ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 2, Number 1, Winter 1972

PRINCIPAL SUBMATRICES. VIII. PRINCIPAL SECTIONS OF A PAIR OF FORMS[1]

R. C. THOMPSON

ABSTRACT. Let A, C be n-square Hermitian matrices, with C positive definite. Let A_i, C_i denote the principal submatrices obtained by deleting row and column i. In this paper new links are obtained between the roots of the determinantal equations $\det(\lambda C - A) = 0$, $\det(\lambda C_i - A_i) = 0$, $i = 1, \cdots, n$.

Let A be an n-square Hermitian matrix. Let $A(i \mid i)$ denote the principal submatrix of A obtained by deleting from A both row i and column i. In certain earlier papers in this series, links between the roots of

$$(1) \qquad \det(\lambda I_n - A) = 0$$

(the eigenvalues of A) and the roots of

$$(2) \qquad \det(\lambda I_{n-1} - A(i \mid i)) = 0, \qquad i = 1, 2, \cdots, n,$$

(the eigenvalues of $A(i \mid i)$) have been studied. It is of course true that, for each fixed i, the roots of (2) interlace the roots of (1). This well-known fact goes back to Cauchy, and for this reason these interlacing inequalities are often called the Cauchy inequalities.

Let C be an n-square positive definite Hermitian matrix. In this paper we study the roots of the equation

$$(3) \qquad \det(\lambda C - A) = 0$$

and their links to the roots of all of the equations

$$(4) \qquad \det(\lambda C(i \mid i) - A(i \mid i)) = 0, \qquad i = 1, 2, \cdots, n.$$

The equation (3) arises when one attempts a simultaneous diagonalization of a pair of quadratic forms having coefficient matrices C and A. (The possibility of this simultaneous diagonalization is important in applied mathematics, especially in mechanics—see [11].) Suppressing the same variable in each of these forms, the correspond-

Received by the editors July 20, 1970.
AMS 1970 *subject classifications*. Primary 15A39, 15A42, 15A63.
Key words and phrases. Linear algebra, principal submatrix, eigenvalue inequalities.
[1]The preparation of this paper was supported in part by the U.S. Air Force Office of Scientific Research, under Grant 698-67.

ing problem in the reduced forms leads to one of the equations (4). Thus, when studying the roots of the equations (3) and (4), we are studying simultaneously the principal $(n-1)$-sections of a pair of n-variable forms.

The roots of (3) are real (they are the eigenvalues of $C^{-1/2}AC^{-1/2}$). It is a fact, not quite as well known as the Cauchy inequalities, that the roots of any one of the equations (4) interlace the roots of (3) (see [**11**]). In this paper we shall show that some of the new results obtained in [3], [5] concerning the Cauchy inequalities may be extended to the roots of (3) and (4).

NOTATION. Throughout this paper we shall let $\lambda_1 \leqq \cdots \leqq \lambda_n$ be the roots of (3) and let $\eta_{i1} \leqq \cdots \leqq \eta_{i,n-1}$ be the roots of

$$\det(\lambda C(i \mid i) - A(i \mid i)) = 0.$$

The numbers $\lambda_1, \cdots, \lambda_n$ need not be distinct, so let μ_i with multiplicity e_i, for $i = 1, 2, \cdots, s$ be the distinct numbers among $\lambda_1, \cdots, \lambda_n$. We arrange the numbering such that $\mu_1 < \mu_2 < \cdots < \mu_s$. Let $\gamma_1 \leqq \cdots \leqq \gamma_n$ be the eigenvalues of C. Of course $\gamma_1 > 0$. Let $C = XX^*$. Let $f(\lambda) = \det(\lambda C - A)$ and $f_{(i)}(\lambda) = \det(\lambda C(i \mid i) - A(i \mid i))$.

We now develop the formulas upon which our results will be based. These formulas are generalizations of formulas presented in [3]. Let V be a unitary matrix such that $V^*X^{-1}AX^{*-1}V = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$. Then $\lambda C - A = XV \operatorname{diag}(\lambda - \lambda_1, \cdots, \lambda - \lambda_n)V^*X^*$. Set $T = XV$. Then

$$\lambda C - A = T \operatorname{diag}(\lambda - \lambda_1, \cdots, \lambda - \lambda_n)T^*. \tag{5}$$

Hence (regarding λ as a polynomial indeterminate) we obtain

$$(\lambda C - A)^{-1} = T^{*-1} \operatorname{diag}((\lambda - \lambda_1)^{-1}, \cdots, (\lambda - \lambda_n)^{-1})T^{-1}. \tag{6}$$

Now

$$f(\lambda) = \det C(\lambda - \lambda_1) \cdots (\lambda - \lambda_n). \tag{7}$$

Multiplying each side of (6) by $f(\lambda)$, we obtain

$$\operatorname{adj}(\lambda C - A) = T^{*-1} \operatorname{diag}\left(\frac{f(\lambda)}{\lambda - \lambda_1}, \cdots, \frac{f(\lambda)}{\lambda - \lambda_n}\right)T^{-1}. \tag{8}$$

Here adj denotes adjugate. The (i, i) diagonal element of the left side of (8) is $f_{(i)}(\lambda)$. Let $T^{*-1} = (t_{ij})$. Comparing the main diagonal positions in (8), we get

$$f_{(i)}(\lambda) = \sum_{j=1}^{n} |t_{ij}|^2 \frac{f(\lambda)}{\lambda - \lambda_j}. \tag{9}$$

The formula (9) is the basic formula from which all of our results will follow. Since

$$(10) \qquad f(\lambda) = \det C \prod_{j=1}^{s} (\lambda - \mu_j)^{e_j}$$

it follows from (9) that

$$(11) \qquad \prod_{j=1}^{s} (\lambda - \mu_j)^{e_j - 1}$$

is a divisor of $f_{(i)}(\lambda)$. Hence the numbers μ_j, with multiplicity $e_j - 1$, $1 \leqq j \leqq s$, are always roots of $f_{(i)}(\lambda)$. We call these roots the *trivial* roots of $f_{(i)}(\lambda)$. The remaining roots, denoted by $\xi_{i1} \leqq \cdots \leqq \xi_{i,s-1}$, are called the *nontrivial* roots. Cancelling the common factor (11) from each side of (9), we obtain

$$(12) \qquad \hat{f}_{(i)}(\lambda) = \sum_{j=1}^{s} \theta_{ij} \hat{f}(\lambda)/(\lambda - \mu_j),$$

where

$$\hat{f}(\lambda) = \det C \prod_{j=1}^{s} (\lambda - \mu_j),$$

$$(13) \qquad \hat{f}_{(i)}(\lambda) = \det C \left(\sum_{j=1}^{n} |t_{ij}|^2 \right) \prod_{j=1}^{s-1} (\lambda - \xi_{ij}),$$

$$\theta_{ij} = \sum_{r;\lambda_r = \mu_j} |t_{ir}|^2.$$

The sum in the last formula of (13) extends over all the e_j values of r for which $\lambda_r = \mu_j$.

We are now ready to establish the *first interlacing principle.* Another proof of this interlacing property, based on extremal arguments, may be found in the book of Gantmaher and Kreĭn [**11**, Theorem 19, p. 86].

THEOREM 1. *Let C be positive definite, and let $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_n$ be the roots of* (3). *For fixed i, let $\eta_{i1} \leqq \cdots \leqq \eta_{i,n-1}$ be the roots of* (4). *Then $\eta_{i1}, \cdots, \eta_{i,n-1}$ interlace $\lambda_1, \cdots, \lambda_n$; that is,*

$$(14) \qquad \lambda_1 \leqq \eta_{i1} \leqq \lambda_2 \leqq \eta_{i2} \leqq \cdots \leqq \lambda_{n-1} \leqq \eta_{i,n-1} \leqq \lambda_n.$$

PROOF. To prove (14), it suffices to prove that $\xi_{i1}, \cdots, \xi_{i,s-1}$ interlace $\mu_1, \cdots, \mu_s$. If none of $\theta_{i1}, \cdots, \theta_{is}$ is zero, we argue as follows. Setting λ in (12) equal in turn to $\mu_s, \mu_{s-1}, \mu_{s-2}, \cdots, \mu_1$, we see that $\hat{f}_{(i)}(\lambda)$ is alternately positive and negative. Thus $\hat{f}_{(i)}(\lambda)$ has at least one root in each of the intervals $(\mu_1, \mu_2), (\mu_2, \mu_3), \cdots, (\mu_{s-1}, \mu_s)$. Since $\hat{f}_{(i)}(\lambda)$ has degree $s - 1$, we have accounted for all roots of $\hat{f}_{(i)}(\lambda)$, and hence $\hat{f}_{(i)}(\lambda)$ has exactly one root in each of these intervals. Thus the interlacing property is established if none of $\theta_{i1}, \cdots, \theta_{is}$ is zero. By continuity, the interlacing property remains valid when some of $\theta_{i1}, \cdots, \theta_{is}$ are zero. (If one wishes, the continuity argument may be avoided by using a more detailed discussion in which one writes out what happens when some of $\theta_{i1}, \cdots, \theta_{is}$ are zero.)

Assume now that C and $\lambda_1, \cdots, \lambda_n$ are fixed, but that A becomes variable, and in fact assume that A varies over all Hermitian matrices subject to the constraint that the roots of (3) are to be $\lambda_1, \cdots, \lambda_n$.

THEOREM 2. *Let positive definite C and real numbers $\lambda_1 \leqq \cdots \leqq \lambda_n$ be given. Suppose also that $\lambda_1, \cdots, \lambda_n$ are not all equal. Let k be a fixed integer with $1 \leqq k \leqq n$. Then the following properties* I *and* II *are equivalent.*

I. *For each choice of $k(n-1)$ numbers*

$$\eta_{i1} \leqq \cdots \leqq \eta_{i,n-1}, \qquad 1 \leqq i \leqq k, \tag{15}$$

possessing the interlacing property (14), *there exists a Hermitian matrix A such that the roots of* (3) *are $\lambda_1, \cdots, \lambda_n$ and the roots* $\det(\lambda C(i \mid i) - A(i \mid i)) = 0$ *are $\eta_{i1}, \cdots, \eta_{i,n-1}$, for all $i = 1, 2, \cdots, k$.*

II. *Each distinct number among $\lambda_1, \cdots, \lambda_n$ has multiplicity at least k, and the $k \times k$ submatrix standing in the upper left corner of C^{-1} is diagonal.*

PROOF. We first write down a formula needed in this and subsequent proofs. In (12), set $\lambda = \mu_j$. We obtain

$$\frac{\sum_{r;\lambda_r=\mu_j} |t_{ir}|^2}{\sum_{r=1}^n |t_{ir}|^2} = \left(\frac{\mu_j - \xi_{i1}}{\mu_j - \mu_1}\right) \cdots \left(\frac{\mu_j - \xi_{i,j-1}}{\mu_j - \mu_{j-1}}\right)\left(\frac{\xi_{ij} - \mu_j}{\mu_{j+1} - \mu_j}\right) \cdots \left(\frac{\xi_{i,s-1} - \mu_j}{\mu_s - \mu_j}\right). \tag{16}$$

Now suppose that I is valid. Let p be fixed, with $1 \leqq p \leqq s$. By I, we may find A such that

$$\xi_{i1} = \mu_1, \cdots, \xi_{i,p-1} = \mu_{p-1}, \qquad \xi_{ip} = \mu_{p+1}, \cdots, \xi_{i,s-1} = \mu_s, \tag{17}$$

for $i = 1, 2, \cdots, k$. Using (17) and taking $j = 1, \cdots, p-1, p+1, \cdots, s$, in (16), we see that $t_{ir} = 0$ if $\lambda_r \neq \mu_p$. This means that the first k rows of T^{*-1} have all nonzero entries confined to the e_p columns whose index r satisfies $\lambda_r = \mu_p$. These first k rows of T^{*-1} act therefore as row e_p-tuples, and hence will be dependent if $k > e_p$. Since T^{*-1} is nonsingular, this is a contradiction. Hence $e_p \geqq k$ for each $p = 1, 2, \cdots, s$. This proves the first part of II.

Continuing to suppose the validity of I, let α and β be fixed integers with $1 \leqq \alpha < \beta \leqq k$. By I, A exists such that

$$\begin{aligned} &\xi_{\alpha 1} = \mu_1, \qquad \xi_{\alpha 2} = \mu_2, \cdots, \xi_{\alpha, s-1} = \mu_{s-1}, \\ &\xi_{\beta 1} = \mu_2, \qquad \xi_{\beta 2} = \mu_3, \cdots, \xi_{\beta, s-1} = \mu_s. \end{aligned} \tag{18}$$

Using (16) exactly as before, we find that the nonzero entries in row α of T^{*-1} are confined to the last e_s positions, and that the nonzero entries in row β are confined to the first e_1 positions. Because $s > 1$, we see that the rows α and β of T^{*-1} are orthogonal n-tuples. Hence the entry in position (α, β) of $T^{*-1}T^{-1}$ is zero. But $T^{*-1}T^{-1} = X^{*-1}VV^{-1}X^{-1} = (XX^*)^{-1} = C^{-1}$. Hence, for each pair α, β with $1 \leqq \alpha < \beta \leqq k$, the (α, β) entry of C^{-1} is zero. This means: the upper left k-square block in C^{-1} is diagonal.

This completes the proof that I implies II.

Now assume that II is satisfied. Let arbitrary numbers $\xi_{i1}, \cdots, \xi_{i,s-1}$ interlacing $\mu_1, \cdots, \mu_s$ be given, for $i = 1, 2, \cdots, k$. We wish to construct A such that the roots of $\hat{f}_{(i)}(\lambda)$ are $\xi_{i1}, \cdots, \xi_{i,s-1}$, for $i = 1, 2, \cdots, k$. For this construction, let $X_i = (x_{i1}, \cdots, x_{in})$ be row i of X^{*-1}, and let

$$g_i(\lambda) = \det C \cdot (\lambda - \xi_{i1}) \cdots (\lambda - \xi_{i,s-1}). \tag{19}$$

We shall use the following fact, proved in [10]:

$$\sum_{j=1}^{s} \frac{g_i(\mu_j)}{\hat{f}'(\mu_j)} = 1. \tag{20}$$

Define numbers θ_{ij}, for $1 \leqq i \leqq k$, $1 \leqq j \leqq s$, by

$$\begin{aligned} \theta_{ij} &= \|X_i\|^2 \frac{\mu_j - \xi_{i1}}{\mu_j - \mu_1} \cdots \frac{\mu_j - \xi_{i,j-1}}{\mu_j - \mu_{j-1}} \frac{\xi_{ij} - \mu_j}{\mu_{j+1} - \mu_j} \cdots \frac{\xi_{i,s-1} - \mu_j}{\mu_s - \mu_j} \\ &= \|X_i\|^2 g_i(\mu_j)/\hat{f}'(\mu_j). \end{aligned} \tag{21}$$

Owing to (20),

$$\sum_{j=1}^{s} \theta_{ij} = \|X_i\|^2. \tag{22}$$

By interlacing, $\theta_{ij} \geqq 0$. For each fixed j, we may (because $e_j \geqq k$)

find k pairwise orthogonal row e_j-tuples $\tau_{1j}, \cdots, \tau_{kj}$ such that $\|\tau_{ij}\|^2 = \theta_{ij}$, $1 \leqq i \leqq k$; $1 \leqq j \leqq s$. We now write down the first k rows of an $n \times n$ matrix:

$$(23) \qquad \begin{bmatrix} \tau_{11} & \tau_{12} & \cdots & \tau_{1s} \\ \tau_{21} & \tau_{22} & \cdots & \tau_{2s} \\ \cdot & \cdot & \cdots & \cdot \\ \tau_{k1} & \tau_{k2} & \cdots & \tau_{ks} \end{bmatrix}.$$

Because of the construction of the τ_{ij}, these k rows are pairwise orthogonal, and because of (22), row i has the same norm as row i of X^{*-1}, for $i = 1, 2, \cdots, k$.

Next, notice that because the upper left k-square block in $C^{-1} = X^{*-1}X^{-1}$ is diagonal, the first k rows $X_1, \cdots, X_k$ of X^{*-1} are pairwise orthogonal.

Since any set of k pairwise orthogonal row vectors may be mapped to any other set of k pairwise orthogonal row vectors by a unitary V, provided the norms of paired vectors are equal, we see from the results of the last two paragraphs that a unitary V exists such that the first k rows of $X^{*-1}V$ are the rows of (23). Use this V to define an $n \times n$ matrix T by setting $T = XV$. Hence we have found a unitary V such that $X^{*-1}V = T^{*-1}$, and such that the first k rows of T^{*-1} are the rows of (23). Using this V, define A by

$$A = XV \operatorname{diag}(\lambda_1, \cdots, \lambda_n)V^*X^*.$$

It is now easy to see that $V^*X^{-1}AX^{*-1}V = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$. By the calculations of the first parts of this paper, $\hat{f}_{(i)}(\lambda)$ is given by (12) and the last equation of (13). (Owing to our choice of T and V, the quantity θ_{ij} defined by the last formula of (13) equals $\|\tau_{ij}\|^2$, and hence coincides with the θ_{ij} defined in (21).) All that remains to be done is to prove that

$$(24) \qquad \hat{f}_{(i)}(\lambda) = \|X_i\|^2 g_i(\lambda).$$

This is so since (24) shows that $\xi_{i1}, \cdots, \xi_{i,s-1}$ are the roots of $\hat{f}_{(i)}(\lambda)$. To prove (24), observe that (using (12) and (21)),

$$\begin{aligned} \hat{f}_{(i)}(\mu_t) &= \theta_{it}\hat{f}(\lambda)/(\lambda - \mu_t)\,|_{\lambda=\mu_t} \\ &= \theta_{it}\hat{f}'(\mu_t) = \|X_i\|^2 g_i(\mu_t), \end{aligned}$$

for $t = 1, 2, \cdots, s$. Thus $\hat{f}_{(i)}(\lambda)$ and $\|X_i\|^2 g_i(\lambda)$ are two polynomials of degree $s - 1$ which are equal for s distinct values of λ. Hence $\hat{f}_{(i)}(\lambda) = \|X_i\|^2 g_i(\lambda)$. This completes the proof that II implies I.

There is an easily derived inequality which shows how a nonzero offdiagonal element $c_{\alpha\beta}$ in C^{-1} restricts the behavior of the $\xi_{\alpha j}$ and $\xi_{\beta j}$.

COROLLARY 1. *Let* $C^{-1} = (c_{ij})$. *Then*

$$\frac{|c_{\alpha\beta}|}{(c_{\alpha\alpha}c_{\beta\beta})^{1/2}} \leqq \left(\frac{(\xi_{\alpha 1} - \mu_1)(\xi_{\beta 1} - \mu_1)}{\mu_2 - \mu_1}\right)^{1/2}$$

$$+ \sum_{j=2}^{s-1} \left(\frac{(\mu_j - \xi_{\alpha,j-1})(\mu_j - \xi_{\beta,j-1})(\xi_{\alpha j} - \mu_j)(\xi_{\beta j} - \mu_j)}{(\mu_j - \mu_{j-1})(\mu_{j+1} - \mu_j)}\right)^{1/2} \tag{25}$$

$$+ \left(\frac{(\mu_s - \xi_{\alpha,s-1})(\mu_s - \xi_{\beta,s-1})}{\mu_s - \mu_{s-1}}\right)^{1/2}.$$

PROOF. By interlacing, each of the fractions

$$(\mu_j - \xi_{ip})/(\mu_j - \mu_p), \qquad p = 1, \cdots, j-2, \tag{26}$$

and

$$(\xi_{ip} - \mu_j)/(\mu_{p+1} - \mu_j), \qquad p = j+1, \cdots, s-1, \tag{27}$$

lies between 0 and 1. Also $\sum_{r=1}^{n} |t_{ir}|^2 = c_{ii}$. Hence we see from (16) that

$$\sum_{r;\lambda_r = \mu_j} |t_{ir}|^2 \leqq c_{ii} \frac{\mu_j - \xi_{i,j-1}}{\mu_j - \mu_{j-1}} \frac{\xi_{ij} - \mu_j}{\mu_{j+1} - \mu_j}. \tag{27$'$}$$

(The first fraction on the right is absent if $j = 1$ and the last fraction is absent if $j = s$.) But

$$|c_{\alpha\beta}| = \left|\sum_{j=1}^{n} t_{\alpha j}\bar{t}_{\beta j}\right| \leqq \sum_{j=1}^{s} \left|\sum_{r;\lambda_r = \mu_j} t_{\alpha r}\bar{t}_{\beta r}\right|$$

$$\leqq \sum_{j=1}^{s} \left(\sum_{r;\lambda_r = \mu_j} |t_{\alpha r}|^2\right)^{1/2} \left(\sum_{r;\lambda_r = \mu_j} |t_{\beta r}|^2\right)^{1/2}.$$

Combining this inequality with (27′), we get (25).

If we set $\xi_{\alpha 1} = \mu_1$, $\xi_{\beta 1} = \mu_2$, $\cdots$, $\xi_{\beta,s-1} = \mu_s$, then (25) clearly shows that $c_{\alpha\beta} = 0$.

Although it is basically trivial, for completeness we include the discussion of the case $\lambda_1 = \cdots = \lambda_n$.

THEOREM 3. *The roots of* (3) *satisfy* $\lambda_1 = \cdots = \lambda_n = \alpha$ *if and only if* $A = \alpha C$. *When this is so,* $f_{(i)}(\lambda) = c_{ii} \det C(\lambda - \alpha)^{n-1}$, $1 \leqq i \leqq n$, *where* c_{ii} *is the ith diagonal element of* C^{-1}.

PROOF. It is easy to see that $V^*X^{-1}AX^{*-1}V = \alpha I$ if and only if $A = \alpha XX^* = \alpha C$. When this is so, the equation (9) and $\sum_{j=1}^n |t_{ij}|^2 = c_{ii}$ yield the formula for $f_{(i)}(\lambda)$.

We next turn our attention to new generalizations of the interlacing principle (to be called the *second interlacing principle*) that was derived in [4] and [7]. Let j be fixed, with $1 \leqq j < s$. Define polynomials $G_j(\lambda)$ and $H_j(\lambda)$ by

$$G_j(\lambda) = \sum_{t=1}^{j} (\gamma_n^{-1} + \cdots + \gamma_{n+1-e_t}^{-1}) \frac{\hat{f}(\lambda)}{\lambda - \mu_t} + \sum_{t=j+1}^{s} (\gamma_1^{-1} + \cdots + \gamma_{e_t}^{-1}) \frac{\hat{f}(\lambda)}{\lambda - \mu_t},$$

$$H_j(\lambda) = \sum_{t=1}^{j} (\gamma_1^{-1} + \cdots + \gamma_{e_t}^{-1}) \frac{\hat{f}(\lambda)}{\lambda - \mu_t} + \sum_{t=j+1}^{s} (\gamma_n^{-1} + \cdots + \gamma_{n+1-e_t}^{-1}) \frac{\hat{f}(\lambda)}{\lambda - \mu_t}.$$

It is easy to see that $G_j(\lambda)$ (respectively $H_j(\lambda)$) is alternately positive and negative when evaluated at $\mu_s, \cdots, \mu_1$. Hence $G_j(\lambda)$ (resp. $H_j(\lambda)$) has exactly one root in each interval $(\mu_1, \mu_2), (\mu_2, \mu_3), \cdots, (\mu_{s-1}, \mu_s)$. Let α be the unique root of $G_j(\lambda)$ in $[\mu_j, \mu_{j+1}]$ and let β be the unique root of $H_j(\lambda)$ in $[\mu_j, \mu_{j+1}]$.

LEMMA. $\alpha \leqq \beta$.

PROOF. Notice that if $\lambda \in [\mu_j, \mu_{j+1}]$, then $(-1)^{s-j}\hat{f}(\lambda)/(\lambda - \mu_t) \geqq 0$ if $t \leqq j$, and $(-1)^{s-j}\hat{f}(\lambda)/(\lambda - \mu_t) \leqq 0$ if $t > j$. Since

$$(-1)^{s-j}(H_j(\lambda) - G_j(\lambda))$$

has the form $\sum_{t=1}^s d_t(-1)^{s-j}\hat{f}(\lambda)/(\lambda - \mu_t)$, with $d_t \geqq 0$ if $t \leqq j$ and $d_t \leqq 0$ if $t > j$, it follows that $(-1)^{s-j}H_j(\lambda) \geqq (-1)^{s-j}G_j(\lambda)$ for all $\lambda \in [\mu_j, \mu_{j+1}]$. Furthermore, $(-1)^{s-j}H_j(\mu_j)$ and $(-1)^{s-j}G_j(\mu_j)$ are both positive, whereas $(-1)^{s-j}H_j(\mu_{j+1})$ and $(-1)^{s-j}G_j(\mu_{j+1})$ are both negative. Thus the graphs of $(-1)^{s-j}H_j(\lambda)$ and $(-1)^{s-j}G_j(\lambda)$ both descend from positive to negative values as λ increases from μ_j to μ_{j+1}, with the graph of $(-1)^{s-j}H_j(\lambda)$ everywhere above the graph of $(-1)^{s-j}G_j(\lambda)$. This plainly implies that $\alpha \leqq \beta$.

THEOREM 4 (SECOND INTERLACING PRINCIPLE). *Let j be fixed, $1 \leqq j < s$. With α and β as above, we have*:

(i) *either $\xi_{1j} = \xi_{2j} = \cdots = \xi_{nj} = \alpha$ or for at least one i, we have $\xi_{ij} > \alpha$;*

(ii) *either* $\xi_{1j} = \xi_{2j} = \cdots = \xi_{nj} = \beta$, *or for at least one* i, *we have* $\xi_{ij} < \beta$.

PROOF. Let us deny (i). Suppose $\xi_{ij} \leqq \alpha$ for all i, with strict inequality at least once. Since

$$\hat{f}_{(i)}(\alpha) = c_{ii} \det C \left(\prod_{t=1}^{j-1} (\alpha - \xi_{it}) \right) (\alpha - \xi_{ij}) \left(\prod_{t=j+1}^{s-1} (\alpha - \xi_{it}) \right),$$

we see that sign $\hat{f}_{(i)}(\alpha)$ is $(-1)^{s-1-j}$, with $\hat{f}_{(i)}(\alpha)$ nonzero for at least one i. Then

$$0 < (-1)^{s-1-j} \sum_{i=1}^{n} \hat{f}_{(i)}(\alpha) = (-1)^{s-1-j} \sum_{i=1}^{n} \left(\left(\sum_{p=1}^{s} \theta_{ip} \right) f(\alpha)/(\alpha - \mu_p) \right)$$

$$= \sum_{p=1}^{s} \left(\left(\sum_{i=1}^{n} \theta_{ip} \right) (-1)^{s-1-j} \hat{f}(\alpha)/(\alpha - \mu_p) \right).$$

But

$$\sum_{i=1}^{n} \theta_{ip} = \sum_{i=1}^{n} \sum_{r;\lambda_r = \mu_p} |t_{ir}|^2 = \sum_{r;\lambda_r = \mu_p} \sum_{i=1}^{n} |t_{ir}|^2$$

is the sum of e_p diagonal elements of $V^*X^{-1}X^{*-1}V = V^*(X^*X)^{-1}V$. Now X^*X has the same eigenvalues as $XX^* = C$. Hence the eigenvalues of $V^*X^{-1}X^{*-1}V$ are $\gamma_1^{-1}, \cdots, \gamma_n^{-1}$. Since (by a result of Ky Fan [2]) the sum of e_p diagonal elements of a Hermitian matrix lies between the sum of the e_p smallest and the sum of the e_p largest eigenvalues, it follows that

$$\gamma_n^{-1} + \cdots + \gamma_{n+1-e_p}^{-1} \leqq \sum_{i=1}^{n} \theta_{ip} \leqq \gamma_1^{-1} + \cdots + \gamma_{e_p}^{-1}.$$

Since $(-1)^{s-1-j}\hat{f}(\alpha)/(\alpha - \mu_p)$ is negative if $p \leqq j$, and positive if $p > j$, we see that

$$\sum_{p=1}^{s} \left(\sum_{i=1}^{n} \theta_{ip} \right) (-1)^{s-1-j} \hat{f}(\alpha)/(\alpha - \mu_p)$$

$$\leqq \sum_{p=1}^{j} (\gamma_n^{-1} + \cdots + \gamma_{n+1-e_p}^{-1}) \hat{f}(\alpha)/(\alpha - \mu_p)(-1)^{s-1-j}$$

$$+ \sum_{p=j+1}^{s} (\gamma_1^{-1} + \cdots + \gamma_{e_p}^{-1}) \hat{f}(\alpha)/(\alpha - \mu_p)(-1)^{s-1-j}$$

$$= G_j(\alpha)(-1)^{s-1-j}.$$

Thus $0 \neq G_j(\alpha)$. But by definition, $G_j(\alpha) = 0$. This is the contradiction that proves (i). The proof of (ii) is similar.

It is likely that the first possibility in either case (i) or case (ii) can occur only in rather restrictive circumstances. This question will be examined in a subsequent paper.

When $C = I_n$, it is easy to see that $\alpha = \beta$, and that Theorem 4 reduces to the interlacing principle mentioned in [4, Theorem 1] and [6, Theorem 1], and heavily used in [4] and [5].

Our next task is to obtain the generalizations of the *upper and lower quadratic inequalities* and the *linear inequalities* that were discussed in [4] and [5].

THEOREM 5. *Let $s > 1$. For $1 \leqq j \leqq s$, we have*

$$\sum_{i=1}^{n} \frac{\mu_j - \xi_{i,j-1}}{\mu_j - \mu_{j-1}} \frac{\xi_{ij} - \mu_j}{\mu_{j+1} - \mu_j} \geqq \gamma_1 \left(\sum_{r=1}^{e_j} \gamma_{n+1-r}^{-1} \right), \tag{28}$$

and

$$\sum_{i=1}^{n} \frac{\mu_j - \xi_{i,j-1}}{\mu_j - \mu_1} \frac{\xi_{ij} - \mu_j}{\mu_s - \mu_j} \leqq \gamma_n \left(\sum_{r=1}^{e_j} \gamma_r^{-1} \right). \tag{29}$$

REMARK. Formula (28) is called the *upper quadratic inequality* and (29) is called the *lower quadratic inequality*. When $j = 1$, the factors involving $\mu_j - \xi_{i,j-1}$ are understood to be absent in these formulas, and when $j = s$, the factors involving $\xi_{ij} - \mu_j$ are similarly understood to be absent.

PROOF. Delete the factors (26) and (27) from the right-hand side of (16) (each of which lies between 0 and 1). These deletions increase the right-hand side of (16). The denominator $\sum_{r=1}^{n} |t_{ir}|^2$ in the left-hand side of (16) is the (i, i) element of C^{-1}. Hence the left-hand side of (16) is decreased if we replace this denominator with γ_1^{-1}. Summing the resulting inequality over i, we see (as in the proof of Theorem 4) that the numerator on the left is bounded below by $\gamma_n^{-1} + \cdots + \gamma_{n+1-e_j}^{-1}$. This proves (28).

To prove (29), we rewrite (16) as

$$\frac{\sum_{r;\lambda_r=\mu_j} |t_{ir}|^2}{\sum_{r=1}^{n} |t_{ir}|^2} = \left(\prod_{p=1}^{j-2} \frac{\mu_j - \xi_{ip}}{\mu_j - \mu_{p+1}} \right) \left(\frac{\mu_j - \xi_{i,j-1}}{\mu_j - \mu_1} \right) \cdot \left(\frac{\xi_{ij} - \mu_j}{\mu_s - \mu_j} \right) \left(\prod_{p=j+1}^{s-1} \frac{\xi_{ip} - \mu_j}{\mu_p - \mu_j} \right). \tag{30}$$

By interlacing, each of the fractions in the first and last factors on the right-hand side is at least one. Deleting these products therefore decreases the right-hand side. The denominator of the left-hand side is bounded below by γ_n^{-1}. Making these replacements in (30), then summing on i, the numerator of the left-hand side is bounded above by $\gamma_1^{-1} + \cdots + \gamma_{e_j}^{-1}$. This proves (29).

We increase the left-hand side of (28) if we delete either of the factors inside the first summation. If we delete $(\xi_{ij} - \mu_j)/(\mu_{j+1} - \mu_j)$, we obtain (after some elementary algebra) the inequality (31) below. If we delete instead $(\mu_j - \xi_{i,j-1})/(\mu_j - \mu_{j-1})$, we obtain (32). These inequalities (31) and (32) bound the *arithmetic mean* of the ξ_{ij}.

THEOREM 6. *Let* $s > 1$. *Then*:

$$\frac{1}{n}\sum_{i=1}^{n} \xi_{i,j-1} \leqq \varphi_j\mu_{j-1} + (1 - \varphi_j)\mu_j, \qquad j = 2, 3, \cdots, s, \tag{31}$$

$$\frac{1}{n}\sum_{i=1}^{n} \xi_{ij} \geqq \varphi_j\mu_{j+1} + (1 - \varphi_j)\mu_j, \qquad j = 1, 2, \cdots, s-1. \tag{32}$$

Here

$$\varphi_j = \frac{1}{n}\gamma_1\left(\sum_{r=1}^{e_j} \gamma_{n+1-r}^{-1}\right),$$

and $0 < \varphi_j < 1$.

REMARK. In (31) we have a convex combination of μ_{j-1} and μ_j which serves as an upper bound for the arithmetic mean of the $\xi_{i,j-1}$ (by interlacing, we only know that this arithmetic mean lies in $[\mu_{j-1}, \mu_j]$). In (32) we have a similar convex combination of μ_j and μ_{j+1} which bounds from below the arithmetic mean of the ξ_{ij}.

COROLLARY 2.

$$\frac{1}{n}\sum_{i=1}^{n} \eta_{ij} \leqq \varphi\lambda_j + (1 - \varphi)\lambda_{j+1}, \tag{33}$$

$$\frac{1}{n}\sum_{i=1}^{n} \eta_{ij} \geqq \varphi\lambda_{j+1} + (1 - \varphi)\lambda_j, \tag{34}$$

where $\varphi = (1/n)(\gamma_1/\gamma_n)$.

PROOF. If $\lambda_j = \lambda_{j+1}$ then $\eta_{ij} = \lambda_j$ and (34) is a triviality. If $\lambda_j < \lambda_{j+1}$ let $\lambda_j = \mu_a, \lambda_{j+1} = \mu_{a+1}$. Then $\eta_{ij} = \xi_{ia}$ and using (32) we have

$$\frac{1}{n}\sum_{i=1}^{n} \eta_{ij} \geqq \varphi_a\lambda_{j+1} + (1 - \varphi_a)\lambda_j.$$

However, $\varphi_a\lambda_{j+1} + (1 - \varphi_a)\lambda_j \geqq \varphi\lambda_{j+1} + (1 - \varphi)\lambda_j$, since this is equivalent to $(\varphi_a - \varphi)(\lambda_{j+1} - \lambda_j) \geqq 0$. This proves (34), and (33) is proved similarly.

As our final result, we show how the inequalities proved in Theorem 10 of [3] and discussed further in [8] may be generalized to the situation under discussion in this paper.

Let Q_{nk} denote the set of all sequences $\omega = \{i_1, \cdots, i_k\}$ of strictly increasing positive integers not exceeding n. The number of sequences in Q_{nk} is $\binom{n}{k}$. Let $A[\omega \mid \omega]$ denote the principal submatrix of A lying at the intersecion of rows and columns $i_1, \cdots, i_k$. Let $\zeta_{\omega 1} \leqq \zeta_{\omega 2} \leqq \cdots \leqq \zeta_{\omega k}$ denote the roots of

$$\det(\lambda C[\omega \mid \omega] - A[\omega \mid \omega]) = 0. \tag{35}$$

It follows from Theorem 1 that if $\tau \in Q_{n,k+1}$ and $\tau \supset \omega$, then the roots of

$$\det(\lambda C[\tau \mid \tau] - A[\tau \mid \tau]) = 0 \tag{36}$$

are interlaced by the roots of (35). It therefore follows that

$$\lambda_j \leqq \zeta_{\omega j} \leqq \lambda_{j+n-k}, \qquad j = 1, 2, \cdots, k,$$

for any $\omega \in Q_{nk}$. We now generalize Theorem 10 of [3] by finding convex combinations of $\lambda_j, \lambda_{j+1}, \cdots, \lambda_{j+n-k}$, which serve as upper and lower bounds for the arithmetic mean

$$\binom{n}{k}^{-1} \sum_{\omega \in Q_{nk}} \zeta_{\omega j}$$

of the jth root (fixed j) of all the different equations (35). Let $E_r(a_1, \cdots, a_k)$ denote the elementary symmetric function of degree r on k variables.

THEOREM 7. *For fixed j and k, $1 \leqq k \leqq n - 1$, $1 \leqq j \leqq k$, we have*

$$\sum_{r=0}^{n-k} \psi_r \lambda_{n-k+j-r} \leqq \binom{n}{k}^{-1} \sum_{\omega \in Q_{nk}} \zeta_{\omega j} \leqq \sum_{r=0}^{n-k} \psi_r \lambda_{j+r} \tag{37}$$

where

$$\psi_r = \frac{(\gamma_1/\gamma_n)^{n-k-r} E_r(n - \gamma_1/\gamma_n, n - 1 - \gamma_1/\gamma_n, \cdots, k + 1 - \gamma_1/\gamma_n)}{\prod_{i=k+1}^{n} i}, \quad 0 \leqq r \leqq n - k, \tag{38}$$

$$\sum_{r=0}^{n-k} \psi_r = 1. \tag{39}$$

PROOF. The proof parallels, in part, the proof of Theorem 10 in [3]. The proof is a descending induction on k. By Corollary 2, the result is valid when $k = n - 1$. Suppose the result established for $k + 1$. Then we have

$$\sum_{r=0}^{n-(k+1)} \theta_r \lambda_{n-(k+1)+j-r} \leqq \binom{n}{k+1}^{-1} \sum_{\tau \in Q_{n,k+1}} \zeta_{\tau j} \leqq \sum_{r=0}^{n-(k+1)} \theta_r \lambda_{j+r}, \qquad 1 \leqq j \leqq k+1, \tag{40}$$

with

$$\theta_r = \left(\frac{\gamma_1}{\gamma_n}\right)^{n-(k+1)-r} E_r\left(n - \frac{\gamma_1}{\gamma_n}, \cdots, k+2-\frac{\gamma_1}{\gamma_n}\right)\left\{\prod_{i=k+2}^{n} i\right\}^{-1}, \qquad 0 \leqq r \leqq n-k-1. \tag{41}$$

For a given $\tau \in Q_{n,k+1}$, there exist exactly $k+1$ sequences $\omega \in Q_{nk}$ for which $\omega \subset \tau$. Moreover, if $\gamma_1{}'$ and γ_{k+1}' are the minimum and maximum eigenvalues of $C[\tau \mid \tau]$, we have

$$\frac{k+1-\gamma_1{}'/\gamma_{k+1}'}{k+1}\zeta_{\tau j} + \frac{1}{k+1}\frac{\gamma_1{}'}{\gamma_{k+1}'}\zeta_{\tau,j+1} \leqq \frac{1}{k+1}\sum_{\omega \in Q_{nk};\omega \subset \tau} \zeta_{\omega j} \leqq \frac{1}{k+1}\frac{\gamma_1{}'}{\gamma_{k+1}'}\zeta_{\tau j} + \frac{k+1-\gamma_1{}'/\gamma_{k+1}'}{k+1}\zeta_{\tau,j+1}. \tag{42}$$

Since $\gamma_1{}' \geqq \gamma_1$ and $\gamma_{k+1}' \leqq \gamma_n$, we see that $\gamma_1{}'/\gamma_{k+1}' \geqq \gamma_1/\gamma_n$. Thus (42) yields the weaker inequality,

$$\frac{(k+1-\gamma_1/\gamma_n)}{k+1}\zeta_{\tau j} + \frac{1}{k+1}\frac{\gamma_1}{\gamma_n}\zeta_{\tau,j+1} \leqq \frac{1}{k+1}\sum_{\omega \in Q_{nk};\omega \subset \tau} \zeta_{\omega j} \leqq \frac{1}{k+1}\frac{\gamma_1}{\gamma_n}\zeta_{\tau j} + \frac{k+1-\gamma_1/\gamma_n}{k+1}\zeta_{\tau,j+1}. \tag{43}$$

As in [3], we sum (43) over all sequences $\tau \in Q_{n,k+1}$ and then divide by $\binom{n}{k+1}$. Call the resulting inequality $*$. The central member of $*$ now becomes, as in [3],

$$\binom{n}{k}^{-1} \sum_{\omega \in Q_{nk}} \zeta_{\omega j}.$$

We now use (40) for j and $j+1$ in the left and right sums in our inequality $*$. We then obtain (37) on recognizing that

$$\psi_0 = \left(\frac{\gamma_1}{\gamma_n}\right)\frac{\theta_0}{k+1}, \qquad \psi_r = \frac{(\gamma_1/\gamma_n)\,\theta_r + (k+1-\gamma_1/\gamma_n)\,\theta_{r-1}}{k+1},$$

$$1 \leqq r < n-k,$$

$$\psi_{n-k} = \frac{k+1-\gamma_1/\gamma_n}{k+1}\,\theta_{n-k-1}.$$

Therefore (37) is established. To show that (39) holds, set $\lambda = 1$ in the polynomial identity

$$\left(\frac{\gamma_1}{\gamma_n}\lambda + n - \frac{\gamma_1}{\gamma_n}\right)\left(\frac{\gamma_1}{\gamma_n}\lambda + n - 1 - \frac{\gamma_1}{\gamma_n}\right) \cdots \left(\frac{\gamma_1}{\gamma_n}\lambda + k + 1 - \frac{\gamma_1}{\gamma_n}\right)$$

$$= \sum_{r=0}^{n-k}\left(\frac{\gamma_1}{\gamma_n}\lambda\right)^{n-k-r} E_r\left(n - \frac{\gamma_1}{\gamma_n}, \cdots, k+1-\frac{\gamma_1}{\gamma_n}\right).$$

The proof is complete.

References

1. Ky Fan and Gordon Pall, *Imbedding theorems for Hermitian and normal matrices*, Canad. J. Math. **9** (1957), 298–304. MR **19**, 6.

2. Ky Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations*. I, Proc. Nat. Acad. Sci. U.S.A. **35** (1949), 652–655. MR **11**, 600.

3. R. C. Thompson, *Principal submatrices of normal and Hermitian matrices*, Illinois J. Math. **10** (1966), 296–308. MR **32** #7565.

4. R. C. Thompson and P. McEnteggert, *Principal submatrices*. II. *The upper and lower quadratic inequalities*, Linear Algebra and Appl. **1** (1968), 211–243. MR **38** #5813.

5. R. C. Thompson, *Principal submatrices*. III. *Linear inequalities*, J. Res. Nat. Bur. Standards Sect. B **72B** (1968), 7–22. MR **38** #5814.

6. ———, *Principal submatrices*. IV. *On the independence of the eigenvalues of different principal submatrices*, Linear Algebra and Appl. **2** (1969), 355–374. MR **39** #6907.

7. ———, *Principal submatrices*. V. *Some results concerning principal submatrices of arbitrary matrices*, J. Res. Nat. Bur. Standards Sect. B **72B** (1968), 115–125. MR **39** #5597.

8. ———, *Principal submatrices*. VI. *Cases of equality in certain linear inequalities*, Linear Algebra and Appl. **2** (1969), 375–379. MR **39** #6908.

9. ———, *Principal submatrices*. VII. *Further results concerning matrices with equal principal minors*, J. Res. Nat. Bur. Standards Sect. B **72B** (1968), 249–252. MR **39** #5598.

10. *Problem E-1623, proposed by R. C. Thompson*, Amer. Math. Monthly **71** (1964), 682–683.

11. F. R. Gantmaher and M. G. Kreĭn, *Oscillation matrices and kernels and small vibrations of mechanical systems*, GITTL, Moscow, 1950; English transl., U.S. Atomic Energy Commission AEC Tr 4481, 1961; available from the Clearinghouse for Federal Scientific and Technical Information, Springfield, Va., 22151. MR **14**, 178.

University of California, Santa Barbara, California 93106