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)
c-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

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

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

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

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

$$K'(A) := \bigcup_{i,j=1 \atop i \neq j}^n K'_{i,j}(A),$$  \hspace{1cm} (1.3)

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

$$\sigma(A) = \{ \lambda \in \mathbb{C} : \det(\lambda I - A) = 0 \} .$$  \hspace{1cm} (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 Gershgorin [4] from 1931. If we define the associated Gershgorin row disks by

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

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

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

then Gershgorin'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 $\mathcal{C}'_j(A) := \sum_{i=1 \atop j \neq i}^{n} |a_{i,j}|$.

For $n \geq 2$, it is the case that it is generally easier to apply Gershgorin'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 Gershgorin, 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 Gershgorin 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^r_{i,j}(A) \subseteq G^r_i(A) \cup G^r_j(A), \text{ for all } i \neq j \text{ (} 1 \leq i, j \leq n \). \quad (1.10)$$

The case of equality in (1.10) is covered in

$$K^r_{i,j}(A) = G^r_i(A) \cup G^r_j(A) \text{ for } i \neq j \text{ only if } R_i(A) = R_j(A) = 0,$$

or if $R_i(A) = R_j(A) > 0$ 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 $i_1, i_2, \cdots, i_m$ are distinct positive integers in $[1, n]$. Then define

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

(1.12)

and set

$$H'_{(m)}(A) := \bigcup_{1 \leq i_1, i_2, \cdots, i_m \leq n} H'_{i_1, \ldots, i_m}(A) \quad (i_1, i_2, \cdots, i_m \text{ distinct}).$$

(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 $H'_{(m)}(A)$ now consists of $\binom{n}{m}$ sets $H'_{i_1, \ldots, i_m}(A)$.

But, for $m \geq 3$, the set $H'_{(m)}(A)$ 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

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

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 Brualdi-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:

\begin{align*}
(a_{i,i})_{i=1}^{n} \quad \text{and} \quad (R'_i(A))_{i=1}^{n},
\end{align*}

(2.1)

as do the sets $H'_{(m)}(A)$ in (1.13). If

$$\Omega(A) := \{D - [b_{i,j}] \in \mathbb{C}^{n \times n}, \, n \geq 2 : b_{i,i} = a_{i,i} \text{ and}$$

\begin{align*}
R'_i(B) = R'_i(A), & \quad 1 \leq i \leq n \},
\end{align*}

(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 Gershgorin disks for any $A \in \mathbb{C}^{n \times n}$.

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

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

$$\alpha_i \in \mathbb{C} \text{ and } r_i \geq 0 \quad (1 \leq i \leq n), \quad (2.5)$$

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

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

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

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

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

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

which is dependent on $\Delta_n$ 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 $\mu_m$ of $\mathcal{P}_n$, defined by

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

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 Gershgorin 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

\[
\sigma(\Omega(\Delta_n)) \subseteq F(\Delta_n, \mu), \quad \text{for any choice of } \Delta_n, \quad (2.11)
\]

where \( \Delta_n \) is defined in (2.5). As we know, (2.11) is valid, for any \( \Delta_n \), for the subsets \( \mu_1 \) and \( \mu_2 \) of (2.10), but not for \( \mu_m \) of (2.10) for \( m \geq 3 \). We say that a subset \( \mu \) of \( \mathcal{P}_n \) is an \textit{eigenvalue inclusion subset} of (2.11) is valid for any choice of \( \Delta_n \). Note that if \( \mu \) is such an eigenvalue inclusion subset of \( \mathcal{P}_n \), then so is

\[
\mu \cup q, \quad \text{for any } q \in \mathcal{P}_n, \quad (2.12)
\]

since adding \( q \) to \( \mu \) 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 \( \mu \) of \( \mathcal{P}_n \) which give the \textit{tightest} inclusion in (2.11). One such subset is given in

\textbf{Theorem 2.1} With \( n \geq 2 \) and with the definition of \( \mu_2 \) in (2.10),

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

and

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

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

\textit{Proof.} The inequalities of (2.13) and (2.14) are a direct consequence of Theorem 1.2 of Varga and Krautstengel [7].

From Theorem 2.1, we thus have our \textit{vindication} of Alfred Brauer's ovals of Cassini!

\section{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 \textit{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 \textit{path} \( \pi \) from vertex \( v_i \) to vertex \( v_j \) is a sequence \( i = i_0, i_1, \cdots, i_k = j \) of distinct vertices for which \( v_{i_0}v_{i_1}, v_{i_1}v_{i_2}, \cdots, v_{i_{k-1}}v_{i_k} \) are abutting arcs, and the \textit{length} of \( \pi \) is said to be \( k \). A directed graph is \textit{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 \textit{circuit} \( \gamma \) of \( \Gamma(A) \) is a sequence \( i_1, \cdots, i_p, i_{p+1} = i_1, \) where \( p \geq 2, i_1, \cdots, i_p \) are distinct, and \( v_{i_1}v_{i_2}, \cdots, v_{i_{p+1}}v_{i_1} \) are arcs of \( \Gamma(A) \). If we write \( \gamma = (i_1, i_2, \cdots, i_p) \) to identify this circuit \( \gamma \), then, noting that these \( \{i_j\}_{j=1}^p \) are distinct, each circuit \( \gamma = (i_1, i_2, \cdots, i_p) \) defines a partition in \( \mathcal{P}_n \). Thus, if \( \mathcal{C}(A) \) denotes the set of all circuits in \( \Gamma(A) \), we thus have

\[
\mathcal{C}(A) \subseteq \mathcal{P}_n \quad (3.1)
\]
(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 \textit{weakly irreducible} if each vertex \( v_i \) of \( \Gamma(A) \) belongs to \textit{some} circuit in \( \mathcal{C}(A) \). (Note that the matrix \( A \) of (1.14) is \textit{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 \textit{weakly irreducible}. If \( \mathcal{C}(A) \) denotes the circuits \( \gamma \) of the directed graph \( \Gamma(A) \) for \( A \), then

\[
\sigma(A) \subseteq \bigcup_{\gamma \in \mathcal{C}(A)} \left\{ z \in \mathbb{C} : \prod_{i \in \gamma} |z - a_{i,i}| \leq \prod_{i \in \gamma} R_i(A) \right\}. \tag{3.2}
\]

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

\[
\mathcal{B}^*(A) := \bigcup_{\gamma \in \mathcal{C}(A)} \left\{ z \in \mathbb{C} : \prod_{i \in \gamma} |z - a_{i,i}| \leq \prod_{i \in \gamma} R_i(A) \right\}. \tag{3.3}
\]

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

\[
\hat{\Omega}(A) := \{ B \in \Omega(A) : \text{the circuits } \gamma(B) \text{ from } \Gamma(B) \text{ are identical with those of } \Gamma(A) \}, \tag{3.4}
\]

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

\[
\sigma(\hat{\Omega}(A)) \subseteq \mathcal{B}^*(A). \tag{3.5}
\]

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

\[
\sigma(\hat{\Omega}(A)) \subseteq \sigma(\Omega(A)) = K^*(A), \tag{3.6}
\]

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

\[
\mathcal{B}^*(A) \subseteq K^*(A). \tag{3.7}
\]

To investigate this, we extend the definition of the Brualdi set \( \mathcal{B}^*(A) \) 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 \( \gamma \) in the directed graph \( \Gamma(A) \), we use the convention that

\[
\mathcal{B}^*(A) := \bigcup_{i=1}^{n} \{ a_{i,i} \} \tag{3.8}
\]

**Theorem 3.2** For any \( A = [a_{i,j}] \in \mathbb{C}^{n \times n}, \ n \geq 2 \), let the sets \( K^*(A) \) and \( \mathcal{B}^*(A) \) 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 \( \gamma \) in \( \Gamma(A) \), then \( B^r(A) \) is defined by (3.8). In this case, as each Cassini oval \( K^r_{i,j}(A) \) of (1.2), for \( i \neq j \), contains the points \( z = a_{i,i} \) and \( z = a_{j,j} \), then \( B^r(A) \subseteq K^r(A) \). Then, consider, any cycle \( \gamma = (i_1, i_2, \ldots, i_p) \) of \( C(A) \), whose associated term in \( B^r(A) \), from (3.3), defines the associated closed set in \( \mathbb{C} \):

\[
B^r_{\gamma}(A) := \left\{ z \in \mathbb{C} : \prod_{j=1}^{p} |z - a_{i_j,i_j}| \leq \prod_{j=1}^{p} R^r_{i_j}(A) \right\}. \tag{3.9}
\]

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

\[
\prod_{j=1}^{p} \left( \frac{|z - a_{i_j,i_j}|}{R^r_{i_j}(A)} \right) \leq 1,
\]

and squaring the above expression gives

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

The above product can also be expressed as

\[
\left( \frac{|z - a_{i_1,i_1}| \cdot |z - a_{i_2,i_2}|}{R^r_{i_1}(A) \cdot R^r_{i_2}(A)} \right) \cdot \left( \frac{|z - a_{i_2,i_2}| \cdot |z - a_{i_3,i_3}|}{R^r_{i_2}(A) \cdot R^r_{i_3}(A)} \right) \cdots \cdot \left( \frac{|z - a_{i_p,i_1}| \cdot |z - a_{i_1,i_1}|}{R^r_{i_p}(A) \cdot R^r_{i_1}(A)} \right) \leq 1.
\]

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

\[
\left( \frac{|z - a_{i_{\ell-1},i_{\ell}| \cdot |z - a_{i_{\ell+1},i_{\ell+1}}|}{R^r_{i_{\ell}}(A) \cdot R^r_{i_{\ell+1}}(A)} \right) \leq 1, \tag{3.10}
\]

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

Note that the inclusion of (3.7) does not require that \( A \) be weakly irreducible, so that \( B^r(A) \) 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 \( B^r(A) \) 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 \( B^r(A) \) 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 \( \gamma \) of \( \Gamma(A) \). For \( n = 10 \), the number of distinct circuits is 1,013, as opposed to 45 ovals of Cassini!
References


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 Guggenhein 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).