Integr. equ. oper. theory 57 (2007), 513–520

0378-620X/040513-8, *published online* December 26, 2006
DOI 10.1007/s00020-006-1472-y

Integral Equations and Operator Theory

The Brauer–Ostrowski Theorem for Matrices of Operators

Gerd Herzog and Christoph Schmoeger

Abstract. The classical Brauer-Ostrowski Theorem gives a localization of the spectrum of a matrix by a union of Cassini ovals. In this paper we prove a corresponding result for operator matrices.

Mathematics Subject Classification (2000). Primary 47A10; Secondary 47A05.

Keywords. Operator matrices, localization of the spectrum, Cassini ovals.

1. Introduction

In [5] and [1] Ostrowski and Brauer independently observed that each eigenvalue of a matrix $A = (a_{jk}) \in \mathbb{C}^{n\times n}$, $n \geq 2$ is contained in a Cassini oval

$$\left\{ \lambda \in \mathbb{C} : |\lambda - a_{rr}|\ |\lambda - a_{ss}| \leq \left(\sum_{r\neq k=1}^{n} |a_{rk}| \right) \left(\sum_{s\neq k=1}^{n} |a_{sk}| \right) \right\}$$

with $r \neq s$. In [2] Feingold and Varga obtained the corresponding result for block matrices. In several cases these results lead to a better localization of the spectrum of a matrix than Gershgorin's Theorem, compare [8] and the references given there. Affected by Gil''s and Salas' devolvements of Gershgorin's Theorem [3],[7], we study in this paper the Brauer-Ostrowski Theorem in the frame of operator matrices.

2. Notations

Let X be a complex Banach space, and $T : X \to X$ linear and bounded. In the sequel we consider:
the spectrum

$$\sigma(T) = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is bijective}\},$$

the resolvent set

$$\rho(T) = \mathbb{C} \setminus \sigma(T),$$

the point spectrum

$$\sigma_p(T) = \{\lambda \in \mathbb{C} : \exists x \in X : x \neq 0, (\lambda I - T)x = 0\},$$

the continuous spectrum

$$\sigma_c(T) = \{\lambda \in \mathbb{C} : \lambda \notin \sigma_p(T),\ (\lambda I - T)(X) \neq X,\ \overline{(\lambda I - T)(X)} = X\},$$

the residual spectrum

$$\sigma_r(T) = \{\lambda \in \mathbb{C} : \lambda \notin \sigma_p(T),\ \overline{(\lambda I - T)(X)} \neq X\},$$

the approximate point spectrum

$$\sigma_{ap}(T) = \{\lambda \in \mathbb{C} : \exists (x_n) \subseteq X : \|x_n\| = 1 \text{ and } (\lambda I - T)x_n \to 0\ (n \to \infty)\},$$

and the compression spectrum

$$\sigma_{com}(T) = \{\lambda \in \mathbb{C} : \overline{(\lambda I - T)(X)} \neq X\}.$$

Note, that $\sigma_p(T)$, $\sigma_c(T)$ and $\sigma_r(T)$ are pairwise disjoint, that

$$\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T),$$

and that

$$\sigma_r(T) = \sigma_{com}(T) \setminus \sigma_p(T).$$

Moreover let X^* denote the dual space of X, let T^* denote the adjoint of T, and note that

$$\sigma(T) = \sigma(T^*),\ \ \sigma_{com}(T) = \sigma_p(T^*) \text{ and } \|T\| = \|T^*\|.$$

3. Matrices of operators

Let $n \in \mathbb{N}$, $n \geq 2$, and $(X_1, \|\cdot\|_1), \ldots, (X_n, \|\cdot\|_n)$ complex Banach spaces. We consider the complex Banach space

$$X = X_1 \times \cdots \times X_n, \quad \|x\|_\infty = \max_{i=1}^{n} \|x_i\|_i\ (x = (x_1, \ldots, x_n) \in X).$$

Now, let $A : X \to X$ be linear and bounded. Then

$$A = (A_{jk}) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix},$$

where $A_{jk} : X_k \to X_j$ is linear and bounded $(j, k = 1, \ldots, n)$. For each $j \in \{1, \ldots, n\}$ we set,

$$p_j(A) = \sum_{j \neq k=1}^{n} \|A_{jk}\|, \quad q_j(A) = \sum_{j \neq k=1}^{n} \|A_{kj}\|.$$

For $r, s \in \{1, \dots, n\}$ with $r \neq s$ we define the following sets, corresponding to the ovals of Cassini:

$$C_{rs}^{(p)}(A) = \sigma(A_{rr}) \cup \sigma(A_{ss})$$
$$\cup \left\{ \lambda \in \rho(A_{rr}) \cap \rho(A_{ss}) : (\|(\lambda I - A_{rr})^{-1}\| \|(\lambda I - A_{ss})^{-1}\|)^{-1} \leq p_r(A) p_s(A) \right\},$$

and

$$C_{rs}^{(q)}(A) = \sigma(A_{rr}) \cup \sigma(A_{ss})$$
$$\cup \left\{ \lambda \in \rho(A_{rr}) \cap \rho(A_{ss}) : (\|(\lambda I - A_{rr})^{-1}\| \|(\lambda I - A_{ss})^{-1}\|)^{-1} \leq q_r(A) q_s(A) \right\}.$$

Since $A^* = (A_{jk})^* = (A_{kj}^*)$ we have

$$C_{rs}^{(p)}(A^*) = C_{rs}^{(q)}(A). \tag{3.1}$$

Next, let

$$C^{(p)}(A) := \bigcup_{r,s=1, r\neq s}^{n} C_{rs}^{(p)}(A), \quad C^{(q)}(A) := \bigcup_{r,s=1, r\neq s}^{n} C_{rs}^{(q)}(A),$$

and note that by means of (3.1), $C^{(p)}(A^*) = C^{(q)}(A)$. Moreover, observe that if $n = 2$, then $C_{12}^{(p)}(A) = C_{12}^{(q)}(A)$, hence $C^{(p)}(A) = C^{(q)}(A)$.

4. Localization of the spectrum

Theorem 4.1. *Let $A = (A_{jk}) : X \to X$ be linear and bounded. Then*

$$\sigma_{ap}(A) \subseteq C^{(p)}(A).$$

Proof. Let $\lambda \in \sigma_{ap}(A)$. For each $m \in \mathbb{N}$ there exists $x^{(m)} = (x_1^{(m)}, \dots, x_n^{(m)}) \in X$ such that

$$\|x^{(m)}\|_\infty = 1, \quad \|\lambda x^{(m)} - Ax^{(m)}\|_\infty \leq \frac{1}{m}.$$

For $m \in \mathbb{N}$ let $r(m), s(m) \in \{1, \dots, n\}$ be such that $r(m) \neq s(m)$, and

$$\|x_{r(m)}^{(m)}\|_{r(m)} \geq \|x_{s(m)}^{(m)}\|_{s(m)} \geq \|x_i^{(m)}\|_i \quad (i \in \{1, \dots, n\} \setminus \{r(m), s(m)\}).$$

By means of the pigeon hole principle we can assume without loss of generality that the sequences $(r(m))_{m=1}^\infty$ and $(s(m))_{m=1}^\infty$ are constant. Hence let $r = r(m)$ and $s = s(m)$. Then $r \neq s$ and, since $\|x^{(m)}\|_\infty = 1$,

$$\|x_i^{(m)}\|_i \leq \|x_s^{(m)}\|_s \leq \|x_r^{(m)}\|_r = 1 \quad (i \notin \{r, s\},\ m \in \mathbb{N}). \tag{4.1}$$

We define

$$z^{(m)} = (z_1^{(m)}, \dots, z_n^{(m)}) := \lambda x^{(m)} - Ax^{(m)} \quad (m \in \mathbb{N}), \tag{4.2}$$

and consider the following cases:

1. Let $\lambda \in \sigma(A_{rr}) \cup \sigma(A_{ss})$. Then $\lambda \in C_{rs}^{(p)}(A)$, thus $\lambda \in C^{(p)}(A)$.
2. Let $\lambda \in \rho(A_{rr}) \cap \rho(A_{ss})$. From (4.2) we get

$$z_r^{(m)} = (\lambda I - A_{rr}) x_r^{(m)} - \sum_{r\neq k=1}^{n} A_{rk} x_k^{(m)} \tag{4.3}$$

and

$$z_s^{(m)} = (\lambda I - A_{ss})x_s^{(m)} - \sum_{s \neq k=1}^{n} A_{sk} x_k^{(m)}. \tag{4.4}$$

By means of (4.3) we have

$$x_r^{(m)} = (\lambda I - A_{rr})^{-1} \left(z_r^{(m)} + \sum_{r \neq k=1}^{n} A_{rk} x_k^{(m)} \right),$$

thus, according to (4.1),

$$\begin{aligned} 1 = \|x_r^{(m)}\|_r &\leq \|(\lambda I - A_{rr})^{-1}\| \left(\|z_r^{(m)}\|_r + \sum_{r \neq k=1}^{n} \|A_{rk}\| \, \|x_k^{(m)}\|_k \right) \qquad (4.5) \\ &\leq \|(\lambda I - A_{rr})^{-1}\| \left(\frac{1}{m} + \|x_s^{(m)}\|_s p_r(A) \right). \end{aligned}$$

From (4.4) we get

$$x_s^{(m)} = (\lambda I - A_{ss})^{-1} \left(z_s^{(m)} + \sum_{s \neq k=1}^{n} A_{sk} x_k^{(m)} \right),$$

hence

$$\begin{aligned} \|x_s^{(m)}\|_s &\leq \|(\lambda I - A_{ss})^{-1}\| \left(\|z_s^{(m)}\|_s + \sum_{s \neq k=1}^{n} \|A_{sk}\| \, \|x_k^{(m)}\|_k \right) \qquad (4.6) \\ &\leq \|(\lambda I - A_{ss})^{-1}\| \left(\tfrac{1}{m} + \|x_r^{(m)}\|_r p_s(A) \right) \\ &= \|(\lambda I - A_{ss})^{-1}\| \left(\tfrac{1}{m} + p_s(A) \right). \end{aligned}$$

We proceed by proving that there exist $\alpha > 0$ and $m_0 \in \mathbb{N}$ such that

$$\|x_s^{(m)}\|_s \geq \alpha \quad (m \geq m_0). \tag{4.7}$$

If not, then there is a subsequence $(x_s^{(m_\nu)})$ of $(x_s^{(m)})$ with $x_s^{(m_\nu)} \to 0$ $(\nu \to \infty)$, thus (4.1) gives

$$x_i^{(m_\nu)} \to 0 \ (\nu \to \infty) \quad (i \in \{1, \ldots, n\} \setminus \{r\}),$$

and, by (4.3),

$$(\lambda I - A_{rr})x_r^{(m_\nu)} \to 0 \quad (\nu \to \infty).$$

Since $\|x_r^{(m_\nu)}\| = 1$ for all $\nu \in \mathbb{N}$, we get the contradiction

$$\lambda \in \sigma_{ap}(A_{rr}) \subseteq \sigma(A_{rr}).$$

Thus (4.7) holds. According to (4.5) and (4.6) we have

$$\begin{aligned} &\|x_s^{(m)}\|_s = 1 \cdot \|x_s^{(m)}\|_s \\ &\leq \|(\lambda I - A_{rr})^{-1}\| \, \|(\lambda I - A_{ss})^{-1}\| \left(\frac{1}{m} + \|x_s^{(m)}\|_s p_r(A) \right) \left(\frac{1}{m} + p_s(A) \right), \end{aligned}$$

and therefore, by (4.7),

$$1 \le \|(\lambda I - A_{rr})^{-1}\| \, \|(\lambda I - A_{ss})^{-1}\| \left(\frac{1}{m\|x_s^{(m)}\|_s} + p_r(A) \right) \left(\frac{1}{m} + p_s(A) \right)$$
$$\le \|(\lambda I - A_{rr})^{-1}\| \, \|(\lambda I - A_{ss})^{-1}\| \left(\frac{1}{\alpha m} + p_r(A) \right) \left(\frac{1}{m} + p_s(A) \right)$$

for $m \ge m_0$. With $m \to \infty$ we derive $\lambda \in C_{rs}^{(p)}(A) \subseteq C^{(p)}(A)$. □

In particular we have:

Corollary 4.2. $\sigma_p(A) \cup \sigma_c(A) \subseteq C^{(p)}(A)$.

Proof. Follows from Theorem 4.1 and $\sigma_p(A) \cup \sigma_c(A) \subseteq \sigma_{ap}(A)$. □

Corollary 4.3. $\sigma_{com}(A) \subseteq C^{(q)}(A)$.

Proof. Since $\sigma_{com}(A) = \sigma_p(A^*)$, Corollary 4.2 shows that

$$\sigma_{com}(A) \subseteq C^{(p)}(A^*) = C^{(q)}(A).$$ □

In case $n = 2$ we have seen that $C^{(p)}(A) = C^{(q)}(A)$. Hence we have:

Corollary 4.4. *If* $n = 2$*, then* $\sigma(A) \subseteq C^{(p)}(A)$.

5. Weighted norms

Let $w_1, \ldots, w_n > 0$. We define equivalent norms on $X_1, \ldots, X_n$, respectively, by setting

$$|||\xi|||_i = w_i \|\xi\|_i \quad (\xi \in X_i, \ i = 1, \ldots, n).$$

For the operators $A_{jk} : X_k \to X_j$ we have

$$\|A_{jk}\| = \sup_{\|\xi\|_k = 1} \|A_{jk}\xi\|_j,$$

hence

$$|||A_{jk}||| := \sup_{|||\xi|||_k = 1} |||A_{jk}\xi|||_j = \frac{w_j}{w_k} \|A_{jk}\| \quad (j, k = 1, \ldots, n).$$

By application of Theorem 4.1 to this situation we obtain:

Theorem 5.1. *Let* $w_1, \ldots, w_n > 0$*. Then*

$$\sigma_{ap}(A) \subseteq \bigcup_{r,s=1, r\ne s}^{n} \sigma(A_{rr}) \cup \sigma(A_{ss}) \cup \Big\{ \lambda \in \rho(A_{rr}) \cap \rho(A_{ss}) : \big(\|(\lambda I - A_{rr})^{-1}\| \, \|(\lambda I - A_{ss})^{-1}\| \big)^{-1} \le \left(\sum_{r \ne k=1}^{n} \frac{w_r}{w_k} \|A_{rk}\| \right) \left(\sum_{s \ne k=1}^{n} \frac{w_s}{w_k} \|A_{sk}\| \right) \Big\}.$$

Remark 5.2. Theorem 5.1 can be extended to the case that $W_i : X_i \to X_i$ is linear, bounded and invertible ($i = 1, \dots, n$) and

$$|||\xi|||_i = \|W_i \xi\|_i \quad (\xi \in X_i,\ i = 1, \dots, n).$$

Then

$$|||A_{jk}||| = \|W_j A_{jk} W_k^{-1}\| \quad (j, k = 1, \dots, n),$$

and the corresponding inclusion for $\sigma_{ap}(A)$ is valid.

Now, consider the scalar matrix

$$B = \begin{pmatrix} 0 & \|A_{12}\| & \|A_{13}\| & \dots & \|A_{1n}\| \\ \|A_{21}\| & 0 & \|A_{23}\| & \dots & \|A_{2n}\| \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \|A_{n1}\| & \dots & \dots & \|A_{n(n-1)}\| & 0 \end{pmatrix},$$

and its spectral radius $r(B) = \max_{\mu \in \sigma(B)} |\mu|$. For each $\tau \geq 0$ we define

$$C_\tau := \bigcup_{r,s=1, r\neq s}^{n} \sigma(A_{rr}) \cup \sigma(A_{ss}) \cup \Big\{ \lambda \in \rho(A_{rr}) \cap \rho(A_{ss}) : \big(\|(\lambda I - A_{rr})^{-1}\|\ \|(\lambda I - A_{ss})^{-1}\|\big)^{-1} \leq \tau^2 \Big\}.$$

Theorem 5.3. *Let B and $r(B)$ be as above. Then*

$$\sigma(A) \subseteq C_{r(B)}.$$

Proof. Let P denote the $n \times n$ matrix with each entry equals 1. Let $\varepsilon > 0$. According to [6, Theorem 10.20] there exists $\delta > 0$ such that $r(B + \delta P) \leq r(B) + \varepsilon$. Now, $B + \delta P$ is irreducible and therefore has a strictly positive Perron eigenvector $v = (v_1, \dots, v_n) \in (0, \infty)^n$.

Set $w_k = v_k^{-1}$ ($k = 1, \dots, n$) and let $j \in \{1, \dots, n\}$. From $(B + \delta P)v = r(B + \delta P)v \leq (r(B) + \varepsilon)v$ (coordinatewise) we derive

$$(r(B) + \varepsilon)v_j \geq \delta v_j + \sum_{j \neq k=1}^{n} (\|A_{jk}\| + \delta) v_k,$$

hence

$$\sum_{j \neq k=1}^{n} \frac{w_j}{w_k} \|A_{jk}\| = \sum_{j \neq k=1}^{n} \frac{v_k}{v_j} \|A_{jk}\| \leq r(B) + \varepsilon.$$

Now Theorem 5.1 shows that $\sigma_{ap}(A) \subseteq C_{r(B)+\varepsilon}$, and with $\varepsilon \to 0+$ we obtain $\sigma_{ap}(A) \subseteq C_{r(B)}$.

By replicating this proof with A^* instead of A we obtain

$$\sigma_{com}(A) = \sigma_p(A^*) \subseteq C_{r(B^\top)} = C_{r(B)},$$

since B and its transposed $B^\top$ have the same spectral radius. So, finally

$$\sigma(A) = \sigma_{ap}(A) \cup \sigma_{com}(A)$$

proves $\sigma(A) \subseteq C_{r(B)}$. □

6. Examples

In order to apply Theorem 4.1,5.1 or 5.3 it is comfortable if the expressions

$$\|(\lambda I - A_{jj})^{-1}\|^{-1} \quad (j = 1, \ldots, n)$$

have a simple structure.

If T is a normal operator on a complex Hilbert space, then

$$\|(\lambda I - T)^{-1}\|^{-1} = \operatorname{dist}(\lambda, \sigma(T)) \quad (\lambda \in \rho(T))$$

see [4, p. 277].

If T is a multiplication operator on a space of complex valued continuous function $C(K)$ (K a compact metric space, say, and $C(K)$ endowed with the maximum norm), that is $(T\xi)(t) = g(t)\xi(t)$ $(t \in K)$ for some $g \in C(K)$, then $\sigma(T) = g(K)$ and likewise

$$\|(\lambda I - T)^{-1}\|^{-1} = \operatorname{dist}(\lambda, \sigma(T)) = \operatorname{dist}(\lambda, g(K)) \quad (\lambda \in \rho(T)).$$

For example, let $X = C(K)^n$, and let $A = (A_{jk}) : X \to X$ be such that $(A_{jj}\xi)(t) = g_j(t)\xi(t)$ with $g_j \in C(K)$ $(j = 1, \ldots, n)$. Let B be as in section 5. Then, according to Theorem 5.3

$$\sigma(A) \subseteq \bigcup_{r,s=1, r\neq s}^{n} \{\lambda : \operatorname{dist}(\lambda, g_r(K))\operatorname{dist}(\lambda, g_s(K)) \leq r(B)^2\}.$$

In the following example let $X_3 = C([0, 1])$ be endowed with the maximum norm $\|\cdot\|_3$, and $X_2 = C^1([0, 1])$, $X_1 = C^2([0, 1])$ endowed with the norms $\|\xi\|_2 = \max\{\|\xi\|_3, \|\xi'\|_3\}$ and $\|\xi\|_1 = \max\{\|\xi\|_3, \|\xi'\|_3, \|\xi''\|_3\}$, respectively. Let $\alpha \geq 0$, and let $A : X \to X$ be defined by

$$(Ax)(t) = \begin{pmatrix} x_1(t) + \alpha \int_0^1 \cos(ts)x_3(s)ds \\ x_1'(t) - x_2(t) \\ x_1''(t) + x_2'(t) + \exp(2\pi i t)x_3(t) \end{pmatrix}.$$

Note that $\sigma(A) = \{\lambda : |\lambda| = 1\}$ if $\alpha = 0$. Application of Theorem 5.3 proves that

$$\begin{aligned} \sigma(A) &\subseteq C_{r(B)} = \{\lambda : |\lambda^2 - 1| \leq r(B)^2\} \\ &\quad \cup\{\lambda : |\lambda - 1|\ ||\lambda| - 1| \leq r(B)^2\} \cup \{\lambda : |\lambda + 1|\ ||\lambda| - 1| \leq r(B)^2\} \\ &= \{\lambda : |\lambda - 1|\ ||\lambda| - 1| \leq r(B)^2\} \cup \{\lambda : |\lambda + 1|\ ||\lambda| - 1| \leq r(B)^2\}, \end{aligned}$$

with

$$B = \begin{pmatrix} 0 & 0 & \alpha \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

It is easy to check that $r(B) = 1$ if $\alpha = 1/2$, and if $\alpha < 1/2$, then $r(B) < 1$ and $0 \notin C_{r(B)}$. Thus A is invertible in this case. Figure 1 shows $C_{r(B)}$ with $r(B) \approx 0.915$ for $\alpha = 0.4$, and $r(B) \approx 0.231$ for $\alpha = 0.01$.

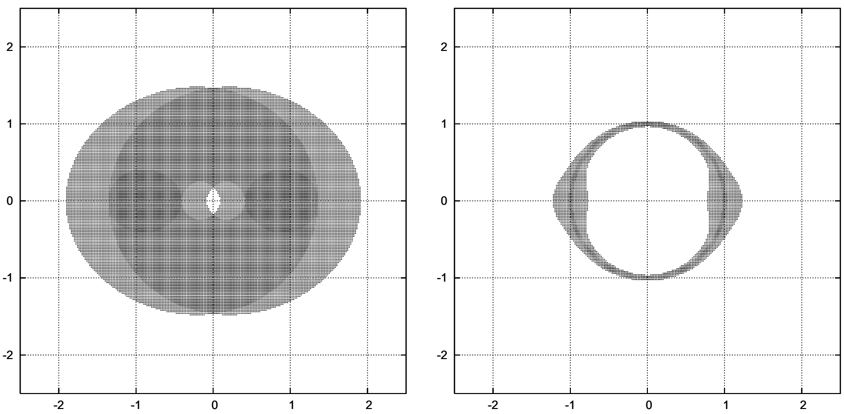

FIGURE 1. $\alpha = 0.4$, $\alpha = 0.01$

References

[1] A. Brauer, *Limits for the characteristic roots of a matrix. II.* Duke Math. J. **14** (1947), 21-26.

[2] D.G. Feingold; R.S. Varga, *Block diagonally dominant matrices and generalizations of the Gerschgorin circle theorem.* Pacific J. Math. **12** (1962), 1241-1250.

[3] M.I. Gil', *Invertibility conditions and bounds for spectra of operator matrices.* Acta Sci. Math. (Szeged) **67** (2001), 299-314.

[4] T. Kato, *Perturbation theory for linear operators.* Die Grundlehren der mathematischen Wissenschaften, Band 132 Springer-Verlag New York, Inc., New York, 1966.

[5] A. Ostrowski, *Über die Determinanten mit überwiegender Hauptdiagonale.* Comment. Math. Helv. **10** (1937), 69-96.

[6] W. Rudin, *Functional analysis.* McGraw-Hill Series in Higher Mathematics. McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1973.

[7] H.N. Salas, *Gershgorin's theorem for matrices of operators.* Linear Algebra Appl. **291** (1999), 15-36.

[8] R.S. Varga, *Geršgorin and his circles.* Springer Series in Computational Mathematics, 36. Springer-Verlag, Berlin, 2004.

Gerd Herzog and Christoph Schmoeger
Institut für Analysis, Universität Karlsruhe
D-76128 Karlsruhe
Germany
e-mail: `Gerd.Herzog@math.uni-karlsruhe.de`
`christoph.schmoeger@math.uni-karlsruhe.de`

Submitted: April 19, 2006
Revised: November 17, 2006