

ON A PROBLEM OF O. TAUSSKY

BERNARD W. LEVINGER AND RICHARD S. VARGA

Recently, O. Taussky raised the following question. Given a nonnegative $n \times n$ matrix $A = (a_{i,j})$, let $\dot{\Omega}_A$ be the set of all $n \times n$ complex matrices defined by

$$(1.1) \quad \dot{\Omega}_A \equiv \{B = (b_{i,j}) \mid |b_{i,j}| = a_{i,j} \text{ for all } 1 \leq i, j \leq n\}.$$

Then, defining the spectrum $S(\mathfrak{M})$ of an arbitrary set \mathfrak{M} of $n \times n$ matrices B as

$$(1.2) \quad S(\mathfrak{M}) \equiv \{\sigma \mid \det(\sigma I - B) = 0 \text{ for some } B \in \mathfrak{M}\},$$

what can be said in particular about $S(\dot{\Omega}_A)$? It is not difficult to see that $S(\dot{\Omega}_A)$ consists of possibly one disk and a series of annular regions concentric about the origin, but our main result is a precise characterization of $S(\dot{\Omega}_A)$ in terms of the minimal Gerschgorin sets for A .

Introduction. We shall distinguish between two cases. If there is a diagonal matrix $D = \text{diag}(x_1, \dots, x_n)$ with $\mathbf{x} \geq \mathbf{0}$ and $\mathbf{x} \neq \mathbf{0}$ such that AD is diagonally dominant, then A is called essentially diagonally dominant. In this case, the set $S(\dot{\Omega}_A)$ is just the minimal Gerschgorin set $G(\Omega_A)$ of [6], rotated about the origin (Theorem 1 and Corollary 2). Determining $S(\dot{\Omega}_A)$ in this case is quite easy, since it suffices to determine those points of the boundary of $G(\Omega_A)$ which lie on the positive real axis (Theorem 2). This is discussed in § 2.

In the general case when A is not essentially diagonally dominant, we must use permutations and intersections (Theorem 3) to fully describe $S(\dot{\Omega}_A)$, in the spirit of [3]. These results are described in § 3. Also in this section is a generalization (Theorems 3 and 4) of a recent interesting result by Camion and Hoffman [1]. Our proof of this generalization differs from that of [1].

Finally, in § 4 we give several examples to illustrate the various possibilities for $S(\dot{\Omega}_A)$.

Before leaving this section, we point out that the question posed by O. Taussky [5, p. 129] has an immediate answer in terms of the results of [3]. In [3], the authors completely characterized the spectrum $S(\Omega_C)$ of a related set Ω_C of matrices, where $C = (c_{i,j})$ was an arbitrary $n \times n$ complex matrix and

$$(1.3) \quad \Omega_C \equiv \{B = (b_{i,j}) \mid |b_{i,j}| = |c_{i,j}| \text{ and } b_{i,j} = c_{i,j} \text{ for all } 1 \leq i, j \leq n\}.$$

Clearly, $\Omega_A \subset \dot{\Omega}_A$. On the other hand, if $D(\theta)$ represents an $n \times n$ diagonal matrix all of whose diagonal entries have modulus unity:

$d_{j,j} = \exp(i\theta_j)$, $1 \leq j \leq n$, then $AD(\theta) \subset \mathring{\Omega}_A$ and $\mathring{\Omega}_A = \bigcup_{\theta} \Omega_{AD(\theta)}$, where the union is over all possible choices of $D(\theta)$. Thus,

$$(1.4) \quad S(\mathring{\Omega}_A) = \bigcup_{\theta} S(\Omega_{AD(\theta)}).$$

While this answers the question posed, it neither gives an insight into the nature of $S(\mathring{\Omega}_A)$, nor allows $S(\mathring{\Omega}_A)$ to be effectively calculated. We shall show that in fact $S(\mathring{\Omega}_A)$ is more easily determined than $S(\Omega_A)$.

2. The essentially diagonally dominant case. Let $A = (a_{i,j})$ be given $n \times n$ nonnegative matrix. In order to develop the material of this section, we recall some definitions and results concerning the *minimal Gerschgorin set* $G(\Omega_A)$ associated with A . In [3, 6], a continuous real-valued function $\nu(\sigma)$, defined for all complex numbers σ , was characterized by

$$(2.1) \quad \nu(\sigma) \equiv \inf_{u > 0} \max_i \left\{ \frac{1}{u_i} \left[\sum_{j \neq i} a_{i,j} u_j - |\sigma - a_{i,i}| u_i \right] \right\}.$$

Using the Perron-Frobenius theory of nonnegative matrices [7, § 2.4 and § 8.2], it can be shown that there exists a nonnegative vector $\mathbf{x} \neq \mathbf{0}$ such that

$$(2.1') \quad -|\sigma - a_{i,i}| x_i + \sum_{j \neq i} a_{i,j} x_j = \nu(\sigma) x_i, \quad 1 \leq i \leq n.$$

From $\nu(\sigma)$, $G(\Omega_A)$ is defined by

$$(2.2) \quad G(\Omega_A) = \{\sigma \mid \nu(\sigma) \geq 0\}.$$

In view of (2.1') and (2.2), a complex number σ is contained in $G(\Omega_A)$ if and only if there is a nonnegative vector $\mathbf{x} \neq \mathbf{0}$ such that

$$(2.3) \quad |\sigma - a_{i,i}| x_i \leq \sum_{j \neq i} a_{i,j} x_j, \quad 1 \leq i \leq n.$$

The set $G(\Omega_A)$ is a closed bounded set, and its boundary, denoted by $\partial G(\Omega_A)$, satisfies,

$$(2.4) \quad \partial G(\Omega_A) \subset S(\Omega_A) \subset G(\Omega_A).$$

We first prove a result concerning $G(\Omega_A)$ which will have later applications.

LEMMA 1. *If, for $z_0 > 0$, $z_0 e^{i\theta} \in G(\Omega_A)$ for all real θ , then all z with $|z| \leq z_0$ are in $G(\Omega_A)$, and $z=0$ is an interior point of $G(\Omega_A)$.*

Proof. This is a simple application of (2.3). By assumption,

$-z_0 \in G(\Omega_A)$. Since $z_0 > 0$ and $a_{i,i} \geq 0, 1 \leq i \leq n$, then

$$|-z_0 - a_{i,i}| = z_0 + a_{i,i}.$$

Thus, for any z with $|z| \leq z_0$,

$$|z - a_{i,i}| \leq |z| + a_{i,i} \leq z_0 + a_{i,i},$$

and (2.3) holds for z with the same vector $\mathbf{x} \geq \mathbf{0}$ which satisfies (2.3) for $-z_0$, which completes the proof.

We next introduce the notion of rotating a given point set P about the origin. Let

$$(2.5) \quad \text{rot } P \equiv \{\sigma \mid \sigma e^{i\theta} \in P \text{ for some real } \theta\}.$$

With this notation, we have

LEMMA 2. $\text{rot } S(\dot{\Omega}_A) = S(\dot{\Omega}_A)$.

Proof. It is clear that $S(\dot{\Omega}_A) \subset \text{rot } S(\dot{\Omega}_A)$. If $\sigma \in \text{rot } S(\dot{\Omega}_A)$, then $\sigma e^{i\theta}$ is an eigenvalue of some B in $\dot{\Omega}_A$ and thus σ is an eigenvalue of $e^{-i\theta}B$. But $e^{-i\theta}B \in \dot{\Omega}_A$ and hence $\sigma \in S(\dot{\Omega}_A)$, which completes the proof.

This elementary result already establishes that the spectrum $S(\dot{\Omega}_A)$ can be described as the union of a family of circles concentric about the origin.

LEMMA 3. If $\sigma \in S(\dot{\Omega}_A)$, then $|\sigma| \in G(\Omega_A)$.

Proof. For any $\sigma \in S(\dot{\Omega}_A)$, there is a matrix $B = (b_{i,j})$ in $\dot{\Omega}_A$ and a vector $\mathbf{y} \neq \mathbf{0}$ such that $B\mathbf{y} = \sigma\mathbf{y}$. Equivalently, we have

$$(2.6) \quad (\sigma - b_{i,i})y_i = \sum_{j \neq i} b_{i,j}y_j, \quad 1 \leq i \leq n.$$

If we take absolute values in (2.6) and note that

$$|\sigma - b_{i,i}| \geq ||\sigma| - |b_{i,i}|| = ||\sigma| - a_{i,i}|,$$

we obtain

$$(2.7) \quad ||\sigma| - a_{i,i}|| |y_i| \leq |\sigma - b_{i,i}| |y_i| = \left| \sum_{j \neq i} b_{i,j}y_j \right| \leq \sum_{j \neq i} a_{i,j} |y_j|,$$

so that $|\sigma|$ satisfies (2.3) with the nonnegative vector $\mathbf{x} = |\mathbf{y}|$, which completes the proof.

From the definition (2.5), it follows that, if P and R are any sets with $P \subset R$, then $\text{rot } P \subset \text{rot } R$. Thus, (2.4) and Lemma 3 combine to give

COROLLARY 1. $\text{rot } \partial G(\Omega_A) \subset S(\dot{\Omega}_A) \subset \text{rot } G(\Omega_A)$.

We now study the case for which the inclusions of Corollary 1 become equalities.

THEOREM 1. *Let A be a nonnegative $n \times n$ matrix. Then, $\text{rot } \partial G(\Omega_A) = S(\mathring{\Omega}_A) = \text{rot } G(\Omega_A)$ if and only if $z = 0$ is not an interior point of $G(\Omega_A)$.*

Proof. First, assume that $z = 0 \notin \text{int } G(\Omega_A)$, and let σ be an arbitrary nonzero point of $\text{rot } G(\Omega_A)$, so that $\sigma e^{i\theta_0} \in G(\Omega_A)$ for some real θ_0 . The circle $|z| = |\sigma|$ cannot lie entirely in $G(\Omega_A)$. For otherwise, by Lemma 1, the entire disk $|z| \leq |\sigma|$ would be contained in $G(\Omega_A)$ and $z = 0$ would be an interior point of $G(\Omega_A)$. Thus, the circle $|z| = |\sigma|$ necessarily intersects the boundary $\partial G(\Omega_A)$, and there exists a real θ_1 such that $\sigma e^{i\theta_1} \in \partial G(\Omega_A)$. It follows that $\sigma \in \text{rot } \partial G(\Omega_A)$, and thus from Corollary 1, σ is also a point of $S(\mathring{\Omega}_A)$. To complete this part of the proof, we need only examine the point $z = 0$. Clearly, the statement that $0 \notin \text{int } G(\Omega_A)$ is equivalent to the statement that either $0 \in G'(\Omega_A)$, the complement of $G(\Omega_A)$, or $0 \in \partial G(\Omega_A)$. Thus, if $0 \in \text{rot } G(\Omega_A)$, i.e., $0 \in G(\Omega_A)$, then the previous remark shows that $0 \in \partial G(\Omega_A)$, which completes the proof of the first part. Now, assume that $\text{rot } \partial G(\Omega_A) = S(\mathring{\Omega}_A) = \text{rot } G(\Omega_A)$, and call this common set of points H . If $0 \in H$, then $0 \in \partial G(\Omega_A)$, and hence $0 \notin \text{int } G(\Omega_A)$. If $0 \notin H$, then $0 \notin G(\Omega_A)$, which implies that $0 \in G'(\Omega_A)$, and again $0 \notin \text{int } G(\Omega_A)$, which completes the proof.

The statement $z = 0 \notin \text{int } G(\Omega_A)$ can be seen to be equivalent to $\nu(0) \leq 0$, and this has an interesting connection with *diagonally dominant matrices*, i.e., $n \times n$ matrices $B = (b_{i,j})$ satisfying

$$(2.8) \quad |b_{i,i}| \geq \sum_{j \neq i} |b_{i,j}|, \quad 1 \leq i \leq n.$$

Obviously, if $\nu(0) \leq 0$, then from (2.1'), there is a nonnegative vector $\mathbf{y} \neq \mathbf{0}$ such that

$$(2.9) \quad a_{i,i}y_i \geq \sum_{j \neq i} a_{i,j}y_j, \quad 1 \leq i \leq n.$$

Thus, if D is the diagonal matrix $D \equiv \text{diag}(y_1, \dots, y_n)$, then (2.9) asserts that the product AD is diagonally dominant. Conversely, if $D = \text{diag}(y_1, \dots, y_n)$ where $\mathbf{y} \geq \mathbf{0}$ and $\mathbf{y} \neq \mathbf{0}$ and AD is diagonally dominant, then it follows from (2.3) that $\nu(0) \leq 0$.

The statement that $\nu(0) \leq 0$ can also be coupled with results of Ostrowski [4] on *H-matrices*, which are defined as follows. Let $B = (b_{i,j})$ be an arbitrary $n \times n$ complex matrix, and associate with B the new matrix $C = (c_{i,j})$, where $c_{i,j} = -|b_{i,j}|$, $i \neq j$, and

$$c_{i,i} = |b_{i,i}|, \quad 1 \leq i \leq n.$$

Then, B is an H -matrix if and only if all the principal minors of C are nonnegative. [That is, the matrix C is a possibly degenerate M -matrix.] In [4], it is shown that B is an H -matrix if and only if there exists a diagonal matrix $D = \text{diag}(y_1, \dots, y_n)$ with $\mathbf{y} \geq \mathbf{0}$, $\mathbf{y} \neq \mathbf{0}$, such that BD is diagonally dominant. Thus we have

COROLLARY 2. *Let A be a nonnegative $n \times n$ matrix. Then, $\text{rot } \partial G(\Omega_A) = S(\mathring{\Omega}_A) = \text{rot } G(\Omega_A)$ if and only if A is an H -matrix.*

Summarizing, we have shown that the sets $\text{rot } \partial G(\Omega_A)$, $S(\mathring{\Omega}_A)$, and $\text{rot } G(\Omega_A)$ are equal in the case that A is an H -matrix, and this might logically be called the essentially diagonally dominant case, the title of this section.

We have already shown that $S(\mathring{\Omega}_A)$ is a collection of annuli and disks concentric about the origin. It is now logical to ask how the radii of these regions can be determined. For convenience, we will assume that A is *irreducible* (cf. [7, p. 20]). The reducible case requires only minor modifications.

We consider the function $\nu(t)$ along the nonnegative real axis $t \geq 0$. Let $\{t_i\}_{i=1}^m$ define the finite sequence of points $t_1 > t_2 > \dots > t_m > 0$, such that $\nu(t_i) = 0$ and $\nu(t_i + \varepsilon) \cdot \nu(t_i - \varepsilon) < 0$ for all sufficiently small $\varepsilon > 0$. Then, these points t_i indicate strong sign changes in $\nu(t)$. In [6], it was shown that the spectral radius of A ,

$$\rho(A) \equiv \max_i \{ |\lambda_i| \mid \det(\lambda_i I - A) = 0 \},$$

is such a point, and since it was further shown that $\nu(\rho(A) + \delta) < 0$ for all $\delta > 0$, it is evidently the largest such point, i.e., $t_1 = \rho(A)$ and $m \geq 1$. We define $t_{m+1} = 0$, and now show that the points t_i divide the nonnegative real axis into intervals in which $\nu(t) \geq 0$.

LEMMA 4. *For $t \geq 0$, $\nu(t) \geq 0$ if and only if $t_{2i} \leq t \leq t_{2i-1}$ for some i with $1 \leq i \leq [(m + 1)/2]$.*

Proof. Since $\nu(t)$ is continuous for $t \geq 0$, it suffices to show that there is no $\mu > 0$, corresponding to a degenerate change of signs, with $\nu(\mu) = 0$ such that $\nu(\mu - \varepsilon) < 0$ and $\nu(\mu + \varepsilon) < 0$ for all sufficiently small $\varepsilon > 0$. This assertion is basically a consequence of the assumption that A is irreducible. For, if such a $\mu > 0$ exists, then $\mu \in \partial G(\Omega_A)$. Moreover, since $|te^{i\theta} - a_{i,i}| > |t - a_{i,i}|$ for any $t > 0$ and any real θ with $0 < |\theta| \leq \pi$, it follows from (2.1) that $\nu(te^{i\theta}) < \nu(t)$ and hence that $\nu(z) < 0$ for all complex $z \neq \mu$ in a neighborhood of μ . Thus, μ is an *isolated* point of $G(\Omega_A)$. As such, it follows [6] that μ is necessarily a diagonal entry of A , i.e., $\mu = a_{j,j}$ for some j . But, since

A is irreducible, it is known [6] that $\nu(a_{k,k}) > 0$ for every $1 \leq k \leq n$. This contradiction establishes the desired result.

THEOREM 2. *Let A be a nonnegative irreducible $n \times n$ matrix, and let $t_1 > t_2 > \cdots > t_m > 0$ be positive real numbers such that $\nu(t_i) = 0$ and $\nu(t_i + \varepsilon) \cdot \nu(t_i - \varepsilon) < 0$ for all sufficiently small $\varepsilon > 0$. If $m > 1$ and z is any complex number with $|z| \geq t_{2[m/2]}$, then $z \in S(\Omega_A)$ if and only if $t_{2i} \leq |z| \leq t_{2i-1}$ for some i with $1 \leq i \leq [m/2]$.*

Proof. If z_0 is any complex number with $|z_0| \geq t_{2[m/2]}$ and $t_{2i} \leq |z_0| \leq t_{2i-1}$ for some $1 \leq i \leq [m/2]$, then from Lemma 4, $\nu(|z_0|) \geq 0$. Also, from Lemma 4 it follows that $\nu(|z|) < 0$ for any $|z|$ with $t_{2i+1} < |z| < t_{2i}$. Thus, all points in the disk $|z| \leq |z_0|$ are not points of $G(\dot{\Omega}_A)$, and we deduce from Lemma 1 that $|z_0| e^{i\theta} \in \partial G(\Omega_A)$ for some real θ . Thus, $z_0 \in \text{rot } \partial G(\Omega_A)$, and thus from Corollary 1, $z_0 \in S(\dot{\Omega}_A)$, which proves one part of this result. Conversely, for any $z_0 \in S(\dot{\Omega}_A)$ with $|z_0| \geq t_{2[m/2]}$, $\nu(|z_0|) \geq 0$ from Lemma 3. Then from Lemma 4, it follows that $t_{2i} \leq |z_0| \leq t_{2i-1}$ for some i with $1 \leq i \leq [m/2]$, which completes the proof.

Using the results of [6], it is now simple to determine the exact number of eigenvalues of any matrix $B \in \dot{\Omega}_A$ which lie in each of the outer annuli: $t_{2i} \leq |z| \leq t_{2i-1}$ for $1 \leq i \leq [m/2]$.

COROLLARY 3. *Let A be a nonnegative irreducible $n \times n$ matrix with $m > 1$. Then, for any $B \in \dot{\Omega}_A$, B has p_i eigenvalues in the annulus $t_{2i} \leq |z| \leq t_{2i-1}$, $1 \leq i \leq [m/2]$, if and only if A has p_i diagonal entries in this annulus.*

Proof. By a familiar continuity argument, going back to Gerschgorin, each connected component of $S(\dot{\Omega}_A)$ contains the same number of eigenvalues for each $B \in \dot{\Omega}_A$, and hence, the same number as A . But from [6], A has p_i eigenvalues in this annulus if and only if A has p_i diagonal entries in this annulus, which completes the proof.

As final remarks in this section, we mention that Theorem 2 precisely gives $S(\dot{\Omega}_A)$ and the radii of its associated concentric annuli in the case that m (the number of strong sign changes in $\nu(t)$ for $t \geq 0$) is *even*. In this regard, it is interesting to point out that the geometrical result of Theorem 1 and Corollary 2 is basically contained in Theorem 2, since it can be obtained by applying Theorem 2 to a family of nonnegative irreducible matrices $A(\varepsilon)$, $\varepsilon \geq 0$, where $A(\varepsilon) \rightarrow A$

as $\varepsilon \downarrow 0$, for which m is again even for each $A(\varepsilon)$ for all sufficiently small $\varepsilon > 0$. We also mention that computing the points t_i or Theorem 2, whether m is even or odd, is not difficult because of the inclusion relationships of (2.1).

In the case that $m = 2l + 1$ is odd, Theorem 2 gives no information about the final disk $0 \leq |z| \leq t_{2l+1}$, and different techniques are necessary to decide which points of this disk are points of $S(\mathring{\Omega}_A)$. This will be discussed in § 3.

3. $\nu(0) > 0$. If $z = 0$ is an interior point of $G(\Omega_A)$, i.e., $\nu(0) > 0$, we can still give a precise characterization of $S(\mathring{\Omega}_A)$ using the methods of [3], but these results are considerably more complicated than those given in § 2. We shall show by means of examples in § 4 that these complications cannot, unfortunately, be avoided.

We first give a more or less well known result.

LEMMA 5. *Let $0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$ be nonnegative real numbers, and ρ an arbitrary complex number. Then, there exist real numbers $\theta_1, \dots, \theta_n$ such that $\rho = \sum_{j=1}^n \alpha_j e^{i\theta_j}$ if and only if*

$$(3.1) \quad \sum_{j=1}^n \alpha_j \geq |\rho| \geq \alpha_n - \sum_{j=1}^{n-1} \alpha_j .$$

Proof. This lemma is precisely Lemma 1 of [1] applied to the $n + 1$ nonnegative numbers $\alpha_1, \dots, \alpha_n, |\rho|$. However, for completeness, we give a proof by induction.

Only the fact that (3.1) implies the existence of the θ_j is nontrivial. For $n = 2$, $|\alpha_2 + \alpha_1 e^{i\theta}| = \sqrt{\alpha_2^2 + 2\alpha_1\alpha_2 \cos \theta + \alpha_1^2}$ which varies continuously from $\alpha_2 + \alpha_1$ to $\alpha_2 - \alpha_1$ as θ varies from 0 to π .

For $n + 1$, we distinguish two cases. Consider first the case where $|\rho| \geq |\alpha_{n+1} - \sum_{i=1}^n \alpha_i|$. Then, as in the previous case for $n = 2$, for some θ we can write $|\rho| = |\alpha_{n+1} + e^{i\theta} \sum_{i=1}^n \alpha_i|$. Otherwise, if $|\rho| < |\alpha_{n+1} - \sum_{i=1}^n \alpha_i|$, then from (3.1) we deduce that $|\rho| < \sum_{i=1}^n \alpha_i - \alpha_{n+1}$, which gives us the inequalities

$$\alpha_n - \sum_{i=1}^{n-1} \alpha_i \leq \alpha_n \leq |\rho| + \alpha_{n+1} \leq \sum_{i=1}^n \alpha_i .$$

Thus, from the inductive hypothesis, $\alpha_{n+1} + |\rho|$, and hence also ρ , have the representations of the desired form.

With this, we now characterize $S(\mathring{\Omega}_A)$ by a set of linear inequalities.

LEMMA 6. *Let σ be an arbitrary complex number. Then $\sigma \in S(\mathring{\Omega}_A)$ if and only if there exists a nonnegative vector $\mathbf{x} \neq \mathbf{0}$ such that*

$$(3.2) \quad \sum_{j=1}^n a_{i,j}x_j \geq |\sigma| x_i \geq a_{i,k}x_k - \sum_{j \neq k} a_{i,j}x_j$$

for each i and k with $1 \leq i, k \leq n$.

Proof. If $\sigma \in S(\mathring{\Omega}_A)$, there exists a matrix $B \in \mathring{\Omega}_A$ and a vector $\mathbf{z} \neq \mathbf{0}$ with $B\mathbf{z} = \sigma\mathbf{z}$. Taking absolute values and setting $|z_j| = x_j$, we obtain for the i -th component

$$\sum_{j=1}^n a_{i,j}x_j \geq \left| \sum_{j=1}^n b_{i,j}x_j \right| = |\sigma| x_i \geq a_{i,k}x_k - \sum_{j \neq k} a_{i,j}x_j,$$

for each $1 \leq k \leq n$, which establishes the first part of this theorem. Conversely, if (3.2) is satisfied by a nonnegative vector $\mathbf{x} \neq \mathbf{0}$ for each i and $k, 1 \leq i, k \leq n$, we can repeatedly apply Lemma 5 to find real constants $\theta_{k,j}$ such that $\sigma x_k = \sum_{j=1}^n a_{k,j}e^{i\theta_{k,j}}x_j$ for $1 \leq k \leq n$, so that $\sigma \in S(\mathring{\Omega}_A)$, which completes the proof.

We now remark that the inequalities of (3.2) are equivalent to the following set of n^2 linear inequalities

$$(3.3) \quad \sum_{j \neq i} (-1)^{\delta_{j,k}} a_{i,j}x_j + (-1)^{\delta_{i,k}} |\sigma| + (-1)^{\delta_{i,k}} a_{i,i} |x_i| \geq 0, \\ 1 \leq i, k \leq n,$$

where $\delta_{i,k}$ is the Kronecker delta function. For $k \neq i$, the second inequality of (3.2) is identical with (3.3). For $k = i$, (3.2) yields

$$\sum_{j \neq i} a_{i,j}x_j \geq (|\sigma| - a_{i,i})x_i \geq - \sum_{j \neq i} a_{i,j}x_j,$$

which is equivalent to

$$\sum_{j \neq i} a_{i,j}x_j - (|\sigma| - a_{i,i})x_i \geq 0.$$

In order to develop the material of this section, we recall some definitions and results [3] concerning the minimal Gerschgorin set $G^\varphi(\Omega_\sigma)$ associated with a matrix C relative to the permutation φ . Let $C = (c_{i,j})$ be an arbitrary $n \times n$ complex matrix, and let φ be any permutation of the first n positive integers. If σ is any complex number, we can define a continuous real valued function $\nu_{\varphi,\sigma}(\sigma)$ by

$$(3.4) \quad \nu_{\varphi,\sigma}(\sigma) = \inf_{u > 0} \max_i \left\{ \frac{1}{u_{\varphi(i)}} \left[\sum_{j \neq i} (-1)^{\delta_{j,\varphi(i)}} |c_{i,j}| u_j \right. \right. \\ \left. \left. + (-1)^{\delta_{i,\varphi(i)}} |\sigma - c_{i,i}| u_i \right] \right\}.$$

The minimal Gerschgorin set $G^\varphi(\Omega_\sigma)$ is given as in (2.2) by

$$(3.5) \quad G^\varphi(\Omega_\sigma) = \{ \sigma \mid \nu_{\varphi, \sigma}(\sigma) \geq 0 \} .$$

Equivalently, $\sigma \in G^\varphi(\Omega_\sigma)$ if and only if there exists a nonnegative vector $\mathbf{x} \neq \mathbf{0}$ such that

$$(3.6) \quad \sum_{j \neq i} (-1)^{\delta_{j, \varphi(i)}} |c_{i,j}| x_j + (-1)^{\delta_{i, \varphi(i)}} | \sigma - c_{i,i} | x_i \geq 0 , \quad 1 \leq i \leq n .$$

In order to couple the inequalities (3.3) to those of (3.6), let $A^\varphi = (a_{i,j}^\varphi)$ be an $n \times n$ matrix derived from A as follows:

$$(3.7) \quad a_{i,j}^\varphi = \begin{cases} a_{i,j} , & j \neq i \\ (-1)^{1+\delta_{i, \varphi(i)}} a_{i,i} , & j = i \end{cases} , \quad 1 \leq i, j \leq n .$$

It is clear from Lemma 6 and the definition of A_φ that $\sigma \in S(\mathring{\Omega}_A)$ implies that $|\sigma| \in G^\varphi(\Omega_{A^\varphi})$ for each permutation φ . Note that this result generalizes Lemma 3 of § 2 to arbitrary permutation. Hence, it follows that $|\sigma| \subset \bigcap_\varphi G^\varphi(\Omega_{A^\varphi})$, so that

$$(3.8) \quad S(\mathring{\Omega}_A) \subset \text{rot} \left(\bigcap_\varphi G^\varphi(\Omega_{A^\varphi}) \right) .$$

We now show that equality is valid in (3.8).

THEOREM 3. *Let $A = (a_{i,j})$ be a nonnegative $n \times n$ matrix. Then,*

$$S(\mathring{\Omega}_A) = \text{rot} \left(\bigcap_\varphi G^\varphi(\Omega_{A^\varphi}) \right) .$$

Proof. From (3.8), it suffices to show that $|\sigma| \in \bigcap_\varphi G^\varphi(\Omega_{A^\varphi})$ implies that $|\sigma| \in S(\mathring{\Omega}_A)$. To prove this, we define the sets $M_{i,k}(|\sigma|)$ from (3.3) by

$$(3.9) \quad M_{i,k}(|\sigma|) = \left\{ \mathbf{x} \geq \mathbf{0} \mid \sum_{j=1}^n x_j = 1 ; \sum_{j \neq i} (-1)^{\delta_{j,k}} a_{i,j} x_j + (-1)^{\delta_{i,k}} | |\sigma| + (-1)^{\delta_{i,k}} a_{i,i} | x_i \geq 0 \right\} .$$

By (3.3), $|\sigma| \in S(\mathring{\Omega}_A)$ is equivalent to the existence of a vector \mathbf{x} with

$$\mathbf{x} \in \bigcap_{1 \leq i, k \leq n} M_{i,k}(|\sigma|) ,$$

and thus we must prove that $\bigcap_{1 \leq i, k \leq n} M_{i,k}(|\sigma|)$ is nonempty. We shall show that the hypothesis, $|\sigma| \in \bigcap_\varphi G^\varphi(\Omega_{A^\varphi})$, implies that any n of the sets $M_{i,k}(|\sigma|)$ have a nonempty intersection. Then, the conclusion will follow from Helly's Theorem [2, p. 33], which states that *if K is a family of at least n convex sets in Euclidean $(n - 1)$ -space,*

R^{n-1} , such that every subclass containing n members has a common point in R^{n-1} , there is a point common to all members of K . Since the $M_{i,k}(|\sigma|)$ are convex and of dimension at most $(n - 1)$, this implies our theorem.

It remains to show that any collection $\{M_{i_j,k_j}(|\sigma|)\}_{j=1}^n$ has a nonempty intersection. This is always true if the second subscript k_j fails to take on the integer value k_0 , $1 \leq k_0 \leq n$. For, if \mathbf{y} is the vector with components $y_{k_0} = 1$, $y_j = 0$ for $j \neq k_0$, we see that (3.3) is satisfied and thus $\mathbf{y} \in \bigcap_{j=1}^n M_{i_j,k_j}(|\sigma|)$. By (3.6) and (3.7), the condition $|\sigma| \in G^\varphi(\Omega_{A^\varphi})$ is equivalent to the assertion that $\bigcap_\varphi M_{i,\varphi(i)}(|\sigma|)$ is nonempty. Thus, $|\sigma| \in \bigcap_\varphi G^\varphi(\Omega_{A^\varphi})$ implies that $\bigcap_{j=1}^n M_{i_j,k_j}(|\sigma|)$ is nonempty whenever $k_j = \varphi(i_j)$ for some permutation φ . Finally, consider a collection $\{M_{j(k),k}\}_{k=1}^n$ where $j(k)$ is not one-to-one. In this case, there is evidently a repeated first index, and for convenience, we assume that $1 = j(1) = j(2) = \dots = j(r)$, $r \geq 2$. Then let \mathbf{y} be any nonnegative vector with $y_1 + y_2 = 1$, $y_j = 0$ for $2 < j \leq n$. For such vectors, it follows from (3.9) that

$$(3.10) \quad \mathbf{y} \in M_{1,1} \text{ if and only if } a_{1,2}y_2 - (|\sigma| - a_{1,1})y_1 \geq 0,$$

$$(3.10') \quad \mathbf{y} \in M_{1,2} \text{ if and only if } -a_{1,2}y_2 + (|\sigma| + a_{1,1})y_1 \geq 0,$$

$$(3.10'') \quad \mathbf{y} \in M_{j(k),k}, k > 2 \text{ if and only if } a_{j(k),1}y_1 + a_{j(k),2}y_2 \geq 0.$$

Clearly, from (3.10'') all such vectors \mathbf{y} are in $\bigcap_{k>2} M_{j(k),k}$. If $a_{1,2} > 0$, then the vector \mathbf{y} with $y_2 = (|\sigma| - a_{1,1})y_1/a_{1,2}$ is in $M_{1,1} \cap M_{1,2}$, and if $a_{1,2} = 0$, then the vector \mathbf{y} with $y_2 = 1$, $y_1 = 0$ is in $M_{1,1} \cap M_{1,2}$. Thus, $\bigcap_{k=1}^n M_{j(k),k}$ is nonempty, and we conclude that any collection of n sets $M_{i,j}$ has a nonempty intersection, which completes the proof.

We can further show that, if $\sigma \in S(\mathring{\Omega}_A)$, then as in [1] there is a unique permutation φ such that $|\sigma| \in G^\varphi(\Omega_{A^\varphi})$. This will permit us to show that at most $(n + 1)$ permutations are necessary to characterize $S(\mathring{\Omega}_A)$ in Theorem 3.

THEOREM 4. *If $\sigma \in S(\mathring{\Omega}_A)$, then there exists a unique permutation φ such that $|\sigma| \in G^\varphi(\Omega_{A^\varphi})$.*

Proof. If $\sigma \in S(\mathring{\Omega}_A)$, then, by Theorem 3, there is at least one permutation φ with $|\sigma| \in G^\varphi(\Omega_{A^\varphi})$. Thus, if $|\sigma| \in G^\psi(\Omega_{A^\psi})$, we must show that $\psi = \varphi$, i.e., $\psi(i) = \varphi(i)$ for $1 \leq i \leq n$.

To prove this, we introduce the sets

$$(3.11) \quad N_{i,k} = \left\{ \mathbf{x} \geq \mathbf{0} \mid \sum_{j=1}^n x_j = 1; \sum_{j \neq i} (-1)^{\delta_j k} a_{i,j} x_j + (-1)^{\delta_{i,k}} (|\sigma| + (-1)^{\delta_{i,k}} a_{i,i}) x_i < 0 \right\},$$

with $1 \leq i, k \leq n$. Clearly, $N_{i,k}$ is the complement of $M_{i,k}(|\sigma|)$ relative to the $(n-1)$ -simplex $S \equiv \{x \geq 0 \mid \sum_{j=1}^n x_j = 1\}$. It is also clear that $N_{i,k}$ is empty if and only if $a_{i,k} = 0$ when $i \neq k$, and $|\sigma| - a_{i,i} = 0$ when $i = k$, and $N_{i,k}$ does not intersect the face of the simplex S defined by $x_k = 0$. Further, it is readily verified that $N_{i,k} \cap N_{i,k'}$ is empty if $k \neq k'$.

If $|\sigma| \notin G^\varphi(\Omega_{A^\varphi})$, it follows from (3.6) and (3.7) that $S = \bigcap_{i=1}^n N_{i,\varphi(i)}$. On the other hand, $|\sigma| \notin G^\varphi(\Omega_{A^\varphi})$ implies from (3.5) that $\nu_{\varphi,A^\varphi}(|\sigma|) < 0$, and hence, from the definition of (3.4), there must exist (by continuity) a positive vector $u > 0$ with $u \in N_{i,\varphi(i)}$ for all $1 \leq i \leq n$, i.e., if u is normalized, then $u \in \bigcap_{i=1}^n N_{i,\varphi(i)}$. Similarly, $|\sigma| \notin G^\psi(\Omega_{A^\psi})$ implies that $S = \bigcap_{i=1}^n N_{i,\psi(i)}$.

Now, let $I = \{j \mid \psi(j) = \varphi(j), 1 \leq j \leq n\}$. Assuming that $\psi \neq \varphi$, then I is a proper subset of the first n positive integers. From the vector $u > 0$ above, form the vector $v \in S$ as follows: $v_{\varphi(j)} = 0, j \in I$; $v_{\varphi(j)} = u_{\varphi(j)}, j \notin I$. Since $u \in N_{i,\varphi(i)}$ for all $1 \leq i \leq n$, it is easy to verify that $v \in N_{i,\varphi(i)}$ for any $i \notin I$, and thus $v \in \bigcap_{i \notin I} N_{i,\varphi(i)}$. Furthermore, $v \in \bigcup_{i \in I} N_{i,\psi(i)}$ since the union of the $N_{i,\psi(i)}$ covers the simplex S , and $N_{j,\psi(j)}$ does not intersect the face $v_{\varphi(j)} = 0$ for $j \in I$. Thus, there is a $k \notin I$ such that $v \in N_{k,\psi(k)} \cap N_{k,\varphi(k)}$. But since $N_{i,k} \cap N_{i,k'}$ is empty if $k \neq k'$, then it follows that $\psi(k) = \varphi(k)$, i.e., $k \in I$, which contradicts the assumption that I is a proper subset of the first n positive integers. Hence, $\varphi(i) = \psi(i)$ for all $1 \leq i \leq n$, which completes the proof.

We remark that the special case $\sigma = 0$ of Theorems 3 and 4 corresponds to the main results of [1].

Letting R' denote the complement of any set R in the complex plane, then Theorem 4 implies:

COROLLARY 4. *If K is an open connected component of $(S(\dot{\Omega}_A))'$, the complement of $S(\Omega_A)$, then there is a unique permutation ψ for which $K \subset (G^\psi(\Omega_{A^\psi}))'$.*

Proof. Since $\bigcap_\varphi G^\varphi(\Omega_{A^\varphi}) \subset S(\dot{\Omega}_A)$ by Theorem 3, then obviously $(S(\Omega_A))' \subset (\bigcap_\varphi G^\varphi(\Omega_{A^\varphi}))' = \bigcup_\varphi (G^\varphi(\Omega_{A^\varphi}))'$. Next, we remark that if $|\sigma|$ were replaced by σ in the definition of $N_{i,k}$ in (3.11), all subsequent arguments remain valid. In particular, from the proof of Theorem 4, it follows that the $(G^\varphi(\Omega_{A^\varphi}))'$ are nonintersecting open sets. Thus, the open connected component K can be in only one set $(G^\psi(\Omega_{A^\psi}))'$, which completes the proof. We remark that in general $K \neq (G^\psi(\Omega_{A^\psi}))'$ because of the rotational invariance of any connected component of $(S(\dot{\Omega}_A))'$.

We now consider the closed connected components of $S(\dot{\Omega}_A)$.

THEOREM 5. *Every connected component of $S(\mathring{\Omega}_A)$ contains the same number of eigenvalues for each matrix B in $\mathring{\Omega}_A$.*

Proof. This is basically a continuity argument. For, given any matrix $B \in \mathring{\Omega}_A$, we can construct a matrix $B(t) \in \mathring{\Omega}_A$ whose entries are continuous functions of t , $0 \leq t \leq 1$, such that $B(0) = A$ and $B(1) = B$. Since the eigenvalues of $B(t)$ then vary continuously with t , each matrix $B \in \mathring{\Omega}_A$ must have the same number of eigenvalues as A in each connected component of $S(\mathring{\Omega}_A)$, which completes the proof.

Theorem 3 states that $S(\mathring{\Omega}_A)$ can be determined from the $n!$ sets $G^\varphi(\Omega_{A^\varphi})$. The next result shows that at most $(n+1)$ permutations are necessary for the determination of $S(\mathring{\Omega}_A)$.

THEOREM 6. *There exist permutations $\varphi_1, \varphi_2, \dots, \varphi_r$ with $r \leq n+1$ such that $S(\mathring{\Omega}_A) = \text{rot}(\bigcap_{i=1}^r G^{\varphi_i}(\Omega_{A^{\varphi_i}}))$.*

Proof. Since the matrix A has n eigenvalues, then $S(\mathring{\Omega}_A)$ can have at most n closed connected components by Theorem 6. Because each closed connected component of $S(\mathring{\Omega}_A)$ is either a (possibly degenerate) disk or an annulus centered at the origin, then it is clear that the complement of $S(\mathring{\Omega}_A)$ consists of at most $(n+1)$ similar regions. By Corollary 3, exactly one permutation corresponds to each open connected component of $(S(\mathring{\Omega}_A))'$, and thus at most $(n+1)$ permutations are necessary to describe $S(\mathring{\Omega}_A)$.

We remark that, since $(S(\mathring{\Omega}_A))'$ always contains the unbounded connected component $\{z \mid |z| > \rho(A)\}$, the identity permutation must always occur as one of the r permutations of Theorem 6. This follows from the fact [3] that $G^\varphi(\Omega_{A^\varphi})$ is a bounded set only for the identity permutation. Of course, if A is essentially diagonally dominant, then $r = 1$ from Theorem 1. We now remark that the results of Theorem 2 and Corollary 3 can be used to obtain an improved upper bound for r . For, if t_m is, as in Theorem 2, the smallest positive number such that $\nu(t_m) = 0$, then by Corollary 3, the number of eigenvalues σ for each $B \in \mathring{\Omega}_A$ with $|\sigma| \geq t_m$ is equal to the number, k , of diagonal entries $a_{i,i}$ of A with $a_{i,i} \geq t_m$, and clearly $k \geq [m/2]$. Thus, by the same argument as above,

$$r \leq n + 1 - k.$$

In § 4, we give an example of a 3×3 matrix for which 3 permutations are required to determine $S(\mathring{\Omega}_A)$. In general, examples can similarly be given where n permutations are required for the $n \times n$ case, and we conjecture that the result of Theorem 6 is valid with

$n + 1$ reduced to n .

To actually calculate $S(\Omega_A)$ in the general case, it is necessary from Corollary 4 to work with the complements of the sets $G^\varphi(\Omega_{A^\varphi})$, i.e., to determine those intervals of the positive real axis ($t \geq 0$) for which $\nu_{\varphi, A^\varphi}(t) < 0$ for some permutation φ . However, it is in general not easy to determine a priori which $r(\leq n + 1)$ of the $n!$ permutations suffice to characterize $S(\mathring{\Omega}_A)$ in Theorem 6. For this reason, the analogue of Theorem 2 which could be stated for the general case seems computationally unattractive.

4. Examples. To illustrate the results of § 2, consider the following diagonally dominant matrix A :

$$(4.1) \quad A = \begin{bmatrix} 1 & 1/2 & 0 \\ 1/2 & 3 & 1/2 \\ 0 & 1/2 & 5 \end{bmatrix}.$$

For this matrix, the minimal Gerschgorin set $G(\Omega_A)$ is given by

$$(4.2) \quad G(\Omega_A) = \{z : 4 |z - 1| \cdot |z - 3| \cdot |z - 5| \leq |z - 5| + |z - 1|\}.$$

From this, it can be verified that the intervals of the nonnegative real axis for which $\nu(t) \geq 0$ are given by

$$(4.3) \quad 0.88 \leq t \leq 1.14; \quad 2.75 \leq t \leq 3.25; \quad 4.86 \leq t \leq 5.12.$$

From Theorem 2, $S(\mathring{\Omega}_A)$ then consists of three concentric annuli, and from Corollary 3, each $B \in \mathring{\Omega}_A$ has exactly one eigenvalue in each annulus.

To illustrate the results of § 3, consider the matrix $A(\varepsilon)$ where

$$(4.4) \quad A(\varepsilon) = \begin{bmatrix} \varepsilon & 1 & 0 \\ 0 & \varepsilon & 1 \\ 1 & 2 & \varepsilon \end{bmatrix},$$

and $\varepsilon \geq 0$. Note that $A(0)$ is the companion matrix for the polynomial $x^3 - 2x - 1$. It is not difficult to show that at most three permutations¹, $\varphi_1 = I$, $\varphi_2 = (23)$, $\varphi_3 = (123)$, are necessary to describe $S(\mathring{\Omega}_{A(\varepsilon)})$, i.e., $G^\varphi(\Omega_{A(\varepsilon)^\varphi})$ is the entire complex plane for all other permutations for every $\varepsilon \geq 0$. Thus, from Theorem 3, $S(\mathring{\Omega}_{A(\varepsilon)})$ is determined by the sets $G^{\varphi_i}(\Omega_{A(\varepsilon)^\varphi_i})$, which turn out to be

$$(4.5) \quad \begin{aligned} G^{\varphi_1}(\Omega_{A(\varepsilon)^\varphi_1}) &= \{\sigma : 1 + 2|\sigma - \varepsilon| - |\sigma - \varepsilon|^3 \geq 0\} \\ &= \{\sigma : |\sigma - \varepsilon| \leq 1.62\}, \end{aligned}$$

¹ Here, we are describing permutations by their disjoint cycles.

$$(4.6) \quad G^{\varphi_2}(\Omega_{A(\varepsilon)\varphi_2}) = \{\sigma : 1 - 2|\sigma - \varepsilon| - |\sigma - \varepsilon| \cdot |\sigma + \varepsilon|^2 \geq 0\},$$

$$(4.7) \quad G^{\varphi_3}(\Omega_{A(\varepsilon)\varphi_3}) = \{\sigma : -1 + 2|\sigma + \varepsilon| + |\sigma + \varepsilon|^3 \geq 0\} \\ = \{\sigma : |\sigma + \varepsilon| \geq 0.45\}.$$

The basic reason for considering such an example is that, for suitable choices of ε , the actual number r of permutations in Theorem 6 which are necessary to describe $S(\Omega_{A(\varepsilon)})$ can be made to vary from one to three. More precisely, for $0 \leq \varepsilon < 0.045$, $r = 3$; for $0.045 \leq \varepsilon < 0.45$, $r = 2$; and for $0.45 \leq \varepsilon$, $r = 1$. The first two cases are illustrated in Figures 1 and 2.

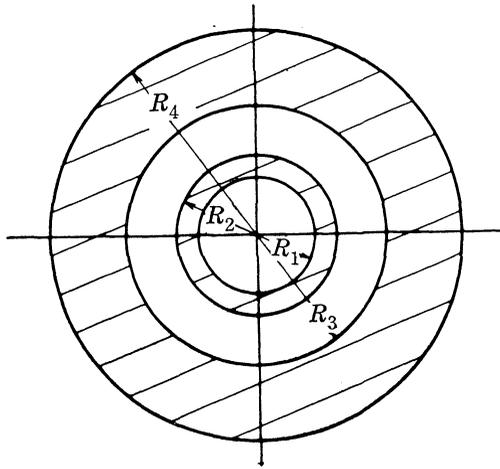


FIG. 1

$\varepsilon = 0$; $R_1 = 0.45$, $R_2 = 0.62$, $R_3 = 1.00$, $R_4 = 1.62$

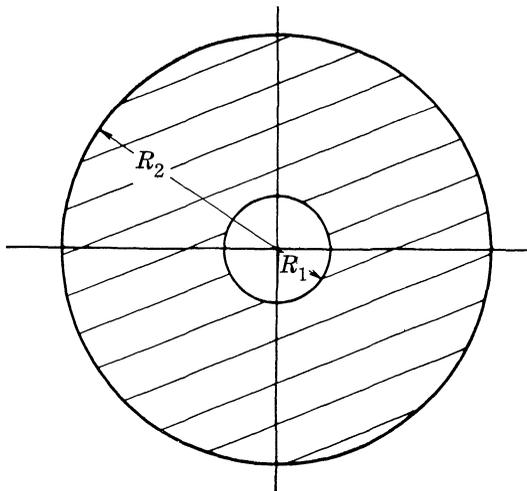


FIG. 2

$\varepsilon = 0.05$; $R_1 = 0.40$, $R_2 = 1.67$

This last example serves to answer some questions which might naturally arise in reading the previous sections. First, it shows that $n \times n$ matrices A exist for which at least n permutations φ are necessary to determine $S(\mathring{\Omega}_A)$. On the other hand, it shows that it is *not* necessary for A to be essentially diagonally dominant in order that $S(\mathring{\Omega}_A)$ coincide with $\text{rot } G(\Omega_A)$ (cf. Theorem 1), since choosing $\varepsilon = 0.5$ in (4.4) gives this condition. Finally, it demonstrates that, in general, it is not possible to find a *single* matrix $B \in \mathring{\Omega}_A$ for which $S(\mathring{\Omega}_A)$ is $\text{rot } S(\Omega_B)$. This fact follows quite easily from the last example with $\varepsilon = 0.05$, in particular.

BIBLIOGRAPHY

1. Paul Camion and A. J. Hoffman, *On the nonsingularity of complex matrices*, Pacific J. Math. **17** (1966), 211-214.
2. H. G. Eggleston, *Convexity*, Cambridge at the University Press, 1958.
3. B. W. Levinger and R. S. Varga, *Minimal Gerschgorin sets II*, Pacific J. Math. **17** (1966), 199-210.
4. A. M. Ostrowski, *Über die Determinanten mit Überwiegender Hauptdiagonale*, Comment. Math. Helv. **10** (1937), 69-96.
5. Olga Taussky, *On the variation of the characteristic roots of a finite matrix under various changes of its elements*, *Recent advances in matrix theory*, edited by H. Schneider, University of Wisconsin Press, 1964, 125-138.
6. Richard S. Varga, *Minimal Gerschgorin sets*, Pacific J. Math. **15** (1965), 119-729.
7. ———, *Matrix iterative analysis*, Prentice-Hall, Inc. 1962.

Received October 29, 1964.

CASE INSTITUTE OF TECHNOLOGY
CLEVELAND, OHIO

