

QUADRATIC FORMS, MATRIX EQUATIONS AND THE
MATRIX EIGENVALUE PROBLEM

A Thesis submitted

by

Biswanath Datta M.Sc.(Cal.) M.Sc.(McM.)

to

The School of Graduate Studies

of the University of Ottawa

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the

Subject of Mathematics

University of Ottawa

March, 1972

To my Parents

ACKNOWLEDGEMENTS

This thesis has been written under the supervision of Dr. J. L. Howland, professor, Department of Mathematics.

The author is deeply grateful to Professor Howland for introducing him to this field of research by offering two supervised reading courses on the subject, which have been valuable assets for the whole work. It was Professor Howland, who also suggested the topic of the thesis and whose unfailing guidance and valuable criticisms in various phases of the work made the work successful.

The author also expresses his gratitude to Dr. V. Linis, who as a chairman of the Department of Mathematics gave him an opportunity to study in his Department and to the Government of the province of Ontario for providing an adequate financial assistance during his period of study.

The author's sincere thanks are also due to Miss C. Royer who typed the manuscript with great patience.

ABSTRACT

This thesis includes the results of investigations into two aspects of the General Matrix Eigenvalue Problem.

In the first chapter information on the Jordan structure of a given matrix is derived from two sources - a Hankel matrix of Newton sums and the solutions of a nonlinear matrix equation. This nonlinear equation is solved theoretically and a few important, well-known results of Matrix Theory are derived as consequences of more general results obtained in this connection. In addition, numerical methods of obtaining the required information from this matrix equation are proposed.

In the second chapter, matrix equations whose solutions separate the eigenvalues of a given matrix are discussed and a numerical method for solving the matrix equations associated with the classical stability problems is proposed. Computationally, the proposed method is simple in the sense that it only requires the solution of an $n \times n$ linear algebraic system and does not involve the difficulties of series solution as do most of the methods available in the literature for the solution of the Lyapunov equation.

In the third chapter a new method for solving the classical stability problems is introduced. This method eliminates the necessity for solving the matrix equations and has decided computational advantages over existing procedures such as using Bezoutiants, bigradients and Schur-Cohn determinants.

The thesis concludes with a result which can be considered as an alternative to the Schur-Cohn Theorem. A polynomial matrix is constructed, whose leading principal minors are the same as the

Schur-Cohn determinants. Similar results were obtained by P. C. Parks and Barnett in case of the Routh-Hurwitz problem.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
ABSTRACT.....	iii
INTRODUCTION.....	v
CHAPTER I QUADRATIC FORMS, MATRIX EQUATIONS AND THE JORDAN STRUCTURE OF A MATRIX.....	1
1.1 SOME BASIC CONCEPTS.....	3
1.2 NONDEROGATORY MATRICES AND JORDAN STRUCTURES.....	13
1.3 QUADRATIC FORMS AND JORDAN STRUCTURE.....	13
1.3.1 JORDAN STRUCTURE AND THE ASSOCIATED HANKEL MATRIX.....	13
1.3.2 HANKEL FORM AND SIMILARITY TO A NORMAL MATRIX.....	17
1.4 MATRIX EQUATIONS AND JORDAN STRUCTURE.....	18
1.4.1 FORMULATION OF A NON-LINEAR MATRIX EQUATION.....	18
1.4.2 THE MATRIX EQUATION $SA = A*S$	20
1.4.3 THE MATRIX EQUATION $SP(A) = A*S$	34
CHAPTER II QUADRATIC FORMS, MATRIX EQUATIONS AND THE LOCATION OF MATRIX EIGENVALUES.....	48
2.1 THE EIGENVALUES IN A HALF PLANE	50
2.1.1 THE HERMITE PROBLEM.....	50
2.1.2 THE ROUTH-HURWITZ PROBLEM.....	59
2.2 THE EIGENVALUES IN THE UNIT CIRCLE.....	72
2.2.1 THE SCHUR-COHN PROBLEM.....	72
2.3 THE EIGENVALUES IN A SECTOR.....	77
CHAPTER III A NEW APPROACH TO THE SOLUTION OF THE CLASSICAL STABILITY PROBLEMS.....	80
3.1 SOLUTION OF THE HERMITE-PROBLEM.....	80

3.2 SOLUTION OF THE ROUTH-HURWITZ PROBLEM.....	87
3.3 SOLUTION OF THE SCHUR-COHN PROBLEM.....	91
APPENDIX AN ALTERNATIVE THEOREM.....	96
CONCLUSIONS.....	101
BIBLIOGRAPHY.....	104
SUPPLEMENTARY BIBLIOGRAPHY.....	109

INTRODUCTION

An arbitrary matrix A is completely determined by its Jordan canonical form and the matrix transforming it to that form. Thus, problems in numerical linear algebra are most logically attacked by transforming the matrices involved to their canonical forms, and thus simplifying the problems to an extent which in many cases, renders them almost trivial.

Unfortunately, it is not easy, in general, to find the Jordan form of an arbitrary matrix, and usually less so to find the appropriate transformation converting it to this form. Moreover, even when it is reasonably feasible to find the Jordan form, as a result, perhaps, of special properties or prior knowledge of the matrices involved, the computations required are lengthy, expensive, and, of course, a source of arithmetic truncation error.

At the same time, many problems of numerical linear algebra do not require, for their solution, all the information which is supplied by the Jordan form. In particular, eigenvalue separation problems, such as those arising in connection with the study of the stability of systems of differential or difference equations, do not require knowledge of the actual numerical values of the eigenvalues of a matrix, but simply an accounting of the numbers of them which lie in specified regions of the complex plane, or even the information that none of them can lie in a specified region. For such problems, it seems reasonable to seek algorithms which are simpler, more widely applicable and more efficient than methods based upon the determination of the Jordan form - i.e.

upon the numerical evaluation of the eigenvalues in question.

Matrix equations are relevant to the problem of separating matrix eigenvalues by methods which, when contrasted with available matrix eigenvalue methods, have advantages of efficiency, reliability and simplicity in return for the smaller amount of information provided. Matrix equations are also relevant to the basic problem of determining the Jordan form of a given matrix.

This thesis is devoted to a study of the application of matrix equations to the numerical solution of certain aspects of the Jordan canonical form problem, and to the numerical solution of various eigenvalue separation problems.

CHAPTER I

QUADRATIC FORMS, MATRIX EQUATIONS AND THE JORDAN STRUCTURE OF A MATRIX

The problem of knowing the Jordan structure of an arbitrary matrix A is the problem of determining the structures of the individual Jordan submatrices appearing in a Jordan canonical representation of A , that is, it is the problem of counting the number of distinct eigenvalues, the number of different pairs of complex eigenvalues, the number of distinct real eigenvalues, etc. with the respective multiplicities in each case. This problem is, therefore, the first aspect of the problem of separating the eigenvalues of A , the second being that of locating the eigenvalues in a specified region. Both aspects can be handled efficiently, once the characteristic or minimal polynomial of A is known, because, the problem then becomes just that of separating the zeros of the characteristic polynomial of A and it is well-known that the method of quadratic forms is an efficient tool for the latter.

Again, by the introduction of matrix equations to define the relevant quadratic forms, the problem of determining the minimal polynomial can be replaced with that of solving matrix equations and recently a theory has been developed by Howland [23] showing how the technique of matrix equations can be applied successfully to solve the second aspect of the problem. The question, therefore, naturally arises if these two apparently different methods, the method of quadratic forms and the technique of matrix equations can

be applied to solve the first aspect of the problem as well. An attempt is made in this chapter to answer this question.

It is shown that the Hankel forms defined by a Hankel matrix of Newton sums may be employed to obtain some information on the Jordan structure of A . This Hankel matrix can be obtained directly from A by computing traces of several powers of A . Regarding the applications of matrix equations to solve the problem, it is shown that a linear matrix equation whose non singular solutions transform the given matrix A into its transpose (conjugate transpose in the complex case) may be used to obtain information relevant to this problem. Methods of computing the solutions of this matrix equation numerically, to obtain the required information have been proposed in some cases. Again, the problem as to when or whether a given matrix A is similar to a normal matrix has been formulated in terms of a non-linear matrix equation which is again satisfied by the solutions of the linear matrix equation mentioned above. This non-linear matrix equation is solved theoretically and a few results on its numerical solutions are obtained. Thus, it is shown that the problem of computing a solution of some specified nature of this equation is equivalent to the problem of computing a suitable polynomial $P(A)$ in A . However, these polynomials, except in some special cases are not constructed.

1.1

SOME BASIC CONCEPTS

A brief account of the classical theory of canonical forms and of several other topics needed for later developments will be introduced in this section.

(i) Elementary Divisors, Invariant Polynomials and Similarity

Let $A = (a_{ij})$ be a given matrix of order n , then the matrix

$$A - \lambda I = \begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{pmatrix}$$

is called the Characteristic Matrix of A and $\det(A - \lambda I) = 0$ is called the Characteristic Equation of A . Let D_j be the greatest common divisor of all the minors of order j in $A - \lambda I$, $j = 1, 2, \dots, n$. Then the polynomials

$$i_1(\lambda) = \frac{D_n(\lambda)}{D_{n-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)}, \quad \dots \quad i_n(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)}$$

$$(D_0(\lambda) = 1)$$

are called the Invariant Polynomials of A . Let the invariant polynomials $i_1(\lambda), i_2(\lambda) \dots i_n(\lambda)$ be decomposed into irreducible factors as follows:

$$\begin{aligned} i_1(\lambda) &= [\phi_1(\lambda)]^{c_1} [\phi_2(\lambda)]^{c_2} \dots [\phi_s(\lambda)]^{c_s}, \\ i_2(\lambda) &= [\phi_1(\lambda)]^{d_1} [\phi_2(\lambda)]^{d_2} \dots [\phi_s(\lambda)]^{d_s}, \\ &\vdots \\ i_n(\lambda) &= [\phi_1(\lambda)]^{l_1} [\phi_2(\lambda)]^{l_2} \dots [\phi_s(\lambda)]^{l_s} \end{aligned}$$

$$(c_k \geq d_k \geq \dots \geq \ell_k \geq 0, \quad k = 1, 2, \dots, s)$$

Then the polynomials among $[\phi_1(\lambda)]^{c_1} \dots [\phi_s(\lambda)]^{\ell_s}$ which are not equal to 1, are called the Elementary Divisors of A.

A matrix A of order n is said to be Similar to a matrix B of the same order if there exists a nonsingular matrix T such that

$$B = TAT^{-1}.$$

B in this case is said to be obtained from A by a Similarity Transformation. A criterion for similarity of two matrices is: Two matrices A and B are similar to each other if and only if they have the same invariant polynomials, or what is the same, the same elementary divisors.

(ii) The Canonical Forms of a Matrix.

Let $f(\lambda) = \lambda^n - a_n \lambda^{n-1} - a_{n-1} \lambda^{n-2} \dots - a_2 \lambda - a_1$ be a monic polynomial of degree n, then the matrix

$$A = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & \dots & 0 & 1 \\ a_1 & a_2 & \dots & a_n \end{pmatrix} \dots \dots \dots (1.1)$$

is such that $f(\lambda) = \det(\lambda I - A)$. The matrix A is said to be the Companion Matrix of the polynomial $f(\lambda)$. It is interesting to note that $f(\lambda)$ is the only non-unit invariant polynomial of A, the others all being equal to one.

The famous Cayley-Hamilton Theorem is: "Every matrix A satisfies

its characteristic equation". Those which also satisfy a polynomial equation of lower degree are called Derogatory. The monic polynomial equation of the least degree which is satisfied by a matrix A is called the Minimal Polynomial of A. A matrix A is said to be Nonderogatory if its characteristic polynomial is equal to its minimal polynomial. An important characteristic of a nonderogatory matrix is: A is nonderogatory if and only if its first n-1 invariant polynomials are 1. Thus: A companion matrix is nonderogatory.

The set of matrices TAT^{-1} constitute a class of matrices similar to A. In this class, there are certain ones which have structures simpler than that of A and therefore are convenient for the study of some properties of A. These matrices are known as Canonical Forms of A. The following forms are of special importance:

(a) Jordan Canonical Form. For every arbitrary matrix A, there always exists a nonsingular matrix T such that

$$T^{-1}AT = \begin{pmatrix} J_1 & 0 & & 0 \\ 0 & J_2 & 0 & 0 \\ \vdots & & & \\ 0 & 0 & 0 & J_k \end{pmatrix} \dots\dots\dots(1.2)$$

where each J_i is of the form

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 \\ \vdots & & & \\ 0 & 0 & \lambda_i & 1 \\ 0 & 0 & \dots\dots\dots 0 & \lambda_i \end{pmatrix} \dots\dots\dots(1.3)$$

The matrix in (1.2) is called the Jordan Canonical Form of A and the numbers λ_i appearing along the main diagonal of this form are Eigenvalues of A. The matrices J_i , $i = 1, \dots, k$ are called the Jordan Submatrices of A and the matrix A is said to be Diagonable if and only if each J_i is order 1.

(b) Rational or Frobenius Canonical Form. For every matrix A, there exists a nonsingular matrix T such that

$$T^{-1}AT = \begin{pmatrix} C_1 & 0 & \dots & 0 \\ 0 & C_2 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & C_k \end{pmatrix}$$

where each C_i , $i = 1, \dots, k$ is a companion matrix of an invariant polynomial of A. This form is called the Frobenius or Rational Canonical Form of A. In the case of a nonderogatory matrix, $k=1$ and therefore the only companion matrix appearing in the rational canonical form is of degree n. Thus: A nonderogatory matrix is similar to the companion matrix of its own characteristic polynomial.

(c) The Hessenberg and Tridiagonal Forms. For every matrix A, there always exists a unitary matrix U such that

$$UAU^* = H$$

where H is null below the first subdiagonal (or above the first superdiagonal). The form H is called the Hessenberg Form of A. If H is such that it is null both below the first subdiagonal and above the first superdiagonal, then H is said to be the Tridiagonal

Form. Thus, the Hessenberg and tridiagonal forms are intermediate between the original unreduced matrix and its rational canonical form.

(d) The Schur Canonical Form. For every matrix A , there exists a unitary matrix U such that $T = UAU^*$ is an upper triangular matrix with the eigenvalues of A along the main diagonal. The matrix T is said to be a Schur Canonical Form of A . If A is real and has real eigenvalues, then U may be chosen to be a real orthogonal matrix. The matrix A is Normal, if and only if T is diagonal.

(iii) Hermitian Forms and the Law of Inertia ([16], vol. 1, p. 297-308).

Any homogeneous polynomial H of the second degree in n variables x_1, x_2, \dots, x_n is called a Hermitian Form and can be represented as

$$H(x, x) = \sum h_{i, k} x_i \bar{x}_k$$

where the coefficient matrix H is hermitian ($H^* = H$). The rank of the matrix $H = (h_{i, k})$ is said to be the Rank of the Hermitian Form.

A Hermitian form can be represented in infinitely many ways in the form

$$H(x, x) = \sum_{i=1}^n a_i X_i \bar{X}_i \dots \dots \dots (1.4)$$

where $a_i, i=1, \dots, n$ are real numbers and

$$X_i = \sum \alpha_{i, k} x_k \quad (i=1, \dots, n)$$

are independent complex linear forms in variables x_1, x_2, \dots, x_n .

Since

$$X_i \bar{X}_i = |X_i|^2$$

every term of the right-hand side of (1.4) is a positive, a negative or a zero square according as $a_i > 0$, < 0 or $= 0$, respectively, and the number of non-zero squares is equal to the rank of the hermitian form. This representation of a hermitian form H as the sum of squares is known as a Canonical or Normal Representation of H . A fundamental result in the theory of hermitian forms now follows:

Sylvester's Law of Inertia: In a normal representation of a hermitian form H , the number of positive and the number of negative squares are independent of the choice of representation.

The triplet (π, ν, δ) of numbers of positive, negative and zero squares respectively in a normal representation of H is called the Inertia of the form and the difference σ between the number of positive and negative squares is called its Signature. The inertia of H is denoted by $\text{In}(H)$.

An important method of determining the rank and signature of a hermitian form is due to Jacobi and can be stated as follows:

(Adapted from [16], vol. i, p. 303).

Let D_1, D_2, \dots, D_r be the first r non vanishing leading principal minors of the determinant of the coefficient matrix H of a hermitian form of rank r , then the number π of positive squares and the number ν of negative squares are respectively equal to the number P of permanences and the number V of variations of sign in the sequence

$$1, D_1, D_2, \dots, D_r$$

i.e., $\pi = P(1, D_1, \dots, D_r)$, $\nu = V(1, D_1, \dots, D_r)$ and the signature

$$\sigma = r - 2V(1, D_1, D_2, \dots, D_r)$$

Note: In view of the uncertainties which may arise in the application of Jacobi's theorem, due to the vanishing of some of the principal minors, a modification of the method, in the real case, has been recently suggested by Howland ([27], p. 4).

A Hermitian form $H(x, x) = \sum h_{i, k} x_i \bar{x}_k$ is said to be Positive (Negative) definite, if

$$H(x, x) > 0 (<0) \text{ for } x_i \neq 0, i = 1, \dots, n.$$

Thus, it follows from the above result of Jacobi that H is positive definite if and only if $D_i, i=1, \dots, n$, are all positive.

(iv) The Bezoutiant (Adapted from [31], p. 15-20).

Let

$$f(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n \quad a_0 \neq 0$$

$$g(x) = b_0 x^n + b_1 x^{n-1} + \dots + b_n \quad b_0 \neq 0$$

be two polynomials of degree n . Then the form

$$B(f, g; x_0, \dots, x_{n-1}) = -B(g, f; x_0, \dots, x_{n-1})$$

which has

$$K(f) = \frac{f(x)g(y) - f(y)g(x)}{x - y} = \sum_{k, \ell=0}^{n-1} c_{k\ell} x^k y^\ell$$

as its generating function, is called the Bezoutiant of f and g .

The coefficient matrix $B = (c_{k, \ell})$ is called the Bezout matrix of the associated Bezoutiant. The numbers $c_{k, \ell}$ can be generated from a recursive relation as follows:

$$c_{k, \ell} = d_{0, k+\ell+1} + d_{1, k+\ell} \dots + d_{k, \ell+1}$$

where $d_{k, \ell} = a_{n-\ell} b_{n-k} - a_{n-k} b_{n-\ell}$ and $d_{k, \ell} = 0$ if k or $\ell > n$

A fundamental property of the Bezoutiant is: The rank of the Bezoutiant is equal to the order of the last non-vanishing principal minor of the Bezou-matrix $B = (c_{k\ell})$ if, in constructing the consecu-

tive principal minors, one starts from the lower right hand corner.

(v) Hankel Matrix and Hankel Form

Let $A = (a_{ij})$ be a matrix of order n . Then the sum of the diagonal elements of A is defined to be the Trace of A:

$$\text{tr}A = \sum_{i=1}^n a_{ii}$$

It is easy to see that

$$\text{tr}A = \sum \lambda_i$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A . Since A^k has eigenvalues $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ ($k = 0, 1, 2, \dots, n$), one immediately obtains:

$$\text{tr}A^k = S_k = \sum_{i=1}^n \lambda_i^k, \quad k = 0, 1, 2, \dots$$

The sums S_k are known as Newton Sums and are connected with the coefficients of the characteristic polynomial, as follows:

$$kp_k = S_k - p_1 S_{k-1} \dots - p_{k-1} S_1$$

$$(k = 1, 2, \dots, n)$$

where

$$\Delta(\lambda) = \lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} \dots - p_n$$

is the characteristic polynomial of A .

The matrix

$$H = \begin{pmatrix} S_0 & S_1 & \dots & S_{n-1} \\ S_1 & S_2 & \dots & S_n \\ \dots & \dots & \dots & \dots \\ S_{n-1} & S_n & \dots & S_{2n-2} \end{pmatrix} \dots \dots \dots (1.5)$$

is defined to be the Hankel Matrix associated with A and the quadratic form

$$H_n(x, x) = \sum_{i, k=0}^{n-1} S_{i+k} x_i \bar{x}_k$$

is called the Hankel form.

(vi) Kronecker Products and Linear Matrix Equations

Let $A = (a_{ij})$ be a matrix of order m and let $B = (b_{ij})$ be another matrix of order n . Then the partitioned matrix

$$K = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{22}B & \dots & a_{2m}B \\ \dots & \dots & \dots & \dots \\ a_{m1}B & a_{m2}B & \dots & a_{mm}B \end{pmatrix}$$

of order mn is defined to be the Kronecker Product of A and B and written as

$$A \otimes B = (a_{ij}B).$$

The following result about kronecker product, relating the eigenvalues of A and B to those of K is of fundamental importance and plays an essential role in the study of general linear matrix equations, to be described in a moment. (Adapted from [7], p. 235).

Let $\lambda_1, \lambda_2, \dots, \lambda_m$ be the eigenvalues of A and those of B be $\mu_1, \mu_2, \dots, \mu_n$. Then the eigenvalues of $K = A \otimes B$ are the mn products $\lambda_i \mu_j$, $i = 1, \dots, m$, $j = 1, \dots, n$.

Let A_1, A_2, \dots, A_n and B_1, B_2, \dots, B_n and C be matrices (real or complex) of order n , then the general linear matrix equation is

of the form

$$A_1 X B_1 + A_2 X B_2 + \dots + A_h X B_h = C$$

where X is also a matrix of order n , to be found. This matrix equation was first formulated by Sylvester [44] in 1884 and was then subsequently studied by Wedderburn [46] and Hitchcock [18] in the early twentieth century. Certain special cases of the equation arise frequently in applications and therefore have been studied by various authors since then. An excellent account of various methods of solving this equation has been recently published by Lancaster [32].

"The problem in its full generality is far from tractable", although the concept of kronecker product, just introduced, facilitates the study of the existence of solutions, to a certain extent.

The equation is essentially equivalent to a system of n^2 equations for n^2 elements x_{rs} of X . The matrix of this system, when x_{rs} are arranged in proper order, is

$$G = (A_1 \otimes B_1^T + A_2 \otimes B_2^T + \dots + A_h \otimes B_h^T)$$

The theory of systems of linear equations, therefore, immediately gives the following result [Adapted from [34], p. 89).

THEOREM 1.1 A necessary and sufficient condition that the matrix equation have a solution X is that the matrix G have the same rank as the $n^2 \times (n^2 + 1)$ array obtained by bordering G with the elements of C .

1.2 NONDEROGATORY MATRICES AND JORDAN STRUCTURES

The Hessenberg form is a very convenient starting point for the solution of many problems in numerical linear algebra, especially for the solution of the general matrix eigenvalue problem, and therefore the reduction of an arbitrary matrix to Hessenberg form is a standard elementary procedure for most of the modern algorithms designed for solving this problem. It is noted that there exists at least one stable and accurate procedure for transforming an arbitrary matrix to Hessenberg form ([47], p. 347-349). This preliminary transformation has the effect of splitting the given matrix into nonderogatory parts in the sense that the eigenvalues of the given matrix are those of the nonderogatory parts taken together. Therefore, as far as the problem of knowing the Jordan structure of a matrix is concerned, there is no loss of generality in restricting attention to nonderogatory matrices.

1.3 QUADRATIC FORMS AND JORDAN STRUCTURE

1.3.1 Jordan Structure and the Associated Hankel Matrix

The Hankel matrix H defined in (1.5) plays an important role in separating the zeros of a polynomial (see e.g., [16], vol. II p. 202-204).¹ But it does not seem to be generally known that it is equally effective for determining the jordan structure of an arbitrary matrix. The purpose of this section is to show how H can be employed to solve this problem.

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A . Then H can be written as

$$H = \begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_n \\ \vdots & \dots & \dots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \lambda_n^{n-1} \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{n-1} \\ \vdots & \dots & \dots & \dots \\ 1 & \lambda_n & \dots & \lambda_n^{n-1} \end{pmatrix} = V^T V,$$

whence

$$\det H = (\det V)^2.$$

Since V is a Vandermonde matrix, $\det V = 0$ if and only if $\lambda_i = \lambda_j$ for some i and j . One therefore obtains the following result.

immediately:

THEOREM 1.2 An arbitrary matrix A , real or complex has all its eigenvalues distinct if and only if the Hankel matrix H in (1.5) is non-singular.

Let exactly $r \leq n$ of the eigenvalues of A be distinct and let λ_i ($i = 1, \dots, r$) be of multiplicity v_i , so that $v_1 + v_2 + \dots + v_r = n$. Then the matrix

¹ These results may also be found in ([48], p. 407), in [49] and in [56].

$$U_r = \begin{pmatrix} s_0 & s_1 & \cdots & s_{r-1} \\ s_1 & s_2 & \cdots & s_r \\ \vdots & \vdots & \cdots & \vdots \\ s_{r-1} & s_r & \cdots & s_{2r-2} \end{pmatrix}$$

standing in the upper left hand corner of H is equal to

$$U_r = \begin{pmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_r \\ \vdots & \cdots & \vdots \\ \lambda_1^{r-1} & \cdots & \lambda_r^{r-1} \end{pmatrix} \begin{pmatrix} v_1 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & v_r \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{r-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & \lambda_r & \cdots & \lambda_r^{r-1} \end{pmatrix}$$

Since λ_i , $i = 1, \dots, r$ are all distinct, it follows from the above expression of U_r that U_r is non-singular. Moreover, the rank of the Hankel matrix H cannot exceed the rank r of V and so, the rank of H is exactly equal to r . Thus theorem 1.3 follows:

THEOREM 1.3 The number of distinct eigenvalues of an arbitrary square matrix A is equal to the rank of its associated Hankel matrix H.

In particular, when all the eigenvalues of A are real, since each v_i is positive, U_r is positive definite. One therefore obtains:

COROLLARY TO THEOREM 1.3 Suppose that all the eigenvalues of an arbitrary square matrix A are real, and that r of them are distinct. Then the r -rowed leading principal minor submatrix U_r of the associated Hankel matrix H is positive definite.

Next, let A be a real matrix, so that its complex eigenvalues occur in conjugate pairs. Let

$$Q = x^T Hx = \sum (x_1 + \lambda_j x_2 + \dots + \lambda_j^{n-1} x_n)^2$$

be a quadratic form associated with H . If λ_1 and $\lambda_2 = \bar{\lambda}_1$ are complex eigenvalues, then

$$x_1 + \lambda_1 x_2 + \dots + \lambda_1^{n-1} x_n = u + iv$$

$$x_1 + \lambda_2 x_2 + \dots + \lambda_2^{n-1} x_n = u - iv$$

and when these terms are squared and added, the result is $2u^2 - 2v^2$. Again, since for each real λ_j , the corresponding square in Q must be positive, one arrives at the following:

THEOREM 1.4 Let A be a real square matrix of order n . Then the number of different pairs of complex conjugate eigenvalues of A is equal to the number of negative terms in a canonical representation of the associated Hankel matrix H .

Combining theorem 1.3 with theorem 1.4 one immediately obtains the following:

THEOREM 1.5 The number of distinct real eigenvalues of a real matrix A is equal to the signature of its associated Hankel matrix H .

Applying Jacobi's method of determining the rank and signature of a real symmetric matrix (p.8), theorem 1.4 and theorem 1.5 can be reformulated in a single theorem as follows:

THEOREM 1.6 Let A be a real square matrix and let r be the rank of its associated Hankel matrix H . Then, if, the first r leading principal minors H_i , $i = 1, \dots, r$ of H are all different from zero, A has p different pairs of complex eigenvalues and $p - q$ distinct real eigenvalues where p and q are respectively equal to the number of variations and the number of permanences of sign in the sequence

$$1, H_1, H_2, \dots, H_r.$$

Example

Let

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

be a real matrix of order 3.

$$S_0 = 3, S_1 = 4, S_2 = 6, S_3 = 10, S_4 = 18.$$

$$H_1 = 3, H_2 = 2, H_3 = 0.$$

Thus the Hankel matrix H has rank 2 and

$$P(1, H_1, H_2) = P(1, 3, 2) = 2$$

$$V(1, H_1, H_2) = V(1, 3, 2) = 0$$

So, all the eigenvalues of A are real and two of them are distinct.

1.3.2 Hankel Form and Similarity to a Normal Matrix

A matrix A is said to be Normal if $AA^* = A^*A$ ($A^* = \overline{A}^T$). A remarkable property of a normal matrix is that it is diagonalizable by a unitary similarity and therefore, in case of a normal matrix, the Jordan structure problem is completely solved. Conversely, if a given matrix A is diagonalizable by any non-singular similarity T (not necessarily unitary) then, since a diagonal matrix is normal, the matrix A is similar to a normal matrix. The question now arises as to when a given matrix (nonderogatory) is diagonalizable. It will be shown in this section how the Hankel matrix H can be employed to answer the question.

Let A be a given nonderogatory matrix and let $\lambda I - A$ denote

the characteristic polynomial of A . Then the matrix A is similar to a diagonal matrix D if and only if all the elementary divisors of $\lambda I - A$ are linear. For, by the definition of similarity, $\lambda I - A$ and $\lambda I - D$ have the same elementary divisors and the elementary divisors of $\lambda I - D$ are all linear. Again, the definition of elementary divisor implies that the elementary divisors of a matrix A are linear if and only if the invariant polynomial of A of highest degree have distinct linear factors. Since for a nonderogatory matrix A , the invariant polynomial of highest degree is just its characteristic polynomial, one arrives at the following result:

A nonderogatory matrix A is diagonalizable if and only if it has all its eigenvalues distinct. Thus, it follows from theorem 1.2 that:

THEOREM 1.7 A nonderogatory matrix A is similar to a normal matrix if and only if the associated Hankel matrix H in (1.5) is non-singular.

1.4 MATRIX EQUATIONS AND JORDAN STRUCTURE

1.4.1 Formulation of a Non-linear Matrix Equation

The problem posed in the previous section may be formulated in terms of a non-linear matrix equation as follows:

Let A be diagonalizable. Then there exists a non-singular matrix P such that

$$P^{-1}AP = D, \text{ a diagonal matrix and since,}$$

$$DD^* = D^*D$$

$$P^{-1}APP^*A^*(P^*)^{-1} = P^*A^*(P^*)^{-1}P^{-1}AP.$$

The matrix

$$S = (PP^*)^{-1}$$

is thus seen to satisfy the non-linear matrix equation

$$A S^{-1} A^* S^{-1} = S^{-1} A^* S^{-1} A \dots \dots \dots (1.6)$$

Conversely, let the above equation (1.6) have a positive definite solution S . Then, by Cholesky decomposition ([21], p. 127),

$$S = LL^*,$$

where L is non-singular lower triangular. The matrix T defined by

$$T = (L^*)^{-1}$$

satisfies

$$T^{-1} A T T^* A^* (T^*)^{-1} = T^* A^* (T^*)^{-1} T^{-1} A T$$

whence $T^{-1} A T$ is normal, according to the definition of normality and therefore A is diagonalizable. Thus:

A matrix A is diagonalizable if and only if the non-linear matrix equation (1.6) admits a positive definite solution.

There exist procedures in the literature ([47], p. 486) for the diagonalization of normal matrices by similarity, which are reliable, accurate and efficient and therefore, the knowledge of the solutions S of the matrix equation would not only help to solve the Jordan structure problem for A , but also would make possible the actual computation of the eigenvalues and eigenvectors of A in a way in which convergence will be guaranteed.

1.4.2 The Matrix Equation $SA = A^*S$

The simplest matrix equation relevant to the problem of knowing the Jordan structure of a given matrix A is

$$SA = A^*S \dots\dots\dots(1.7)$$

The solutions S of this matrix equation satisfy the non-linear equation (1.6) and therefore there is some interest in solving this equation. In what follows, it will be shown what information can be derived from the existence of different types of solutions of the equation and how to compute these solutions numerically.

(i) Existence of Solutions and Jordan Structure

The equation (1.7) is a special case of the general linear matrix equation

$$A_1 X B_1 + A_2 X B_2 + \dots + A_n X B_n = 0$$

discussed earlier and the general theory when applied to this special case shows that the matrix equation is equivalent to a system of algebraic equations of the form

$$Gx = 0$$

where

$$G^T = A \otimes I - I \otimes \bar{A}$$

Since the eigenvalues of G are n^2 numbers

$$\lambda_i - \bar{\lambda}_j, \quad i, j = 1, \dots, n, \quad \text{where } \lambda_1, \lambda_2, \dots, \lambda_n \text{ are}$$

the eigenvalues of A , one obtains the following result

immediately:

THEOREM 1.8 The matrix equation (1.7) has a non-trivial solution S if and only if $\lambda_i = \bar{\lambda}_j$ for some i and j .

Next, let there exist a non-singular solution S of (1.7). Then, since A is similar to A^* , A and A^* have their eigenvalues in common. Thus a necessary condition that the equation (1.7) admit a non-singular solution S is that all the complex eigenvalues of A occur in conjugate pairs.

Conversely, let all the complex eigenvalues of A occur in conjugate pairs and let P be a non-singular matrix such that

$$PAP^{-1} = \text{diag}(J_1, J_2, \dots)$$

where each J_i is of the form (1.3). It can be assumed without any loss of generality that, in this canonical representation of A , all non-real J_i 's appear first and then the real ones. Then, since each J_i is either real or else \bar{J}_i is also one of the Jordan blocks different from J_i , the non real J_i 's may be grouped in pairs such that each pair contains a Jordan block and one of its complex conjugates. With this grouping of the Jordan blocks one can write

$$PAP^{-1} = \text{diag}(C_1, C_2, \dots; D_1, D_2, \dots)$$

where each C_i is a matrix of paired non-real Jordan blocks and D_i 's are real Jordan blocks. For each matrix C_i of order $2v$, it can be shown that there is a matrix

c_1 non zero.

Setting now

$$\bar{S} = \text{diag}(G_1, G_2, \dots; H_1, H_2, \dots)$$

it is seen that

$$\bar{S} P A P^{-1} = (P A P^{-1})^* \bar{S}$$

whence

$$S = P^* \bar{S} P,$$

satisfies

$$S A = A^* S.$$

Clearly, S can be made non-singular and the result follows:

THEOREM 1.9 A necessary and sufficient condition that there exist a non-singular solution S of the equation (1.7) is that any complex eigenvalues of A occur in conjugate pairs.

Since, for a real matrix A , complex eigenvalues always occur in conjugate pairs, one immediately obtains the following:

COROLLARY TO THEOREM 1.9 For every real matrix A , there always exists a non-singular matrix S transforming A into its transpose.

Lastly, let there exist a positive definite solution S of the equation. Then, since S is positive definite, $S^{\frac{1}{2}}$ and $S^{-\frac{1}{2}}$ are properly defined. The equation (1.7) therefore, can be written as

$$S^{+\frac{1}{2}} A S^{-\frac{1}{2}} = S^{-\frac{1}{2}} A^* S^{\frac{1}{2}}.$$

If now

$$P = S^{\frac{1}{2}} A S^{-\frac{1}{2}},$$

then

$$P^* = S^{-\frac{1}{2}} A^* S^{\frac{1}{2}} = S^{\frac{1}{2}} A S^{-\frac{1}{2}} = P$$

and

$$A = S^{-\frac{1}{2}} P S^{\frac{1}{2}}$$

and therefore similar to P .

Conversely, if A is similar to a hermitian matrix, then there exists a non-singular matrix T such that $T^{-1}AT$ is hermitian. So,

$$T^{-1}AT = T^*A^*(T^*)^{-1}.$$

If, now $S = (TT^*)^{-1}$ then S is positive definite and satisfies (1.7).

Thus:

THEOREM 1.10 The matrix equation (1.7) admits a positive definite hermitian solution S if and only if A is similar to a hermitian matrix.

Since the eigenvalues of a hermitian matrix are all real and the similarity transformation preserves the eigenvalues, it follows from theorem 1.10 that the existence of a positive definite hermitian solution S of the equation 1.7 not only implies that A is similar to a normal matrix, but also that it has all its eigenvalues real.

REMARK : The result of the Corollary to theorem 1.9 is true for an arbitrary matrix, real or complex, and is well known.

(ii) Construction of Solutions

Theorems 1.8 to 1.10 are all existence theorems and assume knowledge of the jordan canonical form of A . A theory will now be developed which will provide an alternative criterion of exis-

tence of solutions without any prior information concerning the jordan form of A and at the same time will provide a constructive method of computing these solutions in case A is given in its rational canonical form.

Constructions of Non Trivial Solutions (An Alternative Criterion for Existence)

Let A_0 be a companion matrix of the form (1.1) and let s_1, s_2, \dots, s_n be the n rows of a matrix S_0 . Then the successive rows of $S_0 A_0$ are $s_1 A_0, s_2 A_0, \dots$ and $s_n A_0$ and the successive rows of $A_0^* S$ are $\bar{a}_1 s_n, s_1 + \bar{a}_2 s_n, \dots$ and $s_{n-1} + \bar{a}_n s_n$. The matrix equation

$$S_0 A_0 = A_0^* S_0$$

is therefore completely equivalent to

$$\bar{a}_1 s_n = s_1 A_0$$

$$s_{i-1} + \bar{a}_i s_n = s_i A_0, \quad i = 2, 3, \dots, n$$

Eliminating $s_i, i = 1, \dots, n-1$, one gets

$$s_n (A_0^n - \bar{a}_n A_0^{n-1} \dots - \bar{a}_2 A_0 - \bar{a}_1) = 0$$

This equation gives rise to the following system of linear equations:

$$\bar{\phi}(A_0^T) s_n^T = 0$$

where $\phi(x)$ is the characteristic polynomial of A_0 and $\bar{\phi}(x)$ is the polynomial whose coefficients are the complex conjugates of those of $\phi(x)$.

Again, since A is assumed to be nonderogatory, the matrix equation (1.7) is seen to be completely equivalent to

$$S_0 A_0 = A_0^* S_0$$

where

$$S_0 = (P^*)^{-1} S,$$

P being the matrix transforming A into its rational canonical form A_0 . One thus obtains the following result:

THEOREM 1.11 Let A be a nonderogatory matrix. Then the matrix equation (1.7) has a non-trivial solution S if and only if there exists a non-null vector s such that $s \bar{\phi}(A) = 0$, where $\phi(x)$ is the characteristic polynomial of A .

A Particular Case - A is Real: In case the given matrix A is real, $\bar{\phi}(A) = \phi(A)$ and hence is a null matrix, by the Cayley-Hamilton Theorem. One thus obtains the following well-known result:

COROLLARY TO THEOREM 1.11 Let A be a nonderogatory real matrix. Then the matrix equation

$$SA = A^T S$$

admits an n -parameter family of solutions.

(b) Construction of Non-singular Hermitian and Symmetric Solutions

It has recently been shown by Duke [12] that the matrix equation (1.7) has a non-singular hermitian solution whenever it admits a non-singular solution. In fact, if P is a non-singular matrix satisfying (1.7) then the matrix

$$S = cP + \bar{c}P^*$$

where c is a complex number, also satisfies the equation. Clearly, S is hermitian and can be made non-singular by choosing c properly.

The problem of finding a non-singular solution P of the equation (1.7) is, however, not a trivial one. It is pleasant to note that, in case the given matrix A is a companion matrix, the constructive procedure used to prove theorem 1.11 really gives a constructive method of computing such a solution. Indeed, if s_n is any row n vector which satisfies $\overline{\phi}(A^T) s_n^T = 0$ and makes S non-singular, then, taking s_n as n th row vector of S , all other row vectors s_i , $i = 1, \dots, n-1$ of S can be computed recursively using the relation

$$\overline{a}_1 s_n + s_{i-1} = s_i A, \quad i = 2, 3, \dots, n$$

In the real case, the situation is however different. The equation $SA = A^T S$ has been studied in some details in the literature and consequently, a few results are available on its solution. Any symmetric solution S of the equation

$$SA = A^T S \dots\dots\dots(1.8)$$

is called a Symmetrizer of A and the matrix A is said to be symmetrized by S in this case. It was shown by Desautels [11] that there always exists a n -parameter family of symmetric solutions of (1.8) and every solution of the equation (1.8) is symmetric if and only if A is non-derogatory. This result was also obtained by Tausky and Zassenhaus [45] using different techniques. A numerical constructive procedure of obtaining a symmetrizer of a given arbitrary matrix A was first obtained by Howland and Farrell [28], who showed that, for an arbitrary given matrix A , one could actually construct a tridiagonal matrix T and a diagonal matrix D such that

A is transformed to T by similarity and D symmetrizes T:

$$DT = T^T D.$$

Unfortunately, the algorithm proposed for this purpose was adapted from one of the suspect methods for the general matrix eigenvalue problem, namely the method of Lanczos or minimized iterations ([21], p. 17-24) and therefore it suffers from numerical instability in some practical cases. A stable method has been recently proposed by Howland [25], showing that one could construct a symmetric solution of (1.8) by first transforming the given matrix A to Hessenberg form. Indeed, if A is in upper Hessenberg form and the first row and the first column of S are chosen arbitrarily, then, in view of the facts that the matrices S and SA will have to be symmetric, one finds that the remaining elements of S can be generated systematically. Incidentally, in view of Desautel's result quoted above, one immediately finds that the Corollary to theorem 1.11 gives a constructive method for computing a symmetrizer of a companion matrix. The method can be formulated as follows:

An Algorithm For Computing a Symmetrizer for a Companion Matrix

Let $s_1, s_2 \dots s_n$ be the successive rows of a symmetrizer S. Then

- (i) Choose s_n arbitrarily
- (ii) Compute the remaining rows recursively from the following:

Example $s_{i-1} = s_i A - a_i s_n, i = n, n-1, \dots, 3, 2.$

Let

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_1 & a_2 & a_3 \end{pmatrix}$$

be a given matrix of order 3. Then, if s_1 , s_2 and s_3 are the successive rows of S , choosing $s_3 = (1, 1, 1)$, s_1 and s_2 are given by

$$s_2 = s_3 A - a_3 s_3 = (1, 1, 1) \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_1 & a_2 & a_3 \end{pmatrix}$$

$$-a_3(1, 1, 1) = (a_1 - a_3, 1 + a_2 - a_3, 1)$$

$$s_1 = s_2 A - a_2 s_3 = (a_1 - a_3, 1 + a_2 - a_3, 1) \times$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_1 & a_2 & a_3 \end{pmatrix} - a_2(1, 1, 1)$$

$$= (a_1 - a_2, a_1 - a_3, 1)$$

Thus, a symmetrizer S of A is

$$S = \begin{pmatrix} a_1 - a_2 & a_1 - a_3 & 1 \\ a_1 - a_3 & a_2 - a_3 + 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Hankel Matrix as a Symmetrizer

By the algorithm just stated, if X is a symmetrizer of a companion matrix A with x_1, x_2, \dots, x_n as its row vectors, then x_n can be chosen arbitrarily and x_1, \dots, x_{n-1} are given by

$$x_i = x_n (A^{n-i} - a_n A^{n-i-1} \dots - a_{i+1}),$$

$$i = 1, 2, \dots, n-1$$

In what follows, it will be shown that in case A has distinct eigenvalues there is a unique choice of the vector x_n for which the matrix X becomes just equal to H^{-1} , where H is the Hankel matrix in (1.5).

The first n-1 successive rows of XH are

$$x_n (A^{n-i} - a_n A^{n-i-1} \dots - a_{i+1}) H$$

$$i = 1, 2, \dots, n-1.$$

Therefore, the equation $XH = I$ is true if and only if

$$\left\{ \begin{array}{l} x_n (A^{n-i} - a_n A^{n-i-1} \dots - a_{i+1}) H = e_i, \\ i = 1, \dots, n-1 \\ x_n H = e_n \end{array} \right. \dots\dots(1.9)$$

and

where e_i is the i th row of the identity matrix I. The relation (1.9) is equivalent to

$$\left\{ \begin{array}{l} x_n A^{n-1} H = c_1 e_1 + c_2 e_2 + \dots + c_n e_n \\ x_n A^{n-2} H = c_1 e_2 + c_2 e_3 + \dots + c_{n-1} e_n \\ \vdots \\ x_n A H = c_1 e_{n-1} + c_2 e_n \\ x_n H = c_1 e_n \end{array} \right. \dots\dots(1.9')$$

where c_1, c_2, \dots, c_n are given by the following recursive relation

$$\left\{ \begin{array}{l} c_1 = 1 \\ c_{i+1} = a_n e_i + a_{n-1} c_{i-1} + \dots + a_{n-i+1} c_1, \end{array} \right. \dots\dots(1.10)$$

$$i = 1, 2, \dots n-1$$

The relation (1.9') gives rise to a system of n^2 non homogeneous equations in n unknowns $x_{11}, x_{12} \dots x_{1n}$, which, when arranged properly, can be written in the form

$$Mx'_n = c \dots\dots\dots(1.11)$$

where

$$x'_n = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1n} \end{pmatrix}$$

$$c = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ c_1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ c_1 \\ c_2 \\ \vdots \\ c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{pmatrix} \begin{matrix} \left. \begin{matrix} 0 \\ 0 \\ \vdots \\ 0 \end{matrix} \right\} 1^{st} n \\ \left. \begin{matrix} c_1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ c_1 \\ c_2 \end{matrix} \right\} 2^{nd} n \\ \left. \begin{matrix} \vdots \\ c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{matrix} \right\} last n \end{matrix}$$

The first n coordinates of c are $0, 0, \dots 0, c_1$, the second n coordinates are $0, 0, \dots 0, c_1, c_2$ and so on.

and the coefficient matrix $M(n^2 \times n)$ is

$$M = \left(\begin{array}{ccc} s_0 & s_1 & \dots s_{n-1} \\ s_1 & s_2 & \dots s_n \\ \vdots & \dots & \dots \\ s_{n-1} & \dots & s_{2n-2} \end{array} \right) \begin{array}{l} \text{First } n \\ \dots \\ \dots \end{array}$$

$$\left(\begin{array}{ccc} s_1 & s_2 & \dots s_n \\ s_2 & s_3 & \dots s_{n+1} \\ \vdots & \dots & \dots \\ s_n & s_{n+1} & \dots s_{2n-1} \end{array} \right) \begin{array}{l} \text{Second } n \\ \dots \\ \dots \end{array}$$

$$\left(\begin{array}{ccc} s_{n-1} & s_n & \dots s_{2n-2} \\ s_n & s_{n+1} & \dots s_{2n-1} \\ \vdots & \dots & \dots \\ s_{2n-2} & s_{2n-1} & s_{3n-3} \end{array} \right) \begin{array}{l} \text{Last } n \\ \dots \\ \dots \end{array}$$

It now follows from the identities

$$\begin{cases} s_n = a_n s_{n-1} + a_{n-1} s_{n-2} + \dots + a_2 s_1 + a_1 s_0 \\ s_i = a_n s_{i-1} + a_{n-1} s_{i-2} + \dots + a_2 s_{i-n+1} + a_1 s_{i-n} \quad (1 > n) \end{cases}$$

and the recursive relations (1.10), that the system (1.11) is reduced to the following system of n equations in n unknowns:

$$\begin{pmatrix} s_0 & \dots & s_{n-1} \\ s_1 & \dots & s_n \\ \vdots & & \\ s_{n-1} & \dots & s_{2n-2} \end{pmatrix} x'_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ c_1 \end{pmatrix} \dots \dots \dots (1.12)$$

Since the coefficient matrix of the system (1.12) is non-singular, the system has a unique solution and this solution, when substituted in X , gives the inverse of the matrix H . Thus one obtains the following result: (The result was also obtained by Howland [24] before, using a different technique).

THEOREM 1.12 Let A be a nonderogatory matrix in its rational canonical form with its eigenvalues all distinct. Then the inverse of the Hankel matrix H in (1.5) symmetrizes A .

(c) Construction of a Positive Definite Symmetrizer

Let A be a nonderogatory matrix in its rational canonical form, with real and distinct eigenvalues. Then by theorem 1.10, there is a positive definite solution of the equation (1.8). In fact, the inverse of the associated Hankel matrix H is such a symmetrizer. For, by the result just stated, $S = H^{-1}$ is a symmetrizer of A and since the eigenvalues are all real and distinct, by the Corollary to theorem 1.3, S is positive definite.

Example

Let

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & 3 \end{pmatrix}$$

be the given matrix of order 3. Then, the successive rows of a symmetrizer X of A are

$$x_3 \begin{pmatrix} 2 & -3 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad x_3 \begin{pmatrix} -3 & 1 & 0 \\ 0 & -3 & 1 \\ 0 & -2 & 0 \end{pmatrix} \quad \text{and } x_3,$$

where $x_3 = (x_{11}, x_{12}, x_{13})$ is an arbitrary vector. The system (1.12) is

$$3x_{11} + 3x_{12} + 5x_{13} = 0$$

$$3x_{11} + 5x_{12} + 9x_{13} = 0$$

$$5x_{11} + 9x_{12} + 17x_{13} = 1$$

The unique solution of the system is

$$x_{11} = \frac{1}{2}, x_{12} = -3, x_{13} = \frac{3}{2}$$

The successive rows of the matrix H^{-1} are therefore:

$$\left(1, -\frac{3}{2}, \frac{1}{2}\right), \left(-\frac{3}{2}, \frac{13}{2}, -3\right) \text{ and } \left(1, -3, \frac{3}{2}\right).$$

The matrix

$$S = H^{-1} = \begin{pmatrix} 1 & -\frac{3}{2} & \frac{1}{2} \\ -\frac{3}{2} & \frac{13}{2} & -3 \\ \frac{1}{2} & -3 & \frac{3}{2} \end{pmatrix} \text{ is positive definite and symmetrizes } A.$$

1.4.3 Matrix Equation $SP(A) = A*S$

There clearly exist diagonalizable matrices whose eigenvalues are not all real, and therefore there must correspondingly exist solutions of the non-linear equation (1.6) which are not solutions of the linear equation (1.7). To determine these, the following well-known result about commutativity of matrices is recalled first.

(Adapted from [16], Vol. I, p. 222).

THEOREM 1.13 Every polynomial in a matrix A commutes with A ; and

every matrix that commutes with A is a polynomial in A when, and only when, A is nonderogatory.

The matrix equation (1.6) really states that the matrices A and $S^{-1}A*S$ are commutative and since A is assumed to be nonderogatory it follows from the above theorem that

$$S^{-1}A*S = P(A),$$

where $P(A)$ is a polynomial in A . The matrix equation (1.6) is therefore equivalent to the equation

$$SP(A) = A*S \dots\dots\dots(1.13)$$

A detailed study of the equation (1.13) and in particular, the question of existence and construction of a positive definite solution will now follow:

Construction of Non-Trivial and Non-Singular Solutions

The equation (1.13) is also a special case of the general linear equation and therefore the criterion of existence and uniqueness of the solutions of this equation may be obtained by applying the general theory. Since, this is not useful for the present purpose, a detailed description is omitted. However, it is easy to observe that the constructive method proposed for obtaining the solutions of the equation (1.7) can also be successfully applied to this general equation (1.13).

Let A be a nonderogatory matrix of order n and let $P(A)$ be a polynomial in A . Then the equation (1.13) is equivalent to

$$S_1 P(A) = A_1 * S_1 \dots \dots \dots (1.14)$$

where A_1 is in the companion form (1.1) and

$$S_1 = (T^*)^{-1} S$$

T being the matrix transforming A to the companion form (1.1). The equation (1.13) is therefore completely solved if all the solutions of (1.14) are known, and vice-versa. Attention will thus be concentrated in assuming A_1 in the form (1.1) and finding S_1 satisfying (1.14).

Let s_i , $i = 1, \dots, n$ be the rows of S_1 . Then the successive rows of $S_1 P(A)$ are $s_1 P(A)$, $s_2 P(A)$, \dots , $s_n P(A)$ and those of $A_1 * S_1$ are $\bar{a}_1 s_n$, $s_1 + \bar{a}_2 s_n$, \dots , $s_{n-1} + \bar{a}_n s_n$. These two products are equal if and only if

$$\bar{a}_1 s_n = s_1 P(A)$$

$$s_{i-1} + \bar{a}_i s_n = s_i P(A), \quad i = 2, 3, \dots, n$$

The last relation is equivalent to

$$s_n (P^n(A) - \bar{a}_{n-1} P^{n-1}(A) \dots - \bar{a}_2 P(A) - \bar{a}_1 I) = 0$$

that is,

$$s_n \bar{\phi}(P(A)) = 0$$

where $\phi(x)$ is the characteristic polynomial of A_1 . One thus arrives at the following result:

THEOREM 1.14. Let A be a nonderogatory matrix of order n and $P(A)$ be a polynomial in A . Then the matrix equation (1.13) has a non-trivial solution S if and only if the n th row s_n of S is such that $s_n \bar{\phi}(P(A)) = 0$ where $\phi(x)$ is the characteristic polynomial of A . The other row vectors

of S are given by the following recursive relations

$$s_{i-1} + \bar{a}_i s_n = s_i P(A), \quad i = n, n-1, \dots, 3, 2.$$

Let the rank r of the matrix $\bar{\phi}(P(A))$ be such that $0 < r < n$. Then, since $\bar{\phi}(P(A))$ is not a null matrix, the matrices $P(A)$ and A^* have fewer than n eigenvalues in common and hence $P(A)$ is not similar to A^* , one therefore has

COROLLARY TO THEOREM 1.14 The equation $SP(A) = A^*S$ does not have a non-singular solution if $0 < \text{rank } \bar{\phi}(P(A)) \leq n$.

REMARK It is noted that the above discussion remains valid if $P(A)$ is any given matrix of the same order as A , not necessarily a polynomial in A .

Since the construction of positive definite solutions is of primary interest and every positive definite solution is necessarily non-singular, this corollary gives rise to the following question:

Given a matrix A , does there exist a polynomial $P(A)$ in A such that $P(A) \neq A$ has the same eigenvalues as A^* ? To answer this, one proceeds as follows:

Let $\lambda_1, \lambda_2, \dots, \lambda_r$ be the distinct eigenvalues of A with multiplicities $v_i, i = 1, \dots, r$ and let

$$P(x) = c_0 + c_1 x + c_2 x^2 + \dots + c_{n-1} x^{n-1}$$

be a polynomial of degree not exceeding $n-1$. Then $P(A)$ and A^* have all their eigenvalues in common if the system of equations

$$c_0 + c_1 \lambda_i + c_2 \lambda_i^2 + \dots + c_{n-1} \lambda_i^{n-1} = \bar{\lambda}_i, \dots (1.15)$$

$$i = 1, \dots, r$$

has a solution. In matrix notation, the system (1.15) can be written as

$$Mc = \lambda$$

where

$$c = \begin{pmatrix} c_1 \\ \vdots \\ c_{n-1} \end{pmatrix}, \quad \lambda = \begin{pmatrix} \bar{\lambda}_1 \\ \bar{\lambda}_2 \\ \vdots \\ \bar{\lambda}_r \end{pmatrix}$$

and the coefficient matrix

$$M = \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^r & \lambda_1^{r+1} & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \dots & \lambda_2^r & \lambda_2^{r+1} & \dots & \lambda_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \lambda_j & \dots & \lambda_j^r & \lambda_j^{r+1} & \dots & \lambda_j^{n-1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \lambda_r & \dots & \lambda_r^r & \lambda_r^{r+1} & \dots & \lambda_r^{n-1} \end{pmatrix}$$

If $r < n$, the coefficient matrix M and the augmented matrix (M, λ) are both rectangular and the Vandermonde matrix

$$M_1 = \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{r-1} \\ \dots & \dots & \dots & \dots \\ 1 & \lambda_r & \dots & \lambda_r^{r-1} \end{pmatrix}$$

standing on the upper left hand corner of both of them is non-

singular, since $\lambda_1, \lambda_2, \dots, \lambda_r$ are all distinct. Thus, these matrices have the same rank and the system (1.15) is consistent. When $r = n$, each $v_i = 1$ and the coefficient matrix M is then clearly non-singular. One therefore arrives at the following result:

THEOREM 1.15 Let A be an arbitrary non-scalar matrix of order n . Then there always exists a polynomial $P(A)$ in A such that $P(A) \neq A$, but $P(A)$ and A^* have the same characteristic polynomial.

Note In case all the eigenvalues of A are real and distinct, the system of equations (1.15) is to be set up as follows:

$$\begin{cases} c_0 + c_1\lambda_i + c_2\lambda_i^2 + \dots + c_{n-1}\lambda_i^{n-1} = \lambda_{i-1}, & i = 1, \dots, n \\ \lambda_0 = \lambda_n \end{cases}$$

As before, it can be shown that the system is consistent and has a unique solution and this solution does not correspond to $P(A) = A$.

A Particular Case - A is Real

In case the matrix A is real, $A^* = A^T$ and since A and A^T have the same characteristic polynomial, the following result is easily obtained as a corollary to the above theorem.

COROLLARY TO THEOREM 1.15 (Generalized Cayley-Hamilton Theorem)

For every real non-scalar matrix A , there exists a polynomial $P(A) \neq A$ such that $P(A)$ satisfies the characteristic equation of A .

Construction of Polynomials

Next, it will be shown how polynomials $P(A)$ can actually be constructed in terms of the symmetric functions of the eigenvalues of A , when the given matrix A is real and its eigenvalues are not all distinct.

Since A is real, the system (1.15) reduces to

$$c_0 + c_1 \lambda_i + c_2 \lambda_i^2 + \dots + c_{n-1} \lambda_i^{n-1} = \lambda_i, \dots (1.16)$$

$$i = 1, 2, \dots, r$$

This system is consistent, because $c_0 = c_2 = \dots = c_{n-1} = 0$; $c_1 = 1$ is always a solution. Again, since $\lambda_1, \dots, \lambda_r$ are all distinct, the rank of the system is $r < n$ and therefore, $n-r$ unknowns may be chosen so that the coefficient matrix of the remaining r unknowns is of rank r and whenever these $n-r$ unknowns are assigned any values whatsoever, the remaining r unknowns are uniquely determined. Indeed, if $c_r, c_{r+1}, \dots, c_{n-1}$ are all chosen equal to -1 , then the system (1.6) becomes

$$c_0 + c_1 \lambda_i + c_2 \lambda_i^2 + \dots + c_{r-1} \lambda_i^{r-1} = \lambda_i + \lambda_i^r + \dots + \lambda_i^{n-1} \dots (1.17)$$

$$i = 1, 2, \dots, r$$

The matrix of this system is again a Vandermonde matrix whose determinant is not zero. So, the system can be solved uniquely for the unknowns $c_0, c_1, c_2, \dots, c_{r-1}$. In matrix notation, the system (1.17) can be written as

$$\begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{r-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & \lambda_r & \dots & \lambda_r^{r-1} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{r-1} \end{pmatrix} = \begin{pmatrix} \lambda_1 + \lambda_1^r + \dots + \lambda_1^{n-1} \\ \lambda_2 + \lambda_2^r + \dots + \lambda_2^{n-1} \\ \vdots \\ \lambda_r + \lambda_r^r + \dots + \lambda_r^{n-1} \end{pmatrix}$$

Pre multiplying both sides by the matrix product

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_r \\ \vdots & \dots & \dots & \dots \\ \lambda_1^{r-1} & \lambda_2^{r-1} & \dots & \lambda_r^{r-1} \end{pmatrix} \begin{pmatrix} v_1 & 0 & \dots & 0 \\ 0 & v_2 & \dots & 0 \\ \vdots & \dots & \dots & \dots \\ \dots & \dots & \dots & v_r \end{pmatrix}$$

one obtains

$$\begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_r \\ \vdots & \dots & \dots \\ \lambda_1^{r-1} & \lambda_2^{r-1} & \lambda_r^{r-1} \end{pmatrix} \begin{pmatrix} v_1 & 0 & 0 \\ 0 & v_2 & 0 \\ \vdots & \dots & \dots \\ 0 & \dots & v_r \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{r-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_r & \dots & \lambda_r^{r-1} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{r-1} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_r \\ \vdots & \dots & \dots \\ \lambda_1^{r-1} & \lambda_2^{r-1} & \lambda_r^{r-1} \end{pmatrix} \begin{pmatrix} v_1 & 0 & 0 \\ 0 & v_2 & 0 \\ \vdots & \dots & \dots \\ 0 & 0 & v_r \end{pmatrix} \begin{pmatrix} \lambda_1 + \lambda_1^r + \dots + \lambda_1^{n-1} \\ \lambda_2 + \lambda_2^r + \dots + \lambda_2^{n-1} \\ \vdots & \dots & \dots \\ \lambda_r + \lambda_r^r + \dots + \lambda_r^{n-1} \end{pmatrix} \dots (1.18)$$

The matrix product

$$\begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_r \\ \vdots & \dots & \dots \\ \lambda_1^{r-1} & \lambda_2^{r-1} & \lambda_r^{r-1} \end{pmatrix} \begin{pmatrix} v_1 & 0 & 0 \\ 0 & v_2 & 0 \\ \vdots & \dots & \dots \\ 0 & 0 & v_r \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \lambda_1^{r-1} \\ 1 & \lambda_2 & \lambda_2^{r-1} \\ \vdots & \dots & \dots \\ 1 & \lambda_r & \lambda_r^{r-1} \end{pmatrix}$$

is precisely the $r \times r$ matrix

$$\begin{pmatrix} S_0 & S_1 & \cdots & S_{r-1} \\ S_1 & S_2 & \cdots & S_r \\ \vdots & \dots & \dots & \dots \\ S_{r-1} & S_r & \cdots & S_{2r-2} \end{pmatrix}$$

where $S_i = T_r(A^i) = \lambda_1^i + \lambda_2^i + \dots + \lambda_n^i$, $i = 1, 2, \dots, 2r-2$

$$= v_1 \lambda_1^i + v_2 \lambda_2^i + \dots + v_r \lambda_r^i$$

and the product on the right hand side of (1.18) is just equal to a $(r \times 1)$ column matrix

$$\begin{pmatrix} S_1 + S_r + \dots + S_{n-1} \\ S_2 + S_r + \dots + S_n \\ \vdots \\ S_r + S_{2r-1} + \dots + S_{n+r-2} \end{pmatrix}$$

The system (1.16) is thus completely equivalent to

$$S_0 c_0 + S_1 c_1 + \dots + S_{r-1} c_{r-1} = S_1 + S_r + \dots + S_{n-1}$$

$$S_1 c_0 + S_2 c_1 + \dots + S_r c_{r-1} = S_2 + S_{r+1} + \dots + S_n$$

$$\vdots$$

$$S_{r-1} c_0 + S_r c_1 + \dots + S_{2r-2} c_{r-1} = S_r + S_{2r-1} + \dots + S_{n+r-2}$$

Since the determinant of this system is non-zero, the quantities c_0, c_1, \dots, c_{r-1} can be computed in terms of the symmetric functions of the eigenvalues of A .

Combining theorem 1.14 with theorem 1.15 one concludes the

following:

THEOREM 1.16: The polynomials $P(A)$ defined by

$$P(\lambda_i) = \bar{\lambda}_i, \quad i = 1, \dots, n \dots\dots\dots(1.19)$$

are such that the matrix equations

$$SP(A) = A*S \quad (\text{or } A^T S \text{ in real case})$$

always admit a non-trivial solution S and the last row of every such solution can be chosen arbitrarily. Thus, in particular, a non-trivial solution S may be made non-singular by choosing the last row of S properly.

Example

Let

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{pmatrix}$$

be a matrix of order 3. Then the associated Hankel matrix

$$H = \begin{pmatrix} 3 & 4 & 6 \\ 4 & 6 & 10 \\ 6 & 10 & 18 \end{pmatrix}$$

is singular and therefore the eigenvalues of A are not all distinct. Again, since H has rank 2, by theorem 1.3, there are exactly two distinct eigenvalues of A . Hence $r = 2$. Following the above method of construction, it is seen that

$$P(A) = -A^2 + 4A - 2I$$

has the same characteristic equation as A and therefore the equation

$$SP(A) = A^T S$$

has a non-trivial solution. In fact, choosing $s_3 = (1, 1, 1)$ and computing s_1 and s_2 by using the recursive relation on page 37, it is seen that the non-singular matrix

$$S = \begin{pmatrix} 9 & -20 & 7 \\ -8 & 1 & -2 \\ 1 & 1 & 1 \end{pmatrix}$$

satisfies the equation.

Construction of a Positive Definite Solution

The purpose of this section is to show that the polynomial in (1.19) not only gives rise to a non-singular solution, but also to a positive definite solution, whenever it exists. For this, the following result will be established first:

LEMMA 1.1 Let A be an n -square normal matrix. Then A^* can be expressed as a scalar polynomial in A .

Proof: The result is a classical one, but proofs found in most of the text books (see, e.g. [16], Vol. I. p. 272) assume that the eigenvalues of A are known and the coefficients of the required polynomial are determined by them. It will be shown below how the polynomial can be computed without any prior knowledge of the eigenvalues of the given matrix. The only quantities required are the traces of several matrix products involving A and A^* .

Since A is normal, there exists a unitary matrix U such that

$$U^*AU = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A. Let

$$g(x) = c_0 + c_1x + c_2x^2 + \dots + c_{n-1}x^{n-1}$$

be the required polynomial, whose coefficients are determined by the following system of linear equations

$$c_0 + c_1\lambda_i + c_2\lambda_i^2 + \dots + c_{n-1}\lambda_i^{n-1} = \bar{\lambda}_i, \quad i = 1, \dots, n$$

In matrix notation, the system is equivalent to

$$\begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_n \\ \vdots & & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \lambda_n^{n-1} \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^{n-1} \\ \vdots & & \\ 1 & \lambda_n & \lambda_n^{n-1} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_n \\ \vdots & & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \lambda_n^{n-1} \end{pmatrix} \begin{pmatrix} \bar{\lambda}_1 \\ \bar{\lambda}_2 \\ \vdots \\ \bar{\lambda}_n \end{pmatrix}$$

The matrix product

$$\begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_n \\ \vdots & & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \lambda_n^{n-1} \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{n-1} \\ \vdots & & & \\ 1 & \lambda_n & \dots & \lambda_n^{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \vdots & & & \\ s_{n-1} & s_n & \dots & s_{2n-2} \end{pmatrix}$$

where S_i , $i = 0, \dots, 2n-2$ are given by

$$S_i = \text{Tr}(A^i) = \sum_{k=1}^n \lambda_k^i$$

Again, since A is normal, the eigenvalues of the matrix product $A^k A^*$ are just $\lambda_i^k \bar{\lambda}_i$, $i = 1, \dots, n$, $k = 0, 1, 2, \dots$. Thus, R. H. S. of the system is just equal to

$$\begin{pmatrix} \text{Tr}(A^*) \\ \text{Tr}(AA^*) \\ \vdots \\ \text{Tr}(A^{n-1}A^*) \end{pmatrix}$$

The system thus can be solved for the unknowns c_0, c_1, \dots, c_{n-1} by knowing only traces of the matrix products A^k and $A^k A^*$ ($k = 0, 1, \dots$). Now, since

$$D^* = \text{diag}(\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_n) = \text{diag}(g(\lambda_1), g(\lambda_2), \dots, g(\lambda_n)) = g(D)$$

one obtains

$$A^* = UD^*U^* = Ug(D)U^* = g(UDU^*) = g(A).$$

Turning back to the original question, it is noted that, when the given matrix A has its eigenvalues all distinct, it is similar to a diagonal matrix D by a non-singular similarity T . Since matrices A and D have the same eigenvalues, by the lemma 1.1, the polynomial $P(A)$ defined by (1.19) is such that

$$D^* = (TAT^{-1})^* = (T^*)^{-1}A^*T^* = P(TAT^{-1}) = T(P(A))T^{-1}.$$

Clearly, the matrix $S = (T^*T)^{-1}$ is positive definite and satisfies the equation (1.13).

CHAPTER II

QUADRATIC FORMS, MATRIX EQUATIONS AND THE LOCATION OF MATRIX EIGENVALUES

Let $f(x) = x^n - a_n x^{n-1} \dots - a_2 x - a_1$ be a polynomial of degree n and let it be required to find the numbers of zeros of $f(x)$ in some specified regions. This type of problem arises very often in various branches of engineering, in the field of mathematical economics and in the study of computational algorithms. The specified regions within which the zeros of $f(x)$ are to lie are normally half planes, the unit circle, sectors, ellipses etc. Associated with half planes are two famous classical problems: one, the problem of Hermite, is concerned with the determination of the number of zeros with positive imaginary parts and the other, due to Routh and Hurwitz concerns the number of zeros with negative real parts. The problem of determining the number of zeros in the unit circle is due to Schur and Cohn.

A differential equation is stable if the zeros of the characteristic polynomial are in the interior of the left half plane and similarly, a difference equation is stable if the zeros of the characteristic polynomial are in the interior of the unit circle. Thus, the Routh-Hurwitz problem arises in connection with the study of the stability of a system of differential equations and the Schur-Cohn problem arises in connection with that of a system of difference equations.

Now, $f(x)$ is the characteristic polynomial of the companion matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 1 \\ a_1 & a_2 & \dots & \dots & a_n \end{pmatrix} \dots\dots\dots(2.1)$$

whence the problem of determining the number of zeros of $f(x)$ in a specified region becomes the problem of finding the number of eigenvalues of A in the same region. The latter is the second aspect of the problem of separation of eigenvalues and, as mentioned in Chapter I, it can be solved by using the technique of matrix equations following an interesting method of Howland [23] published recently. It has been proved by Howland that, given an arbitrary matrix A and a polynomial

$$\phi(\lambda, \mu) = \sum_{k, \ell=0}^{n-1} C_{k\ell} \lambda^k \mu^\ell \dots\dots\dots(2.2)$$

in λ and μ such that $\phi(\lambda_j, \bar{\lambda}_j)$ is not pure imaginary for any eigenvalue λ_j of A , the matrix equation.

$$\sum_{k, \ell=0}^{n-1} C_{k\ell} A^k X A^{*\ell} + \sum_{k, \ell=0}^{n-1} \bar{C}_{k\ell} A^\ell X A^{*k} = S \dots\dots(2.3)$$

where S is a given positive definite Hermitian matrix, always admits a Hermitian solution X , the inertia of which specifies the number of eigenvalues of A lying within and outside a specified region. Thus, if π and ν denote the number of positive and negative squares respectively in a normal representation of the solution matrix X , then the required information is supplied by the integers π and ν . Ordinarily X is a nonsingular and $\pi + \nu = n$. Thus, in view of the existence of Howland's method, all the classical stability problems can be

formulated in terms of the matrix equations for the companion matrix A , obtained as special cases of the general matrix equation (2.3).

The application of Howland's method in solving these problems, however, demands that the matrix equations actually be solved. Once a hermitian solution X has been obtained, its signature may be computed by studying the signs of principal minors, according to a method recently described by Howland [27]. The problem of solving these equations is not a trivial one. The purpose of this chapter is to give a brief account of some existing methods of solving the classical stability problems and to propose practical methods of considerable simplicity to solve the associated matrix equations numerically. The proposed methods seem to be closely related to those of Jameson [30] and Molinari [57].

2.1 THE EIGENVALUES IN A HALF PLANE

2.1.1 The Hermite Problem

The Hermite Problem is to find the number of zeros of a given polynomial $f(x)$ in the upper half plane and in particular to derive a necessary and sufficient condition that all the zeros have positive imaginary parts.

(a) Hermitian Forms and the Hermite Problem

(i) Hermite's Method of Solution

To solve the Problem, Hermite [17] considered the Bezoutiant

$$K(f) = -i \frac{f(x) \bar{f}(y) - f(y) \bar{f}(x)}{x - y} = \sum_{k, \ell=0}^{n-1} A_{k, \ell} x^k y^\ell$$

where $\bar{f}(x) = x^n - \bar{a}_n x^{n-1} \dots - \bar{a}_2 x - \bar{a}_1$, a polynomial whose coefficients are the complex conjugates of those of $f(x)$. From this Bezoutiant, Hermite constructed the hermitian form,

$$H(f) = \sum A_{k, \ell} u_k u_\ell$$

whose associated matrix $A_{k, \ell}$ is real, and proved the following theorem:

THEOREM 2.1 (Hermite): If the rank of the Hermitian form H is n , then the number of roots of $f(x) = 0$ with positive imaginary parts and of those with negative imaginary parts are respectively equal to the number of positive and negative squares in a normal representation of H . The roots all have positive imaginary parts if and only if H is positive definite.

Different proofs of the Hermite Theorem are available in the literature (see, e.g., the Survey [31] of Krein and Naimark). Among these, an elegant way of solving the Hermite problem is due to Fujiwara [15] who discovered that a method of Lienard and Chipart proposed for the solution of the Routh-Hurwitz problem can be successfully applied to the problem of Hermite, to the Schur-Cohn problem and to some other problems of similar nature. Using this method, Fujiwara gave an independent proof of the Hermite Theorem. A proof of the theorem slightly different from that of Fujiwara has recently been given by Householder [20].

(ii) Bigradients and the Hermite Problem

Let

$$a_0, a_1, a_2, \dots,$$

$$b_0, b_1, b_2, \dots,$$

be two sequences, either of which may be finite or infinite (if finite it will be assumed extended by zeros). Then the quantities

$$\delta \begin{pmatrix} (a)_i \\ (b)_j \end{pmatrix} = \delta \begin{pmatrix} a_0 & a_1 & & & a_{i+j-1} \\ 0 & a_0 & a_1 & \cdots & a_{i+j-2} \\ 0 & 0 & a_0 & \cdots & a_{i+j-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & b_0 & b_1 & \cdots & b_{i+j-2} \\ b_0 & b_1 & b_2 & \cdots & b_{i+j-1} \end{pmatrix} \dots\dots\dots(2.4)$$

where δ signifies the determinant and there are i rows made up of the a 's and j rows of the b 's, are called Bigradients. Recently, it has been demonstrated by Householder [19] that bigradients may be used in solving the Routh-Hurwitz as well as the Hermite-problem. The following result has been proved:

Let $f(z) = z^n + a_1 z^{n-1} + \dots + a_n = f_0'(z) + i f_1(z)$ be a polynomial of degree n and let $a_j = a_j' + i a_j''$. Then the number of zeros of $f(z)$ above the real axis and that below it are respectively equal to the number of variations and permanences of sign in the sequence

$$1, a_1'', \delta \begin{pmatrix} (a')^1 \\ (a'')^2 \end{pmatrix}, \dots, \delta \begin{pmatrix} (a')^{n-1} \\ (a'')^n \end{pmatrix}, \dots\dots\dots(2.5)$$

provided no two consecutive terms vanish.

(b) Matrix Equation and the Hermite-Problem

The matrix equation that solves the Hermite-problem can be derived from Howland's equation (2.3) by setting $\phi(\lambda, \mu) = i\lambda$

and $S = I$, an identity matrix, whence the equation reduces to

$$iAX - iXA^* = I \dots\dots\dots(2.6)$$

The matrix A is in the form (2.1) and if π and ν are the number of positive squares and the number of negative squares respectively in a normal representation of a hermitian solution X of the equation (2.6) then π is just the number of zeros of $f(x)$ in the lower half plane and ν is the number in the upper half plane, assuming that $\pi + \nu = n$ and $f(x)$ does not have any real zero.

Existence of Solution (A constructive Proof): Let x_1, x_2, \dots, x_n be the row-vectors of the solution matrix X . Then the successive rows of AX are x_2, x_3, \dots, x_n and $a_1 x_1 + a_2 x_2 + \dots + a_n x_n$ and the successive rows of XA^* are $x_1 A^*, x_2 A^*, \dots$ and $x_n A^*$. The matrix equation (2.6) is therefore equivalent to

$$\begin{cases} i x_k - i x_{k-1} A^* = e_{k-1}, k = 2, 3, \dots, n \\ i(a_1 x_1 + a_2 x_2 + \dots + a_n x_n) - i x_n A^* = e_n \end{cases} \dots\dots(2.7)$$

where e_k is the k th row vector of the identity matrix I . Eliminating x_2, \dots, x_n , the relations (2.7) can be written

$$\begin{aligned} -i x_1 (f(A^*)) &= e_1 (A^{*n-1} - a_n A^{*n-2} \dots - a_2) + \dots \\ e_2 (A^{*n-2} - a_n A^{*n-3} \dots - a_3) &+ \dots + e_{n-1} (A^* - a_n) + e_n \end{aligned}$$

This gives rise to the following system of n non-homogeneous equations in n unknowns

$$-i f(\bar{A}) x_1^T = c \dots\dots\dots(2.8)$$

where the vector c is given by

$$c = (\bar{A}^{n-1} - a_n \bar{A}^{n-2} \dots - a_2) e_1^T + (\bar{A}^{n-2} - a_n \bar{A}^{n-3} \dots - a_3) e_2^T \\ + \dots + (\bar{A} - a_n) e_{n-1}^T + e_n^T.$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A . Then the eigenvalues of $f(\bar{A})$ are $f(\bar{\lambda}_1), f(\bar{\lambda}_2), \dots, f(\bar{\lambda}_n)$. So, the determinant of $f(\bar{A})$ is

$$\det[f(\bar{A})] = \prod_{i=1}^n f(\bar{\lambda}_i).$$

$$\text{Now } f(\bar{\lambda}_i) = (\bar{\lambda}_i - \lambda_1)(\bar{\lambda}_i - \lambda_2) \dots (\bar{\lambda}_i - \lambda_n).$$

$$\text{So, } \det[f(\bar{A})] = \prod_{i,j=1}^n (\bar{\lambda}_i - \lambda_j). \text{ The expression } \prod_{i,j=1}^n (\bar{\lambda}_i - \lambda_j)$$

is defined as the Resultant of the polynomials f and \bar{f} and is denoted by $R(f, \bar{f})$. Since the system has a unique solution if and only if $f(\bar{A})$ is non-singular, one immediately gets the following theorem:

THEOREM 2.2: Let A be a companion matrix of the form (2.1) and let $\lambda_1, \lambda_2, \dots, \lambda_n$ be its eigenvalues. Then the matrix equation (2.6) has a unique Hermitian solution X (the unique solution is Hermitian, because X^* is also a solution of the equation whenever X is) if and only if

$$\bar{\lambda}_i \neq \lambda_j \text{ for any } i \text{ and } j.^1$$

Note. The condition in Theorem 2.2 for the existence of the unique solution of the equation (2.6) is completely equivalent to the one derived by Howland [23]. For, according to Howland, the equation

¹ The result is well known ([32] or [34]), but the proof is new and constructive.

if

$$P(A) = \prod_{k, j=1}^n (\phi(\lambda_k, \bar{\lambda}_j) + \bar{\phi}(\lambda_j, \bar{\lambda}_k))$$

is non-zero. Since in this case

$$P(A) = \prod_{k, j=1}^n (\lambda_k - \bar{\lambda}_j)$$

it follows that $P(A)$ is non-zero if and only if

$$\lambda_k \neq \bar{\lambda}_j \text{ for any } k \text{ and } j.$$

Construction of Solution and other Computational Aspects:

The unique solution, whenever it exists, can be computed by solving first the non-homogeneous non-singular system (2.8) for x_1 and then computing the remaining row vectors $x_2 \dots x_n$ recursively from the relations (2.7). The coefficient matrix $(-i f(\bar{A}))$ and the vector c of the system can be computed by using the algorithm of Barnett [6] to be described in the following section.

Example:

Let $f(\lambda) = \lambda^3 + 3\lambda - 2i$ be a given polynomial of degree 3.

$$(a_1 = 2i, a_2 = -3, a_3 = 0)$$

Then

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2i & -3 & 0 \end{pmatrix}$$

and the matrix equation (2.6) gives rise to the following system of equations

$$\begin{pmatrix} -4 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \end{pmatrix} = \begin{pmatrix} 4 \\ -2i \\ -2 \end{pmatrix}$$

where $x_1 = (x_{11}, x_{12}, x_{13})$ is the first row vector of X . Since the system is non-singular, it has a unique solution

$$x_{11} = -1$$

$$x_{12} = \frac{i}{2}$$

$$x_{13} = \frac{1}{2}$$

The remaining two rows x_2 and x_3 are given by

$$\begin{aligned} i x_2 &= (1, 0, 0) + (-i, -\frac{1}{2}, \frac{1}{2}i) \begin{pmatrix} 0 & 0 & -2i \\ 1 & 0 & -3 \\ 0 & 1 & 0 \end{pmatrix} \\ &= (\frac{1}{2}, \frac{1}{2}i, -\frac{1}{2}) \end{aligned}$$

whence

$$x_2 = (-\frac{1}{2}i, \frac{1}{2}, \frac{1}{2}i)$$

$$i x_3 = e_2 + i x_2 A^* = (0, 1, 0) + (\frac{1}{2}, \frac{1}{2}i, -\frac{1}{2}) \begin{pmatrix} 0 & 0 & -2i \\ 1 & 0 & -3 \\ 0 & 1 & 0 \end{pmatrix}$$

$$= (0, 1, 0) + (\frac{1}{2}i, -\frac{1}{2}, -\frac{5i}{2})$$

$$= (\frac{1}{2}i, \frac{1}{2}, -\frac{5i}{2})$$

whence

$$x_3 = (\frac{1}{2}, -\frac{1}{2}i, -\frac{5}{2})$$

so, the unique hermitian solution X of the equation is

$$X = \begin{pmatrix} -1 & \frac{i}{2} & \frac{1}{2} \\ -\frac{i}{2} & \frac{1}{2} & \frac{1}{2}i \\ \frac{1}{2} & -\frac{i}{2} & -\frac{5}{2} \end{pmatrix}$$

The leading principal minors of X are

$$D_1 = -1, D_2 = -\frac{3}{4}, D_3 = \frac{7}{4}$$

Since,

$$V(1, -1, -\frac{3}{4}, \frac{7}{4}) = 2$$

and

$$P(1, -1, -\frac{3}{4}, \frac{7}{4}) = 1$$

by the theorem of Jacobi (p. 8) the number of zeros of $f(\lambda)$ with positive imaginary parts is 2 and that with negative imaginary parts is 1.

(c) Barnett's Method of Solution

The following result has been recently obtained by Barnett [6], using an approach entirely different from the two just described.

Let $f(\lambda) = f_0(\lambda) + i f_1(\lambda)$ be a polynomial of degree n and let $R = f_1(F)$ where F is the companion matrix of f_0 .

THEOREM 2.3 (Barnett): $f(\lambda)$ has p zeros with positive imaginary parts and $n-p$ zeros with negative imaginary parts, where p is the number of variations of sign in the sequence

$$1, \epsilon_1 R_1, \epsilon_2 R_2, \dots, \epsilon_n R_n$$

where $\epsilon_k = (-1)^{\binom{k}{2}} k(k-1)$ and R_k is the minor of order $k \times k$ formed from the first k rows and last k columns of R ; under the assumption that none of R_1, R_2, \dots, R_n is zero.

An algorithm has also been proposed by Barnett to construct the polynomial matrix R . The most distinctive feature of the algorithm is that it does not involve the computation of higher powers of the given matrix. The algorithm is stated as follows:

Let

$$\begin{aligned}\phi(x) &= a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n \quad (a_0 = 1) \\ \psi(x) &= b_0 x^m + b_1 x^{m-1} + \dots + b_{n-1} x + b_m\end{aligned}$$

be two polynomials and let the companion matrix of $\phi(x)$ be

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{pmatrix}$$

Then, when $m < n$, the rows of $\psi(A)$ are given by

$$\begin{aligned}r_1 &= (b_m, \dots, b_1, b_0, 1, 0, \dots, 0) \\ r_2 &= r_1 A \\ r_3 &= r_2 A \\ &\vdots \\ r_n &= r_{n-1} A\end{aligned}$$

and when $m = n$, r_i , $i = 2, \dots, n$ are the same as above, but

$$r_1 = (d_n, d_{n-1}, \dots, d_1), \text{ where}$$

$$d_i = b_i - b_0 a_i$$

2.1.2 The Routh-Hurwitz Problem

The classical problem of Routh and Hurwitz is that of finding the number of zeros of a given polynomial with negative real parts and in particular of obtaining a necessary and sufficient condition that all the zeros lie in the left half plane ($\text{Re}\{z\} < 0$).

(a) Hermitian Forms and the Routh-Hurwitz Problem

(i) Routh's Algorithm

The British mathematician Routh gave a very simple algorithm for determining k , the number of zeros of a real polynomial in the right half plane ($\text{Re}\{z\} > 0$), with the help of Sturm's theorem and the theory of Cauchy indices. The algorithm can be stated as follows: (adapted from [16] vol. II, p. 178-180).

Given a real polynomial

$$f(x) = a_0 x^n + b_0 x^{n-1} + a_1 x^{n-2} + b_1 x^{n-3} + \dots$$

of degree n , the following scheme, known as Routh's Scheme, is formed:

$$\begin{array}{cccc} a_0 & a_1 & a_2 & \dots \\ b_0 & b_1 & b_2 & \dots \\ c_0 & c_1 & c_2 & \dots \\ d_0 & d_1 & d_2 & \dots \\ \vdots & \vdots & \vdots & \vdots \end{array}$$

The quantities c_0, d_0, c_1, d_1 etc can be computed from the coefficients of the given polynomial, as follows:

$$c_0 = a_1 - \frac{a_0}{b_0} b_1, \quad c_1 = a_2 - \frac{a_0}{b_0} b_2 \dots$$

$$d_0 = b_1 - \frac{b_0}{c_0} c_1, \quad d_1 = b_2 - \frac{b_0}{c_0} c_2 \dots$$

Routh's theorem then states that the number of zeros of $f(x)$ lying in the right half plane is equal to the number of variations of sign in the first column of the Routh-scheme and all the zeros have negative real parts if and only if all the elements of the first column of Routh's scheme are different from zero and of same sign.

(ii) Hurwitz's Criterion of Stability

Routh obtained the criterion of stability, unaware of the investigation of Hurwitz, who obtained it in a more beautiful form. Hurwitz's Criterion of Stability can be stated as follows:

Let $f(z) = a_0 z^n + b_0 z^{n-1} + a_1 z^{n-2} + b_1 z^{n-3} + \dots$ be a polynomial of degree n and let the quantities $\Delta_i, i = 1, \dots, n$ be defined as follows:

$$\Delta_1 = b_0, \quad \Delta_2 = \begin{vmatrix} b_0 & b_1 \\ a_0 & a_1 \end{vmatrix}, \quad \dots \quad \Delta_n = \begin{vmatrix} b_0 & b_1 & \dots & b_{n-1} \\ a_0 & a_1 & \dots & a_{n-1} \\ 0 & b_0 & \dots & b_{n-2} \\ 0 & a_0 & \dots & a_{n-2} \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

$$\begin{cases} a_k = 0 \text{ for } k > \lfloor \frac{n}{2} \rfloor \\ b_k = 0 \text{ for } k > \lfloor \frac{n-1}{2} \rfloor \end{cases}$$

Then all the zeros of $f(z)$ have negative real parts if and only if the quantities $\Delta_i, i = 1, \dots, n$ are all positive. The quantities

Δ_i are commonly known as the Hurwitz determinants.

(iii) Equivalence

The equivalence between Hurwitz's Criterion of stability and that of Routh has been established in the literature ([16], vol. II, p. 190-195). In fact, it has been shown that a lower triangular matrix of order n , obtained from the coefficients of Routh's scheme, usually known as Routh-matrix

$$R = \begin{pmatrix} b_0 & b_1 & b_2 & \dots \\ 0 & c_0 & c_1 & \dots \\ 0 & 0 & d_0 & \dots \end{pmatrix}$$

is equivalent to the Hurwitz-matrix

$$H = \begin{pmatrix} b_0 & b_1 & b_2 & \dots & b_{n-1} \\ a_0 & a_1 & a_2 & \dots & a_{n-1} \\ 0 & b_0 & b_1 & \dots & b_{n-2} \\ 0 & a_0 & a_1 & \dots & a_{n-2} \\ 0 & 0 & b_0 & \dots & b_{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \begin{cases} a_k = 0 \text{ for } k > \lfloor \frac{n}{2} \rfloor \\ b_k = 0 \text{ for } k > \lfloor \frac{n-1}{2} \rfloor \end{cases}$$

in the sense that for every $p \leq n$, the corresponding minors of order p in the first p -rows are equal.

This equivalence enables one to reformulate Routh's theorem in the form of the following theorem, known as the Routh-Hurwitz theorem in the literature (adapted from [16], vol. II, p. 194).

THEOREM 2.4 (Routh and Hurwitz): The number k of the roots of the

polynomial equation $f(z) = 0$ in the right half plane ($\text{Re} z > 0$) is equal to the number of variations of sign in the sequence

$$a_0, \Delta_1, \frac{\Delta_2}{\Delta_1}, \dots, \frac{\Delta_n}{\Delta_{n-1}}$$

where $\Delta_i, i = 1, \dots, n$ are Hurwitz-determinants.

(iv) Bezoutiants and the Routh-Hurwitz Problem

It has been shown by Fujiwara [15] that, like Hurwitz's problem, the problem of Routh and Hurwitz can also be handled with the help of Bezoutiants. The procedure is similar to that for the Hermite problem.

Given a polynomial $f(x)$ of degree n with arbitrary real or complex coefficients, a polynomial $f^*(x) = \bar{f}(-x)$ is chosen first and the Bezoutiant

$$\begin{aligned} K(f) &= \frac{f(x) f^*(y) - f(y) f^*(x)}{x - y} \\ &= \sum_{i, k=0}^{n-1} b_{i, k} x^i y^k \end{aligned}$$

associated with $f(x)$ and $f^*(x)$ is then constructed. From the matrix $B = (b_{i, k})$, the Hermitian matrix

$$F = (f_{i, k}), \text{ where}$$

$$f_{i, k} = (-1)^{n-1-i} b_{i, k}$$

is formed, whence the theorem follows:

THEOREM 2.5 (Fujiwara): If π and ν are the numbers of positive and negative squares in a normal representation of the Fujiwara matrix $F = (f_{i, k})$, then π and ν are respectively equal to the number of

zeros of $f(x)$ with negative and positive real parts; under the assumption that the matrix F is non-singular.

(b) Matrix Equations and the Routh-Hurwitz Problem

The Routh-Hurwitz Problem gives rise to the well-known Lyapunov-equation

$$AX + XA^* = -I \dots\dots\dots(2.9)$$

where A is a companion matrix of the form (2.1) and I is the identity matrix. If A has no pure imaginary eigenvalues and if

$\text{In}(X) = (\pi, \nu, \delta)$, then π is the number of eigenvalues of A in the left hand plane and ν is the number in the right. The Lyapunov equation can be derived as a special case of Howland's equation by setting

$$\phi(\lambda, \mu) = -\lambda.$$

In case the given matrix A is real, the matrix equation reduces to

$$AX + XA^T = -I \dots\dots\dots(2.10)$$

This matrix equation plays an important role in the theory of stability of a system of differential equations and was first formulated by A. M. Lyapunov (1892) in his manuscript 'The General Problem of Stability of Motion'. A fundamental result about the equation is: (Adapted from [16], vol. II, p. 189).

THEOREM 2.6 (Lyapunov): The matrix equation (2.10) admits a positive definite solution X if and only if all the eigenvalues of A have negative real parts.

Methods of Solution

Since the Lyapunov equation plays an important role in the theory of stability, the equation has been studied by various authors in the past, not only by mathematicians but by the scientists of many other branches of Applied Science also. As a result of this, numerous papers have been published on the solution of this equation. In a series of interesting papers, Barnett and Storey [1] - [3] have studied the equation rigorously. They have also solved the equation for the discrete case satisfactorily [50]. Among the recent publications works of Barnett [4], of Davison and Man [10], of Howland and Senez [27], of Man [35], of Muller [37] etc. deserve to be mentioned. While most of these methods, like all other classical methods of solution involve the difficulties of Series Solution, the constructive procedure of Howland and Senez does avoid it completely. This method, designed for an arbitrary matrix A demands that the given matrix A be first transformed into an upper Hessenberg form. Since in the present case, the given companion matrix A is itself the transpose of an upper Hessenberg matrix, the method of Howland and Senez solves the Routh-Hurwitz problem with less difficulty. Another constructive method of solving the Lyapunov matrix equation, similar to that of the matrix equation associated with the Hermite problem will be proposed here. The method of Howland and Senez and the one to be proposed may be described as follows:

(i) Method of Howland and Senez

Matrices X_j , T_j and D_j for $1 \leq j \leq n$ are constructed such that X_j is symmetric, T_j is skew-symmetric and D_j is diagonal and they satisfy the relation

$$X_j A = T_j + D_j$$

If the first row and column of X_j are chosen arbitrarily, then the

equation defines explicitly the remaining elements of X_j and hence those of D_j . The required solution matrix X of the equation (2.10) is then given by

$$X = \sum p_j X_j$$

where the quantities p_j are determined by solving an $n \times n$ linear system

$$\sum p_j D_j = -\frac{I}{2}$$

The first rows of X_j can be conveniently chosen, in turn as the rows of the identity matrix I .

(ii) Proposed Method

Let x_i , $i = 1, \dots, n$ be the n rows of the matrix X . Then the Lyapunov equation (2.9) reduces to

$$\left. \begin{aligned} x_{i-1} A^* + x_i &= -e_{i-1} & i = 2, \dots, n \\ (a_1 x_1 + \dots + a_n x_n) + x_n A^* &= -e_n \end{aligned} \right\} \dots\dots\dots(2.11)$$

where e_i , $i = 1, \dots, n$ are the rows of the identity matrix I . These equations give rise to a system of equations of the form

$$\begin{aligned} -\psi(\bar{A})x_1^T &= -[f_{n-1}(-\bar{A})]e_1^T - [f_{n-2}(-\bar{A})]e_2^T \dots \\ &\dots - [f_1(-\bar{A})]e_{n-1}^T - e_n^T = b \dots\dots\dots(2.12) \end{aligned}$$

where

$$f_r(x) = x^r - a_n x^{r-1} \dots - a_{n-r+2} x - a_{n-r+1}$$

$$\psi(x) = f_n(-x) = f(-x)$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A . Then the eigenvalues of $\psi(\bar{A})$ are $\psi(\bar{\lambda}_1), \psi(\bar{\lambda}_2) \dots \psi(\bar{\lambda}_n)$. Since the zeros of $\psi(x)$ are just the negatives of the zeros of $f(x)$,

$$\psi(\bar{\lambda}_i) = (-1)^n (\bar{\lambda}_i + \lambda_1)(\bar{\lambda}_i + \lambda_2) \dots (\bar{\lambda}_i + \lambda_n).$$

Thus, the determinant of the coefficient matrix of the system is

$$\det[-\psi(\bar{A})] = (-1)^{n+n} \prod_{i, j=1}^{2n} (\bar{\lambda}_i + \lambda_j)$$

Since the system has a non-trivial unique solution if and only if $-\psi(\bar{A})$ is non-singular, one immediately obtains the following theorem.

THEOREM 2.7: The Lyapunov equation (2.9) admits an unique Hermitian solution X if and only if

$$\bar{\lambda}_i + \lambda_j \neq 0 \text{ for any } i \text{ and } j.$$

Construction of Solution and other Computational Aspects:

The first row x_1 of the solution matrix X is first determined by solving the non-homogeneous non-singular system (2.12) and then the remaining row vectors from the recursive relations (2.11) in terms of x_1 . The coefficient matrix $-\psi(\bar{A})$ and the vector b on the right hand side of the system may be computed by using the algorithm of Barnett [6].

Example

Let

$$f(x) = x^3 - 3x^2 + 4x - 2$$

be a polynomial of degree 3.

Then

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -4 & 3 \end{pmatrix}$$

and the Lyapunov equation (2.10) gives rise to the following system

of equations:

$$4x_{11} + 6x_{13} = -3$$

$$12x_{11} + 18x_{13} - 20x_{12} = 1$$

$$36x_{11} - 60x_{12} + 34x_{13} = -17$$

where $x_1 = (x_{11}, x_{12}, x_{13})$ is the first row vector of X .

Since the system is non-singular, it has a unique solution

$$x_1 = \left(-\frac{9}{4}, -\frac{1}{2}, 1\right)$$

The other two rows of X , x_2 and x_3 are given by

$$x_2 = -e_1 - x_1 A^* = -(1, 0, 0) - \left(-\frac{9}{4}, -\frac{1}{2}, 1\right) \times$$

$$\begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & -4 \\ 0 & 1 & 3 \end{pmatrix} = \left(-\frac{1}{2}, -1, -\frac{1}{2}\right)$$

$$x_3 = -e_2 - x_2 A^* = -(0, 1, 0) + \left(\frac{1}{2}, 1, \frac{1}{2}\right) \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & -4 \\ 0 & 1 & 3 \end{pmatrix}$$

$$= (0, -1, 0) + \left(1, \frac{1}{2}, -\frac{3}{2}\right)$$

$$= \left(1, -\frac{1}{2}, -\frac{3}{2}\right)$$

So, the unique symmetric solution X of the equation is

$$X = \begin{pmatrix} -\frac{9}{4} & -\frac{1}{2} & 1 \\ -\frac{1}{2} & -1 & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -\frac{3}{2} \end{pmatrix}$$

The leading principal minors of X are

$$D_1 = -\frac{9}{4}, D_2 = \frac{9}{4} - \frac{1}{4} = 2, D_3 = -\frac{15}{16}$$

So, by the theorem of Jacobi, (p. 8), the number of zeros of $f(x)$ with positive real parts = $V(1, -\frac{9}{4}, 2, -\frac{15}{16}) = 3$.

(c) Matrix Equation $AX + XB = -C$

It is interesting to note that the Lyapunov matrix equation is just a special case of the general equation

$$AX + XB = -C \dots \dots \dots (2.13)$$

where A, B and C are known matrices and X is the matrix to be found. The equation 2.13 in its general form arises in various problems of Physics and Mechanics. Specifically, knowledge of its solution is applicable to the numerical solutions of certain boundary value problems in partial differential equations and those of certain linear ordinary differential equation systems [37].

The equation has been studied widely and both algebraic and numerical methods of solving the equations are known to exist. Among the various algebraic solutions, some use the fact that the Jordan forms of A and B are known [40] and some requires the inversion of a matrix of large order [30]. Most of the available numerical procedures are iterative and involve the difficulties of series solutions. Moreover, in many cases, solution X can be determined only in special cases when A and B are both stability (Hurwitzian) matrices [41] (A real (complex) matrix A is said to be a stability (Hurwitzian) matrix if all its eigenvalues have negative real parts). An explicit form of the unique solution ([33], p. 263) is

$$X = \int_0^{\infty} e^{At} C e^{Bt} dt$$

provided that the integral on the right hand side exists for all C . A numerical procedure of evaluating the integral in the special case, when A , B and C are all real and $B = A^T$ (Lyapunov equation) has been recently devised by Man [35]. It will be shown, in this section that the constructive method proposed to solve the Lyapunov equation may as well be applied to solve this general equation, in case the (non derogatory) matrix A is given in its rational canonical form.

If A is in the form (2.1), then the matrix equation 2.13 is equivalent to a set of equations:

$$x_{i+1} + x_i B = -c_i, \quad i = 1, 2, \dots, n-1$$

$$(a_1 x_1 + a_2 x_2 + \dots + a_n x_n) + x_n B = -c_n \dots\dots(2.14)$$

where x_i , $i = 1, \dots, n$ are the rows of X and c_1, c_2, \dots, c_n are the rows of C . These equations give rise to a system of equations of the form:

$$\phi(-B^T) x_1^T = c' \dots\dots\dots(2.15)$$

where

$$c' = c_n^T - [B^T + a_n] c_{n-1}^T + [(B^2)^T + a_n B^T - a_{n-1}] c_{n-2}^T + \dots$$

$$+ \dots + (-1)^{n-1} [(B^{n-1})^T + a_n (B^{n-2})^T - a_{n-1} (B^{n-3})^T + \dots$$

$$+ (-1)^n a_2] c_1^T$$

Thus, the first row vector x_1 can be computed by solving (2.15) whenever the solution exists and then the remaining row vectors x_i , $i = 2, \dots, n$ be obtained from the recursive relation (2.14).

Lappo-Danilevsky Equation

A further special case of the equation (2.13) is

$$AX - XA = \lambda X \dots\dots\dots(2.16)$$

where λ is a parameter. The equation was first formulated by Lappo-Danilevsky and has some role to play in the solution of differential equations. The criterion of the existence of a non-trivial solution was first given by Lappo-Danilevsky and has been reported by Bellman ([7], p. 244). The result has been proven recently by Neudecker [38] in a different fashion. It is, however, clear that the technique just proposed for solving the equation $AX+XB=-C$, when applied to this special case, yield the result in a constructive way. The matrix A , as before, will be assumed to be given in its rational canonical form.

The equation (2.16) when arranged in the form (2.13) becomes

$$AX - (A + \lambda I)X = 0$$

substituting $B = -(A + \lambda I)$ in (2.15), the equation is seen to be equivalent to the system of homogeneous equations:

$$\phi(A^T + \lambda)x_1^T = 0, \text{ where } \phi(x) \text{ is the characteristic polynomial of } A.$$

Since

$$\det[\phi(A^T + \lambda)] = \pm \prod_{i,j=1}^n (\alpha_i - \alpha_j - \lambda)$$

where $\alpha_1, \alpha_2, \dots, \alpha_n$ are the eigenvalues of A , and one obtains the following:

THEOREM (Lappo-Danilevsky): The matrix equation (2.16) with A a companion matrix has a non trivial solution if and only if

$\alpha_i = \alpha_j + \lambda$ for some i and j , where α_i , $i = 1, \dots, n$ are the eigenvalues of A .

(d) Connections between the Lyapunov-Equation and Hermitian Forms

Connections between the Lyapunov-equation and various other existing methods of solving the Routh-Hurwitz Problem, using Hermitian forms, have been recently established by Howland [26]. It has been noted that the Fujiwara matrix F and the solution X of the Lyapunov-equation have the same property in the sense that the inertia of both the matrices supply information on the number of zeros of $f(x)$ in the right and left half planes. It is therefore natural to ask if the Fujiwara matrix F is a solution of the Lyapunov equation. It has been proved by Howland that the matrix F does not satisfy the Lyapunov-equation but is a solution of the matrix equation

$$FA + A^*F = -G$$

where G is a Hermitian matrix of rank 1 and signature 1. In addition, it has been proved that the Fujiwara matrix F is diagonally congruent to the Routh-matrix R by the relation

$$D^*RD = F/2$$

where $D = \text{dg}((i)^k)$, so that R satisfies the matrix equation

$$R(DAD^{-1}) + (D^{*-1}A^*D^*)R = \frac{-(D^*)^{-1}GD^{-1}}{2} \dots\dots\dots(2.17)$$

where the matrix on the right hand side has both rank and signature 1. Thus: The substance of Howland's result is that the methods of Routh, Hurwitz, and others may all be described as applications of

the method of hermitian forms, where the matrices of the forms in question may be defined by matrix equations.

2.2 THE EIGENVALUES IN THE UNIT CIRCLE

2.2.1 The Schur-Cohn Problem

The Schur-Cohn Problem consists in finding the number of zeros of a polynomial $f(x)$ inside the unit circle and in particular in obtaining a necessary and sufficient condition for all the zeros to lie within it.

(a) Hermitian Forms and the Schur-Cohn Problem

(i) Schur-Cohn Criterion

Let $f(x) = x^n - a_n x^{n-1} \dots - a_1$ be a polynomial of degree n . Then using the Hermitian form

$$H = \sum_{j=1}^n |u_j - \bar{a}_n u_{j+1} - \dots - a_{j+1} u_n|^2 - \sum_{j=1}^n |-a_1 u_j - a_2 u_{j+1} - \dots - a_{n-j+1} u_n|^2 \quad (a_{n+1} = 1)$$

Schur and Cohn obtained the following result (adapted from [36], p. 152).

THEOREM 2.8 (Schur-Cohn): If for the polynomial $f(x)$, the determinants

$$\Delta_k = \begin{vmatrix} -a_1 & 0 & \dots & 0 & 1 & -a_n & \dots & -a_{n-k+2} \\ -a_2 & -a_1 & \dots & 0 & 0 & 1 & \dots & -a_{n-k+3} \\ \vdots & \vdots & & & & & & \\ -a_k & -a_{k-1} & \dots & a_1 & 0 & 0 & \dots & 1 \\ 1 & 0 & \dots & 0 & -\bar{a}_1 & -\bar{a}_2 & \dots & -\bar{a}_k \\ -\bar{a}_n & 1 & \dots & 0 & 0 & -\bar{a}_1 & \dots & -\bar{a}_{k-1} \\ \vdots & & & & & & & \\ -\bar{a}_{n-k+2} & \dots & \dots & 1 & 0 & 0 & \dots & -\bar{a}_1 \end{vmatrix} \quad (a_{n+1} = 1)$$

(k = 1, \dots, n)

are all different from zero, then $f(x)$ has no zero on the circle $|z| = 1$ and p zeros in this circle, p being the number of variations of sign in the sequence $1, \Delta_1, \Delta_2, \dots, \Delta_n$.

The determinants Δ_i ($i = 1, \dots, n$) are commonly termed the Schur-Cohn determinants.

(ii) Fujiwara's Method of Solution (Use of the Bezoutiant)

Fujiwara [15] showed that, like the other two problems, the Schur-Cohn problem can be solved by using Bezoutiants.

Let

$$f^*(x) = 1 - \bar{a}_n x - \bar{a}_{n-1} x^2 \dots - \bar{a}_2 x^{n-1} - a_1 x^n$$

and let

$$K(f) = \frac{f(x) f^*(y) - f(y) f^*(x)}{x - y} = \sum_{i, k=0}^{n-1} A_{i, k} x^i y^k$$

be the associated Bezoutiant of $f(x)$ and $f^*(x)$, so that

$\bar{A}_{i, k} = A_{n-1-i, n-k}$. From the Bezoutiant $K(f)$ the Hermitian form

$$H(f) = \sum a_{i, k} u_i \bar{u}_k$$

may be constructed by setting $A_{i, n-1-k} = a_{i, k}$ and replacing x^i by u_i and y^{n-1-k} by \bar{u}_k . Then Fujiwara's result follows:

THEOREM 2.9 (Fujiwara): Let π and π' be the number of positive and negative squares in a normal representation of $H(f)$ and let $\pi + \pi' = n$. Then the number of zeros of $f(x)$ inside the unit circle is π and the number outside it is π' . All the zeros are inside this circle if and only if $H(f)$ is positive definite.

(b) Matrix Equations and the Schur-Cohn Problem

The matrix equation associated with the Schur-Cohn problem is

$$AXA^* - X = I \dots\dots\dots(2.18)$$

where A is in the form (2.1). If $I_n(X) = (\pi, \nu)$ then π is the number of zeros of $f(x)$ outside the circle $|z| = 1$ and ν is the number inside it; under the assumption that no zeros of $f(x)$ lie on this circle.

Existence of Solution (A Constructive Proof):

Let $x_i, i = 1, \dots, n$ be the n -rows of X . Then the matrix equation is equivalent to the set of equations

$$\begin{cases} x_i A^* - x_{i-1} = e_{i-1}, i = 2, 3, \dots, n \\ (a_1 x_1 + a_2 x_2 + \dots + a_n x_n)A^* - x_n = e_n \end{cases} \dots\dots\dots(2.19)$$

where e_1, e_2, \dots, e_n are the n successive rows of I . The equations are equivalent to the single equation (by eliminating x_1, \dots, x_{n-1})

$$x_n(a_1 A^{*n} + a_2 A^{*(n-1)} + \dots + a_n A^* - I) = e_n + e_{n-1}(a_1 A^{*(n-1)} + a_2 A^{*(n-2)} + \dots + a_{n-1} A^*) + \dots + e_1 a_1 A^*$$

This equation gives rise to the following system of non homogeneous linear equations of the form

$$-\psi(\bar{A})x_n^T = e_n^T + (a_1\bar{A}^{n-1} + a_2\bar{A}^{n-2} + \dots + a_{n-1}\bar{A})e_{n-1}^T + \dots + a_1\bar{A}e_1^T = b \dots \dots \dots (2.20)$$

where $\psi(x) = 1 - a_1 x^n - \dots - a_n x$.

Let the eigenvalues of the matrix A be $\lambda_1, \lambda_2 \dots \lambda_n$. Then the eigenvalues of $\psi(\bar{A})$ are $\psi(\bar{\lambda}_i)$, $i = 1, \dots, n$, where $\bar{\lambda}_i$ are the eigenvalues of \bar{A} . Since,

$$\psi(x) = x^n f\left(\frac{1}{x}\right),$$

$$\psi(\bar{\lambda}_i) = (\bar{\lambda}_i)^n f\left(\frac{1}{\bar{\lambda}_i}\right) = (1 - \lambda_1 \bar{\lambda}_i)(1 - \lambda_2 \bar{\lambda}_i) \dots (1 - \lambda_n \bar{\lambda}_i)$$

Hence,

$$\det \psi(\bar{A}) = \prod_{i, j=1}^n (1 - \lambda_i \bar{\lambda}_j)$$

one thus obtains the following existence theorem:

THEOREM 2.10¹: Let A be a companion matrix (2.1) with eigenvalues $\lambda_1, \lambda_2 \dots \lambda_n$. Then the matrix equation (2.18) has a unique Hermitian solution X if and only if

$$\lambda_i \bar{\lambda}_j \neq 1 \text{ for any } i \text{ and } j.$$

Note. As before it is noted that the condition of existence of the unique solution in theorem 2.10 is exactly equivalent to the statement $P(A) \neq 0$, $P(A)$ in this case, being $2(\lambda_i \bar{\lambda}_j - 1)$.

Construction of Solution and other Computational Aspects:

The nth row vector of the solution X is first obtained by

¹ The result is well known ([32] or [34], but proof is new and constructive.

solving the system (2.20) and the other row vectors are then obtained from the recursive relations (2.19). The coefficient matrix $-\psi(\bar{A})$ and the vector b on the right hand side of the system of equation (2.20) can be computed by using the algorithm of Barnett, as before.

Example

Let $f(\lambda) = \lambda^3 - 2\lambda^2$ be a polynomial of degree 3. Then

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

and the matrix equation (2.18) gives rise to the following system of non homogeneous equations:

$$\begin{pmatrix} -1 & 2 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_{31} \\ x_{32} \\ x_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

where $x_3 = (x_{31}, x_{32}, x_{33})$ is the third row of A . Since the system is non singular, it has a unique solution.

$$x_{31} = \frac{4}{3}, \quad x_{32} = \frac{2}{3}, \quad x_{33} = \frac{1}{3}.$$

The remaining two rows x_2 and x_1 , calculated by using the recursive relations (2.19) are

$$x_2 = x_3 A^* - (0, 1, 0) = \left(\frac{2}{3}, -\frac{2}{3}, \frac{2}{3}\right)$$

$$x_1 = x_2 A^* - (1, 0, 0) = \left(-\frac{5}{2}, \frac{2}{3}, \frac{4}{3}\right)$$

Thus, the unique hermitian solution X of the matrix equation is

$$X = \begin{pmatrix} -\frac{5}{3} & \frac{2}{3} & \frac{4}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{4}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

The leading principal minors of X are

$$D_1 = -\frac{5}{3}, D_2 = +\frac{6}{9}, D_3 = \frac{10}{3}$$

So, according to Jacobi-theorem (p. 8)

$$v = V(1, -\frac{5}{3}, \frac{6}{9}, \frac{10}{3}) = 2$$

$$\pi = P(1, -\frac{5}{3}, \frac{6}{9}, \frac{10}{3}) = 1$$

Thus, the number of zeros of $\phi(\lambda)$ outside the unit circle is one of those inside it is 2.

2.3

THE EIGENVALUES IN A SECTOR

Let $f(x) = x^n - a_n x^{n-1} \dots a_2 x - a_1$ be a polynomial with real coefficients. Then the problem is to find the number of zeros of $f(x)$ inside the angle $-\theta < \arg x < \theta$.

(a) Hermitian Forms and Zeros in a Sector

There seems to be known only one result in the literature, of solving this problem using Hermitian Form. This is due to Fujiwara [13], who, using a Hermitian form $H(f)$ which has as its generating function.

$$k(f) = \frac{-[i f(e^{i\theta} x) f(e^{-i\theta} y) - i f(e^{i\theta} y) f(e^{-i\theta} x)]}{x - y}$$

obtained the following:

THEOREM 2.11 (Fujiwara): A necessary and sufficient condition that all zeros of a real polynomial $f(x)$ lie inside the angle $-\theta < \arg x < \theta$ is that the Hermitian form $H(f)$ be positive definite.

(b) Matrix Equations and the Number of Zeros in a Sector

The number of zeros of a real polynomial $f(x)$ in a given sector can be obtained by solving the matrix equation

$$i e^{-i\theta} A X - i e^{i\theta} X A^T = I \dots\dots\dots(2.21)$$

where A is the companion matrix of $f(x)$. Thus, if $\text{In}(X) = (\pi, \nu, \delta)$, then π is the number of zeros in the given sector $-\theta < \arg x < \theta$.

Existence of a Unique Solution (A Constructive Proof):

Let the solution matrix X be given by

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

where x_i is the i th row of X . Then the matrix equation (2.21) is equivalent to the following set of equations:

$$\begin{cases} x_k = e^{2i\theta} x_{k-1} A^T - i e_{k-1}, & k = 2, 3, \dots, n \\ (a_1 x_1 + a_2 x_2 + \dots + a_n x_n) = e^{2i\theta} x_n A^T - i e_n \end{cases} \dots\dots(2.22)$$

where e_k is the k th row of the identity matrix I . The relation (2.22) gives rise to a system of equations of the form

$$f(e^{2i\theta} A)x_1^T = i e_n^T + i f_1(e^{2i\theta} A)e_{n-1}^T + \dots + i f_{n-1}(e^{2i\theta} A)e_1^T$$

where

$$f_r(x) = x^r - a_n x^{r-1} \dots - a_{n-r+2} x - a_{n-r+1}.$$

The system has a unique solution if and only if $f(e^{2i\theta} A)$ is non singular. Let

$$\psi(x) = f(e^{2i\theta} x)$$

Then

$$\psi(A) = f(e^{2i\theta} A)$$

and

$$R = \det(\psi(A)) = \prod_{i=1}^n \psi(\lambda_i),$$

where λ_i are the eigenvalues of A . Again, since

$$\psi(\lambda_k) = (e^{2i\theta} \lambda_k - \lambda_1)(e^{2i\theta} \lambda_k - \lambda_2) \dots (e^{2i\theta} \lambda_k - \lambda_n)$$

$$R = \prod_{k, j=1}^n (e^{2i\theta} \lambda_k - \lambda_j)$$

one thus obtains the following:

THEOREM 2.12: Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of a real companion matrix A . Then the matrix equation (2.21) admits a unique solution X if and only if

$$e^{2i\theta} \lambda_k - \lambda_j \neq 0 \text{ for any } k \text{ and } j.$$

Construction of Solution and other Computational Aspects

The unique solution, whenever it exists, can be computed by solving first the system of equations (2.23) for x_1 and then computing x_2, x_3, \dots, x_n recursively from the relation (2.22). Since, for a given θ , $f(e^{2i\theta} A)$ is a polynomial in A , the coefficient matrix and the vector b on the right hand side of the system (2.23) can be computed, as before, by using the algorithm of Barnett.

CHAPTER III

A NEW APPROACH TO THE SOLUTION OF THE CLASSICAL STABILITY PROBLEMS

Two different approaches - one, the use of Hermitian forms and the other, the use of matrix equations - have been discussed in the previous chapter about the solutions of the stability problems.

In this chapter a third approach will be introduced for certain cases which is different from these two and appears to be more efficient computationally. Let C be the coefficient matrix of the system of equations arising from the associated matrix equation for each problem. It will be shown that C or \bar{C} can be employed to solve the problem in each case. This method eliminates the need to solve the matrix equation. Moreover, C is a polynomial in the given complex matrix A and therefore can be computed using the algorithm of Barnett [reference] which involves no computation of powers of A . Moreover C is of order n in contrast to the Bigradients and the Schur-Cohn determinants which are of order $2n$ ($2n-1$, if the polynomial is monic). The technique used here is similar to that of Barnett [6].

3.1 SOLUTION OF THE HERMITE-PROBLEM

It has been seen that the Hermite problem gives rise to the matrix equation (2.6) and the coefficient matrix of the system of equations arising out of this matrix equation is $-i\phi(\bar{A})$, where $\phi(x)$ is the characteristic polynomial of A . It will be shown below how the

polynomial matrix $i\phi(\bar{A})$ itself can be used to solve the problem.

Let rank of the matrix $\phi(\bar{A})$ be r . Then r is equal to $n-\delta$, where δ is the degree of the greatest common divisor of $\phi(x)$ and $\bar{\phi}(x)$ [5]. So, the nullity of the matrix $\phi(\bar{A}) = n - r =$ the degree of the greatest common divisor of $\phi(x)$ and $\bar{\phi}(x)$. Since $\phi(x)$ and $\bar{\phi}(x)$ can have a common divisor if and only if $\phi(x)$ has a real zero or there is a complex zero of $\phi(x)$ which is one of a conjugate pair, one obtains the following result immediately:

THEOREM 3.1: The nullity of the matrix $\phi(\bar{A})$ is equal to the number of real zeros of $\phi(x)$ increased by the number of complex conjugate zeros counting their multiplicities.

It will now be shown how the signs of the minors of $i\phi(\bar{A})$ determine the number of zeros of $\phi(x)$ in the upper and lower half planes. For this, the following lemma will be established first.

LEMMA 3.1: Let h_1, h_2, \dots, h_n be the successive rows of the Bezout matrix B associated with

$$\phi(x) = x^n - a_n x^{n-1} \dots - a_2 x - a_1$$

and

$$\phi^*(x) = -\bar{\phi}(x) = -x^n + \bar{a}_n x^{n-1} \dots + \bar{a}_2 x + \bar{a}_1.$$

Then

$$h_n = (\bar{a}_1 - a_1, \bar{a}_2 - a_2, \dots, \bar{a}_n - a_n) \dots \dots \dots (3.1)$$

$$h_k = -\bar{a}_{k+1} h_n + h_{k+1} \bar{A}, \quad k = n-1, n-2, \dots, 3, 2, 1 \dots (3.2)$$

where A is the companion matrix of $\phi(x)$.

Proof: It is easy to verify that

$$h_n = (\bar{a}_1 - a_1, \dots, \bar{a}_n - a_n).$$

To prove the second relation, let $B = (b_{ij})$ $i, j = 1, \dots, n$ then setting $b_{k,0} = 0$ one obtains,

$$\begin{aligned} b_{n-r, s} - b_{n-r+1, s-1} &= a_s \bar{a}_{n-r+1} - \bar{a}_s a_{n-r+1} \\ &= a_s \bar{a}_{n-r+1} + \bar{a}_s (b_{n-r+1, n} - \bar{a}_{n-r+1}) \end{aligned}$$

(since, $b_{n-r+1, n} = b_{n, n-r+1} = (\bar{a}_{n-r+1} - a_{n-r+1})$)

$$\begin{aligned} &= \bar{a}_s b_{n-r+1, n} + \bar{a}_{n-r+1} (a_s - \bar{a}_s) \\ &= \bar{a}_s b_{n-r+1, n} - \bar{a}_{n-r+1} (\bar{a}_s - a_s) \\ &= \bar{a}_s b_{n-r+1, n} - \bar{a}_{n-r+1} b_{n, s} \end{aligned}$$

(since $b_{n, s} = \bar{a}_s - a_s$).

$$\text{So, } b_{n-r, s} + \bar{a}_{n-r+1} b_{n, s} = \bar{a}_s b_{n-r+1, n} + b_{n-r+1, s-1} \quad \dots\dots(3.3)$$

Now, $h_{n-r+1} = (b_{n-r+1, 1}, b_{n-r+1, 2}, \dots, b_{n-r+1, n})$

whence

$$h_{n-r+1} \bar{A} = (b_{n-r+1, 1}, b_{n-r+1, 2}, \dots, b_{n-r+1, n}) \begin{pmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \bar{a}_1 & \bar{a}_2 & \dots & \dots & \bar{a}_n \end{pmatrix}$$

$$= (\bar{a}_1 b_{n-r+1, n}, b_{n-r+1, 1} + \bar{a}_2 b_{n-r+1, n}, \dots, b_{n-r+1, n-1} +$$

$$\bar{a}_n b_{n-r+1, n})$$

It therefore follows from (3.3) that

$$h_{n-r} = -\bar{a}_{n-r+1} h_n + h_{n-r+1} \bar{A}, \quad r = 1, 2, \dots, n-1$$

Returning to the original problem, let

$$R = \phi(\bar{A})P$$

where

$$P = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 1 & 0 \\ 0 & & \dots & & 0 \\ \vdots & & \ddots & & \vdots \\ 1 & \dots & \dots & \dots & 0 \end{pmatrix} \dots \dots \dots (3.4)$$

and let the rows of $\phi(\bar{A})$ be $\rho_1, \rho_2, \dots, \rho_n$. Then,

$$\begin{aligned} \rho_1 &= (\bar{a}_1 - a_1, \bar{a}_2 - a_2, \dots, \bar{a}_n - a_n) = h_n \\ &= (\alpha_1, \alpha_2, \dots, \alpha_n) \end{aligned}$$

and

$$\rho_i = \rho_{i-1} \bar{A}, \quad i = 2, 3, \dots, n$$

If the rows of R be $\rho'_1, \rho'_2, \dots, \rho'_n$, then

$$\rho'_1 = (\alpha_n, \alpha_{n-1}, \alpha_{n-2}, \dots, \alpha_2, \alpha_1) = h_n P$$

$$\rho'_2 = \rho_2 P = \rho_1 \bar{A} P = h_n \bar{A} P = h_n P P^{-1} \bar{A} P = \rho'_1 P^{-1} \bar{A} P$$

and similarly,

$$\rho'_i = \rho'_{i-1} P^{-1} \bar{A} P, \quad i = 3, 4, \dots, n.$$

Next, if

$$U = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -\bar{a}_n & 1 & 0 & \dots & 0 \\ -\bar{a}_{n-1} & -\bar{a}_n & 1 & \dots & 0 \\ \vdots & & & 1 & 0 \\ -\bar{a}_2 & -\bar{a}_3 & -\bar{a}_4 & -\bar{a}_n & 1 \end{pmatrix}$$

and if the rows of UR are $\omega_1, \omega_2, \dots, \omega_n$, then

$$\omega_1 = \rho_1' = h_n P$$

$$\begin{aligned} \omega_2 &= -\bar{a}_n \rho_1' + \rho_2' = -\bar{a}_n \omega_1 + \rho_1' P^{-1} \bar{A} P \\ &= -\bar{a}_n \omega_1 + \omega_1 P^{-1} \bar{A} P = -\bar{a}_n h_n P + h_n P P^{-1} \bar{A} P \\ &= -\bar{a}_n h_n P + h_n \bar{A} P = (-\bar{a}_n h_n + h_n \bar{A}) P \\ &= h_{n-1} P \text{ (by lemma 3.1)} \end{aligned}$$

and similarly, $\omega_i = h_{n-i+1} P, i = 3, 4, \dots, n$.

Since,

$$P^{-1} B P = \begin{pmatrix} 0 & 0 & \dots & 1 \\ & \ddots & & \vdots \\ & & \ddots & 0 \\ 1 & \dots & & 0 \end{pmatrix} \begin{pmatrix} h_1 P \\ h_2 P \\ \vdots \\ h_n P \end{pmatrix} = \begin{pmatrix} h_n P \\ h_{n-1} P \\ \vdots \\ h_1 P \end{pmatrix}$$

one gets

$$UR = P^{-1} B P \dots \dots \dots (3.5)^1$$

It has been noted in section 2.1 of Chapter II that the inertia of the hermitian form $H(f)$ determines the number of zeros of $\phi(x)$

¹ Since this thesis has been written, Barnett [54] has obtained a result of which (3.5) is a special case. However the expression (3.5) was derived independently by the author from a different point of view and using techniques different from those of Barnett.

lying strictly above and below the real axis. In fact, since the associated matrices of the hermitian form $H(f)$ and the Bezoutiant $K(f)$ are same, it follows from theorem 2.1 that the number of zeros of $\phi(x)$ with positive imaginary parts and that with negative imaginary parts are respectively equal to the number π of positive squares and that π' of negative squares in a normal representation of the bezoutiant $K(f)$. Therefore if C is the Bezout matrix associated with $K(f)$, then by the theorem of Jacobi (p. 8)

$$\pi = P(1, C_1, C_2, \dots, C_n)$$

$$\pi' = V(1, C_1, C_2, \dots, C_n)$$

where C_1, C_2, \dots, C_n may be assumed to be the leading principal minors of $P^{-1}CP$ (since the number of negative and positive squares remain invariant under a congruent transformation), P being the permutation matrix in 3.4. Now, since $C = iB$ (B is the Bezou-matrix associated with $\phi(x)$ and $-\bar{\phi}(x)$ as in lemma 3.1), it follows from (3.5) that

$$U \text{ i } R = P^{-1}CP.$$

So, by the Cauchy-Binet theorem ([16], vol. 1, p: 11) the leading principal minors of $P^{-1}CP$ and those of iR are the same. Again, since C is real, the leading principal minors of iR are all real. This leads to the formulation of the following theorem:

THEOREM 3.2: Let $\phi(x)$ be a polynomial with complex coefficients whose companion matrix is A ; and let R be the matrix formed from $\phi(\bar{A})$ by interchanging its last column with the first, the last but

one with the second and so on.

Then, the leading principal minors D_i , $i = 1, \dots, n$ of iR are all real and if none of them is zero, $\phi(x)$ has p zeros with positive imaginary parts and q zeros with negative imaginary parts, where p and q are respectively equal to the number of permanences and the number of variations of sign in the sequence

$$1, D_1, D_2, \dots, D_n$$

Example

(a) Let

$$\phi_1(\lambda) = \lambda^3 - 3i\lambda^2 - 3\lambda + i \quad (\lambda_1 = \lambda_2 = \lambda_3 = i)$$

$$a_1 = -i, a_2 = 3, a_3 = 3i$$

Then,

$$\phi_1(\bar{A}) = \begin{pmatrix} 2i & 0 & -6i \\ 6 & -16i & -18 \\ -18i & -48 & -38i \end{pmatrix}$$

and

$$iR = \begin{pmatrix} 6 & 0 & -2 \\ -18i & 16 & 6i \\ -38 & -48i & 18 \end{pmatrix}$$

$$D_1 = 6, D_2 = 96, D_3 = 512$$

So, the number of zeros of $\phi_1(\lambda)$ with positive imaginary parts = $P(1, 6, 96, 512) = 3$.

(b) Let

$$\phi_2(\lambda) = \lambda^3 + 3\lambda - 2i \quad (\lambda_1 = \lambda_2 = i, \lambda_3 = -2i)$$

$$a_1 = 2i, a_2 = 0, a_3 = -3$$

Then,

$$\phi_2(\bar{A}) = \begin{pmatrix} -4i & 0 & 0 \\ 0 & -4i & 0 \\ 0 & 0 & -4i \end{pmatrix}$$

and

$$iR = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 4 & 0 \\ 4 & 0 & 0 \end{pmatrix}$$

$$D_1 = 0, D_2 = 0, D_3 = -64$$

This is a singular case and theorem 3.2 fails. But, however, a result described by Gantmacher ([16], vol. I, Ch. X, § 3 Note 1.) can be adapted in this singular case to obtain the desired information. This gives

$$V(D_0, D_1, D_2, D_3) = V(1, 0, 0, -64) = 1$$

and therefore the number of zeros of $\phi_2(\lambda)$ with negative imaginary parts is one.

3.2 SOLUTION OF THE ROUTH-HURWITZ PROBLEM

As noted earlier, the inertia of the solution X of the Lyapunov equation

$$XA^* + AX = -I$$

specifies the number of zeros of $\phi(x)$ in the right and left planes and hence solve the Routh-Hurwitz Problem. The matrix of the system of linear equations arising out of the Lyapunov equation is $-\psi(\bar{A})$,

where $\psi(x) = \phi(-x)$ and A is the companion matrix of $\phi(x)$. It will now be shown how $(\psi(\bar{A})) = \bar{\psi}(A)$ can be employed to solve the Routh-Hurwitz Problem.

By Theorem 2.5, the number of zeros of $\phi(x)$ with positive and negative real parts are respectively equal to the number π' of negative squares and the number π of positive squares in a normal representation of the Fujiwara matrix F . Let P be a permutation matrix in (3.4), then, since $P^{-1}FP$ has the same number of positive and negative squares as F , by the theorem of Jacobi (p. 8)

$$\begin{aligned}\pi &= P(1, F_1, F_2, \dots, F_n) \\ &\dots\dots\dots(3.6) \\ \pi' &= V(1, F_1, F_2, \dots, F_n)\end{aligned}$$

where F_1, F_2, \dots, F_n may be taken as the leading principal minors of $P^{-1}FP$. Now, the elements $b_{i,k}$ of the bezoutiant B associated with the Bezoutiant $K(f)$ and those $f_{i,k}$ of the Fujiwara matrix F , appearing in section 2.1.2 (iv), p. 62 are connected by the relation

$$f_{i,k} = (-1)^{n-1-i} b_{i,k}$$

whence

$$F_h = (-1)^{h(h-1)/h} B_h, \quad h = 1, 2, \dots, n$$

B_1, B_2, \dots, B_n being the leading principal minors of $P^{-1}BP$. It therefore follows that $B_i, i = 1, \dots, n$ are all real and from (3.6), one obtains

$$\begin{aligned}\pi &= P(1, \epsilon_1 B_1, \epsilon_2 B_2, \dots, \epsilon_n B_n) \\ \pi' &= V(1, \epsilon_1 B_1, \epsilon_2 B_2, \dots, \epsilon_n B_n)\end{aligned}$$

changing the last column with the first, the last but one with the second and so on.

Then, the leading principal minors D_i , $i = 1, \dots, n$ of R are all real, and if none of them is zero, $\phi(x)$ has p zeros with negative real parts and q zeros with positive real parts, where p and q are respectively equal to the number of permanences and variations of sign in the sequence

$$1, \epsilon_1 D_1, \epsilon_2 D_2, \dots, \epsilon_n D_n$$

Example

(a) let

$$\phi_1(\lambda) = \lambda^3 + 6\lambda^2 + 11\lambda + 6 \quad (\lambda_1 = -1, \lambda_2 = -2, \lambda_3 = -3).$$

$$a_1 = -6, a_2 = -11, a_3 = -6.$$

Then,

$$\bar{\phi}(-A) = \begin{pmatrix} 12 & 0 & 12 \\ -72 & -120 & -72 \\ 432 & 720 & 312 \end{pmatrix}$$

and

$$R = \begin{pmatrix} 12 & 0 & 12 \\ -72 & -120 & -72 \\ 312 & 720 & 432 \end{pmatrix}$$

$$D_1 = 12, D_2 = -1440, D_3 = -172800.$$

So, number of zeros of $\phi_1(\lambda)$ with negative real parts =

$$P(1, \epsilon_1 D_1, \epsilon_2 D_2, \epsilon_3 D_3) = P(1, 12, 1440, 172800) = 3.$$

(b) let

$$\phi_2(\lambda) = \lambda^3 - 3\lambda^2 + 4\lambda - 2 \quad (\lambda_1 = 1, \lambda_2 = 1 + i, \lambda_3 = 1 - i)$$

$$a_1 = 2, a_2 = -4, a_3 = 3.$$

Then,

$$\bar{\phi}_2(-A) = \begin{pmatrix} -4 & 0 & -6 \\ -12 & 20 & -18 \\ -36 & 60 & -34 \end{pmatrix}$$

$$R = \begin{pmatrix} -6 & 0 & -4 \\ -18 & 20 & -12 \\ -34 & 60 & -36 \end{pmatrix} \quad D_1 = -6, D_2 = -120, D_3 = +1600$$

Thus, the number of zeros of $\phi_2(\lambda)$ with positive real parts =
 $V(1, \epsilon_1 D_1, \epsilon_2 D_2, \epsilon_3 D_3) = V(1, -6, +120, -1600) = 3.$

3.3 SOLUTION OF THE SCHUR-COHN PROBLEM

The Schur-Cohn Problem gives rise to the matrix equation

$$AXA^* - X = I$$

and the coefficient matrix of the system of linear equations arising out of this matrix equation is $-\psi(\bar{A})$, where $\psi(x) = 1 - a_1 x^n - a_2 x^{n-2} \dots - a_n x$ and A is the companion matrix of $\phi(x) = x^n - a_n x^{n-1} \dots - a_2 x - a_1$. It will be shown below how the principal minors of $\bar{\psi}(A)$ solve the Schur-Cohn Problem.

It has been noted in section 2.2.1 (ii) that the Bezout-matrix $B = (A_i, k)$ and the Fujiwara matrix $F = (a_i, k)$ constructed by Fujiwara to solve the Schur-Cohn problem are connected by the

relation

$$A_{i, n-1-k} = a_{i, k}$$

whence

$$F = BP \dots\dots\dots(3.8)$$

P being the permutation matrix in (3.4). Again, if h_1, h_2, \dots, h_n are the successive rows of B and $\rho_1, \rho_2, \dots, \rho_n$ are the successive rows of $\bar{\psi}(A)$, then, as in the other two cases, it can be verified that

$$\begin{cases} h_n = (1 - a_1 \bar{a}_1, -\bar{a}_1 a_2 - \bar{a}_n, -\bar{a}_1 a_3 - \bar{a}_{n-1} \dots - \bar{a}_1 a_n - \bar{a}_2) = \rho_1 \\ h_k = -a_{k+1} h_n + h_{k+1} A, k = 1, 2, \dots, n-1 \end{cases} \dots\dots(3.9)$$

from which the fundamental relation between the Bezout matrix B and $\bar{\psi}(A)$ can be written down:

$$P^{-1}BP = UR \dots\dots\dots(3.10)$$

where matrices U and P are respectively in (3.7) and (3.4) and

$$R = \bar{\psi}(A)P \dots\dots\dots(3.11)$$

From (3.8) (3.10) and (3.11) one gets

$$P^{-1}FP = U\bar{\psi}(A) \dots\dots\dots(3.12)$$

Applying now Cauchy-Binet theorem to (3.12) one concludes that matrices $P^{-1}FP$ and $\bar{\psi}(A)$ have the same leading principal minors and hence the leading principal minors of $\bar{\psi}(A)$ are all real. Since $P^{-1}FP$ has the same number of positive and negative squares as F,

theorem 2.9 together with the theorem of Jacobi (p. 8) yields the following result:

THEOREM 3.4¹ Let

$$\phi(x) = x^n - a_n x^{n-1} \dots a_2 x - a_1$$

and

$$\psi(x) = -a_1 x^n - a_2 x^{n-1} \dots - a_n x + 1$$

be two polynomials of degree n with real or complex coefficients and let A be the companion matrix of $\phi(x)$.

Then, the leading principal minors D_i , $i = 1, \dots, n$ of $\bar{\psi}(A)$ are all real and if none of them is zero, $\phi(x)$ has p zeros inside the unit circle and q zeros outside it, where p and q are respectively equal to the number of permanences and variations of sign in the sequence

$$1, D_1, D_2, \dots, D_n.$$

Example

(a) Let

$$\phi_1(\lambda) = \lambda^3 - (3 + \frac{1}{2}i)\lambda^2 + (4 + i)\lambda - (i + 2)$$

$$(\lambda_1 = 1 + i, \lambda_2 = 1 - i, \lambda_3 = 1 + \frac{1}{2}i)$$

$$a_1 = i + 2, a_2 = -(4 + i), a_3 = (3 + \frac{1}{2}i).$$

Then

¹ An equivalent form of theorem 3.4 was also obtained by Barnett [51] in a different fashion. The author is grateful to Dr. Barnett for drawing his attention to this fact.

$$\bar{\psi}(A) = \begin{pmatrix} -4 & 6 - \frac{3}{2}i & i - \frac{5}{2} \\ -6 - \frac{1}{2}i & 7 - \frac{3}{2}i & -2 + \frac{1}{4}i \\ -17 - \frac{3}{2}i & \frac{9}{4} + \frac{i}{2} & -i - \frac{107}{2} \end{pmatrix}$$

$$D_1 = -4, D_2 = \frac{35}{4}, D_3 = -\frac{25}{16}.$$

So, the number of zeros of $\phi_1(\lambda)$ outside the unit circle is $V(1, -4, \frac{35}{2}, -\frac{25}{16}) = 3$.

(b) Let

$$\phi_2(\lambda) = \lambda^3 - 2\lambda^2 \quad (\lambda_1 = \lambda_2 = 0, \lambda_3 = 2)$$

$$a_1 = 0, a_2 = 0, a_3 = 2$$

Then,

$$\bar{\psi}(A) = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & -3 \end{pmatrix}$$

$$D_1 = 1, D_2 = 1, D_3 = -3.$$

Thus, the number of zeros of $\phi_2(x)$ outside the unit circle is

$$V(1, 1, 1, -3) = 1$$

and that inside it

$$P(1, 1, 1, -3) = 2.$$

Derivation of the Matrix Equation satisfied by Fujiwara Matrix

As noted earlier, matrix equations satisfied by the Fujiwara-matrix F and the Routh-matrix R associated with the Hermitian Forms employed to solve the Routh-Hurwitz Problem have recently been

discovered by Howland [26]. The matrix equation satisfied by the Fujiwara matrix used to solve the Schur-Cohn Problem is derived below.

Comparing the relations 3.9 with those in p. 28 and choosing $s_n = h_n$, one obtains immediately the following result:

The Bezout-matrix B associated with polynomials $\phi(x)$ and $\psi(x)$ is a symmetrizer of A ;

$$BA = A^T B.$$

Again, since $B = FP$, F satisfies the equation:

$$FPA = A^T FP$$

that is,

$$A^T F = FPAP.$$

Note: There is no difficulty in obtaining similar results in case of previous two problems.

APPENDIX

AN ALTERNATIVE THEOREM

An equivalent form of theorem 3.4 will be presented here. The proof uses the Schur-Cohn determinants, instead of the Bezoutiant and theorem to be presented may be considered as an alternative to the Schur-Cohn theorem (Theorem 2.9).

By theorem 2.9¹, the number of zeros of a given polynomial $\phi(x)$ in the unit circle can be determined by studying the signs of the sequence

$$1, \Delta_1, \Delta_2, \dots, \Delta_n$$

where Δ_k , $k = 1, \dots, n$ are the Schur-Cohn determinants defined in Section 2.2.1 (i). In what follows, it will be shown that these determinants are respectively equal to the leading principal minors of the coefficient matrix of the system of equations (2.20) arising out of the matrix equation (2.18) associated with the Schur-Cohn problem.

It is easy to see that

$$\Delta_k = \begin{vmatrix} A_k^T & (\bar{A}_k') \\ (A_k')^T & \bar{A}_k \end{vmatrix} \dots\dots\dots (1)$$

where A_k and A_k' are triangular matrices

¹ Similar results were obtained by Parks [58] and by Barnett [52] and [53] for the Routh-Hurwitz theorem.

$$A_k = \begin{pmatrix} -a_1 & -a_2 & \cdots & -a_k \\ 0 & -a_1 & \cdots & -a_{k-1} \\ \vdots & & & \\ 0 & 0 & & -a_1 \end{pmatrix} \dots\dots\dots(2)$$

and

$$A'_k = \begin{pmatrix} 1 & -\bar{a}_n & -\bar{a}_{n-1} & \cdots & -\bar{a}_{n-k+2} \\ 0 & 1 & -\bar{a}_n & \cdots & -\bar{a}_{n-k+3} \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \dots\dots(3)$$

Again, since two matrices in either column of (1) commute

$$\Delta_k = \det(A_k^T \bar{A}_k - (A'_k)^T (\bar{A}'_k))$$

(By [22], p. 43)

Further, it is noted that each Δ_k is the leading principal minor of Δ_{k+1} , whence $\Delta_1, \Delta_2, \dots, \Delta_n$ are seen to be equal to the leading principal minors of the matrix $R = (A_n^T \bar{A}_n - (A'_n)^T (\bar{A}'_n))$. By direct multiplication, the matrix is seen to be equal to

$$R = \begin{pmatrix} a_1 \bar{a}_1 - 1 & a_1 \bar{a}_2 + a_n & \cdots & a_1 \bar{a}_n + a_2 \\ a_2 \bar{a}_1 + \bar{a}_n & a_2 \bar{a}_2 + a_1 \bar{a}_1 & \cdots & a_2 \bar{a}_n + a_1 \bar{a}_{n-1} \\ \vdots & -a_n \bar{a}_n - 1 & & -\bar{a}_n a_2 + a_3 \\ a_n \bar{a}_1 + \bar{a}_2 & \bar{a}_2 a_n + a_{n-1} \bar{a}_1 & & a_n \bar{a}_n + \cdots + a_1 \bar{a}_1 \\ & -a_n \bar{a}_2 + \bar{a}_3 & & -a_2 \bar{a}_2 - a_3 \bar{a}_3 \cdots - 1 \end{pmatrix} \dots\dots(4)$$

Let r_1, r_2, \dots, r_n be the rows of R . Then it can be verified that

$$\begin{cases} r_1 = (a_1 \bar{a}_1 - 1, a_1 \bar{a}_2 + a_n, \dots, a_1 \bar{a}_n + a_2) = (\alpha_1, \alpha_2, \dots, \alpha_n) \\ r_i = -\bar{a}_{n-i+2} r_1 + r_{i-1} \bar{A}, i = 2, \dots, n \end{cases} \dots\dots\dots(5)$$

Again, if the rows of $-\psi(\bar{A})$ are $\rho_1, \rho_2, \dots, \rho_n$, then

$$\begin{cases} \rho_1 = (\alpha_1, \alpha_2, \dots, \alpha_n) \\ \text{and by Barnett's algorithm} \dots\dots\dots(6) \\ \rho_i = \rho_{i-1} \bar{A}, i = 2, \dots, n \end{cases}$$

A matrix S with rows s_1, s_2, \dots, s_n are now formed such that

$$\begin{cases} s_1 = r_1 \\ s_2 = r_2 + \bar{a}_n s_1 \\ s_3 = r_3 + \bar{a}_n s_2 + \bar{a}_{n-1} s_1 \dots\dots\dots(7) \\ \vdots \\ s_n = r_n + \bar{a}_n s_{n-1} + \dots + a_2 s_1 \end{cases}$$

Then, using relations (5) - (7), one has

$$s_1 = r_1 = \rho_1$$

$$\begin{aligned} s_2 &= r_2 + \bar{a}_n s_1 = -\bar{a}_n r_1 + r_1 \bar{A} + \bar{a}_n r_1 \\ &= r_1 \bar{A} = \rho_1 \bar{A} = \rho_2 \end{aligned}$$

$$\begin{aligned} s_3 &= r_3 + \bar{a}_n s_2 + \bar{a}_{n-1} s_1 = -\bar{a}_{n-1} r_1 + r_2 \bar{A} + \bar{a}_n s_2 + \bar{a}_{n-1} r_1 \\ &= r_2 \bar{A} + \bar{a}_n s_2 = (s_2 - \bar{a}_n s_1) \bar{A} + \bar{a}_n s_2 = s_2 \bar{A} - \bar{a}_n s_1 \bar{A} + \bar{a}_n s_2 \end{aligned}$$

$$= s_2 \bar{A} - \bar{a}_n \rho_2 + \bar{a}_n \rho_2 = s_2 \bar{A} = \rho_2 \bar{A} = \rho_3.$$

and similarly, $s_i = \rho_i$, $i = 4, 5, \dots, n$.

The leading principal minors of the matrix S are therefore just equal to the leading principal minors of the matrix $-\psi(\bar{A})$. Again, by the Cauchy-Binet theorem, the leading principal minors of S are the same as those of the matrix R and are therefore real (since R is hermitian). Thus the leading principal minors of $-\psi(\bar{A})$ are all real and equal to those of R . This yields the following result:

THEOREM 1 Let

$$\phi(x) = x^n - a_n x^{n-1} \dots - a_1$$

and

$$\psi(x) = 1 - a_1 x^n - a_2 x^{n-1} \dots - a_n x$$

be two polynomials of degree n and let A be the companion matrix of $\phi(x)$.

Then, the leading principal minors D_i , $i = 1, \dots, n$ of $-\psi(\bar{A})$ are all real and, if none of them is zero, $f(x)$ has no zeros on the circle and p zeros in this circle, p being the number of variations of sign in the sequence

$$1, D_1, D_2, \dots, D_n.$$

Example

Let

$$\phi(\lambda) = \lambda^3 - (3 + \frac{i}{2}) \lambda^2 + (4 + i) \lambda - (i + 2)$$

$$a_1 = i + 2, a_2 = -(4 + i), a_3 = (3 + \frac{1}{2}i)$$

Then

$$-\psi(\bar{A}) = \begin{pmatrix} 4 & -6 - \frac{3}{2}i & i + \frac{5}{2} \\ 6 - \frac{i}{2} & -7 - \frac{3}{2}i & 2 + \frac{i}{4} \\ 17 - \frac{3}{2}i & -\frac{9}{4} + \frac{i}{2} & -i + \frac{107}{2} \end{pmatrix}$$

$$D_1 = 4, D_2 = \frac{35}{4}, D_3 = \frac{25}{16}.$$

Since, all D's are non zero, there is no zero of $\phi(\lambda)$ on the unit circle and since,

$$V(D_1, D_2, D_3) = V(4, \frac{35}{4}, \frac{25}{16}) = 0.$$

there is no zero inside it. Thus, all the zeros of $\phi(\lambda)$ are outside the unit circle.

CONCLUSIONS

This thesis includes the results of investigations into two aspects of the General Matrix Eigenvalue Problem.

Assuming that the given matrix A is nonderogatory, it is shown in the first chapter that information about the Jordan Structure of A may be derived from two sources - a Hankel matrix of Newton sums and the solutions of matrix equations. It is well-known that the Hankel form defined by a Hankel matrix of Newton sums plays an important role in the separation of the zeros of a polynomial. It is shown that this Hankel matrix is equally effective in determining the Jordan structure of A . In particular, it is shown how it may be employed to answer a question which is relevant to the determination of the Jordan structure of A : When is a given nonderogatory matrix A similar to a normal matrix? Regarding the application of matrix equations it is shown that the simplest matrix equation relevant to the problem of Jordan structure is $SA = A*S$. After showing what information may be obtained from different types of solutions of this equation, methods of computing these solutions numerically are proposed in some cases. Again, the problem as to when or whether a given matrix A is similar to a normal matrix is shown to be formulated in terms of a non linear matrix equation: $S^{-1}A*SA = A^{-1}SA*S$. This non linear equation is solved theoretically and a few results on its numerical solutions are obtained. Also, some important, well-known results of Matrix Theory are deduced as consequences of the more general results obtained in this connection.

Chapters II and III are devoted to the study of the problem of locating the eigenvalues of A in a specified region, with special emphasis on the zeros of a polynomial. Three classical problems, the Hermite Problem, the Routh-Hurwitz Problem and the Schur-Cohn Problem, arising in connection with the stability of a system of differential or difference equations, are discussed here. Following a recent published result of Howland [23], these three problems and another more general problem that of locating the zero in a given sector, are formulated in terms of matrix equations. In Chapter II, a simple method for solving these matrix equations is proposed. The method consists in reducing the given matrix equation to a $n \times n$ linear algebraic system of the form $Cx = b$, whose solution gives one row vector of the solution matrix. The other row vectors are then obtained in terms of this known row vector from a recursive relation. The chief characteristic of this method is that it avoids the difficulties of series solutions in contrast to most of the methods available in the literature for the solution of the Lyapunov equation associated with the Routh-Hurwitz Problem.

In Chapter III, a new approach to these classical stability problems is introduced. It is shown that in each case the above matrix C is a complete alternative to the Fujiwara matrix F , constructed by Fujiwara to solve the same problem. The construction of C is simpler than that of F and the use of C to solve the problem has decided computational advantages over the bigradients and the Schur-Cohn determinants. Above all, its use eliminates the need to solve the matrix equation.

Finally a result which can be considered as an alternative to the Schur-Cohn theorem is established in the appendix. This consists in constructing a polynomial matrix whose leading principal minors are the same as the Schur-Cohn determinants. Similar results were obtained by P. C. Parks [58] and Barnett [53] in case of the Routh-Hurwitz Problem.

BIBLIOGRAPHY

1. Barnett, S. & Storey, C. Stability Analysis of Constant Linear Systems by Lyapunov's Second Method. Electronics Letters, 2 (May 1966), 165.
2. _____ Solution of the Lyapunov Matrix Equation. Electronics Letters, 2 (Dec. 1966) 466.
3. _____ Remarks on Numerical Solutions of the Lyapunov Matrix Equation. Electronics Letters 3, 9 (Sept. 1967) 117.
4. Barnett, S. Algebraic Methods in Lyapunov Theory. Control, 12 (1968) 545-549.
5. _____ Greatest Common divisor of two polynomials. Linear Algebra and Appl., 3 (1970) 7-9.
6. _____ Location of zeros of a Complex polynomial. Linear Algebra and Appl., 4 (1971) 71-76.
7. Bellman, R. "Introduction to Matrix Analysis". McGraw-Hill, Toronto 1960.
8. Browne, E. T. "Introduction to the Theory of Determinants and Matrices". University of North Carolina, Chapel Hill, 1958.
9. Datta, B. N. On the Solution of Matrix Equations. Essay, Department of Mathematics, University of Ottawa 1971.
10. Davison, E. J. & Man, F. T. $A'Q + QA = -C$. IEEE Trans. Automatic Control., AC - 13 (1968) 448-449.
11. Desantels, E. O. Symmetrizing Matrices. M.Sc. Thesis University of Ottawa, 1967.
12. Duke, J. W. A Note on the Similarity of a Matrix and its Conjugate Transpose. Pacific J. Math., 31 (1969) 321-323.
13. Fujiwara, M. Uber die Nullstellen der ganzen Funktionen von Geschlecht Null Und Eins. Tôhoku Math. J., 25 (1925) 27-35.
14. _____ Ueber die Rezoutiante zweier Polynome. Japan J. Math., 2 (1965) 9-12.

15. _____ On Algebraic Equations whose Roots lie in a Circle or in a Half-plane (German). Math. Z., 24 (1926) 161-169.
16. Gantmacher, F. R. The Theory of Matrices. Vol. I and Vol. II, Chelsea Publishing Co., New York 1959.
17. Hermite, C. Extrait d'une lettre de Mr. Ch. Hermite de Paris à Mr. Borchardt de Berlin sur le Nombre des Racines d'une équation Algèbrique Comprises entre des Limites données. J. Reine Angew. Math., 51 (1856), 39-51.
18. Hitchcock, F. L. A Solution of the Linear Matrix Equation by Double Multiplication Proc. Nat. Acad. Sci. U. S. A., 8 (1922) 78-83.
19. Householder, A. S. Bigradients and the Problem of Routh and Hurwitz. SIAM Rev., 10 (1968) 56-66.
20. _____ Bezoutiants, Elimination and Localization. SIAM Rev., 12, 1 (1970) 73-78.
21. _____ The Theory of Matrices in Numerical Analysis, Blaisdel Pub. Co. New York 1964.
22. _____ The Numerical Treatment of a Single Nonlinear Equation. McGraw-Hill Book Company 1970.
23. Howland, J. L. Matrix Equations and the Separation of Matrix Eigenvalues. J. Math. Anal. Appl., 33 (1971), 683-691.
24. _____ Generating Functions for Symmetrizing Matrices. Manuscript, University of Ottawa, Ottawa.
25. _____ Three Problems in Numerical Linear Algebra. 19th Ontario Mathematical Meeting, Carleton University, Ottawa, 1971.
26. _____ The Method of Quadratic Forms in Dynamic Stability Problems. Third Canadian Conference in Applied Mechanics, Calgary, May, 1971.
27. _____ & Senez, J. A. A Constructive Method for the Solution of the Stability Problem. Numer. Math., 16, 1 (1970), 1-7.

28. Farrel, F. J. & Matrix symmetrizing Methods for the Algebraic Eigenvalue Problem. A. C. M. National Conference, Denver, Colorado, August 28, 1963.
29. Hurwitz, A. Uber die Bedingungen unter Welchen eine Gleichung nur Wurzeln mit negativen Reelen. Teilen besitzt. Math. Ann., 46 (1895) 273-284. Math. Werke. II 533-545.
30. Jameson, A. Solution of the Equation $XA + BX = C$ by inversion of a $M \times N$ or $N \times N$ matrix. SIAM J. Appl. Math., 16 (1968) 1020-1023.
31. Krein, M. G. & Naimark, M. A. The Method of Symmetric and Hermitian Forms in the Theory of the Separation of the Roots of Algebraic Equations. (Russian) Kharkov GNTI, 1936.
32. Lancaster, P. Explicit Solutions of Linear Matrix Equations. SIAM Rev., 12, 4 (1970) 544-566.
33. Lancaster, P. Theory of Matrices, Academic Press, N. Y., 1969.
34. MacDuffee, C. C. "The Theory of Matrices". Chelsea Publishing Co., N. Y., 1946.
35. Man, F. T. A High-order Method of Solution for the Lyapunov Matrix Equation. Comput. J., 14 (1971) 291-292.
36. Marden, M. The Geometry of the Zeros of a Polynomial. in "Math. Surveys 3," Amer. Math. Soc., N. Y., 1949.
37. Muller, P. Chr. Solution of the Matrix Equation $AX + XB = -Q$ and $S^T X + XS = -Q$. SIAM J. Appl. Math., 18, 3 (1970) 682-687.
38. Neudecker, H. A note on Kronecker Matrix Products and Matrix Equation Systems. SIAM J. Appl. Math., 17, 3 (1969) 603-606.
39. Pierpont, J. On a Generalization of the Secular Equation. B. A. M. S., 33 (1927) 294-296.
40. Rutherford, D. E. On the solution of the Matrix Equation $AX + XB = C$. Konigh Ned. Akad. Wetenschap., Proc. Ser. A, 35 (1932) 54-59.
41. Smith, R. A. Matrix Equation $XA + BX = C$. SIAM J. Appl. Math., 16, 1 (1968), 198-201.

42. Snyders, J. & Zakai, M. On Nonnegative Solutions of the Equation $AD + DA' = -C$. SIAM J. Appl. Math. 18, 3 (1970) 704-714.
43. Stephanos, C. Sur une Extension du Calcul des Substitutions Lineae res. J. Math. Pures Appl., 6 (1900) 73-128.
44. Sylvester, J. Sur la Solution du Cas le plus général des équations linéaires en Quantités ou en matrices du second ordre. C. R. Acad. Sci. Paris, 99 (1884) 117-118.
45. Taussky, O. & Zassenhaus, H. On the Similarity Transformation Between a Matrix and its Transpose. Pacific J. Math., 9 (1959) 893-896.
46. Wedderburn, J. Note on the Linear Matrix Equation. Proc. Edinburg Math. Soc., 22 (1904) 49-53.
47. Wilkinson, J. H. The Algebraic Eigenvalue Problem. Clarendon Press, Oxford 1965.
48. Archbold, J. W. Algebra, Pitman, London 1964.
49. Barnett, S. Qualitative Analysis of Polynomials Using Matrices. IEEE. Trans. Aut. Control, Ac - 15 (1970) 380-382.
50. _____ Matrix Methods in Stability Theory. Nelson, London 1970.
51. _____ Number of Zeros of a Complex Polynomial Inside the Unit Circle. Electronics letters, 6 (1970) 164-165.
52. _____ A new Formulation of the Lietaard-Chipart Stability Criterion. Proc. Camb. Phil. Soc., 70 (1971) 269-274.
53. _____ A New Formulation of the Theorems of Hurwitz, Routh and Sturm. J. Inst. Math. Applies., 8 (1971) 240-250.
54. _____ A Note on the Bezoutian Matrix, SIAM J. Appl. Math., 22 (1972) 84-86.
55. Datta, B. N. On the Solutions of Polynomial Equations I and II. Essays, Department of Mathematics, University of Ottawa, 1970.
56. MacDuffee, C. C. Some Applications of Matrices in Theory of Equations. Amer. Math. Monthly, 57 (1950) 154-161.

57. Molinari, B. P. Algebraic Solution of Matrix Linear Equations in Control Theory. Proc. IEEE, 116 (1969) 1748-1754.
58. Parks, P. C. A New Proof of the Routh-Hurwitz Stability Criterion Using the Second Method of Liapunov. Proc. Cam. Phil. Soc., 58 (1962) 694-702.

SUPPLEMENTARY BIBLIOGRAPHY

1. Baker, G. A. A New Derivation of Newton's Identities and their Applications to the Calculation of the Eigenvalues of a Matrix. Soc. Indust. Appl. Math. 7, 2 (1959) 143-148.
2. Ballantine, C. S. Stabilization by a Diagonal Matrix. Proc. Amer. Math. Soc., 25, 4 (1970) 728-734.
3. Duffin, R. J. Algorithms for Classical Stability Problem. SIAM Rev., 11, 2 (1969) 196-213.
4. Eberlein, P. J. On Measures of Non-normality For Matrices. Amer. Math., Mon., 72 (1965), 995-996.
5. Foulkes, H. O. Rational Solutions of the Matrix Equation $XA = BX$. Proc. London Math. Soc. (2), 50 (1948) 196-209.
6. _____ Collineatory Transformation of a Square Matrix into its Transpore. J. London Math. Soc., 17 (1942) 70-80.
7. Fuller, A. T. Conditions for a Matrix to Have only Characteristic Roots with Negative Real Parts. J. Math. Anal. Appl., 2B (1968) 71-98.
8. Hearon, J. Z. Symmetrizable Generalized Inverses of Symmetrizable Matrices. J. Research Nat. Bureau. Standards-B Math. and Math. Physics 71B, # (1967) 229-231.
9. Henrici, P. Bounds for Iterates, Inverses, Spectral Variations and Fields of Values of Non-Normal Matrices. Numer. Math., 4 (1962) 24-40.
10. Howland, J. L. Quadratic Forms, Matrix Equations, and the Separation of Matrix Eigenvalues. Manuscript, University of Ottawa, Ottawa 1972.
11. _____ Note on Normalizable Matrices, *ibid* 1971. A Note on Matrix Equations, *ibid* 1971. A Generalization of Newton's Identities, *ibid*, 1965.
12. _____ The Construction of Matrices Symmetrizing an Arbitrary Matrix, Canadian Math. Congress, Quebec City P.Q. 1963.

13. _____ A Method for Computing the Real Roots of Determinantal Equations. Amer. Math. Monthly (68), 3 (1961) 237-239.
14. Lancaster, P. Explicit Solutions of Linear Matrix Equations. Manuscript, University of Alberta, Calgary.
15. Lehnigk, S. H. On the Hermite Matrix of a Polynomial. Linear Algebra and Appl., 4 (1971) 407-419.
16. Loizou, G. Non-normality and Jordan Condition Number of Matrices. J. Assoc. Comput. Mach., 16, 4 (1969) 580-584.
17. Ma, E. C. A Finite Series Solution of the Matrix Equation $AX - XB = C$. SIAM J. Appl. Math., 14 (1966) 490-495.
18. Macon, N. & Spitzbart, A. Inverses of Vandermonde Matrices. Amer. Math. Monthly, 94-100.
19. Parker, W. V. The Matrix Equation $AX = XB$. Duke Math. J., 17 (1950) 43-51.
20. Penrose, R. On Best Approximate Solutions of Linear Matrix Equations. Camb. Philos., 52 (1956) 17-19.
21. Rosenblum, M. On the Operator Equation $BX - XA = Q$. Duke Math. J., 23 (1956) 263-270.
- The Operator Equation $BX - XA = Q$ With Self Adjoint A and B . Proc. Amer. Math. Soc., 20 (1969) 115-120.
22. Roth, W. E. The Equations $AX = YB = C$ and $AX - XB = C$ in Matrices. Proc. Amer. Math. Soc., 3 (1959) 392-396.
23. Roy, B. C. & Choudhury, A. K. On the Determination of the Inverse of a Vandermonde Matrix. Internat J. Control 12, 3 (1970) 525-527.
24. Thomson, R. C. Some Matrix Factorization Theorems. Pacific J. Math. 23, 3 (1970) 763-822.
25. Ostrowski, A. & Schneider, H. Some Theorems on the Inertia of General Matrices. J. Math. Anal. and Appl., 4 (1962) 72-84.

26. Zaanen, A. C. Normalisable Transformations in Hilbert Space and Systems of Linear Integral Equations. Acta Math., 83 (1950) 197-248.