

Gerschgorin Disks, Brauer Ovals of Cassini (a vindication), and Brualdi Sets

Richard S. Varga
Institute for Computational Mathematics
Kent State University
Kent, Ohio 44240 (USA)
e-mail: (varga@mcs.kent.edu)

Abstract

Here, we show in particular that Alfred Brauer's ovals of Cassini, for determining inclusion regions in the complex plane for the eigenvalues of an arbitrary matrix in $\mathbb{C}^{n \times n}$, $n \geq 2$, are in a certain sense optimal. We also include comparisons with Brualdi sets.

Keywords: Gerschgorin disks, ovals of Cassini, Brualdi sets

1 Introduction

In 1947, Alfred Brauer [1] introduced ovals of Cassini as a means of determining inclusion regions, in the complex plane, for the eigenvalues of an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$, for $n \geq 2$. Let

$$R'_i(A) := \sum_{\substack{j=1 \\ j \neq i}}^n |a_{i,j}| \quad (1 \leq i \leq n) \tag{1.1}$$

denote the deleted row sums of the matrix A , and let

$$K^r_{i,j}(A) := \{z \in \mathbb{C} : |z - a_{i,i}| \cdot |z - a_{j,j}| \leq R'_i \cdot R'_j\}, \quad i \neq j (1 \leq i, j \leq n); \tag{1.2}$$

denote the (i, j) -th row oval of Cassini for the matrix A . Further, set

$$K^r(A) := \bigcup_{\substack{i,j=1 \\ i \neq j}}^n K^r_{i,j}(A), \tag{1.3}$$

and let $\sigma(A)$ denote the eigenvalues of A , i.e.,

$$\sigma(A) = \{\lambda \in \mathbb{C} : \det(\lambda I - A) = 0\}. \tag{1.4}$$

Then in [1], Alfred Brauer established his well-known result of

Theorem 1.1 Let $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$, $n \geq 2$. Then,

$$\sigma(A) \subseteq K^r(A). \quad (1.5)$$

Related to the inclusion result of Alfred Brauer's Theorem 1.1 is the classical theorem of Gerschgorin [4] from 1931. If we define the associated Gerschgorin row disks by

$$G_i^r(A) := \{z \in \mathbb{C} : |z - a_{i,i}| \leq R_i^r\}, \quad (1.6)$$

for all $1 \leq i \leq n$, and if we set

$$G^r(A) := \bigcup_{i=1}^n G_i^r(A), \quad (1.7)$$

then Gerschgorin's Theorem is

Theorem 1.2 Let $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$. Then,

$$\sigma(A) \subseteq G^r(A). \quad (1.8)$$

We remark, as is well known, that Theorems 1.1 and 1.2 are equally valid for the deleted column sums $\mathbb{C}'_i(A) := \sum_{\substack{j=1 \\ j \neq i}}^n |a_{j,i}|$.

For $n \geq 2$, it is the case that it is generally easier to apply Gerschgorin's Theorem 1.2, since this involves n disks, whereas A. Brauer's Theorem 1.1 however involves $\binom{n}{2} = n(n-1)/2$ sets in the complex plane which, from their definition in (1.2), are in general more complicated than are the disks of Gerschgorin, and their number, $\binom{n}{2}$, exceeds n for all $n > 3$. Perhaps for the above reasons, this author has often heard the comment, "I have never used the ovals of Cassini for estimating the eigenvalues of a matrix, since Gerschgorin works just fine for me!" Still, though it seems not well known, it is the case that the Alfred Brauer's ovals of Cassini $K^r(A)$ are *always* at least as good as $G^r(A)$ in estimating $\sigma(A)$, in that

$$K^r(A) \subseteq G^r(A) \text{ for any } A \in \mathbb{C}^{n \times n}. \quad (1.9)$$

These inclusions were mentioned by Alfred Brauer in [2]. The proof of the inclusion in (1.9) simply requires showing, for example, that

$$K_{i,j}^r(A) \subseteq G_i^r(A) \cup G_j^r(A), \text{ for all } i \neq j \ (1 \leq i, j \leq n). \quad (1.10)$$

The case of equality in (1.10) is covered in

$$K_{i,j}^r(A) = G_i^r(A) \cup G_j^r(A) \text{ for } i \neq j \text{ only if } R_i^r(A) = R_j^r(A) = 0, \quad (1.11)$$

$$\text{or if } R_i^r(A) = R_j^r(A) > 0 \text{ and } a_{i,i} = a_{j,j}.$$

Perhaps another deterrent, to the popularity of Alfred Brauer's ovals of Cassini, could stem from the fact that the proof of the inclusion in (1.5) depends on carefully examining

two distinct rows of the matrix A , while Gerschgorin's inclusion in (1.6) examines just *one* particular row of A . Obviously, this was an invitation for mathematicians to consider *three* or more distinct rows of A , in their quest for obtaining other *new* inclusion regions for the spectrum of A . But alas, such a generalization *does not work in general!* That is, for any $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$ with $n \geq m$, assume that i_1, i_2, \dots, i_m are distinct positive integers in $[1, n]$. Then define

$$H_{i_1, \dots, i_m}^r(A) := \left\{ z \in \mathbb{C} : \prod_{k=1}^m |z - a_{i_k, i_k}| \leq \prod_{k=1}^m R'_{i_k}(A) \right\}, \quad (1.12)$$

and set

$$H_{(m)}^r(A) := \bigcup_{1 \leq i_1, i_2, \dots, i_m \leq n} H_{i_1, \dots, i_m}^r(A) \quad (i_1, i_2, \dots, i_m \text{ distinct}). \quad (1.13)$$

The cases $m = 1$ and $m = 2$ of (1.12) thus correspond, respectively, to the Gerschgorin disks of (1.6) and the Cassini ovals of (1.2). We note that $H_{(m)}^r(A)$ now consists of $\binom{n}{m}$ sets $H_{i_1, \dots, i_m}^r(A)$.

But, for $m \geq 3$, the set $H_{(m)}^r(A)$ does *not* always cover the spectrum of A , as the following counter example in Horn and Johnson [5, p. 382] shows. (This counter example has been attributed to Morris Newman in Marcus and Minc [6].) Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{where} \quad \begin{array}{l} R'_1(A) = 1 = R'_2(A), \\ R'_3(A) = R'_4(A) = 0. \end{array} \quad (1.14)$$

Then, $\sigma(A) = \{0, 1, 1, 2\}$. But in the case $m = 3$, any choice of three distinct integers from $(1, 2, 3, 4)$ gives at least one deleted row sum with $R'_k(A) = 0$, so that the union of the sets in (1.13) reduces to a single point $z = 1$, which clearly does not include all eigenvalues of A . The case $m = 4$ similarly fails for the matrix A of (1.14), and this counter example can be extended to all $m \geq 3$.

It is important to remark now that newer results of Brualdi [3] give sufficient conditions, based on graph theory, for these Brauer-like extensions to produce valid eigenvalue inclusion sets in the complex plane. We will treat this more in detail in Section 3.

2 Ovals of Cassini—a Vindication

First, observe that the inclusion regions, for the eigenvalues of $A = [a_{i,j}]$ in $\mathbb{C}^{n \times n}$, $n \geq 2$, in Theorems 1.1 and 1.2, depend *solely* on the $2n$ numbers:

$$\{a_{i,i}\}_{i=1}^n \quad \text{and} \quad \{R'_i(A)\}_{i=1}^n, \quad (2.1)$$

as do the sets $H_{(m)}^r(A)$ in (1.13). If

$$\begin{aligned} \Omega(A) := \{ & B = [b_{i,j}] \in \mathbb{C}^{n \times n}, \quad n \geq 2 : b_{i,i} = a_{i,i} \quad \text{and} \\ & R'_i(B) = R'_i(A), \quad 1 \leq i \leq n \}, \end{aligned} \quad (2.2)$$

it is evident that the eigenvalue inclusions of (1.5) and (1.8) apply to *any* B in $\Omega(A)$. Thus, with

$$\sigma(\Omega(A)) := \bigcup_{B \in \Omega(A)} \sigma(B), \quad (2.3)$$

we have from Theorems 1.1 and 1.2 and (1.9) that

$$\sigma(\Omega(A)) \subseteq K^r(A) \subseteq G^r(A), \quad (2.4)$$

for any $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$. What (2.4) asserts is that $K^r(A)$ is always a "tighter" estimate of $\sigma(\Omega(A))$, than is $G^r(A)$, for *any* $A \in \mathbb{C}^{n \times n}$, i.e., the union of the ovals of Cassini of Brauer are always as good as the union of the Gerschgorin disks for *any* $A \in \mathbb{C}^{n \times n}$.

This brings us to the following reformulation of the above. Let

$$\Delta_n := \{\alpha_i\}_{i=1}^n \cup \{r_i\}_{i=1}^n, \quad n \geq 2, \quad \text{with} \quad (2.5)$$

$$\alpha_i \in \mathbb{C} \text{ and } r_i \geq 0 \quad (1 \leq i \leq n),$$

be n complex numbers and n nonnegative real numbers. Then, associated with Δ_n is the following set of $n \times n$ matrices:

$$\Omega(\Delta_n) := \left\{ B = [b_{i,j}] \in \mathbb{C}^{n \times n} : b_{i,i} = \alpha_i \text{ and } \sum_{\substack{j=1 \\ j \neq i}}^n |b_{i,j}| = r_i \quad (1 \leq i \leq n) \right\}, \quad (2.6)$$

and denote the set of all eigenvalues of all $B \in \Omega(\Delta_n)$ by

$$\sigma(\Omega(\Delta_n)) := \bigcup_{B \in \Omega(\Delta_n)} \sigma(B). \quad (2.7)$$

Next, for $n \geq 2$, let

$$\mathcal{P}_n := \left\{ \text{set of all partitions } (i_1, i_2, \dots, i_p), \text{ with distinct elements of the integers in } \{j\}_{j=1}^n \right\}, \quad (2.8)$$

where the cardinality of \mathcal{P}_n is known to be $2^n - 1$. Now, fix any nonempty subset μ of \mathcal{P}_n , and consider the set

$$F(\Delta_n; \mu) := \bigcup_{(i_1, i_2, \dots, i_s) \in \mu} \left\{ z \in \mathbb{C} : \prod_{j=1}^s |z - \alpha_{i_j}| \leq \prod_{j=1}^s r_{i_j} \right\}, \quad (2.9)$$

which is dependent on Δ_n and on $\mu \subseteq \mathcal{P}_n$. This set is a compact set in the complex plane \mathbb{C} . Note that, for a fixed positive integer m (where $1 \leq m \leq n$), the subset μ_m of \mathcal{P}_n , defined by

$$\mu_m := \text{union of all sets } (i_1, i_2, \dots, i_m) \text{ of distinct integers in } \{j\}_{j=1}^n, \quad (2.10)$$

is such that $F(\Delta_n; \mu_m)$ is exactly the set $H_{(m)}^r$ of (1.13) for the matrices of $\Omega(\Delta_n)$, so that $m = 1$ and $m = 2$ give, respectively, the common Gerschgorin disks and the common ovals of Cassini for the matrices of $\Omega(\Delta_n)$.

We are, of course, interested here in general eigenvalue inclusion sets in the complex plane \mathbb{C} , for arbitrary matrices in $\mathbb{C}^{n \times n}$, $n \geq 2$, and we ask if, for some nonempty $\mu \subseteq \mathcal{P}_n$, there holds

$$\sigma(\Omega(\Delta_n)) \subseteq F(\Delta_n, \mu), \text{ for any choice of } \Delta_n, \tag{2.11}$$

where Δ_n is defined in (2.5). As we know, (2.11) is valid, for any Δ_n , for the subsets μ_1 and μ_2 of (2.10), but not for μ_m of (2.10) for $m \geq 3$. We say that a subset μ of \mathcal{P}_n is an *eigenvalue inclusion subset* of (2.11) is valid for *any* choice of Δ_n . Note that if μ is such an eigenvalue inclusion subset of \mathcal{P}_n , then so is

$$\mu \cup q, \text{ for any } q \in \mathcal{P}_n, \tag{2.12}$$

since adding q to μ gives, from (2.9), that $F(\Delta_n; \mu \cup q) \supseteq F(\Delta_n; \mu)$, thereby still covering $\sigma(\Omega(\Delta_n))$. It is then of interest to find such eigenvalue inclusion subsets μ of \mathcal{P}_n which give the *tightest* inclusion in (2.11). One such subset is given in

Theorem 2.1 *With $n \geq 2$ and with the definition of μ_2 in (2.10),*

$$\sigma(\Omega(\Delta_2)) = \partial F(\Delta_2; \mu_2) \text{ for all } \Delta_2, \text{ for } n = 2, \tag{2.13}$$

and

$$\sigma(\Omega(\Delta_n)) = F(\Delta_n; \mu_2) \text{ for all } \Delta_n, \text{ for all } n > 2. \tag{2.14}$$

Thus, there can be no eigenvalue inclusion subset μ of \mathcal{P}_n which can give tighter eigenvalue inclusions, in (2.13) and (2.14), than the ovals of Cassini.

Proof. The inequalities of (2.13) and (2.14) are a direct consequence of Theorem 1.2 of Varga and Krautstengl [7]. ■

From Theorem 2.1, we thus have our *vindication* of Alfred Brauer's ovals of Cassini!

3 Further Remarks

It is of interest to connect the sets of (2.9) with results of Brualdi [3]. To do this, we need the following definitions. Given $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$, $n \geq 2$, then $\Gamma(A)$ is the *directed graph*, on n vertices $\{v_i\}_{i=1}^n$, for the matrix $A = [a_{i,j}]$, consisting of an *arc* $\overrightarrow{v_i v_j}$ from vertex v_i to vertex v_j only if $i \neq j$ and $a_{i,j} \neq 0$. (This directed graph omits the usual use of loops when $a_{i,i} \neq 0$.) A *path* π from vertex v_i to vertex v_j is a sequence $i = i_0, i_1, \dots, i_k = j$ of *distinct* vertices for which $\overrightarrow{v_{i_0} v_{i_1}}, \overrightarrow{v_{i_1} v_{i_2}}, \dots, \overrightarrow{v_{i_{k-1}} v_{i_k}}$ are abutting arcs, and the *length* of π is said to be k . A directed graph is *strongly connected* if, for each ordered pair of distinct vertices v_i and v_j , there is a path from v_i to v_j . A *circuit* γ of $\Gamma(A)$ is a sequence $i_1, \dots, i_p, i_{p+1} = i_1$, where $p \geq 2, i_1, \dots, i_p$ are distinct, and $\overrightarrow{v_{i_1} v_{i_2}}, \dots, \overrightarrow{v_{i_p} v_{i_1}}$ are arcs of $\Gamma(A)$. If we write $\gamma = (i_1, i_2, \dots, i_p)$ to identify this circuit γ , then, noting that these $\{i_j\}_{j=1}^p$ are distinct, each circuit $\gamma = (i_1, i_2, \dots, i_p)$ defines a partition in \mathcal{P}_n . Thus, if $\mathcal{C}(A)$ denotes the set of all circuits in $\Gamma(A)$, we thus have

$$\mathcal{C}(A) \subseteq \mathcal{P}_n \tag{3.1}$$

(where $\mathcal{C}(A)$ could be the null set). Continuing, a matrix $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$, $n \geq 2$, is said to be *weakly irreducible* if each vertex v_i of $\Gamma(A)$ belongs to *some* circuit in $\mathcal{C}(A)$. (Note that the matrix A of (1.14) is *not* weakly irreducible.)

With his definition of weak irreducibility, Brualdi [3, Corollary 2.4] showed how the counterexamples of (1.14) can be avoided. His result is

Theorem 3.1 *Assume that $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$, $n \geq 2$, is weakly irreducible. If $\mathcal{C}(A)$ denotes the circuits γ of the directed graph $\Gamma(A)$ for A , then*

$$\sigma(A) \subseteq \bigcup_{\gamma \in \mathcal{C}(A)} \left\{ z \in \mathbb{C} : \prod_{i \in \gamma} |z - a_{i,i}| \leq \prod_{i \in \gamma} R'_i(A) \right\}. \quad (3.2)$$

It seems fitting and proper to call the set, on the right of (3.2), the *Brualdi set* for A when A is weakly irreducible, i.e.,

$$\mathcal{B}^r(A) := \bigcup_{\gamma \in \mathcal{C}(A)} \left\{ z \in \mathbb{C} : \prod_{i \in \gamma} |z - a_{i,i}| \leq \prod_{i \in \gamma} R'_i(A) \right\}. \quad (3.3)$$

It is evident from (3.3) that the Brualdi set $\mathcal{B}^r(A)$, a closed set in the complex plane, is unchanged if A is replaced by any matrix in the set

$$\check{\Omega}(A) := \{B \in \Omega(A) : \text{the circuits } \gamma(B) \text{ from } \Gamma(B) \text{ are identical with those of } \Gamma(A)\}, \quad (3.4)$$

where $\Omega(A)$ is defined in (2.2). Thus, if A is weakly irreducible, it follows from (3.2)-(3.4) that

$$\sigma(\check{\Omega}(A)) \subseteq \mathcal{B}^r(A). \quad (3.5)$$

On the other hand, as $\check{\Omega}(A)$ is a subset of $\Omega(A)$, then for $n > 2$,

$$\sigma(\check{\Omega}(A)) \subseteq \sigma(\Omega(A)) = K^r(A), \quad (3.6)$$

the last equality coming from Theorem 2.1. Because of the first inclusion of (3.6), one would expect that

$$\mathcal{B}^r(A) \subseteq K^r(A). \quad (3.7)$$

To investigate this, we extend the definition of the Brualdi set $\mathcal{B}^r(A)$ of (3.3) to any matrix $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$, $n \geq 2$, where if there are no circuits γ in the directed graph $\Gamma(A)$, we use the convention that

$$\mathcal{B}^r(A) := \bigcup_{i=1}^n \{a_{i,i}\} \quad (3.8)$$

Theorem 3.2 *For any $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$, $n \geq 2$, let the sets $K^r(A)$ and $\mathcal{B}^r(A)$ in the complex plane, be defined by (1.3), (3.3) and (3.8). Then, the inclusion of (3.7) is valid.*

Proof. If $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$ has no cycles γ in $\Gamma(A)$, then $\mathcal{B}^r(A)$ is defined by (3.8). In this case, as each Cassini oval $K_{i,j}^r(A)$ of (1.2), for $i \neq j$, contains the points $z = a_{i,i}$ and $z = a_{j,j}$, then $\mathcal{B}^r(A) \subseteq K^r(A)$. Then, consider, any cycle $\gamma = (i_1, i_2, \dots, i_p)$ of $\mathcal{C}(A)$, whose associated term in $\mathcal{B}^r(A)$, from (3.3), defines the associated closed set in \mathbb{C} :

$$\mathcal{B}_\gamma^r(A) := \left\{ z \in \mathbb{C} : \prod_{j=1}^p |z - a_{i_j, i_j}| \leq \prod_{j=1}^p R'_{i_j}(A) \right\}. \quad (3.9)$$

Then, consider the pairs $(i_1, i_2), (i_2, i_3), \dots, (i_p, i_1)$. As γ is a cycle of $\mathcal{C}(A)$, each of these pairs defines a nontrivial Cassini oval $K_{i_j, i_{j+1}}^r(A)$ with $R'_{i_j}(A) > 0$, for all $1 \leq j \leq p$. Assume that z is any point in the set $\mathcal{B}_\gamma^r(A)$ of (3.9), so that

$$\prod_{j=1}^p \left(\frac{|z - a_{i_j, i_j}|}{R'_{i_j}(A)} \right) \leq 1,$$

and squaring the above expression gives

$$\prod_{j=1}^p \left(\frac{|z - a_{i_j, i_j}|}{R'_{i_j}(A)} \right)^2 \leq 1.$$

The above product can also be expressed as

$$\begin{aligned} & \left(\frac{|z - a_{i_1, i_1}| \cdot |z - a_{i_2, i_2}|}{R'_{i_1}(A) \cdot R'_{i_2}(A)} \right) \cdot \left(\frac{|z - a_{i_2, i_2}| \cdot |z - a_{i_3, i_3}|}{R'_{i_2}(A) \cdot R'_{i_3}(A)} \right) \cdots \\ & \cdot \left(\frac{|z - a_{i_p, i_p}| \cdot |z - a_{i_1, i_1}|}{R'_{i_p}(A) \cdot R'_{i_1}(A)} \right) \leq 1. \end{aligned}$$

Hence, as not all of the above factors can exceed unity, there exists an ℓ , with $1 \leq \ell \leq p$, such that

$$\left(\frac{|z - a_{i_\ell, i_\ell}| \cdot |z - a_{i_{\ell+1}, i_{\ell+1}}|}{R'_{i_\ell}(A) \cdot R'_{i_{\ell+1}}(A)} \right) \leq 1, \quad (3.10)$$

where, if $\ell = p$, then $i_{\ell+1} := i_1$. But (3.10) implies that $z \in K_{i_\ell, i_{\ell+1}}^r(A) \subseteq K^r(A)$. As this holds for any $z \in \gamma$ and for any $\gamma \in \mathcal{C}(A)$, we have that $\mathcal{B}^r(A) \subseteq K^r(A)$, the desired result of (3.7). \blacksquare

Note that the inclusion of (3.7) does *not* require that A be weakly irreducible, so that $\mathcal{B}^r(A)$ need not, as in example of (1.14), cover the spectrum of A . The result of Theorem 3.2, however, gives us that if A is indeed weakly irreducible, then the Brualdi set $\mathcal{B}^r(A)$ covers the spectrum of A , and is a subset of the Cassini ovals $K^r(A)$.

We remark that the actual computation of the Brualdi set $\mathcal{B}^r(A)$ can be quite lengthy. Take any $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$, $n \geq 2$, in which each off-diagonal entry of A is nonzero. Then, there are exactly $2^n - n - 1$ distinct circuits γ of $\Gamma(A)$. For $n = 10$, the number of distinct circuits is 1,013, as opposed to 45 ovals of Cassini!

References

- [1] A. Brauer, "Limits for the characteristic roots of a matrix. II," *Duke Math. J.* **14**, (1947), 21–26.
- [2] A. Brauer, "Limits for the characteristic roots of a matrix. VII," *Duke Math. J.* **25**, (1958), 583–590.
- [3] R. Brualdi, "Matrices, eigenvalues, and directed graphs," *Linear Multilinear Algebra* **11**, (1982), 143–165.
- [4] S. Gerschgorin, "Über die Abrenzung der Eigenwerte einer Matrix," *Izv. Akad. Nauk SSSR Ser. Math.* **7**, (1931), 749–754.
- [5] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [6] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, 1964.
- [7] R. S. Varga and Alan Krautstengl, "On Geršgorin-type problems and ovals of Cassini," *Elec. Trans. Numer. Anal.* **8**, (1999), 15–20.

Short Curriculum Vitae for R. S. Varga

Professor Varga obtained his Ph.D. in mathematics at Harvard University in 1954, and has been in academic life since 1960, first at Case Institute of Technology (1960-1969), and Kent State University (1969-present), where he holds the rank of University Professor. His areas of research have been numerical analysis, approximation theory, linear algebra and scientific computing. He has published six books and edited five, has published over 230 papers, and has had 25 Ph.D. students. His honors include a Guggenheim Fellowship, a von Humboldt Prize (Senior U.S. Award), and honorary doctorates from the Universities of Karlsruhe (Germany) and Lille (France). He is currently Editor-in-Chief of the journals *Numerische Mathematik*, and *ETNA* (Electronic Transactions on Numerical Analysis).