluctuations, and the Fluctuations of Clutter Echoes", in D.E. Kerr (ed.), Propagation of Short Radio Waves, MIT Radiation Laboratory Series, vol. 13, secs.6.6-6.21, pp.560-587, McGraw-Hill Book Company, New York, N.Y., 1951.

[22] Nathanson, F.E., "Radar Design Principles", McGraw-Hill Book Company, New York, N.Y., 1969.

[23] Skolnik, M.I., "Radar Handbook", McGraw-Hill Book Company, New York, N.Y., 1970.

[24] Barton, D.K., "Radar Systems Analysis", Prentice-Hall. Inc., Englewood Cliffs, N.J., 1964.

[25] Fishbein, W., Graveline, S.W., and Rittenbach, O.E., "Clutter Attenuation Analysis", Tech. Rep. ECOM-2808, March 1967.

[26] Hawkes, C.D., "Adaptive Digital Filtering for Coherent MTI Radar", Ph.D. Thesis, McMaster University, Hamilton, Ontario, April 1975.

References Continued

[27] Wong, J.L.Y. "Scattering by Randomly Varying Media with Application to Radar", Ph.D. Thesis, University of Southern California, Los Angeles, Calif., 1968.

[28] Wong, J.L.Y., Reed, I.S., and Kaprielian, Z.A., "Model for the Radar Echo from a Random Collection of Rotating Dipole Scatterers", IEEE Trans. Aerosp. Electron. Systems, vol.AES-3, no.2, pp.171-178, March 1967.

[29] Berkowitz, R.S., (ed.), "Modern Radar - Analysis, Evaluation and System Design", John Wiley & Sons, Inc., New York, N.Y., 1965.

[30] Nitzberg, R., "Effect of Oscillator Instability upon MTI Improvement Factor", General Electric Report no.R72EMH23, 11/8/1972.

[31] Skolnik, M.I., "Introduction to Radar Systems", McGraw-Hill Book Company, New York, N.Y., 1962.

[32] McAulay, R.J., "A Theory for Optimal MTI Digital Signal Processing: Part I, Receiver Synthesis", MIT Lincoln Laboratory, Technical Note, 1972-14 (Part I), 22 Feb., 1972.

[33] Linder, R.A., and Kutz, G.H., "Digital Moving Target Indicators", IEEE Trans. Aerosp. Electron.Systems, vol.AES-3, no.6 (supplement), pp.374-385, Nov. 1967.

[34] Zverev, A.I., "Digital MTI Radar Filters", IEEE Trans. Audio Electroacoust., vol.AU-16, no.3, pp.422-432, Sept.1968.

[35] Roecker, W., "The Application of Digital Filters for Moving Target Indication", IEEE Trans. Audio Electroacoust., vol.AU-19, no.1, pp.72-77, March 1971.

[36] Papoulis, A., "Probability, Random Variables, and Stohastic Processes", McGraw-Hill Book Company, New York, N.Y., 1965.

[37] Bendat, J.S., and Piersol, A.G., "Random Data: Analysis and Measurements Procedures", John Wiley & Sons, Inc., New York,N.Y.,1971.

[38] Jenkins, G.M., and Watts, D.G., "Spectral Analysis and Its Applications", Holden-Day, San Francisco, Calif., 1968.

[39] Haykin, S.S., "Signal Processing and Modulation Systems", vol.1, course notes, Faculty of Engineering, McMaster University, Hamilton, July 1976.

[40] Oppenheim, A.V., and Schafer, R.W., "Digital Signal Processing", Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.

References Continued

[41] Cooley, J.W., and Tukey, J.W., "An Algorithm for the Machine Calculation of Complex Fourier Series", Math. Comput., vol.19, no. 90, pp.297-301, April 1965.

[42] Capon, J., "High-Resolution Frequency-Wavenumber Spectrum Analysis", Proc. IEEE, vol.57, no.8, pp.1408-1418, Aug. 1969.

[43] Capon, J., "Signal Processing and Frequency-Wavenumber Spectrum Analysis for a Large Aperture Seismic Array", in: B. Alder, S. Fernbach, and M. Rotenberg (eds.), Methods in Computational Physics, vol.13, no.1, pp.1-59, Academic Press, Inc., New York, N.Y., 1973.

[44] Burg, J.P., "A New Analysis Technique for Time Series Data", paper presented at Advanced Study Institute on Signal Processing, NATO, Enschede, Netherlands, 1968.

[45] Parzen, E., "Statistical Spectral Analysis (Single Channel Case) in 1968", Tech.Rep. 11 on Contract Nonr-225-(80), Stanford University, Stanford, Calif., 1968.

[46] Parzen, E., "Multiple Time Series Modelling", in P.R. Krishnaiah (ed.), Multivariate Analysis 2, pp.389-409, Academic Press, Inc., New York, N.Y., 1969.

[47] Lacoss, R.T., "Data Adaptive Spectral Analysis Methods", Geophysics, vol.36, no.4, pp.661-675, Aug. 1971.

[48] Burg, J.P., "The Relationship Between Maximum Entropy Spectra and Maximum Likelihood Spectra", Geophysics, vol.37, no.2, pp.375-376, April 1972.

[49] Baggeroer, A.B., "Confidence Intervals for Regression (MEM) Spectral Estimates", IEEE Trans. Information Theory, vol.IT-22, no.5, pp.534-545, Sept. 1976.

[50] Capon, J., and Goodman, N.R., "Probability Distributions for Estimators of the Frequency-Wavenumber Spectrum", Proc.IEEE, vol. 58, no.10, pp.1785-1786, Oct. 1970.

[51] Burg, J.P., "Maximum Entropy Spectral Analysis", Ph.D. Thesis, Stanford University, Stanford, Calif., May 1975.

[52] Edward, J.A., and Fitelson, M.M., "Notes on Maximum-Entropy Processing", IEEE Trans. Information Theory, vol.IT-19, no.2, pp. 232-234, March 1973.

References Continued

[53] Kromer, R.E., "Asymptotic Properties of the Autoregressive Spectral Estmator", Ph.D. Thesis, Stanford University, Stanford, Calif., Dec. 1970.

[54] Berk, K.N., "Consistent Autoregressive Spectral Estimates", The Annals of Statistics, vol.2, no.3, pp.489-502, May 1974.

[55] Ulrych, T.J., Smylie, D.E., Jensen, O.G., and Clarke, G.K.C., "Predictive Filtering and Smoothing of Short Records by Using Maximum Entropy", J.Geophys. Res., vol.78, no.23, pp.4959-4964, Aug. 10, 1973.

[56] Ulrych, T.J., "Maximum Entropy Power Spectrum of Truncated Sinusoids", J. Geophys. Res., vol.77, no.8, pp.1396-1400, March 10, 1972.

[57] Chen, W.Y., and Stegen, G.R., "Experiments with Maximum Entropy Spectra of Sinusoids", J. Geophys. Res., vol.79, no.20, pp.3019-3022, July 10, 1974.

[58] Radoski, H.R., Fougere, P.F., and Zawalick, E.J., "A Comparison of Power Spectral Estimates and Applications of the Maximum Entropy Method", J. Geophys. Res., vol.80, no.4, pp.619-625, Feb. 1, 1975.

[59] Ulrych, T.J., and Bishop, T.N., "Maximum Entropy Spectral Analysis and Autoregressive Decomposition", Rev. Geophysics and Space Physics, vol.13, no.1, pp.183-200, Feb. 1975.

[60] Smylie, D.E., Clarke, G.K.C., and Ulrych, T.J., "Analysis of Irregularities in the Earth's Rotation", in: B. Alder, S. Fernbach, and M. Rotenberg (eds.), Methods in Computational Physics, vol.13, no.1, pp.391-430, Academic Press, Inc., New York, N.Y., 1973.

[61] Griffiths, L.J., and Prieto-Diaz, R., "Spectral Analysis of Natural Seismic Events Using Autoregressive Techniques", IEEE Trans. Geosc. Electron., vol.GE-15, no.1, pp.13-25, Jan. 1977.

[62] Landers, T.E., and Lacoss, R.T., "Some Geophysical Applications of Autoregressive Spectral Estimates", vol.GE-15, no.1, pp.26-32, Jan. 1977.

[63] Kaveh, M., "High Resolution Velocity Estimation in the Presence of Extended Clutter", Ph.D. Thesis, Purdue University, West Lafayette, Indiana, Dec. 1974.

References Continued

[64] Kaveh, M., and Cooper, G.R., An Empirical Investigations of the Properties of the Autoregressive Spectral Estimator", IEEE Trans. Information Theory, vol.IT-22, no.3, pp.313-323, May 1976.

[65] McDonough, R.N., "Maximum Entropy Spatial Processing of Array Data", Geophysics, vol.39, no.6, pp.843-853, Dec. 1974.

[66] Fougere, P.F., Zawalick, E.J., and Radoski, H.R., "Spontaneous Line Splitting in Maximum Entropy Power Spectrum Analysis", Phys. Earth and Planet. Inter., vol.12, pp.201-207, 1976.

[67] Ulrych, T.J., and Clayton, R.W., "Time Series Modelling and Maximum Entropy", Phys. Earth and Planet. Inter., vol.12, pp.188-200, 1976.

[68] Fougere, P.F., "A Solution to the Problem of Spontaneous Line Splitting in Maximum Entropy Power Spectrum Analysis", J. Geophys. Res., vol.82, no.7, pp.1051-1054, March 1, 1977.

[69] Fougere, P.F., "A Solution to the Problem of Spontaneous Line Splitting in Maximum Entropy Power Spectrum Analysis", paper presented at 56th Annual AGU Meeting, San Francisco, Calif., 1975.

[70] Pisarenko, V.F., "The Retrieval of Harmonics from a Covariance Function, "Geophys. J. R. Astron. Soc., vol.33, pp.347-366, 1973.

[71] Treitel, S., Gutowski, P.R., and Robinson, E.A., "Empirical Spectral Analysis Revisited", in: Topics in Numerical Analysis III, Academic Press, Inc., New York, N.Y., 1977.

[72] Gutowski, P.R., Robinson, E.A., and Treitel, S., "Novel Aspects of Spectral Estimation," Proceedings JACC 1977, Joint Automatic Control Conference, San Francisco, Calif, June 22-24, 1977.

[73] Box, G.E.P., and Jenkins, G.M., "Time Series Analysis: Forecasting and Control", - revised edition, Holden-Day, San Francisco, Calif., 1976.

[74] van den Bos, A., "Alternative Interpretation of Maximum Entropy Spectral Analysis", IEEE Trans. Information Theory, vol.IT-17, no.4, pp.493-494, July 1971.

[75] Peacock, K.L., and Treitel, S., "Predictive Deconvolution: Theory and Practice", Geophysics, vol.34, no.2, pp.155-169, April 1969.

[76] Robinson, E.A., and Treitel, S., "Principles of Digital Wiener Filtering", Geophys. Prosp., vol.15, pp.311-333, 1967.

References Continued

[77] Levinson, N., "The Wiener RMS (Root Mean Square) Error Criterion in Filter Design and Prediction", Appendix B in: N. Wiener, Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications", The MIT Press, Cambridge, Mass., 1949.

[78] Shannon, C.E., and Weaver, W., "The Mathematical Theory of Communication", The University of Illinois Press, Urbana, Ill., 1959.

[79] Andersen, N., "On the Calculation of Filter Coefficients for Maximum Entropy Method of Spectral Analysis", Geophysics, vol.39, no.1, pp.69-72, Feb. 1974.

[80] Haykin, S., and Kesler, S.K., "The Complex Form of the Maximum Entropy Method of Spectral Estimation", Proc. IEEE, vol.64, no.5, pp.822-823, May 1976.

[81] McGee, T., "An Exposition of the Maximum Entropy Criterion in Spectral Analysis", unpublished report, Department of Geology, University of British Columbia, Vancouver, Canada, Dec. 1971.

[82] Akaike, H., "Fitting Autoregressive Models for Prediction", Ann. Inst. Statist. Math., vol.21, pp.243-247, 1969.

[83] Akaike, H., "Power Spectrum Estimation Through Autoregressive Model Fitting", Ann. Inst. Statist. Math., vol.21, pp.407-419, 1969.

[84] Akaike, H., "Statistical Predictor Identification", Ann. Inst. Statist. Math., vol.22, pp.203-217, 1970.

[85] Akaike, H., "A New Look at the Statistical Model Identification", IEEE Trans. Automatic Control, vol.AC-19, no.6, pp.716-723, Dec. 1974.

[86] Parzen, E., "An Approach to Time Series Modelling and Forecasting Illustrated by Hourly Electricity Demands", Statistical Science Division, State University of New York at Buffalo, Tech. Rep. 37, Jan. 1976.

[87] Hawkes, C.D., and Haykin, S.S., "Modelling of Clutter for Coherent Pulsed Radar", IEEE Trans. Information Theory, vol.IT-21, no.6, pp.703-707, Nov. 1975.

[88] Anderson, T.W., "The Statistical Analysis of Time Series", John Wiley & Sons, Inc., New York, N.Y., 1971.

References Continued

[89] Abramowitz, M., and Segun, I.A., "Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables", pp.932-933, Dover, Inc., New York, N.Y., 1965.

[90] Koopmans, L.H., "The Spectral Analysis of Time Series", Academic Press, Inc., New York, N.Y., 1974.

[91] Haykin, S., and Hawkes, C.D., "Adaptive Digital Filtering for Coherent MTI Radar", Information Sciences, vol.11, pp.335-359, 1976.

[92] Hawkes, C.D., and Haykin, S., "Adaptive Digital Filter for Coherent MTI Radar", IEEE 1975 International Radar Conference, Washington, D.C., pp.57-62, April 1975.

[93] Kesler, S.B., and Haykin, S., "The Maximum Entropy Method Applied to the Spectral Analysis of Computer-Simulated Radar Clutter", accepted for publication in the IEEE Trans. Information Theory, to appear in vol.IT-24, no.2, March 1978.

[94] Kesler, S.B., and Haykin, S., "A Comparison of the Maximum Entropy Method and the Periodogram Method Applied to Spectral Analysis of Computer-Simulated Radar Clutter", accepted for publication in the Canadian Electrical Engineering Journal, to appear in vol. 3, no.1, Jan. 1978.

[95] Kesler, S.B., and Haykin, S., "The Maximum Entropy Method (MEM) with Applications in Radar", in: S. Haykin (ed.), Non-linear Methods of Spectral Analysis, series in "Topics in Applied Physics", Springer Verlag, Berlin, to appear.

[96] "A Brief Description of the ASR-8 Airport Surveillance Radar". Texas Instruments Incorporated, SP05A-EG76, Equipment Group, Aug. 1976.

[97] "Technical Manual for Airborne Wideband Video Recorder/Reproducer, Adviser", vol.1, May 1975 (changed 1 Aug. 1975), RCA (Government and Commercial Systems), Camden, N.J.

[98] Currie, B.W., "Design and Hardware Implementation of Radar Video Display and Sampler Digitizer for use in Radar Clutter Studies", M.Eng. Thesis, McMaster University, Hamilton, Canada, Sept. 1976.

[99] Flock, W.L., "Radar Signature Studies Applicable to Bird-Aircraft Collision Avoidance", Air Force Weapons Laboratory, Air Force System Command, Kirtland Air Force Base, N.M., Final Report for period June 18, 1970 to Nov. 15, 1973.

References Continued

[100] Flock, W.L., and Green, J.L., "The Detection and Identification of Birds in Flight, Using Coherent and Non-Coherent Radars", Proc. IEEE, vol.62, no.6, pp.745-753, June 1974.

[101] Riley, D.C., and Burg, J.P., "Time and Space Adaptive Deconvolution Filters", paper presented at the 42nd Annual Meeting Soc. Explor. Geophys., Anaheim, Calif., 1972.

[102] Alam, M.A., "Adaptive Spectral Estimation", Proceedings of the JACC 1977 - Joint Automatic Control Conference, pp.105-112, San Francisco, Calif., 1977.

[103] Sims, C.S., and D'Mello, M.R., "Adaptive Deconvolution of Seismic Signals", Proceedings of the JACC 1977 - Joint Automatic Control Conference, pp.62-67, San Francisco, Calif., June 1977.

[104] Wang, R.J., "Adaptive Predictive Deconvolution of Seismic Data", Geophys. Prosp., vol.25, pp.342-381, 1977.

[105] Griffiths, L.J., Smolka, F.R., and Trembly, L.D., "Adaptive Deconvolution: A New Technique for Processing Time-Varying Seismic Data", Geophysics, vol.42, no.4, pp.742-759, June 1977.

[106] Widrow, B., McCool, J.M., Larimore, M.G., and Johnson, C.R., Jr., "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter", Proc. IEEE, vol.64, no.8, pp.1151-1162, Aug. 1976.

[107] Alam, M.A., and Sage, A.P., "Sequential Estimation and Identification of Reflection Coefficients by Minimax Entropy Inverse Filtering", Comput. & Elec. Eng., vol.2, pp.315-338, 1975.

[108] Alam, M.A., "Orthonormal Lattice Filter-A Multistage Multichannel Estimator Technique", Geophysics, vol.42, 1977, to be published.

[109] Roy, R., and Lowenschuss, O., "Adaptive Digital MTI Detection Filters", Proceedings of the 1971 IEEE Conference on Decision and Control, Miami Beach, Fla., pp.430-436, Dec. 1971.

[110] Hansen, V.G., Campbell, R.B., Freedman, N., and Schrader, W.W., "Adaptive Digital MTI Signal Processing", EASCON Conference, Washington, D.C., pp.170-176, Sept. 1973.

[111] Benvenuti, P., and Guarguaglini, P.F., "Improvement Factor Evaluation of Open Loop Adaptive Digital MTI", IEEE 1975 International Radar Conference, Washington, D.C., pp.52-56, April 1975.

APPENDIX A

Recursive Calculations of Prediction Error Filter Impulse Response and the Filter Output Power

Starting from the set of normal equations for order M-1

$$\begin{bmatrix} R(0) & R^*(1) & \cdots & R^*(M-1) \\ R(1) & R(0) & \cdots & R^*(M-2) \\ \vdots & \vdots & & \vdots \\ R(M-1) & R(M-2) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{M-1,1} \\ \vdots \\ a_{M-1,M-1} \end{bmatrix} = \begin{bmatrix} P_M \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (A.1)$$

and expanding it to order M, we have

$$\begin{bmatrix} R(0) & R^*(1) & \cdots & R^*(M-1) & R^*(M) \\ R(1) & R(0) & \cdots & R^*(M-2) & R^*(M-1) \\ \vdots & \vdots & & \vdots & \vdots \\ R(M-1) & R(M-2) & \cdots & R(0) & R^*(1) \\ R(M) & R(M-1) & \cdots & R(1) & R(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_{M,1} \\ \vdots \\ a_{M,M-1} \\ a_{M,M} \end{bmatrix} = \begin{bmatrix} P_{M+1} \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}. \quad (A.2)$$

The elements of the last row in the autocorrelation matrix of (A.1) and (A.2) are the complex conjugate of the elements of the first row in a reversed order. The same holds for the row before the last one and the second row, and so on. Thus we can represent the vector of filter coefficients as a sum of two vectors as follows:

$$\begin{bmatrix}1\\ a_{M,1}\\ a_{M,2}\\ \vdots\\ a_{M,M-2}\\ a_{M,M-1}\\ a_{M,M}\end{bmatrix} = \begin{bmatrix}1\\ a_{M-1,1}\\ a_{M-1,2}\\ \vdots\\ a_{M-1,M-2}\\ a_{M-1,M-1}\\ 0\end{bmatrix} + a_{M,M}\begin{bmatrix}0\\ a^*_{M-1,M-1}\\ a^*_{M-1,M-2}\\ \vdots\\ a^*_{M-1,2}\\ a^*_{M-1,1}\\ 1\end{bmatrix} \tag{A.3}$$

where vector $[1, a_{M-1,1}, \ldots, a_{M-1,M-1}, 0]^T$ is the vector of known filter coefficients of order M-1 (Eq. (A.1)) augmented by a zero, and the second vector on the right-hand side is its complex conjugate in reversed order. The multiplicative factor is $a_{M,M}$, since the last equality in Eq. (A.3) must be satisfied.

Using (A.3) in (A.2) and making use of (A.1), we have

$$\begin{bmatrix} R(0) & R^*(1) & \cdots & R^*(M-1) & R^*(M)\\ R(1) & R(0) & \cdots & R^*(M-2) & R^*(M-1)\\ \vdots & \vdots & & \vdots & \vdots\\ R(M-1) & R(M-2) & \cdots & R(0) & R^*(1)\\ R(M) & R(M-1) & \cdots & R(1) & R(0)\end{bmatrix} \left\{\begin{bmatrix}1\\ a_{M-1,1}\\ \vdots\\ a_{M-1,M-1}\\ 0\end{bmatrix} + a_{M,M}\begin{bmatrix}0\\ a^*_{M-1,M-1}\\ \vdots\\ a^*_{M-1,1}\\ 1\end{bmatrix}\right\}$$

$$= \begin{bmatrix}P_M\\ 0\\ \vdots\\ 0\\ D_M\end{bmatrix} + a_{M,M}\begin{bmatrix}D^*_M\\ 0\\ \vdots\\ 0\\ P_M\end{bmatrix} = \begin{bmatrix}P_{M+1}\\ 0\\ \vdots\\ 0\\ 0\end{bmatrix} \tag{A.4}$$

where the quantity D_M is defined by the last equation in (A.4)

$$D_M = \sum_{m=0}^{M-1} a_{M-1,m} \cdot R^*(M-m) \quad . \tag{A.5}$$

We also see from the last equality in (A.4) that

$$D_M + a_{M,M} \cdot P_M = 0 \ , \tag{A.6}$$

and hence, we have the solution for $a_{M,M}$:

$$a_{M,M} = - \frac{D_M}{P_M} \quad . \tag{A.7}$$

Recall from Eq. (2.2.7) that P_M is a Hermitian form and thus it is always real. Moreover, we have shown that the autocorrelation matrix is always positive-definite, resulting in $P_M > 0$. Thus, we will always have a solution for (A.7).

Note that the solution for (A.7) differs from the expression (2.4.11) obtained by minimizing the average error power. However, those two expressions are identical as will be seen later. Solution for (A.7) involves the autocorrelation function values (A.5), and thus it will not be used in the iterative procedure.

Once the value for $a_{M,M}$ is determined we can find the remaining filter coefficients by means of Eq. (A.3). Since the first and the last equations in (A.3) are trivial, we use only M-1 remaining equations to get

$$a_{M,m} = a_{M-1,m} + a_{M,M} \cdot a^*_{M-1,M-m} \quad , \quad m = 1,2,\ldots,M-1 . \tag{A.8}$$

To obtain the value of P_{M+1}, we use the first equation in the last equality in (A.4) to get

$$P_{M+1} = P_M + a_{M,M} D^*_M . \tag{A.9}$$

Substituting for $D^*_M = -a^*_{M,M} P_M$ from (A.6) we get the simple recursion relation for P_{M+1}

$$P_{M+1} = P_M(1 - a_{M,M} a^*_{M,M}) = P_M(1 - |a_{M,M}|^2) . \tag{A.10}$$

As noted by Burg [44], the expression (A.10) gives a simple relation between P_{M+1} and P_M, thus avoiding the problem of round-off error accumulation when P_M becomes quite small. The problem arises in the original Levinson algorithm which calculates P_{M+1} directly from the first equation in (A.2).

APPENDIX B

Recursion Formulae for $p_M(n)$ and $q_M(n)$

Noting that $a_{M-2,0} \equiv 1$ and $a_{M-2,m} \equiv 0$ for $m = M-1, M, \ldots$ and $m = -1, -2, \ldots$ we can rewrite Eq. (2.4.12) for the lower order filter as

$$a_{M-1,m} = a_{M-2,m} + a_{M-1,M-1} \cdot a^*_{M-2,M-1-m}, \quad m = 0,1,\ldots,M. \qquad (B.1)$$

Substituting (B.1) into the expression (2.5.2) we have for $p_M(n)$

$$p_M(n) = \sum_{m=0}^{M} a^*_{M-1,m} \cdot x(n+m)$$

$$= \sum_{m=0}^{M} [a^*_{M-2,m} + a^*_{M-1,M-1} \cdot a_{M-2,M-1-m}] \cdot x(n+m) \qquad (B.2)$$

$$= \sum_{m=0}^{M-1} a^*_{M-2,m} \cdot x(n+m) + a^*_{M-1,M-1} \sum_{m=0}^{M-1} a_{M-2,M-1-m} \cdot x(n+m).$$

The upper limit of the summations in the last row of (B.2) is reduced by 1 since $a_{M-2,M-1} = a_{M-2,-1} \equiv 0$. Now, the first summation is equal to the term on the right-hand side of Eq. (2.5.2), except that M is replaced by M-1. Likewise, the second summation in (B.2) is equal to the term on the right-hand side of Eq. (2.5.3) with the same substitution for M.

Thus, we may rewrite (B.2) as

$$p_M(n) = p_{M-1}(n) + a^*_{M-1,M-1} \cdot q_{M-1}(n) \qquad (B.3)$$

Similarly we have for $q_M(n)$:

$$\begin{aligned} q_M(n) &= \sum_{m=0}^{M} a_{M-1,M-m} \cdot x(n+m) \\ &= \sum_{m=1}^{M} a_{M-2,M-m} \cdot x(n+m) + a_{M-1,M-1} \sum_{m=1}^{M} a_{M-2,m-1} \cdot x(n+m) . \end{aligned} \qquad (B.4)$$

In order to have the same limits as in (B.2) we simply substitute $m'=m-1$ in (B.4) to get

$$q_M(n) = \sum_{m'=0}^{M-1} a_{M-2,M-1-m'} \cdot x(n+m'+1) + a_{M-1,M-1} \sum_{m'=0}^{M-1} a_{M-2,m'} \cdot x(n+m'+1) . \qquad (B.5)$$

After omitting primes we see from Eqs. (2.5.2) and (2.5.3) that $q_M(n)$ may be expressed in terms of lower order quantities as

$$q_M(n) = q_{M-1}(n+1) + a_{M-1,M-1} \cdot p_{M-1}(n+1) . \qquad (B.6)$$

The recursive relations (B.3) and (B.6) indicate that we can calculate $p_M(n)$ and $q_M(n)$ by using only the last filter coefficient in every iteration.

APPENDIX C

Development of Clutter Model for Coherent Pulse Radar

We start from expression (3.2.3) for the transmitted signal

$$\psi_T(t) = \mu_T(t) \cdot e^{j2\pi f_c t} \qquad \text{(C.1)}$$

where the envelope $\mu_T(t)$ is given as a sum of the envelopes of the single pulse, weighted with the appropriate value of the antenna gain, as shown by (3.2.4).

We assume that the received signal is processed by a filter that is matched to the pulse

$$\psi_F(t) = \mu_F(t-\tau_o)\ e^{j2\pi f_c (t-\tau_o)} \qquad \text{(C.2)}$$

where $\mu_F(t)$ is the pulse envelope and τ_o is the reference delay corresponding to the range ring being sampled. The frequency reference is taken as zero Doppler. In general, the envelope $\mu_F(t)$ need not be the same as that of the individual pulse $\mu_S(t)$ of the transmitted pulse train.

The total clutter signal I(p) at the pth sampling instant, is defined in [26] and is given by

$$I(p) = \sum_i \{ [\sum_k s_{i,k}(p,f_r)\, \chi_{SF}(\tau_k, f_d)]\, a_p(i) \} \cdot e^{j2\pi f_d p T_R} \qquad \text{(C.3)}$$

where $\chi_{SF}(\tau,f)$ is the cross-ambiguity function between $\mu_S(t)$ and $\mu_F(t)$, f_r is the dipole rotational frequency, and f_d is the Doppler shift given by the average radial velocity of the entire cloud of scatterers. The autocorrelation function of the clutter signal is defined by

$$R_I(q) = E[I^*(p)\cdot I(p+q)] \tag{C.4}$$

where q is the order of the lag product. For the case of linear polarization in both transmitter and receiver, the value of autocorrelation at zero lag is

$$R_I(0) = \frac{3}{2}E[C^2]\ . \tag{C.5}$$

Using (C.5) and the relation among the random variables C, U, and L (Eq. (3.2.2)) we find that the normalized autocorrelation function is given by [26]

$$R_I(q) = \frac{2}{3}\left[1+\frac{1}{2}\int_{-\infty}^{\infty}\cos(4\pi q f_r T_R)p(f_r)df_r\right]\left[\int_{-\infty}^{\infty}e^{j2\pi q f_d T_R}\cdot p(f_d)df_d\right]e^{-\frac{q^2T_R^2}{2T_o^2}} \tag{C.6}$$

where $p(f_r)$ and $p(f_d)$ denote the probability density functions of f_r and f_d, respectively. The last term on the right-hand side of (C.6) represents the antenna pattern, which is assumed to be of Gaussian shape for analytical convenience, with $1.65T_o$ being the two-way 3 dB beamwidth of the antenna.

The model that was progammed is based on Eq. (C.3). Unlike previous approaches to clutter modelling, which have centered on techniques that synthesize the process by passing white noise through linear filters, the present approach does not involve a previous knowledge of autocorrelation function. The choice of the scatterer array length in a computer model was dictated by several factors. As a compromise between an adequate representation of the scatterer field and the computation time, the dynamic range representing the main lobe of the antenna was chosen to be 30 dB. With 24 scatterers assumed in the 3 dB one-way beamwidth, the array length was 76.

When the Gaussian distribution is assumed for the dipole rotational frequency f_r and Doppler shift f_d, the value for the autocorrelation function of the clutter signal (C.6) is expressed in terms of the corresponding first and second order moments. Let μ_r and σ_r^2 be the normalized mean and variance of f_r, respectively, and μ_d and σ_d^2 be the normalized mean and variance of f_d, respectively. The normalization factor is interpulse period T_R. We also normalize T_o with respect to T_R. For the case of Gaussian distributions, the normalized autocorrelation function of the clutter (C.6) becomes

$$R_I(q) = \frac{2}{3}\left[1+\frac{1}{2}\cos(4\pi q\mu_r)\,e^{-8\pi^2 q^2 \sigma_r^2}\right]\cdot e^{-j2\pi q\mu_d - 2\pi^2 q^2 \sigma_d^2 - \frac{q^2}{2T_o^2}} \quad . \quad \text{(C.7)}$$

The Fourier transform of Eq. (C.7) represents the theoretical power spectral density of the clutter signal. Analytical evaluation based on this relation is straightforward but tedious. However, we do not need the explicit formula for it. We use it as a reference to compare a different methods of spectral estimation. For this purpose, we simply use the fast Fourier transform routine to get the theoretical spectral density.

APPENDIX D

Characteristics of the ASR-8 Radar

ASR-8 is a modern ATC radar with coherent processing, manufactured by Texas Instruments in 1976 [96]. Its principal characteristics are listed in Table D - 1. Figure D - 1 shows the calculated free-space antenna elevation patterns for both high and low beams. The azimuth pattern is presented in Fig. D - 2. The traces under the main lobe are crosspolarized measurements, i.e., responses of vertically polarized receiving pattern to horizontally polarized transmitting one, and vice versa.

Table D-1 RADAR CHARACTERISTICS

Antenna Characteristics	
Aperture size	4.9 meters (16.1 feet) azimuth by 2.7 meters (9 feet) elevation
Gain (minimum)	
Normal beam	33.5 dB
Passive (high) beam	32.5 dB
Sidelobes (maximum)	
Normal beam	−24 dB
Passive beam	−22 dB
Beamwidth:	
Elevation	4.8 degrees minimum at −3 dB, COSECANT to 30 degrees
Azimuth	1.35 degrees minimum at −3 dB point
VSWR	1.4:1
Polarization	Linear vertical or circular, remotely selectable
Integrated cancellation ratio (minimum)	
Normal beam	22 dB (25 dB typical)
Passive beam	22 dB (24 dB typical)
Rotary joint	5-channel, 3 S-band and 2 L-band channels
Data pickoff	
Dual azimuth pulse generators	2 sets reference pulse and 4,096 change pulses
Drive mechanism safety switch	On tower
Tilt angle: referenced to the lower 3 dB point of the normal beam	±2.5 degrees, manually adjusted
Drive motor (dual)	5.0 horsepower (each)
Antenna rotation rate	12.5 rpm (15 rpm optional)
Transmitter Characteristics	
Source	VA-87E air cooled klystron
Frequency	Tunable 2.7 to 2.9 GHz; derived from crystal STALO and COHO
Peak power	1.4 mW
Average power	875 W nominal
Pulsewidth	0.6 μs
Repetition frequency	4-pulse stagger with 1,040 average, or fixed (selectable from 700 to 1,200 pps)
Cooling	Forced air
Waveguide pressurization	Freon-12 for diplexer
Fault detection	Automatic recycling and shutdown sequence
Modulator	12 plug-in modules, fail-soft
HVPS	Solid-state, 250 Vdc
Receiver Characteristics	
Parametric amplifier	
Type	Solid-state with Gunn pump source
Noise figure	1.25 dB maximum
Bandwidth	200 MHz (2.7 to 2.9 GHz) fixed tuned
Gain	15 dB minimum
System noise figure	4.0 dB maximum, measured at noise coupler
Receiver channels	MTI IF Normal IF Log IF
STALO	VHF crystal oscillator with multiplier (2.7 to 2.9 GHz)
COHO	30 MHz crystal

Table D-1 RADAR CHARACTERISTICS (Continued)

Receiver bandwidth	
Normal IF	1.1 MHz
MTI IF	5 MHz
MTI video	585 kHz
Log IF	1.1 MHz
Log IF (dynamic range)	80 dB
STC	Diode RF attenuator before parametric amplifier; 0 to 40 dB
Minimum discernible signal (MDS)	
Normal receiver	–110 dBm
Log receiver	–109 dBm
MTI receiver	–108 dBm
Processor Characteristics	
MTI/Normal Channel Operating Modes (video alignment in all modes)	
Regular video	
Enhanced video	
Log FTC video	
Log FTC with weather background	
Enhanced Log FTC	
Enhanced Log FTC with weather background	
Number of Bits	
MTI channel	10
Normal channel	6
Weather channel	7
Integrators (MTI and normal)	6
Storage (includes alignment)	97,177
Sample rate	2.14 MHz
MTI cancellers	In-phase and quadrature
MTI canceller modes	
Single canceller	
Dual canceller	
Dual canceller with feedback (25, 30, 35, 40 dB cancellation of scanning clutter)	
PRF stagger	4-pulse sequence
First blind speed	Greater than 2,150 knots
Velocity-response nulls	Less than 10 dB peak to peak
Log-CFAR	
MTI	Digital, 10 bits
Normal	Analog
System MTI improvement factor (stagger, scanning)	34 dB
STC function generator	
Initial attenuation	0 to 40 dB
Initial delay	100 μs
Range decay exponent	R^{-1} to R^{-5}
Receiver gain attenuation	0 to 40 dB
Range azimuth gate	
Generator	Field programmable
Controlled functions	MTI/Normal video gating Dual antenna beam switching Receiver noise generator gating Receiver test trigger gating Stagger/nonstagger select Azimuth strobe STC select Receiver gain control

Table D-1 RADAR CHARACTERISTICS (Continued)

Display Site Remote Unit Characteristics	
Line compensator	30 dB minimum common mode rejection Compensates for up to 6,100 meters of coax
Distribution amplifier (91 meters coax with 75-ohm termination)	
Triggers	Rise time <0.1 microsecond Duration 1.0 to 1.5 microseconds Amplitude ≥15 volts
Video	Rise time <0.1 microsecond Droop (200 μs pulse) ≤10 percent Amplitude 0 to 4 volts
ACP/ARP	Pulsewidth 23 ±3.0 microseconds Amplitude 5 ±1 volt
Power requirement	115 Vac, single phase, 50/60 Hz
Power Supplies	Redundant diode floating +5 Vdc, +15Vdc, −15 Vdc, +24 Vdc
Maintenance PPI Characteristics	
Operational characteristics	
CRT size	40-centimeter diameter
Display ranges	Five selectable ranges Position 1: 5 to 20 nautical miles (nmi) preset Position 2: 15 to 20 nmi preset Position 3: 35 to 70 nmi preset Position 4: 40 to 135 nmi preset Position 5: 190 to 275 nmi preset (beacon pretrigger required)
Range marks	1, 2, 5, 10 nautical miles selectable Every fifth mark intensified
Videos	MTI Normal (MTI/normal range gate adjustable out to 60-nautical mile radar range) Beacon Map Intensity of all signals is adjustable
Azimuth synchronization	4,096 APC with ARP
Triggers	Radar and beacon pretrigger
Decentering	1 radius in any direction
Electrical characteristics	
CRT	P7 phosphor
Range mark accuracy	±0.25 nautical mile at 60 nautical miles
Sweep linearity	To ±2.5 millimeters for range marks
Azimuth accuracy	±0.5 degree
Pretrigger delay	45 to 120 μs
Deflection	Fixed yoke magnetic
Video bandwidth	3.5 MHz
Video inputs	75 ohms, 8 volts peak video amplitude
ACP, ARP inputs	Externally terminated to match line. Five volts P-P maximum amplitude, any width between 1 μs and 50 percent of ACP duty cycle; logic level or approximate sinusoid.
Pretrigger inputs	75 ohms, 5 to 70 volts peak amplitude
Voltage	102 to 138 volts, single phase, 47 to 470 Hz
Power	250 watts

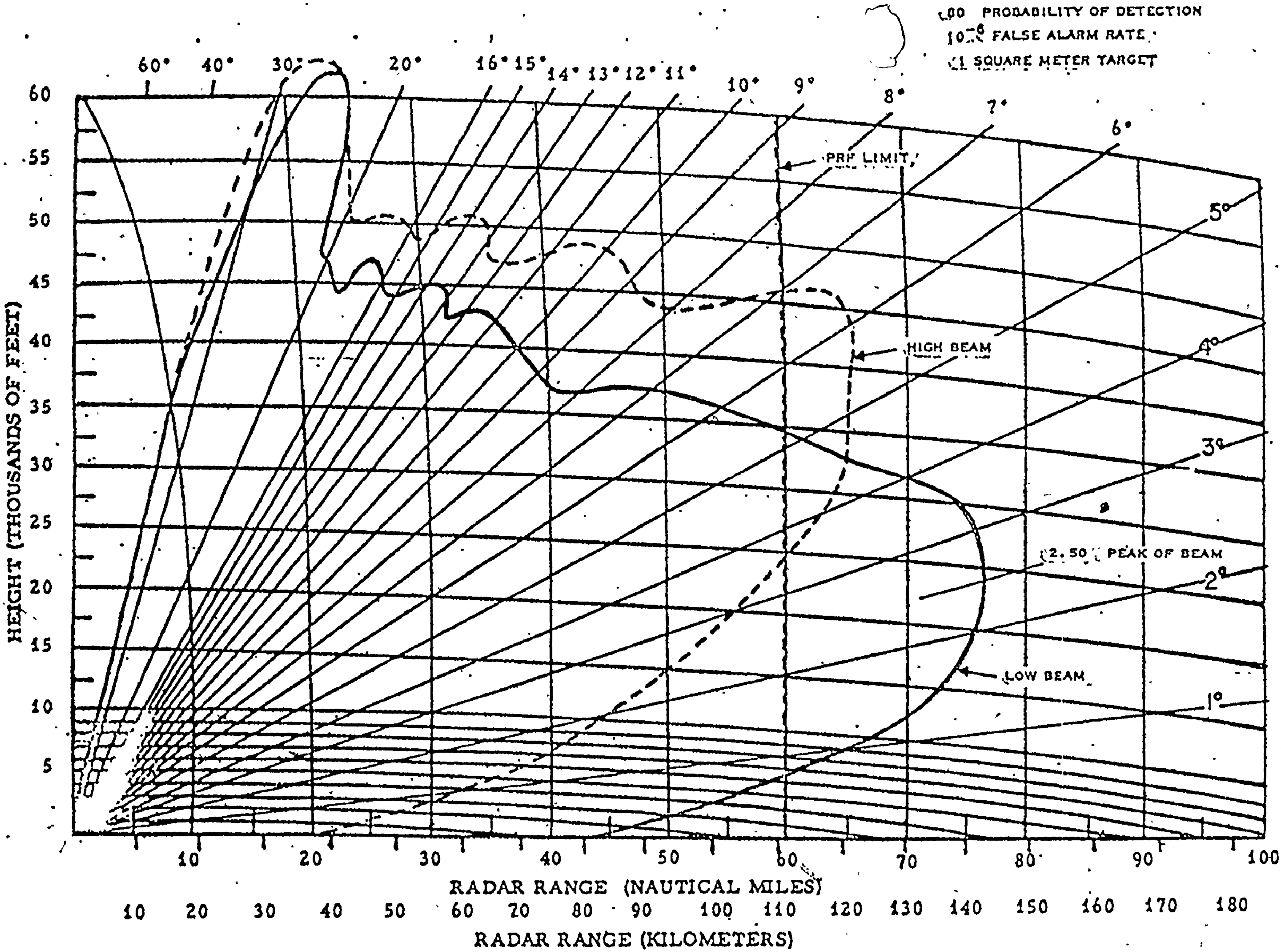


Fig. D-1 Calculated Free Space Antenna Elevation Pattern.

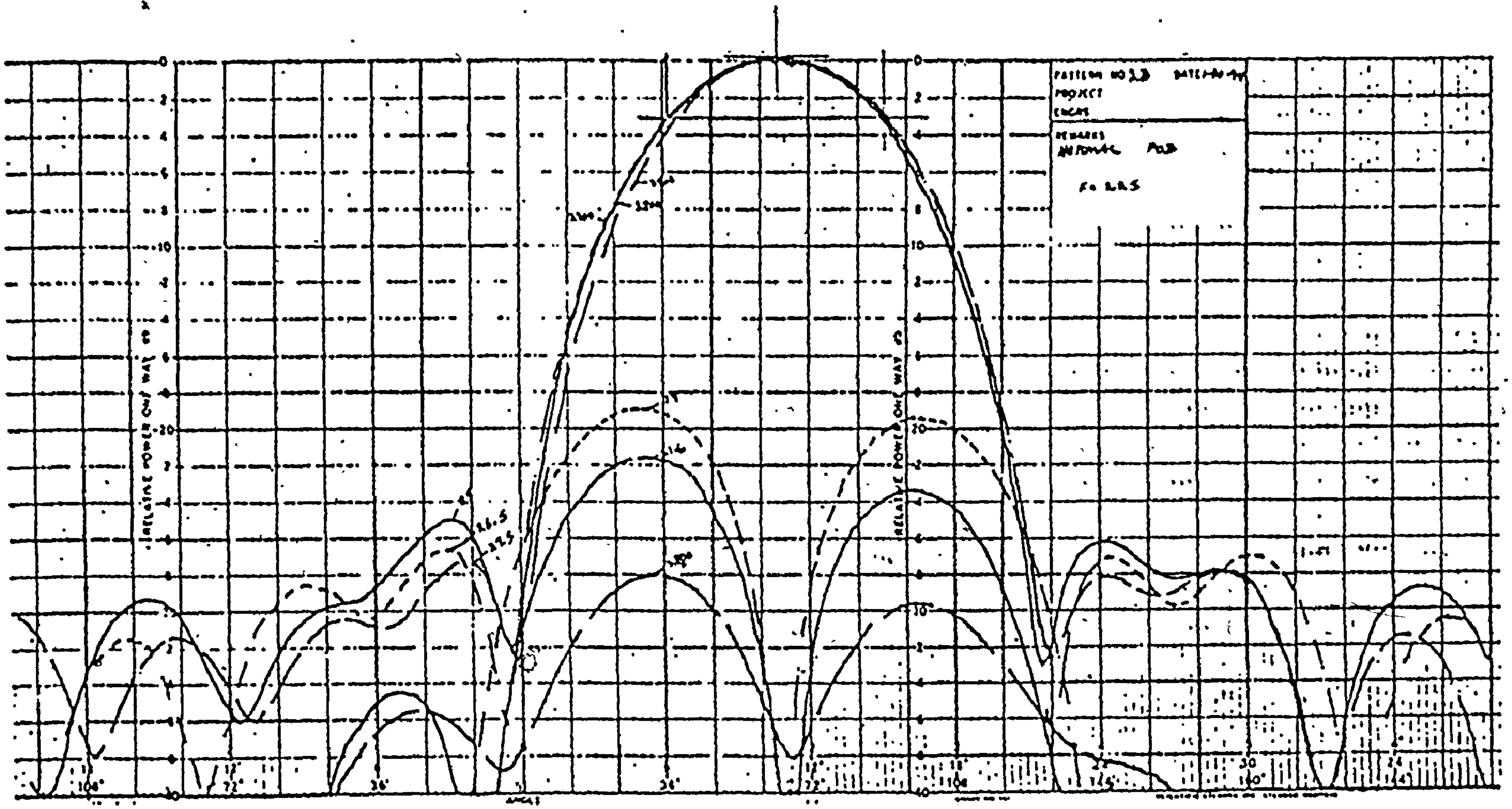


Fig. D-2 Measured Antenna Azimuth Pattern.

APPENDIX E

Documented Listings of Computer Programs

This appendix contains documented listings of the comuter programs used in the analysis of recorded clutter data. They are used for selection and arrangement of samples stored in mass memory and for the spectral analysis of arranged data. Listed programs are written for the CDC 1700 minicomputer. The corresponding programs for spectral analysis of computer-simulated clutter signals (which was done in Chapter 4) have been run on the CDC 6400 computer. Since these are essentially the same as the previous ones, except for the minor adaptations, they are not listed.

Program SAMARR reads the samples from mass memory, separates I and Q samples from the computer word (subroutine DEMIX), and arranges them in the format required for spectral estimation programs. There is an option for the sample mean removal from the data (subroutine MNOPT). Depending on the order of the data samples (see Fig. 5.3 - 1), the options for reordering is provided by subroutine ARROPT.

Program MEMFP1 calculates the maximum entropy spectra of the signals stored on pseudotape LUIN by SAMARR. The order of the PEF is automatically chosen by the final prediction error criterion. Options are provided for choosing linear or logarithmic scale for computed spectra and for rewinding the pseudotape. This last option is convenient for statistical analysis of the stored data. The scale of range for both coordinate axes is also optional.

Program PER64 evaluates the spectra by means of Welch's method of modified periodograms, with 64 data samples. Options are provided for scaling the axes, rewinding the pseudotapes, and also, the window function may be chosen by calling subroutine WINDW1. There is also the possibility of averaging the data up to four 64 points blocks. The data are padded with zeros so that the FFT's are calculated with 128 points. The programs PER32 and PER16, which perform the spectral estimations of 32 and 16 data points, respectively, are identical with PER64 and therefore are not listed. Also, the programs PER128 and PER256 are not listed. They differ from the previous ones only in that they do not have the options for averaging over more data blocks. This would simply require more core storage.

The other supporting routines which existed in the program library have not been listed. They have basically been used for inputing and outputing the data, transferring the data from one logical unit to another, for plotting, and for separating the computer words into I and Q samples (i.e. subroutine DEMIX).

```
      PROGRAM SAMARR
C     ------- ------
C
C   THIS PROGRAM SEPARATES RAM WORDS INTO  I  AND  Q  SIGNAL
C   SAMPLES (DEMIX), ARRANGES THEM INTO COMPLEX POINTS AND
C   STORES THEM IN 512 (4X64 COMPLEX) POINTS ON PSEUDOTAPE
C   L U I N,  AS REQUIRED BY  M E M F P 1.  MAX MEM WORDS
C   =3840 (60 BLOCKS, 160 SECTORS).
C
C   SPECIFY PARAMETERS;
C         LOUT     = DISPLAY L.U.
C         LUIN     = STORING PSEUDOTAPE AS PER   M E M G E N .
C         ISECAD   = STARTING SECTOR ADDRESS ON DISK # 6.
C         NWDS     = # OF WORDS TO BE PICKED PER RUN
C         IOPT     = 0 - PROCESSING DATA IN ORIGINAL FORM
C                  = 1 - CALLING SUBROUTINE  A R R O P T (2*300)
C                  = 2 -        "             "        "   (8*200)
C         MOPT     = 0 - NO MEAN CALCULATION.
C                  = 1 - CALLING SUBROUTINE   M N O P T.
C         LUIN1    = 11 - 15  FILTERS THE DATA AND WRITES ON LUIN1
C                  = 0-10 OR 16,... AUTOMATICALLY SKIPS  M T I 2
C
      DIMENSION IRAM(1929), JRAM(1920), DATA(512)
      DIMENSION DAT1(128), DAT2(128), DAT3(128),DAT4(128)
      EQUIVALENCE(IRAM(10),JRAM(1)),(DAT1(1),DATA(1)),(DAT2(1),
     & DATA(129)),(DAT3(1),DATA(257)),(DAT4(1),DATA(385))
    6 WRITE(4,10)
   10 FORMAT(*LUIN,LUIN1,ISECAD,NWDS,IOPT,MOPT=*)
      CALL HCREAD(4,6,LUIN,LUIN1,ISECAD,NWDS,IOPT,MOPT)
      WRITE(4,20)
   20 FORMAT(*DISPLAY UNIT=*)
      CALL HCREAD(4,1,LOUT)
      IF(NWDS.LT.1.OR.NWDS.GT.1929) GO TO 6
      IF(ISECAD.LT.0.OR.ISECAD.GT.32479) GO TO 6
      CALL DSKRW(6,0,IRAM,NWDS,ISECAD)
      WRITE(LOUT,40) NWDS, ISECAD
   40 FORMAT(1X,I6,*WORDS STARTING AT SECTOR *,I6,//)
      NSAM = NWDS - 9
      IF(LOUT.EQ.16.OR.LOUT.EQ.17) GO TO 60
      WRITE(LOUT,51) (JRAM(J), J=1,NSAM)
   51 FORMAT(10(1X,I6))
      GO TO 7
   60 CONTINUE
      WRITE(LOUT,71) (JRAM(J), J=1,NSAM)
   71 FORMAT(20(I6))
    7 CONTINUE
      WRITE(LOUT,99)
   99 FORMAT(/)
C
C RAM WORDS AUTOMATICALLY OFFSET BY 9, WITH EQUIV. (JRAM(1),
C IRAM(10)).  UNSCRAMBLE  I  AND  Q  USING   D E M I X.
C
```

```
      NL = NSAM/256
      NR = NSAM - NL*256
      LOOP = NL
      IF(NR.GT.0) LOOP = LOOP + 1
C
C  BEGINNING OF THE MAIN LOOP.
C
      DO 15 K=1,LOOP
      DO 1 I=1,256
      KI = (K-1)*256 + I
      CALL DEMIX (JRAM(KI), ITI,ITQ,6)
      I2 = 2*I
      I1 = I2 - 1
      DATA(I1) = FLOAT(ITI)
      DATA(I2) = FLOAT(ITQ)
    1 CONTINUE
C
C  THE ARRAY DATA(512) IS NOW FORMED AND ALSO THE ARRAYS
C  DAT1,2,3,4(128). IF MOPT.NE.0  CALCULATE THE MEAN
C  VALUES OF ALL SUBARRAYS, DAT1,2,3,4, AS WELL AS DATA(512)
C
      IF(MOPT.EQ.0) GO TO 5
      CALL MNOPT(DAT1,DAT2,DAT3,DAT4,LOUT,AMI,AMQ)
      DO 3 I=1,256
      I2 = 2*I
      I1 = I2 - 1
      DATA(I1) = DATA(I1) - AMI
      DATA(I2) = DATA(I2) - AMQ
    3 CONTINUE
    5 CONTINUE
      CALL ARROPT(DAT1,DAT2,DAT3,DAT4,IOPT)
C
C  TRANSFER THE DATA IN BLOCKS OF 64 COMPLEX POINTS ON
C  PSEUDOTAPE L U I N  (UNFORMATED).
C
      WRITE(LUIN) DAT1
      WRITE(LUIN) DAT2
      WRITE(LUIN) DAT3
      WRITE(LUIN) DAT4
      IFL=0
   21 CONTINUE
C
C  DISPLAY THE DATA ON DISPLAY UNIT (L O U T).
C
      IF(LOUT.EQ.16.OR.LOUT.EQ.17) GO TO 80
      WRITE(LOUT,50) (DAT1(I), I=1,128)
      WRITE(LOUT,99)
      WRITE(LOUT,50) (DAT2(I), I=1,128)
      WRITE(LOUT,99)
      WRITE(LOUT,50) (DAT3(I), I=1,128)
      WRITE(LOUT,99)
```

```
      WRITE(LOUT,50) (DAT4(I), I=1,123)
      WRITE(LOUT,99)
      GO TO 15
   80 CONTINUE
      WRITE(LOUT,70) (DAT1(I), I=1,128)
      WRITE(LOUT,99)
      WRITE(LOUT,70) (DAT2(I), I=1,128)
      WRITE(LOUT,99)
      WRITE(LOUT,70) (DAT3(I), I=1,128)
      WRITE(LOUT,99)
      WRITE(LOUT,70) (DAT4(I), I=1,128)
      WRITE(LOUT,99)
   15 CONTINUE
   50 FORMAT(6(E11.4))
   70 FORMAT(10(E12.4))
      IF(IFL.GT.0) GO TO 6
      IF(LUIN1.LT.11.OR.LUIN1.GT.15) GO TO 6
      REWIND LUIN1
C
C   FILTER THE DATA WITH DOUBLE DELAY CANCELLER (SUBROUTINE
C   M T I 2)  AND WRITE THE FILTERED DATA ON PSEUDOTAPE
C   L U I N 1   (UNFORMATED).
C
      CALL MTI2(DATA)
      WRITE(LUIN1) DAT1
      WRITE(LUIN1) DAT2
      WRITE(LUIN1) DAT3
      WRITE(LUIN1) DAT4
C
C   WRITE THE FILTERED DATA ON DISPLAY UNIT  (L O U T).
C
      IFL=IFL+1
      GO TO 21
      GO TO 6
      END
 MON
```

```
      SUBROUTINE ARROPT (DAT1,DAT2,DAT3,DAT4,IOPT)
C     ========== =================================
C
C  THIS SUBROUTINE ARRANGES THE DATA COLLECTED IN
C  FORMATS  "2 AND 8 RANGE BINS".
C
      DIMENSION X(128),D1(128),D2(128),D3(128),D4(128)
      DIMENSION DAT1(1),DAT2(1),DAT3(1),DAT4(1)
      IF(IOPT.LT.1.OR.IOPT.GT.2) RETURN
      GO TO (60,70), IOPT
C
C  ARRANGE THE DATA COLLECTED IN FORMAT 2X800.
C
   60 CONTINUE
      DO 20 I=1,32
      I2=2*I
      I1=I2-1
      I44=4*I
      I43=I44-1
      I42=I44-2
      I41=I44-3
      J1=64+I1
      J2=64+I2
      D1(I1)=DAT1(I41)
      D1(I2)=DAT1(I42)
      D1(J1)=DAT2(I41)
      D1(J2)=DAT2(I42)
      D2(I1)=DAT3(I41)
      D2(I2)=DAT3(I42)
      D2(J1)=DAT4(I41)
      D2(J2)=DAT4(I42)
      D3(I1)=DAT1(I43)
      D3(I2)=DAT1(I44)
      D3(J1)=DAT2(I43)
      D3(J2)=DAT2(I44)
      D4(I1)=DAT3(I43)
      D4(I2)=DAT3(I44)
      D4(J1)=DAT4(I43)
      D4(J2)=DAT4(I44)
   20 CONTINUE
      GO TO 80
C
C  ARRANGE THE DATA COLLECTED IN FORMAT 8X200.
C
   70 CONTINUE
      DO 200 J=1,4
      GO TO (1,2,3,4), J
    1 CONTINUE
      DO 11 I=1,128
   11 X(I)=DAT1(I)
      GO TO 5
    2 CONTINUE
```

```
      DO 22 I=1,128
   22 X(I)=DAT2(I)
      GO TO 5
    3 CONTINUE
      DO 33 I=1,128
   33 X(I)=DAT3(I)
      GO TO 5
    4 CONTINUE
      DO 44 I=1,128
   44 X(I)=DAT4(I)
    5 CONTINUE
      DO 100 I=1,8
      I2=2*I
      I1=I2-1
      J1=16*(J-1)+I1
      J2=16*(J-1)+I2
      JJ1=J1+64
      JJ2=J2+64
      K=16*(I-1)
      D1(J1)=X(K+1)
      D1(J2)=X(K+2)
      D1(JJ1)=X(K+3)
      D1(JJ2)=X(K+4)
      D2(J1)=X(K+5)
      D2(J2)=X(K+6)
      D2(JJ1)=X(K+7)
      D2(JJ2)=X(K+8)
      D3(J1)=X(K+9)
      D3(J2)=X(K+10)
      D3(JJ1)=X(K+11)
      D3(JJ2)=X(K+12)
      D4(J1)=X(K+13)
      D4(J2)=X(K+14)
      D4(JJ1)=X(K+15)
      D4(JJ2)=X(K+16)
  100 CONTINUE
  200 CONTINUE
   80 CONTINUE
      DO 90 I=1,128
      DAT1(I)=D1(I)
      DAT2(I)=D2(I)
      DAT3(I)=D3(I)
      DAT4(I)=D4(I)
   90 CONTINUE
      RETURN
      END
MON.
```

```
      SUBROUTINE MTI2 (DATA)
C     ========== ==========
C
C  FILTERS THE INPUT DATA WITH DOUBLE DELAY CANCELLER
C  USING FORMULA      X(I) = DATA(I+4)-2.*DATA(I+2)+DATA(I).
C  THIS FORMULA PROVIDES MTI FILTERING OF BOTH  I  AND  Q
C  COMPONENTS.   BOTH INPUT AND OUTPUT ARRAYS STORED IN
C  D A T A   WITH  I  AND  Q  ALTERNATIVELLY.
C
      DIMENSION X(512), DATA(1)
      DO 1 I=1,508
      X(I)=DATA(I+4)-2.*DATA(I+2)+DATA(I)
    1 CONTINUE
      DO 2 I=509,512
      X(I)=0.
    2 CONTINUE
      DO 3 I=1,512
      DATA(I)=X(I)
    3 CONTINUE
      RETURN
      END
 MON
```

```
      PROGRAM MEMFPI
      ------- ------
C
C
C GENERAL PROGRAM FOR COMPUTING SPECTRA USING MAXIMUM ENTROPY
C METHOD. HANDLES UP TO 3840 SAMPLES.
C PEF FILTER LENGTH ADJUSTED BY MINIMUM FPE CRITERION.
C DATA ASSUMED TO BE ON PSEUDOTAPE (LUIN) IN 64 COMPLEX BLOCKS.
C    N=NO. OF SAMPLES TO BE PROCESSED
C    IOP = 0   -  LIN SCALE
C    IOP = 1   -  LOG SCALE
C    M = 30   -   MAXIMUM POSSIBLE LENGTH
C
      DIMENSION IA(16),JA(16)
      DIMENSION X(612),Y(612),B1(512),B2(512),B11(100),B21(100)
      DIMENSION DAT1(128),DAT2(128),DAT3(128),DAT4(128)
      DIMENSION A(50),AA(50)
      EQUIVALENCE (X(1),DAT1(1)),(X(129),DAT2(1))
      EQUIVALENCE (X(257),DAT3(1)),(X(385),DAT4(1))
      EQUIVALENCE (X(1),B1(1)),(X(513),B11(1))
      EQUIVALENCE (Y(1),B2(1)),(Y(513),B21(1))
      DATA IA(1),IA(2),IA(3),IA(4),IA(5),IA(6),IA(7),IA(8),IA(9),
     1IA(10),IA(11),IA(12),IA(13),IA(14),IA(15),IA(16)
     2/$0,$10,$20,$30,$40,$50,$60,$70,$80,$90,$100,$110,$120,
     3$130,$140,$150/
      DATA JA(1),JA(2),JA(3),JA(4),JA(5),JA(6),JA(7),JA(8),JA(9),
     1 JA(10),JA(11),JA(12),JA(13),JA(14),JA(15),JA(16)
     2 /$160,$170,$180,$190,$200,$210,$220,$230,$240,$250,$260,
     3$270,$280,$290,$300,$310/
   44 WRITE(4,1)
    1 FORMAT(*NO.OF S,LUIN,LOUT,LOUTI,IOP,IOPRW= *)
      CALL HCREAD(4,6,N,LUIN,LOUT,LOUTI,IOP,IOPRW)
      WRITE(4,22)
   22 FORMAT(*ALIM = *)
      READ(4,23) ALIM
      WRITE(4,24)
   24 FORMAT(*DF= *)
      READ(4,23) DF
   23 FORMAT(F10.5)
      IF(LOUT.GT.10.AND.LOUT.LT.16) REWIND LOUT
      IF(LUIN.GT.10.AND.LUIN.LT.16.AND.IOPRW.EQ.1) REWIND LUIN
      NL=N/256
      NR=N-NL*256
      LOOP=NL
      IF(NR.GT.0)LOOP=LOOP+1
      DO 310 K=1,LOOP
      READ(LUIN)DAT1
      READ(LUIN)DAT2
      READ(LUIN)DAT3
      READ(LUIN)DAT4
C
C STORE THE SIGNAL IN APPROPRIATE FILES
C
```

```
      CALL DK854(X,1024,0,IA(K))
  310 CONTINUE
C
C INITIALIZE THE ARRAY B2(.).
C
      DO 320 K=1,LOOP
      CALL DK854(X,1024,1,IA(K))
      IF(K.EQ.LOOP)GO TO 330
      IDX=K+1
      CALL DK854(B11,200,1,IA(IDX))
  330 LL=512
      IF(K.EQ.LOOP)LL=510
      DO 4 J=1,LL,2
      B2(J)=X(J+2)
      B2(J+1)=X(J+3)
    4 CONTINUE
C
C STORE THE ARRAY B2(.)
C
      CALL DK854(Y,1024,0,JA(K))
  320 CONTINUE
      DO 5 J=1,50
      A(J)=0.
      AA(J)=0.
    5 CONTINUE
      A(1)=1.
      P=0.
      FPE = 0.
      IFL = 0
      DO 410 K=1,LOOP
      CALL DK854(X,1024,1,IA(K))
      N2=512
      IF(K.EQ.LOOP.AND.NR.NE.0)N2=2*NR
      DO 6 J=1,N2,2
      P=P+X(J)**2+Y(J+1)**2
    6 CONTINUE
  410 CONTINUE
      P=P/N
C
C AVERAGED POWER OF THE N SAMPLES
C
      WRITE(LOUT,53)P
   53 FORMAT(*P(0)=*,E13.6)
C
C CALCULATE THE P.E.F. COEFFICIENTS
C
      M=50
      DO 13 I=1,M
      WRITE(LOUT,18)I
      MAXCOF=I+1
      JEND=N-I
      MC2=MAXCOF*2
      MC1=MC2-1
```

```
      M2=2*I
      M1=M2-1
      JL=JEND/256
      JE=JEND-JL*256
      LOOP1=JL
      IF(JE.GT.0)LOOP1=LOOP1+1
      JEND1=256
      IF(I.EQ.1)GO TO 8
      DO 510 K=1,LOOP1
      CALL DK854(X,1024,1,IA(K))
      CALL DK854(Y,1024,1,JA(K))
      IF(K.EQ.LOOP1)GO TO 520
      K1=K+1
      CALL DK854(B11,200,1,IA(K1))
      CALL DK854(B21,200,1,JA(K1))
  520 IF(K.EQ.LOOP1.AND.JE.NE.0)JEND1=JE
      JEND2=2*JEND1
      DO 7 J=1,JEND2,2
      B1(J)=B1(J)+AA(M1)*B2(J)+AA(M2)*B2(J+1)
      B1(J+1)=B1(J+1)+AA(M1)*B2(J+1)-AA(M2)*B2(J)
      B2(J)=B2(J+2)+AA(M1)*B1(J+2)-AA(M2)*B1(J+3)
      B2(J+1)=B2(J+3)+AA(M2)*B1(J+2)+AA(M1)*B1(J+3)
    7 CONTINUE
      CALL DK854(X,1024,0,IA(K))
      CALL DK854(Y,1024,0,JA(K))
  510 CONTINUE
    8 SDEM=0.
      SNUMR=0.
      SNUMI=0.
      DO 610 K=1,LOOP1
      CALL DK854(X,1024,1,IA(K))
      CALL DK854(Y,1024,1,JA(K))
      JEND1=256
      IF(K.EQ.LOOP1.AND.JE.NE.0)JEND1=JE
      JEND2=2*JEND1
      DO 9 J=1,JEND2,2
      SNUMR=SNUMR+B1(J)*B2(J)+B1(J+1)*B2(J+1)
      SNUMI=SNUMI+B1(J)*B2(J+1)-B1(J+1)*B2(J)
      X2=B1(J)*B1(J)+B1(J+1)*B1(J+1)
      Y2=B2(J)*B2(J)+B2(J+1)*B2(J+1)
      SDEM=SDEM+X2+Y2
    9 CONTINUE
  610 CONTINUE
      A(MC1)=-2.*SNUMR/SDEM
      A(MC2)=-2.*SNUMI/SDEM
      DO 10 J=2,I
      J2=2*J
      J1=J2-1
      L2=2*(MAXCOF-J+1)
      L1=L2-1
      A(J1)=AA(J1)+A(MC1)*AA(L1)+A(MC2)*AA(L2)
      A(J2)=AA(J2)+A(MC2)*AA(L1)-A(MC1)*AA(L2)
   10 CONTINUE
```

```
      DO 11 J=1,MC2
      AA(J)=A(J)
   11 CONTINUE
      DO 12 J=1,MAXCOF
      J2=2*J
      J1=J2-1
      WRITE(LOUT,16)J,A(J1),A(J2)
   12 CONTINUE
      P=P*(1.-A(MC1)**2-A(MC2)**2)
      IF(I.EQ.1) GO TO 333
      FTEMP = FPE
  333 FPE = P*FLOAT(N+I+1)/FLOAT(N-I-1)
      WRITE(LOUT,17) I, P, FPE
      IF(I.EQ.1) GO TO 13
      IF(IFL.EQ.1) GO TO 777
      IF(FPE.GT.FTEMP) IFL = 1
   13 CONTINUE
C
C  CALCULATE THE SPECTRAL DENSITY ESTIMATE
C
  777 CONTINUE
      DF=DF/128.
      DT=1.
      W=.5/DT
      PI2=6.2831853
      DO 15 J=1,129
      F=(J-65)*DF
      SUMX=1.
      SUMY=0.
      DO 14 JJ=2,MAXCOF
      ANG=PI2*F*(JJ-1)*DT
      COSX=COS(ANG)
      SINX=SIN(ANG)
      JJ2=JJ*2
      JJ1=JJ2-1
      SUMX=SUMX+A(JJ1)*COSX+A(JJ2)*SINX
      SUMY=SUMY+A(JJ2)*COSX-A(JJ1)*SINX
   14 CONTINUE
      DENOM=SUMX*SUMX+SUMY*SUMY
      PF=P/DENOM/W
      L2=2*J
      L1=L2-1
      X(L1)=PF
      X(L2)=0.
   15 CONTINUE
      CALL FMXMIN(X,256,XMAX,XMIN)
C
C NORMALIZE TO THE MAXIMUM FREQUENCY COMPONENT
C
      DO 50 J=1,257
      X(J)=X(J)/XMAX
   50 CONTINUE
C
```

```
C IF IOP=1 CONVERT TO LOG SCALE
C
      IF(IOP.EQ.1)GO TO 40
      CALL FPLOT(X,257,1.,0.)
C
C OUTPUT THE POWER SPECTRUM
C
      WRITE(LOUT,300)
  300 FORMAT(/*  FREQUENCY SPECTRUM (LINEAR SCALE)*/)
      WRITE(LOUT,301)
  301 FORMAT(2X,*NO.*,6X,*FREQUENCY*,8X,*MAGNITUDE*/)
      DO 302 J=1,257,2
      J1=(J+1)/2
      F=(J1-65)*DF
      WRITE(LOUT,303) J1,F,Y(J)
      IF(X(J).GT.0.009.AND.X(J).LT.0.011) WRITE(LOUT1,303) J1,F,X(J)
  302 CONTINUE
  303 FORMAT(2X,I3,5X,E12.5,5X,E12.5)
   16 FORMAT(*A(*,I2,*)=*,F10.5,*+ J*,F10.5)
   17 FORMAT(/*P.E.F.OUT.POWER*,2X,*P(*,I2,*)=*,E12.5,* FPE=*,E12.5/)
C
   18 FORMAT(/*    M=*,I2//,*    P.E.F. COEFFICIENTS*/)
      GO TO 70
   40 DO 80 J=1,257,2
      X(J)=10.*ALOG10(X(J))
      IF(X(J).LT.ALIM) X(J) = ALIM
   80 CONTINUE
      CALL CPLOT(X,257,0.,ALIM,0.,ALIM,0)
      WRITE(LOUT,90)
   90 FORMAT(/*  FREQUENCY SPECTRUM (LOG SCALE)*/)
      WRITE(LOUT,301)
      DO 91 J=1,257,2
      J1=(J+1)/2
      F=(J1-65)*DF
      WRITE(LOUT,303) J1,F,X(J)
      IF(X(J).GT.-21.0.AND.X(J).LT.-19.0) WRITE(LOUT1,303) J1,F,X(J)
      IF(X(J).GT.-11.0.AND.X(J).LT.-9.0) WRITE(LOUT1,303) J1,F,X(J)
      IF(X(J).EQ.0.) WRITE(LOUT1,303) J1,F,X(J)
   91 CONTINUE
   70 CONTINUE
      GO TO 44
      END
 MON
```

```
      PROGRAM PER64
C     ------- -----
C
C   THIS PROGRAM CALCULATES AVERAGED PERIODOGRAMS OVER
C   N B L  BLOCKS. NBL(MAX) = 4. PERIODOGRAMS ARE
C   CALCULATED FOR 64 DATA POINTS EXTENDED WITH ZEROS UP TO
C   128 POINTS. DATA ASSUMED TO BE ON PSEUDOTAPE  L U I N ,
C   AS PER  M E M E P 1  PROGRAM.
C   INPUT PARAMETERS:
C       LUIN=  PSEUDOTAPE WITH DATA BLOCKS
C       LOUT=  L.U. FOR LISTING RESULTS
C       NBLØCK= NUMBER OF DATA BLOCKS AVERAGED.
C       IOPRW=  1    REWIND   L U I N.
C            =  0    DO NOT REWIND  L U I N.
C       IOPSC=  1    FULL SCALE   (-0.5, +0.5)
C            =  2    HALF SCALE   (-0.25, +0.25)
C            =  3    QUARTER SCALE  (-0.125, +0.125)
C            =  4    TENTH SCALE  (-0.05, +0.05)
C       IOPWN=  1    B A R T L E T T   WINDOW
C            =  2    H A M M I N G   WINDOW
C            = OTHER NO(RECTANGULAR) WINDOW.
C
      DIMENSION PER2(128)
      DIMENSION A1(128),A2(128),A3(128),A4(128)
      DIMENSION DAT(1024),COSX(128),PER1(512),PER(128)
      DIMENSION D1(256),D2(256),D3(256),D4(256)
      EQUIVALENCE (D1(1),DAT(1)),(D2(1),DAT(257)),
     $ (D3(1),DAT(513)),(D4(1),DAT(769))
      EQUIVALENCE (D1(1),A1(1)),(D2(1),A2(1)),(D3(1),A3(1)),
     $ (D4(1),A4(1))
C
   44 WRITE(4,1)
    1 FORMAT(*LUIN,LOUT,NBLOCK,IOPRW,IOPSC,IOPWN=*)
      CALL HCREAD(4,6,LUIN,LOUT,NBL,IOPRW,IOPSC,IOPWN)
      WRITE(4,2)
    2 FORMAT(*ALIM=*)
      READ(4,11) ALIM
   11 FORMAT(F10.5)
      LG2N=6
      N=2**LG2N
      N2=2*N
      N2P1=N2+1
      LGFF=7
      NMAX=2**LGFF
      NMAX2=2*NMAX
      NMAX4=4*NMAX
      IF(NBL.GT.4) GO TO 44
      IF(IOPRW.EQ.1) REWIND LUIN
      LOOP = 0
      INOUT=0
      INCT=2
      READ (LUIN) A1
      CALL WINDW1(A1,N2,IOPWN)
      DO 15 I=N2P1,NMAX2
```

```
15 DI(I)=0.
   CALL FFT42G(DI(1),DI(2),INOUT,LOOP,LGFF,INCT,COSX)
   LOOP=LOOP+1
   IF(NBL.EQ.1) GO TO 3
   READ(LUIN) A2
   CALL WINDWI(A2,N2,IOPWN)
   DO 25 I=N2P1,NMAX2
25 D2(I)=0.
   CALL FFT42G(D2(1),D2(2),INOUT,LOOP,LGFF,INCT,COSX)
   IF(NBL.EQ.2) GO TO 3
   READ (LUIN) A3
   CALL WINDWI(A3,N2,IOPWN)
   DO 35 I=N2P1,NMAX2
35 D3(I)=0.
   CALL FFT42G(D3(1),D3(2),INOUT,LOOP,LGFF,INCT,COSX)
   IF(NBL.EQ.3) GO TO 3
   READ(LUIN) A4
   CALL WINDWI(A4,N2,IOPWN)
   DO 45 I=N2P1,NMAX2
45 D4(I)=0.
   CALL FFT42G(D4(1),D4(2),INOUT,LOOP,LGFF,INCT,COSX)
 3 CONTINUE
C
C  ZERO THE PERIODOGRAM ARRAY.
C
   DO 4 I=1,NMAX4
   PERI(I) = 0.
 4 CONTINUE
C
C  CALCULTE THE PERIODOGRAMS FROM THE  D F T- S OF DATA.
C
   DO 6 K=1,NBL
   KN=(K-1)*NMAX
   KN2=2*KN
   DO 5 J=1,NMAX
   J2 = 2*J
   J1 = J2 - 1
   KNJ=KN+J
   KNJ2=KN2+J2
   KNJ1=KN2+J1
   X=DAT(KNJ1)
   Y=DAT(KNJ2)
   PERI(KNJ)= PERI(KNJ)+X*X+Y*Y
 5 CONTINUE
 6 CONTINUE
C
C  AVERAGE THE PERIODOGRAMS.
C
   DO 7 J = 1,NMAX
   NJ=NMAX+J
   N2J=2*NMAX+J
   N3J=3*NMAX+J
   GO TO (10,20,30,40), NBL
10 PER(J)=PERI(J).
  GO TO 50
```

```
   20 PER(J)=(PER1(J)+PER1(NJ))/2.
      GO TO 50
   30 PER(J)=(PER1(J)+PER1(NJ)+PER1(N2J))/3.
      GO TO 50
   40 PER(J)=(PER1(J)+PER1(NJ)+PER1(N2J)+PER1(N3J))/4.
   50 CONTINUE
    7 CONTINUE
C
C  ORDER THE DATA, NORMALIZE, CONVERT TO DB SCALE,PLOT,
C  AND LIST THEM.
C
      CALL ORDAT(PER,NMAX)
      GO TO (17,27,37,47), IOPSC
   17 DO 18 I=1,NMAX
   18 PER2(I)=PER(I)
      NM=NMAX
      GO TO 57
   27 N2MAX=NMAX/2
      DO 28 I=1,N2MAX
   28 PER2(I)=PER(I+33)
      NM=N2MAX
      GO TO 57
   37 N4MAX=NMAX/4
      DO 38 I=1,N4MAX
   38 PER2(I)=PER(I+48)
      NM=N4MAX
      GO TO 57
   47 N10MAX=NMAX/10+1
      DO 48 I=1,N10MAX
   48 PER2(I)=PER(I+57)
      NM=N10MAX
   57 CONTINUE
      CALL FMXMIN(PER2,NM,P2MAX,P2MIN)
      DO 8 I=1,NM
      PER2(I)=PER2(I)/P2MAX
      PER2(I)=10.*ALOG10(PER2(I))
      IF(PER2(I).LT.ALIM) PER2(I)=ALIM
    8 CONTINUE
      CALL FPLOT(PER2,NM,0.,ALIM)
      DO 9 I=1,NM
      I1=I-1
      F=FLOAT(I1)/FLOAT(NM)
      WRITE(LOUT,12) I,F,PER2(I)
    9 CONTINUE
   12 FORMAT(2X,I3,5X,E14.6,5X,E14.6)
      GO TO 44
      END
 MON
```

```
      SUBROUTINE WINDWI(A,N2,IOPWN)
C     ========== ==================
C
C  THIS SUBROUTINE CALCULATES EITHER ONE OF THE
C  FOLLOWING WINDODS:  B A R T L E T T,  H A M M I N G,
C  OR  P A R Z E N  WINDOW. CHOICE IS MADE BY THE
C  VALUE OF THE PARAMETER  I O P W N.
C
      DIMENSION WIN(64),A(1)
      IF(IOPWN.LT.1.OR.IOPWN.GT.2) RETURN
      N=N2/2
      HM=(FLOAT(N)-1.)/2.
      HP=(FLOAT(N)+1.)/2.
      PI=3.1415927
      GO TO (10,20), IOPWN
C
C  B A R T L E T T   WINDOW
C
   10 CONTINUE
      DO 1 I=1,N
      FI=FLOAT(I)-1.
      WIN(I)=1.-ABS((FI-HM)/HP)
    1 CONTINUE
      GO TO 3
C
C  H A M M I N G   WINDOW
C
   20 CONTINUE
      DO 2 I=1,N
      FI=FLOAT(I)-1.
      WIN(I)=.54+.46*COS(PI*(FI/HM-1.))
    2 CONTINUE
    3 CONTINUE
C
C  MULTIPLY AN ARRAY  A  WITH THE WINDOW.
C
      DO 4 I=1,N
      I2=2*I
      I1=I2-1
      A(I1)=A(I1)*WIN(I)
      A(I2)=A(I2)*WIN(I)
    4 CONTINUE
      RETURN
      END
```

APPENDIX F

Proof of the Equivalence Between the Adaptive Ladder Structure and the Forward and Backward FIR Filters

In this appendix we prove the eqivalence between the filter structures in Figs. 6.4 - 1 and 6.4 - 2 by showing that the output series of both filters are identical.

The output signal of the forward PEF in Fig. 6.4 - 1(a) is given by

$$y^f(n+M) = x(n+M) + a_{M,1}x(n+M-1) + \dots + a_{M,M}x(n)$$

$$= \sum_{m=0}^{M} a_{M,m}x(n+M-m) \quad , \quad n = 1,2,\dots,N-M \, , \qquad \text{(F.1(a))}$$

where $a_{M,0} \equiv 1$. Inverting the order in the last summation, we obtain

$$y^f(n+M) = \sum_{m=0}^{M} a_{M,M-m}x(n+m) \, , \quad n = 1,2,\dots,N-M \, . \qquad \text{(F.1(b))}$$

Similarly, at the output of the backward filter in Fig. 6.4 - 1(b) we have

$$y^b(n+M) = a^*_{M,M}x(n+M) + a^*_{M,M-1}x(n+M-1) + \dots + x(n)$$

$$= \sum_{m=0}^{M} a^*_{M,M-m}x(n+M-m) = \sum_{m=0}^{M} a^*_{M,m}x(n+m) \, , \; n = 1,2,\dots,N-M. \qquad \text{(F.2)}$$

We shall prove by induction that the equations (F.1) and (F.2) are also valid for the circuit in Fig. 6.4 - 2.

The outputs of the first order filters in Fig. 6.4 [illegible] 2 are

$$y_1^f(n+1) = y_o^f(n+1) + a_{1,1}y_o^b(n+1) = x(n+1) + a_{1,1}x(n), \quad n=1,\ldots,N-1, \tag{F.3}$$

and

$$y_1^b(n+1) = y_o^b(n) + a^*_{1,1}y_o^f(n) = x(n) + a^*_{1,1}x(n+1), \quad n=1,\ldots,N-1. \tag{F.4}$$

These equations are identical to Eqs. (F.1) and (F.2) for M=1. We now make an assumption that equations (F.1) and (F.2) are valid for (k-1)th stage of the ladder, i.e., we have

$$\begin{aligned} y_{k-1}^f(n+k-1) &= y_{k-2}^f(n+k-1) + a_{k-1,k-1}y_{k-2}^b(n+k-1) \\ &= \sum_{m=0}^{k-1} a_{k-1,k-1-m}x(n+m), \qquad n = 1,2,\ldots,N-k+1, \end{aligned} \tag{F.5}$$

and

$$\begin{aligned} y_{k-1}^b(n+k-1) &= y_{k-2}^b(n+k-2) + a^*_{k-1,k-1}y_{k-2}^f(n+k-2) \\ &= \sum_{m=0}^{k-1} a^*_{k-1,m}x(n+m), \qquad n = 1,2,\ldots,N-k+1. \end{aligned} \tag{F.6}$$

Starting with Eq. (F.5) we increment the order of the stage from k-1 to k and have for the forward output series

$$\begin{aligned} y_k^f(n+k) &= y_{k-1}^f(n+k) + a_{k,k}y_{k-1}^b(n+k) \\ &= \sum_{m=0}^{k-1} a_{k-1,k-1-m}x(n+1+m) + a_{k,k}\sum_{m=0}^{k-1} a^*_{k-1,m}x(n+m), \quad n=1,\ldots,N-k. \end{aligned} \tag{F.7}$$

We now break the first summation on the right-hand side of Eq. (F.7) into sum of the first k-1 terms and the last term. Similarly, in the second summation, we take the first term and the sum of the last k-1 terms. Thus we have

$$y_k^f(n+k) = \sum_{m=0}^{k-2} a_{k-1,k-1-m} x(n+1+m) + a_{k-1,0} x(n+k)$$

$$+ a_{k,k} a^*_{k-1,0} x(n) + a_{k,k} \sum_{m=1}^{k-1} a^*_{k-1,m} x(n+m)$$

$$= x(n+k) + \sum_{m=1}^{k-1} (a_{k-1,k-m} + a_{k,k} a^*_{k-1,m}) x(n+m) + a_{k,k} x(n),$$

$$n = 1,2,\ldots,N-k, \qquad (F.8)$$

where we have changed the limits in the first sum from 0 and k-2 into 1 and k-1, respectively.

Recalling equation 2.4 – 12 we see that the term in parenthesis is simply $a_{k,k-m}$, so that we have

$$y_k^f(n+k) = x(n+k) + \sum_{m=1}^{k-1} a_{k,k-m} x(n+m) + a_{k,k} x(n) =$$

$$= \sum_{m=0}^{k} a_{k,k-m} x(n+m) , \quad n = 1,2,\ldots,N-k. \qquad (F.9)$$

Similarly, from Eq. (F.6) we have

$$y_k^b(n+k) = \sum_{m=0}^{k} a^*_{k,m} x(n+m) , \quad n = 1,2,\ldots,N-k. \qquad (F.10)$$

For the last stage, M, we simply replace k by M in Eqs. (F.9) and (F.10) and the proof by induction is completed.