Z. Wahrscheinlichkeitstheorie 1, 319—334 (1963)

Results and Problems in the Theory of Doubly-Stochastic Matrices*

By

L. Mirsky

1. Introduction

In a short but penetrating note published in 1923, Schur [57] gave a highly effective method for deriving inequalities between the characteristic roots and the diagonal elements of hermitian matrices**. Now, if $H = (h_{rs})$ is a hermitian matrix with characteristic roots $\omega_1, \ldots, \omega_n$, then there exists a unitary matrix $U = (u_{rs})$ such that

$$H = U \operatorname{diag}(\omega_1, \ldots, \omega_n)\, U^*; \tag{1.1}$$

and the starting point in Schur's argument was the observation of a simple consequence of (1.1), namely

$$\begin{pmatrix} h_{11} \\ \cdot \\ \cdot \\ \cdot \\ h_{nn} \end{pmatrix} = \begin{pmatrix} |u_{11}|^2 & \cdots & |u_{1n}|^2 \\ \cdot & \cdots & \cdot \\ |u_{n1}|^2 & \cdots & |u_{nn}|^2 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \cdot \\ \cdot \\ \cdot \\ \omega_n \end{pmatrix}. \tag{1.2}$$

The matrix $(|u_{rs}|^2)$ appearing on the right-hand side has very special properties. It is square in shape, its elements are real non-negative numbers, and the sum of the elements in each row and each column is equal to 1. Schur referred to such matrices, or more precisely to the linear substitutions associated with them, simply as 'averages' (*Mittelbildungen*) but in modern terminology they are known as *doubly-stochastic* (d.s.) matrices***.

A great deal of work on the properties of d.s. matrices has been carried out since they were first introduced into the literature, but it is probably no over-simplification to say that two fundamental theorems, proved respectively by Hardy, Littlewood, and Pólya in 1929 and by G. Birkhoff in 1946, have been dominant in this field of study. It is therefore inevitable that the present survey should be largely concerned with the territory charted by these two theorems.

---

* This survey is based on an address given to the meeting of the Stochastic Analysis Group held in Southampton in April 1962.

** Schur's investigations in this field were pursued further by A. Ostrowski [51].

*** The term 'stochastic process' (derived from στόχος = target) has come to replace the older term 'random process'. Certain matrices which occur in the theory of stochastic processes are consequently called 'stochastic matrices', and the special stochastic matrices with which we are concerned are now known as 'doubly-stochastic matrices'. As far as I am aware, the term 'doubly-stochastic' was first employed in Feller's book [18].

HARDY, LITTLEWOOD, and PÓLYA established their result by means of an elementary though somewhat involved algebraic procedure, while BIRKHOFF based his proof on P. Hall's theorem about systems of distinct representatives [*20*]. Both arguments are notable examples of ad hoc reasoning, but in recent years it has come to be recognized that the natural setting for the discussion of d.s. matrices is the theory of convex polytopes. We shall, therefore, recall briefly a few familiar definitions and results in this field. For a systematic treatment of convex sets, the reader may be referred to Eggleston's tract [*10*].

We shall be concerned with points (vectors) in real n-dimensional euclidean space E^n. The inner product of x and y will be denoted by (x, y). A point x of a convex set $\mathfrak{C}$ is said to be *extreme* (in $\mathfrak{C}$) if it does not admit a representation of the form $x = \frac{1}{2}(y + z)$, where $y, z \in \mathfrak{C}$ and $y \neq z$.

(a) A *convex polytope* is defined as the convex hull of a finite non-empty set of points, and it is easily demonstrated that a convex polytope is identical with the convex hull of its extreme points. Moreover, it is known that if the intersection of a finite number of closed half-spaces is bounded and non-empty, then it is a convex polytope.

(b) Let $\mathfrak{C}$ be the intersection of a finite number of closed half-spaces, i.e. let it be the set of all vectors x such that $(a_k, x) \leqq q_k$ $(1 \leqq k \leqq N)$, where the a's are given non-zero vectors and the q's are given scalars*. It is then easily seen that, if x_0 is an extreme point of (the convex set) $\mathfrak{C}$, then $(a_k, x) = q_k$ for at least n values of k **.

(c) Again, let $\mathfrak{C}$ be any closed convex set and let z be a point. Then there exists a plane strictly separating z from $\mathfrak{C}$ if and only if $z \notin \mathfrak{C}$. If we formulate this statement in algebraic terms and apply it to the case when $\mathfrak{C}$ is a convex polytope, we infer that z does not belong to the convex hull of the points $z_1, \ldots, z_m$ if and only if

$$(z, u) > \max_{1 \leqq k \leqq m} (z_k, u)$$

for some u. We shall refer to this result as the *separation theorem.*

(d) Finally, we mention Carathéodory's theorem. Let $\mathfrak{L}$ be an m-dimensional linear variety in E^n, and suppose that $\mathfrak{X}$ is a subset of $\mathfrak{L}$. Then any point in the convex hull of $\mathfrak{X}$ lies in the convex hull of at most $m + 1$ suitable points of $\mathfrak{X}$.

2. The theorems of Hardy, Littlewood, and Pólya

Throughout our discussion, Δ_n will stand for the set of all d.s. $n \times n$ matrices. Next, let $\mathfrak{S}_n$ denote the set of all permutations on the symbols $1, 2, \ldots, n$. If $x = (x_1, \ldots, x_n)$ and $\pi \in \mathfrak{S}_n$, we shall write $x_\pi = (x_{\pi 1}, \ldots, x_{\pi n})$. Again, $\mathfrak{H}(x)$ denotes the convex hull of the $n!$ vectors x_π $(\pi \in \mathfrak{S}_n)$.

Let $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ be real vectors; denote by $x_1^*, \ldots, x_n^*$

* Since the inequalities $(a, x) \leqq q$ and $(-a, x) \leqq -q$ are together equivalent to $(a, x) = q$, it follows that the set of relations defining $\mathfrak{C}$ can also contain equalities.

** In fact, more is true: the point $x_0 \in \mathfrak{C}$ is extreme if and only if the set of vectors a_k such that $(a_k, x_0) = q_k$ contains n linearly independent vectors.

the numbers $x_1, \ldots, x_n$ arranged in non-ascending order of magnitude; and let $y_1^*, \ldots, y_n^*$ be defined analogously. If the relations

$$y_1^* + \cdots + y_k^* \leqq x_1^* + \cdots + x_k^* \qquad (1 \leqq k \leqq n) \tag{2.1}$$

are satisfied, we shall write $y \preccurlyeq x$. If, in addition, there is equality in (2.1) for $k = n$, we shall write $y \prec x$.

The principal result of HARDY, LITTLEWOOD, and PÓLYA on d.s. matrices can now be stated [*23*, *24* (Theorem 46); cf. also *1*, *14*, *36*, *43*, *51*, *53*].

Theorem 1a. *Let x, y be any real vectors with n components. Then the following statements are equivalent.* (i) $y \prec x$; (ii) $y \in \mathfrak{H}(x)$; (iii) $y = Dx$ *for some* $D \in \Delta_n$.

In fact, the theorem as originally enunciated asserted only the equivalence of (i) and (iii). Clause (ii) was added much later by R. RADO [*53*] who appears to have been the first mathematician to make explicit use of results on convex sets in the discussion of d.s. matrices.

The implication (ii) $\Rightarrow$ (iii) is immediate, and (iii) $\Rightarrow$ (i) is entirely straightforward. The crux of the argument thus lies in the proof of the implication (i) $\Rightarrow$ (ii). Assume, then, that for certain vectors x and y, (i) is true while (ii) is false. Since $y \notin \mathfrak{H}(x)$, it follows by the separation theorem (§ 1, c) that there exists a vector $u = (u_1, \ldots, u_n)$ such that

$$(y, u) > \max_{\pi \in \mathfrak{S}_n} (x_\pi, u),$$

i.e.

$$\sum_{k=1}^{n} y_k u_k > \max_{\pi \in \mathfrak{S}_n} \sum_{k=1}^{n} x_{\pi k} u_k = \sum_{k=1}^{n} x_k^* u_k^*,$$

and therefore (cf. [*24*], Theorem 368)

$$\sum_{k=1}^{n} y_k^* u_k^* > \sum_{k=1}^{n} x_k^* u_k^*. \tag{2.2}$$

On the other hand, in view of (i), we have

$$\begin{aligned}\sum_{k=1}^{n} y_k^* u_k^* &= \sum_{k=1}^{n-1} (u_k^* - u_{k+1}^*)(y_1^* + \cdots + y_k^*) + u_n^*(y_1^* + \cdots + y_n^*) \\ &\leqq \sum_{k=1}^{n-1} (u_k^* - u_{k+1}^*)(x_1^* + \cdots + x_k^*) + u_n^*(x_1^* + \cdots + x_n^*) \\ &= \sum_{k=1}^{n} x_k^* u_k^*,\end{aligned}$$

and this contradicts (2.2). The proof is therefore complete*.

We may strengthen the result just established by replacing Δ_n in (iii) by the set of 'orthostochastic' $n \times n$ matrices, i.e. matrices of the type (a_{rs}^2), where (a_{rs}) is an orthogonal matrix** [*27*, *44*].

Further, it may be mentioned that there is a rather obvious analogue of Theorem 1a for stochastic matrices***. A more interesting result can be proved

* This argument is due to RADO [*53*].

** It is easy to verify that the set of orthostochastic matrices is properly contained in the set of d.s. matrices.

*** A square matrix is called *stochastic* if its elements are non-negative and the sum of the elements in each row is equal to 1. It will be recalled that the study of these matrices figures prominently in the theory of Markov chains.

for doubly-substochastic (d.s.s.) matrices, i.e. square matrices with real non-negative elements none of whose row-sums or column-sums exceeds 1. Write $a^+ = \max(a, 0)$ and, when $x = (x_1, \dots, x_n)$, put $x^+ = (x_1^+, \dots, x_n^+)$. If x, y are any real vectors with n components, then there exists a d.s.s. $n \times n$ matrix E such that $y = Ex$ if and only if $y \prec x^+$ and $-y \prec (-x)^+$ ([*45*] cf. also [*11*]).

An unsolved and difficult problem* is concerned with the extension of Theorem 1a to the case of *complex* vectors: if z and w are given complex vectors, what conditions are necessary and sufficient for the existence of a d.s. matrix D such that $w = Dz$? We encounter this problem when we seek to establish criteria for the existence of a normal matrix with prescribed diagonal elements and characteristic roots.

In recent years, the theorem of HARDY, LITTLEWOOD, and PÓLYA has been assimilated into much more general investigations. Thus it emerges as a corollary of a result of FAN [*13*] concerning convex functions defined in topological vector spaces. This is not altogether surprising, for there is a natural link between convex functions and d.s. matrices. This was recognized by HARDY, LITTLEWOOD, and PÓLYA who proved the following result ([*23*]; [*24*] (Theorem 108)), the germ of which is already contained in Schur's paper [*57*].

Theorem 1b. *Let x_k, y_k $(1 \leqq k \leqq n)$ be given real numbers. Then the inequality*

$$\Phi(y_1) + \cdots + \Phi(y_n) \leqq \Phi(x_1) + \cdots + \Phi(x_n) \tag{2.3}$$

*is valid for every convex function** Φ if and only if*

$$(y_1, \dots, y_n) \prec (x_1, \dots, x_n). \tag{2.4}$$

It will be observed that the *statement* of this theorem does not involve d.s. matrices. They make their appearance in the proof, and this seems a natural mode of argument, though it should be noted that there also exist proofs depending on quite different ideas [*19, 30*].

The necessity of the condition (2.4) becomes evident if we apply (2.3) in turn to the functions

$$\Phi(t) = t, \quad -t, \quad (t - x_k)^+ \qquad (1 \leqq k \leqq n).$$

To demonstrate its sufficiency, we note that, by Theorem 1a, (2.4) implies the existence of a d.s. matrix $D = (d_{rs})$ such that

$$y_r = d_{r1} x_1 + \cdots + d_{rn} x_n \qquad (1 \leqq r \leqq n).$$

Hence, for any convex function Φ,

$$\Phi(y_r) \leqq d_{r1} \Phi(x_1) + \cdots + d_{rn} \Phi(x_n) \qquad (1 \leqq r \leqq n);$$

and (2.3) follows if we sum for $1 \leqq r \leqq n$.

Numerous modifications of this result are possible. For example, it was shown by PÓLYA [*52*] that the theorem remains valid if (2.4) is weakened to

$$(y_1, \dots, y_n) \preccurlyeq (x_1, \dots, x_n),$$

* See, however, the papers of HORN [*27*] and SHERMAN [*60*].

** The function Φ is said to be convex if, for any numbers x, y and any positive numbers λ, μ with sum 1, we have

$$\Phi(\lambda x + \mu y) \leqq \lambda \Phi(x) + \mu \Phi(y).$$

provided that we restrict ourselves to the class of convex *non-decreasing* functions. Again, instead of considering

$$\Phi(t_1) + \cdots + \Phi(t_n),$$

we can admit more general classes of functions of $t_1, \ldots, t_n$ and obtain several variants of Theorem 1b (see [*46*]). Thus, for instance, the inequality $F(y) \leqq F(x)$ is valid for every function $F(x) = F(x_1, \ldots, x_n)$, which is symmetric with respect to the x_k and convex with respect to the vector variable x, if and only if $y \prec x$. There are many results more loosely linked with Theorem 1b. We content ourselves with a mere mention of inequalities involving 'Schur-convex' functions* [*51, 57*] and of results on 'symmetric gauge functions' and on 'unitarily invariant matrix norms'** [*11, 15, 47*].

Theorem 1b is the source of a whole series of matrix inequalities, of which we give a single example. Let $H = (h_{rs})$ be a hermitian matrix with characteristic roots $\omega_1, \ldots, \omega_n$. Then, by (1.2),

$$(h_{11}, \ldots, h_{nn})^T = D(\omega_1, \ldots, \omega_n)^T,$$

where $D \in \Delta_n$. Hence, by Theorem 1a,

$$(h_{11}, \ldots, h_{nn}) \prec (\omega_1, \ldots, \omega_n)$$

and so, by Theorem 1b,

$$\Phi(h_{11}) + \cdots + \Phi(h_{nn}) \leqq \Phi(\omega_1) + \cdots + \Phi(\omega_n),$$

where Φ is any convex function. In particular, if H is positive definite, then, for any real number $p > 1$, we have***

$$h_{11}^p + \cdots + h_{nn}^p \leqq \omega_1^p + \cdots + \omega_n^p. \tag{2.5}$$

In view of Pólya's modification of Theorem 1b mentioned above, this inequality can be extended: if $\omega_1 \geqq \cdots \geqq \omega_n$ and $p > 1$, then, for $1 \leqq k \leqq n$, we have

$$h_{11}^p + \cdots + h_{kk}^p \leqq \omega_1^p + \cdots + \omega_k^p.$$

3. Birkhoff's theorem

The simplest d.s. matrices are, of course, the permutation matrices, and it is natural to conjecture that they are cast for a special role in the theory we are describing. That this is, indeed, the case is demonstrated by Birkhoff's theorem [*3*].

Theorem 2. *The set Δ_n of doubly-stochastic $n \times n$ matrices is identical with the convex hull of the set of $n \times n$ permutation matrices.*

It is remarkable that so striking and intuitively so simple a result was not discovered till 1946. There exist now, in addition to Birkhoff's original treatment, several other proofs [*2, 9, 22, 26, 43, 50, 61*]. Most of these draw upon combina-

---

* A function F of a vector variable is said to be *Schur-convex* if, for every vector x and every d.s. matrix D, we have $F(Dx) \leqq F(x)$. The class of convex, symmetric functions is contained in the class of Schur-convex functions.

** For definitions of these terms, see e.g. [*47*].

*** The inequality (2.5) was proved in a different way by SCHUR [*57*].

torial or upon geometric ideas, or upon both. In the combinatorial proofs, the essential step often consists in the demonstration that, for any d.s. $n \times n$ matrix $D = (d_{rs})$, there is a permutation $\pi \in \mathfrak{S}_n$ such that $d_{r,\pi r} > 0$ $(1 \leqq r \leqq n)^\star$; from this Theorem 2 follows readily by induction with respect to the number of strictly positive elements of D (see e.g. [*9*]). However, the proof we shall indicate below (due to Hoffman and Wielandt [*26*]) is of the geometric type and depends on properties of convex sets.

We shall interpret real $n \times n$ matrices as points in E^{n^2}. The matrix $X = (x_{rs})$ belongs to Δ_n precisely if

$$x_{rs} \geqq 0 \qquad (1 \leqq r, s \leqq n), \tag{3.1}$$

$$\sum_{s=1}^{n} x_{rs} = 1 \qquad (1 \leqq r \leqq n), \tag{3.2}$$

$$\sum_{r=1}^{n} x_{rs} = 1 \qquad (1 \leqq s \leqq n-1), \tag{3.3}$$

since the relation $x_{1n} + \cdots + x_{nn} = 1$ is a consequence of (3.2) and (3.3). Thus Δ_n is the non-empty, bounded intersection of a finite number of closed half-spaces and so, by § 1 (a), is a convex polytope. Theorem 2 will therefore follow if we can show that the extreme points of Δ_n are precisely the permutation matrices. One half of this statement is trivial, for every permutation matrix is clearly extreme in Δ_n. To prove the converse, suppose that X is extreme in Δ_n. Then, by § 1 (b), it follows that equality must hold in at least n^2 of the relations (3.1), (3.2), and (3.3), and so in at least $n^2 - 2n + 1$ of the relations (3.1). This implies that at least one row of X must consist of $n - 1$ zeros and one unit. In the column containing this unit, all other elements must be equal to zero. We are thus able to reduce our problem to the consideration of Δ_{n-1}; and the proof is now easily completed by induction with respect to n.

The theorem just discussed shows that every d.s. matrix can be expressed as a convex combination of permutation matrices. But the representation of any one d.s. matrix does not require all $n!$ permutation matrices. In fact, the set Δ_n lies in a linear variety of dimension $(n-1)^2$ in E^{n^2}. Hence, by Carathéodory's theorem (§ 1, d), any d.s. $n \times n$ matrix belongs to the convex hull of at most $(n-1)^2 + 1$ suitable permutation matrices, and it is not difficult to show that this result is best possible [*17, 21, 40*]. However, when additional information is given, the number $(n-1)^2 + 1$ can be diminished. If $D \in \Delta_n$, let $\nu(D)$ denote the least number of permutation matrices which contain D in their convex hull. Then, as we have seen,

$$\nu(D) \leqq (n-1)^2 + 1$$

and here the sign '$\leqq$' cannot, in general, be replaced by '$<$'. Now it is plain that all characteristic roots of D lie on the unit disk $|z| \leqq 1$. We shall denote by $c = c(D)$ the number of characteristic roots on the unit circle $|z| = 1$ (so that $c \geqq 1$ since 1 is a characteristic root of every d.s. matrix.) It was shown by

⋆ This is an easy consequence of Hall's theorem [*20*]. Cf. also [*33*].

Marcus, Minc, and Moyls [37] that $\nu(D)$ is small when $c(D)$ is large. More precisely, if D is indecomposable*, then

$$\nu(D) \leqq c\left(\frac{n}{c} - 1\right)^2 + 1 . \tag{3.4}$$

This relation and other, more precise, estimates due to the same authors go some way towards the solution of the problem of characterization of the set of d.s. $n \times n$ matrices D for which $\nu(D)$ has a prescribed value**.

One consequence of Birkhoff's theorem is that any function of a matrix variable, defined and convex on Δ_n, assumes its maximum for a permutation matrix. This principle was stressed by Marcus who used it to obtain inequalities in matrix theory [35] and also to derive afresh known inequalities [36]. For example, he gave a proof, on these lines, of an extremal property of hermitian matrices due to Fan [12]. Again, Hoffman and Wielandt [26] used the same principle to show that, if A, B are normal matrices with characteristic roots $\{\alpha_k\}$, $\{\beta_k\}$ respectively, then, for a suitable numbering of the roots,

$$\|A - B\|^2 \geqq \sum_{k=1}^{n} |\alpha_k - \beta_k|^2 ,$$

where $\|.\|$ denotes the euclidean norm.

Not surprisingly, there are analogues of Theorem 2 for stochastic and also for d.s.s. matrices [17, 45]. The latter result is an easy consequence of the fact, noted by Horn [27], that any d.s.s. $n \times n$ matrix E can be exhibited as a submatrix of a suitable d.s. $N \times N$ matrix D, where $N \leqq 2n$. Since D is a convex combination of $N \times N$ permutation matrices, it follows that E is a convex combination of $n \times n$ sub-permutation matrices***. We infer, therefore, that the set of d.s.s. $n \times n$ matrices is identical with the convex hull of the set of $n \times n$ sub-permutation matrices.

Various writers considered extensions of Birkhoff's theorem. Thus Mendelsohn and Dulmage [42] determined the convex hull of the set of all sub-permutation matrices which possess exactly r non-vanishing elements. An as yet unproved generalization of Birkhoff's theorem has been proposed by Révész [55]. Let $\lambda_1, \ldots, \lambda_n$ be non-negative numbers with sum 1 and denote by Λ_n the set of all $n \times n$ matrices (x_{rs}) with real non-negative elements such that

$$\sum_{s=1}^{n} x_{rs} = 1 \quad (1 \leqq r \leqq n), \qquad \sum_{r=1}^{n} \lambda_r x_{rs} = \lambda_s \quad (1 \leqq s \leqq n).$$

---

* The square matrix A is said to be *indecomposable* if there exists no relation of the form

$$P^T A P = \begin{pmatrix} A_1 & O \\ A_2 & A_3 \end{pmatrix},$$

where P is a permutation matrix and A_1, A_3 are square matrices. The condition of indecomposability in (3.4) is not a serious restriction since it can be shown that, for any d.s. matrix D, there exists a permutation matrix P such that $P^T D P$ is the direct sum of indecomposable d.s. matrices.

** For the determination of an upper bound of $\nu(D)$ in terms of a graph associated with D, see [29].

*** A *sub-permutation matrix* is a square matrix in which at most one element in each row and in each column is equal to 1 while all other elements are equal to zero.

It is then required to prove that the convex polytope Λ_n possesses at most $n!$ extreme points. Theorem 2 corresponds to the case $\lambda_1 = \cdots = \lambda_n = n^{-1}$ of this conjecture.

We next turn to a different kind of refinement of Birkhoff's theorem. Suppose that, instead of considering the entire symmetric group of $n!$ permutation matrices, we restrict our attention to some subgroup $\mathfrak{G}$ and seek to determine the convex hull of the set of permutation matrices in $\mathfrak{G}$. Our aim, then, is to link up the multiplicative and the linear structure of $\mathfrak{G}$, and this problem turns out to be unexpectedly difficult. Even for the case of the alternating group, the answer is not known. The situation is equally obscure with regard to the corresponding extension of Theorem 1a of HARDY, LITTLEWOOD, and PÓLYA. Here the problem is to determine necessary and sufficient conditions for the existence of a (d.s.) matrix D belonging to the convex hull of $\mathfrak{G}$ and such that $y = Dx$, where x and y are given vectors. In view of the intractable nature of these questions, it may be worth while to investigate the easier problem of characterization of diagonal elements of d.s. matrices. For the group of *all* permutations the problem was solved by HORN [*27*]: the numbers $x_1, \dots, x_n$ are the diagonal elements of some d.s. $n \times n$ matrix if and only if

$$0 \leqq x_1, \dots, x_n \leqq 1 \tag{3.5}$$

and

$$\sum_{k=1}^{n} x_k - 2 \min_{1 \leqq j \leqq n} x_j \leqq n - 2\,. \tag{3.6}$$

Next, consider the case of the alternating group. Let a d.s. matrix be called *even* if it belongs to the convex hull of permutation matrices associated with even permutations. Making use of the separation theorem, it is not difficult to show [*49*] that $x_1, \dots, x_n$ are the diagonal elements of some even d.s. $n \times n$ matrix if and only if they satisfy (3.5) and the relation

$$\sum_{k=1}^{n} x_k - 3 \min_{1 \leqq j \leqq n} x_j \leqq n - 3\,. \tag{3.7}$$

Suppose, now, that $D = (d_{rs})$ is an even d.s. matrix; let π be an even permutation; and denote by P_π the permutation matrix corresponding to π *. Then DP_π is again an even d. s. matrix and so, by (3.7),

$$\sum_{k=1}^{n} (DP_\pi)_{kk} - 3 \min_{1 \leqq j \leqq n} (DP_\pi)_{jj} \leqq n - 3;$$

in other words

$$\sum_{k=1}^{n} d_{k,\pi k} - 3 \min_{1 \leqq j \leqq n} d_{j,\pi j} \leqq n - 3\,. \tag{3.8}$$

It is natural to attempt to invert this conclusion by inquiring whether a d.s. matrix D which satisfies (3.8) for every even permutation π is necessarily even, but it is not known whether this is the case.

---

* By this we mean that the (r, s)-th element of P_π is equal to $\delta_{r,\pi s}$.

4. Infinite and continuous analogues of the fundamental theorems

In his book on *Lattice Theory* [*4*, p. 266], Birkhoff proposed the problem of extending his theorem to the set Δ of (denumerably) infinite d.s. matrices. This question has acquired some celebrity as 'Birkhoff's problem 111'. It may be noted, in the first place, that if $D = (d_{rs}) \in \Delta$, then a combinatorial theorem due to de Bruijn ([*6*]; cf. also Isbell [*28*]) guarantees the existence of a permutation π of the positive integers such that $d_{r,\pi r} > 0$ $(r = 1, 2, \ldots)$. In contrast to the finite case, however, this result does not lead to a solution of Birkhoff's problem since the possibility that $\inf_r d_{r,\pi r}$ is zero cannot be discounted.

Indeed, before the problem can be treated successfully, it needs to be formulated more precisely. Now, in the transition from the finite to the infinite case, the purely algebraic notion of a convex hull is naturally superseded by the topological notion of convex closure. Let $\mathfrak{B}$ be a vector space of infinite matrices such that $\Delta \subsetneqq \mathfrak{B}$, and denote by $\mathfrak{P}$ the set of infinite permutation matrices. It is then required to prove that, for a suitable choice of $\mathfrak{B}$ and a suitable topology $\mathscr{T}$ on $\mathfrak{B}$ (which is compatible with the linear structure of $\mathfrak{B}$), the set Δ is the convex closure of $\mathfrak{P}$ under $\mathscr{T}$.

This problem was first discussed by Isbell [*28*] who considered the space $\mathfrak{B}$ of 'boundedly line-summable' infinite real matrices, i.e. the space of matrices such that $X \in \mathfrak{B}$ if

$$\sup_{r \geqq 1} \sum_{s=1}^{\infty} |x_{rs}| < \infty, \qquad \sup_{s \geqq 1} \sum_{r=1}^{\infty} |x_{rs}| < \infty.$$

A solution was given by Rattray and Peck [*54*] and, independently, by Kendall [*32*]. Kendall's choice of $\mathfrak{B}$ is identical with that of Isbell; his treatment depends on a separation theorem for locally convex spaces [*5*, pp. 73—74] and the topology used by him is the weakest (locally convex Hausdorff) topology which makes all elements, row-sums, and column-sums of matrices continuous as linear functionals on $\mathfrak{B}$.

There still remain some obvious questions. For example, how can one characterize those infinite d.s. matrices which can be expressed as finite (or as denumerably infinite) convex combinations of permutation matrices?

For vector spaces of infinitely many dimensions we possess, just as for finite-dimensional vector spaces, the notion of extreme points. It is plain that every permutation matrix is extreme in Δ. Making use of ideas connected with the Krein-Mil'man theorem [*5*, p. 84], Kendall and Kiefer [*32*] established the converse proposition by showing that every extreme point in Δ is a permutation matrix. This is a purely algebraic result and it is very satisfactory that an algebraic proof of this and of related results has been recently discovered by Mauldon [*41*]. A divergence between the finite and the infinite case is, however, to be noted. From the fact that the permutation matrices are the only extreme points in Δ_n, Birkhoff's (finite) theorem follows at once. On the other hand, the result of Kendall and Kiefer does not imply the infinite analogue of Birkhoff's theorem.

As a footnote to the work on Birkhoff's problem 111, we mention a characterization of diagonal elements [*48*]. According to a result stated in § 3, the numbers $x_1, \ldots, x_n$ are the diagonal elements of some d.s. $n \times n$ matrix if and only if they satisfy (3.5) and (3.6); and it is plain that the latter condition can be rewritten in the form

$$2(1 - \min_{1 \leqq j \leqq n} x_j) \leqq \sum_{k=1}^{n} (1 - x_k).$$

It is not difficult to extend this result by proving that $x_1, x_2, x_3, \ldots$ are the diagonal elements of an infinite d.s. matrix if and only if $0 \leqq x_k \leqq 1$ $(k = 1, 2, \ldots)$ and

$$2(1 - \inf_{j \geqq 1} x_j) \leqq \sum_{k=1}^{\infty} (1 - x_k) \leqq \infty.$$

The introduction of infinite d.s. matrices raises the question whether there exists an infinite analogue of Theorem 1a of HARDY, LITTLEWOOD, and PÓLYA. Let $x = (x_1, x_2, \ldots)$, $y = (y_1, y_2, \ldots)$ be real vectors with infinitely many components and, for $k \geqq 1$, write

$$M_k(x) = \sup (x_{i_1} + \cdots + x_{i_k}),$$

where the upper bound is taken with respect to all sets of k distinct positive integers $i_1, \ldots, i_k$. Suppose that, for some matrix $D \in \Delta$,

$$y = Dx. \tag{4.1}$$

It is easy to deduce that, if the sequence $\{x_k\}$ is bounded, then

$$M_k(y) \leqq M_k(x) \qquad (k = 1, 2, \ldots). \tag{4.2}$$

Furthermore, we have

$$\sum_{k=1}^{\infty} y_k = \sum_{k=1}^{\infty} x_k, \tag{4.3}$$

provided that the series on the right-hand side converges absolutely. It is harder to decide whether the converse inference is valid, i.e. whether the relations (4.2) and (4.3) imply the existence of a matrix $D \in \Delta$ satisfying (4.1).

We next turn to consider the possibility of continuous analogues, and here Theorem 1b is a natural starting point since the replacement of sums by integrals at once suggests itself. However, before we can pursue this line of thought, we need some preliminary definitions.

Let $\mathfrak{M}$ denote the class of functions bounded and measurable on the unit interval. Two functions $f, g \in \mathfrak{M}$ are said to be *equimeasurable* if, for every (finite or infinite) interval I, the measures of the two sets

$$\{t : 0 \leqq t \leqq 1, f(t) \in I\}, \quad \{t : 0 \leqq t \leqq 1, g(t) \in I\}$$

are equal. It can be easily shown [*24*, pp. 276—277] that, associated with every function $f \in \mathfrak{M}$, there is a function $f^* \in \mathfrak{M}$ which is non-increasing and equimeasurable with f *. Moreover, f^* is essentially unique**. For any functions $f, g \in \mathfrak{M}$, we shall write $g \prec f$ if

$$\int_0^u g^*(t)\,dt \leqq \int_0^u f^*(t)\,dt$$

for $0 \leqq u \leqq 1$ and if this relation reduces to an equality for $u = 1$. HARDY, LITTLEWOOD, and PÓLYA [23] obtained an analogue of Theorem 1b which states that, for $f, g \in \mathfrak{M}$, the inequality

$$\int_0^1 \Phi(g(t))\,dt \leqq \int_0^1 \Phi(f(t))\,dt \tag{4.4}$$

is valid for every convex function Φ if and only if $g \prec f$. As our knowledge of infinite d.s. matrices is as yet slight, we cannot establish this result by adapting the argument used in proving Theorem 1b. However, the proof does not present any serious difficulties if we make use of the fact that, in a finite interval, a convex function can be approximated uniformly by the sum of a linear function and a finite number of positive multiples of functions of type $(t - c)^+$ ***. For further theorems related to the inequality (4.4), we refer the reader to papers by LORENTZ [34] and FAN and LORENTZ [16], and for an alternative continuous analogue of Theorem 1b to the paper of KARAMATA [30].

The possibility of establishing a continuous analogue of Theorem 1a appears to be more elusive. We naturally wish to replace d.s. matrices by suitable operators on $\mathfrak{M}$. Now it is easily verified that a matrix D is d.s. if and only if it satisfies the following conditions: (i) $x \geqq 0$ implies $Dx \geqq 0$, where inequalities between vectors are interpreted component-wise; (ii) $Du = u$, where $u = (1, \ldots, 1)^T$; (iii) $\sigma(Dx) = \sigma(x)$ for every x, where $\sigma(x)$ denotes the sum of components of x. This characterization of d.s. matrices suggests that an operator $\mathscr{D}$ on $\mathfrak{M}$ should be called *doubly-stochastic* if (a) $\mathscr{D}$ maps $\mathfrak{M}$ linearly into itself; (b) $f \geqq 0$, $f \in \mathfrak{M}$ imply $\mathscr{D}f \geqq 0$; (c) $f \equiv 1$ implies $\mathscr{D}f \equiv 1$; (d) $f \in \mathfrak{M}$ implies

$$\int_0^1 \mathscr{D}f(t)\,dt = \int_0^1 f(t)\,dt.$$

We can now formulate a conjectured analogue of Theorem 1a: if $f, g \in \mathfrak{M}$, then there exists a d.s. operator $\mathscr{D}$ such that $g = \mathscr{D}f$ if and only if $g \prec f$.

The set of d.s. operators is convex if multiplication by scalars and addition are defined in the obvious way. It is natural to extend the notion of a permutation matrix by calling a d.s. operator a *permutator* if it transforms every function in $\mathfrak{M}$ into an equimeasurable function. This definition at once prompts several questions. We content ourselves with mentioning the most interesting of these. Does every d.s. operator belong to the convex closure, in a suitable sense, of the set of permutators?

---

* Thus the statement 'the functions f and g are equimeasurable' is analogous to the statement 'the vectors x and y can be obtained from each other by a permutation of components'. Moreover, the relation between f and f^* is analogous to that between the vectors $(x_1, \ldots, x_n)$ and $(x_1^*, \ldots, x_n^*)$.

** I owe this observation to Dr. H. BURKILL.

*** Needless to say, an alternative proof of Theorem 1b can also be based on this principle.

5. Further problems

In this final section we shall give a brief account of some properties of d.s. matrices which do not fall within the ambit of the two fundamental theorems.

5.1. Van der Waerden's problem. The *permanent*, per A, of the $n \times n$ matrix $A = (a_{rs})$ is defined by the equation

$$\operatorname{per} A = \sum_{\pi \in \mathfrak{S}_n} a_{1,\pi 1} \dots a_{n,\pi n}.$$

More than thirty-five years ago, VAN DER WAERDEN [62] proposed the problem of determining the minimum of per D as D ranges over Δ_n. It has been conjectured that, for $D \in \Delta_n$,

$$\operatorname{per} D \geqq n!\, n^{-n}, \tag{5.1}$$

with equality if and only if $D = J_n$, where J_n denotes the $n \times n$ matrix all of whose elements are equal to n^{-1}. Until recently this problem received no attention and it still remains unsolved, but in the last few years several partial results have been secured.

If (5.1) is valid, then

$$d_{1,\pi 1} \dots d_{n,\pi n} \geqq n^{-n} \text{ for some } \pi \in \mathfrak{S}_n \tag{5.2}$$

and therefore, in view of the inequality of the arithmetic and geometric means, we have, for some $\pi \in \mathfrak{S}_n$,

$$d_{1,\pi 1} + \cdots + d_{n,\pi n} \geqq 1; \quad d_{1,\pi 1} > 0, \dots, d_{n,\pi n} > 0. \tag{5.3}$$

The relation (5.2) has not been proved*, but MARCUS and REE [40] obtained results stronger than (5.3). They showed, for example, that if k characteristic roots of D lie on the unit circle, then, for some $\pi \in \mathfrak{S}_n$,

$$d_{1,\pi 1} + \cdots + d_{n,\pi n} \geqq k; \quad d_{1,\pi 1} > 0, \dots, \quad d_{n,\pi n} > 0.$$

Another result closely related to van der Waerden's problem was found by MARCUS and NEWMAN [38]. Let Δ_n^+ denote the set of all d.s. $n \times n$ matrices with strictly positive elements. Then, if per D attains its lower bound in Δ_n^+, it does so for $D = J_n$ only. More recently, the same authors [39] obtained a host of new inequalities for the permanent by observing that if $x_1, \dots, x_k$ and $y_1, \dots, y_k$ are any vectors in n-dimensional unitary space $\mathfrak{U}$ with inner product $(.,.)$, then per $((x_i, y_j))$ can be interpreted as an inner product on a certain space associated with $\mathfrak{U}$. Hence, by Schwarz's inequality

$$|\operatorname{per}((x_i, y_j))|^2 \leqq \operatorname{per}((x_i, x_j)) \cdot \operatorname{per}((y_i, y_j));$$

and this relation gives rise to a variety of results. For example, let $\tilde{\Delta}_n$ denote the set of all symmetric positive definite or positive semi-definite d.s. $n \times n$ matrices. Then (5.1) holds for all $D \in \tilde{\Delta}_n$ with equality if and only if $D = J_n$.

In spite of recent progress, most problems concerning the permanent are still unsolved. We mention three of these problems, formulated by MARCUS and NEWMAN [39]. Does the inequality

$$\operatorname{per}(AB) \leqq \min(\operatorname{per} A, \operatorname{per} B)$$

* Compare the addendum at the end of the paper.

hold for all $A, B \in \Delta_n$? If $D \in \widetilde{\Delta}_n$, is the following counterpart to Hadamard's determinantal inequality valid:

$$\operatorname{per} D \geqq d_{11} d_{22} \ldots d_{nn} ?$$

Again, does the inequality $\operatorname{per}(I - D) \geqq 0$ hold for all $D \in \Delta_n$? MARCUS and NEWMAN [*39*] showed that it certainly holds when D is d.s. and symmetric.

5.2. Distribution of characteristic roots. Our knowledge of the spectral properties of d.s. matrices is very scanty. The most obvious question that arises in this context is concerned with the determination of the region Ω_n of the complex plane which is specified by the requirement that $z \in \Omega_n$ if and only if z is a characteristic root of some matrix in Δ_n. The analogous problem for stochastic matrices was raised in 1938 by A. N. KOLMOGOROFF, investigated by DMITRIEV and DYNKIN [*7*, *8*], and finally solved by KARPELEVICH [*31*]. Both the statement and the proof of Karpelevich's result are complicated; the problem of specifying Ω_n is probably more tractable but it has not yet been solved or even discussed in the literature. However, a partial result can be obtained quite easily as follows*.

For $k \geqq 2$, denote by Π_k the closed, finite region bounded by the regular k-gon whose vertices are the points θ^m $(m = 0, 1, \ldots, k-1)$, where $\theta = e^{2\pi i/k}$. It is clear that, if $2 \leqq k \leqq n$, then θ is a characteristic root of some $n \times n$ permutation matrix, say P. Now any number $z \in \Pi_k$ can be expressed in the form

$$z = t_0 + t_1 \theta + t_2 \theta^2 + \cdots + t_{k-1} \theta^{k-1},$$

where $t_0, t_1, \ldots, t_{k-1}$ are suitable non-negative numbers with sum 1. It follows that z is a characteristic root of the d.s. $n \times n$ matrix

$$t_0 I + t_1 P + t_2 P^2 + \cdots + t_{k-1} P^{k-1}.$$

Thus $\Pi_k \subseteqq \Omega_n$ and so

$$\Pi_2 \cup \Pi_3 \cup \ldots \cup \Pi_n \subseteqq \Omega_n. \tag{5.4}$$

Whether the sign of inclusion in this relation can be replaced by the sign of equality remains at present an open question.

Denote by Ω the union of all Ω_n, so that $z \in \Omega$ if and only if z is a characteristic root of some d.s. matrix. Then (5.4) implies that the whole interior of the unit disk belongs to Ω while, of course, Ω is contained in the closed unit disk. Finally, it follows at once from the work of DMITRIEV and DYNKIN [*7*] that the only points of Ω on the unit circle are the 'rational' points $e^{2\pi i a/b}$, where a, b are any integers.

Needless to say, the much harder problem of finding necessary and sufficient conditions for a set of n complex numbers to be the set of characteristic roots of some matrix in Δ_n is at present completely inaccessible.

5.3. Miscellaneous problems. In conclusion, we enumerate a few unsolved problems.

(i) If x, y are given real vectors, what conditions are necessary and sufficient for the existence of a non-singular d.s. matrix D such that $y = Dx$?

---

* This result has been communicated to me by Dr. J. F. C. KINGMAN.

(ii) If x, y are given real vectors, what conditions are necessary and sufficient for the existence of a symmetric d.s. matrix D such that $y = Dx$? (Cf. HOFFMAN [*25*]).

(iii) What is the convex hull of all $n \times n$ permutation matrices other than the unit matrix?

(iv) Is there a convenient way of characterizing those d.s. matrices which are unitarily similar to diagonal matrices?

(v) Is the set of orthostochastic $n \times n$ matrices everywhere dense, with respect to the euclidean norm, in Δ_n?

(vi) It was conjectured by S. KAKUTANI (see SHERMAN [*58*]) that, if $X, Y \in \Delta_n$ and $Xu \prec Yu$ for every real vector u, then there exists a matrix $D \in \Delta_n$ such that $X = DY$. This conjecture was shown [*59*] to be generally incorrect for $n \geqq 4$, but SCHREIBER [*56*] proved that it is valid when Y is non-singular. The question of a convenient condition which, in the general case, would ensure the existence of the requisite matrix D is still open.

(vii) It will be recalled that, for any infinite d.s. matrix (d_{rs}), there exist a permutation π of the positive integers such that $d_{r,\pi r} > 0$ $(r = 1,2,\ldots)$. J. R. ISBELL has raised the question whether this result can be proved without recourse to the axiom of choice.

(viii) Let $f(x, y)$ be a non-negative and continuous function defined on the unit square, which satisfies the relations

$$\int_0^1 f(x,y)\,dy = 1 \; (0 \leqq x \leqq 1)\,, \quad \int_0^1 f(x,y)\,dx = 1 \; (0 \leqq y \leqq 1)\,.$$

P. RÉVÉSZ has conjectured the existence of a measurable and measure-preserving transformation T, defined on the unit interval, such that $f(x, Tx) > 0$ for $0 \leqq x \leqq 1$. If this conjecture is correct, it provides a continuous analogue of the result quoted in (vii).

The list of such problems can, of course, be extended almost indefinitely. Indeed, it is evident that in the study of d.s. matrices we have scarcely advanced beyond the fringe of the subject.

Added in proof (22 January 1963). Since this paper was submitted for publication, a number of new results on d.s. matrices have appeared in print. (a) P. RÉVÉSZ, A probabilistic solution of problem 111 of G. Birkhoff, Acta Math. Acad. Scient. Hungaricae **13**, 187—198 (1962) has formulated and proved a probabilistic version of Birkhoff's theorem. (b) M. MARCUS and H. MINC, Some results on doubly stochastic matrices, Proc. Amer. math. Soc. **13**, 571—579 (1962) have now succeeded in establishing the inequality (5.2). (c) J. R. ISBELL, Infinite doubly stochastic matrices, Canad. math. Bull. **5**, 1—4 (1962) has given a short and elementary proof of the theorem, due to KENDALL and KIEFER, to the effect that any matrix which is extreme in the set of infinite d. s. matrices must be a permutation matrix. He has also shown that an infinite d. s. matrix $D = (d_{rs})$ can be expressed as a finite convex combination of permutation matrices if and only if the elements d_{rs} take only finitely many distinct values. The proof of this proposition rests on the axiom of choice.

References

[*1*] BELLMAN, R., and A. HOFFMAN: On a theorem of Ostrowski and Taussky. Arch. der Math. *5*, 123—127 (1954).

[*2*] BERGE, C.: Théorie des Graphes et ses Applications. (Paris: 1958).

[*3*] BIRKHOFF, G.: Tres observaciones sobre el algebra lineal. Univ. nac. Tucumán, Revista Ser. A, **5**, 147—150 (1946).

[4] Birkhoff, G.: Lattice Theory (revised edition; New York: 1948).
[5] Bourbaki, N.: Espaces Vectoriels Topologiques (Ch. I—II). Paris: 1953.
[6] Bruijn, N. G. de: Gemeenschappelijke representantensystemen van twee klassenindeelingen van een verzameling. Nieuw Arch. Wiskunde II. R. **22**, 48—52 (1943).
[7] Dmitriev, N., and E. Dynkin: On the characteristic numbers of a stochastic matrix. Doklady Akad. Nauk SSSR n. Ser. **49**, 159—162 (1945).
[8] — Characteristic roots of stochastic matrices. Izvestija Akad. Nauk SSSR, Ser. Mat. **10**, 167—184 (1946) (In Russian).
[9] Dulmage, L., and I. Halperin: On a theorem of Frobenius-König and J. von Neumann's game of hide and seek. Trans. roy. Soc. Canada, Sect. III, III. Ser., **49**, 23—29 (1955).
[10] Eggleston, H. G.: Convexity. Cambridge: 1958.
[11] Fan, K.: Maximum properties and inequalities for the eigenvalues of completely continuous operators. Proc. nat. Acad. Sci. USA **37**, 760—766 (1951).
[12] — A minimum property of the eigenvalues of a Hermitian transformation. Amer. math. Monthly **60**, 48—50 (1953).
[13] — Existence theorems and extreme solutions for inequalities concerning convex functions or linear transformations. Math. Z. **68**, 205—216 (1957—58).
[14] — Convex Sets and their Applications (Argonne National Laboratory, 1959).
[15] —, and A. J. Hoffman: Some metric inequalities in the space of matrices. Proc. Amer. math. Soc. *6*, 111—116 (1955).
[16] —, and G. G. Lorentz: An integral inequality. Amer. math. Monthly **61**, 626—631 (1954).
[17] Farahat, H. K., and L. Mirsky: Permutation endomorphisms and refinement of a theorem of Birkhoff. Proc. Cambridge philos. Soc. **56**, 322—328 (1960).
[18] Feller, W.: An Introduction to Probability Theory and its Applications, vol. I. New York: 1950.
[19] Fuchs, L.: A new proof of an inequality of Hardy-Littlewood-Pólya. Mat. Tidsskr. B, 53—54 (1947).
[20] Hall, P.: On representatives of subsets. J. London math. Soc. **10**, 26—30 (1935).
[21] Hammersley, J. M.: A short proof of the Farahat-Mirsky refinement of Birkhoff's theorem on doubly-stochastic matrices. Proc. Cambridge philos. Soc. **57**, 681 (1961).
[22] —, and J. G. Mauldon: General principles of antithetic variates. Proc. Cambridge philos. Soc. **52**, 476—481 (1956).
[23] Hardy, G. H., J. E. Littlewood, and G. Pólya: Some simple inequalities satisfied by convex functions. Messenger of Math. **58**, 145—152 (1929).
[24] — Inequalities. Cambridge: 1934.
[25] Hoffmann, A. J.: On an inequality of Hardy, Littlewood, and Pólya. (National Bureau of Standards, Report No. 2977, 1953.)
[26] —, and H. W. Wielandt: The variation of the spectrum of a normal matrix. Duke math. J. **20**, 37—40 (1953).
[27] Horn, A.: Doubly stochastic matrices and the diagonal of a rotation matrix. Amer. J. Math **76**, 620—630 (1954).
[28] Isbell, J. R.: Birkhoff's problem 111. Proc. Amer. math. Soc. **6**, 217—218 (1955).
[29] Johnson, D. M., A. L. Dulmage, and N. S. Mendelsohn: On an algorithm of G. Birkhoff concerning doubly stochastic matrices. Canad. math. Bull. **3**, 237—242 (1960).
[30] Karamata, J.: Sur une inégalité relative aux fonctions convexes. Publ. math. Univ. Belgrade **1**, 145—148 (1932).
[31] Karpelevich, F. I.: On the characteristic roots of matrices with non-negative elements. Izvestija Akad. Nauk SSSR. Ser. Mat. **15**, 361—383 (1951) (In Russian).
[32] Kendall, D. G.: On infinite doubly-stochastic matrices and Birkhoff's problem 111. J. London math. Soc. **35**, 81—84 (1960).
[33] König, D.: Über Graphen und ihre Anwendung auf Determinantentheorie und Mengenlehre. Math. Ann. **77**, 435—465 (1916).
[34] Lorentz, G. G.: An inequality for rearrangements. Amer. math. Monthly **60**, 176—179 (1953).
[35] Marcus, M.: Convex functions of quadratic forms. Duke math. J. **24**, 321—326 (1957).
[36] — Some properties and applications of doubly stochastic matrices. Amer. math. Monthly **67**, 215—222 (1960).

[37] MARCUS, M., H. MINC, and B. MOYLS: Some results on non-negative matrices. J. Res. nat. Bur. Standards **65** B, 205—209 (1961).
[38] —, and M. NEWMAN: On the minimum of the permanent of a doubly stochastic matrix. Duke math. J. **26**, 61—72 (1959).
[39] — — Inequalities for the permanent function. Ann. of Math. II. Ser., **75**, 47—62 (1962).
[40] —, and R. REE: Diagonals of doubly stochastic matrices. Quart. J. Math., Oxford II. Ser., **10**, 296—302 (1959).
[41] MAULDON, J. G.: Extreme points of convex sets of doubly-stochastic matrices. Not yet published.
[42] MENDELSOHN, N. S., and A. L. DULMAGE: The convex hull of sub-permutation matrices. Proc. Amer. math. Soc. **9**, 253—254 (1958).
[43] MIRSKY, L.: Proofs of two theorems on doubly-stochastic matrices. Proc. Amer. math. Soc. **9**, 371—374 (1958).
[44] — Matrices with prescribed characteristic roots and diagonal elements. J. London math. Soc. **33**, 14—21 (1958).
[45] — On a convex set of matrices. Arch. der Math. **10**, 88—92 (1959).
[46] — Inequalities for certain classes of convex functions. Proc. Edinburgh math. Soc. **11**, 231—235 (1959).
[47] — Symmetric gauge functions and unitarily invariant norms. Quart. J. Math., Oxford II. Ser., **11**, 50—59 (1960).
[48] — An existence theorem for infinite matrices. Amer. math. Monthly **68**, 465—469 (1961).
[49] — Even doubly-stochastic matrices. Math. Ann. **144**, 418—421 (1961).
[50] NEUMANN, J. VON: A certain zero-sum two-person game equivalent to the optimal assignment problem. Contributions to the Theory of Games, vol. II (Princeton, 1953), 5—12.
[51] OSTROWSKI, A.: Sur quelques applications des fonctions convexes et concaves au sens de I. Schur. J. Math. pur. appl. IX. Sér., **31**, 253—292 (1952).
[52] PÓLYA, G.: Remark on Weyl's note: Inequalities between the two kinds of eigenvalues of a linear transformation. Proc. nat. Acad. Sci. USA **36**, 49—51 (1950).
[53] RADO, R.: An inequality. J. London math. Soc. **27**, 1—6 (1952).
[54] RATTRAY, B. A., and J. E. L. PECK: Infinite stochastic matrices. Trans. roy. Soc. Canada, Sect. III, III. Ser., **49**, 55—57 (1955).
[55] RÉVÉSZ, P.: Seminar on Random Ergodic Theory. (Mathematical Institute, University of Aarhus, 1961.)
[56] SCHREIBER, S.: On a result of S. Sherman concerning doubly stochastic matrices. Proc. Amer. math. Soc. **9**, 350—353 (1958).
[57] SCHUR, I.: Über eine Klasse von Mittelbildungen mit Anwendung auf die Determinantentheorie. S.-Ber. Berliner math. Ges. **22**, 9—20 (1923).
[58] SHERMAN, S.: On a conjecture concerning doubly stochastic matrices. Proc. Amer. math. Soc. **3**, 511—513 (1952).
[59] — A correction to 'On a conjecture concerning doubly stochastic matrices'. Proc. Amer. math. Soc. **5**, 998—999 (1954).
[60] — Doubly stochastic matrices and complex vector spaces. Amer. J. Math. **77**, 245—246 (1955).
[61] VOGEL, W.: Lineare Programme und allgemeine Vertretersysteme. Math. Z. **76**, 103—115 (1961).
[62] WAERDEN, B. L. VAN DER: Aufgabe 45. J.-Ber. Deutsch. Math.-Verein. **35**, 117 (1926).

Department of Pure Mathematics
The University
Sheffield 10, England

(Received May 30, 1962)