A NONZERO DETERMINANT RELATED TO SCHUR'S MATRIX

BY
STEPHEN SALAFF

For arbitrary complex numbers α_j, β_j, subject only to the restriction that $|\alpha_j|+|\beta_j|>0$, $j=0, 1,\dots, q-1$, we prove

THEOREM. *The* $2q\times 2q$ *determinant*

$$D=\left|\begin{array}{cccc|ccc}
\beta_0\varepsilon_0^0 & \beta_0\varepsilon_1^0 & \cdots & \beta_0\varepsilon_{q-1}^0 & \alpha_0\varepsilon_q^0 & \cdots & \alpha_0\varepsilon_{2q-1}^0 \\
\beta_1\varepsilon_0^1 & \beta_1\varepsilon_1^1 & \cdots & \beta_1\varepsilon_{q-1}^1 & \alpha_1\varepsilon_q^1 & \cdots & \alpha_1\varepsilon_{2q-1}^1 \\
\vdots & \vdots & & \vdots & \vdots & & \vdots \\
\beta_{q-1}\varepsilon_0^{q-1} & \beta_{q-1}\varepsilon_1^{q-1} & \cdots & \beta_{q-1}\varepsilon_{q-1}^{q-1} & \alpha_{q-1}\varepsilon_q^{q-1} & \cdots & \alpha_{q-1}\varepsilon_{2q-1}^{q-1} \\
\hline
\bar{\alpha}_{q-1}\varepsilon_0^q & \bar{\alpha}_{q-1}\varepsilon_1^q & \cdots & \bar{\alpha}_{q-1}\varepsilon_{q-1}^q & \bar{\beta}_{q-1}\varepsilon_q^q & \cdots & \bar{\beta}_{q-1}\varepsilon_{2q-1}^q \\
\vdots & \vdots & & \vdots & \vdots & & \vdots \\
\bar{\alpha}_1\varepsilon_0^{2q-2} & \bar{\alpha}_1\varepsilon_1^{2q-2} & \cdots & \bar{\alpha}_1\varepsilon_{q-1}^{2q-2} & \bar{\beta}_1\varepsilon_q^{2q-2} & \cdots & \bar{\beta}_1\varepsilon_{2q-1}^{2q-2} \\
\bar{\alpha}_0\varepsilon_0^{2q-1} & \bar{\alpha}_0\varepsilon_1^{2q-1} & \cdots & \bar{\alpha}_0\varepsilon_{q-1}^{2q-1} & \bar{\beta}_0\varepsilon_q^{2q-1} & \cdots & \bar{\beta}_0\varepsilon_{2q-1}^{2q-1}
\end{array}\right|$$

where $\varepsilon_k=e^{\pi ik/q}$, *does not vanish.*

The matrix $V=(\varepsilon_k^j)_{j,k=0}^{2q-1}$ is the even order Schur matrix. Schur's matrix of order n is the Vandermonde $V_n=(e^{2\pi ijk/n})_{j,k=0}^{n-1}$. It bears his name in honor of I. Schur's work in the theory of group representations and in number theory. Vandermondes like V_n occur in the computation of the characters of the irreducible integral representations of the full linear group [3, p. 185]. Schur's derivation of the formula for the Gauss sum $\sum_{j=0}^{n-1} e^{2\pi ij^2/n}$ exploits the fact that the sum is the trace of V_n [8].

The nonvanishing of D is of importance in boundary value problems arising in the Birkhoff theory of nonselfadjoint nth order linear ordinary differential operators. Regular boundary conditions for such operators are defined by the nonvanishing of two determinants of order n, which involve the leading coefficients in the boundary forms, and the ε_k. Birkhoff [2] obtained asymptotic estimates for the eigenvalues and eigenfunctions of differential operators determined by regular boundary conditions, and proved an expansion theorem analogous to the Fourier expansion. The question of whether selfadjoint boundary conditions are regular appears open. In a forthcoming paper [7] we make use of an extension of the present theorem, stated in §II.2, to answer this question in the affirmative when $n=2q$.

Received by the editors March 28, 1966.

In the preliminary section I we first review positive definiteness for matrices. We then give a lemma which states that each minor of a unitary matrix is proportional to the complex conjugate of its cofactor. There follows the construction of a class of positive definite matrices, involving a proof that the matrix

$$(\operatorname{cosec} \pi(j+k+1)/2q)_{j,k=0}^{q-1}$$

is totally definite. Facts about the Schur matrix are then summarized. The actual proof of the theorem is in §II.

In this paper $C_k(A)$ will denote the kth compound of the matrix A. For its definition and a number of consequences we refer to [1, p. 90], [4, pp. 16–17]. We state here, however, an equivalent defining property

$$C_k(A)x_1 \wedge \cdots \wedge x_k = Ax_1 \wedge \cdots \wedge Ax_k$$

from which it follows that

$$C_k(AB) = C_k(A)C_k(B). \tag{1}$$

The determinant of A is $|A|$ or $\det A$. The n dimensional identity is written I_n, A^* is the transposed conjugate of A and $A \otimes B$ is the tensor product of A and B. All matrices will be square.

I wish to thank H. O. Cordes for showing me the close relation between the Lemma, §I.3, and the theorem. I am also grateful to Donald Sarason and David Spring for many useful discussions.

I. Preliminaries.

1. *A brief review of positive definiteness.* The hermitian matrix $A=(a_{jk})_{j,k=1}^n$ is positive definite (p.d.) if $(Ax, x)=\sum_{j,k=1}^n a_{jk}x_j\bar{x}_k>0$ for all nonzero vectors $x=(x_1, \ldots, x_n)$. It is a consequence that A is p.d. if and only if all its principal minors (the ones chosen from the same rows as columns) are positive, or if and only if $A=B^2$ for some nonsingular hermitian B. If A and B are hermitian and $a_{jk}=d_jd_k'b_{jk}$, where d_j, d_k' are positive, $j, k=1, \ldots, n$, then A is p.d. if and only if B is.

If A is p.d. then so are $A \otimes A$ and $C_k(A)$, $k=1, \ldots, n$. In fact $B^2 \otimes B^2 =(B \otimes B)^2$ and $C_k(B^2)=(C_k(B))^2$. The matrix of squares $(a_{jk}^2)_{j,k=1}^n$, being a principal submatrix of $A \otimes A$, is p.d. also.

2. *A lemma on unitary matrices.*

Lemma. *Let A be an $n \times n$ unitary matrix. Let M be any minor of A and let m be the cofactor of M in A. Then $M=\bar{m}\det A$.*

This is a well-known result. It was proved by Muir [5].

Corollary. *If $AA^*=\lambda^2 I_n$, λ real, and m is an $r \times r$ minor of A with cofactor M, then $M=\bar{m}\lambda^{-2r}\det A$.*

Just replace A by $\lambda^{-1}A$ in the lemma.

3. *A class of positive definite matrices.* We require a notation for matrices indexed by sets somewhat more general than the natural numbers. Let Q be the ordered set $\{0, 1, \ldots, q-1\}$. For $r=1, \ldots, q$, Q_r is the set of all strictly increasing sequences in Q of length r; ρ and σ will always be elements of Q_r. Order Q_r by setting $\rho > \sigma$ if the first integer in ρ which differs from the corresponding integer in σ is greater than that integer. Let $(a_{\rho\sigma})_{\rho,\sigma\in Q_r}$ be the $\binom{q}{r}\times\binom{q}{r}$ matrix whose elements $a_{\rho\sigma}$ are indexed by Q_r with the given ordering. Thus $(a_{jk})_{j,k=0}^{q-1}$, indexed by Q_1, is the usual notation for a $q\times q$ matrix. Its $r\times r$ submatrix with row and column indices in ρ and σ respectively, is written $(a_{jk})_{j\in\rho,\, k\in\sigma}$.

LEMMA. *The matrices P_r are positive definite for $r=1, 2, \ldots, q$, where*

$$P_r = \Big(\prod_{j\in\rho;k\in\sigma} \operatorname{cosec}(x_j+x_k)\Big)_{\rho,\sigma\in Q_r}$$

and $0<x_0<x_1<\cdots<x_{q-1}<\pi/2$.

Proof. First observe that each P_r is a symmetric matrix with positive elements. Now fix r, $2\leqq r\leqq q$. We define the positive quantities

$$t_j = \tan x_j, \quad j\in Q, \qquad T_\rho = \prod_{j'>j;j,j'\in\rho} (t_{j'}-t_j),$$

$$C_\rho = \prod_{j\in\rho} \cos x_j, \qquad y_{\rho\sigma} = \prod_{j\in\rho;k\in\sigma} \operatorname{cosec}(x_j+x_k).$$

Let $p_{\rho\sigma}$ be the minor of P_1 with row and column indices in ρ and σ. Then

$$\begin{aligned} p_{\rho\sigma} &= \det(\operatorname{cosec}(x_j+x_k))_{j\in\rho,\,k\in\sigma} \\ &= \det(\sec x_j \sec x_k\,(t_j+t_k)^{-1})_{j\in\rho,\,k\in\sigma} \\ &= C_\rho^{-1}C_\sigma^{-1}\det((t_j+t_k)^{-1})_{j\in\rho,\,k\in\sigma} \\ &= C_\rho^{-1}C_\sigma^{-1}T_\rho T_\sigma \prod_{j\in\rho;k\in\sigma}(t_j+t_k)^{-1} \\ &= C_\rho^{r-1}C_\sigma^{r-1}T_\rho T_\sigma y_{\rho\sigma}. \end{aligned}$$

The second equation is a consequence of the identity

$$\operatorname{cosec}(x+y) = \sec x \sec y\,(\tan x+\tan y)^{-1},$$

and the fourth holds by a lemma of Cauchy [6, Vol. 2, pp. 98, 299]. It follows that $p_{\rho\sigma}$ is positive and that

$$P_r = D_r C_r(P_1) D_r. \tag{2}$$

Here D_r is a diagonal matrix with positive diagonal elements $C_\rho^{1-r}T_\rho^{-1}$, $\rho\in Q_r$.

In conclusion, P_1 is p.d. (it is even totally definite). Thus $C_r(P_1)$ is p.d., and from (2), P_r is p.d., $r=2, \ldots, q$.

We remark that this proof generalizes easily to show that if

(i) g and h are positive functions defined for the sequence $x_0<x_1<\cdots<x_{q-1}$ and h is strictly increasing

(ii) f satisfies the functional equation $f(x+y)=g(x)g(y)[h(x)+h(y)]^{-1}$ then

$$\left(\prod_{j\in\rho;k\in\sigma} f(x_j+x_k)\right)_{\rho,\sigma\in Q_r}$$

is p.d. for $r=1,\ldots,q$.

4. *Some facts about the Schur matrix.* From the definition, $V_nV_n^*=nI_n$ and $|\det V_n|=n^{n/2}$. If m is any $r\times r$ minor of V_n and M is its cofactor in V_n, we infer from Corollary 2, that $M=\bar{m}n^{-r}\det V_n$.

Assuming now that $n=2q$, we call the first q columns of V the left columns and the last q the right ones. Likewise, the first and last q rows will be the top and bottom. For $0\leqq j_0<j_1<\cdots<j_{q-1}\leqq 2q-1$, the right minor

$$\begin{vmatrix} \varepsilon_q^{j_0} & \cdots & \varepsilon_{2q-1}^{j_0} \\ \vdots & & \vdots \\ \varepsilon_q^{j_{q-1}} & \cdots & \varepsilon_{2q-1}^{j_{q-1}} \end{vmatrix} = \varepsilon_q^{j_0+\cdots+j_{q-1}} \begin{vmatrix} \varepsilon_0^{j_0} & \cdots & \varepsilon_{q-1}^{j_0} \\ \vdots & & \vdots \\ \varepsilon_0^{j_{q-1}} & \cdots & \varepsilon_{q-1}^{j_{q-1}} \end{vmatrix}.$$

Hence the absolute value of any right minor equals that of the left minor chosen from the same rows. Further, the determinant on the right side of this equality is the Vandermonde $\prod_{0\leqq l<k\leqq q-1}(\varepsilon_{j_k}-\varepsilon_{j_l})$.

If m is any left minor with cofactor M then $M=c_1\bar{m}$, where $c_1=(2q)^{-q}\det V$, $|c_1|=1$. We arrive at the equation

$$mM = c_1|m|^2$$

which is central to the proof of the theorem, in which we develop D by Laplace expansion from the left columns.

II. **Proof of the Theorem.**

1. Fix the integer r, $1\leqq r\leqq q-1$. The mapping $\sim: Q_r\to Q_{q-r}$ sends ρ onto the element $\tilde{\rho}\in Q_{q-r}$ consiting of those $q-r$ integers in Q disjoint from ρ. We define

$$m_{\rho\sigma} = \begin{vmatrix} \varepsilon_0^{j_0} & \cdots & \varepsilon_{q-1}^{j_0} \\ \vdots & & \vdots \\ \varepsilon_0^{j_{r-1}} & \cdots & \varepsilon_{q-1}^{j_{r-1}} \\ \varepsilon_0^{2q-1-k_{q-r-1}} & \cdots & \varepsilon_{q-1}^{2q-1-k_{q-r-1}} \\ \vdots & & \vdots \\ \varepsilon_0^{2q-1-k_0} & \cdots & \varepsilon_{q-1}^{2q-1-k_0} \end{vmatrix}$$

where $\rho=\{j_0,\ldots,j_{r-1}\}$ and $\tilde{\sigma}=\{k_0,\ldots,k_{q-r-1}\}$. In words, $m_{\rho\sigma}$ is the left minor of V chosen from the rows indicated by ρ in the top half and the ones indicated by $\tilde{\sigma}$, in reverse order, in the bottom half. Let $M_{\rho\sigma}$ be the cofactor of $m_{\rho\sigma}$ in V. We showed in §I.4 that

$$|m_{\rho\sigma}| = |M_{\rho\sigma}| \quad\text{and}\quad m_{\rho\sigma}M_{\rho\sigma} = c_1|m_{\rho\sigma}|^2.$$

From the Vandermonde form for $m_{\rho\sigma}$,

$$(3)\quad m_{\rho\sigma} = \prod_{j'>j;j,j'\in\rho}(\varepsilon_{j'}-\varepsilon_j)\prod_{k'>k;k,k'\in\tilde{\sigma}}(\varepsilon_{2q-1-k}-\varepsilon_{2q-1-k'})\prod_{j\in\rho;k\in\tilde{\sigma}}(\varepsilon_{2q-1-k}-\varepsilon_j).$$

Empty products, which arise if $r=1$ or $q-1$ will be interpreted as equal to one. Put $x_j=(j+\frac{1}{2})\pi/2q$,

$$S_\rho = \prod_{j'>j;j,j'\in\rho} \sin(x_{j'}-x_j), \quad \text{and} \quad S'_\rho = S_\rho \prod_{j\in\rho;k\in Q} \sin(x_j+x_k).$$

We define, as in the Lemma, §I.3,

$$y_{\rho\bar{\sigma}} = \prod_{j\in\rho;k\in\bar{\sigma}} \operatorname{cosec}(x_j+x_k) = y_{\bar{\sigma}\rho}.$$

Since $y_{\rho\sigma}y_{\rho\bar{\sigma}}=\prod_{j\in\rho;k\in Q}\operatorname{cosec}(x_j+x_k)$, we get

$$S_\rho y_{\rho\bar{\sigma}}^{-1} = S'_\rho y_{\rho\sigma}.$$

The identity $\varepsilon_t-\varepsilon_s=2i\sin[(t-s)\pi/2q]e^{i(t+s)\pi/2q}$ when applied to (3) yields

$$|m_{\rho\sigma}| = 2^{q(q-1)/2}S_\rho S_{\bar{\sigma}} y_{\rho\bar{\sigma}}^{-1} = 2^{q(q-1)/2}S'_\rho S_{\bar{\sigma}} y_{\rho\sigma}.$$

The cofactor $M_{\rho\sigma}$ is by definition the signed minor of V from the right, the rows indicated by $\bar{\rho}$ in the top and by σ, in reverse order, in the bottom, i.e., the rows omitted by $m_{\rho\sigma}$. By the equality in absolute value of left and right minors from the same rows,

$$|M_{\rho\sigma}| = 2^{q(q-1)/2}S_{\bar{\rho}}S_\sigma y_{\sigma\bar{\rho}}^{-1} = |m_{\sigma\rho}|.$$

We can now assert that

$$m_{\rho\sigma}M_{\rho\sigma} = m_{\sigma\rho}M_{\sigma\rho} = c(S'_\rho S_{\bar{\sigma}} y_{\rho\sigma})^2$$

where $c=2^{q(q-1)}c_1$.

If $\gamma=(\gamma_0,\ldots,\gamma_{q-1})$ is any vector of q complex components, let $\gamma^\rho=\prod_{j\in\rho}\gamma_j$, $\bar{\gamma}^\rho=\prod_{j\in\rho}\bar{\gamma}_j$. In this notation,

$$\begin{vmatrix} \beta_{j_0}\varepsilon_0^{j_0} & \cdots & \beta_{j_0}\varepsilon_{q-1}^{j_0} \\ \vdots & & \vdots \\ \beta_{j_{r-1}}\varepsilon_0^{j_{r-1}} & \cdots & \beta_{j_{r-1}}\varepsilon_{q-1}^{j_{r-1}} \\ \bar{\alpha}_{k_{q-r-1}}\varepsilon_0^{2q-1-k_{q-r-1}} & \cdots & \bar{\alpha}_{k_{q-r-1}}\varepsilon_{q-1}^{2q-1-k_{q-r-1}} \\ \vdots & & \vdots \\ \bar{\alpha}_{k_0}\varepsilon_0^{2q-1-k_0} & \cdots & \bar{\alpha}_{k_0}\varepsilon_{q-1}^{2q-1-k_0} \end{vmatrix} = \beta^\rho\bar{\alpha}^{\bar{\sigma}}m_{\rho\sigma}.$$

The cofactor of this term in the Laplace expansion of D from the first q columns is $\alpha^{\bar{\rho}}\bar{\beta}^\sigma M_{\rho\sigma}$.

The sum $\sum_{\rho,\sigma\in Q_r}\beta^\rho\alpha^{\bar{\rho}}\bar{\beta}^\sigma\bar{\alpha}^{\bar{\sigma}}m_{\rho\sigma}M_{\rho\sigma}$ contains just those terms in the expansion of D obtained by taking r of the rows in the left minor from the top. It is a quadratic form $(A_r\gamma_r,\gamma_r)$ in the variables $\gamma_r=(\beta^\rho\alpha^{\bar{\rho}})_{\rho\in Q_r}$ with matrix of coefficients

$$A_r = (m_{\rho\sigma}M_{\rho\sigma})_{\rho,\sigma\in Q_r} = c((S'_\rho S_{\bar{\sigma}} y_{\rho\sigma})^2)_{\rho,\sigma\in Q_r}.$$

By the Lemma, §I.3, $(y_{\rho\sigma})_{\rho,\sigma\in Q_r}$ is p.d. Since S'_ρ and $S_{\bar{\sigma}}$ are positive, $c^{-1}A_r$ is also p.d. This holds for $r=1,2,\ldots,q-1$.

Let $\gamma_0=\prod_{j=0}^{q-1}\alpha_j\bar{\alpha}_j$, $\gamma_q=\prod_{j=0}^{q-1}\beta_j\bar{\beta}_j$ be vectors of one component. The matrices $A_0=A_q$ accompanying γ_0, γ_q in the expansion have one element

$$\begin{vmatrix} \varepsilon_0^0 & \cdots & \varepsilon_{q-1}^0 \\ \vdots & & \vdots \\ \varepsilon_0^{q-1} & \cdots & \varepsilon_{q-1}^{q-1} \end{vmatrix} \begin{vmatrix} \varepsilon_q^q & \cdots & \varepsilon_{2q-1}^q \\ \vdots & & \vdots \\ \varepsilon_q^{2q-1} & \cdots & \varepsilon_{2q-1}^{2q-1} \end{vmatrix}$$

$$= (-1)^q \begin{vmatrix} \varepsilon_0^q & \cdots & \varepsilon_{q-1}^q \\ \vdots & & \vdots \\ \varepsilon_0^{2q-1} & \cdots & \varepsilon_{q-1}^{2q-1} \end{vmatrix} \begin{vmatrix} \varepsilon_q^0 & \cdots & \varepsilon_{2q-1}^0 \\ \vdots & & \vdots \\ \varepsilon_q^{q-1} & \cdots & \varepsilon_{2q-1}^{q-1} \end{vmatrix}.$$

This product equals c_1 times the square of the absolute value of any one of the four determinants.

We have shown that

$$D = \sum_{r=0}^{q} (A_r\gamma_r, \gamma_r)$$

and that $c^{-1}A_r$ is p.d., $r=0, 1, \ldots, q$. By hypothesis $|\alpha_j|+|\beta_j|>0$, $j\in Q$. Hence there exists an r, $0\leqq r\leqq q$, and a $\rho\in Q_r$ such that $\beta^\rho\alpha^{\bar{\rho}}\neq 0$. The vector γ_r is not zero, and the theorem is established.

As an illustration, when $q=2$,

$$D = -i[2(|\alpha_0\alpha_1|^2+|\beta_0\beta_1|^2)+|\alpha_0\beta_1-\alpha_1\beta_0|^2+3|\alpha_0\beta_1+\alpha_1\beta_0|^2].$$

2. *An extension.* There are many nonzero subdeterminants of D. For each subset $L=\{g_0, g_1, \ldots, g_{l-1}\}$ of Q we define L_r to be the set of all selections of r strictly increasing integers from L, and assign to L_r the ordering of Q_r. Then $L_r\subset Q_r$ and $(y_{\rho\sigma})_{\rho,\sigma\in L_r}$ is a principal submatrix of P_r. Using this fact it can be shown that $D_L\neq 0$, where D_L is the $2l\times 2l$ subdeterminant consisting of the rows g_j from the top, $2q-1-g_j$ from the bottom in reverse order, the first l columns on the left and the first l on the right. That is,

$$D_L = \left|\begin{array}{ccc|ccc} \beta_{g_0}\varepsilon_0^{g_0} & \cdots & \beta_{g_0}\varepsilon_{l-1}^{g_0} & \alpha_{g_0}\varepsilon_q^{g_0} & \cdots & \alpha_{g_0}\varepsilon_{q+l-1}^{g_0} \\ \vdots & & \vdots & \vdots & & \vdots \\ \beta_{g_{l-1}}\varepsilon_0^{g_{l-1}} & \cdots & \beta_{g_{l-1}}\varepsilon_{l-1}^{g_{l-1}} & \alpha_{g_{l-1}}\varepsilon_q^{g_{l-1}} & \cdots & \alpha_{g_{l-1}}\varepsilon_{q+l-1}^{g_{l-1}} \\ \hline \bar{\alpha}_{g_{l-1}}\varepsilon_0^{2q-1-g_{l-1}} & \cdots & \bar{\alpha}_{g_{l-1}}\varepsilon_{l-1}^{2q-1-g_{l-1}} & \bar{\beta}_{g_{l-1}}\varepsilon_q^{2q-1-g_{l-1}} & \cdots & \bar{\beta}_{g_{l-1}}\varepsilon_{q+l-1}^{2q-1-g_{l-1}} \\ \vdots & & \vdots & \vdots & & \vdots \\ \bar{\alpha}_{g_0}\varepsilon_0^{2q-1-g_0} & \cdots & \bar{\alpha}_{g_0}\varepsilon_{l-1}^{2q-1-g_0} & \bar{\beta}_{g_0}\varepsilon_q^{2q-1-g_0} & \cdots & \bar{\beta}_{g_0}\varepsilon_{q+l-1}^{2q-1-g_0} \end{array}\right|.$$

BIBLIOGRAPHY

1. A. Aitken, *Determinants and matrices*, Oliver and Boyd, London, 1961.

2. G. D. Birkhoff, *Boundary value and expansion problems of ordinary linear differential equations*, Trans. Amer. Math. Soc. **9** (1908), 373–395.

3. H. Boerner, *Representations of groups*, North-Holland, Amsterdam, 1963.

4. M. Marcus and H. Minc, *A survey of matrix theory and matrix inequalities*, Allyn and Bacon, Boston, 1964.

5. T. Muir, *Note on hyperorthogonants*, Trans. Roy. Soc. South Africa **14** (1927), 337–341.

6. G. Polya and G. Szego, *Aufgaben und Lehrsätze aus der Analysis*, Springer, Berlin, 1925.

7. S. Salaff, *Regular boundary conditions for ordinary differential operators*, (to appear).

8. I. Schur, *Über die Gaussschen Summen*, Göttingen Nachrichten, 1921, 147–153.

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA