

August 13, 1953

To the Graduate Council:

I am submitting herewith a thesis written by Joseph Frederick Elliott entitled "The Characteristic Roots of Certain Real Symmetric Matrices." I recommend that it be accepted for 9 quarter hours of credit in partial fulfillment of the requirements for the degree of Master of Science, with a major in Mathematics.

Walton L. Lively
Major Professor

We have read this thesis
and recommend its acceptance:

Edgar D. Coates
D. H. Harwood, Jr.

Accepted for the Council:

W. L. Lively
Dean of the Graduate School

THE CHARACTERISTIC ROOTS OF CERTAIN REAL SYMMETRIC MATRICES

A THESIS

Submitted to
The Graduate Council
of
The University of Tennessee
in
Partial Fulfillment of the Requirements
for the degree of
Master of Science

by

Joseph Frederick Elliott

August 1953

ACKNOWLEDGEMENT

This thesis was completed while the author was an Oak Ridge Institute of Nuclear Studies Fellow at the Oak Ridge National Laboratory.

The author wishes to express his sincere appreciation to Professor Wallace Givens of the University of Tennessee for his encouraging and patient supervision of the writing of this thesis and to Dr. A. S. Householder and the Mathematics Panel of the Oak Ridge National Laboratory for their assistance in the final preparation of this paper.

TABLE OF CONTENTS

SECTION	PAGE
I. INTRODUCTION	1
II. PHYSICAL SOURCE OF THE PROBLEM	3
III. BACKGROUND	8
IV. GENERAL RESULTS	16
1. An oscillation matrix	16
2. The Fibonacci numbers	25
3. Variations of the oscillation matrix	26
4. Alternating values on the diagonal	31
5. The inverse of a certain matrix	33
6. Symmetric circulants	36
7. An extension by using the direct product	38
8. Bounds for characteristic roots	41
9. The general case	43
10. Summary	45
BIBLIOGRAPHY	50

I. INTRODUCTION

The main purpose of this thesis is to collect and coordinate some known results in the study of characteristic roots of certain real symmetric matrices. Specifically, most of the matrices considered have their elements $a_{ij} = 0$ except for $|i - j| \leq 1$ and are known as Jacobi matrices. The thesis fills in the details of two papers by D. E. Rutherford [12; 13]¹, with some translation of his results for determinants to the problem of finding characteristic roots of matrices. Some of the results obtained are briefly discussed relative to certain theorems on bounds for characteristic roots.

The first part of the paper is an attempt to show some of the physical sources of the problem and indicate some applications of the results. The aim here is to exhibit matrices of the type considered in the problem and little effort is made to give a detailed theory of the physical laws involved.

A background for the study is the purpose of the next section. A rather general treatment of theorems concerning characteristic roots is given, especially with respect to real symmetric matrices. More recent discoveries concerning bounds of characteristic roots are given. Proofs are either sketched briefly or omitted entirely since most of the theorems are well known.

The main problem is introduced by solving one of the physical problems discussed in the section on applications. The results obtained

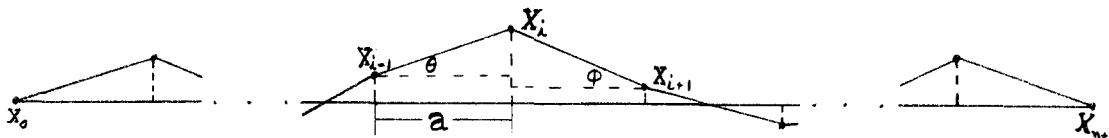
¹References in brackets, [], are to the bibliography at the end of the thesis.

are used as a basis for a more generalized treatment with the aim of finding the characteristic roots of several classes of n by n matrices.

The matrices to be considered are at best somewhat restricted in scope. This seems to limit the usefulness of the results. However, it is feasible from a practical viewpoint to reduce any real symmetric matrix to a Jacobi matrix in such a way as to preserve the characteristic roots. This fact, [6], gives added value to the results obtained. Aside from the actual solution of certain physical problems, it is hoped that results here obtained might be of some importance for at least two reasons: first, it makes available several classes of matrices of order n for which the proper values can be obtained from tables of trigonometric functions and, secondly, the methods used here might suggest an approach to the solution of the problem for more general real symmetric matrices. Much remains to be done, even in the special case of the Jacobi matrix.

II. PHYSICAL SOURCE OF THE PROBLEM

The task of finding the characteristic roots of a matrix is very often the crux of a physical problem. The example below, which deals with a real symmetric matrix, is taken from the theory of oscillations [11, p. 184] .



Consider $n + 2$ particles, each of mass v , distributed equally along a stretched string. Let the distance between two consecutive particles at equilibrium be a so that the length of the string is $(n + 1)a$. Let x_i represent the displacement of the i -th particle along a line perpendicular to the equilibrium line, with positive displacement on one side and negative displacement on the other side, subject to the condition that only plane transverse motion is to be considered. Suppose the tension in the string is K , that gravity is not to be considered and that the end points of the string are fixed; that is, $x_0 = x_{n+1} = 0$. With these hypotheses, it is desired to study the motion of the system after a small displacement from equilibrium, assuming that each particle moves in a line perpendicular to the equilibrium line at the equilibrium position of the particle.

After a small displacement, the tension in the string is still approximately K . Then the resultant force acting on the i -th particle

to restore it to equilibrium is given by

$$v \ddot{x}_1 = -K \sin \theta - K \sin \varphi ,$$

where

$$\ddot{x}_i = \frac{d^2 x_i}{dt^2}$$

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

For small displacement, $\sin \theta$ and $\sin \varphi$ are approximately equal to $\tan \theta$ and $\tan \varphi$ respectively. Hence the equation of motion becomes

$$\begin{aligned} \ddot{x}_i &= -\frac{K}{v} \cdot \frac{x_i - x_{i-1}}{a} - \frac{K}{v} \cdot \frac{x_i - x_{i+1}}{a} \\ &= -\frac{K}{va} (2x_i - x_{i-1} - x_{i+1}) , \quad i = 1, 2, \dots, n . \end{aligned}$$

In matrix notation², this gives

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \\ \vdots \\ \ddot{x}_n \end{bmatrix} = \begin{bmatrix} \frac{-2K}{va} & \frac{K}{va} & & & \\ \frac{K}{va} & \frac{-2K}{va} & \frac{K}{va} & & \\ & \frac{K}{va} & \frac{-2K}{va} & \frac{K}{va} & \\ & & \frac{K}{va} & \frac{-2K}{va} & \frac{K}{va} \\ & & & \frac{K}{va} & \frac{-2K}{va} & \frac{K}{va} \\ & \dots & \dots & \dots & \dots & \dots \\ & & & & \frac{K}{va} & \frac{-2K}{va} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

²For purposes of this thesis, zero elements of displayed matrices will generally be omitted.

It is immediately obvious that the i -th particle does not execute simple harmonic motion unless very special initial conditions are imposed. However, by a suitable linear transformation $y = Tx$ (y and x column vectors), an equation $\ddot{y} = T A T^{-1} y$ can be obtained where $T A T^{-1}$ is a diagonal matrix (Th. 3.3). In this equation the motion is simple harmonic motion and the normal modes of vibration are obtained from the characteristic roots of $T A T^{-1}$ by the relation $T A T^{-1} y = -\omega^2 y$ where $\omega = f/2\pi =$ the angular frequency. It will then follow that $x = T^{-1} y$ gives the motion of the original system.

The characteristic roots of $T A T^{-1}$ are the same as the characteristic roots of A , by similarity properties. So the solution of the problem basically depends on the characteristic roots of A or, in general, on the characteristic roots of a matrix of the form

$$\begin{bmatrix} a & b & & & & \\ b & a & b & & & \\ & b & a & b & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & a & b \\ & & & & b & a \end{bmatrix}.$$

Various matrices of this general type may be obtained by varying the initial conditions. For example, if the distances between the first two particles and the last two particles are changed and denoted by a_1 and a_{n+1} respectively, then

$$\ddot{x}_1 = \left(\frac{-K}{va_1} - \frac{K}{va}\right)x_1 + \frac{K}{va}x_2$$

$$\ddot{x}_i = \frac{-K}{va}(2x_i - x_{i-1} - x_{i+1}), \quad i = 2, 3, \dots, n-1$$

$$\ddot{x}_n = \frac{K}{va}x_{n-1} + \left(\frac{-K}{va} - \frac{K}{va_{n+1}}\right)x_n$$

Now the solution of the problem depends upon the characteristic roots of a matrix of the form

$$\begin{bmatrix} a+c & b & & & & \\ b & a & b & & & \\ & b & a & b & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & a & b \\ & & & & b & a+d \end{bmatrix}$$

In general, if the distance between the i -th and $(i-1)$ st particle is a_i , the equation of motion in matrix notation is

$$\begin{bmatrix} \ddot{x}_1 \\ \cdot \\ \cdot \\ \cdot \\ \ddot{x}_n \end{bmatrix} = \begin{bmatrix} \left(\frac{-K}{va_1} - \frac{K}{va_2}\right) & \frac{K}{va_2} & & & \\ \frac{K}{va_2} & \left(\frac{-K}{va_2} - \frac{K}{va_3}\right) & \frac{K}{va_3} & & \\ & \frac{K}{va_3} & \left(\frac{-K}{va_3} - \frac{K}{va_4}\right) & \frac{K}{va_4} & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & \frac{K}{va_n} & \left(\frac{-K}{va_n} - \frac{K}{va_{n+1}}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

One is therefore interested essentially in the characteristic roots of a matrix of the form

$$\begin{bmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & a_3 & b_3 & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b_{n-2} & a_{n-1} & b_{n-1} \\ & & & & b_{n-1} & a_n \end{bmatrix} .$$

The oscillation problem requires that the a_i and b_i satisfy the conditions $b_i + a_{i+1} + b_{i+1} = 0$, $i = 1, 2, \dots, n-2$ and $b_i > 0$, $i = 1, 2, \dots, n-1$, but these restrictions will be ignored in the following mathematical investigations.

III. BACKGROUND

The matrix obtained from a matrix A by changing rows to columns is called the transpose of A , denoted in this paper by A^T . If the matrix A is such that $A = A^T$, then A is said to be symmetric. One says that A is a real symmetric matrix if $A = A^T$ and all the elements of A are real numbers. It should be noted that a symmetric matrix is necessarily square; that is, the number of rows is the same as the number of columns. Hence one is justified in speaking of a symmetric matrix of order n where n indicates the number of rows or columns.

If, for a square matrix A of order n with elements in the field of complex numbers, there exists a non-zero column vector x (an n by 1 matrix; that is, n rows and 1 column) such that $Ax = \lambda x$ for some complex number λ , then λ is said to be a characteristic root (proper value, latent root, eigenvalue) of A ; and x is said to be a characteristic vector (proper vector, eigenvector) of A corresponding to λ . It is clear that $(A - \lambda I_n)x = 0$ has a non-zero solution if and only if the determinant of the system vanishes; that is, if and only if $|A - \lambda I_n| = 0$. This is called the characteristic equation of the matrix A . From it one obtains the characteristic roots of A .

Theorem 3.1. The characteristic roots of a real symmetric matrix are real.

Proof: Let λ be a characteristic root of a real symmetric matrix A . Then there exists a non-zero column vector x for which

$$(1) \quad A x = \lambda x ,$$

where one does not yet know that λ is real and so must allow the components of x to be complex. Multiplying both sides of (1) by the transposed conjugate of x one obtains

$$(2) \quad \overline{x^T} A x = \lambda \overline{x^T} x .$$

Taking the transposed conjugate of both sides of (2) it follows that

$$\overline{(\overline{x^T} A x)^T} = \overline{x^T} A x = \lambda \overline{x^T} x = \overline{\lambda} \overline{x^T} x$$

which implies that $\lambda = \overline{\lambda}$. Hence λ is real. This proof is suggested by an exercise in Birkhoff and MacLane [1, p. 309] .

Definition 3.1. If x and y are two 1 by n row vectors with real elements, then x and y are said to be orthogonal if and only if $x y^T = 0$. The length of the row vector x is defined to be $(x x^T)^{1/2}$.

Definition 3.2. A square matrix P is called orthogonal if and only if $P^T P = I$; that is, each row of P has length one and any two rows are orthogonal.

Definition 3.3. A square matrix A is called hermitian if $A^T = A^C$ where $A^C = (\overline{a}_{ij})$.

It should be noted that hermitian matrices are important for our purposes because a real symmetric matrix is an hermitian matrix.

Hence results may often be obtained for real symmetric matrices as special hermitian matrices.

Definition 3.4. A square matrix A is called unitary if $A \cdot A^{TC} = I$.

Theorem 3.2. Characteristic vectors corresponding to distinct roots of a real symmetric matrix A are orthogonal.

Proof: Let λ_1 and λ_j be distinct characteristic roots of A . Then there are non-zero column vectors x and y (characteristic vectors of A) such that $Ax = \lambda_1 x$ and $Ay = \lambda_j y$.

From $Ax = \lambda_1 x$ it follows that $x^T A = \lambda_1 x^T$. Multiplying both sides by y , one obtains

$$x^T A y = \lambda_1 x^T y = \lambda_j x^T y.$$

Since $\lambda_1 \neq \lambda_j$, then $x^T y = 0$. Hence x and y are orthogonal.

Definition 3.5. If A and B are n by n matrices over a field F , then A is said to be similar to B over F if there is a non singular matrix P with elements in F such that $A = P^{-1} B P$. If $P^{-1} = P^T$, then A is said to be orthogonally similar to B . If P is a unitary matrix, then A is said to be unitarily similar to B .

Many of the properties of orthogonal similarity are implied by the properties of unitary similarity.

Theorem 3.3. Every real symmetric matrix is orthogonally

similar to a diagonal matrix.

Proof: One proof of the theorem [10, p. 186] shows the existence of an orthonormal basis of the vector space $V_n(R)$, consisting of characteristic vectors of the real symmetric matrix A . If the matrix P is formed by letting these characteristic vectors of A be the columns of P , then it is easily seen that P is orthogonal and $P^T A P = \text{diag. } (k_1, k_2, \dots, k_n)$.

An important property of orthogonal similarity is that two matrices which are orthogonally similar are both similar and congruent. Hence they have the same characteristic roots and if one is symmetric, so is the other.

A similar theorem is true for hermitian matrices, with the substitution of unitarily similar for orthogonally similar and of conjunctive for congruent.

A trivial corollary to Theorem 3.3 is that a real symmetric matrix is orthogonally similar to a Jacobi matrix. This has significance because the reduction to a Jacobi matrix is much more feasible in actual practice than a reduction to diagonal form. Since the characteristic roots of a Jacobi matrix can be calculated without too much difficulty, one has a practical method for obtaining the characteristic roots of real symmetric matrices [6].

Theorem 3.4. Let A be a square matrix of order n over the field F of complex numbers and denote the characteristic roots of A by $\lambda_1, \lambda_2, \dots, \lambda_n$. If $g(x)$ is a polynomial in $F[x]$, then the characteristic roots of $g(A)$ are $g(\lambda_1), g(\lambda_2), \dots, g(\lambda_n)$.

Proof: The proof is indicated as an exercise in Perlis [10, p. 138]. A more concise presentation of the same method is given by Zurmühl [14, p. 228]. The latter proof is as follows:

Let the polynomial $g(x) - y$ be written as

$$g(x) - y = a(x - x_1)(x - x_2) \dots (x - x_k) .$$

Then

$$g(A) - y I = a(A - x_1 I)(A - x_2 I) \dots (A - x_k I)$$

from which it follows that

$$\begin{aligned} |g(A) - y I| &= a^n |A - x_1 I| \cdot |A - x_2 I| \dots |A - x_k I| \\ &= a^n \prod_{i=1}^n (\lambda_i - x_1) \cdot \prod_{i=1}^n (\lambda_i - x_2) \dots \prod_{i=1}^n (\lambda_i - x_k) \\ &= a \prod_{i=1}^k (\lambda_1 - x_i) \cdot a \prod_{i=1}^k (\lambda_2 - x_i) \dots a \prod_{i=1}^k (\lambda_n - x_i) . \end{aligned}$$

It is clear that

$$|g(A) - y I| = \prod_{i=1}^n [g(\lambda_i) - y] .$$

Hence the characteristic roots of $g(A)$ are $g(\lambda_1), g(\lambda_2), \dots, g(\lambda_n)$.

Theorem 3.5. If the characteristic roots of a real symmetric

matrix A are $(\lambda_1 < \lambda_2 < \dots < \lambda_n)$ and the characteristic roots of the matrix B obtained by striking out the last row and the last column of A are $(\delta_1 < \delta_2 < \dots < \delta_{n-1})$ and if $\lambda_i \neq \delta_j$ for every i, j , then $\lambda_1 < \delta_1 < \lambda_2 < \delta_2 < \dots < \delta_{n-1} < \lambda_n$.

Proof: The matrix A has the form

$$A = \begin{pmatrix} B & y^T \\ y & a_{nn} \end{pmatrix} \quad \text{where } y \text{ is a } 1 \text{ by } (n-1)$$

row vector, and a_{nn} is a single element of A . Since B is a real symmetric matrix there is an orthogonal matrix P such that

$$P^T B P = \text{diag. } (\delta_1, \delta_2, \dots, \delta_{n-1}) .$$

The matrix

$$\begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix}$$

is orthogonal and

$$\begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} B & y^T \\ y & a_{nn} \end{pmatrix} \begin{pmatrix} P^T & 0 \\ 0 & 1 \end{pmatrix} = \begin{bmatrix} \delta_1 & & & w_1 \\ & \delta_2 & & \vdots \\ & & \ddots & \vdots \\ & & & \delta_{n-1} & w_{n-1} \\ w_1 & \dots & w_{n-1} & a_{nn} \end{bmatrix} = Q.$$

where w_i is a polynomial in the elements of y and P . Since Q is orthogonally similar to A one knows that the characteristic polynomial of A is given by

$$f(\lambda) = |\lambda I - Q| = (\lambda - a_{nn}) g(\lambda) - \sum_{i=1}^{n-1} w_i^2 \prod_{\substack{j=1 \\ j \neq i}}^{n-1} (\lambda - \delta_j)$$

where $g(\lambda)$ is the characteristic polynomial of B . It is then clear that $w_i \neq 0$ for if $w_i = 0$ then δ_i is a zero of $f(\lambda)$, contrary to the hypothesis. If one lets λ take on values $-\infty, \delta_1, \delta_2, \dots, \delta_{n-1}, +\infty$, the alternation in sign of $f(\lambda)$ implies $\lambda_1 < \delta_1 < \dots < \delta_{n-1} < \lambda_n$. It should be noted that the cases where n is an odd number and where n is an even number must be treated separately.

Removing the restriction that the characteristic roots of A and B be distinct, the result is not a strict separation of roots, since some of the characteristic roots of B might be equal to characteristic roots of A . As an extreme example one might take A to be a diagonal matrix. For a detailed discussion of these aspects of the problem, the reader is referred to Browne [2].

Theorem 3.6. The characteristic roots of $A \oplus B$ are the characteristic roots of A together with those of B .

Proof: The symbol $\oplus$ indicates the direct sum. $A \oplus B$ is defined to be the block matrix

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

The characteristic polynomial of $A \oplus B$ is then $|A - \lambda I| \cdot |B - \lambda I|$ and the theorem follows immediately.

Theorem 3.7. [Parker, p. 105]. If λ is a characteristic root of the matrix A and R is the greatest sum obtained for the absolute values of the elements of a row and T is the greatest sum obtained for the absolute values of the elements of a column, then

$$|\lambda| \leq \min(R, T).$$

Theorem 3.8. [Parker, p. 107]. If $A = (a_{ij})$ is an hermitian matrix with characteristic roots $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, then

$$\frac{1}{2}(\lambda_n - \lambda_1) \geq \max |a_{ij}|, \quad i \neq j.$$

Theorem 3.9. [Brauer, p. 390]. The absolute value of one of the characteristic roots of A equals R (see notation in Theorem 3.7) only if $R_1 = R_2 = \dots = R$ and $\arg(a_{ij}) = \varphi + \varphi_i - \varphi_j$ where φ and $\varphi_1, \varphi_2, \dots, \varphi_n$ are arbitrary real numbers. Then $R e^{i\varphi}$ is a characteristic root of A .

The proofs of the last three theorems are found in the references indicated.

Many other theorems concerning characteristic roots of matrices might be included in this section. It was thought best, however, to include only those which are used most in the thesis. The reader is referred to the bibliography for references to more complete treatments.

IV. GENERAL RESULTS

1. An Oscillation Matrix

It has been briefly indicated that the matrices to be studied arise naturally in certain physical problems. On the other hand, an exact physical application is not always available when one begins to generalize the first basic examples. It is hoped that the extensions will at least be of some mathematical interest.

Consider first, then, the triple diagonal matrix A which arose during the consideration of the modes of vibration of a system consisting of certain masses attached to a stretched string. The solution of the problem depends upon the characteristic roots of the following n by n matrix:

$$A = \begin{bmatrix} a & b & & & & & & \\ b & a & b & & & & & \\ & b & a & b & & & & \\ . & . & . & . & . & . & . & . \\ & & & & b & a & b & \\ & & & & & b & a & \end{bmatrix}$$

where a and b are certain real numbers determined by the conditions of the physical problem.

The task of finding the proper values of A can be reduced by using the theorem that if $A = g(B)$ where $g(x)$ is a rational function, then the characteristic roots of A are given by $g(\lambda_k)$, $k = 1, 2, \dots, n$, where λ_k is a characteristic root of the matrix B . Applying

this theorem (Th. 3.4) here, one sees that $A = a I_n + b B$ where

$$B = \begin{bmatrix} 0 & 1 & & & & \\ 1 & 0 & 1 & & & \\ & 1 & 0 & 1 & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & 1 & 0 & 1 \\ & & & & 1 & 0 \end{bmatrix}.$$

It is then sufficient to know the characteristic roots of B . These can be obtained by considering the characteristic equation of B . Let us denote the characteristic polynomial of B by $\varphi_n(x)$ so that

$$\varphi_n(x) = |xI - B| = \begin{vmatrix} x & -1 & & & \\ -1 & x & -1 & & \\ & -1 & x & -1 & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & -1 & x & -1 \\ & & & & -1 & x \end{vmatrix} = \begin{vmatrix} x & 1 & & & \\ 1 & x & 1 & & \\ & 1 & x & 1 & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & 1 & x & 1 \\ & & & & 1 & x \end{vmatrix}.$$

The last equality follows from elementary properties of determinants. It is also clear from determinant theory that $\varphi_{n+2}(x) - x \varphi_{n+1}(x) + \varphi_n(x) = 0$ where $\varphi_{n+2}(x)$ and $\varphi_{n+1}(x)$ are characteristic polynomials of B when B is of order $(n+2)$ and $(n+1)$ respectively. From the theory of finite differences [7, p. 543] one then knows that

$$\varphi_n(x) = C_1(r_1)^n + C_2(r_2)^n$$

where r_1 and r_2 are roots of the equation $r^2 - x r + 1 = 0$ and C_1 and C_2 are to be determined from two known values of $\varphi_k(x)$. The roots of $r^2 - x r + 1 = 0$ are

$$r_1 = \frac{x + (x^2 - 4)^{\frac{1}{2}}}{2} \quad \text{and} \quad r_2 = \frac{x - (x^2 - 4)^{\frac{1}{2}}}{2}.$$

These are simplified if one makes the substitution $x = 2 \cos \theta$. Then $r_1 = e^{i\theta}$ and $r_2 = e^{-i\theta}$ and the value for $\varphi_n(x)$ becomes

$$\varphi_n(x) = C_1 e^{ni\theta} + C_2 e^{-ni\theta}.$$

It is easily seen that $\varphi_1(x) = x = e^{i\theta} + e^{-i\theta}$ and $\varphi_2(x) = x^2 - 1 = e^{2i\theta} + 1 + e^{-2i\theta}$. Using these values, one finds that

$$C_1 = \frac{e^{i\theta}}{e^{i\theta} - e^{-i\theta}} \quad \text{and} \quad C_2 = -\frac{e^{-i\theta}}{e^{i\theta} - e^{-i\theta}}.$$

Hence

$$\varphi_n(x) = \frac{e^{(n+1)i\theta} - e^{-(n+1)i\theta}}{e^{i\theta} - e^{-i\theta}} = \frac{\sin (n+1)\theta}{\sin \theta}$$

which is a result that will be referred to throughout the paper. One should note that it assumes the indeterminate form $\frac{0}{0}$ when $|x| = 2$. However, if one writes $\varphi_n(x)$ as

$$\varphi_n(x) = \sum_{r=0}^n e^{(n-2r)i\theta} = \frac{e^{(n+1)i\theta} - e^{-(n+1)i\theta}}{e^{i\theta} - e^{-i\theta}}$$

then it is clear that $\varphi_n(2) = n + 1$ and $\varphi_n(-2) = (-1)^n (n + 1)$.

These are never zero so that all the zeroes of $\varphi_n(x)$ can be obtained from the expression

$$\varphi_n(x) = \frac{\sin(n+1)\theta}{\sin \theta}, \quad ,$$

where $x = 2 \cos \theta$.

So the characteristic roots of the matrix B , where

$$B = \begin{bmatrix} 0 & 1 & & & & & \\ 1 & 0 & 1 & & & & \\ & 1 & 0 & 1 & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & 1 & 0 & 1 & \\ & & & & 1 & 0 & \end{bmatrix}$$

are the roots of its characteristic equation

$$\varphi_n(x) = \frac{\sin(n+1)\theta}{\sin \theta} = 0.$$

It is easily seen that $\varphi_n(x)$ is zero for $\theta = \frac{k\pi}{n+1}$, $k = 1, 2, \dots, n$. Hence the characteristic roots of B are $2 \cos \frac{k\pi}{n+1}$, $k = 1, 2, \dots, n$. It is then immediate that the characteristic roots of the matrix A , where

$$A = \begin{bmatrix} a & b & & & \\ b & a & b & & \\ & b & a & b & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & a & b \\ & & & & b & a \end{bmatrix}$$

are $\lambda_k = a + 2b \cos \frac{k\pi}{n+1}$, $k = 1, 2, \dots, n$.

This result was obtained by Gantmakher and Kreĭn [4, p. 134] who also found an explicit formula for the characteristic vectors of A . If

$$v^j = \begin{bmatrix} v_1^j \\ v_2^j \\ \vdots \\ v_n^j \end{bmatrix}$$

is the characteristic vector corresponding to the characteristic root λ_j , then it is shown that $v_k^j = C \sin k \theta_j$ where

$$\cos \theta_j = \frac{a - \lambda_j}{2b},$$

and C is some constant. When one substitutes the value already found for λ_j , it follows that

$$v_k^j = C \sin \frac{k j \pi}{n+1}.$$

Actually, this is shown in Rutherford's paper [13, p. 241] .
 Before considering this section, however, it is necessary to discuss
 some concepts of power series of matrices.

It is well known that the series

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

converges for every number x . If A is a square matrix, it is
 natural to define $\cos A$ as

$$\cos A = I - \frac{1}{2!} A^2 + \frac{1}{4!} A^4 - \dots$$

One would then ask what is meant by the convergence of such a series
 as $\cos A$. In general, the answer is that a series $S_p = \sum_{r=0}^p \alpha_r A^r$,
 where A is a square matrix and α_r is a real number, is said to
 converge as $p \rightarrow \infty$ if every element of S_p converges as $p \rightarrow \infty$.
 It can then be shown that the series $\cos A$ converges for every
 square matrix A . Furthermore, since $A = \text{diag. } (a_1, a_2, \dots, a_n)$
 implies $A^k = \text{diag. } (a_1^k, a_2^k, \dots, a_n^k)$, it follows that if $A =$
 $\text{diag. } (a_1, a_2, \dots, a_n)$ then $\cos A = \text{diag. } (\cos a_1, \cos a_2, \dots, \cos a_n)$.

One can now develop a formula for the characteristic vectors of
 the matrix

$$A = \begin{bmatrix} a & b & & & & & \\ b & a & b & & & & \\ & b & a & b & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & a & b & \\ & & & & b & a & \end{bmatrix} .$$

Let

$$A(0) = \begin{bmatrix} 0 & b & & & & & \\ b & 0 & b & & & & \\ & b & 0 & b & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & 0 & b & \\ & & & & b & 0 & \end{bmatrix} .$$

Then $A = a I + A(0)$. It is asserted that

$$(3) \quad A = a I + 2 b \cos \theta_n$$

where $\theta_n = B_n \psi_n B_n$ in which B_n is a matrix whose r, s -th element is $(2/n+1)^{1/2} \sin \frac{r s \pi}{n+1}$ and

$$\psi_n = \frac{\pi}{n+1} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & n \end{bmatrix} .$$

The proof of (3) begins by proving $B_n^2 = I$. From the definition

of B_n it follows that if c_{ij} is the i, j -th element of $C = B_n^2$ then

$$c_{ij} = \frac{2}{n+1} \sum_{x=1}^n \sin \frac{i x \pi}{n+1} \sin \frac{j x \pi}{n+1} .$$

By a trigonometric substitution one obtains

$$\begin{aligned} c_{ij} &= \frac{1}{n+1} \sum_{x=1}^n \cos \frac{(i-j)x\pi}{n+1} \\ &\quad - \frac{1}{n+1} \sum_{x=1}^n \cos \frac{(i+j)x\pi}{n+1} . \end{aligned}$$

One now needs the fact (which is easily proved) that

$$\sum_{x=1}^n \cos \frac{x r \pi}{n+1} = \begin{cases} 0 & \text{if } r \text{ is an odd number} \\ -1 & \text{if } r \text{ is an even number} . \end{cases}$$

Using this it is immediate that

$$c_{ij} = \frac{n}{n+1} + \frac{1}{n+1} = 1 \quad \text{if } i = j .$$

$$c_{ij} = 0 - 0 = 0 \quad \text{if } (i-j) \text{ is an odd number.}$$

$$c_{ij} = -\frac{1}{n+1} + \frac{1}{n+1} = 0 \quad \text{if } (i-j) \text{ is an even number.}$$

It has now been shown that $B_n^2 = I$. From this it follows that

$$\cos \theta_n = \cos (B_n \psi_n B_n) = B_n (\cos \psi_n) B_n ,$$

which is obtained by examining the cosine series for θ_n . The proof of assertion (3) will be completed by showing that $A(0) = B_n (2b \cos \psi_n) B_n$. If the p, q -th element of $B_n (2b \cos \psi_n) B_n$ is d_{pq} then

$$(4) \quad d_{pq} = \frac{4b}{n+1} \sum_{x=1}^n \sin \frac{p x \pi}{n+1} \cos \frac{x \pi}{n+1} \sin \frac{x q \pi}{n+1} ,$$

which is obtained in a straightforward manner and makes use of the fact that

$$2b \cos \psi_n = \begin{bmatrix} 2b \cos \frac{\pi}{n+1} & & & \\ & 2b \cos \frac{2\pi}{n+1} & & \\ & & 2b \cos \frac{3\pi}{n+1} & \\ \dots & \dots & \dots & \dots \\ & & & 2b \cos \frac{n\pi}{n+1} \end{bmatrix} .$$

By a trigonometric substitution, d_{pq} can be written as

$$d_{pq} = \frac{b}{n+1} \sum_{x=1}^n \left[\cos \frac{(p-q+1)x\pi}{n+1} + \cos \frac{(p-q-1)x\pi}{n+1} - \cos \frac{(p+q+1)x\pi}{n+1} - \cos \frac{(p+q-1)x\pi}{n+1} \right] .$$

From this it follows that

$$d_{pq} = 0 \quad \text{if} \quad p = q$$

$$d_{pq} = \frac{b}{n+1}(n+1) = b \quad \text{if } |p - q| = 1$$

$$d_{pq} = 0 \quad \text{if } |p - q| > 1.$$

Thus $A(0) = B_n(2b \cos \psi_n)B_n = 2b \cos \theta_n$. Hence $A = aI + A(0) = aI + 2b \cos \theta_n$ and the proof of assertion (3) is complete. It follows that

$$AB_n = B_n(aI + 2b \cos \psi_n).$$

Since the diagonal elements of the diagonal matrix $(aI + 2b \cos \psi_n)$ are exactly the characteristic roots of A , then the columns of B_n are the characteristic vectors of A . But the elements of the j -th column of B_n are $v_k^j = (2/(n+1))^{1/2} \sin \frac{kj\pi}{n+1}$, $k = 1, 2, \dots, n$, which is the result obtained by Gantmakher and Krein [4, p. 134].

2. The Fibonacci Numbers

As a curiosity, the result that

$$\varphi_n(x) = \begin{vmatrix} x & 1 & & & \\ 1 & x & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & x & 1 \\ & & & 1 & x \end{vmatrix} = x \varphi_{n-1}(x) - \varphi_{n-2}(x) = \prod_{k=1}^n (x - 2 \cos \frac{k\pi}{n+1})$$

can be related to Fibonacci numbers. Consider the numbers 0, 1, 2, 3, 5, 8, 13, ... where the i -th number ($i > 3$) is the sum of the two

numbers immediately preceding it and the first three numbers are 0, 1 and 2. It is interesting to note that the $(n + 1)$ st number in the sequence is $i^n \varphi_n(-i)$. For if one lets $x = -i$ and multiplies $\varphi_n(x) = x \varphi_{n-1}(x) - \varphi_{n-2}(x)$ through by i^n , it follows that

$$i^n \varphi_n(-i) = (i)^{n-1} \varphi_{n-1}(-i) - i^n \varphi_{n-2}(-i) =$$

$$i^{n-1} \varphi_{n-1}(-i) + (i)^{n-2} \varphi_{n-2}(-i) \quad .$$

The facts that $i \varphi_1(-i) = -i^2 = 1$ and $i^2 \varphi_2(-i) = i^4 - i^2 = 2$ complete the inductive process. Hence the n -th Fibonacci number could be written as

$$(i)^{n-1} \varphi_{n-1}(-i) = (i)^{n-1} \prod_{k=1}^{n-1} (-i - 2 \cos \frac{k\pi}{n}) =$$

$$\prod_{k=1}^{n-1} (1 - 2i \cos \frac{k\pi}{n}) \quad .$$

3. Variations of the Oscillation Matrix

The next question might be a consideration of matrices which are slight variations of the matrix A where

$$A = \begin{bmatrix} a & b & & & & \\ & b & a & b & & \\ & & b & a & b & \\ & & & & b & a & b \\ & & & & & b & a \end{bmatrix} \quad .$$

If some of the elements are changed, could the characteristic roots still be found? These changes might be due to changes in initial conditions in the physical problem or to merely an arbitrary extension to more general cases. In any case one might suspect that variations in the elements of A would lead to quite different results.

First, suppose the first and last elements on the main diagonal of A are changed, thus obtaining

$$R_n = \begin{bmatrix} a+d & b & & & \\ & b & a & b & \\ & & b & a & b \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & a & b \\ & & & & b & a+c \end{bmatrix}.$$

Proceeding as before, we note that $R_n = a I + b C$, where

$$C = \begin{bmatrix} \frac{d}{b} & 1 & & & \\ 1 & 0 & 1 & & \\ & 1 & 0 & 1 & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & 1 & 0 & 1 \\ & & & & 1 & \frac{c}{b} \end{bmatrix}.$$

The proper values of C are obtained from the zeroes of

$$|x I - C| = \varphi_n(x) - \frac{c+d}{b} \varphi_{n-1}(x) + \frac{cd}{b^2} \varphi_{n-2}(x)$$

where $\varphi_k(x)$ has the same meaning as before. Substituting the values

already found for $\phi_k(x)$ it follows that

$$|x I - C| = \frac{\sin(n+1)\theta - \frac{c+d}{b} \sin n\theta + \frac{cd}{b^2} \sin(n-1)\theta}{\sin \theta}.$$

In particular, consider the following cases as examples:

1) If $c = -d = b$, then

$$|x I - C| = \frac{\sin(n+1)\theta - \sin(n-1)\theta}{\sin \theta} = 2 \cos n\theta$$

which vanishes if $\theta = (2k-1)\pi/2n$, $k = 1, 2, \dots, n$. The proper values of the matrix C are therefore $\gamma_k = 2 \cos \frac{(2k-1)\pi}{2n}$, $k = 1, 2, \dots, n$. This implies in turn that the characteristic roots of R_n , subject to the restrictions noted, are $\alpha_k = a + 2b \cos \frac{(2k-1)\pi}{2n}$, $k = 1, 2, \dots, n$.

2) If $c = 0$ and $d = -b$, then

$$|x I - C| = \frac{\sin(n+1)\theta + \sin n\theta}{\sin \theta} = \frac{\sin(\frac{2n+1}{2}\theta)}{\sin \frac{\theta}{2}}$$

which is zero if $\theta = 2k\pi/(2n+1)$, $k = 1, 2, \dots, n$. For this example, the matrix C has characteristic roots $\gamma_k = 2 \cos \frac{2k\pi}{2n+1}$, $k = 1, 2, \dots, n$ and the corresponding R_n has characteristic roots $\alpha_k = a + 2b \cos \frac{2k\pi}{2n+1}$, $k = 1, 2, \dots, n$.

3) If $c = d = b$, then

$$\begin{aligned}
 |x I - C| &= \frac{\sin (n+1)\theta - 2 \sin n\theta + \sin (n-1)\theta}{\sin \theta} \\
 &= \frac{2 \sin n\theta (\cos \theta - 1)}{\sin \theta}
 \end{aligned}$$

which vanishes if $\theta = k\pi/n$, $k = 1, 2, \dots, n$. It should be noted that $\theta = \pi$ gives an indeterminate form but $x = -2$ is easily verified directly to be a zero of $|x I - C|$. Hence the characteristic roots of

$$R_n = \begin{bmatrix} a+b & b & & & & \\ & b & a & b & & \\ & & b & a & b & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & a & b \\ & & & & b & a+b \end{bmatrix}$$

are $\lambda_k = a + 2b \cos \frac{k\pi}{n}$, $k = 1, 2, \dots, n$.

One special combination of results already obtained may be mentioned. Suppose

$$\varphi_n(x) = \begin{vmatrix} x & 1 & & & \\ 1 & x & 1 & & \\ & 1 & x & 1 & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & 1 & x & 1 \\ & & & & 1 & x \end{vmatrix}$$

and

$$\varphi_n(y) = \begin{vmatrix} y & 1 & & & & \\ 1 & y & 1 & & & \\ & 1 & y & 1 & & \\ . & . & . & . & . & . \\ & & & 1 & y & 1 \\ & & & & 1 & y \end{vmatrix}$$

where $\varphi_n(x)$ is the determinant of a matrix $P_n(x)$ and $\varphi_n(y)$ is the determinant of a matrix $P_n(y)$. Then

$$|P_n(x) P_n(y)| = \begin{vmatrix} (z-1), & 2a, & 1 & & & \\ 2a, & z, & 2a, & 1 & & \\ 1, & 2a, & z, & 2a, & 1 & \\ . & . & . & . & . & . \\ & & & & 1, & 2a, & z, & 2a \\ & & & & & 1, & 2a, & (z-1) \end{vmatrix}$$

where $z = xy + 2$ and $2a = (x + y)$. Since it has already been shown that

$$\varphi_n(r) = \prod_{k=1}^n (r - 2 \cos \frac{k\pi}{n+1}),$$

then it follows that

$$\begin{aligned} |P_n(x) P_n(y)| &= \prod_{k=1}^n (x - 2 \cos \frac{k\pi}{n+1}) \prod_{k=1}^n (y - 2 \cos \frac{k\pi}{n+1}) = \\ &\prod_{k=1}^n [z - (2 + a)^2 + (a - 2 \cos \frac{k\pi}{n+1})^2] . \end{aligned}$$

This implies that the characteristic roots of

$$\begin{bmatrix} 1, & -2a, & -1 & & & \\ -2a, & 0, & -2a, & -1 & & \\ -1, & -2a, & 0, & -2a, & -1 & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ & & & & -1, & -2a, & 0, & -2a \\ & & & & & -1, & -2a, & 1 \end{bmatrix} \quad n \text{ by } n$$

are $2 + a^2 - (a - 2 \cos \frac{k\pi}{n+1})^2$, $k = 1, 2, \dots, n$.

4. Alternating values on the diagonal

At least one case in which the terms on the main diagonal alternate in value can be solved by the results already obtained. Let

$$S_{2n} = \begin{bmatrix} a & 1 & & & & \\ 1 & b & 1 & & & \\ & 1 & a & 1 & & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ & & & 1 & a & 1 \\ & & & & 1 & b \end{bmatrix} \quad 2n \text{ by } 2n$$

Then

$$|\lambda I_{2n} - S_{2n}| = \begin{vmatrix} (\lambda-a), & -1 & & & & \\ -1, & (\lambda-b), & & & & \\ & -1, & (\lambda-a), & -1 & & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ & & & -1, & (\lambda-a), & -1 \\ & & & & -1, & (\lambda-b) \end{vmatrix}.$$

It is easily seen that $\lambda = a$ or $\lambda = b$ are not zeroes of $|\lambda I_{2n} - S_{2n}|$. Hence one is permitted to multiply each row containing $(\lambda - a)$ by $(\lambda - b)^{1/2}$, multiply each row containing $(\lambda - b)$ by $(\lambda - a)^{1/2}$, divide each column containing $(\lambda - a)$ by $(\lambda - a)^{1/2}$ and divide each column containing $(\lambda - b)$ by $(\lambda - b)^{1/2}$. One thus obtains the result that

$$|\lambda I_{2n} - S_{2n}| = \begin{vmatrix} z, & -1 & & & \\ -1, & z, & -1 & & \\ & -1, & z, & -1 & \\ \dots & \dots & \dots & \dots & \dots \\ & & & -1, & z, & -1 \\ & & & & -1, & z \end{vmatrix}$$

$$= \frac{\sin (2n + 1)\theta}{\sin \theta}$$

where $z = \left[(\lambda - a)(\lambda - b) \right]^{1/2} = 2 \cos \theta$. It follows easily that the characteristic roots of S_{2n} are

$$\frac{a + b \pm \left[(a - b)^2 + 16 \cos^2 \frac{k\pi}{2n+1} \right]^{1/2}}{2},$$

$k = 1, 2, \dots, n$.

By a similar argument one finds that the characteristic roots of the matrix

$$S_{2n+1} = \begin{bmatrix} a & 1 & & & \\ 1 & b & 1 & & \\ & 1 & a & 1 & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & 1 & b & 1 \\ & & & & 1 & a \end{bmatrix}$$

are

$$\frac{a + b \pm \left[(a - b)^2 + 16 \cos^2 \frac{k\pi}{2n+2} \right]^{1/2}}{2},$$

$k = 1, 2, \dots, n$, together with the characteristic root $\lambda = a$ which may be verified directly.

5. The Inverse of a Certain Matrix

Another possible means of extending the results thus far obtained is to find inverses for some of the matrices for which characteristic roots have already been found. This can be done in a straightforward manner in at least one case. Consider the matrix

$$R_n(x, a, b) = \begin{bmatrix} x+b & 1 & & & \\ 1 & x & 1 & & \\ & 1 & x & 1 & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & 1 & x & 1 \\ & & & & 1 & x+a \end{bmatrix}.$$

The inverse of $R_n(x, a, b)$ is found by making use of the relationship

$$|R_n(x, a, b)| = \varphi_n(x) + (a + b) \varphi_{n-1}(x) + a b \varphi_{n-2}(x)$$

where the value of $\varphi_k(x)$ is the same as before.

If the r, s -th element of $[R_n(x, a, b)]^{-1}$ is denoted by h_{rs} , it is not too difficult to see that

$$h_{rs} = h_{sr} = \frac{(-1)^{r+s} |R_{r-1}(x, 0, b)| \cdot |R_{n-s}(x, a, 0)|}{|R_n(x, a, b)|}, \quad r \leq s.$$

This is obtained by considering the cofactors of the elements of $R_n(x, a, b)$ divided by the determinant of the matrix. If one substitutes the value found previously for $\varphi_n(x)$, then

$$h_{rs} = (-1)^{r+s} \frac{[\sin r\theta + b \sin(r-1)\theta] [\sin(n-s+1)\theta + a \sin(n-s)\theta]}{\sin \theta [\sin(n+1)\theta + (a+b) \sin n\theta + a b \sin(n-1)\theta]}.$$

This expression can be simplified by letting $a = -\frac{\sin \beta \theta}{\sin(\beta+1)\theta}$ and $b = -\frac{\sin \alpha \theta}{\sin(\alpha+1)\theta}$. By certain trigonometric manipulations one gets

$$\begin{aligned} h_{rs} &= \frac{\sin^2 \theta \sin(r + \alpha)\theta \sin(n + \beta - s + 1)\theta}{\sin^3 \theta \sin(n + \beta + \alpha + 1)\theta} \\ &= \frac{\sin(r + \alpha)\theta \sin(n + \beta - s + 1)\theta}{\sin \theta \sin(n + \beta + \alpha + 1)\theta} \\ &= \frac{(-1)^{r+s} \varphi_{\alpha+r-1}(x) \varphi_{n+\beta-s}(x)}{\varphi_{n+\beta+\alpha}(x)}. \end{aligned}$$

It should be noted that the last form of h_{rs} is not valid

unless we extend our definition of $\varphi_n(x)$ so that even when n is not an integer, we have

$$\varphi_n(2 \cos \theta) = \frac{\sin (n+1)\theta}{\sin \theta} \quad , \quad \theta \neq 0, \pi$$

and $\varphi_n(2) = n+1$, $\varphi_n(-2) = (-1)^n (n+1)$. This must be done because, in general, α and β would not be integers.

As a simple example to illustrate the method suppose one wishes to find the inverse of $R_3(2, 0, 0)$. Then

$$h_{rs} = h_{sr} = (-1)^{r+s} \frac{(r)(4-s)}{4} \quad , \quad r \leq s$$

from which it follows that

$$\left[R_3(2, 0, 0) \right]^{-1} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}^{-1} = \frac{1}{4} \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 3 \end{bmatrix} .$$

Since the characteristic roots of $R_3(2, 0, 0)$ are $2 + 2 \cos \frac{k\pi}{4}$, $k = 1, 2, 3$, then the characteristic roots of

$$Q = \begin{bmatrix} 3 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 3 \end{bmatrix}$$

are $q_k = 2/(1 + \cos \frac{k\pi}{4}) = \sec^2 \frac{k\pi}{8}$, $k = 1, 2, 3$.

Another case where the matrix is of order n gives the result that

$$-2 [R_n(-2,1,-1)]^{-1} = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 3 & 3 & \dots & 3 \\ 1 & 3 & 5 & \dots & 5 \\ \vdots & \vdots & \vdots & & \\ 1 & 3 & 5 & & 2n-1 \end{bmatrix} = H \quad .$$

It follows that the characteristic roots of H are $\frac{1}{2} \csc^2 \frac{(2k-1)\pi}{4n}$, $k = 1, 2, \dots, n$.

6. Symmetric Circulants

Another class of matrices for which it is possible to find exact solutions for characteristic roots are the matrices which are called circulants. A circulant matrix is a matrix of the form

$$E = \begin{bmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \dots & a_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & a_2 & a_3 & \dots & a_0 \end{bmatrix}$$

where the elements of the i -th row are the elements of the $(i-1)$ st row permuted in the cyclic manner indicated. If

$$K = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

then

$$E = a_0 I + a_1 K + a_2 K^2 + \dots + a_{n-1} K^{n-1}.$$

Hence the characteristic roots of E can be obtained if the characteristic roots of K are known. The characteristic equation of K is easily seen to be

$$\begin{vmatrix} x & -1 & 0 & 0 & \dots & 0 \\ 0 & x & -1 & 0 & \dots & 0 \\ 0 & 0 & x & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & x & -1 \\ -1 & 0 & 0 & 0 & \dots & x & \end{vmatrix} = x^n - 1 = 0.$$

It follows that the characteristic roots of K are $e^{\frac{2k\pi i}{n}}$, $k = 1, 2, \dots, n$. Hence the characteristic roots of the circulant E are

$$\alpha_k = a_0 + a_1 e^{\frac{2k\pi i}{n}} + a_2 e^{\frac{4k\pi i}{n}} + \dots + a_{n-1} e^{\frac{2(n-1)k\pi i}{n}},$$

$k = 1, 2, \dots, n$. If E is symmetric so that $a_r = a_{n-r}$, then the proper values are

$$\begin{aligned} \alpha_k = a_0 + a_1 \left(e^{\frac{2k\pi i}{n}} + e^{\frac{2(n-1)k\pi i}{n}} \right) + a_2 \left(e^{\frac{4k\pi i}{n}} + e^{\frac{2(n-2)k\pi i}{n}} \right) \\ + \dots + L, \quad k = 1, 2, \dots, n \end{aligned}$$

Hence $\alpha_k = a_0 + 2a_1 \cos \frac{2k\pi}{n} + 2a_2 \cos \frac{4k\pi}{n} + \dots + L$, $k = 1, 2, \dots, n$

where L is $2 a_s \cos \frac{2sk\pi}{n}$ if $n = 2s + 1$ and $L = (-1)^k a_s$ if $n = 2s$.

As a special case, choose $a_1 = a_{n-1} = 1$ and $a_i = 0$ for $i \neq 1, n-1$. This gives a matrix of the form

$$G = \begin{bmatrix} 0 & 1 & & & 1 \\ 1 & 0 & 1 & & \\ & 1 & 0 & 1 & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & & & 1 & 0 \end{bmatrix}.$$

It is immediate that the characteristic roots of G are $2 \cos \frac{2k\pi}{n}$, $k = 1, 2, \dots, n$. Actually, one could just as easily obtain the more general result that the characteristic roots of a matrix

$$H = \begin{bmatrix} a & b & & & b \\ b & a & b & & \\ & b & a & b & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ b & & & b & a \end{bmatrix}$$

are $a + 2b \cos \frac{2k\pi}{n}$, $k = 1, 2, \dots, n$.

7. An Extension by using the Direct Product

A further extension to other matrices for which the characteristic roots can be found requires another theorem.

Theorem 4.1. If G and H are real symmetric matrices of order n and m respectively, then the characteristic roots of $G(H) = G \otimes I_m + I_n \otimes H$ are $\lambda_i + \alpha_j$ where $\lambda_1, \dots, \lambda_n$ are the characteristic

roots of G and $\alpha_1, \dots, \alpha_m$ are the characteristic roots of H .

The symbol $\otimes$ is used to indicate the direct product. For a discussion of direct product the reader is referred to MacDuffee [8, p. 82].

Proof: Since G and H are real symmetric matrices it is possible to find orthogonal matrices P_1 and P_2 such that $P_1 G P_1^T = \text{diag. } (\lambda_1, \dots, \lambda_n)$ and $P_2 H P_2^T = \text{diag. } (\alpha_1, \dots, \alpha_m)$. If $P_1 G P_1^T$ and $P_2 H P_2^T$ are denoted by D_1 and D_2 respectively, then

$$(P_1 \otimes P_2) [G(H)] (P_1 \otimes P_2)^T = (P_1 \otimes P_2)(G \otimes I_m)(P_1^T \otimes P_2^T) +$$

$$(P_1 \otimes P_2)(I_n \otimes H)(P_1^T \otimes P_2^T) = D_1 \otimes I_m + I_n \otimes D_2 = D_p$$

where D_p is a diagonal matrix with the characteristic roots of $G(H)$ along its diagonal. The matrix D_p is of the following form:

$$D_p = \text{diag. } (B_1, B_2, \dots, B_n) + \text{diag. } (C_1, C_2, \dots, C_n)$$

where

$$B_1^{m \text{ by } m} = \text{diag. } (\lambda_1, \lambda_1, \dots, \lambda_1)$$

and

$$C_1^{m \text{ by } m} = \text{diag. } (\alpha_1, \alpha_2, \dots, \alpha_m) .$$

Hence

$$D_p^{mn \text{ by } mn} = \text{diag. } (K_1, K_2, K_3, \dots, K_n)$$

where

$$K_i = \text{diag. } (\lambda_i + \alpha_1, \lambda_i + \alpha_2, \dots, \lambda_i + \alpha_m)$$

and the proof is complete.

As an example of an application of this theorem, suppose

$$G^{n \text{ by } n} = \begin{bmatrix} a & b & & & & & \\ b & a & b & & & & \\ & b & a & b & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & a & b & \\ & & & & b & a & \end{bmatrix}$$

and

$$H^{m \text{ by } m} = \begin{bmatrix} d & c & & & & & \\ c & d & c & & & & \\ & c & d & c & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & c & d & c & \\ & & & & c & d & \end{bmatrix}.$$

Then $G(H)$ is the matrix in which the diagonal elements of G are replaced by $(a I_m + H)$ and the elements b are replaced by $b I_m$, with zeroes elsewhere to complete a matrix of order $m n$. We know that the characteristic roots of G are

$$\lambda_k = a + 2b \cos \frac{k\pi}{n+1}, \quad k = 1, 2, \dots, n$$

and the characteristic roots of H are

$$\alpha_j = d + 2c \cos \frac{j\pi}{m+1}, \quad j = 1, 2, \dots, m.$$

It follows that the characteristic roots of $G(H)$ are

$$a + d + 2b \cos \frac{k\pi}{n+1} + 2c \cos \frac{j\pi}{m+1}$$

where $k = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$, with all possible combinations being taken.

8. Bounds for Characteristic Roots

Let us now consider the results obtained for at least one case with reference to bounds for characteristic roots. Specifically, take

$$A = \begin{bmatrix} a & b & & & & \\ b & a & b & & & \\ & b & a & b & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & a & b \\ & & & & b & a \end{bmatrix}$$

for which the characteristic roots are $\lambda_k = a + 2b \cos \frac{k\pi}{n+1}$,
 $k = 1, 2, \dots, n$ and

$$B = \begin{bmatrix} a+b & b & & & \\ & b & a & b & \\ & & b & a & b \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & b & a & b \\ & & & & b & a+b \end{bmatrix}$$

with characteristic roots $\alpha_k = a + 2b \cos \frac{k\pi}{n}$, $k = 1, 2, \dots, n$.

For matrix A we have that

$$|\lambda_k| = \left| a + 2b \cos \frac{k\pi}{n+1} \right| \leq |a| + 2|b| \cdot \left| \cos \frac{k\pi}{n+1} \right| <$$

$$|a| + 2|b|.$$

Here, then, we have a strict inequality for Theorem 3.7 which concluded that $|\lambda_k| \leq |a| + 2|b|$. That strict inequality does not hold in general is immediately seen in the case of matrix B. For matrix B we have

$$|\alpha_k| \leq |a| + 2|b| \left| \cos \frac{k\pi}{n} \right| \leq |a| + 2|b|$$

with the equality holding when $a > 0$, $b < 0$ and $k = n$. General conditions necessary for the equality to hold are given in Theorem 3.9.

Let us examine the spread of the characteristic roots of A and B relative to Theorem 3.8, where the spread is indicated as

$$\frac{1}{2}(\lambda_n - \lambda_1) \text{ and } \frac{1}{2}(\alpha_n - \alpha_1) \text{ if } \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \text{ and}$$

$\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$. Then

$$\frac{1}{2}(\lambda_n - \lambda_1) = b(2 \cos \frac{\pi}{n+1}) \geq b \quad \text{if } b > 0$$

and

$$\frac{1}{2}(\alpha_n - \alpha_1) = (-b) (1 + \cos \frac{\pi}{n}) > |b| \quad \text{if } b < 0 .$$

One then notes that the equality holds in the first case for $n = 2$ but that strict inequality holds for the characteristic roots of B .

Much work is currently being done in this area of bounds for characteristic roots. Results that are sharper than those listed in this thesis are contained in the papers of Brauer [3] together with references to other writers in this field. However, the notation and background that must be built up limits us here to this brief mention of a topic that seems to offer opportunities for much further exploration.

9. The General Case

It would clearly be desirable to consider the matrix

$$A^{n \text{ by } n} = \begin{bmatrix} a_1 & b_1 & & & & \\ b_1 & a_2 & b_2 & & & \\ & b_2 & a_3 & b_3 & & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ & & & b_{n-2} & a_{n-1} & b_{n-1} \\ & & & & b_{n-1} & a_n \end{bmatrix}$$

One can at least prove the following theorem:

Theorem 4.2. Given $[5]$. If A has a characteristic root of multiplicity k , then at least $(k - 1)$ of the b_i 's must be zero.

Proof: The theorem is shown true for $k = 2$ and extended by induction. If λ_j is a double characteristic root of A , then $(A - \lambda_j I)$ has rank at most $(n - 2)$. By striking out the first row and last column of $(A - \lambda_j I)$ one obtains immediately an $(n - 1)$ by $(n - 1)$ submatrix whose determinant is $\prod b_i$. If the rank of $(A - \lambda_j I)$ is at most $(n - 2)$, then $\prod b_i$ equals zero. Hence some b_i is zero. One can then use the theorem that the characteristic roots of the direct sum of two matrices B and C are the characteristic roots of B together with those of C . For if λ_j is a characteristic root of A of multiplicity k , then certainly at least one b_i is zero. Then A can be represented as a direct sum of two matrices which we may call B and C . If $k > 2$ then either B or C (or both) must have at least a double characteristic root; hence, another b_i is zero. Continuing in this way it follows that if λ_j is of multiplicity k , then at least $(k - 1)$ of the b_i 's are zero.

Another result applicable to the general case is contained in Gantmakher and Kreĭn $[4, p. 134]$ and concerns bounds for the largest and smallest characteristic roots of A . If the characteristic roots of A are $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, then the results obtained are

$$\min a_i - 2\left(\cos \frac{\pi}{n+1}\right)(\max b_i) \leq \lambda_1 \leq \max a_i - 2\left(\cos \frac{\pi}{n+1}\right)(\min b_i)$$

$$\min a_i + 2\left(\cos \frac{\pi}{n+1}\right)(\min b_i) \leq \lambda_n \leq \max a_i + 2\left(\cos \frac{\pi}{n+1}\right)(\max b_i) .$$

A general method for calculating the characteristic roots of A for arbitrary a_i and b_i is given in [6] .

10. Summary

Characteristic roots were found for several classes of real symmetric matrices. It was thought desirable to tabulate these results in order to make them more readily accessible.

<u>Matrix</u>	<u>Characteristic Roots</u>
$\begin{bmatrix} a & b & & & \\ & b & a & & \\ & & b & a & \\ \dots & & & & \dots \\ & & & b & a & b \\ & & & & b & a \end{bmatrix}$	$\lambda_k = a + 2b \cos \frac{k\pi}{n+1} , \quad k = 1, 2, \dots, n .$

$\begin{bmatrix} a-b & b & & & \\ & b & a & & \\ & & b & a & b \\ \dots & & & & \dots \\ & & & b & a & b \\ & & & & b & a \end{bmatrix}$	$\lambda_k = a + 2b \cos \frac{2k\pi}{2n+1} , \quad k = 1, 2, \dots, n .$
--	---

$\begin{bmatrix} a-b & b & & & \\ & b & a & & \\ & & b & a & b & \dots \\ \dots & & & & & \dots \\ & & & & b & a & b \\ & & & & & b & a+b \end{bmatrix}$	$\lambda_k = a + 2b \cos \frac{(2k-1)\pi}{2n} , \quad k = 1, 2, \dots, n .$
--	---

$$\begin{bmatrix}
 a+b & b & & & \\
 b & a & b & & \\
 & b & a & b & \\
 \dots & \dots & \dots & \dots & \dots \\
 & & & b & a & b \\
 & & & & b & a+b
 \end{bmatrix}^{n \text{ by } n} \quad \lambda_k = a + 2b \cos \frac{k\pi}{n}, \quad k = 1, 2, \dots, n.$$

$$\begin{bmatrix}
 -1 & 2a & 1 & & \\
 2a & 0 & 2a & & \\
 1 & 2a & 2a & 1 & \\
 \dots & \dots & \dots & \dots & \dots \\
 & & & 1 & 2a & 0 & 2a \\
 & & & & 1 & 2a & -1
 \end{bmatrix}^{n \text{ by } n} \quad \lambda_k = (a - 2 \cos \frac{k\pi}{n+1})^2 - (a^2 + 2), \quad k = 1, 2, \dots, n.$$

$$\begin{bmatrix}
 a & 1 & & & \\
 1 & b & 1 & & \\
 & 1 & a & 1 & \\
 \dots & \dots & \dots & \dots & \dots \\
 & & & 1 & a & 1 \\
 & & & & 1 & b
 \end{bmatrix}^{2n \text{ by } 2n} \quad \lambda = \frac{a+b \pm \left[(a-b)^2 + 16 \cos^2 \frac{k\pi}{2n+1} \right]^{1/2}}{2}, \quad k = 1, 2, \dots, n.$$

$$\begin{bmatrix}
 a & 1 & & & \\
 1 & b & 1 & & \\
 & 1 & a & 1 & \\
 \dots & \dots & \dots & \dots & \dots \\
 & & & 1 & b & 1 \\
 & & & & 1 & a
 \end{bmatrix}^{(2n+1) \text{ by } (2n+1)} \quad \lambda = \frac{a+b \pm \left[(a-b)^2 + 16 \cos^2 \frac{k\pi}{2n+2} \right]^{1/2}}{2}, \quad k = 1, 2, \dots, n \text{ and } \lambda=a.$$

$$\begin{bmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \dots & a_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & a_2 & a_3 & \dots & a_0 \end{bmatrix}^{n \text{ by } n}$$

If $a_r = a_{n-r}$ then

$$\lambda_k = a_0 + 2a_1 \cos \frac{2k\pi}{n} + 2a_2 \cos \frac{4k\pi}{n} + 2a_3 \cos \frac{6k\pi}{n} + \dots + L, \quad k = 1, 2, \dots, n$$

where $L = 2a_s \cos \frac{2sk\pi}{n}$ if $n = 2s + 1$ and $L = (-1)^k a_s$ if $n = 2s$.

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 3 & 3 & \dots & 3 \\ 1 & 3 & 5 & \dots & 5 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 3 & 5 & \dots & (2n-1) \end{bmatrix}^{n \text{ by } n}$$

$$\lambda_k = \frac{1}{2} \csc^2 \frac{(2k-1)\pi}{4n}, \quad k = 1, 2, \dots, n$$

$$\begin{bmatrix} (a I_m + H) & b I_m & & & \\ b I_m & (a I_m + H) & b I_m & & \\ & b I_m & (a I_m + H) & b I_m & \\ \dots & \dots & \dots & \dots & \dots \\ & & & b I_m & (a I_m + H) \end{bmatrix}^{n \text{ by } n \text{ (blocks)}}$$

If

$$H^{m \text{ by } m} = \begin{bmatrix} d & c & & & \\ c & d & c & & \\ & c & d & c & \\ \dots & \dots & \dots & \dots & \dots \\ & & & c & d & c \\ & & & & c & d \end{bmatrix}$$

and

I_m is the unit matrix of order m , then

$$\alpha_{r,k} = a + d + 2b \cos \frac{k\pi}{n+1} + 2c \cos \frac{r\pi}{m+1} ,$$

$$k = 1, 2, \dots, n ; \quad r = 1, 2, \dots, m .$$

BIBLIOGRAPHY

1. Birkhoff, Garrett and MacLane, Saunders. A Survey of Modern Algebra. New York: MacMillan, 1947.
2. Browne, E. T. "On the Separation Property of the Roots of the Secular Equation". Amer. J. Math. 52, 843-850 (1930).
3. Brauer, Alfred. "Limits for the Characteristic Roots of a Matrix". Duke Math. J. 13, 387-395 (1946). (First of a series in this Journal.)
4. Gantmakher, F. and Kreĭn, M. G. Ostsilliat͡sinnye matritsy i i͡adra. (Russian).
5. Givens, Wallace. Private communication.
6. _____. "Numerical Computation of Characteristic Values of a Real Symmetric Matrix". Forthcoming report at the Oak Ridge National Laboratory.
7. Jordan, Charles. Calculus of Finite Differences. New York: Chelsea, 1950.
8. MacDuffee, C. C. The Theory of Matrices. New York: Chelsea, 1946.
9. Parker, W. V. "Characteristic Roots and Field of Values of a Matrix". Bull. Amer. Math. Soc. (2) 57, 103-108 (1951).
10. Perlis, Sam. Theory of Matrices. Cambridge 42, Mass.: Addison-Wesley, 1952.
11. Rutherford, D. E. Classical Mechanics. London: Oliver and Boyd, 1951.
12. _____. "Some Continuant Determinants arising in Physics and Chemistry". Proc. Roy. Soc. Edinburgh. (A) 62, 229-236, 1945.
13. _____. "Some Continuant Determinants arising in Physics and Chemistry II". Proc. Roy. Soc. Edinburgh. (A) 63, 232-241, 1950.
14. Zurmühl, R. Matrizen. Berlin: 1950.