DOUBLY STOCHASTIC MATRIX EQUATIONS

BY

J. S. MONTAGUE AND R. J. PLEMMONS†

ABSTRACT

It is shown that for real $m \times n$ matrices A and B the system of matrix equations $AX = B$, $BY = A$ is solvable for X and Y doubly stochastic if and only if $A = BP$ for some permutation matrix P. This result is then used to derive other equations and to characterize the Green's relations on the semigroup Ω_n of all $n \times n$ doubly stochastic matrices. The regular matrices in Ω_n are characterized in several ways by use of the Moore-Penrose generalized inverse. It is shown that a regular matrix in Ω_n is orthostochastic and that it is unitarily similar to a diagonal matrix if and only if it belongs to a subgroup of Ω_n. The paper is concluded with extensions of some of these results to the convex set S_n of all $n \times n$ nonnegative matrices having row and column sums at most one.

All matrices considered in this paper are real. Most of the definitions and notation are found in [12], although some are given below.

A square matrix $A = (a_{ij})$ is *doubly stochastic* if

$$a_{ij} \geqq 0 \text{ and } \sum_{k=1}^{n} a_{ik} = \sum_{k=1}^{n} a_{kj} = 1 \text{ for each } i \text{ and } j.$$

Let Ω, denote the set of all $n \times n$ doubly stochastic matrices and let $\mathscr{P}_n$ be the set of all $n \times n$ permutation matrices, that is, doubly stochastic matrices $P = (p_i)$ with $p_{ij} = 0$ or 1 for each i and j. Algebraically, Ω_n forms a semigroup under matrix multiplication and $\mathscr{P}_n$ is a subgroup of Ω_n. Moreover, Ω_n is a compact Hausdorff semigroup with respect to the natural topology [11]. Geometrically, Ω_n forms a convex polyhedron with the permutation matrices as vertices [14].

If $A_1, A_2, \cdots, A_s$ are square matrices then $A_1 \oplus A_2 \oplus \cdots \oplus A_s$ will denote the *direct sum*. The matrix A is said to be *decomposable* if there exists a permutation matrix P such that

†His research was supported by the N. S. F. Grant GP-15943.

Received October 15, 1972

$$PAP^T = \begin{bmatrix} B & 0 \\ 0 & D \end{bmatrix},$$

where B and D are square; otherwise A is *indecomposable.* Now A is said to be *partly decomposable* if there exist permutation matrices P, Q such that

$$PAQ = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix},$$

where B and D are square; otherwise A is *totally indecomposable.* It is easy to see that a doubly stochastic matrix A is either totally indecomposable or else there exist permutation matrices P and Q so that

$$PAQ = A_1 \oplus \cdots \oplus A_s$$

where each A_i is a totally indecomposable doubly stochastic matrix.

By the *norm* of an n-vector $x^T = (x_1, \cdots, x_n)$ we mean the Euclidean norm: $\| x \| = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$. By the norm of an $m \times n$ matrix A we mean the *spectral norm*

$$\| A \| = \sup_{x \neq 0} \frac{\| Ax \|}{\| x \|}.$$

The *singular values* of A are the square roots of the eigenvalues of $A^T A$. It follows [12] that $\| A \|$ is just the maximum singular value of A. Moreover there exist orthogonal matrices U and V such that

$$UAV^T = D = \operatorname{diag}\{\lambda_1, \cdots, \lambda_n\} \tag{1}$$

where the λ_i are the singular values of A. The form (1) is called a *singular value decomposition of A.*

Notice that if A is doubly stochastic then $\| A \| \leqq 1$, since the eigenvalues of $A^T A$ are less than or equal to 1.

Finally, by the *convex hull* $H(A)$ of A we mean the *convex hull* of the column vectors of A; that is, $H(A)$ is the set of all Ax where for $x^T = (x_1, \cdots, x_n)$, $x_i \geqq 0$ and $\sum_{i=1}^n x_i = 1$.

In Section I, necessary and sufficient conditions are given for the solvability of the equations $AX = B$, $BY = A$ for $X, Y \in \Omega_n$ and A, B arbitrary $m \times n$ real matrices. This characterization is used in Section II to derive the Green's relations on the semigroup Ω_n. In the final section, the regular matrices in Ω_n are investigated.

I. The equations $AX = B,\ BY = A$

S. Sherman [18] and S. Schreiber [16] have considered the matrix equation $AX = B$, where A, B and X are all doubly stochastic. A characterization of A and B in order that a solution X exist was given by Schreiber in the special case where A is nonsingular. More recently, D. J. Hartfiel [10] has considered the matrix equation $AXB = X$, with A, B and X doubly stochastic.

The purpose of this section is to consider the system of matrix equations

$$AX = B,\ BY = A, \quad X,\ Y \in \Omega_n$$

where A and B are *arbitrary* $m \times n$ real matrices. It is shown that solutions X and Y in Ω_n exist if and only if there is a permutation matrix P such that $A = BP$.

Let $x = (x_1, \cdots, x_n)$ be a real n-vector. By the *multiplicity* of x_i in x is meant the number of elements of x equal to x_i. Also $x_1^*, \cdots, x_n^*$ denote the elements of x arranged in nonincreasing order. For $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$ the relation $\preccurlyeq$ is defined by $a \preccurlyeq b$ if and only if

$$(2) \qquad a_1^* + \cdots + a_k^* \leqq b_1^* + \cdots b_k^* \text{ for } 1 \leqq k \leqq n, \text{ and } \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i.$$

We say that $a \approx b$ if $a \preccurlyeq b$ and $b \preccurlyeq a$.

The following classical result is fundamental to this section.

LEMMA 1 (Hardy, Littlewood and Pólya [9, Th. 46]). *Let a and b be n-vectors. Then the following statements are equivalent.*

i. $a \preccurlyeq b$.

ii. $a = bD$ *for some* $D \in \Omega_n$.

LEMMA 2. *Let a and b be n-vectors. Then the following statements are equivalent.*

i. $a \approx b$.

ii. *The equations* $aX = b,\ bY = a$ *are solvable for* $X,\ Y \in \Omega_n$.

iii. $a = bP$ *for some permutation matrix* P.

PROOF. The equivalence of (i) and (ii) follows from Lemma 1. Now if (i) holds, so that $a \preccurlyeq b$ and $b \preccurlyeq a$, then it follows from (2) that each component of a is a component of b with the same multiplicity, and conversely. Thus (i) implies (iii). The implication (iii) implies (i) is immediate.

LEMMA 3. *Let $a = (a, \cdots, a_n)$ be a real n-vector with the property that if $a_i < a_j$ and $i \leqq s \leqq j$, then $a_i = a_s$. (That is, all the multiple components in a* *occur in blocks.) Let* $a'_1, \cdots, a'_k$ *be the distinct elements of* a, *in order of appearance, and let* n_i *be the multiplicity of* a_i *in* a. *Then if* $aX = a$ *for some* $X \in \Omega_n$, *then* $X = X_1 \oplus \cdots \oplus X_k$, *where* X_i *is an* $n_i \times n_i$ *doubly stochastic matrix for* $1 \leqq i \leqq k$.

PROOF. Since $aX = a$ implies $aPP^TXP = aP$ for any $P \in \mathscr{P}_n$, no generality is lost by assuming that $a'_i > a'_{i+1}$ for each i. The proof is by induction on k, the result being trivial if $k = 1$.

Assume the lemma holds for all vectors with $k - 1$ distinct components. Suppose a has k distinct components and let $X = (x_{ij}) \in \Omega_n$, where $aX = a$. Then for $1 \leqq j \leqq n_1$,

$$a'_1 = a_1 = a_j = \sum_{i=1}^{n} a_i x_{ij} = \sum_{i=1}^{n} a_i x_{ij} + \sum_{i=n_1+1}^{n} a_i x_{ij}$$

$$\leqq a'_1 \sum_{i=1}^{n_1} x_{ij} + a'_2 \sum_{i=n_1+1}^{n} x_{ij} \leqq a_1.$$

Since equality must hold throughout and since $a'_2 < a'_1$, $\sum_{i=n_1+1}^{n} x_{ij} = 0$, $j = 1, \cdots, n_1$. Thus the nonzero entries in the first n_1 columns of X are confined to the first n_1 rows. Then since

$$\sum_{i=1}^{n_1} \sum_{j=1}^{n_1} x_{ij} = \sum_{j=1}^{n_1} \sum_{i=1}^{n_1} x_{ij} = \sum_{j=1}^{n_1} \sum_{i=1}^{n} x_{ij} = \sum_{j=1}^{n_1} 1 = n_1,$$

the nonzero entries of the first n_1 rows of X are confined to the first n_1 columns. Then $X = X_1 \oplus Y$, where X_1 and Y are doubly stochastic. Moreover if b and c are vectors consisting of the first n_1 and last $n - n_1$ entries of a, respectively, so that in block form $a = (b, c)$, then

$$aX = (b, c)\begin{pmatrix} X_1 & 0 \\ 0 & Y \end{pmatrix} = (bX_1, cY).$$

Since c has only $k - 1$ distinct entries, the result now follows by induction.

LEMMA 4. *Let* b *and* c *be* n-*vectors with* $b \approx c$. *If* $X \in \Omega_n$ *is totally indecomposable and* $bX = c$, *then all the elements of* b *are equal and* $b = c$.

PROOF. Let P and Q be permutation matrices such that bP^T and cQ have their components in nonincreasing order. Then $bP^T = cQ$ since $b \approx c$. Now let $a = bP^T = cQ$. Then a has the form given in Lemma 3 and $bX = c$ implies that $bP^TPXQ = cQ$, so that $aPXQ = a$. Thus $PXQ = X_1 \oplus \cdots \oplus X_k$ where each X_i is $n_i \times n_i$, by Lemma 3. Consequently, since X is totally indecomposable, a, b and c have only one distinct element. Hence $b = c$ is, in fact, a constant vector.

THEOREM 1. *Let A and B denote arbitrary $m \times n$ real matrices. Then the equations $AX = B$, $BY = A$ are solvable for X, Y in Ω_n if and only if $A = BP$ for some permutation matrix P.*

PROOF. Suppose $AX = B$ and $BY = A$ for X, $Y \in \Omega_n$ and let A_i, B_i denote the i'th rows of A and B. Then $A_i \approx B_i$ for $1 \leqq i \leqq m$. Consequently if X is totally indecomposable, $A = B$ by Lemma 4. Otherwise choose permutation matrices R and Q such that $RXQ = X_1 \oplus \cdots \oplus X_k$, where each X_i is totally indecomposable. Then

$$AR^TRXQ = AXQ = BQ$$

so that $A_iR^T(RXQ) = B_iQ$, for $1 \leqq i \leqq m$. Let $A_i' = A_iR^T$ and $B_i' = B_iQ$. Then $A_i' \approx B_i'$. Let $A_i' = (A_{i1}', \cdots, A_{ik}')$ and $B_i' = (B_{i1}', \cdots, B_{ik}')$ denote the block decomposition of A_i and B_i corresponding to the decomposition of X. Then $A_{ij}'X_j = B_{ij}'$ for $1 \leqq j \leqq k$.

Now let $a_1', \cdots, a_k'$ be the distinct elements of A_i' (and hence of B_i'), with $a_1' > a_2' > \cdots > a_k'$. Further, let n_{jl} be the multiplicity of a_l' in A_{ij}, let m_{jl} be the multiplicity of a_l' in B_{ij} and let n_l be the multiplicity of a_l' in A_i (and hence in B_i). Then

$$n_l = \sum_{j=1}^{k} n_{jl} = \sum_{j=1}^{k} m_{jl} \text{ for each } l.$$

Now since $B_{ij}' \preccurlyeq A_{ij}'$, then $m_{j1} \leqq n_{j1}$ for each j. Since $\sum_{j=1}^{k} m_{j1} = \sum_{j=1}^{k} n_{j1}$, this means $m_{j1} = n_{j1}$ for each j. Thus the multiplicity of a_1' in A_{ij}' is equal to its multiplicity in B_{ij}' for each j. This argument can now be applied again to get $m_{j2} = n_{j2}$ for each j, etc., eventually giving $m_{jl} = n_{jl}$ for each l. Thus $A_{ij}' \approx B_{ij}'$ for each j. Then since X_j is totally indecomposable, $A_{ij}' = B_{ij}'$ is a constant vector by Lemma 4. This means that $A_i = B_i$, for $1 \leqq i \leqq m$, so that $AR^T = BQ$. Thus $A = BP$ where $P = QR$.

The converse is immediate.

The proof of the following corollary is obtained by taking transposes in the previous theorem.

COROLLARY 1. *Let A and B be arbitrary $m \times n$ real matrices. Then the matrix equations $XA = B$, $YB = A$ are solvable for X, Y in Ω_m if and only if $A = PB$ for some permutation matrix P.*

II. The Green's relations in Ω_n

In this section some of the algebraic properties of the semigroup Ω_n are explored. In particular, the Green's relations on Ω_n are derived by use of Theorem 1 and its corollary. The following concepts from [4, Chapter II] will be needed.

Let S denote a semigroup with identity and let $a, b \in S$. The relation $\mathscr{R}[\mathscr{L},\mathscr{J}]$ is defined on S by $a\mathscr{R}b$ $[a\mathscr{L}b, a\mathscr{J}b]$ if and only if a and b generate the same principal right [left, two-sided] ideal in S. The relation $\mathscr{H}$ is defined on S by $\mathscr{H} = \mathscr{L} \cap \mathscr{R}$. Then each of $\mathscr{L}$, $\mathscr{R}$, $\mathscr{J}$ and $\mathscr{H}$ is an equivalence relation. Finally, the intersection of all the equivalence relations on S containing $\mathscr{L} \cup \mathscr{R}$ is denoted by $\mathscr{D}$. These are known as the Green's relations and they play a fundamental role in the study of the algebraic structure of semigroups.

From the definition of $\mathscr{R}$, if $a, b \in S$ then $a\mathscr{R}b$ if and only if there exist x and $y \in S$ such that $ax = b$ and $by = a$. Dually, $a\mathscr{L}b$ if and only if there exist x and $y \in S$ such that $xa = b$ and $yb = a$. The $\mathscr{R}-$ $[\mathscr{L}-, \mathscr{J}-, \mathscr{H}-, \mathscr{D}-]$ class of S containing a will be denoted by $R_a[L_a, J_a, H_a, D_a]$.

In particular then, the problem of characterizing the Green's relations on Ω_n may be solved by characterizing solutions to certain matrix equations which define the appropriate equivalence classes. The next theorem follows immediately from Theorem 1 and. ts Corollary.

THEOREM 2. *Let $A, B \in \Omega_n$. Then*

i. *$A\mathscr{R}B$ if and only if $A = BP$ for some $P \in \mathscr{P}_n$.*
ii. *$A\mathscr{L}B$ if and only if $A = QB$ for some $Q \in \mathscr{P}_n$.*
iii. *$A\mathscr{H}B$ if and only if $A = BP = QB$ for some $P, Q \in \mathscr{P}_n$.*
iv. *$A\mathscr{D}B$ if and only if $A = PBQ$ for some $P, Q \in \mathscr{P}_n$.*

Now since the topological semigroup Ω_n is compact, the relations $\mathscr{J}$ and $\mathscr{D}$ are the same on Ω_n [11, p. 30].

REMARKS

1. It follows from Theorem 1 that each $\mathscr{R}-[\mathscr{L}-, \mathscr{H}-, \mathscr{D}-]$ class of Ω_n is finite. Thus each principal right [left, two-sided] ideal of Ω_n has only a finite number of generators.

2. If $A, B \in \Omega_n$ and $A\mathscr{R}B$, then the convex hulls of the columns of A and B are of course the same. That the converse of this statement does not always hold is evident by the following example.

Let

$$A = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{2} & 0 & \frac{1}{3} & \frac{1}{6} \\ 0 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \end{bmatrix} \quad \text{and } B = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{8} & \frac{3}{8} \\ \frac{1}{2} & 0 & \frac{1}{8} & \frac{3}{8} \\ 0 & \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \\ 0 & \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \end{bmatrix}.$$

Then the convex hulls of the columns of A and B consist of all nonnegative vectors of the form

$$\begin{bmatrix} a \\ a \\ b \\ b \end{bmatrix}, \ a + b = \tfrac{1}{2}.$$

Clearly A and B are not $\mathscr{R}$—related.

III. Regular matrices in Ω_n

An element a in a semigroup S is said to be *regular* if the equation $a = axa$ is solvable for $x \in S$. If in addition $x = xax$, then a and x are said to be *semi-inverses*. Notice that if $a = axa$ then a and xax are semi-inverses. In this section the characterizations of the Green's relations given in Section II are used to investigate regularity in Ω_n.

Clearly not every matrix in Ω_n is regular. For example, the only regular doubly stochastic nonsingular matrices are the permutations matrices.

If an element in a $\mathscr{D}$-class D of a semigroup S is regular then each element in D is regular [4] and D is called a *regular $\mathscr{D}$-class*. In this case, associated with D is a maximal subgroup of S, isomorphic to each $\mathscr{H}$-class of S in D that contains an idempotent.

The maximal subgroup of Ω_n containing the $n \times n$ identity matrix I is the group $\mathscr{P}_n$ of all $n \times n$ permutation matrices. At the other extreme, the only idempotent in Ω_n of rank 1 is $\Lambda = (1/n)$ and the maximal subgroup of Ω_n containing Λ is $H_\Lambda = \{\Lambda\}$. Stefan Schwarz [17] and H. K. Farahat [7] have shown that the maximal subgroups of Ω_n are direct products of symmetric groups. This result can also be obtained by using our characterization of the Green's relation $\mathscr{H}$ on Ω_n in conjunction with the following well known result.

LEMMA 5 (Doob [5]). *An $n \times n$ doubly stochastic matrix E is idempotent if and only if there exists a permutation matrix P such that PEP^T has the form*

$$PEP^T = E_1 \oplus \cdots \oplus E_k$$

where each E_i is the $\lambda_i \times \lambda_i$ matrix $(1/\lambda_i)$, $n = \lambda_1 + \cdots + \lambda_k$ and $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_k \geqq 1$.

If E itself has the form given in the lemma then it will be called a *canonical idempotent*. Clearly each idempotent in Ω_n is symmetric and, moreover, Ω_n contains only a finite number of idempotents.

LEMMA 6. *An $\mathscr{R}$-[$\mathscr{L}$-] class of Ω_n contains at most one idempotent. If a $\mathscr{D}$-class D of Ω_n contains an idempotent it contains exactly one canonical idempotent.*

PROOF. Let E, F be idempotents in an $\mathscr{R}$-class R of Ω_n. Then E and F are left identities for the elements in R so that

$$E = E^T = (FE)^T = E^T F^T = EF = F.$$

If a $\mathscr{D}$-class D contains an idempotent E then $PEP^T \in D$ for each permutation matrix P so that D contains a canonical idempotent. However, from the form given in Lemma 5, if J, J' are distinct canonical idempotents, then there do not exist permutation matrices P, Q such that $J' = PJQ$. Thus D contains at most one canonical idempotent by Theorem 2.

The regular matrices $A \in \Omega_n$, that is, the matrices A for which the equation $A = AXA$ is solvable in Ω_n, will now be characterized in several ways. The following concepts will facilitate these characterizations.

By the *Moore-Penrose generalized inverse* of an arbitrary $m \times n$ complex matrix A is meant the unique solution to the equations

$$A = AXA, \; X = XAX \text{ with } AX, \; XA \text{ Hermitian.}$$

The solution always exists and is unique. It will be denoted by A^+. The properties and applications of such inverses are described in a number of papers including those of Penrose [15], Ben Israel and Charnes [1], and Greville [8]. The following discussions will be restricted to $n \times n$ real matrices A, in which case the requirement on the idempotents AA^+ and A^+A is that they be symmetric.

Consider the singular value decomposition

$$UAV^T = D = \operatorname{diag}\{\lambda_1, \cdots, \lambda_n\}. \tag{3}$$

It follows that if D' is obtained from D by inverting the nonzero λ_i then

$$A^+ = V^T D' U \tag{4}$$

Now the matrix A is called a *partial isometry* if the linear transformation $y = Ax$ preserves Euclidean distances on the range $R(A^T)$ of A^T; that is, if

$$\| Ax_1 - Ax_2 \| = \| x_1 - x_2 \|, \text{ for all } x_1, x_2 \in R(A^T).$$

This condition may be readily shown to be equivalent to

$$\| Ax \| = \| x \| \text{ for all } x \in R(A^T).$$

It has been shown by Erdelyi [6] that the matrix A is a partial isometry if and only if $A^+ = A^T$.

Finally, A is said to be *row-monotone* if

$$Ax \geqq 0, \; x \in R(A^T) \text{ implies } x \geqq 0.$$

It was shown in [2, Th. 3] that if A is nonnegative then A^+ is nonnegative if and only if A and A^T are row-monotone. Regular matrices in Ω_n are now characterized.

THEOREM 3. *Let $A \in \Omega_n$. Then the following statements are equivalent.*

1. *A is regular.*
2. *A^T is the unique semi-inverse of A in Ω_n.*
3. *The singular values of A are* 0 *and* 1.
4. *A is a partial isometry.*
5. *A and A^T are row-monotone.*
6. *A^+ is nonnegative.*
7. $\| A^+ \| = 1$.
8. $A^+ = A^T$.

PROOF.

($1 \Rightarrow 2$). Let X be any semi-inverse of A in Ω_n. Then AX, in the $\mathscr{R}$-class R_A containing A, is idempotent and $AX = AP$ for some permutation matrix P, by Theorem 2. Now $(AP)^T = AP$ so that $A^T = PAP$. Then $AA^T = APAP = AP = AX$. Also $(PA)^2 = PAPA = PA$ so that $PA = XA$, since the idempotent in the $\mathscr{L}$-class containing A is unique by Lemma 6. Then

$$X = XAX = PAX = PAP = A^T.$$

($2 \Rightarrow 3$). Since A^TA is idempotent its eigenvalues are 0 and 1, so that the singular values of A are 0 and 1.

($3 \Rightarrow 4$). If the singular values of A are 0 and 1 then A^TA is idempotent since there exists an orthogonal matrix U such that $UA^TAU^T = \text{diag}\{\lambda_1, \cdots, \lambda_n\}$, where $\lambda_i = 0$ or 1. Now A^TA and A^T have the same range. Thus since an idempotent is the identity on its range, $x \in R(A^T)$ implies that $A^TAx = x$. Then for $x \in R(A^T) = R(A^TA)$

$$\| Ax \| = x^TA^TAx = x^Tx = \| x \|.$$

($4 \Rightarrow 5$). By [6], $A^T = A^+$ so that from [2], A and A^T are row-monotone.

($5 \Rightarrow 6$). This also follows from [2].

($6 \Rightarrow 7$). Assume that A is nonnegative. To show that $\| A^+ \| = 1$, it suffices to show that $A^+ \in \Omega_n$. Let e denote the column n-vector of ones. Now since $A \in \Omega_n$, $Ae = A^Te = e$.

Then

$$e = A^Te = A^T(A^+)^TA^Te = A^T(A^+)^Te = (A^+A)^Te = A^+Ae = A^+e.$$

Dually $(A^+)^Te = e$, so that $A^+ \in \Omega_n$ and therefore $\| A^+ \| = 1$.

$(7) \Rightarrow (8)$. If $\| A^+ \| = 1$ then since the nonzero singular values of A are the reciprocals of those of A^+ by (3) and (4), the singular values of A are 0 and 1. Consequently AA^T and A^TA are idempotent so that $A^+ = A^T$.

($8 \Rightarrow 1$). This follows from the definition of A^+.

It is shown next that regular doubly stochastic matrices are obtained from unitary matrices. If $U = (u_{ij})$ is unitary, then the matrix $A = (a_{ij})$, where $a_{ij} = |u_{ij}|^2$ for each i and j, is doubly stochastic. Matrices obtained in this way are said to be *orthostochastic*. Not every doubly stochastic matrix has this property [14].

THEOREM 4. *Each regular doubly stochastic matrix is orthostochastic.*

PROOF. Let A be regular in Ω_n and let E denote the canonical idempotent in D_A. Then by Theorem 2 there exist permutation matrices P and Q such that $E = PAQ$. In particular then, it suffices to show that E is orthostochastic.

Now $E = E_1 \oplus \cdots \oplus E_k$ where E_i is $\lambda_i \times \lambda_i$ with each entry $1/\lambda_i$ for $1 \leqq i \leqq k$. To show that E is orthostochastic it suffices to show that each E_i is orthostochastic. However, if λ is any positive integer and if ω is a primitive λ'th root of 1, then the matrix $M = (m_{ij})$ where

$$m_{ij} = \frac{1}{\sqrt{\lambda}} \omega^{(i-1)(j-1)}$$

is unitary and thus the matrix

$$(|m_{ij}|^2) = (1/\lambda)$$

is orthostochastic. Therefore A is orthostochastic.

The converse of the theorem is not true. Notice that

$$A = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

is orthostochastic but not regular.

The final result provides a partial solution to Mirsky's Problem No. (iv) in [14, p. 243], asking for a characterization of those doubly stochastic matrices that are unitarily similar to a diagonal matrix. The characterization is given for the regular matrices in Ω_n.

THEOREM 5. *Let $A \in \Omega_n$. Then A is regular and unitarily similar to a diagonal matrix if and only if A belongs to a subgroup of Ω_n, in which case the nonzero eigenvalues of A are on the unit circle.*

PROOF. Suppose A is regular and $UAU^* = D = \mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$ for some unitary matrix U. Then $A^+ = U^*D'U$ where D' is obtained from D by inverting the nonzero entries. However, since A is regular it follows that $A^+ = A^T$ by Theorem 3. Then $AA^T = A^TA$ is idempotent and thus A belongs to a subgroup of Ω_n. Also, since $U^*D'U \in \Omega_n$, the nonzero eigenvalues of A must lie on the unit circle.

Conversely if A belongs to a subgroup H of Ω_n with identity E, then $AA^T = A^TA = E$ by Theorem 3. Thus A is regular and, in particular, normal, so that A is unitarily similar to a diagonal matrix.

IV. Appendix: Sub-doubly stochastic matrices

Let S_n denote the set of all $n \times n$ nonnegative matrices having row and column sums at most one. Then S_n is a multiplicative semigroup containing Ω_n as a subsemigroup. In addition, S_n is a convex polyhedron with Ω_n on its boundary. The set of all $n \times n$ sub-permutation matrices (that is, all $0-1$ matrices in S_n) forms the vertices of S_n [13].

Many of the results given in the previous sections also hold with Ω_n replaced by S_n. The purpose of this appendix is to indicate how the previous results for Ω_n can be modified in order to obtain similar theorems for S_n. First, it is shown how Theorem 1 can be restated for S_n by modifying Lemmas 1–4.

If r is any real number then let

$$\bar{r} = \begin{cases} r & \text{if } r \geqq 0, \\ 0 & \text{if } r < 0. \end{cases}$$

For an n-vector $a = (a_1, \cdots, a_n)$, $\bar{a}$ is defined to be $(\bar{a}_1, \cdots, \bar{a}_n)$. Then by applying [13, Th. 6], Lemma 1 can be replaced by the following.

LEMMA 1′. *Let a and b be n-vectors. Then the following are equivalent.*

(i) $a \preccurlyeq \bar{b}$ *and* $-a \preccurlyeq -\bar{b}$.

(ii) $a = bS$ *for some* $S \in S_n$.

Now Lemma 2 remains valid when Ω_n is replaced by S_n. For the proof one notes that if $a \preccurlyeq \bar{b}$, $-a \preccurlyeq -\bar{b}$, $b \preccurlyeq \bar{a}$ and $-b \preccurlyeq -\bar{a}$, then $a \approx b$.

In Lemma 3, Ω_n is replaced by S_n and the conclusion is changed to read "$X = X_1 \oplus \cdots \oplus X_k$, where all but possibly one of the X_i are doubly stochastic, the exceptional one corresponding to $a_i' = 0$". The proof is essentially the same except that the ordering of the elements of a is different.

Finally, Lemma 4 remains valid when Ω_n is replaced by S_n and Theorem 1 can be restated as follows.

THEOREM 1′. *Let A and B denote arbitrary $m \times n$ real matrices. Then the equations $AX = B$, $BY = A$ are solvable for X, Y in S_n if and only if $A = BP$ for some permutation matrix P.*

The proof of Theorem 1′ is exactly the same as the proof of Theorem 1, provided that Ω_n is replaced by S_n. A similar restatement of Corollary 1 can be given.

The entire Section II can be stated in terms of S_n as well as Ω_n. Thus Theorem 2 with Ω_n replaced by S_n, characterizes the Green's relations on the semigroup S_n.

In Section III, Lemma 5 can be modified to give a characterization of the doubly sub-stochastic idempotents by replacing $PEP^T = E_1 \oplus \cdots \oplus E_k$ by

$$PEP^T = E_1 \oplus \cdots \oplus E_k \oplus Z$$

where Z is the zero matrix of order m where E has exactly m nonzero rows [columns]. Thus Lemma 6 remains valid with Ω_n replaced by S_n.

The following theorem is immediate from the restatements of Corollary 2 and Lemma 6. It shows that regularity is S_n can be studied in terms of regularity in Ω_m, for some $m \leqq n$.

THEOREM 6. *Let $A \in S_n$. Then A is regular in S_n if and only if there exist permutation matrices P and Q such that*

$$PAQ = B \oplus Z$$

where B is regular in Ω_m, *for some* $m \leqq n$, *and Z is the zero matrix of order* $n - m$.

The maximal subgroups of S_n can now be determined by applying Theorem 6 and the remarks preceding Lemma 5.

COROLLARY 3. *The semigroup* S_n *contains finitely many maximal subgroups, each of which is isomorphic to a direct product of full symmetric groups.*

Finally, we note that Theorem 3 and Theorem 5 hold with Ω_n replaced by S_n, whenever A is not the zero matrix, while Theorem 4 is given only for Ω_n.

REFERENCES

1. A. Ben-Israel and A. Charnes, *Contributions to the theory of generalized inverses*, J. Soc. Indust. Appl. Math. **XI** (1963), 667–699.

2. A. Berman and R. Plemmons, *Monotonicity and the generalized inverse*, SIAM J. Appl. Math., **22** (1972), 155–161.

3. T. L. Boullion and P. L. Odell, *Generalized Inverse Matrices*, John Wiley and Sons, New York, 1971.

4. A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups*, Amer. Math. Soc. Surveys, Vol. I, 1967.

5. J. L. Doob, *Topics in the theory of Markov chains*, Trans. Amer. Math. Soc. **52** (1942), 37–64.

6. I. Erdelyi, *On partial isometries in finite dimensional Euclidean spaces*, J. Soc. Indust. Appl. Math. **XIV** (1966), 453–467.

7. H. K. Farahat, *The semigroup of doubly stochastic matrices*, Proc. Glasgow Math. Assoc. **7** (1966), 178–183.

8. T. N. E. Greville, *Some applications of the pseudoinverse of a matrix*, SIAM Rev. **II** (1960), 15–22.

9. G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge Press, Cambridge, England, 1934.

10. D. J. Hartfiel, *The matrix equation* $AXB = X$, Pacific J. Math. **36** (1971), 659–669.

11. K. H. Hofmann and P. S. Mostert. *Elements of compact Semigroups*, Charles E. Merrill, Columbus, Ohio, 1966.

12. A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.

13. L. Mirsky, *On a convex set of matrices*, Arch. Math. **10** (1959), 88–92.

14. L. Mirsky, *Results and problems in the theory of doubly stochastic matrices*, Z. Wahrscheinlichkeitstheorie und Veiw. Gebiete **1** (1963), 319–334.

15. R. Penrose, *A generalized inverse for matrices*, Proc. Camb. Philos. Soc. **51** (1955), 406–413.

16. S. Schreiber, *On a result of S. Sherman concerning doubly stochastic matrices*, Proc. Amer. Math. Soc. **9** (1958), 350—353.

17. S. Schwarz, *A note on the structure of the semigroup of doubly stochastic matrices*, Mat. Casopis Sloven. Akad. Vied. **17** (1967), 308–316.

18. S. Sherman, *A correction to "On a conjecture concerning doubly stochastic matrices"*, Proc. Amer. Math. Soc. **5** (1954), 988–999.

UNIVERSITY OF TENNESSEE,
KNOXVILLE, TENNESSEE, U. S. A.