DOUBLY STOCHASTIC MATRICES AND THE DIAGONAL OF A ROTATION MATRIX.* [1]

By Alfred Horn.

An elegant theorem of Hardy, Littlewood and Polya (Theorem 1 below) has been generalized and applied in various ways in recent literature. In this paper we shall present an extension of their theorem (Theorm 4). As an application we prove that the set of all diagonals of rotation matrices of order n is equal to the convex hull of those points $(\pm 1, \cdots, \pm 1)$ of which an even number of coordinates are -1. From this, we determine the set of all diagonals of orthogonal or of unitary matrices. It would be interesting to prove without Theorem 4 that the set of diagonals of rotation matrices is convex.

1. Definitions and notation. If A is a matrix, then A_{ij} will denote its i, j component and we write $A = (A_{ij})$. A similar convention will be used for vectors. An n-vector is a vector with n components. A unit vector is a vector of length one. The diagonal of a matrix A is the vector $(A_{11}, \cdots, A_{nn})$. A diagonal matrix is one whose non-diagonal elements vanish. The identity matrix is the diagonal matrix with diagonal $(1, \cdots, 1)$.

A permutation matrix is a matrix with elements all 0 or 1 such that every row and every column contains exactly one element equal to 1. A doubly stochastic (d. s.) matrix is a matrix P such that $P_{ij} \geqq 0$, $\sum_i P_{ij} = \sum_j P_{ij} = 1$ for all i and j. We use the symbols A', $\bar{A}$ and A^* for transpose, conjugate and conjugate-transpose respectively. By an orthogonal matrix, we mean a real matrix A such that $A' = A^{-1}$, A rotation matrix, or rotation, is an orthogonal matrix with determinant $+1$. A unitary matrix is a matrix A such that $A^* = A^{-1}$.

If x is a real n-vector, then $H(x)$ is the convex hull of all the points $(x_{\alpha_1}, \cdots, x_{\alpha_n})$, α varying over all permutations of $(1, \cdots, n)$. By $S^n{}_k$, we mean the set of all k termed sequences σ of integers for which $1 \leqq \sigma_1 < \cdots < \sigma_k \leqq n$. The only member of $S^n{}_0$ is the empty sequence.

* Received April 16, 1953.

[1] This paper was written while the author received partial support from the Office of Naval Research.

If x is a complex number, $\mathcal{R}x$ denotes its real part. Occasionally we shall use i both as a summation index and to denote the imaginary unit $(-1)^{\frac{1}{2}}$. But we shall do this in such a way as to avoid confusion.

2. Doubly stochastic and ortho-stochastic matrices.

THEOREM 1. *Let x, y be real n-vectors. Then the following statements are equivalent:*

(1a) $y = Px$, *where P is a d. s. matrix.*

(1b) $y \,\varepsilon\, H(x)$.

(1c) $\max\limits_{\sigma \varepsilon S^n_k} \sum_{i=1}^{k} y_{\sigma_i} \leqq \max\limits_{\sigma \varepsilon S^n_k} \sum_{i=1}^{k} x_{\sigma_i}, \quad 1 \leqq k \leqq n$ and $\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i$.

(1d) $\sum_{i=1}^{n} f(y_i) \leqq \sum_{i=1}^{n} f(x_i)$ *for any convex function f whose domain contains all the numbers x_i, y_i, $1 \leqq i \leqq n$.*

Proof. The equivalence of (1a), (1c) and (1d) is proved in [1], pp. 49 and 89. Also (1b) is equivalent to (1a) by virtue of the following theorem of G. Birkhoff [2].

THEOREM 2. *A matrix is d. s. if and only if it lies in the convex hull of the set of all permutation matrices.*

In case $x_1 \geqq \cdot \cdot \cdot \geqq x_n$, the conditions (1c) are equivalent to

(1e) $$\sum_{i=n-k+1}^{n} x_i \leqq \sum_{i=1}^{k} y_{\sigma_i} \leqq \sum_{i=1}^{k} x_i, \quad \sigma \,\varepsilon\, S^n_k, \quad 1 \leqq k \leqq n.$$

We take this opportunity to point out a companion to Theorem 1, various parts of which are scattered through the literature.

THEOREM 3. *Let x, y be n-vectors for which $x_i \geqq 0$, $y_i \geqq 0$, $1 \leqq i \leqq n$. Then the following statements are equivalent:*

(2a) $y = Px$, *where P is a matrix such that* $P_{ij} \geqq 0$, $\sum_{i=1}^{n} P_{ij} \leqq 1$, $\sum_{j=1}^{n} P_{ij} \leqq 1$, $1 \leqq i \leqq n$, $1 \leqq j \leqq n$.

(2b) $\max\limits_{\sigma \varepsilon S^n_k} \sum_{i=1}^{k} y_{\sigma_i} \leqq \max\limits_{\sigma \varepsilon S^n_k} \sum_{i=1}^{k} x_{\sigma_i}, \quad 1 \leqq k \leqq n.$

(2c) $\sum_{i=1}^{n} f(y_i) \leqq \sum_{i=1}^{n} f(x_i)$ *for any convex non-decreasing function f whose domain contains all the numbers x_i, y_i, $1 \leqq i \leqq n$.*

Proof. The proof given in [1] for the implication (1d) $\rightarrow$ (1c) also proves (2c) $\rightarrow$ (2b). The fact that (2b) $\rightarrow$ (2c) is proved in [3]. Ky Fan [4] proves (2a) $\rightarrow$ (2b). He also states (2b) $\rightarrow$ (2a) and quotes a lemma used in his proof. Using the fact that (1c) $\rightarrow$ (1a), a simple proof may be obtained. It is also possible to derive (2a) $\rightarrow$ (2b) from (1a) $\rightarrow$ (1b) by using the fact that any matrix of the kind described in (2a) may be augmented to a d. s. matrix (of order $\leqq 2n$).

A special class of d. s. matrices consists of those matrices Q such that $Q_{ij} = U_{ij}^2$ where U is an orthogonal matrix. Let us call such matrices orthostochastic (o. s.) matrices. Not every d. s. matrix is an o. s. matrix. In fact it is not hard to see that the matrix

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

cannot even be expressed as a product of o. s. matrices. (This example is due to A. J. Hoffman.) Nevertheless we have the following theorem.

THEOREM 4. *If x, y are n-vectors satisfying* (1c), *then there exists an o. s. matrix Q such that $y = Qx$.*

The proof of Theorem 4 will be simplified if we first prove a lemma.

DEFINITON. If $x_1 \geqq \cdot\cdot\cdot \geqq x_n$, and $1 \leqq m \leqq n$, let $T^m(x_1, \cdot\cdot\cdot, x_n)$ be the set of points $(z_1, \cdot\cdot\cdot, z_m)$ such that $\sum_{i=n-k+1}^{n} x_i \leqq \sum_{i=1}^{k} z_{\sigma_i} \leqq \sum_{i=1}^{k} x_i$ whenever $\sigma \in S^n{}_k$, $1 \leqq k \leqq m$. Also let $T^0(x_1, \cdot\cdot\cdot, x_n)$ be the empty set.

LEMMA. *If $x_1 \geqq \cdot\cdot\cdot \geqq x_n$ and $(y_1, \cdot\cdot\cdot, y_{n-1}) \in T^{n-1}(x_1, \cdot\cdot\cdot, x_n)$ and $n \geqq 2$, then there exists a set of $n-1$ real ortho-normal n-vectors $u^1, \cdot\cdot\cdot, u^{n-1}$ such that*

$$y_j = \sum_{i=1}^{n} (u^j{}_i)^2 x_i, \quad 1 \leqq j \leqq n-1. \tag{3}$$

Proof. The statement is obvious for $n = 2$. We proceed by induction. Suppose the lemma is true for $2 \leqq n < N$. Clearly $T^{N-1}(x_1, \cdot\cdot\cdot, x_N)$ consists of the points $(y_1, \cdot\cdot\cdot, y_{N-1})$ such that $(y_1, \cdot\cdot\cdot, y_{N-2}) \in T^{N-2}(x_1, \cdot\cdot\cdot, x_N)$ and such that $b \leqq y_{N-1} \leqq a$, where b is the largest of the numbers $b_\sigma = \sum_{i=N-k}^{N} x_i - \sum_{i=1}^{k} y_{\sigma_i}$, and a is the smallest of the numbers $a_\sigma = \sum_{i=1}^{k+1} x_i - \sum_{i=1}^{k} y_{\sigma_i}$ as σ ranges over $S^{N-2}{}_k$ and k varies from 0 to $N-2$. (When σ is the empty sequence, $b_\sigma = x_N$ and $a_\sigma = x_1$).

For each $\sigma \varepsilon S^{N-2}{}_k$, $0 \leqq k \leqq N-2$, let T_σ be the subset of $T^{N-2}(x_1,\cdots,x_N)$ on which $a = a_\sigma$, and let T'_σ be the subset on which $b = b_\sigma$. It is not hard to verify that T_σ consists of the points $(y_1,\cdots,y_{N-2})$ such that

$$(y_{\sigma_1},\cdots,y_{\sigma_k}) \,\varepsilon\, T^k(x_1,\cdots,x_{k+1}) \text{ and } (y_{\sigma'_1},\cdots,y_{\sigma'_{N-2-k}}) \,\varepsilon$$

$T^{N-2-k}(x_{k+2},\cdots,x_N)$, where σ' is the sequence consisting of those integers from 1 to $N-2$ which are not among the terms of σ. Verify also that $T_\sigma = T'_{\sigma'}$. It follows that $T^{N-1}(x_1,\cdots,x_N) = \bigcup_{k=0}^{N-2} \bigcup_{\sigma \varepsilon S^{N-2}{}_k} S_\sigma$, where S_σ is the set of points $(y_1,\cdots,y_{N-1})$ such that $(y_1,\cdots,y_{N-2}) \,\varepsilon\, T_\sigma$ and $b_{\sigma'} \leqq y_{N-1} \leqq a_\sigma$.

Now let $(y_1,\cdots,y_{N-1})$ be a fixed point of $T^{N-1}(x_1,\cdots,x_N)$. Then $(y_1,\cdots,y_{N-1}) \,\varepsilon\, S_\sigma$ for some fixed $\sigma \varepsilon S^n{}_k$ and some fixed $k \leqq N-2$. We must find an ortho-normal sequence $u^1,\cdots,u^{N-1}$ for which (3) holds. By the induction hypothesis, there exist k real ortho-normal $k+1$-vectors $v^{\sigma_1},\cdots,v^{\sigma_k}$ and $N-2-k$ real ortho-normal $N-1-k$-vectors $v^{\sigma'_1},\cdots,$ $v^{\sigma'_{N-2-k}}$ such that $y_{\sigma_m} = \sum_{i=1}^{k+1} (v^{\sigma_m}{}_i)^2 x_i$, $1 \leqq m \leqq k$, $y_{\sigma'_m} = \sum_{i=1}^{N-1-k} (v^{\sigma'_m}{}_i)^2 x_{k+1+i}$, $1 \leqq m \leqq N-2-k$. We define $N-2$ ortho-normal N-vectors $u^1,\cdots,u^{N-2}$ as follows: $u^{\sigma_m} = (v^{\sigma_m}{}_1,\cdots,v^{\sigma_m}{}_{k+1},0,\cdots,0)$, $1 \leqq m \leqq k$; $u^{\sigma'_m} = (0,\cdots,0,$ $v^{\sigma'_m}{}_1,\cdots,v^{\sigma_m'}{}_{N-1-k})$, $1 \leqq m \leqq N-2-k$. We have $y_j = \sum_{i=1}^{N} (u^j{}_i)^2 x_i$, $1 \leqq j \leqq N-2$. If u ranges over the circle consisting of all real N-vectors orthogonal to $u^1,\cdots,u^{N-2}$, then the values of the sum

$$\sum_{i=1}^{N} (u_i)^2 x_k \tag{4}$$

fill out some interval. If we take $u = (v_1,\cdots,v_{k+1},0,\cdots,0)$, where $(v_1,\cdots,v_{k+1})$ is a real unit $k+1$-vector orthogonal to $v^{\sigma_1},\cdots,v^{\sigma_k}$, then (4) takes on the value $\sum_{i=1}^{k+1} x_i - \sum_{j=1}^{k} y_{\sigma_j} = a_\sigma$. Also if we take $u = (0,\cdots,0,$ $v_1,\cdots,v_{N-1-k})$, where $(v_1,\cdots,v_{N-1-k})$ is a real unit vector orthogonal to $v^{\sigma'_1},\cdots,v^{\sigma'_{N-2-k}}$, then (4) takes on the value $\sum_{i=k+2}^{N} x_i - \sum_{j=1}^{N-2-k} y_{\sigma'_j} = b_{\sigma'}$. Therefore if $b_{\sigma'} \leqq y_{N-1} \leqq a_\sigma$, there exists a real unit N-vector u^{N-1} orthogonal to $u^1,\cdots,u^{N-2}$ such that $y_{N-1} = \sum_{i=1}^{N} (u^{N-1}{}_i)^2 x_i$.

Proof of Theorem 4. There is no loss of generality if we assume $x_1 \geqq \cdots \geqq x_n$. As mentioned above, condition (1c) is equivalent with (1e).

Furthermore it is easily verified that (1e) is equivalent to the conditions $(y_1, \cdots, y_{n-1}) \varepsilon T^{n-1}(x_1, \cdots, x_n)$, and $y_n = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n-1} y_i$. By the preceding lemma, there exists an ortho-normal system $u^1, \cdots, u^{n-1}$ such that (3) holds. If u^n is a real unit vector orthogonal to $u^1, \cdots, u^{n-1}$, it follows that $\sum_{i=1}^{n} (u^n{}_i)^2 x_i = \sum_{i=1}^{n} x_i - \sum_{j=1}^{n-1} y_j = y_n$. Thus if we set $Q_{ij} = (u^j{}_i)^2$, then $Q = (Q_{ij})$ is the desired o. s. matrix.

An immediate consequence of Theorem 4 is the following.

THEOREM 5. *A vector y can be the diagonal of a Hermitian matrix with eigenvalues $x_1, \cdots, x_n$ if and only if* (1c) *holds.*

Proof. A Hermitian matrix A has eigenvalues $x_1, \cdots, x_n$ if and only if $A = UBU^*$, where B is the diagonal matrix with diagonal $(x_1, \cdots, x_n)$, and U is a unitary matrix. Therefore y is the diagonal of such a matrix A if and only if $y_i = \sum_{j=1}^{n} | U_{ij} |^2 x_j$. Theorems 1 and 4 now yield the conclusion.

The neecssity of (1c) in Theorem 5 was pointed out by Schur [5], who showed that this leads to a simple proof of Hadamard's inequality for determinants.

It is interesting to compare Theorem 5 with the following theorem, proved in [6].

THEOREM 6. *Let $x_i > 0$, $1 \leqq i \leqq n$. Then there exists a Hermitian matrix A with eigenvalues $x_1, \cdots, x_n$ such that $\prod_{i=1}^{k} y_i = M_k$, $1 \leqq k \leqq n$, where M_k is the determinant formed from the first k rows and columns of A, if and only if $\max_{\sigma \varepsilon S^n{}_k} \prod_{i=1}^{k} y_{\sigma_i} \leqq \max_{\sigma \varepsilon S^n{}_k} \prod_{i=1}^{k} x_{\sigma_i}$ and $\prod_{i=1}^{n} y_i = \prod_{i=1}^{n} x_i$.*

3. Remarks on the complex case. If $y = Px$, where x, y are *complex* n-vectors, and P is a d. s. matrix, then there is an analogue to the implication (1a) $\rightarrow$ (1e). We can show that for each $\sigma \varepsilon S^n{}_k$, $\sum_{i=1}^{k} y_{\sigma_i}$ lies in the convex hull of all the points $\sum_{i=1}^{k} x_{\tau_i}$, $\tau \varepsilon S^n{}_k$. More generally, we have:

THEOREM 7. *If $y = Px$, where x, y are complex n-vectors, and P is a d. s. matrix, and if $c_1, \cdots, c_n$ are any complex numbers, then $\sum_{i=1}^{n} c_i y_i$ lies in* *the convex hull of all the points* $\sum_{i=1}^{n} c_i x_{\alpha_i}$, $\alpha \varepsilon R^n$, *where* R^n *is the set of all permutations of* $(1, \cdots, n)$.

Proof. By Theorem 2, we may write $P = \sum_{\alpha \varepsilon R^n} a_\alpha P^\alpha$, where $P^\alpha{}_{ij} = 1$ when $\alpha_i = j$, and $P^\alpha{}_{ij} = 0$ otherwise, and $a_\alpha \geqq 0$, $\sum_\alpha a_\alpha = 1$. Therefore

$$\sum_{i=1}^{n} c_i y_i = \sum_{i,j=1}^{n} P_{ij} c_i x_j = \sum_\alpha a_\alpha \sum_{i,j=1}^{n} P^\alpha{}_{ij} c_i x_j$$
$$= \sum_\alpha a_\alpha \sum_{i=1}^{n} c_i \sum_{j=1}^{n} P^\alpha{}_{ij} x_j = \sum_\alpha a_\alpha \sum_{i=1}^{n} c_i x_{\alpha_i}.$$

We do not know whether the converse of Theorem 7 is true. However, there cannot be any simple analogue of (1e) $\rightarrow$ (1a). In fact there exists a pair x, y of 4-vectors such that $\sum_{i=1}^{4} c_i y_i$ lies in the convex hull of the points $\sum_{i=1}^{4} c_i x_{\alpha_i}$ for all *real* $c_1 \cdots, c_4$ and yet there exists no d. s. matrix P such that $y = Px$. For example take $y = (1/4 + 2i/3, -1/4 + 2i/3, 0, 0)$ and $x = (1, -1, i, i/3)$. Theorem 4 also breaks down in the complex case. In fact if $y = (0, 1/2 + i/2, -1/2 + i/2)$, $x = (1, -1, i)$, then there is one and only one d. s. matrix P for which $y = Px$, namely

$$P = \tfrac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

However the equivalence (1d) $\leftrightarrow$ (1a) still holds for complex vectors [7].

4. The diagonal of a rotation.

THEOREM 8. *A vector* $(d_1, \cdots, d_n)$ *is the diagonal of a rotation of order n if and only if it lies in the convex hull of those points* $(\pm 1, \cdots, \pm 1)$ *of which an even number (possibly* 0) *of coordinates are* -1.

Proof. If R is a rotation, then its eigenvalues are all of modulus 1. The complex eigenvalues occur in conjugate pairs and an even number of the real eigenvalues are -1. Therefore we may say that if n is odd, say $n = 2m + 1$, then the eigenvalues are of the form 1, $\exp(\pm i\theta_1), \cdots, \exp(\pm i\theta_m)$, where $0 \leqq \theta_j < 2\pi$. If $n = 2m$, the eigenvalues have the form $\exp(\pm i\theta_1), \cdots,$ $\exp(\pm i\theta_m)$. In what follows, the case n even is practically identical with the case n odd. Accordingly we will only treat the case $n = 2m + 1$. It is well known that there exists an orthogonal matrix U such that $R = U'AU$, where

$$A = \begin{bmatrix} 1 & & & 0 \\ & B_1 & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & B_m \end{bmatrix}, \quad \text{and} \quad B_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}, \ 1 \leqq i \leqq m.$$

Therefore

$$R_{jj} = U_{j1}{}^2 + U_{j2}{}^2 \cos\theta_1 + U_{j3}{}^2 \cos\theta_1 + \cdots + U_{j2m}{}^2 \cos\theta_m + U_{j2m+1}{}^2 \cos\theta_m, \qquad 1 \leqq j \leqq n.$$

By Theorem 4, it follows that $(d_1, \cdots, d_n)$ is the diagonal of a rotation with eigenvalues 1, $\exp(\pm i\theta_1), \cdots, \exp(\pm i\theta_m)$ if and only if it lies in $H(1, \cos\theta_1, \cos\theta_1, \cdots, \cos\theta_m, \cos\theta_m)$. Since $H(x)$ is unchanged if we permute the coordinates of x, we find that the set S of all diagonals of n-th order rotations has the form

$$(5) \qquad S = \bigcup\nolimits_{1 \geqq a_1 \geqq \cdots \geqq a_m \geqq -1} H(1, a_1, a_1, \cdots, a_m, a_m).$$

Let T be the convex hull of those points $(\pm 1, \cdots, \pm 1)$ of which an even number of coordinates are -1. To prove $S \subset T$, we need only show that if $1 \geqq a_1 \geqq \cdots \geqq a_m \geqq -1$, then $(1, a_1, a_1, \cdots, a_m, a_m) \, \varepsilon \, T$. This follows immediately from the formula

$$\sum\nolimits_{x_i = \pm 1, 1 \leqq i \leqq m} \prod_{j=1}^{m} \tfrac{1}{2}(1 + x_j a_j)(1, x_1, x_1, \cdots, x_m, x_m)$$
$$= \sum\nolimits_{x_i = \pm 1, 1 \leqq i \leqq m-1} \prod_{j=1}^{m-1} \tfrac{1}{2}(1 + x_j a_j)(1, x_1, x_1, \cdots, x_{m-1}, x_{m-1}, a_m, a_m)$$
$$= \cdots = (1, a_1, a_1, \cdots, a_m, a_m).$$

To show that $T \subset S$, we first note that each of the vertices of T is obviously in S, by (5). Therefore to complete the proof, we need only show that S is convex. Suppose x, y are distinct points of S, and $0 < \lambda < 1$. By (5), there exist sequences $\{a_i\}$, $\{b_i\}$ such that $1 \geqq a_1 \geqq \cdots \geqq a_m \geqq -1$, $1 \geqq b_1 \geqq \cdots \geqq b_m \geqq -1$,

$$x \, \varepsilon \, H(1, a_1, a_1, \cdots, a_m, a_m), \quad \text{and} \quad y \, \varepsilon \, H(1, b_1, b_1, \cdots, b_m, b_m).$$

By using the equivalence of (1b) and (1c) it is easy to show that

$$\lambda x + (1 - \lambda) y \, \varepsilon \, H(1, \lambda a_1 + (1-\lambda) b_1, \lambda a_1 + (1-\lambda) b_1, \cdots, \lambda a_m + (1-\lambda) b_m, \ \lambda a_m + (1-\lambda) b_m) \subset S.$$

THEOREM 9. *Let x be a real n-vector, $n \geqq 2$. The following statements are equivalent*

(6a) $\sum_{j \neq i} x_j \leqq n - 2 + x_i$, *and* $0 \leqq x_i \leqq 1, 1 \leqq i \leqq n$.

(6b) *x lies in the convex hull of B, where B is the set of those points with coordinates all* 0 *or* 1 *for which the number of* 0 *coordinates is different from one.*

(6c) *x is the diagonal of a rotation, and* $x_i \geqq 0$, $1 \leqq i \leqq n$.

(6d) *x is the diagonal of a d. s. matrix.*

Proof. Given any set A of k integers between 1 and n, there exists a permutation which leaves invariant the members of A, and only these, as long as $k \neq n - 1$. Therefore the set of diagonals of permutation matrices of order n is exactly B. Theorem 2 now shows that (6d) and (6b) are equivalent. We conclude the proof by deriving the implications (6b) $\rightarrow$ (6c) $\rightarrow$ (6a) $\rightarrow$ (6d).

(6b) $\rightarrow$ (6c) : Since the set S of diagonals of rotations is convex, we need only prove $B \subset S$. If $x \varepsilon B$, then some permutation of x has the form $(y_1, \cdots, y_n)$, in which $y_i = 0$, $1 \leqq i \leqq k$, $y_i = 1$, $k + 1 \leqq i \leqq n$, and $k \neq 1$. If k is even, then $y \varepsilon S$ by (5). (When $n = 2m$, (5) remains valid if we drop the 1 in $H(1, a_1, a_1, \cdots, a_m, a_m)$). If k is odd, then $k \geqq 3$, and using the equivalence of (1b) and (1c), it is easy to prove that $y \varepsilon H(z)$, where $z_i = 1$, $1 \leqq i \leqq n - k + 1$, $z_{n-k+2} = \frac{1}{2}$, $z_{n-k+3} = -\frac{1}{2}$, and $z_i = 0$ otherwise. Again $H(z) \subset S$ by (5).

(6c) $\rightarrow$ (6a) : Let x be the diagonal of a rotation of order n. If $n = 2m + 1$, then by (5), $x \varepsilon H(1, a_1, a_1, \cdots, a_m, a_m)$ for some sequence $\{a_i\}$ with $1 \geqq a_1 \geqq \cdots \geqq a_m \geqq -1$. Therefore by (1e),

$$\sum_{j \neq i} x_j \leqq 1 + 2a_1 + \cdots + 2a_{m-1} + a_m \leqq n - 2 + x_i.$$

The case n even is analogous.

(6a) $\rightarrow$ (6d) : Suppose x satisfies (6a). We may exclude the obvious case $x_i = 1$ for all i. Let $y_i = 1 - x_i$. Then $0 \leqq y_i \leqq 1$ and $2y_i \leqq \sum_{j=1}^{n} y_j \neq 0$. Let $\beta = \sum_{i=1}^{n} y_i$ and set $z_i = 2y_i/\beta$. Then $0 \leqq z_i \leqq 1$ and $\sum_{i=1}^{n} z_i = 2$. By the equivalence of (1b) and (1c), we have $(z_1, \cdots, z_n) \varepsilon H(1, 1, 0, \cdots, 0)$.

Therefore $(z_1, \cdots, z_n) = \sum_{1 \leqq i < j \leqq n} \alpha_{ij} u^{ij}$, where u^{ij} is the n-vector whose i-th and j-th components are 1, and whose other components are 0, and $\alpha_{ij} \geqq 0$, $\sum_{i<j} \alpha_{ij} = 1$. Therefore $(1 - x_1, \cdots, 1 - x_n) = \sum_{i<j} \gamma_{ij} u^{ij}$, where $\gamma_{ij} = \beta \cdot \alpha_{ij}/2$. If we define $P_{ii} = x_i$ and $P_{ij} = P_{ji} = \gamma_{ij}$ for $1 \leqq i \leqq j \leqq n$, we obtain a symmetric d. s. matrix $P = (P_{ij})$ with the desired property.

THEOREM 10. *For $n \geqq 2$, a real vector $(d_1, \cdots, d_n)$ is the diagonal of an orthogonal matrix if and only if the vector $x = (| d_1 |, \cdots, | d_n |)$ satisfies any one of the conditions of Theorem* 9.

Proof. Suppose x satisfies (6c). If x is the diagonal of a rotation, then d is the diagonal of an orthogonal matrix, since an orthogonal matrix remains orthogonal when we multiply any row by -1.

Now suppose d is the diagonal of an orthogonal matrix. Then at least one of the two vectors $(| d_1 |, \cdots, | d_{i-1} |, \pm d_i, | d_{i+1} |, \cdots, | d_n |)$ is the diagonal of a rotation. The proof of (6c) $\rightarrow$ (6a) did not use the hypothesis $x_i \geqq 0$. Therefore we have $\sum_{j \neq i} | d_j | \leqq n - 2 \pm d_i \leqq n - 2 + | d_i |$, which proves (6a).

THEOREM 11. *For $n \geqq 2$, a complex vector $(d_1, \cdots, d_n)$ is the diagonal of a unitary matrix if and only if the vector $x = (| d_1 |, \cdots, | d_n |)$ satisfies any one of the conditions of Theorem* 9.

Proof. Suppose x satisfies (6c). Then x is the diagonal of a rotation R. If we multiply the i-th row of R by $\arg d_i$, whenever $d_i \neq 0$, then R becomes a unitary matrix with diagonal $(d_1, \cdots, d_n)$.

To prove the converse (the idea of this proof is due to J. von Neumann), let $\alpha = \sup (| d_1 | + \cdots + | d_{n-1} | - | d_n |)$ as d varies over all diagonals of unitary matrices of order n. In order to prove (6a), we need only show $\alpha = n - 2$. Clearly $\alpha \geqq n - 2$, as is seen by considering the identity matrix. Now let U be a unitary matrix for which $| U_{11} | + \cdots + | U_{n-1\, n-1} | - | U_{nn} | = \alpha$. By multiplying each row by a proper factor, we may assume that $U_{ii} \geqq 0$, $1 \leqq i \leqq n$. By the maximal property of U, we have $U_{ii} \geqq U_{nn}$, since otherwise an interchange of rows of U would increase the value of $\sum_{i=1}^{n-1} U_{ii} - U_{nn}$. Thus it involves no loss of generality to assume $U_{11} \geqq \cdots \geqq U_{nn} \geqq 0$.

If $U_{n-1\, n-1} = 0$, then $\alpha = \sum_{i=1}^{n-2} U_{ii} \leqq n - 2$ and (6a) is proved. Henceforth assume $U_{n-1\, n-1} > 0$. For $1 \leqq p < q \leqq n$, let V^{pq} be the matrix of order n for which $V^{pq}{}_{pp} = V^{pq}{}_{qq} = \cos\theta$, $V^{pq}{}_{pq} = -\bar{V}^{pq}{}_{qp} = \sin\theta \exp(i\beta)$, where θ, β are arbitrary real numbers, and $V^{pq}{}_{ij} = \delta_{ij}$ for $i \neq p, q$, and $j \neq p, q$. If $W = UV^{12}$, we have $\sum_{i=1}^{n-1} | W_{ii} | - | W_{nn} | \leqq \sum_{i=1}^{n-1} U_{ii} - U_{nn}$. Therefore

$$(7) \quad | U_{11} \cos\theta - U_{12} \sin\theta \exp(-i\beta) | + | U_{21} \sin\theta \exp(i\beta) + U_{22} \cos\theta | \leqq U_{11} + U_{22}.$$

The expansion of the left side of (7) in powers of θ begins with

$$(U_{11} + U_{22}) + \theta[\mathfrak{R}(\exp(i\beta)(U_{21} - \bar{U}_{12}))] + \cdots.$$

Consequently the coefficient of θ must vanish for all β. This can occur only if $U_{21} = \bar{U}_{12}$. Similarly $U_{ji} = \bar{U}_{ij}$ for $1 \leqq i \leqq j \leqq n-1$. From this it follows that $| U_{in} | = | U_{ni} |$, $1 \leqq i \leqq n$.

If $U_{nn} > 0$, then by multiplying U on the right by V^{in} and using the same reasoning as before, we find $U_{ni} = -\bar{U}_{in}$, $1 \leqq i \leqq n-1$. Therefore, if we multiply the last row of U by -1, we obtain a symmetric unitary matrix B for which $B_{nn} < 0$ and such that $\alpha = \sum_{i=1}^{n} B_{ii}$.

In case $U_{nn} = 0$, let k be an integer for which $U_{kn} \neq 0$, and let B be the matrix obtained by multiplying the n-th column of U by arg (U_{kn}) and the n-th row of U by arg (U_{nk}). Then $B_{nk} = B_{kn} > 0$. If $j \neq k$, $1 \leqq j \leqq n-1$, we have

$$B_{jn}B_{kn} = -\sum_{i=1}^{n-1} B_{ji}\bar{B}_{ki} = -\sum_{i=1}^{n-1} \bar{B}_{ij}B_{ik} = \bar{B}_{nj}B_{nk}.$$

Therefore $B_{jn} = \bar{B}_{nj}$. Thus in both cases, we have obtained a symmetric unitary matrix B for which $\alpha = \sum_{i=1}^{n} B_{ii}$ and $B_{nn} \leqq 0$. But the eigenvalues of a symmetric unitary matrix are all ± 1. Also α is equal to the sum of the eigenvalues of B. Since B is not the identity matrix, B_{nn} being non-positive, the trace of B can be at most $n-2$. The proof is complete.

By using Theorem 11, we can prove many curious inequalities in the elements of a unitary matrix. For example, if we apply Theorem 11 to the product

$$\begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix} \cdot \tfrac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

where U is an arbitrary unitary matrix of order 3, we find

$$| -U_{11} + U_{12} + U_{13} | + | U_{21} - U_{22} + U_{23} | + | U_{31} + U_{32} - U_{33} | \leqq 5.$$

The bound is attained for

$$U = \tfrac{1}{3}\begin{pmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}.$$

THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] G. H. Hardy, J. E. Littlewood, and G. Polya, "*Inequalities*," Cambridge University Press, 1952.

[2] G. Birkhoff, "Three observations on linear algebra," *Universidad Nacional de Tucumán, Revista, Serie* A, vol. 5 (1946), pp. 147-151 (In Spanish).

[3] G. Polya, "Remark on Weyl's note 'Inequalities between the two kinds of eigenvalues of a linear transformation'," *Proceedings of the National Academy of Sciences*, vol. 36 (1950), pp. 49-51.

[4] Ky Fan, "Maximum properties and inequalities for the eigenvalues of completely continuous operators," *ibid.*, vol. 37 (1951), pp. 760-766.

[5] I. Schur, "Über eine klasse von mittelbildungen mit anwendungen auf der determinantentheorie," *Sitzungsberichte der Berliner Mathematischen Gesellschaft*, vol. 22 (1923), pp. 9-20.

[6] A. Horn, "On the eigenvalues of a matrix with prescribed singular values," to appear in *Proceedings of the American Mathematical Society*, vol. 5 (1954).

[7] S. Sherman, "On a theorem of Hardy, Littlewood, Polya, and Blackwell," *Proceedings of the National Academy of Sciences*, vol. 37 (1951), pp. 826-831.