PARAMETRIC MODELING OF LINEAR AND NONLINEAR SYSTEMS

by

Francis Anthony Perry

June 1980

Thesis Advisor: S. R. Parker

Approved for public release; distribution unlimited
Parametric Modeling of Linear and Nonlinear Systems

Francis Anthony Perry

Naval Postgraduate School
Monterey, California 93940

Approved for public release; distribution unlimited


The problem of obtaining parametric models for linear and nonlinear systems based on observations of the input and output of the system is one of wide ranging interest. For linear systems, moving average (MA) and autoregressive (AR) models have received considerable attention and based on the Levinson algorithm, a number of very powerful methods involving lattice filter structures have been developed to obtain the model solutions. For nonlinear systems the Volterra series model which is a nonlinear extension of the moving...
average model is frequently used.

The purpose of this research is to extend these techniques to more general linear and nonlinear models. Using the equation error formulation, lattice solution methods in batch processing and adaptive form are developed for both single and multichannel autoregressive moving average (ARMA) models for linear systems and Volterra series models for nonlinear systems. A nonlinear extension of the ARMA model is also considered and is shown in some cases to remedy problems encountered in Volterra modeling of nonlinear systems. Lattice methods are also developed for the nonlinear ARMA model and it is shown that the methods obtained for linear ARMA modeling follow as a special case of the nonlinear results.

Experimental verification of the methods developed for single channel linear ARMA modeling is presented and used to explore the characteristics of the lattice solution techniques. The results clearly indicate that the lattice methods are extremely powerful, capable of producing highly accurate system models using short runs of data.
ABSTRACT

The problem of obtaining parametric models for linear and nonlinear systems based on observations of the input and output of the system is one of wide ranging interest. For linear systems, moving average (MA) and autoregressive (AR) models have received considerable attention and, based on the Levinson algorithm, a number of very powerful methods involving lattice filter structures have been developed to obtain the model solutions. For nonlinear systems the Volterra series model which is a nonlinear extension of the moving average model is frequently used.

The purpose of this research is to extend these techniques to more general linear and nonlinear models. Using the equation error formulation, lattice solution methods in batch processing and adaptive form are developed for both single and multichannel autoregressive moving average (ARMA) models for linear systems and Volterra series models for nonlinear systems. A nonlinear extension of the ARMA model is also considered and is shown in some cases to remedy problems encountered in Volterra modeling of nonlinear systems. Lattice methods are also developed for the nonlinear ARMA model and it is shown that the methods obtained for linear ARMA modeling follow as a special case of the nonlinear results.

Experimental verification of the methods developed for single channel linear ARMA modeling is presented and used
to explore the characteristics of the lattice solution techniques. The results clearly indicate that the lattice methods are extremely powerful, capable of producing highly accurate system models using short runs of data.
# TABLE OF CONTENTS

## I. INTRODUCTION
- The Problem Statement ................................................................. 10
- Overview .......................................................................................... 16

## II. DISCRETE TIME LINEAR SYSTEM MODELING; BACKGROUND THEORY
- Moving Average Models ........................................................................ 22
- Autoregressive Models .......................................................................... 27
- Recursive in Order Solutions for AR and MA Models ............................ 33
  1. The Levinson Algorithm for AR Modeling ........................................... 34
  2. The Levinson Algorithm for MA Modeling ........................................... 36
- Lattice Form AR and MA Models ............................................................ 38
  1. The AR Modeling Lattice Structures .................................................. 38
  2. The MA Modeling Lattice Structures .................................................. 46
- Multichannel AR and MA Modeling ...................................................... 48

## III. ARMA MODELING
- Linear ARMA Modeling and Its Relation to AR and MA Modeling ......... 72
  1. Model Transition Relationships ......................................................... 76
  2. Modeling Input Signal Requirements ................................................. 82
- A Recursive in Order Solution for the ARMA Model ............................ 86
- Lattice Form ARMA Modeling .............................................................. 97
- ARMA Model Simulations; Batch Processing ....................................... 101
- Adaptive Lattice ARMA Modeling ....................................................... 134
- A Lattice Approach for Multichannel ARMA Modeling ....................... 147

## IV. NONLINEAR SYSTEM MODELING .................................................. 152
A. VOLTERRA NONLINEAR MODELING 152

1. Lattice Filter Methods for Volterra Modeling 160

B. NONLINEAR ARMA MODELING 165

1. Identifiability Conditions and Memory Requirements 171

2. Lattice Filter Methods for the Nonlinear ARMA Model 180

V. APPLICATIONS, CONCLUSIONS AND OPEN QUESTIONS 188

A. APPLICATIONS 188

1. Reduced Order Modeling 188

2. Modeling for Fault Detection 196

B. CONCLUSIONS AND OPEN QUESTIONS 198

APPENDIX A - Alternate Multichannel Model Forms 203

APPENDIX B - A Key Relationship for the Multichannel Levinson Algorithm 219

APPENDIX C - The multichannel Levinson Algorithm Solution for AR Modeling 222

APPENDIX D - The Multichannel Levinson Algorithm Solution for MA Modeling 227

APPENDIX E - Prony's Method for ARMA Modeling 230

APPENDIX F - Parabolic Surfaces in n-Dimensions 235

APPENDIX G - Multichannel ARMA Modeling 238

APPENDIX H - Delay Free Loops 241

APPENDIX I - Nonlinear ARMA Modeling Examples 243

I.A - Cascade of Linear and Nonlinear Subsystems 243

I.B - A Nonlinear ARMA Model for a Phase Locked Loop 246

APPENDIX J - Model Simulation Program Listings 250

LIST OF REFERENCES 276

INITIAL DISTRIBUTION LIST 282
ACKNOWLEDGEMENT

It is with deepest appreciation that I wish to acknowledge my advisor, Professor Sydney R. Parker for his countless hours of patient assistance and guidance over the course of this research. I would like to thank the members of my doctoral committee who have followed this effort, contributing their time and valuable suggestions. I also wish to acknowledge the contributions of Commander Stacey Holmes and Doctor Joel Morris of the Office of Naval Research, and Mr. Nate Butler of the Naval Electronic Systems Command for their support of this research. Finally, I wish to thank my wife, Beth, and daughter, Karen, for their forebearance and encouragement.
I. INTRODUCTION

Traditionally, man has attempted to create models of portions of his environment for two principal reasons:

1. To gain insight and understanding as to their functioning;
2. As a prelude to taking some action such as attempting to exercise control over them.

The field of physics for instance, is replete with examples where men have created models to study and explain phenomenon from planetary motion to the motion and even origin of subatomic particles. In designing machines, engineers routinely rely on models of the components they use to describe how they will function, and to obtain the desired results in the final product. Economics is another field where the use of models abounds for such purposes as identifying, forecasting or trying to direct trends.

The scope of the modeling problem is quite broad beginning with a decision on the type of model to be used, what physical quantities to measure, how to estimate the parameters of the model from the measurement, and finally a verification of the model. In the chapters that follow, one facet of this problem, that of estimating model parameters, will be explored in detail for a number of linear and non-linear models.
A. THE PROBLEM STATEMENT

The primary concern of this work is the determination of discrete time models for both linear and nonlinear, time invariant systems from sampled observations of the system inputs and outputs. The general approach underlying all of the models examined here assumes that the system to be modeled is described by the equation

\[ y(k) = F_{10}[u(k)] + F_{20}[y(k-1)] + F_{30}[u(k), y(k-1)] \]  \hspace{1cm} (1.1)

where \( F_{10}, F_{20} \) and \( F_{30} \) are functions of past and present values of their arguments, and \( u(k) \) and \( y(k) \) are the system input and output respectively. This is depicted in Figure (1.1). A possible method for modeling this type of system is to create a model of exactly the same configuration with functions \( F_{10}, F_{20} \) and \( F_{30} \), assume a form for these functions, operate the system and model in parallel with the same input and adjust the parameters of the model functions to minimize the mean square error (MSE) between the model output \( \hat{y}(k) \) and the system output. The symbol "\(^\wedge\)" is used here to indicate that the signal is an estimate of \( y(k) \). This is depicted in Figure 1.2 and is often referred to as direct form modeling since the assumed topology of the system is directly copied.

---

1 Script notation will be used to refer to quantities associated with the system while nonscript notation will be used for their corresponding approximants in the model.
Figure 1.1. The assumed form for systems to be modeled. $T$ represents a unit delay.

Figure 1.2. A direct approach to system modeling.

by the model. Here the model output is given by

$$\hat{y}(k) = F_{10}[u(k)] + F_{20}[\hat{y}(k-1)] + F_{30}[u(k), \hat{y}(k-1)]$$  \hspace{1cm} (1.2)
and the error signal is referred to as the "output error". As an example, if the system is linear an appropriate model is

$$F_{10}[u(k)] = \sum_{i=0}^{N} a(i)u(k-i) \quad (1.3a)$$

$$F_{20}[\hat{y}(k-1)] = \sum_{i=1}^{N} b(i) \hat{y}(k-i) \quad (1.3b)$$

$$F_{30}[u(k),\hat{y}(k-1)] = 0 \quad (1.3c)$$

(for linear models, $F_{30}[\cdot]$ will be zero). In general however, the direct form approach will have serious difficulties if either $F_{20}[\cdot]$ or $F_{30}[\cdot]$ are nonzero since the past values of $\hat{y}(k)$ used in these functions are themselves dependent on the model parameters. A minimum mean square output error approach results in a system of highly nonlinear simultaneous equations which must be solved to obtain the model parameters.

To avoid these difficulties, the equation error approach [Refs. 34 and 23] to system modeling (which uses different model forms in the analysis and synthesis phases) will be applied to the problem. The analysis model is depicted in Figure 1.3 and differs from the direct form model in that $F_{20}$ and $F_{30}$ are functions of past and present values of the delayed system output rather than the analysis model output.
For each of the models studied, a general form for the three functions is assumed and the parameters of the model (coefficients of the functions) are set to obtain a MMSE solution. In each case, the MSE cost function is a quadratic function of the model parameters (due to both the equation error formulation and to the form chosen for the functions) with a unique minimum and therefore the solution is given by a system of linear equations. The synthesis model is of the same form assumed for the system in Figure 1.1 and uses the functions $F_{10}$, $F_{20}$ and $F_{30}$ determined during the analysis phase.
As an alternative to the topology shown in Figure 1.3 it will occasionally be convenient to consider the error signal $e(k)$ as the output of the analysis model rather than the prediction $\hat{y}(k)$. This can be accomplished in one of two ways by defining

$$ F_2[y(k)] = y(k) - P_{20}[y(k-1)] $$ (1.4a)

or

$$ F_3[u(k),y(k)] = y(k) - P_{30}[u(k),y(k-1)] $$ (1.4b)

and reformulating the analysis model as shown in Figures 1.4a or 1.4b. These model forms are often referred to as prediction error models since their outputs are the errors in predicting $y(k)$ rather than the predictions themselves. There are, however, no substantive differences between the modeling approaches depicted in Figure 1.3, 1.4a and 1.4b.

The equation error formulation can be generalized to multiple input multiple output systems as well (henceforth referred to as multichannel systems) by considering $u(k)$ and $y(k)$ as vectors of $Q_1$ input signals and $Q_0$ output signals and $F_{10}$, $F_{20}$, $F_{30}$, $F_2$ and $F_3$ as vector functions. The prediction error signal $e(k)$ becomes a $Q_0$-vector of signals and the model parameters can be set to minimize the trace of the prediction error covariance matrix $P = \varepsilon\{e(k)e(k)^T\}$. Such generalizations have been developed to a degree in the multichannel filtering literature and will be investigated further here.
Figure 1.4. Prediction error model forms.
It is important to keep in mind however, that while the equation error formulation can be used to find a model solution, it is an indirect method as opposed to the direct form method which minimizes the mean square value of output error. The direct form model has been modified to obtain the equation error analysis model so that the parameters can be obtained via the solution of systems of linear equations. The price paid for this simplification in the model analysis problem is that additive noise on the measured system output will introduce a bias in the model coefficient estimates.

B. OVERVIEW

Chapter II along with appendices A through F provide a unified review of the existing background theory on minimum mean square equation error modeling of linear systems. The moving average (MA) and autoregressive (AR) models are presented and their relative merits compared. In Section II.C, the Levinson algorithm [Refs. 9, 10 and 27] for the AR and MA models is developed, greatly simplifying the solution process for these models. Section II.D, then shows that the Levinson algorithm defines the AR and MA models in terms of lattice filter structures.

These lattice structures have received widespread attention and have led to a host of new developments in modeling, spectral estimation, filter structures and adaptive filtering. Examination of the properties of these forms have suggested a number of new methods for calculating model coefficients
that offer increased accuracy, and in some cases guarantee model stability even in the presence of correlation estimates obtained by averaging over short time intervals [Refs. 5, 20, 29 and 36]. Applied by Burg [Ref. 5] to spectral estimation, these methods allow the determination of power spectra via AR modeling from very short runs of data without any need for the use of a window function. In finite precision arithmetic implementations, the lattice structures have been shown by Markel and Gray [Ref. 33] to be less sensitive to roundoff noise and coefficient quantization than direct structures and have led to the design of other structures that offer improved performance over conventional parallel realizations. Griffiths has shown that these lattices can be implemented adaptively [Refs. 16, 17 and 18] and that they offer the potential for more rapid convergence than conventional LMS adaptive filters. Recently Morf [Refs. 36, 37 and 38] has also used these lattice structures to implement a recursively updated deterministic least squares adaptive scheme. It is readily apparent therefore, that the original work of Levinson and the lattice structures that have evolved from it have had an important impact on the field of digital signal processing.

In Section II.E., the multichannel generalization of many of the single channel AR and MA modeling results is presented. After a discussion of the basic multichannel AR and MA models [Refs. 26 and 45], the multichannel version of the Levinson algorithm originally developed by Whittle [Ref. 56],
and Wiggins and Robinson [Ref. 61] is presented. A new
form for the models is introduced and used here however,
to facilitate the application of these results later in
various other modeling problems. Multichannel lattice
structures are then derived and from them alternative
solution methods for the modeling problems are developed.

Finally, in Section II.F. the LMS adaptive algorithm
[Ref. 58] is reviewed and the adaptive implementations of
the lattice structures due to Griffiths are presented.

In Chapter III, the more general autoregressive moving
average model is presented using the equation error formu-
lation attributed to Kalman [Ref. 23]. After a brief
discussion of the model, new model transition formulas
are developed showing how the ARMA model is related to the
simpler and less general AR and MA models. System input
signal requirements for the ARMA modeling process are ex-
plored and it is shown that the power spectrum of the input
signal can be considered as a frequency dependent weighting
function in the model optimization. Then the main result
of the chapter is presented. With suitable assumptions,
a recursive in order solution method for ARMA modeling
(the \((n+1)\)st order solution is obtained from the \(n\)-th
order solution) is obtained based on the Levinson algorithm for
multichannel AR models. From this, lattice solution methods
for the ARMA model are developed in both batch processing
and adaptive form. (Batch processing here refers to assuming
ergodicity and estimating correlations with time averaging).
A similar result has recently been presented by Morf [Refs. 37 and 38] with the assumption of a white noise input signal to the system. The results presented here follow from a different approach without the assumption of a white noise input. Experimental results are also presented verifying the methods and theory, and showing their advantages (and disadvantages) over conventional ARMA modeling methods.

The programs used in these simulations are listed in Appendix J. In Section III.F., and Appendix G, it is shown that these single channel methods readily extend to the multi-channel ARMA model, and as one would expect, can be obtained as a special case.

In Chapter IV two types of nonlinear models, the Volterra series model and the new nonlinear ARMA model recently proposed by Parker [Ref. 64], are considered. After a brief discussion of the Volterra model, it is shown that the solution can be obtained using multichannel MA lattice methods if the regular form of the Volterra kernels is used in place of the conventional symmetric form. Then the nonlinear ARMA model is presented in Section IV.B. and it is shown that for many systems, this model can remedy the problem of the large number of terms (ideally infinite) required by the Volterra model to represent the system in much the same way that the ARMA model solved the problem arising in the MA model. In Section IV.B.2 it is also shown that by using the regular form, the solution for the nonlinear ARMA model can be obtained using multichannel AR lattice methods and that the linear ARMA model solutions developed in Chapter III follow
as a special case. Appendix I then presents two examples of nonlinear ARMA modeling. First a somewhat academic example of a cascade of linear and nonlinear subsystems is given then a nonlinear ARMA model is proposed for the tracking behavior of a phase locked loop.

Finally, in Chapter V, two applications for the linear and nonlinear ARMA modeling methods developed in Chapters III and IV are discussed briefly. (They are reduced order modeling of complex systems and modeling for fault detection and diagnosis.) Then in Section V.B. conclusions are drawn on the results of this work and a list of significant open questions (both old unanswered questions and new ones raised here) is compiled.
II. DISCRETE TIME LINEAR SYSTEM MODELING: BACKGROUND THEORY

While few physical systems are absolutely linear, linear models often suffice to accurately describe their behavior under normal operating conditions. A rich body of theory has therefore been developed for the analysis and modeling of linear systems [Ref. 22] and a thorough knowledge of this theory is vitally important to anyone interested in understanding the functioning of these systems. The continuing expansion in the availability of powerful, inexpensive digital computing capabilities has also made discrete time techniques take on a special prominence. With this as motivation, the portion of the background theory in discrete time linear modeling upon which much of the remainder of this work depends, is developed here from the unifying standpoint of a minimum mean square equation error model solution.

The moving average and the autoregressive models are developed first for single input single output systems. Their solution via the Levinson algorithm is presented and from this algorithm alternate solution methods based on lattice filter structures are derived. It is shown that almost all of these results can be generalized to the multiple input multiple output case and the corresponding multichannel modeling methods are developed. Finally, adaptive implementation of the modeling methods for both
the conventional filter structures and the lattice filter structures is presented as an alternative means of obtaining model solutions.

A. MOVING AVERAGE MODELS

The moving average (MA) model was among the earliest discrete models developed. [Refs. 4, 11 and 19] It estimates the current value of the output of a system as a weighted combination of the present value and \( N \) past values of the system input where \( N \) is the order of the model. The problem then is to estimate the weighting function or impulse response of the MA model in some fashion. Since the MA model characterizes a system in terms of a finite duration approximation of its impulse response and since any linear time invariant, single input single output system is completely specified by its impulse response, the MA model is quite general and can be used for a wide class of systems. Defining \( (N+1) \)-vectors of model weights and input data as

\[
\mathbf{a}^+ = [a(0) \ldots a(N)]^T \quad (2.1a)
\]

1 A superscript "+" is used to indicate that in spite of the fact that these vectors are used for a \( N \)-th order model, they are \((N+1)\)-vectors with elements indexed from zero to \( N \) rather than from one to \( N \). Superscript \( T \) denotes transpose.

2 Superscripts in parenthesis will later be added to the model coefficient vectors to explicitly indicate their dependence on the order of the problem being solved. They are omitted for simplicity however whenever doing so does not result in ambiguous notation.
the MA estimate of the system output becomes

\[ \hat{y}(k) = \sum_{n=0}^{N} a(n)u(k-n) = u^+(k) a^+ \]

In terms of the modeling approach of Figure 1.3, \( F_{20} \) and \( F_{30} \) are assumed to be zero. \( F_1 \) is a linear time invariant function of past and present values of \( u(k) \). Assuming stationarity, an expression for the mean square value of the error as a quadratic function of the weights \( \{a(n)\} \) is given by

\[ E_2 = a^+_u R_{uu} + a^+_y R_{uy} + 2 a^+_u R_{uy} + R_{uu}(0) \]

where in general \( R_{vw}(n) = \varepsilon\{v(k)w(k+n)\} \), \( R_{vw} = \varepsilon\{v(k)w(k+n)\} \), \( R_{vw} = \varepsilon\{v(k)w(k)^T\} \) and \( \varepsilon\{ \} \) denotes expectation.

\[ R_{uu} = \begin{bmatrix} R_{uu}(0) & \cdots & R_{uu}(N) \\ \vdots & \ddots & \vdots \\ \vdots & & \vdots \\ R_{uu}(N) & \cdots & R_{uu}(0) \end{bmatrix} \]

\[ R_{uy} = \begin{bmatrix} R_{uy}(0) & \cdots & R_{uy}(N) \end{bmatrix}^T \]

The surface described by equation 2.2 can be pictured as a concave hyperparaboloid with a unique minimum and the
characteristics of such a surface are described in Appendix F. For example when \( N=1 \), the MSE as a function of \( a(0) \) and \( a(1) \) appears as shown in Figure 2.1.

\[ E^2 \]

\[ a(0) \]

\[ a(1) \]

Figure 2.1. MSE as a function of model weights for a first order (\( N=1 \)) MA model.

The minimum mean square error solution for the coefficients is given by

\[ \frac{R}{u^+ u} + a^+ \text{OPT} = \frac{R}{u^+ y} \]  \hspace{1cm} (2.3)
and the corresponding minimum value of the cost function $E_2$ is

$$E_{2\text{min}} = R_{yy}(0) - a_{0\text{OPT}}^T E_{u+y}$$  \hspace{1cm} (2.4)$$

Equation (2.3) is a discrete time matrix form of the Wiener Hopf equation

$$R_{uy}(\tau) = \int_{-\infty}^{\infty} R_{uu}(\tau-\lambda) h(\lambda) \, d\lambda$$  \hspace{1cm} (2.5)$$

where $u(\tau)$ and $y(\tau)$ are the continuous input and output signals and $h(\tau)$ is the system impulse response. The process of finding $a_{0\text{OPT}}^+$ in equation (2.3) is the discrete time equivalent of deconvolving the input autocorrelation function from the cross correlation of input and output to obtain the system impulse response in equation (2.5).

Consequently the MA modeling process has been called discrete Wiener filtering or stochastic deconvolution.

This model constitutes a direct form approach as defined in section 1.1 but does not encounter difficulty in obtaining the model weights since both $F_{20}$ and $F_{30}$ are assumed to be zero. As such, it possesses the advantage that the estimates of the model parameters will not be biased by the presence of additive noise on the output of the system as shown in Figure 2.2, as long as the noise is uncorrelated with the input signal. This can readily be seen by replacing $v$ in equation 2.3 by $y + v$. Additive noise on the input signal
however will adversely affect the modeling process, and introduce a bias in the solution for the model coefficients.

In the transform domain, the model can be represented by a polynomial in powers of $z^{-1}$ and has therefore been referred to as an **all zero model**

$$A(z) = \sum_{n=0}^{N} a(n) z^{-n} \quad (2.6)$$

In terms of this transfer function relationship, any bias introduced in one or more of the model coefficients has the effect of shifting the zero locations of the model.

In summary a discussion of the advantages and disadvantages of MA modeling is instructive.

**Advantages:**

1) The solution for the model parameters involves only linear equations.
2) The solution is unbiased in the presence of additive noise on the system output as long as the noise and system input are uncorrelated.

3) Since the model is nonrecursive it is always stable.

Disadvantages:

1) The number of terms (N+1) needed for sufficient model accuracy may be quite large.

2) The solution of a large system of equations is required.

3) The required correlation terms are usually not known and must be estimated by assuming ergodicity and averaging in time. This requires the data to be windowed and set to zero outside the averaging interval in order to maintain the even symmetry of the autocorrelation functions.

4) The modeling process is restricted to linear time invariant systems.

B. AUTOREGRESSIVE MODELS

The autoregressive (AR) model attempts to deal with one of the difficulties (1) encountered in MA modeling; the need for a large number of coefficients to accurately describe the model. [Refs. 2, 4, 11, 19 and 28] In AR modeling, which is sometimes referred to as linear prediction, a prediction error approach is considered where

\[ e(k) = y(k) - \sum_{n=1}^{N} b(n) y(k-n) \]  

(2.7a)
This can be written as

\[ e(k) = y(k) - \hat{y}(k) \]

\[ = y(k) - y(k)^T b \]  \hspace{1cm} (2.7b)

with

\[ y(k) = [y(k-1) \cdots y(k-N)]^T \]  \hspace{1cm} (2.7c)

and

\[ b = [b(1) \cdots b(N)]^T \]  \hspace{1cm} (2.7d)

Here \( F_{10} \) and \( F_{30} \) are assumed to be zero (this assumption will be modified later to allow a dependence on the input signal in the synthesis phase) and \( F_{20} \) provides an estimate of the current value of the system output as a weighted sum of \( N \) past outputs. The mean square value of prediction error as a quadratic function of the weights \( \{b(n)\} \) is given by

\[ E_2 = b^T R_{yy} b - 2 b^T R_{yy} + R_{yy}(0) \]  \hspace{1cm} (2.8a)

and the corresponding MMSE solution for the weights is given by

\[ R_{yy} b_{OPT} = E_{yy} \]  \hspace{1cm} (2.8b)
Using equation 2.7a, an expression can be written in the transform domain for the prediction error model which accepts $y(k)$ as its input and produces the error sequence $e(k)$ as its output.

$$\frac{E(z)}{Y(z)} = 1 - \sum_{n=1}^{N} b(n)z^{-n} = B(z)$$

If it is assumed that the system input output relationship can be represented in transfer function form with

$$\frac{Y(z)}{U(z)} = H(z) = \frac{a(0)}{1 - \sum_{n=1}^{N} b(n) z^{-n}} = \frac{a(0)}{B(z)}$$

and that the model parameters can be determined so that $B(z) = B(z)$, then the prediction error output will be exactly $e(k) = a(0) u(k)$. For this reason AR prediction error modeling has often been called inverse filtering since the prediction error filter essentially reverses the actions of the system (with the exception of a gain). Since the analysis model is in this inverse form rather than in the direct form, the presence of additive noise on the measurement of the system output as shown in Figure 2.3 will introduce a bias in the solution for the model parameters. This
Figure 2.3. Autoregressive prediction error modeling as an inverse filtering process.

has the effect of shifting the roots of $B(z)$ which are estimates of the poles of the system and is the price paid for the ability to obtain the model solution from a set of linear equations.

Thus far, only the analysis portion of the AR modeling process has been discussed. With the inverse filtering interpretation of the prediction error analysis model, a reasonable synthesis model is given in transfer function form as

$$H(z) = \frac{a(0)}{B(z)}$$  \hspace{1cm} (2.11)

with the gain term set so that the mean square value of $a(0) u(k)$ is the same as that of the prediction error signal. Thus it follows that

$$a(0)^2 = \frac{\mathbb{E}\{e(k)^2\}}{R_{uu}(0)}$$  \hspace{1cm} (2.12)
Since the synthesis model is in the form of an all pole filter, an appropriate impulse response with infinite duration might be obtained using a low order model (small N), a result that is impossible to obtain in any finite order MA model. This is not to say however that a low or even finite order AR model will always be an appropriate model for any linear system. If the transfer function representation for a system contains both zeros and poles, no finite order AR or MA model can serve to exactly represent it. This fact can be understood by considering the form of a geometric series

$$\frac{1}{1-Cz^{-1}} = \sum_{n=0}^{\infty} (Cz^{-1})^n \quad \text{for } |Cz^{-1}| < 1 \quad (2.13)$$

which shows that a single pole can be represented by an infinite number of zeros and vice versa. Thus if the system has a single zero, a high order AR model may be required to represent it with sufficient accuracy.

In summary, the advantages and disadvantages of AR modeling may be listed as follows:

Advantages:

1) The solution for the model parameters involves only linear equations.

2) Sometimes an appropriate infinite impulse response can be obtained with a small number of parameters in the model.

3) Direct knowledge or measurement of the system input is not required for determining the system poles. Only a knowledge of its mean square...
value is necessary for determining the gain factor.

Disadvantages:

1) The model is biased by the presence of additive noise on the measured system output signal.

2) The number of terms required for sufficient model accuracy may be quite large if zeros are present in the system. If this occurs, the inversion of a large matrix will be required.

3) The required correlation terms are usually not known and must be estimated by assuming ergodicity and using time averages.

4) The modeling process is restricted to linear time invariant systems.

This list of advantages and disadvantages is quite similar to the one compiled for MA models with two notable differences; the bias in the model and the absence of a requirement for input measurements. This second point is significant in that it has led to the application of AR modeling to many problems where an input signal is unmeasurable or indeed does not exist including speech modeling and spectral estimation. [Refs. 2, 5, 12, 15, 21, 32 and 44] The noise problem has restricted the process to applications where measurements with sufficiently high signal to noise ratio are available, making the effects of the bias minimal. [Refs. 24 and 43]
C. RECURSIVE IN ORDER SOLUTIONS FOR AR AND MA MODELS

The preceding discussions of the AR and MA modeling problems tacitly assumed an apriori knowledge of the correct model order. If this knowledge is not available a reasonable approach for determining the correct model order must be developed [Ref. 53]. A commonly employed strategy is to successively increment the model order while observing the MSE until further increases fail to substantially reduce the MSE. This requires solving for a number of different models and can be an arduous task if equations (2.8b) or (2.3) are employed directly.

The autocorrelation matrices appearing in the AR and MA model equations (2.8b) and (2.3) are highly structured matrices (both Toeplitz and symmetric) and this fact can be exploited to facilitate the solution of these equations. The Levinson algorithm [Refs. 9, 10 and 27] makes use of this structure to obtain model solutions recursively in order, that is, the solution for the n-th order model is assumed to be known and the solution for the (n+1)-st order model is then obtained from it. In this manner it is possible to start with a first order AR or a zeroth order MA solution given by a single equation and build up the desired order solution. The AR model will be treated first since it is a special case that simplifies the analysis. The simplifications arise due to the fact that the r_{yy} vector on the right hand side of equation (2.8b) is made up primarily of terms also appearing on the left hand side in R_{yy}.
Superscripts in parenthesis are used to explicitly indicate the order of the problem when specifically needed.

1. The Levinson Algorithm For AR Modeling

The $n$-th order AR model solution of equation 2.8b is given by

$$R_{yy}^{(n)} b^{(n)} = r_{yy}^{(n)} \quad (2.14)$$

The Levinson algorithm assumes a relationship between the $n$-th and $(n+1)$-st order solutions given by

$$b^{(n+1)} = \begin{bmatrix} b^{(n)} \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon^{(n)} \\ \vdots \\ k^{(n+1)} \end{bmatrix} \quad (2.15)$$

and solves for the vector $\varepsilon^{(n)}$ and the coefficient $k^{(n+1)}$. Define permuted versions of the vectors $b^{(n)}$ and $r_{yy}^{(n)}$ as $f^{(n)}$ and $p_{yy}^{(n)}$ by reversing the order of their elements.

$$f^{(n)} = \begin{bmatrix} b(n)(n) \\ \vdots \\ b(1)(n) \end{bmatrix} \quad p_{yy}^{(n)} = \begin{bmatrix} r_{yy}(n) \\ \vdots \\ r_{yy}(1) \end{bmatrix} \quad (2.16)$$

Because of the Toeplitz symmetric structure of the auto-correlation matrix, equation (2.14) can also be written as

$$R_{yy}^{(n)} f^{(n)} = p_{yy}^{(n)} \quad (2.17)$$
and this relationship is essential in the development of the Levinson algorithm. (To apply the algorithm therefore when time averaged estimates of the needed correlations are used, the data must be windowed prior to averaging to maintain the even symmetry in the autocorrelation function estimates produce the required structure in the autocorrelation matrix.) Making use of equation (2.15), in the (n+1)-st order version of equation (2.14), and using the relationship of (2.17) to solve for \( e(n) \) and \( k(n+1) \) results in

\[
\begin{align*}
\hat{e}(n) &= -\hat{f}(n)k(n+1) \\
\end{align*}
\]  

(2.18a)

and

\[
\begin{align*}
k(n+1) &= \frac{R_{yy}(n+1) - \hat{\rho}(n)\hat{b}(n)}{R_{yy}(0) - \hat{\rho}(n)\hat{b}(n)\hat{r}_{yy}} \\
\end{align*}
\]  

(2.18b)

Therefore, in using equations 2.15 and 2.18 to obtain \( \hat{b}^{(n+1)} \) from \( \hat{b}^{(n)} \) via the Levinson algorithm only one new piece of information, \( k^{(n+1)} \), need be calculated. The denominator of equation (2.18b) can also be recognized from equation (2.8c) as the MMSE for the n-th order AR model \( E^{(n)}_2 \), and thus there is little concern over the possibility of it being zero. If the n-th order solution produces a perfect prediction (zero MSE) there is no point in trying to find a better prediction by increasing the order to n+1. The evaluation of equation (2.18b) can be further simplified by observing that the MSE also follows a recursion from one order to the next.
given by

\[ E_2^{(n+1)} = E_2^{(n)} [1 - k^{(n+1)}]^2 \]  \hspace{1cm} (2.19)

making it unnecessary to evaluate the denominator at each value of \( n \). (Details of this derivation are omitted here but included in Appendix A in the derivation of the more general multichannel Levinson algorithm.) This relation for the propagation of mean square prediction error also leads to the restriction that \( k^{(n+1)} \) must be bounded in magnitude by unity.

2. The Levinson Algorithm For MA Modeling

Next consider the n-th order MA model given by

\[ R_{uu}^{(n)} + a_+^{(n)} = R_{uy}^{(n)} \]  \hspace{1cm} (2.20)

and again, assume a relationship between the \((n+1)\)-st order and n-th order solutions given by

\[ a_+^{(n+1)} = \begin{bmatrix} a_+^{(n)} \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} y^{(n+1)} \\ \vdots \\ g^{(n+1)} \end{bmatrix} \]  \hspace{1cm} (2.21)

Notice that in this n-th order problem, \( R_{uu}^{(n)} \) is actually a \( n+1 \) by \( n+1 \) matrix and could be written as \( R_{uu}^{(n+1)} \). Define
Using equations (2.21) and (2.22) in the $n+1$ order MA model equation it follows that

$$\gamma^{(n+1)} = -f^{(n+1)}g^{(n+1)}$$ \hspace{1cm} (2.23a)

where $f^{(n+1)}$ is defined in a manner similar to (2.16) and is comprised of the coefficients that arise in an $(n+1)$-st order autoregression on the input signal $u(k)$. Therefore to obtain a moving average model relating the system output signal $y(k)$ to the input signal $u(k)$, an autoregressive model for the system input must first be solved. Furthermore,

$$g^{(n+1)} = \frac{R_{uy}^{(n+1)} - \rho_{uu}^{(n+1)}T}{R_{uu}(0) - \rho_{uu}^{(n+1)}Tf^{(n+1)}}$$ \hspace{1cm} (2.23b)

and the denominator of equation (2.23b) is the MMSE in the $(n+1)$-st order autoregression on the signal $u(k)$.

It is significant to note that in applying the Levinson algorithm to find a given order AR or MA model, all lower order models along with their MSE's are obtained. Also, intermediate quantities emerge (the $k^{(n)}$) in the AR
model and the \( \{ k(n) \} \) and \( \{ g(n) \} \) in the MA model) which fully characterize the models and could be used as an alternative to the \( \{ a(n) \} \) or \( \{ b(n) \} \) coefficients. This point will be developed further in subsequent sections.

D. LATTICE FORM AR AND MA MODELS

The Levinson algorithm derived in the previous section can be used to derive lattice structures to implement the MA model and the AR analysis and synthesis models as alternatives to a tapped delay line type of implementation using the coefficients \( a(n) \) or \( b(n) \) directly. [Refs. 29, 30, 32 and 33]

1. The AR Modeling Lattice Structures

From the relationship between the \( (n+1) \)-st and \( n \)-th order solutions to the AR modeling problem determined in equations (2.15) and (2.18a) it follows that the transfer function of the prediction error model can be written recursively in order as

\[
B(n+1)(z) = B(n)(z) - k(n+1)z^{-n-1}B(n)(z^{-1}) \quad (2.24)
\]

Defining a new transfer function

\[
B(n)(z) = z^{-n}B(n)(z^{-1}) \quad (2.25a)
\]

equation (2.24) can be written as

\[
B(n+1)(z) = B(n)(z) - k(n+1)z^{-1}B(n)(z) \quad (2.25b)
\]
and an expression can also be written for $B^{(n+1)}(z)$ recursively in order by rewriting equation (2.25a) for order $(n+1)$ and substituting equation (2.24) yielding,

$$B^{(n+1)}(z) = z^{-1}B^{(n)}(z) - k^{(n+1)}B^{(n)}(z) \quad (2.25c)$$

As discussed earlier in connection with equation (2.9), $B^{(n)}(z)$ describes the $n$-th order prediction error model and when its input is the system output $Y(z)$, it produces the $n$-th order prediction error signal.

$$E^{(n)}(z) = B^{(n)}(z) Y(z) \quad (2.26)$$

In the time domain this signal can be interpreted as the error in predicting $y(k)$ forward in time from a weighted combination of the $n$ past values \{y(k-1)\cdots y(k-n)\}.

To understand the significance of $E^{(n)}(z)$ consider the output signal when this model is excited by $Y(z)$.

$$E^{(n)}(z) = B^{(n)}(z) Y(z)$$

$$= z^{-n}[1 - \sum_{i=1}^{n} B^{(n)}(i)z^i]Y(z) \quad (2.27)$$

In the time domain, $E^{(n)}(k)$ can be interpreted as the error in predicting $y(k-n)$ backward in time from a weighted combination of the future signals \{y(k-n+1)\cdots y(k)\}. These $n$-th order forward and backward prediction processes at time
\( k_o \) are illustrated in Figure 2.4. Henceforth, an overbar will always be used to denote quantities associated with backward in time predictions.

\[ y(k) \uparrow \]

\begin{align*}
\text{Forward Prediction} \\
\text{Backward Prediction}
\end{align*}

\[ k_0 - n \quad k_0 \quad k \]

\text{Figure 2.4. Forward and Backward Prediction Error Filtering.}

From equations (2.25b) and (2.25c) equations can be written recursively in order for these forward and backward prediction error sequences as:

\[ e^{(n+1)}(k) = e^{(n)}(k) - k^{(n+1)}e^{(n)}(k-1) \quad (2.28a) \]

\[ \bar{e}^{(n+1)}(k) = \bar{e}^{(n)}(k-1) - k^{(n+1)}e^{(n)}(k) \quad (2.28b) \]
Noting that the prediction error for a zero-order AR predictor of \( y(k) \) (or no predictor at all) is just the signal \( y(k) \) itself,

\[
e^{(0)}(k) = e^{(0)}(k) = y(k) \tag{2.28c}
\]

the prediction error filter can be drawn in lattice form as shown in Figure 2.5 for a second order model.

![Figure 2.5. Lattice Form Of A Second Order Prediction Error Model.](image)

This structure has many interesting properties, among the most important of which is the successive decoupling property. In going from one order AR model to the next, all of the previously determined transfer function coefficients \( \{b(n)\} \) will generally change. The Levinson algorithm shows however that only one new piece of information is needed to obtain the optimum \((n+1)\)-st order solution from the optimum \(n\)-th order solution (see equation (2.24)). In terms of the lattice filter of Figure 2.4 this
means that given the optimum n-th order model in lattice form, one need only add another stage to the structure, setting the coefficient of that stage \( k^{(n+1)} \) to minimize the mean square value of \( e^{(n+1)}(k) \). Nothing in the first \( n \) stages need be changed. The overall high order minimization problem is in this fashion decomposed into a sequence of first order minimizations, one at each lattice stage.

Another important property of the lattice which can be proven and will be of use later is the orthonormalization of the backward prediction error sequence [Ref. 32] which states that

\[
\epsilon\{e^{(i)}(k)e^{(j)}(k)\} = \begin{cases} 
0 & i \neq j \\
\sum_i^\infty & i = j 
\end{cases} \quad (2.29)
\]

Thus it is seen that a set of orthogonal signals (the backward prediction errors at the various stages) are generated as a by-product of the lattice model.

As a consequence of the successive decoupling property of the lattice, a number of alternatives to equation (2.18b) for determining the lattice coefficients can be found. The most obvious method is to set \( k^{(n+1)} \) to explicitly minimize the mean square value of forward prediction error in equation (2.28a) at the \((n+1)\)-st order stage given the best lattice of order \( n \). This is termed the forward method and is denoted by a subscript \( F \) on the
lattice coefficients. The resulting solution is given by

\[ k_F^{(n+1)} = \frac{\mathbb{E}\{e^{(n)}(k)e^{(n)}(k-1)\}}{\mathbb{E}\{e^{(n)}(k)^2\}} \]  

(2.30)

Alternately, the mean square value of the backward prediction error signal in equation (2.28b) could be minimized to determine the coefficient resulting in the backward method solution given by

\[ k_B^{(n+1)} = \frac{\mathbb{E}\{e^{(n)}(k)e^{(n)}(k-1)\}}{\mathbb{E}\{e^{(n)}(k)^2\}} \]  

(2.31)

Since, however, the forward and backward prediction error transfer functions are given by \( B^{(n)}(z) \) and \( z^{-1}B^{(n)}(z^{-1}) \), it follows that

\[ |B^{(n)}(z)| = |B^{(n)}(z)| \]  

(2.32)

and since they are both driven by the same input, \( Y(z) \), the mean square values of both the forward and backward prediction error signals at a given stage are the same making equations (2.30) and (2.31) equivalent. It is also possible to show that they are equivalent to equation (2.18b). Recognizing that the required expectations will eventually have to be estimated by using time averages, these two methods for calculating \( k^{(n+1)} \) will not in general be exactly equivalent and it might be preferable to use the arithmetic mean of the mean square values of forward and
backward prediction error as a cost function

$$\mathcal{E}[\varepsilon (n+1)(k)^2 + \varepsilon (n+1)(k)^2]$$  \hspace{1cm} (2.33a)

This leads to a third method derived by Burg [Ref. 5] in his work on maximum entropy spectral analysis and given by

$$k_{BG}^{(n+1)} = \frac{2\varepsilon \{e(n)(k)e(n)(k-1)\}}{\varepsilon \{e(n)(k)^2\} + \varepsilon \{e(n)(k-1)^2\}}$$  \hspace{1cm} (2.33b)

Notice that $k_{BG}^{(n+1)}$ is the harmonic mean of $k_F^{(n+1)}$ and $k_B^{(n+1)}$. A fourth method due to Itakura and Saito [Ref. 20] can also be derived which results in

$$k_{IS}^{(n+1)} = \frac{\varepsilon \{e(n)(k)e(n)(k-1)\}}{\sqrt{\varepsilon \{e(n)(k)^2\}\varepsilon \{e(n)(k-1)^2\}}}$$  \hspace{1cm} (2.34)

and $k_{IS}^{(n+1)}$ is simply the geometric mean of the forward and backward coefficients.

Since equation (2.34) is of the form of a normalized correlation $k_{IS}$ will always be bounded by unity in magnitude as required by equation (2.19). Furthermore since

$$|\text{Harmonic Mean}| < |\text{Geometric Mean}|$$

it follows that $k_{BG}^{(n+1)}$ will be similarly bounded. These bounds are significant since Markel [Ref. 32] has shown that $|k^{(n)}| < 1$ is a necessary and sufficient condition to ensure
that the roots of $B^{(n)}(z)$ be within the unit circle guaranteeing the stability of the $n$-th order all pole model. No such guarantees of model stability exist when the forward or backward solution methods of equations (2.30) and (2.31) are used with the correlation estimates obtained by averaging for finite time intervals.

To determine the AR synthesis model in lattice form it is only necessary to rewrite equation (2.28a) as

$$e^{(n)}(k) = e^{(n+1)}(k) + k^{(n+1)}e^{(n)}(k-1) \quad (2.35)$$

Together with equations (2.28b) and (2.28c), this describes the structure shown in Figure 2.6 for a second order case and when it is driven by the second order prediction error signal, it will reconstruct $y(k)$ exactly. Thus it implements the transfer function $\frac{1}{B^{(2)}(z)}$ or, in general, when stages are used $\frac{1}{B^{(N)}(z)}$. Recognizing that if the prediction error model is an accurate model of the system denominator polynomial, $e^{(N)}(k) = a(o)u(k)$, this input signal is used in the synthesis model. Because of analogies with transmission lines and wave propagation models, the lattice coefficients have been referred to as reflection coefficients.
2. The MA Modeling Lattice Structure

A similar lattice form is applicable to the MA modeling problem. From equations (2.21) and (2.23a) the transfer function of the MA model can be written recursively in order as

\[ A(n+1)(z) = A(n)(z) + g(n+1)B_n+1(z) \]  \hspace{1cm} (2.36)

where, as discussed in connection with equation (2.23a), \( B_n+1(z) \) is the backward prediction error transfer function for an autoregressive model of the input signal \( u(k) \). Multiplying both sides of equation (2.36) by \( U(z) \) and transforming into the time domain it follows that

\[ \hat{y}(n+1)(k) = \hat{y}(n)(k) + g(n+1)e_{n+1}(k) \]  \hspace{1cm} (2.37)
where \( e_{n+1}(k) \) is the backward prediction error signal from the autoregression on the input signal \( u(k) \), and can be obtained by operating a prediction error lattice with \( u(k) \) as its input. Then with the additional term in equation (2.37) the lattice form of the MA model can be drawn as shown in Figure 2.7.

![Lattice Form Of The MA Model (Second Order Case)](image)

It was stated earlier that the AR prediction error lattice, as a by-product, forms a set of orthogonal or uncorrelated backward prediction error signals from its input. Here in the MA model, these orthogonal signals are linearly combined to form the MA estimate of the system output. If the input signal \( u(k) \) is a white process, an
examination of any of the solution methods previously discussed will show that all of the \( \{k^{(n)}\} \) lattice coefficients will be zero since the delayed samples of a while process are already orthogonal. Otherwise the \( \{k^{(n)}\} \) lattice coefficients will be set to orthogonalize the backward prediction error signals. As a consequence of this, the weighting coefficients \( g^{(n)} \) can be set independently of each other; that is \( g^{(n)} \) can be set to minimize

\[
\varepsilon \{e^{(n)}(k)^2\} = \varepsilon \{[y(k) - \hat{y}^{(n)}(k)]^2\}
\]  

(2.38)
given the best prediction of order \( n-1, \hat{y}^{(n)}(k) \). This results in an alternative expression to equation (2.23b) for \( g^{(n)} \) given by

\[
g^{(n)} = \frac{\varepsilon \{e^{(n-1)}(k) e^{(n)}(k)\}}{\varepsilon \{e^{(n)}(k)^2\}} = \frac{\varepsilon \{y(k) e^{(n)}(k)\}}{\varepsilon \{e^{(n)}(k)^2\}}
\]  

(2.39)

Here \( e^{(n)}(k) \) is the error between the system output and its \( n \)-th order MA estimate.

E. MULTICHANNEL AR AND MA MODELING

Both the AR and MA modeling problems previously discussed, as well as their solution via the Levinson recursion and lattice filter methods, can be generalized to the multichannel case by replacing the various signals with signal vectors and replacing the weighting coefficients with appropriately dimensioned matrices of coefficients.
A discussion of this appears in Robinson [Ref. 45]. The equations that describe the AR and MA models and their MMSE solutions are repeated here for convenience.

\[ \hat{y}(k) = \sum_{i=1}^{N} b(i) y(k-i) \quad (2.40a) \]

\[
\begin{bmatrix}
R_{yy}(0) & R_{yy}(-1) & \ldots & R_{yy}(1-N) \\
R_{yy}(1) & R_{yy}(0) & \ldots & R_{yy}(2-N) \\
\vdots & \vdots & \ddots & \vdots \\
R_{yy}(N-1) & R_{yy}(N-2) & \ldots & R_{yy}(0)
\end{bmatrix}
\begin{bmatrix}
b(1) \\
b(2) \\
\vdots \\
b(N)
\end{bmatrix}
= 
\begin{bmatrix}
R_{yy}(1) \\
R_{yy}(2) \\
\vdots \\
R_{yy}(N)
\end{bmatrix} \quad (2.40b)
\]

\[ \hat{y}(k) = \sum_{i=0}^{N} a(i) u(k-i) \quad (2.41a) \]

\[
\begin{bmatrix}
R_{uu}(0) & R_{uu}(-1) & \ldots & R_{uu}(-N) \\
R_{uu}(1) & R_{uu}(0) & \ldots & R_{uu}(1-N) \\
\vdots & \vdots & \ddots & \vdots \\
R_{uu}(N) & R_{uu}(N-1) & \ldots & R_{uu}(0)
\end{bmatrix}
\begin{bmatrix}
a(0) \\
a(1) \\
\vdots \\
a(N)
\end{bmatrix}
= 
\begin{bmatrix}
R_{uy}(0) \\
R_{uy}(1) \\
\vdots \\
R_{uy}(N)
\end{bmatrix} \quad (2.41b)
\]

In a multichannel generalization, \( y(k) \) becomes a vector of \( Q_0 \) output signals, \( u(k) \) becomes a vector of \( Q_i \) input signals, \( b(i) \) becomes a square matrix of \( Q_0 \times Q_0 \) coefficients and \( a(i) \) becomes a \( Q_0 \times Q_i \) matrix of coefficients so that the \( N \)-th
order multichannel model equations can be written as

\[
\hat{\gamma}(k) = \sum_{n=1}^{N} b(n) \gamma(k-n) \quad (2.42a)
\]

\[
\hat{\nu}(k) = \sum_{n=0}^{N} a(n) u(k-n) \quad (2.42b)
\]

The equations for the MMSE solutions (2.40b) and (2.41b) generalize directly as well by replacing each correlation coefficient \( R_{\nu\nu}(n) \) by matrices of correlation coefficients given by

\[
R_{\nu\nu}(n) = \varepsilon \{ \nu(k) \, w(k+n) \}^T \quad (2.43)
\]

where \( \nu(k) \) and \( w(k) \) are signal vectors. This causes the overall correlation matrices to take on a block Toeplitz structure. The transfer function relationships of the AR prediction error model and the MA model take the form of matrix polynomials

\[
B(z) = b(1) z^{-1} + \cdots + b(N) z^{-N} \quad (2.44a)
\]

\[
A(z) = a(0) + a(1) z^{-1} + \cdots + a(N) z^{-N} \quad (2.44b)
\]

so that

\[
E(z) = [I - B(z)] \nu(z) \quad (2.44c)
\]
\[ \hat{Y}(z) = A(z) U(z) \quad (2.44d) \]

where \( E(z) \) is the transform of the multichannel AR prediction error vector \( e(k) = y(k) - \hat{y}(k) \). Alternately, equations (2.44a) and (2.44b) can be written as single matrices whose entries are polynomials in \( z \) rather than scalars. \( B(z) \) is of necessity a \( Q_0 \times Q_0 \) square matrix polynomial while the dimensions of \( A(z) \) (\( Q_0 \times Q_1 \)) depend upon the number of inputs and outputs which need not be the same.

In the single channel AR problem, \( B(z) \) provides the transfer function of the prediction error, or inverse filter, and must be inverted to obtain the all pole synthesis filter. The stability of the synthesis model therefore depends upon the roots of this polynomial. The matrix polynomial \([I-B(z)]\) in the multichannel AR problem is, in like fashion, an inverse filter representation and must be inverted to obtain the synthesis model. This inversion of a matrix with polynomial entries is defined in the same manner as the inversion of a square matrix with scalar entries. To see what this inverse matrix polynomial looks like consider as an example a two channel autoregression with a prediction error filter given by

\[
[I-B(z)] = \begin{bmatrix}
  1 - b_{11}(z) & -b_{12}(z) \\
  -b_{21}(z) & 1 - b_{22}(z)
\end{bmatrix}
\]
Applying Cramer's rule, the inverse matrix polynomial is written as

$$[I-B(z)]^{-1} = \frac{1}{\det[I-B(z)]} \begin{bmatrix} 1-b_{22}(z) & b_{21}(z) \\ b_{12}(z) & 1-b_{11}(z) \end{bmatrix}$$

and it is apparent that the stability of the multichannel synthesis model is dependent upon the locations of the zeros of the polynomial $\det[I-B(z)]$.

This straightforward generalization of the AR and MA models is what has customarily been used in the literature to develop the multichannel models and similarly derived generalizations of the Levinson algorithm to solve them recursively in order are available as well. [Refs. 57 and 61]

The multichannel AR and MA modeling problems and their solutions via the Levinson algorithm can however be recast as shown in Appendix A to make them compatible with the form of other linear and nonlinear modeling problems. To avoid confusion later in the application of the results, the derivations in Appendix A have been carried out in a generic form with $x$ and $d$ used to represent some of the signals and coefficients respectively. The symbols $u$, $y$, $a$ and $b$ have been reserved to denote system input, output and weighting coefficients.

Equations (A.7) and (A.26) provide the MMSE solutions to the multichannel AR and MA modeling problems in forms
different than (although entirely equivalent to) those resulting from the straightforward generalizations of equations (2.40b) and (2.41b). The multichannel generalization of the Levinson algorithm derived in Appendix A can, with one exception, be seen as a matrix algebra generalization of the single channel algorithm and as, one would suspect, the single channel algorithm results as a special case of Appendix A. The one exception is that in the multichannel case, the n-th order forward and backward prediction error models are not simply related to one another. The single channel AR backward predictor is given by \( z^{-n}b(n)(z^{-1}) \) but in the multichannel case the backward prediction is not \( z^{-n}[I-D(n)(z^{-1})] \).

Because of this, two reflection coefficient matrices \( K \) and \( \overline{K} \) are required at each stage in the recursion to relate the n-th and (n+1)-st order solutions rather than just one as in the single channel problem. Also, in the single channel case, the fact that \( |b(n)(z)| = |b(n)(z)| \) and therefore \( \varepsilon\{e(n)(k)^2\} = \varepsilon\{e(n)(k)^2\} \) made possible the derivation of the Burg algorithm and the Itakaura-Saito algorithm, which ensured the magnitude of the reflection coefficient was bounded by unity and that equation (2.19) would result in nonnegative values of MSE. In the multichannel algorithm however, the forward and backward prediction error covariance matrices and their traces are not the same (except for \( P(0) \) and \( \overline{P}(0) \)) and as a result, straightforward generalizations of the Burg and Itakaura-Saito algorithms are not possible.
Consequently with correlation estimates obtained by averaging over finite time intervals, there are no guarantees that equations (A.18a) and (A.19b) maintain the positive definiteness of the prediction error covariance matrices.

Multichannel generalizations of Burg's algorithm due to Nuttal [Refs. 40 and 41], Morf [Ref. 39], and Strand [Ref. 50] which guarantee the positivity of the covariance matrices are available but are not explored here because of their complexity and because they would take the discussion too far afield.

The relations of equations (A.20) and (A.30) which are repeated here for convenience permit the construction of the multichannel AR analysis and synthesis lattice structures and the MA lattice structure. For the multichannel AR model,

\[ e^{(n+1)}(k) = e^{(n)}(k) - K^{(n+1)} e^{(n)}(k) \]  
(2.45a)

\[ e^{(n+1)}(k) = e^{(n)}(k-l) - K^{(n+1)} e^{(n)}(k) \]  
(2.45b)

and the corresponding prediction error lattice is shown in Figure 2.8.
To obtain the multichannel AR synthesis lattice, equation (2.45a) can simply be rewritten as

\[ e(n)(k) = e(n+1)(k) + \mathbf{K}(n+1)^T \mathbf{\overline{e}}(n)(k-1) \]  

(2.46) resulting in the structure shown in Figure 2.9.
For the multichannel MA model, equations (2.45) and
\[
\hat{\nu}^{(n+1)}(k) = \hat{\nu}^{(n)}(k) + G^{(n+1)} T \hat{e}^{(n+1)}(k)
\] (2.47)
describe the lattice structure shown in Figure 2.10.

![Diagram of multichannel MA lattice structure](image)

**Figure 2.10.** Multichannel MA lattice structure for a second order model.

As in the single channel case, the multichannel prediction error lattice exhibits the successive decoupling property and orthogonalizes the backward prediction errors at the various stages so that
\[
\varepsilon[\hat{e}^{(i)}(k) \hat{e}^{(j)}(k)^T] = \begin{cases} 
0 & \text{i} \neq \text{j} \\
\sigma^2(i) & \text{i} = \text{j}
\end{cases} \] (2.48)
As a consequence of the successive decoupling, the forward and backward reflection coefficient matrices at the \((n+1)\)-st stage can be set to minimize the trace of the forward and backward prediction error covariance matrices respectively, given the best lattice of order \(n\). This provides alternatives to equations (A.17) for calculating \(K\) and \(\bar{K}\) and also generalizes the forward and backward single channel solutions discussed previously, resulting in

\[
K(n+1) = P(n)^{-1} \Delta(n)^T \tag{2.49a}
\]

\[
\bar{K}(n+1) = P(n)^{-1} \Delta(n) \tag{2.49b}
\]

where

\[
\Delta(n) = \varepsilon\{e(n)(k-1)^T\} \tag{2.49c}
\]

It is also possible to show that these relationships are entirely equivalent to equation (A.17). In the multi-channel MA lattice, the orthogonality of the backward prediction error signals also allows the \(G\) matrices to be calculated in succession providing an alternative to equation (A.28c) and generalizing the single channel solution given by equation (2.39). Setting \(G(n)\) to minimize the trace of the error covariance matrix

\[
P_0(n) = \varepsilon\{e_0(n)(k)\ e_0(n)(k)^T\} \tag{2.50a}
\]
where

\[ e_0^{(n)}(k) = \chi(k) - \hat{\chi}^{(n)}(k) \]  

(2.50b)

results in a solution given by

\[ \hat{e}^{(n)} = \bar{P}(n)^{-1} \left\{ \bar{e}^{(n)}(k) \chi(k)^T \right\} \]  

(2.50c)

Another important characteristic of the lattice solutions to the AR and MA modeling problems given by equations (2.49) and (2.50) and their single channel counterparts is that they do not impose any requirements to window the data when finite time averages are used to estimate correlations. The autocorrelation function of a signal is inherently an even function so that \( R_{vv}(n) = R_{vv}(-n) \). This fact is responsible for much of the special structure of the correlation matrices that appear in the model solution equations, and was also used in the derivation of the Levinson algorithm. In estimating the autocorrelation function via time averaging over a finite interval, a window function that is nonzero over only a given interval must be applied to the data to retain this even symmetry property in the estimate. If the data is not set to zero outside a given interval, end effects will destroy the symmetry so that \( \hat{R}_{vv}(n) \neq \hat{R}_{vv}(-n) \) as depicted in Figure 2.11.
In the lattice solutions of equation (2.49) however, there is no requirement to make such an artificial assumption about the data (that it is zero outside some interval).

These properties of the lattice solution methods were responsible for their initial use by Burg in his work on maximum entropy spectral analysis [Ref. 5] and have continued to generate interest in the application of lattice methods to other types of modeling problems.

Figure 2.11. Time averaging to estimate correlations without windowing.
F. ADAPTIVE MODELING

The LMS adaptive algorithm provides a well known alternative method for obtaining the solution to the AR or MA modeling problems which does not require the estimation of correlations or the inversion of a matrix [Refs. 58, 59 and 60]. This algorithm updates an estimate of the model solution vector at each time instant by an amount proportional to the negative of the instantaneous gradient of the cost function; i.e., in a MA model,

\[ a^+(k+1) = a^+(k) - \mu \nabla^+(k) \quad (2.51a) \]

where \( \mu \) is a proportionality constant or adaptive gain. Since the actual gradient is usually not known, it is approximated by using the square of a single sample of the error as an estimate of the MSE so that

\[ \nabla^+(k) = \frac{\partial e(k)^2}{\partial a^+} \bigg|_{a^+=a^+(k)} = -2 \mu^+(k) e(k) \quad (2.51b) \]

and

\[ a^+(k+1) = a^+(k) + \mu \nabla^+(k) e(k) \quad (2.51c) \]

In each of the models considered here, the cost function (MSE or trace \( P \)) is a quadratic function of the model weights and defines a concave hyperparabola with a unique minimum. The functioning of the LMS algorithm under these conditions...
can easily be understood by considering the scalar case of equation (2.51) illustrated in Figure 2.12.

As this illustration shows, the algorithm can actually diverge for too large a value of adaptive gain. The rate of convergence is also dependent on the size of the adaptive gain. Widrow has shown that for stability, the gain must be set so that

\[ 0 < \mu < \frac{1}{\lambda_{\text{max}}} \]  

(2.52a)

While, in the mean, the weight vector converges with an exponential time constant of

\[ \tau = \frac{1}{2\mu \lambda_{\text{min}}} \]  

(2.52b)

where \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \) are the smallest and largest eigenvalues of the input autocorrelation matrix \( \Sigma_{uu} \). From the standpoint of stability, \( \mu \) should be made relatively small but for rapid convergence, equation (2.52b) dictates that it should be large. Setting

\[ \mu = \frac{\alpha}{\lambda_{\text{max}}} \]  

(2.53a)

where \( \alpha \) is a normalized gain and \( 0 < \alpha < 1 \), equation (2.52b) becomes

\[ \tau = \frac{1}{2\alpha} \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \]  

(2.53b)
Figure 2.2. Behavior of the LMS adaptive algorithm for various adaptive gains.

1. Small adaptive gain; steady convergence toward the solution.
2. Intermediate adaptive gain; oscillatory convergence toward the solution.
3. Large adaptive gain; divergence away from the solution.
and for a wide disparity between the largest and smallest
eigenvalues (\(\lambda_{\text{min}} \ll \lambda_{\text{max}}\)), convergence will be quite slow.
This consideration becomes increasingly important when high
order model solutions are obtained adaptively since the
dimensionality of the input autocorrelation matrix will be
high and the possibility of a wide eigenvalue disparity
greater.

These same adaptive techniques have been applied by
Griffiths [Refs. 16, 17, 18 and 31], to the AR and MA
lattice filter structures derived in the last section. The
key difference between the conventional adaptive filter and
the adaptive lattice is that in the lattice, the adaptation
is carried out on a stage by stage basis for each of the
reflection coefficient matrices, while in the more conven-
tional approach, the entire weight vector is adapted. It
has already been established that the lattice structure
makes the model solutions recursive in order. Implementing
the lattice adaptivity makes the solution recursive in time
as well since the estimate of the solution at each instant
is dependent upon prior estimates of the solution.

The conventional adaptive filter algorithm forms an
error signal as the difference between some desired signal
and its estimate; i.e.,

\[
e(k) = y(k) - a^+(k)T u^+(k)
\]  

(2.54)
where \( y(k) \) is the desired signal, \( a^+ \) is the weight vector and \( u^+(k) \) is the input vector, and the time update for the weight vector is given by equation (2.51c). To derive the adaptive AR lattice consider equations (2.45) for a single stage. The lattice in general has vector error and desired signals and coefficient matrices as opposed to scalar error and desired signals and a coefficient vector in equation (2.54) but such a generalization is straightforward. Comparing equation (2.45a) to (2.54) it is clear that:

1) \( e^{(n+1)}(k) \) is analogous to the error signal;
2) \( e^{(n)}(k) \) is analogous to the desired signal;
3) \( \tilde{e}^{(n)}(k-1) \) is analogous to the input signal vector.

Using the trace of \( P^{(n+1)} \) as a cost function and applying a LMS adaptive algorithm to determine the forward reflection coefficient matrix it follows that

\[
K^{(n+1)}(k+1) = K^{(n+1)}(k) + 2 \mu^{(n+1)} \tilde{e}^{(n)}(k-1) e^{(n+1)}(k)^T
\]

(2.55a)

With these analogies, equations (2.51c) and (2.55a) are seen to be virtually identical with the exception that (2.51c) uses a scalar error to adapt a weight vector and (2.55a) uses a vector error signal to adapt a coefficient matrix.

Proceeding in a similar fashion with equation 2.45b it is clear that:

1) \( \tilde{e}^{(n+1)}(k) \) is analogous to the error signal;
2) \( \tilde{e}^{(n)}(k-1) \) is analogous to the desired signal;
3) \( e^{(n)}(k) \) is analogous to the input signal vector. With the trace of \( P^{(n+1)} \) as a cost function, the time update relation for the backward reflection coefficient matrix is

\[
P^{(n+1)}(k+1) = P^{(n+1)}(k) + 2P^{(n+1)} e^{(n)}(k) e^{(n+1)}(k) T
\]

(2.55b)

For a MA lattice, equation (2.47) must also be considered. Multiplying both sides of (2.47) by minus one and adding \( y(k) \) results in

\[
e^{(n+1)}(k) = e^{(n)}(k) - G^{(n+1)} T e^{(n+1)}(k)
\]

(2.56)

where \( e^{(n+1)}(k) \) is defined as in equation (2.30). It is evident that:

1) \( e^{(n+1)}(k) \) is analogous to the error signal;
2) \( e^{(n)}(k) \) is analogous to the desired signal;
3) \( e^{(n+1)}(k) \) is analogous to the input signal vector.

With the trace of \( P_0^{(n+1)} \) as a cost function, the time update relation for \( G^{(n+1)} \) is given by

\[
G^{(n+1)}(k+1) = G^{(n+1)}(k) + 2 \mu e^{(n+1)}(k) e^{(n+1)}(k) T
\]

(2.57)
It is significant to note that three different adaptive gains have been used in equations (2.55) and (2.57) and that the gains have been superscripted indicating that they vary from one lattice stage to the next. For stability considerations the adaptive gain used in the LMS algorithm must satisfy equation (2.52a) and therefore is related to the largest eigenvalue of the input autocorrelation matrix by equation (2.53a). In developing the time update relations for the lattice coefficients, three different input signals were used and these inputs also differ from one lattice stage to the next. Indeed, even for the case where the input $\mathbf{X}(k)$ to the lattice structure is stationary, the inputs to all lattice stages except the first are nonstationary since these inputs are outputs of other lattice stages. This fact indicates that time varying adaptive gains are appropriate as well in equations (2.55) and (2.57). Equation (2.53a) is of no direct usefulness however in setting the adaptive gains since the time varying eigenvalues are not known. Recognizing that the largest eigenvalue is always less than the trace of the input autocorrelation matrix (which is a measure of the power in the input signal vector) the gains can be set as

$$
\mu^{(n+1)}(k) = \frac{\alpha}{\sigma^{(n+1)}(k)^2} \tag{2.58a}
$$
where \( \sigma_{n+1}(k)^2 \), \( \overline{\sigma}_{n+1}(k)^2 \) and \( \gamma_{n+1}(k)^2 \) are estimates of the power in the three input signal vectors and \( \alpha \) is a normalized adaptive step size with \( 0 < \alpha < 1 \). A method that has commonly been applied to obtain these estimates is to employ a first order low pass filter so that

\[
\sigma_{n+1}(k+1)^2 = [1-\alpha]\sigma_{n+1}(k)^2 + \alpha \overline{e}^2(n)(k-1) T_e^2(n-1)(k-1) \quad (2.59a)
\]

\[
\overline{\sigma}_{n+1}(k+1)^2 = [1-\alpha]\overline{\sigma}_{n+1}(k)^2 + \alpha \overline{e}^2(n)(k) T_e^2(n)(k) \quad (2.59b)
\]

\[
\gamma_{n+1}(k+1)^2 = [1-\alpha]\gamma_{n+1}(k)^2 + \alpha \overline{e}_0^2(n)(k) T_{e_0}^2(n)(k) \quad (2.59c)
\]

Taken together, equations (2.55), (2.57), (2.58) and (2.59) define the adaptive solutions for the AR and MA multichannel lattice models.

To understand the potential advantage offered by the adaptive lattice form, recognize that while the conventional approach solves a high order minimization problem by adapting all the coefficients at once, the lattice breaks the problem down into a succession of lower order minimizations at each stage and solves these lower order problems adaptively. The
dimensionality of the input autocorrelation matrices at each lattice stage in general is significantly less than that of the large input autocorrelation matrix in the conventional adaptive algorithm and consequently it is hoped that the possibility of a large eigenvalue disparity with its attendant slow convergence is reduced.

This advantage is most evident in a single channel adaptive lattice where the inputs at each stage are single signals and their corresponding autocorrelation matrices are 1x1 in dimension. In this case the ratio of smallest to largest eigenvalues is unity and the convergence of each stage is quite rapid while the convergence of the overall model is independent of the eigenvalue ratio for the overall higher dimension input autocorrelation matrix. This has been demonstrated by Satorius [Refs. 46 and 47] who has shown that the single channel adaptive lattice converges much more rapidly than the corresponding conventional adaptive filter, and does so independently of the eigenvalue ratio on the overall input channel autocorrelation matrix.

Furthermore in a single channel adaptive lattice, the time update relations are simplified by the fact that the forward and backward reflection coefficients are the same. Using the average of the mean square values of backward and forward prediction errors as a cost function and applying an adaptive algorithm it follows that
\[ k^{(n+1)}(k+1) = k^{(n+1)}(k) + \frac{\alpha}{\sigma_{n+1}(k)} [e^{(n)}(k)\bar{e}^{(n+1)}(k) + e^{(n)}(k-1)\bar{e}^{(n+1)}(k)] \]  

(2.60a)

where

\[ \sigma_{n+1}(k+1)^2 = [1-\alpha] \sigma_{n+1}(k)^2 + \frac{\alpha}{2} [e^{(n)}(k)^2 + e^{(n)}(k-1)^2] \]  

(2.60b)

The very nature of the lattice structure however with the output of one stage providing an input to the next stage, greatly complicates the analysis of the convergence properties of the adaptive lattice model. Even when \( x(k) \), the input to the lattice, is stationary, inputs to all stages except the first are nonstationary. An approximate analysis of convergence and stability on a stage by stage basis is possible if it is assumed that all prior stages have converged and are providing stationary inputs to the stage under investigation. With this assumption, the adaptive solution for the \( k^{(n+1)} \), \( \bar{k}^{(n+1)} \) and \( \bar{e}^{(n+1)} \) matrices are obtained from the operation of three independent LMS algorithms as shown earlier, with inputs given by \( \bar{e}^{(n)}(k-1) \), \( e^{(n)}(k) \) and \( \bar{e}^{(n+1)}(k) \), respectively. Stability limits on the adaptive gains used in the stage and the convergence properties of the stage are then determined by the eigenvalues of the \( F^{(n)} \), \( P^{(n)} \) and \( F^{(n+1)} \) matrices. A more exact analysis of the properties of the adaptive lattice that
considers the nonstationary character of the inputs to the second and subsequent stages is not currently available.
III. ARMA MODELING

One of the most serious disadvantages of either AR or MA modeling is the fact that to adequately represent even simple linear systems, both methods may require a large number of parameters (a high order model). This problem arises since, from a transfer function standpoint, AR and MA models attempt to model the system using only poles or only zeros, in spite of the fact that the physical system may have both zeros and poles. While modeling the effects of a zero with a number of poles and visa versa can be analytically justified as shown in the previous chapter, it makes far more sense (both from the viewpoint of model accuracy and efficient use of model parameters) to let the model represent the system as it really is with both zeros and poles if this is at all possible. The ARMA model is a generalization of the AR and MA models and accomplishes exactly this, representing the system in rational transfer function form.

It is worth noting that the titles of all pole and all zero modeling that have been associated with AR and MA modeling are misnomers. Both have equal numbers of zeros and poles. In the AR model however, all the zeros occur at the origin of the z-plane as do the poles of a MA model. The ARMA model removes these constraints.
After a brief discussion of two alternate ARMA modeling methods due to Shanks and Prony, the equation error formulation for ARMA modeling is developed and the new results presented. Model transition formulas relating the ARMA model to the MA and AR models are developed and the input signal requirements of the modeling process explored. It is shown that after suitable modification, the Levinson algorithm can be applied to solve the ARMA modeling problem recursively in order and lattice solution methods are also developed for both a batch processing and an adaptive model solution. The results of experimental simulations of both of these modeling solution methods are presented and discussed, and comparisons are made with conventional means of ARMA modeling using the equation error formulation. Finally it is also shown that the lattice solution methods can be generalized to solve for the multichannel ARMA model with arbitrary numbers of inputs and outputs.

A. LINEAR ARMA MODELING AND ITS RELATION TO AR AND MA MODELING

The ARMA model for linear systems assumes the current value of the output of the system is given by a weighted combination of present and past values of the input and past values of the output. In terms of the discussions of Chapter I, \( F_{30} \) is assumed to be zero and \( F_{10} \) and \( F_{20} \) take on the following forms

\[
F_{10}[u(k)] = \sum_{n=0}^{N} a(n) u(k-n)
\]  

(3.1a)
leading to a transfer function representation for the system given by

$$H(z) = \frac{\sum_{n=0}^{N} a(n) z^{-n}}{1 - \sum_{n=1}^{N} b(n) z^{-n}} = \frac{A(z)}{B(z)}$$  \hspace{1cm} (3.1c)

A number of methods exist for finding the model coefficients \{a(n)\} and \{b(n)\}. As stated in Chapter I, a MMSE solution via the direct form modeling approach requires the solution of a system of highly nonlinear equations and in general is untractable. An alternative is to first obtain an estimate of the denominator polynomial \(B(z)\) by some means such as AR modeling and then using this in the system shown in Figure 3.1, estimate the numerator polynomial \(A(z)\) by setting its coefficients to minimize the mean square value of the error. This method was first explored by Shanks. [Ref. 49]

Another alternative is to apply the Prony method [Refs. 8, 52 and 56] derived in Appendix E which obtains the model parameters by matching the impulse response of the system and model over the first \(N+M+1\) sample intervals. Both of these techniques share a common characteristic. They both start by independently estimating the denominator coefficients (or model poles) and then, given this estimate, solve for
Figure 3.1. Shanks method for ARMA modeling

The numerator coefficients (or model zeros). This is intuitively unappealing in that one would expect these two estimation problems to be more closely coupled with the zero estimates also affecting the estimates of the poles.

The application of the equation error formulation to the ARMA modeling problem permits simultaneous estimation of the model zeros and poles and was first used by Kalman [Ref. 23] in work on self-optimizing control systems. The prediction error form of the model is considered here where $P_{30}$ is set to zero while

$$F_{10}[u(k)] = \sum_{n=0}^{M} a(n) u(k-n) \quad (3.2a)$$

$$F_{2}[y(k)] = y(k) - \sum_{n=1}^{N} b(n) y(k-n) \quad (3.2b)$$
The analysis model is depicted in Figure 3.2 where

$$A(z) = \sum_{n=0}^{M} a(n)z^{-n} \quad (3.3a)$$

and

$$B(n) = 1 - \sum_{n=1}^{N} b(n)z^{-n} \quad (3.3b)$$

and these polynomials are the estimates of the system transfer function numerator and denominator.

Figure 3.2. The equation error formulation for ARMA modeling.
The expression for the model error can be written as

\[ e_0(k) = y(k) - [y(k)^T \mid u^+(k)^T] \begin{bmatrix} b \\ a^+ \end{bmatrix} \]  

(3.4a)

where \( y(k) \) and \( b \) are defined as in equation (2.7) and

\[ u^+(k) = [y(k) \cdots u(k-M)]^T \]  

(3.4b)

\[ a^+ = [a(0) \cdots a(M)]^T \]  

(3.4c)

This results in an expression for the mean square error which is a quadratic function of the model coefficients, with the MMSE solution for those coefficients given by

\[ \begin{bmatrix} R_{yy} & R_{yu}^+ \\ \hline R_{u+y} & R_{uu}^+ \end{bmatrix} \begin{bmatrix} b \\ a^+ \end{bmatrix} = \begin{bmatrix} R_{yy} \\ \hline R_{u+y} \end{bmatrix} \]  

(3.5a)

and

\[ E_{2_{\text{min}}} = R_{yy}(0) - [b^T \mid a^+^T] \begin{bmatrix} R_{yy} \\ \hline R_{u+y} \end{bmatrix} \]  

(3.5b)

1. **Model Transition Relationships**

Comparing equations (3.4) and (3.5) with their AR and MA counterparts, it is clear that the ARMA model provides a generalization of these other models since they can be obtained from the ARMA model by assuming that either \( a^+ \) or \( b \) is zero. Consequently it is susceptible to the same type of
bias introduced in the AR and MA models by the presence of additive noise on either the system input or output signals. To develop the relationships between these models further, consider the inversion of the correlation matrix in equation (3.5a) in terms of its component matrices. Since the left and right inverses of a nonsingular square matrix are the same, either

\[
\begin{bmatrix}
A & B \\
C & D
\end{bmatrix}
\begin{bmatrix}
R_{yy} & R_{yu} \\
R_{uy} & R_{uu}
\end{bmatrix}
= 
\begin{bmatrix}
I & 0 \\
0 & I
\end{bmatrix}
\]

or

\[
\begin{bmatrix}
R_{yy} & R_{yu} \\
R_{uy} & R_{uu}
\end{bmatrix}
\begin{bmatrix}
A & B \\
C & D
\end{bmatrix}
= 
\begin{bmatrix}
I & 0 \\
0 & I
\end{bmatrix}
\]

(3.6a)

(3.6b)

can be used to find the required inverse in partitioned form. Solving for the right inverse of equation (3.6b) yields

\[
A = \left( R_{yy} - \frac{1}{R_{yu}} + \frac{1}{R_{u}R_{u}} \right)^{-1}
\]

(3.7a)

\[
B = -\frac{1}{R_{yu}} \frac{1}{R_{yu}} + \left( \frac{1}{R_{u}R_{u}} - \frac{1}{R_{yu}} \frac{1}{R_{yu}} \right)^{-1}
\]

(3.7b)

\[
C = -\frac{1}{R_{u}R_{u}} + \frac{1}{R_{yu}} \left( R_{yy} - \frac{1}{R_{yu}} + \frac{1}{R_{yu}} \frac{1}{R_{yu}} \right)^{-1}
\]

(3.7c)

\[
D = \left( \frac{1}{R_{u}R_{u}} - \frac{1}{R_{yu}} \frac{1}{R_{yu}} \right)^{-1}
\]

(3.7d)
Solving for the left inverse of equations (3.6a) gives identical results for \( A \) and \( D \) while equivalent but different forms are obtained for \( B \) and \( C \) given by

\[
B = -(R_{yy} - R_{yu} + R_{uy} + R_{uy}^{-1})^{-1} R_{yu} + R_{uy}^{-1}
\]

\[
C = -(R_{uu} + R_{uy} + R_{yy} R_{yu} + R_{yy}^{-1})^{-1} R_{uy} + R_{yy}^{-1}
\]

Using equations (3.7) in equation (3.5a), the solutions for the ARMA coefficient vectors are given by

\[
b = (R_{yy} - R_{yu} + R_{uy} + R_{uy}^{-1})^{-1} R_{yy}
\]

\[
- R_{yy}^{-1} R_{yu} + (R_{uu} + R_{uy}^{-1} R_{yy} R_{yu} + R_{yy}^{-1})^{-1} R_{uy}
\]

and

\[
a^+ = - R_{uu}^{-1} + R_{uy} + (R_{yy} - R_{yu} + R_{uy} + R_{uy}^{-1})^{-1} R_{yy}
\]

\[
+ (R_{uu} + R_{uy} - R_{yy} R_{yu}^{-1} R_{yy}^{-1} R_{uu} + R_{uy})^{-1} R_{uy}
\]

The matrix inversion lemma [Refs. 11 and 19] which states that

\[
(E + FGH)^{-1} = E^{-1} - E^{-1} F (G^{-1} + HE^{-1} F)^{-1} HE^{-1}
\]

\[ (3.10) \]
for nonsingular square matrices $E$, $G$ and $E + F G H$, can be
used in equation (3.9) to rewrite them as

$$b = R^{-1} + R^{-1} R_{y y} R_{y y} + (R_{u u} + R_{u y} R_{y y} R_{y u})^{-1}$$

$$[R_{u y} + R_{y y} R_{y y} R_{u y} - R_{u y}]$$

(3.11a)

$$a^+ = R^{-1} + R_{u y} + R_{u y} + (R_{u y} R_{y y} R_{u y} R_{y u})^{-1}$$

$$[R_{u y} + R_{u y} + R_{u y} R_{y y} - R_{u y}]$$

(3.11b)

The all zero and all pole model solutions of corresponding
orders however are

$$a^+_{AZ} = R^{-1} + R_{u y} + R_{u y}$$

and $b_{AP} = R^{-1} R_{y y}$

where subscripts are used to distinguish these solutions
from their counterparts in the ARMA zero pole model.

From equations (3.11) it follows that

$$b_{ZP} = b_{AP} + R_{y y} R_{y y} + (R_{u u} + R_{u y} R_{y y} R_{y u})^{-1}$$

$$[R_{u y} + b_{AP} - R_{u y}]$$

(3.12a)
Following a similar development, the left inverse relationships of equations (3.8) can be used to write

\[
\begin{align*}
\mathbf{a}_{ZP} &= \mathbf{a}_{AZ} + \mathbf{R}^{-1}_{u} \mathbf{R} + \mathbf{R}_{u} \mathbf{y} \mathbf{y}^{-1} \mathbf{R}_{uu}^{-1} \mathbf{R}_{u} \mathbf{y}^{-1} \mathbf{R}_{uy} \mathbf{y}^{-1} \mathbf{R}_{yy}^{-1} \\
\mathbf{b}_{ZP} &= (\mathbf{R}_{yy} - \mathbf{R}_{yu} \mathbf{u}^{-1} \mathbf{R}_{u}^{+} \mathbf{R}_{uy}^{+})^{-1} [\mathbf{R}_{yy} - \mathbf{R}_{yu} \mathbf{R}_{yy}^{-1} \mathbf{a}_{AZ}] \quad (3.12c) \\
\mathbf{a}_{ZP} &= (\mathbf{R}_{u} \mathbf{u}^{-1} \mathbf{R}_{uy}^{+} \mathbf{R}_{yy}^{-1} \mathbf{R}_{uy}^{+} \mathbf{R}_{yu}^{-1} \mathbf{R}_{u}^{+})^{-1} [\mathbf{R}_{u} \mathbf{u}^{-1} \mathbf{R}_{u} \mathbf{y}^{-1} \mathbf{R}_{uy} \mathbf{y}^{-1} \mathbf{R}_{uy} \mathbf{B}_{AP}] \quad (3.12d)
\end{align*}
\]

Equations (3.12) are termed the "Zero Pole Model Transition Formulas" and specify the relationships between the various models. It is interesting to note that equations (3.12a) and (3.12b) take the form of a linear observer with a new estimate of the solution given by the old estimate plus a gain times an error term. To gain some insight into the functioning of these formulas, consider the form of \( \mathbf{R}_{yy}(n) \) and \( \mathbf{R}_{uy}(n) \) for the linear system described by the transfer function of equation (3.1c)

\[
\mathbf{R}_{yy}(-n) = \sum_{L=1}^{N} b(i) \mathbf{R}_{yy}(-i-n) + \sum_{L=0}^{M} a(i) \mathbf{R}_{uy}(-i-n) \quad (3.13a)
\]
\[ R_{uy}(-n) = \sum_{L=1}^{N} b(i) R_{uy}(-i-n) + \sum_{i=0}^{M} a(i) R_{uu}(-i-n) \]  

(3.13b)

Assuming that \( N=N \) and \( M=M \), and writing equation (3.13a) for \(-N\leq n<1\) and equation (3.13b) for \(-M<n<0\) results in

\[ R_{yy} b + R_{yu}^{-} a^{+} = R_{yy} \]  

(3.14a)

\[ R_{u+y} b + R_{u+u}^{-} a^{+} = R_{u+y} \]  

(3.14b)

These constraints on the system input and output auto and cross correlation coefficients are the ARMA modeling equations of (3.5a) with the model coefficients replaced by the system parameters. In AR modeling, \( b_{ap} \) is set to satisfy the constraints of equation (3.14a) with the assumption that \( a^{+} \) is zero. The error term in the model transition formula (3.12a) then checks this solution to see if it also satisfies the constraints of equation (3.14b) still assuming that \( a^{+} \) is zero.

\[ \text{error} = R_{u+y} b_{ap} - R_{u+y} \]  

(3.15)
If this error is zero and the constraints of equation (3.14b) are satisfied, equation (3.12a) sets \( b_{zp} = b_{AP} \) and equation (3.12d) sets \( a_+ \) to zero. If however the error is nonzero, (3.12a) adjusts \( b_{AP} \) in proportion to the error to obtain \( b_{zp} \) and (3.12d) then provides a nonzero \( a_{zp}^+ \). Thus equations (3.12a) and (3.12d) are complementary, specifying the ARMA or zero pole model solution of order M over N when given the N-th order AR or all pole model solution.

In like manner, equations (3.12b) and (3.12c) give the ARMA model solution of order M over N when given the M-th order MA or all zero model solution. The all zero solution is obtained from equation (3.14b) assuming that \( b \) is zero. Equation (3.12b) checks this solution against the constraints imposed by equation (3.14a) with the same assumptions and adjusts \( a_+ \) appropriately to determine \( a_{zp}^+ \). Equation (3.12c) then sets \( b_{zp} \), completing the zero pole model solution.

2. Modeling Input Signal Requirements

Another aspect of the modeling problem that must be considered is that of system identifiability. If, from the available measurements of signals, a model can be obtained that accurately represents the system's operation, the system is considered identifiable. The two issues that arise therefore are the measurement requirements (which signals must be measured) and requirements on the input signal used to excite the system during the modeling process. In the equation error formulation of the ARMA
model, both existing signals (system input and output) must be observed (or at least a knowledge of their auto and cross correlation functions must be available). Most discussions of input signal requirements for identifiability simply state that the system can be identified if the input signal is sufficiently rich, persistent or exciting, eg [Ref. 19].

To explore the question of input signal requirements further, consider the mean square equation error cost function being minimized. Assuming that the equation error signal is ergodic and has finite energy its mean square value is obtained via time averaging as

$$E_2 = \varepsilon(e(k)^2) = \sum_{n=-\infty}^{\infty} e(n)^2$$  \hspace{1cm} (3.16)

Applying Parseval's relation this becomes

$$E_2 = \sum_{n=-\infty}^{\infty} e(n)^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E}(e^{j\theta})\mathbb{E}^*(e^{j\theta}) \, d\theta$$  \hspace{1cm} (3.17)

where $\theta = \omega T$ and $\ast$ indicates the complex conjugate. The equation error is represented in the transform domain as

$$E(z) = [B(z)H(z) - A(z)]U(z)$$  \hspace{1cm} (3.18)

and using this in equation (3.17), the cost function becomes

$$E_2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |[B(e^{j\omega})H(e^{j\omega}) - A(e^{j\omega})] |^2 \cdot |U(e^{j\omega})|^2 \, d\theta$$  \hspace{1cm} (3.19)
showing that the power spectrum of the input signal acts as a frequency dependent weighting function on a transfer function error term. Therefore to identify the system equally well at all frequencies, the input must have a flat spectrum as will be the case for a white noise input or an impulse function input. Otherwise, the model transfer function will only be matched to the system transfer function over the range of frequencies where the input signal has significant power.

As an example consider the equation error ARMA model for a fourth order system driven by a single sine wave input at a frequency of π/3. According to equation (3.19), the model transfer function will only be required to match the system at this single frequency, and to accomplish this, only a first order model is needed. Any increase in model order above first order therefore should have no effect. Figure 3.3a shows a comparison of the magnitude spectrum of the fourth order system and its first order ARMA model obtained using the sinusoidal input and as anticipated they match at the frequency π/3 (coincidentally they also match at one other frequency as well). Figure 3.3b shows the same comparison but with a fourth order ARMA model. It is clear that increasing the model order failed to improve its accuracy and that the model accurately represents the system only at the frequency of the input signal and, by coincidence, at one other frequency.
Figure 3.3. ARMA models for a system driven by a single sinusoid.
It should be noted that this type of analysis to determine input signal requirements could be applied to the AR and MA models as well resulting in the same conclusions.

B. A RECURSIVE IN ORDER SOLUTION FOR THE ARMA MODEL

Since the equation error formulation for the ARMA model is a generalization of similar formulations for AR and MA modeling, it is reasonable to assume that a Levinson-type algorithm could be devised to obtain the ARMA solution recursively in order, and that from that algorithm, lattice filter methods applicable to the ARMA modeling problem could be derived. Attempts to develop such an algorithm directly for the ARMA modeling equation (3.5a), however, fail to provide useful results. The first problem that arises is in deciding which model order to make recursive; the order of the numerator polynomial, the order of the denominator polynomial or both. If it is assumed that the numerator and denominator are of equal orders \( M=N \), the ARMA modeling equations become as shown in equation (3.20). However, efforts to develop a Levinson-type algorithm for this system of equations, where the numerator and denominator polynomials of the model are incremented simultaneously to obtain recursive in order solutions, are still frustrated by the presence of the \((N+1)\)-st row and column in equation (3.20). This arises because the numerator coefficient vector \( \mathbf{a}^* \) is a \((N+1)\)-vector while the denominator coefficient vector \( \mathbf{b} \) is a \(N\)-vector. If it is further assumed that the coefficient
\[
\begin{bmatrix}
R_{yy}(0) & R_{yy}(-1) & \ldots & R_{yy}(1-N) & R_{yu}(1) & R_{yu}(0) & \ldots & R_{yu}(1-N) \\
R_{yy}(1) & R_{yy}(0) & \ldots & R_{yy}(2-N) & R_{yu}(2) & R_{yu}(1) & \ldots & R_{yu}(2-N) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
R_{yy}(N-1) & R_{yy}(N-2) & \ldots & R_{yy}(-N) & R_{yu}(0) & R_{yu}(-1) & \ldots & R_{yu}(0) \\
\end{bmatrix}
\begin{bmatrix}
b(1) \\
b(2) \\
\vdots \\
b(N) \\
\end{bmatrix}
= 
\begin{bmatrix}
R_{yy}(1) \\
R_{yy}(2) \\
\vdots \\
R_{yy}(N) \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
R_{uy}(1) & R_{uy}(2) & \ldots & R_{uy}(1-N) \\
R_{uy}(1) & R_{uy}(0) & \ldots & R_{uy}(1-N) \\
\vdots & \vdots & \vdots & \vdots \\
R_{uy}(N-1) & R_{uy}(N-2) & \ldots & R_{uy}(0) \\
\end{bmatrix}
\begin{bmatrix}
a(0) \\
a(1) \\
\vdots \\
a(N) \\
\end{bmatrix}
= 
\begin{bmatrix}
R_{uy}(0) \\
R_{uy}(1) \\
\vdots \\
R_{uy}(N) \\
\end{bmatrix}
\]

(3.20)
a(0) is known in advance or can be estimated in some other fashion, equation (3.20) for the solution for the remaining 2N coefficients becomes as written in equation (3.21).

\[
\begin{bmatrix}
R_{yy}(0) & \ldots & R_{yy}(1-N) \\
\vdots & & \vdots \\
R_{yy}(N-1) & \ldots & R_{yy}(0)
\end{bmatrix}
\begin{bmatrix}
R_{yu}(0) & \ldots & R_{yu}(1-N) \\
\vdots & & \vdots \\
R_{yu}(N-1) & \ldots & R_{yu}(0)
\end{bmatrix}
\begin{bmatrix}
b(1) \\
b(N)
\end{bmatrix}
= 
\begin{bmatrix}
R_{yy}(1) \\
R_{yu}(1) \\
R_{uu}(1)
\end{bmatrix}
\begin{bmatrix}
a(0) \\
a(1) \\
a(N)
\end{bmatrix}
\]

The coefficient a(0) essentially has the role of a gain for the model, and a method for estimating it after all the other coefficients are obtained (as was done in AR modeling) will be discussed later. Equation (3.21) can be written as

\[
\begin{bmatrix}
R_{yy} & R_{yu} \\
R_{uy} & R_{uu}
\end{bmatrix}
\begin{bmatrix}
b \\
a
\end{bmatrix}
= 
\begin{bmatrix}
\Sigma_{yy} & \Sigma_{yu} \\
\Sigma_{uy} & \Sigma_{uu}
\end{bmatrix}
\begin{bmatrix}
1 \\
-a(0)
\end{bmatrix}
\]

(3.22)
Now consider the form of a two channel autoregressive model where the two input channels are \( y(k) \) and \( u(k) \); that is

\[
x_1(k) = y(k)
\]

\[
x_2(k) = u(k)
\]

Using equations (A.4a) and (A.7b) the two channel AR modeling equations are

\[
\begin{bmatrix}
R_{yy} & R_{yu} \\
R_{uy} & R_{uu}
\end{bmatrix}
\begin{bmatrix}
d_{11} & d_{12} \\
d_{21} & d_{22}
\end{bmatrix}
= 
\begin{bmatrix}
R_{yy} & R_{yu} \\
R_{uy} & R_{uu}
\end{bmatrix}
\tag{3.23}
\]

and comparing equations (3.22) and (3.23) it is clear that with the exception of the gain term \( a(0) \), the ARMA modeling solution can be obtained from the two channel AR solution of (3.23). Furthermore equation (3.23) can be solved independently of the gain term via the Levinson algorithm as shown in Appendix A or via the multichannel lattice methods developed in the previous chapter. Then all that remains to complete the solution for the ARMA model is to estimate the gain term \( a(0) \) and solve for the other model coefficients using

\[
\begin{bmatrix}
b \\
\hat{a}
\end{bmatrix}
= 
\begin{bmatrix}
d_{11} & d_{12} \\
d_{21} & d_{22}
\end{bmatrix}
\begin{bmatrix}
1 \\
-a(0)
\end{bmatrix}
\tag{3.24}
\]
From this it follows that the transfer functions $A(z)$ and $B(z)$ for the ARMA prediction error analysis model can be related to the two-channel AR prediction error matrix polynomial transfer function by

$$
\begin{bmatrix}
B(z) \\
-A(z)
\end{bmatrix}
= \begin{bmatrix} 1 \\ -a(0) \end{bmatrix} \begin{bmatrix} 1 \\ [I - D(z)] \end{bmatrix}
$$

(3.25)

It is also easy to relate the ARMA equation error to the two-channel AR prediction error vector. Using equations (A.3) and (A.5), the two-channel AR prediction error vector is written as

$$
e(k)^T = \begin{bmatrix} y(k)^T \\ u(k)^T \end{bmatrix} = \begin{bmatrix} y(k)^T \\ u(k)^T \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix}
$$

(3.26)

$$
= \begin{bmatrix} e_y(k) \\ e_u(k) \end{bmatrix}
$$

Defining

$$
\Psi = \begin{bmatrix} 1 \\ -a(0) \end{bmatrix}
$$

(3.27a)
and postmultiplying equation (3.26) by $\psi$ and using (3.24) it follows that

$$e(k)^T \psi = y(k) - a(0)u(k) - [y(k)^T \ u(k)^T] \begin{bmatrix} b \\ -a \end{bmatrix} \quad (3.27b)$$

But from equation (3.4a) this is exactly the ARMA model equation error so that

$$e_0(k) = e(k)^T \psi \quad (3.27c)$$

and

$$\sigma\{e_0(k)^2\} = \psi^T P \psi \quad (3.27d)$$

where $P$ is the forward prediction error covariance matrix for the 2 channel AR model. Equation (3.27d) also provides a means of estimating the gain term $a(0)$ after the two channel AR solution has been obtained by setting it to minimize the mean square value of equation error resulting in

$$a(0) = \frac{\sigma\{e_0(k)e_\psi(k)\}}{\sigma\{e_\psi(k)^2\}} \quad (3.28)$$

and completing the ARMA model solution.
The portion of the ARMA model solution in equation (3.24) given by the two channel AR solution can be found recursively in order using either the Levinson algorithm or the lattice filter techniques. If the desired ARMA model order is not known in advance, a model gain term a(0) can be estimated for each order two channel AR solution to find the ARMA model of corresponding order along with its MSE. In this fashion, the entire family of ARMA models for the system from order zero to order N, along with their mean square errors are obtained. If, on the other hand, the desired ARMA model order is known apriori, the gain term need not be calculated at each stage. Only one gain term must be calculated to obtain the ARMA solution after the appropriate order two channel AR solution has been found.

It has already been shown that to fully identify the system using the equation error ARMA formulation, the input signal must have a flat spectrum as in the case of white noise. When white noise is used as the system input u(k), simplifications emerge in the solution of the two channel AR model via the Levinson algorithm or lattice methods. For a white sequence with variance $\sigma_u^2$,

$$ R_{uu}(n) = \begin{cases} 0 & , n \neq 0 \\ \sigma_u^2 & , n = 0 \end{cases} \quad (3.29a) $$
and

\[
R_{uy}(n) = \begin{cases} 
0 & ; n < 0 \\
\sigma_u^2 h(n) & ; n \geq 0 
\end{cases} 
\] (3.29b)

where \(h(n)\) is the sampled impulse response of the system.

Consider equation (A.16c) for \(K(1)\).

\[
K(1) = \begin{bmatrix}
R_{yy}(0) & R_{yu}(0) \\
R_{uy}(0) & R_{uu}(0)
\end{bmatrix}^{-1} \begin{bmatrix}
R_{yy}(1) & 0 \\
R_{uy}(0) & 0
\end{bmatrix} 
\] (3.29c)

This shows that \(k_{12}^{(1)}\) and \(k_{22}^{(1)}\) are zero and furthermore since \(R_{yu}^{(n)}\) and \(R_{uu}^{(n)}\) are zero for all \(n\) it is seen that

\[
k_{12}^{(n)} = k_{22}^{(n)} = 0 
\] (3.29d)

for all \(n\) as well. This can readily be understood by considering the role of these two coefficients at each stage in the AR prediction of \(y(k)\) and \(u(k)\). \(k_{12}\) and \(k_{22}\) are the coefficients used in trying to predict \(u(k)\) from past values of \(y(k)\) and \(u(k)\), and when \(u(k)\) is a white sequence it cannot be predicted, forcing these coefficients to be zero.

No such simplifications occur in the backward prediction problem (and therefore in the \(K\) matrices) since even for a white \(u(k)\), a backward prediction of \(u(k-n)\) from subsequent values of \(u\) and \(y\) is possible. This is because in
general, a linear dependence of $y(k)$ upon past and present values of $u(k)$ can occur (and certainly will occur when the relationship between $y(k)$ and $u(k)$ is described by an ARMA model).

As a result of these simplifications, it is seen from equation (A.19a) that the polynomials $d_{12}(z)$ and $d_{22}(z)$ are zero when $u(k)$ is white and the ARMA model is given by

$$
\begin{bmatrix}
B(z) \\
-A(z)
\end{bmatrix} =
\begin{bmatrix}
1 - d_{11}(z) & 0 \\
-d_{21}(z) & 1
\end{bmatrix}
\begin{bmatrix}
1 \\
-a(0)
\end{bmatrix}
$$

(3.30)

In this special case it follows that

$$
B(z) = \det[I - D(z)]
$$

(3.31)

and therefore stability of the ARMA model and of the two channel AR model are equivalent. In general, however, no such connection exists for arbitrary input signals. Furthermore, even when the system input $u(k)$ is white, solutions for $k_{12}$ and $k_{22}$ will not in general be exactly zero since the required correlations are usually not known and must be estimated.

This development showing that the ARMA model solution can be obtained from a two channel autoregressive model depends on two assumptions:

1) The numerator and denominator polynomial orders in the model transfer function are assumed to be the
same and are incremented simultaneously to build up the desired solution recursively in order;

2) It is assumed that the coefficient of \( z \) to the zero power \( (a(0)) \) in the numerator polynomial of the model is either known in advance, or that another means of estimating it can be found so that it need not be estimated directly in the modeling equations of (3.20).

The second assumption causes no concern since the two channel autoregressive solution is obtained independently of \( a(0) \), and given that solution, it has been shown that \( a(0) \) can indeed be estimated in another fashion in equation (3.28).

The first assumption however, warrants further consideration since it seems somewhat restrictive (at first) to require that the numerator and denominator polynomials of the model have the same order when in fact, the system being modeled may have different order numerator and denominator polynomials. To see that this assumption is not restrictive in general, consider what is occurring as the model is built up recursively in order. At each model order \( n \), the procedure finds the best \( n \)-th order model (with \( n \) zeros and \( n \) poles) in a minimum mean square equation error sense. Making the model numerator and denominator orders different (or equivalently, forcing some of the coefficients to zero in the model where the orders are the same) places a priori constraints on the model, forcing some of the poles or zeros to the origin in the \( z \)-plane, rather than allowing the
model to place them at will to minimize the cost function. As an example, consider the process of obtaining an ARMA model for a system given by

\[
H(z) = \frac{\sum_{n=0}^{2} a(n) z^{-n}}{1 - \sum_{n=1}^{3} b(n) z^{-n}}
\]

where two of the system's four zeros actually occur at the origin of the z plane. Constraining any of the model zeros to the origin at orders one, two or three will result in a model with higher cost (MSE) than if they were not constrained. Even at order three, a model without constraints can be expected to use the "extra" zero to help in approximating the effects of the system's fourth pole as shown in equation (2.13) yielding a lower cost and more accurate model than would result if one zero were forced to the origin. Only at order four are such constraints reasonable but even then, they are not necessary since the modeling procedure itself should recognize that the best fourth order model will have two zeros at z=0.

Therefore, it is seen that assuming equal orders for the model numerator and denominator is entirely reasonable as a general approach in obtaining MMSE models for unknown systems or even reduced order models for known systems. When it is known in advance that the best MMSE model for a system has zeros at z=0, imposing such a constraint on the model can
reduce the computational complexity of obtaining the solution, but even here the assumption of equal orders is not restrictive.

C. LATTICE FORM ARMA MODELING

In chapter two it was shown that the lattice structure of Figure 2.7 could be applied to solve the multichannel AR modeling problem in terms of the reflection coefficient matrices given by equations (2.43). Since the ARMA model solution can be obtained from a two channel AR solution with \( y(k) \) and \( u(k) \) as the input channels, only the structure described by equation (3.27c) need be added to a two channel AR lattice to obtain the lattice form of the ARMA analysis model. It is interesting to look at the exact structure of this lattice model as shown in Figure 3.4 for a second order case. This structure is seen as a lattice interconnection of two single channel AR lattices operating on the input signals \( y(k) \) and \( u(k) \). The coefficients on the main diagonals of the \( K \) and \( \bar{K} \) matrices specify the single channel lattices while the off diagonal elements specify the interconnections. (This will also be the case for extensions to lattices with any number of channels.)

The ARMA synthesis model implementing the transfer function \( A(z)/B(z) \) can also be put in lattice form. The forward prediction in the two channel AR analysis lattice is described by
Figure 3.4. Lattice analysis structure for ARMA modeling.
\[
\begin{bmatrix}
  e_y(k) \\
  e_u(k)
\end{bmatrix}^{-(n+1)} =
\begin{bmatrix}
  e_y(k) \\
  e_u(k)
\end{bmatrix}^{-n} -
\begin{bmatrix}
  k_{11} & k_{21} \\
  k_{12} & k_{22}
\end{bmatrix} \begin{bmatrix}
  e_y(k-1) \\
  e_u(k-1)
\end{bmatrix}^{-n}
\]

(3.32)

and the equation for \( e_k(k)^{(n+1)} \) can be rewritten as

\[
e_y(k)^{(n)} = e_y(k)^{(n+1)} + k_{11} e_y(k-1)^{(n)} + k_{21} e_u(k-1)^{(n)}
\]

(3.33)

Equation (3.33) along with the equation of \( e_u \) in (3.32), and
the equations for the backward prediction (2.45b) describe
the structure shown in Figure 3.5 for a second order case.
To provide the required input at \( e_y(k)^{(2)} \), recognize from
equation (3.27c) that

\[
e_y(k)^{(2)} = e_y(k) + a(0) e_u(k)^{(2)}
\]

(3.34a)

If the ARMA model is an accurate representation of the
system, the equation error \( e_y(k) \) will be quite small (ideally
zero) so that in general for the \( N \)-th order case

\[
e_y(k)^{(N)} = a(0) e_u(k)^{(N)}
\]

(3.34b)

This is indicated by the dashed feedback path in Figure 3.5.
Figure 3.5. Lattice synthesis structure for ARMA modeling.
D. ARMA MODEL SIMULATIONS; BATCH PROCESSING

ARMA modeling procedures for linear systems using both the lattice filter method and a brute force matrix inversion with equation (3.20) have been implemented and the results of these two approaches have been compared in over a thousand model simulations of more than thirty different linear systems. The experimental results which follow are a representative sampling of these simulations.

In the lattice filter method, equations (2.49) were used to calculate the forward and backward reflection coefficient matrices, with time averages over a specified interval used to estimate the required correlations in $\hat{P}^{(0)} = \overline{P}^{(0)}$ and $\Delta^{(n)}$ for $0 \leq n \leq N-1$. Equations (A.15), (A.16a) and (A.16b) were used to obtain the two channel AR model coefficients from the $K$ and $\bar{K}$ matrices. and with the gain calculated in equation (3.28), equation (3.24) was used to obtain the desired ARMA model coefficients. Equations (A.18) were used to update the forward and backward prediction error covariance matrices from one lattice stage to the next. Figure 3.6a provides a flow diagram of the procedure.

In the brute force matrix inversion method with equation (3.20) time averaging was again employed to estimate the required correlation coefficients. A rectangular window was applied to the data, however, to retain the even symmetry of the autocorrelation function in these estimates. The ARMA model coefficients were then obtained using a general purpose library subroutine (which employed gaussian
elimination) to solve equation (3.20). Figure 3.6b provides a flow diagram of this procedure.

In both cases zero mean, unit variance gaussian white noise was used as the system input. In the simulation results that follow, a description of each system discussed (transfer function coefficients, zero locations and pole locations) is listed in tabular form and root locations in the z plane as well as transfer function magnitudes are plotted for the system, and various models obtained for it. In each case, models were obtained for averaging intervals of 200, 500, 1000, 2000 and 4000 data points. Only the results for the two extremes of 200 and 4000 points are included here.

The first system considered has a second order numerator in z inverse and a fourth order denominator, and its characteristics are listed in Table 3.1. Figure 3.7 shows a comparison of the root locations and transfer function magnitude of this system with those of fourth order lattice filter and brute force models obtained when correlations were estimated by averaging over 200 samples. Figure 3.8 provides the same comparison for a longer averaging interval of 4000 samples of data. While both methods perform comparably with the longer averaging interval, the lattice method produces a far more accurate model with the short averaging interval. It is also interesting to compare the performance of the two methods when the system is overmodeled; that is, when the model order is increased beyond that of
Figure 3.6a. Flow diagram of batch processing lattice solution for all ARMA models orders 1 to N.
Estimate $R_{yy}(0)$, $R_{uu}(0)$ and $R_{yu}(0)$.

Estimate $R_{yy}(n) = R_{yy}(-n)$, $R_{uu}(n) = R_{uu}(-n)$, $R_{yu}(n) = R_{uy}(n)$ from windowed data.

Calculate n-th order ARMA coefficients and MSE from equations (3.20) and (3.5b).

Figure 3.6b. Flow diagram of batch processing brute force solution for all ARMA models of order 1 to N.
TABLE 3.1
SYSTEM A
TRANSFER FUNCTION COEFFICIENTS

<table>
<thead>
<tr>
<th>NUMERATOR</th>
<th>DENOMINATOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A(0)$ = 0.25000</td>
<td>$B(1)$ = 1.14000</td>
</tr>
<tr>
<td>$A(1)$ = 0.35000</td>
<td>$B(2)$ = -1.45490</td>
</tr>
<tr>
<td>$A(2)$ = 0.24500</td>
<td>$B(3)$ = 0.88490</td>
</tr>
<tr>
<td>$A(3)$ = 0.0</td>
<td>$B(4)$ = -0.40745</td>
</tr>
<tr>
<td>$A(4)$ = 0.0</td>
<td></td>
</tr>
</tbody>
</table>

ROOT LOCATIONS

<table>
<thead>
<tr>
<th>ZEROS</th>
<th>POLYES</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE</td>
<td>IM</td>
</tr>
<tr>
<td>-0.70000</td>
<td>0.70000</td>
</tr>
<tr>
<td>-0.70000</td>
<td>-0.70000</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Figure 3.7

(a) $|H(\phi)|_{dB}$ vs $\phi$ - SYSTEM
- 4th ORDER LATTECE MODEL
200 POINT AVERAGES

(b) $|H(\phi)|_{dB}$ vs $\phi$ - SYSTEM
- 4TH ORDER BRUTE FORCE MODEL
200 POINT AVERAGES
Figure 3.7 cont
Figure 3.8
Figure 3.8 con't
the system. Figures 3.9 and 3.10 show such a comparison when sixth order models are obtained for this fourth order system. Ideally, the extra zeros and poles should lie at the origin in the z plane making the additional transfer function coefficients zero. Figure 3.9 shows that for a 200 point averaging interval, the lattice locates the extra roots in the vicinity of the origin while the brute force model does not. When the averaging interval is increased to 4000 points, the extra roots of the lattice model move in toward the origin while those of the brute force model do not. Instead, a zero pole cancellation at some arbitrary location occurs in the brute force model. In all cases investigated during this effort, the lattice method clustered the extra roots in the vicinity of the origin and as the averaging interval was increased to take in more data, these roots were consistently moved in closer to the origin. This property is further evidenced by the plot of mean square equation error as a function of model order shown in Figure 3.11 for a 200 point averaging interval. The MSE for the lattice model flattens out at order four indicating that further increases in model order fail to increase its accuracy. Meanwhile the MSE for the brute force model continues to decrease beyond fourth order as it uses the additional roots to reduce modeling errors caused by inaccuracies in the fourth order model.
Figure 3.9
Figure 3.9 con't
Figure 3.10
Figure 3.10 con't
Figure 3.11. Mean square value of equation error (as a percentage of the mean square value of system output) vs model order for lattice and brute force models of system A with 200 point averages.
To investigate the ability of these modeling methods to distinguish roots located near one another in the z plane, a pair of zeros were added to the previous system in close proximity to one of the pole pairs. The characteristics of this system are listed in Table 3.2. Figures 3.12 and 3.13 show the lattice method and brute force modeling results for 200 and 4000 point averaging intervals. With data over only 200 sampling instants, neither method is able to accurately model the effects of the adjacent roots. When the averaging interval is increased, the lattice correctly models the plant while the brute force method does not, and even results in an unstable model. (Figure 3.13b comparing the transfer function magnitudes has been plotted in spite of the model instability.)

To investigate the ability of these zero pole modeling methods to model systems that are actually all zero or all pole, the systems listed in Tables 3.3 and 3.4 were used. The modeling results are shown in Figures 3.14, 3.15, 3.16 and 3.17.

The conclusions drawn from the results of this experimental study are as follows:

1) For short data lengths, the lattice filter method provides more accurate models than does the brute force modeling method.

2) When the system is overmodeled using the lattice method, the excess roots are clustered about the origin in the z plane and as the averaging interval is increased and more data taken in, these roots move
### TABLE 3.2

**SYSTEM B**

**TRANSFER FUNCTION COEFFICIENTS**

<table>
<thead>
<tr>
<th>Numerator A((c))</th>
<th>DENOMINATOR B((1))</th>
<th>DENOMINATOR B((2))</th>
<th>DENOMINATOR B((3))</th>
<th>DENOMINATOR B((4))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00000</td>
<td>1.14000</td>
<td>-1.45490</td>
<td>0.88490</td>
<td>-0.40745</td>
</tr>
<tr>
<td>1.34000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.79940</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.20400</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.88533</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ROOT LOCATIONS**

<table>
<thead>
<tr>
<th>Zeros</th>
<th>Poles</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE</td>
<td>IM</td>
</tr>
<tr>
<td>-0.70000</td>
<td>0.70000</td>
</tr>
<tr>
<td>-0.70000</td>
<td>-0.70000</td>
</tr>
<tr>
<td>0.03000</td>
<td>0.95000</td>
</tr>
<tr>
<td>0.93000</td>
<td>-0.95000</td>
</tr>
</tbody>
</table>
Figure 3.12
Figure 3.12 cont
Figure 3.13
Figure 3.13 con't
### TABLE 3.3

**SYSTEM C**

**TRANSFER FUNCTION COEFFICIENTS**

<table>
<thead>
<tr>
<th>NUMERATOR</th>
<th>DENOMINATOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A(0) = 1.00000$</td>
<td>$B(0) = 0.0$</td>
</tr>
<tr>
<td>$A(1) = -1.50000$</td>
<td>$B(1) = 0.6$</td>
</tr>
<tr>
<td>$A(2) = 0.62500$</td>
<td>$B(2) = 0.0$</td>
</tr>
</tbody>
</table>

**ROOT LOCATIONS**

<table>
<thead>
<tr>
<th>ZEROS</th>
<th>POLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE</td>
<td>IM</td>
</tr>
<tr>
<td>0.75000</td>
<td>0.25000</td>
</tr>
<tr>
<td>0.75000</td>
<td>-0.25000</td>
</tr>
</tbody>
</table>

### TABLE 3.4

**SYSTEM D**

**TRANSFER FUNCTION COEFFICIENTS**

<table>
<thead>
<tr>
<th>NUMERATOR</th>
<th>DENOMINATOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A(0) = 1.00000$</td>
<td>$B(0) = 1.50000$</td>
</tr>
<tr>
<td>$A(1) = 0.0$</td>
<td>$B(1) = 0.62500$</td>
</tr>
<tr>
<td>$A(2) = 0.0$</td>
<td>$B(2) = 0.0$</td>
</tr>
</tbody>
</table>

**ROOT LOCATIONS**

<table>
<thead>
<tr>
<th>ZEROS</th>
<th>POLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE</td>
<td>IM</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.75000</td>
<td>0.25000</td>
</tr>
<tr>
<td>0.75000</td>
<td>-0.25000</td>
</tr>
</tbody>
</table>
Figure 3.14
Figure 3.14 cont
Figure 3.15
Figure 3.15 con't
Figure 3.16
Figure 3.16 cont
Figure 3.17
Figure 3.17 con't
consistently in toward the origin. This also forces the excess transfer function coefficients and reflection coefficients to be very small (ideally zero), clearly indicating that the model order is higher than necessary. The brute force method, on the other hand, scatters the excess roots throughout the $z$ plane and produces cancellations of the excess zeros and poles. This results in nonzero values for the excess transfer function coefficients.

3) The MSE as a function of model order is generally better behaved for the lattice method than for the brute force method, decreasing rapidly until the correct model order is reached and then failing to decrease substantially as the model order is increased further.

Two qualitative explanations for the improved performance of the lattice filter modeling method are offered. The first is that the data is not windowed in the lattice method while a window that is nonzero only over the span of the averaging interval must be used in the brute force method to maintain even symmetry in the autocorrelation function estimates. The effects of this difference between the two modeling methods should be most noticeable for short data lengths, and become less evident as the length of the averaging interval is increased. The second possible cause for the lattice method's better performance is that the actual output sequences of the $n$-th order lattice are used to calculate the coefficients at the $(n+1)$-st order stage. In this manner,
modeling errors that have occurred in the \( n \)-th order lattice can be compensated for to some extent in the \((n+1)\)-st order stage. No similar phenomenon is evident in the brute force modeling approach. Consider the differences between the Levinson algorithm (which is equivalent to the matrix inversion method) of equations (A.16) and the lattice method given by equations (2.49). Both methods calculate the corrections that must be made to the \( n \)-th order model to obtain the \((n+1)\)-st order model in terms of \( K^{(n+1)} \) and \( \hat{K}^{(n+1)} \) and theoretically,

\[
\varepsilon \{ e^{(n)}(k) e^{(n)}(k-1)^T \} = R_{xx}^{(n+1)} - \rho^{(n)} \rho^{(n)} \hat{D}^{(n)}
\]

When correlations are estimated by averaging over finite intervals however this equality will not in general be satisfied making the two methods different. The Levinson algorithm will estimate the correction terms to be added to the optimum \( n \)-th order model while the lattice method estimates the correction terms to be added to the estimated \( n \)-th order model actually obtained.

The improved performance of the lattice method is not achieved without cost, however. The method is made computationally expensive by the need to store the system input and output sequences and the lattice prediction error sequences and pass them through successive stages of the lattice as it is built up in order. The computational complexity of the two methods is compared in Table 3.5.
Table 3.5. Computational requirements for batch processing ARMA modeling of an N-th order system using P samples of the system input and output.

| Number of Correlation Estimates Required | \(4N+3\) averaged over \(P\) data points | \(4N+3\) \\
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix Inversions</td>
<td>(1)-dimension (2N+1)</td>
<td>(2N)-dimension (2) (1)-dimension (1)</td>
</tr>
<tr>
<td>Data Storage Requirements</td>
<td>N.A.</td>
<td>2P samples</td>
</tr>
<tr>
<td>Computations To Pass Data Through The Filter</td>
<td>N.A.</td>
<td>8NP multiplications (4NP) additions</td>
</tr>
</tbody>
</table>
E. ADAPTIVE LATTICE ARMA MODELING

In addition to the batch processing method described in the previous section, the lattice ARMA analysis model can be implemented adaptively as well. The adaptive lattice solution for the multichannel AR lattice, which solves most of the ARMA modeling problem, has already been described in chapter II. To make the lattice ARMA model adaptive, only an adaptive solution for the gain term \( a(0) \) need be added. To avoid ambiguity, the time varying adaptive estimate of this term at time \( k \) is denoted by \( a_0(k) \). Applying an LMS adaptive algorithm it follows that

\[
a_0(k+1) = a_0(k) - \mu_0 \nabla(k)
\]

(3.35)

and using equation (3.27d) to form an instantaneous estimate of the gradient yields

\[
\hat{\nabla}(k) = -2 \ e_u(k) [e_y(k) - a_0(k) e_u(k)]
\]

(3.36)

\[
= -2 \ e_u(k) e_0(k)
\]

so that

\[
a_0(k+1) = a_0(k) + 2\mu_0 \ e_u(k) e_0(k)
\]

(3.37)
Here it is clear that

1) \( e_0(k) \) is analogous to the error signal.
2) \( e_u(k) \) is analogous to the input signal.
3) \( e_y(k) \) is analogous to the desired signal.

so that for stability \( \mu_0 \) must satisfy

\[
0 < \mu_0 < \frac{1}{\mathbb{E}\{e_u(k)^2\}} \tag{3.38}
\]

Once again, however, the mean square value of \( e_u(k) \) may vary from stage to stage and over time as the two channel AR lattice adapts making it appropriate to apply relations similar to equations (2.58) and (2.59)

\[
\mu_0(k) = \frac{\alpha}{\sigma_0(k)} \tag{3.39a}
\]

\[
\sigma_0(k) = (1-\alpha) \sigma_0(k-1) + \alpha e_u(k)^2 \tag{3.39b}
\]

where \( \alpha \) is the normalized adaptive step size and the dependence on order is implicit (a superscript \((n)\) could be used on \( a_0, \sigma_0, \psi \), and all the error terms in equations (3.35) through (3.39) to explicitly denote their dependence on the order of the solution).

This adaptive lattice ARMA modeling scheme was implemented and the results of its use in modeling system A described in Table 3.1 are presented here. A normalized adaptive step size of \( \alpha = .05 \) was used in the following simulations and
the results represent an ensemble average of one hundred trials. Unity variance white noise was used as the system input. A flow diagram of the procedure is shown in Figure 3.18.

Figure 3.19 shows a plot of the mean square equation error as a function of time for a fourth order model while it adapts and Figure 3.20 shows the behavior of one term in the K matrix, $k_{11}^{(n)}$, and one term in the $\overline{K}$ matrix, $\overline{k}_{11}^{(n)}$, at the first five lattice stages, $1 \leq n \leq 5$. These graphs of the $k_{11}$ terms clearly show the successive stage by stage manner in which the lattice model adapts. Figure 3.21 shows a comparison of the transfer function magnitude and root locations of the system with those of the fourth order adaptive model after 1000 and 2000 iterations. Figure 3.22 shows the same comparison in the overmodeled case when a sixth order model is obtained for this fourth order system.

While these results show that the adaptive solution performs well and is a viable alternative to the batch processing solution, Figures 3.22c and 3.22d show that one advantageous characteristic of the batch solution has not carried over. The excess roots in the overmodeled case are not tightly clustered in the vicinity of the origin indicating that the excess transfer function coefficients are not near zero. In examining Figure 3.20 it is also evident that the convergence of the reflection coefficients at the overmodeled lattice stage ($k_{11}^{(5)}$ and $\overline{k}_{11}^{(5)}$) towards their true value of zero is quite slow. This relatively slow tracking
Initialize the parameter and step size normalizing factors

Sample system input and output

Find current value of $\varepsilon(n)(n)$ and $\varepsilon(n)(k)$ for $0 \leq n \leq N$

Update the step size normalizing factors using equations (2.59)

Update the $K(n)$ and $\bar{K}(n)$ estimates using equations (2.55) and (2.58) for $1 \leq n \leq N$

Update $\alpha(n)$ using equations (3.37) and (3.39a) for $n$ equal to the desired order ARMA solution

If desired, calculate ARMA coefficients from $K$, $\bar{K}$ and $a_0$ estimates using equations (A.15), (A.16) and (3.24).

Figure 3.18. Flow diagram of adaptive lattice ARMA solution.
Figure 3.19

Mean square value of equation error for the fourth order adaptive lattice model as a function of time.
Figure 3.20a

$K(1,1)$ term from the forward reflection coefficient matrices at the first five stages of the adaptive lattice model as a function of time.
Figure 3.20b

$K(1,1)$ term from the backward reflection coefficient matrices at the first five stages of the adaptive lattice model as a function of time.
Figure 3.21
Figure 3.21 con't
Figure 3.22

143
Figure 3.22 con't
of overmodeled, zero valued parameters has been found to be a general characteristic of the adaptive lattice algorithm and has also been noted briefly by Morf. [Ref. 38]

To understand the reason for this behavior, consider what occurs as the overmodeled fifth stage in the previous example adapts. Initially, before the coefficients of the first four lattice stages converge to their optimum values, the prediction error sequences out of the fourth stage are large and suboptimum. These signals provide incorrect inputs to the fifth stage driving its parameter estimates to some values other than their optimum zero values. As the first four stages converge, the prediction error sequences going into the fifth stage get small and since they drive the gradient estimates, convergence back toward zero is quite slow in these parameter estimates.

The cost functions being minimized at the overmodeled fifth stage are the trace of $P^{(5)}$ and $\bar{P}^{(5)}$ given by

$$P^{(5)} = P^{(4)} - \Delta^{(4)} K^{(5)} - K^{(5)T} \Delta^{(4)T} + K^{(5)T} P^{(4)} K^{(5)}$$

$$\bar{P}^{(5)} = \bar{P}^{(4)} - \Delta^{(4)T} \bar{K}^{(5)} - \bar{K}^{(5)T} \Delta^{(4)} + \bar{K}^{(5)T} \bar{P}^{(4)} \bar{K}^{(5)}$$

Applying the results of Appendix F, the parabolic surfaces defined by these cost functions are described by the eigenvalues and eigenvectors of $P^{(4)}$ and $\bar{P}^{(4)}$, the prediction error covariance matrices at the fourth stage. Consider the forward
predictions. The actual system output is given by

\[ y(k) = \sum_{L=1}^{4} b(i) y(k-i) + \sum_{i=0}^{4} a(i) u(k-i) \tag{3.41} \]

and for a white input signal, the minimum errors in the fourth order two channel autoregressive predictions of \( y \) and \( u \) are

\[ e_y^{(4)}(k) = a(0) u(k) \tag{3.42a} \]

\[ e_u^{(4)}(k) = u(k) \tag{3.42b} \]

This results in an optimal prediction error covariance matrix given by

\[ P^{(4)} = \begin{bmatrix} a(0)^2 & a(0) \\ a(0) & 1 \end{bmatrix} R_{uu}(0) \tag{3.43} \]

with eigenvalues of 0 and 1+a(0)^2. Also in the case of the system described by Table 3.1 where \( a(3) = a(4) = 0 \), it is seen from equation (3.41) that a perfect backward two channel AR prediction of \( y(k-4) \) is possible resulting in an optimal backward prediction error covariance matrix of

\[ P^{(4)} = \begin{bmatrix} 0 & 0 \\ 0 & \epsilon \{ e_u^{(4)}(k) \} \end{bmatrix} \tag{3.44} \]
which also has a zero eigenvalue.

Since the ellipses obtained by passing a plane through the parabolic cost surface have axes whose half lengths given by $1/\sqrt{\lambda_i}$ and since $P^{(4)}$ and $\bar{P}^{(4)}$ each have one eigenvalue of zero, it is seen that the parabolic bowls along which $K^{(5)}$ and $\bar{K}^{(5)}$ adapt, degenerate toward infinitely long parabolic troughs as the first four stages converge toward their optimum values. This is responsible for the slow convergence of the overmodeled parameters back toward zero. To avoid this problem, some means of detecting this degeneration of the cost surface and then resetting the appropriate parameters to zero must be found and this certainly provides an interesting area for future study.

**F. A LATTICE APPROACH FOR MULTICHANNEL ARMA MODELING**

The lattice filter solution methods for the ARMA model can readily be generalized to multiple input multiple output ARMA models for linear systems. The equations for the multichannel ARMA model of the system shown in Figure 3.21 are developed in Appendix G with the solution for the model coefficients given by equation (G.7) repeated here for convenience.

\[
\begin{bmatrix}
R_{YY} & R_{Y+U^+} \\
R_{U^+Y} & R_{U^+U^+}
\end{bmatrix}
\begin{bmatrix}
B \\
A^+
\end{bmatrix}
= 
\begin{bmatrix}
R_{YY} \\
R_{U^+Y}
\end{bmatrix}
\]

(3.45)
This is clearly a generalization of the single channel ARMA equation error solution given by equation (3.20) and just as in the single channel case some assumptions are required to apply a Levinson type algorithm.

If it is assumed that all the $a_{ij}(0)$ are known or can be estimated in another manner, they can be incorporated into a matrix given by

$$ A_0 = \begin{bmatrix} a_{11}(0) & \cdots & a_{1Q_0}(0) \\ \vdots & \ddots & \vdots \\ a_{Q_11}(0) & \cdots & a_{Q_1Q_0}(0) \end{bmatrix} $$

(3.48)

and equation (3.45) becomes

$$ \begin{bmatrix} R_{YY} & R_{YU} \\ R_{UY} & R_{UU} \end{bmatrix} \begin{bmatrix} B \\ A \end{bmatrix} = \begin{bmatrix} R_{YY} & R_{YU} \\ R_{UY} & R_{UU} \end{bmatrix} \begin{bmatrix} I \\ -A_0 \end{bmatrix} $$

(3.47)
This is similar in form to the equations for the solution of a \((Q_i + Q_0)\) channel autoregression with input channels \(y_1(k), \ldots, y_Q(k), u_1(k), \ldots, u_{Q_0}(k)\) so that the multichannel ARMA solution is related to the multichannel AR solution by

\[
\begin{bmatrix}
  B \\
  A
\end{bmatrix} = D \begin{bmatrix}
  I \\
  -A_0
\end{bmatrix}
\]

The multichannel AR prediction error vector is given by

\[
e(k)^T = \begin{bmatrix}
  y(k)^T \\
  u(k)^T
\end{bmatrix} - [y^T, u^T] D = \begin{bmatrix}
  e_y(k)^T \\
  e_u(k)^T
\end{bmatrix}
\]

and defining

\[
\psi = \begin{bmatrix}
  I \\
  -A_0
\end{bmatrix}
\]

it follows that

\[
e(k)^T \psi = y(k)^T - u(k) A_0 - [y^T, u^T] \begin{bmatrix}
  B \\
  A
\end{bmatrix} = e_0(k)^T
\]
establishing the generalization of equation 3.27c relating the multichannel AR prediction error vector to the multichannel ARMA error vector $e_0(k)$. The ARMA prediction error covariance then is given by

$$P_0 = \psi^T P \psi$$  \hspace{1cm} (3.52)

and the coefficient matrix $A_0$ can be set to minimize the trace of $P_0$ resulting in a solution given by

$$A_0 = \varepsilon\{e_u(k) e_u(k)^T\}^{-1} \varepsilon\{e_u(k) e_y(k)^T\}$$  \hspace{1cm} (3.53)

completing the multichannel generalization of the single channel results. The portion of the solution given by the multichannel autoregression can be solved as before using the lattice methods in either batch or adaptive fashion. Then the matrix of gains can be obtained from equation (3.53) by batch processing or equations (3.37) and (3.39) can be generalized to yield an adaptive solution given by

$$A_0(k+1) = A_0(k) + 2 \frac{\alpha}{\sigma_0^2(k)} e_u(k) e_0(k)^T$$  \hspace{1cm} (3.54a)

where

$$\sigma_0^2(k) = (1-\alpha) \sigma_0^2(k-1) + \alpha e_u(k)^T e_u(k)$$  \hspace{1cm} (3.54b)
It is clear that the equations and methods developed earlier for the single channel ARMA model are a special case of these results with $Q_1 = Q_0 = 1$. 
IV. NONLINEAR SYSTEM MODELING

The modeling of nonlinear systems is a far more complex problem than linear systems modeling. No attempt is made here to provide a comprehensive treatment of the problem. Rather, two specific models for systems comprised of the interconnection of linear and memoryless nonlinear subsystems are considered. Both of these models, the Volterra model and the new nonlinear ARMA model, are shown to be generalizations of the MA and ARMA modeling problems explored previously so that with appropriate modifications, the Levinson algorithm and lattice methods can be used to solve for the model parameters.

A. VOLTERRA NONLINEAR MODELING

The Volterra series model characterizes nonlinear systems using a generalization of convolution where the system output is approximated as a summation (possibly infinite) of outputs of degree $m$ systems.

$$\hat{y}(k) = \sum_{m=1}^{M} y_m(k)$$

(4.1a)

This is shown pictorially in Figure 4.1.
A number of representations for these $m$-th degree systems are possible. The most commonly used representation is in terms of symmetric Volterra kernels and is given by

$$y_m(k) = \sum_{n_m=0}^{\infty} \ldots \sum_{n_1=0}^{\infty} a_s(n_1 \ldots n_m)u(k-n_1)\ldots u(k-n_m) \quad (4.1b)$$

and $a_s(n_1 \ldots n_m)$ is the $m$-th degree symmetric Volterra kernel. (Any permutation of the indices results in the same value for this kernel giving a high degree of symmetry.)

This model arises quite naturally for a linear system in cascade with a power series nonlinearity as shown in Figure 4.2 for a quadratic nonlinearity where the output can be written as
\[ y(k) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} a(n_1) a(n_2) u(k-n_1) u(k-n_2) \]  
\text{(4.2a)}

and

\[ a_{s}(n_1n_2) = a(n_1) a(n_2) \]  
\text{(4.2b)}

Figure 4.2. A quadratic nonlinear system.

The Volterra series model in this form has been widely discussed in the literature [e.g. Ref. 1, 6, 13, 14, 25, 26, 48 and 54] since Wiener [Ref. 62] first applied it to systems analysis and modeling. It has the added benefit of treating linear systems as a special case of the model since the first degree kernel is exactly the convolutional representation of the MA model.

The number of kernels \((M)\) required for an accurate model depends on the nature of the nonlinearity in the system and as long as the nonlinearity is soft, a relatively low degree
model will suffice, keeping the problem manageable in that regard. The primary difficulty associated with Volterra nonlinear modeling arises from the fact that it uses a nonrecursive MA representation for the linear portion of the system, requiring in general, an infinite memory as indicated by the upper limits on the summations in equation (4.1b). In practice, these summations need to be truncated as shown in equation (4.3)

\[ y_m(k) = \sum_{n_1=0}^{N} \sum_{n_m=0}^{N} a_s(n_1 \ldots n_m) u(k-n_1) \ldots u(k-n_m) \]  

but a large number of terms may still be required to accurately model the system. One method of solving for the model is to set the parameters (terms of the Volterra kernels) to obtain a minimum mean square equation error where the equation error is defined as the difference between the system output and \( \hat{y}(k) \).

To simplify the solution and reduce the number of parameters that must be obtained, the symmetry of the kernels can be exploited by rewriting equation (4.1b) as

\[ y_m(k) = \sum_{n_1=0}^{N} \sum_{n_2=n_1}^{N} \sum_{n_m=n_{m-1}}^{N} a_t(n_1 \ldots n_m) u(k-n_1) \ldots u(k-n_m) \]  

where \( a_t(n_1 \ldots n_m) \) is the m-th degree triangular Volterra kernel. For a finite upper limit of \( N \) on the summations the m-th degree symmetric kernel will contain \( (N+1)^m \) terms.
but only \( \frac{(N+m)!}{N!\ m!} \) of them are distinct with the remainder determined by symmetry considerations.

Still another representation for \( y_m(k) \) uses the regular form of the Volterra kernel (this terminology has recently been introduced by Mitzel, Clancy and Rugh [Refs. 7 and 35]). It is given by

\[
y_m(k) = \sum_{h_1=0}^{\infty} \ldots \sum_{h_m=0}^{\infty} a_r(h_1\ldots h_m)u(k-h_1)u(k-h_1-h_2) \]

\[
\ldots u(k-h_1-\ldots-h_m) \quad (4.5)
\]

where \( a_r(h_1\ldots h_m) \) is the \( m \)-th degree regular Volterra kernel. With infinite upper limits on the summations, the symmetric, triangular and regular forms of the Volterra kernels are equivalent, however, when finite upper limits are used they cover the field of the model kernels in different ways.

Because of its symmetry, it is reasonable to have equal upper limits on all the summations on the symmetric kernel as was done in equation (4.3). This is shown in Figure 4.3a for a second degree kernel. The equivalent triangular kernel is given by

\[
y_m(k) = \sum_{n_1=0}^{N} \ldots \sum_{n_m=0}^{N} a_t(n_1\ldots n_m)u(k-n_1)\ldots u(k-n_m) \quad (4.6a)
\]

and it covers the region shown in Figure 4.3b for a second degree case. The corresponding regular form expansion, however,
requires variable upper limits on the summations to cover
the equivalent kernel space as shown in equation (4.6b), and
Figure 4.3c for a second degree kernel.

\[ y_m(x) = \sum_{h_1=0}^{N} \sum_{h_2=0}^{N-h_1} \sum_{h_m=0}^{N-h_1-1} a_r(h_1 \ldots h_m)u(k-h_1)u(k-h_1-h_2) \]

\[ \ldots u(k-h_1-\ldots-h_m) \]  

(4.6b)

Thus it is seen that only half the field needs to be covered
by the triangular and regular expansions to identify the
kernel associated with the square field of the regular ex-
pansion. When the regular form expansion is used with con-
stant finite upper limits there is no inherent reason to
make all the upper limits equal since there is no symmetry
in the kernel. Considering the regular expansion

\[ y_m(k) = \sum_{h_1=0}^{N_1} \ldots \sum_{h_m=0}^{N_m} a_r(h_1 \ldots h_m)u(k-h_1)\ldots (u(k-h_1-\ldots-h_m) \]

(4.7a)

a triangular expansion given by

\[ y_2(k) = \sum_{n_1=0}^{N_1} \sum_{n_2=n_1}^{n_1+N_2} \sum_{n_m=n_1}^{n_1+\ldots+n_m} a_t(n_1 \ldots n_m)u(k-n_1)\ldots u(k-n_m) \]

(4.7b)

is required to cover the corresponding field. For a quadratic
expansion this is shown in Figures 4.4c and 4.4b. The equi-
ivalent region in the symmetric field is shown in Figure 4.2a.
Figure 4.3. Range of variables corresponding to a square field in the symmetric kernel space.
Figure 4.4. Range of variables corresponding to a rectangular field in the regular kernel space.
Therefore, identifying a rectangular kernel in the regular kernel space is equivalent to obtaining a symmetric kernel in the arrow shaped region of Figure 4.4a.

Which type of expansion is more appropriate for a given system depends on the shape of the kernels for that system. For example, if a quadratic nonlinear system has a kernel with a relatively square shape in the symmetric kernel space, a regular form expansion with constant upper summation limits will have to estimate many zero valued terms and is inefficient. On the other hand if the system has a kernel similar to the arrow shaped region of Figure 4.4a in the symmetric space, a regular expansion will be efficient while a symmetric expansion over a larger field would be required with many zero valued parameters. Not enough is known about how to relate types of nonlinear systems with various shaped kernels however, and the question of kernel shape and the best type of expansion is not pursued further here.

1. Lattice Filter Methods For Volterra Modeling

Lattice filter methods derived earlier can be applied to obtain a minimum mean square equation error solution for the Volterra model if the regular form of the expansion is used. The Volterra model can be put in the form of a multiple input single output MA model by defining a new family of signals as nonlinear combinations of delayed values of the system input u(k).
\[ u_{h_2 \ldots h_m}(k) = u(k) \, u(k-h_2) \, u(k-h_2-h_3) \ldots u(k-h_2-\ldots-h_m) \]  

(4.8)

For finite summations the regular form of the expansion becomes

\[
y_m(k) = \sum_{h_1=0}^{N_{m1}} \sum_{h_2=0}^{N_{m2}} \sum_{h_m=0}^{N_{mm}} a_r(h_1 \ldots h_m) u_{h_2 \ldots h_m}(k-h_1) 
\]

(4.9)

and can be regarded as the sum of the outputs of a large number of linear filters whose inputs are given by \( u_{h_2 \ldots h_m}(k) \).

Equation (4.9) is exactly the form of a \( Q_i \) input, single output MA model where

\[
Q_i = (N_{m2}+1)(N_{m3}+1) \ldots (N_{mm}+1) 
\]

(4.10)

Furthermore, if the same upper limit on the summations over \( h_1 \) is used in each of the various \( m \)-th degree systems, \( (N_{11}=N_{21}=\ldots=N_{m1}) \), the overall Volterra model given by equations (4.1a) and (4.9) is in a form suitable for solution via the multichannel Levinson algorithm or the multichannel MA lattice methods. The requirement to use the same upper limit on all summations over \( h_1 \) arises because the Levinson algorithm and lattice methods assume the same amount of memory in each channel of the model.

As a specific example consider a second degree expansion where \( N_{11}=N_{21}=N_1 \) and \( N_{21}=N_2 \).
\[ y(k) = \sum_{h_1=0}^{N_1} a_{r_1}(h_1)u(k-h_1) + \sum_{h_2=0}^{N_2} \sum_{h_1=0}^{N_1} a_{r_2}(h_1, h_2)u_{h_2}(k-h_1) \]  

\[ u_{h_2}(k) = u(k)u(k-h_2) \]

Defining data and coefficient vectors for each channel given by

\[ u^+_{h_2}(k) = [u_{h_2}(k) \ldots u_{h_2}(k-N_1)]^T \]  

\[ a^+_{h_2} = [a_{r_2}(0, h_2) \ldots a_{r_2}(N_1, h_2)]^T \]

and embedding them into single data and coefficient vectors written as

\[ X^+(k) = [u^+(k)^T \mid u^+_0(k)^T \mid \ldots \mid u^+_N(k)^T]^T \]  

\[ d^+ = [a^+T \mid a^+_0 \mid \ldots \mid a^+_N]^T \]

the equation for the Volterra model output becomes

\[ \hat{y}(k) = X^+(k)^T d^+ \]
which is clearly of the same form as equation (A.24a) and represents a $N_2 + 2$ input, single output MA model. All the nonlinearities in the Volterra model are external to this MA model, in the formation of its various input signals from the system input $u(k)$. This model is illustrated in Figure 4.5.

It is interesting to consider what the recursive in order nature of the Levinson algorithm or lattice methods mean to the nonlinear Volterra problem. In building up the MA model solution recursively in order, the upper limit of the summations over $h_1$ is increased until the desired value is reached. In terms of the Volterra model kernels, this means allowing each of the regular form kernels to grow in size in the $h_1$ dimension while holding their boundaries fixed at prespecified values in all the other dimensions. In the linear MA model, the kernel has only one dimension and therefore the recursive in order solution eliminates any requirement to prespecify its upper limit (the order of the model). For the higher degree nonlinear kernels, these methods reduce by one the number of kernel boundaries that must be prespecified for each kernel. Allowing the kernel to grow in any of the other $M-1$ dimensions ($h_2$ through $h_m$) corresponds to adding additional channels to an existing lattice and simple methods of accomplishing this are not available.
Figure 4.5. A second degree nonlinear Volterra model in multichannel MA form. The $a_n^+$ coefficients are coefficients of the single input single output MA models shown.
3. NONLINEAR ARMA MODELING

As was previously mentioned, the primary difficulty associated with the Volterra series model arises from the fact that it is a nonlinear generalization of the MA model and as such, a large number of terms may be required to accurately represent even a mildly nonlinear system. In linear system modeling this difficulty was remedied by using the more general ARMA model. It is reasonable to assume therefore that a nonlinear generalization of the ARMA model could remedy the problem in the nonlinear modeling case (for at least certain types of nonlinear systems). Such a generalization called the nonlinear ARMA model has recently been proposed. [Ref. 64] This model forms an estimate of the current value of the system output as follows:

\[ \hat{y}(k) = \sum_{n_1=0}^{\infty} a_s(n_1)u(k-n_1) + \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} a_s(n_1n_2)u(k-n_1)u(k-n_2) \]

\[ + \sum_{n_1=0}^{\infty} \sum_{n_p=0}^{\infty} a_s(n_1...n_p)u(k-n_1)...u(k-n_p) \]

\[ + \sum_{m_1=1}^{\infty} b_s(m_1)y(k-m_1) + \sum_{m_1=1}^{\infty} \sum_{m_2=1}^{\infty} b_s(m_1m_2)y(k-m_1)y(k-m_2) \]

\[ + \sum_{m_1=1}^{\infty} \sum_{m_q=1}^{\infty} b_s(m_1...m_q)y(k-m_1)...y(k-m_q) \]
\[ + \sum_{n_1=0}^{\infty} \sum_{m_1=1}^{\infty} C(n_1 m_1) u(k-n_1) y(k-m_1) + \ldots + \sum_{n_1=0}^{\infty} \sum_{n_p=0}^{\infty} \sum_{m_1=1}^{\infty} \ldots \]

\[ \ldots \sum_{m_q=1}^{\infty} C(n_1 \ldots n_p m_1 \ldots m_q) u(k-n_1) \ldots u(k-n_p) y(k-m_q) \]

\[ \ldots y(k-M_q) \]  

(4.13)

The first three terms of equation (4.13) are a discrete Volterra expansion of the input signal \( u(k) \) and represent \( F_{10}[u(k)] \) in terms of the discussion of Chapter I. The second three terms are a discrete Volterra expansion on the system output delayed one sample interval and represent \( F_{20}[y(k-1)] \). The final two terms are bivariate expansions of the system input and delayed output and represent \( F_{30}[u(k), y(k-1)] \). (This is the first and only model considered where \( F_{30}[*] \) is not assumed to be zero.) Equation (4.13) is clearly a nonlinear extension of the linear ARMA model contained in the first and fourth terms. As was the case in the Volterra model, the number of multiple summations required in (4.13) is dependent upon the nature of the nonlinearity in the system being modeled. The upper limits on the summations in (4.13) have been written as infinity indicating a requirement for infinite memory or model order. As is discussed subsequently, however, the required model order (memory) may in fact be finite due to the nature of the system being modeled, thereby alleviating the difficulty encountered in the Volterra nonlinear model previously presented.
The kernels of the input expansion and output expansion $a_s(\cdot)$ and $b_s(\cdot)$ are symmetric since any permutation of the indices results in the same value for the kernel. They have therefore been labeled with a subscript "s" and can also be written in triangular and regular form.

$$\sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_p=0}^{\infty} a_s(n_1 \ldots n_p) u(k-n_1) \ldots u(k-n_p)$$

$$= \sum_{n_1=0}^{\infty} \sum_{n_2=n_1}^{\infty} \cdots \sum_{n_p=n_{p-1}}^{\infty} a_t(n_1 \ldots n_p) u(k-n_1) \ldots u(k-n_p)$$

$$= \sum_{h_1=0}^{\infty} \sum_{h_2=0}^{\infty} \cdots a_r(h_1 \ldots h_p) u(k-h_1) u(k-h_1-h_2) \ldots u(k-h_1-\ldots-h_p) \quad (4.14)$$

$$\sum_{m_1=1}^{\infty} \sum_{m_2=1}^{\infty} \cdots \sum_{m_q=1}^{\infty} b_s(m_1 \ldots m_q) y(k-m_1) \ldots y(k-m_q)$$

$$= \sum_{m_1=0}^{\infty} \sum_{m_2=m_1}^{\infty} \cdots \sum_{m_q=m_{q-1}}^{\infty} b_t(m_1 \ldots m_q) y(k-1-m_1) \ldots (k-1-m_q)$$

$$= \sum_{h_1=0}^{\infty} \sum_{h_2=0}^{\infty} \cdots b_r(h_1 \ldots h_q) y(k-1-h_1) y(k-1-h_1-h_2) \ldots y(k-1-h_1-\ldots-h_q) \quad (4.15)$$
In writing the triangular and regular forms in equation \((4.15)\) the lower index on the summations has been shifted to zero.

In the case of the bivariate expansion terms in equation \((4.13)\) the kernels do not possess symmetry so that triangular expansions are not possible but regular form expansions are possible.

\[
\sum_{n_1=0}^{\infty} \ldots \sum_{n_p=0}^{\infty} \sum_{m_1=1}^{\infty} \ldots \sum_{m_q=1}^{\infty} C(n_1 \ldots n_p m_1 \ldots m_q) u(k-n_1) \\
\ldots u(k-n_p) y(k-m_1) \ldots y(k-m_q).
\]

\[
= \sum_{h_1=0}^{\infty} \ldots \sum_{h_{p+q}=0}^{\infty} c_r(h_1 \ldots h_{p+q}) u(k-h_1) \ldots u(k-h_1 \ldots -h_p) \\
y(k-1-h_1 \ldots -h_{p+1}) \ldots y(k-1-h_1 \ldots -h_{p+q})
\]

\[
= \sum_{h_1=0}^{\infty} \ldots \sum_{h_{p+q}=0}^{\infty} c_r(h_1 \ldots h_{p+q}) y(k-1-h_1) \ldots y(k-1-h_1 \ldots -h_q) \\
u(k-h_1 \ldots -h_{q+1}) \ldots u(k-h_1 \ldots -h_{p+q})
\]

\((4.16)\)

Thus two regular forms are possible for the bivariate expansions. Figures 4.6 and 4.7 illustrate the manner in which the regular form of the bivariate expansion covers the field.
Figure 4.6. Range of variables for a rectangular field in the original form of the bivariate expansion of degree two.
Figure 4.7. Range of variables for a rectangular field in the regular form of the bivariate expansion of degree two.
of model kernels for a second degree case for finite upper limits of $N_1$ and $N_2$ on the summations. It is interesting to note that because of a lack of symmetry in the original form of the bivariate expansion, the causal region in the regular form extends outside the quarter plane.

As was the case with the Volterra model, it will be shown that in regular form, a minimum mean square equation error solution can be obtained for the nonlinear ARMA model coefficients using either the Levinson algorithm or the lattice methods. This will provide the nonlinear generalization of the results presented in Chapter III on linear ARMA modeling. Before developing this method of solution, however, the applicability of the nonlinear ARMA model to various types of systems and its memory requirements will be considered.

1. **Identifiability Conditions and Memory Requirements**

   In the previous chapter on linear ARMA modeling it was stated that there were two facets to the question of model identifiability; input signal requirements and measurement requirements. The question of measurement requirements was not discussed, however, since it was assumed that all signals (input and output) were observed. In the study of systems comprised of interconnected linear and nonlinear subsystems, however, various internal signals exist and the effect of either observing or not observing them on the modeling process must be explored. To do so, it will be assumed that the system under study can be put into the form of Figure 4.8 fulfilling the following equations.
\[
\begin{align*}
\mathbf{x}_L(k) &= \mathbf{\Gamma}_1 \mathbf{x}_N(k) + \mathbf{C}_1 \mathbf{u}(k) \\
\mathbf{x}_N(k) &= \mathbf{\Gamma}_2 \mathbf{x}_L(k) + \mathbf{C}_2 \mathbf{u}(k) \\
\mathbf{y}_N(k) &= \mathbf{F}[\mathbf{x}_N(k)] \\
\mathbf{y}_L(k) &= \mathbf{T}[\mathbf{x}_L(x)]
\end{align*}
\]

where

\[
\mathbf{x}_L(k) = [x_{L1}(k) \ldots x_{LP}(k)]^T
\]

a vector of inputs to P linear subsystems;

\[
\mathbf{x}_N(k) = [x_{N1}(k) \ldots x_{NQ}(k)]^T
\]

a vector of inputs to Q nonlinear memoryless subsystems;

\[
\mathbf{y}_L(k) = [y_{L1}(k) \ldots y_{LP}(k)]^T
\]

a vector of outputs from P linear subsystems;

\[
\mathbf{y}_N(k) = [y_{N1}(k) \ldots y_{NQ}(k)]
\]

a vector of outputs from Q nonlinear memoryless subsystems.
Figure 4.8. A General Nonlinear System
The $z$ - transforms of these signals are also defined as $X_L(z), X_N(z), Y_L(z), Y_N(z)$. $\Gamma_1, \Gamma_2, C_1$ and $C_2$ are matrices whose elements are either 0, -1 or +1 indicating the interconnections of the various subsystems. $T[.]$ and $F[.]$ are diagonal matrices specifying the linear and nonlinear subsystems as follows.

$$Y_L(z) = T(z) X_L(z) \quad (4.19a)$$

where

$$T(z) = [T_{ij}(z)] = \begin{cases} A_i(z) & i=j \\ 1-B_i(z) & i\neq j \\ 0 & \end{cases} \quad (4.19b)$$

and

$$A_i(z) = \sum_{n=0}^{N} a_i(n) z^{-n} \quad (4.19c)$$

$$B_i(z) = \sum_{n=1}^{N} b_i(n) z^{-n} \quad (4.19d)$$

$$Y_N(k) = F[x_N(k)] = \begin{bmatrix} F_1[.] & 0 & \cdots & 0 \\ F_2[.] & \vdots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & F_Q[.] & \end{bmatrix} \begin{bmatrix} x_{N_1}(k) \\ x_{N_2}(k) \\ \vdots \\ x_{N_Q}(k) \end{bmatrix}$$

$$= \begin{bmatrix} F_1[x_{N_1}(k)] \\ \vdots \\ F_Q[x_{N_Q}(k)] \end{bmatrix} \quad (4.20)$$
This overall system given by Figure 4.8 is adequate to represent a broad class of systems comprised of interconnections of linear and nonlinear subsystems including cascades, parallel connections and feedback systems. Equations (4.19) and (4.20) assume all the subsystems are single input single output noninteracting systems. If desired, the collection of linear subsystems given by \( T(z) \) can readily be put into the general multichannel ARMA form of Section III.F to allow each output to be a function of past values of all outputs, and past and present values of all inputs.

Alternate representations of these linear and nonlinear subsystems are also useful. Equations (4.19) are equivalent to

\[
Y_L(z) = [A_i(z)]X_L(z) + [B_i(z)]Y_L(z) \tag{4.21a}
\]

or in the time domain

\[
y_L(k) = [a_i(k)]^\ast x_L(k) + [b_i(k)]^\ast y_L(k) \tag{4.21b}
\]

Here \(^\ast\) represents convolution and is carried out in the same sense as matrix multiplication and the matrices.

\([a_i(k)], [b_i(k)], [A_i(z)] \text{ and } [B_i(z)]\) are diagonal matrices whose \(i\)-th entries are the time domain functions \(a_i(k)\) or \(b_i(k)\) or the polynomials \(A_i(z)\) or \(B_i(z)\). The time domain functions \(a_i(k)\) and \(b_i(k)\) are the inverse transforms of the polynomials \(A_i(z)\) and \(B_i(z)\). This can also be represented
in nonrecursive form as

\[ y_L(k) = [h_i(k)]^*x_L(x) \]  \hspace{1cm} (4.22a)

where \([h_i(k)]\) is a diagonal matrix of impulse responses defined by

\[ h_i(k) = \sum_{n=0}^{\infty} h_i(n) \delta(k-n) \]  \hspace{1cm} (4.22b)

The nonlinear systems can also be represented in terms of
inverse functions assuming they exist over the necessary
ranges of the variables so that

\[ \begin{bmatrix} F_1^{-1}[\cdot] & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ddots & F_Q^{-1}[\cdot] \end{bmatrix} \begin{bmatrix} y_{N1}(k) \\ \vdots \\ y_{NQ}(k) \end{bmatrix} = \begin{bmatrix} x_N(k) \end{bmatrix} \]

So that equations (4.17) are iterative it is necessary that no delay free loops exist in the system of
Figure 4.8. A necessary and sufficient condition for the
absence of delay free loops is developed in Appendix H and
requires that the terms of the determinant of the matrix
\[
\alpha = \lim_{z \to \infty} \Gamma_2 T(z) \Gamma_1
\]  
(4.24)

and all its principal minors must be zero.

The various signal vectors and equations (4.17) can now be partitioned as

\[
\begin{bmatrix}
\mathbf{x}_1'(k) \\
\mathbf{x}_2'(k) \\
\mathbf{x}_3'(k) \\
\mathbf{x}_4'(k)
\end{bmatrix} =
\begin{bmatrix}
\Gamma_{1a} & \Gamma_{1b} \\
\Gamma_{1c} & \Gamma_{1d}
\end{bmatrix}
\begin{bmatrix}
\mathbf{x}_N'(k) \\
\mathbf{x}_N''(k)
\end{bmatrix} +
\begin{bmatrix}
\mathbf{c}_{1a} \\
\mathbf{c}_{1b}
\end{bmatrix} u(k)
\]  
(4.25a)

\[
\begin{bmatrix}
\mathbf{x}_1''(k) \\
\mathbf{x}_2''(k) \\
\mathbf{x}_3''(k) \\
\mathbf{x}_4''(k)
\end{bmatrix} =
\begin{bmatrix}
\Gamma_{2a} & \Gamma_{2b} \\
\Gamma_{2c} & \Gamma_{2d}
\end{bmatrix}
\begin{bmatrix}
\mathbf{x}_L'(k) \\
\mathbf{x}_L''(k)
\end{bmatrix} +
\begin{bmatrix}
\mathbf{c}_{2a} \\
\mathbf{c}_{2b}
\end{bmatrix} u(k)
\]  
(4.25b)

\[
\begin{bmatrix}
\mathbf{y}_1'(k) \\
\mathbf{y}_2'(k) \\
\mathbf{y}_3'(k) \\
\mathbf{y}_4'(k)
\end{bmatrix} =
\begin{bmatrix}
\mathbf{F}_1[.] \\
\mathbf{F}_2[.]
\end{bmatrix}
\begin{bmatrix}
\mathbf{x}_N'(k) \\
\mathbf{x}_N''(k)
\end{bmatrix}
\]  
(4.25c)

\[
\begin{bmatrix}
\mathbf{y}_1''(k) \\
\mathbf{y}_2''(k) \\
\mathbf{y}_3''(k) \\
\mathbf{y}_4''(k)
\end{bmatrix} =
\begin{bmatrix}
\mathbf{T}_1[.] \\
\mathbf{T}_2[.]
\end{bmatrix}
\begin{bmatrix}
\mathbf{x}_L'(k) \\
\mathbf{x}_L''(k)
\end{bmatrix}
\]  
(4.25d)

Equation (4.25c) can also be written in terms of inverse functions as in equation (4.23), and (4.25d) can be written using either recursive or nonrecursive representations of the linear systems as was done in (4.21) and (4.22). The single primed signal vectors in equations (4.25) represent those signals which are observed and the double primed signals are those which are not. It is assumed that all input signals
in $u(k)$ are observed. It is possible to rewrite equations (4.25) as a single composite matrix equation as done in (4.26a). As written, equation (4.26a) is an infinite memory or model order representation because of the presence of the $h(k)^*$ operators but a finite memory version can be written by expressing rows 4 and 8 as

$$ y_L'(k) = a_a(k)*x_L'(k) + b_a(k)*y_L'(k) + a_b(k)*x_L''(k) $$

$$ + b_b(k)*y_L''(k) $$

(4.23b)

$$ y_L''(k) = a_c(k)*x_L'(k) + b_c(k)*y_L'(k) + a_d(k)*x_L''(k) $$

$$ + b_d(k)*y_L''(k) $$

(4.26c)

The third and seventh rows can also be written in terms of inverse functions if desired as

$$ x_N'(k) = F^{-1}_a[y_N'(k)] + F^{-1}_b[y_N''(k)] $$

(4.26d)

$$ x_N''(k) = F^{-1}_c[y_N'(k)] + F^{-1}_d[y_N''(k)] $$

(4.26e)

Now the problem of writing a system of iterative equations for the observed signals in terms of only observed signals consists of rewriting equation (4.26a) so that the upper right 4 by 5 partition is a null matrix. If this can be done, the general form of the nonlinear ARMA model can be
\[
\begin{bmatrix}
    u(k) \\
    x^1_L(k) \\
    x^1_N(k) \\
    y^1_L(k) \\
    y^1_N(k) \\
    x''_L(k) \\
    x''_N(k) \\
    y''_L(k) \\
    y''_N(k)
\end{bmatrix} =
\begin{bmatrix}
    I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    C_{1a} & 0 & 0 & \Gamma_{1a} & 0 & 0 & 0 & \Gamma_{1b} \\
    C_{2a} & 0 & \Gamma_{2a} & 0 & 0 & 0 & \Gamma_{2b} & 0 \\
    0 & h_d(k)^* & 0 & \gamma & h_b(k)^* & 0 & 0 & 0 \\
    0 & 0 & \Gamma_d(\cdot) & 0 & 0 & \Gamma_b(\cdot) & 0 & 0 \\
    C_{1b} & 0 & 0 & \Gamma_{1c} & 0 & 0 & 0 & \Gamma_{1d} \\
    C_{2b} & 0 & \Gamma_{2c} & 0 & 0 & 0 & \Gamma_{2d} & 0 \\
    0 & h_p(k)^* & 0 & 0 & h_d(k)^* & 0 & 0 & 0 \\
    0 & 0 & \Gamma_c(\cdot) & 0 & 0 & \Gamma_d(\cdot) & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
    u(k) \\
    x^1_L(k) \\
    x^1_N(k) \\
    y^1_L(k) \\
    y^1_N(k) \\
    x''_L(k) \\
    x''_N(k) \\
    y''_L(k) \\
    y''_N(k)
\end{bmatrix}
\]

(4.26a)
used to identify the composite effects of the operators appearing in the upper left 5 by 5 partition. In some cases this will only be possible of the infinite memory version of the model (with \( h(k)* \) operators) is used and in other cases a finite memory model will suffice.

The process of determining whether infinite or finite memory nonlinear ARMA models can be used to identify a given system is illustrated for two examples in Appendix I. First a system consisting of a cascade of linear and nonlinear systems is considered. Then a model of the tracking behavior of a phase locked loop is put in the form of a nonlinear ARMA representation.

2. Lattice Filter Methods for the Nonlinear ARMA Model

As was the case with Volterra modeling, lattice filter methods can be applied to the nonlinear ARMA modeling problem if the regular form of the kernels is used. A family of signals is defined as nonlinear combinations of delayed values of \( y(k) \) and \( u(k) \) as follows.

\[
u_{h_2 \ldots h_m}(k) = u(k) u(k-h_2)u(k-h_2-h_3) \ldots u(k-h_2-\ldots-h_m)
\]  
\[4.27a\]

\[
y_{h_2 \ldots h_m}(k) = y(k-1)y(k-1-h_2)y(k-1-h_2-h_3) \ldots y(k-1-h_2-\ldots-h_m)
\]  
\[4.27b\]
\[ y(k-l-h_2\ldots-h_{p+q}) = u(k) u(k-h_2)\ldots u(k-h_2\ldots-h_p) \]

or alternately

\[ y_{h_2\ldots h_q} u_{h_{q+1}\ldots h_{q+p}} = y(k-l) y(k-l-h_2)\ldots y(k-l-h_2\ldots-h_q) \]

\[ u(k-h_2\ldots-h_{q+1}) u(k-h_2\ldots-h_{q+p}) \]

(4.27d)

With finite summations, the regular form of equation (4.14) becomes

\[
\sum_{h_p=0}^{N_p} \sum_{h_2=0}^{N_2} \left( \sum_{h_1=0}^{N_1} a_r(h_1\ldots h_p) u_{h_2\ldots h_p} (k-h_1) \right) 
\]

(4.28a)

Equation (4.15) becomes

\[
\sum_{h_q=0}^{M_q} \sum_{h_2=0}^{M_2} \left( \sum_{h_1=0}^{M_1} b_r(h_1\ldots h_q) y_{h_2\ldots h_q} (k-h_1) \right) 
\]

(4.28b)
and equations (4.16) become

\[
\begin{align*}
\sum_{h_p+q=0}^{L_p+q} \cdots \sum_{h_2=0}^{L_p+q,2} \sum_{h_1=0}^{L_p+q,1} C_{r1}(h_1 \ldots h_p+q) u_{h_2} \cdots h_p \\
(y_{h_{p+1}} \cdots h_{p+q})^{(k-h_1)}
\end{align*}
\]

or

\[
\begin{align*}
\sum_{h_q+p=0}^{L_q+p} \cdots \sum_{h_2=0}^{L_q+p,2} \sum_{h_1=0}^{L_q+p,1} C_{r2}(h_1 \ldots h_q+p) y_{h_2} \cdots h_q \\
u_{h_{q+1}} \cdots h_{q+p}^{(k-h_1)}
\end{align*}
\]

These terms can be viewed as the summation of the outputs of a large number of linear filters whose inputs are the signals defined in equations (4.27). In the context of multichannel filtering, each of these filters can be considered as a single channel and each of the input signals can be associated with one of the channels. Since the lattice models use the same amount of memory in each channel, the upper limits on all summations over \(h_1\) will be made the same \((N_{p,1} = N_{q,1} = L_{p+q,1} = N_1)\). The upper limits on the summations over the other indices determine the number of channels required.
For a quadratic nonlinear system, the present value of the system output is estimated as

\[
\hat{y}(k) = \sum_{h_1=0}^{N_1} a_r(h_1) u(k-h_1) + \sum_{h_2=0}^{N_2} \sum_{h_1=0}^{N_1} a_r(h_1, h_2) u_{h_2}(k-h_1)
\]

\[
+ \sum_{h_1=1}^{N_1} b_r(h_1) y(k-h_1) + \sum_{h_2=0}^{M_2} \sum_{h_1=0}^{N_1} b_r(h_1, h_2) y_{h_2}(k-h_1)
\]

\[
+ \sum_{h_2=0}^{L_2} \sum_{h_1=0}^{N_1} c_r(h_1, h_2) u h_{h_2}(k-h_1)
\]

(4.29)

where \( u_{h_2}(k) = u(k) u(k-h_2) \), \( y_{h_2}(k) = y(k-l) y(k-l-h_2) \)

and \( u h_{h_2}(k) = u(k) y(k-l-h_2) \). Signal and coefficient vectors can be defined for the various quantities in equation (4.29) following the conventions established earlier; for example

\[
y(k) = [y(k-1) \ldots y(k-N_1)]^T
\]

(4.30a)

\[
y_{h_2}^T(k) = [y_{h_2}(k) \ldots y_{h_2}(k-N_1)]^T
\]

(4.30b)

and

\[
b = [b_r(1) \ldots b_r(N_1)]^T
\]

(4.30c)

\[
b_{h_2}^+ = [b_r(0, k_2) \ldots b_r(N_1, h_2)]^T
\]

(4.30d)
Embedding all these vectors into single data and coefficient vectors, equation (4.29) becomes

\[ \hat{y}(k) = \mathbf{X}_1(k)^T \mathbf{d}_1 \]  \hspace{1cm} (4.31a)

where

\[ \mathbf{X}_1(k) = [\mathbf{y}(k)^T \mathbf{u}_0^+(k)^T \ldots \mathbf{u}_{N_{22}}^+(k)^T \mathbf{v}_0^+(k)^T \ldots \mathbf{v}_{L_{22}}^+(k)^T] \]  \hspace{1cm} (4.31b)

and

\[ \mathbf{d}_1 = [\mathbf{b}^T \mathbf{a}_0^+ \mathbf{c}_0^+ \ldots \mathbf{b}_{M_{22}}^+ \mathbf{a}_{N_{22}}^+ \mathbf{c}_{L_{22}}^+] \]  \hspace{1cm} (4.31c)

The minimum mean square equation error solution for the model coefficient vector \( \mathbf{d}_1 \) is given by

\[ \varepsilon\{ \mathbf{X}_1(k) \mathbf{X}_1(k)^T \} \mathbf{d}_1 = \varepsilon\{ \mathbf{X}_1(k) \mathbf{y}(k) \} \]  \hspace{1cm} (4.32)

where the equation error is defined as

\[ e(k) = y(k) - \hat{y}(k) \]
While equation (4.32) is similar in form to equations (A.7a) and (A.26a) for AR and MA models, it lacks the necessary form for the application of the Levinson algorithm because of the fact that some component data and coefficient vectors are $N_1$-vectors while others are $(N_1+1)$-vectors. This is the same problem that arose in linear ARMA modeling and can be handled here in a similar manner. Defining

$$\psi = \begin{bmatrix} 1 & -a_r(0) & -b_r(0,0) & \cdots & -b_r(0,M_{22}) \\ -a_r(0,0) & \cdots & -a_r(0,N_{22}) \\ -c_{r1}(0,0) & \cdots & -c_{r1}(0,L_{22}) \end{bmatrix}^T$$

(4.33a)

and

$$x(k) = \begin{bmatrix} y(k) & u(k) & y_0(k) & \cdots & y_{M_{22}}(k) \\ u_0(k) & \cdots & u_{N_{22}}(k) \\ uy_0(k) & \cdots & uy_{L_{22}}(k) \end{bmatrix}^T$$

(4.33b)

and assuming that the coefficients in $\psi$ can be estimated in some other manner, the MMSE solution for the remaining model coefficients becomes

195
\[ \varepsilon\{X(k)X(k)^T\} \psi = \varepsilon\{X(k)x(k)^T\} \]  \hspace{1cm} (4.34)

Here \(X(k)\) and \(d\) are defined as \(X_1(k)\) and \(d_1\) with all superscripts"+" removed from their component vectors (indicating they are all indexed from 1 to \(N_1\) and are all \(N_1\)-vectors). Note that just as in the linear ARMA case, the coefficients in \(\psi\) essentially correspond to gains on their respective channels. Comparing equation (4.34) with (A.7a) for a multichannel autoregression it is clear that \(d\) can be obtained from the multichannel AR solution by

\[ d = D \psi \]  \hspace{1cm} (4.35)

where the signals in \(x(k)\) comprise the channels in the autoregression. Therefore, either the Levinson algorithm or the lattice methods can be used to solve for \(D\) and from it and a knowledge of \(\psi\), the nonlinear ARMA model coefficients can be obtained.

By analogy with equation (3.27) it also follows that the nonlinear ARMA equation error is related to the AR prediction error vector by

\[ e(k) = \psi^T e(k) \]  \hspace{1cm} (4.36a)

so that

\[ \varepsilon\{e(k)^2\} = \psi^T P \psi \]  \hspace{1cm} (4.36b)
where $P$ is the AR prediction error covariance matrix. The coefficients in $\psi$ can therefore be set to minimize the mean square value of the nonlinear ARMA equation error in (4.36b) resulting in a solution given by

$$\begin{bmatrix} P_{22} & \cdots & P_{2N} \\ \vdots & & \vdots \\ P_{N2} & P_{NN} \end{bmatrix} \begin{bmatrix} \phi_1(0) \\ \vdots \\ \phi_N(0, L_{22}) \end{bmatrix} = \begin{bmatrix} P_{21} \\ \vdots \\ P_{N1} \end{bmatrix}$$

(4.37a)

where $N$ is the total number of channels in the model

$$N = 1 + 1 + (M_{22} + 1) + (N_{22} + 1) + (L_{22} + 1)$$

and the $p_{ij}$ are the elements of the prediction error covariance matrix. It is readily apparent that the linear ARMA model and its solution via the Levinson algorithm or lattice methods are a special case of this formulation of the nonlinear ARMA model just as one would expect.
V. APPLICATIONS, CONCLUSIONS, AND OPEN QUESTIONS

In the previous four chapters, existing methods for AR and MA modeling were reviewed and from them new methods for linear and nonlinear ARMA modeling were developed. Here, two applications for these new methods in reduced order modeling and modeling for fault detection and diagnosis are examined briefly. Then the results of this research are summarized, conclusions drawn and significant open questions for the continuation and extension of this work are listed.

A. APPLICATIONS

1. Reduced Order Modeling

Oftentimes, complex physical systems, both linear and nonlinear, can be approximated quite closely using simple models. The lattice solution methods developed here provide a very natural and efficient means of determining reduced order models for complex, high order systems especially in the case of linear systems. In Chapter III it was argued that for linear systems it is reasonable to build up ARMA models by simultaneously incrementing the order of the numerator and denominator polynomials as the lattice method does. When this method is used to build up a given order model, all lower order models and their mean square values of equation error are readily obtained as well (the only additional
calculations needed are for the MSE and the gain term or gain matrix in the multichannel case) making it easy to compare the various models and decide if reduced order models provide sufficient accuracy.

Consider for example the seventh order system whose characteristics are listed in Table 5.1. The magnitude spectra of second, third, sixth and seventh order lattice models obtained using batch processing with 4000 point averages are compared to that of the system in Figure 5.1. It is apparent that a second order model is unable to approximate the system well, however a third order model does provide a good approximation. Furthermore, increasing the model order to four, five and six fails to significantly improve its performance as evidenced by the sixth order plot in Figure 5.1c. The model accuracy is not significantly increased until seventh order (which corresponds to the order of the actual system) when a very good fit is achieved. This is further illustrated by the plot of the normalized mean square equation error as a function of model order shown in Figure 5.2. The cost drops rapidly going from second to third order but then fails to decrease significantly until seventh order is reached. The roots of the system and the various order models are also plotted in Figure 5.3.

The benefits of the lattice method for reduced order nonlinear ARMA modeling are not quite so pronounced however, since adding extra stages to the lattice corresponds to allowing the kernels to grow in only one of their multiple
TABLE 5.1

SYSTEM E
TRANSFER FUNCTION COEFFICIENTS

<table>
<thead>
<tr>
<th>Numerator</th>
<th>Denominator</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A(1) = 1.00000 )</td>
<td>( B(1) = 2.18950 )</td>
</tr>
<tr>
<td>( A(1) = -2.16510 )</td>
<td>( B(2) = -2.21260 )</td>
</tr>
<tr>
<td>( A(2) = 1.56250 )</td>
<td>( B(3) = 1.51740 )</td>
</tr>
<tr>
<td>( A(3) = 0.0 )</td>
<td>( B(4) = -0.85350 )</td>
</tr>
<tr>
<td>( A(4) = 0.0 )</td>
<td>( B(5) = 0.45144 )</td>
</tr>
<tr>
<td>( A(5) = 0.0 )</td>
<td>( B(6) = -0.23462 )</td>
</tr>
<tr>
<td>( A(6) = 0.0 )</td>
<td>( B(7) = 0.09331 )</td>
</tr>
</tbody>
</table>

Root Locations

<table>
<thead>
<tr>
<th>Zeros</th>
<th>Poles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re</td>
<td>Im</td>
</tr>
<tr>
<td>1.0E250</td>
<td>0.62500</td>
</tr>
<tr>
<td>1.0E250</td>
<td>-0.62500</td>
</tr>
<tr>
<td>C.C</td>
<td>0.0</td>
</tr>
<tr>
<td>C.C</td>
<td>0.0</td>
</tr>
<tr>
<td>C.C</td>
<td>0.0</td>
</tr>
<tr>
<td>C.C</td>
<td>0.0</td>
</tr>
<tr>
<td>C.C</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Figure 5.1.
Figure 5.1 cont'
Figure 5.2. Mean square value of equation error (as a percentage of the mean square value of system output) vs. model order for lattice models of system E obtained using batch processing and 4000 point averages.
Figure 5.3.
Figure 5.3 cont'd.
dimensions. (The linear "kernel" had only one dimension.) Here some a priori information would be useful in determining the values to prespecify for the boundaries of the kernels in the other dimensions. Still however, when compared to a brute force matrix inversion approach where all kernel boundaries must be prespecified, the lattice method provides a viable alternative in obtaining reduced order models especially for low degree systems.

2. Modeling For Fault Detection

The problem of fault detection and possible diagnosis can be formulated as follows.

a. Obtain a parameterization that describes the current functioning of the system under test.

b. From this parameterization, determine if the system is functioning normally or if a fault has occurred by comparison with a fault dictionary.

It is the first part of this problem that has been addressed in this work. The parameterization can be as simple as sampled measurements of the response to specific inputs however, the large volume of data that would generally be involved in such an approach would greatly complicate the second part of the problem. A more efficient approach in terms of utilization of parameters is to model the system and use the model parameters as a description of its current functioning.

For linear systems, ARMA models provide a very general framework with a number of possible parameter sets.
Three candidates are polynomial coefficients, root locations and lattice reflection coefficients with the latter offering many advantages. In addition to the advantages demonstrated by the experimental results of Chapter III, reflection coefficients provide a very effective and methodical way to build up knowledge of a system's characteristics. As model order is increased to more accurately represent a system, reflection coefficients already determined don't change making them ideal candidates for use in a dictionary lookup scheme (a characteristic not shared by the other parameterizations). This is made more important since reduced order models could be adequate to detect and perhaps diagnose some faults, especially catastrophic ones. While more parameters are required when reflection coefficients are used, these same coefficients also provide all reduced order models. For a single channel ARMA example, 8N reflection coefficients and N gains provide all models from order 1 to N while $N^2+2N$ parameters would be required using either polynomial coefficients or roots to provide the same information. (6N reflection coefficients are needed if the input is white noise since $k_{12}=k_{22}=0$.)

A similar argument could be made for the use of lattice reflection coefficients with the nonlinear ARMA model for fault detection and diagnosis of nonlinear systems.
B. CONCLUSIONS AND OPEN QUESTIONS

The purpose of this research was to extend existing theories and methods in the modeling of linear and nonlinear systems to broader, more general types of models. After a discussion of available results in AR and MA modeling of linear systems with particular emphasis on the Levinson algorithm and lattice filter methods, model transition formulas were developed to relate the more general ARMA model for linear systems to the AR and MA models. It was shown that with suitable assumptions, the ARMA model solution could also be obtained recursively in order using either a modified Levinson algorithm or lattice filter methods. These results were developed extensively in both theory and practice for single channel linear ARMA modeling with experimental verification of both the batch processing and adaptive lattice methods presented. Portions of these results have already been published. [Refs. 65 and 66]

The theory was also developed to generalize these results to the multichannel ARMA case.

Based on the simulation results it was concluded that the lattice methods offer the following advantages over a conventional brute force matrix inversion approach to ARMA modeling using windowed correlation estimates.

1. For short runs of data the batch lattice methods provide much more accurate results than the brute force method.
2. The batch lattice method performs much better than the brute force method when the system is overmodeled.

3. The MSE as a function of model order is well behaved for the lattice method.

4. The adaptive lattice method has difficulty tracking zero valued overmodeled parameters.

The cost of these advantages is the extra computational burden of passing the data through the lattice filter during the modeling process.

In the discussion of nonlinear system models the Volterra model was considered as a nonlinear extension of MA modeling and it was shown that lattice methods could be used to obtain the model solution if the problem was recast, using the regular form of the Volterra kernels. Then the new nonlinear ARMA model was considered and it was shown that this representation in some cases solves the problem of requiring a very large number of model parameters encountered in Volterra modeling. Then lattice methods were developed for the nonlinear ARMA problem and it was shown that the linear ARMA techniques presented earlier are a special case of the nonlinear ARMA methods. For both types of nonlinear models, the recursive in order nature of the lattice methods was shown to allow the various model kernels to grow in one dimension while holding their boundaries fixed at pre-specified values in the other dimensions. The use of the model nonlinear ARMA was also illustrated with two examples.
and a nonlinear ARMA model was proposed for the tracking behavior of a phase locked loop.

Several significant questions remain for the continuation and extension of this work and are listed here,

1. Stability of the linear and nonlinear ARMA models must be considered. In the linear problem, stability is dependent on the roots of the denominator polynomial of the synthesis model. The methods developed do not guarantee stability of the resulting model. (This was not found to be a problem in practice, however, unless extremely short runs of data in the range of 30 to 50 samples were used. Even then, model instability was not a frequent occurrence.) Stability for the nonlinear ARMA model remains to be clearly defined.

2. Input signal requirements in the nonlinear ARMA modeling process need to be investigated. In linear ARMA modeling the power spectrum of the input signal was found to play an important role. No requirements emerged however, on the probability density function (pdf) of the input. In nonlinear systems where the behavior is inherently level dependent, it is intuitively appealing to use an input signal with a flat power spectrum across the frequency range of interest and whose amplitude is uniformly distributed over the range of interest.
3. The inability of the adaptive lattice method to track zero valued overmodeled parameters is an interesting problem warranting further consideration. If some means of detecting the degeneration of the cost surface towards an infinite trough can be found, the problem could be remedied by simply resetting the appropriate parameters to zero.

4. Experimental experience needs to be gained with the nonlinear ARMA model itself and the lattice methods developed for it.

5. The characteristics of the lattice solution methods need to be further quantified to gain a comprehensive understanding of how and why it performs as it does. Also, further comparisons should be made between the lattice methods and conventional methods. Some comparisons were made here for batch processing methods but only a rectangular window was used on the data. Comparisons should be made using other types of window functions in the brute force method and the adaptive lattice method should also be compared with a conventional LMS adaptive algorithm applied to the equation error model in which all of the a(i) and b(i) coefficients are adapted simultaneously.

6. In the adaptive lattice method, scaling of the lattice input signals needs to be investigated. It was noted that the ratio of largest to smallest eigenvalue
of the input autocorrelation matrix was related to the speed of convergence of the adaptive algorithm. For the first lattice stage this matrix is

\[ P(0) = \varepsilon \begin{bmatrix} y^2(k) & y(k) & u(k) \\ u(k) & y(k) & u^2(k) \end{bmatrix} \]

If the system has a high gain such that the mean square value of the output \( y(k) \) is very much greater than that of the input \( u(k) \), convergence will be slow. This could be remedied by implementing an adaptive scaling scheme at the lattice input (perhaps similar to the first order low pass filter estimates used for adaptive step size normalization).
APPENDIX A
Alternate Multichannel Model Forms

Multichannel generalization of the MA and AR models were discussed in Chapter II along with their solution via the Levinson algorithm. Here the multichannel models and the Levinson algorithm are developed in an alternate form more compatible with other linear and nonlinear modeling problems to be explored later.

Consider first an N-th order, Q₀-channel, AR model where the current value of the signal vector \( x(k) \)

\[
x(k) = [x_1(k) \ldots x_{Q_0}(k)]^T
\]  

\( (A.1a) \)

\[
X(z) = [X_1(z) \ldots X_{Q_0}(z)]^T
\]  

\( (A.1b) \)

is to be predicted from a weighted combination of N past values of each of the component signals. For each signal this estimate can be written as

\[
x_j(k) = \sum_{i=1}^{Q_0} \sum_{n=1}^{N} d_{ij}(n) x_i(k-n)
\]  

\( (A.2a) \)

or

\[
\hat{x}_j(k) = \sum_{i=1}^{Q_0} d_{ij}(z) X_i(z)
\]  

\( (A.2b) \)

where

\[
d_{ij}(z) = \sum_{n=1}^{N} d_{ij}(n) z^{-n}
\]  

\( (A.2c) \)
Define an \( N \)-vector for each of the \( Q_0 \) channels to contain their required time histories as

\[
x_i(k) = [x_i(k-1) \ldots x_i(k-N)]^T
\]

for \( 1 \leq i \leq Q_0 \) and embed all of these vectors into a single \( NQ_0 \)-data vector given by:

\[
X(k) = [x_1(k)^T \ldots x_{Q_0}(k)^T]^T
\]

Define a \( NQ_0 \times Q_0 \) matrix of weights as

\[
D = \begin{bmatrix}
  d_{11} & \ldots & d_{1Q_0} \\
  \vdots & \ddots & \vdots \\
  d_{Q_01} & \ldots & d_{Q_0Q_0}
\end{bmatrix}
\]

where the \( N \)-vectors \( d_{ij} \) are given by

\[
d_{ij} = [d_{ij}(1) \ldots d_{ij}(N)]^T
\]

and contain the coefficients of the polynomials \( d_{ij}(z) \). These polynomials can be combined in a matrix polynomial defined by

\[
D(z) = \begin{bmatrix}
  d_{11}(z) & \ldots & d_{1Q_0}(z) \\
  \vdots & \ddots & \vdots \\
  \vdots & \ddots & \vdots \\
  d_{Q_01}(z) & \ldots & d_{Q_0Q_0}(z)
\end{bmatrix}
\]
With these definitions, the N-th order prediction of \( x(k) \) and its associated prediction error vector becomes

\[
\hat{x}(k)^T = X(k)^T D \quad \text{(A.5a)}
\]

\[
e(k)^T = x(k)^T - \hat{x}(k)^T \quad \text{(A.5b)}
\]

or in the transform domain

\[
E(z)^T = X(z)^T [I - D(z)] \quad \text{(A.5c)}
\]

Comparison of equation (A.5c) and (2.44c) show that this matrix polynomial differs from the more generally used form of (2.44c) by a transposition. The coefficient matrix \( D \) can be found by minimizing the trace of the prediction error covariance matrix

\[
P = \varepsilon\{e(k) e(k)^T\} \quad \text{(A.6)}
\]

leading to a system of linear equations given by

\[
\varepsilon\{X(k) X(k)^T\} D = \varepsilon\{X(k) x(k)^T\} \quad \text{(A.7a)}
\]

or

\[
\begin{bmatrix}
R_{x_1 x_1} & \cdots & R_{x_1 x_Q0} \\
\vdots & \ddots & \vdots \\
R_{x_Q0 x_1} & \cdots & R_{x_Q0 x_Q0}
\end{bmatrix}
\begin{bmatrix}
D_{11} & \cdots & D_{1Q0} \\
\vdots & \ddots & \vdots \\
D_{Q1} & \cdots & D_{Q0 Q0}
\end{bmatrix}
\]

\[
A.7b)
\]
Adopting a shorthand notation, this becomes

$$\mathbf{R} \mathbf{D} = \mathbf{r}$$  \hspace{1cm} (A.7c)

As in the case of a single channel autoregression, the multichannel generalization of the Levinson algorithm also requires that the backward prediction problem of estimating \(x(k-N)\) backward in time from the values of \(x(k-N+1)\) through \(x(k)\), be solved simultaneously. Using an overscore to indicate quantities associated with the backward problem it follows that

$$\mathbf{X}(k)^T \mathbf{D} = \mathbf{\hat{x}}(k)^T$$  \hspace{1cm} (A.8a)

$$\mathbf{\hat{\epsilon}}(k)^T = x(k-N)^T - \mathbf{\hat{x}}(k)^T$$  \hspace{1cm} (A.8b)

or

$$\mathbf{\hat{\epsilon}}(z)^T = \mathbf{X}(z)^T z^{-N} [\mathbf{I} - \mathbf{D}(z)]$$  \hspace{1cm} (A.8c)

where

$$\mathbf{X}(k) = [\mathbf{X}_1(k)^T \ldots \mathbf{X}_Q(k)^T]^T$$  \hspace{1cm} (A.8d)

$$\mathbf{X}_i(k) = [x_i(k-N+1) \ldots x_i(k)]^T$$  \hspace{1cm} (A.8e)

and

$$\mathbf{D} = \begin{bmatrix}
\mathbf{D}_{11} & \ldots & \mathbf{D}_{1Q_0} \\
\vdots & \ddots & \vdots \\
\mathbf{D}_{Q_01} & \ldots & \mathbf{D}_{Q_0Q_0}
\end{bmatrix}$$  \hspace{1cm} (A.8f)
\[ \bar{d}_{ij} = [\bar{d}_{ij}(1) \ldots \bar{d}_{ij}(N)]^T \]  
(A.8g)

with \[ \bar{D}(z) = \begin{bmatrix} \bar{d}_{11}(z) & \bar{d}_{10}(z) \\ \vdots & \vdots \\ \bar{d}_{01}(z) & \bar{d}_{00}(z) \end{bmatrix} \]  
(A.8h)

and \[ \bar{d}_{ij}(z) = \sum_{n=1}^{N} \bar{d}_{ij}(n) z^n \]  
(A.8i)

Setting the coefficients of this backward predictor to minimize the trace of the backward prediction error covariance matrix

\[ P = \varepsilon\{\overline{e}(k) \overline{e}(k)^{T}\} \]  
(A.9)

leads to a MMSE solution given by

\[ \varepsilon\{\overline{X}(k) \overline{X}(k)^{T}\} \bar{D} = \varepsilon\{\overline{X}(k) x(k-N)^{T}\} \]  
(A.10a)

and since \[ \varepsilon\{\overline{x}_i(k) \overline{x}_j(k)^{T}\} = R_{x_i x_j} \]

and \[ \varepsilon\{\overline{x}_i(k) x_j(k-N)^{T}\} = R_{x_j x_i} \]

this can be written as
Adopting a shorthand notation, this becomes

$$\mathbf{RD} = \mathbf{f}$$ (A.11)

At this point it is important to take note of a subtle difference between the single and multichannel problems that has arisen. While the single channel equivalent of equation (A.10b) for the backward predictor was not written previously, it is obviously

$$\mathbf{R}_{\mathbf{x}_{1}x_{1}}^T \mathbf{D} = \mathbf{R}_{\mathbf{x}_{1}x_{1}}$$ (A.12a)

and since

$$\mathbf{R}_{\mathbf{x}_{1}x_{1}}^T = \mathbf{R}_{\mathbf{x}_{1}x_{1}}$$ (A.12b)

it follows that the single channel forward and backward prediction problems have exactly the same solution (as was found to be the case in Section II.D. when the AR prediction error lattice was developed). This fact is responsible for
a number of simplifications in the single channel case:

a.) Only a single reflection coefficient $k^{(n+1)}$ was needed to link the $n$-th and $(n+1)$-st order solutions for both the backward and forward problems (see equations (2.25b) and (2.25c)).

b.) $|B(z)| = |\mathbf{B}(z)|$

c.) $\epsilon(e(k)^2) = \epsilon(\mathbf{e}(k)^2)$

d.) The development of the Burg method and the Itakura-Saito method for calculating the reflection coefficients are a direct consequence of c above.

Unfortunately however, none of these simplifications carry over into the more general multichannel AR problem because by comparing equations (A.7b) and (A.10b) it is evident that in general $D \neq \overline{D}$.

Proceeding as in the single channel case, it is shown in Appendix B that because of the structure in the correlation matrices $\mathbf{R}$ and $\overline{\mathbf{R}}$, equations (A.7c) and (A.11) can be re-written as

\[
\overline{\mathbf{R}} \mathbf{F} = \rho \quad \text{(A.13a)}
\]

and

\[
\mathbf{R} \overline{\mathbf{F}} = \overline{\rho} \quad \text{(A.13b)}
\]

where $\mathbf{F}$, $\overline{\mathbf{F}}$, $\rho$ and $\overline{\rho}$ are obtained from $\mathbf{D}$, $\overline{\mathbf{D}}$, $\mathbf{r}$ and $\overline{\mathbf{r}}$ respectively by taking their component vectors and turning them upside down in place; i.e.
where

$$F = \begin{bmatrix} f_{11} & \ldots & f_{1Q_0} \\ \vdots & & \vdots \\ f_{Q_01} & \ldots & f_{Q_0Q_0} \end{bmatrix}$$

(A.14a)

and

$$\phi = \begin{bmatrix} \phi_{x_1x_1} & \ldots & \phi_{x_1x_{Q_0}} \\ \vdots & & \vdots \\ \phi_{x_{Q_0}x_1} & \ldots & \phi_{x_{Q_0}x_{Q_0}} \end{bmatrix}$$

(A.14c)

with

$$\phi_{x_1x_j} = [R_{x_i^T}x_j^T(N) \ldots R_{x_i^T}x_j^T(1)]^T$$

(A.14d)

As will soon be evident, the relationships of equations (A.13) form the cornerstone for the Levinson algorithm and are made possible because the blocks comprising the $R$ and $\bar{R}$ matrices are themselves Toeplitz and because $R_{x_i^T}x_j^T = R_{x_j^T}x_i^T$.

Assume that the $(n+1)$-st order solutions can be related to the $n$-th order solutions by

$$a_{ij}^{(n+1)} = \begin{bmatrix} d_{ij}^{(n)} \\ 0 \end{bmatrix} + \begin{bmatrix} e_{ij}^{(n)} \\ - e_{ij}^{(n+1)} \end{bmatrix}$$

(A.15a)
and

\[ \mathbf{d}_{ij}^{(n+1)} = \begin{bmatrix} \mathbf{d}_{ij}^{(n)} \\ \mathbf{e}_{ij}^{(n)} \\ 0 \\ \mathbf{K}_{ij}^{(n+1)} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_{ij}^{(n)} \\ \mathbf{K}_{ij}^{(n+1)} \end{bmatrix} \]  
(A.15b)

Embedding the vectors \( \mathbf{e}_{ij}^{(n)} \) and \( \mathbf{d}_{ij}^{(n)} \) and the coefficients \( k_{ij}^{(n+1)} \) and \( \mathbf{K}_{ij}^{(n+1)} \) into matrices designated \( \mathbf{e}_{ij}^{(n)} \) and \( \mathbf{F}_{ij}^{(n)} \), \( K_{ij}^{(n+1)} \) and \( \mathbf{K}_{ij}^{(n+1)} \) and solving for them in the \((n-1)\)st order problem it is shown in Appendix C that

\[ \mathbf{e}_{ij}^{(n)} = - \mathbf{F}_{ij}^{(n)} \mathbf{K}_{ij}^{(n+1)} \]  
(A.16a)

\[ \mathbf{e}_{ij}^{(n)} = - \mathbf{F}_{ij}^{(n)} \mathbf{K}_{ij}^{(n+1)} \]  
(A.16b)

\[ K_{ij}^{(n+1)} = [R_{xx}(0) - \mathbf{F}_{ij}^{(n)} \mathbf{D}_{ij}^{(n)}]^{-1} \begin{bmatrix} R_{xx}(n+1) - \mathbf{F}_{ij}^{(n)} \mathbf{D}_{ij}^{(n)} \\ \mathbf{F}_{ij}^{(n)} \mathbf{D}_{ij}^{(n+1)} \end{bmatrix} \]  
(A.16c)

\[ K_{ij}^{(n+1)} = [R_{xx}(0) - \mathbf{F}_{ij}^{(n)} \mathbf{D}_{ij}^{(n)}]^{-1} \begin{bmatrix} R_{xx}(n+1) - \mathbf{F}_{ij}^{(n)} \mathbf{D}_{ij}^{(n)} \\ \mathbf{F}_{ij}^{(n)} \mathbf{D}_{ij}^{(n+1)} \end{bmatrix} \]  
(A.16d)

Also the inverted terms on the right hand side of equations (A.16c) and (A.16d) are just the backward and forward prediction error covariance matrices for the optimum \( n \)-th order models so that (A.16c) and (A.16d) become
\begin{align}
    K^{(n+1)} &= P(n)^{-1} [R_{xx}(n+1) - \bar{\varrho}(n)^T D(n)] \quad \text{(A.17a)} \\
    R^{(n+1)} &= P(n)^{-1} [R_{xx}(n+1)^T - \bar{\varrho}(n)^T D(n)] \quad \text{(A.17b)}
\end{align}

As in the case of the single channel problem, the prediction error covariance matrices obey the recursions.

\begin{align}
    P^{(n+1)} &= P(n) [I - K^{(n+1)} R^{(n+1)}] \quad \text{(A.18a)} \\
    \bar{P}^{(n+1)} &= \bar{P}(n) [I - \bar{K}^{(n+1)} \bar{R}^{(n+1)}] \quad \text{(A.18b)}
\end{align}

completing the multichannel generalization of the Levinson algorithm for AR models.

Comparing equations (A.16), (A.17) and (A.18) with their single channel counterparts (2.18) and (2.19) it is clear that the multichannel Levinson algorithm simply represents a matrix algebra generalization of the single channel algorithm. Once again, predictors of all orders 0 ≤ n ≤ N are obtained in the process of finding the N-th order predictor along with all their prediction error covariance matrices and the overall minimization requiring the inversion of a \(Q_0N \times Q_0N\) matrix is replaced by a sequence of N minimizations, each requiring the inversion of two \(Q_0 \times Q_0\) matrices \(P(n)\) and \(\bar{P}(n)\).
Using the relationships given in equations (A.15) and (A.16), successive orders of matrix polynomials can also be related to one another by

\[ [I-b(n+1)(z)] = [I-b(n)(z)]-z^{-1}[I-b(n)(z)]K(n+1) \]  
(A.19a)

\[ z^{-n}[I-b(n+1)(z)] = z^{-1}z^{-n}[I-b(n)(z)]-z^{-n}[I-b(n)(z)]K(n+1) \]  
(A.19b)

Premultiplying both sides of these equations by \( X(z)^T \), transposing, and transforming into the time domain provides relationships between the prediction error signals at each stage.

\[ e(n+1)(k) = e(n)(k) - K(n+1)^T \bar{e}(n)(k-1) \]  
(A.20a)

\[ \bar{e}(n+1)(k) = \bar{e}(n)(k-1) - \bar{K}(n+1)^T \bar{e}(n)(k) \]  
(A.20b)

Recognizing that for a zeroth order prediction, the forward and backward prediction error vectors are just the input vector itself, it follows that

\[ e(0)(k) = \bar{e}(0)(k) = x(k) \]  
(A.20c)

and the multichannel equivalents of equations (2.28) have been obtained.
Next consider the N-th order multichannel MA model in which the current value of a \( Q_0 \)-vector of output signals \( y(k) \) is to be predicted from the present value and \( N \) past values of a \( Q_i \)-vector of input signals \( x(k) \).

\[
\hat{y}_j(k) = \sum_{i=1}^{Q_i} \sum_{n=0}^{N} d_{ij}(n) x_i(k-n) \quad \text{(A.21a)}
\]

or

\[
\hat{y}_j(z) = \sum_{i=1}^{Q_i} d_{ij}^+(z) X_i(z) \quad \text{(A.21b)}
\]

where

\[
d_{ij}^+(z) = \sum_{n=0}^{N} d_{ij}(n) z^{-n} \quad \text{(A.21c)}
\]

Using a superscript "+" to indicate the fact that the vectors are indexed from 0 to \( N \) rather than from 1 to \( N \), define \((N+1)\)-vectors for each of the input channels to contain their required time histories as

\[
x_j^+(k) = [x_j(k) \ldots x_j(k-n)]^T \quad \text{(A.22a)}
\]

and embed these vectors into a single \( Q_i(N+1) \)-data vector given by

\[
X^+(k) = [x_1^+(k)^T \ldots x_{Q_i}^+(k)^T]^T \quad \text{(A.22b)}
\]
Define a $Q_i(N+1) \times Q_0$ matrix of weights as

$$D^+ = \begin{bmatrix}
    d_{11}^+ & \cdots & d_{1Q_0}^+ \\
    \vdots & \ddots & \vdots \\
    d_{Q_i1}^+ & \cdots & d_{Q_iQ_0}^+
\end{bmatrix} \tag{A.23a}$$

where the $(N+1)$-vectors $d_{ij}^+$ are given by

$$d_{ij}^+ = [d_{ij}(0) \ldots d_{ij}(N)]^T \tag{A.23b}$$

and contain the coefficients of the polynomial $d_{ij}(z)$. These polynomials can be combined into a single matrix polynomial given by

$$D^+(z) = \begin{bmatrix}
    d_{11}^+(z) & \cdots & d_{1Q_0}^+(z) \\
    \vdots & \ddots & \vdots \\
    d_{Q_i1}^+(z) & \cdots & d_{Q_iQ_0}^+(z)
\end{bmatrix} \tag{A.23c}$$

With these definitions, the $N$-th order MA prediction of $y(k)$ is given by

$$\hat{y}(k)^T = X^+(k)^T D^+ \tag{A.24a}$$
or in the transform domain,

\[ \hat{\gamma}(z)^T = \hat{X}(z)^T D^+(z) \]  

(A.24b)

Defining a prediction error vector as

\[ e_0(k)^T = y(k)^T - \hat{y}(k)^T \]  

(A.25a)

and setting the coefficients in \( D^+ \) to minimize the trace of the prediction error covariance matrix

\[ P_0 = \varepsilon \{ e_0(k) e_0(k)^T \} \]  

(A.25b)

results in a solution given by

\[ \varepsilon \{ X^+(k) \ X^+(k)^T \} D^+ = \varepsilon \{ X^+(k) \ y(k)^T \} \]  

(A.26a)

or

\[
\begin{bmatrix}
R_{x_1x_1}^+ & \cdots & R_{x_1x_Q}^+ \\
R_{x_1x_1}^+ & \cdots & R_{x_1x_Q}^+ \\
\vdots & \ddots & \vdots \\
R_{x_Qx_1}^+ & \cdots & R_{x_Qx_Q}^+
\end{bmatrix}
\]

\[
\begin{bmatrix}
P_{x_1y_1}^+ & \cdots & P_{x_1y_Q}^+ \\
P_{x_1y_1}^+ & \cdots & P_{x_1y_Q}^+ \\
\vdots & \ddots & \vdots \\
P_{x_Qy_1}^+ & \cdots & P_{x_Qy_Q}^+
\end{bmatrix}
\]

\[ D^+ = \]  

(A.26b)

\[ \begin{bmatrix}
P_{x_1y_1}^+ & \cdots & P_{x_1y_Q}^+ \\
P_{x_1y_1}^+ & \cdots & P_{x_1y_Q}^+ \\
\vdots & \ddots & \vdots \\
P_{x_Qy_1}^+ & \cdots & P_{x_Qy_Q}^+
\end{bmatrix}
\]
Adopting a shorthand notation this becomes

\[
\mathbf{R}^+ \mathbf{D}^+ = \mathbf{r}^+ \quad (A.26c)
\]

Assume a relationship between the components of the \((n+1)\)-st and \(n\)-th order solutions given by

\[
d^{+(n+1)}_{ij} = \begin{bmatrix} d^{+(n)}_{ij} \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} Y_{ij}^{(n+1)} \\ \vdots \\ g_{ij}^{(n+1)} \end{bmatrix} \quad (A.27)
\]

Embedding the vectors \(Y_{ij}^{(n+1)}\) and the coefficients \(g_{ij}^{(n+1)}\) into matrices, designated \(\mathbf{J}^{(n+1)}\) and \(\mathbf{G}^{(n+1)}\), and solving for them in the \((n+1)\)-st order MA problem it is shown in Appendix D that

\[
\mathbf{J}^{(n+1)} = -\mathbf{F}^{(n+1)} \mathbf{G}^{(n+1)} \quad (A.28a)
\]

\[
\mathbf{G}^{(n+1)} = \mathbf{P}^{(n+1)^{-1}} \left[ \mathbf{R}_{xy}^{(n+1)} - \mathbf{\rho}^{(n+1)^T} \mathbf{D}^+(n) \right] \quad (A.28b)
\]

where \(\mathbf{F}^{(n+1)}\), \(\mathbf{\rho}^{(n+1)}\), and \(\mathbf{P}^{(n+1)}\) are matrices that emerge from the back prediction problem in a multichannel \((n+1)\)-st order autoregression on the input signal vector \(\mathbf{x}(k)\). Again, it is clear that the multichannel MA solutions given by equations (A.28) are a matrix algebra generalization of equations (2.23) for the single channel MA model.
To relate successive order MA matrix polynomials to one another, equations (A.27) and (A.28a) can be used to write

\[ D^+(z)^{(n+1)} = D^+(z)^{(n)} + z^{-n} \{ -D(z)^{(n)} G^{(n+1)} \} \]  
(A.29)

where the second term on the right hand side that premultiplies \( G \) is the backward prediction error matrix polynomial from the autoregression on the input signal \( x(k) \). Pre-multiplying by \( X^T(z) \), transposing, and transforming into the time domain results in the multichannel equivalent of equation (2.37)

\[ \hat{y}^{(n+1)}(k) = \hat{y}^{(n)}(k) + G^{(n+1)} e^{(n+1)}(k) \]  
(A.30)

and completes the derivation of the recursive in order solutions for the multichannel MA model.
APPENDIX B

A Key Relationship For The Multichannel Levinson Algorithm

Equation (A.13) is a key relationship in the development of the Levinson algorithm responsible for much of the algorithm's simplicity. Without this relationship, more than just the $K$ and $\tilde{K}$ matrices would be needed to obtain the new model from the previous model.

In equation (A.7), consider the multiplication of the $i$-th block row of $\mathbf{R}$ by the $j$-th block column of $\mathbf{D}$ to form $r_{x_i x_j}$.

$$r_{x_i x_j} \mathbf{d}_{ij} + \ldots + r_{x_i x_{Q_0}} \mathbf{d}_{Q_0 j} = r_{x_i x_j} \quad (B.1)$$

In particular, consider a general term on the left hand side of equation (B.1) in detail.

\[
\begin{bmatrix}
  r_{x_i x_{m}}(0) & \ldots & r_{x_i x_{m}}(1-N) \\
  \vdots & \ddots & \vdots \\
  r_{x_i x_{m}}(N-1) & r_{x_i x_{m}}(0)
\end{bmatrix}
\begin{bmatrix}
  d_{m_j}(1) \\
  \vdots \\
  d_{m_j}(N)
\end{bmatrix}
\]

\[
= \begin{bmatrix}
  r_{x_i x_j}(1) \\
  \vdots \\
  r_{x_i x_j}(N)
\end{bmatrix}
\]

(B.2)

Define upside down versions of the $d_{ij}$ and $r_{x_i x_j}$ vectors as

\[
\mathbf{d}_{ij} = \begin{bmatrix}
  d_{ij}(N) \\
  \vdots \\
  d_{ij}(1)
\end{bmatrix}
\]

(B.3a)
and
\[ \mathbf{p}_{x_i x_j} = \begin{bmatrix} R_{x_i x_j}^{(N)} \\ \vdots \\ R_{x_i x_j}^{(1)} \end{bmatrix} \]  
(B.3b)

Using these permuted vectors in equation (B.1) in place of the \( d \) and \( r \) vectors, the relationship is still satisfied if the \( R \) matrices are permuted as well. In particular, from (B.2) it is evident that

\[
\begin{bmatrix} R_{x_i x_m}^{(0)} & \cdots & R_{x_i x_m}^{(N-1)} \\ R_{x_i x_m}^{(1-N)} & \cdots & R_{x_i x_m}^{(0)} \end{bmatrix} \begin{bmatrix} d_{ij}^{(N)} \\ d_{ij}^{(1)} \end{bmatrix} 
\]

or equivalently

\[
\mathbf{R}^T \mathbf{f}_{ij} + \cdots + \mathbf{R}^T \mathbf{f}_{Q_0 j} = \mathbf{p}_{x_i x_j} \]  
(B.4b)

Embedding the \( \mathbf{f}_{ij} \) and \( \mathbf{p}_{x_i x_j} \) vectors into matrices designated \( \mathbf{F} \) and \( \mathbf{P} \) respectively and using the definition of \( \mathbf{R} \) it follows that

\[
\mathbf{R} \mathbf{F} = \mathbf{P} \]  
(B.5)
Defining \( f_{ij} \) and \( \bar{x}_i x_j \) as upside down versions of their corresponding vectors in \( \bar{D} \) and \( \bar{r} \) in equation (A.11) and embedding them in matrices designated \( \bar{F} \) and \( \bar{\rho} \), it also follows from a similar development that

\[
R \bar{F} = \bar{\rho} \tag{B.6}
\]

Equations (B.5) and (B.6) are essential to the development of the Levinson algorithm and are made possible by the Toeplitz structure of the block components of the \( R \) and \( \bar{R} \) matrices.
APPENDIX C

The Multichannel Levinson Algorithm Solution
For AR Modeling

In the \((n+1)\)-st order versions of equations (A.7) and (A.10), the component matrices that make up \(\mathbf{R}\) and \(\overline{\mathbf{R}}\) can be written as follows

\[
\begin{align*}
\mathbf{R}_{x_ix_j}^{(n+1)} &= \begin{bmatrix}
\mathbf{R}_{x_ix_j}^{(n)} & \mathbf{R}_{x_ix_j}^{(-N)} \\
\mathbf{R}_{x_ix_j}^{(-N)} & \vdots \\
\mathbf{R}_{x_ix_j}(N) \cdots & \mathbf{R}_{x_ix_j}(1) \\
\mathbf{R}_{x_ix_j}(1) & \mathbf{R}_{x_ix_j}(0)
\end{bmatrix} \\
\mathbf{R}_{x_ix_j} &= \begin{bmatrix}
\mathbf{R}_{x_ix_j}^{(n)} & \mathbf{\overline{R}}_{x_ix_j}^{(n)} \\
\mathbf{\overline{R}}_{x_ix_j}^{(n)} & \mathbf{\overline{R}}_{x_ix_j}^{(n)} \\
\mathbf{\overline{R}}_{x_ix_j}(n) & \mathbf{\overline{R}}_{x_ix_j}(0)
\end{bmatrix}
\end{align*}
\]

and

\[
\begin{align*}
\mathbf{R}_{x_ix_j}^{(n+1)} &= \mathbf{R}_{x_ix_j}^{(n+1)^T} = \begin{bmatrix}
\mathbf{R}_{x_ix_j}^{(n)^T} & \mathbf{\overline{R}}_{x_ix_j}^{(n)^T} \\
\mathbf{\overline{R}}_{x_ix_j}^{(n)^T} & \mathbf{\overline{R}}_{x_ix_j}^{(n)^T} \\
\mathbf{\overline{R}}_{x_ix_j}(n) & \mathbf{\overline{R}}_{x_ix_j}(0)
\end{bmatrix}
\end{align*}
\]
Additionally,

\[ r_{x_i x_j}^{(n+1)} = \begin{bmatrix} r_{x_i x_j}^{(n)} \\ r_{x_i x_j}^{(n+1)} \end{bmatrix} \] (C.3)

With these matrices and vectors written in this partitioned form, and with the relationships assumed in equations (A.15) between the n-th and (n+1)-st order solutions, the (n+1)-st order modeling equations become:

**Forward Model:**

\[ R^{(n)} D^{(n)} + R^{(n)} \epsilon^{(n)} + \rho^{(n)} K^{(n+1)} = r^{(n)} \] (C.4a)

\[ \rho^{(n)^T} D^{(n)} + \rho^{(n)^T} \epsilon^{(n)} + R_{xx}(0) K^{(n+1)} = R_{xx}^{(n+1)} \] (C.4b)

**Backward Model:**

\[ \overline{R}^{(n)} \overline{D}^{(n)} + \overline{R}^{(n)^T} \overline{\epsilon}^{(n)} + \overline{\rho}^{(n)} \overline{K}^{(n+1)} = \overline{r}^{(n)} \] (C.5a)

\[ \overline{\rho}^{(n)^T} \overline{D}^{(n)} + \overline{\rho}^{(n)^T} \overline{\epsilon}^{(n)} + R_{xx}(0) \overline{K}^{(n+1)} = R_{xx}^{(n+1)^T} \] (C.5b)
Equations (C.4a) and (C.5a) contain the n-th order modeling equations within them however, and therefore can be written as

\[ \mathbf{R}^{(n)} \mathbf{e}^{(n)} = - \mathbf{\tilde{R}}^{(n)} \mathbf{K}^{(n+1)} \]  
\[ \mathbf{\tilde{R}}^{(n)} \mathbf{\tilde{e}}^{(n)} = - \mathbf{\rho}^{(n)} \mathbf{\tilde{K}}^{(n+1)} \]

(C.6a)

(C.6b)

and applying the relationship developed in Appendix B (equations (B.5) and (B.6)) yields

\[ \mathbf{e}^{(n)} = - \mathbf{\bar{F}}^{(n)} \mathbf{K}^{(n+1)} \]  
\[ \mathbf{\bar{e}}^{(n)} = - \mathbf{\bar{F}}^{(n)} \mathbf{\bar{K}}^{(n+1)} \]

(C.7a)

(C.7b)

Thus the \( \mathbf{e} \) and \( \mathbf{\bar{e}} \) vectors have been found in terms of the known quantities \( \mathbf{f} \) and \( \mathbf{\bar{f}} \) and the unknown \( \mathbf{K} \) and \( \mathbf{\bar{K}} \) matrices. Substituting equations (C.7) into (C.4b) and (C.5b) completes the solution resulting in expressions for the \( \mathbf{K} \) and \( \mathbf{\bar{K}} \) matrices given by

\[ \mathbf{K}^{(n+1)} = \left[ \mathbf{R}_{xX}(0) - \mathbf{\bar{F}}^{(n)T} \mathbf{D}^{(n)} \right]^{-1} \left[ \mathbf{R}_{xX}(n+1) - \mathbf{\rho}^{(n)T} \mathbf{D}^{(n)} \right] \]

(C.8a)
The terms on the right hand side of these equations that are inverted are just the backward and forward prediction error covariance matrices for the optimum n-th order models given by

\[
\bar{\mathbf{P}}(n+1) = [\mathbf{R}_{xx}(0) - \mathbf{r}(n)^T \mathbf{D}(n)]^{-1} [\mathbf{R}_{xx}(n+1) T - \mathbf{\phi}(n)^T \mathbf{D}(n)]
\]

(C.9b)

Using equations (C.3), (A.15) and (C.7), the forward prediction error covariance matrix for the optimum (n+1)-st order model can therefore be written as

\[
\mathbf{P}(n+1) = \mathbf{R}_{xx}(0) - \mathbf{r}(n)^T \mathbf{D}(n) + \mathbf{r}(n)^T \mathbf{F}(n) \mathbf{K}(n+1)
\]

(C.10a)

\[
= \mathbf{P}(n) + [\mathbf{r}(n)^T \mathbf{F}(n) - \mathbf{R}_{xx}(n+1) T] \mathbf{K}(n+1)
\]

(C.10b)

\[
= \mathbf{P}(n) [\mathbf{I} - \mathbf{P}(n)^{-1} [\mathbf{R}_{xx}(n+1) T - \mathbf{\phi}(n)^T \mathbf{D}(n)] \mathbf{K}(n+1)]
\]

(C.10c)
\[ p(n+1) = p(n) [I - K(n+1)K(n+1)] \quad (C.10d) \]

and following a similar development for the backward prediction error covariance matrix results in

\[ \tilde{p}(n+1) = \tilde{p}(n) [I - K(n+1)K(n+1)] \quad (C.11) \]
APPENDIX D

The Multichannel Levinson Algorithm Solution
For MA Modeling

First note that because of the definitions of the $x_i^+(k)$ and $x_i(k)$ vectors (indexed from 0 to n and 1 to n respectively) in the n-th order models, the matrices $R_{x_i^+x_j^+}^{(n)}$ in the n-th order MA problem could also be written in terms of $x_i(k)$ and $x_j(k)$ as (assuming stationarity)

$$R_{x_i^+x_j^+}^{(n)} = R_{x_i^+x_j}^{(n+1)}$$  \hspace{1cm} (D.1a)

so that

$$R_{x_i^+x_j}^{(n+1)} = R_{x_i^+x_j}^{(n)}$$  \hspace{1cm} (D.1b)

The components of $R_{x_i^+x_j}^{(n+1)}$ and $R_{x_i^+x_j}^{(n)}$ in the (n+1)-st order MA modeling problem can be written in partitioned form as

$$R_{x_i^+x_j}^{(n+1)} = \begin{bmatrix} R^{(n)}_{x_i^+x_j^+} & \overline{R}^{(n+1)}_{x_i^+x_j} \\ \overline{R}^{(n+1)}_{x_i^+x_j} & R_{x_i^+x_j}^{(n)} \end{bmatrix}$$ \hspace{1cm} (D.2a)
Using these partitioned forms, and the relationship in equation (A.27) between the n-th and (n+1)-st order solutions, the (n+1)=st order modeling equations become

\[
\begin{align*}
R^+(n) D^+(n) & + R^+(n) \tilde{\eta}^{(n)} + \overline{\sigma}(n+1) G^{(n+1)} = R^+(n) \\
\overline{\sigma}(n+1)^T D^+(n) & + \overline{\sigma}(n+1)^T \tilde{\eta}^{(n+1)} + R_{xx}(0) G^{(n+1)} \\
& = R_{xy}(n+1)
\end{align*}
\]

Equation (D.3a) contains within it, the modeling equation for the n-th order MA model and using equation (D.1b) can be rewritten as

\[
\begin{align*}
R(n+1) \tilde{\eta}^{(n+1)} & = - \overline{\sigma}(n+1) G^{(n+1)} \\
& = R_{xy}(n+1)
\end{align*}
\]

A comparison of this result with equation (C.6a) shows that

\[
\tilde{\eta}^{(n+1)} = - \overline{F}^{(n+1)} G^{(n+1)}
\]
where $\mathbf{F}^{(n+1)}$ is the permuted version of the backward solution in an $(n+1)$-st order AR model of the input signal vector $x(k)$. Furthermore, substituting equation (D.5) into (D.3b) results in a solution for $g^{(n+1)}$ given by

$$g^{(n+1)} = \left[ R_{xx}(0) - \rho^{(n+1)T} \mathbf{F}^{(n+1)} \right]^{-1}$$

$$\left[ R_{xy}^{(n+1)} - \rho^{(n+1)T} \mathbf{D}^{+}(n) \right]$$

(D.6)

Since

$$\rho^{(n+1)T} \mathbf{F}^{(n+1)} = \rho^{(n+1)T} \mathbf{D}^{(n+1)}$$

it follows that the inverted term on the right hand side of equation (D.6) is just the backward prediction error covariance matrix for the optimum $(n+1)$-st order AR model on the input signal vector $x(k)$. 

229
APPENDIX F

Prony's Method For ARMA Modeling

Prony's method [Refs. 8, 52 and 56] obtains a zero pole model for a system by matching the impulse responses of the system and model over the first \( M+N+1 \) sample intervals where \( M \) and \( N \) are the orders of the model transfer function numerator and denominator polynomials. Assume that a signal \( y(k) \) is available that represents the impulse response of a causal system and that a rational transfer function model for this systems is desired. Using a \( \hat{y} \) to denote the model output and \( u(k) \) to denote the input to both the system and model, the model transfer function is given by

\[
H(z) = \frac{\hat{Y}(z)}{U(z)} = \frac{\sum_{n=0}^{M} a(n)z^{-n}}{1 + \sum_{n=1}^{N} b(n)z^{-n}} \quad (E.1)
\]

For a unit sample input it follows that

\[
B(z) \hat{Y}(z) = A(z) \quad (E.2)
\]

Equating like powers of \( z \) in this relationship results in

\[
\sum_{i=0}^{N} b(i) \hat{y}(n-i) = \begin{cases} 
   a(n) ; & 0 \leq n \leq M \\
   0 & \text{else}
\end{cases} \quad (E.3)
\]

where \( b(0)=1 \).
Equating \( \hat{y}(k) \) and \( y(k) \) over the interval \( 0 \leq k \leq M+N \) produces a set of \( M+N+1 \) equations which can be expressed in matrix form as

\[
\begin{bmatrix}
 y(0) & 0 & 0 & \ldots & 0 \\
 y(1) & y(0) & 0 & \ldots & 0 \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 y(M) & y(M-1) & y(M-2) & \ldots & 0 \\
 y(M+1) & y(M) & y(M-1) & \ldots & 0 \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 y(M+N) & \ldots & \ldots & \ldots & y(M)
\end{bmatrix}
\begin{bmatrix}
 a(0) \\
 b(1) \\
 \vdots \\
 b(N) \\
 a(M)
\end{bmatrix} =
\begin{bmatrix}
 1 \\
 b(1) \\
 \vdots \\
 b(N) \\
 0
\end{bmatrix}
\]  
(E.4a)

or adopting a shorthand notation

\[
\begin{bmatrix}
 y_1^+ & y_1^- \\
 y_2 & y_2 \\
\end{bmatrix}
\begin{bmatrix}
 1 \\
 b
\end{bmatrix} =
\begin{bmatrix}
 a^+ \\
 0
\end{bmatrix}
\]  
(E.4b)

Assuming that the \( N \times N \) matrix \( Y_2 \) is not singular it follows that

\[
b = -Y_2^{-1} Y_2
\]  
(E.5a)

\[
a^+ = Y_1^+ - Y_1 Y_2^{-1} Y_2
\]  
(E.5b)
The original Prony method goes on to form a partial fraction expansion and inverse transformation on the model transfer function \( H(z) \) resulting in a model for the impulse response of the system given as a sum of complex exponentials. This is unnecessary here however, since a rational transfer function model is the form sought.

Prony's method inherently assumes that matching a sufficient portion of the impulse response of the system results in an accurate model but this is not necessarily the case unless the impulse response damps out quickly or unless the system can be represented exactly by a low order model. Otherwise a very high order model may be required to obtain sufficient accuracy. Other difficulties associated with this technique are:

1. The system impulse response must be available;
2. There are no built in mechanisms to test for or ensure stability of the model;
3. There is no averaging of noise in the data;
4. Only a small portion of the available data points \((M+N+1)\) are actually used.

These difficulties can be partially overcome by modifying the procedure to obtain an approximate match of the system and model responses over their entire duration rather than an exact match over the first \( M+N+1 \) points. Consider equation (E.4) modified to include the entire signal \( y(k) \) for \( 0 \leq k \leq \infty \).
Adopting a shorthand notation this becomes

\[
\begin{bmatrix}
  y(0) & 0 & \ldots & 0 \\
y(1) & y(0) & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
y(M) & y(M-1) & \ldots & y(M) \\
y(M+1) & y(M) & \ldots & \vdots \\
\end{bmatrix}
\begin{bmatrix}
1 \\
b \\
\vdots \\
0 \\
\vdots \\
\end{bmatrix}
= \begin{bmatrix}
a^+ \\
0 \\
\ddots \\
\vdots \\
\end{bmatrix}
\]  

(E.6a)

This yields two equations

\[
\begin{align*}
y_1^+ + y_1 b &= a^+ \\
y_3 + y_3 b &= 0
\end{align*}
\]  

(E.7a) 

(E.7b)

but with \(y_3, y_3\) and 0 having an infinite number of rows, equation (E.7b) will in general have no solution. In practice
only a finite portion of the system impulse response \( y(k) \) can be considered but equation (E.7b) will still be inconsistent in general. A least squares estimate of \( b \) can however, be obtained by minimizing \( e^T e \) where

\[
e = Y_3 \ b - (-Y_3)
\]

resulting in

\[
b = (Y_3^T Y_3)^{-1} Y_3^T Y_3
\]

which in turn can be used in equation (E.7a) to find \( a^+ \). This approximate version of Prony's method is the one most commonly applied.
APPENDIX F

Parabolic Surfaces In n-Dimensions

Multi-dimensional parabolic surfaces are described by an equation of the form

\[ y = x^T A x - 2 x^T b + c \quad (F.1) \]

where:
- \( y \) is the independent variable;
- \( x \) is a vector of dependent variables;
- \( A \) is a symmetric constant matrix;
- \( b \) is a constant vector;
- \( c \) is a constant.

\((x, b \text{ and } c \text{ can also be considered as matrices with the trace of the right hand side set equal to the independent variable but the problem remains essentially unchanged.)}\)

Completing the square for nonsingular \( A \) this becomes

\[ y = (x - A^{-1} b)^T A (x - A^{-1} b) + c - b^T A^{-1} b \quad (F.2) \]

so that for positive definite \( A \) it is clear that the minimum value of \( y \) is obtained for \( x = A^{-1} b \) and this minimum is \( c - b^T A^{-1} b \). It is also clear that nonzero values of \( b \) and \( c \) simply raise and lower the surface and move the minimum point away from \( x = 0 \). The shape of the surface (its relative concavity or flatness) is determined by the matrix
in the quadratic term of equation (F.1). Therefore, to study the shape of this surface, consider the simpler problem with \( b \) and \( c \) set to zero.

\[
y = x^T A x \tag{F.3}
\]

One way to examine the relative flatness or concavity is to look at the locus of points on the surface for constant values of \( y \); in particular when \( y = 1 \). Recognize that \( A \) can be rewritten as \( Q \Lambda Q^T \) where \( \Lambda \) is a diagonal matrix of eigenvalues and \( Q \) is a matrix whose columns are the unit length eigenvectors of \( A \). Now (F.3) becomes

\[
x^T Q \Lambda Q^T x = 1 \tag{F.4}
\]

and introducing a new set of variables \( w = Q^T x \) (which are simply a rotation of the variables in \( x \)), equation (F.4) reduces to

\[
\lambda_1 w_1^2 + \ldots + \lambda_n w_n^2 = 1 \tag{F.5}
\]

This equation describes an ellipsoid in \( n \) dimensions whose axes half lengths are given by \( 1/\sqrt{\lambda_i} \) for \( 1 \leq i \leq n \). This follows from letting all but one of the \( w \)'s equal to zero and solving for the nonzero variable so that one point on the surface is for example \( w_1 = 1/\sqrt{\lambda_1} \) with \( w_2 = \ldots = w_n = 0 \).
This point is just a multiple of the first eigenvector of $A$ so that in general, the axes of the ellipsoid point in the direction of the eigenvectors of $A$ with half lengths given by the reciprocal of the square root of the eigenvalues.
APPENDIX G
Multichannel ARMA Modeling

Consider a system with $Q_0$ output signals $[y_1(k) \ldots y_{Q_0}(k)]$ and $Q_i$ input signals $[u_1(k) \ldots u_{Q_i}(k)]$. The multichannel ARMA analysis model forms an estimate of the present value of each output as a weighted combination of past values of all output signals and past and present values of all input signals.

\[
y_n(k) = \sum_{j=1}^{Q_0} \sum_{i=1}^{N} b_{jn}(i) y_j(k-i) + \sum_{j=1}^{Q_i} \sum_{i=0}^{N} a_{jn}(i) u_j(k-i)
\]  

Define data vectors for all the input and output channels as

\[
y_n(k) = [y_n(k-1) \ldots y_n(k-N)]^T
\]  

\[
u_n^+(k) = [u_n(k) \ldots u_n(k-N)]^T
\]  

and embed them into $Q_0N$ and $Q_i(N+1)$ vectors given by

\[
\mathbf{Y} = [y_1(k)^T \ldots y_{Q_0}(k)^T]^T
\]  

\[
\mathbf{U}^+ = [u_1^+(k)^T \ldots u_{Q_i}^+(k)^T]^T
\]
Define $NQ_0 \times Q_0$ and $(N+1)Q_i \times Q_0$ matrices of coefficients given by

$$B = \begin{bmatrix} b_{11} & \cdots & b_{1Q_0} \\ \vdots & \ddots & \vdots \\ b_{Q_01} & \cdots & b_{Q_0Q_0} \end{bmatrix} \quad (G.4a)$$

and

$$A^+ = \begin{bmatrix} a_{11}^+ & \cdots & a_{1Q_0}^+ \\ \vdots & \ddots & \vdots \\ a_{Q_i1}^+ & \cdots & a_{Q_iQ_0}^+ \end{bmatrix} \quad (G.4b)$$

where

$$b_{ij} = [b_{ij}(1) \ldots b_{ij}(N)]^T \quad (G.4c)$$

$$a_{ij}^+ = [a_{ij}(0) \ldots a_{ij}(N)] \quad (G.4d)$$

With these definitions, the multichannel ARMA estimate for the vector of output signals becomes

$$\hat{\chi}(k)^T = [\chi^T \ U^+^T] \begin{bmatrix} B \\ A^+ \end{bmatrix} \quad (G.5)$$

Forming a prediction error vector as

$$e_8(k) = \chi(k) - \hat{\chi}(k) \quad (G.6)$$
and setting the model coefficients to minimize the trace of the prediction error covariance matrix results in a solution given by

\[
\begin{bmatrix}
R_{yy} & R_{yu}^+
\end{bmatrix}
\begin{bmatrix}
B
\end{bmatrix}
=
\begin{bmatrix}
R_{yy}
\end{bmatrix}
\]

(G.7)

In the transform domain, the prediction error model is represented by

\[
E_0(z) = Y(z) [I - B(z)] - U(z) A(z)
\]

(G.8)

where \(E\), \(Y\) and \(U\) are the transforms of the error, output and input vectors and the coefficients of the \(i,j\)-th elements of the matrix polynomials \(B(z)\) and \(A(z)\) are the elements of the vectors \(b_{ij}\) and \(a_{ij}^+\). The multichannel ARMA synthesis model is then given by

\[
Y^T(z) = U(z)^T A(z) [I - B(z)]^{-1}
\]

(G.9)

with the matrix polynomial fraction serving as the generalization of the zero pole transfer function.
APPENDIX H

Delay Free Loops

To develop equation (4.24) guaranteeing the absence of delay free loops in the system of Figure 4.8, consider the equation for $y_N(k)$.

$$y_N(k) = F[x_N(k)]$$  \hspace{1cm} \text{(H.1)}

Since $F[.]$ is a memoryless nonlinear function, proving that $x_N(k)$ is not a function of $y_N(k)$ is equivalent to proving that $y_N(k)$ is not a function of itself, and therefore no delay free loops exist. From equations (4.17b) and (4.17d) with $u(k)=0$ it follows that

$$x_N(z) = \Gamma_2 T(z) \Gamma_1 y_N(z)$$  \hspace{1cm} \text{(H.2)}

and the coefficient of $y_N$ at time $k$ on the right hand side is given by

$$a = \lim_{z \to \infty} \Gamma_2 T(z) \Gamma_1$$  \hspace{1cm} \text{(H.3)}

A nonzero $a_{ij}$ indicates a dependence of $x_{Ni}(k)$ upon $y_{Nj}(k)$ and clearly therefore all the main diagonal elements of $a$ must be zero to avoid delay free loops. These elements are the terms of the $(Q-1)$-st principal minors of $a$. While this is a necessary condition it is not sufficient to avoid delay free loops since loops may exist through two or more signals.
The condition that \( a_{ij} a_{ji} = 0 \) for all \( i, j \) such that \( 1 \leq i, j \leq Q \) and \( i \neq j \), ensures that no delay free loop exist through 2 signals and is equivalent to requiring all terms of the \((Q-2)\)-nd principal minors of \( \alpha \) are zero. A term of a determinant [Ref. 63] is defined as the product of elements of the matrix taken one from each row and one from each column and the determinant of the matrix is the sum of all possible terms. It is clear therefore that every term must contain a cycle such as \( a_{ij} a_{jk} \ldots a_{li} \) and must therefore be zero. Since the determinant consists of every possible term, requiring that all terms of the determinants of the \((Q-i)\)-th principal minors are zero ensures that no delay free loops exist through any combination of \( i \) signals. Examining all terms of the determinant of \( \alpha \) and all its principal minors ensures that all possible loops through the \( Q \) signals in \( x_N(k) \) are examined. If any delay free loop exists, then at least one of the terms of one of the determinants will be nonzero.
APPENDIX I

Nonlinear ARMA Modeling Examples

The determination of memory requirements for the nonlinear ARMA model as well as its applicability to systems consisting of interconnections of linear and memoryless nonlinear subsystems is illustrated here using two examples. First a cascade of linear and nonlinear subsystems is considered. Then a real world example is considered and a nonlinear ARMA model is proposed for the tracking behavior of a phase locked loop.

A. A CASCADE OF LINEAR AND NONLINEAR SUBSYSTEMS

Consider the system shown in Figure I.1 where the signals $u(k)$ and $y_{LZ}(k)$ are observed. In terms of the topology of Figure 4.8, seven signals can be identified ($x_{L1}, x_{L2}, y_{L1}, y_{L2}, x_{N1}, y_{N1}$, and $u$) however for convenience three of the seven equations in (4.26) which specify

\begin{align*}
  x_{L1}(k) &= u(k) \\  x_{L2}(k) &= y_{N1}(k) \\  x_{N1}(k) &= y_{L1}(k)
\end{align*}

will not be explicitly written. In this case, equation (4.26) becomes

\begin{equation}
  \text{(I.1a)}
\end{equation}

\begin{equation}
  \text{(I.1b)}
\end{equation}

\begin{equation}
  \text{(I.1c)}
\end{equation}
where the finite memory representations of $T_1(z)$ and $T_2(z)$ have been used. The objective now is to eliminate $a_2(k)^*$ from the upper right partition. Using the equations of rows three and four, the first row can be rewritten as

\begin{equation}
y_{L2}(k) = b_2(k)^* y_{L2}(k) + a_2(k)^* F_1 [a_1(k) u(k) + b_1(k)^* y_{L1}(k)]
\end{equation}

Notationally it is difficult to write equation (I.3) in the operator matrix form of equation (I.2) because of the nonlinear function however, the following representation is adopted.
\[
\begin{bmatrix}
  y_{L2}(k) \\
y_{L1}(k) \\
y_{N1}(k)
\end{bmatrix} =
\begin{bmatrix}
  b_2(k)^* & a_2(k)^* & b_1(k)^* & 0 \\
 0 & 1 & 0 & 0 \\
0 & a_1(k)^* & b_1(k)^* & 0 \\
0 & 0 & F_1[\cdot] & 0
\end{bmatrix}
\begin{bmatrix}
u(k) \\
y_{L2}(k) \\
y_{L1}(k) \\
y_{N1}(k)
\end{bmatrix}
\]
\[(I.4)\]
No further reductions are possible to make the upper right partition a null matrix leading to the conclusion that for a finite memory nonlinear ARMA model to be appropriate, either $b_1(k)$ must be zero (the first linear system must be nonrecursive) or $y_{L1}(k)$ must also be observed. Alternately, if $b_1(k)$ is nonzero, an infinite memory representation can be used for the first linear system by replacing $a_1(k) h(.)$ with $h_1(k) h(.)$ and $b_1(k) h(.)$ with zero in equation (1.4) indicating that an infinite memory nonlinear ARMA model is appropriate when only $u(k)$ and $y_{L2}(k)$ are observed.

B. A NONLINEAR ARMA MODEL FOR A PHASE LOCKED LOOP

A continuous time model for the tracking behavior of a phase locked loop [Ref. 55] is shown in Figure I.2 where:
- $\theta_1(t)$ is the phase of the incoming signal
- $\theta_2(t)$ is the estimate of the phase of the incoming signal
- $e(t)$ is the phase error signal
- $F(s)$ is the transfer function of the loop filter
- $K_1$ and $K_2$ are constants

![Diagram](image)

Figure I.2. A nonlinear model for the tracking behavior of a phase-locked loop.
The model is nonlinear because of the sin function in the loop. Often the assumption is made that \( e(t) \ll \pi/2 \) so that \( \sin e(t) \approx e(t) \) in which case a linearized model is obtained as

\[
\frac{\theta_2(s)}{\theta_1(s)} = \frac{K_1 K_2 F(s)}{K_1 K_2 F(s) + s}
\]

(A.5)

A nonlinear ARMA model for the system can be obtained by first discretizing the model of Figure I.2 as shown in Figure I.3 where \( F(z) \) represents the discrete loop filter.

The integration has been approximated as \( \frac{z^{-1}}{1-z^{-1}} \). (It is necessary to use this Euler forward approximation for integration rather than one such as the trapezoid rule to avoid a delay free loop). Defining a single linear system as the cascade of \( F(z) \) and the discrete integration

\[
K_1 K_2 \frac{z^{-1}}{1-z^{-1}} F(z) = \frac{A(z)}{1-B(z)}
\]

(I.6)
Equation (4.26) for the phase locked loop becomes

\[
\begin{bmatrix}
\theta_2(k) \\
\theta_1(k) \\
y_N(k) \\
e(k)
\end{bmatrix} =
\begin{bmatrix}
b_1(k)^* & 0 & a_1(k)^* & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & \sin(.) \\
-1 & 1 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\theta_2(k) \\
\theta_1(k) \\
y_N(k) \\
e(k)
\end{bmatrix}
\]

(1.7)

where it is assumed that \( \theta_1(k) \) and \( \theta_2(k) \) are observed. Using the relationships specified by rows 3 and 4, this can be written as

\[
\begin{bmatrix}
\theta_2(k) \\
\theta_1(k) \\
y_N(k) \\
e(k)
\end{bmatrix} =
\begin{bmatrix}
b_1(k)^* & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & \sin(.) \\
-1 & 1 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\theta_2(k) \\
\theta_1(k) \\
y_N(k) \\
e(k)
\end{bmatrix}
\]

(1.8)

from which it is clear that a finite memory nonlinear ARMA model is appropriate. The model can be obtained from the first row in equation (I.8) by substituting a series expansion.
for the sin function and truncating it at the degree desired resulting in

\[ \theta_2(k) = b_1(k) \ast \Theta_2(k) + a_1(k) \ast \sum_{m=0}^{M} \frac{(-1)^m}{(2m+1)} \Gamma(\Theta_2(k) - \Theta_1(k)) 2^{m+1} \]

(I.9)

where \(2M+1\) is the degree of the series approximation to \(\sin(.)\). (Note that \(\lim_{z \to \infty} A_1(z) = \lim_{z \to \infty} B_1(z) = 0\) so that the right hand side of equation (I.9) involves only past values of the output \(\theta_2(k)\).) An infinite memory Volterra series model for this system has been considered by Van Trees [Ref. 55].
APPENDIX J

Model Simulation Program Listings

This appendix provides a listing of the fortran model simulation programs used in the experimental study of the lattice characteristics. Included are the main programs for the batch processing ARMA lattice, the adaptive ARMA lattice and the brute force model solution. Each main program is followed by a collection of subroutines used only by that specific main program. Then a collection of common subroutines called from two or more locations in the batch lattice, adaptive lattice or brute force method is listed.
C** THIS IS THE MAIN PROGRAM FOR THE ECHIC PROCESSING ARMA LATTICE.
C** TOGETHER WITH THE AFOREMENTIONED SUBROUTINES, IT FINDS ALL ARMA
C** MODELS FROM ORDER 1 TO THE SYSTEM CRITERIA PLUS 2 USING FIRST 200
C** PCFAT AVERAGES THEN 4000 POINT AVERAGES. RESULTS ARE CLTPT ON THE
C** SYSTEM PRINTER AND A FILE CONTAINING MODEL BCOEFFICIENTS ARE ROOTS
C** IS CREATED FOR LATER PLOTTING ON THE VERSATCE PLOTTING.
C*********************************************************************************
C DIMENSION U(5000),V(5000),PA(10),FE(10),BMA(1C),BME(1C)
C GENERATE THE SYSTEM INPUT SIGNAL
C WRITE (6,1CC)
C CALL SNCRM(28,1,5CO)
C GIVEN THE DATA IN SUBROUTINE PCOEF, SET UP THE PLANT
C CALL PCCEF(FA,PF,FG,IFD,N)
C FLA THE PLANT AND GENERATE THE OUTPUT SIGNAL
C CALL FRUN(PA,PB,FG,IFL,U,Y,N,IFL,5CO)
C FINC THE LATTICE MODELS WITH 200 AND 4000 POINT AVERAGES
C DC 1200 IR=1.2
C IF(IR.EQ.1) NP=2CC
C IF(IR.EQ.2) NP=4CC
C MAX=10
C MROD=MINC(MAX,N+2)
C CALL LATFM(L,Y,PA,PB,FG,MROD,NP)
C CONTINUE
C FORMAT(1,*)
C STOP
C END
C*********************************************************************************
C** THIS SUBROUTINE CALCULATES ALL LATTICE ARMA MODELS FOR THE SYSTEM
C** FROM FIRST CRITER TO ORDER NACE USING NP POINT AVERAGES WHEN GIVEN
C** SYSTEM INPUT AND OUTPUT DATA.
C SLERCUINE LATFM(L,Y,FLTA,PLTE,FLTE,MORD,NP)
C DIMENSION U(11),Y(11),PLTA(11),PLTE(11)
C DIMENSION A(10),B(10),EVF(5000,2),EUF(5000,2),EY(5000,2),EYF(5000,2)
C DIMENSION FLB(5000,2),PF(2,2),PE(2,2),DEL(2,2)
C DIMENSION FK(2,2,1C),PK(2,2,1C),TFK(2,2),TBK(2,2)
DIMENSION WR(10)
DIMENSION ALF(10,2),AYF(10,2),BLF(10,2),BYF(10,2)
WRITE (6,1310)
WRITE (6,1310)
WRITE (6,1310) NF
WRITE (6,1310)
INTST=500
DC 2 II=1,1C
A(II)=0.0
B(II)=0.0
IGCT=0
CC 10 CC I=1,WORD
IF(I.LT.1) CC TO 2C
CC TIME AVERAGE TC FIND THE PRECICTION ERROR COVAIANCE MATRICIES
CC AT THE ZERCTH ORDER STAGE.
CC
CC 10 J=1,5C00
EVE(J,J)=Y(J)
EVE(J,J)=Y(J)
EVE(J,J)=L(J)
EVE(J,J)=U(J)
CONTINUE
PF(1,1)=0.0
PF(1,2)=0.0
PF(2,2)=0.0
AFF=INTST*NP
CC 15 J=INTST,NPF
PF(1,1)=PF(1,1)+Y(J)**2
PF(1,2)=PF(1,2)+Y(J)*L(J)
PF(2,2)=PF(2,2)+U(J)**2
CONTINUE
PF(1,1)=PF(1,1)/NP
PF(1,2)=PF(1,2)/NP
PF(2,2)=PF(2,2)/NP
PF(1,1)=PF(1,1)
PF(1,2)=PF(1,2)
PF(2,1)=PF(2,1)
PF(2,2)=PF(2,2)
IZ=0
CC
CC LIST THE ZERCTH ORDER ERROR COVAIANCE MATRICIES ALONG WITH
CC EIGENVALUES, EIGENVECTORS AND ELLIPSCIC AXIS HALF LENGTHS.
CC
WRITE (6,1000) IZ,PF(1,1),PF(1,2)
WRITE (6,1010) PF(2,1),PF(2,2)
CALL EIG(PF)
WRITE (6,1020) IT, EB(1,1), PE(1,1)
WRITE (6,1010) PB(1,1), PB(2,1)
CALL EIG(PB)
DET=PF(1,1)*PF(2,2)-EF(1,2)*PF(2,1)
DET=DET, DETF
GC TO 31

C IF AT THE SECOND STAGE OR BEYOND, CALCULATE THE PRECISION ERRORS
C SEQUENCES BY PASSING THE OLD ERRORS SEQUENCES TO THE LATEST
C LATTICE STAGE.

CC 25 J=1:5000
EYF(J,2)=EYF(J,1)
EYE(J,2)=EYE(J,1)
ELE(J,2)=EUE(J,1)

CC CONTINUE
CC 30 J=J+1:5000
IF(J,NE.1) GO TO 27
EYF(J,1)=EYF(J,2)
ELE(J,1)=EUE(J,2)

CC CONTINUE

CC 27 EYF(J,1)=EYF(J,2)-FK(1,1,1-1)*EYE(J-1,2)-FK(2,1,1-1)*EUE(J-1,2)
ELE(J,1)=ELE(J,2)-FK(1,2,1-1)*EYE(J-1,2)-FK(2,2,1-1)*EUE(J-1,2)
EYE(J,1)=EYE(J-1,1)-FK(1,1,-1)*EYF(J,2)-FK(2,1,-1)*EUF(J,2)
EUF(J,1)=EUF(J-1,2)-FK(1,-1,1)*EYE(J,2)-FK(2,-1,1)*EUF(J,2)

CC CONTINUE

CC 30 TIME AVERAGE TO FIND THE DEL MATRIX.
CC
NC=NC+1:INT(*T
DEL(1,1)=0.0
DEL(1,2)=0.0
DEL(2,1)=0.0
DEL(2,2)=0.0

CC 40 J=INT(TST,NCY
DEL(1,1)=DEL(1,1)*EYF(J,1)*EYB(J-1,1)
DEL(1,2)=DEL(1,2)*EYF(J,1)*EUB(J-1,1)
DEL(2,1)=DEL(2,1)*EUF(J,1)*EYB(J-1,1)
DEL(2,2)=DEL(2,2)*EUF(J,1)*EUB(J-1,1)

CC CONTINUE

DEL(1,1)=DEL(1,1)/NC
DEL(1,2)=DEL(1,2)/NC
DEL(2,1)=DEL(2,1)/NC
DEL(2,2)=DEL(2,2)/NC
TFK(1,1)=DEL(1,1)  
TFK(1,2)=DEL(2,1)  
TFK(2,1)=DEL(1,2)  
TFK(2,2)=DEL(2,2)  
TEK(1,1)=DEL(1,1)  
TEK(1,2)=DEL(2,1)  
TEK(2,1)=DEL(1,2)  
TEK(2,2)=DEL(2,2)  

CALCULATE THE NEXT REFLECTION COEFFICIENT MATRICES FOR THE LATTICE.

CALL SCLK(PF,FK,TFK,1)  
CALL SLOK(PF,FK,TEK,1)  

CALCULATE THE PRECINCT ERROR COVARIANCE MATRICES USING THE CLC  
CCOVARIANCE MATRICES AND THE LATEST REFLECTION COEFFICIENT MATRICES.

CALL LOCW1(FP,PF,FK,TEK,1)  
G1T=PF(1,1)*PF(2,2)-PP(1,1)*PF(2,2)  
DET=PF(1,1)*PF(2,2)-PF(1,2)*PF(2,1)  

CALCULATE THE ARMA MCCEL GAIN AT THIS ORDER SOLUTION.

G=PF(1,2)/PF(2,2)  

CALCULATE THE ARMA MCCEL TRANSFER FUNCTION POLYNOMIALS AT THIS  
ORDER SOLUTION.

CALL UCPLY1(A,B,FK,TEK,1)  
CALCULATE THE ARMA MEAN SQUARE ERROR FOR THIS ORDER.

ESC=PF(1,1)*PF(2,2)*GG-PP(1,2)*PF(2,1)*GG  

LIST LATEST LATTICE COEFFICIENTS, COVARIANCE MATRICES, EIGENVALUES,  
EIGENVECTORS AND AXIS HALF LENGTHS.

WRITE (6,1100) 1,FK(1,1),1,FK(2,1)  
WRITE (6,1100) 1,FK(1,2),1,FK(2,2)  
WRITE (6,1100) 1,FK(1,1),1,FK(2,1)  
WRITE (6,1100) 1,FK(1,2),1,FK(2,2)  
WRITE (6,1100) 1,FP(1,1),1,FP(2,1)  
WRITE (6,1100) 1,FP(1,2),1,FP(2,2)  
WRITE (6,1100) 1,FP(1,1),1,FP(2,1)  
WRITE (6,1100) 1,FP(1,2),1,FP(2,2)  
CALL EIG(PF)  
WRITE (6,1100) 1,FB(1,1),1,FB(2,1)  
WRITE (6,1100) 1,FB(1,2),1,FB(2,2)  
CALL EIG(PB)  
WRITE (6,11030) DETF,DETB
IF((DETF.LT.0.0).AND.(CETB.LT.0.0)) WRITE (6,1200)
WRITE (6,1400) I,ESQ
WRITE (6,1000) C

LIST ARM NUCLEAR TRANSFER FUNCTION COEFFICIENTS

WRITE (6,1010) C
WRITE (6,1020) INC,B,INC,INC,

FINE AND LISTS NUCLEAR AND FLCT NUCLEAR AND SYSTEM TRANSFER

FUNCTION MAGNITUDES.

CALL SPECIP(E,A,C,PLTE,PLTA,PLTC,MCRO,I)

CONTINUE

FORMAT('DPF',*I2,*E14.5,5x,E14.5)
FORMAT('EX,E14.5,5X,E14.5')
FORMAT('PF = E14.5,5X,DET FB = E14.5')
FORMAT('PSE FOR THE *I2,* CREP 2-P MODEL = E14.5')
FORMAT('COEFFICIENTS ARE:')
FORMAT('E14.5,5X,*)
FORMAT('25X,1A(0)=E14.5')
FORMAT('E14.5,5X,KE(1,12)=E14.5')

FORMAT('********** NEGATIVE DEFINITE COVARIANCE***********)
FORMAT('********** EXPLICIT LATTICE SOLUTIONS***********)
FORMAT('********** WITHOUT EXPLICIT LATTICE SOLUTIONS***********)
FORMAT('********** PCOM AVERAGES***********)
RETURN
END

**************************************************************************
THIS SUBROUTINE CALCULATES THE REFLECTION COEFFICIENT MATRICES
WHEN GIVEN THE APPROPRIATE CEL AND F MATRICES.
**************************************************************************

SUBROUTINE SCLK(AF,AK,DEL,I)
DIMENSION AF(2,2),AK(2,2),DEL(2,2)

CALL LEXT2F(AP,2,2,CEL,IGT,WK,IER)
CONTINUE
I+1 = 1+2
CONTINUE
I0 = CONTINUE
RETURN
END

******************************************************************************

THIS IS THE MAIN ADAPTIVE ARMA LATTICE PROGRAM TOGETHER WITH THE
APPROPRIATE SUBROUTINES IT WILL RUN THE LATTICE 1CC TIMES AND
AVERAGE THE RESULTS OVER THE ENSEMBLE. RESULTS WILL BE PRINTED ON
THE SYSTEM PRINTER AND TWO OUTPUT DATA FILES WILL ALSO BE CREATED
CONTAINING MODEL COEFFICIENTS, RCOIS AND SPECIFIED TIME SEQUENCES
FOR LATER FACITING ON THE VERSATEC FLIPPER.

******************************************************************************

!DIMENSION XI(5000),2),AER(5000,5)
!DIMENSION AFK(2,2,10,3),ABK(2,2,10,3),AA0(10,3)
!DIMENSION AEZPSQ(I0,3),PA(I0),PE(I0),IS(I0),XX(65CC),AK(65CC,5)
NCL=2
NE=100
EPS=FLOAT(NE)

GIVEN THE DATA IN SUBROUTINE PCCEF, SET UP THE SYSTEM TO BE PCCELEC.
CALL PCCEF(FA,PA,PE,IC,N)
ICF=NE
CALL START(2E,IS,1CC)
CC SC I=1,6500
TF=I-1
X(11)=FLOAT(IM)

RUN THE ADAPTIVE LATTICE NE5 TIMES WITH INDEPENDENT INPUTS FOR
ENSEMBLE AVERAGING.
CC IC 100 I=1,NE5
CC GENERATE AN INPUT SIGNAL FOR THE 1-TH RUN USING IS(I) AS A SEECC.
CALL SNCRM(IS(I),X(1,2),6500)
CC RUN THE SYSTEM TO GENERATE THE OUTPUT FOR THE 1-TH RUN.
CALL PRLA(FA,PA,PE,IC,X(1,2),X(1,1),N,IFL,65CC)
CC RUN THE ADAPTIVE ARMA LATTICE.
CALL ADLAT(3000,10RC,AER,AFK,AEK,A0,AEZPSQ,AK)
WRITE (6,10CC) I
CC CONTINUE
C AVERAGE THE RESULTS OVER THE ENSEMBLE AND OUTPUT THEM.
C
C 110 I=1,3000
C 105 AER(I,INCI)=AER(I,INCI)/ENS
C 106 INCI=1.5
C 107 IF(I.GT.3CC(I) GO TO 1C6
C 108 IF(I.RB)=AER(I,IRB)/ENS
C 109 CONTINUE

1C6
C 112 I=1,3
C 113 CONTINUE

110 CALL FUNCTION(AER(I,1))
C 115 I=1:3000
C 116 INCI=1.5
C 117 IF(I.RB)=AER(I,IRB).
C 118 CALL FUNCTION(AER(I,IRB)/ENS
C 119 AEZPSO(I,J,K)=AEZPSO(I,J,K)/ENS
C 120 IF(I.RB)=AER(I,IRB)
C 121 CALL FUNCTION(AER(I,IRB)
C 122 AEK(I,K,L,J,1)=AEK(I,K,L,J,1)/ENS
C 123 CALL FUNCTION(AER(I,IRB)
C 124 CALL FUNCTION(AER(I,IRB)
C 125 CONTINUE

1C7
C 126 WRITE (CIC) NES,1
C 127 CALL PRINT(I,1),AER(I,1),3000,C
C 128 CALL PRINT(I,1),AK(I,1),3000,C
C 129 CALL FUNCTION(I,1)
C 130 CONTINUE
C 131 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

135 CONTINUE
C 136 CALL FUNCTION(AER(I,IRB)
C 137 CALL FUNCTION(AER(I,IRB)
C 138 CONTINUE

1C8
C 139 WRITE (CIC) NES,1
C 140 CALL PRINT(I,1),AER(I,1),3000,C
C 141 CALL PRINT(I,1),AK(I,1),3000,C
C 142 CALL FUNCTION(I,1)
C 143 CONTINUE
C 144 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

145 CONTINUE
C 146 CALL FUNCTION(AER(I,IRB)
C 147 CALL FUNCTION(AER(I,IRB)
C 148 CONTINUE

1C9
C 149 WRITE (CIC) NES,1
C 150 CALL PRINT(I,1),AER(I,1),3000,C
C 151 CALL PRINT(I,1),AK(I,1),3000,C
C 152 CALL FUNCTION(I,1)
C 153 CONTINUE
C 154 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

155 CONTINUE
C 156 CALL FUNCTION(AER(I,IRB)
C 157 CALL FUNCTION(AER(I,IRB)
C 158 CONTINUE

1C10
C 159 WRITE (CIC) NES,1
C 160 CALL PRINT(I,1),AER(I,1),3000,C
C 161 CALL PRINT(I,1),AK(I,1),3000,C
C 162 CALL FUNCTION(I,1)
C 163 CONTINUE
C 164 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

165 CONTINUE
C 166 CALL FUNCTION(AER(I,IRB)
C 167 CALL FUNCTION(AER(I,IRB)
C 168 CONTINUE

1C11
C 169 WRITE (CIC) NES,1
C 170 CALL PRINT(I,1),AER(I,1),3000,C
C 171 CALL PRINT(I,1),AK(I,1),3000,C
C 172 CALL FUNCTION(I,1)
C 173 CONTINUE
C 174 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

175 CONTINUE
C 176 CALL FUNCTION(AER(I,IRB)
C 177 CALL FUNCTION(AER(I,IRB)
C 178 CONTINUE

1C12
C 179 WRITE (CIC) NES,1
C 180 CALL PRINT(I,1),AER(I,1),3000,C
C 181 CALL PRINT(I,1),AK(I,1),3000,C
C 182 CALL FUNCTION(I,1)
C 183 CONTINUE
C 184 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

185 CONTINUE
C 186 CALL FUNCTION(AER(I,IRB)
C 187 CALL FUNCTION(AER(I,IRB)
C 188 CONTINUE

1C13
C 189 WRITE (CIC) NES,1
C 190 CALL PRINT(I,1),AER(I,1),3000,C
C 191 CALL PRINT(I,1),AK(I,1),3000,C
C 192 CALL FUNCTION(I,1)
C 193 CONTINUE
C 194 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

195 CONTINUE
C 196 CALL FUNCTION(AER(I,IRB)
C 197 CALL FUNCTION(AER(I,IRB)
C 198 CONTINUE

1C14
C 199 WRITE (CIC) NES,1
C 200 CALL PRINT(I,1),AER(I,1),3000,C
C 201 CALL PRINT(I,1),AK(I,1),3000,C
C 202 CALL FUNCTION(I,1)
C 203 CONTINUE
C 204 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

205 CONTINUE
C 206 CALL FUNCTION(AER(I,IRB)
C 207 CALL FUNCTION(AER(I,IRB)
C 208 CONTINUE

1C15
C 209 WRITE (CIC) NES,1
C 210 CALL PRINT(I,1),AER(I,1),3000,C
C 211 CALL PRINT(I,1),AK(I,1),3000,C
C 212 CALL FUNCTION(I,1)
C 213 CONTINUE
C 214 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

215 CONTINUE
C 216 CALL FUNCTION(AER(I,IRB)
C 217 CALL FUNCTION(AER(I,IRB)
C 218 CONTINUE

1C16
C 219 WRITE (CIC) NES,1
C 220 CALL PRINT(I,1),AER(I,1),3000,C
C 221 CALL PRINT(I,1),AK(I,1),3000,C
C 222 CALL FUNCTION(I,1)
C 223 CONTINUE
C 224 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

225 CONTINUE
C 226 CALL FUNCTION(AER(I,IRB)
C 227 CALL FUNCTION(AER(I,IRB)
C 228 CONTINUE

1C17
C 229 WRITE (CIC) NES,1
C 230 CALL PRINT(I,1),AER(I,1),3000,C
C 231 CALL PRINT(I,1),AK(I,1),3000,C
C 232 CALL FUNCTION(I,1)
C 233 CONTINUE
C 234 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

235 CONTINUE
C 236 CALL FUNCTION(AER(I,IRB)
C 237 CALL FUNCTION(AER(I,IRB)
C 238 CONTINUE

1C18
C 239 WRITE (CIC) NES,1
C 240 CALL PRINT(I,1),AER(I,1),3000,C
C 241 CALL PRINT(I,1),AK(I,1),3000,C
C 242 CALL FUNCTION(I,1)
C 243 CONTINUE
C 244 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

245 CONTINUE
C 246 CALL FUNCTION(AER(I,IRB)
C 247 CALL FUNCTION(AER(I,IRB)
C 248 CONTINUE

1C19
C 249 WRITE (CIC) NES,1
C 250 CALL PRINT(I,1),AER(I,1),3000,C
C 251 CALL PRINT(I,1),AK(I,1),3000,C
C 252 CALL FUNCTION(I,1)
C 253 CONTINUE
C 254 CALL CPTZP(AFK,ABK,AKC,AEZPSO,PA,PB,PG,IORC)

255 CONTINUE
C 256 CALL FUNCTION(AER(I,IRB)
C 257 CALL FUNCTION(AER(I,IRB)
C 258 CONTINUE

C******************************************************************************************
C THIS SUBROUTINE RUNS THE ACAPTIVE ARPA LATTICE OVER A FIFTY POINT * 
C INTERVAL GIVEN RECORDS OF THE SYSTEM INPUT AND OUTPUT DATA. *
C******************************************************************************************

SUBROUTINE ADLAT(I,NET,ITORC,AER,AER,AER,AKC,AEZPSO,PK)
DIMENSION X(500,1),EFT1(10),EFT2(10),EFT2(10),EFT2(10),EFT2(10)
DIMENSION SIGF(10),AER(5000,5),AK(2,10,2),AK(2,10,2),AK(2,10,2)
DIMENSION A(10,2), A0(10), SA(10), E2P(10), AEZFSQ(1C,2), SIGB(1C)

C INITIALIZE THE VARIABLES

MAX=10
NCF=2
ALP=0.05
DEL=0.1
T=500
[...]

6 IF(I,J,IK,1)=0.0
6 IF(I,J,I,1)=C0
6 SIGF(I)=DEL
6 SIGB(I)=DEL
6 E(1)=0.0
6 E2P(I)=C0

10 CALL THE ADAPTIVE LATTICE OVER AN NPT FCINT INTERVAL STARTING AT IST

10 NFF=NPT+IST
10 IP=IST
K=IST, NPP
10 IF(I,K=IST+1

15 FIND THE ERROR SEQUENCES AT TIME K USING THE LATEST COEFFICIENTS

15 EE(0,J,1)=EE(J,1,1)
15 CONTINUE
15 IF(I=E0,1) GO TO 3C
15 IP=I,1
15 IF(I,E0,1) GO TO 3C
15 N=1, NCF
15 IF(A=0.0
15 P=C0
15 IF(L=1,NCF
15 FA=FA+FK(L,P,1)*EE(L,IM,2)
15 E=EA+BK(L,1)*EE(L,IM)
15 EF(I,1)=EF(I,IM)-F
LOCATE THE STEP SIZE NORMALIZING FACTORS AT EACH STAGE

CC 70 I=1,ICRD
PI=I-1
WSF=0.0
WSG=0.0
SLF(I)=(1.0-ALP)*SA(I)+ALP*(EF(I,I)**2)
IF(I.EQ.1) CC TO 66
CC 55 J=1,ACH
WSF=WSF+EF(J,J)**2
66 CC 65 J=1,ACH
WSG=WSG+X(K,J)**2
IF(K.EQ.IST) GO TO 65
WSF=WSF+X(K+,J)**2
CC 1N
CS(I)=SIGF(I)+SIGB(I)**2
CC 71 I=1,ICRD
SIGF(I)=SIGF(I)
CC 2N
LOCATE THE LATTICE COEFFICIENTS
CC 100 I=1,ICRD
PI=I-1
AC(I)=AC(I)+(ALP/SA(I))*EF(I,I)**2
IF(I.EQ.1) CC TO 66
CC 66 I=1,ACH
EE 75 I=1,ACH
EFACE(IJ,IK,I)=EB(IJ,IK,I)**2+EF(IK,I)
CFACE(IJ,IK,I)=EF(IJ,IK,I)**2+EB(IK,I)**2
FK(I,J,IK,I)=FK(I,J,IK,I)+(ALF/SIGF(I))*EE(IJ,IK,I)**2+EF(IK,I)**2
AVERAGE THE INSTANTANEOUS ZP MSE AC --- THE ENSEMBLE FOR THE CORRECT C CECC AND AN EVERMODEL OF THE CECECC C
IF(KML. CT. 5CCD) EC TO 1C4
CC 1C3. IAC=I, 5
AE(KPL, IACX)=AE(KPL, IACX)+E2F(IACX)**2
1C4
C CECC
IF AT TIE 1CC, 2CC OR 3000 TIME POINTS SINCE ADAPTATION STARTEC,
CECC AVERAGE THE CECECCUES ARE THE INSTANTANEOUS ZP MSE
CECC ESTIMATES AT EACH CECECC FACT 1 TO 1 CC

TST LST+1CCC
TST LST+2CCC
TST LST+2CCC
IF(KL. EQ. TST) GO TO 1C5
IF(KL. EQ. TST) GO TO 1C5
IF(KL. EQ. TST) GO TO 1C5
GO TO 2CC
1C5
L=1
12C
L=2
1C3
L=3
1C4
EC TO 1C6
CC 125 L=1, IORD
CC 126 L=1, INCH
CC 126 L=1, INCH
AER(IJL, IKL, L)=AER(IJL, IKL, L)+E2F(IJL, IKL, L)
AER(IJL, IKL, L)=AER(IJL, IKL, L)+E2F(IJL, IKL, L)
AER(IJL, IKL, L)=AER(IJL, IKL, L)+E2F(IJL, IKL, L)
CC CONTINUE
AEZPS(I, L)=AEZPS(I, L)+E2F(I)**2
ACC(I, L)=ACC(I, L)+E2(I)
CC CONTINUE
RETURN
I14C FCPMAT(*C,25X,'A(0)=','E14.5)
I15C FCPMAT(*B,12,'=','E14.5,5X,'A(1,12)=','E14.5)
RETURN
END

******************************************************************************
C* THIS SUBROUTINE OUTLITS DATA SEQUENCES TO A FILE FOR LATER
C* PLATING.
******************************************************************************
C
C SUBROUTINE FUNOF(CAT)
C DIMENSION DAT(1)
C CC 10 I=1,3000,6
I2=I+2
I3=I+3
I4=I+4
I5=I+5
WRITE (7,1000) DAT(I),DAT(I),CAT(I2),DAT(I3),CAT(I4),DAT(I5)
CONTINUE
C00C FCPMAT(6,E10.4,1X))
200C RETURN
END

******************************************************************************
C* THIS IS THE MAIN PROGRAM FOR THE ELITE FORCE PCEILING PROCEDURE
C* USING MINQUEE DATA TO ESTIMATE CORRELATION FRACTIONS TOGETHER
C* WITH THE APPROPRIATE SLEROUTINES, IT CALCULATES ALL ARMA MODELS
C* CF ORDER 10 THE SYSTEM ORDER FILL TWO USING FIRST 200 FCINT
C* AVERAGE 14000 FCINT AVERAGES, RESULTS ARE OUPUT ON THE
C* SYSTEM PRINTER AND A FILE OF MODEL COEFFICIENTS AND FCETS IS ALSO
C* CREATED FOR LATER PLATING ON THE VERSATOC FLITER.
******************************************************************************
C
C DIMENSION U(5000),1(5000),PA(10),FE(10),BEQIC,BEKIC
C DIMENSION DLCUN(20),RYY(20),RYUM(2C),RYUP(2G)
C
C GENERATE THE SYSTEM INPUT SIGNAL
C WRITE (6,1000)
C CALL SNCRM(128,U,5000)
C
C GIVEN THE DATA I N SUBROUTINE PCE, SET UP THE SYSTEM
C CALL PCEEF(FA,PE,FC,IFC,N)
C
C RUN THE SYSTEM AND GENERATE THE OUTPUT SIGNAL
CALL FRLN(D1,PB,PF,IFC,U,Y,A,IFL,5000)

C CETAIN THE BLITE FORCE MODELS FOR 20C AND 400C PCINT AVERAGES
C
CC 1200 IR=1,2
IF(IR.EQ.1) NFR=20C
IF(IR.EQ.2) NFR=400C
C
C ESTIMATE THE CORRELATION FUNCTIONS USING WINCODEC DATA
C
WRITE (6,1000)
CALL CORR(I,Y,RRU,RRY,RRUP,RYC,RUC,RYUC,AF)
C
C FIND THE ARMA MODELS WITH BRUTE FORCE MATRIX INVERSION
C
NFR=10
MFR=MNA(MAX,N+2)
CC 10 I=1,MFR
CALL BPDL(FUU,RRY,RYLM,RYUF,RYC,FLO,RYUO,I,FB,FA,FC,MFR,Y,U,NP)
120C CCIBUTE
100C FCMPAT(*11)
1CC STOP
1D PRINT

***********************************************************************
C THIS SUBROUTINE ESTIMATES THE CORRELATIONS NEEDED IN THE BLITE
C FORCE MODELS. THE FORCE BY TIME AVERAGING, A RECTANGULAR WINDOW
C IS USED ON THE DATA. THE ESTIMATES ARE ALSO PLOTTED AND LISTED ON
C THE SYSTEM PRINTER.
***********************************************************************
C
C SUBROUTINE CORR(U,Y,FLU,RRU,RRY,RRUP,RUPY,RYC,RUC,RLC,RYLC,AF)
DIMENSION C(L),Y(L),RRU(L),RRY(L),RYY(L),RYUM(L),RYLP(L)
C MCONNECT X(1:41),YY(41)
INTST=500
WRITE (6,102C) NF
CC 10 I=1,2C
RYY(L)=C*C
RYC(L)=C*C
RYM(L)=C*C
RYP(L)=C*C
RLC=0.C
RLC=0.C
CC 50 N=1,2C
AF=AF-M*INTST
CC 20 I=INTST,NP
RYY(N)=RYY(N)+YY(I)*Y(I+N)
```c
RLL(N)=FLL(A)+UL[I]*UL[I+1]
FYL(N)=RYUF[N]+Y[I]*UL[I+1]
FSL(N)=RVLF[N]+UL[I]*UL[I+1]
CCATAGE
RLY[N]=EY[N]/(NP-A)
RLL[N]=RUU[N]/(NP-A)
RSL[N]=RVLF[N]/(NP-A)
FF=NP/INT
FIC=RYO+Y[I]*NY[I]
RLC=RUC+UL[I]*UL[I]
RYO=RYOC+Y[I]*UL[I]
RLC=RUC/AF
RLC=PYLC/NF
CE TO =-241
YY[I]=RYO
IF(AE=21) YY[I]=RYY(IABS[I-21])
WRITE (*,10CC)
CALL FPLT(X,Y,Y,41,0)
CE EC =-1,41
YY[I]=RLO
IF(AE=21) YY[I]=RUU(IABS[I-21])
WRITE (*,10CC)
CALL FPLT(X,Y,Y,41,0)
CE 90 I=1,41
YY[I]=RYU
IF(GT=21) YY[I]=RYUF[I-21]
WRITE (*,10D0)
CALL FPLT(X,Y,Y,41,0)
10CC FORMAT(3*RYU(0)=E14.5,RUU(C)=E14.5,RYL(C)=E14.5)
CE 100 I=1,42
WRITE (*,10D0) I,RYY(I),RUU(I),FYL(I),RYUF(I)
166 FORMAT(2*RYY,RYU,RYUM,RYUL='0.4(E14.5,2X))
100C FORMAT:* THE ACF OF THE OUTPUT Y[I]*
100C FORMAT:* THE ACF OF THE [YPLT U I]*
100C FORMAT:* THE ACF OF THE SIGNALS Y AND U [I]*
100C FORMAT*(0.15) THE ENTRIES WERE AVERAGED TO ESTIMATE CC RELATIONS*;
RETURN
C
```
SUBROUTINE EFPDL (FLY, FY, RYUM, RYLF, RYF, RYU, RYLP, RY)

IF (IR+1 .LT. ICR) GO TO 50

IF (IR+1 .GT. ICR) GO TO 40

IF (IR+1 .EQ. ICR) GO TO 30

IF (IR+1 .GE. ICR+1) GO TO 20

GO TO 10

10 IF (IR+1 .GT. ICR) GO TO 30

50 CONTINUE

40 CONTINUE

30 CONTINUE

20 CONTINUE

10 CONTINUE
**CC** 110 I=1, A
   IF* = T**P
   EEE(I) = B(I)
   WRITE (6, 120) 1, EEE(I), I, B(I)
110 CALL ESC(E(I), 1, E(I), U, N, P)
   ESC = RQ * E(I)
   RYCO
120 I = 1 A
120 B = EEE-C-E(I) * RYY(I) - EA(I) * RYLP(I)
   WRITE (6, 110) B=E
   CALL SCPEP(E(I), EA, E(NF), P6, PA, PA, P, C, N)
   ECR = RCR * E
113C FCMPAT(* R MATRIX BY CCLS WAS:*)
114C FCMPAT(* R')134'13=-',E14.5)
115C FCMPAT(* ORBITAL FORCE INVERSION SOLUTON OF CREER *',I2,' IS*)
116C FCMPAT(* AC = ',E14.5)
117C FCMPAT(* MATE IT; EFF(TL),131)
118C FCMPAT(* MS EQUATION ERROR = ',E14.5)
   RETURN
   EAE

C**********************************************************************
C THIS SUBROUTINE ESTIMATES THE PEAK SQUARE VALUE OF CUTPLT ERROR
C FOR THE ARNOW MODEL.
C**********************************************************************

C SUBROUTINE ESO(A, B, Y, U, N, NF)
C CIPESUDA (I), B(I), RYV(I), PUP(I)
C CIPESUDA Y(I), U(I), TF(5000), EF(I)
C INVST = 500
C 1C I = 1 A
10 B(I) = B(I)
   CALL PRU(A, B, E, C, U, Y, N, IFL, SCCC)
   IF(1FL = 0) GO TO TC 7C
   AFF = AF * INVST
   ERR = 0.0
   CC 2C I = INTST, NFF
   ERR = ERR + (Y(I) - Y(I) * 11) * 42
   ERR = ERR/AF
   WRITF (6, 100C) N, ERR
   GC TC 8C
7C WRITE (6, 101C) N
80 CONTINUE
100C FCMPAT(* OMS CUTPLT ERROR FOR THE ',I2,' CREER Z-P MODEL = ',E14.5)
110C FCMPAT(* OMS CUTPLT ERROR FOR ',I2,' CREER NET CALC; MODEL UNSTAB)
C******************************************************************************
C* THE FOLLOWING SUBROUTINES ARE COMMON ROUTINES CALLED FROM THE OR
C* NER EQUATIONS IN THE EATCF LATTICE, ADAPTIVE LATTICE OR ERUITE FORCE
C* CCHELING PROGRAMS.
C******************************************************************************
C
C******************************************************************************
C* THIS SUBROUTINE RUNS THE SYSTEM WHEN GIVEN ITS TRANSFER FUNCTION
C* COEFFICIENTS AND THE INPUT SIGNAL AND PRODUCES THE SYSTEM CCTFLT
C******************************************************************************
C
C******************************************************************************
C
C******************************************************************************
C
C******************************************************************************
C CROUTINE FRUN(A,E,G,ID,LY,NN,IFL,IP)
DIMENSION A(I),E(I),U(I),Y(I),X(I)
EC 10 I=1,N
X(I)=0.0
10 Y(I)=0.0
J=IF-IC
IFL=0
CC 50 I=1,J;
IF(IFL.EQ.1) GO TO 50
SPR=0.0
SPH=0.0
EC 20 II=1,N
SPR=SMR-B(II)*X(II)
SPH=SMR+A(II)*X(II)
EC 10 II=1,N
20 Y(I+ID)=SPR+SPH
IF(Y(I+ID).GT.1CC00000000.C) IFL=1
AP=A-1
CC 30 II=1,NN
30 X(I+I-1)=X(I-N-1)
X(I)=SMR
EC CONTINUE
RETURN
END
C
C******************************************************************************
C* THIS SUBROUTINE CALCULATES POLYNOMIAL COEFFICIENTS WHEN GIVEN THE
C******************************************************************************
C* EQUATIONS.
C******************************************************************************
C
C******************************************************************************
C
C******************************************************************************
C CROUTINE MAKPOL(N,R,C)
CCPLEX#16 F(N),C(I)
IF(N.LE.0) RETURN
EC 10 I=1,N
```c
C(1)=R(1)
K=1
K=K-1
C(1)=C(1)+C(1-1)
K=K-1
C(20) I=1,K
J=1+L
C(1)=R(J)*C(1)
K=K/2
K=2*K/(A-K)
C(4) I=K/2
C(1)=C(1)
RETURN
END
```

This subroutine updates the TBC channel or transfer function for the next given time and the latest lattice reflection coefficients matrices.

Subroutine LDLY(A, B, FK, BK, AYF, ALF, AYF, AVE, AVE, BLF, EYE, EYE, I, C)

Dimensions FK(2, 2), BK(2, 2), AVE(10, 1), BLF(12, 1), EYE(12, 1), ELE(12, 1)

Parameters FK(I, J), BK(I, J), AVE(I, J), BLF(I, J), EYE(I, J), ELE(I, J)

```
10 = I - 1
C(24) J=1,TF
EYE(J, J) = BUF(J, J)
AYF(J, J) = AVE(J, J)
ALF(J, J) = ALF(J, J)
EYE(J, J) = BAYE(J, J)
ELE(J, J) = AVE(J, J)
ALF(J, J) = AVE(J, J)
C(25) J=1,TF
```
Y1MAX=0.0
Y2MAX=0.0
CC 50 I=1,2C1
LIM=FLCAT(I-I)
S(I)=F(I)/2CC.O-A1N
FP=PG
FA=AO
I1N=0.,0
I2N=0.,0
FCP=1.,C
FCF=0.0
FCM=0.,C
CC 20 K=1,ACRD
In=K*N[I]
XAF=MXAF+FA(K)*FG*CSS(TH)
YAF=RYAF+AA(K)*CG*CSS(TH)
XAIN=MAIN-A[K]*GB*SIN(TH)
B1N=AINB-K)*GB*SIN(TH)
C1N=ACB-BK)*GK*CSS(TH)
C1N=ACB-BK)*GK*CSS(TH)
C1N=ACB-BK)*GK*CSS(TH)
Y1[I]=SCRT((RNP**2+AINP**2)/(RCF**2+AINP**2))
Y2[I]=SCRT((RMN**2+AINP**2)/(RCF**2+AINP**2))
IF(Y1[I] .LT. 0000001) Y1[I]=0.0000001
IF(Y2[I] .LT. 00000001) Y2[I]=0.0000001
Y1[I]=200+ALOG10(Y1[I])
Y2[I]=200+ALOG10(Y2[I])
IF(Y1[I] .GT. Y1MAX) Y1MAX=Y1[I]
IF(Y2[I] .GT. Y2MAX) Y2MAX=Y2[I]
CONTINUE
WRITE(6,1023)
CALL ZEROS(AO,A,ICR,0,IPFL)
WRITE(6,1024)
CALL ZEROS(100,B,ICR,1,IPFL)
IF(ICR .NE. ACRD) WRITE(2,1140)
IF(YMAX .GT. Y1MAX) GC TC 60
WRITE(6,111C)
CALL FPLT(X,Y1,201,1)
CALL FPLT(X,Y2,201,3)
WRITE(6,1010)
RETURN
WRITE(6,111C)
CALL FPLT(X,Y2,201,1)
CALL FPLT(X,Y2,201,3)
WRITE(6,1010)
RETURN
WRITE(6,111C)
CALL FPLT(X,Y2,201,1)
CALL FPLT(X,Y2,201,3)
WRITE(6,1010)
```fortran
* This subroutine calculates the zeros of a polynomial and lists
** in a file and call the system printer.
*
* SUBROUTINE ZEROS(C,C,N,IFL,IFL1)
*
CALL ZERPOLY(A,N,Z,IER)
*
** This subroutine updates the forward and backward prediction error
** covariance matrices from one lattice stage to the next given the
** covariance matrices and the latest lattice reflection.
**
C* CCEFFICIENT MATRICIES.
C
C SLERCUTIINE LCCOV(FF,PF,FK,BK1)
C GENSICA PF[2,2],FB[2,2],FK[2,2],BK[2,2,1]
C GENSECION TEM[2,2]
C TEM[1,1] = F4K1,1,1] - EK[1,2,1] * FK[2,1,1]
C TEM[1,2] = EK[1,2,1] - BK[2,2,1] * FK[1,1,1]
C W1=PF[1,1] * TEM[1,1] + PF[1,2] * TEM[1,2]
C W2=PF[2,1] * TEM[1,1] + PF[2,2] * TEM[1,2]
C W3=PF[1,1] * TEM[2,1] + PF[1,2] * TEM[2,2]
C
C C* CERIUTIINE EIG(P)
C GENSECION P[2,2],FF(3),EVAL(2),EVEC(2,2),BK(5)
C PF[1]=P[1,1]
C FF[2]=P[2,1]
C PF[3]=P[2,2]
C CALL EIGS(FP,2,1,EVAL,EVEC,2,BK,IER)
C A=0.0/SCRT(FEES(EVAL(1))
C B=0.0/SCRT(FEES(EVAL(2))
C WIFE (6,10C) EVAL(1),EVAL(2)
**FILE (6,102C) A,B
FILE (6,101C) EVEC(1,1),EVEC(2,1),EVEC(1,2),EVEC(2,2)**

**100C FORMAT(*,10X,I4,E14.5,5X)**
**101C FORMAT(*,10X,E14.5,10X,E14.5)**
**102C FORMAT(*,10X,I4)**
**RETURN**
**END**

****************************************************************************************************

C THIS SUBROUTINE TAKES THE ROOTS, GAIN AND BUDLEN DATA FROM EVEC DATA, IN IT AND CALCULATES THE SYSTEM TRANSFER FUNCTION. THE DATA IS
C ALSD LISTED ON THE SYSTEM TRANSFER FUNCTION IN A FILE.

****************************************************************************************************

SUBROUTINE ECFCEF(A,B,C,IC,N)
COMPLEX*16 R(10),C(10)
REAL A(1),B(1),Z(2,10),PR(2,10)
EQUIVALENCE (R,RR),(C,CC)

SET THE PLANT ORDER
N=7

SET THE GAIN
B=1.0

SET THE PLANT DELAY
C=IC

ENTER THE COMPLEX ZEROS IN R(1) THRU R(N)

R(1)=(1.02523,.625)
R(2)=(1.02523,-.625)
R(3)=(0.C.C,0.0)
R(4)=(0.C.C,0.0)
R(5)=(0.0,0.0)
R(6)=(0.0,0.0)
R(7)=(0.0,0.0)
R(8)=(0.0,0.0)
R(9)=(0.0,0.0)
R(10)=(0.0,0.0)
R(11)=(0.0,0.0)
R(12)=(0.0,0.0)

WRITE (6,1000) I,IC
WRITE (6,1001) N
WRITE (6,1002) 6,IC
CC IC I=1,N
C

C THIS SUBLTLINE PROduces THE FCNTS OR THE SYSTEM PRINTER IF FCFTF
C IS CALLED OR OR THE TIME SHARING TERMINAL IF FCFTT IS CALLED
C
C
C
C SUBROUTINE FFCTT(YY,YY,YY,YY,YY)
C DIMENSION X(II),Y(II)
C WRITE (6,1111)
C CALL FCFTF(Y,YY,YY,YY)
1111 FORMAT (*)
RETURN
END
LIST OF REFERENCES


<table>
<thead>
<tr>
<th>No.</th>
<th>Initial Distribution List</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Defense Technical Information Center</td>
</tr>
<tr>
<td></td>
<td>Cameron Station</td>
</tr>
<tr>
<td></td>
<td>Alexandria, Virginia 22314</td>
</tr>
<tr>
<td>2.</td>
<td>Library, Code 0142</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, California 93940</td>
</tr>
<tr>
<td>3.</td>
<td>Department Chairman, Code 62</td>
</tr>
<tr>
<td></td>
<td>Department of Electrical Engineering</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, California 93940</td>
</tr>
<tr>
<td>4.</td>
<td>Professor S. R. Parker, Code 82Px</td>
</tr>
<tr>
<td></td>
<td>Department of Electrical Engineering</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, California 93940</td>
</tr>
<tr>
<td>5.</td>
<td>Professor R. W. Hamming, Code 52Hg</td>
</tr>
<tr>
<td></td>
<td>Department of Computer Science</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, California 93940</td>
</tr>
<tr>
<td>6.</td>
<td>Professor R. Vanholsen, Code 82Pz</td>
</tr>
<tr>
<td></td>
<td>Department of Electrical Engineering</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, California 93940</td>
</tr>
<tr>
<td>7.</td>
<td>Professor H. A. Titus, Code 62Te</td>
</tr>
<tr>
<td></td>
<td>Department of Electrical Engineering</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, California 93940</td>
</tr>
<tr>
<td>8.</td>
<td>Professor T. C. Wang, Code 63Wg</td>
</tr>
<tr>
<td></td>
<td>Department of Mathematics</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, California 93940</td>
</tr>
<tr>
<td>9.</td>
<td>Professor D. N. Kirs, Code 62KI</td>
</tr>
<tr>
<td></td>
<td>Department of Electrical Engineering</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, California 93940</td>
</tr>
<tr>
<td>10.</td>
<td>Professor D. Thaler, Code 62Tr</td>
</tr>
<tr>
<td></td>
<td>Department of Electrical Engineering</td>
</tr>
<tr>
<td></td>
<td>Naval Postgraduate School</td>
</tr>
<tr>
<td></td>
<td>Monterey, California 93940</td>
</tr>
</tbody>
</table>
11. Professor T. Tao, Code 6
Department of Electrical Engineering
Naval Postgraduate School
Monterey, California 93940

12. Lieutenant Francis A. Perry
124 Belle Drive
Marina, California 93933

13. Captain H. Cox, USN
Naval Electronic Systems Command
Naval Electronic Systems Command Headquarters
PME-124
Washington, D.C. 20350

14. Commander
Naval Surface Weapons Center
White Oak Laboratory
ATTN: LCDR John Gauss, USN
Silver Springs, Maryland 20910

15. Mr. Martin Mandelberg
U.S. Coast Guard Research and Development Center
Avery Point
Groton, Connecticut 06340

16. Lieutenant Commander John Martinos
Hellenic Navy
Elefant 22
Chalandri
Athens, Greece

17. Lieutenant Martin Romeo
SMC 2250
Naval Postgraduate School
Monterey, California 93940

18. Lieutenant Vassilios Xidomas
SMC 2500
Naval Postgraduate School
Monterey, California 93940

19. Dr. Joel Novles
Code 427-06R
Office of Naval Research
800 North Quincy
Arlington, Virginia 22217

20. Commander Stacey Holman, USN
Naval Electronic Systems Command
Code 310
2311 Jefferson Davis Highway
Arlington, Virginia 22202
21. Mr. Nate Butler  
Naval Electronic Systems Command  
Naval Electronic Systems Command Headquarters  
Washington, D.C. 20360

22. Dr. A. S. Milton  
Naval Research Laboratory  
Code 6550  
Washington, D.C. 20375

23. Professor Janice Jenkins  
Northwestern University Medical School  
Reingold ECG Center  
310 E. Superior Street  
Chicago, Illinois 60611

24. Professor Sammuel Bedrosean  
Department of Systems Engineering  
Moore School DZ  
University of Pennsylvania  
Philadelphia, Pennsylvania 19174

25. Dr. Stephan Horvath  
E.T.H. Zurich  
CH-8093 Zurich  
Switzerland

26. Professor Jim McClellan  
M.I.T.  
Department of E.E. and C.S.  
Cambridge, Massachusetts 02139

27. Commander  
Naval Ocean Systems Center  
ATTN: Lieutenant Frank Perry, U.S.N.  
San Diego, California 92132

28. Professor Martin Mott  
Stanford University  
Information Systems Laboratory  
Durand 100  
Stanford, California 94305

29. Professor Candid Widrow  
Stanford University  
840 Lathrop Drive  
Stanford, California 94305

30. Professor Lloyd Griffiths  
M.I.T.  
Department of E.E. and C.S.  
Room 7-614  
77 Massachusetts Avenue  
Cambridge, Massachusetts 02139
31. Professor A. V. Oppenheim
M.I.T.
Department of E.E. and C.S.
Room 36-393
Cambridge, Massachusetts 02139

32. Dr. Ben Friedlander
Systems Control Inc.
1801 Page Mill Road
Palo Alto, California 94304

33. Dr. John Triechler
Argo Systems
884 Hermosa Court
Sunnyvale, California 94086

34. Dr. Bill Hodgkiss
University of California at San Diego
Marine Physical Laboratory
Scripps Institution of Oceanography
San Diego, California 92152

35. Commander
Naval Ocean Systems Center
ATTN: Dr. Harper Whitehouse
San Diego, California 92152