

Some Inequalities for Combinatorial Matrix Functions*

MARVIN MARCUS AND GEORGE W. SOULES

*Department of Mathematics,
University of California,
Santa Barbara, California*

Communicated by Mark Kac

1. INTRODUCTION

In [4] it was conjectured that an analog for permanents of the Fischer theorem for determinants is true. In a paper appearing currently [1], E. Lieb settles this question. One of the results in the present paper is a strengthened version of the conjecture as well as several other variations and extensions of the classical Hadamard-Fischer types of theorems.

Our first result will be a formulation of the Laplace expansion theorem for a very general class of matrix functions. This result is entirely combinatorial in structure and will subsequently be used to prove our main results. In the latter part of the paper we prove a general theorem on products of generalized functions of principal submatrices that contains many of the classical theorems on determinants. The class of generalized matrix functions includes, e.g., the permanent, the determinant, and the product of the main diagonal entries, which are of substantial interest in proving inequalities of combinatorial significance (see, e.g., [6], [7], [8]).

Let G be a subgroup of S_n , the symmetric group of degree n , and let χ be a complex valued character on G of degree 1, i.e., χ is a non-zero homomorphism of G into the complex numbers C . A generalized matrix function of the n -square matrix A is defined by

* The research of both authors was supported by the U.S. Air Force Office of Scientific Research under Grant 698-65.

$$d_{\chi}^G(A) = \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^n a_{i\sigma(i)}. \quad (1)$$

For example, if $G = S_n$, $\chi = 1$ then d_{χ}^G is the permanent function. Observe that if φ and τ are permutations in G and $B = (b_{ij}) = (a_{\varphi(i), \tau(j)})$ then

$$\begin{aligned} d_{\chi}^G(B) &= \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^n b_{i\sigma(i)} \\ &= \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^n a_{\varphi(i), \tau\sigma(i)} \\ &= \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^n a_{i, \tau\sigma\varphi^{-1}(i)} \\ &= \chi(\tau^{-1}\varphi) d_{\chi}^G(A). \end{aligned} \quad (2)$$

In subsequent statements we shall have occasion to use a slight extension of the definition (1). Suppose then that $r + s = n$, $r, s \geq 1$, and G_r is a subgroup of G each member of which holds the integers $r + 1, \dots, r + s = n$ individually fixed. Also, let G_s be a subgroup of G each member of which holds the integers $1, \dots, r$ individually fixed. Then if A and B are respectively r -square and s -square matrices define

$$d_{\chi}^r(A) = d_{\chi}^{G_r}(A \dot{+} I_s) = \sum_{\sigma \in G_r} \chi(\sigma) \prod_{i=1}^r a_{i\sigma(i)}, \quad (3)$$

$$d_{\chi}^s(B) = d_{\chi}^{G_s}(I_r \dot{+} B) = \sum_{\sigma \in G_s} \chi(\sigma) \prod_{i=1}^s b_{i, \sigma(r+i)-r}. \quad (4)$$

We next define a general notion of submatrix. Let $I_{m,n}^r$ denote the totality of n^m sequences $\omega = (\omega_1, \dots, \omega_m)$, $1 \leq \omega_i \leq n$, $i = 1, \dots, m$. If A is any n -square matrix and ω and γ are in $I_{m,n}^r$ then $A[\omega \mid \gamma]$ is the m -square matrix whose i, j entry is $a_{\omega_i \gamma_j}$, $i, j = 1, \dots, m$. Note that if X is n -square and $B = X[r + 1, \dots, n \mid r + 1, \dots, n]$ then from (4) (since $b_{ij} = x_{r+i, r+j}$)

$$\begin{aligned} d_{\chi}^s(B) &= \sum_{\sigma \in G_s} \chi(\sigma) \prod_{i=1}^s b_{i, \sigma(r+i)-r} \\ &= \sum_{\sigma \in G_s} \chi(\sigma) \prod_{i=1}^s x_{r+i, \sigma(r+i)} \\ &= \sum_{\sigma \in G_s} \chi(\sigma) \prod_{i=r+1}^n x_{i, \sigma(i)}. \end{aligned} \quad (5)$$

2. STATEMENT OF RESULTS

Our first result is a generalization of the classical Laplace expansion theorem.

THEOREM 1. *Let A be an n -square matrix over an arbitrary field, G a subgroup of S_n and χ a character of degree 1 on G . Suppose G_r (resp. G_s) is a subgroup of G , $r + s = n$, which leaves the integers $r + 1, \dots, n$ (resp. $1, \dots, r$) elementwise fixed. Let R be a system of left coset representatives of the direct product $G_r \times G_s$ in G , i.e.,*

$$G = \bigcup_{\sigma \in R} \sigma(G_r \times G_s).$$

Then

$$d_\chi^G(A) = \sum_R \chi(\sigma^{-1}\tau) d_\chi^r(A[\sigma(1), \dots, \sigma(r) \mid \tau(1), \dots, \tau(r)]) \times d_\chi^s(A[\sigma(r + 1), \dots, \sigma(n) \mid \tau(r + 1), \dots, \tau(n)]). \tag{6}$$

The summation can be either over all $\sigma \in R$ with τ fixed in G or over all $\tau \in R$ with σ fixed in G .

Using (6) we prove the following analog of the Fischer inequality for determinants (see (14)).

THEOREM 2. *If A is an n -square positive definite Hermitian matrix, $1 \leq r \leq n$, and*

$$\begin{aligned} A_1 &= A[1, \dots, r \mid 1, \dots, r], \\ A_2 &= A[r + 1, \dots, n \mid r + 1, \dots, n], \\ B &= A[1, \dots, r \mid r + 1, \dots, n] \end{aligned}$$

then

$$\text{per}(A) - \text{per}(A_1) \text{per}(A_2) \geq \mu^{n-2} \|B\|^2 \tag{7}$$

where μ is the minimum eigenvalue of A and $\|B\|$ is the Euclidean norm of the matrix B , i.e., $\text{tr}(BB^*)^{1/2}$.

In [1] Lieb proves that

$$\text{per}(A) - \text{per}(A_1) \text{per}(A_2) \geq 0 \tag{8}$$

with equality if and only if $A = A_1 \dot{+} A_2$. The inequality (7) gives an estimate on the size of the difference in (8). The case $r = 1$ of (8) was

proved in [2] and for $r > 1$, (8) was conjectured in [4]. Also, (7) is proved for $r = 1$ in a paper to appear in the Proceedings of the Symposium on Inequalities (Wright-Patterson Air Force Base, August 1965).

With the aid of Theorem 2 the Hadamard determinant theorem can be sharpened as follows.

THEOREM 3. *Let A be an n -square positive definite Hermitian matrix and let f be either*

$$\text{per}(A) = \prod_{i=1}^n a_{ii} \text{ or } \prod_{i=1}^n a_{ii} = \det(A).$$

Then

$$\eta^{n-2} \sum_{i < j} |a_{ij}|^2 \geq f \geq \mu^{n-2} \sum_{i < j} |a_{ij}|^2 \quad (9)$$

where μ and η are respectively the minimum and maximum eigenvalues of A . The upper (resp. lower) inequality holds if and only if A is a diagonal matrix, or by a simultaneous permutation of its rows and columns (i.e., $P^T A P$, P a permutation matrix), A may be brought to the form

$$A = \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix} \oplus \lambda J_{n-2} \quad (10)$$

where $\lambda = \eta$ (resp. $\lambda = \mu$).

By applying Theorem 3 we have the following result.

COROLLARY. *If $G = A_n$, the alternating group, and $\chi = 1$ then*

$$|d_{\chi}^A(A) - \prod_{i=1}^n a_{ii}| \leq \eta^{n-2} \sum_{1 \leq i < j \leq n} |a_{ij}|^2 \quad (11)$$

and equality occurs in (11) if and only if A is a diagonal matrix.

Our next result is another extension of the Fischer inequality and contains new results even in the case of the determinant function. Before going on we introduce several preliminary combinatorial notions.

The group $G \subset S_m$ operates on $\Gamma_{m,n}$ in the following obvious way: if $\sigma \in G$ and $\omega = (\omega_1, \dots, \omega_m) \in \Gamma_{m,n}$ then $\omega^\sigma = (\omega_{\sigma(1)}, \dots, \omega_{\sigma(m)})$. From each orbit in $\Gamma_{m,n}$ select a sequence which is first in lexicographic order and denote the resulting system of distinct representatives by Δ . Let G_ω denote the stabilizer of ω , i.e., $\sigma \in G_\omega$ if and only if $\omega^\sigma = \omega$,

and let $\nu(\omega)$ denote the order of G_ω . Let $\bar{\Delta}$ denote the subset of Δ consisting of those $\omega \in \Delta$ for which the character χ is identically equal to 1 on G_ω . For example, if $G = S_n$ and $\chi = \varepsilon$, the signum function, then Δ is the set of non-decreasing sequences in $\Gamma_{m,n}$ and $\bar{\Delta}$ is the subset of strictly increasing sequences. For this choice of G and χ we shall use the following special notation: $G_{m,n} = \Delta$, $Q_{m,n} = \bar{\Delta}$. If ω is any sequence in $\Gamma_{m,n}$ we will let $m_i(\omega)$ denote the multiplicity of occurrence of the integer i in ω , $i = 1, \dots, n$.

THEOREM 4. *For each $\omega \in \bar{\Delta}$ let b_ω be a non-negative number. If A is an n -square positive definite Hermitian matrix then*

$$\prod_{\omega \in \bar{\Delta}} \left(\frac{d_\chi^G(A[\omega | \omega])}{\nu(\omega)} \right)^{b_\omega} \geq \min_{\sigma \in S_n} \prod_{i=1}^n \lambda_{\sigma(i)}^{q_i} \tag{12}$$

where

$$q_i = \sum_{\omega \in \bar{\Delta}} b_\omega m_i(\omega), \quad i = 1, \dots, n \quad \text{and} \quad \lambda_1 \geq \dots \geq \lambda_n$$

are the eigenvalues of A .

For the determinant and permanent functions we have the following results for the n -square positive definite Hermitian matrix A .

THEOREM 5. *Let $\omega^{(1)}, \dots, \omega^{(r)}$ be strictly increasing sequences of arbitrary lengths not exceeding n . Suppose s distinct integers $\alpha_1, \dots, \alpha_s$ appear among $\omega^{(1)}, \dots, \omega^{(r)}$ and α_t occurs a total of p_t times, $p_1 \geq \dots \geq p_s \geq 1$. Then*

$$\prod_{j=1}^r \det(A[\omega^{(j)} | \omega^{(j)}]) \geq \prod_{j=1}^s \lambda_{n-j+1}^{p_j} . \tag{13}$$

In particular if $r = 2$, and $\omega^{(1)}$ and $\omega^{(2)}$ are disjoint and exhaust all of $1, \dots, n$ then

$$\det(A[\omega^{(1)} | \omega^{(1)}]) \det(A[\omega^{(2)} | \omega^{(2)}]) \geq \det(A), \tag{14}$$

the classical Fischer inequality. If the sequences $\omega^{(1)}, \dots, \omega^{(r)}$ are non-decreasing then

$$\prod_{j=1}^r \frac{\text{per}(A[\omega^j | \omega^j])}{r_j} \geq \prod_{j=1}^s \lambda_{n-j+1}^{p_j} \tag{15}$$

where

$$r_j = \prod_{l=1}^j m_l(\omega^{(j)})!$$

In particular if ω is any non-decreasing sequence of length not exceeding n then

$$\text{per}(A[\omega]) \geq \prod_{l=1}^u m_l(\omega)! \prod_{j=1}^s \lambda_{n-j+1}^{p_j}. \tag{16}$$

Various choices of $\omega^{(1)}, \dots, \omega^{(r)}$ in Theorems 4 and 5 yield many standard results. For example, if $\omega = (1, \dots, n)$ then (12) becomes

$$d_{\mathbb{Z}}^G(A) \geq \prod_{j=1}^u \lambda_{n-j+1} = \det(A),$$

the inequality of Schur [9]. Again, suppose $\omega^{(1)}, \dots, \omega^{(r)}$ are strictly increasing sequences of lengths not exceeding n and moreover each of $1, \dots, n$ appears k times among all the $\omega^{(j)}, j = 1, \dots, r$. Then (13) becomes

$$\prod_{j=1}^r \det(A[\omega^{(j)} \mid \omega^{(j)}]) \geq \det(A)^k.$$

3. PROOFS

To prove Theorem 1 we write the left coset decomposition of G modulo $H = G_r \times G_s$:

$$G = \bigcup_{\tau \in R} \tau H.$$

Then summing over the individual cosets we have

$$\begin{aligned} d_{\mathbb{Z}}^G(A) &= \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^n a_{i\sigma(i)} \\ &= \sum_{\tau \in R} \sum_{\pi \in H} \chi(\tau\pi) \prod_{i=1}^n a_{i,\tau\pi(i)}. \end{aligned} \tag{17}$$

Any summation over H can be effected by summing separately over G_r and G_s since H is the direct product of these groups. We recall also that the permutations in G_s leave $1, \dots, r$ elementwise fixed and similarly the permutations in G_r leave $r+1, \dots, r+s = n$, elementwise fixed. Thus continuing (17) we have

$$\begin{aligned}
 d_\chi^G(A) &= \sum_{\tau \in R} \sum_{\varphi \in G_r} \sum_{\theta \in G_s} \chi(\tau\varphi\theta) \prod_{i=1}^n a_{i,\tau\varphi\theta(i)} \\
 &= \sum_{\tau \in R} \chi(\tau) \sum_{\varphi \in G_r} \chi(\varphi) \prod_{i=1}^r a_{i,\tau\varphi(i)} \sum_{\theta \in G_s} \chi(\theta) \prod_{i=r+1}^n a_{i,\tau\theta(i)} \\
 &= \sum_{\tau \in R} \chi(\tau) d_\chi^r(A[1, \dots, r \mid \tau(1), \dots, \tau(r)]) \\
 &\quad \times d_\chi^s(A[r+1, \dots, n \mid \tau(r+1), \dots, \tau(n)]).
 \end{aligned} \tag{18}$$

In the last equality in (18) we have used (3) and (5) with

$$B = A[r+1, \dots, n \mid \tau(r+1), \dots, \tau(n)].$$

To obtain the expansion of $d_\chi^G(A)$ by columns (i.e., summation over τ) in (6), replace A in (18) by $C = A[\sigma(1), \dots, \sigma(n) \mid 1, \dots, n]$, $\sigma \in G$, and use (2) to conclude that

$$\begin{aligned}
 d_\chi^G(A) &= \chi(\sigma^{-1}) d_\chi^G(C) \\
 &= \chi(\sigma^{-1}) \sum_{\tau \in R} \chi(\tau) d_\chi^r(C[1, \dots, r \mid \tau(1), \dots, \tau(r)]) \\
 &\quad \times d_\chi^s(C[r+1, \dots, n \mid \tau(r+1), \dots, \tau(n)]) \\
 &= \sum_{\tau \in R} \chi(\sigma^{-1}\tau) d_\chi^r(A[\sigma(1), \dots, \sigma(r) \mid \tau(1), \dots, \tau(r)]) \\
 &\quad \times d_\chi^s(A[\sigma(r+1), \dots, \sigma(n) \mid \tau(r+1), \dots, \tau(n)]).
 \end{aligned} \tag{19}$$

To establish the expansion by rows (i.e., summation over σ in (6)) we argue as follows. By taking the inverses of the elements in each coset σH , we obtain

$$\begin{aligned}
 G &= \bigcup_{\sigma \in R} \sigma H \\
 &= \bigcup_{\sigma \in R} H\sigma^{-1}.
 \end{aligned}$$

Thus as in (18) we have

$$\begin{aligned}
 d_\chi^G(A) &= \sum_{\sigma \in R} \sum_{\varphi \in G_r} \sum_{\theta \in G_s} \chi(\varphi\theta\sigma^{-1}) \prod_{i=1}^n a_{i,\varphi\theta\sigma^{-1}(i)} \\
 &= \sum_{\sigma \in R} \sum_{\varphi \in G_r} \sum_{\theta \in G_s} \chi(\varphi\theta\sigma^{-1}) \prod_{i=1}^n a_{\sigma(i),\varphi\theta(i)} \\
 &= \sum_{\sigma \in R} \chi(\sigma^{-1}) d_\chi^r(A[\sigma(1), \dots, \sigma(r) \mid 1, \dots, r]) \\
 &\quad \times d_\chi^s(A[\sigma(r+1), \dots, \sigma(n) \mid r+1, \dots, n])
 \end{aligned}$$

and we complete the proof by applying (2) again.

The choices of G_r and G_s in Theorem 1 that are pertinent to our needs are the following. Let $G = S_n$ and let G_r be the entire subgroup of S_n holding the integers $r + 1, \dots, n$ individually fixed, i.e., G_r is isomorphic to S_r . Similarly let G_s be the entire subgroup of S_n isomorphic to S_{n-r} holding the integers $1, \dots, r$ individually fixed. There are two choices of a system of left coset representatives that are useful.

$$(a) \quad R = \left\{ \tau_\gamma = \begin{pmatrix} 1 & \dots & r & r+1 & \dots & n \\ \gamma_1 & \dots & \gamma_r & \gamma'_1 & \dots & \gamma'_{n-r} \end{pmatrix} \mid \gamma = (\gamma_1, \dots, \gamma_r) \in Q_{r,n} \right\} \quad (20)$$

where $\gamma'_1 < \dots < \gamma'_{n-r}$ is the sequence in $Q_{n-r,n}$ complementary to γ in $1, \dots, n$. There are $\binom{n}{r}$ permutations in R and we can see that no two are equivalent modulo $G_r \times G_s$ as follows: if $\tau_\alpha^{-1}\tau_\beta \in G_r \times G_s$ then

$$\begin{pmatrix} \alpha_1 & \dots & \alpha_r & \alpha'_1 & \dots & \alpha'_{n-r} \\ 1 & \dots & r & r+1 & \dots & n \end{pmatrix} \begin{pmatrix} 1 & \dots & r & r+1 & \dots & n \\ \beta_1 & \dots & \beta_r & \beta'_1 & \dots & \beta'_{n-r} \end{pmatrix}$$

must map the set $\{1, \dots, r\}$ onto itself. Thus the integers in α and β are the same and since α and β are in $Q_{r,n}$ it follows that $\alpha = \beta$.

$$(b) \quad R = \bigcup_{m=0}^M R_m \quad \text{where } M = \min(r, s), R_0 = \{e\} \text{ and for } m > 0 \quad (21)$$

$$R_m = \{ \sigma_{\alpha\beta} = (\alpha_1\beta_1) (\alpha_2\beta_2) \dots (\alpha_m\beta_m) \mid \alpha_1 < \dots < \alpha_m \leq r < \beta_1 < \dots < \beta_m \leq n \}.$$

That is, for $m > 0$, R_m is the set of all products of transpositions of the form $\prod_{i=1}^m (\alpha_i\beta_i)$ where each α_i is at most r and each β_i is greater than r .

Clearly there are $\binom{r}{m} \binom{s}{m}$ permutations in R_m and moreover $R_i \cap R_j$ is empty if $i \neq j$. Hence, the number of elements in R is

$$\sum_{m=0}^M \binom{r}{m} \binom{s}{m} = \binom{n}{r}.$$

On the other hand, suppose two elements $\sigma_{\alpha\beta} \in R_m$ and $\sigma_{\gamma\delta} \in R_p$ are equivalent modulo $G_r \times G_s$. Then $\sigma_{\alpha\beta}\sigma_{\gamma\delta} = \sigma_{\alpha\beta}^{-1}\sigma_{\gamma\delta} \in G_r \times G_s$ and hence must map the sets $\{1, \dots, r\}$ and $\{r + 1, \dots, n\}$ onto themselves. But

$$\sigma_{\alpha\beta}\sigma_{\gamma\delta} = (\alpha_1\beta_1) (\alpha_2\beta_2) \dots (\alpha_m\beta_m) (\gamma_1\delta_1) (\gamma_2\delta_2) \dots (\gamma_p\delta_p),$$

and hence

$$\sigma_{\alpha\beta}\sigma_{\gamma\delta}(\gamma_t) = \sigma_{\alpha\beta}(\delta_t) \in \{1, \dots, r\}, \quad t = 1, \dots, p. \tag{22}$$

But $\delta_t \in \{r + 1, \dots, n\}$ and hence unless $\delta_t \in \{\beta_1, \dots, \beta_m\}$, $\sigma_{\alpha\beta}$ would leave δ_t fixed, in contradiction to (22). Thus $\{\delta_1, \dots, \delta_p\} \subset \{\beta_1, \dots, \beta_m\}$. Also, since

$$(\sigma_{\alpha\beta}\sigma_{\gamma\delta})^{-1} = \sigma_{\gamma\delta}\sigma_{\alpha\beta} \in G_r \times G_s,$$

we can conclude that $\{\beta_1, \dots, \beta_m\} \subset \{\delta_1, \dots, \delta_p\}$. Thus $p = m$ and $\beta = \delta$. Similarly $\alpha = \gamma$ and hence $\sigma_{\alpha\beta} = \sigma_{\gamma\delta}$.

With the choices of $G = S_n, G_r, G_s$, and R as given in (a), (19) becomes, with $\sigma \in S_n$ and τ_γ as in (20),

$$\begin{aligned} d(A) &= \sum_{\tau_\gamma \in R} \chi(\sigma^{-1}\tau_\gamma) d(A[\sigma(1), \dots, \sigma(r) \mid \tau_\gamma(1), \dots, \tau_\gamma(r)]) \\ &\quad \times d(A[\sigma(r + 1), \dots, \sigma(n) \mid \tau_\gamma(r + 1), \dots, \tau_\gamma(n)]) \\ &= \sum_{\gamma \in Q_{r,n}} \chi(\sigma^{-1}\tau_\gamma) d(A[\sigma(1), \dots, \sigma(r) \mid \gamma_1, \dots, \gamma_r]) \\ &\quad \times d(A[\sigma(r + 1), \dots, \sigma(n) \mid \gamma'_1, \dots, \gamma'_{n-r}]) \end{aligned} \tag{23}$$

where $d(A)$ is either $\text{per}(A)$ or $\text{det}(A)$.

The summation (23) is of course the usual Laplace column expansion for determinants or permanents according as $\chi = \varepsilon$ or $\chi \equiv 1$.

We can also investigate (6) for the choice of R in (b). Thus (6) becomes, with $\tau \in S_n$ and $\sigma_{\alpha\beta}$ as in (21),

$$\begin{aligned} d(A) &= \sum_{m=0}^M \sum_{\sigma_{\alpha\beta} \in R_m} \chi(\sigma_{\alpha\beta}\tau) d(A[\sigma_{\alpha\beta}(1), \dots, \sigma_{\alpha\beta}(r) \mid \tau(1), \dots, \tau(r)]) \\ &\quad \times d(A[\sigma_{\alpha\beta}(r + 1), \dots, \sigma_{\alpha\beta}(n) \mid \tau(r + 1), \dots, \tau(n)]). \end{aligned} \tag{24}$$

Let $Q'_{m,n-r} = \{\beta = (\beta_1, \dots, \beta_m) \mid r + 1 \leq \beta_1 < \beta_2 < \dots < \beta_m \leq n\}$, $m = 1, \dots, M$, and let $f(\sigma_{\alpha\beta})$ be the summand in (24) so that

$$\begin{aligned} d_\chi^{S_n}(A) &= \sum_{m=0}^M \sum_{\sigma_{\alpha\beta} \in R_m} f(\sigma_{\alpha\beta}) \\ &= \sum_{m=0}^M \sum_{\alpha \in Q_{m,r}} \sum_{\beta \in Q'_{m,n-r}} f(\sigma_{\alpha\beta}). \end{aligned} \tag{25}$$

Now the sequence $(\sigma_{\alpha\beta}(1), \dots, \sigma_{\alpha\beta}(r))$ is obtained from the sequence $(1, \dots, r)$ by replacing α_t with β_t , $t = 1, \dots, m$. Similarly the sequence

$(\sigma_{a_i}(r + 1), \dots, \sigma_{a_j}(n))$ is obtained from the sequence $(r + 1, \dots, n)$ by replacing β_t with α_t , $t = 1, \dots, m$. In the case $\chi = 1$ and thus $d(A) = \text{per}(A)$, (25) becomes

$$\text{per}(A) = \sum_{m=0}^M F(m) \tag{26}$$

where

$$F(m) = \sum_{\sigma_{a_j} \in R_m} f(\sigma_{a_j}).$$

Observe that

$$\begin{aligned} F(0) &= f(e) \\ &= \text{per}(A[1, \dots, r \mid \tau(1), \dots, \tau(r)]) \\ &\quad \times \text{per}(A[r + 1, \dots, n \mid \tau(r + 1), \dots, \tau(n)]) \end{aligned} \tag{27}$$

and that for $m > 0$

$$\begin{aligned} F(m) &= \sum_{\alpha \in Q_{m,r}} \sum_{\beta \in Q'_{m,n-r}} \text{per}(A[\beta_1, \dots, \beta_m, \alpha'_1, \dots, \alpha'_{r-m} \mid \tau(1), \dots, \tau(r)]) \\ &\quad \times \text{per}(A[\alpha_1, \dots, \alpha_m, \beta'_1, \dots, \beta'_{s-m} \mid \tau(r + 1), \dots, \tau(n)]) \end{aligned} \tag{28}$$

where $(\alpha'_1, \dots, \alpha'_{r-m})$ is the sequence in $Q_{r-m,r}$ complementary to α in $1, \dots, r$, and $(\beta'_1, \dots, \beta'_{s-m})$ is the sequence in $Q'_{s-m,n-r}$ complementary to β in $r + 1, \dots, n$.

Before proving Theorem 2 we list certain properties of the induced power matrix $P_k(A)$ where A is an n -square matrix [7]. The entries of $P_k(A)$ are the numbers

$$\text{per}(A[\alpha \mid \beta]) / \sqrt{v(\alpha)v(\beta)}, \alpha, \beta \in G_{k,n}$$

where

$$v(\alpha) = \prod_{t=1}^n m_t(\alpha)!,$$

and these are arranged doubly lexicographically in α and β in the $\binom{n+k-1}{k}$ -square matrix $P_k(A)$. If A is positive definite Hermitian so is $P_k(A)$ and the eigenvalues of $P_k(A)$ are the numbers $\prod_{t=1}^n \lambda_t^{m_t(\omega)}$, $\omega \in G_{k,n}$, where $\lambda_1 \geq \dots \geq \lambda_n$ are the eigenvalues of A . Observe that $P_k(A)$ has as a principal submatrix the $\binom{n}{k}$ -square matrix $H_k(A)$

lying in positions (α, β) , $\alpha, \beta \in Q_{k,n}$. Accordingly the entries of $H_k(A)$ are the numbers

$$\text{per}(A[\alpha | \beta]), \alpha, \beta \in Q_{k,n}.$$

It follows from a standard theorem on Hermitian matrices that the eigenvalues of $H_k(A)$ are in the interval $\mu^k \leq x \leq \eta^k$ where $\mu = \lambda_n$ and $\eta = \lambda_1$ are the smallest and largest eigenvalues of A respectively.

To prove Theorem 2 we first observe that for $\tau = e$ in (27) we have

$$\begin{aligned} F(0) &= \text{per}(A_1) \text{per}(A_2) \\ &> 0, \end{aligned}$$

since A_1 and A_2 are positive definite [9]. For $0 < m < M$ we show $F(m) \geq 0$ by using (28) with $\tau = e$. We first expand the permanents in the summand in (28) by use of (23) with $\sigma = e$. Thus

$$\begin{aligned} &\text{per}(A[\beta_1, \dots, \beta_m, \alpha'_1, \dots, \alpha'_{r-m} | 1, \dots, r]) \\ &= \sum_{\gamma \in Q_{m,r}} \text{per}(A[\beta_1, \dots, \beta_m | \gamma]) \text{per}(A[\alpha'_1, \dots, \alpha'_{r-m} | \gamma']) \end{aligned} \tag{29}$$

where γ' is the strictly increasing sequence in $1, \dots, r$ complementary to γ . Similarly

$$\begin{aligned} &\text{per}(A[\alpha_1, \dots, \alpha_m, \beta'_1, \dots, \beta'_{s-m} | r+1, \dots, n]) \\ &= \sum_{\delta \in Q'_{m,n-r}} \text{per}(A[\alpha_1, \dots, \alpha_m | \delta]) \text{per}(A[\beta'_1, \dots, \beta'_{s-m} | \delta']). \end{aligned} \tag{30}$$

Thus using (29) and (30), (28) becomes

$$\begin{aligned} F(m) &= \sum_{\alpha, \gamma \in Q_{m,r}} \sum_{\beta, \delta \in Q'_{m,n-r}} \text{per}(A[\beta | \gamma]) \text{per}(A[\alpha' | \gamma']) \text{per}(A[\alpha | \delta]) \\ &\quad \times \text{per}(A[\beta' | \delta']). \end{aligned} \tag{31}$$

Let K be the $\binom{r}{m}$ -square matrix whose (α, γ) entry, $\alpha, \gamma \in Q_{m,r}$, is

$$K_{\alpha\gamma} = \text{per}(A[\alpha' | \gamma']) \tag{32}$$

and observe that by reversing the order of both the rows and columns of K that we obtain $H_{r-m}(A_1)$. Hence it follows that the eigenvalues of K and $H_{r-m}(A_1)$ are the same and K is positive definite Hermitian. Further,

let L be the $\binom{s}{m}$ -square matrix whose (β, δ) entry, $\beta, \delta \in Q'_{m,n-r}$, is

$$L_{\beta\delta} = \text{per}(A[\beta' \mid \delta']) \tag{33}$$

and similarly L has the same eigenvalues as $H_{s-m}(A_2)$ and is positive definite Hermitian.

Now let C be the $\binom{r}{m} \times \binom{s}{m}$ matrix whose (γ, β) entry is

$$C_{\gamma\beta} = \text{per}(A[\beta \mid \gamma]), \tag{34}$$

$\gamma \in Q_{m,r}$, $\beta \in Q'_{m,n-r}$. Note that

$$\text{per}(A[\alpha \mid \delta]) = \overline{\text{per}(A[\delta \mid \alpha])} = \bar{C}_{\alpha\delta} \tag{35}$$

From (31), (32), (33), (34), and (35) we have

$$\begin{aligned} F(m) &= \sum_{\alpha, \gamma \in Q_{m,r}} \sum_{\beta, \delta \in Q'_{m,n-r}} C_{\gamma\beta} K_{\alpha\gamma} \bar{C}_{\alpha\delta} L_{\beta\delta} \\ &= \sum_{\alpha, \beta, \gamma, \delta} K_{\alpha\gamma} C_{\gamma\beta} L_{\beta\delta} (C^*)_{\delta\alpha} \\ &=: \sum_{\alpha} (KCLC^*)_{\alpha\alpha} \\ &= \text{tr}(KCLC^*) \\ &\geq 0 \end{aligned} \tag{36}$$

because both K and CLC^* are positive semi-definite. Hence $F(m) \geq 0$ for $m = 1, 2, \dots, M - 1$. We next compute $F(M)$. From (28) with $M = s < r$ we have

$$\begin{aligned} F(M) &= \sum_{\alpha \in Q_{s,r}} \text{per}(A[r + 1, \dots, n, \alpha'_1, \dots, \alpha'_{r-s} \mid 1, \dots, r]) \\ &\quad \times \text{per}(A[\alpha_1, \dots, \alpha_s \mid r + 1, \dots, n]). \end{aligned} \tag{37}$$

Now from (23) applied to $\text{per}(A[r + 1, \dots, n, \alpha'_1, \dots, \alpha'_{r-s} \mid 1, \dots, r])$ we obtain

$$\begin{aligned} F(M) &= \sum_{\alpha, \beta \in Q_{s,r}} \text{per}(A[r + 1, \dots, n \mid \beta]) \text{per}(A[\alpha' \mid \beta']) \\ &\quad \times \text{per}(A[\alpha \mid r + 1, \dots, n]) \\ &= (Ku, u) \\ &\geq 0, \end{aligned} \tag{38}$$

where K is defined in (32) and u is the $\binom{r}{s}$ -tuple whose β entry is

$$\text{per}(A[r + 1, \dots, n \mid \beta]), \beta \in Q_{s,r}.$$

If $r = s = M$, then in (37) $(\alpha_1, \dots, \alpha_s)$ is $(1, \dots, r)$ and

$$\begin{aligned} F(M) &= \text{per}(A[r + 1, \dots, n \mid 1, \dots, r]) \text{per}(A[1, \dots, r \mid r + 1, \dots, n]) \\ &= |\text{per}(B)|^2 \\ &\geq 0. \end{aligned} \tag{39}$$

Hence from (26), (36), (38), and (39) we have

$$\begin{aligned} \text{per}(A) - \text{per}(A_1)\text{per}(A_2) &= \sum_{m=1}^M F(m) \\ &\geq F(1). \end{aligned} \tag{40}$$

We compute from (36) that

$$F(1) = \text{tr}(KCLC^*)$$

where in this case $C = \bar{B}$. Thus

$$\begin{aligned} F(1) &= \text{tr}(K\bar{B}L\bar{B}^*) \\ &\geq \lambda_{\min}(K)\text{tr}(\bar{B}L\bar{B}^*) \\ &= \lambda_{\min}(K)\text{tr}(L\bar{B}^*\bar{B}) \\ &\geq \lambda_{\min}(K)\lambda_{\min}(L) \| B \|^2 \\ &= \lambda_{\min}(H_{r-1}(A_1))\lambda_{\min}(H_{s-1}(A_2)) \| B \|^2 \\ &\geq \lambda_{\min}^{r-1}(A_1)\lambda_{\min}^{s-1}(A_2) \| B \|^2 \\ &\geq \mu^{r-1}\mu^{s-1} \| B \|^2 \\ &= \mu^{n-2} \| B \|^2, \end{aligned}$$

where $\lambda_{\min}(X)$ denotes the minimum eigenvalue of the Hermitian matrix X . This, together with (40), completes the proof of Theorem 2. We note that for $r = s > 1$ we have also proved that $\text{per}(A) \geq \text{per}(A_1)\text{per}(A_2) + |\text{per}(B)|^2 + \mu^{n-2} \| B \|^2$, which improves Lieb's result.

An inductive argument applied to Theorem 2 also proves that if A is partitioned into submatrices with square blocks down the main diagonal, i.e.,

$$A = (A_{ij}), \quad i, j = 1, \dots, k,$$

then

$$\text{per}(A) - \prod_{i=1}^k \text{per}(A_{ii}) \geq \mu^{n-2} \sum_{1 \leq i < j \leq k} |A_{ij}|^2.$$

Of course if all A_{ii} are 1-square we have

$$\text{per}(A) - \prod_{i=1}^n a_{ii} \geq \mu^{n-2} \sum_{1 \leq i < j \leq n} |a_{ij}|^2, \tag{41}$$

one of the inequalities in Theorem 3.

We now prove the upper inequality in (9) for the permanent and obtain the cases of equality. First define

$$A_m = A[1, \dots, m \mid 1, \dots, m], \quad 1 \leq m \leq n,$$

and use (23) with $r = m$, $\sigma = e$, to obtain

$$\text{per}(A_{m+1}) = \sum_{j=1}^{m+1} \text{per}(A[1, \dots, m \mid 1, \dots, \hat{j}, \dots, m+1]) a_{m+1,j}$$

where \hat{j} means that the integer j is deleted. Similarly we have

$$\begin{aligned} & \text{per}(A[1, \dots, m \mid 1, \dots, \hat{j}, \dots, m+1]) \\ &= \sum_{i=1}^m \text{per}(A[1, \dots, \hat{i}, \dots, m \mid 1, \dots, \hat{j}, \dots, m]) a_{i,m+1} \end{aligned}$$

so that

$$\text{per}(A_{m+1}) = a_{m+1,m+1} \text{per}(A_m) + \sum_{i,j=1}^m \bar{a}_{m+1,i} a_{m+1,j} \text{per}(A_m(i \mid j))$$

where $X(i \mid j)$ is the submatrix of X obtained by deleting row i and column j from X . Thus we can write

$$\text{per}(A_{m+1}) = a_{m+1,m+1} \text{per}(A_m) + (K_m u_m, u_m) \tag{42}$$

where K_m is the m -square matrix whose (i, j) entry is $\text{per}(A_m(i \mid j))$ and $u_m = (a_{m+1,1}, \dots, a_{m+1,m})$. As in the proof of Theorem 2 we know that

$$\lambda_{\max}(K_m) \leq \eta^{m-1}. \tag{43}$$

By an obvious induction on (42) we can write

$$\text{per}(A) - \prod_{i=1}^n a_{ii} = (K_{n-1}u_{n-1}, u_{n-1}) + \sum_{s=1}^{n-2} \prod_{j=s+2}^n a_{jj}(K_s u_s, u_s) \quad (44)$$

and hence since $a_{ii} \leq \eta$, $i = 1, \dots, n$, we have

$$\begin{aligned} \text{per}(A) - \prod_{i=1}^n a_{ii} &\leq \eta^{n-2} \|u_{n-1}\|^2 + \sum_{s=1}^{n-2} \eta^{n-s-1} \eta^{s-1} \|u_s\|^2 \\ &= \eta^{n-2} \sum_{s=1}^{n-1} \|u_s\|^2 \\ &= \eta^{n-2} \sum_{1 \leq i < j \leq n} |a_{ij}|^2. \end{aligned} \quad (45)$$

If A is diagonal or A has the form (10) it is clear that the upper inequality (9) is equality. Conversely suppose equality holds in (9) and (45). Then if A is not diagonal let t be the smallest integer for which $u_{t-1} \neq 0$, $2 \leq t \leq n$, i.e., $a_{tk} \neq 0$ for some k , $1 \leq k \leq t-1$, and from (44) and (45)

$$\prod_{j=t+1}^n a_{jj}(K_{t-1}u_{t-1}, u_{t-1}) = \eta^{n-t}(K_{t-1}u_{t-1}, u_{t-1}).$$

Since A is positive definite K_{t-1} is positive definite and hence

$$(K_{t-1}u_{t-1}, u_{t-1}) \neq 0.$$

Thus

$$a_{t+1,t+1} = \dots = a_{nn} = \eta.$$

But since η is the maximum eigenvalue of A it follows that A is a direct sum of the form

$$A = A_t \dot{+} \eta I_{n-t},$$

and

$$A_t = \begin{bmatrix} a_{11} & 0 & \cdot & \cdot & \cdot & 0 & \bar{a}_{11} \\ 0 & \cdot & & & & & \cdot \\ \vdots & & & & & 0 & \vdots \\ 0 & 0 & \cdot & & & a_{t-1,t-1} & \bar{a}_{t,t-1} \\ a_{t1} & \cdot & \cdot & \cdot & \cdot & a_{t,t-1} & a_{tt} \end{bmatrix}.$$

If $t = 2$, A has the form (10). If t were greater than 2 we would have

$$\begin{aligned} \text{per}(A_t) &= \prod_{i=1}^t a_{ii} + \sum_{s=1}^{t-1} \prod_{j=1, j \neq s}^{t-1} a_{jj} |a_{ts}|^2 \\ &\leq \prod_{i=1}^t a_{ii} + \eta^{t-2} \sum_{s=1}^{t-1} |a_{ts}|^2. \end{aligned} \quad (46)$$

But equality in (45) implies that equality holds in (46) and since $a_{tk} \neq 0$,

$$\prod_{j=1, j \neq k}^{t-1} a_{jj} = \eta^{t-2}$$

and as before

$$a_{11} = \dots = a_{k-1, k-1} = a_{k+1, k+1} = \dots = a_{t-1, t-1} = \eta.$$

Hence A_t has only a_{tk} and \bar{a}_{tk} as non-zero off diagonal elements and thus A_t and A may be brought into the form (10) by a simultaneous permutation of rows and columns. Virtually the same argument gives the case of equality in (41). Thus (9) is proved for the permanent, including the cases of equality.

The formula analogous to (42) for the determinant is

$$a_{m+1, m+1} \det(A_m) - \det(A_{m+1}) = (K_m u_m, u_m) \tag{47}$$

where K_m denotes the adjugate of A_m . Thus, if the eigenvalues of A_m are $\gamma_1 \geq \dots \geq \gamma_m$ and the eigenvalues of A are $\lambda_1 \geq \dots \geq \lambda_n$ then for any unit vector x ,

$$\begin{aligned} \lambda_n^{m-1} \leq \lambda_{n-m+2} \dots \lambda_n \leq \gamma_2 \dots \gamma_m \leq (K_m x, x) \leq \\ \gamma_1 \dots \gamma_{m-1} \leq \lambda_1 \dots \lambda_{m-1} \leq \lambda_1^{m-1}. \end{aligned} \tag{48}$$

Thus to prove the upper inequality in (9) for the determinant, we obtain (43) from (48), and from (47) we obtain (44) with $\prod_{i=1}^n a_{ii} - \det(A)$ replacing $\text{per}(A) - \prod_{i=1}^n a_{ii}$. The proof then proceeds as in (45), and the case of equality is identical also. The lower inequality in (9) is obtained for the determinant using the same proof.

To prove the corollary we observe, for the alternating group A_n and $\chi \equiv 1$,

$$d_\chi^n(A) = \frac{1}{2}(\text{per}(A) + \det(A)).$$

An application of the upper inequalities in (9) proves (11). Equality in (11) would imply A is of the form (10) from Theorem 3, with $\lambda = \eta$. For such an A , (11) becomes

$$0 \leq \eta^{n-2} |b|^2$$

thus equality implies $b = 0$, and A is diagonal.

To prove Theorem 4 we use the notation and techniques of [3, Theorem 2]. Thus on the space of n -tuples, V_n define $T: V_n \rightarrow V_n$ by $Tx = A^T x$, and let v_1, \dots, v_n be an orthonormal basis of eigenvectors of T corresponding respectively to $\lambda_1, \dots, \lambda_n$. Then if e_1, \dots, e_n is the standard basis in V_n we compute that

$$\begin{aligned} d_\chi^G((Te_{\omega_i}, e_{\omega_j})) &= d_\chi((A^T e_{\omega_i}, e_{\omega_j})) \\ &= d_\chi(A[\omega \mid \omega]). \end{aligned}$$

But according to equation (11) in [3] we have

$$\begin{aligned} d_\chi^G((Te_{\omega_i}, e_{\omega_j})) &= \nu(\omega) \left(K(T) \sqrt{\frac{h}{\nu(\omega)}} e_{\omega^*}, \sqrt{\frac{h}{\nu(\omega)}} e_{\omega^*} \right) \\ &= \nu(\omega) \sum_{\gamma \in \bar{\mathcal{A}}} c_{\omega, \gamma} \prod_{t=1}^n \lambda_t^{m_t(\gamma)}, \end{aligned} \tag{49}$$

in which

$$c_{\omega, \gamma} = \left| \left(\sqrt{\frac{h}{\nu(\omega)}} e_{\omega^*}, \sqrt{\frac{h}{\nu(\gamma)}} v_\gamma^* \right) \right|^2.$$

Equations (2) and (3) in [3] tell us that

$$\sum_{\gamma \in \bar{\mathcal{A}}} c_{\omega, \gamma} = 1, \quad \sum_{\omega \in \bar{\mathcal{A}}} c_{\omega, \gamma} = 1$$

for each ω and γ in $\bar{\mathcal{A}}$ respectively and moreover

$$\sum_{i=1}^n m_i(\omega) s_{it} = \sum_{\gamma \in \bar{\mathcal{A}}} m_t(\gamma) c_{\omega, \gamma} \tag{50}$$

for any $\omega \in \bar{\mathcal{A}}$ and each $t, 1 \leq t \leq n$, where $s_{it} = |(e_i, v_t)|^2, i, t = 1, \dots, n$. Set $d_\omega = d_\chi^G([\omega \mid \omega])/\nu(\omega)$ and use the concavity of the log function to obtain from (49) and (50),

$$\begin{aligned} \log d_\omega &\geq \sum_{\gamma \in \bar{\mathcal{A}}} c_{\omega, \gamma} \sum_{t=1}^n m_t(\gamma) \log \lambda_t \\ &= \sum_{t=1}^n \log \lambda_t \sum_{\gamma \in \bar{\mathcal{A}}} c_{\omega, \gamma} m_t(\gamma) \\ &= \sum_{t=1}^n \log \lambda_t \sum_{i=1}^n m_i(\omega) s_{it}. \end{aligned} \tag{51}$$

Then since the scalars b_{ω} are non-negative we have from (51) that

$$\sum_{\omega \in \mathcal{A}} b_{\omega} \log d_{\omega} \geq \sum_{t=1}^n \sum_{i=1}^n \sum_{\omega \in \mathcal{A}} b_{\omega} m_i(\omega) s_{it} \log \lambda_t. \quad (52)$$

The expression on the right in (52) is linear in the doubly stochastic matrix $S = (s_{it})$ and hence by Birkhoff's theorem [5] assumes its minimum value when S is a permutation matrix. Thus in the notation of Theorem 4

$$\begin{aligned} \sum_{\omega \in \mathcal{A}} b_{\omega} \log d_{\omega} &\geq \sum_{i=1}^n \sum_{\sigma \in \mathcal{A}} b_{\omega} m_i(\omega) \log \lambda_{\sigma(i)} \\ &= \sum_{i=1}^n q_i \log \lambda_{\sigma(i)} \end{aligned} \quad (53)$$

where σ is the permutation minimizing the expression in (52). Taking exponentials in (53) produces the result (12).

To prove Theorem 5 we need only modify the above argument slightly. Suppose that we look at (51) in the case of the determinant function. Then for any strictly increasing sequence ω of length not exceeding n (51) becomes

$$\log \det(A[\omega \mid \omega]) \geq \sum_{t=1}^n \log \lambda_t \sum_{i=1}^n m_i(\omega) s_{it}. \quad (54)$$

In (54) replace ω by $\omega^{(j)}$, $j = 1, \dots, r$ (see the statement of Theorem 5), and add the resulting inequalities. This then gives

$$\log \prod_{j=1}^r \det(A[\omega^{(j)} \mid \omega^{(j)}]) \geq \sum_{t=1}^n \log \lambda_t \sum_{i=1}^n \sum_{j=1}^r m_i(\omega^{(j)}) s_{it}. \quad (55)$$

Now $\sum_{j=1}^r m_i(\omega^{(j)})$ is just the total number of times the integer i occurs among $\omega^{(1)}, \dots, \omega^{(r)}$; we have called the s distinct integers that appear among these sequences $\alpha_1, \dots, \alpha_s$ and their multiplicities of occurrence are respectively p_1, \dots, p_s , $p_1 \geq \dots \geq p_s \geq 1$. Now let

$$M_i = \sum_{j=1}^r m_i(\omega^{(j)}), \quad i = 1, \dots, n,$$

(some M_i may be zero) so that minimizing (55) as before we have for some $\sigma \in S_n$,

$$\log \prod_{j=1}^r \det(A[\omega^{(j)} \mid \omega^{(j)}]) \geq \sum_{i=1}^n M_i \log \lambda_{\sigma(i)}$$

and

$$\prod_{j=1}^r \det(A[\omega^{(j)} \mid \omega^{(j)}]) \geq \prod_{i=1}^n \lambda_{\sigma(i)}^{M_i} \quad (56)$$

$$\geq \prod_{j=1}^s \lambda_{n-j+1}^{p_j}.$$

The last inequality in (56) is the well-known result that states that the thing to do to make a product of powers small is to put the largest exponent on the least factor, the second largest exponent on the next smallest factor, etc.

The inequalities (15) and (16) are proved in very much the same way from (51).

REFERENCES

1. E. H. LIEB, *Proofs of Some Conjectures on Permanents*, *J. Math. Mech.* **16** (1966), 127–134.
2. M. MARCUS, The Hadamard Theorem for Permanents, *Proc. Amer. Math. Soc.* **15** (1964), 967–973.
3. M. MARCUS, Matrix Applications of a Quadratic Identity for Decomposable Symmetrized Tensors, *Bull. Amer. Math. Soc.* **71** (1965), 360–364.
4. M. MARCUS AND H. MINC, Permanents, *Amer. Math. Monthly* **72** (1965), 577–591.
5. M. MARCUS AND H. MINC, A Survey of Matrix Theory and Matrix Inequalities, Allyn and Bacon, Boston, 1964.
6. H. MINC, Upper Bounds for Permanents of (0,1)-Matrices, *Bull. Amer. Math. Soc.* **69** (1963), 789–791.
7. H. J. RYSER, Compound and Induced Matrices in Combinatorial Analysis, *Amer. Math. Soc., Proc. Symp. Appl. Math.* **10** (1960), 149–167.
8. H. J. RYSER AND W. B. JURKAT, Matrix Factorizations of Determinants and Permanents, *J. Algebra* **3** (1966), 1–27.
9. I. SCHUR, Über endliche Gruppen und Hermitesche Formen, *Math. Z.* **1** (1918), 184–207.