Electronic Transactions on Numerical Analysis.
Volume 8, 1999, pp. 15-20.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

ON GERŠGORIN-TYPE PROBLEMS AND OVALS OF CASSINI *

RICHARD S. VARGA † AND ALAN KRAUTSTENGL ‡

Abstract. Recently, two Geršgorin-type matrix questions were raised. These are answered here, using ovals of Cassini.

Key words. Geršgorin circle theorem, ovals of Cassini.

AMS subject classification. 15A18.

1. Introduction.

Given a matrix $A = [a_{i,j}] \in \mathbb{C}^{n\times n}$, its usual Geršgorin disks are given by

$$\left\{ z \in \mathbb{C} : |z - a_{i,i}| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{i,j}| =: r_i(A) \right\} =: \Gamma_i(A) \quad (1 \leq i \leq n), \tag{1.1}$$

and, if $\sigma(A)$ denotes the spectrum of A, i.e.,

$$\sigma(A) := \{\lambda \in \mathbb{C} : \lambda \text{ is an eigenvalue of A}\},$$

the famous Geršgorin circle theorem [3] (or Horn and Johnson [4, p. 344] and [7, p. 16]) is that

$$\sigma(A) \subseteq \bigcup_{i=1}^{n} \Gamma_i(A) =: \Gamma(A). \tag{1.2}$$

At this moment, only the n diagonal entries $\{a_{i,i}\}_{i=1}^n$ and the n row sums $\{r_i(A)\}_{i=1}^n$ of (1.1) determine the n Geršgorin disks $\{\Gamma_i(A)\}_{i=1}^n$ and the inclusion set $\Gamma(A)$, of (1.2), in the complex plane. Then, with the notation

$$N := \{1, 2, \cdots, n\},$$

these $2n$ pieces of information are used to define the following set of matrices:

$$\Omega := \left\{ B = [b_{i,j}] \in \mathbb{C}^{n\times n} : b_{i,i} = a_{i,i} \text{ and } r_i(B) = r_i(A) \text{ for all } i \in N \right\}. \tag{1.3}$$

It is evident that $\sigma(B)$, for each $B \in \Omega$, must also satisfy the inclusion of (1.2), and, with the notation

$$\sigma(\Omega) := \bigcup_{B\in\Omega} \sigma(B), \tag{1.4}$$

* Received March 30, 1998. Accepted for publication August 13, 1998. Recommended by V. Mehrmann.

† Institute for Computational Mathematics, Department of Mathematics and Computer Science, Kent State University, Kent, Ohio 44242 (U.S.A.) (varga@msc.kent.edu). This research supported by National Science Foundation Grant DMS-9707359.

‡ Department of Mathematics, Case Western Reserve University. 10900 Euclid Ave., Cleveland, OH 44106 (U.S.A.) (axk66@po.cwru.edu).

ETNA
Kent State University
etna@mcs.kent.edu

we have

$$\sigma(\Omega) \subseteq \Gamma(A). \tag{1.5}$$

As we shall see, the inclusion in (1.5) is not always one of equality.

Recently, the following questions were asked of one of us:

Question 1. Can a precise description of $\sigma(\Omega)$ be given?

Question 2. When is it the case that $\sigma(\Omega)$ is a single closed disk (with a nonempty interior) in the complex plane?

The point of this note is to answer the above questions.

We remark that the set $\sigma(\Omega)$ of (1.4) is, for $n > 2$, in general *larger* than the *minimal Geršgorin set* (cf. [8]) for any matrix A, simply because the minimal Geršgorin set uses more specific information about the matrix $A = [a_{i,j}] \in \mathbb{C}^{n\times n}$ in determining its minimal Geršgorin set. Specifically, the associated set of matrices, for the minimal Geršgorin set of a given matrix $A = [a_{i,j}] \in \mathbb{C}^{n\times n}$, is

$$\Omega_{MG} := \left\{B = [b_{i,j}] \in \mathbb{C}^{n\times n} : b_{i,i} = a_{i,i} \text{ and } |b_{i,j}| = |a_{i,j}| \text{ for all } i,j \in N\right\}, \tag{1.6}$$

i.e., in addition to knowing the n diagonal entries $\{a_{i,i}\}_{i=1}^{n}$, one is also given the moduli of *all* $n(n-1)$ nondiagonal entries $\{|a_{i,j}|\}_{\substack{i,j=1\\ i\neq j}}^{n}$, for a total of n^2 pieces of information. Though it is not essential for this paper, we remark that it is evident from (1.3) and (1.6) that

$$\Omega_{MG} \subseteq \Omega, \text{ so that } \sigma(\Omega_{MG}) \subseteq \sigma(\Omega) \text{ for any } A \in \mathbb{C}^{n\times n}, \quad n > 2, \tag{1.7}$$

where it can be shown that the equalities of containment in (1.7) cannot in general hold for all $A \in \mathbb{C}^{n\times n}$, $n > 2$, but in the case $n = 2$, there holds

$$\Omega_{MG} = \Omega, \text{ so that } \sigma(\Omega_{MG}) = \sigma(\Omega) \text{ for any } A \in \mathbb{C}^{2\times 2}. \tag{1.8}$$

2. On Question 1. We need some additional notations and results. Given the 2×2 matrix

$$B = \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \in \mathbb{C}^{2\times 2}, \tag{2.1}$$

its associated *oval of Cassini* is defined as

$$K(B) := \{z \in \mathbb{C} : |(z - b_{1,1})(z - b_{2,2})| \leq |b_{1,2} \cdot b_{2,1}|\}. \tag{2.2}$$

Then, for a matrix $A = [a_{i,j}] \in \mathbb{C}^{n\times n}$, $n \geq 2$, we denote its ovals of Cassini by

$$K_{i,j}(A) := \{z \in \mathbb{C} : |(z - a_{i,i})(z - a_{j,j})| \leq r_i(A) \cdot r_j(A)\}, \text{ for all } i \neq j \text{ in } N. \tag{2.3}$$

There are $\binom{n}{2} = \frac{n(n-1)}{2}$ ovals of Cassini $K_{i,j}(A)$ associated with the matrix A, and from a well-known result of A. Brauer [1] (see also [4, p. 380]), we have that if

$$K(A) := \bigcup_{\substack{i,j=1 \\ i \neq j}}^{n} K_{i,j}(A), \tag{2.4}$$

then

$$\sigma(A) \subseteq K(A). \tag{2.5}$$

But, as the ovals of Cassini in (2.3) use only the information given by $\{a_{i,i}\}_{i=1}^n$ and $\{r_i(A)\}_{i=1}^n$, it is also true from (1.3) that

$$\sigma(\Omega) \subseteq K(A). \tag{2.6}$$

We remark from (2.3) that each oval of Cassini $K_{i.j}(A),\ i \neq j$, is a bounded and closed (hence, compact) set in the complex plane $\mathbb{C}$, as is their finite union $K(A)$ in (2.4). Next, if T is a set in $\mathbb{C}$, then $\overline{T}$ denote its closure and $T' := \mathbb{C} \backslash T$ denotes its complement. The boundary of T is defined, as usual, by $\partial T := \overline{T} \bigcap \overline{T'}$.

Our first result, which sharpens the inclusion in (2.6), shows exactly how $\sigma(\Omega)$ fills out $K(A)$.

THEOREM 2.1. *Given any* $A = [a_{i,j}] \in \mathbb{C}^{n \times n},\ n \geq 2$, *then*

$$\sigma(\Omega) = \partial K(A) = \partial K_{1,2}(A),\ \textit{when}\ n = 2, \tag{2.7}$$

and

$$\sigma(\Omega) = K(A),\ \textit{when}\ n \geq 3. \tag{2.8}$$

Proof. For $n = 2$, each matrix B in Ω is, from (1.3), necessarily of the form

$$B = \begin{bmatrix} a_{1,1} & r_1(A)e^{i\psi_1} \\ r_2(A)e^{i\psi_2} & a_{2,2} \end{bmatrix},\ \text{with } \psi_1, \psi_2 \text{ arbitrary real numbers.} \tag{2.9}$$

If λ is any eigenvalue of B, then det $(B - \lambda I) = 0$, so that from (2.9),

$$(a_{1,1} - \lambda)(a_{2,2} - \lambda) = r_1(A) \cdot r_2(A)e^{i(\psi_1 + \psi_2)}.$$

Hence,

$$|a_{i,i} - \lambda| \cdot |a_{2,2} - \lambda| = r_1(A) \cdot r_2(A). \tag{2.10}$$

As (2.10) corresponds to the case of equality in (2.3), we see that $\lambda \in \partial K_{1,2}(A)$. Since this is true for any eigenvalue λ of any B in Ω and since, from (2.4), $K_{1,2}(A) = K(A)$ in this case $n = 2$, then $\sigma(\Omega) \subseteq \partial K_{1,2}(A) = \partial K(A)$. Conversely, it is easily seen that each point of $\partial K_{1,2}(A)$ is, for suitable choices of real ψ_1 and ψ_2, an eigenvalue of some B in (2.9), so that $\sigma(\Omega) = \partial K_{1,2}(A) = \partial K(A)$, the desired result of (2.7).

To establish (2.8), first assume that $n \geq 4$, and consider a matrix $B = [b_{i,j}]$ in $\mathbb{C}^{n\times n}$, which has the partitioned form

$$B = \left[\begin{array}{c|c} B_{1,1} & B_{1,2} \\ \hline O & B_{2,2} \end{array}\right], \tag{2.11}$$

where

$$B_{1,1} := \begin{bmatrix} a_{1,1} & se^{i\psi_1} \\ te^{i\psi_2} & a_{2,2} \end{bmatrix}, \text{ with } 0 \leq s \leq r_1(A),\ 0 \leq t \leq r_2(A), \tag{2.12}$$

with ψ_1 and ψ_2 arbitrary real numbers, and with $b_{j,j} = a_{j,j}$ for all $1 \leq j \leq n$. Now, for any choices of s and t with $s \in [0, r_1(A)]$ and $t \in [0, r_2(A)]$, the entries of the block $B_{1,2}$ can be chosen so that the row sums, $r_1(B)$ and $r_2(B)$, in the first two rows of B, equal those of A. Similarly, because $n \geq 4$, the row sums of the matrix $B_{2,2}$ of (2.11) can be chosen to be the same as those in the remaining row sums of A. Thus, by our construction, the matrix B of (2.11) is an element of Ω. (We remark that this construction fails to work in the case $n = 3$, unless $r_3(A) = 0$). But from the partitioned form in (2.11), it is evident that

$$\sigma(B) = \sigma(B_{1,1}) \bigcup \sigma(B_{2,2}). \tag{2.13}$$

However, from the parameters s, t, ψ_1, ψ_2 in $B_{1,1}$ in (2.12), it can be seen from (2.3), that, for *each* $z \in K_{1,2}(A)$, there are choices for these parameters such that z is an eigenvalue of $B_{1,1}$. In other words, the eigenvalues of $B_{1,1}$, on varying s and t with $0 \leq s \leq r_1(A)$ and $0 \leq t \leq r_2(A)$, fill out $K_{1,2}(A)$, where we note that the remaining eigenvalues of B (namely, those of $B_{2,2}$) must still lie, from (2.6), in $K(A)$. As this applies to *all* ovals of Cassini $K_{i,j}(A)$, for $i \neq j$, upon suitable permutations of the rows and columns of B of (2.11), then $\sigma(\Omega) = K(A)$, for all $n \geq 4$.

For the remaining case $n = 3$ of (2.8), any matrix B in Ω can be expressed as

$$B = \begin{bmatrix} a_{1,1} & se^{i\psi_1} & (r_1(A)-s)e^{i\psi_2} \\ te^{i\psi_3} & a_{2,2} & (r_2(A)-t)e^{i\psi_4} \\ ue^{i\psi_5} & (r_3(A)-u)e^{i\psi_6} & a_{3,3} \end{bmatrix}, \tag{2.14}$$

where

$$\begin{array}{l} 0 \leq s \leq r_1(A),\ 0 \leq t \leq r_2(A), \text{ and } 0 \leq u \leq r_3(A), \text{ and} \\ \{\psi_i\}_{i=1}^{6} \text{ are arbitrary real numbers.} \end{array} \tag{2.15}$$

Now, choose any complex number z in the oval of Cassini $K_{1,2}(A)$, i.e., (cf. (2.3)), let z satisfy

$$|z - a_{1,1}| \cdot |z - a_{2,2}| \leq r_1(A) \cdot r_2(A). \tag{2.16}$$

The above inequality implies that either $|z - a_{1,1}| \leq r_1(A)$ or $|z - a_{2,2}| \leq r_2(A)$ is valid. Assuming, without loss of generality, that $|z - a_{1,1}| \leq r_1(A)$, choose the parameters s and ψ_1 such that $|z - a_{1,1}| = s$ and $z - a_{1,1} = se^{i\psi_1}$, and similarly, choose the parameters t and

ψ_3 such that $|z - a_{2,2}| = t$ and $z - a_{2,2} = te^{i\psi_3}$, where, from (2.16), $s \cdot t \leq r_1(A) \cdot r_2(A)$. Then with the vector $\mathbf{x} = [1, 1, 0]^T$ in $\mathbb{R}^3$, it can be verified that $B\mathbf{x} = \boldsymbol{\eta}$, where

$$\begin{cases} \eta_1 = a_{1,1} + se^{i\psi_1} = z, \\ \eta_2 = te^{i\psi_3} + a_{2,2} = z, \\ \eta_3 = ue^{i\psi_5} + (r_3(A) - u)e^{i\psi_6}. \end{cases} \tag{2.17}$$

Then, on setting $u := r_3(A)/2$ and $\psi_5 := \psi_6 + \pi$, we have $\eta_3 = 0$. Thus, $B\mathbf{x} = z\mathbf{x}$, and z is an eigenvalue of B. Hence, each z in $K_{1,2}(A)$ is an eigenvalue of some B in Ω. Consequently, as this construction for $n = 3$ can be applied to *any* point of *any* oval of Cassini $K_{i,j}(A)$ with $i \neq j$, then $\sigma(\Omega) = K(A)$, which completes the proof of (2.8). □

We remark that there is a way to enlarge the set Ω, in the spirit of what has been done in the treatment of minimal Geršgorin sets (cf. [8]), which gives a unified inclusion result for *any* $n \geq 2$, which does not distinguish between the cases $n = 2$ and $n \geq 3$, as in Theorem 2.1. To this end, set

$$\hat{\Omega} := \left\{ B = [b_{i,j}] \in \mathbb{C}^{n \times n} : b_{i,i} = a_{i,i} \text{ and } 0 \leq r_i(B) \leq r_i(A) \text{ for all } i \in N \right\}, \tag{2.18}$$

so that $\Omega \subseteq \hat{\Omega}$, and this of course implies

$$\sigma(\Omega) \subseteq \sigma(\hat{\Omega}). \tag{2.19}$$

But from the definitions in (2.3), we also have

$$\sigma(\hat{\Omega}) \subseteq K(A). \tag{2.20}$$

THEOREM 2.2. *Given any* $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$ *with* $n \geq 2$*, then*

$$\sigma(\hat{\Omega}) = K(A). \tag{2.21}$$

Proof. From Theorem 2.1 and the inclusions of (2.19) and (2.20), it is only necessary to establish (2.21) in the case $n = 2$, i.e., for $n = 2$, it suffices to show that $\sigma(\hat{\Omega}) = K(A) = K_{1,2}(A)$. The proof of this is similar to the proof in the first part of Theorem 2.1. For $n = 2$, each matrix B in $\hat{\Omega}$ is, from (2.18), of the form

$$B = \begin{bmatrix} a_{1,1} & se^{i\psi_1} \\ te^{i\psi_2} & a_{2,2} \end{bmatrix}, \quad \text{with } 0 \leq s \leq r_1(A),\ 0 \leq t \leq r_2(A), \text{ and } \psi_1, \psi_2 \text{ arbitrary real numbers.} \tag{2.22}$$

If λ is any eigenvalue of B, then det $(B - \lambda I) = 0$ implies that

$$(a_{1,1} - \lambda)(a_{2,2} - \lambda) = ste^{i(\psi_1 + \psi_2)},$$

so that

$$|(a_{1,1} - \lambda)(a_{2,2} - \lambda)| = st, \text{ for all real } \psi_1, \psi_2.$$

But as s and t run, respectively, through the intervals $[0, r_1(A)]$ and $[0, r_2(A)]$, and as ψ_1 and ψ_2 run through all real numbers, it is evident that the eigenvalues of the matrices B in $\hat{\Omega}$ fill out (cf. (2.3)) the set $K_{1,2}(A) = K(A)$, i.e., in this case $n = 2,\ \sigma(\hat{\Omega}) = K_{1,2}(A) = K(A)$. □

3. On Question 2. The answer to Question 2 is straight-forward, based on the results of Section 2 and the following observations. Given a matrix $A = [a_{i,j}] \in \mathbb{C}^{n\times n},\ n \geq 2$, then, with the definitions of $\Gamma_i(A)$ of (1.1) and $K_{i,j}(A)$ of (2.3), it is easy to verify that

$$\text{(3.1)} \qquad \begin{cases} K_{i,j}(A) & \subseteq \Gamma_i(A) \bigcup \Gamma_j(A) \text{ for all } i,j \in N \text{ with } i \neq j, \\ & \text{with equality holding only if } r_i(A) = r_j(A) = 0, \\ & \text{ or if } a_{i,i} = a_{j,j} \text{ and } r_i(A) = r_j(A) > 0. \end{cases}$$

The set inclusions in (3.1) imply, with the definitions of $\Gamma(A)$ of (1.2) and $K(A)$ of (2.4), that one always has

$$\text{(3.2)} \qquad K(A) \subseteq \Gamma(A).$$

So, to answer Question 2, it can be seen from Theorems 2.1 and 2.2 that, if $A = [a_{i,j}] \in \mathbb{C}^{n\times n},\ n \geq 2$, then

a) for $n = 2,\ \sigma(\Omega)$ is never a single closed disk with a nonempty interior.
b) for $n = 2,\ \sigma(\hat{\Omega})$ is a single closed disk with a nonempty interior only if $a_{1,1} = a_{2,2}$ and $r_1(A) \cdot r_2(A) > 0$.
c) for $n > 2,\ \sigma(\Omega) = \sigma(\hat{\Omega})$ is a simple closed disk with nonempty interior only if there are s, t in N with $s \neq t$ such that $a_{s,s} = a_{t,t}$ and $r_s(A) \cdot r_t(A) > 0$, and all remaining ovals of Cassini $K_{i,j}(A)$ lie in $K_{s,t}(A)$.

4. Final Remarks. It was a surprise for us to see that the determination of $\sigma(\Omega)$, in Theorem 2.1, was completely in terms of ovals of Cassini, a topic rarely seen in the current research literature in linear algebra. In fact, on consulting many (16) books on linear algebra, we could find only three books where ovals of Cassini are even mentioned: in an exercise in Varga ([7, p. 22]), in a lengthier discussion in Horn and Johnson [4, pp. 380-381], and in Marcus and Minc [5, p. 149]. But, in none of these books was it stated that the ovals of Cassini are in general *at least as good* (cf. (3.2)) as the Geršgorin disks, in estimating the spectrum $\sigma(A)$ of a given matrix in $\mathbb{C}^{n\times n}$. This appears in Brauer's paper [1] of 1947 (which curiously does not mention Geršgorin's earlier paper [3] of 1931.) What apparently was of more interest in the literature is the fact that Brauer's use of *two* rows at a time, to determine an inclusion set for the eigenvalues of a given matrix, cannot in general be extended in the same manner to k rows at a time, for $k \geq 3$, and explicit counterexamples to this are given in [4, p. 382] and [5, p. 149]. However, a "correct" generalization of Brauer's results, using graph theory, is nicely given in Brualdi [2].

REFERENCES

[1] A. BRAUER, *Limits for the characteristic roots of a matrix II*, Duke Math. J., 14 (1947), pp. 21–26.
[2] R. BRUALDI, *Matrices, eigenvalues, and directed graphs*, Lin. Multilin. Alg. 11 (1982), pp. 143–165.
[3] S. GERSCHGORIN, *Über die Abgrenzung der Eigenwerte einer Matrix*, Izv. Akad. Nauk SSSR Ser. Mat., 7 (1931), pp. 749–754.
[4] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
[5] M. MARCUS AND M. MINC, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, 1964.
[6] O. TAUSSKY, *Bounds for characteristic roots of matrices*, Duke Math. J. 15 (1948), pp. 1043-1044.
[7] R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.
[8] R. S. VARGA, *Minimal Gerschgorin Sets*, Pacific J. Math, 15 (1965), pp. 719–729.
[9] X. ZHANG AND D. GU, *A note on Brauer's Theorem*, Lin. Algebra. Appl. 196 (1994), pp. 163-174.