THE PERMANENT FUNCTION

by

FRANK COLIN MAY

B.A., University of British Columbia, 1959

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

Master of Arts

in the Department of

Mathematics

We accept this thesis as conforming to the

required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April, 1961

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Mathematics

The University of British Columbia,
Vancouver 8, Canada.

Date April 12, 1961.

ABSTRACT

Let X be a square matrix of order k over a field F. The permanent of X is given by

$$\mathrm{per}(X) = \sum_{\sigma} (x_{1\sigma(1)} x_{2\sigma(2)} \cdots x_{k\sigma(k)})$$

where σ ranges over all the permutations of 1,2,...,k. The original object of this investigation was to characterize those linear maps which leave the permanent unaltered ; that is, per(X) = per(T(X)), all X.

Let $M_{m,n}$ denote the vector space of all matrices having m rows and n columns with entries taken from F. Fix an integer r, $2 \leq r \leq \min(m,n)$. The r-th permanental compound of $X \in M_{m,n}$ is defined in an analogous way to the r-th compound of X, and is denoted by $P_r(X) \in M_{\binom{m}{r},\binom{n}{r}}$.

Subject to mild restrictions on F, the following theorem can be proved. Let T be a linear map on $M_{m,n}$ into itself, let S_r be a non-singular linear map on $M_{\binom{m}{r},\binom{n}{r}}$ onto itself. Suppose that $P_r(T(X)) = S_r(P_r(X))$, all $X \in M_{m,n}$. Then for $\max(m,n) > 2$, we have T(X) = DPXQK when $m \neq n$; when m = n , we have either T(X) = DPXQK, allX, or T(X) = DPX'QK, all X. Here P,Q are permutation matrices and D,K are diagonal matrices, of appropriate orders. For the case r = m = n = 2 , there is a certain non-singular linear map B on $M_{2,2}$ onto itself such that BTB(X) = UXV, all X, or BTB(X) = UX'V, all X. Here U,V are non-singular.

The original problem arises in the case r = m = n , with $S_r = 1$, the unit of F.

I hereby certify that this abstract is satisfactory.

. . . .

. . . .

TABLE OF CONTENTS

ACKNOWLEDGEMENT

The author wishes to express his appreciation for the very generous assistance given him by Dr. Marvin Marcus in the preparation of this thesis. He is also pleased to acknowledge the financial support given him by the National Research Council of Canada.

INTRODUCTION

Let X be an n-square matrix with elements in a field F. The permanent of X is defined by

$$(1)\qquad \operatorname{per}(X) = \sum_{\sigma} x_{1\sigma(1)}\, x_{2\sigma(2)} \cdots x_{n\sigma(n)}$$

where σ runs over the symmetric group of permutations on the integers $1,2,\ldots,n$. This function makes its appearance in certain combinatorial applications [13], and is involved in a conjecture of van der Waerden [6], [9]. Certain formal properties of per (X) are known [1], and an old paper of Pólya [12] shows that for $n > 2$ one cannot multiply the elements of X by constants in any uniform way so as to convert the permanent into the determinant. Indeed, it can be shown that no linear operation on X (for $n > 2$) will transform the permanent into the determinant.

The purpose of this thesis is to characterize those linear operations on matrices which leave the permanent unaltered. This problem and its generalizations have been considered for the determinant function by Frobenius [3] and Kantor [5], later by Schur [14], Morita [11], Dieudonné [2], Marcus and Moyls [8], Marcus and Purves [10], Marcus and May [7]. In view of the result of Pólya [12], it does not seem likely that many of the techniques used in the above papers can be used to investigate the permanent function. Most of these rely heavily on certain properties of the determinant function which are no longer valid for the permanent function. For example, it is in general false that per (AB) = per (A) per (B).

DEFINITIONS AND NOTATION.

Let $M_{m,n}$ denote the vector space of all $m \times n$ matrices over a field F, with the natural basis of unit matrices E_{ij}, where E_{ij} is the matrix with 1 in position (i,j) and 0 elsewhere. In the sequel, r will denote a fixed integer satisfying $2 \leq r \leq \min(m,n)$. When dealing with index sets, the following notation will be used. $Q_{n,r}$ denotes the totality of strictly increasing sequences of integers satisfying $1 \leq i_1 < i_2 < \ldots < i_r \leq n$. As usual, $\alpha = (i_1,\ldots,i_r)$ precedes $\beta = (j_1,\ldots,j_r)$ in the lexicographic ordering of $Q_{n,r}$, $\alpha < \beta$, if there is t such that $i_t < j_t$ and $i_s \leq j_s$, all $s < t$.

Let $X \in M_{m,n}$. We define the r-th permanental compound of X, denoted by $P_r(X) \in M_{\binom{m}{r},\binom{n}{r}}$ as follows : if $\omega = (i_1,\ldots,i_r) \in Q_{m,r}$ and $\delta = (j_1,\ldots,j_r) \in Q_{n,r}$, then the (ω,δ) entry (in the doubly lexicographic ordering) of $P_r(X)$ is $X_{\omega\delta}$, where $X_{\omega\delta}$ is the permanent of the matrix in $M_{r,r}$ whose (s,t) entry is $x_{i_s j_t}$, $(s,t = 1,\ldots,r)$. We denote the (ω,δ) unit matrix in $M_{\binom{m}{r},\binom{n}{r}}$ by $E_{\omega\delta}$.

Let $x_\alpha = (x_{\alpha 1},\ldots,x_{\alpha n})$, $(\alpha = 1,\ldots,r)$, be any vectors over F. Then the permanental product of the vectors x_α , $(\alpha = 1,\ldots,r)$, denoted by $x_1 \vee x_2 \vee \cdot \cdot \cdot \vee x_r$, is defined to be the $\binom{n}{r}$- vector whose $\delta = (j_1,\ldots,j_r) \in Q_{n,r}$ coordinate is $\operatorname{per}(x_{\alpha j_\beta})$, $(\alpha = 1,\ldots,r \,;\, \beta = 1,\ldots,r)$, in the lexicographic ordering.

We denote the rank of X by $\varrho(X)$, the transpose of X by X', the i^{th} row of X by $X_{(i)}$, the j^{th} column of X by $X^{(j)}$, and the determinant of X by det (X). Let A be an s x t matrix. If $j > 0$, $k > 0$, we define

$$A \dotplus 0_{j,k} = \begin{vmatrix} A & 0_{s,k} \\ 0_{j,t} & 0_{j,k} \end{vmatrix} ,$$

where $0_{j,k}$ denotes the j x k zero matrix. If $j = k = 0$, we let the matrix $A \dotplus 0_{j,k}$ be A. If $j = 0$, $k > 0$, or $j > 0$, $k = 0$, then we let $A \dotplus 0_{j,k}$ be

$$\begin{vmatrix} A & 0_{s,k} \end{vmatrix} \quad \text{or} \quad \begin{vmatrix} A \\ 0_{j,t} \end{vmatrix} \quad \text{respectively.}$$

If $u = (u_1,\ldots,u_n)$ and $v = (v_1,\ldots,v_n)$ are n-vectors, the symbols $u \perp v$ and $u // v$ will indicate respectively that $\sum_i u_i v_i = 0$ and that u and v are linearly dependent. If $C \in M_{m,n}$ and $X \in M_{m,n}$, we define the Hadamard product of C and X to be the matrix $Y = C * X \in M_{m,n}$ given by $y_{ij} = c_{ij}x_{ij}$, $(i = 1,\ldots,m \ ; \ j = 1,\ldots,n)$.

Next, let T be a linear map of $M_{m,n}$ into itself. If T is non-singular, the inverse of T is denoted by T^{-1}. Let P and Q be permutation matrices in $M_{m,m}$ and $M_{n,n}$ respectively. In the sequel, we shall have occasion to use maps H obtained from T as follows :

$$H(X) = P\,T(X)\,Q \ , \ \text{all } X \in M_{m,n}.$$

We shall say that such a map H is the same as T to within permutation.

In the case $m = n = 2$ we shall need the special map B defined on $M_{2,2}$ as follows :

$$(2) \qquad \begin{aligned} B(E_{ij}) &= E_{ij} \text{ if } i \leq j \\ B(E_{21}) &= - E_{21} \end{aligned}$$

Clearly B is non-singular, $B = B^{-1}$, and

$$\text{per } (B(X)) = \det (X)$$

for all $X \in M_{2,2}$.

RESULTS.

Our main results are contained in the

THEOREM. Let T be a linear map of $M_{m,n}$ into itself, and let r be an integer satisfying $2 \leq r \leq \min(m,n)$. Suppose that the ground field F contains at least r elements, and is not of characteristic 2. Assume that there exists a non-singular linear map S_r of $M_{\binom{m}{r},\binom{n}{r}}$ into itself such that

(3) $$P_r(T(X)) = S_r(P_r(X))$$

for all $X \varepsilon M_{m,n}$.

Then, for $m + n > 4$, there are permutation matrices $P \varepsilon M_{m,m}$, $Q \varepsilon M_{n,n}$ and non-singular diagonal matrices $D \varepsilon M_{m,m}$, $K \varepsilon M_{n,n}$ such that if $m \neq n$,

(4) $$T(X) = D\ P\ X\ Q\ K$$

for all $X \varepsilon M_{m,n}$; if $m = n$ (>2), T has the form (4) or

(5) $$T(X) = D\ P\ X'\ Q\ K$$

for all $X \varepsilon M_{m,n}$.

For $m = n = 2$, there are non-singular matrices U and V in $M_{2,2}$ such that

(6) $$[BTB]\ (X) = U\ X\ V$$

for all $X \varepsilon M_{2,2}$, or else

(7) $$[BTB]\ (X) = U\ X'\ V$$

for all $X \varepsilon M_{2,2}$.

We note here that in case $r = m = n > 2$ and $S_n = 1$, then this result tells us that the only linear operations which hold the permanent fixed, i.e.,

(8) $$\text{per}\,(T(X)) = \text{per}\,(X)\ ,\ \text{all } X \in M_{n,n}\ ,$$

must be obtainable, to within taking the transpose, by pre- and post-multiplication of X by diagonal matrices whose product has permanent 1 together with pre- and post-multiplication of X by permutation matrices.

We shall prove the theorem in a sequence of lemmas, some of which may be of interest in themselves.

Lemma 1. Let $X \in M_{m,n}$, let $Q \in M_{m,m}$ be a permutation matrix, and let $D \in M_{m,m}$ be a diagonal matrix. Then

(a) $P_r\,(QX) = P_r\,(Q)\ P_r\,(X)$

(b) $P_r\,(DX) = P_r\,(D)\ P_r\,(X)$

(c) $P_r\,(X') = P_r'(X)$

where $P_r'(X)$ denotes the transpose of $P_r\,(X)$.

Proof : First note that if $x_\mu = (x_{\mu 1},\ldots,x_{\mu n})$, $(\mu = 1,\ldots,r)$, are any n-vectors, then

$$x_1 \vee \cdots \vee x_r = x_{\lambda(1)} \vee \cdots \vee x_{\lambda(r)}$$

for any permutation λ on $1,\ldots,r$. In particular, if $\omega = (i_1,\ldots,i_r) \in Q_{m,r}$ then

$$X_{(i_1)} \vee \cdots \vee X_{(i_r)} = X_{(\lambda(i_1))} \vee \cdots \vee X_{(\lambda(i_r))}$$

for any permutation λ on $i_1,\ldots,i_r$. This is an immediate consequence of the fact that the permanent of a matrix is unaltered by a row (or column) permutation. Let σ be the permutation corresponding to Q. The rows of QX are $X_{(\sigma(1))},\ldots,X_{(\sigma(m))}$. Let e_k denote the unit vector (of appropriate length) with 1 in position k, and 0 elsewhere. Now row ω of $P_r(Q)$ is $e_{\sigma(i_1)} \vee \ldots \vee e_{\sigma(i_r)}$. Let $i_{\alpha_1},\ldots,i_{\alpha_r}$ be the rearrangement of $i_1,\ldots,i_r$ such that $\sigma(i_{\alpha_1}) < \sigma(i_{\alpha_2}) < \ldots < \sigma(i_{\alpha_r})$. Then $e_{\sigma(i_1)} \vee \ldots \vee e_{\sigma(i_r)}$ $= e_{\sigma(i_{\alpha_1})} \vee \ldots \vee e_{\sigma(i_{\alpha_r})}$ is the unit $\binom{m}{r}$-vector with 1 in position $(\sigma(i_{\alpha_1}),\ldots,\sigma(i_{\alpha_r})) \in Q_{m,r}$ and 0 elsewhere. Thus row ω of the product $P_r(Q)\ P_r(X)$ is $X_{(\sigma(i_{\alpha_1}))} \vee \ldots \vee X_{(\sigma(i_{\alpha_r}))}$ $= X_{(\sigma(i_1))} \vee \ldots \vee X_{(\sigma(i_r))}$, which is clearly row ω of $P_r(QX)$. Thus (a) is established.

Let $\delta = (j_1,\ldots,j_r) \in Q_{n,r}$. Then row δ of $P_r(X')$ is $X^{(j_1)} \vee \ldots \vee X^{(j_r)}$. On the other hand, row δ of $P_r'(X)$ is column δ of $P_r(X)$ which is certainly $X^{(j_1)} \vee \ldots \vee X^{(j_r)}$. This proves (c).

Let d_k be the diagonal element in row k of D. Let $\omega = (i_1,\ldots,i_r) \in Q_{m,r}$. Now $P_r(D)$ is again a diagonal matrix whose diagonal element in row ω is $d_{i_1}d_{i_2}\ldots d_{i_r}$. Part (b) follows at once from the fact that the permanent function is linear in each row (and column). In particular, $(d_{i_1}\ldots d_{i_r})\ X_{(i_1)} \vee \ldots \vee X_{(i_r)} = d_{i_1}X_{(i_1)} \vee \ldots \vee d_{i_r}X_{(i_r)}$, which is row ω of $P_r(DX)$. The lemma is proved.

Corollary. Let $X \in M_{m,n}$, let $Q \in M_{n,n}$ be a permutation matrix, and let $D \in M_{n,n}$ be a diagonal matrix. Then

(a') $\quad P_r(XQ) = P_r(X)\, P_r(Q)$

(b') $\quad P_r(XD) = P_r(X)\, P_r(D)$

Proof : An identical computation proves both (a') and (b'). We prove (a').

$$P_r(XQ) = P_r{}'(Q'X') = (P_r(Q')\, P_r(X'))' = P_r(X)\, P_r(Q)$$

Lemma 2. T is non-singular.

Proof : Suppose that $T(U) = 0$. Then for any $X \in M_{m,n}$, we have, using (3),

$$S_r(P_r(U + X)) = P_r(T(U + X)) = P_r(T(U) + T(X))$$
$$= P_r(T(X)) = S_r(P_r(X)).$$

Since S_r is non-singular,

(9) $\quad P_r(U + X) = P_r(X)$

holds for all $X \in M_{m,n}$. For any permutation matrices P and Q of appropriate sizes, Lemma 1 and its corollary tell us that

$$P_r(PUQ + PXQ) = P_r(P(U + X)Q)$$
$$= P_r(P)\, P_r(U + X)\, P_r(Q) = P_r(P)\, P_r(X)\, P_r(Q)$$
$$= P_r(PXQ).$$

Now as X runs over $M_{m,n}$ so does PXQ. It suffices then to show that (9) implies $u_{11} = 0$.

Choose $X \in M_{m,n}$ such that

$x_{11} = 0$

$x_{kk} = t - u_{kk}$, $2 \leq k \leq r$

$x_{ij} = - u_{ij}$, $i \neq j$ and $1 \leq i,j \leq r$

$x_{ij} = 0$, otherwise .

Then the (1,1) entry of $P_r(U + X)$ is $u_{11}t^{r-1}$. On the other hand, the (1,1) entry of $P_r(X)$ is a polynomial in t of degree at most $r-2$. Since F contains at least r elements, we conclude that $u_{11} = 0$.

Lemma 3. Let s be an integer satisfying $1 \leq s \leq \min(m,n)$. Then there is a basis for $M_{\binom{m}{s},\binom{n}{s}}$ of the form $P_s(X)$, with $X \in M_{m,n}$.

Proof : Let $\omega = (i_1,\ldots,i_s) \in Q_{m,s}$ and let $\delta = (j_1,\ldots,j_s) \in Q_{n,s}$. If $X \in M_{m,n}$ is the matrix with $x_{i_t j_t} = 1$, $(t = 1,\ldots,s)$, and $x_{ij} = 0$ otherwise, then $P_s(X) = E_{\omega\delta}$.

Lemma 4. There exists a non-singular linear map S_2 of $M_{\binom{m}{2},\binom{n}{2}}$ into itself such that

(10) $\quad P_2(T(X)) = S_2(P_2(X))$

for all $X \in M_{m,n}$. That is, if (3) holds for $r > 2$, it holds for $r = 2$ as well.

Proof : Let $Y = T(X)$. Using (3) we can write

(11) $\quad Y_{\omega\delta} = \sum_{\alpha,\beta} s^{\alpha,\beta}_{\omega,\delta} X_{\alpha,\beta}$

for any $\omega \in Q_{m,r}$ and $\delta \in Q_{n,r}$; here we sum over all $\alpha \in Q_{m,r}$, $\beta \in Q_{n,r}$. In (11) the scalars $s^{\alpha,\beta}_{\omega,\delta}$ are the entries in the matrix representation of S_r with respect to the natural basis in $M_{\binom{m}{r},\binom{n}{r}}$, ordered doubly lexicographically. Since T is non-singular, we may write

$$x_{st} = \sum_{p=1,q=1}^{m\ ,\ n} g^{p,q}_{s,t}\, y_{pq}$$

where the scalars $g^{p,q}_{s,t}$ are the entries in the matrix representation of T^{-1} with respect to the natural basis in $M_{m,n}$. Now (11) may be regarded as a polynomial identity in the variables y_{pq}.

We compute that

$$\frac{\partial Y_{\omega\delta}}{\partial y_{pq}} = \sum_{\alpha,\beta} s^{\alpha,\beta}_{\omega,\delta} \frac{\partial X_{\alpha\beta}}{\partial y_{pq}}$$

$$= \sum_{\alpha,\beta} s^{\alpha,\beta}_{\omega,\delta} \sum_{u=1,v=1}^{m\ ,\ n} \frac{\partial x_{uv}}{\partial y_{pq}} \frac{\partial X_{\alpha\beta}}{\partial x_{uv}}$$

$$= \sum_{\alpha,\beta} \sum_{u=1,v=1}^{m\ ,\ n} \left(s^{\alpha,\beta}_{\omega,\delta}\, g^{p,q}_{u,v} \right) \frac{\partial X_{\alpha\beta}}{\partial x_{uv}} ,$$

where we take $p \in \omega$ and $q \in \delta$. Now $s^{\alpha,\beta}_{\omega,\delta}\, g^{p,q}_{u,v}$, the coefficient of $\dfrac{\partial X_{\alpha\beta}}{\partial x_{uv}}$ in the last expression of this equation, is a scalar independent of X and Y .

We conclude that any (r-1)-order permanental minor of $Y = T(X)$ is expressible as a fixed linear combination of the (r-1)-order permanental minors of X. In other words, there is a linear map R_o of $M_{\binom{m}{r-1},\binom{n}{r-1}}$ into itself such that

(12) $$P_{r-1}(T(X)) = R_o(P_{r-1}(X))$$

for all $X \varepsilon M_{m,n}$.

Since T is non-singular, we see from (3) that

(13) $$P_r(T^{-1}(X)) = S_r^{-1}(P_r(X))$$

for all $X \varepsilon M_{m,n}$. If we apply the above reasoning to (13), we conclude that there is a linear map R^o of $M_{\binom{m}{r-1},\binom{n}{r-1}}$ into itself such that for all $X \varepsilon M_{m,n}$,

$$P_{r-1}(T^{-1}(X)) = R^o(P_{r-1}(X)) .$$

That is, for all $X \varepsilon M_{m,n}$, we have

(14) $$P_{r-1}(X) = R^o(P_{r-1}(T(X))) .$$

Combining (12) and (14) we have

(15) $$P_{r-1}(X) = R^oR_o(P_{r-1}(X))$$

for all $X \varepsilon M_{m,n}$. Lemma 3, with $s = r-1$, tells us that R^oR_o is the identity map of $M_{\binom{m}{r-1},\binom{n}{r-1}}$ onto itself. Consequently R_o is non-singular in (12), and we set $S_{r-1} = R_o$. Then, using (12), we proceed to reduce r-1 to r-2, etc., finally obtaining (10) .

Let $A \varepsilon M_{m,n}$. If A has at most one non-zero row (column), we shall call A a row (column) matrix. If A is a row (column) matrix, then the number of non-zero entries in A is denoted by $\varphi(A)$.

Lemma 5. Let $A \varepsilon M_{m,n}$, and suppose that $P_2(A) = 0$. Then $\varrho(A) = 0$, 1, or 2. Moreover, if A has rank 1, then A is a row (or column) matrix ; if A has rank 2, then to within permutation of the rows and columns of A, A has the form

$$(16) \qquad \begin{vmatrix} \alpha & \beta \\ \lambda & \mu \end{vmatrix} \dot{+} 0_{m-2,n-2} ,$$

where $\alpha\mu + \beta\lambda = 0$, and $\alpha\mu - \beta\lambda \neq 0$.

Proof : Assume that $A \neq 0$. Suppose first that $\varrho(A) = 1$. We may write row t of A as some multiple of a fixed vector $z = (z_1,\dots,z_n)$, say $c_t z$, $(t = 1,\dots,m)$. Since $P_2(A) = 0$, we see that $2c_t c_s z_i z_j = 0$ if $t \neq s$ and $i \neq j$. Since F is not of characteristic 2, we have $c_t c_s z_i z_j = 0$ if $t \neq s$ and $i \neq j$. Since $A \neq 0$, some $c_{t_o} \neq 0$, and some $z_{i_o} \neq 0$. If there is $j \neq i_o$ for which $z_j \neq 0$, then $c_s = 0$ whenever $s \neq t_o$.

Suppose next that $\varrho(A) > 1$. By a suitable permutation we may bring A to the form

$$A = \begin{bmatrix} \alpha & \beta & a_1 & a_2 & \cdot & \cdot & \cdot & a_{n-2} \\ \lambda & \mu & b_1 & b_2 & \cdot & \cdot & \cdot & b_{n-2} \\ c_1 & d_1 & & & & & & \\ c_2 & d_2 & & & & & & \\ \cdot & \cdot & & & & & & \\ \cdot & \cdot & & & H & & & \\ \cdot & \cdot & & & & & & \\ c_{m-2} & d_{m-2} & & & & & & \end{bmatrix}$$

where $H \in M_{m-2,n-2}$, $\alpha\mu \neq 0$, and $\alpha\mu - \beta\lambda \neq 0$. We have

$$\left.\begin{aligned} \alpha b_t + \lambda a_t &= 0 \\ \beta b_t + \mu a_t &= 0 \end{aligned}\right\}, \quad (t = 1,\ldots,n-2)$$

$$\left.\begin{aligned} \alpha d_s + \beta c_s &= 0 \\ \lambda d_s + \mu c_s &= 0 \end{aligned}\right\}, \quad (s = 1,\ldots,m-2) \quad .$$

But $\alpha\mu - \beta\lambda \neq 0$. Hence $c_s = d_s = 0$, $(s = 1,\ldots,m-2)$, and $a_t = b_t = 0$, $(t = 1,\ldots,n-2)$. Therefore $\mu h_{ij} = 0$ for each element h_{ij} of H. Since $\mu \neq 0$, we have $H = 0$. Also, we note that $\alpha\beta\lambda\mu \neq 0$. This proves Lemma 5.

Corollary. Let $F_{ij} = T(E_{ij})$. Then $\varrho(F_{ij}) = 1$ or 2.

Proof : From (10) we see that

$$P_2(F_{ij}) = S_2(P_2(E_{ij})) = S_2(0) = 0.$$

Lemma 5, together with its corollary, enables us to describe partially the structure of the images F_{ij} of the unit matrices E_{ij} . We now make the additional assumption that $m + n \geq 5$. In this case we are able to obtain the exact structure of the F_{ij} .

Lemma 6. $\varrho(F_{ij}) = 1$.

Proof : By Lemma 2, $F_{ij} \neq 0$. Suppose that $\varrho(F_{ij}) = 2$. We lose no generality in assuming that $i = j = 1$ and that F_{11} has the form (16). Consider F_{1t}, $(2 \leq t \leq n)$. From $P_2(E_{11} + \sigma E_{1t}) = 0$, all $\sigma \in F$, we have using (10), $P_2(F_{11} + \sigma F_{1t}) = 0$, all $\sigma \in F$. Since $\alpha\beta\lambda\mu \neq 0$ in (16), we see at once that $\varphi(F_{1t}) = 2$ if $\varrho(F_{1t}) = 1$; moreover, F_{1t} would be zero outside of positions (1,1), (1,2), (2,1), and (2,2). If $\varrho(F_{1t}) = 2$, then by letting σ vary over F, we see again that F_{1t} is zero outside these same positions. A similar argument leads to the same conclusions concerning F_{s1}, $(2 \leq s \leq m)$.

We next show that F_{1t}, $(t = 1,\ldots,n)$ and F_{s1}, $(s = 1,\ldots,m)$, all lie in the space spanned by the following three matrices :

$$G_1 = \begin{vmatrix} \alpha & \beta \\ 0 & 0 \end{vmatrix} \dotplus 0_{m-2,n-2} ,$$

$$G_2 = \begin{vmatrix} 0 & \beta \\ 0 & \mu \end{vmatrix} \dotplus 0_{m-2,n-2} , \text{ and}$$

$$G_3 = \begin{vmatrix} \alpha & 0 \\ \lambda & 0 \end{vmatrix} \dotplus 0_{m-2,n-2} .$$

Observe that

$$G_4 = \begin{vmatrix} 0 & 0 \\ \lambda & \mu \end{vmatrix} \dotplus 0_{m-2,n-2} = G_2 + G_3 - G_1 ,$$

and that

$$F_{11} = G_1 + G_4 = G_2 + G_3 .$$

First let us assume that $\varrho(F_{1t}) = 1$. We may further assume without loss of generality that $b_{11}b_{21} \neq 0$ and $b_{12} = b_{22} = 0$, where we have set

$$F_{1t} = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} \dotplus 0_{m-2,n-2} .$$

Now $P_2(F_{11} + F_{1t}) = 0$ implies that $b_{11}\mu + b_{21}\beta = 0$, and hence F_{1t} is a multiple of G_3. For $(b_{11},b_{21}) \perp (\mu,\beta)$ $\perp (\alpha,\lambda)$, and so $(b_{11},b_{21})//(\alpha,\lambda)$.

Next assume that $\varrho(F_{1t}) = 2$. We have

$$b_{11}b_{12}b_{21}b_{22} \neq 0$$

and $P_2(F_{11} + F_{1t}) = 0$ shows that

$$b_{11}\mu + b_{22}\alpha + b_{12}\lambda + b_{21}\beta = 0 .$$

Now

$$\alpha\mu + \beta\lambda = b_{11}b_{22} + b_{12}b_{21} = 0 .$$

So there are non-zero constants c and d such that $\lambda = c\alpha$, $\mu = -c\beta$, $b_{21} = db_{11}$, and $b_{22} = -db_{12}$.

Consequently we have

$$0 = b_{11}\mu + b_{22}\alpha + b_{12}\lambda + b_{21}\beta = (c - d)(\alpha b_{12} - \beta b_{11}).$$

Thus either $c = d$ or $\alpha b_{12} = \beta b_{11}$. If $c = d$ we see that

$(b_{12},b_{22}) \perp (\lambda,\alpha) \perp (\beta,\mu)$ and $(b_{11},b_{21}) \perp (\mu,\beta) \perp (\alpha,\lambda)$, whence $(b_{12},b_{22})//(\beta,\mu)$ and $(b_{11},b_{21})//(\alpha,\lambda)$. Therefore if $c = d$, F_{1t} is a linear combination of G_2 and G_3. One shows similarly that in case $\alpha b_{12} = \beta b_{11}$, F_{1t} is a linear combination of G_1 and G_4. Thus the matrices F_{1t}, $(t = 1,\ldots,n)$, and, similarly, the matrices F_{s1}, $(s = 1,\ldots,m)$, all lie in the space of dimension 3 spanned by G_1, G_2, and G_3. But $m + n - 1 > 3$. We have contradicted Lemma 2. Hence $\varrho(F_{ij}) = 1$.

Lemmas 5 and 6 tell us that each F_{ij} is either a row or column matrix.

Lemma 7. $\varphi(F_{ij}) = 1$.

Proof : We lose no generality in assuming that $i = j = 1$ and that F_{11} is a row matrix with its non-zero row in row 1. By a suitable permutation of columns we may assume that row 1 of F_{11} has the form

$$(a_1,a_2,\ldots,a_\varphi,0,\ldots,0) \quad ,$$

where we have set $\varphi = \varphi(F_{11})$ for the sake of brevity. Note that $a_t \neq 0$, $(t = 1,\ldots,\varphi)$.

If $\varphi \geq 3$, then Lemma 5 tells us at once that F_{1t}, $(t = 1,\ldots,n)$, and F_{s1}, $(s = 1,\ldots,m)$, would all be row matrices each with its non-zero row in row 1. For $P_2(F_{11} + F_{1t}) = P_2(F_{11} + F_{s1}) = 0$. Since $m + n - 1 > n$, we have contradicted Lemma 2.

Suppose then that $\varphi = 2$. We have

$$F_{11} = |a_1 \quad a_2| \dot{+} 0_{m-1,n-2}$$

with $a_1a_2 \neq 0$. We first show that F_{12} is a row matrix with its non-zero row in row 1. If not, then by permuting the last $m - 1$ rows of F_{12}, we can take F_{12} in the form

$$\begin{vmatrix} 0 & 0 \\ b_1 & b_2 \end{vmatrix} \dot{+} 0_{m-2,n-2}, \tag{17}$$

where $b_1b_2 \neq 0$ and $a_1b_2 + a_2b_1 = 0$. We next remark that

$$P_2(T^2(X)) = S_2(P_2(T(X))) = S_2^2(P_2(X)) \tag{18}$$

for all $X \varepsilon M_{m,n}$. Consequently all our results concerning the nature of T apply equally well to T^2. In particular, $T^2(E_{11})$ is either a row matrix or a column matrix. But

$$T^2(E_{11}) = T(a_1E_{11} + a_2E_{12}) = a_1F_{11} + a_2F_{12}$$

$$= \begin{vmatrix} a_1^2 & a_1a_2 \\ a_2b_1 & a_2b_2 \end{vmatrix} \dot{+} 0_{m-2,n-2}.$$

However, $a_1a_2b_1b_2 \neq 0$. This contradiction shows that F_{12} is a row matrix lying in row 1.

Consider F_{1t}, $(t > 2)$. If F_{1t} is not a row matrix lying in row 1, then we may assume that F_{1t} has the form (17). Again, $(a_1,a_2) \angle (b_2,b_1)$. From $P_2(F_{12} + F_{1t}) = 0$ we see immediately that F_{12} has the form

$$F_{12} = |c_1 \quad c_2| \dot{+} 0_{m-1,n-2},$$

with $c_1c_2 \neq 0$. So $(c_1,c_2) \not\perp (b_2,b_1)$. But this implies that F_{12} is a multiple of F_{11} and we contradict Lemma 2.

Now again by Lemma 2, F_{21} cannot lie entirely in row 1. Since $P_2(F_{11} + F_{21}) = 0$, we may assume that F_{21} has the form (17). By an argument exactly analogous to that given above, we see that each of F_{2t}, $(t = 1,\ldots,n)$, is a row matrix lying in row 2.

There are two cases to consider :

(i) $m = 2$, $n \geq 3$,

(ii) $m \geq 3$.

In case (i) there is $j_o > 1$ such that F_{1j_o} has a non-zero entry in column 3. Now from $P_2(F_{1j_o} + F_{2j_o}) = 0$, we see that the non-zero entries of F_{2j_o} lie in precisely the same columns as do those of F_{1j_o}. Moreover, we have $\varphi(F_{1j_o}) = \varphi(F_{2j_o}) = 1$, or 2. Now $P_2(E_{11} + E_{21} + \sigma E_{1j_o} - \sigma E_{2j_o}) = 0$, all $\sigma \in F$. Consequently $P_2(F_{11} + F_{21} + \sigma F_{1j_o} - \sigma F_{2j_o}) = 0$, all $\sigma \in F$. But this contradicts Lemma 5. In case (ii), note that $P_2(E_{11} + E_{21} + \sigma E_{31}) = 0$, all σ, and so we must have $P_2(F_{11} + F_{21} + \sigma F_{31}) = 0$, all σ. By Lemma 5, this implies that all the non-zero entries of F_{31} are in its first two rows. This contradicts Lemma 2 once again. Thus $\varphi = 1$.

Lemma 7 tells us that for $m + n \geq 5$, we can write $T(E_{ij}) = c_{ij}E_{i'j'}$. By Lemma 2, $c_{ij} \neq 0$, and moreover, $(i,j) \neq (s,t)$ implies that $(i',j') \neq (s',t')$. We set

$$i' = \sigma(i,j) \quad \text{and} \quad j' = \omega(i,j)$$

so that $T(E_{ij}) = c_{ij}E_{\sigma(i,j)\omega(i,j)}$.

Lemma 8. (Let $m + n \geq 5$.) If $m \neq n$, then there are permutation matrices $P \in M_{m,m}$ and $Q \in M_{n,n}$, and a matrix $C = (c_{ij}) \in M_{m,n}$ with each $c_{ij} \neq 0$, such that for all $X \in M_{m,n}$

(19) $\quad T(X) = C * (PXQ)$.

If $m = n$ (>2) , then T has the form (19) or else

(20) $\quad T(X) = C * (PX'Q)$

for all $X \in M_{m,n}$.

Proof : We may assume without loss of generality that $m \leq n$. Now by a suitable permutation of the rows and columns, we may take $\sigma(1,1) = \omega(1,1) = 1$. Then $P_2(E_{11} + E_{22}) \neq 0$ shows that $P_2(F_{11} + F_{22}) \neq 0$, and so $\sigma(2,2) > 1$, $\omega(2,2) > 1$. By a suitable permutation of the last $m - 1$ rows and the last $n - 1$ columns we may take $\sigma(2,2) = \omega(2,2) = 2$. In a similar way, the conditions $P_2(E_{11} + E_{33}) \neq 0$ and $P_2(E_{22} + E_{33}) \neq 0$ show that $\sigma(3,3) > 2$ and $\omega(3,3) > 2$. Proceeding in this way, it is clear that we may assume that $\sigma(k,k) = \omega(k,k) = k$,($k = 1,\ldots,m$).

Fix $\alpha \leq m$, $\beta \leq m$, so that $\alpha \neq \beta$. Now from $P_2(E_{\alpha\alpha} + E_{\alpha\beta}) = 0$ we see that $\sigma(\alpha,\beta) = \alpha$ or $\omega(\alpha,\beta) = \alpha$. Also $P_2(E_{\beta\beta} + E_{\alpha\beta}) = 0$ implies $\sigma(\alpha,\beta) = \beta$ or $\omega(\alpha,\beta) = \beta$. Therefore we must have either

(21) $\qquad \sigma(\alpha,\beta) = \alpha$ and $\omega(\alpha,\beta) = \beta$, $\qquad$ or

(22) $\qquad \sigma(\alpha,\beta) = \beta$ and $\omega(\alpha,\beta) = \alpha$,

for the non - singularity of T shows that we cannot have $\sigma(\alpha,\beta) = \omega(\alpha,\beta)$.

Suppose first that (21) holds. Let $\delta \leq n$, $\delta \neq \alpha$, $\delta \neq \beta$. From $P_2(E_{\alpha\beta} + E_{\alpha\delta}) = 0$ we have $\sigma(\alpha,\delta) = \alpha$ or $\omega(\alpha,\delta) = \beta$. From $P_2(E_{\alpha\alpha} + E_{\alpha\delta}) = 0$ we have $\sigma(\alpha,\delta) = \alpha$ or $\omega(\alpha,\delta) = \alpha$. It follows that $\sigma(\alpha,\delta) = \alpha$. If in addition we have $\delta \leq m$, then $P_2(E_{\alpha\delta} + E_{\delta\delta}) = 0$ shows that $\sigma(\alpha,\delta) = \delta$ or $\omega(\alpha,\delta) = \delta$. Hence $\omega(\alpha,\delta) = \delta$.

Let $k \neq \alpha$ and consider $E_{k\beta}$. From $P_2(E_{\alpha,\beta} + E_{k\beta}) = 0$ we conclude that $\sigma(k,\beta) = \alpha$ or $\omega(k,\beta) = \beta$. But $\sigma(k,\beta) \neq \alpha$ because $\sigma(\alpha,t) = \alpha$, $(t = 1,\ldots,n)$, and T is non - singular. Hence $\omega(k,\beta) = \beta$. Also $P_2(E_{kk} + E_{k\beta}) = 0$ shows that $\sigma(k,\beta) = k$ or $\omega(k,\beta) = k$. Hence $\sigma(k,\beta) = k$, $\omega(k,\beta) = \beta$. If we repeat this argument now with k replacing α in (21) we conclude that

(23) $\qquad \sigma(i,j) = i$, $\omega(i,j) = j$, $\qquad (i = 1,\ldots,m; j = 1,\ldots,m)$.

Moreover, if $j > m$, the non - singularity of T ensures that $\omega(i,j) > m$. Now we already know that $\sigma(i,j) = i$ for such j . Furthermore, $P_2(E_{sj} + E_{tj}) = 0$ shows that $\omega(s,j) = \omega(t,j)$. Thus if (21) holds, T may be reduced to the form (19) by a suitable permutation of the last n - m columns of X .

Suppose next that (22) holds. We shall show that actually m = n and that

(24) $\qquad \sigma(i,j) = j$, $\omega(i,j) = i$, $\qquad (i = 1,\ldots,m; j = 1,\ldots,m)$.

From $P_2(E_{\alpha\beta} + E_{\alpha k}) = 0$ we have $\sigma(\alpha,k) = \beta$ or $\omega(\alpha,k) = \alpha$. Also $P_2(E_{\alpha\alpha} + E_{\alpha k}) = 0$ shows that $\sigma(\alpha,k) = \alpha$ or

$\omega(\alpha,k) = \alpha$. It follows that $\omega(\alpha,k) = \alpha$, $(k = 1,\ldots,n)$, because $\alpha \neq \beta$. Thus $m = n$, for T maps the n - dimensional space spanned by $E_{\alpha t}$, $(t = 1,\ldots,n)$, into the space spanned by $E_{s\alpha}$, $(s = 1,\ldots,m)$, an m - dimensional space, and $m \leq n$.

We conclude also, from $P_2(E_{\alpha k} + E_{kk}) = 0$, that $\sigma(\alpha,k) = k$ or $\omega(\alpha,k) = k$. Since $\omega(\alpha,k) = \alpha$, it follows that $\sigma(\alpha,k) = k$, $(k = 1,\ldots,m)$, which establishes (24). So T has the form (20).

Lemma 9. $\varrho(C) = 1$.
Proof : Let $1 \leq i < s \leq m$, $1 \leq j < t \leq n$. If (19) holds, choose X so that $PXQ = E_{ij} + E_{it} - E_{sj} + E_{st}$; if (20) holds, choose X so that $PX'Q$ has this same form. We can certainly do this because T is non - singular. In either case, $P_2(X) = P_2(PXQ) = 0$, by Lemma 1 together with its corollary. Therefore $0 = P_2(T(X)) = P_2(c_{ij}E_{ij} + c_{it}E_{it} - c_{sj}E_{sj} + c_{st}E_{st})$, and so $c_{ij}c_{st} - c_{it}c_{js} = 0$. Thus each second order subdeterminant of C vanishes. We recall that each $c_{ij} \neq 0$.

Using Lemma 9 we can write $c_{ij} = d_i k_j$, $(i = 1,\ldots,m; j = 1,\ldots,n)$. We set $D = \mathrm{diag}(d_1,\ldots,d_m) \varepsilon M_{m,m}$ and $K = \mathrm{diag}(k_1,\ldots,k_n) \varepsilon M_{n,n}$. Using Lemma 8 we can write (4) for $m \neq n$ and (4) or (5) for $m = n$ (>2) . The proof of the theorem is complete for the case $m + n \geq 5$.

Suppose that $m = n = 2$. Then (3) reduces to the equation

(25) $\quad \mathrm{per}(T(X)) = \alpha. \mathrm{per}(X)$

for all $X \varepsilon M_{2,2}$, where α is some nonzero element of F. Using (2) together with (25) we see that

$$\det[BTB(X)] = \operatorname{per}[B^2TB(X)] = \operatorname{per}[TB(X)]$$
$$= \alpha.\operatorname{per}[B(X)] = \alpha.\det[X]$$

for all $X \varepsilon M_{2,2}$. Thus BTB preserves the rank of each matrix in $M_{2,2}$. Moreover, $(BTB)^{-1} = BT^{-1}B$ exists and has the same property. Consequently we may appeal to a theorem of Jacob [4] to conclude that BTB has the desired form. The proof of the theorem is complete.

We observe that if $m \neq n$, we have

$$P_r(T(X)) = P_r(DPXQK) = P_r(D)P_r(P)P_r(X)P_r(Q)P_r(K)$$
$$= S_r(P_r(X))$$

for all $X \varepsilon M_{m,n}$. It follows from Lemma 3 that

$$S_r(Y) = P_r(D)P_r(P)YP_r(Q)P_r(K) = D_oP_oYQ_oK_o$$

for all $Y \varepsilon M_{\binom{m}{r},\binom{n}{r}}$, and so S_r has the same form as T . Similarly, if $m = n \geqslant 2$, and T has the form (4), then S_r has the above form. Also, if $m = n > 2$, and T has the form (5), then

$$S_r(Y) = D_oP_oY'Q_oK_o$$

for all $Y \varepsilon M_{\binom{m}{r},\binom{n}{r}}$.

In conclusion, we present an example to show that neither (4) nor (5) need hold if $r = m = n = 2$. We put

$$T(E_{11}) = E_{11} + E_{12} + E_{22} - E_{21}$$

$$T(E_{12}) = E_{11} + E_{12}$$

$$T(E_{21}) = E_{12} + E_{22}$$

$$T(E_{22}) = - E_{12}$$

Then for any $X \varepsilon M_{22}$ we have

$$T\left(\begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix}\right) = \begin{vmatrix} (x_{11} + x_{12}) & (x_{11} + x_{12} + x_{21} - x_{22}) \\ (- x_{11}) & (x_{11} + x_{21}) \end{vmatrix}$$

and an easy computation shows that $\text{per}(T(X)) = \text{per}(X)$. Observe that $T(E_{11})$ has rank 2. It is obvious that $T(X)$ cannot be put into either of the forms (4) or (5). However, we can write $BTB(X) = UX'V$ where

$$U = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} \quad \text{and} \quad V = \begin{vmatrix} 1 & 1 \\ 0 & -1 \end{vmatrix} \quad .$$

BIBLIOGRAPHY

1. A.C.Aitken, Determinants and Matrices, (5th edition), Edinburgh, Oliver and Boyd, (1948).

2. J.Dieudonné, Sur une généralisation du groupe orthogonal à quatre variables, Archiv der Math., 1, (1948), 282 - 287.

3. G.Frobenius, Über die Darstellung der endlichen Gruppen durch linearer Substitutionen, Sitzungber. der Berliner Akademie, 994 - 1015, (7).

4. H.G.Jacob, Coherence invariant mappings on Kronecker products, Amer. J. Math., Vol.77, (1955), 177 - 189.

5. S.Kantor, Theorie der Äquivalenz von linearen ∞^{λ}-Scharen bilinearer Formen, Sitzungber. der Munchener Akademie, (1897), 367 - 381, (2).

6. D.König, Theorie der Graphen, New York, Chelsea, (1950), 238.

7. M.Marcus and F.May, On a theorem of I.Schur concerning matrix transformations, accepted for publication, Archiv der Math.

8. M.Marcus and B.N.Moyls, Linear transformations on algebras of matrices, Can. J. Math., Vol.11, (1959), 61 - 66.

9. M.Marcus and M.Newman, Permanents of doubly stochastic matrices, Proc. of Symposia of Applied Math., Vol.10, Amer. Math. Soc., (1960), 169 - 174.

10. M.Marcus and R.Purves, Linear transformations on algebras of matrices;the invariance of the elementary symmetric functions, Can. J. Math., Vol.11, (1959), 383 - 396.

11. K.Morita, Schwarz's lemma in a homogeneous space of higher dimensions, Japanese J. Math., Vol.19, (1944), 45 - 56.

12. G.Pólya, Aufgabe 424, Archiv der Math. und Phys., Vol.20(3), 271.

13. H.J.Ryser, Compound and induced matrices in combinatorial analysis, Proc. of Symposia of Applied Math., Vol.10, Amer. Math. Soc., (1960), 166.

14. I.Schur, Einige Bemerkungen zur Determinantentheorie, Sitzungber. der Preussischen Akademie der Wissemschaften zu Berlin, Vol.25, (1925), 454 - 463.