ON A COMBINATORIAL RESULT OF R. A. BRUALDI AND M. NEWMAN

MARVIN MARCUS[1] AND STEPHEN PIERCE

1. Introduction. Let H be a subgroup of S_n and let A be an n-square matrix over a field F. Following Schur (7) we define the generalized matrix function $d_H(A)$ by

$$d_H(A) = \sum_{\sigma \in H} \prod_{t=1}^{n} a_{t\sigma(t)}.$$

For example, if $H = S_n$, then $d_H(A)$ is the permanent function, per(A); if H is the identity group, then $d_H(A)$ is the product of the main diagonal elements of A, etc. For $1 \leqq r \leqq n$ define H_r to be those elements of H which leave $r+1, \ldots, n$ individually fixed, and let H'_r be the subgroup of S_r obtained by restricting the permutations in H_r to $\{1, \ldots, r\}$. Our main result is contained in the following theorem.

THEOREM. *Let A be an n-square row substochastic matrix, i.e., a non-negative matrix in which every row sum is at most* 1. *Then*

$$(1) \qquad d_H(A) \leqq \max_{\sigma, \tau \in H} d_{H'_r}(A[\sigma(1), \ldots, \sigma(r) \mid \tau(1), \ldots, \tau(r)]),$$

where $A[\sigma(1), \ldots, \sigma(r) \mid \tau(1), \ldots, \tau(r)]$ is the r-square matrix in which the i, j entry is $a_{\sigma(i), \tau(j)}$, $i, j = 1, \ldots, r$.

To discuss cases of equality in (1), we place a mild restriction on H which henceforth will be referred to as condition (M): for every pair, i, j, $1 \leqq i, j \leqq n$, there exists $\sigma \in H$ such that $i, j \in [\sigma(r+1), \ldots, \sigma(n)]$. We remark that condition (M) is certainly true for any doubly transitive group but is substantially weaker than double transitivity. Clearly, if $r = n-1$, H cannot satisfy (M), and thus we state the following necessary and sufficient conditions for equality to hold in (1).

If $r \leqq n-2$, H satisfies (M), *and $d_H(A) \neq 0$, then equality holds in* (1) *if and only if A is a permutation matrix corresponding to some ϕ in H. If d_H is the permanent function, then $H = S_n$ and* (M) *is satisfied by H. Moreover, in this case, $H'_r = S_r$ and hence* (1) *becomes*

$$\operatorname{per}(A) \leqq \max_{\sigma, \tau \in H} \operatorname{per}(A[\sigma(1), \ldots, \sigma(r) \mid \tau(1), \ldots, \tau(r)]);$$

Received November 2, 1966.

[1]The research of this author was supported by the Air Force Office of Scientific Research under grants AFOSR 698-65 and 698-67.

if $r \leqq n - 2$, *equality holds if and only if* A *is a permutation matrix.* This inequality was recently proved by M. Newman and R. Brualdi (**1**). The result is interesting in that it relates a generalized function of the matrix A to induced generalized functions of submatrices of A.

It is clear that the above theorem will apply to any matrix with non-negative entries and positive row sums by suitable normalization. We state this result in Theorem 4 below.

2. Theorems and proofs. Before proceeding we introduce some notation. Let G be any subgroup of S_r, and for $1 \leqq r \leqq n$ let $\Gamma_{r,n}$ denote the set of n^r sequences $\omega = (\omega_1, \ldots, \omega_r)$, $1 \leqq \omega_i \leqq n$, $i = 1, \ldots, r$. If ω and γ are in $\Gamma_{r,n}$, we say that ω is equivalent to γ, $(\omega \sim \gamma)$ (modulo G), if there is some $\sigma \in G$ such that $\omega_i = \gamma_{\sigma(i)}$, $i = 1, \ldots, r$: $\omega = \gamma^\sigma$. Clearly, $\sim$ is an equivalence relation. From each equivalence class, choose the representative that is first in lexicographic order. Denote the resulting system of distinct representatives (S.D.R.), ordered lexicographically, by $\Delta_{r,n}^G$. We shall sometimes abbreviate $\Delta_{r,n}^G$ to Δ_G. If $\alpha \in \Delta_G$, let $\Gamma(\alpha)$ be the equivalence class in $\Gamma_{r,n}$ to which it belongs. Note that $Q_{r,n} \subset G_{r,n} \subset \Delta_G$, where $Q_{r,n}$ is the set of all strictly increasing sequences in $\Gamma_{r,n}$ and $G_{r,n}$ is the set of all non-decreasing sequences in $\Gamma_{r,n}$. For $\alpha \in \Gamma_{r,n}$, let $m_t(\alpha)$ be the number of times t occurs in α, and for $1 \leqq p \leqq r$, let $\Delta_G{}^p$ be the lexicographically ordered subset of Δ_G consisting of all sequences α satisfying $m_t(\alpha) \leqq p$, $t = 1, \ldots, n$. For α and β in Δ_G, let $K_G(A)$ be the matrix whose α, β entry is

$$\frac{d_G(A^{\mathrm{T}}[\beta|\alpha])}{\nu(\alpha)}, \tag{2}$$

where $\nu(\alpha)$ is the number of σ in G satisfying $\alpha^\sigma = \alpha$. For example, if all the α_i are different, $\nu(\alpha) = 1$; if all α_i are the same, $\nu(\alpha) = g$, the order of G; if $G = S_r$, then

$$\nu(\alpha) = \prod_{t=1}^{n} m_t(\alpha)!.$$

Also observe that if $G = S_r$,

$$(K_G(A))_{\alpha,\beta} = \frac{\mathrm{per}(A^{\mathrm{T}}[\beta|\alpha])}{\gamma(\alpha)},$$

and hence $K_G(A)$ is $P_r(A)$, the rth induced power matrix of A (**6**). A result in (**8**) states that

$$K_G(AB) = K_G(A)K_G(B). \tag{3}$$

Observe that

$$A^{\mathrm{T}}[\beta|\alpha] = (A[\alpha|\beta])^{\mathrm{T}}$$

and thus

$$d_G(A^{\mathrm{T}}[\beta|\alpha]) = d_G((A[\alpha|\beta])^{\mathrm{T}}) = d_G(A[\alpha|\beta]).$$

For $\alpha, \beta \in \Delta_G{}^p$, let $K_G{}^p(A)$ be the matrix whose α, β entry is

(4) $$\frac{d_G(A[\alpha|\beta])}{\nu(\beta)}.$$

Observe that $(K_G{}^r(A))^{\mathrm{T}} = K_G(A^{\mathrm{T}})$.

THEOREM 1. *Let A be an n-square row substochastic matrix. Then for $1 \leqq p \leqq r \leqq n$, $K_G{}^p(A)$ is also row substochastic.*

Proof. Since A is row substochastic, we may write

$$AJ = DJ,$$

where J is the n-square matrix, all of whose entries are 1, and

$$D = \mathrm{diag}(d_{11}, \ldots, d_{nn}),$$

in which d_{ii} is the ith row sum of A. Then from (3)

(5) $$K_G(A)K_G(J) = K_G(D)K_G(J).$$

Computing the α, β entry of both sides of (5), we obtain

(6) $$\sum_{\gamma\in\Delta_G} \frac{d_G(A[\alpha|\gamma])}{\nu(\alpha)} \frac{d_G(J[\gamma|\beta])}{\nu(\gamma)} = \sum_{\gamma\in\Delta_G} \frac{d_G(D[\alpha|\gamma])}{\nu(\alpha)} \frac{d_G(J[\gamma|\beta])}{\nu(\gamma)}.$$

Now $d_G(J[\gamma|\beta]) = g$, the order of G, for any γ and β in Δ_G. Hence, after suitable cancellation, (6) becomes

(7) $$\sum_{\gamma\in\Delta_G} \frac{d_G(A[\alpha|\gamma])}{\nu(\gamma)} = \sum_{\gamma\in\Delta_G} \frac{d_G(D[\alpha|\gamma])}{\nu(\gamma)}.$$

By computation,

(8) $$d_G(D[\alpha|\gamma]) = \sum_{\sigma\in G} \prod_{t=1}^{r} d_{\alpha_t,\gamma_{\sigma(t)}}.$$

Unless $\alpha = \gamma^\sigma$,

$$\prod_{t=1}^{r} d_{\alpha_t,\gamma_{\sigma(t)}} = 0.$$

But α and γ come from an S.D.R., and hence $\alpha = \gamma^\sigma$ implies $\gamma = \alpha$, and $\alpha^\sigma = \alpha$. Thus,

$$d_G(D[\alpha|\gamma]) = 0$$

unless $\gamma = \alpha$, and, from (8),

$$d_G(D[\alpha|\alpha]) = \sum_{\sigma\in G} \prod_{t=1}^{r} d_{\alpha_t,\alpha_{\sigma(t)}}.$$

There are $\nu(\alpha)$ such σ in G for which $\alpha^\sigma = \alpha$, and hence

(9) $$\sum_{\sigma\in G} \prod_{t=1}^{r} d_{\alpha_t,\alpha_{\sigma(t)}} = \nu(\alpha) \prod_{t=1}^{r} d_{\alpha_t,\alpha_t} = \nu(\alpha) \prod_{t=1}^{n} d_{tt}^{m_t(\alpha)}.$$

https://doi.org/10.4153/CJM-1968-102-1 Published online by Cambridge University Press

Using (9) in (7) and the fact that A is row substochastic ($d_{tt} \leqq 1$), we obtain

(10)
$$\sum_{\gamma \in \Delta_G} \frac{d_G(A[\alpha|\gamma])}{\nu(\gamma)} = \prod_{t=1}^{n} d_{tt}^{m_t(\alpha)} \leqq 1.$$

Now all terms on the left of (10) are non-negative and hence, summing over $\gamma \in \Delta_G^p$, we obtain

(11)
$$\sum_{\gamma \in \Delta_G^p} \frac{d_G(A[\alpha|\gamma])}{\nu(\gamma)} \leqq 1$$

for all $\alpha \in \Delta_G^p$. But the left side of (11) is just the αth row sum of $K_G^p(A)$. This proves the theorem.

If d_G is the permanent function, and $p = 1$, then $\Delta_G^1 = Q_{r,n}$ and $\nu(\gamma) = 1$ for all $\gamma \in Q_{r,n}$. Then (11) becomes

(12)
$$\sum_{\gamma \in Q_{r,n}} \operatorname{per}(A[\alpha|\gamma]) \leqq 1$$

for all $\alpha \in Q_{r,n}$. This is the Newman-Brualdi result (**1**).

We next prove the following purely combinatorial result that will enable us to discuss the cases of equality in (1).

THEOREM 2. *Let G be a subgroup of S_r. Let A be an $r \times n$ matrix with no zero rows. If $1 \leqq p \leqq r \leqq n$, then*

(13)
$$d_G(A[1, \ldots, r|\gamma]) = 0$$

for all $\gamma \in \Delta_G$, $\gamma \notin \Delta_G^p$, if and only if every column of A has at most p non-zero entries. (Recall that $\gamma \notin \Delta_G^p$ implies that $m_t(\gamma) > p$ for some t.)

Proof. Suppose no column of A has more than p non-zero entries, and $m_t(\gamma) = k > p$, where $\gamma \in \Delta_G$. Then $A[1, \ldots, r|\gamma]$ clearly has an $(r-p)\times k$ zero submatrix. Since $k > p$, $r - p + k > r$ and hence, by the Frobenius-König Theorem, (13) must hold.

Conversely, suppose the sth column of A, $A^{(s)}$, has k non-zero entries, $k > p$. Choosing $\gamma = (i, i, \ldots, i)$, $i = 1, \ldots, n$, we see from (13) that every column of A has at least one zero entry. Thus, $p < k \leqq r - 1$. Determine $\omega = (\omega_1, \ldots, \omega_k) \in Q_{k,r}$ such that $a_{\omega_1 s} \ldots a_{\omega_k s} \neq 0$. Let $\beta = (\beta_1, \ldots, \beta_{r-k})$ be the sequence complementary to ω in $\{1, \ldots, r\}$. Then $a_{\beta_i s} = 0$, $i = 1, \ldots, r-k$. Choose any sequence $\gamma \in \Gamma_{r,n}$ with the following properties: $m_s(\gamma) = k$; $\gamma_{\omega_1} = \ldots = \gamma_{\omega_k} = s$. We assert that for any such γ

(14)
$$d_G(A[1, \ldots, r|\gamma]) = 0.$$

First, if $\gamma \in \Delta_G$, then γ satisfies (13) and hence (14) holds. If $\gamma \notin \Delta_G$, then $\gamma \in \Gamma(\alpha)$ for some $\alpha \in \Delta_G$, i.e., $\gamma = \alpha^\sigma$ for some $\sigma \in G$. Then, since

$$m_s(\alpha) = m_s(\alpha^\sigma) = k > p,$$

we have

https://doi.org/10.4153/CJM-1968-102-1 Published online by Cambridge University Press

$$d_G(A[1,\ldots,r|\gamma]) = d_G(A[1,\ldots,r|\alpha^\sigma]) = \sum_{\tau\in G}\prod_{i=1}^{r} a_{i\alpha_{\sigma\tau(i)}} =$$
$$\sum_{\mu\in G}\prod_{i=1}^{r} a_{i\alpha_{\mu(i)}} = d_G(A[1,\ldots,r|\alpha]) = 0.$$

The above procedure shows, in fact, that if $\alpha \sim \beta$ and $\gamma \sim \delta$, then

$$d_G(A[\alpha|\gamma]) = d_G(A[\beta|\delta]). \tag{15}$$

Returning to the proof, we expand (14) as follows:

$$0 = d_G(A[1,\ldots,r|\gamma]) = \sum_{\sigma\in G}\prod_{i=1}^{r} a_{i\gamma_{\sigma(i)}} = \sum_{\sigma\in G}\prod_{i=1}^{k} a_{\omega_i\gamma_{\sigma(\omega_i)}}\prod_{i=1}^{r-k} a_{\beta_i\gamma_{\sigma(\beta_i)}}. \tag{16}$$

Let G_{r-k} be the subgroup of G consisting of all σ in G which map $\{\omega_1,\ldots,\omega_k\}$ onto itself. If $\sigma \notin G_{r-k}$, then for some i and j, $1 \leqq i \leqq r-k$, $1 \leqq j \leqq k$, $\sigma(\beta_i) = \omega_j$. In this case the second product in (16) contains the term

$$a_{\beta_i\gamma_{\sigma(\beta_i)}} = a_{\beta_i\gamma_{\omega_j}} = a_{\beta_i s} = 0.$$

Thus, we may assume that (16) is summed over G_{r-k}:

$$\sum_{\sigma\in G_{r-k}}\prod_{i=1}^{k} a_{\omega_i\gamma_{\sigma(\omega_i)}}\prod_{i=1}^{r-k} a_{\beta_i\gamma_{\sigma(\beta_i)}} = 0. \tag{17}$$

Now

$$\prod_{i=1}^{k} a_{\omega_i\gamma_{\sigma(\omega_i)}} = \prod_{i=1}^{k} a_{\omega_i s}$$

is a constant $q \neq 0$ for $\sigma \in G_{r-k}$. Hence, (17) becomes

$$\sum_{\sigma\in G_{r-k}}\prod_{i=1}^{r-k} a_{\beta_i\gamma_{\sigma(\beta_i)}} = 0. \tag{18}$$

Since $\sigma \in G_{r-k}$, σ maps $\{\beta_1,\ldots,\beta_{r-k}\}$ onto itself, and thus $\sigma(\beta_i) = \beta_{\phi(i)}$, $i = 1,\ldots,r-k$, for some $\phi \in S_{r-k}$ (ϕ depends on σ). The set of all ϕ thus determined is clearly a subgroup S of S_{r-k}, and, of course, as σ runs over G_{r-k}, ϕ runs over S. Hence (18) has the form

$$\sum_{\phi\in S}\prod_{i=1}^{r-k} a_{\beta_i\gamma_{\beta_{\phi(i)}}} = d_S(A[\beta_1,\ldots,\beta_{r-k}|\gamma_{\beta_1},\ldots,\gamma_{\beta_{r-k}}]) = 0.$$

It is clear from the choice of γ that γ_{β_i} can be any of $1,\ldots,n$, except s, for $i = 1,\ldots,r-k$. Thus, setting $B = A[\beta_1,\ldots,\beta_{r-k}|\,1,\ldots,s-1,s+1,\ldots,n]$, we have

$$d_S(B[1,\ldots,r-k|\gamma]) = 0 \tag{19}$$

for any $\gamma \in \Gamma_{r-k,n-1}$. We now assert that (19) implies that B has a zero row. To see this, let V be an $(n-1)$-dimensional space over F with basis $e_1,\ldots,e_{n-1}$. Let

$$x_i = \sum_{j=1}^{n-1} b_{ij} e_j, \qquad i = 1, \ldots, r - k.$$

Consider (**2**, p. 322) the symmetric product of $x_1, \ldots, x_{r-k}$ with respect to S, denoted by

$$x_1 * \ldots * x_{r-k}. \tag{20}$$

Using the multilinearity of the symmetric product, we find that (20) has the form

$$x_1 * \ldots * x_{r-k} = \sum_{\omega \in \Delta^S_{r-k,n-1}} \frac{d_S(B[1, \ldots, r-k|\omega])}{\nu(\omega)} e_{\omega_1} * \ldots * e_{\omega_{r-k}}. \tag{21}$$

Clearly, (19) and (21) imply that $x_1 * \ldots * x_{r-k} = 0$. Hence, some $x_i = 0$, and thus the ith row of B, $B_{(i)}$, is zero. But the ith row of B is the β_ith row of $A[\beta_1, \ldots, \beta_{r-k}| 1, \ldots, s-1, s+1, \ldots, n]$. However, $a_{\beta_i s}$ is also 0; hence, the β_ith row of A is 0, a contradiction. This completes the proof.

THEOREM 3. *Suppose $r \neq 1$ and $1 \leqq p \leqq r$. Let A be an n-square row sub-stochastic matrix. Then $K_G^p(A)$ is row stochastic if and only if A is a permutation matrix.*

Proof. If $K_G^p(A)$ is row stochastic, we have from (11) that

$$1 = \sum_{\gamma \in \Delta_G^p} \frac{d_G(A[\alpha|\gamma])}{\nu(\gamma)} \leqq \prod_{t=1}^{n} d_{tt}^{m_t(\alpha)} \leqq 1 \tag{22}$$

for all $\alpha \in \Delta_G^p$. Since $Q_{r,n} \subset \Delta_G^p$, (22) implies that all the d_{tt} are 1. Therefore, $D = I_n$ and $AJ = J$. Hence, A is row stochastic. Referring to (10), we see that if (22) holds, then

$$d_G(A[\alpha|\gamma]) = 0$$

for any $\alpha \in \Delta_G^p$, and any $\gamma \notin \Delta_G^p$, i.e., any $\gamma \in \Delta_G$, with $m_t(\gamma) > p$ for some t. Therefore, for each $\alpha \in \Delta_G^p$, the matrix $A[\alpha|1, \ldots, n]$ satisfies the hypothesis of Theorem 2, and we conclude that $A[\alpha|1, \ldots, n]$ has at most p non-zero entries in each column. We now claim that in fact A has at most one non-zero entry in each column. Suppose $A^{(s)}$ had non-zero entries in rows i and j. Construct a non-decreasing sequence α using i exactly p times and j at least once but not more than p times and such that $m_t(\alpha) \leqq p$ for all t. Since $G_{r,n} \subset \Delta_G$, it is clear that $\alpha \in \Delta_G^p$. Then the sth column of $A[\alpha|1, \ldots, n]$ will have at least $p + 1$ non-zero entries, a contradiction.

Now since A is row stochastic, the sum of its entries is n. Furthermore, since every column has at most one non-zero entry, each column sum can be at most one. But, if the column sums are to add up to n, each column sum must be one. Hence every column of A has exactly one non-zero entry which must be one, i.e., A is a permutation matrix.

Now suppose A is a permutation matrix corresponding to some $\tau \in S_n$. If $\alpha, \beta \in \Delta_G^p$ we compute (δ is the Kronecker delta)

$$(23)\quad d_G(A[\alpha|\beta]) = \sum_{\sigma\in G}\prod_{i=1}^{r} a_{\alpha_i\beta_{\sigma(i)}} = \sum_{\sigma\in G}\prod_{i=1}^{r}\delta_{\alpha_i\tau(\beta_{\sigma(i)})} = \sum_{\sigma\in G}\delta_{\alpha\tau(\beta^\sigma)} = \sum_{\sigma\in G}\delta_{\tau^{-1}(\alpha)\beta^\sigma}.$$

Now τ^{-1} is a one-to-one function; therefore, if $m_t(\alpha) \leqq p$ for all t, the same must be true for the sequence $\tau^{-1}(\alpha)$. Therefore, $\tau^{-1}(\alpha) \in \Gamma(\omega)$ for some ω in $\Delta_G{}^p$, and hence, $\tau^{-1}(\alpha) = \omega^\phi$ for some $\phi \in G$. Then (23) becomes

$$(24)\qquad \sum_{\sigma\in G}\delta_{\omega^\phi\beta^\sigma} = \sum_{\sigma\in G}\delta_{\omega\beta^{\sigma\phi^{-1}}}.$$

Hence, (24) is zero unless $\beta = \omega$ and $\omega^{\sigma\phi^{-1}} = \omega$. Thus,

$$(25)\qquad d_G(A[\alpha|\beta]) = \nu(\tau^{-1}(\alpha))\delta_{(\tau^{-1}(\alpha))^{\phi^{-1}},\beta}.$$

Using (25), we obtain for $\alpha \in \Delta_G{}^p$,

$$\sum_{\gamma\in\Delta_G{}^p}\frac{d_G(A[\alpha|\gamma])}{\nu(\gamma)} = \sum_{\gamma\in\Delta_G{}^p}\frac{\nu(\tau^{-1}(\alpha))\delta_{(\tau^{-1}(\alpha))^{\phi^{-1}},\gamma}}{\nu(\gamma)} = 1,$$

and hence $K_G{}^p(A)$ is row stochastic. In fact, $K_G{}^p(A)$ is also a permutation matrix. Observe that if $r = 1$, then p must be 1 and hence $K_G{}^1(A) = A$. Therefore, for $r = 1$ we can only conclude that A is row stochastic.

Before proving our main result, we state a Laplace expansion theorem for $d_H(A)$ as developed in (**4**). Let H_r be defined as in the introduction. For $r + s = n$ let H_s be those elements of H which leave $1, \ldots, r$ individually fixed. For X an r-square matrix, and Y an s-square matrix, define

$$(26)\qquad d^r(X) = d_{H_r}(X \dotplus I_s)$$

and

$$(27)\qquad d^s(Y) = d_{H_s}(I_r \dotplus Y).$$

Since $H_r \cap H_s$ is the identity, the product $H_r \times H_s$ is direct. Let H be expressed as a union of left cosets of $H_r \times H_s$ and let R be a system of distinct representatives for these cosets. Then the result in (**4**) states that for any $\sigma \in R$

$$(28)\qquad d_H(A) = \sum_{\tau\in R} d^r(A[\sigma(1), \ldots, \sigma(r)|\tau(1), \ldots, \tau(r)]) \times d^s(A[\sigma(r+1), \ldots, \sigma(n)|\tau(r+1), \ldots, \tau(n)]).$$

The summation may also be taken over σ in R with τ in R fixed. Recall that H'_r is the restriction of H_r to $\{1, \ldots, r\}$. If $\sigma \in H_s$, then $\sigma(r + i) = r + \phi(i)$, $i = 1, \ldots, s$, for some $\phi \in S_s$. It is clear that the set of all ϕ thus determined is a subgroup H'_s of S_s. We now assert that

$$(29)\qquad d^r(X) = d_{H'_r}(X)$$

and

https://doi.org/10.4153/CJM-1968-102-1 Published online by Cambridge University Press

$$d^s(Y) = d_{H'_s}(Y). \tag{30}$$

We prove (30); the proof of (29) is essentially the same. From (27),

$$d^s(Y) = d_{H_s}(I_r \dotplus Y) = \sum_{\sigma\in H_s} \prod_{i=1}^{n} (I_r \dotplus Y)_{i\sigma(i)}. \tag{31}$$

If $i \leqq r$, $\sigma(i) = i$ and $(I_r \dotplus Y)_{i\sigma(i)} = 1$, so that (31) becomes

$$\sum_{\sigma\in H_s} \prod_{i=r+1}^{n} (I_r \dotplus Y)_{i,\sigma(i)} = \sum_{\sigma\in H_s} \prod_{j=1}^{s} (I_r \dotplus Y)_{r+j,\sigma(r+j)} =$$

$$\sum_{\phi\in H'_s} \prod_{j=1}^{s} (I_r \dotplus Y)_{r+j,r+\phi(j)} = \sum_{\phi\in H'_s} \prod_{j=1}^{s} Y_{j,\phi(j)} = d_{H'_s}(Y).$$

Now (28) has the form

$$d_H(A) = \sum_{\tau\in R} d_{H'_r}(A[\sigma(1), \ldots, \sigma(r)|\tau(1), \ldots, \tau(r)]) \tag{32}$$

$$\times\, d_{H'_s}(A[\sigma(r+1), \ldots, \sigma(n)|\tau(r+1), \ldots, \tau(n)]).$$

Denote the ordered sets $(1, \ldots, r)$ and $(r+1, \ldots, n)$ by $\mathscr{I}_r$ and $\mathscr{I}'_r$, respectively. We rewrite (32) as

$$d_H(A) = \sum_{\tau\in R} d_{H'_r}(A[\sigma(\mathscr{I}_r)|\tau(\mathscr{I}_r)])d_{H'_s}(A[\sigma(\mathscr{I}'_r)|\tau(\mathscr{I}'_r)]). \tag{33}$$

We next prove that

$$\sum_{\tau\in R} d_{H'_s}(A[\sigma(\mathscr{I}'_r)|\tau(\mathscr{I}'_r)]) \leqq 1 \tag{34}$$

for all $\sigma \in R$. In (11), let $G = H'_s$, and let $p = 1$. Let $\Delta_{H'_s}{}^1 = \Delta_s$. Note that $\Delta_s = \Delta_{H'_s} \cap D_{s,n}$, where $D_{s,n}$ is the set of all $\omega = (\omega_1, \ldots, \omega_s) \in \Gamma_{s,n}$ in which the ω_i are distinct. Then, since $\nu(\gamma) = 1$ for any $\gamma \in D_{s,n}$, (11) becomes

$$\sum_{\gamma\in\Delta_s} d_{H'_s}(A[\alpha|\gamma]) \leqq 1 \tag{35}$$

for all $\alpha \in \Delta_s$. If we knew that $\tau(\mathscr{I}'_r)$ ran through a subset of Δ_s as τ runs through R (hitting no element of Δ_s twice), we could use (35) to obtain (34). However, $\tau(\mathscr{I}'_r)$ does not necessarily have this property. Nevertheless, since τ is a one-to-one function, $\tau(\mathscr{I}'_r) \in D_{s,n}$, and hence $\tau(\mathscr{I}'_r)$ is equivalent to some $\alpha_\tau \in \Delta_s$. Thus, by (15),

$$d_{H'_s}(A[\sigma(\mathscr{I}'_r)|\tau(\mathscr{I}'_r)]) = d_{H'_s}(A[\alpha_\sigma|\alpha_\tau])$$

for suitable $\alpha_\sigma, \alpha_\tau \in \Delta_s$. Therefore, we can use (35) to obtain (34) if we can show that $\tau_1, \tau_2 \in R$, $\tau_1 \neq \tau_2$ implies that $\tau_1(\mathscr{I}'_r)$ and $\tau_2(\mathscr{I}'_r)$ are in different equivalence classes of $D_{s,n}$ as determined by H'_s, i.e., we need to show that the mapping $\tau \to \alpha_\tau$ is one-to-one on R. Suppose $\tau_1(\mathscr{I}'_r) \sim \tau_2(\mathscr{I}'_r)$. Let $\gamma_i = \tau_1(r+i)$ and $\delta_i = \tau_2(r+i)$, $i = 1, \ldots, s$. Then there exists $\phi \in H'_s$ such that $\gamma_i = \delta_{\phi(i)}$, $i = 1, \ldots, s$. Thus,

$$\tau_1(r+i) = \tau_2(r+\phi(i)), \qquad i = 1, \ldots, s.$$

We know that there is some $\sigma \in H_s$ such $r + \phi(i) = \sigma(r+i)$, $i = 1, \ldots, s$, so we have

$$\tau_1(r+i) = \tau_2\sigma(r+i), \qquad i = 1, \ldots, s.$$

Thus,

$$\sigma^{-1}\tau_2^{-1}\tau_1(r+i) = r+i, \qquad i = 1, \ldots, s.$$

Hence, $\sigma^{-1}\tau_2^{-1}\tau_1 \in H_r$, i.e., $\sigma^{-1}\tau_2^{-1}\tau_1 = \theta$ for some $\theta \in H_r$. Therefore, $\tau_2^{-1}\tau_1 = \sigma\theta \in H_r \times H_s$. Therefore, $\tau_2 = \tau_1$, a contradiction. This establishes (34). Letting

$$d_{H'_s}(A[\sigma(\mathscr{I}'_r)|\tau(\mathscr{I}'_r)]) = c_{\sigma\tau},$$

we write $\sum_{\tau\in R} c_{\sigma\tau} \leqq 1$ for all $\sigma \in R$. We now have

$$(36) \qquad \begin{aligned} d_H(A) &= \sum_{\tau\in R} c_{\sigma\tau} d_{H'_r}(A[\sigma(\mathscr{I}_r)|\tau(\mathscr{I}_r)]) \\ &= \left(\sum_{\tau\in R} c_{\sigma\tau}\right)\left[\sum_{\tau\in R} \frac{c_{\sigma\tau}}{\sum_{\mu\in R} c_{\sigma\mu}} d_{H'_r}(A[\sigma(\mathscr{I}_r)|\tau(\mathscr{I}_r)])\right] \\ &\leqq \sum_{\tau\in R} c_{\sigma\tau} \max_{\tau\in R} d_{H'_r}(A[\sigma(\mathscr{I}_r)|\tau(\mathscr{I}_r)]) \end{aligned}$$

for all $\sigma \in R$. Thus, (36) is at most

$$(37) \qquad \max_{\sigma,\tau\in R} d_{H'_r}(A[\sigma(\mathscr{I}_r)|\tau(\mathscr{I}_r)]) \leqq \max_{\sigma,\tau\in H} d_{H'_r}(A[\sigma(\mathscr{I}_r)|\tau(\mathscr{I}_r)]).$$

This proves (1). Suppose equality holds in (1), $r \neq n-1$, $d_H(A) \neq 0$, and H satisfies (M). We claim that R also satisfies (M). To see this, let $1 \leqq i, j \leqq n$. We know that there exists some $\theta \in H$ such that $i, j \in \theta(\mathscr{I}'_r)$. The permutation θ is in some left coset of $H_r \times H_s$; let ϕ be the element of R representing this coset. Then $\phi^{-1}\theta \in H_r \times H_s$ and setwise we have

$$(38) \qquad \phi^{-1}\theta(\mathscr{I}'_r) = \mathscr{I}'_r.$$

Since ϕ is one-to-one, we may apply ϕ to both sides of (38) to obtain

$$\phi(\mathscr{I}'_r) = \theta(\mathscr{I}'_r)$$

and hence, $i, j \in \phi(\mathscr{I}'_r)$. Now, referring to (36), we see that if equality holds in (1), we must have

$$(39) \qquad \textstyle\sum_{\tau\in R} c_{\sigma\tau} = 1$$

for all $\sigma \in R$. Hence, from (35) and (39) we have

$$1 = \sum_{\gamma\in\Delta_s} d_{H'_s}(A[\alpha_\sigma|\gamma]) = \prod_{t=1}^{n} d_{tt}^{m_t(\alpha^\sigma)} = 1.$$

https://doi.org/10.4153/CJM-1968-102-1 Published online by Cambridge University Press

Since R satisfies (M), every t between 1 and n is in some α_σ. Hence, all d_{tt} are 1, and A is row stochastic. Moreover, from (10) and (39) we have

$$d_{H's}(A[\sigma(\mathscr{I}'_r)|\gamma]) = 0$$

whenever, $\gamma \in \Delta_{H's}$ and $\gamma \notin \Delta_s$, i.e., whenever $m_t(\gamma) \geqq 2$ for some t. Therefore, for all $\sigma \in R$, the matrix $A[\sigma(\mathscr{I}'_r)|1, \ldots, n]$ satisfies the hypothesis of Theorem 2 for the case $p = 1$. Hence, $A[\sigma(\mathscr{I}'_r)| 1, \ldots, n]$ has at most one non-zero entry in each column. This implies that A has at most one non-zero entry in each column. For, suppose $A^{(t)}$ had non-zero entries in rows i and j. Pick $\sigma \in R$ so that $i, j \in \sigma(\mathscr{I}'_r)$. Then the tth column of $A[\sigma(\mathscr{I}'_r)|1, \ldots, n]$ would have at least two non-zero entries, a contradiction. Since A is row stochastic, we conclude, as in Theorem 3, that A is a permutation matrix and, since $d_H(A) \neq 0$, A must be a permutation matrix corresponding to some $\phi \in H$. Conversely, suppose A is a permutation matrix corresponding to some $\phi \in H$. Then $d_H(A) = 1$. For $\sigma, \tau \in H$, let

$$X_{\sigma\tau} = A[\sigma(\mathscr{I}_r)|\ \tau(\mathscr{I}_r)].$$

Clearly, $d_{H'_r}(X_{\sigma\tau})$ must be 0 or 1 since $X_{\sigma\tau}$ will have at most one non-zero entry in each row and column. We shall find $\sigma, \tau \in H$ such that $d_{H'_r}(X_{\sigma\tau}) = 1$. Let σ be the identity and let $\tau = \phi$. We know that $a_{ij} = \delta_{i\phi(j)}$, so

$$(40) \qquad d_{H'_r}(A[\sigma(\mathscr{I}_r)|\tau(\mathscr{I}_r)]) = d_{H'_r}(A[1, \ldots, r|\phi(1), \ldots, \phi(r)]) = \sum_{\theta \in H'_r} \prod_{i=1}^{r} a_{i\phi\theta(i)} = \sum_{\theta \in H'_r} \prod_{i=1}^{r} \delta_{i\theta(i)} = 1.$$

3. Counterexamples. In dealing with the question of equality in (1), we had to exclude the case $r = n - 1$. We give two examples to show that in this case the result fails. In both examples, d_H will be the permanent function. Thus, $H = S_n$ and $H'_r = H'_{n-1} = S_{n-1}$. Hence, $d_{H'_{n-1}}$ is also the permanent function. First, let $A = J_n$, the matrix all of whose entries are n^{-1}. Clearly,

$$\operatorname{per}(A[\sigma(\mathscr{I}_r)|\tau(\mathscr{I}_r)]) = \frac{1}{n^{n-1}} \cdot (n-1)! = \frac{n!}{n^n} = \operatorname{per}(A)$$

for any $\sigma, \tau \in H$. Thus equality holds in (1), but A is certainly not a permutation matrix. Observe that H satisfies (M) and $d_H(A) \neq 0$. Next, consider the matrix

$$A = \begin{Vmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{Vmatrix}.$$

Clearly, $\operatorname{per}(A) = 1/8$. It is easy to check that any 3-square matrix of the form $A[\sigma(1), \sigma(2), \sigma(3)|\ \tau(1), \tau(2), \tau(3)]$, for any $\sigma, \tau \in S_4$, can have at most one non-zero diagonal. Thus we have equality in (1) and again A is not a permutation matrix.

We now give an example to show that (M) cannot in general be omitted. Consider the matrix

$$A = \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix}.$$

Let $H = [e, (12)]$ be a subgroup of S_4. Let $r = 2$. We see that H'_r is S_2. Moreover, $d_H(A) = \frac{1}{2}$. Clearly, for any choice of $\sigma, \tau \in H$, we obtain

$$d_{H'_r}(A[\sigma(1), \sigma(2) | \tau(1), \tau(2)]) = d_{H'_r} \begin{Vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{Vmatrix} = \frac{1}{2}.$$

Thus equality holds in (1) and again A is not a permutation matrix.

We now obtain a generalization of (1) for arbitrary non-negative matrices.

THEOREM 4. *Let A be an n-square non-negative matrix with positive row sums $d_1, \ldots, d_n$. Let H be a subgroup of S_n and let H'_r be defined as before. Then*

$$d_H(A) \leqq \max_{\sigma, \tau \in H} d_{\sigma(r+1)} \ldots d_{\sigma(n)} d_{H'_r}(A[\sigma(\mathscr{I}_r) | \tau(\mathscr{I}_r)]). \tag{41}$$

If $r \leqq n - 2$, $d_H(A) \neq 0$, and H satisfies (M), *then equality holds in* (41) *if and only if A is a generalized permutation matrix corresponding to some $\phi \in H$.* (*A generalized permutation matrix is a matrix which has exactly one non-zero entry in each row and column.*)

Proof. Let B be an n-square matrix defined as follows:

$$B_{(i)} = (d_i)^{-1} A_{(i)}. \tag{42}$$

Then B is row stochastic and hence (1) applies to B. Therefore,

$$d_H(B) \leqq \max_{\sigma, \tau \in H} d_{H'_r}(B[\sigma(\mathscr{I}_r) | \tau(\mathscr{I}_r)]). \tag{43}$$

Clearly,

$$d_H(B) = (d_1 \ldots d_n)^{-1} d_H(A) \tag{44}$$

and we compute

$$\begin{aligned} d_{H'_r}(B[\sigma(\mathscr{I}_r) | \tau(\mathscr{I}_r)]) &= \sum_{\theta \in H'_r} \prod_{t=1}^{r} b_{\sigma(t), \tau\theta(t)} \\ &= \sum_{\theta \in H'_r} \prod_{t=1}^{r} \left(\frac{a_{\sigma(t), \tau\theta(t)}}{d_{\sigma(t)}} \right) \\ &= (d_{\sigma(1)} \ldots d_{\sigma(r)})^{-1} d_{H'_r}(A[\sigma(\mathscr{I}_r) | \tau(\mathscr{I}_r)]). \end{aligned} \tag{45}$$

Using (44) and (45) in (43), we obtain (41).

https://doi.org/10.4153/CJM-1968-102-1 Published online by Cambridge University Press

Now suppose equality holds in (41). Clearly, if the inequality were strict in (43), it would also be strict in (41). Therefore equality must hold in (43). Hence, B is a permutation matrix corresponding to some $\phi \in H$. Using (42), we see that A must be a generalized permutation matrix corresponding to the same ϕ. The non-zero entry in row i of A is, of course, d_i.

References

1. R. A. Brualdi and M. Newman, *Inequalities for the permanental minors of non-negative matrices*, Can. J. Math. *18* (1966), 608–615.

2. M. Marcus and H. Minc, *Generalized matrix functions*, Trans. Amer. Math. Soc. *116* (1965), 316–329.

3. ——— *A survey of matrix theory and matrix inequalities* (Allyn and Bacon, Boston, 1964).

4. M. Marcus and G. W. Soules, *Inequalities for combinatorial matrix functions*, J. Combinatorial Theory *2* (1967), 145–163.

5. H. J. Ryser, *Combinatorial mathematics* (Carus Math. Monograph No. 14, 1963).

6. ——— *Compound and induced matrices in combinatorial analysis*, Proc. Sympos. Appl. Math., Vol. 10: Combinatorial Analysis (Amer. Math. Soc., Providence, R.I., 1960).

7. I. Schur, *Über endliche Gruppen und Hermitesche Formen*, Math. Z. *1* (1918), 184–207.

8. J. H. M. Wedderburn, *Lectures on matrices* (Amer. Math. Soc. Colloq. Publ., Vol. 17, Providence, R.I., 1934).

University of California,
Santa Barbara, California