A NONZERO DETERMINANT RELATED TO SCHUR'S MATRIX

BY

STEPHEN SALAFF

For arbitrary complex numbers \( a_j, \beta_j \), subject only to the restriction that \( |a_j| + |eta_j| > 0, j = 0, 1, \ldots, q-1 \), we prove

**Theorem.** The \( 2q \times 2q \) determinant

\[
D = \begin{vmatrix}
\beta_0 e_0^0 & \beta_0 e_1^0 & \cdots & \beta_0 e_{q-1}^0 & \alpha_0 e_q^0 & \cdots & \alpha_0 e_{2q-1}^0 \\
\beta_1 e_0^1 & \beta_1 e_1^1 & \cdots & \beta_1 e_{q-1}^1 & \alpha_1 e_q^1 & \cdots & \alpha_1 e_{2q-1}^1 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\
\beta_{q-1} e_{q-1}^{q-1} & \beta_{q-1} e_q^{q-1} & \cdots & \beta_{q-1} e_{2q-1}^{q-1} & \alpha_{q-1} e_{q+1}^{-1} & \cdots & \alpha_{q-1} e_{2q-1}^{-1} \\
\alpha_{q-1} e_0^{-1} & \alpha_{q-1} e_1^{-1} & \cdots & \alpha_{q-1} e_{q-1}^{-1} & \beta_{q-1} e_q^{q} & \cdots & \beta_{q-1} e_{2q-1}^{q} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\
\alpha_{2q-1} e_0^{2q-2} & \alpha_{2q-1} e_1^{2q-2} & \cdots & \alpha_{2q-1} e_{q-1}^{2q-2} & \beta_{2q-2} e_q^{2q-2} & \cdots & \beta_{2q-2} e_{2q-1}^{2q-2} \\
\alpha_{2q} e_0^{2q-1} & \alpha_{2q} e_1^{2q-1} & \cdots & \alpha_{2q} e_{q-1}^{2q-1} & \beta_{2q-1} e_q^{2q-1} & \cdots & \beta_{2q-1} e_{2q-1}^{2q-1}
\end{vmatrix}
\]

where \( e_k = e^{i\pi k/2} \), does not vanish.

The matrix \( V = (e_k)^{2q-1}_{j,k=0} \) is the even order Schur matrix. Schur's matrix of order \( n \) is the Vandermonde \( V_n = (e^{2\pi k/2n})^{n-1}_{j,k=0} \). 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 \( \sum_{j=0}^{2n-1} e^{2\pi j^2/n} \) 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 \( n \)th 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 \( e_k \). 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 §11.2, 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
\[(\csc \pi(j + k + 1)/(2q))_{j,k=0}^{n-1}\]
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 \(k\)th 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
\[(1)\quad C_k(AB) = C_k(A)C_k(B)\]
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 \(A\) 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 \((a_{jk})_{j,k=1}^n\), 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 \(AA^* = \lambda^2 I_n, \lambda\) 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 \( r \); \( \rho \) and \( \sigma \) will always be elements of \( Q_r \). Order \( Q_r \) by setting \( \rho > \sigma \) if the first integer in \( \rho \) which differs from the corresponding integer in \( \sigma \) is greater than that integer. Let \((a_{\rho \sigma})_{\rho \in Q_r, \sigma \in Q_r}\) be the \((q) \times (q)\) matrix whose elements \( a_{\rho \sigma} \) are indexed by \( Q_r \) with the given ordering. Thus \((a_{\rho \sigma})_{\rho \in Q_1, \sigma \in Q_1}\), 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 \( \rho \) and \( \sigma \) respectively, is written \((a_{\rho \sigma})_{\rho \in Q_r, \sigma \in Q_r}\).

**Lemma.** The matrices \( P_r \) are positive definite for \( r = 1, 2, \ldots, q \), where

\[
P_r = \left( \prod_{\rho \in Q_r, \sigma \in Q_r} \sec \left( x_j + x_k \right) \right)_{\rho \in Q_r, \sigma \in Q_r}
\]

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 \( r, 2 \leq r \leq q \). We define the positive quantities

\[
t_j = \tan x_j, \quad j \in Q, \quad T_\rho = \prod_{j' \neq j, j' \in \rho} (t_{j'} - t_j),
\]

\[
C_\rho = \prod_{j \in \rho} \cos x_j, \quad y_{\rho \sigma} = \prod_{j \in \rho, k \in \sigma} \sec (x_j + x_k).
\]

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

\[
p_{\rho \sigma} = \det \left( \sec \left( x_j + x_k \right) \right)_{j \in \rho, k \in \sigma}
\]

\[
= \det \left( \sec x_j \sec x_k (t_j + t_k)^{-1} \right)_{j \in \rho, k \in \sigma}
\]

\[
= C_\rho^{-1} C_\sigma^{-1} \prod_{j \in \rho, k \in \sigma} (t_j + t_k)^{-1}
\]

\[
= C_\rho^{-1} C_\sigma^{-1} T_\rho T_\sigma \prod_{j \in \rho, k \in \sigma} (t_j + t_k)^{-1}
\]

\[
= C_\rho^{-1} C_\sigma^{-1} T_\rho T_\sigma y_{\rho \sigma}.
\]

The second equation is a consequence of the identity

\[
\sec (x+y) = \sec x \sec y \left( \tan x + \tan y \right)^{-1},
\]

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

\[
P_r = D_r C_r (P_1) D_r.
\]

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

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

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

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

4. Some facts about the Schur matrix. From the definition, $V_n V_n^* = n I_n$ and $|\det V_n| = n^{q/2}$. 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 $M = m n^{-r} \det V_n$.

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 $0 \leq j_0 < j_1 < \cdots < j_{q-1} \leq 2q - 1$, the right minor

$$
\begin{vmatrix}
  e_{j_0}^0 & \cdots & e_{j_0}^{2q-1} \\
  \vdots & & \vdots \\
  e_{j_{q-1}}^0 & \cdots & e_{j_{q-1}}^{2q-1}
\end{vmatrix} = \begin{vmatrix}
  e_{j_0}^0 & \cdots & e_{j_0}^{q-1} \\
  \vdots & & \vdots \\
  e_{j_{q-1}}^0 & \cdots & e_{j_{q-1}}^{q-1}
\end{vmatrix}
$$

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 $\prod_{0 \leq k < l \leq 2q-1} (e^k - e^l)$.

If $m$ is any left minor with cofactor $M$ then $M = c_1 m$, 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 $r$, $1 \leq r \leq q - 1$. The mapping $\sim : \mathbb{Q}_r \to \mathbb{Q}_{q-r}$ sends $\rho$ onto the element $\bar{\rho} \in \mathbb{Q}_{q-r}$ consisting of those $q-r$ integers in $Q$ disjoint from $\rho$. We define

$$
m_{\rho \sigma} = \begin{vmatrix}
  e_{j_0}^0 & \cdots & e_{j_0}^{q-1} \\
  \vdots & & \vdots \\
  e_{j_{q-r-1}}^0 & \cdots & e_{j_{q-r-1}}^{q-1}
\end{vmatrix}
$$

where $\rho = \{j_0, \ldots, j_{r-1}\}$ and $\sigma = \{k_0, \ldots, k_{q-r-1}\}$. In words, $m_{\rho \sigma}$ is the left minor of $V$ chosen from the rows indicated by $\rho$ in the top half and the ones indicated by $\sigma$, in reverse order, in the bottom half. Let $M_{\rho \sigma}$ be the cofactor of $m_{\rho \sigma}$ in $V$. We showed in §1.4 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 $m_{\rho \sigma}$,

$$
(3) \quad m_{\rho \sigma} = \prod_{I' > j_{r-1}, I' \in \rho} \left( e_{j' - e_j} \right) \prod_{k' > k_{q-r-1}, k' \in \sigma} \left( e_{2q - 1 - k - e_{2q - 1 - k'}} \right) \prod_{j \in \rho; k \in \delta} \left( e_{2q - 1 - k - e_j} \right).
$$
Empty products, which arise if \( r = 1 \) or \( q - 1 \) will be interpreted as equal to one.

Put \( x_j = (j + \frac{1}{2})\pi/2q \),

\[
S_\rho = \prod_{j > i, j \neq \rho} \sin(x_j - x_i), \quad \text{and} \quad S'_\rho = S_\rho \prod_{j \neq \rho, k \neq \rho} \sin(x_j + x_k).
\]

We define, as in the Lemma, §1.3,

\[
y_{\rho \beta} = \prod_{i \neq \rho, k \neq \rho} \cosec(x_j + x_k) = y_{\beta \rho}.
\]

Since \( y_{\rho \beta} y_{\rho \beta} = \prod_{i \neq \rho, k \neq \rho} \cosec(x_j + x_k) \), we get

\[
S_\rho y_{\rho \beta}^{-1} = S'_\rho y_{\beta \rho}.
\]

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

\[
|m_{\rho \beta}| = 2^{q(q-1)/2} S_\rho S_\beta y_{\rho \beta}^{-1} = 2^{q(q-1)/2} S'_\rho S_\beta y_{\beta \rho}.
\]

The cofactor \( M_{\rho \beta} \) is by definition the signed minor of \( V \) from the right, the rows indicated by \( \rho \) in the top and by \( \sigma \), 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,

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

We can now assert that

\[
m_{\rho \sigma} M_{\rho \sigma} = m_{\rho \sigma} M_{\sigma \rho} = c(S'_\rho S_\beta y_{\rho \beta})^2
\]

where \( c = 2^{q(q-1)/2} \).

If \( \gamma = (\gamma_0, \ldots, \gamma_{q-1}) \) is any vector of \( q \) complex components, let \( \gamma' = \prod_{\rho \neq \sigma} \gamma_\rho \), \( \gamma'' = \prod_{\rho \neq \sigma} \gamma_\sigma \). In this notation,

\[
\begin{vmatrix}
\beta_{j0} e_0^\rho & \cdots & \beta_{j0} e_{q-1}^\rho \\
\vdots & \ddots & \vdots \\
\beta_{j-1} e_{j-1}^\rho & \cdots & \beta_{j-1} e_{q-1}^\rho \\
\tilde{a}_{kq} e_{q-1}^k \cdots \tilde{a}_{kq} e_{q-1}^{k-1} & \cdots & \tilde{a}_{kq} e_{q-1}^{q-1} \cdots \tilde{a}_{kq} e_{q-1}^{q-1} \\
\tilde{a}_{k0} e_{q-1}^0 \cdots \tilde{a}_{k0} e_{q-1}^{q-1} & \cdots & \tilde{a}_{k0} e_{q-1}^{q-1} \cdots \tilde{a}_{k0} e_{q-1}^{q-1}
\end{vmatrix} = \beta^\rho \tilde{a}^\rho m_{\rho \sigma}.
\]

The cofactor of this term in the Laplace expansion of \( D \) from the first \( q \) columns is \( \beta^\rho \tilde{a}^\rho M_{\rho \sigma} \).

The sum \( \sum_{\rho, \sigma \neq \rho} \beta^\rho \tilde{a}^\rho \beta^\sigma \tilde{a}^\sigma m_{\rho \sigma} M_{\rho \sigma} \) 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 \( (A, \gamma, \gamma) \) in the variables \( \gamma_r = (\beta^\rho \tilde{a}^\rho)_{\rho \neq \rho} \), with matrix of coefficients

\[
A_r = (m_{\rho \sigma} M_{\rho \sigma})_{\rho, \sigma \neq \rho} = c((S'_\rho S_\beta y_{\rho \beta})^2)_{\rho, \sigma \neq \rho}.
\]

By the Lemma, §1.3, \( (y_{\rho \beta})_{\rho, \sigma \neq \rho} \) is p.d. Since \( S'_\rho \) and \( S_\beta \) are positive, \( c^{-1}A_r \) is also p.d. This holds for \( r = 1, 2, \ldots, q-1 \).
Let \( \gamma_0 = \prod_{j=0}^{q} \alpha_j \beta_j \), \( \gamma_q = \prod_{j=0}^{q} \beta_j \beta_j \) be vectors of one component. The matrices \( A_0 = A_q \) accompanying \( \gamma_0, \gamma_q \) in the expansion have one element

\[
A_0 = A_q = \begin{bmatrix}
\varepsilon^0_0 & \cdots & \varepsilon^0_{q-1} & \varepsilon^q_0 & \cdots & \varepsilon^q_{2q-1} \\
\vdots & & \vdots & \vdots & & \vdots \\
\varepsilon^0_{q-1} & \cdots & \varepsilon^q_{2q-1} & \varepsilon^q_0 & \cdots & \varepsilon^q_{2q-1}
\end{bmatrix}
\]

\[
= (-1)^q \begin{bmatrix}
\varepsilon^0_0 & \cdots & \varepsilon^q_0 & \varepsilon^0_0 & \cdots & \varepsilon^q_{2q-1} \\
\vdots & & \vdots & \vdots & & \vdots \\
\varepsilon^0_{q-1} & \cdots & \varepsilon^q_{2q-1} & \varepsilon^q_0 & \cdots & \varepsilon^q_{2q-1}
\end{bmatrix}.
\]

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 \( r, 0 \leq r \leq q \), and a \( \rho \in Q_r \) such that \( \beta^\rho a^\rho \neq 0 \). The vector \( \gamma_r \) 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 \( L=\{g_0, g_1, \ldots, g_{2q-1}\} \) 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 \( (\gamma_{p \sigma})_{p \in L_r} \) 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,

\[
D_L = \begin{bmatrix}
\beta_{g_0} \varepsilon^{g_0}_0 & \cdots & \beta_{g_0} \varepsilon^{g_0}_{q-1} & \alpha_{g_0} \varepsilon^{g_0}_0 & \cdots & \alpha_{g_0} \varepsilon^{g_0}_{2q-1} \\
\vdots & & \vdots & \vdots & & \vdots \\
\beta_{g_{q-1}} \varepsilon^{g_{q-1}}_0 & \cdots & \beta_{g_{q-1}} \varepsilon^{g_{q-1}}_{q-1} & \alpha_{g_{q-1}} \varepsilon^{g_{q-1}}_0 & \cdots & \alpha_{g_{q-1}} \varepsilon^{g_{q-1}}_{2q-1} \\
\beta_{g_{q-1}} \varepsilon^{g_{q-1}}_0 & \cdots & \beta_{g_{q-1}} \varepsilon^{g_{q-1}}_{q-1} & \alpha_{g_{q-1}} \varepsilon^{g_{q-1}}_0 & \cdots & \alpha_{g_{q-1}} \varepsilon^{g_{q-1}}_{2q-1} \\
\alpha_{g_0} \varepsilon^{g_0}_0 & \cdots & \alpha_{g_0} \varepsilon^{g_0}_{q-1} & \beta_{g_0} \varepsilon^{g_0}_0 & \cdots & \beta_{g_0} \varepsilon^{g_0}_{2q-1}
\end{bmatrix}
\]

BIBLIOGRAPHY


*University of California,*  
*Berkeley, California*