SPECTRUM ESTIMATION USING ADAPTIVE FILTERS

by

Emin ANARIN

Submitted to the Faculty of Engineering
in Partial Fulfillment for the
Requirements of the Degree of

MASTER OF SCIENCE
in
ELECTRICAL ENGINEERING

BOĞAZİÇİ UNIVERSITY
February 1983

This thesis has been approved by:

Doç. Dr. Yorgo ISTEFANOPULOS
(Thesis Supervisor)

Doç. Dr. Erdal PANAYIRCI

Y. Doç. Dr. Uğur YARAMANCI

Date: February 1983

ACKNOWLEDGEMENTS

This thesis has been prepared for the partial fulfillment of the requirements of Boğaziçi University, School of Engineering for the degree of Master of Schience in Electrical Engineering.

I would like to express my sincere gratitude to those who have helped me during the development of this study, especially to my thesis advisor Doç. Dr. Yorgo ISTEFANOPULOS for his invaluable aid.

I also wish to extend my sincere thanks to Doç. Dr. Bülent SANKUR for his encouragement and always willingness to be of help during this thesis study.

I am again grateful to Mustafa AKTAR for his invaluable aid.

ABSTRACT

In this work, the adaptive discrete-time, linear non-recursive filters (or estimators) designed by the Least Mean Square (LMS) algorithm are investigated.

The development of adaptive techniques for estimating the parameters of sinusoidal signals in white noise is important in many applications. Therefore, a signal enhancing technique for statistically stationary signals based on conventional Least Mean Square (LMS) adaptive filtering and some other newly developed procedures of adaptive spectral estimation of discrete time series are presented in this thesis.

An adaptive filter configuration known as the Adaptive Line Enhancer (ALE) originally suggested by Widrow [1] for the detection of sinusoidal signals in wide band noise is studied in detail. New expressions related to the decorrelation parameter for the cases of one, two and multiple sinusoids are obtained.

This thesis also investigates the method in [10] for eliminating sinusoidal or other periodic interference corrupting a signal. This task is typically accomplished by expli-

citly measuring the frequency of the interference and implementing a notch filter at that frequency.

For the colored noise case, the optimal filter length for ALE is obtained by maximizing the SNR ratio of ALE.

The → estimation in LMS algorithm will be better if the estimates of the tap gain coefficients are better. Better estimates are obtained by running the LMS algorithm longer. Therefore, it is useful to have a rapidly convergent algorithm and so called Ladder or Lattice filter. For that reason we introduce the Lattice Filter Implementation of the general ALE as in [36]. Also a class of stable and efficient recursive lattice methods for linear prediction depending on the choosen reflection coefficients. Computer simulations are also performed to discuss everything in the thesis.

ÖZETÇE

Bu tezde küçük kareler (KK) algoritması ile çalışan uyarlamalı, kesikli zamanlı, doğrusal, transversal süzgeçler incelenmektedir.

Beyaz gürültü içindeki sinüsoidal işaretlerin parametrelerinin kestiriminde kullanılan uyarlamalı tekniklerdeki gelişmeler, birçok uygulamada önemli olmaktadır. Bu sebeble; bu tezde istatistiksel bakımdan durağan işaretlerin küçük kareler ile uyarlamalı süzgeçlenmesi ve diğer yeni gelişen kesikli zamanlı serilerin uyarlamalı görünge kestirimleri sunulmaktadır.

Widrow tarafından beyaz gürültü içindeki sinüsoidal işaretlerin sezmesinde kullanılan, uyarlamalı çizgi kuvvetlendirici olarak bilinen, bir çeşit uyarlamalı süzgeçi üzerinde çalışmaktadır. Bir, iki ve çoklu sinüsoidal işaretler için, yeni ilintisizlik değiştirgen ifadeleri elde edilmektedir.

Bu tez ayni zamanda sinüsoidal ve diğer dönemsel girişimleri eleme methodlarını incelemektedir. Bu girişimin sıklığının ölçülmesi ve bulunan sıklıkta notch süzgeç gerçekleştirilmesi ile elde edilebilmektedir.

İşaretin gürültüye oranını enbüyükleyerek, renklendirilmiş gürültü ortamında, uyarlamalı çizgi kuvvetlendiricisinin eniyi süzgeç boyu elde edilmektedir.

Şayet kazanç katsayıları mükemmel ise, küçük kareler kestirimide mükemmel olacaktır. İdeal kestirim, (KK) algoritmasını uzun süre geçiştirmekle elde edilebilmektedir. Bu sebebten, örü süzgeç diye adlanırılan ve süratle yakınsayan bir algoritma kullanışlı olmaktadır. Uyarlamalı çizgi kuvvetlendiricisinin örü süzgeç olarak gerçekleştirilmesi yapılabilmektedir. Ayni zamanda verimli yansıma katsayıları ve onların özyineli denklemleri verilmektedir. Bilgisayar benzetimleri keza tez içinde yer alan düşüncelere fikir vermek amacıyla yapılmaktadır.

TABLE OF CONTENTS

INTRODUCTION

Carrier detection and estimation is based on the spectrum or power spectral density (PSD). Estimation of the power spectral density or simply spectrum of discretely sampled deterministic and stochastic processes is usually based on procedures employing the Fast Fourier Transform (FFT). This approach to spectrum analysis is computationally efficient and produces reasonable results for a large class of signal processes. In spite of these advantages there are several inherent performance limitations of the FFT approach. The most important limitation is that of the frequency resolution, i.e., the ability to distinguish the spectral responses of two or more signals. The frequency resolution in hertz is roughly the receiprocal of the time interval in seconds over which sampled data is available. These performance limitations of the FFT approach are particularly troublesome when analyzing short data records. Short data records occur frequently in practice because many measured processes are brief in duration or have slowly time varying spectra that may be considered constant only for short records.

In an attempt to alleviate the inherent limitations of the FFT approach many alternative spectral estimation procedures have been proposed within the last decade.

Modern spectrum estimation techniques are based on modeling of the data by a small set of parameters. When the model is an accurate representation of the data, spectral estimates can be obtained whose performance exceed that of the classical FFT, estimator. The improvement in performance is manifested by higher resolution and a lack of side lobes. It should also be emphasized that in addition to an accurate model of the data, one must base the spectral estimator on a good estimator of the model paramters. Usually this entails a maximum likelihood parameter estimator. If the model is inappropriate, as in the case of an AR model for an AR process with additive observation noise, poor (biased) spectral estimates will result. If the model is accurate but a poor statistical estimator of the parameters is employed as in the case of the ARMA spectral estimate using the modified Yule-Walker equations poor (inflated variance) spectral estimates will also result.

However, the most common analysis techniques have been the autocorrelation and covariance methods of linear prediction in which the observed signal is modeled as an AR (all pole) process. As typically implemented these are block data structured approaches which create a whitening or inverse filter for the available data block. These techniques also assume that the data are stationary during the time window in which autocorrelation measurement are taken. However, signal statistics may not remain stationary. Also instead of block processing time series data in the method of linear predictior the inverse filter can be implemented as a continuously updated all zero adaptive transversal or adaptive lattice structure. These structures have received considerable attention recently. The usual approach to their derivation has been to use a noisy gradient descent algorithm to adapt the filter coefficients toward their "optimal" values under a minimum mean square error perfromance criterion.

Adaptive filter is a learning machine. In the design of optimum systems, a complete knowledge of the model is assumed. In most realistic situations such a priori knowledge is not available and one faces the design of optimum systems with an incomplete model knowledge. Since the design is done while data is being taken, it constitutes an adaptive problem In adaptive problems we want to build a system (filter) to operate efficiently in an unknown or changing environment. The adaptive systems have the unique capability of operating without a total priori knowledge of their input signal statistics and thus have been of continuing interest to scientists for the last years.

The traditional form of the adaptive LMS filter is the tapped delay-line prediction-error filter. The function of the LMS algorithm is to adjust the weights adaptively in the absence of the a priori knowledge of the input statistics toward their optimum values. In this respect LMS filters are adaptive Wiener filters [30] or as in [36] adaptive line enhancers (ALE).

ALE is a prefilter or an adaptive digital transversal filter that is designed to supress broad-band components in its input while passing narrow band components with little attenuation.

In Chapter 1, the adaptive transversal filter is introduced. The fundamentals of discrete-time transversal filters and the related Wiener filter theory results are investigated. The operation of LMS filters with stationary stochastic inputs is studies and the recursive equation of the weights is obtained.

In Chapter 2, the steady state behavior of the adaptive line enhance (ALE) and its implementation for detecting the sinusoidal signals in broad band noise is analyzed. The decorrelation parameter Δ is analyzed and its optimum value for the cases of one, two and multiple sinusoids is obtained.

In Chapter 3, a class of notch filters is derived to eliminate sinusoidal or other periodic interferences corrupting a signal, while analyzing LMS-ALE adaptive notch filter, the optimum value of Δ which was found in the previous chapter is used. We also investigate a constrained recursive adaptive filter and its advantages. At the end of this chapter, we introduce sequential regression (SER) adaptive notch filter and we made a comparison between LMS and SER adaptive notch filters.

In Chapter 4 we investigate the optimal filter length for ALE by considering different methods. First method is based on the maximization of the SNR ratio for white noise case. Second method is studied by means of the weights of the ALE. The properties of the weights of the ALE are used to determine the detection system. The optimal filter length is found so as to optimize the detection performance of ALE. The last method is based again on the maximization of the SNR ratio of ALE for the colored noise case.

In Chapter 5 we introduce the adaptive lattice filter configuration. We also point out the advantages and necessities of using the adaptive lattice filter. The derivation of Δ step predictor in lattice form is given. Also we investiga the case of whitening or inverse filtering.

In Chapter 6 a class of reflection coefficients is discussed. Also their effects on the stability of the filter is investigated.

In Chapter 7 we analyze the recursive estimation of the reflection coefficients. The aim of this chapter is to make an adaptive filter very sensitive to the changes in the signal. Two methods dealing with this situation are presen-ted.

Chapter 8 summarizes the conclusions of this study and gives some suggestions for further research.

CHAPTER 1

ADAPTIVE TRANSVERSAL FILTER

1.1. INTRODUCTION

The term "filter" is often applied to any device or system that processes incoming signals or other data in such a way as to eliminate noise or smooth the signals or identify each signal as belonging to a particular class or predict the next input signal from moment to moment.

In the design of optimum systems a complete knowledge of the system model is assumed. In most realistic situations, however, such a priori knowledge is not available and this fact necessitates the design of optimum systems with an incomplete model knowledge. Since the design is done while data is being taken, it constitutes an adaptive problem. In adaptive problems we want to build a system (filter) to operate efficiently in an unknown or changing environment.

This thesis presents an approach to signal filtering using an adaptive filter that is in some sense self-designing (really self optimizing). The filter to be considered here consists of a tapped delay line with variable weight (variable gain) whose input signals are the signals at the delay line taps, a summer to add the weighted signals and a mechanism to adjust them automatically. The filter is adjusted so as to provide the best estimation of a given signal as a weighted sum of a set of inputs. This is achieved by continuously updating the filter weights in such a way as to reduce the average estimation error power in each iteration.

Among the stochastic approximation methods used in adaptive filtering the simplest and the most commonly used is least mean squares (LMS) algorithm in which the weights are updated in the negative direction of the gradient of the square of a single error sample. Two kinds of processes take place in the adaptive filter, training and operating. The training (adaptation process) is concerned with adjusting the weights. The operating process consists in forming the output signal as a weighted sum of the delay line tap signals using the weights resulting from the training process.

1.2. THE FILTER STRUCTURE

The analysis of the adaptive filter can be developed by considering the adaptive linear systems as shown in Figure 1.2.1.

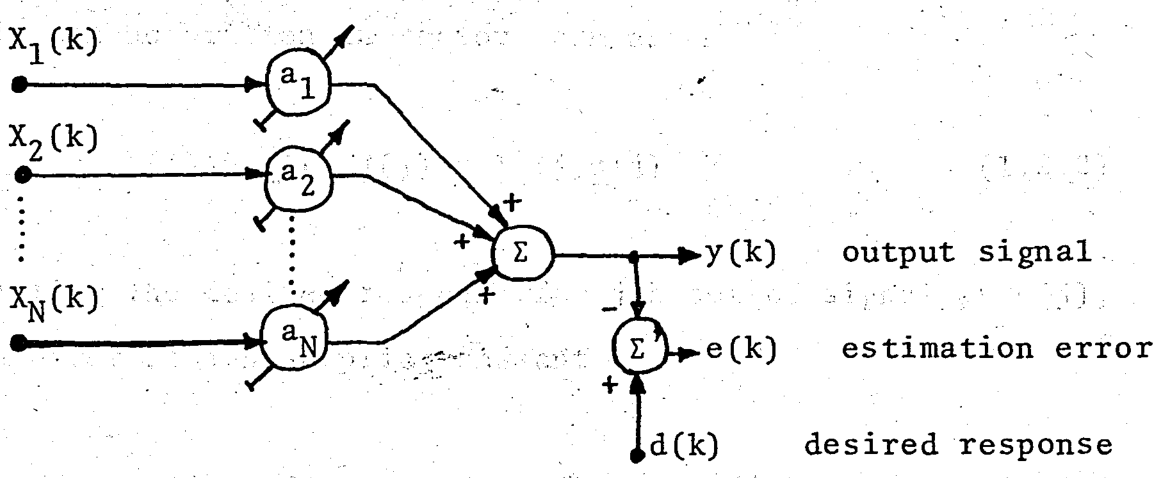


FIGURE 1.2.1. The Adaptive Linear Combiner.

In the systems of Figure 1.2.1 a set of stationary input signals is weighted and summed to form an output signal. The input signals in the set are assumed to occur simultaneously and discretely in time.

The set of input signals at the jth sampling instant are given by

$$X^T(j) = \left[X_1(j) \; \dots \; X_N(j) \right]$$

The set of weights is designated by the vector

$$a^T(j) = \left[a_1(j), \; a_2(j) \; \dots \; a_N(j) \right] .$$

The jth output signal

$$Y(j) = \sum_{i=1}^{N} a_i(j) X_i(j) \qquad (1.2.1.)$$

This can be written in vector form as

$$Y(j) = \underline{a}^T(j)\underline{X}(j) = \underline{X}^T(j)\underline{a}(j) \quad (1.2.2)$$

Denoting the desired response for jth set of signal as d(j), the error at jth sampling instant

$$e(j) = d(j) - y(j) = d(j) - \underline{a}^T(j)\underline{X}(j) \quad (1.2.3)$$

The square of this error

$$e^2(j) = d^2(j) - 2d(j)X^T(j)a(j) + a^T(j)X(j)X^T(j)a(j) \quad (1.2.4)$$

Assuming that d(j) and $\underline{X}(j)$ are stationary processes, the mean square error (MSE) is given by

$$\bar{e}^2(j) = E\{e^2(j)\} = E\{d^2(j)\} - 2P^Ta + a^TRa \quad (1.2.5)$$

where P is the cross correlation vector between X(j) and d(j) given by

$$P = E\{d(j)X(j)\} \quad (1.2.6)$$

and R is the symmetric and positive definite input correlation matrix

$$R = E\{X(j)X^T(j)\} \tag{1.2.7}$$

It can be observed from (1.2.5) that the MSE is a quadratic function of the weights.

The MSE performance function may be visualized as a bowl shaped surface namely, a parabolic function of the weight variables. The LMS adaptive process constitutes of continously searching the minimum point of this parabolic surface. This can be accomplished by means of the method of steepest descent. The method of steepest descent uses the gradient of the performance function in seeking its minimum. The gradient at any point on the parabolic surface may be obtained by differentiating the MSE function of equation (1.2.5) with respect to the weight vector. The gradient is

$$\nabla(\bar{e}^2(j)) = - 2P + 2Ra \tag{1.2.8}$$

The optimal weight vector a* which yields the minimum MSE (MMSE) is obtained by setting the gradient to zero:

$$a^* = R^{-1}P \tag{1.2.9}$$

Equation (1.2.9) is the Wiener-Hopf equation in the discrete-time case. An expression for the minimum MSE may be obtained by substituting (1.29) into (1.2.5)

$$e_{min} = E\{d^2(j)\} - P^T a^* \tag{1.2.10}$$

Defining

$$V = a - a^* \tag{1.2.11}$$

as the weight error vector and inserting (1.2.10) into (1.2.5) and using (1.2.11) one can express the MSE as

$$e = e_{min} + V^T R V \tag{1.2.12}$$

Since R is symmetric and positive definite, it can be expressed as

$$R = Q \Lambda Q^{-1} = Q \Lambda Q^T \tag{1.2.13}$$

where Q is the orthonormal modal matrix of R, and Λ is the diagonal matrix which consists of the eigenvalue of R which are real and positive:

$$\Lambda = \text{diag}\ (\lambda_1, \lambda_2, \ldots \lambda_N) \tag{1.2.14}$$

Hereafter the matrix $Q^T = Q^{-1}$ will be used to transform the vectors X(j), a,V into the "primed coordinates" whenever it will be convenient to do so.

The transformed wiehgt error vector is given by

$$V' = Q^T V \qquad (1.2.15)$$

Substituting (1.2.15) in (1.212) and using (1.2.13) the MSE is obtained as

$$e = e_{min} + V'^T \Lambda V' \qquad (1.2.16)$$

or

$$e = e_{min} + \sum_{p=1}^{N} \lambda_p V_p'^2 \qquad (1.2.17)$$

where V_p' is the p'th entry of V'.

1.3. THE LMS ALGORITHM

The purpose of the adaptation process is to find an exact or at least an approximate solution of the Wiener-Hopf equation (1.2.9). One way of finding the optimum weight vector is simply to solve (1.2.9). Although this solution is generally straight forward, it could present serious computation problems when the number of weights N is large and when input data arrival rates are high. In addition to the necessity of inverting an NxN matrix, this method may require as many as n(n+1)/2 autocorrelation and cross correlation measurements to be made in order to obtain the element of R and P.

The LMS algorithm, first proposed by Widrow and Hopf [33] is a well-known stochastic approximation algorithm which resembles the steepest descent method. The algorithm utilizes the estimated gradient for updating, since true gradients are not avilable in adaptive filtering. The estimate of the gradient in the LMS algorithm is the gradient of the square of the single error sample at the instant j.

One method for obtaining the estimated gradient of the MSE function is to take the gradient of a single time sample of the squared error, that is

$$\nabla(\bar{e}^2(j)) = \nabla[e^2(j)] = 2e(j)\nabla(\varepsilon(j)) \qquad (1.3.1)$$

From (1.2.3) we have

$$\nabla(e(j)) = \nabla[d(j) - a^T(j)X(j)] = - X(j) \qquad (1.3.2)$$

Thus

$$\tilde{\nabla}(\bar{e}^2(j)) = - 2e(j)X(j) \qquad (1.3.3)$$

The gradient estimate of (1.3.3) is unbiased as will be shown by the following argument. For a given weight vector $\bar{a}(j)$ the expected value of the gradient estimate is

$$E[\tilde{\nabla}(\bar{e}^2(j))] = -2\ E\{X(j)(d(j) - X^T(j)a(j))\}$$

$$= -\ 2\ [P - Ra] \qquad (1.3.4)$$

Comparing (1.2.8) and (1.3.4) we see that

$$E[\tilde{\nabla}(\tilde{e}^2(j))] = \nabla[\bar{e}^2(j)] \qquad)1.3.5)$$

and therefore for a given weight vector, the gradient estimate $\tilde{\nabla}[\bar{\varepsilon}^2(j)]$ is unbiased.

When using the LMS algorithm, changes in the weight vector occur along the direction of the estimated gradient vector. Accordingly,

$$a(j+1) = a(j) + \mu\ \tilde{\nabla}\ (\bar{e}^2(j)) \qquad (1.3.6)$$

where

$a(j) \triangleq$ weight vector before adaptation

$a(j+1) \triangleq$ weight vector after adaptation

$\mu \triangleq$ scalor constant controlling rate of convergence and stability.

Therefore the filter weights can be computed using (1.3.6). Further details with the filter parameters will be taken in the next chapters.

1.4. ADAPTIVE LINE ENHANCER (ALE)

In recent years there has been increasing interest in adaptive filters for various signal processing applications. Here we describe an adaptive device, known as an adaptive line enhancer (ALE), for detecting sinusoidal signals in wide-band noise. The ALE was first proposed by Widrow [1] and since then has been studied by Zeidler [7], Griffiths [2], Treichler [12], Glover [10], Nehorai and Malah [62], and others.

The generally used form of the ALE is shown in Figure 1.4.1. In the ALE the second or reference input, instead of being separately derived, is a delayed version of the input signal. The delayed input is processed with an adaptive transversal filter and subtracted from the original input signal to produce the error signal. The weighting coefficients of the filter are recursively adjusted by means of (1.3.6) so as to minimize the expected error power.

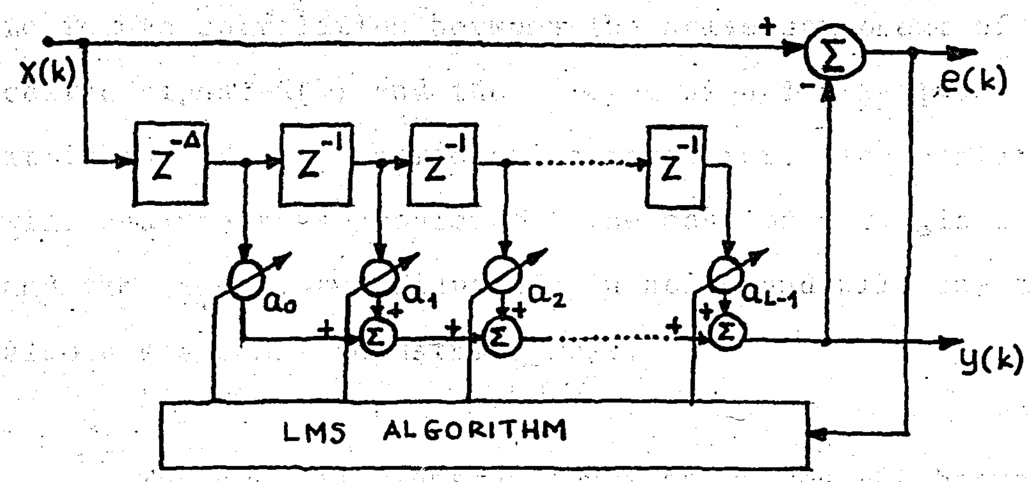


FIGURE 1.4.1. Block Diagram of ALE.

The input signal is assumed to be of the form

$$X(k) = S(k) + n(k) \qquad k = 0,1,... \qquad (1.4.1)$$

where the signal is the sum of a number os sinusoids

$$S(k) = \sum_{i=1}^{M} C_i \, \mathrm{Sin}(w_i k + \phi_i)$$

and n(k) is a zero mean white noise with

$$E\{n(k)n(m)\} = \sigma_o^2 \delta(k-m)$$

Since a reference signal in ALE is obtained by delaying the received signal, therefore

$$X(k-\Delta) = S(k-\Delta) + n(j-\Delta) \qquad (1.4.2)$$

fro some $\Delta \geq 1$. Actually the choice of $\Delta = 1$ is sufficient to remove correlation between the noise component of the received signal X(k) and the delayed signal $X(k-\Delta)$ and therefore it is called the decorrelation parameter. In Chapter 2, we will analyze this problem for the case of a single sinusoid and the case of two sinusoids in noise and will see that Δ also has a phase adjustment role.

The overall transfer function of the ALE between the input and the error output is

$$G(z) = 1 - H(z) \qquad (1.4.3)$$

where

$$H(z) = \sum_{k=0}^{L-1} a_k \, Z^{-(\Delta+k)}$$

and the $\{a_k\}$ are the estimated tap-gain coefficients obtained via the LMS algorithm.

If we consider the inverse of the overall transfer function by putting $Z = e^{j\omega}$

$$P(\omega) = \frac{1}{1-H(\omega)} \qquad (1.4.4)$$

The value of that yields $\max_{\omega}|P(\omega)|^2$ is taken as the estimate of the frequency of the sinusoid. That is ALE is also used as a carrier detector.

CHAPTER 2

DERIVATION OF OPTIMAL VALUE OF Δ

2.1. INTRODUCTION

During the operation of the adaptive line enhance (ALE) the delay causes decorrelation between the noise components of the input data in two processor channels while introducing a simple phase difference between the sinusoidal components. The adaptive filter responds by forming a transfer function equivalent to that of a narrowband filter centered at the frequency of the sinusoidal components. The noise component of the delayed input is rejected while the phase difference of the sinusoidal components is readjusted so that they cancel each other at the summing function, producing a minimum error signal composed of the noise component of the instantaneous input data alone.

In the use of the ALE to detect sinusoidal singals in uncorrelated or white noise any value of Δ of delay can be choosen. But in [36] Δ has a phase adjustment role which is better served by a choice $\Delta > 1$. It is normal to take $\Delta > 1$ becuase the coice of $\Delta = 1$ is only sufficient to remove correlation between the white noise component of the original observed waveform y(t) and that of the delayed reference waveform $y(t-\hat{\Delta})$.

The ALE can also be used to detect sinusoidal signals in correlated or colored noise. In this case it is often necessary to choose a large value of Δ to ensure decorrelation between the noise components and phase adjustment between the sinusoidal components in the two processor channels.

2.2. THE FREQUENCY RESPONSE OF ALE

From [7] it is seen that the frequency response of the steady state ALE which will be denoted by H(ω) can be expressed as follows:

$$H(\omega) = \sum_{k=0}^{L-1} a^*_k \, Z^{-(\Delta+k)} \Big|_{Z=e^{j\omega}} \qquad (2.2.1)$$

where a^*_k is the Wiener-Hopf solution of matrix equation.

The form of the assumed solution for a^*_k for N sinusoidal inputs of the form is given

$$a^*_k = \sum_{n=1}^{2N} A_n e^{j\omega_n k} \qquad (2.2.2)$$

where for notational convenience W_{n+N} is defined as $-W_n$ $(n=1,2,\ldots,N)$, the W_{n+N} are thus the negative frequency components of the input sinusoids. In [7] the equation which is related to the A_n was given as follows;

$$A_r + \sum_{\substack{n=1 \\ n\neq r}}^{2N} \gamma_{rn} A_n = \frac{e^{jW_r\Delta}}{L+2\delta_o^{\,2}/\delta_1^{\,2}} \qquad r=1,2,\ldots,2N \tag{2.2.3}$$

It is 2N equations in the 2N constants $A_1, A_2, \ldots, A_{2N}$. In (2.2.3) $\sigma^2{}_{n+N}$ is defined as $\sigma^2{}_n$ $(n=1,2,\ldots,N)$ and γ_{rn} is given by (2.2.4)

$$\gamma_{rn} = \frac{1}{L+2\dfrac{\delta_o^{\,2}}{\delta_r^{\,2}}} \; \frac{1 - e^{j(W_n-W_r)L}}{1 - e^{j(W_n-W_r)}} \tag{2.2.4}$$

A number of interesting analytic properties of a^*_k can be observed through (2.2.2) and (2.2.4). First (2.2.2) implies that when the input to the ALE consists of N sinusoids and additive white noise, the mean steady state impulse response of the ALE can be expressed as a weighted sum of the input sinusoids. From (2.2.4) it is seen that the coefficients γ_{rn} are proportional to $(1 - e^{j(W_n-W_r)L})/(1 - e^{j(W_n-W_r)})$ which is the L point Fourier transfrom of $\exp(jW_n k)$ evaluated at W_r. Note that from the form of γ_{rn} it follows that $A_{n+N} = A_n$ $(n=1,2,\ldots,N)$. This relation is of course necessary to ensure that a^*_k is real.

2.3. CHOICE OF Δ FOR ONE SINUSOIDAL SIGNAL

For one sinusoidal we can simplify the (2.2.2), (2.2.3) and (2.2.4) as follows

$$a^*_k = A_1 e^{jW_1 k} + A_2\, e^{-jW_1 k} \tag{2.3.1}$$

$$A_1 = \bar{A}_2 = \frac{1}{[L+2\delta_o^2/\delta_1^2][1 - |\gamma_{12}|^2]} \left[e^{jW_1\Delta} - e^{-jW_1\Delta}\gamma_{12}\right] \tag{2.3.2}$$

$$\gamma_{12} = \frac{1}{L+2\delta_o^2/\delta_1^2}\,\frac{1 - e^{-2jW_1 L}}{1 - e^{-2jW_1}} = \frac{1}{L+2\delta_o^2/\delta_1^2}\,\frac{\operatorname{Sin}W_1 L}{\operatorname{Sin}W_1}\, e^{-jW_1(L-1)} \tag{2.3.3}$$

Since there is one sinusoidal therefore the transfer function of ALE at W_q frequency must be maximum, i.e., unity gain. We can formulate this situation as follows:

$$H(w)\Big|_{w=w_1} \rightarrow 1$$

or

$$\min|1 - H(w_1)| \text{ and } \lim_{w\to w_1} |1 - H(w)| \cong 0 \tag{2.3.4}$$

Now let us find the expression for $H(w_1)$. From (2.2.1) we have

$$H(w)_{w=w_1} = \sum_{k=0}^{L-1} (A_1 L + \bar{A}_1\, e^{-j2W_1 k_o})\, e^{-jW_1\Delta} \tag{2.3.5}$$

By putting the values of A_1 and γ_{12}, we have

$$H(w_1) = A_1 L + \bar{A}_1 \frac{1 - e^{-j2W_1 L}}{1 - e^{-j2W_1}} e^{-jW_1 \Delta}$$

$$= \frac{1}{L+2\delta_o^2/\delta_1^2} \frac{e^{-j2W_1\Delta}}{1 - |\gamma_{12}|^2} \left(e^{jW_1\Delta} - \gamma_{12} e^{-jW_1\Delta}\right)L$$

$$+ \left(e^{-jW_1\Delta} - \gamma_{12} e^{jW_1\Delta}\right) \frac{1 - e^{-j2W_1 L}}{1 - e^{-j2W_1}} \qquad (2.3.6)$$

Since

$$\gamma_{12} = \frac{1}{L+2\delta_o^2/\delta_1^2} \frac{1 - e^{j2W_1 L}}{1 - e^{j2W_1}} \qquad (2.3.7)$$

$$|\gamma_{12}|^2 = \frac{1}{(L+2\delta_o^2/\delta_1^2)^2} \frac{1 - \mathrm{Cos}2W_1 L}{1 - \mathrm{Cos}2W_1} \qquad (2.3.8)$$

Let

$$T = L + 2\ \delta_o^2/\delta_1^2$$

$$P = \frac{1 - e^{-j2W_1 L}}{1 - e^{-j2W_1}}$$

$$K = \frac{1 - \mathrm{Cos}2W_1 L}{1 - \mathrm{Cos}2W_1}$$

Then

$$H(w_1) = \frac{T}{T^2-K} \quad (L - K) + e^{-j2W_1\Delta} P(1 - \frac{L}{T}) \tag{2.3.9}$$

Let

$$T^2-K = \beta$$

$$R = \frac{SinW_1L}{SinW_1}$$

Then

$$R_e\{H(W_1)\} = \frac{T}{\beta} [(L-K) + R(1 - \frac{L}{T}) CosW_1(2\Delta + L-1)] \tag{2.3.10}$$

$$I_m\{H(W_1)\} = \frac{-R}{\beta} (T - L) SinW_1(2\Delta + L-1) \tag{2.3.11}$$

After some manipulations we have

$$R_e\{H(W_n)\} = \frac{4\delta_o^2}{\delta_1^2} \frac{Cos(2\Delta + L-1)W_1 \; SinW_1L \; SinW_1}{\beta (1 - Cos2W_1)}$$

$$- \frac{2L}{\beta} \frac{\delta_o^2}{\delta_1^2} - \frac{4}{\beta} (\frac{\delta_o^2}{\delta_1^2})^2 + 1 \tag{2.3.12}$$

and

$$I_m\{H(W_1)\} = - \frac{4SinW_1 \; SinW_1L \; SinW_1(2\Delta + L-1)}{\beta (1 - Cos2W_1)} (\frac{\delta_o^2}{\delta_1^2}) \tag{2.3.13}$$

At the frequency W_1. $J(W_1)$ must be real. Therefore the imaginary part of $H(W_1)$ and $\frac{d}{d\Delta}$ $(1 - H(W_1))$ must be zero to minimize $[1 - H(W_1)]$. This condition gives us the following relation.

$$\mathrm{Sin}W_1(2\Delta + L-1) = \mathrm{Sin}k\pi \qquad (2.3.14)$$

and

$$W_1(2\Delta + L-1) = k\pi \qquad \text{for } k=1,2,\ldots \qquad (2.3.15)$$

Also in [36] the same condition was demonstrated by a different procedure which can be summarized as follows.

Consider the average error variance expression

$$V = \frac{1}{2\pi}\,\delta_o^2 \int_o^{2\pi} |1 - H(W)|^2 dw + \delta_1^2 |1 - H(W_1)|^2 \qquad (2.3.16)$$

To find the minimum value of V we must compute the stationary points given by

$$0 = \frac{dv}{d\Delta} = \delta_o^2 \frac{1}{2\pi} \int_o^{2\pi} R_e\{[1-\bar{H}(W)]\frac{d}{d\Delta}[1-H(W)]\}d$$

$$+ 2\delta_1^2 R_e\{[1-\bar{H}(W_1)]\frac{d}{d\Delta}[1 - H(W_1)]\} \qquad (2.3.17)$$

At this point we make an approximation. Assume that δ_1^2 is large compared to δ_o^2 and compute the stationary points from the second term only, giving

$$R_e\{[1-\tilde{H}(W_1)]\frac{d}{d\Delta}[1-H(W_1)]\} = R_e[1-H(W_1)]\frac{d}{d\Delta} R_e[1-H(W_1)]$$

$$+ I_m[1-H(W_1)]\frac{d}{d\Delta} I_m[1-H(W_1)] = 0 \qquad (2.3.18)$$

Combining (2.3.12), (2.3.13) and (2.3.18) yields

$$\frac{1}{\beta^2} \quad 4 \left(\frac{\delta_o^2}{\delta_1^2}\right)^2 + 2L \frac{\delta_o^2}{\delta_1^2} \quad 4 \frac{\delta_o^2}{\delta_1^2} \cdot \frac{2W_1 \mathrm{Sin}W_1 \mathrm{Sin}W_1 L \ \mathrm{Sin}W_1(2\Delta+L-1)}{1 - \mathrm{Cos}\ 2W_1} = 0$$

$$(2.3.19)$$

The stationary points are thus given by

$$W_o(L + 2\Delta - 1) = k\pi \qquad k \text{ integer} \qquad (2.3.20)$$

It is seen from (2.3.13) that the above condition gives $H(W_n)$ real. Furthermore, it follows from (2.3.10) that solutions with k even make the real part closer to unity when $W_1L < \pi$ while odd k should be choosen when $\pi \leq W_nL < 2\pi$. However there is still freedom in the choice of k. Since we want only integer values of Δ, it is natural to choose k so that Δ given by (2.3.20) is an integer.

The ideal performance of the ALE would of course be obtained if the input sinusoid appeared at the predictor output with the same amplitude and phase, thus yielding minimum error variance. This means that H(W) should be equal to unity at $W=W_1$ and zero elsewhere. It is however clear that this is in general impossible to achieve when the noise variance is non-zero or when the observations are finite in number.

2.4. CHOICE OF Δ FOR TWO SINUDOIDAL SIGNALS

For the case of two sinusoids in white noise, from (2.2.2) the filter coefficients are given by

$$a^*_k = A_1 e^{jW_1 k} + A_2 e^{jW_2 k} + A_3 e^{jW_3 k} + A_4 e^{jW_4 k} \qquad (2.4.1)$$

where

$$W_3 = -W_1$$
$$W_4 = -W_2$$

$$A_1 = \bar{A}_3 = \frac{1}{1-\gamma_{12}\gamma_{21}} \quad \frac{e^{jW_1\Delta}}{L+2\delta_o^2/\delta_1^2} - \frac{\gamma_{12} e^{jW_2\Delta}}{L+2\delta_o^2/\delta_1^2}$$

$$A_2 = \bar{A}_4 = \frac{1}{1-\gamma_{12}\gamma_{21}} \quad \frac{e^{jW_2\Delta}}{L+2\delta_o^2/\delta_1^2} - \frac{\gamma_{21} e^{jW_1\Delta}}{L+2\delta_o^2/\delta_1^2}$$

Therefore the transfer function of ALE is given by

$$H(W) = \sum_{k=0}^{L-1} \left[A_1 e^{jW_1 k} + \bar{A}_1 e^{-jW_1 k} + A_2 e^{jW_2 k} + \bar{A}_2 e^{-jW_2 k}\right] e^{-jWk} e^{-jW\Delta} \qquad (2.4.2)$$

If we assume two sinusoids with equal power, the transfer function must have a deep null at $(W_1+W_2)/2$. But if the power of each sinusoid is different the above condition is not valid any more. We formulate a new condition which is related to the power content of the sinusoids. We can write this condition as follows; [7]

$$H\left(\frac{\delta_1^2}{\delta_1^2+\delta_2^2} W_1 + \frac{\delta_2^2}{\delta_1^2+\delta_2^2} W_2\right) = \min\{H(W)\} \qquad (2.4.3)$$

If $\delta_1^2 = \delta_2^2$ then $H(\frac{W_1+W_2}{2}) = \min\{H(W)\}$.

The other way to minimize the following average error expression

$$V = \frac{1}{2\pi} \delta_o^2 \int_o^{2\pi} |1-H(W)|^2 dw + |1-H(W_1)|^2 \delta_1^2$$

$$+ |1 - H(W_2)|^2 \delta_2^2 \qquad (2.4.4)$$

The value of Δ which satisfies the above conditions was found as follows:

$$\Delta + \frac{(L-1)}{2} = \frac{(2k+1)\pi}{\Delta W} = (k + 1/2)/\Delta f \qquad (2.4.5)$$

where k is any non-negative integer such that $(k+1/2)/\Delta f$ $(L-1)/2$. The results expressed by (2.3.20) and (2.4.5) indicate that it may be possible to improve the resolution of one and two sinusoids in H(W) by varying the delay Δ so that (2.3.20) and (2.4.5) are satisfied. We observed the above conditions by means of computer simulations. This variation of resolution with Δ is similar to the dependence of the periodogram resolution and FFT resolution of sinusoids on their initial phase and zero appending.

For the details of the derivation, see Appendix A-1 and Appendix A-2. The procedures presented so far were based mainly on enhancing the signal. In a similar manner we can enhance the noise by using the adaptive notch filter. In Chapter 3 we will see this approach.

CHAPTER 3

ADAPTIVE NOTCH FILTERS

3.1. INTRODUCTION

This section investigates a method for eliminating sinusoidal or other periodic interference corrupting a signal. In general this problem can be solved by measuring the frequency of the interference and using a notch filter at that frequency. In [10] Glover uses an adaptive filter to eliminate interference. The procedure is called the adaptive noise cancelling and it is applicable when a reference input (desired input) is available which contains the interference alone. The reference input is filtered in such a way that it closely matches the interfering sinusoid and is then subtracted from the primary input leaving the signal alone.

In this procedure, one of the basic needs is to have a very narrow notch which is usually desired in order to filter out the interference without distorting signal. However, if the interference is not precisely known and if the notch is very narrow, the center of the notch may not fall exactly over the interference frequency. Also there are many applications where the interfering sinusoid drifts slowly in frequency. A fixed notch can not work here at all unless it is designed wide enough to cover the range of the drift. In such a situation it is often necessary to measure the frequency of the interference and then use a notch filter at that frequency. However, the estimation of frequency of several sinusoids can require a great deal of calculations.

Glover [10] proposed an alternative simpler method which can be used when a reference for the interference is available and makes measurement of its frequency unnecessary. This reference is adaptively filtered to match the interfering sinusoids as closely as possible, allowing them to be subtracted out.

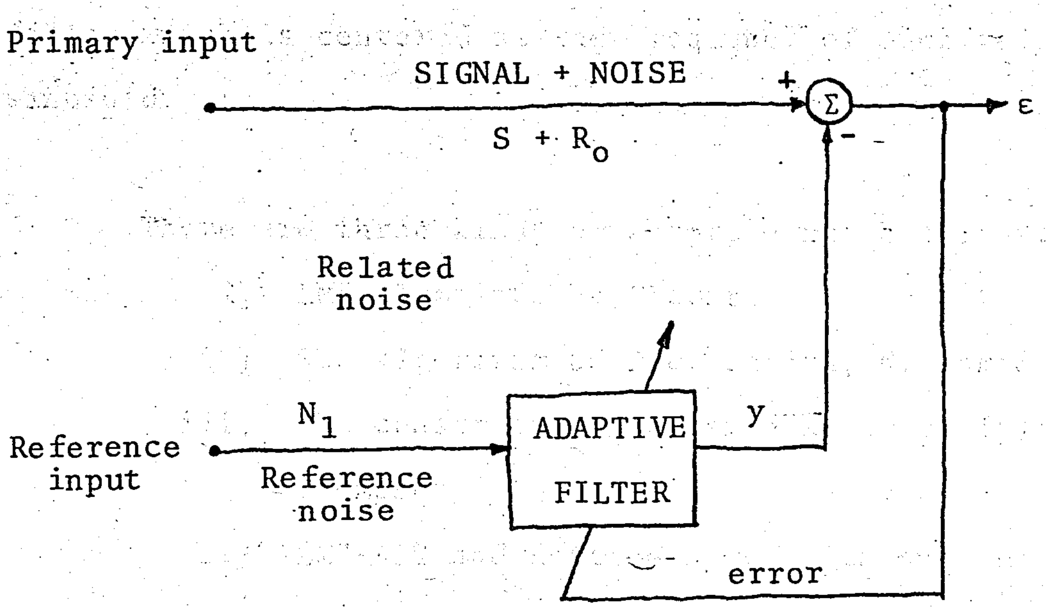


FIGURE 3.1.1. Adaptive Noise Cancelling System.

An adaptive filter is used in adaptive noise cancelling (ANC) as shown in Figure 3.1.1. The primary input consists of the signal plus noise $S+R_o$. The reference input is the related noise n_1. The reference n_1 is filtered to match R_o and then subtracted from the primary input. The error signal to the adaptation algorithm is therefore the output of the ANC system.

In the broad band case, the solution for the adaptive filter is a constant set of filter weights. Any deviation in the weights after convergence to this solution is considered to be simply noise in the adaptive process.

When the reference is sinusoidal, significant time varying components in the weights give rise to a tunable notch filter which is centered at the frequency of each reference sinusoid.

There are three kinds of adaptive notch filter:

i) LMS algorithm by Glover

ii) SER algorithm by D.D. Parikh, N. Ahmed

iii) The constrained recursive adaptive filter by Thompson

iv) LMS-ALE and Lattice by us with optimum Δ.

3.2. LMS-ALE NOTCH FILTER

Notch filters are capable of eliminating (or reducing) sinusoidal interferences by creating notches at appropriate places in the overall transfer function. The adaptive filter which is used as a notch filter in here is a transversal filter. The filter input is the delayed version of the primary input. This sequence is then applied to an N stage tapped-delay-line (TDL). The values at the N taps of the TDL at time k constitute the elements of the reference as a vector.

The adaptation algorithm most often used to set the weights of the filter is the LMS algorithm [1] given by the following equation for the weights.

$$a_k(j+1) = a_k(j) + 2\mu \left[X(j)X(j-\Delta-k) - X(j-\Delta-k) \sum_{i=0}^{L-1} X(j-\Delta-i)a_i(j)\right]$$

$$\text{for } k = 0,1,\ldots,L-1 \qquad (3.2.1)$$

where $a_k(j)$ is the jth update of the kth weight of the ALE, μ is a scalar representing the influence of the input $X(j)$ on the $(j+1)$st update of a_k and L and Δ are respectively the number of weights and the decorrelation parameter.

Since

$$\varepsilon(j) = X(j) - \sum_{i=0}^{L-1} X(j-\Delta-i)\, W_i(j) \qquad (3.2.2)$$

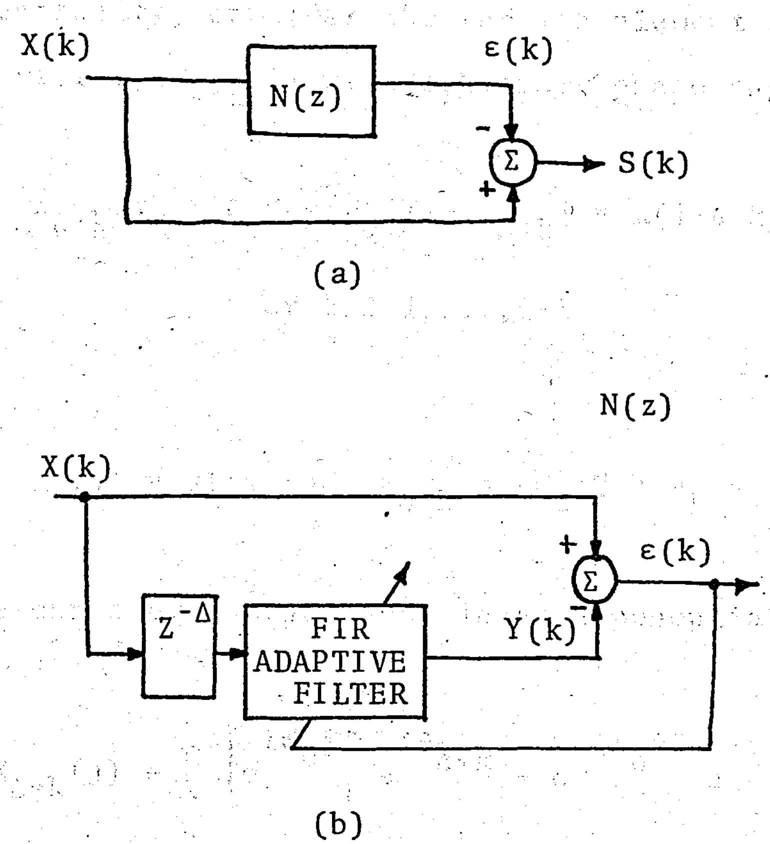


FIGURE 3.2.1. (a) Signal detection with Notch Filter
(b) Detail for adaptive Notch Filter

Therefore

$$a_k(j+1) = a_k(j) + 2\mu\varepsilon(k)X(j-\Delta-k) \qquad (3.2.3)$$

Let's take the Z transform of $W_k(j)$

$$Z\{a_k(j+1) - a_k(j)\} = 2\mu Z\{\varepsilon(k)X(j-\Delta-k)\} \qquad (3.2.4)$$

Let the input be of the form

$$X(j) = C \text{ Cos } [W_o jT + \theta] \qquad (3.2.5)$$

For generality, consider the $(\Delta+k)$th element of a general input X vector, $X_{\Delta+k}(j)$ with arbitrary phase angle $\theta_{\Delta+k}$

$$X_{\Delta+k}(j) = C\,\mathrm{Cos}\left[W_o jT + \theta_{\Delta+k}\right] = X(j-\Delta-k) \qquad (3.2.6)$$

$$\text{for } k=0,1,\ldots,L-1$$

where

$$X_o(j) = X(j) \text{ and } \theta_{\Delta+k} = -W_o\Delta T + \theta_k = -W_o T[k+\Delta] + \theta$$

Now we can express the input in an exponential form as follows

$$X_{\Delta+k}(j) = \frac{C}{2}\left[e^{iW_o jT}\, e^{i\theta_{\Delta+k}} + e^{-iW_o jT}\, e^{-i\theta_{\Delta+k}}\right] \qquad (3.2.7)$$

For the scalor form we can write [k takes any value between 0 and L-1]

$$a_k(j+1) = a_k(j) + 2\mu\,\varepsilon(j)X(j-\Delta-k)$$

or

$$a_k(j+1) = a_k(j) + 2\mu\,\varepsilon(j)X_{\Delta+k}(j) \qquad (3.2.8)$$

Therefore the Z transform of kth weight is then

$$a_k(z) = 2\mu\, U(z) \quad Z\{\varepsilon(j)X_{\Delta+k}(j)\} \qquad (3.2.9)$$

$$\text{for } k = 0,1,\ldots,L-1$$

where

$$Z\{\varepsilon(j)X_{\Delta+k}(j)\} = \frac{C}{2}\left[e^{i\theta_{\Delta+k}} E[Ze^{-iW_oT}] + e^{-i\theta_{\Delta+k}} E[Ze^{iW_oT}\right.$$

and

$$U(z) = \frac{1}{z-1} \qquad E[z] = Z\{\varepsilon(j)\}$$

Now let us calculate the output of the filter Y(z). Since

$$Y(k) = \sum_{i=0}^{L-1} a_i(j)\ X(j-\Delta-i)$$

$$= a_1(j)X(j-\Delta) + a_2(j)X(j-\Delta-1) + \ldots$$

$$+ a_{i-1}(j)X(j-\Delta-L-1) \qquad (3.2.10)$$

and Z transform of this sequences can be given as follows:

$$Y(z) = \frac{\mu C^2}{2}\left[\sum_{i=0}^{L-1} U(Ze^{-jW_oT})\ \{e^{2j\theta_{\Delta+i}} E(Ze^{-j2W_oT}) + E(z)\}\right.$$

$$\left. + \sum_{i=0}^{L-1} U(Ze^{jW_oT})\ \{e^{-2j\theta_{\Delta+i}} E(Ze^{j2W_oT}) + E(z)\}\right]$$

$$(3.2.11)$$

By rearranging and collecting terms we have

$$Y(z) = \frac{\mu C^2 L}{2} E(z) \left[U(Ze^{-jW_oT}) + U(Ze^{jW_oT}) \right]$$

$$+ \frac{\mu C^2}{2} U(Ze^{-jW_oT}) E(Ze^{-2jW_oT}) \sum_{i=0}^{L-1} e^{2j\theta_{\Delta+i}}$$

$$+ \frac{\mu C^2}{2} U(Ze^{jW_oT}) E(Ze^{2jW_oT}) \sum_{i=0}^{L-1} e^{-2j\theta_{\Delta+i}} \qquad (3.2.12)$$

The second and third terms in the expression for Y(z) are time varying terms and introduce at Y(z) unwanted frequency shifted component of E(z). The first term represents the time invariant part of the response from E(z) to Y(z), since only frequencies of E(z) appear at the output.

Now let us look at the exponential summation terms. Since we are using TDL filter the $\theta_{\Delta+i}$ arbitrary phase shift for the ith element of the X vector is written as

$$\theta_{\Delta+i} = \theta - W_oT[i+\Delta] \qquad \text{for } i=0,1,\ldots,L-1 \quad (3.2.13)$$

Substituting for $\theta_{\Delta+i}$, the summations are easily found to be

$$\sum_{i=0}^{L-1} e^{\pm j2\theta_{\Delta+i}} = e^{\pm 2j[\theta - W_oT(L-1-\Delta)]} \beta(W_{o_1}L) \qquad (3.2.14)$$

where

$$\beta(W_{o_1}L) = \frac{\text{Sin } L\, W_oT}{\text{Sin } W_oT}$$

From Chapter 2 we found the optimal decorrelation parameter Δ is equal to $\frac{1}{2}\{\frac{k\pi}{W_o} - (L-1)\}$ therefore by replacing it with equation (3.2.13) we have

$$= e^{\mp j2\theta}\, e^{\mp j\{W_oT[-3L+3]\}}\, e^{\mp j\{W_oT[2\Delta+L-1]\}}\, \beta(W_o,L)$$

$$= e^{\mp j2\theta}\, e^{\mp j\{3W_oT(1-L)\}}\, e^{\mp jk\pi}\, \beta(W_o,L)$$

$$= A \cdot \beta(W_o,L) \qquad (3.2.14)$$

By rewriting Y(z) we have

$$Y(z) = \frac{\mu C^2 L}{2} E(z)\, [U(Ze^{-jW_oT}) + U(Ze^{jW_oT})]$$

$$+ \frac{\mu C^2}{2}\, \beta(W_o,L)\, A[U(Ze^{-jW_oT})\, E(Ze^{-j2W_oT})$$

$$+ U(Ze^{jW_oT})\, E(Ze^{j2W_oT})\,] \qquad (3.2.15)$$

Since A is exponential term therefore it has unity amplitude. Now we can build up the relation between wanted and unwanted term and make an approximation for Y(z). It is clear that the following statement is true for approximation.

$$Y(z) = \begin{cases} Y(z) = f\{U(Ze^{\mp jW_oT}),\, E(z)\} & \text{If } \dfrac{\beta(W_o,L)}{L} << 1 \\ Y(z) = f\{U(Ze^{\mp jW_oT}), E(z), E(Ze^{\mp j2W_oT})\} & \text{If } \dfrac{\beta(W_o,L)}{L} \lessgtr 1 \end{cases}$$

From here the number of weights L in the adaptive filter can be increased to obtain a better $\beta(W_o,L)/L$ ratio. If the proper choice of parameters is made, the transfer function between E(z) and Y(z) is approximated by an LTI filter. In Chapter 4 we analyze this problem which is related to the filter length.

If $\beta(W_o,L)/L$ is very small we can write the notch filter expression as follows |10|:

$$H(z) = \frac{1}{1 + \frac{\mu C^2 L}{2}\left[U(Ze^{-jW_oT}) + U(Ze^{jW_oT})\right]}$$

$$= \frac{Z^2 - 2Z \cos W_oT + 1}{Z^2 - 2\left(1 - \frac{L\mu C^2}{2}\right) Z \cos W_oT + \left(1 - \frac{L\mu C^2}{2}\right)} \qquad (3.2.16)$$

It is clear that this is the transfer function for a 2nd order digital notch filter at the frequency W_o. The zeros of H(z) are at $Z = e^{\mp jW_oT}$, precisely on the unit circle. If $\frac{\mu LC^2}{2} << 1$, the pole locations are approximated by

$$Z \approx \left(1 - \frac{\mu C^2 L}{2}\right) e^{\pm jW_oT} \qquad (3.2.17)$$

The zeros lie on the unit circle at frequencies $\mp W_o$ with the poles a distance approximately $\mu C^2 L/2$ behind them radially toward the center of the circle.

ward the center of the circle. Near the frequency $W=W_o$ $H(z)$ can be approximated by the nearby pole and zero

$$H(z) = \frac{Z - e^{jW_oT}}{Z - (1 - \frac{\mu C^2 L}{2})\, e^{jW_oT}} \tag{3.2.18}$$

The 3db bandwidth (BW) is then obtained by finding the two points on the unit circle which are $\sqrt{2}$ times as far from the pole as they are from the zero and is given by [10]

$$BW = \frac{\mu C^2 L}{T} \tag{3.2.19}$$

3.3. A CONSTRAINED RECURSIVE ADAPTIVE FILTER FOR ENHANCEMENT OF NARROWBAND SIGNALS IN WHITE NOISE

3.3.1. The Constrained Recursive Filter

A constrained recursive adaptive filter can be used as a notch filter and enhance the narrowband signals in white noise. Among the most popular of such filters is adaptive line enhancer (ALE) which consists of a linear predictor with a tapped delay line (TDL) introduced by Widrow and studied in the previous section as an adaptive notch filter.

A recursive filter structure offers the significant advantage of an arbitrary narrowband frequency response with only a few memory elements and weighting coefficients, but the adaptation of those coefficients is much more difficult than for a TDL filter.

The recursive filter is as shown in Figure 3.3.1 with the transfer function of the signal enhancement filter taking the form

$$G(z) = 1 - H(z) \tag{3.3.1}$$

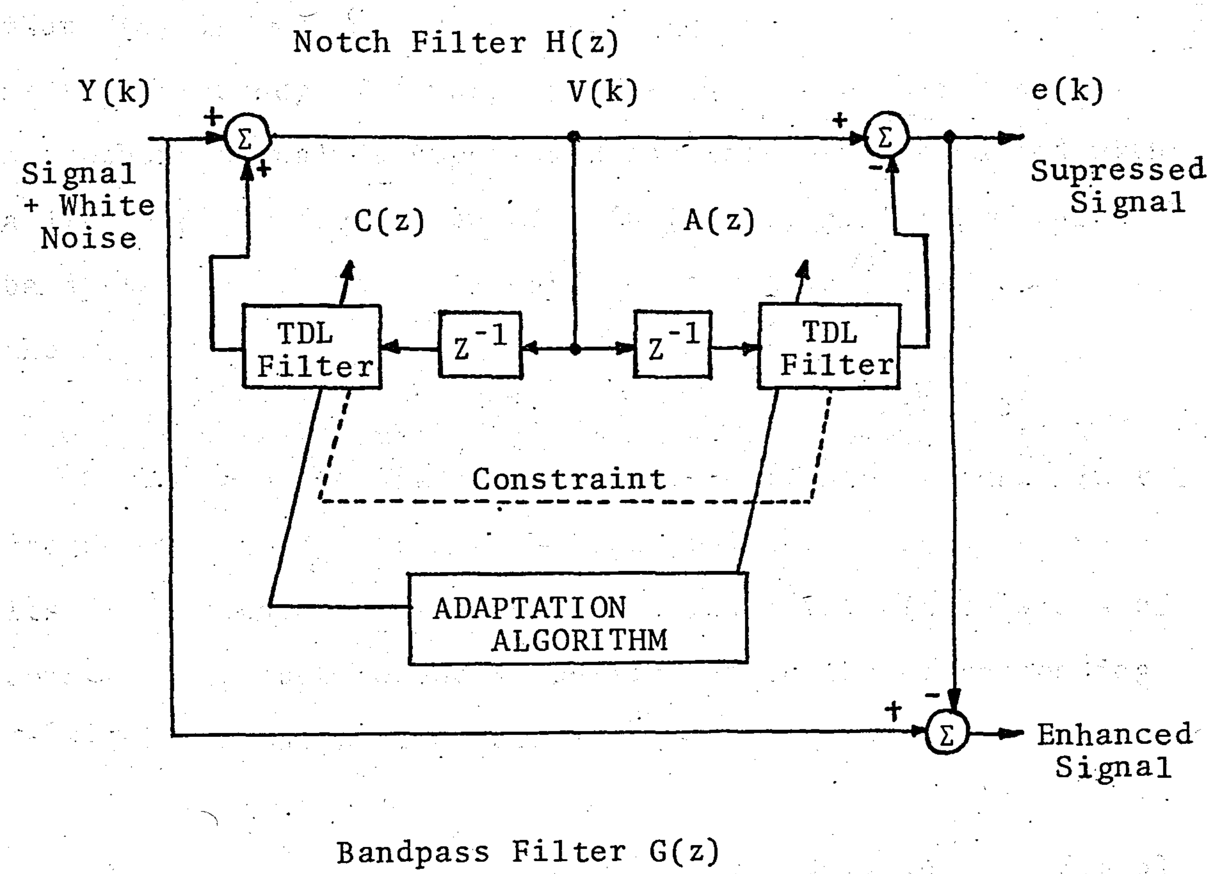


FIGURE 3.3.1. A Constrained Recursive Adaptive Line Enhancer.

where

$$H(z) = \frac{1 - A(z)}{1 - C(z)} \qquad (3.3.2)$$

with

$$A(z) = a_1 z^{-1} + a_2 z^{-2} + \ldots + a_n z^{-n} \qquad (3.3.3)$$

and

$$C(z) = C_1 z^{-1} + C_2 z^{-2} + \ldots + C_n z^{-n} \qquad (3.3.4)$$

The motivation for this filter structure stems from the fact that H(z) is desired to form a notch in its frequency response at the frequency of a narrowband signal. In this manner the narrowband signal is suppressed and the noise is passed with a little distrotion, then G(z) in (3.3.1) will represent a bandpass filter that will enhance the signal with respect to the noise.

In order to facilitate the formation of notches in the frequency response of H(w), a constraint is imposed between its feed-forward and feed-back coefficients. It consists of constraining each feedback coefficient to the corresponding feed-forward one by the relation [63]

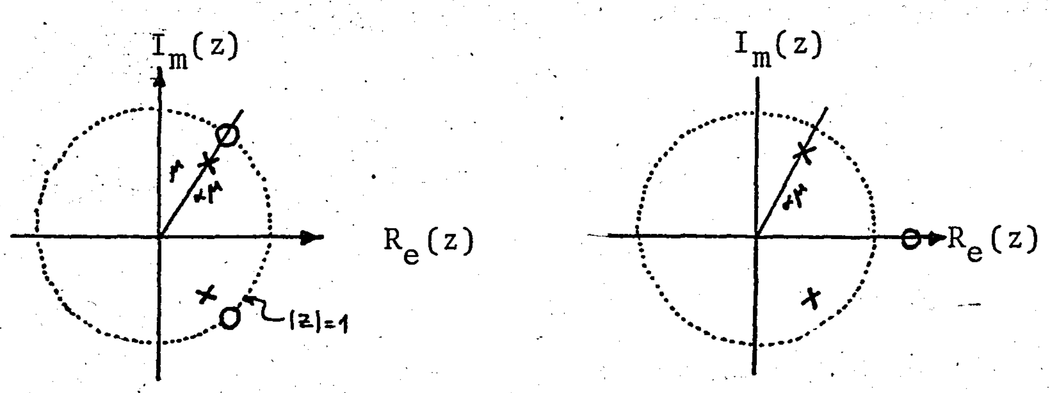
$$C_i = \alpha^i a_i \qquad i=1,2,\ldots,n \qquad (3.3.5)$$

suggested in [63], in which α is a selectable parameter which is chosen close to, but slightly less than one. The reason for choosing this particular constraint becomes clear when we observe the location of poles and zeros of H(z) under this constraint.

By substituting (3.3.5) into (3.3.3) and (3.3.4) implies that $C(z) = A(z|\alpha)$. Therefore the zeros and poles of H(z) which are denoted by (ξ_i) and (Π_i) $(i=1,2,...,n)$ respectively, must satisfy the relation:

$$\Pi_i = \alpha\xi_i \qquad i=1,2,...,n \qquad (3.3.6)$$

From (3.3.6) we see that the constraint places the poles of H(z) at the same polar-coordinate angles as its zeros but with slightly reduced magnitudes, causing H(z) to form the desired notch response when its zeros are located on or near the unit circle as shown in Figure 3.3.2.



a. Poles and zeros of H(z)

b. Poles and zeros of G(z)

FIGURE 3.3.2. Pole/Zero Patterns for G(z) and H(z).

The role of the parameter α in the constraint is to control the notch width of H(z) by controlling pole-zero separation. Also in ALE the notch width depends on μ called the adaptive step size.

As a measure of the signal-enhancement capability of the constrained filter the signal enhancement factor SEF defined as the ratio of signal power gain to noise power gain for filter G(z), is used. For a sine wave signal whose frequency coincides with the peak response of G(z), the SEF is simply the reciprocal of the (equivalent-noise) bandwidth of G(z). When the nagnitude μ of a conjugate pair of zeros of H(z) is near one, then G(z) whose poles are constrained to have magnitude $\alpha\mu$, forms a bandpass response with bandwidth approximately $1 - \alpha\mu$, making

$$\mathrm{SEF} = \frac{1}{1 - \alpha\mu} \tag{3.3.7}$$

3.3.2. The Bootstrap Adaptation Algorithm

The filter represented by the transfer function H(z) in (3.3.2), (3.3.3) and (3.3.4) can be represented in the time domain by the equations

$$V(k) = Y(k) + \underline{X}^T(k)\ \underline{C}(k) \tag{3.3.8}$$

$$e(k) = V(k) - \underline{X}^T(k)\, a(k) \qquad (3.3.9)$$

where Y(k) and e(k) represent the input and output respectively;

$$\underline{X}(k) \triangleq [V(k-1),\ V(k-2), \ldots,\ V(k-n)]^T \qquad (3.3.10)$$

represents a state vector; and

$$\underline{a}(k) \triangleq [a_1(k),\ a_2(k), \ldots,\ a_n(k)] \qquad (3.3.11)$$

and

$$\underline{C}(k) = [C_1(k),\ C_2(k), \ldots,\ C_n(k)] \qquad (3.3.12)$$

represent feed-forward and feedback parameter vectors respectively. In addition, the parameter constraint (3.3.5) can be represented by

$$\underline{C}(k) = \underline{\underline{M}}\ \underline{a}(k) \qquad (3.3.13)$$

in which M is the diagonal matrix

$$M = \text{diag}\ [\alpha,\ \alpha^2,\ \ldots,\ \alpha^n]$$

The bootstrap adaptation algorithm is motivated by the observation that the feed-forward portion of the filter H(z), represented by (3.3.9) has exactly the form of an ordinary linear predictor for which there exist adaptation algorithms for minimizing mean square error. The bootstrap consists of utilizing one of these algorithms for updating the feed-forward parameter vector $\underline{a}(k)$ and then computing the feedback parameter vector $\underline{C}(k)$ simply to maintain the constraint (3.3.13).

The simplest form of the bootstrap algorithm involves the use of a normalized version of the Widrow-Hoff LMS algorithm represented by the recursions

$$\underline{a}(k+1) = \underline{a}(k) + \frac{\gamma}{r(k)} \underline{X}(k)\ e(k) \qquad (3.3.14)$$

$$r(k) = (1-\gamma) r(k-1) + \gamma\ \underline{X}^T(k)\ \underline{X}(k) \qquad (3.3.15)$$

in which $r(k)$ is an on-line estimate of $E\{\underline{X}^T(k)\underline{X}(k)\}$, and γ is a selectable scalar constant satisfying $0 < \gamma << 1$.

3.4. SEQUENTIAL REGRESSION ADAPTIVE NOTCH FILTERS

3.4.1. Introduction

The main objective of this part is to present a class of adaptive notch filters which are derived using an SER approach [66]. In [10] the notion of using Widrow's LMS algorithm to derive a class of notch filters was introduced.

In [66] it was shown that the SER adaptive notch filters have the following advantages, relative to the LMS counterparts, when each of the filters has the same number of coefficients (weights): (1) The rate of adaptation is substantially faster and (2) a sharper notch is realizable over a large bandwidth. The advantage of the LMS approach however, is that it results in filters that are easier to implement.

3.4.2. SER Algorithm

SER algorithm cost function is defined as follows:

$$R(a_{r+1}) = q \sum_{k=1}^{r} [d(k) - q^T{}_{r+1} X_k]^2 + a^T{}_{r+1} a_{r+1} \qquad (3.4.1)$$

where q is a scalar and

$$a_k^T = [q_o(k) \ a_1(k) \ \ldots \ a_N(k)]$$

$$X_k^T = [X(k) \ X(k-1) \ \ldots \ X(-N)]$$

d_k denotes the desired output at time k (see Figure 3.4.1.)

The filter weights can be computed using the relation as in 66

$$a_{k+1} = a_k + q \ P_k^{-1} \ X_k \ e(k) \qquad (3.4.2)$$

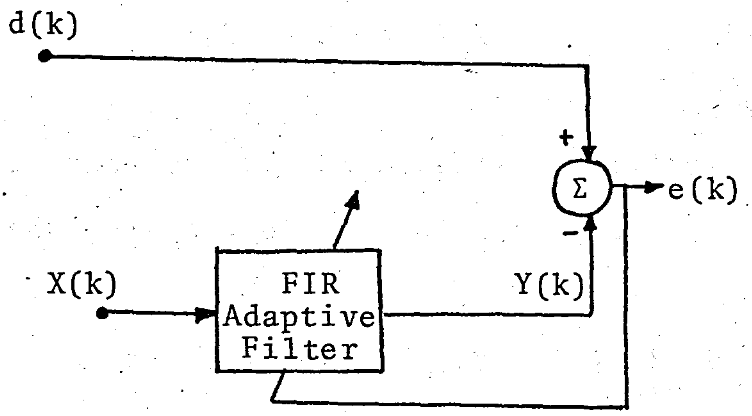


FIGURE 3.4.1. SER Adaptive Noise Cancelling Mode.

where e(k) = d(k) - Y(k) is the error at the kth iteration and

$$P_r = I + q \sum_{k=1}^{r} X_k X_k^T$$

The (N+1) x (N+1) matrix P_1^{-1} can also be computed recursively using the matrix inverse Lemma

$$P_k^{-1} = P_{k-1}^{-1} - \frac{1}{\gamma} P_{k-1}^{-1} X_k X_k^T P_{k-1}^{-1} \qquad (3.4.3)$$

where $\gamma = 1/q + X_k^T P_{k-1}^{-1} X_k$ is a scalar and implies that $P_o^{-1} = I$.

3.4.3. Derivation of Notch Filter

For the input be form

$$x(k) = C \cos(W_o kT + \theta_o) \tag{3.3.4}$$

the Z transform of the weights and filter output can be given respectively as follows [66] and [10]

$$A_i(z) = \frac{qC}{2} U(Z) \{E(Ze^{-jW_oT}) \sum_{n=0}^{N} P_{i,n} e^{j\theta_n}$$

$$+ E(Ze^{jW_oT}) \sum_{n=0}^{N} P_{i,n} e^{-j\theta_n} \tag{3.3.5}$$

for $i = 0,1,\ldots,N$

$$Y(z) = \sum_{i=0}^{N} \frac{C}{2} [A_i(Ze^{-jW_oT})e^{j\theta_i} + A_i(Ze^{jW_oT})e^{-j\theta_i}] \tag{3.4.6}$$

where

$P_{i,n}$ denotes the (i,n)th element of P_k^{-1}

$$U(z) = \frac{1}{z-1}$$

and

$$E(z) = Z\{e(k)\}$$

Again after some modification, Y(z) can be approximated by discarding the time-varying term to obtain [66]

$$Y(z) \cong \frac{qC^2\xi}{4} E(z) \{U(Ze^{-jW_oT}) + U(Ze^{jW_oT})\} \tag{3.4.7}$$

where

$$\xi = \sum_{i=0}^{N} \sum_{n=i}^{N} b_{i,n} P_{i,n} \mathrm{Cos}\left[(i-n)W_o T\right]$$

and

$$b_{i,n} = \begin{cases} 1 & i=n \\ 2 & i \neq n \end{cases}$$

The notch filter expression is given as follows

$$G(z) = \frac{1}{1 + \frac{qC^2\xi}{4} \{U(Ze^{-jW_oT}) + U(Ze^{jW_oT})\}} \qquad (3.4.8)$$

and whose 3db bandwidth is given by

$$BW = \frac{q\xi C^2}{2T} \text{ rad/s.} \qquad (3.4.9)$$

We see that the expression for G(z) is different from the previous one. In three derivations and previous chapter we can see that the filter length is a very important parameter during the design. In Chapter 4, by taking some criterion we try to find the optimal filter length to use in all applications.

CHAPTER 4

DERIVATION OF THE OPTIMUM FILTER LENGTH

4.1. INTRODUCTION

During the derivation of the optimal value of filter length for an adaptive line enhancer, two critical points must be considered: stability and optimal operation. There are two ways to derive the optimal length. The first method arises from the following observation: By improving the estimate of the steady state mean squared error (MSE) a tighter stability is obtained and at the same time the SNR gain attained by the ALE is also improved. The MSE is minimized by using the LMS algorithm to adapt the ALE weights. The SNR is optimized by chossing the filter length optimally. Since the transversal filter implements a bandpass filter, the number of weights L determines the bandwidth of this filter and improves the gain in the signal to noise ratio.

In particular it can be shown that for a given step-size parameter μ which satisfies the stability constraint ther exists an optimal number of weights which maximizes the SNR gain that is used as a performance measure.

Another method can be summarized as follows. The coefficients which are adapted by using LMS algorithm converge to a set of zeros when no sinusoid is present in the input data and to a sinusoidal distribution when a sinusoid is present. Therefore one can obtain a detection system for the sinusoid by computing the Fourier Transform of the weights and comparing the magnitude of the transform with a fixed threshold. The detection performance can be improved by employing optimal filter length.

4.2. MAXIMUM SNR METHOD FOR WHITE NOISE

The purpose of this method is to present a better estimate for the steady state MSE which enables the derivation of more accurate expression for the SNR gain achieved by the ALE as well as a more accurate stability constraint.

Since the LMS algorithm uses an estimate of the MSE gradient for adapting the weights, the actual instantaneous values of the a(k)'s fluctuate after convergence about their mean value causing a degredation in the performance of the adaptive filter. Assuming that the weights have converged, let

$$a(k) = E\{a(k)\} + V(k) = a^* + V(k) \qquad (4.2.1)$$

Then the output from the transversal filter y(k) can be described as the sum of two terms

$$y(k) = a^T(k)X(k) = a^{*T}X(k) + V^T(k)X(k)$$

$$= y^*(k) + y^V(k) \qquad (4.2.2)$$

where $y^*(k)$ is the output expected from the optimal Wiener, filter and $y^V(k)$ is a noise component added due to the weights' fluctuations. With the assumption of no correlation between $y^*(k)$ and $y^V(k)$

$$E[y^2(k)] = E[(y^*(k))^2] + E[(y^V(k))^2] \qquad (4.2.3)$$

Using the derivation in Adaptive Transversal filter section, we have

$$E[[y^V(k)]^2] = \mu \text{ trace}[Rxx] \; \xi min. \qquad (4.2.4)$$

where ξmin is the minimum MSE achieved by the Wiener solution. Thus using (4.2.4) the steady state MSE, ξss is given by

$$\xi ss = \xi min + E[(y^V(k))^2] = \xi min + \mu trace[Rxx] \; \xi min$$

$$= [1 + \mu trace(Rxx)] \; \xi min. \qquad (4.2.5)$$

In [56] A. Nehorat and D. Malah pointed out the interesting problem which was related to μ and misadjustment. In particular the misadjustment is defined as the ratio of the excess MSE to be minimum MSE and is given by

$$M = \mu \text{trace}[Rxx] \qquad (4.2.6)$$

But this result is proper only for very small values of μ. In an attempt to extend the above results for larger values of μ as well as to adequately predict the divergence of the adaptation process, let us derive the upper and lower limits for μ which is important for stability constraint. From the previous section; the weights expression is given by

$$a(k+1) = a(k) - 2\mu R_{xx} X(k) \qquad (4.2.7)$$

Subtracting a* from both sides of (4.2.7) yields

$$V(k+1) = V(k) - 2\mu R_{xx} V(k) = [I - 2\mu R_{xx}]V(k) \qquad (4.2.8)$$

Equation (4.2.8) is a linear homogeneous vector difference equation whose solution characterizes the dynamic behavior of the weight vector as it begins at a(o) and if the process is convergent, relaxes toward a*, as seen by Equation (4.2.1).

The solution of (4.2.8) is given by

$$V(k) = [I - 2\mu R_{xx}]^k V(o) \tag{4.2.9}$$

This solution is stable (convergent) if

$$\lim_{k\to\infty} [I - 2\mu R_{xx}]^k = 0 \tag{4.2.10}$$

Since

$$[I - 2\mu R_{xx}] = Q(I - 2\mu\Lambda)Q^{-1} \tag{4.2.11}$$

and

$$[I - 2\mu R_{xx}]^k = Q[I - 2\mu\Lambda]^k Q^{-1} \tag{4.2.12}$$

Condition (4.2.10) will be satisfied if

$$\lim_{k\to\infty} [I - 2\mu\Lambda]^k = 0 \tag{4.2.13}$$

Condition (4.2.13) will be met when

$$[1 - 2\mu\lambda p] < 1 \tag{4.2.14}$$

for p=1,2,..., n. Since all eigenvalues are positive

$$\frac{1}{\lambda_{max}} > \mu > 0 \tag{4.2.15}$$

where λ_{max} is the largest eigenvalue of R. Equation (4.2.15) gives the stable range for μ.

The upper limit in (4.2.15) was found to be too high by A. Neharoi and D. Malch with computer simulation. In (4.2.4) ξ_{min} is replaced by the actual steady-state MSE ξ_{ss}, and in place of (4.2.5) we obtain

$$\xi_{ss} = \xi_{min} + \mu \text{trace } R_{xx} \ \xi_{ss} \tag{4.2.16}$$

and hence

$$\xi_{ss} = \xi_{min}/(1 - \mu \text{trace } R_{xx}) \tag{4.2.17}$$

resulting in a misadjustment of

$$M = \mu \text{trace}(R_{xx})/[1 - \mu \text{trace}(R_{xx})] \tag{4.2.18}$$

Clearly, if μ is sufficiently small ($\mu \text{trace}(R_{xx}) << 1$) the results in (4.2.17) and (4.2.12) coincide with those in (4.2.5) and (4.2.6) respectively. However, (4.2.16) and (4.2.18) are proper for higher values of μ, even up to divergence which is predicted from (4.2.17) to occur when μ reaches $1/\text{trace}(R_{xx})$. Thus, the stability constraint on μ which replaces (4.2.15) is given by

$$0 < \mu < \frac{1}{\text{trace } R_{xx}} \tag{4.2.19}$$

It is interesting to note that (4.2.19) is usually used as a sufficient condtion for stability since $\text{trace}\{R_{xx}\} \geq \lambda_{max}$ and is usually easier to evaluate. The above shows that (4.2.19) is also a necessary condition.

Now let us continue with the derivation of L by using the above results. In ALE operation, for a given step-size parameter μ which satisfies the stability constraint there exists an optimal number of weights which maximizes the SNR gain.

Let the total power of the input signal be P_x. Then since the reference input signal $X(k-\Delta)$ is a delayed version of the input signal and the transversal filter has L taps we can write the following formula

$$\text{trace}\left[R_{xx}\right] = L\, r_{xx}(o) = L\, P_x \qquad (4.2.20)$$

Assuming an input signal of the form

$$x(k) = S(k) + n(k) = \sum_{m=1}^{N} C_m \text{Cos}(W_m k + \phi_m) + n(k) \qquad (4.2.21)$$

i.e., N sinusoidal signals with an addive zero mean white noise sequence n(k), the autocorrelation sequence $r_{xx}(\ell)$ which determines R_{xx} is given by

$$r_{xx}(\ell) = \sum_{m=1}^{N} \frac{C_m^{\,2}}{2} \text{Cos}\, W_m L + \sigma_n^{\,2}\, \delta(\ell) \qquad (4.2.22)$$

where σ_n^2 is the noise power and $\delta(\ell)$ is kronecker δ function. Since the Wiener solution for a single sinusoidal in white noise at frequency W_o is given by [7]

$$a^*_k = \frac{SNR_{im}}{1 + SNR_{im}\ L/2} \mathrm{Cos} W_o(k+\Delta) \qquad (4.2.23)$$

where

$$a^*_k = [a^*_1, \ldots, a^*L-1]^T$$ and SNR_i is the input SNR,

$$SNR_{im} = \frac{C_m^{\ 2}}{2\sigma_n^{\ 2}} \qquad m=1,2,\ldots,N$$

Then the optimal Wiener solution for the case of N sinusoidal signals can be given by

$$a^* = \sum_{k=1}^{N} a^*_k \qquad (4.2.24)$$

The corresponding output of the transversal filter is given by

$$y^*(k) = \sum_{i=0}^{L-1} a^*_i(k)\ X(k-\Delta-i) = [a^*_k] \begin{bmatrix} X(k-\Delta) \\ \vdots \\ X(k-\Delta-L-1) \end{bmatrix} \qquad (4.2.25)$$

The total power of the output signal from the transversal filter is given by

$$E\{(y^*(k))^2\} = \sigma_n^2 \ \frac{2}{L} \sum_{m=1}^{N} (b^*_m)^2 + \sum_{m=1}^{N} (b^*_m C_m)^2/2 \tag{4.2.26}$$

where

$$b^*_m = (\frac{L}{2}) \ \frac{C_m^2}{2\sigma_n^2} \ / \ (1 + \frac{C_m^2}{2\sigma_n^2} \ \frac{L}{2}) \tag{4.2.27}$$

The overall output SNR is given therefore by

$$SNR_o = \frac{L}{2} \sum_{m=1}^{N} (b^*_m C_m)^2/2 \ \ / \ \ \sigma_n^2 \sum_{m=1}^{N} (b^*_m)^2$$

$$= \frac{L}{2} \sum_{m=1}^{N} (\frac{C_m^2}{2\sigma_n^2}) \ (b^*_m)^2 \ \ / \ \sum_{m=1}^{N} (b^*_m)^2 \tag{4.2.28}$$

and we define

$$SNR_{ALE} = \frac{SNR_o}{\sum_{m=1}^{N} SNR_{im}} = \frac{SNR_o}{SNR_T}$$

SNR_{ALE} is the gain in SNR achieved by the ALE which has the Wiener solution weights.

The decrease in SNR_{ALE} with the increase in number of sinusoidal signals is due to the corresponding larger number of bondbass filters, each passing not only the desired signal but also a band of the noise, thus increasing the overall output noise.

For the particular case of equal power N sinusoids; SNR_{ALE} is given by

$$SNR_{ALE} = \frac{L}{2N} \tag{4.2.29}$$

We turn now to the performance of the ALE with the actual weights a as obtained with the LMS algorithm. From (4.2.16), (4.2.17) and (4.2.18) we conclude that in order to find the actual total output power one has to add to the right hand side of (4.2.25) an additional term which is equal to the excess MSE given by $\bar{M}\ \xi min$. Thus, (4.2.28) is replaced by

$$SNR_o = \frac{L}{2} \sum_{m=1}^{N} \left(\frac{C_m^2}{2\sigma_n^2}\right)(b^*_m)^2 \Big/ \sum_{m=1}^{N} (b^*_m)^2 + \bar{M}\ \xi min \tag{4.2.30}$$

Now let us find the expression for ξmin. The output e(k) has three components: the desired wide-band component n(k), its filtered version from the predictor output which is a distortion component, and the attenuated sinusoids. We find that the sinusoids at e(k) are given by

$$S_{ek} = \sum_{m=1}^{N} (1-b^*_m) C_m \operatorname{Cos}(W_m k + \phi_m) \tag{4.2.31}$$

Now let us consider the average power of e(k). Noting that all the component of e(k) are uncorrelated we find that

$$n_{1ek} = \sum_{m=1}^{N} b^*_m \, n(k) \tag{4.2.32}$$

$$n_{2ek} = n(k) \tag{4.2.33}$$

$$E\{e^2(k)\} = E\{n_{2ek}^2\} + E\{n_{1ek}^2\} + E\{S_{ek}^2\} \tag{4.2.34}$$

$$E\{e^2(k)\} = \sigma_n^{\,2} + \sigma_n^{\,2} \left(\frac{2}{L}\right) \sum_{m=1}^{N} (b^*_m)^2 + \sum_{m=1}^{N} (1 - b^*_m)^2 \frac{C_m^{\,2}}{2}$$

Therefore we can express

$$\xi_{min} = E\{e^2(k)\} \tag{4.2.35}$$

For the particular case of equal-power sinusoids so that $SNR_T = N\,SNR_{im}$ and $b^* = b^*_m$ $m=1,2,\ldots,N$ we obtain

$$SNR_o = \frac{\frac{L}{2}\, SNR_T\, (b^*)^2}{Nb^{*2} + \bar{M}\,\xi_{min}} \tag{4.2.36}$$

$$\frac{SNR_o}{SNR_T} = \frac{(b^*)^2}{\frac{2N}{L}(b^*)^2 + \frac{\mu LP_x}{1-\mu LP_x}\left[1 + \frac{2N}{L}(b^*)^2 + SNR_T(1-b^*)^2\right]}$$

With the substitution of (4.2.27) for b* in (4.2.36) we have

$$\frac{SNR_o}{SNR_T} = \frac{1}{\frac{2N}{L} + \frac{\mu LP_x}{1-\mu LP_x}\left|1 + \frac{2N}{L}\left(1 + \frac{2}{SNR_T}\right) + \frac{4N^2}{L^2 SNR_T}\left(1 + \frac{1}{SNR_T}\right)\right|} \qquad (4.2.37)$$

Since the LMS algorithm attempts to minimize the MSE it does not maximize, in general, the output SNR as would be desired for the ALE. This can be seen from

$$\min_a E\{e^2(k)\} = E\{n^2(k)\} + \min_a E\{(S(k) - y(k))^2 + n_{1ek}^2\} \qquad (4.2.38)$$

which is clearly not equivalent to maximizing SNR_o where

$$SNR_o = \frac{E\{\tilde{y}^2(k)\}}{E\{n_{1ek}^2\}} \qquad (4.2.39)$$

It is therefore of importance to properly choose the number of weights L and the step-size parameter μ in order to optimize the performance of the ALE for a given application. In practice, L can not be increased beyond a certain L_{max} and μ cannot be decreased below a certain $\mu_{min} > 0$. By selecting $\mu_o = \mu_{min}$ the optimal value for L is found by differentiating (4.2.37) with respect to L, to be

$$L_{opt} = \left[\frac{2N}{\mu_o P_x} + \frac{4N}{(SNR_T)^2} + \frac{4N}{SNR_T}\right]^{\frac{1}{2}} \qquad (4.2.40)$$

If also $2\mu_o P_x << (SNR_T)^2/(SNR_T + 1)$, (4.2.40) is simplified to

$$L_{opt} = [2N/(\mu_o P_x)]^{\frac{1}{2}} \qquad (4.2.41)$$

The maximum SNR gain is then given by

$$\left(\frac{SNR_o}{SNR_T}\right)_{max} \cong \frac{L_{opt}}{4N} \qquad (4.2.42)$$

Equation (4.2.42) is the half of the Equation (4.2.28) which was derived from the optimal Wiener solution.

4.3. OPTIMAL DETECTOR METHOD (DETECTION PERFORMANCE METHOD)

This method is concerned with the application of a linear predictive filter which employs time-varying coefficients, to sets of data consisting of white noise which may or may not contain a sinusoid. The coefficients are adapted using the LMS algorithm. It has been shown [2] that the set of ocefficients converges to a set of zero mean, independent values when no sinusoid is present in the input data and to a sinusoidal distribution when a sinusoid is present. One can therefore obtain a detection system for the sinusoid by computing the Fourier Transform of the weights and comparing the magnitude of the transform with a fixed threshold.

Adaptive linear predictors used in this manner have been termed "Adaptive Line Enhancers" are ALE's. This section descusses the detection performance of an ALE containing L coefficients which adapt on N samples of the input data. The performance is compared with the optimal detector for a sinusoid in white noise which consists of a Fourier Transform of the entire N data samples.

It has been shown previously 50 that under certain assumptions, the probability density function of the detection statistic used in the ALE weight transform detector can be modelled using the non-central chi distribution. Briefly if $W_e(k)$ denotes the Lth ALE weight after k adaptations and the frequency of interest is W_o, we define real and imaginary parts of the DFT of the ALE weights at time k as $U_w(k)$ and $V_w(k)$ respectively, which can be written as

$$U_w(k) = \sum_{L=0}^{L-1} W_e(k) \text{ Cos } W_o \ell \qquad (4.3.1)$$

$$V_w(k) = \sum_{L-0}^{L-1} W_e(k) \text{ Sin } W_o \ell \qquad (4.3.2)$$

Detection consists of computing U_w and V_w at a time k corresponding to the last data sample processed and then comparing the sum of the squares of U_w and V_w with a fixed threshold. In order to avoid adverse start transients, we assume that the filter is initially filled with data prior to the onset of adaptation. With this assumption, a total of N-L samples are available for adaptation and the detection statistics $Z_w{}^2$ becomes

$$Z_w{}^2 = U_w{}^2(N-L) + V^2{}_w(N-L) \qquad (4.3.3)$$

The mean value of the weights at time N-L when a sinusoidal signal is present are given by

$$E\{W_\ell(N-1)\} = \frac{2a^*(N-L)}{L} \operatorname{Cos}(W_o\ell + \psi) \qquad (4.3.4)$$

where

$$a^*(N-L) = [1 - (1 - \mu\lambda^*)^{N-L}] \frac{\frac{L}{2}\,\mathrm{SNR}}{1 + \frac{L}{2}\,\mathrm{SNR}}$$

and

$$\lambda_{max} = \lambda^* = n_o + \frac{LA^2}{4} = n_o(1 + \frac{L}{2}\,\mathrm{SNR})$$

In these expressions, n_o is the white noise power level at the ALE input, A is amplitude of the sinusoid and SNR = $A^2/2n_o$ is the input signal to noise ratio. When the signal is not present, the weights have zero mean value.

Since LMS algorithm uses an estimate of the MSE gradient for adapting the weights, the actual instantaneous values of W_e fluctuate (after convergence) about their mean value $E\{W_e(N-L)\} = W^*$ causing a degradation in the performance of the adaptive filter, therefore the weight vector also contains misadjustment noise (weight noise).

Therefore (4.3.4) can be rewritten as follows:

$$E[W_e(N-L)] = E[W_e(N-L) - W_n] \tag{4.3.5}$$

and it is clear that

$$E\{W_N\} = 0 \tag{4.3.6}$$

$$E[(W_e(N-L) - E\{W_e(N-L)\})^2] = \mu n_o \tag{4.3.7}$$

Under the assumptions used above in ALE analysis, the weights are modelled as Gaussian with a variance of

$$Var[W_e(N-L)] = \mu n_o \tag{4.3.8}$$

The terms $U_w(N-L)$ and $V_w(N-L)$ are then also Gaussian with variance $\mu n_o L/2$ and means $a^*(N-L)Sin\psi$ and $a^*(N-L)Cos\psi$, respectively, under H_1, i.e., the signal present hypothesis. Under the null hypothesis H_o, both terms have zero mean and variance $\mu n_o L/2$.

Given these statistical descriptions, the probability density functions for $|Z_w|^2$ in (4.3.3) can be derived |51,54,55|

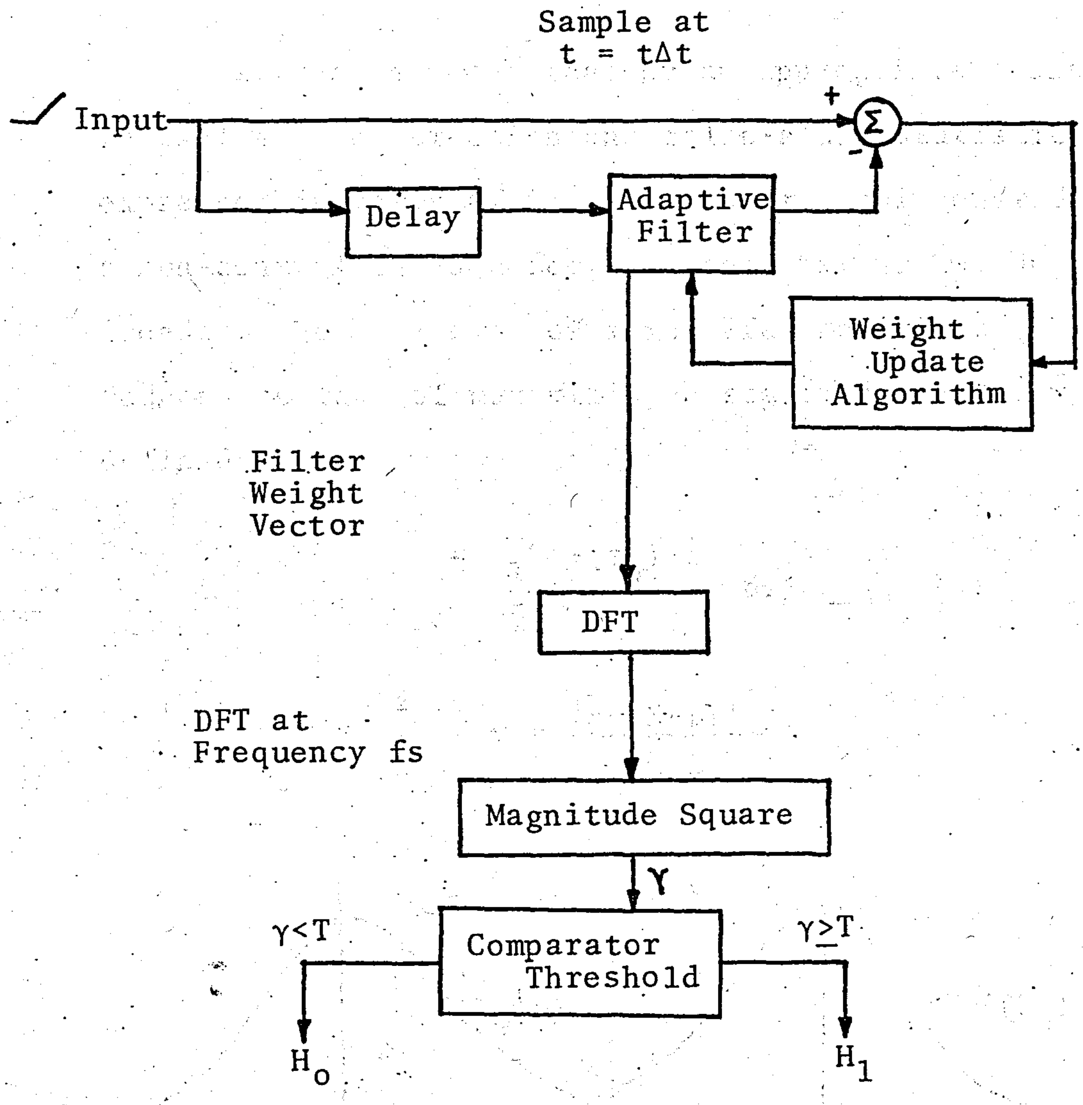


FIGURE 4.3.1. General Detection System

The theoretical density function for the squared magnitude of the DFT of the weights has the form of a two degree of freedom chi-square density function. That is,

$$P(z) = \frac{1}{\bar{z}} \exp\left(- \frac{z_w}{\bar{z}}\right) \tag{4.3.9}$$

where $z_w = [W_k]^2$ and $\bar{z} = E[[W_k]]^2 = \delta_z$.

It can be shown that by an appropriate substitution of variables, the detection and false-alarm statistics can be expressed in terms of integrals over a chi-squared pdf and a non-central chi-squared pdf, each having two degrees of freedom. For the case of fixed Pfa, maximization of P_d then reduces to that of maximizing a scalar parameter $\bar{\gamma}$ which is defined as

$$\int_{\xi}^{\infty} \frac{1}{\bar{z}} e^{-(z_w)/\bar{z}} \, dz_w$$

$$\frac{z_w}{\bar{z}} \triangleq \gamma = \frac{|a^*(N-L)|^2}{\mu n_o L} \tag{4.3.10}$$

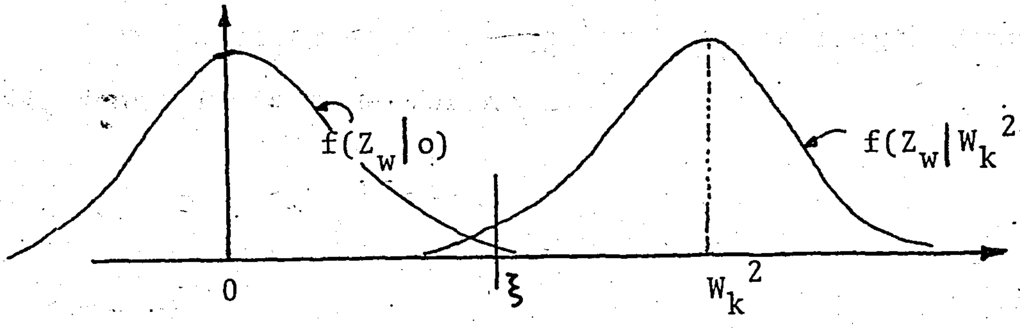


FIGURE 4.3.2. Density Functions for Detection Problem.

By allowing μ and L to vary simultaneously such that $\tilde{\gamma}$ and consequently the ALE detection performance, are both maximized. The obvious approach is to evaluate the partial derivatives of $\tilde{\gamma}$ with respect to μ and L and then set these partials to zero so as to obtain two equations in the two unknowns. But this method does not yield closed form analytical expressions. An alternative approach is to assume that the final solution satisfies the condition N>>L.

Under these conditions, Reeves [49] has shown that the optimal value of adaptive step-size μ_o given by

$$\mu_o = \frac{1.25643}{n_o(N-L)(1 + \frac{L}{2} SNR)} \qquad (4.3.11)$$

Substituting this value into Equations (4.3.10) results in an expression for $\tilde{\gamma}$ in (4.3.11) which depends only on L, N and SNR. The resulting value L_o which maximizes $\tilde{\gamma}$ is

$$L_o = \frac{\sqrt{1 + N.\ SNR/2} - 1}{SNR/2} \qquad (4.3.12)$$

It is clear that the optimal filter length depnds on SNR, N and it is true for N>>L.

4.4. MAXIMUM SNR METHOD FOR COLORED NOISE

We use the matrix formulation of the enhancement of sinusoids in colored noise to obtain new expressions for the optimal least squares coefficients and frequency response of the Δ step predictor. From this analysis we can approach the similar results to obtain optimum filter length.

Main notation

$a^* \triangleq |a_o, a_1, \ldots, a_{L-1}|^T$ optimal coefficient vector (L×1) (4.4.1a)

$X_k \triangleq |d_{k-\Delta}, d_{k-\Delta-1}, \ldots, d_{k-\Delta-L+1}|$ data vector (L×1) (4.4.1b)

$R_{xx} \triangleq E\{X_k X_k^H\}$ data correlation matrix (L×1) (4.4.1c)

$P \triangleq E\{X_k \bar{d}_k\}$ cross correlation matrix (L×1) (4.4.1d)

$\gamma(w) \triangleq |1, e^{-jW}, \ldots, e^{-j(L-1)W}|^T$ (L×1) (4.4.1e)

$\gamma_m \triangleq A_m \gamma(W_m)$ (L×1) (4.4.1f)

$\Gamma = |\gamma_1, \ldots, \gamma_N|$ observability type matrix (L×N) (4.4.1g)

where T denotes transpose, H is the Hermitian transpose, and (-) is the complex conjugate, W is the frequency and A_m will denote the amplitude of the mth complex sinusoid at the input and C_m for real sinusoids. Notice that while A_m and C_m are scalars, all other capital letters are used for matrices. In this notation, the predictor output is

$$Y_k = W^{*H} X_k = \sum_{i=0}^{L-1} \bar{W}_i \, d_{k-\Delta-i} \qquad (4.4.2)$$

and its error

$$e_k = d_k - Y_k \qquad (4.4.3)$$

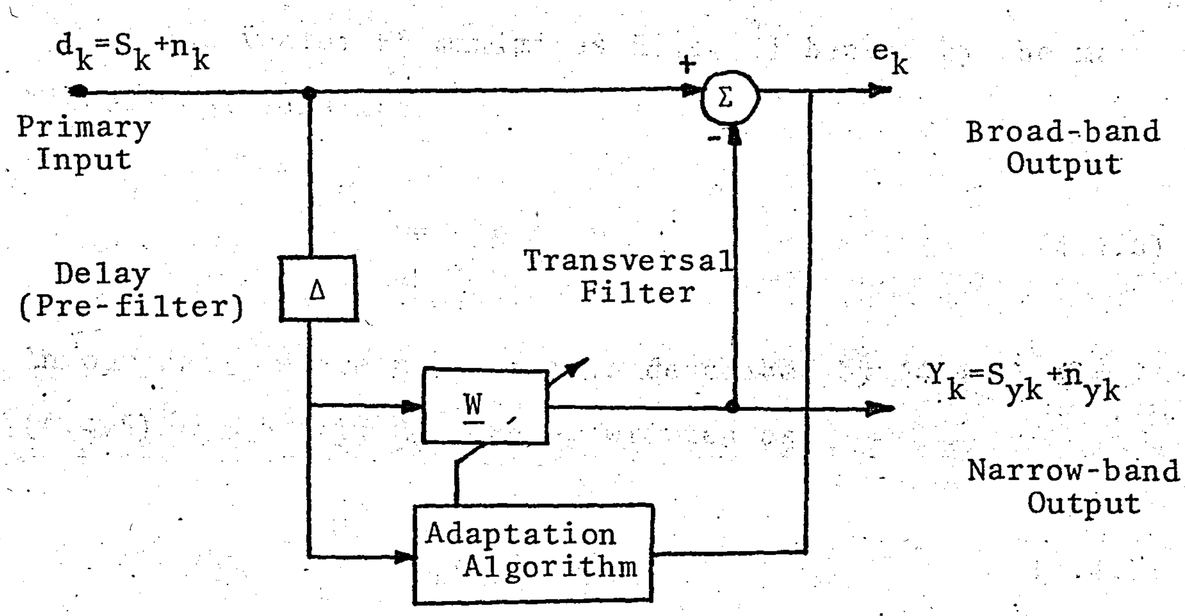


FIGURE 4.4.1. Block Diagram of the Δ-step Predictor or (Prefiltered) ALE.

To determine the optimal coefficients, assume that the input consists of N complex sinusoids with additive zero-mean colored noise, i.e.,

$$d_k = S_k + n_k = \sum_{m=1}^{N} A_m e^{j(W_m k + \psi_m)} + n_k \quad (4.4.4)$$

where $\{\psi_m\}$ are independent and uniformly distributed over $|0,2\pi|$ and n_k is not necessarily white. The autocorrelation sequence of the input (4.4.4) is

$$r_{dd}(q) \triangleq E\{d_k \bar{d}_{k-q}\} = \sum_{m=1}^{N} A_m^2 e^{jW_m q} + r_{nn}(q) \quad (4.4.5)$$

where $r_{nn}(q)$ denotes the noise correlation.

The vector W* minimizes $E\{|\varepsilon_k|^2\}$ hence, by the matrix Wiener-Hopf equation

$$W^* = R_{xx}^{-1} P \quad (4.4.6)$$

In our case, where the input is described by (4.4.4) and (4.4.5) the matrix R_{xx} can be written as

$$R_{xx} = \sum_{m=1}^{N} \gamma_m \gamma_m^H + R_{nn} = \Gamma\Gamma^H + R_{nn} \quad (4.4.7)$$

where R_{nn} is the covariance matrix of n_k. Applying the well known matrix Inversion lemma for (4.4.7) then in [62]

$$R_{xx}^{-1} = R_{nn}^{-1} [I_L - \Gamma R^{-1} \Gamma^H R_{nn}^{-1}] \quad (4.4.8)$$

where I_L is the L×L identity matrix and

$$R = I_N + \Gamma^H R_{nn}^{-1} \Gamma$$

To find the vector P, we assume that the dealy Δ has been chosen correctly, i.e., large neough to sufficiently decorrelate n_k, the wide band component of the input. In this case P includes only the sinusoidal part given by

$$P = \sum_{m=1}^{N} A_m e^{-jW_m \Delta} \gamma_m = \Gamma V \Delta \qquad (4.4.9)$$

where the delays vector V_Δ is defined by

$$V_\Delta \triangleq |A_1 e^{-jW_1 \Delta}, \ldots, A_N e^{-jW_N \Delta}|^T$$

The optimal (complex weight vector (4.4.6) can now be rewritten as

$$\begin{aligned} W^* &= R_{nn}^{-1} [I_L - \Gamma R^{-1} \Gamma^H R_{nn}^{-1}] \quad \Gamma V_\Delta \\ &- R_{nn}^{-1} \quad \Gamma R^{-1} [R - \Gamma^H R_{nn}^{-1} \Gamma] V_\Delta \end{aligned} \qquad (4.4.10)$$

or using the definition (4.4.9) of R, we finally get

$$W^* = R_{nn}^{-1} \Gamma R^{-1} \quad V_\Delta \qquad (4.4.11)$$

or the optimal Wiener solution W* can be described by the sum

$$W^* = \sum_{m=1}^{N} W^*_m \tag{4.4.12}$$

where W^*_m is the Wiener solution for a single sinusoidal signal in colored noise at frequency W_m given by [62]

$$W^* = \frac{A_m^2}{1 + A_m^2 \psi(W_m, W_m)} e^{-jW_m \Delta} R_{nn}^{-1} \gamma(W_m)$$

and

$$\psi(W_m, W_m) = \frac{1}{A_m^2} [\Gamma^H R_{nn}^{-1} \Gamma]_{m,m} = \gamma^H(W_m) R_{nn}^{-1} \gamma(W_m)$$

Since

$$H(W) \Big|_{z=1} = \sum_{i=0}^{L-1} \bar{W}_i z^{-(\Delta+i)} \Big|_{z=1} = W^{*H} \gamma(W) e^{-jW\Delta} \Big|_{W=0} = W^{*H}$$

$$H(o) = V_\Delta^H R^{-1} \Gamma^H R_{nn}^{-1} \gamma(W) e^{-jW\Delta} \Big|_{W=0} = W^{*H} \tag{4.4.13}$$

The sinusoidal component at the predictor output is

$$S_{yk} = \sum_{m=1}^{N} H(W_m) A_m e^{j(W_m k + \psi_m)} = W^{*H} S_k \tag{4.4.14}$$

where S_k denotes the sinusoidal component of the data vector x_k given by

$$S_k = \sum_{m=1}^{N} A_m e^{j(W_m(k-\Delta)+\psi_m)} \quad (W_m) = V_{S_k} \qquad (4.4.15)$$

where

$$V_{S_k} = \left| e^{j(W_1(k-\Delta)+\psi_1)}, \ldots, e^{j(W_N(k-\Delta)+\psi_n)} \right|^T$$

Therefore the output signal component is

$$S_{yk} = V_\Delta^H R^{-1} \Gamma^H R_{nn}^{-1} \Gamma V_{S_k} \qquad (4.4.16)$$

For the particular input of real sinusoids in 56 and 62 S_{yk} is given by

$$S_{yk} = \sum_{m=1}^{M} \frac{\rho_m \eta(W_m) L/2}{1 + \rho_m \eta(W_m) L/2} C_m \operatorname{Cos}(W_m + \psi_m) \qquad (4.4.17)$$

where the input SNR and the likelihood variable of the m'th sinusoidal component are defined by

$$\rho_m \triangleq C_m^2 / 2\delta_n^2$$

$$\eta(W_m) = \frac{\delta_n^2}{L} \psi(W_m, W_m)$$

Thus, each real sinusoid has amplitude gain given by

$$b^*_m \triangleq \frac{\rho_m \eta(W_m) L/2}{1 + \rho_m \eta(W_m) L/2} \qquad (4.4.18)$$

Formula (4.4.18) generalizes previous results which were found for a white noise merely by introducing the likelihood weighting factor $\eta(W_m)$. It shows that amplitude distortions may occur at the predictor output when the noise is colored through the dependency of $\eta(W_m)$ on the sinusoid frequencies and noise spectrum. (Note that $\eta(W_m) = 1$ for white noise.)

The total power of the output signal from the transversal filter having the ideal weights is therefore given by

$$E\{Y^*_k{}^2|\} = \sum_{m=1}^{M} (b^*_m C_m)^2/2 + (r_{nn})^2 \frac{2}{L} \sum_{m=1}^{M} (b^*_m)^2 \qquad (4.4.19)$$

The overall output SNR is given by

$$\rho_o^* = \frac{L}{2} [(b^*_m C_m)^2/2] \; / \; [r_{nn}{}^2 \sum_{m=1}^{M} (b^*_m)^2] \qquad (4.4.20)$$

$$\rho_o^* = \frac{L}{2} \sum_{m=1}^{M} \rho_{im} (b^*_m)^2 \; / \; \sum_{m=1}^{M} (b^*_m)^2$$

where $\rho_{im} = \dfrac{C_m{}^2}{2r_{nn}{}^2}$

The overall input SNR $_i$ is given by

$$\rho_i = \sum_{i=1}^{M} \rho_{im} \qquad (4.4.21)$$

and we define

$$P^* = \frac{\rho_o^*}{\rho_i} = \frac{L}{2} \sum_{m=1}^{M} \rho_{im}(b^*_m)^2 \; / \; \rho_i \sum_{m=1}^{M} (b^*_m)^2 \qquad (4.4.22)$$

P^* is the gain in SNR achieved by the ALE which has the Wiener solution weights. [There is no difference between colored and white noise case.]

For the actual weight case (LMS algorithm) we can replace the value of b_m^* in (4.4.18) to the (4.2.36) in white noise case.

$$P^*_L = 1 \; / \; \{\frac{2N}{L} + \frac{\mu L P_x}{1-\mu L P_x} [1 + \frac{2N}{L} (1 + \frac{2}{\rho_i}) + \frac{4N^2}{L^2 \rho_i} (1 + \frac{1}{\rho_i})]\} \qquad (4.4.23)$$

With the practical assumption that $\mu_o LPx << 1$ the optimal value for L is found by differentiating P_L^* with respect to L, to be

$$L_{opt} = \left[\frac{2N}{\mu o P_x} + \frac{4N}{\rho_i^2} + \frac{4N}{\rho_i}\right]^{\frac{1}{2}} \qquad (4.4.24)$$

which is similar for the case of white noise.

If the noise is colored through the dependency of $\eta(W_m)$ the L_{opt} will be different from (4.4.24).

CHAPTER 5

ADAPTIVE LATTICE FILTER

5.1. INTRODUCTION

In the field of signal processing it is sometimes desirable to make use of a filter which adapts itself to the input signal in such a way that the error output of the filter is minimized (i.e., the filter is designed to eliminate noise, interference echos or other unwanted signals). Such an adaptive filter is one aspect of linear prediction, the basic assumption of which is that the signal in question can be modeled as a linear combination of previous inputs and/or outputs of the filter. The traditional form of the adaptive filter is the tapped-delay-line prediction error filter (TDL) [36], [31].

However, depending on the form of calculation used this PEF may suffer from either poor resolution or lack of stability as well as a number of other calculation limitations [37], [25] For example, in LMS algorithm the identification will be better if the estimates of the tap gain coefficients are better. Better estimates are obtained by running the LMS algorithm longer. However, the signal statistics may not remain stationary over such longer intervals. Therefore it is useful to have a rapidly convergent algorithm and so called ladder or lattice filter implementations have been suggested for such purposes [36].

Another interesting difference between TDL and lattice structures for approximately the same amount of signal distortion is that the lattice algorithm will produce considerably less harmonic distortion than the TDL (LMS) algorithm [64].

In addition to these there are a number of important advantages to using the lattice structure. One of the most important advantages is the fact for each stage the backward prediction error at the output is orthogonal to both prediction erros at the input. This decouples successive stages, thereby enabling independent optimization of each stage of the lattice. This is in contrast to the TDL structure where the coefficients are adjusted jointly, leading to poor convergence properties. The convergence time of the TDL structure is determined by the ratio of largest to smallest eigenvalue of the correlation matrix of the signal set in the filters. However, no analytical studied of the convergence properties of the adaptive lattice structure.

Since the input-output relations of the TDL and lattice structure are identical their transfer functions in steady state will be the same. However, steady state will in general be attained much more rapidly with the lattice structure.

There is also a difference between TDL and lattice structure which is related to the optimization technique. For TDL the usual approach for the derivation of coefficients has been to use a noisy gradient descent algorithm to adapt the filter coefficients toward their "optimal" values under a minimum mean square error performance criterion. The coefficients of the lattice structure proposed by Morf [35] have been derived in a significantly different manner in that they satisfy a global least squares optimality criterion at every point in time.

Also a more recent form of adaptive filter providing a solution, is the lattice prediction error filter originally proposed by Burg for use in spectral estimation and independently derived by Itakura and Saito and they guarantee the stability of the estimated all pole filter without requiring windowing of the observed signal [15], [8], [9].

5.2. DERIVATION OF Δ STEP PREDICTOR IN LATTICE FORM

As mentioned in the introduction, a lattice form implementation of the TDL will be considered due to its potentially superior convergence properties. All derivations are performed for the case of known statistics.

Let {y(.)} be a zero mean stochastic process and {y(t)} be random variables from this process.

Let $\hat{y}(t \; t-1,t-n)$ be the linear least squares estimate (LLSE) of y(t) given y(t-1),..., y(t-n).

Define the nth-order forward and backward prediction errors as;

$$e_n(t) = y(t) - \hat{y}(t|t-1,t-n) \quad (1) \qquad (5.2.1)$$

and

$$r_n(t) = y(t-n) - \hat{y}(t-n|t-n+1,t) \quad (2) \qquad (5.2.2)$$

respectively. Let

$$e_n(t+\Delta-1) = y(t+\Delta-1) - \hat{y}(t+\Delta-1|t-1,t-n) \qquad (5.2.3)$$

Suppose that we have one more random variable y(t-n-1) and we wish to obtain the LLSE of y(t+Δ-1) given y(t-1),..., y(t-n-1). From the innovation approach to linear least square estimation we have [65], [36]

$$\hat{y}[t+\Delta-1|t-1,t-n-1] = \hat{y}[t+\Delta-1|t-1,t-n]$$

$$+ \text{[LLSE of } y(t+\Delta-1) \text{ given the new information received with } y(t-n-1)] \quad (5.2.4)$$

Since the new information received with $y(t-n-1)$ is given by $y(t-n-1) - \hat{y}(t-n-1|t-n,t-1)$ and from Equation (5.2.2) this is equal to $r_n(t-1)$ we can rearrange (5.2.4) as follows:

$$\hat{y}(t+\Delta-1|t-1,t-n-1) = \hat{y}(t+\Delta-1|t-1,t-n) + \text{[LLSE of } y(t+\Delta-1)|r_n(t-1)] \quad (5.2.5)$$

which can be expressed as [36], [57].

$$\hat{y}[t+\Delta-1|t-1,t-n-1] = \hat{y}(t+\Delta-1|t-1,t-n) + \frac{E(y(t+\Delta-1)r_n(t-1)}{E(|r_n(t-1)|^2)} r_n(t-1) \quad (5.2.6)$$

Let us substract $y(t+\Delta-1)$ from both sides of (5.2.6)

$$\hat{y}(t+\Delta-1|t-1,t-n-1) - y(t+\Delta-1) = \hat{y}(t+\Delta-1|t-1,t-n) + \frac{E[y(t+\Delta-1)r_n(t-1)]}{E[|r_n(t-1)|^2]} r_n(t-1) - y(t+\Delta-1) \quad (5.2.7)$$

From (5.2.3) we have

$$e_{n+1}(t+\Delta-1) = e_n(t+\Delta-1) + \frac{E[y(t+\Delta-1)r_n(t-1)]}{E(|r_n(t-1)|^2)} r_n(t-1) \tag{5.2.8}$$

From the definition of LLSE $r_n(t-1)$ is orthogonal to $y(t-1) \ldots y(t-n)$. Hence (5.2.8) can be computed as

$$e_{n+1}(t+\Delta-1) = e_n(t+\Delta-1) - \frac{E[e_n(t+\Delta-1)r_n(t-1)]}{R_n(t-1)} r_n(t-1) \tag{5.2.9}$$

where

$$R_n(t-1) = E[(r_n(t-1)^2]$$

Similarly we can derive the following relations for the (n+1) th order forward and backward prediction errors:

$$e_{n+1}(t) = e_n(t) - \frac{E(e_n(t)r_n(t-1))}{R_n(t-1)} r_n(t-1) \tag{5.2.10}$$

$$r_{n+1}(t) = r_n(t-1) - \frac{E(r_n(t-1)e_n(t))}{E_n(t)} e_n(t) \tag{5.2.11}$$

where

$$E_n(t) = E(|e_n(t)|^2)$$

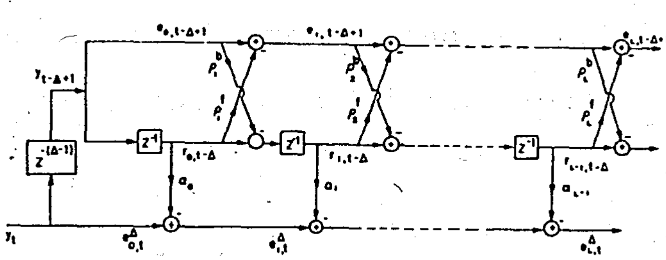
Changing the time index $t+\Delta-1$ to t in (5.2.9), (5.2.10) and (5.2.11) and varying the value of n from zero to $L-1$, we obtain the Lth order lattice filter structure of Figure 5.2.1 where

$$\alpha_n = E(e_n(t)\ r_n(t-\Delta))\ /\ R_n(t-\Delta)$$

$$\rho^f_{n+1} = E(e_n(t-\Delta+1)\ r_n(t-\Delta))\ /\ R_n(t-\Delta)$$

$$\rho^b_{n+1} = E(r_n(t-\Delta)\ e_n(t-\Delta+1))\ /\ E_n(t-\Delta+1)$$

In TDL the input and error output at time t are given by y(t) and $y(t) - \hat{y}(t|t-\Delta,t-\Delta-L+1)$. Since $e_o(t) = y(t)$ and $e_L(t) = y(t) - \hat{y}(t|t-\Delta,t-|-L+1)$ the structure of Figure 5.2.1 is the lattice form structure filter.



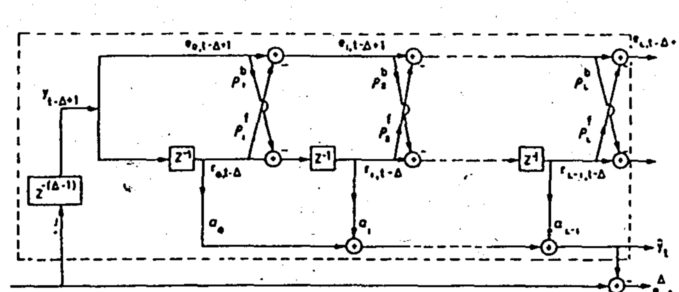


FIGURE 5.2.1. Lattice Form of TDL (ALE).
FIGURE 5.2.2. Equivalent representation of the Lattice form of TDL (AL

Redrawing the circuit of Figure 5.2.1 as shown in Figure 5.2.2, we see that the structure shown in the dotted box acts as a Δ-step predictor. Not that when $\Delta=1, \alpha_n = \rho^f_{n+1}$ and hence, the lattice form TDL (ALE) reduces to the well known lattice form linear prediction error filter [36].

5.3. LATTICE FORM LINEAR PREDICTION ERROR FILTER

Several lattice and ladder structures have been proposed for the implementation of all pole and pole-zero digital filters. However, only a single lattice structure due to Itakura and Saito [9] is available for the implementation of all zero filters. The lattice of Itakura and Saito had two multipliers in each stage. There are also one, two, three and four multiplier lattice structures. In particular the proper one is of the course the one multiplier form because of decreased number of multiplications.

In linear prediction, the signal spectrum is modeled by an all pole spectrum with a transfer function given by in [8], [3] and [9]

$$H(z) = \frac{G}{A(z)} \qquad (5.3.1)$$

where

$$A(z) = \sum_{k=0}^{P} a_k Z^{-k} \qquad a_o = 1$$

is known as the inverse filter. G is a gain factor, a_k are the predictor coefficients, and P is the number of poles or predictor coefficients in the model.

In order to analyze the spectral properties of the lattice filtering algorithm, it is useful to first consider the relationship of the reflection coefficients to the coefficients of the TDL. The TDL coefficients obey the constraints

$$a_{m,i} = 1 \qquad \text{for} \quad i=0 \qquad (5.3.2)$$

$$-1 \leq a_{m,m} \leq 1$$

and

$$a_{m.i} = 0 \text{ for } \quad i > m \qquad \text{or} \quad i < 0$$

The basic relationship between lattice and TDL types filters is that the reflection coefficient $\rho_i(n)$ equals the final coefficient $a_{i,i}$ of an ith order TDL for $1 \leq i \leq m$.

The filter coefficients of this TDL are then calculated from the Levinson recursion algorithm [8].

$$a_{m,i} = a_{m-1,i} + a_{m,m} \, a^*_{m-1,m-i} \qquad \text{for } (1 \leq i \leq m) \qquad (5.3.3)$$

by starting with m=2 and working up to the order of the filter. After each recursion the coefficients $a_{m,i}$ $1 \leq i \leq m$ are the desired coefficients for the m'th order predictor. The algorithm proceeds recursively to compute the following parameter sets

$$
\begin{array}{ll}
i=1 & \{a_{11}\} \\
i=2 & \{a_{21}, a_{22}\} \\
i=3 & \{a_{31}, a_{32}, a_{33}\} \\
i=4 & \{a_{41}, a_{42}, a_{43}, a_{44}\} \\
\vdots & \\
i=m & \{a_{m1}, a_{m2}, a_{m3}, \ldots, a_{mm}\}
\end{array}
$$

The parameters $\{a_{11}, a_{22}, a_{33}, \ldots, a_{mm}\}$ are often called the reflection coefficients and are designated as $\{\rho_1, \rho_2, \ldots, \rho_m\}$.

Therefore desired coefficients are $\{a_{m1}, a_{m2}, \ldots, a_{mm}\}$.

The theory of linear prediction lends an important interpretation to the Levinson-Durbin algorithm. Denote the prediction error for a mth order linear predictor as $f_m(n)$

$$f_m(n) = X_n + \sum_{k=1}^{m} a_m k \; X(n-k)$$

$$= \sum_{k=0}^{m} a_m k \; X(n-k) \qquad (5.3.4)$$

By using the Levinson-Durbin algorithm we have

$$f_m(n) = \sum_{k=1}^{m-1} (a_{m-1,k} + a_{mm}\, a^*_{m-1,m-k}) X(n-k)$$

$$+ a_{m,m} X(n-m) + X(n) = X(n) + \sum_{k=1}^{m-1} a_{m-1,k} X(n-k)$$

$$+ a_{m,m} X(n-m) + \sum_{k=1}^{m-1} a^*_{m-1,m-k} X(n-k)$$

Let

$$b_m(n) = X(n-m) + \sum_{k=1}^{m} a^*_{m,k} X(n-m+k) \qquad (5.3.5)$$

Therefore

$$f_o(n) = b_o(n) = X(n)$$

$$f_m(n) = f_{m-1}(n) + a_{m,m} b_{m-1}(n-1)$$

$$b_m(n) = a_{m,m} f_{m-1}(n) + b_{m-1}(n-1) \qquad (5.3.6)$$

The term $b_m(n)$ is the backward prediction error, i.e., the error when one attempts to predict X(n-m) on the basis of samples X(n-m+1) ... X(n). The relationships of (5.3.4) and (5.3.5) give again the lattice filter structure as shown in Figure 5.3.1.

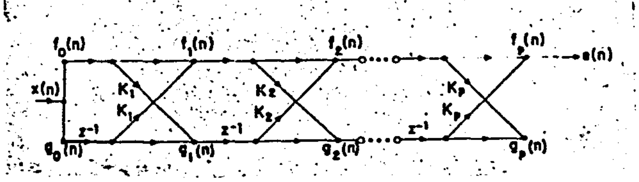


FIGURE 5.3.1. Lattice Formulation of Prediction error (Whitening or inverse filter).

Note that the transfer function of the entire filter is just

$$H(z) = \frac{1}{1 + \sum_{k=1}^{m} a_{mk} Z^{-k}} \qquad (5.3.7)$$

This filter is often called either the "inverse" filter or "prediction error" filter. If X(n) is the input signal, $f_m(n)$ is the forward residual at stage m and $b_m(n)$ is the backward residual at stage m. In z transform notation (5.3.6) can be written as

$$F_o(z) = B_o(z) = X(z)$$

$$F_m(z) = F_{m-1}(z) + a_{mm} Z^{-1} B_{m-1}(z)$$

$$B_m(z) = a_{mm} F_{m-1}(z) + Z^{-1} B_{m-1}(z) \qquad (5.3.8)$$

Let the forward and backward transfer functions at stage m be defined by

$$A_m(z) = \frac{F_m(z)}{X(z)} = \frac{F_m(z)}{F_o(z)}$$

and

$$G_m(z) = \frac{B_m(z)}{X(z)} = \frac{B_m(z)}{B_o(z)} \qquad (5.3.9)$$

Then from (5.3.8) and (5.3.9) it is easy to see that $A_m(z)$ and $G_m(z)$ obey the recursion relations

$$A_o(z) = G_o(z) = 1$$

$$A_m(z) = A_{m-1}(z) + a_{mm} Z^{-1} G_{m-1}(z)$$

$$G_m(z) = a_{mm} A_{m-1}(z) + Z^{-1} G_{m-1}(z) \qquad (5.3.10)$$

Furthermore one can show from (5.3.10) that

$$G_m(z) = Z^{-m} A_m(z^{-1}) \qquad (5.3.11)$$

Thus, if $A_m(z)$ is given by

$$A_m(z) = \sum_{k=0}^{m} a_m(k) Z^{-k} \qquad (5.3.12)$$

where $a_m(k)$ are the polynomial coefficients for an m stage lattice then

$$G_m(z) = \sum_{k=0}^{m} a_m(m-k) Z^{-k} \qquad (5.3.13)$$

and $G_m(z)$ is the reverse polynomial corresponding to $A_m(z)$. From (5.3.10) and (5.3.12) we also have

$$a_m(o) = 1$$

$$a_m(m) = a_{m.m} \qquad (5.3.14)$$

Now, given some polynomial $A_p(z)$ with $a_p(o) = 1$ one can generate all the polynomials $A_m(z)$, $m<p$ and the coeffi-

cients $a_{m,m}$ using the following reverse recursion derived from (5.3.10)

$$a_{m,m} = a_{m(m)}$$

$$A_{m-1}(z) = \frac{A_m(z) - a_{m,m}\, G_m(z)}{1 - a_{m,m}^2} \qquad (5.3.15)$$

along with (5.3.11) and beginning with m=p. It is clear from (5.3.15) that should $|a_{m',m'}| = 1$ for some m' = m, then the solution for $A_{m'-1}(z)$ is indeterminate. Therefore the reverse recursion (5.3.15) is possible iff $|a_{m,m}| \neq 1$ for all m.

It also follows from (5.3.11), (5.3.12) and (5.3.13) that the zeros of $G_m(z)$ are the reciprocal of the zeros of $A_m(z)$. In particular if all the zeros of $A_m(z)$ fall inside the unit circle, in which case $A_m(z)$ is minimum phase, then $G_m(z)$ is maximum phase. One can show that the minimum phase condition for $A_m(z)$ is guaranteed iff

$$-1 < a_{i,i} < 1 \qquad 1 \leq i \leq m \qquad (5.3.16)$$

The coefficients $a_{m,m}$ are taken as reflection coefficients or partial correlation coefficients. Therefore from (5.3.16) $A_m(z)$ and $G_m(z)$ are minimum and maximum phase respectively.

CHAPTER 6

ALGORITHMS FOR THE CALCULATION OF LATTICE FILTERS

6.1. INTRODUCTION

The algorithms suggested for the calculation of the reflection coefficients $\rho_i(n)$ all have in common the basic objective of minimizing the mean square forward and backward errors (the output of each filter stage) i.e., to obtain the lowest values of $F_i(n)$ and $B_i(n)$ defined by the expectations

$$F_i(n) = E[|f_i(n)|^2] \tag{6.1.1}$$

and

$$B_i(n) = E[|b_i(n)|^2] \tag{6.1.2}$$

Differentiating these quantities with respect to the reflection coefficient gives two values for the coefficient by minimizing the forward and backward mean square errors separately. The equation

$$\rho_i^F(n) = \frac{C_{i-1}(n)}{B_{i-1}(n-1)} \tag{6.1.3}$$

minimizes the forward error, and

$$\rho_i^B(n) = \frac{C_{i-1}(n)}{F_{i-1}(n)} \qquad (6.1.4)$$

minimizes the backward error. The factor $C_i(n)$ is the expectation of the negative cross-power of forward and backward errors, given by

$$C_i(n) = - E[f_i(n) \cdot b_i^*(n-1)] \qquad (6.1.5)$$

(where * denotes complex conjugation). This section looks at four algorithms suggested for minimizing both forward and backward error expectations.

6.2. FORWARD AND BACKWARD (F+B) ALGORITHM

The most direct of these algorithms was suggested by Griffiths and simply uses $\rho_i^F(n)$ and $\rho_i^B(n)$ as the forward and backward reflection coefficients respectively or

$$\text{(F+B) Algorithms} \qquad \begin{aligned} \rho_i^f(n) &= \rho_i^F(n) \\ \rho_i^b(n) &= \rho_i^B(n) \end{aligned} \qquad (6.2.1)$$

This is the only algorithm for which the forward and backward reflection coefficients are not the complex conjugates of each other.

The problem with this approach is that as $\rho^F(\rho^B)^* = 1$

under almost all circumstances either $\rho_i^f(n)$ or $\rho_i^b(n)$ will be greater than one, whereas for a stable filter the reflection coefficient should have a value less than one. Note that since $F_{i-1}(n)$ and $B_{i-1}(n-1)$ are both non-negative and the numerators in (6.1.3) and (6.1.4) are identical ρ_i^f and ρ_i^b always have the same sign S

$$S = \text{sign } \rho^f = \text{sign } \rho^b \qquad (6.2.2)$$

6.3. FORWARD/BACKWARD - MINIMUM (M) ALGORITHM

It follows that if either $\rho_i^F(n)$ or $\rho_i^B(n)$ is greater than one, then the other will be less than one. Thus an alternative to the (F+B) approach (in order to guarantee stability) is to choose the value with the smaller magnitude as $\rho_i^M(n)$ for all values of i and n. Such an algorithm was suggested by Makhoul |8| and is formulated as |9|

$$\rho_i^f(n) = \rho_i^M(n) = \frac{C_{i-1}(n)}{\max\{F_{i-1}(n),\ B_{i-1}(n-1)\}}$$

M
Algorithm

$$\rho_i^b(n) = \{\rho_i^M(n)\}^* \qquad (6.3.1)$$

or we can write

$$\rho_i^M(n) = S \min \{|\rho_i^f|, \ |\rho_i^b|\} \tag{6.3.2}$$

Since $(\rho_i^f)(\rho_i^b)^* = 1$

$$\text{If } |\rho_i^f| > 1 \qquad \text{then } |\rho_i^b| < 1$$

$$\text{or if } |\rho_i^b| > 1 \qquad \text{then } |\rho_i^f| < 1$$

It satisfies again $|\rho_i^b| < 1$ and $|\rho_i^f| < 1$.

This says that at each stage compute ρ_i^b and ρ_i^f and choose as the reflection coefficient the one with the smaller magnitude.

6.4. GEOMETRIC-MEAN (G) ALGORITHM

There are two major algorithms presently in use which attempt to minimize the forward and backward error expectation jointly. These algorithms were developed independently at about the same time. The algorithm originated by Itakura and Saito uses the geometric mean of the forward and backward expectations and is given by

G Algorithm

$$\rho_i^f(n) = \rho_i^G(n) = \frac{C_{i-1}(n)}{|F_{i-1}(n)B_{i-1}(n-1)|^{1/2}}$$

$$\rho_i^b(n) = |\rho_i^G(n)|^* \tag{6.4.1}$$

$\rho_i^G(n)$ is the negative of the statistical correlation between $f_i(n)$ and $b_i(n-1)$. From the properties of the geometric mean, it follows that

$$\min\left|\,|\rho_i^f|\,|\rho_i^b|\,\right| \le |\rho_i^G| \le \max\left|\,|\rho_i^f|\,|\rho_i^b|\,\right| \qquad (6.4.2)$$

Now since $|\rho_i^G| < 1$ it follows that if the magnitude of either ρ_i^f or ρ_i^b is greater than one, the magnitude of the other is necessarily less than one. This property brings to mind another possible definition for the reflection coefficient that guarantees stability.

6.5. HARMONIC MEAN (H) ALGORITHM

The other major algorithm was developed by Burg for use in spectral estimation and uses the harmonic mean of the forward and backward values

H Algorithm

$$\rho_i^f(n) = \rho_i^H(n) = \frac{2C_{i-1}(n)}{|F_{i-1}(n) + B_{i-1}(n-1)|}$$

$$\rho_i^b(n) = |\rho_i^H(n)|^* \qquad (6.5.1)$$

and one can show that

$$|\rho_i^m| \le |\rho_i^H| \le |\rho_i^G| \qquad (6.5.2)$$

One of the important property of ρ_i^H that is not shared by ρ_i^G and ρ_i^M is that ρ_i^H results directly from the minimization of error criterion.

In addition to the algorithms presented here, there are an infinite number of possible algorithms falling into a class for which the forward or backward error minimum, geometric mean and harmonic mean algorithms are special cases. However, Burg's harmonic-mean algorithm can be seen to result directly from the minimization of a well defined error criterion. This criterion minimizes the sum of the variances of the forwards and backwards residuals.

The error is defined as the sum of the variances of the forward and backward residuals.

$$E_{i+1}(n) = F_{i+1}(n) + B_{i+1}(n) \tag{6.5.3}$$

Using the recursive equation for $f_i(n)$ and $b_i(n)$ one can show that the forward and backward minimum errors at stage (i+1) are related to those at stage i by the following

$$F_{i+1}(n) = [1 - (\rho_{i+1}^H)^2]\, F_i(n) \tag{6.5.4}$$

$$B_{i+1}(n) = [1 - (\rho_{i+1}^H)^2]\, B_i(n-1) \tag{6.5.5}$$

This formulation is originally due to Burg.

6.6. GENERAL METHOD

Between ρ_i^M and ρ_i^G there are infinity of values that can be choosen as valid reflection coefficients (i.e., $|\rho|<1$). These can be conveniently defined by taking the generalized rth mean of ρ_i^f and ρ_i^b.

$$\rho_i^r = [\tfrac{1}{2} (|\rho_i^f|^r + |\rho_i^b|^r)]^{1/r} \qquad (6.6.1)$$

As $r \to 0$ $\rho_i^r \to \rho_i^G$, the geometric mean. For $r > 0$, ρ_i^r can not be guaranteed to satisfy $|\rho|<1$. Therefore for ρ_i^r to be a reflection coefficient, we must have $r \leq 0$. In particular

$$\rho_i^o = K^G \qquad \rho_i^{-\infty} = \rho_i^M \qquad (6.6.2)$$

If the signal is stationary one can show that

$$\rho_i^f = \rho_i^b \qquad (6.6.3)$$

and that

$$= \rho_i^f = \rho_i^b \quad \text{for all } r \qquad (6.6.4)$$

CHAPTER 7

RECURSIVE ESTIMATION OF THE REFLECTION COEFFICIENTS

7.1. INTRODUCTION

When dealing with adaptive filtering of signals whose statistics are expected to change (either continuously or abruptly), it is desirable to design the filter to be continuously adaptive so that the filter characteristics may change along with those of the signal. The general approach to make the system adaptive is to modify the reflection coefficients by making them recursive (i.e., updated with each sample) at the same time by allowing them to forget past samples as they become more distant in time. The forgetting feature of the algorithm is controlled by an adaptive weighting constant that is exponential in nature, giving more weight to the more recent samples which better represent the current signal statistics. It is a kind of a sliding exponential window technique.

The adaptive constant alone sets the rate at which the parameters of the lattice structure filter converge to a new set of values unlike the traditional tapped delay line adaptive algorithms (e.g., least mean squares) where the signal statistics also play a part in convergence behavior.

There are two basic methods for recursive estimation of the adaptive form of the reflection coefficients. These methods are presented here using Burg's harmonic mean algorithm, but they can be equally well used with any of the other available algorithms which were studied in Chapter 6 before.

The first method adds an update term directly to the reflection coefficients at each recursion while the second method updates the summation of $\rho_i(n)$ separately.

7.2. METHOD 1

The simplest approach to the recursive estimation of the reflection coefficients is to consider the new coefficient as being the sum of the old coefficient and a correction term. The correction term is just the difference between the new and old values of the coefficients as given by [68].

$$\rho_{m+1}(n) - \rho_{m+1}(n-1) = \frac{-2 \sum_{i=1}^{n} |f_m(i) b_m^*(i-1)|}{\sum_{i=1}^{n} [|f_m(i)|^2 + |b_m(i-1)|^2]} + \frac{2 \sum_{i=1}^{n} |f_m(i) b_m^*(i-1)|}{\sum_{i=1}^{n-1} [|f_m(i)|^2 + |b_m(i-1)|^2]} \qquad (7.2.1)$$

Note that the difference between n and n-1 as the limits on the summations. This equation can be written as the sum of the old coefficient and a new update term which contains only information from the present time interval (i.e., the input to that filter stage) both multiplied by a third term. This results in the equation

$$\rho_{m+1}(n) - \rho_{m+1}(n-1) = \left[\frac{-2|f_m(n) b_m^*(n-1)}{|f_m(n)|^2 + |b_m(n-1)|^2} - \rho_{m+1}(n-1) \right] \cdot \frac{|f_m(n)|^2 + |b_m(n-1)|^2}{\sum_{i=1}^{n} [|f_m(i)|^2 + |b_m(i-1)|^2]} \quad (7.2.2)$$

Rearranging (7.2.2) gives

$$\rho_{m+1}(n) = |1 - \gamma(n)| \, \rho_{m+1}(n-1) - \frac{2\gamma(n) \; f_m(n) b_m^*(n-1)}{|f_m(n)|^2 + |b_m(n-1)|^2} \quad (7.2.3)$$

where

$$\gamma(n) = \frac{|f_m(n)|^2 + |b_m(n-1)|^2}{\sum_{i=1}^{n} ||f_m(i)|^2 + |b_m(i-1)|^2|}$$

It can be seen that for the steady state (constant power) case $\gamma(n) \cong \frac{1}{n}$ where n is the number of data samples processed. If however $\gamma(n) = \gamma$ is held constant in the calculation, then it may be replaced by using the weighting factor ω as defined by the formula

$$\omega = 1 - \gamma = 1 - 1/n' \cong e^{-1/n'} \quad \text{(for } n' >> 0) \tag{7.2.4}$$

where n' is the theoretical data adaptive length of the filtering action. (For $n' \geq 10$, the exponential form of (7.2.4) is less than 0.5 percent from the actual value.)

Also in |67| there is such a situation which is summarized as follows. In deterministic least squares algorithm we choose the adaptation criterion for the filter as the minimization of

$$V = \frac{1}{2} \sum_{s=0}^{t} e^2(s) \tag{7.2.5}$$

with respect to the filter parameters. When the statistics of the observed process vary slowly, an exponential weighting is applied to the data so as to track the slowly varying parameters of the process. Weighting of the data with a sliding exponential window is equivalent to minimizing

$$V = \frac{1}{2} \sum_{s=0}^{t} \lambda^{t-s} e^2(s) \qquad \lambda \leq 1 \qquad (7.2.6)$$

where λ is a so called forgetting factor. The effect of λ reflects itself in the recursion of error covariance.

Rewriting (7.2.3) with ω gives

$$\rho_{m+1}(n) = \omega\rho_{m+1}(n-1) + \alpha_m(n) f_m(n) b_m^*(n-1) \qquad (7.2.7)$$

where

$$\rho_{m+1}(o) \qquad \text{(for the normal case)}$$

and the adaptive step size $\alpha_m(n)$ is given as

$$\alpha_m(n) = -2(1 - \omega) / [|f_m(n)|^2 + |b_m(n-1)|^2] \qquad (7.2.8)$$

The recursive relationship in (7.2.7) can also be written as the sum

$$\rho_{m+1}(n) = \sum_{i=1}^{n} [\omega^{(n-i)} \alpha_m(i) f_m(i) b_m^*(i-1)] \qquad (7.2.9)$$

Again (7.2.9) is similar to (7.2.6). An implicit condition on this recursive relationship is that the power of the prediction error $f_m(n)$ or $b_m(n-1)$ is not a time varying function.

7.3. METHOD 2

A second approach is to retain both summations as in $\rho_{m+1}(n)$ and enlarge them at each time interval. Thus the equation becomes

$$\rho_{m+1}(n) = \frac{V_{m+1}(n)}{Y_{m+1}(n)} \tag{7.3.1}$$

where

$$V_{m+1}(n) = \mu V_{m+1}(n-1) - 2f_m(n)b_m^*(n-1)$$

and

$$Y_{m+1}(n) = \mu Y_{m+1}(n-1) + |f_m(n)|^2 + |b_m(n-1)|^2$$

The initial conditions are $V_{m+1}(o) = Y_{m+1}(o) = 0$. The weighting factor is introduced to regulate the importance of the new term in the summation with respect to the previous term and thus control the adaptive speed of the filter. Normally, is in the range of 0 _1. This recursive relationship is equivalent to the equation

$$\rho_{m+1}(n) = \frac{-2 \sum_{i=1}^{n} [\mu^{(n-i)} f_m(i)b_m^*(i-1)]}{\sum_{i=1}^{n} \mu^{(n-i)} [|f_m(i)|^2 + |b_m^*(i-1)|^2]} \tag{7.3.2}$$

which, in turn is equivalent to actual $\rho_{m+1}(n)$ with the forward prediction error $f_m(i)$ and the delayed backward prediction error $b_m(i-1)$ weighted by the factor $\mu^{(n-i)/2}$. This form of weighting does not affect the stationarity of the input. Method 2 has the advantage over Method 1 of not assuming constant power. However, Method 2 is more complex computationally.

Here it should be noted that the factors ω and μ have no relationship to each other except that they both approach to zero.

7.4. CONVERGENCE PROPERTIES OF METHOD 1

An important characteristic of the adaptive filter is the rate at which the reflection coefficients convergence to their optimum values for given (stationary) input signal statistics. This rate of convergence is controlled by the adaptive weighting parameter (ω or μ).

The instantaneous estimate of the first reflection coefficient at time n can be defined as

$$\rho_1'(n) = \frac{-2\ f_o(n)\ b_o{}^*(n-1)}{|f_o(n)|^2 + |b_o(n-1)|^2} \qquad (7.4.1)$$

Combining (7.2.7), (7.2.8) and (7.5.1) we have

$$\rho_1(n) = \omega\ \rho_i(n-1) + (1 - \omega)\ \rho_1'(n) \qquad (7.4.2)$$

For a truly stationary process beginning at time n=0 the instantaneous estimation of (7.4.1) for n>1 will in fact be equal to the optimum value of the reflection coefficient ρ_1. Using this fact and given the initial value of the reflection coefficient $\rho_1(o)$ (for example, the filter's start-up values, or the value for a previous time series to which the filter has adapted) the filter's convergence equation can be computed by repeated application of the recursion equation (7.4.2) as

$$\rho_1(n) = \omega^n \rho_1(o) + (1 - \omega)\omega^{n-1}\rho_1' + (1 - \omega) \sum_{i=0}^{n-2} |W^i \hat{\rho}_1|$$

$$= \omega^n \rho_1(o) + (1 - \omega)\ {}^{n-1}\rho_1' + (1 - \omega^{n-1})\hat{\rho}_1 \qquad (7.4.3)$$

From this, the fractional error in the reflection coefficient at time n can be computed as

$$\varepsilon_1(n) = \frac{\hat{\rho}_1 - \rho_1(n)}{\hat{\rho}_1} = W^{n-1}\ [1 - \frac{\omega\rho_1(o) + (1 - \omega)\rho_1'(1)}{\hat{\rho}_1}] \qquad (7.4.4)$$

The factor $\rho_1'(1)$ need not be known for the most practical applications of this filter. Indeed, for the initial start-up case where $\rho_1(o) = X(o) = 0$, we have $\rho_1'(1) = 0$, resulting in the simplified versions

$$\rho_1(n) = (1 - \omega^{n-1})\,\hat{\rho}_1 \qquad (7.4.5)$$

and

$$\varepsilon_1(n) = \omega^{n-1} \qquad (7.4.6)$$

for (7.4.3) and (7.4.4) respectively.

For the transition case where $\rho_1(o)$ is known but not equal to zero, given values of ω approaching unity (which is the common case) and therefore $\rho_1'(1) \approx \rho_1(o)$, (7.4.3) and (7.4.4.) can be simplified respectively as follows:

$$\rho_1(n) \cong (1 - \omega^{n-1})\hat{\rho}_1 + \omega^{n-1}\rho_1(o) \qquad (7.4.7)$$

and

$$\varepsilon_1(n) \cong \omega^{n-1}\,(1 - \rho_1(o)/\hat{\rho}_1) \qquad (7.4.8)$$

This measure of convergence error can also be written in terms of the ratio of the data length actually processed to the theoretical data adaptive length n' by applying (7.2.4). Thus (7.4.6) becomes

$$\varepsilon_1(n) = e^{-(n-1)/n'} \quad \text{for } n>0 \text{ and } n'>>0 \qquad (7.4.9)$$

7.5. CONVERGENCE PROPERTIES OF METHOD 2

Similar to the discussion for Method 1, instantaneous estimates can be made for the numerator and denominator terms used in the calculation of the first reflection coefficient by Method 2. For a truly stationary process, these estimates $V_1'(n)$ and $y_1'(n)$, as defined by

$$V_1'(n) = -2\ f_o(n) b_o^*(n-1) \qquad (7.5.1)$$

and

$$Y_1'(n) = |f_o(n)|^2 + |b_o(n-1)|^2 \qquad (7.5.2)$$

are equal to the optimum vales $\hat{V}_1$ and $\hat{Y}_1$ (for n 1) such that $\hat{V}_1/\hat{Y}_1 = \hat{\rho}_1$. Combining denominator and numerator of (7.3.1) with (7.5.1) and (7.5.2) give the recursion relationships

$$V_1(n) = \mu V_1(n-1) + V_1'(n) \qquad (7.5.3)$$

and

$$Y_1(n) = \mu Y_1(n-1) + Y_1'(n) \qquad (7.5.4)$$

Repeated applications of these recursions results in the following formulas for the reflection coefficient

$$\rho_1 = \frac{V_1(n)}{Y_1(n)} = \frac{\sum_{i=0}^{n-2} [\mu^i \hat{V}_1] + \mu^{n-1} V_1'(1) + \mu^n V_1(o)}{\sum_{i=0}^{n-2} |\mu^i \hat{Y}_1| + \mu^{n-1} Y_1'(1) + \mu^n Y_1(o)} \qquad (7.5.5)$$

or for $\mu \neq 1$

$$\rho_1(n) = \frac{\frac{1-\mu^{n-1}}{1-\mu}\hat{V}_1 + \mu^{n-1}V_1'(1) + \mu^n V_1(o)}{\frac{1-\mu^{n-1}}{1-\mu}\hat{Y}_1 + \mu^{n-1}Y_1'(1) + \mu^n Y_1(o)} \qquad (7.5.6)$$

These equations are difficult to simplify significantly, except for the initial start-up where $V_1(o) = Y_1(o) = X(o) = 0$. Then $b_o(o) = 0$ and therefore $V_1'(1) = 0$ and $Y_1'(1) = |f_o(1)|^2$. In a stationary environment, the forward and backward prediction error powers are equal, so that $Y_1'(1) = Y_1/2$ simplifying (7.5.6) to

$$\rho_1(n) = \frac{\frac{1-\mu^{n-1}}{1-\mu}\hat{V}_1}{\left(\frac{1-\mu^{n-1}}{1-\mu} + \frac{\mu^{n-1}}{2}\right)\hat{Y}_1} = \frac{2\,(1-\mu^{n-1})}{(2-\mu^{n-1}-\mu^n)} \cdot \hat{\rho}_1$$

(7.5.7)

From this, the fractional error in the reflection coefficient at time n can be computed as

$$\varepsilon_1(n) = \frac{\hat{\rho}_1 - \rho_1(n)}{\hat{\rho}_1} = \frac{\mu^{n-1} - \mu^n}{2-\mu^{n-1}-\mu^n} \qquad (7.5.8)$$

Another special case of interest is when $\mu=1$, for which (7.5.5) simplies to

$$\rho_1(n) = \frac{(n-1)\hat{V}_1 + V_1'(1) + V_1(o)}{(n-1)\hat{Y}_1 + Y_1'(1) + Y_1(o)} \qquad (7.5.9)$$

For the initial start-up case as described above, this simpliefies further to

$$\rho_1(n) = \frac{(n-1)\,\hat{V}_1}{(n-1)\,\hat{Y}_1} = \frac{2n-2}{2n-1}\;\hat{\rho}_1 \qquad (7.5.10)$$

The corresponding value of the fractional error in $\rho_1(n)$ then becomes

$$\varepsilon_1(n) = \frac{\hat{\rho}_1 - \rho_1(n)}{\hat{\rho}_1} = \frac{1}{2n-1} \qquad (7.5.11)$$

As with Method 1, these convergence rates can also be applied to the relevant signal component at the filter stage output.

CHAPTER 8

CONCLUSIONS

For the cases of one and two sinusoids, we showed that substantial improvements could be obtained by choosing a suitable value of the delay parameter rather than the usual choice of $\Delta=1$. But there is a problem which is related to the computation of the optimal value of Δ. Calculation of the optimal value of Δ requires knowledge of W_i and L. This problem can be solved by considering the following discussion. Choosing the inital value of Δ as unity, carry out the recursions of (1.3.6) for a desired number of iterations and compute $|P(w)|^2$ from the resulting coefficient estimates. From the computed value of $|P(w)|^2$, estimate the value of W_i and use the formula which is related to the optimal value of Δ.

Comparing the simulations figure, it is seen that the simulation results agree very closely with the theory. The ALE with near optimum value of Δ gives a sharper spectral estimate. It is clear that the sharpness indicates how accurate the estimate is. This situation can also beeseen easily by observing a deep null for the case of two sinusoids. By taking the near the optimum value gives more information than the choice of $\Delta=1$.

In ALE $Z^{-\Delta}|e^{jw\Delta} = e^{-jw\Delta}$ acts like an all pass filter and consists only of poles and zeros at $Z=0$ or at $Z=\infty$, input and output of it both have the same magnitude on the unit circle and the transfer function must be entirely all-pass with unity magnitude.

For this reason we can determine the finite impulse response (FIR) filter such that the output energy is minimized subject to the following constraints. First constraint is $a^T a = 1$ and the second constraint includes the dynamic behavior of ALE which is given by (1.3.6). With this minimization, the performance of ALE in noise concelling will be better than the previous case.

Also the change of the position of $Z^{-\Delta}$ will change the performance of the ALE. By putting $Z^{-\Delta}$ in the first processor channel that is in the primary input we can change the performance but we can guarantee the decorrelation process in the noise components for two channels.

As the decorrelation parameter Δ is increased, a time window is produced within which the error process may be correlated at lag Δ and beyond its correlation remains zero. In general Δ plays a role for stability. With suitable time delays in filter design, causal approximations to delayed version of noncausal impulse responses are realizable.

In Chapter 4 guidelines for the optimal selection of the ALE parameters, namely the number of weights L and the adaptation step-size parameter μ are given by considering two different methods.

By using the optimal value of L we can get the more accurate expressions for Δ.

The results in Chapter 4 have clearly shown that the longest ALE filter is not necessarily the best and that significant performance reductions can be expected if incorrect filter lengths are employed.

In Chapter 6 different algorithms results yield different spectrum as shown in the simulations. The best one is Burg algorithm which specify the peak more clearly than others in the spectrum.

REFERENCES

1. B. Widrow, et at., "Adaptive Noise Cancelling: Principles and Applications", Proc. IEEE, Vol. 63, pp. 1692-1717, Dec. 1975.

2. L. Griffiths, "Rapid Measurement of Digital Instantaneous Frequency", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-23, pp. 209-222, April 1975.

3. J. Makhoul, "Linear Prediction: A Tutorial Review", Proc. IEEE, Vol. 63, pp. 561-580, April, 1975.

4. B. Widrow, "Adaptive Filters", in Aspects of Network and System Theory, R. Kalman and N. DeClaris, Eds, New York: Holt Rinehart and Winston, 1971, pp. 563-587.

5. B. Widrow, P. Mantey, L. Griffiths and B. Goode, "Adaptive Antenna Systems", Proc. IEEE, Vol. 55, pp. 2143-2159, Dec., 1967.

6. B. Widrow, J. McCool, M. Larimore and C. Johnson, Jr., "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter", Proc. IEEE, Vol. 64, pp. 1151-1162, Aug., 1976.

7. J.R. Zeidler, E.H. Satorious, et.al., "Adaptive Enhancement of Multiple Sinusoids in Uncorrelated Noise", IEEE Trans. ASSP, Vol. ASSP-26, June 1978, pp. 240-254.

8. J. Makhoul, "Stable and Efficient Lattice Methods for Linear Prediction", IEEE Trans. ASSP, Vol. ASSP-25, October 1977, pp. 423-428.

9. J. Makhoul, "A Class of All-Zero Lattice Digital Filters: Properties and Applications", IEEE Trans. ASSP, Vol. ASSP-26, August 1978, pp. 304-314.

10. J.R. Glover, Jr., "Adaptive Noise Cancelling Applied to Sinusoidal Interferences", IEEE Trans. Acoust. Speech, Signal Processing, Vol. ASSP-25, Dec., 1977.

11. D. Lee, M. Morf, and B. Friedlander, "Recursive Square-Root Ladder Estimation Algorithms", IEEE Trans. Acoust. Speech Signal Processing, Vol. ASSP-29, pp. 627-641, June 1981.

12. J. Treichler, "Transient and Convergent Behavior of the Adaptive Line Enhancer", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-27, pp. 53-62, Feb. 1979.

13. J. Glover, "High Order Algorithms for Adaptive Filters", IEEE Trans. Commun., Vol. COM-27, pp. 216-221, Jan. 1979.

14. L.J. Griffiths, "Adaptive Structure for Multi-input Noise Cancelling Applications", in Proc. ICASSP, Washington D.C., pp. 925-928, 1979.

15. I.L. Ayala, "On a New Adaptive Lattice Algorithm for Recursive Filters", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-30, pp. 316-319, April 1982.

16. D. Parikh, N. Ahmed, and S.D. Stearns, "An Adaptive Lattice Algorithm for Recursive Filters", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-28, pp. 110-111, Feb. 1980.

17. B.R. Davis and W.G. Cowley, "Bias and Variance of Spectral Estimates from an all-pole Digital Filter," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-30, pp. 322-329, April 1982.

18. W.S. Hodgkiss and J.A. Presley, Jr., "The Complex Adaptive Least Squares Lattice", IEEE Trans. Acoust, Speech, Signal Processing, Vol. ASSP-30, pp. 330-333, April 1982.

19. D.R. Morgan and S.E. Craig, "Real-time Adaptive Linear Prediction Using the Least-Mean Square Gradient Algorithm", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-24, pp. 494-507.

20. M. Morf, A.Vieira and D.R. Lee, "Ladder Forms for Identification and Speech Processing", in Proc. IEEE Conf. Decision Contr., Dec. 1977, pp. 1074-1078.

21. W.S. Hodgkiss and J.A. Presley, "Adaptive Tracking of Multiple Sinusoids whose Power Levels are Widely Separated", IEEE Trans. Circuit Syst., Vol. CAS-28, pp. 550-561, June 1981.

22. C.J. Gibson and S. Haykin, "A Comparison of Algorithms for the Calculation of Adaptive Lattice Filters", in Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing, Denver, Col., April 9-11, 1980, pp. 978-983.

23. M. Dentino, J. McCool and B. Widrow, "Adaptive Filtering in the Frequency Domain", Proc. IEEE, Vol. 66, pp. 1658-1659, Dec., 1978.

24. B.Widrow, J. McCool and B. Widrow, "The Complex LMS Algorithm", Proc. IEEE, Vol. 63, pp. 719-720, Apr. 1975.

25. F.W. Symons, "The Complex Adaptive Lattice Structure", IEEE Trans. Acoust. Speech, Signal Processing, Vol. ASSP-270, pp. 292-295, June 1979.

26. E.H. Satorious and S.T. Alexander, "Channel Equalization Using Adaptive Lattice Algorithms", IEEE Trans. Commun. Vol. COM-27, pp. 899-903, June 1979.

27. E.H. Satorius and J.D. Pack, "Application of Least Squares Lattice Algorithms to Adaptive Equalization", IEEE Trans. Commun., Vol. COM-29, pp. 136-142, Feb. 1981.

28. Ferrara, E.R., "Fast Implementation of LMS Adaptive Filters", IEEE Trans. ASSP, Vol. ASSP-28, August 1980.

29. Dennis R. Morgan, Samuel E. Craig, "Real-Time Adaptive Linear Prediction Using the Least Mean Square Gradient Algorithm", IEEE Trans. ASSP, Vol. ASSP-24, pp. 494-506, December 1976.

30. B. Widrow, "Adaptive Filters I: Fundamentals", Stanford Electronics Labs, Stanford, Calif., Rep. SEL-66-126 (Tech. Rep. 6754-6) Dec., 1966.

31. J.R. Treichler, "Response of the Adaptive Line Enhancer to Chirped and Doppler-Shifted Sinusoids", IEEE Trans. ASSP, Vol. ASSP-28, pp. 343-348, June 1980.

32. F.W. Symons, Jr., "Narrow Band Interference Rejection Using the Complex Digital Inverse Filter", IEEE Trans. ASSP, Vol. ASSP-26, pp. 94, Feb. 1978.

33. B. Widrow and M.E. Hoff, Jr., "Adaptive Switching Circuits", in IRG Wescon Conv. Rec., pt. 4, 1960, pp. 96-104.

34. B. Friedlander, T. Kailath, M. Morf, and L. Ljung, "Extended Levinson and Chandrasekhar Equation for General Discrete-Time Linear Estimation Problems", IEEE Trans. A.C., Vol. AC-23, pp. 653-659, August 1978.

35. M. Morf and T. Kailath, "Square-root Algorithms for Least-Squares Estimation", IEEE Trans. AC, Vol. AC-200, pp. 487-497, August 1975.

36. V.U. Reddy, B. Egardt and T. Kailath, "Optimized Lattice-Form Adaptive Line Enhancer for a Sinusoidal Signal in Broadband Noise", IEEE Trans. ASSP., Vol. ASSP-29, pp. 702-710, June 1981.

37. D.G. Messerschmitt, "A Class of Generalized Lattice Filters", IEEE Trans ASSP, Vol. ASSP-28, pp. 198-204, April 1980.

38. P.L. Chu and D.G. Messerschmitt, "Zero-Sensitivity Analysis of the Digital Lattice Filter", in Proc. 1980 ICASSP, (Denver), April 1980.

39. C.R. Johnson, Jr., "A Stable Family of Adaptive IIR Filters", in Proc. 1980 IEEE Int. Conf. on Acoustic, Speech, Signal, Processing, (Denver, Col.), pp. 1001-1004, April 1980.

40. I.D. Landau, "Unbiased Recursive Identification Using Model Reference Adaptive Techniques", IEEE Trans. AC, Vol. AC-21, pp. 194-202, April 1976.

41. I.D. Landau and H.M. Silveira, "A Stability Theorem with Applications to Adaptive Control", IEEE Trans. AC, Vol. AC-24, pp. 305-312, April 1979.

42. Y.H. Lin and K.S. Narendra, "A New Error Model for Adaptive Systems", IEEE Trans. AC, Vol. AC-24, pp. 305-312, April 1979.

43. A. Weiss and D. Mitra, "Digital Adaptive Filters: Conditions for Convergence, Rates of Convergence, Effects of Noise and Errors Arising from the Implementation", IEEE Trans. Inform. Theory, Vol. IT-25, pp. 637-652, Nov. 1979.

44. E.R. Ferrara and B. Widrow, "Multi-channel Adaptive Filtering for Signal Enhancement", IEEE Trans. ASSP, Vol. ASSP-29, pp. 766-770, June 1981.

45. E.R. Ferrara, "Fast Implementation of LMS Filters", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-28, pp. 474-475, August 1980.

46. N.J. Bershad and P.L. Feintuck, "Analysis of the Frequenc Domain Adaptive Filter", Proc. IEEE, Vol. 67, pp. 1658-1659, Dec. 1979.

47. J.T. Rickard and G.M. Dillard, "Adaptive Detection Algorithms for Multiple Target Situations", IEEE Trans. AES, Vol. AES-13, pp. 338-343, July 1977.

48. W.S. Burdic, "Detection of Narrow-band Signals Using the Time Domain Adaptive Filters", IEEE Trans. AES, Vol. AES-14, pp. 578-591, July 1978.

49. P.M. Reeves, "Detection of Narrow-band Signals Using Time Domain Adaptive Filters", in Proc. 1979 IEEE Conf. ASSP (Washington, D.C.), pp. 945-950, April 1979.

50. R. Medaugh and L.J. Griffiths, "Optimum Filter Length for an Adaptive Linear Predictor", in Proc. 12th Asimolar Conf. Circuits, Systems and Computers (Pacific Grove, Calif), pp. 175-180, Nov. 1978.

51. J.T. Richard, M. Dentino and J. Zeidler, "Detection Performance of an Adaptive Processor in Non-Stationary Noise", in Proc. 1979 IEEE Conf. ASSP (Washington, D.C.), pp. 136-139, April 1979.

52. M. Shensa, "The Spectral Dynamics of Evolving LMS Adaptive Filters", in Proc. 1979 IEEE Conf. ASSP (Washington D.C.), pp. 950-953, April 1979.

53. E.H. Satorius, J.R. Zeidler, and S.T. Alexander, "Linear Predictive Filtering of Narrow-band Processes in Additive Broad-band Noise", in Proc. 1979 IEEE Conf. ASSP (Washington, D.C.) pp. 937-941, April 1979.

54. J.T. Richard and J.R. Zeidler, "Second-order Output Statistics of the Adaptive Line Enhancer", IEEE Trans. ASSP, Vol. ASSP-27, pp. 31-39, Feb. 1979.

55. M. Dentino, H. Huey, W.S. Burdic and J.R. Zeidler, "Statistical Properties of the Adaptive Line Enhanced (ALE) Processor", in Proc. 12th Asimolar Conf. Circuits, Systems and Computers, pp. 184-189, Nov. 1978.

56. A. Nehorai and D. Malah, "On the Stability and Performance of the Adaptive Line Enhancer", in Proc. IEEE Int. Conf. ASSP (Denver, Col.), pp. 478-481, April 1980.

57. J. Makhoul, "Linear Prediction: A Tutorial Review", Proc. IEEE, Vol. 63, pp. 561-580, April 1975.

58. J.R. Treichler, " -LMS and Its Use in a Noise-Compensating Adaptive Spectral Analysis Technique", in 1979 Proc. IEEE Int. Conf. ASSP (Washington, D.C.), pp. 933-936, April 1979.

59. B. Widrow and J.M. McCool, "A Comparison of Adaptive Algorithms Based on the Methods of Steepest Descent and Random Search", IEEE Trans. Antennas Propagat., Vol. AP-24, pp. 615-637, September 1976.

60. E.I. Jury, Theory and Application of the Z Transform Method, New York: Wiley, 1964.

61. M.L. Honig and D.G. Messerschmitt, "Convergence Properties of an Adaptive Digital Lattice Filter", IEEE Trans. ASSP, Vol. ASSP-29, pp. 642-654, June 1981.

62. A. Nehorai and M. Morf, "Enhancement of Sinusoids in Colored Noise and Whitening Performance of Exact Least Squares Predictors", IEEE Trans. ASSP, Vol. ASSP-30, pp. 353-363, June 1982.

63. P.A. Thompson, "A Constrained Recursive Adaptive Filter for Enhancement of Narrow-band Signals in White Noise", in Proc. 12th Asimolar Conf. Circuits, Systems and Computers, (Pacific Grove, Calif.), pp. 214-219, Nov. 1978.

64. E.H. Satorius, J.D. Smith, and P.M. Reeves, "Adaptive Noise Cancelling Interference Using a Lattice Structure", in Proc. 1979 IEEE Conf. ASSP (Washington, D.C.), pp. 929-932, April 1979.

65. J.S. Meditch, Stochastic Optimal Linear Estimation and Control, New York: McGraw-Hill, 1969.

66. Dakshesh D. Parikh and Nasir Ahmed, "Sequential Regression Consideration of Adaptive Notch Filters", IEEE Trans. ASSP, Vol. ASSP-28, pp. 313-317, June 1980.

67. V.U. Reddy, B. Egardt and T. Kailath, "Least Squares Type Algorithm for Adaptive Implementation of Piserenko's Harmonic Retrieval Method", IEEE Trans. ASSP, Vol. ASSP-30, pp. 399-405, June 1982.

68. Carey J. Gibson and S. Haykin, "Learning Characteristics of Adaptive Lattice Filtering Algorithms", IEEE Trans. ASSP, Vol. ASSP-28, pp. 681-691, December 1980.

APPENDIX A1

Let $$m_1 = \frac{\delta_1^{\,2}}{\delta_1^{\,2} + \delta_2^{\,2}} \qquad m_2 = \frac{\delta_2^{\,2}}{\delta_1^{\,2} + \delta_2^{\,2}} \tag{A1.1}$$

$$\underset{W_a}{H(m_1 w_1 + m_2 w_2)} = \sum_{k=0}^{L-1} A_1 e^{jW_1 k} e^{-j(W_a)\Delta} e^{-j(W_a)k}$$

$$+ A_2 e^{jW_2 k} e^{-j(W_a)k} e^{-j(W_a)\Delta}$$

$$+ A_3 e^{-jW_1 k} e^{-j(W_a)k} e^{-j(W_a)\Delta}$$

$$+ A_4 e^{-jW_2 k} e^{-j(W_a)k} e^{-j(W_a)\Delta}$$

$$= e^{-j(W_a)\Delta} \Big[\sum_{k=0}^{L-1} A_1 e^{j(W_1 - W_a)k} + A_2 e^{j(W_2 - W_a)k}$$

$$+ A_3 e^{-j(W_1 + W_a)k} + A_4 e^{-j(W_2 + W_a)k} \Big]$$

$$= e^{-j(W_a)\Delta} \Big[A_1 \cdot T_1 e^{-j(W_1 - W_a)(L-1)/2}$$

$$+ A_2 T_2 e^{-j(W_2 - W_a)(L-1)/2} + A_3 T_3 e^{j(W_1 + W_a)(L-1)/2}$$

$$+ A_4 T_4^{\,j(W_2 + W_a)(L-1)/2} \Big] \qquad \text{(A1.2}$$

where

$$T_1 = \frac{\sin(W_1 - W_a)L/2}{\sin(W_1 - W_a)/2}$$

$$T_2 = \frac{\sin(W_2 - W_a)L/2}{\sin(W_2 - W_a)/2}$$

$$T_3 = \frac{\sin(W_1 + W_a)L/2}{\sin(W_1 + W_a)/2}$$

$$T_4 = \frac{\sin(W_2 + W_a)L/2}{\sin(W_2 + W_a)/2}$$

Since $$m_1 + m_2 = 1$$

Therefore, we can write

$$W_1 - W_a = m_2 \Delta W = m_2(W_1 - W_2)$$

$$W_2 - W_a = m_2 \Delta W - W = -m_1(W_1 - W_2) = -m_1 \Delta W$$

$$W_1 + W_a = 2W_1 - m_2 \Delta W$$

$$W_2 + W_a = 2W_2 + m_1 \Delta W \qquad (A1.3)$$

$$H(W_a) = e^{-jW_a\Delta}\Big\{A_1 \cdot T_1\, e^{-jm_2\Delta W(L-1)/2} + A_2T_2\, e^{jm_1\Delta W(L-1)/2}$$

$$+ A_3T_3\, e^{j(2W_1)(L-1)/2}\, e^{-j(+m_2\Delta W)(L-1)/2}$$

$$+ A_4T_4\, e^{j(2W_2)(L-1)/2}\, e^{+j(m_1\Delta W)(L-1)/2}\Big\} \qquad \text{(A1.4)}$$

$$H_1(W_a) = e^{-jW_a\Delta}\big[A_1T_1\, e^{-jm_2\Delta W(L-1)/2}$$

$$+ A_2T_2\, e^{j\Delta W(L-1)/2}\, e^{-jm_2\Delta W(L-1)/2}\big]$$

$$= e^{jW_a\Delta}\big[e^{-jm_2\Delta W(L-1)/2}\,(A_1T_1 + A_2T_2\, e^{j\Delta W(L-1)/2}\big]$$

$$H_2(W_a) = e^{-jW_a\Delta}\big[(A_3T_3\, e^{j2W_1(L-1)/2}$$

$$+ A_4T_4\, e^{j2W_2(L-1)/2}\, e^{+j\Delta W(L-1)/2}\, e^{-jm_2\Delta W(L-1)/2}\big]$$

$$H(W_a) = e^{-jW_a\Delta}\, e^{-jm_2\Delta W(L-1)/2}\Big\{A_1T_1 + A_2T_2\, e^{j\Delta W(L-1)/2}$$

$$+ A_3T_3\, e^{jW_1(L-1)} + A_4T_4\, e^{jW_2(L-1)}\, e^{j\Delta W(L-1)/2}\Big\}$$

$$A_1T_1 = \frac{1}{1-\gamma_{12}\gamma_{21}}\left[\frac{e^{jW_1\Delta}}{L+2\delta_o^{\,2}/\delta_1^{\,2}} - \frac{\gamma_{12}e^{jW_2\Delta}}{L+2\delta_o^{\,2}/\delta_2^{\,2}}\right]\frac{\mathrm{Sin}(m_2\Delta W)L/2}{\mathrm{Sin}(m_2\Delta W)/2}$$

Since $$\gamma_{12} = \frac{1}{L+2\delta_o^2/\delta_1^2} \cdot \frac{e^{j\Delta W(L-1)/2}}{\mathrm{Sin}(\Delta W)/2} \cdot \mathrm{Sin}(\Delta WL)/2$$

$$\gamma_{21} = \frac{1}{L+2\delta_o^2/\delta_2^2} \cdot \frac{e^{-j\Delta W(L-1)/2}}{\mathrm{Sin}(\Delta W/2)} \mathrm{Sin}(\Delta WL/2)$$

Let $$L + 2\delta_o^2/\delta_2^2 = M_2 \qquad M_3 = \frac{\mathrm{Sin}\ (\Delta WL/2)}{\mathrm{Sin}(\Delta W/2)}$$

$$L + 2\delta_o^2/\delta_1^2 = M_1$$

Therefore

$$\gamma_{12} \cdot M_1 = \bar{\gamma}_{21} M_2$$

$$\gamma_{12}\gamma_{21} = \frac{1}{M_1M_2}\left[\frac{\mathrm{Sin}(\Delta WL/2)}{\mathrm{Sin}(\Delta W/2)}\right]^2 = \frac{M_3^2}{M_1M_2}$$

$$A_1T_1 = \frac{1}{1 - \frac{M_3^2}{M_1M_2}}\left[\frac{e^{jW_1\Delta}}{M_1} - \frac{\gamma_{12}\ e^{jW_2\Delta}}{M_2}\right] T_1$$

$$A_1T_1 = \frac{1}{M_1M_2 - M_3^2}\left[M_2\ e^{jW_1\Delta} - M_1\gamma_{12}\ e^{jW_2\Delta}\right] T_1$$

Since

$$\gamma_{12} = \frac{M_3}{M_1}\ e^{j\Delta W(L-1)/2}$$

$$A_1T_1 = \frac{1}{M_1M_2 - M_3^2} \left[M_2\, e^{jW_1\Delta} - M_3\, e^{j(\Delta W(L-1)/2\, +\, W_2\Delta} \right] T_1$$

$$e^{j\Delta W(L-1)/2}\, A_2T_2 = \frac{1}{1 - \frac{M_3^2}{M_1M_2}} \left[\frac{e^{jW_2\Delta}}{M_1} - \frac{\gamma_{21}\, e^{jW_1\Delta}}{M_1} \right] T_2\, e^{j\Delta W(L-1)/2}$$

$$= \frac{1}{M_1M_2 - M_3^2} \left[M_1e^{jW_2\Delta} - M_2\gamma_{21}\, e^{jW_1\Delta} \right] T_2\, e^{j\Delta W(L-1)/2}$$

$$= \frac{1}{M_1M_2 - M_3^2} \left[M_1e^{jW_2\Delta} - M_3\, e^{j(W_1\Delta\, -\, \Delta W(L-1)/2} \right] T_2e^{j\Delta W(L-1)/2}$$

$$e^{jW_1(L-1)}\, A_3T_3 = \frac{1}{M_1M_2 - M_3^2} \left[M_2e^{-jW_1\Delta} - M_3e^{-j(\Delta W(L-1)/2 + W_2\Delta} \right] T_3e^{jW_1(L-1)}$$

$$A_4T_4 = \frac{1}{M_1M_2 - M_3^3} \left[M_1e^{-jW_2\Delta} - M_3e^{-j(W_1\Delta - \Delta W(L-1)/2} \right] T_4$$

$$H(W_a) = \frac{e^{-jW_a\Delta} \; e^{-jm_2\Delta W(L-1)/2}}{(M_1M_2 - M_3^2)} \Bigg\{ M_2T_1e^{jW_1\Delta}$$

$$- M_3T_1e^{j(\Delta W(L-1)/2+W_2\Delta)} + M_1T_2e^{j\Delta W_2 + W(L-1)/2}$$

$$- M_3T_2e^{j(W_1\Delta)} + M_2T_3e^{j\Delta W_1(L-1-\Delta)}$$

$$- M_3T_3e^{-j(\Delta W(L-1)/2+W_2\Delta-W_1(L-1))}$$

$$+ M_1T_4e^{j(W_2(L-1)+\Delta W(L-1)/2-W_2\Delta)}$$

$$- M_3T_4e^{j(W_2(L-1)+\Delta W(L-1)/2 \; - \; W_1 \; +\Delta W)L-1)/2} \Bigg\}$$

$$H(W_a) = \frac{e^{-j(W_a\Delta+m_2\Delta W(L-1)/2)}}{(M_1M_2 - M_3^2)} \Bigg\{ M_2T_1e^{j\theta_1} + M_3T_1e^{j(\theta_2+\theta_3)}$$

$$+ M_1T_2e^{j|\theta_2+\theta_3|} - M_3T_2e^{j\theta_1} + M_2T_3e^{-j(\theta_4+\theta_1)}$$

$$- M_3T_3e^{-j(\theta_3+\theta_2-\theta_4)} + M_1T_4e^{j(\theta_5+\theta_3-\theta_2)}$$

$$- M_3T_4e^{j(\theta_5+\theta_3-\theta_1)} \Bigg\}$$

where

$$\theta_1 = W_1\Delta$$
$$\theta_2 = W_2\Delta$$
$$\theta_3 = \Delta W(L-1)/2$$
$$\theta_4 = W_1(L-1)$$
$$\theta_5 = W_2(L-1)$$

$$H(W_a) = \frac{e^{-j(W_a\Delta + m_2\Delta W(L-1)/2)}}{M_1M_2 - M_3^2}\Big\{(M_2T_1 - M_3T_2)e^{j\theta_1}$$

$$+ (M_1T_2 - M_3T_1)e^{j(\theta_2+\theta_3)} + M_2T_3e^{j(\theta_4-\theta_1)}$$

$$- M_3T_3e^{j(\theta_4-\theta_2-\theta_3)} + (M_1T_4e^{-j\theta_2} - M_3T_4e^{j(\theta_3-\theta_1)})e^{j(\theta_5+\theta_3}$$

Note that; since

$$\theta_4-\theta_2+\theta_3 = W_1(L-1) - W_2\Delta - W_1\frac{(L-1)}{2} + \frac{W_2}{2}(L-1)$$

$$= \frac{(W_1+W_2)}{2}(L-1) - W_2\Delta$$

and

$$\theta_5+\theta_3-\theta_2 = W_2(L-1) + \frac{W_1}{2}(L-1) - \frac{W_2}{2}(L-1) - W_2\Delta$$

$$= (\frac{W_1+W_2}{2})(L-1) - W_2\Delta$$

therefore

$$\theta_5+\theta_3-\theta_2 = \theta_4-\theta_2-\theta_3$$

and we can rewrite $H(W_a)$;

$$H(W_a) = \frac{e^{-j(W_a\Delta+m_2\Delta W(L-1)/2)}}{(M_1M_2 - M_3^2)}\Big\{(M_2T_1-M_3T_2)e^{j\theta_1}$$

$$+ (M_1T_2-M_3T_1)e^{j(\theta_2+\theta_3)} + M_2T_3e^{j(\theta_4-\theta_1)}$$

$$+ (M_1T_4-M_3T_3)e^{j(\theta_4-\theta_2-\theta_3)} - M_3T_4e^{j(\theta_5+2\theta_3-\theta_1)}\Big\}$$

Similarly,

$$\theta_4-\theta_1 = \theta_5+2\theta_3-\theta_1$$

$$H(W_a) = \frac{e^{-j(W_a\Delta+m_2\Delta W(L-1)/2)}}{(M_1M_2 - M_3^2)}\Big\{(M_2T_1-M_3T_2)e^{j\theta_1}$$

$$+ (M_1T_2-M_3T_1)e^{j(\theta_2+\theta_3)} + (M_2T_3-M_3T_4)e^{j(\theta_4-\theta_1)}$$

$$+ (M_1T_4-M_3T_3)e^{j(\theta_4-\theta_2-\theta_3)}\Big\}$$

$$H(W_a) = \frac{e^{-j\alpha}}{R}\left[\rho_1 e^{j\theta_1} + \rho_2 e^{j(\theta_2+\theta_3)} + \rho_3 e^{j(\theta_4-\theta_1)} + \rho_4 e^{j(\theta_4-\theta_2-\theta_3)}\right]$$

where

$$\alpha = W_a\Delta + m_2\Delta W(L-1)/2$$
$$R = M_1M_2 - M_3^2$$
$$\rho_1 = M_2T_1 - M_3T_2$$
$$\rho_2 = M_1T_2 - M_3T_1$$
$$\rho_3 = M_2T_3 - M_3T_4$$
$$\rho_4 = M_1T_4 - M_3T_3$$

$$|H(W_a)|^2 = \rho_1^2 + \rho_2^2 + \rho_3^2 + \rho_4^2 + 2[\rho_1\rho_2 \operatorname{Cos}(\theta_1-\theta_2+\theta_3) + \rho_3\rho_4 \operatorname{Cos}(\theta_1-\theta_4-\theta_3-\theta_2+\theta_4) + \rho_1\rho_3 \operatorname{Cos}(\theta_1+\theta_1-\theta_4) + \rho_1\rho_4 \operatorname{Cos}(\theta_1+\theta_3+\theta_2-\theta_4) + \rho_2\rho_4 \operatorname{Cos}(\theta_2+\theta_3+\theta_3+\theta_2-\theta_4) + \rho_2\rho_3 \operatorname{Cos}(\theta_2+\theta_3+\theta_1-\theta_4)]$$

By neglecting the terms T_3 and T_4 we found the value of Δ as follows

$$\Delta = \frac{(2k+1)}{\Delta W} - \frac{(L-1)}{2}$$

APPENDIX A2

For two sinusoidal signals, the transfer function of ALE can be given as follows

$$H(w) = \sum_{k=0}^{L-1} [A_1 e^{jW_1 k} + A_2 e^{jW_2 k} + A_3 e^{-jW_1 k}$$

$$+ A_4 e^{-jW_2 k}] \, e^{-jW(k+\Delta)} \qquad (A2.1)$$

By neglecting the contribution of negative frequency components we can approximate (A2.1) as

$$H(w) = \sum_{k=0}^{L-1} [A_1 e^{jW_1 k} + A_2 e^{jW_2 k}] \, e^{-jW(k+\Delta)} \qquad (A2.2)$$

The error which is caused by ALE consists of three components. The first component is due to white noise spreading, the second and third components are the attenuation of the first and second sinusoidal signals. This situation can be formulated by (A2.3).

$$V = \frac{\delta_o^2}{\pi} \int_o^{\pi} |1-H(w)|^2 dw + |1-H(w_1)|^2 \delta_1^2 + |1-H(w_2)|^2 \delta_2^2 \qquad (A2.3)$$

where δ_o^2 is the power of the white noise. δ_1^2 and δ_1^2 are the power of the first and second sinusoidal signals respectively.

Our aim is tomminimize V with respect to Δ. At this point we make an approximation again. Assume that $\delta_o^2 << \delta_1^2$ and $\delta_o^2 << \delta_2^2$. Therefore by neglecting the first term of (A2.3) we find

$$\frac{dv}{d\Delta} = 0 = R_e[1 - H(W_1)] \frac{d}{d\Delta} R_e[1 - H(W_1)]$$

$$+ I_m[1 - H(W_1)] \frac{d}{d\Delta} I_m[1 - H(W_1)]$$

$$+ R_e[1 - H(W_2)] \frac{d}{d\Delta} R_e[1 - H(W_2)]$$

$$+ I_m[1 - H(W_2)] \frac{d}{d\Delta} I_m[1 - H(W_2)] \qquad \text{(A2.4)}$$

The real and imaginary parts of the transfer function can be given by (A2.5)

$$R_e[H(W_1)] = C_1 + C_2 \text{ Cos } [\Delta W[(L-1)/2 + \Delta]]$$
$$I_m[H(W_1)] = - C_2 \text{ Sin } \Delta W([L-1]/2 + \Delta)$$
$$R_e[H(W_2)] = C_3 + C_4 \text{ Cos } \Delta W[(L-1)/2 + \Delta]$$
$$I_m[H(W_2)] = C_4 \text{ Sin } \Delta W[(L-1)/2 + \Delta] \qquad \text{(A2.5)}$$

where

$$C_1 = \frac{L(L + 2\delta_o^2/\delta_2^2) - \left[\frac{\text{Sin } \Delta W \ L/2}{\text{Sin } \Delta W/2}\right]^2}{[L + 2\delta_o^2/\delta_2^2][L + 2\delta_o^2/\delta_1^2] - \left[\frac{\text{Sin}\Delta WL/2}{\text{Sin}\Delta W/2}\right]^2}$$

$$C_2 = \frac{\frac{\text{Sin}\Delta WL/2}{\text{Sin}\Delta W/2}(L + 2\delta_o^2/\delta_1^2) - \frac{\text{Sin}\Delta WL/2}{\text{Sin}\Delta W/2}}{(L + 2\delta_o^2/\delta_2^2)(L + 2\delta_o^2/\delta_1^2) - \left[\frac{\text{Sin}\Delta WL/2}{\text{Sin}\Delta W/2}\right]^2}$$

$$C_3 = \frac{L(L + 2\delta_o^2/\delta_1^2) - \left[\frac{\text{Sin}\Delta WL/2}{\text{Sin}\Delta W/2}\right]^2}{(L + 2\delta_o^2/\delta_1^2)(L + 2\delta_o^2/\delta_2^2) - \left[\frac{\text{Sin}\Delta WL/2}{\text{Sin}\Delta W/2}\right]^2}$$

$$C_4 = \frac{\frac{\text{Sin}\Delta WL/2}{\text{Sin}\Delta W/2}[2\delta_o^2/\delta_2^2]}{(L + 2\delta_o^2/\delta_2^2)(L + 2\delta_o^2/\delta_1^2) - \left[\frac{\text{Sin}\Delta WL/2}{\text{Sin}\Delta W/2}\right]^2}$$

From (A2.4) we have

$$\frac{dv}{d\Delta} = 0 = -[C_1C_2 + C_3C_4] \text{ Sin } \Delta W|(L-1)/2 + \Delta| \qquad \text{(A2.6)}$$

Therefore, the optimum value of Δ can be given

$$\Delta + (L-1)/2 = (2k+1)\frac{\pi}{\Delta W} \qquad \text{(A2.7)}$$

(A2.7 is identical with (A1.17).

APPENDIX A3

The optimal value of Δ for the case of multiple sinusoidal average error variance for N sinusoidal signals can be written as

$$V = \frac{\delta_o^2}{2\pi} \int_o^{2\pi} |1 - H(w)|^2 dw + \sum_{i=1}^{N} |1 - H(w_i)|^2 \delta_i^2 \qquad (A3.1)$$

In a similar manner as in Appendix A2, we try to minimize V with respect to V by neglecting the first term in (A3.1). The transfer function of ALE for N sinusoidal signals case is as follows

$$H(w) = \sum_{k=0}^{L-1} a_k Z^{-(\Delta+k)} \qquad (A3.2)$$

where

$$a_k = \sum_{n=1}^{2N} A_n e^{jW_n k}$$

From the formula which is related to the A_n and γ_{rn} we can find H(w). But in here we want to make an approximation by assuming L is very large. Then $\gamma_{rn} \rightarrow 0$ for all n and r, the A_n uncouple and are given to a good approximation by [7]

$$A_n = \frac{e^{jW_n\Delta}}{L + 2\delta_o^2/\delta_n^2} \qquad n=1,2,...,2N \qquad (A3.3)$$

Equation (A3.3) is identical to the expression for the amplitude of the mean steady state ALE impluse response for one sinusoid at W_n in white noise. Therefore the frequency response of the steady state ALE which will be denoted by $H^*(w)$ can be simply expressed in the term of the A_n:

$$H^*(w) = \sum_{k=0}^{L-1} a_k \, e^{-jW(k+\Delta)}$$

$$= \sum_{n=1}^{2N} A_n \, e^{-jW\Delta} \, \frac{1 - e^{j(W_n-W)L}}{1 - e^{j(W_n-W)}} \qquad (A3.4)$$

As L becomes large, so that (A3.3) is valid $H^*(w)$ is given to a good approximation by

$$H^*(w) = \sum_{n=1}^{N} \frac{e^{-j(W_n+W)\Delta}}{L + 2\delta_o^2/\delta_n^2} \, \frac{1 - e^{-j(W_n+W)L}}{1 - e^{-j(W_n+W)}}$$

$$+ \sum_{n=1}^{N} \frac{e^{j(W_n-W)\Delta}}{L + 2\delta_o^2/\delta_n^2} \, \frac{1 - e^{j(W_n-W)L}}{1 - e^{j(W_n-W)}} \qquad (A3.5)$$

Equation (A3.5) corresponds to a sum of bandpass filters (centered at $\mp W_n$) each having a peak value given by

$$(L/2)\ SNR_n/((L/2)SNR_n + 1) \qquad (A3.6)$$

where $SNR_n = \delta_n^2/\delta_o^2$. As $L \to \infty$ all of the peak values in (17) approach 1, and the ALE becomes a linear superposition of perfectly resolved bandpass filters, each with unity gain at its frequency. Caution must be exercised in choosing L, however, because as L is increased, the weight vector noise is also increased. Therefore, in practice, a value of for L which provides a trade-off between weight vector noise and enhancement abilities should be chosen as in Capter 4.

Again by returning to (A3.5) we have

$$H^*(w) = \sum_{n=1}^{N} \frac{e^{-j(W_n+W)[\Delta+L-1/2]}}{L + 2\delta_o^2/\delta_n^2} \frac{\sin(W_n+W)/2}{\sin(W_n+W)/2}$$

$$+ \sum_{n=1}^{N} \frac{e^{j(W_n-W)[\Delta+L-1/2]}}{L + 2\delta_o^2/\delta_n^2} \frac{\sin(W_n-W)L/2}{\sin(W_n-W)/2} \quad \text{(A3.7)}$$

From (A3.7) we can find the real and imaginary part of H(w) respectively, as follows:

$$R_e\{H(w)\} = \sum_{n=1}^{N} \frac{\sin(W_n+W)L/2}{\sin(W_n+W)/2} \frac{\cos(W_n+W)[\Delta+\frac{L-1}{2}]}{L + 2\delta_o^2/\delta_n^2}$$

$$+ \sum_{n=1}^{N} \frac{\sin(W_n-W)L/2}{\sin(W_n-W)/2} \frac{\cos(W_n-W)(\Delta+\frac{L-1}{2}}{L + 2\delta_o^2/\delta_n^2} \quad \text{(A3.8)}$$

$$I_m\{H(w)\}\Gamma = \sum_{n=1}^{N} - \frac{\mathrm{Sin}(W_n+W)\left[\Delta + \frac{L-1}{2}\right]}{L + 2\delta_o^2/\delta_n^2} \frac{\mathrm{Sin}(W_n+W)L/2}{\mathrm{Sin}(W_n+W)/2}$$

$$+ \sum_{n=1}^{N} \frac{\mathrm{Sin}(W_n-W)\left[\Delta + \frac{L-1}{2}\right]}{L + 2\delta_o^2/\delta_n^2} \frac{\mathrm{Sin}(W_n-W)L/2}{\mathrm{Sin}(W_n-W)/2} \tag{A3.9}$$

Now our problem is a simply traditional minimization of (A3.1) with respect to Δ and it is given in general (i.e., not neglecting the first term in (A3.1)) by (A3.10)

$$\frac{dv}{d\Delta} = 0 = \frac{\delta_o^2}{2} \int_o^{2\pi} R_e\left[1 - \bar{H}(w)\right] \frac{d}{d\Delta}\left[1 - H(w)\right]dw$$

$$+ 2 \sum_{i=1}^{N} R_e\{[1 - \bar{H}(w_i)] \frac{d}{d\Delta} [1 - H(w_i)]\} \tag{A3.10}$$

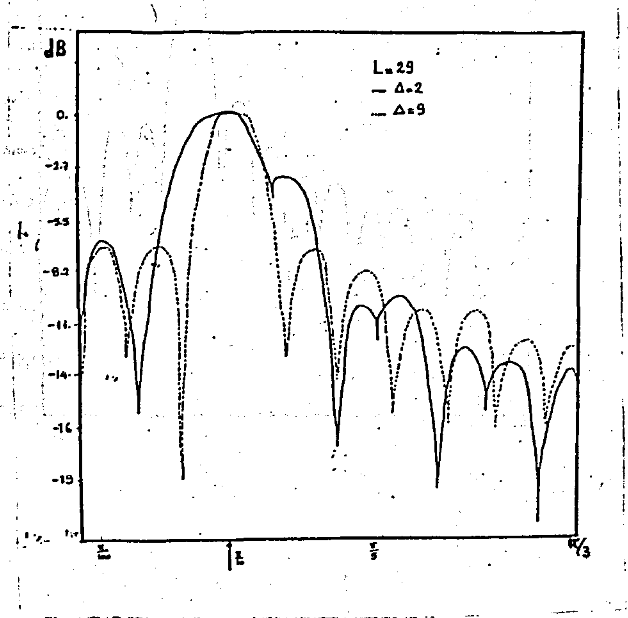
and by neglecting the first term (A3.11)

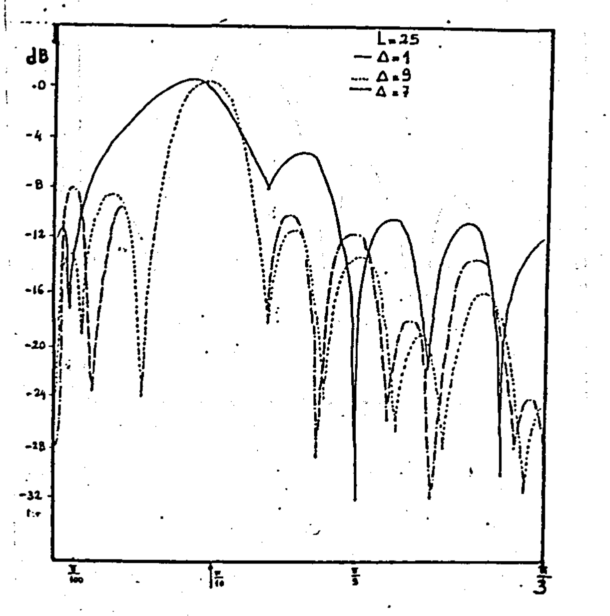
$$\frac{dv}{d\Delta} = 0 = \sum_{n=1}^{N} R_e\left[1 - H(w_i)\right] \frac{d}{d\Delta} \left[1 - H(w_i)\right]$$

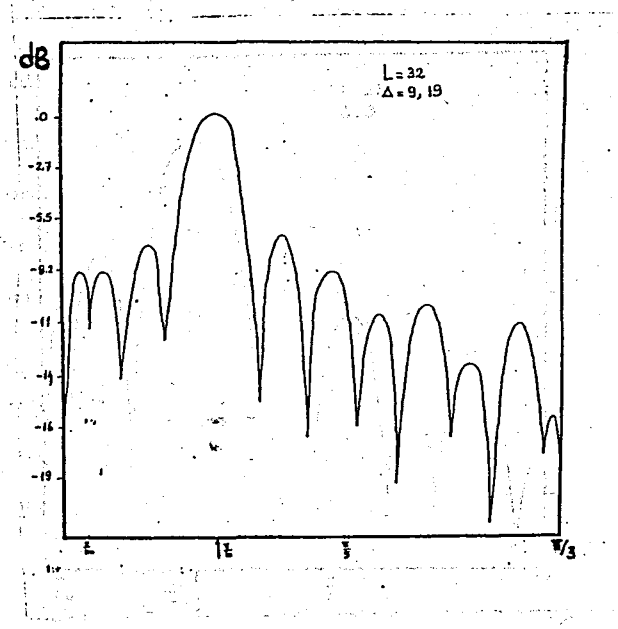
$$+ I_m[1 - H(w_i)] \cdot \frac{d}{d\Delta} I_m[1 - H(w_i)] \tag{A3.11}$$

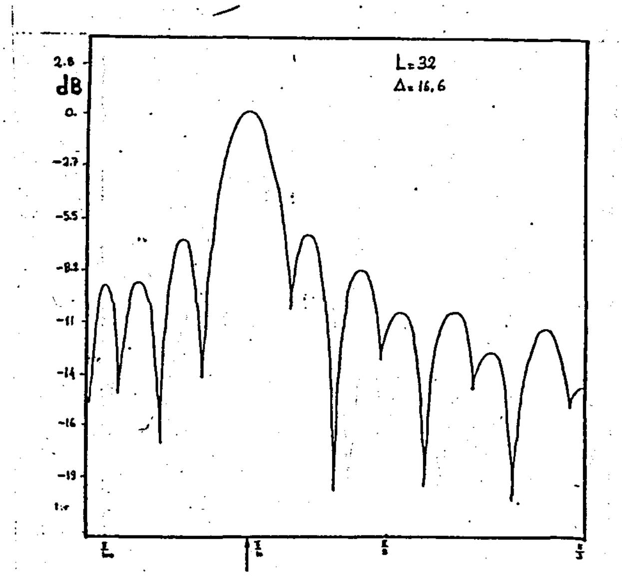
and from (A3.11) we find the stationary point of it.

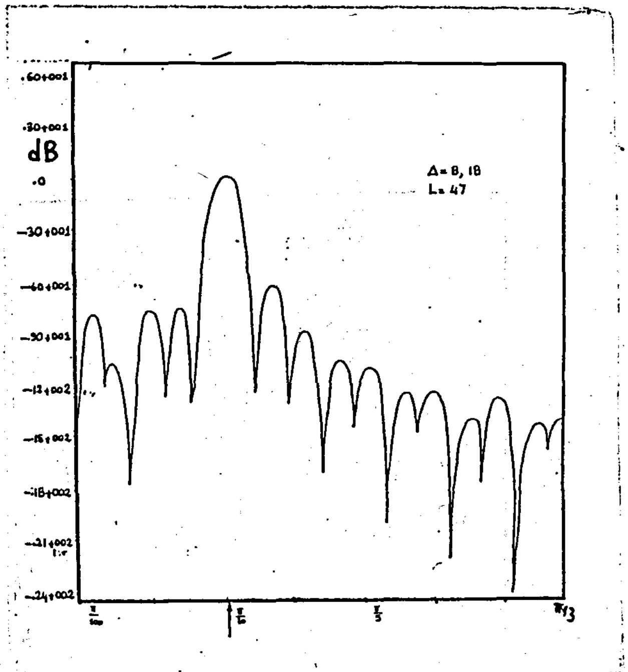
APPENDIX A4

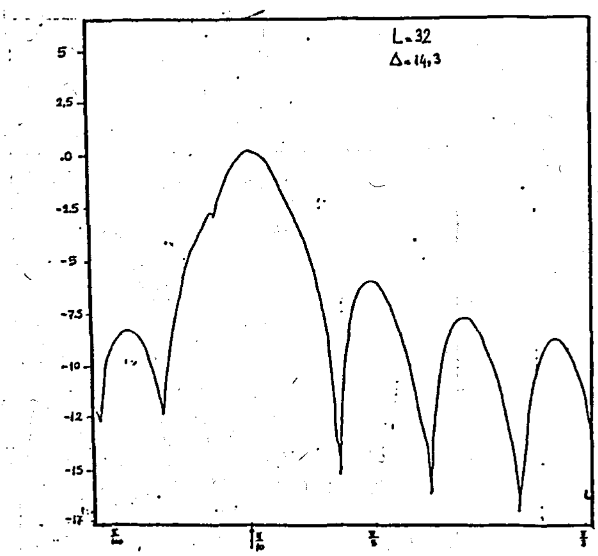


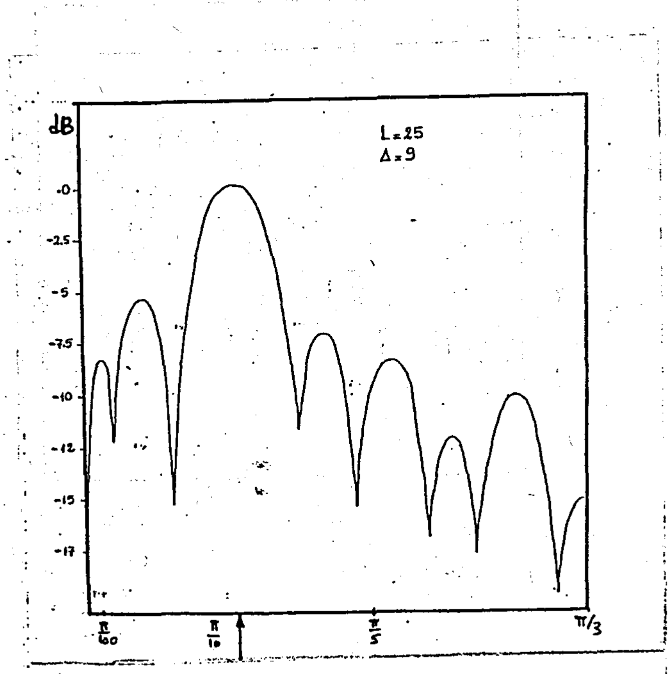


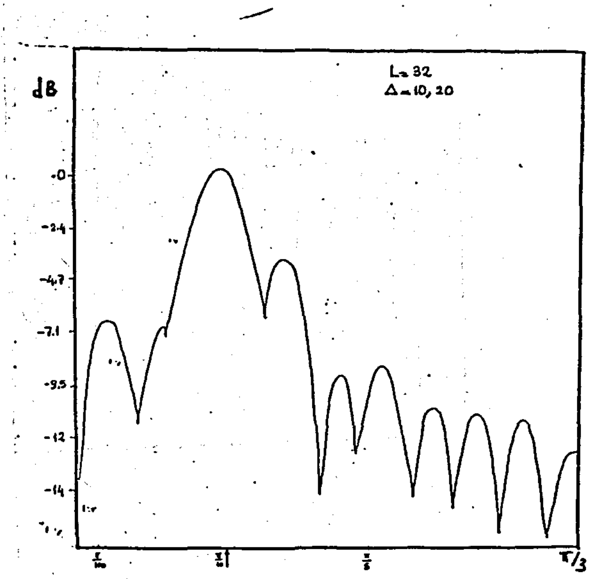


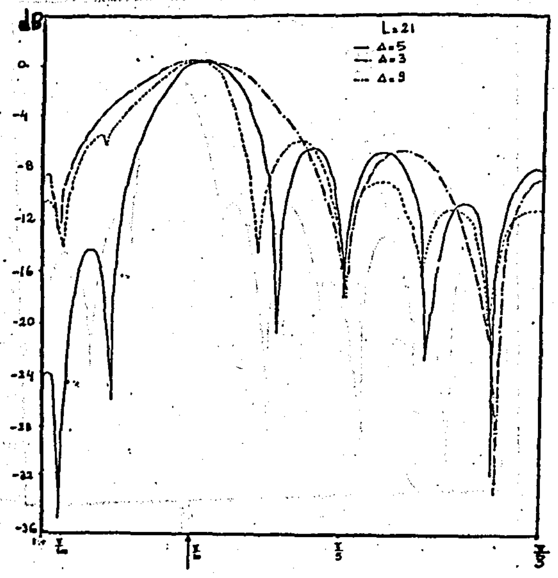


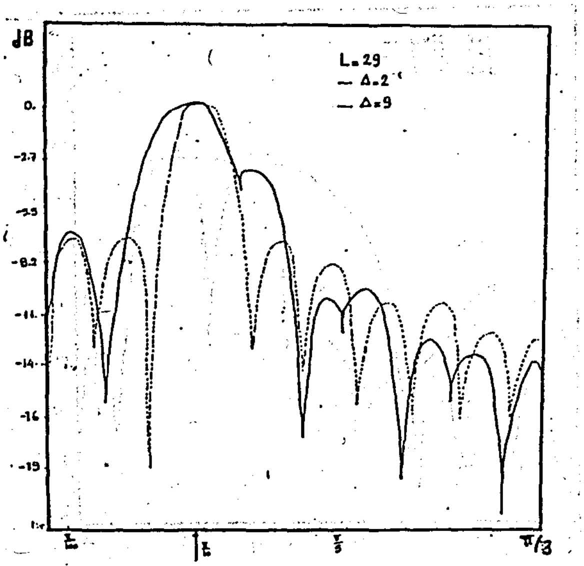

dB
Δ = 8, 18
L = 47
π/3

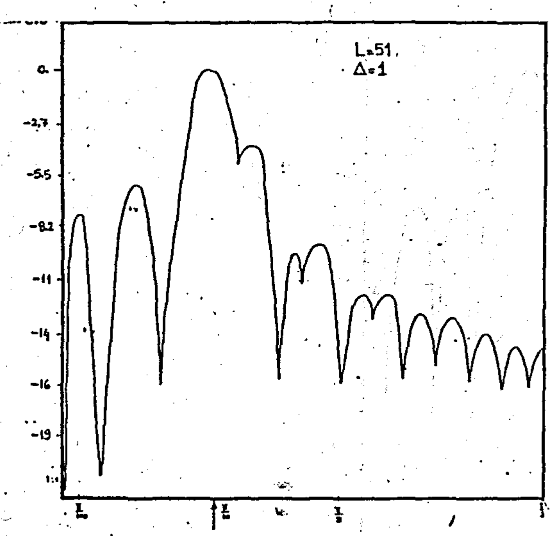

L = 32
Δ = 14,3
5
2,5
-10
-12
-15

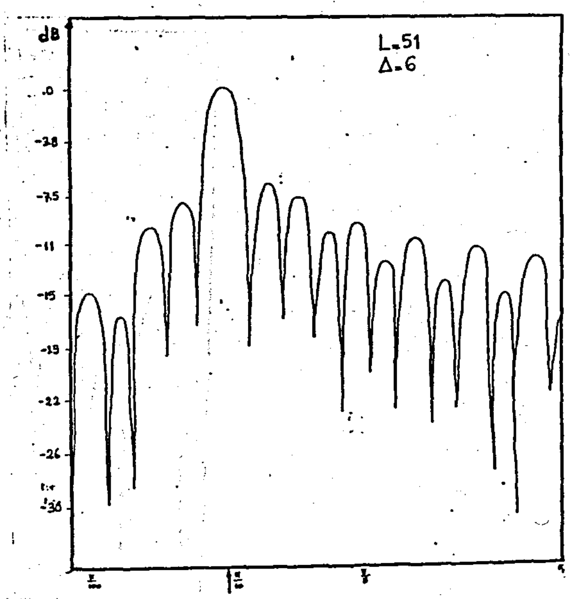

dB
L = 25
Δ = 9
-2.5
-5
-7.5
-10
-12
-15
-17
π/100
π/10
π/5
π/3

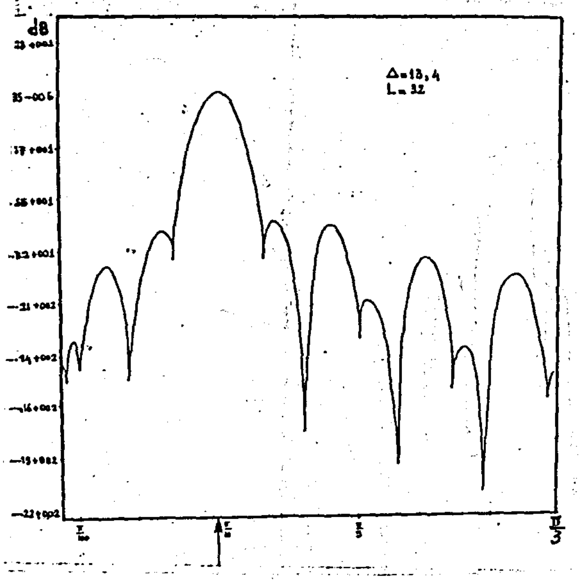

dB
L = 32
Δ = 10, 20
-2.4
-4.7
-7.1
-9.5
-12
-14
π/3

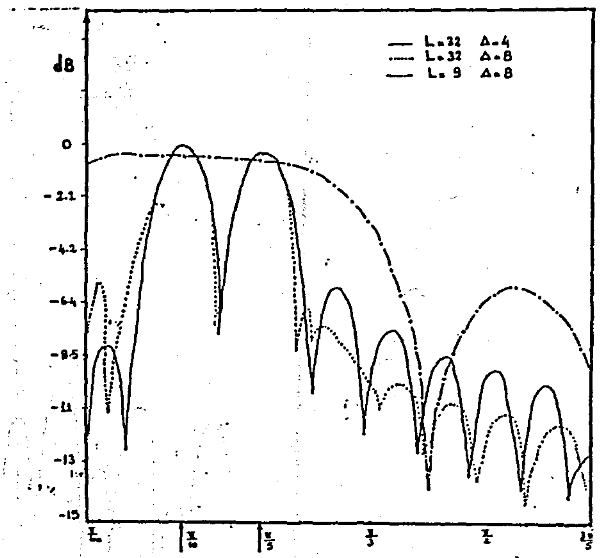

dB
L=21
Δ=5
Δ=3
Δ=9
0
-4
-8
-12
-16
-20
-24
-28
-32
-36

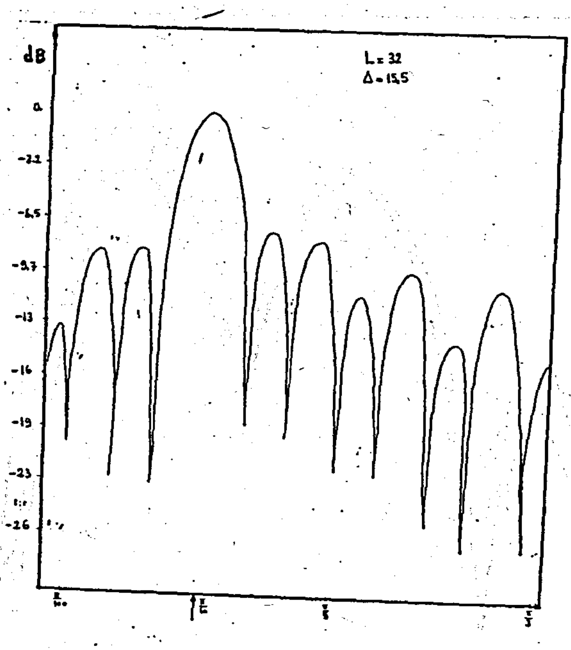

dB
L=29
Δ=2
Δ=9
0.
-2.7
-5.5
-8.2
-11
-14
-16
-19
π/3

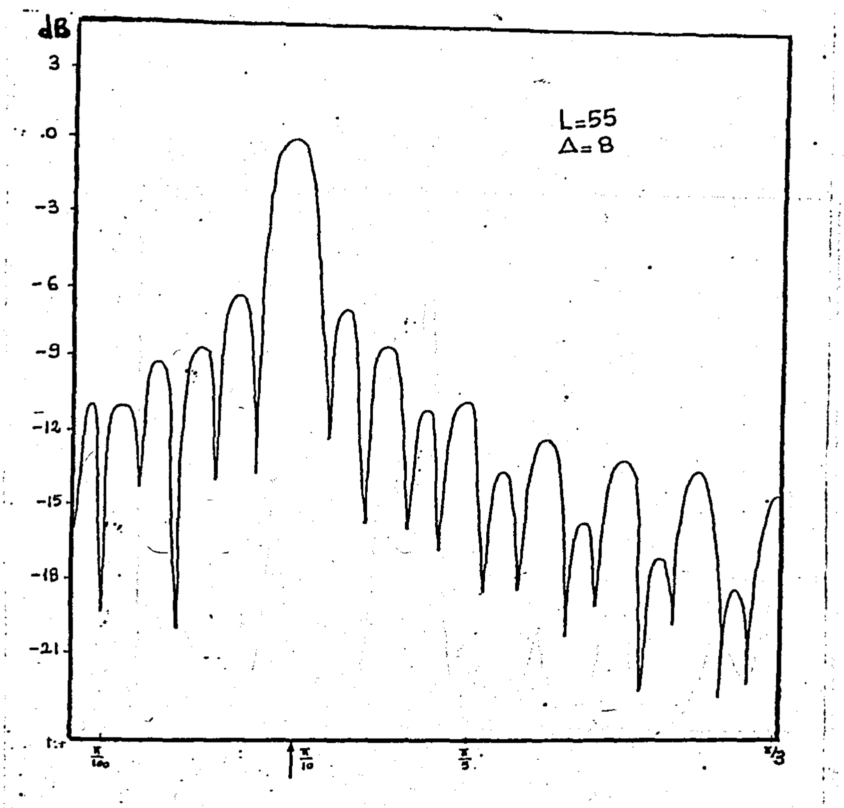

L=51
Δ=1
0.
-2,7
-5.5
-8.2
-11
-14
-16
-19

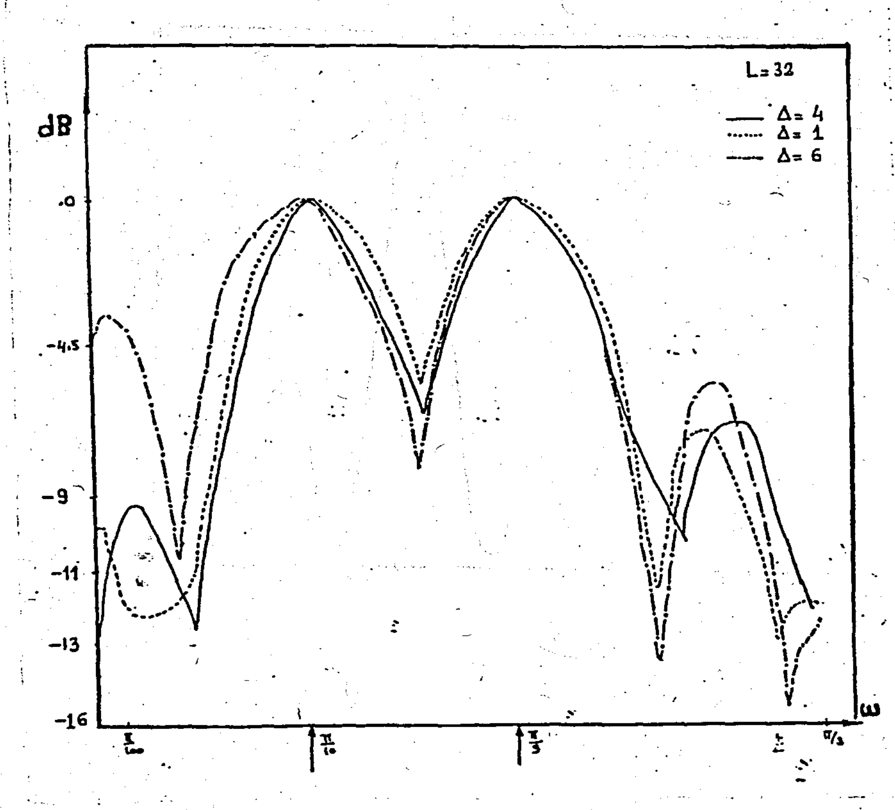

dB
L=51
Δ=6
0
-3.8
-7.5
-11
-15
-19
-22
-26
-30

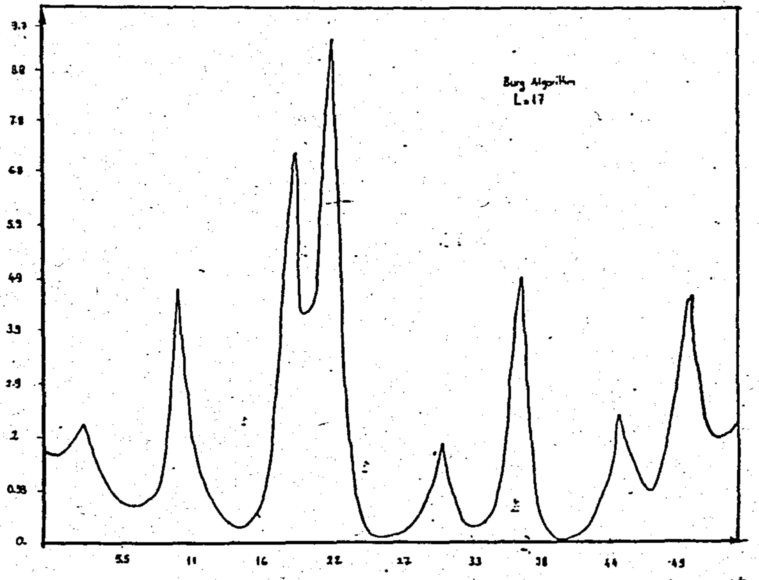

dB
Δ=13,4
L=32

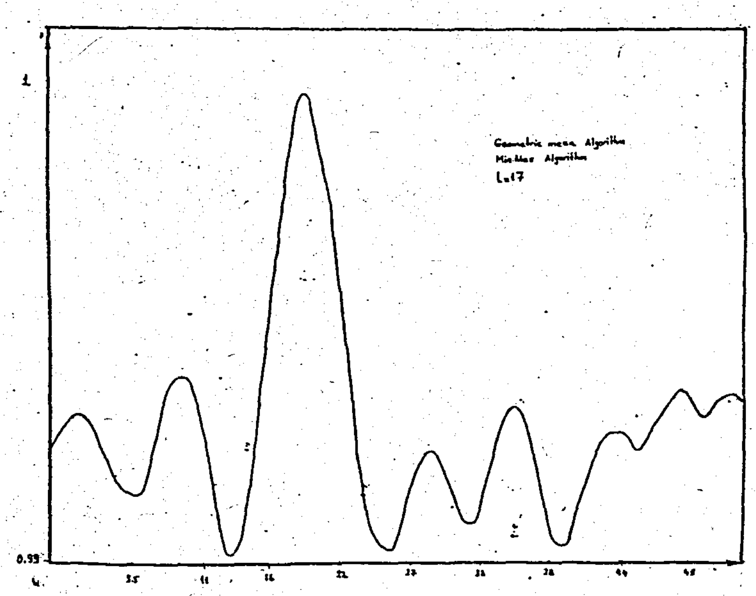

dB
L=32 Δ=4
L=32 Δ=8
L=9 Δ=8
0
-2.1
-4.2
-6.4
-8.5
-11
-13
-15


dB
L=32
Δ=15,5
0
-3.2
-6.5
-9.7
-13
-16
-19
-23
-26


dB
3
.0
-3
-6
-9
-12
-15
-18
-21
L=55
Δ= 8
π/100
π/10
π/5
π/3


dB
.0
-4.5
-9
-11
-13
-16
L=32
Δ= 4
Δ= 1
Δ= 6
π/100
π/10
π/5
π/3
ω


Burg Algorithm
L=17
5.5
11
16
22
33
0.
2


Geometric mean Algorithm
Min-Max Algorithm
L=17
1
5.5
11
16