Werk

Titel: Spectral Properties of Doubly-Stochastic Matrices.

Autor: MIRSKY, L.; Perfect, Hazel

Jahr: 1965

PURL: https://resolver.sub.uni-goettingen.de/purl?362162050_0069|log7

Kontakt/Contact

Digizeitschriften e.V.
SUB Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen

info@digizeitschriften.de

Spectral Properties of Doubly-Stochastic Matrices

By

Hazel Perfect and **L. Mirsky**, Sheffield

(Received June 10, 1964)

§ 1. Introduction

1.1. A real square matrix with non-negative elements all of whose row-sums and column-sums are equal to 1 is a said to be *doubly-stochastic* (d.s.). The class of d.s. matrices has been the subject of numerous investigations, especially in recent years[1], but very little is as yet known about the spectral properties of this class. The present paper is intended as a contribution to the study of this question.

In discussing the spectral theory of d.s. matrices, we find it helpful to formulate the following central problems.

Problem A. To determine 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 a d.s. $n \times n$ matrix. This problem may, of course, be modified in a variety of ways; for example, we might seek to describe the region Ω_n^0 consisting of those complex numbers which can serve as characteristic roots of d.s. $n \times n$ matrices with zero trace.

Problem B. To determine the region $\mathfrak{D}^n$ of *complex* n-dimensional space which is specified by the requirement that $(z_1, \ldots, z_n) \in \mathfrak{D}^n$ if and only if there exists a d.s. $n \times n$ matrix with characteristic roots $z_1, \ldots, z_n$. Analogously, we define the region D^n in relation to *real* n-dimensional space[2].

It is plain that Problem A is contained in Problem B and that this latter problem admits of a number natural modifications in which we restrict our attention to, say, indecomposable or to primitive d.s. matrices.

---

[1] For a reasonably comprehensive survey of known results, see [7].

[2] The definition implies that, if $(z_1, \ldots, z_n)$ belongs to $\mathfrak{D}^n$, or to D^n, then so does every vector obtained from $(z_1, \ldots, z_n)$ by a permutation of the components.

Problem C. To find necessary and sufficient conditions for a given real monic polynomial of degree n to be the characteristic polynomial of a d.s. $n \times n$ matrix.

None of the problems just stated appears to be easy, and we shall only be able to solve them completely for $n=3$ [3]. In all other cases we shall obtain partial results. We mention, by way of comparison, that, for the class of *stochastic* matrices, Problem A was solved by *Karpelevich* [3] and Problem B was discussed by several writers (see e.g. [10] and [12]).

1.2. A brief description of the contents of the paper is as follows. In § 2 we state some consequences of the *Perron-Frobenius* theorem for the class of d.s. matrices. These consequences form the basis of much of the subsequent discussion. In § 3 we investigate the structure of the regions Ω_n and Ω_n^0, while in § 4 we restrict the discussion to characteristic roots of d.s. matrices which lie on the unit cricle. § 5 is devoted principally to the investigation of the n-dimensional region D^n. Here the main result is Theorem 8 which implies, in particular, that any set of n real non-negative numbers less than or equal to 1 and containing 1 is the set of characteristic roots of a d.s. $n \times n$ matrix. Next, in § 6, we show how the problem of determining the region $\mathfrak{D}^n$ can be reduced to the corresponding problem for primitive matrices. The question of convexity of a region closely related to D^n is discussed in § 7. Finally, in §§ 8 and 9 we examine the special cases $n = 3$ and $n = 4$ and obtain, among other things, a complete description of the regions $\mathfrak{D}^3$ and D^3.

§ 2. The Perron-Frobenius theorem

2.1. Two almost obvious but useful facts may be stated at the outset of our discussion.

Lemma 1. (a) *Every d.s. matrix has 1 as a characteristic root.* (b) *Every characteristic root of every d.s. matrix is less than or equal to 1 in modulus.*

Statement (a) is immediate, a characteristic vector associated with the root 1 being $(1,1,\ldots,1)^T$ [4]. Statement (b) follows easily from *Hadamard*'s inclusion theorem (see e.g. [6], Theorem 7.5.4).

2.2. In the course of our discussion we shall lean fairly heavily on parts of the celebrated *Perron-Frobenius* theorem. We therefore use the

[3] The case $n = 2$ is trivial and need not detain us.

[4] A raised T is used to denote transposition.

present section to state the relevant parts of the theorem and to point out certain consequences for the special case of d.s. matrices.

To begin with, we need to explain our terminology. Two square matrices will be said to be *cogredient* if one of them can be obtained from the other as the result of a permutation of the rows and the same permutation of the columns. Thus A, B are cogredient if and only if there exists a permutation matrix P such that $B = P^T AP$. It is clear, then, that cogredient matrices have the same characteristic roots.

A square matrix is said to be *decomposable* if it is cogredient to a matrix of the form

$$\begin{Vmatrix} A_1 & O \\ A_3 & A_2 \end{Vmatrix},$$

where A_1 and A_2 are square matrices. Otherwise it is said to be *indecomposable.*

We shall say that a characteristic root ω of a matrix A is *maximal* if, for every characteristic root ω' of A, the relation $|\omega'| \leqq |\omega|$ is valid. Thus a characteristic root of a d.s. matrix is maximal if and only if it lies on the unit circle.

Lemma 2. (Perron-Frobenius)

If an indecomposable non-negative matrix A possesses exactly r maximal characteristic roots, then these are of the form $\varrho \exp(2\pi ik/r)$, $0 \leqq k < r$, where ϱ is a positive number. If $r > 1$, then A is cogredient to a matrix of the form

$$\begin{Vmatrix} O & A_{12} & O & \dots & O \\ O & O & A_{23} & \dots & O \\ \dots & \dots & \dots & \dots & \dots \\ O & O & O & \dots & A_{r-1,r} \\ A_{r1} & O & O & \dots & O \end{Vmatrix},$$

where the diagonal blocks are square matrices.

A particularly simple proof of the Perron-Frobenius theorem (including parts not stated here) has been given by *Wielandt* [13]. For a very thorough discussion of the entire spectral theory of non-negative matrices, we refer the reader to *Gantmacher*'s book ([2], vol. II, chap. XIII).

We now turn to the special case of d.s. matrices.

Lemma 3. Every d.s. matrix is cogredient to a direct sum of indecomposable d.s. matrices.

This result is almost certainly well-known. As, furthermore, it follows very easily from the definitions, we omit the details of the proof.

Making use of Lemma 2, *Marcus, Minc*, and *Moyls* ([5], Theorem 1) obtained the following result.

Lemma 4. If an indecomposable d.s. $n \times n$ matrix D has exactly r characteristic roots of unit modulus, then these are precisely the r-th roots of unity. If $r > 1$, then r is a divisor of n (in symbols: $r \mid n$), and D is cogredient to a matrix of the form

$$\begin{Vmatrix} O & D_1 & O & \dots & O \\ O & O & D_2 & \dots & O \\ \dots & \dots & \dots & \dots & \dots \\ O & O & O & \dots & D_{r-1} \\ D_r & O & O & \dots & O \end{Vmatrix},$$

where all blocks are of type $n/r \times n/r$ and the D_i are d.s. matrices.

Finally, we recall that a non-negative indecomposable matrix which possesses only one maximal characteristic root is called *primitive.*

§ 3. The regions Ω_n and Ω_n^0

3.1: For $k \geqq 2$, we shall denote by Π_k the region of the complex plane specified as the convex hull of the k-th roots of unity.

Theorem 1. For $n \geqq 2$, we have

$$\Pi_2 \cup \Pi_3 \cup \dots \cup \Pi_n \subseteq \Omega_n. \tag{3.1}$$

This relation was pointed out by *J. F. C. Kingman* [7]. We give a proof for the sake of completeness which is a slight modification of that stated in [7].

It is clearly sufficient to show that

$$\Pi_k \subseteq \Omega_n \qquad (2 \leqq k \leqq n). \tag{3.2}$$

Let $z \in \Pi_k$ and write $\eta = \exp(2\pi i/k)$. Then

$$z = a_0 + a_1 \eta + \dots + a_{k-1} \eta^{k-1} \tag{3.3}$$

for suitable non-negative a's whose sum is 1. Write

$$D = I_{n-k} \oplus C,$$

where $\oplus$ signifies the formation of a direct sum of matrices, I_{n-k} is the unit matrix of order $n - k$, and C is the circulant $k \times k$ matrix whose first row is $a_0, a_1, \dots, a_{k-1}$. Then D is a d.s. $n \times n$ matrix and, by virtue of (3.3), z is a characteristic root of D. Thus $z \in \Omega_n$ and (3.2)

follows. The assertion (3.1) is therefore established. In § 8, we shall show that the sign of inclusion in (3.1) can be replaced by the sign of equality for $n = 3$. Whether this is so generally, remains an open question.

Corollary. Let ϱ_n denote the upper bound of numbers α with the property that the disc $|z| \leqq \alpha$ is contained in Ω_n. Then $\varrho_n \geqq \cos \pi/n$.

Virtually the same argument as that used to prove Theorem 1 leads to a result concerning Ω_n^0. We shall denote by Π_n^0 the union of the point 1 and the convex hull of the n-th roots of unity different from 1.

Theorem 2. For $n \geqq 2$, we have

$$\Pi_n^0 \subseteq \Omega_n^0. \tag{3.4}$$

Let $z \in \Pi_n^0$. If $z = 1$, then $z \in \Omega_n^0$ by Lemma 1(a). If $z \neq 1$, then z belongs to the convex hull of the n-th roots of unity different from 1. Hence

$$z = a_1 \eta + a_2 \eta^2 + \ldots + a_{n-1} \eta^{n-1},$$

where $\eta = \exp(2\pi i/n)$ and $a_1, \ldots, a_{n-1}$ are suitable non-negative numbers whose sum is 1. Denoting by C the $n \times n$ circulant matrix whose first row is $0, a_1, \ldots, a_{n-1}$, we observe that z is a characteristic root of the d.s. matrix C with zero trace. Hence $z \in \Omega_n^0$ and the assertion follows. We shall subsequently note that (3.4) reduces to equality for $n = 3$.

We next discuss a result of *Dmitriev* and *Dynkin* [1] relating to the structure of Ω_n.

Theorem 3. If $\omega \in \Omega_n$ and $0 \leqq t \leqq 1$, then $t\omega \in \Omega_n$.

To prove this result[5], we shall assume (as may be done in view of the relation $\Pi_2 \subseteq \Omega_n$) that $\omega \neq 1$. Since $\omega \in \Omega_n$, there exist a d.s. $n \times n$ matrix D and a vector $x \neq 0$ such that $Dx = \omega x$. Denoting by $\sigma(x)$ the sum of the components of x, we infer that

$$\sigma(x) = \sigma(Dx) = \omega \, . \, \sigma(x)$$

so that $\sigma(x) = 0$. Let J be the $n \times n$ matrix all of whose elements are equal to n^{-1}. Then $Jx = 0$ and so

$$\{tD + (1 - t) J\} x = t\omega \, . \, x.$$

[5] We are not able to follow the argument in [1] and the proof given here is due to Dr. *H. K. Farahat*, to whom we are also indebted for a number of stimulating discussions.

Thus $t\omega$ is a characteristic root of the d.s. matrix $tD + (1 - t)J$, and consequently $t\omega \in \Omega_n$, as asserted.

3.2. Finally, we give a further method for generating points of Ω_n from given points of that region. We shall denote by $\mathfrak{F}$ the class of functions φ of a complex variable with the following properties: (i) $\varphi(z)$ is regular for $|z| \leqq 1$; (ii) $\varphi(1) = 1$; (iii) $\varphi^{(k)}(0) \geqq 0$ for $k \geqq 0$. In view of (i), $\varphi(z)$ can be expanded as a power series in z whose radius of convergence is greater than 1. Hence, by Lemma 1(b) and the fact that the product of d.s. matrices is again d.s., we infer that, if $\varphi \in \mathfrak{F}$, then $\varphi(D)$ is defined for every d.s. matrix D and is itself d.s.

We note that the function e^{z-1} belongs to $\mathfrak{F}$ and that so does every polynomial in z whose coefficients are non-negative real numbers with sum 1. For the record, we state the following result.

Theorem 4. If $\omega \in \Omega_n$ and $\varphi \in \mathfrak{F}$, then $\varphi(\omega) \in \Omega_n$. In particular, $e^{\omega-1} \in \Omega_n$.

For, if ω is a characteristic root of the d.s. $n \times n$ matrix D, then $\varphi(\omega)$ is a characteristic root of the d.s. $n \times n$ matrix $\varphi(D)$.

§ 4. Characteristic roots of d.s. matrices on the unit circle

4.1. We begin with a rather obvious result.

Theorem 5. If all characteristic roots of a d.s. matrix D lie on the unit circle, then D is a permutation matrix.

This result can be exhibited as an easy consequence of the Perron-Frobenius theorem. An alternative argument runs as follows. Let $D = (d_{rs})$ be of type $n \times n$ and denote its characteristic roots by $\omega_1, \ldots, \omega_n$, so that $|\omega_1| = \ldots = |\omega_n| = 1$. Then, by *Schur*'s inequality ([11], Theorem 2),

$$n = \sum_{r=1}^{n} |\omega_r|^2 \leqq \sum_{r,s=1}^{n} d_{rs}^2 \leqq \sum_{r,s=1}^{n} d_{rs} = n.$$

Hence

$$\sum_{r,s=1}^{n} d_{rs}^2 = \sum_{r,s=1}^{n} d_{rs}$$

and so $d_{rs}^2 = d_{rs} (1 \leqq r, s \leqq n)$. Thus every d_{rs} is 1 or 0, and the assertion is proved. It should be noted, however, that Theorem 5 is contained in a much more powerful result due to *Marcus, Minc,* and *Moyls* ([5], Theorems 2 and 3).

4.2. To begin a more systematic study of characteristic roots of d.s. matrices on the unit circle,we define the subset Θ_n of the complex plane by the requirement that $\zeta \in \Theta_n$ if and only if $|\zeta| = 1$ and ζ is a characteristic root of a d.s. $n \times n$ matrix. If in this definition the expression 'd.s. matrix' is replaced by 'd.s. matrix with zero trace', 'indecomposable d.s. matrix', or 'indecomposable d.s. matrix with zero trace' respectively, then the set so defined in place of Θ_n will be denoted by Θ_n^0, $\hat{\Theta}_n$, $\hat{\Theta}_n^0$ respectively. Furthermore, we shall denote by Z_n the set of n-th roots of unity.

Theorem 6. For $n \geqq 2$, we have

$$\text{(i)}\quad \hat{\Theta}_n = \hat{\Theta}_n^0 = Z_n;$$

$$\text{(ii)}\quad \Theta_n = \bigcup_{2 \leq k \leq n} Z_k;$$

$$\text{(iii)}\quad \Theta_n^0 = \bigcup_{\substack{2 \leq k \leq n \\ k \neq n-1}} Z_k.$$

The second part of this theorem was established, in a different way, by *Dmitriev* and *Dynkin* [1]; their proof is reproduced in [8].

To prove (i), let $\zeta \in Z_n$. If $\zeta = 1$, then $\zeta \in \hat{\Theta}_n^0$ trivially. If $\zeta \neq 1$, write $\zeta = \exp(2\pi ik/n)$, where $0 < k < n$. Denoting by P the circulant $n \times n$ matrix with first row $\delta_{1,k+1}, \ldots, \delta_{n,k+1}$, we note that ζ is a characteristic root of P and that P is d.s., has zero trace, and is indecomposable. Thus $\zeta \in \hat{\Theta}_n^0$ and consequently

$$Z_n \subseteq \hat{\Theta}_n^0. \tag{4.1}$$

Next, we have trivially the relation

$$\hat{\Theta}_n^0 \subseteq \hat{\Theta}_n. \tag{4.2}$$

Again, let $\zeta \in \hat{\Theta}_n$, so that ζ is a characteristic root of an indecomposable d.s. $n \times n$ matrix D. Denote by h the number of characteristic roots of D (including ζ) which lie on the unit circle. Then, by Lemma 4, these are precisely the h-th roots of unity, so that $\zeta^h = 1$. Moreover, also by Lemma 4, n is a multiple of h, so that $\zeta^n = 1$, i.e. $\zeta \in Z_n$. Thus

$$\hat{\Theta}_n \subseteq Z_n; \tag{4.3}$$

and (i) now follows by (4.1), (4.2), and (4.3).

Next, let $\zeta \in Z_k$, where $2 \leqq k \leqq n$. Then, by (i), ζ is a characteristic root of a d.s. $k \times k$ matrix D and so of the d.s. $n \times n$ matrix $D \oplus I_{n-k}$. Thus $Z_k \subseteq \Theta_n (2 \leqq k \leqq n)$ and consequently

$$\bigcup_{2 \leq k \leq n} Z_k \subseteq \Theta_n. \tag{4.4}$$

Again, let $\zeta \in \Theta_n$, i.e. let $|\zeta| = 1$ and suppose that ζ is a characteristic root of a d.s. $n \times n$ matrix E. Now, by Lemma 3, E is cogredient to a direct sum of indecomposable d.s. matrices and so ζ is a characteristic root of an indecomposable d.s. $k \times k$ matrix, where k is some integer in the range $2 \leqq k \leqq n$. Hence $\zeta \in \hat{\Theta}_2 \cup \ldots \cup \hat{\Theta}_n$ and so, by (i), $\zeta \in Z_2 \cup \ldots \cup Z_n$. We have, therefore,

$$\Theta_n \subseteq \bigcup_{2 \leq k \leq n} Z_k; \tag{4.5}$$

and (4.4) and (4.5) imply (ii).

Now let $\zeta \in Z_k$, where k satisfies $2 \leqq k \leqq n$, $k \neq n-1$. Then, by the construction carried out in the proof of (i), ζ is a characteristic root of a d.s. $k \times k$ matrix D with zero trace. Denote by D' any d.s. $(n-k) \times (n-k)$ matrix with zero trace. Then ζ is a characteristic root of the d.s. $n \times n$ matrix $D \oplus D'$ with zero trace, and hence

$$Z_k \subseteq \Theta_n^0 \quad (2 \leqq k \leqq n;\ k \neq n-1).$$

It follows that

$$\bigcup_{\substack{2 \leq k \leq n \\ k \neq n-1}} Z_k \subseteq \Theta_n^0. \tag{4.6}$$

Again, let $\zeta \in \Theta_n^0$, i.e. let $|\zeta| = 1$ and suppose that ζ is a characteristic root of a d.s. $n \times n$ matrix E with zero trace. But E is cogredient to a direct sum of indecomposable d.s. matrices; and so ζ is a characteristic root of E_1, where E is cogredient to $E_1 \oplus E_2$, the matrices E_1, E_2 are d.s. and have zero trace, and E_1 is indecomposable. Denoting the order of E_1 by k, we see that $\zeta \in \hat{\Theta}_k^0$ and it is plain that $2 \leqq k \leqq n$, $k \neq n-1$. Hence, by (i), $\zeta \in Z_k$ and so

$$\Theta_n^0 \subseteq \bigcup_{\substack{2 \leq k \leq n \\ k \neq n-1}} Z_k. \tag{4.7}$$

Statement (iii) now follows from (4.6) and (4.7).

4.3. By the *boundary set* of a matrix we shall understand the set of its maximal characteristic roots. We shall next give a complete description of the boundary sets of all d.s. matrices.

Lemma 5. The complex numbers $\zeta_1, \ldots, \zeta_r$ constitute the boundary set of an indecomposable d.s. $n \times n$ matrix if and only if

$$\{\zeta_1, \ldots, \zeta_r\} = Z_r, \quad r \mid n. \tag{4.8}$$

If $\zeta_1, \ldots, \zeta_r$ is the boundary set of an indecomposable d.s. $n \times n$ natrix, then the relations (4.8) hold by Lemma 4. Assume, then, that the relations (4.8) are valid and write $n = rs$. Denote by J_s the $s \times s$ matrix all of whose elements are equal to s^{-1} and let the d.s. $n \times n$ matrix D be defined by

$$D = \begin{Vmatrix} O & J_s & O & \ldots & O \\ O & O & J_s & \ldots & O \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ O & O & O & \ldots & J_s \\ J_s & O & O & \ldots & O \end{Vmatrix},$$

where each row and each column contains r blocks. Write $\eta = \exp(2\pi i/r)$, and denote by T the partitioned matrix of $r \times r$ blocks in which the (k, l)-th block is equal to $\eta^{(k-1)(l-1)} I_s$. Then, as is easily verified,

$$T^{-1} DT = J_s \oplus \eta J_s \oplus \ldots \oplus \eta^{r-1} J_s.$$

Now the characteristic roots of J_s are 0 ($s - 1$ times) and 1. Hence the characteristic roots of D are 0 ($n-r$ times), $\zeta_1, \ldots, \zeta_r$[6] and so $\{\zeta_1, \ldots, \zeta_r\}$ is the boundary set of D. Moreover, D is indecomposable; for otherwise it would be cogredient to the direct sum of two d.s. matrices and so would have 1 as a repeated characteristic root.

In stating our next result, we shall employ a notation resembling that of union. Let $S_1, \ldots, S_k$ be sets of objects. For any object x, denote by $\nu(x)$ the number of S_i for which $x \in S_i$. We now define $S_1 \dot{\cup} \ldots \dot{\cup} S_k$ ($= S$, say) as the set of all objects which belong to at least one S_i, with the convention that each x in S occurs exactly $\nu(x)$ times.

Theorem 7. The complex numbers $\zeta_1, \ldots, \zeta_r$ constitute the boundary set of a d.s. $n \times n$ matrix if and only if there exist positive integers k, $r_1, \ldots, r_k, s_1, \ldots, s_k$ such that

$$n = r_1 s_1 + \ldots + r_k s_k, \tag{4.9}$$

$$\{\zeta_1, \ldots, \zeta_r\} = Z_{r_1} \dot{\cup} \ldots \dot{\cup} Z_{r_k}. \tag{4.10}$$

Let the conditions (4.9) and (4.10) be satisfied. Then, for $1 \leqq i \leqq k$, there exists, by Lemma 5, an (indecomposable) d.s. matrix D_i of order $r_i s_i$ with boundary set Z_{r_i}. It follows that $D_1 \oplus \ldots \oplus D_k$ is a d.s. $n \times n$ matrix with boundary set $Z_{r_1} \dot{\cup} \ldots \dot{\cup} Z_{r_k}$, i.e. $\{\zeta_1, \ldots, \zeta_r\}$.

---

[6] This conclusion also follows easily from Lemma 7 below.

Next, let D be a d.s. $n \times n$ matrix with boundary set $\{\zeta_1, \ldots, \zeta_r\}$. Now D is cogredient to a direct sum $D_1 \oplus \ldots \oplus D_k$, where $D_1, \ldots, D_k$ are indecomposable d.s. matrices. Denoting by n_i the order of D_i, we have

$$n = n_1 + \ldots + n_k.$$

Now, by Lemma 5, the boundary set of D_i is Z_{r_i}, where r_i is a divisor of n_i. Writing $n_i = r_i s_i$, we see that (4.9) is satisfied. Moreover, the boundary set of D is $Z_{r_1} \dot{\cup} \ldots \dot{\cup} Z_{r_k}$ and (4.10) follows.

§ 5. The regions D^n and $\mathfrak{D}^n$

5.1. Problem B is certainly very much harder than Problem A and has not, to our knowledge, been previously treated in the literature. At present, we are only able to offer incomplete results.

A set of $n \times n$ matrices $A_1, \ldots, A_k$ such that

$$A_r^2 = A_r \neq O, \quad A_r A_s = O \quad (1 \leqq r, s \leqq k; r \neq s)$$

will be called an *idempotent system.*

Lemma 6. If $E_1, \ldots, E_n$ is an idempotent system of n matrices of type $n \times n$ and if $\alpha_1, \ldots, \alpha_n$ are any (complex) numbers, then the set of characteristic roots of the matrix

$$M = \alpha_1 E_1 + \ldots + \alpha_n E_n \tag{5.1}$$

is identical with $\{\alpha_1, \ldots, \alpha_n\}$.

We assume, in the first place, that the α's are distinct. Since $E_r \neq O$, there exists a vector x such that $E_r x \neq 0$. From (5.1) ,we infer that

$$M E_r x = \alpha_r E_r^2 x = \alpha_r E_r x.$$

Thus $\alpha_1, \ldots, \alpha_n$ are characteristic roots of M and, since they are distinct, $\{\alpha_1, \ldots, \alpha_n\}$ is precisely the set of characteristic roots of M. The case of arbitrary α's now follows by continuity[7].

Theorem 8. Let $n \geqq 2$ and let $\lambda_1, \ldots, \lambda_{n-1}$ be real numbers such that

$$1 \geqq \lambda_1 \geqq \ldots \geqq \lambda_{n-1} \tag{5.2}$$

and

$$\frac{1}{n} + \frac{\lambda_1}{n(n-1)} + \frac{\lambda_2}{(n-1)(n-2)} + \ldots + \frac{\lambda_{n-1}}{2.1} \geqq 0. \tag{5.3}$$

Then $(1, \lambda_1, \ldots, \lambda_{n-1}) \in \mathrm{D}^n$.

[7] We can avoid an appeal to considerations of continuity if we make use of the known fact that matrices of an idempotent system can be reduced to diagonal form by the same similarity transformation.

We shall write $J_0 = J_0(n)$ for the $n \times n$ matrix all of whose elements are equal to n^{-1}. Further, we shall define the $n \times n$ matrix J_k by the equation

$$J_k = I_k \oplus J_0(n-k) \ (1 \leqq k \leqq n-1).$$

It is then clear that $J_{n-1} = I_n$. Let the matrix $D = (d_{rs})$ be defined by the equation

$$D = (1-\lambda_1)J_0 + (\lambda_1-\lambda_2)J_1 + \ldots + (\lambda_{n-2}-\lambda_{n-1})J_{n-2} + \lambda_{n-1}J_{n-1}. \quad (5.4)$$

Every row-sum and every column-sum of each of the matrices

$$J_0, J_1, \ldots, J_{n-1}$$

is equal to 1, and the same relations therefore hold for D. Again, it follows from (5.2) and (5.4) that all elements of D except possibly those involving λ_{n-1} are non-negative. The exceptional elements are

$$d_{n-1,n} = d_{n,n-1} = \frac{1-\lambda_1}{n} + \frac{\lambda_1-\lambda_2}{n-1} + \ldots + \frac{\lambda_{n-2}-\lambda_{n-1}}{2},$$

$$d_{n-1,n-1} = d_{nn} = \frac{1-\lambda_1}{n} + \frac{\lambda_1-\lambda_2}{n-1} + \ldots + \frac{\lambda_{n-2}-\lambda_{n-1}}{2} + \lambda_{n-1}$$

$$= \frac{1}{n} + \frac{\lambda_1}{n(n-1)} + \frac{\lambda_2}{(n-1)(n-2)} + \ldots + \frac{\lambda_{n-1}}{2.1},$$

and it follows by (5.2) and (5.3) that these elements, too, are non-negative. Thus D is a d.s. matrix.

By (5.4) we have at once

$$D = J_0 + \lambda_1(J_1-J_0) + \lambda_2(J_2-J_1) + \ldots + \lambda_{n-1}(J_{n-1}-J_{n-2}). \quad (5.5)$$

Now it is easily verified that

$$J_pJ_q = J_{\min(p,q)} \quad (0 \leqq p, q < n)$$

and from this it follows that the matrices

$$J_0, J_1-J_0, J_2-J_1, \ldots, J_{n-1}-J_{n-2}$$

constitute an idempotent system. Hence, by (5.5) and Lemma 6, the characteristic roots of the d.s. (and symmetric) matrix D are $1, \lambda_1, \ldots \ldots, \lambda_{n-1}$. Our assertion is thereby established.

5.2. We append some remarks on the theorem just proved.

(i) For $r > 1$, denote by K_r the $r \times r$ matrix all of whose diagonal elements are equal to 0 and all of whose off-diagonal elements are equal to $(r-1)^{-1}$. It can be shown that the set of matrices (5.4), where the

λ's are variable but subject to (5.2) and (5.3), is identical with the convex hull of the d.s. matrices

$$I_n, I_{n-r} \oplus K_r \qquad (2 \leqq r \leqq n).$$

(ii) In § 8 we shall see that the conditions shown here to be sufficient for the validity of the relation $(1, \lambda_1, \ldots, \lambda_{n-1}) \in \mathrm{D}^n$ are also necessary when $n=3$ [8]. However, these conditions are no longer necessary for $n \geqq 4$. Thus the characteristic roots of the d.s. matrix

$$\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} \oplus \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} \oplus I_{n-4}$$

are $1, \lambda_1, \ldots, \lambda_{n-1}$, where

$$\lambda_1 = \ldots = \lambda_{n-3} = 1, \quad \lambda_{n-2} = \lambda_{n-1} = -1;$$

and these numbers satisfy (5.2) but not (5.3).

5.3. We conclude the present section with some easy results on the region $\mathfrak{D}^n$.

Theorem 9. If $\{\xi_i : 1 \leqq i \leqq m\} \in \mathfrak{D}^m$ *and* $\{\eta_i : 1 \leqq i \leqq n\} \in \mathfrak{D}^n$ *and if* φ *is any function of the class* $\mathfrak{F}$, *then*

$$\{\varphi(\xi_i \eta_j) : 1 \leqq i \leqq m, \quad 1 \leqq j \leqq n\} \in \mathfrak{D}^{mn}.$$

To prove this proposition, let D_1, D_2 be d.s. matrices of order m, n with characteristic roots ξ_i $(1 \leqq i \leqq m)$ and η_i $(1 \leqq i \leqq n)$ respectively. Then the left direct product $E=D_1 \,.\times D_2$(cf. [4], § 43) is a matrix of order mn and is obviously again d.s. Since its characteristic roots are $\xi_i \eta_j (1 \leqq i \leqq m; 1 \leqq j \leqq n)$, we have

$$\{\xi_i \eta_j : 1 \leqq i \leqq m; \ 1 \leqq j \leqq n\} \in \mathfrak{D}^{mn}.$$

Now $\varphi(E)$ is again d.s. and its characteristic roots are $\varphi(\xi_i \eta_j)$ $(1 \leqq i \leqq m, 1 \leqq j \leqq n)$. The assertion is therefore proved.

Corollary. If $(\xi_1, \ldots, \xi_n) \in \mathfrak{D}^n$, $0 \leqq \alpha \leqq 1$, $-1 \leqq \sigma \leqq 1$, *then*

$$\{1 - \alpha + \alpha \xi_i, 1 - \alpha + \alpha \sigma \xi_i : 1 \leqq i \leqq n\} \in \mathfrak{D}^{2n}. \tag{5.6}$$

In particular

$$(\xi_1, \ldots, \xi_n, \sigma \xi_1, \ldots, \sigma \xi_n) \in \mathfrak{D}^{2n}. \tag{5.7}$$

Since $-1 \leqq \sigma \leqq 1$, we have $(1, \sigma) \in \mathfrak{D}^2$ and so, by Theorem 9,

[8] We assume that the λ's are numbered in non-ascending order of magnitude.

$$\{\varphi(\xi_1), \ldots, \varphi(\xi_n), \varphi(\sigma\xi_1), \ldots, \varphi(\sigma\xi_n)\} \in \mathfrak{D}^{2n}$$

for any φ in $\mathfrak{F}$. Taking $\varphi(x) = 1 - \alpha + \alpha x$, we obtain (5.6). For $\alpha = 1$, this reduces to (5.7).

§ 6. Reduction of the problem of $\mathfrak{D}^n$ to the primitive case

6.1. We shall need some additional notation. If in the definition of $\mathfrak{D}^n$ (see § 1) the term 'd.s. matrix' is replaced by 'indecomposable d.s. matrix' or by 'primitive d.s. matrix', then we shall write $\hat{\mathfrak{D}}^n$ or $\hat{\hat{\mathfrak{D}}}^n$ respectively in place of $\mathfrak{D}^n$.

Let $\mathfrak{V}_1, \ldots, \mathfrak{V}_k$ be sets of vectors of order $n_1, \ldots, n_k$ respectively. We shall denote by $\mathfrak{V}_1 \oplus \ldots \oplus \mathfrak{V}_k$ the set of all vectors of order $n_1 + \ldots + n_k$ in which the first n_1 components form a vector in $\mathfrak{V}_1$, the next n_2 a vector in $\mathfrak{V}_2$, and so on.

Let $\mathfrak{S}$ be a set of s-vectors. Then $\mathfrak{S}_{(r)}$ will denote the set of sr-vectors of the form $\{(z_i^{1/r})_j : 1 \leqq i \leqq s,\ 1 \leqq j \leqq r\}$, where $\{z_i : 1 \leqq i \leqq s\} \in \mathfrak{S}$. Here $(z^{1/r})_1, \ldots, (z^{1/r})_r$ denote the set (in some order) of all r-th roots of z[9].

6.2. We recall that every d.s. matrix is cogredient to the direct sum of indecomposable d.s. matrices. It therefore follows that $\mathfrak{D}^n$ is the set of vectors obtained by permutations of components in vectors belonging to

$$\hat{\mathfrak{D}}^{n_1} \oplus \ldots \oplus \hat{\mathfrak{D}}^{n_r},$$

where $k, n_1, \ldots, n_k$ are positive integers such that $n_1 + \ldots + n_k = n$. Thus the problem of determining $\mathfrak{D}^n$ is reduced to that of $\hat{\mathfrak{D}}^n$. And this, in turn, will be reduced to the analogous problem involving primitive matrices only.

Lemma 7. Let

$$M = \begin{Vmatrix} O & A_1 & O & \ldots & O \\ O & O & A_2 & \ldots & O \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ O & O & O & \ldots & A_{r-1} \\ A_r & O & O & \ldots & O \end{Vmatrix},$$

where all blocks are of type $s \times s$. If the set of characteristic roots of $A_1 A_2 \ldots A_r$ is $\{\omega_1, \ldots, \omega_s\}$, then the set of characteristic roots of M is $\{\omega_1, \ldots, \omega_s\}_{(r)}$, i.e. $(\omega_k^{1/r})_j$ $(1 \leq k \leq s,\ 1 \leq j \leq r)$.

[9] We have not specified the order of components in the vectors of $\mathfrak{S}_{(r)}$ but for the discussion below this is irrelevant.

For a proof of this result on the characteristic roots of partitioned matrices, we refer the reader to [9].

Theorem 10. We have $\hat{\mathfrak{D}}^n = \bigcup_{r \mid n} (\hat{\hat{\mathfrak{D}}}^{n/r})_{(r)}$.

Let $(z_1, \ldots, z_n) \in \hat{\mathfrak{D}}^n$, and denote by E an indecomposable d.s. $n \times n$ matrix with characteristic roots $z_1, \ldots, z_n$. If E is primitive, then $(z_1, \ldots, z_n) \in \hat{\hat{\mathfrak{D}}}^n$ and so

$$(z_1, \ldots, z_n) \in \bigcup_{r \mid n} (\hat{\hat{\mathfrak{D}}}^{n/r})_{(r)}. \tag{6.1}$$

If, on the other hand, E is imprimitive, then we denote by $r\,(>1)$ the number of its maximal characteristic roots. Then, by Lemma 4, $r \mid n$ and we write $n = rs$. Further, E is cogredient to a matrix of the form

$$\begin{Vmatrix} O & E_1 & O & \ldots & O \\ O & O & E_2 & \ldots & O \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ O & O & O & \ldots & E_{r-1} \\ E_r & O & O & \ldots & O \end{Vmatrix},$$

where each block is of type $s \times s$ and the E_i are d.s. Denote by $\xi_1, \ldots, \xi_s$ the characteristic roots of $E_1 \ldots E_r$. Then, by Lemma 7, the set of characteristic roots of E is $\{\xi_1, \ldots, \xi_s\}_{(r)}$. Moreover, E has, by hypothesis, exactly r characteristic roots of unit modulus. Hence $E_1 \ldots E_r$ has exactly one characteristic root of unit modulus and so is primitive. Thus $(z_1, \ldots, z_n) \in (\hat{\hat{\mathfrak{D}}}^s)_{(r)} = (\hat{\hat{\mathfrak{D}}}^{n/r})_{(r)}$, and (6.1) is again valid. We therefore have

$$\hat{\mathfrak{D}}^n \subseteq \bigcup_{r \mid n} (\hat{\hat{\mathfrak{D}}}^{n/r})_{(r)}.$$

Next, let $1 < r \leqq n$ and $r \mid n$; and write $rs = n$. Let E_1 be a d.s. primitive $s \times s$ matrix and denote its characteristic roots by $\xi_1, \ldots, \xi_s$. Put

$$E = \begin{Vmatrix} O & E_1 & O & \ldots & O \\ O & O & I & \ldots & O \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ O & O & O & \ldots & I \\ I & O & O & \ldots & O \end{Vmatrix},$$

where each block is of type $s \times s$ and r is the number of non-zero blocks.

Then E is a d.s. $n \times n$ matrix and, in view of Lemma 7, its characteristic roots are

$$\{\xi_1, \ldots, \xi_s\}_{(r)}. \tag{6.2}$$

Since E_1 is primitive, just one of the numbers $\xi_1, \ldots, \xi_s$ has unit modulus. Hence E has exactly r characteristic roots of unit modulus and these are the r-th roots of unity. Therefore E is indecomposable and so the vector with components (6.2) belongs to $\hat{\mathfrak{D}}^n$. Thus

$$(\hat{\hat{\mathfrak{D}}}^{n/r})_{(r)} = (\hat{\hat{\mathfrak{D}}}^{s})_{(r)} \subseteq \hat{\mathfrak{D}}^n;$$

and this is plainly still valid for $r = 1$. Hence

$$\bigcup_{r \mid n} (\hat{\hat{\mathfrak{D}}}^{n/r})_{(r)} \subseteq \hat{\mathfrak{D}}^n,$$

and the assertion follows.

§ 7. A question of convexity

We shall denote by $\overline{\mathrm{D}}^n$ the region of real n-dimensional space specified by the requirement that $(\lambda_1, \ldots, \lambda_n) \in \overline{\mathrm{D}}^n$ if and only if $(1, \lambda_1, \ldots, \lambda_n) \in \mathrm{D}^{n+1}$. The convexity of $\overline{\mathrm{D}}^1$ is trivial and the convexity of $\overline{\mathrm{D}}^2$ will follow as a consequence of the discussion in § 8. However, for $n \geqq 3$, $\overline{\mathrm{D}}^n$ is non-convex; and the object of the present section is to establish this result.

Lemma 8. Let $n \geqq 3$ and $1 > \alpha_1, \ldots, \alpha_{n-1} > -1$. If $(\alpha_1, \ldots, \alpha_{n-1}, -1) \in \overline{\mathrm{D}}^n$, then $\alpha_1 + \ldots + \alpha_{n-1} = 0$.

By hypothesis, there exists a d.s. $(n+1) \times (n+1)$ matrix M with characteristic roots $1, \alpha_1, \ldots, \alpha_{n-1}, -1$. If M were decomposable, then, by Lemma 3, it would be cogredient to the direct sum of two d.s. matrices and so would have 1 as a repeated characteristic root. Hence M is indecomposable and, since it has two maximal characteristic roots, its trace vanishes by virtue of Lemma 4. The assertion therefore follows.

Theorem 11. The region $\overline{\mathrm{D}}^n$ is non-convex for $n \geqq 3$.

Let $k \geqq 2$. The characteristic roots of the d.s. matrix

$$\begin{Vmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{Vmatrix} \oplus \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} \oplus \ldots \oplus \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix},$$

where the second summand is repeated $k-1$ times, are $1^k, 0, (-1)^{k-1}$. (We use superscripts to indicate multiplicity). Thus

$$(1^k, 0, (-1)^{k-1}) \in D^{2k},$$

i.e.
$$(1^{k-1}, 0, (-1)^{k-1}) \in \overline{D}^{2k-1}.$$

Hence

$$(1, 0, 1, -1, 1, -1, \ldots 1, -1, -1) \in \overline{D}^{2k-1}$$

and

$$(0, 1, -1, 1, -1, 1, \ldots -1, 1, -1) \in \overline{D}^{2k-1}.$$

If $\overline{D}^{2k-1}$ were convex, it would follow that

$$(1/2, 1/2, 0, 0, \ldots, 0, -1) \in \overline{D}^{2k-1},$$

and this relation is false by virtue of Lemma 8. Thus $\overline{D}^{2k-1}$ is non-convex for $k \geqq 2$.

Again, the characteristic roots of the d.s. matrix

$$\left\| \begin{matrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{matrix} \right\| \oplus \left\| \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\| \oplus \ldots \oplus \left\| \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\|,$$

where the second summand is repeated $k-1$ times, are 1^k, 0^2, $(-1)^{k-1}$. Hence

$$(1^{k-1}, 0^2, (-1)^{k-1}) \in \overline{D}^{2k}$$

and therefore

$$(1, 0, 0, 1, -1, 1, -1, \ldots 1, -1, -1) \in \overline{D}^{2k},$$

$$(0, 1, 0, -1, 1, -1, 1, \ldots -1, 1, -1) \in \overline{D}^{2k}.$$

If $\overline{D}^{2k}$ were convex, it would follow that

$$((1/2)^2, 0^{2k-3}, -1) \in \overline{D}^{2k};$$

and this is again impossible in view of Lemma 8. Dence $\overline{D}^{2k}$ is non-convex for $k \geqq 2$, and the proof is complete.

§ 8. D.s. matrices of order 3

8.1. We begin the discussion of d.s. matrices of order 3 by sharpening Theorems 1 and 2 for the case $n = 3$.

Theorem 12. We have $\Omega_3 = \Pi_2 \cup \Pi_3$.

In view of Theorem 1, it suffices to show that $\Omega_3 \subseteq \Pi \cup \Pi_3$. If $z \in \Omega_3$ and z is real, then, by Lemma 1(b), $-1 \leqq z \leqq 1$, i.e. $z \in \Pi_2$. If $z \in \Omega_3$ and z is non-real, write $z = \sigma + i\tau$ (σ, τ real). Then there exists a d.s. 3×3 matrix $D = (d_{rs})$ with characteristic roots 1, $\sigma + i\tau$, $\sigma - i\tau$. We have

$$1 + 2\sigma = \operatorname{tr} D \geqq 0$$

and so

$$\sigma \geqq -1/2. \tag{8.1}$$

Again, the sum of products of pairs of characteristic roots of D is $2\sigma + \sigma^2 + \tau^2$, and therefore

$$2\sigma + \sigma^2 + \tau^2 = \sum_{1 \leq r < s \leq 3} \begin{vmatrix} d_{rr} & d_{rs} \\ d_{sr} & d_{ss} \end{vmatrix} \leqq \sum_{1 \leq r < s \leq 3} d_{rr}\, d_{ss}$$
$$\leqq 1/3 \left(\sum_{r=1}^{3} d_{rr} \right)^2 = 1/3 (\operatorname{tr} D)^2 = 1/3 (1 + 2\sigma)^2.$$

Hence $3\tau^2 \leqq (1 - \sigma)^2$ and, since $1 - \sigma \geqq 0$, we obtain

$$|\tau| \leqq \frac{1 - \sigma}{\sqrt{3}}. \tag{8.2}$$

From (8.1) and (8.2) we infer that $\sigma + i\tau \in \Pi_3$. The proof is therefore complete.

Theorem 13. We have $\Omega_3^0 = \Pi_3^0$.

A d.s. 3×3 matrix with zero trace must have the form

$$\begin{Vmatrix} 0 & a & 1 - a \\ 1 - a & 0 & a \\ a & 1 - a & 0 \end{Vmatrix},$$

where $0 \leqq a \leqq 1$. The characteristic roots of this matrix are

$$1, \; -1/2 \pm i(a - 1/2)\sqrt{3},$$

and the assertion follows.

8.2. Next, we give a complete characterization of the regions $\mathfrak{D}^3$ and D^3.

Theorem 14. Three given numbers can serve as the characteristic roots of a d.s. 3×3 *matrix if and only if they are either* $1, z, \bar{z}$, *where* z *is a non-real number in* Π_3, *or else* $1, \lambda, \mu$, *where* λ, μ *are real numbers such that*[10]

$$1 \geqq \lambda \geqq \mu, \tag{8.3}$$

$$\lambda + 3\mu + 2 \geqq 0. \tag{8.4}$$

[10] Naturally we need not assume that $\lambda \geqq \mu$. This is merely a notational convenience, and we can dispense with it if we replace (8.3) and (8.4) by the inequalities

$$\lambda \leqq 1, \; \mu \leqq 1, \; \lambda + 3\mu + 2 \geqq 0, \; 3\lambda + \mu + 2 \geqq 0.$$

By Theorem 12 we know that, if z is non-real and $z \in \Pi_3$, then $1, z, \bar{z}$ are the characteristic roots of a d.s. 3×3 matrix. Conversely, if a d.s. 3×3 matrix has a non-real characteristic root z, then the set of its characteristic roots is $1, z, \bar{z}$ and $z \in \Pi_3$.

Next, let λ, μ be real numbers satisfying (8.3) and (8.4). Then, by Theorem 8, there exists a d.s. 3×3 matrix with characteristic roots $1, \lambda, \mu$. Indeed, one such matrix is

$$\frac{1}{6}\begin{Vmatrix} 2+4\lambda & 2-2\lambda & 2-2\lambda \\ 2-2\lambda & 2+\lambda+3\mu & 2+\lambda-3\mu \\ 2-2\lambda & 2+\lambda-3\mu & 2+\lambda+3\mu \end{Vmatrix}.$$

Again, if $1, \lambda, \mu$ (where $\lambda \geqq \mu$) are the characteristic roots of a d.s. 3×3 matrix D, then plainly $1 \geqq \lambda$ and it remains to show that the inequality (8.4) is valid. Let t, d denote respectively the trace and determinant of D and write

$$D = \begin{Vmatrix} p & q & 1-p-q \\ r & s & 1-r-s \\ 1-p-r & 1-q-s & p+q+r+s-1 \end{Vmatrix},$$

so that $p, q, r, s \geqq 0$ and $p+q+r+s \geqq 1$ (and four other linear inequalities are satisfied). Then

$$t = q + r + 2p + 2s - 1,$$
$$d = q + r - p - s - 3qr + 3ps,$$

and therefore

$$3t^2 + 2t + 4d - 1 = 3(q-r)^2 + 12(p+s)(p+q+r+s-1) + 12ps.$$

Consequently

$$3t^2 + 2t + 4d - 1 \geqq 0. \tag{8.5}$$

Now write

$$\theta = \lambda + 3\mu + 2, \quad \varphi = 3\lambda + \mu + 2.$$

Then $\theta + \varphi = 4t$ and so

$$\theta + \varphi \geqq 0. \tag{8.6}$$

Moreover

$$\theta\varphi = 3(\lambda+\mu+1)^2 + 2(\lambda+\mu+1) + 4\lambda\mu - 1$$
$$= 3t^2 + 2t + 4d - 1$$

and so, by (8.5),

$$\theta\varphi \geqq 0. \tag{8.7}$$

From (8.6) and (8.7) we infer that $\theta \geqq 0$ (and also $\varphi \geqq 0$). Thus (8.4) is satisfied and the proof of the theorem is complete.

We note, then, that $(1, \lambda, \mu) \in \mathrm{D}^3$ if and only if (λ, μ) belongs to the closed quadrilateral Q with vertices $(1, 1)$, $(1, -1)$, $(-1, 1)$, $(-\frac{1}{2}, -\frac{1}{2})$. This implies, in particular, that $\overline{\mathrm{D}}^2$ is convex.

8.3. Finally, we proceed to describe the characteristic polynomials of all d.s. 3×3 matrices. It will be convenient to say that a d.s. 3×3 matrix is of *type I* if all its characteristic roots are real and of *type II* if its characteristic roots are of the form $1, z, \bar{z}$. A d.s. matrix with characteristic roots $1, \alpha, \alpha$ (α real) belongs, of course, to both types. We shall write

$$\chi(x) = x^3 - c_1 x^2 + c_2 x - c_3.$$

Theorem 15. The real polynomial $\chi(x)$ is the characteristic polynomial of a d.s. 3×3 matrix of type I if and only if

$$c_2 = c_1 + c_3 - 1; \tag{8.8}$$

$$c_1 \geqq 0; \tag{8.9}$$

$$\max\{c_1 - 2, \tfrac{1}{4}(1 - 2c_1 - 3c_1^2)\} \leqq c_3 \leqq \tfrac{1}{4}(c_1 - 1)^2. \tag{8.10}$$

It is the characteristic polynomial of a d.s. 3×3 matrix of type II if and only if the relations (8.8) *and* (8.9) *are valid and*

$$\tfrac{1}{4}(c_1 - 1)^2 \leqq c_3 \leqq \tfrac{1}{3}c_1^2 - c_1 + 1.$$

Consequently, $\chi(x)$ is the characteristic polynomial of a d.s. 3×3 matrix if and only if (8.8) *and* (8.9) *are valid and*

$$\max\{c_1 - 2, \tfrac{1}{4}(1 - 2c_1 - 3c_1^2)\} \leqq c_3 \leqq \tfrac{1}{3}c_1^2 - c_1 + 1.$$

It is clear that $\chi(x)$ is the characteristic polynomial of a d.s. 3×3 matrix of type I if and only if

$$x^3 - c_1 x^2 + c_2 x - c_3 \equiv (x - 1)(x - \lambda)(x - \mu),$$

where $(\lambda, \mu) \in Q$, so that

$$c_1 = 1 + \lambda + \mu,\ c_2 = \lambda\mu + \lambda + \mu,\ c_3 = \lambda\mu.$$

Now the equations $c_1 = 1 + \lambda + \mu$, $c_3 = \lambda\mu$ map Q onto the region of the (c_1, c_3)-plane bounded by the lines $c_3 = c_1 - 2$, $c_1 = 0$ and the parabolas

$$c_3 = \tfrac{1}{4}(c_1 - 1)^2,\ c_3 = \tfrac{1}{4}(1 - 2c_1 - 3c_1^2).$$

Hence the relations (8.8), (8.9), and (8.10) are satisfied; and the converse inference is also clear.

Again, $\chi(x)$ is the characteristic polynomial of a d.s. 3×3 matrix of type II if and only if

$$x^3 - c_1 x^2 + c_2 x - c_3 \equiv (x-1)(x-\sigma-i\tau)(x-\sigma+i\tau),$$

where $\sigma + i\tau \in \Pi_3$. Hence

$$c_1 = 1 + 2\sigma,\ c_2 = 2\sigma + \sigma^2 + \tau^2,\ c_3 = \sigma^2 + \tau^2.$$

The equations $c_1 = 1 + 2\sigma$, $c_3 = \sigma^2 + \tau^2$ map the triangle Π_3 onto the region of the (c_1, c_3)-plane bounded by the line $c_1 = 0$ and the parabolas

$$c_3 = {}^1/_4(c_1 - 1)^2,\ c_3 = {}^1/_3 c_1^2 - c_1 + 1.$$

The assertion therefore follows. We note, in passing, that Theorem 15 contains necessary and sufficient conditions for two real numbers to serve as trace and determinant of a d.s. 3×3 matrix. We are not able, however, to solve the corresponding problem for the case of $n \times n$ matrices.

§ 9. D.s. matrices of order 4

9.1. Our next result is convenient for the computation of characteristic roots of certain partitioned matrices.

Lemma 9. Let A, B be $n \times n$ matrices. Then the characteristic roots of the matrix

$$M = \begin{Vmatrix} A & B \\ B & A \end{Vmatrix}$$

are identical with the characteristic roots of $A + B$ together with those of $A - B$.

Denoting by $\chi(x; M)$ the characteristic polynomial of M, we have

$$\begin{aligned}
\chi(x; M) &= \det\left\{ \begin{Vmatrix} I_n & I_n \\ O & I_n \end{Vmatrix} \cdot \begin{Vmatrix} x I_n - A & -B \\ -B & x I_n - A \end{Vmatrix} \cdot \begin{Vmatrix} I_n & -I_n \\ O & I_n \end{Vmatrix} \right\} \\
&= \det \begin{Vmatrix} x I_n - A - B & O \\ -B & x I_n - A + B \end{Vmatrix} \\
&= \det(x I_n - A - B).\ \det(x I_n - A + B) \\
&= \chi(x; A + B).\ \chi(x; A - B),
\end{aligned}$$

as required.

9.2. We now turn to the discussion of D^4. By Theorem 8, we know that, if λ, μ, ν are real numbers such that $1 \geqq \lambda \geqq \mu \geqq \nu$ and

$$3 + \lambda + 2\mu + 6\nu \geqq 0, \tag{9.1}$$

then $(1, \lambda, \mu, \nu) \in D^4$. We recall, however, that not all points in D^4 are characterized by these inequalities. We shall now determine a further subregion of D^4.

Theorem 16. Let λ, μ, ν be real numbers and suppose that

$$1 \geqq \lambda \geqq \mu \geqq \nu \tag{9.2}$$

and

$$|\lambda + \mu| \leqq 1 + \nu. \tag{9.3}$$

Then $(1, \lambda, \mu, \nu) \in D^4$.

We note, first of all, that the point $(1, 1, -1)$ of $\overline{D}^3$ belongs to the region specified by the equations (9.1) and (9.2) but not to that specified by (9.3) and (9.2), while for the point $(1, -1, -1)$ of $\overline{D}^3$ the situation is reversed. Thus neither of the Theorems 8 (for $n = 3$) and 16 implies the other. At the same time they do not, between them, provide a complete description of all sets of real characteristic roots of d.s. 4×4 matrices. Thus the characteristic roots of the matrix

$$\begin{Vmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{Vmatrix} \oplus \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$$

are $1, \lambda = 1, \mu = 0, \nu = -1$, but the numbers λ, μ, ν satisfy neither (9.1) nor (9.3).

To prove Theorem 16, we write $a = 1 + \lambda + \mu + \nu, b = 1 + \lambda - \mu - \nu$, $c = 1 + \mu - \nu - \lambda$, $d = 1 + \nu - \lambda - \mu$. Then $a, b, c, d \geqq 0$ and the matrix

$$\frac{1}{4} \begin{Vmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{Vmatrix}$$

is d.s. Its characteristic roots, which are easily computed by means of Lemma 9, are $1, \lambda, \mu, \nu$. The assertion is therefore proved.

Corollary. Let λ, μ, ν be real numbers. Then $1, \lambda, \mu, \nu$ are the characteristic roots of a d.s. 4×4 matrix with zero trace if and only if

$$1 \geqq \lambda, \mu, \nu \geqq -1, \quad 1 + \lambda + \mu + \nu = 0.$$

The stated conditions are obviously necessary. To show that they are also sufficient we shall assume, as may be done without loss of generality, that $\lambda \geqq \mu \geqq \nu$. Then the relations (9.2) and (9.3) are satisfied, and the Corollary follows by virtue of Theorem 16. In fact

$$\frac{1}{2}\begin{Vmatrix} 0 & 1+\lambda & 1+\mu & 1+\nu \\ 1+\lambda & 0 & 1+\nu & 1+\mu \\ 1+\mu & 1+\nu & 0 & 1+\lambda \\ 1+\nu & 1+\mu & 1+\lambda & 0 \end{Vmatrix}$$

is a d.s. matrix with zero trace and characteristic roots $1, \lambda, \mu, \nu$.

9.3. We have now determined all sets of four real numbers which can serve as characteristic roots of a d.s. 4×4 matrix with zero trace. If the given numbers are not all real, then our knowledge is less complete, but we state briefly such information as we possess. It will be convenient to say that the non-real number $\sigma + i\tau (\sigma, \tau \text{ real})$ is *admissible* if it is a characteristic root of a d.s. 4×4 matrix with zero trace. Thus $\sigma + i\tau$ (with $\tau \neq 0$) is admissible if and only if

$$(1, -1-2\sigma, \sigma + i\tau, \sigma - i\tau) \in \mathfrak{D}^4. \tag{9.4}$$

It is plain that a necessary condition for the validity of (9.4) is

$$-1 \leqq \sigma \leqq 0, \ \sigma^2 + \tau^2 \leqq 1.$$

To derive another necessary condition, let $\sigma + i\tau$ be admissible and denote by D a d.s. 4×4 matrix with characteristic roots $1, -1-2\sigma, \sigma \pm i\tau$. Then D^2 is also d.s., and we have

$$0 \leqq \operatorname{tr}(D^2) = 2 + 4\sigma + 6\sigma^2 - 2\tau^2,$$

so that

$$0 < |\tau| \leqq 1 + 2\sigma + 3\sigma^2.$$

Again, to obtain a sufficient condition for the validity of (9.4) we recall that, by Theorem 2, $\Pi_4^0 \subseteq \Omega_4^0$. Hence, if

$$-1 \leqq \sigma \leqq 0, \ 0 < |\tau| \leqq 1 + \sigma,$$

then $\sigma + i\tau$ is admissible. In particular, then, the numbers $\pm i$ are admissible and, in view of Theorem 6 (iii), they are the only numbers of unit modulus with this property.

References

1. *N. A. Dmitriev* and *E. B. Dynkin*, On the characteristic numbers of a stochastic matrix. Doklady Akad. Nauk SSSR N.S. **49**, 159–162 (1945).

2. *F. R. Gantmacher*, The Theory of Matrices (New York, 1959), 2 volumes.

3. *F. I. Karpelevich*, On the characteristic roots of matrices with non-negative elements. Izv. Akad. Nauk SSSR Ser. Mat. **15**, 361–383 (1951) (Russian).

4. *C. C. MacDuffee*, The Theory of Matrices (New York, 1946).

5. *M. Marcus, H. Minc,* and *B. Moyls,* Some results on non-negative matrices. J. Res. Nat. Bur. Standards **65B**, 205—209 (1961).

6. *L. Mirsky,* An Introduction to Linear Algebra (Oxford, 1955).

7. *L. Mirsky,* Results and problems in the theory of doubly-stochastic matrices. Z. Wahrscheinlichkeitstheorie **1**, 319—334 (1963).

8. *L. Mirsky,* Inequalities and existence theorems in the theory of matrices. J. Math. Analysis Appl. **9**, 99—118 (1964).

9. *L. Mirsky,* Inequalities for characteristic roots and singular values of complex matrices. Not yet published.

10. *H. Perfect,* Methods of constructing certain stochastic matrices. I. Duke Math. J. **20**, 395—404 (1953). II. Ibid. **22**, 305—311 (1955).

11. *I. Schur,* Über die charakteristischen Wurzeln einer linearen Substitution mit einer Anwendung auf die Theorie der Integralgleichungen. Math. Annalen **66**, 488—510 (1909).

12. *H. R. Suleimanova,* Stochastic matrices with real characteristic numbers. Doklady Akad. Nauk SSSR N.S. **66**, 343—345 (1949) (Russian).

13. *H. Wielandt,* Unzerlegbare, nicht-negative Matrizen. Math. Zeitschrift **52**, 642—648 (1950).

Department of Pure Mathematics,

University of Sheffield.