PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 3, 1971

THE MATRIX EQUATION $AXB = X$

D. J. HARTFIEL

This paper considers the solutions of the matrix equation $AXB = X$ where we specify A and B to be n-square and doubly stochastic. Solutions are found explicitly and do not depend on either the Jordan or Rational canonical forms. We further find all doubly stochastic solutions of this equation, by noting that $J_n = (1/n)$, the n-square doubly stochastic matrix in which each entry is $1/n$, is always a solution and that the doubly stochastic solutions form a compact convex set. We solve the equation by characterizing the vertices of this convex set.

Matrices considered in this paper are real matrice unless otherwise stated. Most of the definitions and notation may be found in [5], although some will be presented below.

If $A_1, A_2, \cdots, A_s$ are square matrices, by $\sum_{k=1}^{\cdot s} A_k$ we mean the direct sum of the A_k's. If $s = 2$ we may write $A_1 \oplus A_2$ for this direct sum. We say that a square matrix A is reducible if there exists a permutation matrix P so that $PAP^t = \begin{pmatrix} X & O \\ Y & Z \end{pmatrix}$ where X and Z are square and P^t denotes the transpose of P. If A is not reducible, then it is said to be irreducible. A square matrix $A = (a_{ij})$ is doubly stochastic if $a_{ij} \geqq 0$ and $\sum_k a_{ik} = \sum_k a_{kj} = 1$ for all i, j. It readily follows that if A is doubly stochastic, then there exists a permutation matrix P such that $PAP^t = \sum_{k=1}^{\cdot s} A_k$ where each A_k is doubly stochastic and irreducible.

The following two celebrated theorems in matrix theory are used in the paper.

BIRKHOFF'S THEOREM. *The set of all n-square doubly stochastic matrices, Ω_n, forms a convex polyhedron with the permutation matrices as vertices* [5, p. 97].

PERRON-FROBENIUS THEOREM. *Let A be an n-square nonnegative irreducible matrix. Then:*

(i) *A has a real positive characteristic root r which is simple. If λ is any characteristic root of A, then $|\lambda| \leqq r$.*

(ii) *If A has h characteristic roots of modulus*

$$r: \lambda_0 = r, \lambda_1, \cdots, \lambda_{h-1}$$

then these are h distinct roots of $\lambda^h - r^h = 0$, h is called the index of imprimitivity of A. If $h = 1$ the matrix is called primitive.

(iii) *If* $\lambda_0, \lambda_1, \cdots, \lambda_{n-1}$ *are all the characteristic roots of* A, *and* $\theta = e^{i(2\pi/n)}$ *then* $\lambda_0\theta, \cdots, \lambda_{n-1}\theta$ *are* $\lambda_0, \cdots, \lambda_{n-1}$ *in some order.*

(iv) *If* $h > 1$, *then there exists a permutation matrix* P *such that*

$$PAP^t = \begin{pmatrix} 0 & A_{12} & 0 & \cdots & 0 & 0 \\ 0 & 0 & A_{23} & \cdots & 0 & 0 \\ \vdots & & & & \vdots & \\ 0 & 0 & 0 & \cdots & 0 & A_{h-1,h} \\ A_{h,1} & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} .$$

where the zero blocks down the main diagonal are square [5, p. 125].

If A is a nonnegative matrix and

$$a_{i_1j_1}, a_{i_1j_2}, a_{i_2j_2}, \cdots \cdot a_{i_{m-1}j_m}, a_{i_mj_m} = a_{i_1j_1}$$

are all positive elements in A, then A is said to have a loop of length m. If $A = (a_{ij})$ is such that all a_{ij} are equal, then we say that A is flat. If A is partitioned into block matrices A_{ij}, i.e., $A = (A_{ij})$, and each A_{ij} is flat, then a block loop is defined similarly.

1. Preliminary results. First we note that if P and Q are permutation matrices then $AXB = X$ if and only if

$$PAP^tPXQQ^tBQ = PXQ .$$

Since A and B can each be put into a direct sum of irreducible matrices by simultaneous row and column permutations we may assume by the Perron-Frobenius Theorem that

$$A = \sum_{\alpha=1}^{s}{}^{\cdot} A_\alpha , \qquad B = \sum_{\beta=1}^{r}{}^{\cdot} B_\beta$$

$$A_\alpha = \begin{pmatrix} 0 & A_1^\alpha & 0 & \cdots & 0 \\ 0 & 0 & A_2^\alpha & \cdots & 0 \\ & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & \cdots & A_{s_\alpha-1}^\alpha \\ A_{s_\alpha}^\alpha & 0 & 0 & \cdots & 0 \end{pmatrix}, \qquad B_\beta = \begin{pmatrix} 0 & B_1^\beta & 0 & \cdots & 0 \\ 0 & 0 & B_2^\beta & \cdots & 0 \\ & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & \cdots & B_{r_\beta-1}^\beta \\ B_{r_\beta}^\beta & 0 & 0 & \cdots & 0 \end{pmatrix}$$

where A_α is irreducible with index of imprimitivity s_α; B_β is irreducible with index of imprimitivity r_β. Further the 0 blocks down the main diagonal on A_α and B_β are all square.

Note that the dimension of each A_k^α $(k = 1, 2, \cdots, s_\alpha)$ is the same for each fixed α. For fixed β the dimensions of the B_k^β $(k = 1, 2, \cdots, r_\beta)$ are also equal. Hence

$$A_{\alpha}^{s_\alpha} = \begin{pmatrix} C_1 & 0 & 0 & \cdots & 0 \\ 0 & C_2 & 0 & \cdots & 0 \\ & & \cdots\cdots & & \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & C_{s_\alpha} \end{pmatrix}, \qquad B_{\beta}^{r_\beta} = \begin{pmatrix} D_1 & 0 & 0 & \cdots & 0 \\ 0 & D_2 & 0 & \cdots & 0 \\ & & \cdots\cdots & & \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & D_{r_\beta} \end{pmatrix}$$

where each C_k $(k = 1, 2, \cdots, s_\alpha)$, D_k $(k = 1, 2, \cdots, r_\beta)$ is a primitive doubly stochastic matrix. Now let p be a sufficiently large integer so that A^p and B^p are direct sums of primitive matrices.

LEMMA 1.1. *If T is a linear operator on a convex set S whose vertices are X_i $(i = 1, 2, \cdots, m)$, then $T(S)$ is a convex set whose vertices are in $\{T(X_i) \mid i = 1, \cdots, m\}$.*

THEOREM 1.2. *The set of doubly stochastic solutions of the matrix equation $A^pXB^p = X$ (p previously defined) is the convex hull of*

$$\left\{\lim_{k\to\infty} (A^p)^k P_l(B^p)^k \mid P_l \text{ is a permutation matrix, } l = 1, 2, \cdots, n!\right\}.$$

Proof. If V is an $m \times m$ primitive doubly stochastic matrix, then $\lim_{k\to\infty} V^k = J_m$, the flat $m \times m$ doubly stochastic matrix.

$$\lim_{k\to\infty} (A^p)^k \quad \text{and} \quad \lim_{k\to\infty} (B^p)^k$$

exist, their limits being direct sums of flat doubly stochastic matrices. Let $L(X) = \lim_{k\to\infty} (A^p)^k X (B^p)^k$. This is a linear operator defined on the set of $n \times n$ matrices.

By Lemma 1.1, $L(\Omega_n)$ is the convex hull of $\{L(P_l) \mid P_l$ is a permutation matrix$\}$ i.e., of $\{\lim_{k\to\infty} (A^p)^k P_l (B^p)^k \mid P_l$ is a permutation matrix$\}$.

Now if $A^pXB^p = X$, $X \in \Omega_n$, then $L(X) = X$ and by Birkhoff's Theorem, X is in the convex hull of the $\{L(P_l) \mid P_l$ is a permutation matrix$\}$. Furthermore, if X is in the convex hull of the $\{L(P_l) \mid P_l$ is a permutation matrix$\}$ i.e., $X = \Sigma\lambda_l L(P_l)$ where $\lambda_l \geqq 0$ and $\Sigma\lambda_l = 1$, then

$$\begin{aligned} X &= \Sigma\lambda_l L(P_l) = \Sigma\lambda_l \lim_{k\to\infty} (A^p)^k P_l (B^p)^k \\ &= A^p\left[\Sigma\lambda_l \lim_{k\to\infty} (A^p)^{k-1} P_l (B^p)^{k-1}\right]B^p = A^pXB^p , \end{aligned}$$

and X is a solution of the matrix equation.

THEOREM 1.3. *$Y \in \Omega_n$ is a solution of $AXB = X$ if and only if $Y = \sum_{k=0}^{p-1} A^k W B^k / p$ where $W \in \Omega_n$ is a solution of $A^pXB^p = X$.*

Proof. If $Y = \sum_{k=0}^{p-1} A^k W B^k / p$, W a solution of $A^pXB^p = X$, then $AYB = Y$.

Further if Y is solution of $AXB = X$ then Y is a solution of $A^pXB^p = X$ and so $Y = \sum_{k=0}^{p-1} A^k YB^k/p$.

Let $M(Z) = \sum_{k=0}^{p-1} A^kZB^k/p$. Then M is a linear operator defined on the set of $n \times n$ matrices.

COROLLARY 1.4. *The vertices of the set of doubly stochastic solutions of $AXB = X$ is a subset of $\{M[L(P_l)] \mid P_l$ is a permutation matrix$\}$.*

Proof. The proof follows from Lemma 1.1, Theorem 1.2, and Theorem 1.3.

COROLLARY 1.5. *If one of A or B is primitive, then the only doubly stochastic solution of the equation $AXB = X$ is J_n.*

Proof. Either $\lim_{k\to\infty} (A^p)^k$ or $\lim_{k\to\infty} (B^p)^k$ is J_n. Thus if X is doubly stochastic, then $L(X) = J_n$.

2. The operator L. Our primary aim here is to investigate the structure of the convex set $L(\Omega_n)$: in particular its vertices.

From § 1 we know for P_l a permutation matrix

$$L(P_l) = \lim_{k\to\infty} (A^p)^k P_l (B^p)^k = \Big({\sum_\gamma}^{\cdot} J_\gamma^A\Big) P_l \Big({\sum_\sigma}^{\cdot} J_\sigma^B\Big)$$

where J_τ^A and J_σ^B are flat doubly stochastic matrices whose dimensions correspond to the dimension of the primitive matrices in the direct sums A^p and B^p respectively.

Suppose $a_\gamma \times a_\gamma$ is the dimension of J_γ^A and $b_\sigma \times b_\sigma$ is the dimension of J_σ^B. Set $({\sum}^{\cdot}_\gamma J_\gamma^A)P_l({\sum}^{\cdot}_\sigma J_\sigma^B) = V_l$. Partition V_l into blocks $V_{\gamma\sigma}$ of dimension $a_\gamma \times b_\sigma$.

LEMMA 2.1. *If $X \in L(\Omega_n)$ is partitioned into block matrices $X_{\gamma\sigma}$ of dimension $a_\gamma \times b_\sigma$, then each $X_{\gamma\sigma}$ is flat.*

THEOREM 2.2. *If $X \in L(\Omega_n)$ is partitioned into block matrices $X_{\gamma\sigma}$ of dimension $a_\gamma \times b_\sigma$, then X is a vertex of $L(\Omega_n)$ if and only if X does not have a block loop.*

Proof. Suppose X has a block loop

$$X_{\gamma_1\sigma_1}, X_{\gamma_1\sigma_2}, X_{\gamma_2\sigma_2}, \cdots, X_{\gamma_m\sigma_m} = X_{\gamma_1\sigma_1} .$$

Add $\varepsilon > 0$ to each element in the $\gamma_1\sigma_1$ block. Subtract $(b_{\sigma_1}/b_{\sigma_2})\varepsilon$ from each element in the $\gamma_1\sigma_2$ block. All the row sums of the matrix are now one. Now add $(a_{\gamma_1}b_{\sigma_1}/a_{\gamma_2}b_{\sigma_2})\varepsilon$ to each element in the $\gamma_2\sigma_2$ block.

All the column sums of the matrix are now one. Now subtract

$$\frac{b_{\sigma_2}a_{\gamma_1}b_{\sigma_1}}{b_{\sigma_3}a_{\gamma_2}b_{\sigma_2}}\varepsilon$$

from each element in the $\gamma_2\sigma_3$ block. All the row sums of the matrix are now one. Continuing in this manner we see that in the $\gamma_m\sigma_m$ block we add $(a_{\gamma_{m-1}}b_{\sigma_{m-1}}\cdots b_{\sigma_1}/a_{\gamma_m}b_{\sigma_m}\cdots b_{\sigma_2})\varepsilon = \varepsilon$. This is exactly what is in the $\gamma_m\sigma_m$ or $\gamma_1\sigma_1$ block. Now all rows and columns sum to one. Call this generated matrix X'. Now considering the same block loop we generate X'' by replacing ε by $-\varepsilon$ in X'. Again all rows and columns sum to one. Now $X = \frac{1}{2}(X' + X'')$, and since X' and $X'' \in L(\Omega_n)$ for ε sufficiently small, X is an interior point.

On the other hand if $X \in L(\Omega_n)$ and interior to it, there are X' and X'' in $L(\Omega_n)$ so that $X = \frac{1}{2}(X' + X'')$. We may pick X' and X'' in $L(\Omega_n)$ so that they have zero blocks in the block position if and only if X does. Now if $X' \neq X''$ then there is a $\gamma_1\sigma_1$ block so that $X'_{\gamma_1\sigma_1} < X''_{\gamma_1\sigma_1}$ where $X'_{\gamma_1\sigma_1}$ is a block in X', $X''_{\gamma_1\sigma_1}$ is a block in X'' and the relation is elementwise. Hence there is a $X'_{\gamma_1\sigma_2} > X''_{\gamma_1\sigma_2}$ and so on. This generates a block loop in X.

COROLLARY 2.3. *X is a vertex of the convex set of doubly stochastic matrices if and only if X does not have a loop.*

Proof. Consider the matrix equation $IXI = X$ and apply the Theorem 2.2.

We are now in a position to find the vertices of $L(\Omega_n)$. Partition each permutation matrix P_l into blocks $P^l_{\gamma\sigma}$ of dimension $a_\gamma \times b_\sigma$. Let $n_{\gamma\sigma}$ be the number of ones in the $\gamma\sigma$ block of P_l. Then

$$\left(\sum_\gamma{}^{\cdot}\, J^A_\gamma\right)P_l\left(\sum_\sigma{}^{\cdot}\, J^B_\sigma\right) = V_l$$

and $V_{\gamma\sigma}$ has all its elements equal to $n_{\gamma\sigma}/a_\gamma b_\sigma$. We may now use Theorem 2.2 on this finite set to establish exact vertices.

EXAMPLE.

$$\begin{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{pmatrix} P_l \begin{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{pmatrix}.$$

Partitioning the matrices P_l we have

$$(1)\qquad L\left(\begin{array}{c|c} 10 & 00 \\ 01 & 00 \\ \hline 00 & 10 \\ 00 & 01 \end{array}\right) = \left(\begin{array}{c|c} \frac{1}{2}\frac{1}{2} & 0 \\ \frac{1}{2}\frac{1}{2} & \\ \hline 0 & \frac{1}{2}\frac{1}{2} \\ & \frac{1}{2}\frac{1}{2} \end{array}\right), \quad \text{a vertex},$$

$$(2)\qquad L\left(\begin{array}{c|c} 00 & 10 \\ 00 & 01 \\ \hline 10 & 00 \\ 01 & 00 \end{array}\right) = \left(\begin{array}{c|c} 0 & \frac{1}{2}\frac{1}{2} \\ & \frac{1}{2}\frac{1}{2} \\ \hline \frac{1}{2}\frac{1}{2} & 0 \\ \frac{1}{2}\frac{1}{2} & \end{array}\right), \quad \text{a vertex}.$$

All vertices are of the form $L(P_l)$ for some permutation P_l. However, $L(P_l)$ is not always a vertex for every l. For example,

$$(3)\qquad L\left(\begin{array}{c|c} 10 & 00 \\ 00 & 10 \\ \hline 01 & 00 \\ 00 & 01 \end{array}\right) = \left(\begin{array}{c|c} \frac{1}{4}\frac{1}{4} & \frac{1}{4}\frac{1}{4} \\ \frac{1}{4}\frac{1}{4} & \frac{1}{4}\frac{1}{4} \\ \hline \frac{1}{4}\frac{1}{4} & \frac{1}{4}\frac{1}{4} \\ \frac{1}{4}\frac{1}{4} & \frac{1}{4}\frac{1}{4} \end{array}\right) \quad \text{an interior point}.$$

We can further note by Theorem 2.2 that 1 and 2 are the only vertices of $L(\Omega_n)$.

3. General solutions of $A^pXB^p = X$. We already know from Theorem 1.2 that for each $W \in \Omega_n$, $L(W)$ is a solution of $A^pXB^p = X$. Actually we have shown that if W is any $n \times n$ matrix then $L(W)$ is a solution of $A^pXB^p = X$. Further if W is a solution of the equation then $L(W) = W$. i.e., $(\sum_\gamma^{\cdot} J_\gamma^A) W (\sum_\sigma^{\cdot} J_\sigma^B) = W$. Partition W into blocks $W_{\gamma\sigma}$ as in § 2. Now $J_\gamma^A W_{\gamma\sigma} J_\sigma^B = W_{\gamma\sigma}$ implies that $W_{\gamma\sigma}$ is flat. Also if each $W_{\gamma\sigma}$ of W is flat, then W is a solution. Hence we know all solutions of the matrix equation $A^pXB^p = X$.

4. Orbits in matrices. Let $C = (c_{ij})$ be a $p \times q$ matrix. Suppose we pick some $c_{i_1j_1}$. Then by the orbit of $c_{i_1j_1}$ we mean the set of positions $(i_1 - k, j_1 + k)$ $[k = 0, 1, \cdots]$ where the row index is modulo p and the column index is modulo q.

EXAMPLE.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \qquad \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$$

The numbers in the positions of the orbit of

(1) 5 are 5, 3, 7
(2) 2 are 2, 9, 4
(3) 1 are 1, 8, 6
(4) a are a, e, c, d, b, f.

Consider the group $Z/p \oplus Z/q$ where Z is the additive group of integers. Note that $K = \{(-k \bmod p, k \bmod q) \mid k \in Z\}$ is a subgroup of $(Z/p \oplus Z/q)$. Hence we can consider orbits as cosets in $(Z/p \oplus Z/q)/K$ by looking at indices. We now see:

1. The number of elements in each orbit is the same.
2. If two orbits intersect, they are the same.
3. If one orbit contains a row index k times then all orbits contain that row index k times. The same property holds for columns.
4. Each row index and column index appear at least once in each orbit.
5. If p and q are relatively prime, then there is only one distinct orbit.

Finally we note that since orbits are defined by indices, we may consider block orbits in partitioned matrices.

5. The operator M.

Our aim here is to investigate the structure of the convex set $M[L(\Omega_n)]$: in particular to find its vertices. Let $X \in L(\Omega_n)$. Partition X into blocks $X_{\gamma\sigma}$ of dimension $a_\gamma \times b_\sigma$, then

$$M(X) = \frac{1}{p}\sum_{k=0}^{p-1} A^k X B^k = \frac{1}{p}\sum_{k=0}^{p-1}\sum_\alpha{}^{\cdot} \begin{pmatrix} 0 & A_1^\alpha & 0 & 0 & \cdots & 0 \\ 0 & 0 & A_2^\alpha & 0 & \cdots & 0 \\ & & \cdots\cdots & & & \\ A_{s_\alpha}^\alpha & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}^k X \sum_\beta{}^{\cdot} \begin{pmatrix} 0 & B_1^\beta & 0 & 0 & \cdots & 0 \\ 0 & 0 & B_2^\beta & 0 & \cdots & 0 \\ & & \cdots\cdots & & & \\ B_{r_\beta}^\beta & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}^k$$

and since the blocks $X_{\gamma\sigma}$ of X are flat we may write

$$M(X) = \frac{1}{p}\sum_{k=0}^{p-1}\sum_\alpha{}^{\cdot} \begin{pmatrix} 0 & J_1^\alpha & 0 & 0 & \cdots & 0 \\ 0 & 0 & J_2^\alpha & 0 & \cdots & 0 \\ & & \cdots\cdots & & & \\ J_{s_\alpha}^\alpha & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}^k X \sum_\beta{}^{\cdot} \begin{pmatrix} 0 & J_1^\beta & 0 & 0 & \cdots & 0 \\ 0 & 0 & J_1^\beta & 0 & \cdots & 0 \\ & & \cdots\cdots & & & \\ J_{r_\beta}^\beta & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}^k$$

where $J_k^\alpha\,(k = 1, 2, \cdots, s_\alpha)$ and $J_k^\beta\,(k = 1, 2, \cdots, r_\beta)$ are flat doubly stochastic matrices whose dimensions are the same as those of A_k^α and B_k^β, respectively. Suppose the irreducible blocks A_α of $\sum^{\cdot}_\alpha A_\alpha$ have dimension $p_\alpha \times p_\alpha$ and the irreducible blocks B_β of $\sum^{\cdot}_\beta B_\beta$ have the dimension $q_\beta \times q_\beta$. Partition X into blocks $X'_{\alpha\beta}$ of dimension $p_\alpha \times q_\beta$. We call these blocks the major blocks of X. Now since X is already partitioned into blocks of dimension $a_\gamma \times b_\sigma$, we see that the major blocks are partitioned into the $X_{\gamma\sigma}$ blocks in the first partitioning. We call each block in the original partition a minor block. Note that inside each major block, all minor blocks are of the same dimension.

Now suppose $X'_{\alpha\beta}$ is a major block of X. Then we see the sequence

$$X'_{\alpha\beta}, \begin{pmatrix} 0 & J_1^\alpha & 0 & \cdots & 0 \\ 0 & 0 & J_2^\alpha & \cdots & 0 \\ & & \cdots\cdots & & \\ J_{s_\alpha}^\alpha & 0 & 0 & \cdots & 0 \end{pmatrix} X'_{\alpha\beta} \begin{pmatrix} 0 & J_1^\beta & 0 & \cdots & 0 \\ 0 & 0 & J_2^\beta & \cdots & 0 \\ & & \cdots\cdots & & \\ J_{r_\beta}^\beta & 0 & 0 & \cdots & 0 \end{pmatrix}, \cdots,$$
$$\begin{pmatrix} 0 & J_1^\alpha & 0 & \cdots & 0 \\ 0 & 0 & J_2^\alpha & \cdots & 0 \\ & & \cdots\cdots & & \\ J_{s_\alpha}^\alpha & 0 & 0 & \cdots & 0 \end{pmatrix}^{p-1} X'_{\alpha\beta} \begin{pmatrix} 0 & J_1^\beta & 0 & \cdots & 0 \\ 0 & 0 & J_2^\beta & \cdots & 0 \\ & & \cdots\cdots & & \\ J_{r_\beta}^\beta & 0 & 0 & \cdots & 0 \end{pmatrix}^{p-1}$$

is such that each minor block in $X'_{\alpha\beta}$ moves through its orbit in $X'_{\alpha\beta}$ at least once.

By the definition of M and the remarks made above we see that $M(X)$, $X \in L(\Omega_n)$, is found as follows. Let X be partitioned into major and minor blocks. Consider the orbit of the minor blocks in each major block. Sum the blocks in each orbit with sufficiently many copies in order that there are p blocks. Then divide the sum by p and replace each block in the orbit by this block. From this we see that $X \in M[L(\Omega_n)]$ if and only if

1. $X \in L(\Omega_n)$.
2. If $X_{\gamma_1\sigma_1}$ and $X_{\gamma_2\sigma_2}$ are in the same major block and in the same orbit in the major block, then they are equal.

We now find necessary and sufficient conditions for X to be a vertex of $M[L(\Omega_n)]$.

DEFINITION. If $X_{\alpha_1\beta_1}, X_{\alpha_1\beta_2}, \cdots, X_{\alpha_m\beta_m} = X_{\alpha_1\beta_1}$ are major blocks of X, $X \in M[L(\Omega_n)]$ and each $X_{\alpha_k\beta_k}$ $(k = 1, 2, \cdots, m)$, $X_{\alpha_k\beta_{k+1}}$ $(k = 1, 2, \cdots, m-1)$ has exactly one positive minor block orbit, then

$$X_{\alpha_1\beta_1}, X_{\alpha_1\beta_2}, \cdots, X_{\alpha_m\beta_m}$$

is an orbital block loop in X.

THEOREM 5.1. *$X \in M[L(\Omega_n)]$ is a vertex if and only if*

1. *there do not exist two different positive minor block orbits in any major block of X, and*
2. *there does not exist an orbital block loop in X.*

Proof. First suppose $X \in M[L(\Omega_n)]$ and X has two positive block orbits in a major block $X_{\alpha\beta}$ of X. Then we add $\varepsilon > 0$ to each element in each block of one of these orbits and subtract ε from each element of each block in the other orbit. Call this matrix X'. To generate the matrix X'' replace ε by $-\varepsilon$ in X'. Now for ε sufficiently small, X' and $X'' \in M[L(\Omega_n)]$. Since $X = \frac{1}{2}(X' + X'')$, X is interior and therefore if X is a vertex it must satisfy 1.

Now suppose $X \in M[L(\Omega_n)]$ satisfies 1 but not 2. This means X

has an orbital block loop, say $X_{\alpha_1\beta_1}, X_{\alpha_1\beta_2}, \cdots, X_{\alpha_m\beta_m} = X_{\alpha_1\beta_1}$. Each of these major blocks has a positive orbit by definition. Flatten each major block; i.e., if $X_{\alpha\beta}$ is a block in the orbital block loop and has s different orbits, divide the element c in the positive orbit by s and replace all elements in the major block by c/s. If we call this matrix X' then $X' \in M[L(\Omega_n)]$. We may now use the scheme of Theorem 2.2 to alternately add and subtract $\varepsilon > 0$ from this major block loop, thereby generating X_1' and $X_2' \in M[L(\Omega_n)]$ and $X' = \frac{1}{2}(X_1' + X_2')$. Now absorb the flat major blocks back into the original orbits, i.e., if $X_{\alpha\beta}$ is a major block in the orbital block loop with s different orbits then replace each element c in each block of the original positive orbit by sc. Put zero blocks in all other orbits in this major block. Doing this to X', X_1', and X_2' we generate X, X_1, and X_2, respectively. Note $X_1, X_2 \in M[L(\Omega_n)]$. Further $X = \frac{1}{2}(X_1 + X_2)$. Hence X is interior.

Finally suppose X satisfies 1 and 2. Suppose that there exist $X_1, X_2 \in M[L(\Omega_n)]$ so that $X = \frac{1}{2}(X_1 + X_2)$. We may suppose X_1 and X_2 have the same zero pattern as X. If $X_1 \neq X_2$ and X_1, X_2 satisfy 1 we can see by an argument similar to Theorem 2.2, that X has an orbital block loop. This contradicts X having property 2. Hence we see that X is a vertex.

Using this theorem and the remarks preceeding this theorem we see that we have characterized the vertices of $M[L(\Omega_n)]$.

EXAMPLE.

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} X \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

There are three orbits for X given in the following diagram.

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

They are the positions occupied by 1, 2 and 3 respectively. Consider the vertices of $L(\Omega_n)$. Using 1 of Theorem 5.1 we see

(a) $$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

has a one in each orbit; hence

$$M\left[\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}\right] = \begin{pmatrix}1/3 & 1/3 & 1/3\\ 1/3 & 1/3 & 1/3\\ 1/3 & 1/3 & 1/3\end{pmatrix}$$

which is interior.

(b)
$$\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}$$

has 3 ones in the same orbit, hence

$$M\left[\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\right] = \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}$$

which is a vertex. The other vertices are

$$\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}.$$

6. General solutions of $AXB = X$. Partition X into the major and minor blocks. Since $AXB = X$ would imply $A^pXB^p = X$ we see that each minor block of X must be flat. If we add the further condition that minor blocks on the same orbit are all equal then we see from § 5 that X is a solution and all solutions are of this form.

7. General remarks. It is interesting to note that in order to obtain solutions of $AXB = X$ it is only necessary to know the block form of A and B, i.e., if A_1 is doubly stochastic and has the same block form as A and B_1 is doubly stochastic and has the same block form as B then $AXB = X$ if and only if $A_1XB_1 = X$.

From § 4, property 5, we see that if A and B are irreducible, where the index of imprimitivity of A and the index of imprimitivity of B are relatively prime, then J_n is the only doubly stochastic solution. The only general solution is flat. This follows since there is only one orbit in X. Each block in the orbit is flat and all blocks in the orbit are equal.

Finally we point out that our result can be extended to a more general setting by considering the following result due to Sinkhorn (7):

THEOREM. *Let D be the set of all $n \times n$ matrices with row and column sums equal to* 1, *M_{n-1} the set of $(n-1) \times (n-1)$ matrices.*

Let $R = 1 \oplus M_{n-1}$. *Then there is a nonsingular matrix* P *so that* $PDP^{-1} = R$.

From this we know that if A_1 and A_2 are $(n-1) \times (n-1)$ matrices then there are nonsingular matrices P and Q so that $P^{-1}(1 \oplus A_2)P$ and $Q(1 \oplus A_1)Q^{-1}$ have row and column sums equal to 1. If $P^{-1}(1 \oplus A_2)P$ and $Q(1 \oplus A_1)Q^{-1}$ are nonnegative and real and hence doubly stochastic, then since

$$A_1XA_2 = X$$

if and only if

$$(1 \oplus A_1)(1 \oplus X)(1 \oplus A_2) = 1 \oplus X$$

if and only if

$$Q(1 \oplus A_1)Q^{-1}Q(1 \oplus X)PP^{-1}(1 \oplus A_2)P = Q(1 \oplus X)P ,$$

we can also find the solutions to the matrix equation

$$A_1XA_2 = X .$$

References

1. R. A. Brualdi, *Convex sets of nonnegative matrices*, to appear.
2. H. O. Foulkes, *Rational solutions of the matrix equation* $XA = BX$, Proceedings of the London Mathematical Society (2), **50** (1948), 196-209.
3. F. R. Gantmacher, *The Theory of Matrices*, Chelsea Publishing Co., New York. 1960.
4. W. B. Jurkat and H. J. Ryser, *Term ranks and permanents of nonnegative matrices*, J. of Algebra, **5**, #3, March 1967.
5. Marvin Marcus and Henryk Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, 1964.
6. W. V. Parker, *The Matrix Equation* $AX = XB$, Duke Math. J. **17** (1950), 43-51.
7. Richard Sinkhorn, *On the Factor Space of the Complex Doubly Stochastic Matrices*. Abstract 62T-243, Notices American Mathematical Society, 9, (1962), 334-335.

Received February 17, 1970.

TEXAS A&M UNIVERSITY