

T-1332

A COMPUTER METHOD OF SOLVING
A SYSTEM OF SIMULTANEOUS EQUATIONS
WITH QUADRATIC COEFFICIENTS

ARTHUR LAKES LIBRARY
COLORADO SCHOOL OF MINES
GOLDEN, COLORADO

By

Thomas L. Ross

ProQuest Number: 10781695

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10781695

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science in Mathematics.

Signed: Thomas L. Ross
Thomas L. Ross

Golden, Colorado

Date: Dec 16, 1970

ARTHUR LAKES LIBRARY
COLORADO SCHOOL OF MINES
GOLDEN, COLORADO

Approved: W. W. Whitman
W. W. Whitman
Thesis Advisor

J. R. Lee
J. R. Lee
Head of Department

Golden, Colorado

Date: Dec 16, 1970

TABLE OF CONTENTS

ABSTRACT 1

INTRODUCTION 3

METHOD OF SOLUTION 9

 Expansion by Minors 11

 Algebraic Eigenvalue Problem 13

 Conversion to Eigenvalue Problem 14

 Solution of Eigenvalue Problem 29

ERROR ANALYSIS 32

CONCLUSIONS 34

BIBLIOGRAPHY 36

LIBRARY LANCEL MERRILL
COLORADO SCHOOL OF MINES
GOLDEN, COLORADO

ABSTRACT

The solution of a system of equations with quadratic coefficients is used in the aerospace industry. Systems of second-order linear differential equations conform to the above when the Laplace transformations are computed. These differential equations can be used to describe the motion of various space vehicles. A stability analysis check is performed on these vehicles using their equations of motion in the S (Laplace) plane. Therefore, the solutions of these systems of equations in the S plane are needed by the engineers for use in the stability analysis of these vehicles.

RECEIVED LIBRARY
COLORADO SCHOOL OF MINES
GOLDEN, COLORADO

When examined, the normal triangularization procedure proved to be very inaccurate when used to form the polynomial represented by the coefficient matrix. Taking the Laplace transform of the given set of differential equations, gives rise to a set of simultaneous equations with, at most, quadratic coefficients. All initial conditions were assumed to be zero in the Laplace transforms. At this point, we expanded the determinant of the coefficient matrix by minors, single elements in a row or column, and removed what roots we could by using the quadratic formula. Next the process of converting the problem to an eigenvalue problem was

discussed and a 2 x 2 example was used to illustrate this procedure. Finally, the eigenvalues were computed using the QR transform developed by Francis (2) and programmed by Van Ness (3) .

INTRODUCTION

The rigid body equations of a Titan rocket form a system of second-order linear differential equations. The system of equations was developed to study the lateral motion of the airframe. Bending variables, engine deflection variables, roll moment variables, normal force variables and gyro variables are considered. Fuel slosh modes, considered to be the same as structural bending modes, are automatically included in the system of equations.

The following assumptions were made:

1. The engine is considered a part of the airframe but with an added degree of freedom in that it is gimballed to the airframe.
2. The fluid is considered sloshing, and the sloshing fluid modes are treated in the same manner as bending modes.
3. We have the capability of using 13 bending equations, but the system may not require all of the bending equations. Usually 8 bending modes are sufficient to describe the bending forces on the vehicle.

The system of second-order linear differential

equations has the following form:

$$\begin{aligned} a_{11}\ddot{X}_1 + b_{11}\dot{X}_1 + c_{11}X_1 + \dots + a_{1n}\ddot{X}_n + b_{1n}\dot{X}_n + c_{1n}X_n &= d_1\ddot{\delta} + e_1\dot{\delta} + f_1\delta \\ a_{21}\ddot{X}_1 + b_{21}\dot{X}_1 + c_{21}X_1 + \dots + a_{2n}\ddot{X}_n + b_{2n}\dot{X}_n + c_{2n}X_n &= d_2\ddot{\delta} + e_2\dot{\delta} + f_2\delta \\ &\vdots \\ a_{n1}\ddot{X}_1 + b_{n1}\dot{X}_1 + c_{n1}X_1 + \dots + a_{nn}\ddot{X}_n + b_{nn}\dot{X}_n + c_{nn}X_n &= d_n\ddot{\delta} + e_n\dot{\delta} + f_n\delta \end{aligned}$$

where the X_i 's are the functions (functions of time),
 δ is the forcing function (function of time), a_{ij} 's and d_i 's
 are the coefficients of the second derivatives with respect
 to time of the X_i 's and δ , respectively, b_{ij} 's and e_i 's are
 the coefficients of the first derivatives with respect to
 time of the X_i 's and δ , respectively and the c_{ij} 's and f_i 's
 are the coefficients of the function variables, X_i 's and δ
 respectively.

Taking the Laplace transformations of the above set of
 equations and assuming all initial conditions are equal to
 zero, we obtain the following set of transformed equations:

$$\begin{aligned} (a_{11}S^2 + b_{11}S + c_{11})Y_1 + \dots + (a_{1n}S^2 + b_{1n}S + c_{1n})Y_n &= (d_1S^2 + e_1S + f_1)\sigma \\ (a_{21}S^2 + b_{21}S + c_{21})Y_1 + \dots + (a_{2n}S^2 + b_{2n}S + c_{2n})Y_n &= (d_2S^2 + e_2S + f_2)\sigma \\ &\vdots \\ (a_{n1}S^2 + b_{n1}S + c_{n1})Y_1 + \dots + (a_{nn}S^2 + b_{nn}S + c_{nn})Y_n &= (d_nS^2 + e_nS + f_n)\sigma \end{aligned}$$

where $Y_i = L(X_i)$ and $\delta = L(\sigma)$, L is the Laplace operator,
 Y_i and σ are functions of S and X_i and δ are functions
 time.

The coefficients of the above equations are fixed for each rocket. These coefficients are determined by the structural engineers through the use of stress and strain tests performed on the rocket after it is built. The stability engineers are responsible for the flight stability of the rocket once it leaves the ground on its mission. The flight of the rocket is controlled by an on-board computer called an autopilot. This autopilot is fed signals from the airframe equations given above and then it sends a command to the different modules of the rocket. For example, assume that Y_1 is the engine deflection in the above system, The signal for Y_1 is fed into the on-board computer and this computer feeds the signal through a filter and with an appropriate gain factor controls this signal to stabilize the flight. All of the filtering is done in the Laplace plane using the on-board computer.

The on-board computer is hard wired, and therefore, it must be checked out on the ground through the use of digital and analog computers before it can be implemented in the rocket. For this purpose, the term, transfer function, is defined as the ratio between input and output of 2 polynomials in the Laplace plane. From the above system of equations, we see that the solution for Y_1 is the ratio of 2 polynomials in S . Therefore, Y_1 is called a transfer function.

Applying Cramer's Rule to the system given above, the denominator polynomial represents the determinant of the matrix formed by the quadratic coefficients on the left side of the system. The numerator polynomials represent the determinants of the matrices formed by the quadratic coefficients on the left side of the system with the exception that the coefficients of the variable Y_i being solved for are replaced by the coefficients of the right side of the system. The transfer function Y_i will have the following form:

$$Y_i = J \frac{\prod_{i=1}^m (S - \alpha_i)}{K \prod_{i=1}^n (S - B_i)} \quad (J = \text{constant}).$$

The α_i 's are the zeroes of the transfer function and the B_i 's are the poles of the transfer function. The zeroes and poles are either real or complex conjugates. Since the polynomial coefficients are real, the complex roots will always be conjugate pairs.

Using a modern large-scale digital computer to solve for the α_i 's and B_i 's presents a difficult problem. Assuming that n (number of differential equations) is relatively large ($n \geq 10$), finding the determinants of the coefficient matrices is very difficult. Using a normal

triangularization procedure to introduce zeroes below the main diagonal of the matrix by equation elimination is a tedious problem and difficult to program. As the elimination process continues by columns, the size of the eliminating polynomials keeps increasing and programming is impossible. Even if we succeeded in computing the polynomial, the coefficients would be less than accurate due to truncation and roundoff error in the computer. This method would also be very time consuming and the cost would be prohibitive.

In a paper by James R. Winkelman ⁽⁵⁾ it is shown that in a triangularization process the elements in rows i , $i+1$, \dots , n , $i \geq 3$, are exactly divisible by the diagonal element in the $i-2$ row. The following example illustrates Winkelman's method. Given the following set of simultaneous equations:

$$\begin{aligned} 5X_1 + 3X_2 + 2X_3 + 5X_4 &= 3 \\ 1X_1 + 2X_2 + 1X_3 + 3X_4 &= 1 \\ 4X_1 + 2X_2 + 1X_3 + 3X_4 &= 2 \\ 3X_1 + 4X_2 + 5X_3 + 2X_4 &= 0 \end{aligned}$$

Elimination by equation 1 produces:

$$\begin{aligned} 5X_1 + 3X_2 + 2X_3 + 5X_4 &= 3 \\ 0 + 7X_2 + 3X_3 + 10X_4 &= 2 \\ 0 - 22X_2 - 13X_3 - 5X_4 &= -2 \\ 0 + 11X_2 + 19X_3 - 5X_4 &= -9. \end{aligned}$$

Elimination by equation 2 produces:

$$\begin{aligned} 5X_1 + 3X_2 + 2X_3 + 5X_4 &= 3 \\ 7X_2 + 3X_3 + 10X_4 &= 2 \\ - 25X_3 + 185X_4 &= 30 \\ 100X_3 + 145X_4 &= -85 \end{aligned}$$

Notice that rows 3 and 4 are exactly divisible by the diagonal element in row 1. Dividing rows 3 and 4 by this element we obtain:

$$\begin{array}{rcl} 5X_1 + 3X_2 + 2X_3 + 5X_4 & = & 3 \\ 7X_2 + 3X_3 + 10X_4 & = & 2 \\ - 5X_3 + 37X_4 & = & 6 \\ 20X_3 - 29X_4 & = & -17. \end{array}$$

Elimination by equation 3 produces:

$$\begin{array}{rcl} 5X_1 + 3X_2 + 2X_3 + 5X_4 & = & 3 \\ 7X_2 + 3X_3 + 10X_4 & = & 2 \\ - 5X_3 + 37X_4 & = & 6 \\ - 595X_4 & = & -35. \end{array}$$

Notice that row 4 is exactly divisible by the diagonal element in row 2. This example illustrates Winkelman's method on a very simple matrix. This method would be an aid in solving for the Y_i 's in the Laplace system if we were doing it by hand calculation. It would keep the polynomials formed during triangularization from getting too large just as it kept the integers small in the above example. In a digital computer, not all numbers can be represented exactly using the allowable bit patterns in the machine. This inexact representation, together with roundoff and truncation problems will seldom give an exact division using Winkelman's method. One might program a tolerance into the machine, and if the division is within this tolerance, call it exact. With certain restrictions a tolerance might be calculated but this could give rise to a fictitious polynomial the roots of which would be of little use.

METHOD OF SOLUTION

Since none of the methods discussed give accurate polynomial coefficients, a different procedure must be found. Since there are several fast and accurate methods to compute eigenvalues, the above problem will be converted to an eigenvalue problem. First, the matrix with quadratic elements can be compressed using expansion by minors. The expansion by minors will be limited to rows and columns containing a single element. The term, element, is used here to define the coefficient of a Y_i or σ . By using the limitation described above, the quadratic formula is the most complex root solver to be used. Another factor to consider is that the roots removed during the expansion by minors are exact or in the same degree of accuracy as the coefficients. Next, the compressed matrix will be converted, using a series of matrix transformations, to an eigenvalue problem. Then the eigenvalues shall be computed. Finally, an error analysis will be performed on the eigenvalues. This is accomplished by letting $S = \pm \gamma$, where $\pm \gamma \neq \alpha_i$ or $\pm \gamma \neq B_i$ and evaluating the polynomial and matrix at these S values and checking the ratio of these numbers.

Let M denote a coefficient matrix (S plane) for either

a numerator polynomial or a denominator polynomial of n equations. The method of solution used in this paper is to expand the coefficient matrix, M , by minors, using single elements in the rows and columns. Element has the same meaning as previously defined. We shall define this compressed matrix to be M' . M' may be the same size as M if there are no columns or rows with single elements. The above procedure can reduce the order of M by eliminating roots using the quadratic formula and, also, eliminate some of the future work required.

Next, we shall convert the problem to one of finding the eigenvalues. This will be accomplished by expanding M' using the identity matrix and then breaking it down into two matrices.

Finally, the eigenvalues will be computed. There are several standard methods used to calculate eigenvalues, but the one used in this paper was developed by Francis (2) and will be discussed later.

Expansion By Minors

We will check the coefficient matrix, M , for rows and columns that contain only one non-zero element. M , M' and element will be used in the same context as previously defined. Let us assume that row i contains only one non-zero element and this element is located in column j . Define this non-zero element to be m_{ij} . Since we are actually working with a determinant, we can expand M by removing row i and column j from M forming a new matrix, M' , which is a cofactor of M . Let $|M|$ denote the determinant of matrix M . Then $|M| = m_{ij} |M'|$. We shall repeat the above procedure until no more cancellation can be accomplished.

At this point, there are three possibilities for the matrix, M' . First, all of the rows and columns could be cancelled in which case all roots are found using the quadratic formula. An example of this situation would be a diagonal matrix. This case is very elementary and we will assume that it does not exist. Second, if M' contains a row or column of all zeroes, then $|M'| = 0$ and we have a trivial solution. We will also eliminate this problem from our discussion. Third, M' does not fall into either

of the other two categories. This is the matrix, M' , which we will discuss. We must also realize that the roots removed during the above process must be combined with the eigenvalues to give the true solution.

ALGEBRAIC EIGENVALUE PROBLEM

This section of the paper will be used to give the basic definition of the algebraic eigenvalue problem. This will be needed as a background for further sections of the paper.

"The algebraic eigenvalue problem is the determination of those values for which the set of N homogeneous linear equations in N unknowns

$$AX = \lambda X \quad (1)$$

has a non-trivial solution. Equation (1) may be written in the form

$$(A - \lambda I) X = 0 \quad (2)$$

and for arbitrary λ this set of equations has only the solution $X = 0$. The general theory of simultaneous linear algebraic equations shows that there is a non-trivial solution if, and only if, the matrix $(A - \lambda I)$ is singular, that is

$$\det (A - \lambda I) = 0. \quad (3)$$

The determinant on the left side of equation (3) may be expanded to give the explicit polynomial equation

$$\alpha_0 + \alpha_1 \lambda + \dots + \alpha_{n-1} \lambda^{n-1} + (-1)^n \lambda^n = 0. \quad (4)$$

Equation (4) is called the characteristic equation of the matrix A and the polynomial on the left side of the equation is called the characteristic polynomial. Since the coefficient of λ^n is non-zero and we are assuming throughout that we are working in the complex number field, the equation always has n roots. In general, the roots may be complex, even if matrix A is real, and there may be roots of any multiplicities up to n . The n roots are called the eigenvalues, latent roots, characteristic values or proper values of the matrix A . (Wilkinson, 1965, p. 2)"

Conversion to Eigenvalue Problem

The preceding section was used to define the algebraic eigenvalue problem. It is the purpose of this section to explain how we will convert our problem to an eigenvalue problem, thereby enabling us to use Francis' method of finding the eigenvalues.

At this point, we are only given a matrix, M' , which has as its elements quadratic polynomials in S . To convert our quadratic system to a linear system, as required for the general eigenvalue problem, we will need to use several basic theorems pertaining to determinants of matrices. The statements of the theorems along with proofs and examples will be given.

Theorem 1: For any given $n \times n$ matrix, A , i rows and i columns with elements a_{ij} , such that $a_{ij} = 0$ except along the diagonal of matrix A where these elements $a_{ij} = 1$, may be added to A to form a new $m \times m$ matrix A' where $m=n+i$ and the value of the determinant is unchanged.

$$\text{i.e.} \quad |A| = |A'|$$

Proof: Assume we are given matrix A with i rows and columns added to it. These additional rows and columns conform to the definition in the theorem.

Expansion by minors along the diagonal

gives

$$\prod_{\substack{j = n + 1 \\ i = j}}^m a_{ij} |A| = 1^{(m-n)} |A| = |A|$$

The sign of a_{ij} is always positive.

Example: $|A| = \begin{vmatrix} 4 & 3 \\ 1 & 2 \end{vmatrix}$

Adding two rows and columns:

$$|A'| = \begin{vmatrix} 4 & 3 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

$$|A| = 5 = |A'|$$

Expanding the above discussion to matrices with quadratic elements, we will convert the quadratic system to a linear system. Assume we are given the following matrix:

$$A = \begin{bmatrix} a_{11}S^2 + b_{11}S + c_{11} & a_{12}S^2 + b_{12}S + c_{12} \\ a_{21}S^2 + b_{21}S + c_{21} & a_{22}S^2 + b_{22}S + c_{22} \end{bmatrix}$$

For each column that contains an $a_{ij} \neq 0$ we will add to the matrix a column and row with 1 on the diagonal and zeroes elsewhere. This gives the

following matrix:

$$A' = \begin{bmatrix} a_{11}s^2 + b_{11}s + c_{11} & a_{21}s^2 + b_{21}s + c_{21} & 0 & 0 \\ a_{21}s^2 + b_{21}s + c_{21} & a_{22}s^2 + b_{22}s + c_{22} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

From Theorem 1, we have shown that the

$$\det A = |A| = |A'| = \det A',$$

therefore, the problem remains unchanged.

We now need to state and prove several theorems:

Theorem 2: If all of the elements of a certain row (or column) of a determinant are multiplied by a number K , the value of the determinant is multiplied by K .

Proof: In a given determinant (D) let each element of a certain row be multiplied by K , and denote the resulting determinant by D' . Expand each determinant according to the elements of the row in question. The two determinants are exactly alike except that each term of the expansion for D' includes the factor K .

Therefore $D' = KD$.

Example:

$$|\Delta'| = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ Ka_3 & Kb_3 & Kc_3 & Kd_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix}, \text{ then } |\Delta'| = K|\Delta|$$

$$|\Delta'| = Ka_3A_3 + Kb_3B_3 + Kc_3C_3 + Kd_3D_3 = K$$

$$(a_3A_3 + b_3B_3 + c_3C_3 + d_3D_3) = K|\Delta|$$

where A, B, C, D denote the cofactors.

Corollary: Any factor common to all of the elements of a row (or column) may be removed and written before the determinant.

Theorem 3: If each element of a specified row (or column) of a determinant is written as the sum of two terms, the determinant can be written as the sum of two determinants.

Proof: In the determinant

$$|D| = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2+a_2' & b_2+b_2' & c_2+c_2' \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Denote the cofactors of the elements of the second row by A_2 , B_2 , C_2 , respectively. If we expand D according to the elements of this row, we obtain:

$$|D| = (a_2+a_2')A_2 + (b_2+b_2')B_2 + (c_2+c_2')C_2$$

$$|D| = (a_2A_2 + b_2B_2 + c_2C_2) + (a_2'A_2 + b_2'B_2 + c_2'C_2)$$

$$|D| = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2' & b_2' & c_2' \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Theorem 4: If K is any number, the value of a determinant remains unchanged if to each element of any row (or column) we add K times the corresponding element by some other row (or column).

Example:

$$|D| = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1+Ka_3 & b_1+Kb_3 & c_1+Kc_3 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Proof: Using Theorem 3 on the preceding example, we can write:

$$\begin{vmatrix} a_1+Ka_3 & b_1+Kb_3 & c_1+Kc_3 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} Ka_3 & Kb_3 & Kc_3 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

By the corollary to Theorem 2, we can write:

$$\begin{vmatrix} Ka_3 & Kb_3 & Kc_3 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = K \begin{vmatrix} a_3 & b_3 & c_3 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Since row 1 and row 3 are identical, the determinant is equal to zero. Hence the theorem is established.

Recalling the discussion of the definition of the algebraic eigenvalue problem from Wilkinson, we shall try to get our problem into the form

$$\det (A - \lambda I) = |A - \lambda I| = 0.$$

Replacing λ by S , we obtain

$$|A - S I| = 0. \quad (5)$$

This, as shown previously, when expanded is nothing more than a polynomial in S and the roots of this polynomial are the eigenvalues of the matrix A . Using the theorems and corollaries given before we shall convert the problem to one of finding the eigenvalues.

We are given

$$M' = \begin{bmatrix} a_{11}S^2 + b_{11}S + c_{11} & a_{12}S^2 + b_{12}S + c_{12} \\ a_{21}S^2 + b_{21}S + c_{21} & a_{22}S^2 + b_{22}S + c_{22} \end{bmatrix} \quad (6)$$

M' is derived from two equations. We will use this matrix for illustration purposes only. The method to be used is easily adaptable for any number of equations, but a simplified example will be used for ease of discussion. Using Theorem 1, we will add a row and column to the matrix for every column that contains an $a_{ij} \neq 0$.

This gives

$$M' = \begin{bmatrix} a_{11}S^2 + b_{11}S + c_{11} & a_{12}S^2 + b_{12}S + c_{12} & 0 & 0 \\ a_{21}S^2 + b_{21}S + c_{21} & a_{22}S^2 + b_{22}S + c_{22} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7)$$

Using Theorem 4, we shall multiply column 3
 by $-a_{11}S$ and add it to column 1, column 3
 by $-a_{12}S$ and add it to column 2, column 4
 by $-a_{21}S$ and add it to column 1, and column 4
 by $-a_{22}S$ and add it to column 2. We obtain

$$M' = \begin{bmatrix} a_{11} S^2 + b_{11}S + c_{11} & a_{12} S^2 + b_{12}S + c_{12} & 0 & 0 \\ a_{21} S^2 + b_{21}S + c_{21} & a_{22} S^2 + b_{22}S + c_{22} & 0 & 0 \\ -a_{11}S & -a_{12}S & 1 & 0 \\ -a_{21}S & -a_{22}S & 0 & 1 \end{bmatrix} \quad (8)$$

Note that the value of the determinant of M' remains unchanged. Also using Theorem 4, multiply rows 3 and 4 by S and add them to rows 1 and 2, respectively. Therefore,

$$M' = \begin{bmatrix} b_{11} S + c_{11} & b_{12} S + c_{12} & S & 0 \\ b_{21} S + c_{21} & b_{22} S + c_{22} & 0 & S \\ -a_{11}S & -a_{12}S & 1 & 0 \\ -a_{21}S & -a_{22}S & 0 & 1 \end{bmatrix}, \quad (9)$$

At this point, we will evaluate the determinants of the following two matrices to show that the determinants are equal.

Let:

$$|M_1'| = \begin{vmatrix} a_{11}s^2 + b_{11}s + c_{11} & a_{12}s^2 + b_{12}s + c_{12} \\ a_{21}s^2 + b_{21}s + c_{21} & a_{22}s^2 + b_{22}s + c_{22} \end{vmatrix} .$$

$$\begin{aligned} |M_1'| &= (a_{11}a_{22} - a_{12}a_{21})s^4 + (a_{11}b_{22} + a_{22}b_{11} - a_{12}b_{21} - a_{21}b_{12})s^3 \\ &+ (a_{11}c_{22} + b_{11}b_{22} + a_{22}c_{11} - a_{12}c_{21} - b_{12}b_{21} - a_{21}c_{12})s^2 \\ &+ (b_{11}c_{22} + b_{22}c_{11} - b_{12}c_{21} - b_{21}c_{12})s + (c_{11}c_{22} - c_{12}c_{21}) . \end{aligned}$$

Let:

$$|M_2'| = \begin{vmatrix} b_{11}s + c_{11} & b_{12}s + c_{12} & s & 0 \\ b_{21}s + c_{21} & b_{22}s + c_{22} & 0 & s \\ -a_{11}s & -a_{12}s & 1 & 0 \\ -a_{21}s & -a_{22}s & 0 & 1 \end{vmatrix}$$

$$|M_2'| = s \begin{vmatrix} b_{21}s + c_{21} & b_{22}s + c_{22} & s \\ -a_{11}s & -a_{12}s & 0 \\ -a_{21}s & -a_{22}s & 1 \end{vmatrix} + \begin{vmatrix} b_{11}s + c_{11} & b_{12}s + c_{12} & 0 \\ b_{21}s + c_{21} & b_{22}s + c_{22} & s \\ -a_{21}s & -a_{22}s & 1 \end{vmatrix}$$

$$\begin{aligned} |M_2'| &= s(-a_{12}b_{21}s^2 - a_{12}c_{21}s + a_{11}a_{22}s^3 - a_{12}a_{21}s^3 + a_{11}b_{22}s^2 + a_{11}c_{22}s) \\ &+ b_{11}b_{22}s^2 + (c_{11}b_{22} + c_{22}b_{11})s + c_{11}c_{22} - a_{21}b_{12}s^3 - a_{21}c_{12}s^2 \\ &+ a_{22}b_{11}s + a_{22}c_{11}s - b_{12}b_{21}s - (b_{12}c_{21} + b_{21}c_{12})s - c_{11}c_{21} \end{aligned}$$

$$\begin{aligned}
 |M_2'| &= (a_{11}a_{22}-a_{12}a_{21})S^4 + (a_{11}b_{22}+a_{22}b_{11}-a_{12}b_{21}-a_{21}b_{12})S^3 \\
 &\quad + (a_{11}c_{22}+b_{11}b_{22}+a_{22}c_{11}-a_{12}c_{21}-b_{12}b_{21}-a_{21}c_{12})S^2 \\
 &\quad + (b_{11}c_{22}+b_{22}c_{11}-b_{12}c_{21}-b_{21}c_{12})S + (c_{11}c_{22}-c_{12}c_{21})
 \end{aligned}$$

Therefore, $|M_1'| = |M_2'| = |M'|$ where

$$|M'| = \begin{vmatrix} b_{11}S+c_{11} & b_{12}S+c_{12} & S & 0 \\ b_{21}S+c_{21} & b_{22}S+c_{22} & 0 & S \\ -a_{11}S & -a_{12}S & 1 & 0 \\ -a_{21}S & -a_{22}S & 0 & 1 \end{vmatrix}.$$

Recalling the definition of matrix addition, we see that the sum of two matrices is merely the sum of their corresponding elements. If matrix A contains elements a_{ij} and matrix B contains elements b_{lk} , then matrix C, where $C = A + B$, has elements c_{ij} where $c_{ij} = a_{ij} + b_{lk}$ where $i = 1$ and $j = k$.

Using the corollary of matrix addition, we will break the matrix M' into two matrices, one matrix will contain only "S" terms and zero terms and the rest of the elements will be put into the other matrix. Now

$$|M'| = \begin{vmatrix} b_{11}S & b_{12}S & S & 0 \\ b_{21}S & b_{22}S & 0 & S \\ -a_{11}S & -a_{12}S & 0 & 0 \\ -a_{21}S & -a_{22}S & 0 & 0 \end{vmatrix} + \begin{vmatrix} c_{11}c_{12} & 0 & 0 \\ c_{21}c_{22} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (10)$$

and the determinant of M' is equal to the determinant of the sum of the matrices as shown above.

Another property of matrices that will be used is that of scalar multiplication. If the matrix A with elements a_{ij} is multiplied by the scalar T , a new matrix with elements Ta_{ij} is formed. Using the corollary of this definition, we can factor an S out of the first matrix in the above equation. Then

$$|M'| = \left| S \begin{bmatrix} b_{11} & b_{12} & 1 & 0 \\ b_{21} & b_{22} & 0 & 1 \\ -a_{11}-a_{12} & 0 & 0 & 0 \\ -a_{21}-a_{22} & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} & 0 & 0 \\ c_{21} & c_{22} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right| \quad (11)$$

Let

$$A = \begin{bmatrix} b_{11} & b_{12} & 1 & 0 \\ b_{21} & b_{22} & 0 & 1 \\ -a_{11}-a_{12} & 0 & 0 & 0 \\ -a_{21}-a_{22} & 0 & 0 & 0 \end{bmatrix} \quad (12)$$

and

$$B = \begin{bmatrix} c_{11} & c_{12} & 0 & 0 \\ c_{21} & c_{22} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

Therefore, we now have

$$|M'| = |S \ A + B| . \quad (14)$$

Recalling the definition of the algebraic eigenvalue problem, we are trying to get the problem into the form

$$|X - S \ I| = 0,$$

where the eigenvalues of the matrix X are the roots of the polynomial which is formed with the expansion of the above equation as was shown in the discussion of the eigenvalue problem. Setting the equation to zero, we obtain

$$|S \ A + B| = 0. \quad (15)$$

Getting equation (15) into proper form involves three different cases:

Case 1: A is non singular.

A non-singular implies that A^{-1} (A inverse) exists since the $|A| \neq 0$. Finding the inverse of A, (A^{-1}), and post multiplying each term of the left hand side of equation (15) by A^{-1} we obtain

$$\begin{aligned} & |S \ A \ A^{-1} + B \ A^{-1}| = 0 \\ \text{or} & \quad |S \ I + B \ A^{-1}| = 0 \end{aligned} \quad (16)$$

$$\text{Letting} \quad X = B \ A^{-1}, \quad (17)$$

$$\text{we get} \quad |S \ I + X| = 0$$

and need to calculate the eigenvalues of X.

Case 2: A is singular, B is non-singular.

A singular, B non-singular implies that A^{-1} does not exist, but B^{-1} does exist. Finding B^{-1} and post-multiplying equation (15) by B^{-1} gives

$$|S \quad A \quad B^{-1} \quad + \quad B \quad B^{-1}| = 0$$

or

$$|S \quad A \quad B^{-1} \quad + \quad I| = 0$$

Letting $X = A B^{-1}$ and dividing both terms by S, we obtain

$$|X + \frac{1}{S} I| = 0. \quad (18)$$

After computing the eigenvalues of X, we need to invert them to get the correct eigenvalues for our problem. If we did not wish to invert the eigenvalues, we could compute X^{-1} and find the eigenvalues of X^{-1} . These would satisfy our problem.

Case 3: A is singular, B is singular.

A singular, B singular implies that neither inverse exists so it is impossible to convert the problem to an eigenvalue problem using the same procedures defined in Cases 1 and 2. Recalling that

$$|A \quad S + B| = 0 = \sum_{i=0}^n a_i S^i, \quad (19)$$

we see that for

$$\sum_{i=0}^n a_i S^i = 0 \quad (20)$$

to have a non-trivial solution not all of the a_i 's = 0. There exists at most n values of S for which

$$\sum_{i=0}^n a_i S^i = 0 = |A S + B|.$$

This is by the definition of the roots of a polynomial. Assuming λ is a real number, then

$$|A S + B| = |A S - A \lambda + B + A \lambda|.$$

We have done nothing more than add and subtract the term $A \lambda$. Writing the above equation in another form gives

$$|A S + B| = |A (S - \lambda) + (A \lambda + B)|. \quad (21)$$

This determinant of the matrix $A \lambda + B$ is a polynomial in λ and there are at most n values of λ for which

$$|A \lambda + B| = 0.$$

We will now choose an element γ such that

$$|A \gamma + B| \neq 0. \quad (22)$$

This can only be done by trial and error, but is relatively easy because the γ

picked is not likely to have the exact bit pattern in the computer as one of the λ 's.

Since

$$|A \gamma + B| \neq 0,$$

then $(A \gamma + B)^{-1}$ exists. Using equation (21) and substituting $\gamma = \lambda$, we obtain

$$|A S + B| = |A (S - \gamma) + (A \gamma + B)|. \quad (23)$$

We have already shown that $(A \gamma + B)^{-1}$ exists and pre-multiplying by this inverse gives

$$|A S + B| = |(A \gamma + B)^{-1} A (S - \gamma) + (A \gamma + B)^{-1} (A \gamma + B)|$$

or

$$|A S + B| = |(A \gamma + B)^{-1} A (S - \gamma) + I| \quad (24)$$

Dividing the right side of equation (24) by the term $(S - \gamma)$, where $S - \gamma$ is a scalar, gives

$$|A S + B| = \left| (A \gamma + B)^{-1} A + \frac{1}{S - \gamma} I \right|. \quad (25)$$

Letting

$$X = (A \gamma + B)^{-1} A, \quad (26)$$

then

$$|A S + B| = \left| X + \frac{1}{S - \gamma} I \right|. \quad (27)$$

Equating the right side of equation (27) to zero, we obtain

$$\left| X + \frac{1}{S - \gamma} I \right| = 0.$$

Now we can compute the eigenvalues of X . These eigenvalues are the roots of the polynomial

$$\sum_{i=0}^n a_i \left[\frac{1}{S - \gamma} \right]^i.$$

To find the correct roots for the polynomial in S , we must invert the eigenvalues and add γ to each. Assume that x_i are the eigenvalues of matrix X . Then

$$x_i = \frac{1}{S - \gamma}$$

or

$$S + \frac{1}{x_i} + \gamma. \quad (28)$$

Next we shall discuss the method used to compute the eigenvalues of the matrix X for all three cases.

Solution of Eigenvalue Problem

"The QR transformation consists of forming a sequence of matrices $A^{(k)}$, where

$A^{(1)} = A$, the given matrix is Hessenberg form, and

$$A^{(k+1)} = Q^{(k)} * A^{(k)} Q^{(k)}.$$

The matrix $Q^{(k)}$ is unitary and is chosen so that

$$Q^{(k)} * A^{(k)} = R^{(k)}$$

where $R^{(k)}$ is an upper triangular matrix (Francis, 1962, p. 265)."

The preceding discussion gives the basic concept of the QR transformation for finding the eigenvalues of the matrix X . X corresponds to A in the above discussion.

"Before applying the QR transform, the X matrix is transformed into Hessenberg form. In this form, all of the elements below the first sub-diagonal are zero:

$$X = \begin{bmatrix} X & X & X & X & X & X \\ X & X & X & X & X & X \\ 0 & X & X & X & X & X \\ 0 & 0 & X & X & X & X \\ 0 & 0 & 0 & X & X & X \\ 0 & 0 & 0 & 0 & X & X \end{bmatrix}.$$

The purpose of the QR transform is to make elements on the subdiagonal of the Hessenberg matrix go to zero. For example, if the matrix

can be transformed to

$$X = \begin{bmatrix} X & X & X & X & X & X \\ X & X & X & X & X & X \\ 0 & X & X & X & X & X \\ 0 & 0 & X & X & X & X \\ 0 & 0 & 0 & X & X & X \\ 0 & 0 & 0 & 0 & 0 & a \end{bmatrix},$$

then it is known that the matrix has a real eigenvalue of "a", and the last row and column can be dropped, reducing the order of the matrix by one. If the matrix transforms into the form

$$X = \begin{bmatrix} X & X & X & X & X & X \\ X & X & X & X & X & X \\ 0 & X & X & X & X & X \\ 0 & 0 & X & X & X & X \\ 0 & 0 & 0 & 0 & X & X \\ 0 & 0 & 0 & 0 & X & X \end{bmatrix}$$

then a pair of eigenvalues, possibly complex, can be found directly from the lower right 2 by 2 sub-matrix, and the order of the original matrix is reduced by two by dropping the last two rows and columns. Actually, in the computer program, if any element on the subdiagonal is reduced far enough to be considered a zero (with floating point computer operations, a true zero is practically never obtained), the matrix is divided into two sub-matrices and these are handled independently. It is this feature of breaking the large matrices into smaller pieces that makes this method so powerful with large systems. (Van Ness, 1965, pp. 22-24)."

Once the eigenvalues are computed, the roots removed using the expansion by minors techniques must be combined with the eigenvalues to give the total solution to our

problem. Now we can proceed to an error analysis of our solution to determine if the eigenvalues are accurate within the tolerances of the computer.

ERROR ANALYSIS

Referring to matrix M' , equation (6), we see that

$$|M'| = m_n S^n + m_{n-1} S^{n-1} + \dots + m_1 S + m_0 \quad (29)$$

where the coefficients m_i , $i=0, n$, are combinations of the coefficients a , b , c . Normalizing the polynomial we obtain

$$|M'| = m_n (S^n + p_{n-1} S^{n-1} + \dots + p_1 S + p_0) \quad (30)$$

where $p_i = \frac{m_i}{m_n}$, $i = 0, n-1$. Assume B_i , $i = 1, n$, are

the roots or eigenvalues found for the solution. We see that

$$\prod_{i=1}^n (S - B_i) = S^n + b_{n-1} S^{n-1} + \dots + b_1 S + b_0. \quad (31)$$

Taking the ratio of equation (30) to equation (31) gives

$$\frac{|M'|}{\prod_{i=1}^n (S - B_i)} = m_n \quad (32)$$

since $b_i = p_i$, $i = 0, n-1$. This theoretically is true, but to check this fact we can choose some positive value, α_1 , and some negative value, α_2 , for S such that $\alpha_1 \neq B_i$, $\alpha_2 \neq B_i$, $i = 1, n$. Using these two values, α_1 and α_2 , we can

evaluate the ratio in equation (32) and arrive at two values of m_n , m_{n1} and m_{n2} . Since we do not know the value of m_n without forming the polynomial, we can take the ratio m_{n1}/m_{n2} and see if the ratio is equal to 1. In practice, if we are within 10^{-5} of 1, the roots will be accurate. If this is not the case, there are several convergent criteria in the eigenvalue package which can be tightened to give better accuracy on the eigenvalues.

Another method which could be utilized for an error analysis is to compute the determinant of the matrix X, equations (17), (18), or (26), and check to see if it is equal to the product of the eigenvalues. If one of the eigenvalues is equal to zero, this method could indicate that the eigenvalues are accurate when, in fact, some of the non-zero eigenvalues may be very inaccurate.

CONCLUSIONS

The example used for illustration in this paper was a single 2×2 matrix. In practice, most of the systems will be 10×10 or greater. However, there are no changes needed in the technique described in this paper to find the solutions of larger systems of equations. All computer work was done using a CDC 6000 Series Computer and all math subroutines utilized double precision arithmetic to achieve greater accuracy. The CDC machine is an ideal machine to use on problems such as this because of the 60 bit words used by the machine. If an IBM 360 Series Computer is to be used with this technique, it might be necessary to incorporate quadruple precision arithmetic to get the desired accuracy in the eigenvalues.

The method described in this paper is adaptable to higher order linear differential equations. If, for example, we were given a system of fourth-order linear differential equations, we would have a system of simultaneous equations whose coefficients are, at most, fourth order in S . In this case, one would introduce a new column and row to the coefficient matrix for every column that

contained a S^4 term. Using the properties of determinants, the S^4 terms could be eliminated. Then we could work on the S^3 terms and eliminate them, and finally we would eliminate the S^2 terms and follow the rest of the techniques described in this paper. One of the problems encountered with high order systems (derivatives greater than second order) is that for each order increase the size of the matrix may be doubled. Therefore, it might become necessary to partition the matrices because of storage problems in the computer, and consequently, computer time will probably increase exponentially making the solution too costly in terms of money.

The total program is very fast with systems of up to 45 differential equations. Also, the results obtained from the sample problems were very accurate. It takes about 2 minutes of computer time to solve a 27 x 27 set of equations.

BIBLIOGRAPHY

1. Brink, R. W., 1954, A First Year of College Mathematics, Appleton-Century-Crofts, Inc., New York, 725 pages.
2. Francis, J. G. F., 1962, The QR Transformation: The Computer Journal, v. 4, No. 3.
3. Van Ness, J. E., 1965, Sensitivities of Load-Frequency Control Systems: Northwestern University.
4. Wilkinson, J. H., 1965, The Algebraic Eigenvalue Problem, Oxford University Press, Ely House, London.
5. Winkelman, J. R., 1958, A Method of Solving Simultaneous Linear Algebraic Equations with Polynomial Coefficients, Mathematics Department, University of Denver, Denver, Colorado.
6. Wylie, C. R., Jr., 1960, Advanced Engineering Mathematics, McGraw-Hill Book Co., New York, 696 pages.