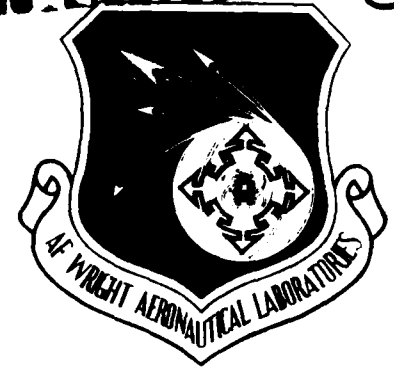


DTIC FILE COPY

2



AD-A196 498

AFWAL-TR-88-3042

**DETERMINATION OF COMPLEX  
EXPONENTIALS, LEAST SQUARES  
AND PREDICTION METHODS**

**CHARLES L. KELLER**

Aeroelastic Group  
Structures Division

February 1988

Final Report for the Period February 1985 to December 1987

Approved for public release; distribution is unlimited.

FLIGHT DYNAMICS LABORATORY  
AIR FORCE WRIGHT AERONAUTICAL LABORATORIES  
AIR FORCE SYSTEMS COMMAND  
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6553

DTIC  
ELECTE  
AUG 01 1988  
S D  
C/H

17 207 0

## NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report has been reviewed by the Office of Public Affairs (ASD/PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

Maxwell Bas for

DR CHARLES KELLER  
Project Manager  
Aeroelastic Group

Nelson D. Wolf

NELSON D. WOLF  
Acting Chief  
Analysis & Optimization Branch

FOR THE COMMANDER

R. M. Bader

ROBERT M. BADER  
Acting Chief, Structures Division

If your address has changed, if you wish to be removed from our mailing list, or if the addressee is no longer employed by your organization please notify AFWAL/FIBRC, WPAFB, OH 45433-6553 to help us maintain a current mailing list.

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

Form Approved  
OMB No. 0704-0188

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFWAL-TR-88-3042		
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFWAL-TR-88-3042			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
6a. NAME OF PERFORMING ORGANIZATION Flight Dynamics Laboratory AF Wright Aeronautical Lab		6b. OFFICE SYMBOL (If applicable) AFWAL/FIBRC	7a. NAME OF MONITORING ORGANIZATION		
6c. ADDRESS (City, State, and ZIP Code) Wright Patterson AFB OH 45433-6553			7b. ADDRESS (City, State, and ZIP Code)		
8a. NAME OF FUNDING/SPONSORING ORGANIZATION		8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER  In-House		
8c. ADDRESS (City, State, and ZIP Code)			10. SOURCE OF FUNDING NUMBERS		
PROGRAM ELEMENT NO. 611021		PROJECT NO. 2304	TASK NO. N1	WORK UNIT ACCESSION NO. 02	
11. TITLE (Include Security Classification)  Determination of Complex Exponentials, Least Squares and Prediction Methods					
12. PERSONAL AUTHOR(S) Charles L. Keller					
13a. TYPE OF REPORT Final		13b. TIME COVERED FROM Feb 85 to Dec 87		14. DATE OF REPORT (Year, Month, Day) February 1988	15. PAGE COUNT 42
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Prony, Progressive Algorithm, Least Squares, Stationary Sequences, Lattice Filters, Correlation Analysis, Singular Value Decomposition		
01	01				
20	04				
19. ABSTRACT (Continue on reverse if necessary and identify by block number) An efficient method for determining a sum of complex exponentials and their complex amplitudes is presented in detail. The principal feature is the application of Lanczos' Progressive Algorithm which eliminates the need for knowing or assuming the number of such quantities in the sum. The method is applicable to appropriate averages of the sampled data also. The algebraic properties of the method of least squares for the linear case are presented and provide the basis for the presentation of the 1-step-ahead prediction for stationary sequences using lattice filters and canonical variate analysis.  <i>Canonical variate analysis</i>					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a. NAME OF RESPONSIBLE INDIVIDUAL Maxwell Blair			22b. TELEPHONE (Include Area Code) (513) 255-7384		22c. OFFICE SYMBOL AFWAL/FIBRC

DD Form 1473, JUN 86

Previous editions are obsolete.

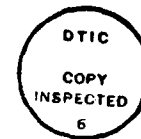
SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

## PREFACE

This report describes work performed in the Aeroelastic Group, Analysis and Optimization Branch, Structures Division, Air Force Wright Aeronautical Laboratories, Air Force Systems Command under Program Element 61102F, Project No.2304, Task N1, Computational Aspects of Fluid and Structural Mechanics, Work Unit 2. This is a final report on work performed at intervals during the period February 1985 through December 1987.

The author thanks Dr. K. G. Guderley for listening and his helpful suggestions.



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

## TABLE OF CONTENTS

Section		Page
I	INTRODUCTION	1
II	DETERMINATION OF COMPLEX EXPONENTIALS	3
III	THE PROGRESSIVE ALGORITHM	14
IV	THE METHOD OF LEAST SQUARES	21
V	LATTICE FILTERS	29
VI	CANONICAL VARIATE ANALYSIS	37
	REFERENCES	42

## SECTION I

### INTRODUCTION

For some lightly damped structure the free response may be expressed as a sum of a number of complex exponentials with complex amplitudes. It is clear that the number of such terms is not necessarily the same for different experiments on the same structure. The number of these terms depends upon the initial conditions. From a practical point of view one would be satisfied, usually, with a free response which contains all the frequencies in some limited range of interest.

In Section II we give the details for efficiently and accurately carrying out Prony's method, [1], for determining the complex exponentials and complex amplitudes from evenly spaced sampled data assuming that the number of such terms present is known. The method given is also applicable when the data is averaged in an appropriate but natural way. There is no difficulty in applying the method to multidegree of freedom cases. We believe that one could possibly treat the case of a complex frequency with multiplicity greater than 1.

In Section III a procedure, the Progressive Algorithm [2], is given which essentially enables us to determine the number of complex frequencies present. Actually the progressive algorithm performs the first step of the method given in Section II. Thus, it is no longer necessary to know in advance the number of complex frequencies present. The progressive algorithm is also applicable to the averaged data. Given a real valued function of a finite number of distinct complex frequencies with their associated complex amplitudes it is clear that a method comprised of the procedures given in Sections II and III is a theoretically exact method for determining these frequencies and amplitudes from equally spaced function values.

In Section IV we derive the algebraic properties associated with the method of Least Squares for the case of one quantity depending linearly on another. This in turn provides

the background for developing the method of Lattice Filters, [3], for stationary sequences and the method of Canonical Variate Analysis, [4].

## SECTION II

### DETERMINATION OF COMPLEX EXPONENTIALS

Consider a sum of terms of the form  $\mathbf{a}_k \exp(\lambda_k t)$ . The coefficient  $\mathbf{a}_k$  may be either a vector or scalar. For every such term there is a companion term which is the complex conjugate, so that the components of the sum are real. Such a sum arises as the homogeneous solution of a linear system of ordinary differential equations with constant coefficients. In particular dynamical systems with symmetric mass, stiffness and damping matrices. Regardless of origin, the problem of determining the parameters  $\mathbf{a}_k$  and  $\lambda_k$  from experimental data is frequently encountered.

In this section we examine and discuss a very old method associated with the name Prony. The method or procedure described is efficient, easily programmed and very accurate when tested on numerically produced data sampled at an appropriate rate. In a practical problem the number of the terms  $\mathbf{a}_k \exp(\lambda_k t)$  is not known. In the next section we give a method for dealing with this difficulty. There is the problem of multiplicity for a particular  $\lambda_k$  which we will discuss also.

Consider a real valued function  $x(t)$  defined by

$$x(t) = \sum_{k=1}^n a_k \exp(\lambda_k t) \quad (1)$$

where  $n = 2m$ ,

$$a_{m+k} = \bar{a}_k$$

and

$$\lambda_{m+k} = \bar{\lambda}_k.$$

The bar denotes the complex conjugate. The Eq (1) can be written also as

$$x(t) = 2 \sum_{k=1}^m \exp(\operatorname{Re}[\lambda_k] t) \left( \operatorname{Re}[a_k] \cos(\operatorname{Im}[\lambda_k] t) - \operatorname{Im}[a_k] \sin(\operatorname{Im}[\lambda_k] t) \right). \quad (2)$$

Now

$$\exp(\lambda(t+h)) = \exp(\lambda t) \exp(\lambda h).$$

Set

$$\eta_k = \exp(\lambda_k h). \quad (3)$$

From Eq (1) we obtain

$$\begin{aligned} x(t) &= a_1 e^{\lambda_1 t} + \dots + a_n e^{\lambda_n t} \\ x(t+h) &= \eta_1 a_1 e^{\lambda_1 t} + \dots + \eta_n a_n e^{\lambda_n t} \\ &\vdots \\ x(t+nh) &= \eta_1^n a_1 e^{\lambda_1 t} + \dots + \eta_n^n a_n e^{\lambda_n t}. \end{aligned} \quad (4)$$

The Eqs (4) may be written as

$$\begin{bmatrix} x(t) & -1 & \dots & -1 \\ x(t+h) & -\eta_1 & \dots & -\eta_n \\ \vdots & \vdots & \ddots & \vdots \\ x(t+nh) & -\eta_1^n & \dots & -\eta_n^n \end{bmatrix} \begin{bmatrix} 1 \\ a_1 e^{\lambda_1 t} \\ \vdots \\ a_n e^{\lambda_n t} \end{bmatrix} = 0. \quad (5)$$

Since this homogeneous system of equations has a nontrivial solution the determinant of the system matrix is zero.

Let  $A_{jk}$  denote the cofactor of the  $jk^{th}$  element in the determinant of the system matrix. Eq (5). We have

$$A_{n-1,1} x(t+nh) + A_{n,1} x(t+(n-1)h) + \dots + A_{1,1} x(t) = 0. \quad (6)$$

Assume for the present that the  $\lambda_k$  and hence also the  $\eta_k$  are different from one another. Observe, Eq (5), that the minor of the element  $x(t+nh)$  is a Vandermonde determinant. Hence the cofactor  $A_{n-1,1} \neq 0$  and Eq (6) can be written as

$$x(t+nh) + \frac{A_{n,1}}{A_{n-1,1}} x(t+(n-1)h) + \dots + \frac{A_{1,1}}{A_{n-1,1}} x(t) = 0. \quad (7)$$

The Eqs (6) and (7) show that the function values  $x(t)$ ,  $x(t+h)$ , ... satisfy a linear  $n^{th}$  order difference relation.

In Eq (7) set the coefficient

$$\frac{A_{j,1}}{A_{n-1,1}} = p_j \quad (8)$$

and rewrite Eq (7) as

$$x(t + nh) + p_n x(t + (n - 1)h) + \dots + p_1 x(t) = 0. \quad (9)$$

In Eq (5) replace the first column of the coefficient matrix by the symbols

$$-1, -\eta, -\eta^2, \dots, -\eta^n$$

Then in the same way that we obtained Eq (9), the expansion of the determinant of the modified coefficient matrix will give the polynomial

$$\eta^n + p_n \eta^{n-1} + \dots + p_1$$

and it is clear from elementary determinant theory that the  $\eta_k$ , Eq (3), are the roots of the polynomial equation

$$P(\eta) = \eta^n + p_n \eta^{n-1} + \dots + p_1 = 0. \quad (10)$$

Now from elementary college algebra we have

$$\begin{aligned} P(\eta) &= \prod_{k=1}^n (\eta - \eta_k) \\ &= \prod_{k=1}^m (\eta - \eta_k)(\eta - \bar{\eta}_k) \\ &= \prod_{k=1}^m (\eta^2 - 2\eta \operatorname{Re}[\eta_k] + (\operatorname{Re}[\eta_k])^2 + (\operatorname{Im}[\eta_k])^2) \end{aligned}$$

since the  $n$  roots of Eq (10) occur in complex conjugate pairs. From this it is clear that the coefficients,  $p_k$ , of Eq (9), and also in Eq (10), are real.

Set  $x(t) = x_1$ ,  $x(t + h) = x_2$ , ...,  $x(t + nh) = x_{n+1}$ , and so on. From Eq (9) we obtain the system of equations

$$\begin{bmatrix} x_1 & x_2 & \dots & x_n \\ x_2 & x_3 & \dots & x_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n+1} & \dots & x_{2n-1} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} = \begin{bmatrix} -x_{n-1} \\ -x_{n-2} \\ \vdots \\ -x_{2n} \end{bmatrix}. \quad (11)$$

Thus, in principle, from a set of  $2n$  values of the function  $x(t)$  the coefficients  $p_1, p_2, \dots, p_n$  can be determined, if the coefficient matrix in Eq (11) is nonsingular. From Eq (4) we obtain

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \eta_1 & \eta_2 & \dots & \eta_n \\ \vdots & \vdots & \ddots & \vdots \\ \eta_1^{n-1} & \eta_2^{n-1} & \dots & \eta_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 e^{\lambda_1 t} \\ a_2 e^{\lambda_2 t} \\ \vdots \\ a_n e^{\lambda_n t} \end{bmatrix}$$

and similarly

$$\begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \eta_1 & \eta_2 & \dots & \eta_n \\ \vdots & \vdots & \ddots & \vdots \\ \eta_1^{n-1} & \eta_2^{n-1} & \dots & \eta_n^{n-1} \end{bmatrix} \begin{bmatrix} \eta_1 a_1 e^{\lambda_1 t} \\ \eta_2 a_2 e^{\lambda_2 t} \\ \vdots \\ \eta_n a_n e^{\lambda_n t} \end{bmatrix}$$

and so on. That is, we obtain

$$\begin{bmatrix} x_1 & x_2 & \dots & x_n \\ x_2 & x_3 & \dots & x_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n-1} & \dots & x_{2n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \eta_1 & \eta_2 & \dots & \eta_n \\ \vdots & \vdots & \ddots & \vdots \\ \eta_1^{n-1} & \eta_2^{n-1} & \dots & \eta_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 e^{\lambda_1 t} & \eta_1 a_1 e^{\lambda_1 t} & \dots & \eta_1^{n-1} a_1 e^{\lambda_1 t} \\ a_2 e^{\lambda_2 t} & \eta_2 a_2 e^{\lambda_2 t} & \dots & \eta_2^{n-1} a_2 e^{\lambda_2 t} \\ \vdots & \vdots & \ddots & \vdots \\ a_n e^{\lambda_n t} & \eta_n a_n e^{\lambda_n t} & \dots & \eta_n^{n-1} a_n e^{\lambda_n t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \eta_1 & \eta_2 & \dots & \eta_n \\ \vdots & \vdots & \ddots & \vdots \\ \eta_1^{n-1} & \eta_2^{n-1} & \dots & \eta_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & a_2 e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} 1 & \eta_1 & \dots & \eta_1^{n-1} \\ 1 & \eta_2 & \dots & \eta_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \eta_n & \dots & \eta_n^{n-1} \end{bmatrix} \quad (12)$$

Now the determinant of the left hand side is equal to the product of the determinants of the right hand side. By our assumptions above, it is clear that the value of each of the determinants on the right hand side is different from zero. It follows from our observations, Eq (12), that we may assume that the coefficients  $p_j$  are determined by the Eqs (11) or known.

We consider next the second task, namely determining the zeros of the polynomial  $P(\eta)$ . The zeros  $\eta_k$  of the polynomial  $P(\eta)$  occur in complex conjugate pairs. Hence we only need to determine the zeros lying in the upper half plane. For  $\lambda_k = \alpha_k + i\beta_k$  we have from Eq (3)

$$\eta_k = e^{\alpha_k h} e^{i\beta_k h} \quad (13)$$

In the present discussion we are only considering lightly damped systems. Thus  $\alpha_k < 0$  and is numerically small. Hence  $0 < \exp(\alpha_k h) < 1$  and it is clear from Eq (13) that all the zeros  $\eta_k$  lie inside the unit circle. In particular, if  $\alpha_k$  is sufficiently small numerically and if  $h$  is sufficiently small also then the zeros  $\eta_k$  will be close to 1 in magnitude. That is the roots  $\eta_k$  lie close to the circumference of the unit circle.

If the zeros  $\eta_k$  lie close to the unit circle we are able to determine them to a high degree of accuracy in two "steps". The first step is a linear search along the circumference of the unit circle which gives an initial approximation  $\hat{\eta}_k$  to the zero  $\eta_k$ . The second step is a Newton iteration process starting with this initial approximation.

To implement the first step we have

$$\frac{1}{P(\eta)} \frac{dP}{d\eta} = \frac{1}{\eta - \eta_1} + \dots + \frac{1}{\eta - \eta_n}. \quad (14)$$

Set

$$f(\eta) = \left| \frac{dP/d\eta}{P(\eta)} \right|. \quad (15)$$

By means of a search, determine approximately local maxima  $\hat{\eta}_k$  of  $f(\eta)$  along the circumference of the unit circle in the upper half-plane. One should find  $m$  values of  $\hat{\eta}_k$ ,  $1 \leq k \leq m$ . It is clear that the search is not restricted to the circumference of the unit circle. One may search along the circumference of circles of radii different from 1.

The recurrence relation for a Newton Process for determining an  $\eta$  satisfying  $P(\eta) = 0$  is

$$z_{j+1} = z_j - \frac{P(z_j)}{P'(z_j)} \quad (16)$$

$$z_0 = \hat{\eta}_k$$

where  $P'(\eta)$  denotes the derivative of  $P(\eta)$  with respect to  $\eta$ .

Let  $\eta_k$  denote one of these roots. We have

$$\eta_k = e^{\lambda_k h}$$

which we write as

$$\eta_k = r_k e^{i(\theta_k + 2\pi l)} = e^{\lambda_k h} = e^{(\alpha_k + i\beta_k)h}.$$

From this we obtain

$$\alpha_k = (1/h)\log(\tau_k) \quad (17)$$

$$\beta_k = (\theta_k + 2\pi l)/h. \quad (18)$$

It is clear from Eq (18) that  $\lambda_k$  is not uniquely defined. One takes  $l = 0$  in Eq (18) unless there is good reason to do otherwise.

Once the  $\lambda_k$ 's are determined, it only remains to determine the  $a_k$ 's. The following system of equations is available for determining the  $a_k$ 's

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \eta_1 & \eta_2 & \dots & \eta_n \\ \vdots & \vdots & \ddots & \vdots \\ \eta_1^{n-1} & \eta_2^{n-1} & \dots & \eta_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 e^{\lambda_1 t} \\ a_2 e^{\lambda_2 t} \\ \vdots \\ a_n e^{\lambda_n t} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \quad (19)$$

This system can be solved by a linear systems solver. However, because of the special form of the coefficient matrix one can determine the value of the product  $a_k e^{\lambda_k t}$  in a more efficient way.

Here now we develop the procedure for solving the system, Eqs (19), which utilizes the special form of the coefficient matrix. Denote the cofactor of the  $j k^{\text{th}}$  element of the coefficient matrix by  $B_{jk}$ . Then by Cramer's rule

$$\begin{aligned} \left( \sum_{j=1}^n \eta_k^{j-1} B_{jk} \right) a_k e^{\lambda_k t} &= \sum_{j=1}^n x_j B_{jk} \\ &= \sum_{j=1}^n x(t + (j-1)h) B_{jk}. \end{aligned} \quad (20)$$

Divide Eq (20) through by  $B_{nk}$ . The coefficient of  $a_k e^{\lambda_k t}$  in Eq (20) is a polynomial in  $\eta$  of degree  $n-1$  evaluated at  $\eta = \eta_k$ . We have

$$\begin{aligned} \hat{P}(\eta_k) &= \eta_k^{n-1} + \frac{B_{n-1k}}{B_{nk}} \eta_k^{n-2} + \dots + \frac{B_{1k}}{B_{nk}} \\ &= \sum_{j=1}^n \frac{B_{jk}}{B_{nk}} \eta_k^{j-1}. \end{aligned} \quad (21)$$

It is clear from Eq (19) that the roots of the polynomial equation  $\hat{P}(\eta) = 0$  are the  $\eta_j$ ,  $j \neq k$ . Hence

$$\hat{P}(\eta) = \prod_{\substack{j=1 \\ j \neq k}}^n (\eta - \eta_j). \quad (22)$$

The polynomial, Eq (10), is

$$\begin{aligned} P(\eta) &= \eta^n + p_n \eta^{n-1} + \dots + p_1 \\ &= \prod_{j=1}^n (\eta - \eta_j). \end{aligned} \quad (23)$$

Then

$$\begin{aligned} \frac{dP}{d\eta} &= n\eta^{n-1} + (n-1)p_n \eta^{n-2} + \dots + p_2 \\ &= \sum_{k=1}^n \prod_{\substack{j=1 \\ j \neq k}}^n (\eta - \eta_j) \end{aligned} \quad (24)$$

and

$$\begin{aligned} \left. \frac{dP}{d\eta} \right|_{\eta=\eta_k} &= n\eta_k^{n-1} + (n-1)p_n \eta_k^{n-2} + \dots + p_2 \\ &= \prod_{\substack{j=1 \\ j \neq k}}^n (\eta_k - \eta_j) = \hat{P}(\eta_k). \end{aligned} \quad (25)$$

Thus the coefficient of  $a_k e^{\lambda_k t}$  is

$$\hat{P}(\eta_k) = \left. \frac{dP}{d\eta} \right|_{\eta=\eta_k}. \quad (26)$$

The right hand side, *RHS*, of Eq (20) after division by  $B_{nk}$  is the quantity

$$RHS = \sum_{j=1}^n x(t + (j-1)h) \frac{B_{jk}}{B_{nk}}. \quad (27)$$

That is, *RHS* is the polynomial  $\hat{P}(\eta)$  with  $\eta^j$  replaced by the value  $x(t + jh)$ , where  $j = 0, \dots, n-1$ . We make the following observations:

$$\begin{aligned} \hat{P}(\eta) &= P(\eta)/(\eta - \eta_k) \\ &= \frac{P(\eta) - P(\eta_k)}{\eta - \eta_k}, \text{ since } P(\eta_k) = 0 \\ &= \frac{\eta^n - \eta_k^n}{\eta - \eta_k} + p_n \frac{\eta^{n-1} - \eta_k^{n-1}}{\eta - \eta_k} + \dots + p_2. \end{aligned} \quad (28)$$

For the case  $n = 4$  Eq (28) becomes

$$\begin{aligned} \hat{P}(\eta) &= \eta^3 + \eta^2 \eta_k + \eta \eta_k^2 + \eta_k^3 \\ &\quad + p_4 (\eta^2 + \eta \eta_k + \eta_k^2) \\ &\quad + p_3 (\eta + \eta_k) \\ &\quad + p_2. \end{aligned} \quad (29)$$

Now we rewrite  $\dot{P}(\eta)$  given by Eq (29) as a polynomial in  $\eta_k$ . We have

$$\begin{aligned}\dot{P}(\eta) &= \eta_k^3 \\ &+ (\eta + p_4)\eta_k^2 \\ &+ (\eta^2 + p_4\eta + p_3)\eta_k \\ &+ (\eta^3 + p_4\eta^2 + p_3\eta + p_2).\end{aligned}\tag{30}$$

Finally to obtain the value which we called *RHS* replace  $\eta^j$  in Eq (30) by  $x(t + jh)$ , where  $j = 0, \dots, n - 1$  and obtain, for the case  $n = 4$ ,

$$\begin{aligned}RHS &= x(t)\eta_k^3 \\ &+ (x(t + h) + p_4x(t))\eta_k^2 \\ &+ (x(t + 2h) + p_4x(t + h) + p_3x(t))\eta_k \\ &+ x(t + 3h) + p_4x(t + 2h) + p_3x(t + h) + p_2x(t).\end{aligned}\tag{31}$$

The expression for the value of *RHS* for a general value of  $n$  is readily obtained from Eq (31). We have

$$\begin{aligned}RHS &= x(t)\eta_k^{n-1} \\ &+ (x(t + h) + p_nx(t))\eta_k^{n-2} \\ &+ (x(t + 2h) + p_nx(t + h) + p_{n-1}x(t))\eta_k^{n-3} \\ &\vdots \\ &+ x(t + (n-1)h) + p_nx(t + (n-2)h) + \dots + p_2x(t).\end{aligned}\tag{32}$$

Now we have shown that

$$a_k e^{\lambda_k t}\tag{33}$$

is equal to the quotient of two polynomials of degree  $n - 1$ , each of which is evaluated at  $\eta_k$ . Polynomials are readily and rapidly evaluated. The numerator is given by Eq (32) and the denominator by Eq (26). The polynomials are the same, have the same coefficients, for all values of  $k$ , for  $1 \leq k \leq m$  and, of course,

$$a_{k,m} = a_k.$$

Usually there is no reason for not taking  $t = 0$  in Eq (33).

There is no difficulty in going from the scalar case just discussed to the case where the coefficients  $\mathbf{a}_k$  are p-dimensional vectors. One applies the procedure described above to a component. The  $\lambda_k$  do not have to be determined for each component. Thus the determination of the polynomial  $P(\eta)$  is done for only one component. The coefficients for the polynomial, Eq (32), have to be determined for each component using the observed values for that component.

Reviewing the above development of a procedure for the determination of complex exponentials we see that the first task is the determination of the coefficients  $p_k$  of a difference equation, Eq (9). These coefficients are also the coefficients for a polynomial equation, Eq (10). The second task is the determination of the roots  $\eta_k$  of this polynomial equation. Once the  $\eta_k$  are known the  $\lambda_k$  are also determined by Eqs (17) and (18). The last task is the determination of the  $a_k$ 's or the  $\mathbf{a}_k$ 's using Eqs (26) and (32). If the number  $n$  of complex exponentials is known the first task can be accomplished by solving a system of linear equations, Eq (11). For the case  $n$  not known an alternative to the use of Eq (11) is given in SECTION III.

For the vector case there is the possibility of two or more linearly independent vectors  $\mathbf{a}_k$  as the coefficient of the same complex exponential in the *true* expression for the function  $\mathbf{f}(t)$ , Eq (1). It is clear, however, that the experimental data can only express the sum, that is, these two or more coefficients of the same complex exponential are summed and regarded as one. If one suspects that there are multiple roots one must process other sets of data with different initial conditions. Let  $\mathbf{a}_{k1}$  and  $\mathbf{a}_{k2}$  denote the coefficients corresponding to the value  $\lambda_k$  obtained from two different sets of data. If  $\lambda_k$  is not a multiple root  $\mathbf{a}_{k2}$  will be some scalar multiple of  $\mathbf{a}_{k1}$ . On the other hand if  $\lambda_k$  is a multiple root then

$$\mathbf{a}_{k2} - c\mathbf{a}_{k1} \neq 0$$

for any value of the scalar  $c$ . Hence one must process a number of data sets equal to the multiplicity of the  $\lambda_k$ .

We have shown that a set of  $2n$  equally spaced values of the function  $x(t)$ , Eq (1), completely determine the function  $x(t)$  in principle. Next we show that the algorithms developed for determining the function  $x(t)$  apply also to appropriate *averages* of the evenly spaced values of  $x(t)$ .

Suppose that we have very many, say  $N \gg 2n$ , equally spaced values of the function  $x(t)$ . From Eq (11) we have, for example

$$\begin{aligned} x_1 p_1 + x_2 p_2 + \cdots + x_n p_n &= -x_{n+1} \\ x_2 p_1 + x_3 p_2 + \cdots + x_{n+1} &= -x_{n+2} \\ &\vdots \\ x_r p_1 + x_{r+1} p_2 + \cdots + x_{n+r-1} p_n &= -x_{n+r}. \end{aligned}$$

Adding these equations together and dividing by  $r$  we obtain

$$\left(\frac{1}{r} \sum_{k=1}^r x_k\right) p_1 + \left(\frac{1}{r} \sum_{k=1}^r x_{k+1}\right) p_2 + \cdots + \left(\frac{1}{r} \sum_{k=1}^r x_{n+k-1}\right) p_n = -\frac{1}{r} \sum_{k=1}^r x_{n+k}$$

Set

$$\hat{x}_j = \frac{1}{r} \sum_{k=1}^r x_{k+j-1}. \quad (34)$$

Thus replacing the value  $x_j$  in Eq (11) by the average  $\hat{x}_j$  we obtain equations of exactly the same form as Eqs (11) for determining the coefficients  $p_k$ .

From Eq (4), see also the equations immediately preceding and leading up to Eq (12), we have

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_r \\ x_2 & x_3 & \cdots & x_{r+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n+1} & \cdots & x_{n+r-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \eta_1 & \eta_2 & \cdots & \eta_n \\ \vdots & \vdots & \ddots & \vdots \\ \eta_1^{n-1} & \eta_2^{n-1} & \cdots & \eta_n^{n-1} \end{bmatrix} \begin{bmatrix} a_1 e^{\lambda_1 t} & a_1 e^{\lambda_1 t} \eta_1 & \cdots & a_1 e^{\lambda_1 t} \eta_1^{r-1} \\ a_2 e^{\lambda_2 t} & a_2 e^{\lambda_2 t} \eta_2 & \cdots & a_2 e^{\lambda_2 t} \eta_2^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n e^{\lambda_n t} & a_n e^{\lambda_n t} \eta_n & \cdots & a_n e^{\lambda_n t} \eta_n^{r-1} \end{bmatrix} \quad (35)$$

Multiplying this equation, both sides, on the right by an  $r$  vector with all components

equal to 1 and dividing through by  $r$  we obtain

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \eta_1 & \eta_2 & \dots & \eta_n \\ \vdots & \vdots & \ddots & \vdots \\ \eta_1^{n-1} & \eta_2^{n-1} & \dots & \eta_n^{n-1} \end{bmatrix} \begin{bmatrix} \frac{1}{r}(1 + \eta_1 + \dots + \eta_1^{r-1})a_1 e^{\lambda_1 t} \\ \frac{1}{r}(1 + \eta_2 + \dots + \eta_2^{r-1})a_2 e^{\lambda_2 t} \\ \vdots \\ \frac{1}{r}(1 + \eta_n + \dots + \eta_n^{r-1})a_n e^{\lambda_n t} \end{bmatrix}. \quad (36)$$

Thus when using average values as given by Eq (34) we obtain, instead of Eq (19) which we solved for the product  $a_k e^{\lambda_k t}$ , Eq (36) which is solved for the quantity

$$\frac{1}{r}(1 + \eta_k + \dots + \eta_k^{r-1})a_k e^{\lambda_k t}.$$

The algorithm given for solving for the value of  $a_k e^{\lambda_k t}$  is also applicable for solving for the quantity

$$\frac{1}{r}(1 + \eta_k + \dots + \eta_k^{r-1})a_k e^{\lambda_k t}.$$

## SECTION III

### THE PROGRESSIVE ALGORITHM

In this section we consider the *progressive* algorithm which was used by C. Lanczos in [2]. This procedure is useful when the number of complex exponentials in the function, Eq (1), is not known. To develop the algorithm we first describe the procedure from the point of view of systems of equations. Then we repeat the description using matrix notation. The essential computations are summarized in Eqs (55-60).

For discussion purposes we assume a set of numbers  $x_k$ , and suppose there is a least value  $n$  and a vector having components  $\eta_1, \dots, \eta_{n-1}, 1$  satisfying the condition

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ x_2 & x_3 & \cdots & x_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n+1} & \cdots & x_{2n-1} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ 1 \end{bmatrix} = 0. \quad (37)$$

Such a system of homogeneous equations arises in the obvious way from Eq (9). Typically for a problem of this nature one might delete a row, say the last row of this system, Eq (37), take the entries remaining in the last column to the other side of the equal sign and in this way obtain the system of equations, Eq (11). The difficulty is that the value of  $n$  is not known. We could still use such an attack by taking, in turn,  $n = 3, 4, \dots$  until we obtained a value of  $n$  for which Eq (37) was satisfied in some numerical sense. The special form of the coefficient matrix, Eq (37), enables one to develop an efficient procedure for carrying out this approach.

Suppose we have obtained, some how, for a value of  $k < n$  the following conditions

$$\begin{aligned} x_1 \xi_1 + x_2 \xi_2 + \cdots + x_k \xi_k + x_{k+1} &= 0 \\ x_2 \xi_1 + x_3 \xi_2 + \cdots + x_{k-1} \xi_k + x_{k-2} &= 0 \\ &\vdots \\ x_{k+1} \xi_1 + x_{k-2} \xi_2 + \cdots + x_{2k} \xi_k + x_{2k-1} &= h_1 \end{aligned} \quad (38)$$

and

$$\begin{aligned}
 x_2 \zeta_1 + x_3 \zeta_2 + \cdots + x_{k+1} \zeta_k + x_{k+2} &= 0 \\
 x_3 \zeta_1 + x_4 \zeta_2 + \cdots + x_{k+2} \zeta_k + x_{k+3} &= 0 \\
 &\vdots \\
 x_{k+2} \zeta_1 + x_{k+3} \zeta_2 + \cdots + x_{2k+1} \zeta_k + x_{2k+2} &= h_2.
 \end{aligned} \tag{39}$$

Observe that by multiplying each of the Eqs (38) by  $-h_2/h_1$  and adding the resulting system of equations to the Eqs (39) one obtains

$$\begin{aligned}
 x_1 \eta_1 + x_2 \eta_2 + \cdots + x_{k-1} \eta_{k+1} + x_{k+2} &= 0 \\
 x_2 \eta_1 + x_3 \eta_2 + \cdots + x_{k+2} \eta_{k+1} + x_{k+3} &= 0 \\
 &\vdots \\
 x_{k+1} \eta_1 + x_{k+2} \eta_2 + \cdots + x_{2k+1} \eta_{k+1} + x_{2k+2} &= 0
 \end{aligned} \tag{40}$$

where

$$\begin{aligned}
 \eta_1 &= (-h_2/h_1) \xi_1 \\
 \eta_2 &= (-h_2/h_1) \xi_2 + \zeta_1 \\
 &\vdots \\
 \eta_k &= (-h_2/h_1) \xi_k + \zeta_{k-1} \\
 \eta_{k+1} &= (-h_2/h_1) \xi_{k+1} + \zeta_k.
 \end{aligned} \tag{41}$$

Next compute

$$x_{k+2} \eta_1 + x_{k+3} \eta_2 + \cdots + x_{2k+2} \eta_{k+1} + x_{2k+3} = h. \tag{42}$$

If  $h = 0$  then the set of equations comprised of the Eqs (40) and Eq (42) is the system Eq (37) and the desired solution vector is given by the Eqs (41). If  $h \neq 0$  then the system, Eqs (40) and Eq (42), will play the role of the system, Eq (38). We need to obtain the system which will play the role of the system, Eq (39). To obtain the new system Eq (39) multiply the old Eq (39) by  $-h/h_2$  and add to the new system Eq (38) with the first row

deleted. We have

$$\begin{aligned}
 x_2(\eta_1 - \zeta_1 h/h_2) + x_3(\eta_2 - \zeta_2 h/h_2) + \cdots + x_{k+2}(\eta_{k+1} - h/h_2) + x_{k+3} &= 0 \\
 x_3(\eta_1 - \zeta_1 h/h_2) + x_4(\eta_2 - \zeta_2 h/h_2) + \cdots + x_{k+3}(\eta_{k+1} - h/h_2) + x_{k+4} &= 0 \\
 &\vdots
 \end{aligned} \tag{43}$$

$$x_{k+2}(\eta_1 - \zeta_1 h/h_2) + x_{k+3}(\eta_2 - \zeta_2 h/h_2) + \cdots + x_{2k+2}(\eta_{k+1} - h/h_2) + x_{2k+3} = 0.$$

From the Eqs (43) set

$$\begin{aligned}
 \hat{\eta}_1 &= \eta_1 - \zeta_1 h/h_2 \\
 \hat{\eta}_2 &= \eta_2 - \zeta_2 h/h_2 \\
 &\vdots \\
 \hat{\eta}_k &= \eta_k - \zeta_k h/h_2 \\
 \hat{\eta}_{k+1} &= \eta_{k+1} - h/h_2.
 \end{aligned} \tag{44}$$

Using the values obtained in the Eqs (44) compute

$$x_{k+3}\hat{\eta}_1 + x_{k+4}\hat{\eta}_2 + \cdots + x_{2k+3}\hat{\eta}_{k+1} + x_{2k+4} = \hat{h}. \tag{45}$$

Replacing the  $\eta$ 's by  $\zeta$ 's,  $h$  by  $h_1$ , the  $\hat{\eta}$ 's by  $\zeta$ 's and  $\hat{h}$  by  $h_2$  the Eqs (40) and (42) are of the form Eqs (38) and Eqs (43) and (45) are of the form Eqs (39). We are now ready to repeat the process described above.

The Eqs (38-45) completely describe the *Progressive Algorithm*, but from these equations it is somewhat difficult to see what is going on. Here now we will develop the progressive algorithm using matrix notation. The Eqs (38) and (39) can be written as

$$\begin{aligned}
 X_{k1} U_k &= H_{k1} \\
 X_{k2} V_k &= H_{k2}
 \end{aligned} \tag{46}$$

where

$$\begin{aligned}
 X_{k1} &= \begin{bmatrix} x_1 & x_2 & \cdots & x_{k+1} \\ x_2 & x_3 & \cdots & x_{k+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k+1} & x_{k+2} & \cdots & x_{2k+1} \end{bmatrix} \\
 X_{k2} &= \begin{bmatrix} x_2 & x_3 & \cdots & x_{k+2} \\ x_3 & x_4 & \cdots & x_{k+3} \\ \vdots & \vdots & \ddots & \vdots \\ x_{k-2} & x_{k-3} & \cdots & x_{2k-2} \end{bmatrix}
 \end{aligned} \tag{47}$$

and

$$U_k = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_k \\ 1 \end{bmatrix}, H_{k1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ h_{k1} \end{bmatrix}, V_k = \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_k \\ 1 \end{bmatrix}, H_{k2} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ h_{k2} \end{bmatrix}. \quad (48)$$

We think of the matrix  $X_{k1}$  as partitioned into two block submatrices

$$X_{k11} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k+1} \end{bmatrix}, X_{k12} = \begin{bmatrix} x_2 & \cdots & x_{k+1} \\ x_3 & \cdots & x_{k+2} \\ \vdots & \ddots & \vdots \\ x_{k-2} & \cdots & x_{2k+1} \end{bmatrix}.$$

The block submatrices of the matrix  $X_{k2}$  are

$$X_{k21} = \begin{bmatrix} x_2 & \cdots & x_{k+1} \\ x_3 & \cdots & x_{k+2} \\ \vdots & \ddots & \vdots \\ x_{k+2} & \cdots & x_{2k+1} \end{bmatrix}, X_{k22} = \begin{bmatrix} x_{k+2} \\ x_{k+3} \\ \vdots \\ x_{2k+2} \end{bmatrix}.$$

Observe that

$$X_{k12} = X_{k21}.$$

The vectors  $U_k$  and  $V_k$  are partitioned in the obvious way so as to be compatible with the partitioning of the matrices  $X_{k1}$  and  $X_{k2}$ . We have

$$U_{k1} = [\xi_1], U_{k2} = \begin{bmatrix} \xi_2 \\ \vdots \\ \xi_k \\ 1 \end{bmatrix}$$

and

$$V_{k1} = \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_k \end{bmatrix}, V_{k2} = [1].$$

We have now

$$(-h_{k2}/h_{k1})X_{k1}U_k = X_{k1}(-h_{k2}/h_{k1})U_k = (-h_{k2}/h_{k1})H_{k1} = -H_{k2}.$$

Using the block submatrices we now write this equation and the equation

$$X_{k2}V_k = H_{k2}$$

as

$$X_{k11}(-h_{k2}/h_{k1})U_{k1} + X_{k12}(-h_{k2}/h_{k1})U_{k2} = -H_{k2}$$

$$X_{k21}V_{k1} + X_{k22} = H_{k2}.$$

Adding these two equations we obtain

$$[X_{k11} \quad X_{k12} \quad X_{k22}] \begin{bmatrix} (-h_{k2}/h_{k1})\xi_1 \\ V_{k1} - (h_{k2}/h_{k1})U_{k2} \\ 1 \end{bmatrix} = 0.$$

Set

$$U_{k+1} = \begin{bmatrix} (-h_{k2}/h_{k1})\xi_1 \\ V_{k1} - (h_{k2}/h_{k1})U_{k2} \\ 1 \end{bmatrix} \quad (49)$$

and

$$h_{k+11} = [x_{k+2} \quad \dots \quad x_{2k+3}]U_{k+1} \quad (50),$$

the scalar product of the row vector  $[x_{k+2} \quad \dots \quad x_{2k+3}]$  and the column vector  $U_{k+1}$ .

Adjoining the row vector

$$[x_{k+2} \quad \dots \quad x_{2k+3}]$$

to the matrix

$$[X_{k11} \quad X_{k12} \quad X_{k22}]$$

we have

$$X_{k+11} = \begin{bmatrix} x_1 & x_2 & \dots & x_{k-1} & x_{k-2} \\ x_2 & x_3 & \dots & x_{k-2} & x_{k-3} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{k+2} & x_{k+3} & \dots & x_{2k-2} & x_{2k+3} \end{bmatrix}.$$

The Eqs (49) and (50) show the computations involved in going from the  $k^{th}$  stage to the  $(k+1)^{th}$  stage for the matrix  $X_{k+1}$ . If  $h_{k+11} = 0$  we are finished, if not, we have to obtain the vector  $V_{k+1}$  and the value  $h_{k+12}$  corresponding to the matrix  $X_{k+12}$ .

Delete the first row of the matrix  $X_{k+11}$  and denote the matrix arising from this deletion by  $X^*$ . Set

$$H = X^* U_{k-1} \quad (51)$$

and note that

$$X_{k2}(-h_{k+11}/h_{k2})V_k = -H. \quad (52)$$

Note also that

$$X^* = [X_{k2} \quad X]$$

where

$$X = \begin{bmatrix} x_{k+3} \\ \vdots \\ x_{2k+3} \end{bmatrix}.$$

Now we can write

$$U_{k+1} = \begin{bmatrix} U_{k+11} \\ 1 \end{bmatrix}$$

and adding Eqs (51) and (52) we obtain

$$X^* \begin{bmatrix} U_{k+11} - (h_{k+11}/h_{k2})V_k \\ 1 \end{bmatrix} = 0.$$

Set

$$V_{k+1} = \begin{bmatrix} U_{k+11} - (h_{k+11}/h_{k2})V_k \\ 1 \end{bmatrix} \quad (53)$$

and

$$h_{k+12} = [x_{k+3} \quad \dots \quad x_{2k+4}]V_{k+1}. \quad (54)$$

Adjoining the row vector  $[x_{k+3} \quad \dots \quad x_{2k+4}]$  to the matrix  $X^*$  we obtain the matrix  $X_{k+12}$ .

Once the progressive algorithm is started the essential computations are those given by Eqs (49) and (50) and Eqs (53) and (54). On examining these equations one sees that the progressive algorithm could *fail* if for  $h_{k1} \neq 0$  the value  $h_{k2} = 0$ . For our application it is clear, in principle, at least, that such a condition cannot arise, since for an appropriate value of  $k$  both  $X_{k1}$  and  $X_{k2}$  should determine the coefficients of the polynomial  $P(\eta)$ . Eq (10).

Now it is clear also that in numerical work we cannot expect to obtain  $h_{k1} = 0$ . Thus in practice the algorithm will terminate when the value  $h_{k1}$  is sufficiently small.

We will now summarize the above algorithm, using superscripts to distinguish one stage of the computations from the following stage. The algorithm is as follows:

For  $k = 1$

$$\xi_1^{(1)} = -x_2/x_1 \quad \text{and} \quad x_2\xi_1^{(1)} + x_3 = h_1^{(1)} \quad (55)$$

$$\zeta_1^{(1)} = -x_3/x_2 \quad \text{and} \quad x_3\zeta_1^{(1)} + x_4 = h_2^{(1)}. \quad (56)$$

For  $k > 1$

$$\begin{aligned} \xi_1^{(k+1)} &= (-h_2^{(k)}/h_1^{(k)})\xi_1^{(k)} \\ \xi_2^{(k+1)} &= (-h_2^{(k)}/h_1^{(k)})\xi_2^{(k)} + \zeta_1^{(k)} \\ &\vdots \\ \xi_k^{(k+1)} &= (-h_2^{(k)}/h_1^{(k)})\xi_k^{(k)} + \zeta_{k-1}^{(k)} \\ \xi_{k+1}^{(k+1)} &= (-h_2^{(k)}/h_1^{(k)}) + \zeta_k^{(k)} \end{aligned} \quad (57)$$

and

$$h_1^{(k+1)} = x_{k+2}\xi_1^{(k+1)} + \cdots + x_{2k+2}\xi_{k+1}^{(k+1)} + x_{2k+3}. \quad (58)$$

If  $h_1^{(k+1)}$  is numerically sufficiently small, finished. If not, then

$$\begin{aligned} \zeta_1^{(k+1)} &= \xi_1^{(k+1)} - \zeta_1^{(k)}h_1^{(k+1)}/h_2^{(k)} \\ \zeta_2^{(k+1)} &= \xi_2^{(k+1)} - \zeta_2^{(k)}h_1^{(k+1)}/h_2^{(k)} \\ &\vdots \\ \zeta_k^{(k+1)} &= \xi_k^{(k+1)} - \zeta_k^{(k)}h_1^{(k+1)}/h_2^{(k)} \\ \zeta_{k+1}^{(k+1)} &= \xi_{k+1}^{(k+1)} - h_1^{(k+1)}/h_2^{(k)} \end{aligned} \quad (59)$$

and

$$h_2^{(k+1)} = x_{k+3}\zeta_1^{(k+1)} + \cdots + x_{2k+3}\zeta_{k-1}^{(k+1)} + x_{2k+4}. \quad (60)$$

We note that the progressive algorithm is equally applicable if the  $x_j$  in Eq (37) are replaced by the averages  $\hat{x}_j$ , Eq (34).

## SECTION IV

### THE METHOD OF LEAST SQUARES

In this section we examine in some detail the method of *least squares*. For points  $\mathbf{x}_k = (x_{1k}, \dots, x_{nk})$  in  $n$ -space and  $\mathbf{y}_k = (y_{1k}, \dots, y_{mk})$  in  $m$ -space, we suppose

$$\mathbf{y}_k = g(\mathbf{x}_k).$$

The function  $g(\mathbf{x})$  is either not known or not readily evaluated. For some simpler function

$$\hat{\mathbf{y}} = f(\mathbf{x}, \mathbf{a})$$

depending on a parameter set  $\mathbf{a}$ , the parameter set  $\mathbf{a}$  is determined so that the sum of the squares of the magnitude of the error,  $\mathbf{y}_k - \hat{\mathbf{y}}_k = \mathbf{y}_k - f(\mathbf{x}_k, \mathbf{a})$  is least. Because of our special interest this review only examines the case  $\mathbf{y}$  depending linearly on  $\mathbf{x}$ .

For this case we regard the points  $\mathbf{x}_k$  and  $\mathbf{y}_k$  as  $n$ -vectors and  $m$ -vectors respectively. Thus

$$\mathbf{x}_k = [x_{1k} \quad \dots \quad x_{nk}]^T,$$

$$\mathbf{y}_k = [y_{1k} \quad \dots \quad y_{mk}]^T$$

and, for example,

$$\|\mathbf{y}_k\|^2 = \sum_{j=1}^m y_{jk}^2.$$

We will limit our discussion to the case where the number  $N$  of such vectors is finite but much greater than either  $m$  or  $n$ . For an  $(m \times n)$  matrix  $A$  set

$$\hat{\mathbf{y}}_k = A\mathbf{x}_k \tag{61}$$

and the error

$$\mathbf{e}_k = \mathbf{y}_k - \hat{\mathbf{y}}_k. \tag{62}$$

Desired is a matrix  $A$  for which the average

$$\begin{aligned} E[\|\mathbf{e}_k\|^2] &= E[\|\mathbf{y}_k - \hat{\mathbf{y}}_k\|^2] \\ &= \frac{1}{N} \sum_{k=1}^N \|\mathbf{e}_k\|^2 \end{aligned} \tag{63}$$

is minimal. For  $k = 1, 2, \dots$  we obtain from Eqs (61) and (62) the system of equations for the first component of the error

$$e_{1k} = y_{1k} - a_{11}x_{1k} - a_{12}x_{2k} - \dots - a_{1n}x_{nk} \quad (64)$$

and then

$$\sum_{k=1}^N e_{1k}^2 = \sum_{k=1}^N \left( y_{1k} - \sum_{j=1}^n a_{1j}x_{jk} \right)^2. \quad (65)$$

In this equation we have a function of the parameters  $a_{11}, \dots, a_{1n}$ . We know from the calculus that to find an extremum, we differentiate with respect to  $a_{1l}$ , equate to zero for  $l = 1, \dots, n$ , and solve the resulting system of equations for the  $a_{1j}$ . We have on differentiating Eq (65) with respect to  $a_{1l}$ , for  $1 \leq l \leq n$

$$\sum_{k=1}^N \left( y_{1k}x_{lk} - \sum_{j=1}^n a_{1j}x_{jk}x_{lk} \right) = 0. \quad (66)$$

These equations can be rewritten as

$$\left( \frac{1}{N} \sum_{k=1}^N x_{1k}x_{lk} \right) a_{11} + \dots + \left( \frac{1}{N} \sum_{k=1}^N x_{nk}x_{lk} \right) a_{1n} = \frac{1}{N} \sum_{k=1}^N y_{1k}x_{lk}. \quad (67)$$

Set

$$B_{xx}(j, l) = \frac{1}{N} \sum_{k=1}^N x_{jk}x_{lk} \quad (68)$$

$$B_{yx}(1, l) = \frac{1}{N} \sum_{k=1}^N y_{1k}x_{lk}. \quad (69)$$

The functions  $B_{xx}(j, l)$  and  $B_{yx}(j, l)$  are called the correlation function and the cross-correlation function respectively. When there is no danger of confusion the subscript  $xx$  will be omitted.

The Eqs (67) determine the first row of the matrix  $A$  so that the average of the square of the first component of the error is minimal. In the same way the remaining rows of  $A$  can be determined to minimize the average of the square of the corresponding component of the error. Using the notation just introduced for the sums, the equation for the desired matrix  $A$  can be written as

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} B_{xx}(1, 1) & \dots & B_{xx}(1, n) \\ \vdots & \ddots & \vdots \\ B_{xx}(n, 1) & \dots & B_{xx}(n, n) \end{bmatrix} = \begin{bmatrix} B_{yx}(1, 1) & \dots & B_{yx}(1, n) \\ \vdots & \ddots & \vdots \\ B_{yx}(m, 1) & \dots & B_{yx}(m, n) \end{bmatrix} \quad (70)$$

It is clear from Eq (68) that the coefficient matrix of  $A$  is symmetric.

It is convenient at this point to obtain a relation which will be useful later. Let us write

$$\begin{aligned} e_{1k}^2 &= \left( y_{1k} - \sum_{j=1}^n a_{1j} x_{jk} \right) \left( y_{1k} - a_{11} x_{1k} - \cdots - a_{1n} x_{nk} \right) \\ &= y_{1k}^2 - \sum_{j=1}^n a_{1j} x_{jk} y_{1k} \\ &\quad - a_{11} \left( y_{1k} x_{1k} - \sum_{j=1}^n a_{1j} x_{jk} x_{1k} \right) \\ &\quad - \cdots - a_{1n} \left( y_{1k} x_{nk} - \sum_{j=1}^n a_{1j} x_{jk} x_{nk} \right). \end{aligned}$$

Then

$$\begin{aligned} \sum_{k=1}^N e_{1k}^2 &= \sum_{k=1}^N \left( y_{1k}^2 - \sum_{j=1}^n a_{1j} x_{jk} y_{1k} \right) \\ &\quad - a_{11} \sum_{k=1}^N \left( y_{1k} x_{1k} - \sum_{j=1}^n a_{1j} x_{jk} x_{1k} \right) \\ &\quad - \cdots - a_{1n} \sum_{k=1}^N \left( y_{1k} x_{nk} - \sum_{j=1}^n a_{1j} x_{jk} x_{nk} \right). \end{aligned} \tag{71}$$

It follows from Eqs (66) and (71) that

$$\begin{aligned} E[e_{1k}^2] &= E[y_{1k} y_{1k}] - \sum_{j=1}^n a_{1j} E[x_{jk} y_{1k}] \\ &= B_{yy}(1, 1) - \sum_{j=1}^n a_{1j} B_{zy}(j, 1) \\ &= B_{yy}(1, 1) - \sum_{j=1}^n a_{1j} B_{yz}(1, j). \end{aligned} \tag{72}$$

Hence for the average of the  $i^{\text{th}}$  component of the error

$$\begin{aligned} E[e_{ik}^2] &= E[y_{ik} y_{ik}] - \sum_{j=1}^n a_{ij} E[x_{jk} y_{ik}] \\ &= B_{yy}(i, i) - \sum_{j=1}^n a_{ij} B_{zy}(j, i) \\ &= B_{yy}(i, i) - \sum_{j=1}^n a_{ij} B_{yz}(i, j). \end{aligned} \tag{73}$$

We want to obtain the same results in another way. Let us regard the vectors  $\mathbf{x}_k$  as the columns of an  $(n \times N)$  matrix. Denote the transpose of this matrix by  $X$ . In the same

way the vectors  $\mathbf{y}_k$  are regarded as the columns of an  $(m \times N)$  matrix and the transpose of this matrix is  $Y$ . Let  $X_j$  and  $Y_j$  denote the  $j^{\text{th}}$  columns of  $X$  and  $Y$  respectively.

$$X_j = [x_{j1} \quad x_{j2} \quad \dots \quad x_{jN}]^T$$

$$Y_j = [y_{j1} \quad y_{j2} \quad \dots \quad y_{jN}]^T.$$

That is the  $k^{\text{th}}$  component of  $X_j$  is the  $j^{\text{th}}$  component of  $\mathbf{x}_k$ , and similarly, the  $k^{\text{th}}$  component of  $Y_j$  is the  $j^{\text{th}}$  component of  $\mathbf{y}_k$ . The vectors  $X_j$  and  $Y_j$  are  $N$ -dimensional vectors. A scalar product,  $\langle X_j, X_l \rangle$  or  $\langle Y_j, Y_l \rangle$ , for these vectors is given by Eqs (68) and (69). It is obvious that the conditions for a scalar product are satisfied, since the scalar product defined here differs from the customary scalar product only by the factor  $1/N$ . Set

$$\|X_j\|_m^2 = \langle X_j, X_j \rangle. \quad (74)$$

Consider the system of equations

$$Y_j = a_{j1}X_1 + \dots + a_{jn}X_n. \quad (75)$$

This is an over determined system. It has a strict solution if and only if  $Y_j$  lies in the space  $\mathcal{G}$  spanned by the vectors  $X_1, \dots, X_n$ . If  $Y_j$  is not in the space  $\mathcal{G}$  then for the vector  $\hat{Y}_j$  in  $\mathcal{G}$  closest to  $Y_j$ , the difference  $Y_j - \hat{Y}_j$  is orthogonal to every vector  $X_l$ , for  $1 \leq l \leq n$ .

The proof of this assertion is by contradiction. Suppose  $\hat{X}$  is a vector of the spanning set  $X_1, \dots, X_n$  which is not orthogonal to  $Y_j - \hat{Y}_j$ . Set

$$b = \langle Y_j - \hat{Y}_j, \hat{X} \rangle \quad (76)$$

and

$$Y_0 = \hat{Y}_j + \frac{b}{\|\hat{X}\|_m^2} \hat{X}. \quad (77)$$

Then

$$\begin{aligned} \|Y_j - Y_0\|_m^2 &= \langle Y_j - \hat{Y}_j - \frac{b}{\|\hat{X}\|_m^2} \hat{X}, Y_j - \hat{Y}_j - \frac{b}{\|\hat{X}\|_m^2} \hat{X} \rangle \\ &= \langle Y_j - \hat{Y}_j, Y_j - \hat{Y}_j \rangle \\ &\quad - \frac{2b}{\|\hat{X}\|_m^2} \langle Y_j - \hat{Y}_j, \hat{X} \rangle + \frac{b^2}{\|\hat{X}\|_m^2} \\ &= \|Y_j - \hat{Y}_j\|_m^2 - \frac{b^2}{\|\hat{X}\|_m^2}. \end{aligned} \quad (78)$$

From this we see that if some member  $\hat{X}$  of the spanning set  $X_1, \dots, X_n$  is not orthogonal to  $Y_j - \hat{Y}_j$ , then there is a vector  $Y_0$  in  $\mathcal{G}$  closer to  $Y_j$  than  $\hat{Y}_j$ .

Hence, instead of trying to solve the over determined system, Eq(75), we want to determine the vector

$$\hat{Y}_j = a_{j1}X_1 + \dots + a_{jn}X_n \quad (79)$$

closest to  $Y_j$ . We have then

$$Y_j - \hat{Y}_j = Y_j - a_{j1}X_1 - \dots - a_{jn}X_n. \quad (80)$$

Note, for  $j = 1$  this equation is just the system, Eq (64), above. Since, as has just been shown, this difference must be orthogonal to the spanning vectors  $X_l$  for  $1 \leq l \leq n$ , we obtain the system of equations

$$a_{j1}\langle X_l, X_1 \rangle + \dots + a_{jn}\langle X_l, X_n \rangle = \langle X_l, Y_j \rangle. \quad (81)$$

It is clear that we can rewrite Eq (81) as

$$[a_{j1} \quad \dots \quad a_{jn}] \begin{bmatrix} B_{xr}(1,l) \\ \vdots \\ B_{xr}(n,l) \end{bmatrix} = B_{yr}(j,l).$$

Doing so for  $l = 1, \dots, n$  and  $j = 1, \dots, m$  we obtain again the Eq (70).

A difficulty, albeit minor, with this last derivation is that it focuses attention on the vectors  $X_j$  and  $Y_j$ . There is a tendency to lose track of the relation of the result obtained to the vectors  $\mathbf{x}_k$  and  $\mathbf{y}_k$ . The Eqs (81) determine coefficients  $a_{j1}, \dots, a_{jn}$  and in turn by Eq (79) a  $\hat{Y}_j$  for which  $\|Y_j - \hat{Y}_j\|_m^2$  is minimal. The  $k^{th}$  component of  $Y_j$  is  $y_{jk}$ , hence denote the  $k^{th}$  component of  $\hat{Y}_j$  by  $\hat{y}_{jk}$ . Then

$$\|Y_j - \hat{Y}_j\|_m^2 = \frac{1}{N} \sum_{k=1}^N (y_{jk} - \hat{y}_{jk})^2$$

and

$$\begin{aligned} \sum_{j=1}^m \|Y_j - \hat{Y}_j\|_m^2 &= \frac{1}{N} \sum_{k=1}^N \sum_{j=1}^m (y_{jk} - \hat{y}_{jk})^2 \\ &= \frac{1}{N} \sum_{k=1}^N \|\mathbf{y}_k - \hat{\mathbf{y}}_k\|^2 \\ &= E[\|\mathbf{y}_k - \hat{\mathbf{y}}_k\|^2]. \end{aligned}$$

From Eq (79) we have for the  $k^{\text{th}}$  component of  $\hat{Y}_j$ ,

$$\hat{y}_{jk} = a_{j1}x_{1k} + \dots + a_{jn}x_{nk}.$$

Thus for  $j = 1, \dots, m$  the system, Eqs (81) determine the elements of a matrix  $A$  for which

$$\hat{y}_k = Ax_k$$

and  $E[\|y_k - \hat{y}_k\|^2]$  is minimal as desired.

Again from Eqs (61) and (62) we have

$$\begin{aligned} e_k &= y_k - Ax_k \\ &= [I_m \quad -A] \begin{bmatrix} y_k \\ x_k \end{bmatrix}. \end{aligned} \quad (82)$$

Post multiplying this equation with  $[y_k^T \quad x_k^T]$  we obtain a matrix equation in block form

$$[e_k y_k^T \quad e_k x_k^T] = [I_m \quad -A] \begin{bmatrix} y_k y_k^T & y_k x_k^T \\ x_k y_k^T & x_k x_k^T \end{bmatrix}$$

This gives a system of matrix equations

$$e_k y_k^T = y_k y_k^T - Ax_k y_k^T \quad (83)$$

$$e_k x_k^T = y_k x_k^T - Ax_k x_k^T. \quad (84)$$

For the left side of Eq (84) we have

$$e_k x_k^T = \begin{bmatrix} e_{1k}x_{1k} & \dots & e_{1k}x_{nk} \\ \vdots & \ddots & \vdots \\ e_{mk}x_{1k} & \dots & e_{mk}x_{nk} \end{bmatrix}$$

and

$$E[e_k x_k^T] = \begin{bmatrix} \langle Y_1 - \hat{Y}_1, X_1 \rangle & \dots & \langle Y_1 - \hat{Y}_1, X_n \rangle \\ \vdots & \ddots & \vdots \\ \langle Y_m - \hat{Y}_m, X_1 \rangle & \dots & \langle Y_m - \hat{Y}_m, X_n \rangle \end{bmatrix}.$$

Now we have just seen that for the  $\hat{Y}_j$  closest to  $Y_j$  the difference, that is, the least error, must be orthogonal to the spanning vectors  $X_1, \dots, X_n$ . It follows that

$$E[e_k x_k^T] = 0 \quad (85)$$

and hence

$$AE[\mathbf{x}_k \mathbf{x}_k^T] = E[\mathbf{y}_k \mathbf{x}_k^T]. \quad (86)$$

That is

$$A \begin{bmatrix} E[x_{1k}x_{1k}] & \dots & E[x_{1k}x_{nk}] \\ \vdots & \ddots & \vdots \\ E[x_{nk}x_{1k}] & \dots & E[x_{nk}x_{nk}] \end{bmatrix} = \begin{bmatrix} E[y_{1k}x_{1k}] & \dots & E[y_{1k}x_{nk}] \\ \vdots & \ddots & \vdots \\ E[y_{mk}x_{1k}] & \dots & E[y_{mk}x_{nk}] \end{bmatrix} \quad (87)$$

and this is just Eq (70) again.

From Eqs (87) and (70) let us set

$$\begin{aligned} \mathcal{B}_{xx} &= \begin{bmatrix} E[x_{1k}x_{1k}] & \dots & E[x_{1k}x_{nk}] \\ \vdots & \ddots & \vdots \\ E[x_{nk}x_{1k}] & \dots & E[x_{nk}x_{nk}] \end{bmatrix} = \begin{bmatrix} B_{xx}(1,1) & \dots & B_{xx}(n,1) \\ \vdots & \ddots & \vdots \\ B_{xx}(1,n) & \dots & B_{xx}(n,n) \end{bmatrix} \\ \mathcal{B}_{yx} &= \begin{bmatrix} E[y_{1k}x_{1k}] & \dots & E[y_{1k}x_{nk}] \\ \vdots & \ddots & \vdots \\ E[y_{mk}x_{1k}] & \dots & E[y_{mk}x_{nk}] \end{bmatrix} = \begin{bmatrix} B_{yx}(1,1) & \dots & B_{yx}(1,n) \\ \vdots & \ddots & \vdots \\ B_{yx}(m,1) & \dots & B_{yx}(m,n) \end{bmatrix}. \end{aligned} \quad (88)$$

Then we have

$$A = \mathcal{B}_{yx} \mathcal{B}_{xx}^{-1} \quad (89)$$

and for this matrix  $A$  the average

$$E[\|\mathbf{e}_k\|^2] = E[\|\mathbf{y}_k - \hat{\mathbf{y}}_k\|^2] = E[\|\mathbf{y}_k - A\mathbf{x}_k\|^2]$$

is minimal.

From Eq (82) we obtain

$$\begin{aligned} \mathbf{e}_k \mathbf{e}_k^T &= [I_m \quad -A] \begin{bmatrix} \mathbf{y}_k \mathbf{y}_k^T & \mathbf{y}_k \mathbf{x}_k^T \\ \mathbf{x}_k \mathbf{y}_k^T & \mathbf{x}_k \mathbf{x}_k^T \end{bmatrix} \begin{bmatrix} I_m \\ -A^T \end{bmatrix} \\ &= [\mathbf{y}_k \mathbf{y}_k^T - A\mathbf{x}_k \mathbf{y}_k^T \quad \mathbf{y}_k \mathbf{x}_k^T - A\mathbf{x}_k \mathbf{x}_k^T] \begin{bmatrix} I_m \\ -A^T \end{bmatrix}. \end{aligned}$$

It follows from Eqs (86) and (83) that

$$\begin{aligned} E[\mathbf{e}_k \mathbf{e}_k^T] &= E[\mathbf{y}_k \mathbf{y}_k^T - A E[\mathbf{x}_k \mathbf{y}_k^T] \\ &= \mathcal{B}_{yy} - A \mathcal{B}_{yx} = E[\mathbf{e}_k \mathbf{y}_k^T]. \end{aligned} \quad (90)$$

Now from Eq (73) the  $i^{\text{th}}$  diagonal element of the right hand side of this equation is the average of the square of the  $i^{\text{th}}$  component of the error. Hence we have

$$\text{trace}(\mathcal{B}_{yy} - A \mathcal{B}_{yx}) = E[\|\mathbf{e}_k\|^2]. \quad (91)$$

There is a difficulty with the method of least squares which we will now try to show.

For any  $n$ -vector  $\mathbf{a}$

$$(\mathbf{a}^T \mathbf{x}_k)^2 = \mathbf{a}^T \mathbf{x}_k \mathbf{x}_k^T \mathbf{a} \geq 0. \quad (92)$$

From this it is clear that the matrix

$$\mathcal{B}_{xx} = E[\mathbf{x}_k \mathbf{x}_k^T] \quad (93)$$

is symmetric and nonnegative. Hence there are eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$  and associated eigenvectors  $\mathbf{u}_1, \dots, \mathbf{u}_n$  satisfying the conditions

$$\begin{aligned} \mathcal{B}_{xx} \mathbf{u}_j &= \lambda_j \mathbf{u}_j \\ \mathbf{u}_i^T \mathbf{u}_j &= \delta_{ij} \\ &= 1 \text{ for } i = j \\ &= 0 \text{ otherwise.} \end{aligned} \quad (94)$$

Then we can write

$$\mathcal{B}_{xx} = \sum_{k=1}^n \lambda_k \mathbf{u}_k \mathbf{u}_k^T \quad (95)$$

and if  $\lambda_n > 0$  then

$$\mathcal{B}_{xx}^{-1} = \sum_{k=1}^n \frac{1}{\lambda_k} \mathbf{u}_k \mathbf{u}_k^T. \quad (96)$$

Consider the product  $\mathcal{B}_{xx}^{-1} \mathbf{a}$ . Since in practice the vector  $\mathbf{a}$  is experimentally determined there is error in  $\mathbf{a}$ . If now, for example,  $\lambda_n$  is much less than 1 then the error in the component of  $\mathbf{a}$  in the direction  $\mathbf{u}_n$  will be greatly magnified. A similar remark holds for all values of  $k$  for which  $\lambda_k$  is much less than 1.

SECTION V  
LATTICE FILTERS

In the previous section we discussed the method of least squares. In this section we will see the simplifications to this method when the data is stationary. Let  $x(t)$  denote a real *stationary random sequence* then the mean is  $m(t) = E[x(t)]$  and for a stationary process

$$m(t + \tau) = m(t) \quad \text{for all } \tau. \quad (97)$$

From this it follows that  $m(t)$  is a constant and it is sufficient to determine  $m(t)$  for  $t = 0$ . In the following we shall suppose for the sequence  $x(t)$  that the mean is zero for if not, we can consider instead the stationary random sequence

$$x(t) - E[x(t)].$$

Thus, here and in the following we have an *ensemble* or set of sequences and  $x(t)$  is a representative member of this set. The mean,  $m(t)$  is the average over the sequences in this set of sequences at a fixed value of  $t$ .

The *correlation function*  $B(t, s)$  is

$$B(t, s) = E[x(t)x(s)] = E[x(s)x(t)] = B(s, t) \quad (98)$$

and for a stationary process

$$B(t + \tau, s + \tau) = B(t, s) \quad \text{for all } \tau. \quad (99)$$

Taking, for example,  $\tau = -s$  we see that  $B(t, s)$  depends on the difference  $t - s$  alone. Hence for a stationary process we can write

$$B(t, s) = B(t - s, 0) = B(\tau) \quad \text{where } \tau = t - s. \quad (100)$$

The correlation function has the properties:

$$B(0) > 0 \quad (101a)$$

$$B(-\tau) = B(\tau) \quad (101b)$$

$$|B(\tau)| \leq B(0) \quad (101c)$$

Properties (a) and (b) follow immediately from Eqs (98-100) and (c) follows from

$$E[(x(t + \tau) \pm x(t))^2] = E[x^2(t + \tau)] \pm 2E[x(t + \tau)x(t)] + E[x^2(t)] \geq 0.$$

Let us note also that for any  $n, a_1, \dots, a_n, t_1, \dots, t_n$  that

$$\begin{aligned} & (a_1 x(t_1) + \dots + a_n x(t_n))^2 \\ &= [a_1 \quad \dots \quad a_n] \begin{bmatrix} x(t_1) \\ \vdots \\ x(t_n) \end{bmatrix} [x(t_1) \quad \dots \quad x(t_n)] \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}. \end{aligned}$$

This equation can be written as

$$\begin{aligned} & (a_1 x(t_1) + \dots + a_n x(t_n))^2 \\ &= [a_1 \quad \dots \quad a_n] \begin{bmatrix} x(t_1)x(t_1) & \dots & x(t_1)x(t_n) \\ \vdots & \ddots & \vdots \\ x(t_n)x(t_1) & \dots & x(t_n)x(t_n) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}. \end{aligned}$$

Taking the expected value of both sides we obtain

$$[a_1 \quad \dots \quad a_n] \begin{bmatrix} B(0) & \dots & B(t_1 - t_n) \\ \vdots & \ddots & \vdots \\ B(t_n - t_1) & \dots & B(0) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \geq 0. \quad (102)$$

We shall refer to the matrix of this quadratic form as the correlation matrix. From the properties of the correlation function  $B(t)$  it is clear that the correlation matrix for a stationary process is symmetric. Usually  $t_1, \dots, t_n$  are successive values, for example,  $t + 1, \dots, t + n$ , and the correlation matrix becomes

$$\begin{bmatrix} B(0) & \dots & B(1 - n) \\ \vdots & \ddots & \vdots \\ B(n - 1) & \dots & B(0) \end{bmatrix}. \quad (103)$$

Note that the element in the  $j^{\text{th}}$  row and  $k^{\text{th}}$  column is

$$B(t + j, t + k) = B(j - k)$$

and the element in the  $(j + 1)^{\text{th}}$  row and the  $(k + 1)^{\text{th}}$  column is

$$B(t + j + 1, t + k + 1) = B(j - k).$$

From these observations it follows that the correlation matrix determined from a stationary random sequence is a symmetric, nonnegative Toeplitz matrix. In this case the correlation matrix is known if one knows the elements in the first row or column.

Consider a *scalar stationary random sequence*  $x(t)$  with zero mean. Set

$$x_{(t \pm j)} = x(t \pm j).$$

Thus  $\dots x_{(t-n-1)}, x_{(t-n)}, x_{(t-n+1)}, \dots, x_{(t-1)}, x_{(t)}, x_{(t+1)}, \dots$  is a set of scalar values. A linear forward predictor of order  $n$ , LFP( $n$ ), is defined by the equation

$$\hat{x}_t = a_{n1}x_{(t-1)} + a_{n2}x_{(t-2)} + \dots + a_{nn}x_{(t-n)}. \quad (104)$$

Here  $\hat{x}_t$  denotes the predicted value of  $x_t$  based on the  $n$  values  $x_{t-1}, \dots, x_{t-n}$ . Let  $e_{nt}$  stand for the prediction error of order  $n$ .

$$e_{nt} = x_t - \hat{x}_t = x_t - \sum_{j=1}^n a_{nj}x_{t-j}. \quad (105)$$

The prediction coefficients  $a_{n1}, \dots, a_{nn}$  are determined so as to minimize the prediction error in the least squares sense.

To relate this to our previous discussion of the method of least squares set

$$\mathbf{x} = [x_{t-1} \quad \dots \quad x_{t-n}]^T.$$

This  $n$ -dimensional vector  $\mathbf{x}$ , which is a representative of many such  $n$ -vectors, plays the role of the  $n$ -vectors  $\mathbf{x}_k$ . The representative value  $x_t$  plays the role of the  $m$ -vectors  $\mathbf{y}_k$  and the  $n$ -vector

$$\mathbf{a}_n = [a_{n1} \quad \dots \quad a_{nn}]^T$$

the role of the matrix  $A$ . Having noted this connection to our previous discussions we write Eq (105) as

$$e_{nt} = x_t - \hat{x}_t = [1 \quad -a_{n1} \quad \dots \quad -a_{nn}] \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-n} \end{bmatrix} \quad (106)$$

$$= [1 \quad -\mathbf{a}_n^T] \begin{bmatrix} x_t \\ \mathbf{x} \end{bmatrix}. \quad (107)$$

Post multiply Eq (107) with the row vector  $[x_t \dots x_t \ n] = [x_t \ \mathbf{x}^T]$  and obtain, as in the previous section,

$$[e_{nt}x_t \ e_{nt}\mathbf{x}^T] = [1 \ -\mathbf{a}_n^T] \begin{bmatrix} x_t x_t & x_t \mathbf{x}^T \\ \mathbf{x} x_t & \mathbf{x} \mathbf{x}^T \end{bmatrix}. \quad (108)$$

Taking the expectation we know from our discussion in SECTION IV that we can obtain the two equations

$$E[x_t x_t] - \mathbf{a}_n^T E[\mathbf{x} x_t] = E[e_{nt} x_t] \quad (109)$$

$$E[x_t \mathbf{x}^T] - \mathbf{a}_n^T E[\mathbf{x} \mathbf{x}^T] = E[e_{nt} \mathbf{x}^T] = 0. \quad (110)$$

Eq (110) determines the vector  $\mathbf{a}_n$  and once this vector is known Eq (109) determines the value  $E[e_{nt} x_t]$ .

From Eq (107) we obtain

$$e_{nt} e_{nt} = [1 \ -\mathbf{a}_n^T] \begin{bmatrix} x_t x_t & x_t \mathbf{x}^T \\ \mathbf{x} x_t & \mathbf{x} \mathbf{x}^T \end{bmatrix} \begin{bmatrix} 1 \\ -\mathbf{a}_n \end{bmatrix}. \quad (111)$$

Taking the expectation and performing the indicated multiplication on the left we have

$$E[e_{nt}^2] = [E[x_t^2] - \mathbf{a}_n^T E[\mathbf{x} x_t] \quad E[x_t \mathbf{x}^T] - \mathbf{a}_n^T E[\mathbf{x} \mathbf{x}^T]] \begin{bmatrix} 1 \\ -\mathbf{a}_n \end{bmatrix}.$$

It follows from Eqs (110) and (109) that this equation becomes

$$E[e_{nt}^2] = E[x_t^2] - \mathbf{a}_n^T E[\mathbf{x} x_t] = E[e_{nt} x_t]. \quad (112)$$

Set

$$r_n^2 = E[e_{nt}^2]. \quad (113)$$

Finally, then, taking the expectation of Eq (108) we can write the two equations, Eqs (109) and (110) as a single matrix equation

$$[r_n^2 \ 0 \ \dots \ 0] = [1 \ -\mathbf{a}_n^T] \begin{bmatrix} B(0) & B(1) & \dots & B(n) \\ B(-1) & B(0) & \dots & B(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ B(-n) & B(1-n) & \dots & B(0) \end{bmatrix}. \quad (114)$$

Let us denote the correlation matrix in Eq (114) by  $\mathbf{B}_{n+1}$ . We observe that except for a normalization factor the row vector  $[1 \ -a_{n1} \ \dots \ -a_{nn}]$  is the first row of the inverse of  $\mathbf{B}_{n+1}$ . Also from examination of Eq (114) we see that

$$\begin{bmatrix} 1 & -a_{n1} & \dots & -a_{nn} \\ -a_{nn} & \dots & -a_{n1} & 1 \end{bmatrix} \mathbf{B}_{n+1} = \begin{bmatrix} r_n^2 & 0 & \dots & 0 \\ 0 & \dots & 0 & r_n^2 \end{bmatrix}. \quad (115)$$

Our object is to show that if we have the prediction coefficients  $a_{n1}, \dots, a_{nn}$  of order  $n$  then the prediction coefficients of order  $n+1$  are readily obtained.

We suppose the correlation matrix  $\mathbf{B}_n$  is known for every value  $n$  of interest. We suppose the prediction coefficients  $a_{n1}, \dots, a_{nn}$  are also known. Then from Eq (115) the value  $r_n^2$  is known and we can compute

$$\begin{bmatrix} 1 & -a_{n1} & \dots & -a_{nn} & 0 \\ 0 & -a_{nn} & \dots & -a_{n1} & 1 \end{bmatrix} \mathbf{B}_{n+2} = \begin{bmatrix} r_n^2 & 0 & \dots & 0 & q_n \\ q_n & 0 & \dots & 0 & r_n^2 \end{bmatrix}. \quad (116)$$

The right side of Eq (116) is not quite of the right form. There is a value  $q_n$  which would be zero if we had the right prediction coefficients of order  $n+1$ . This value of  $q_n$  is given by

$$q_n = B(n+1) - \sum_{k=1}^n a_{nk} B(n+1-k) \quad (117)$$

It is clear, however, that if we multiply Eq (116) by the matrix

$$K = \begin{bmatrix} 1 & -q_n/r_n^2 \\ -q_n/r_n^2 & 1 \end{bmatrix} \quad (118)$$

the right hand side takes on the right form. The prediction coefficients of order  $n+1$  are given by

$$\begin{aligned} a_{n+1,j} &= a_{nj} - (q_n/r_n^2) a_{n,n+1-j} \text{ for } 1 \leq j \leq n \\ a_{n+1,n+1} &= q_n/r_n^2. \end{aligned} \quad (119)$$

One has also

$$r_{n+1}^2 = r_n^2 - q_n^2/r_n^2. \quad (120)$$

Starting values are readily obtained. We have

$$\begin{bmatrix} 1 & -a_{11} \\ -a_{11} & 1 \end{bmatrix} \begin{bmatrix} B(0) & B(1) \\ B(-1) & B(0) \end{bmatrix} = \begin{bmatrix} r_1^2 & 0 \\ 0 & r_1^2 \end{bmatrix}. \quad (121)$$

From this equation we obtain

$$-a_{11}B(0) + B(1) = 0$$

or

$$a_{11} = \frac{B(1)}{B(0)} \quad (122)$$

and

$$\begin{aligned} r_1^2 &= B(0) - a_{11}B(1) \\ &= \frac{B^2(0) - B^2(1)}{B(0)} \geq 0. \end{aligned} \quad (123)$$

Consider now a stationary random sequence  $\mathbf{x}(t)$  of  $m$ -vectors. A linear forward predictor of order  $p$  is of the form

$$\begin{aligned} \hat{\mathbf{x}}(t) &= A_{p1}\mathbf{x}(t-1) + \dots + A_{pp}\mathbf{x}(t-p) \\ &= [A_{p1} \ \dots \ A_{pp}] \begin{bmatrix} \mathbf{x}(t-1) \\ \vdots \\ \mathbf{x}(t-p) \end{bmatrix}. \end{aligned} \quad (124)$$

Here the lengthy  $mp$ -vector  $[\mathbf{x}^T(t-1) \ \dots \ \mathbf{x}^T(t-p)]^T$  plays the role of the  $n$ -vector  $\mathbf{x}_k$  of Section IV. The  $(m \times mp)$  matrix  $[A_{p1} \ \dots \ A_{pp}]$  plays the role of the matrix  $A$  and  $\mathbf{x}(t)$  the role of the vector  $\mathbf{y}_k$ . Again, for the stationary case the mean

$$E[\mathbf{x}(t)] = E[\mathbf{x}(t+\tau)] \text{ for all } \tau.$$

Hence as for the scalar case we may suppose  $\mathbf{x}(t)$  has mean zero.

The error

$$\begin{aligned} \mathbf{e}_p(t) &= \mathbf{x}(t) - \hat{\mathbf{x}}(t) \\ &= [I_m \ -A_{p1} \ \dots \ A_{pp}] \begin{bmatrix} \mathbf{x}(t) \\ \vdots \\ \mathbf{x}(t-p) \end{bmatrix}. \end{aligned}$$

Post multiplying this equation by the vector  $[\mathbf{x}^T(t) \ \dots \ \mathbf{x}^T(t-p)]$  obtain

$$[\mathbf{e}_p(t)\mathbf{x}^T(t) \ \mathbf{e}_p(t)\mathbf{x}^T(t-1) \ \dots \ \mathbf{e}_p(t)\mathbf{x}^T(t-p)]$$

for the left side and the matrix

$$\begin{bmatrix} \mathbf{x}(t)\mathbf{x}^T(t) & \mathbf{x}(t)\mathbf{x}^T(t-1) & \dots & \mathbf{x}(t)\mathbf{x}^T(t-p) \\ \mathbf{x}(t-1)\mathbf{x}^T(t) & \mathbf{x}(t-1)\mathbf{x}^T(t-1) & \dots & \mathbf{x}(t-1)\mathbf{x}^T(t-p) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}(t-p)\mathbf{x}^T(t) & \mathbf{x}(t-p)\mathbf{x}^T(t-1) & \dots & \mathbf{x}(t-p)\mathbf{x}^T(t-p) \end{bmatrix}$$

multiplies the matrix  $[I_m \quad -A_{p1} \quad \dots \quad -A_{pp}]$  on the right. Set

$$R_p(t) = E[\mathbf{x}(t)\mathbf{x}^T(r)]. \quad (125)$$

Taking the expectation of the equation just described we obtain

$$\begin{bmatrix} R_p(t) & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} I_m & -A_{p1} & \dots & -A_{pp} \end{bmatrix} \begin{bmatrix} B_m(0) & B_m(1) & \dots & B_m(p) \\ B_m(-1) & B_m(0) & \dots & B_m(p-1) \\ \vdots & \vdots & \ddots & \vdots \\ B_m(-p) & B_m(1-p) & \dots & B_m(0) \end{bmatrix}. \quad (126)$$

The subscript,  $m$ , is to indicate that the entries  $B_m(j)$  are matrices. This matrix in this block form is a block symmetric Toeplitz matrix and a procedure like that described by Eqs (115-118) can be used to determine the prediction coefficients  $A_{p1}, \dots, A_{pp}$

Starting values are readily obtained. We have

$$A_{11} = B_m(1)B_m^{-1}(0) \quad (127)$$

and

$$R_1(t) = B_m(0) - A_{11}B_m(-1). \quad (128)$$

Set  $\mathbf{B}_{m,p+1}$  equal to the matrix with block elements  $B_m(j)$ . Eq (126). Then as in Eq (116)

$$\begin{bmatrix} I_m & -A_{p1} & \dots & -A_{pp} & 0 \\ 0 & -A_{pp} & \dots & -A_{p1} & I_m \end{bmatrix} \mathbf{B}_{m,p+2} = \begin{bmatrix} R_p(t) & 0 & \dots & 0 & Q_p \\ Q_p & 0 & \dots & 0 & R_p(t) \end{bmatrix} \quad (129)$$

From this equation we have

$$Q_p = B_m(p+1) - \sum_{k=1}^p A_{pk} B_m(p+1-k). \quad (130)$$

Multiply Eq (129) on the left by

$$\mathbf{K}_p = \begin{bmatrix} I_m & -Q_p R_p^{-1}(t) \\ -Q_p R_p^{-1}(t) & I_m \end{bmatrix}. \quad (131)$$

From this we obtain

$$A_{p \cdot 1 j} = A_{pj} + Q_p R_p^{-1}(t) A_{pp \cdot 1 j} \quad (132)$$

for  $1 \leq j \leq p$  and

$$A_{p+1,p+1} = Q_p R_p^{-1}(t). \quad (133)$$

In practice, since  $R_p(t)$  and  $Q_p$  are known one first solves

$$A_{p+1,p+1} R_p(t) = Q_p \quad (134)$$

for  $A_{p+1,p+1}$ . One then uses this matrix in Eq (132) to compute the matrices  $A_{p+1,j}$

Let us review first the scalar case. The attractive features of the *stationary process* are that one does not have to choose in advance the order of the system for determining the predictor coefficients. It is not necessary to compute, store and solve a system of equations involving the matrix  $\mathbf{B}_n$ . It is sufficient to know the values  $B(0), \dots, B(n)$  of the correlation function for a sufficiently large value of  $n$ . As the prediction coefficients  $a_{n1}, \dots, a_{nn}$  are computed for  $n = 1, 2, \dots$  one is able to monitor the results by means of the value  $r_n^2$ . The value  $r_n^2$  is the average of the square of the error. This value can be used as a criteria for determining the order of the linear forward predictor. If for some value of  $n$  one obtains  $r_n^2 < 0$  we infer that the experimental data does not support the determination of prediction coefficients for a value of  $n$  this large.

Corresponding remarks hold for the vector case. Here again we do not need the matrix  $\mathbf{B}_{m,p}$  but its entries  $B_m(0), \dots, B_m(p)$  for  $p$  sufficiently large. The process of determining the prediction coefficient matrices  $A_{p1}, \dots, A_{pp}$  for  $p = 1, 2, \dots$  is monitored by means of the diagonal entries of the matrix  $R_p(t)$ .

## SECTION VI

### CANONICAL VARIATE ANALYSIS

Here we review a method, *canonical variate analysis*, for determining a one step ahead linear forward predictor which is optimal with respect to a prescribed norm. As above, SECTION IV, we are concerned with a set of  $n$ -vectors  $\mathbf{x}_1, \dots, \mathbf{x}_N$  and a set of  $m$ -vectors,  $\mathbf{y}_1, \dots, \mathbf{y}_N$ . The method does not require that the  $\mathbf{x}_k$  and  $\mathbf{y}_k$  be stationary and jointly stationary as in the previous section. However, stationarity, if present, will provide some simplifications.

Let  $G$  denote a symmetric positive definite matrix of order  $m$ . Then

$$\begin{aligned}\|\mathbf{y}_k\|_G^2 &= \mathbf{y}_k^T G \mathbf{y}_k \\ &= \mathbf{y}_k^T G^{1/2} G^{1/2} \mathbf{y}_k \\ &= \|G^{1/2} \mathbf{y}_k\|^2.\end{aligned}\tag{135}$$

Observe for any orthogonal matrix  $U$  of order  $m$  and for

$$\mathbf{z}_k = U^T G^{1/2} \mathbf{y}_k\tag{136}$$

that

$$\|\mathbf{z}_k\|^2 = \|\mathbf{y}_k\|_G^2.\tag{137}$$

Let us suppose that the matrix

$$\mathcal{B}_{xx} = E[\mathbf{x}_k \mathbf{x}_k^T]\tag{138}$$

is positive definite. Then  $\mathcal{B}_{xx}$  is nonsingular. Here we want to determine a matrix  $J$  of order  $n$  which transforms the vectors  $\mathbf{x}_k$  into vectors  $\mathbf{w}_k$ ,

$$\mathbf{w}_k = J \mathbf{x}_k.\tag{139}$$

for which the covariance

$$C(\mathbf{w}_k, \mathbf{w}_k) = \mathcal{B}_{ww} = I_n.\tag{140}$$

Now

$$\begin{aligned}
 C(\mathbf{w}_k, \mathbf{w}_k) &= E[\mathbf{w}_k \mathbf{w}_k^T] \\
 &= E[J \mathbf{x}_k \mathbf{x}_k^T J^T] \\
 &= J \mathcal{B}_{xx} J^T.
 \end{aligned} \tag{141}$$

Since the matrix  $\mathcal{B}_{xx}$  is symmetric and positive definite our objective, Eq (140), will be realized, if for any orthogonal matrix  $V$  we take

$$J = V^T \mathcal{B}_{xx}^{-1/2}. \tag{142}$$

The covariance

$$\begin{aligned}
 C(\mathbf{w}_k, \mathbf{z}_k) &= E[\mathbf{w}_k \mathbf{z}_k^T] = \mathcal{B}_{wz} \\
 &= E[V^T \mathcal{B}_{xx}^{-1/2} \mathbf{x}_k \mathbf{y}_k^T G^{1/2} U] \\
 &= V^T \mathcal{B}_{xx}^{-1/2} \mathcal{B}_{xy} G^{1/2} U.
 \end{aligned} \tag{143}$$

If, now,  $U$  and  $V$  are the orthogonal matrices in the *singular value decomposition* of the matrix  $\mathcal{B}_{xx}^{-1/2} \mathcal{B}_{xy} G^{1/2}$  then

$$V^T \mathcal{B}_{xx}^{-1/2} \mathcal{B}_{xy} G^{1/2} U = S \tag{144}$$

where  $S$  is an  $(n \times m)$  matrix. All the entries in  $S$  are zeros except for the diagonal elements  $s_{jj}$  for  $1 \leq j \leq r$ , where  $r$  is the rank of the matrix  $\mathcal{B}_{xy}$ , and  $r \leq l = \min[m, n]$ . These diagonal elements satisfy the inequalities  $s_{11} \geq s_{22} \geq \dots \geq s_{rr} > 0$ .

We now want the matrix  $A$  for which

$$\begin{aligned}
 \hat{\mathbf{z}}_k &= A \mathbf{w}_k \\
 \mathbf{e}_k &= \mathbf{z}_k - \hat{\mathbf{z}}_k [I_m \quad -A] \begin{bmatrix} \mathbf{z}_k \\ \mathbf{w}_k \end{bmatrix}
 \end{aligned} \tag{145}$$

and  $E[\|\mathbf{e}_k\|^2]$  is least. As we know, from what has been said above, taking the average of

$$[\mathbf{e}_k \mathbf{z}_k^T \quad \mathbf{e}_k \mathbf{w}_k^T] = [I_m \quad -A] \begin{bmatrix} \mathbf{z}_k \mathbf{z}_k^T & \mathbf{z}_k \mathbf{w}_k^T \\ \mathbf{w}_k \mathbf{z}_k^T & \mathbf{w}_k \mathbf{w}_k^T \end{bmatrix}$$

obtain the system of equations

$$\begin{aligned}
 \mathcal{B}_{zz} - A \mathcal{B}_{wz} &= E[\mathbf{e}_k \mathbf{z}_k^T] \\
 \mathcal{B}_{zw} - A \mathcal{B}_{ww} &= E[\mathbf{e}_k \mathbf{w}_k^T]
 \end{aligned}$$

For the desired matrix  $A$ ,  $E[\mathbf{e}_k \mathbf{w}_k^T] = 0$  and

$$A \mathbb{B}_{wv} = \mathbb{B}_{zw} \quad (146)$$

and since  $\mathbb{B}_{vv} = I_n$

$$A = \mathbb{B}_{zw} = S^T \quad (147)$$

by Eq (144). Thus the predicted value  $\hat{\mathbf{z}}_k$  is given by

$$\begin{aligned} \hat{\mathbf{z}}_k &= S^T \mathbf{w}_k \\ &= \sum_{j=1}^r s_{jj} w_{jk} \end{aligned} \quad (148)$$

and one solves the equation

$$U^T G^{1/2} \hat{\mathbf{y}}_k = \hat{\mathbf{z}}_k \quad (149)$$

for the predicted value  $\hat{\mathbf{y}}_k$ .

Also, as we have already seen,

$$\mathbf{e}_k \mathbf{e}_k^T = \begin{bmatrix} I_m & -A \end{bmatrix} \begin{bmatrix} \mathbf{z}_k \mathbf{z}_k^T & \mathbf{z}_k \mathbf{w}_k^T \\ \mathbf{w}_k \mathbf{z}_k^T & \mathbf{w}_k \mathbf{w}_k^T \end{bmatrix} \begin{bmatrix} I_m \\ -A^T \end{bmatrix}.$$

From this equation we have

$$\begin{aligned} E[\mathbf{e}_k \mathbf{e}_k^T] &= \begin{bmatrix} \mathbb{B}_{zz} - A \mathbb{B}_{wz} & \mathbb{B} - A \mathbb{B}_{ww} \end{bmatrix} \begin{bmatrix} I_m \\ -A^T \end{bmatrix} \\ &= \mathbb{B}_{zz} - A \mathbb{B}_{wz} \end{aligned} \quad (150)$$

because of Eq (146). Now

$$A \mathbb{B}_{wz} = \mathbb{B}_{zw} \mathbb{B}_{zw} = S S^T \quad (151)$$

and then

$$\begin{aligned} \text{Trace } E[\mathbf{e}_k \mathbf{e}_k^T] &= E[\|\mathbf{e}_k\|^2] \\ &= \text{Trace } \mathbb{B}_{zz} - \text{Trace } S S^T \\ &= \text{Trace } \mathbb{B}_{zz} - \sum_{j=1}^r s_{jj}^2 \end{aligned} \quad (152)$$

We want to examine the observation, Eq (13), in greater detail by considering the scalar case. Application of the method to the scalar case is relatively straight forward.

Thus, for a random sequence with zero mean and for a provisional value of  $n$  we want to determine values  $a_{n1}, \dots, a_{nn}$  for which the predicted value

$$\hat{x}_t = a_{n1}x_{(t-1)} + a_{n2}x_{(t-2)} + \dots + a_{nn}x_{(t-n)}$$

is best.

For a value  $n$

$$\mathcal{B}_{xx} = E \begin{bmatrix} x_{(t-1)}x_{(t-1)} & \dots & x_{(t-1)}x_{(t-n)} \\ \vdots & \ddots & \vdots \\ x_{(t-n)}x_{(t-1)} & \dots & x_{(t-n)}x_{(t-n)} \end{bmatrix}$$

and we can write

$$\mathcal{B}_{xx} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \dots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T$$

where  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ . Here we want a rather large value of  $n$  for which the smallest eigenvalue  $\lambda_n > 0$  and not too small. For such a value of  $n$  the matrix  $\mathcal{B}_{xx}$  is nonsingular of order  $n$ . The matrix

$$\mathcal{B}_{xy} = E \begin{bmatrix} x_{(t-1)}x_t \\ \vdots \\ x_{(t-n)}x_t \end{bmatrix}$$

in this case is an  $(n \times 1)$  matrix or an  $n$ -vector. For the scalar case the matrix  $G$  is a positive constant which for our purposes here we take as 1.

We readily see that the singular value decomposition of the matrix  $\mathcal{B}_{xx}^{-1/2} \mathcal{B}_{xy}$  is

$$V^T \mathcal{B}_{xx}^{-1/2} \mathcal{B}_{xy} U = [s_{11} \quad 0 \quad \dots \quad 0]^T$$

where  $U = 1$  and hence the  $A$  matrix is the row vector

$$A = [s_{11} \quad 0 \quad \dots \quad 0]. \quad (153)$$

Here one must keep in mind that in the present case  $x_t$  plays the role of the vectors  $\mathbf{y}_k = \mathbf{z}_k$  and  $\mathbf{x}_k = [x_{(t-1)} \quad \dots \quad x_{(t-n)}]^T$ . Doing so, it is clear then that

$$\hat{x}_t = A V^T \mathcal{B}_{xx}^{-1/2} \begin{bmatrix} x_{(t-1)} \\ \vdots \\ x_{(t-n)} \end{bmatrix}. \quad (154)$$

From this equation one could obtain the coefficients  $a_{n1}, \dots, a_{nn}$ . On the other hand, if we set

$$\mathbf{w}_k = V^T B_{xx}^{-1/2} \begin{bmatrix} x(t-1) \\ \vdots \\ x(t-n) \end{bmatrix} \quad (155)$$

then

$$\hat{x}_t = s_{11} u_{1k} \quad (156)$$

that is, the product of  $s_{11}$  and the first component of  $\mathbf{w}_k$ .

In review and conclusion then the main feature of this report is the application of Lanczos' Progressive algorithm in the method proposed for the determination of complex exponentials. From our limited numerical experimentation with the method we believe the method merits further investigation and testing.

The method of lattice filters for prediction in stationary sequences is attractive because of the simplicity of the computations and the ability to monitor how good the predicted value is by means of the quantity  $r_n^2$  or  $R_p(t)$ . The matrix square roots and the singular value decomposition in the canonical variate analysis method indicates that the computations in this method are more extensive.

## REFERENCES

1. Lanczos, C. "Applied Analysis", Prentice Hall Inc. 1956, Englewood Cliffs, NJ.
2. Lanczos, C., "An Iterative Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators" Research Paper 2133, Journal of Research of the National Bureau of Standards, Vol. 45, No.4, Oct. 1950.
3. Friedlander, B. "Lattice Filters for Adaptive Processing", Proceedings of the IEEE, vol. 70, No. 8, Aug.1982.
4. Larimore, W. E., "System Identification, Reduced-order Filtering and Modeling via Canonical Variate Analysis", Appendix B, Proceedings of the 1983 American Control Conference, June 22-24, San Francisco, CA.
5. Rao, C. R. "Linear Statistical Inference and Its Applications", Wiley, 1973, New York, NY.