Gershgorin Domains for Partitioned Matrices

A. van der Sluis
Mathematisch Instituut
Rijksuniversiteit
Utrecht-Uithof, The Netherlands

Submitted by R. S. Varga

ABSTRACT

Inclusion domains for the eigenvalues of a partitioned matrix are specified in terms of perturbations of its diagonal blocks. The size of such perturbations is measured using the Kantorovitch-Robert-Deutsch vectorial norms. The inclusion domains obtained thereby are compared with inclusion domains otherwise obtainable. A few new properties of vectorial norms are proved as well. A result of Bonsall, Schneider, Strang and Ljubič is generalized.

1. INTRODUCTION AND SOME RESULTS

In [5] the well-known Gershgorin circle theorem is generalized to partitioned matrices. With the conventions (see also Sec. 2)

(1) A is partitioned into blocks A_{ij}, $i,j=1,\dots,N$,

(2) $\|A_{ij}\|$ is the operator norm of A_{ij} associated with vector norms in source and image space,

(3) $\sum_{j,\neq i}$ denotes $\sum_{j=1,\, j\neq i}^{N}$ for any given value of i,

the following theorem is a reformulation of [5, Theorem 2]:

THEOREM 1.1. *All eigenvalues of A are contained in the set $G=\bigcup G_i$, where G_i is the set of all $\lambda\in\mathbb{C}$ satisfying*

$$\|(A_{ii}-\lambda I_i)^{-1}\|^{-1}\leqslant\sum_{j,\neq i}\|A_{ij}\|. \tag{1.2}$$

The expression in the left-hand member of (1.2) *is to be given the value* 0 *if* λ *is an eigenvalue of* A_{ii}.

 0024-3795/79/040265+16$01.75

The following theorem (cf. [7, Theorem 1]; for a proof see Sec. 2.7) gives an alternative characterization of the sets G_i:

THEOREM 1.3. *For any i the set G_i is identical with the set of eigenvalues of all matrices $A_{ii}+X$, with $\|X\| \leqslant \Sigma_{j,\neq i}\|A_{ij}\|$, X complex.*

This theorem specifies the inclusion domain G for the eigenvalues of A in terms of *perturbations of the diagonal blocks*. It is the purpose of this paper to give a more general class of inclusion domains (to be called *Gershgorin domains*) in this fashion, i.e. as the set of eigenvalues of all matrices $A_{ii}+X$, $i=1,2,\dots,N$, where the complex matrix X is majorized in some sense by the ith row of off-diagonal blocks of A. This will be done by means of *vectorial norms* for matrices (to be introduced in Sec. 3).

An example of such a vectorial norm and the corresponding Gershgorin domain is given by the following theorem, which is a special case of our general Theorem 4.1:

THEOREM 1.4. *For any matrix B let $\|B\|'$ denote the column vector whose kth coordinate is the sum of the moduli of the elements in the kth row of B. For any i let G_i denote the set of eigenvalues of all matrices $A_{ii}+X$ with $\|X\|' \leqslant \Sigma_{j,\neq i}\|A_{ij}\|'$ (this vector inequality to hold coordinatewise). Then the set $G=\bigcup G_i$ contains all eigenvalues of A.*

It will be investigated whether the Gershgorin domains derived from Theorem 4.1 for various choices of the vectorial norm may improve those obtainable from Theorem 1.1. On the whole the outcome of this will be somewhat disappointing (cf. Remark 5.9). Yet it will turn out (cf. Remark 5.10) that for all matrices A the Gershgorin domain given by Theorem 1.4 is a subset of the one given by Theorem 1.1 for the important case of the Hölder ∞-norm, whereas for some matrices A it is a strict subset, provided the partitioning is submaximal (for the definition see Sec. 2.1).

Our results will also allow a unified approach to Wilkinson's perturbation theory (cf. [11, Chapter 2, Secs. 15–21], and our Theorem 5.4 and Remark 5.5) and allow a generalization of this theory to partitioned matrices, where the off-diagonal blocks may be measured in a variety of ways.

2. PARTITIONING AND NORMS

2.1

Let any (column) vector $v\in\mathbb{C}^n$ be partitioned as $(v_1,\dots,v_N)^T$, where v_i has n_i coordinates (so $\Sigma n_i=n$). This means that $\mathbb{C}^n$ itself is considered as a direct sum $S_1 \oplus \cdots \oplus S_N$. We always suppose $N \geqslant 2$. The partitioning will be called *maximal* if $n_i = 1$ for all values of i, *submaximal* if $n_i > 1$ for at least one value of i, and *strongly submaximal* if $n_i > 1$ for at least two values of i.

2.2

Each of the S_i is supposed to be provided with a norm $\|\cdot\|$. If a norm is written with a lower case letter or number subscript, like $\|\cdot\|_p$ or $\|\cdot\|_2$, this denotes a Hölder norm: $\|v_i\|_p = (\Sigma_k |\xi_k|^p)^{1/p}$ if $\xi_1, \xi_2, \ldots$ are the coordinates of v_i.

2.3

Corresponding to the given partitioning of vectors, any $n \times n$ matrix A is partitioned as (A_{ij}).

2.4

We denote the set of linear mappings of S_j into S_i by L_{ij}. The norms on S_i and S_j then induce (operator) norms for the elements of L_{ij}, likewise denoted by $\|\cdot\|$. We may consider A_{ij} (cf. Sec. 2.3) as an element of L_{ij}, and thereby $\|A_{ij}\|$ is defined.

2.5

For any j we may also consider *row* vectors u_j of n_j coordinates. They may be considered as linear functionals on S_j by having them assign the value $u_j v_j = \Sigma \eta_k \xi_k$ to the vector $v_j \in S_j$ if u_j and v_j have coordinates η_k and ξ_k respectively. Then those row vectors form the dual space S_j^D of S_j, and have operator norms

$$\|u_j\| = \max_{v_j \in S_j} \frac{|u_j v_j|}{\|v_j\|}.$$

If, in particular, $\|\cdot\|$ is the Hölder p-norm on S_j, then S_j^D gets the Hölder q-norm with $1/p + 1/q = 1$, i.e., $\|u_j\| = \|u_j\|_q = (\Sigma |\eta_k|^q)^{1/q}$ if $\eta_1, \eta_2, \ldots$ are the coordinates of $u_j \in S_j^D$ (cf. [6, Chapter III, Example 1.25]).

2.6

We recall (cf. [6, Chapter I, (2.27), or Chapter III, Corollary 1.24]) that for any $v_j \in S_j, v_j \neq 0$, there exists at least one element $v_j^D \in S_j^D$ such that $v_j^D v_j = \|v_j^D\| \, \|v_j\| = 1$ (implying that v_j is a maximizing vector of v_j^D). According to the terminology of Bauer [1], v_j^D and v_j then might be called a *dual pair*.

Defining $M = v_i v_j^D$ for $v_i \in S_i$, $v_j \in S_j$, we have $M \in L_{ij}$, $M v_j = v_i$ and $\|M\| = \|v_i\| / \|v_j\|$.

2.7

We now prove Theorem 1.3. Recalling the definition $\text{glb}(B) = \min_{v\neq 0}\|Bv\|/\|v\|$ for any matrix B, we note that the left-hand side of (1.2) equals $\text{glb}(A_{ii}-\lambda I_i)$. Let c denote the right-hand side of (1.2). Then (1.2) implies the existence of a vector $v_i, v_i\neq 0$, with $\|(A_{ii}-\lambda I_i)v_i\| \leqslant c\|v_i\|$. Now take $X = -(A_{ii}-\lambda I_i)v_i v_i^D$. Then $\|X\| \leqslant c\|v_i\|/\|v_i\| = c$ (cf. the norm of M in Sec. 2.6), and $(A_{ii}-\lambda I_i + X)v_i = 0$, proving that λ is an eigenvalue of $A_{ii}+X$.

Conversely, if $(A_{ii}+X-\lambda I_i)v_i = 0$ for a certain $\|X\| \leqslant c$ and $v_i \neq 0$, then $\|(A_{ii}-\lambda I_i)v_i\| = \|Xv_i\| \leqslant c\|v_i\|$; hence $\text{glb}(A_{ii}-\lambda I_i) \leqslant c$.

3. VECTORIAL NORMS FOR LINEAR MAPPINGS

3.1

We consider the following expressions for measuring the elements of L_{ij} for any i and j:

(a) $\|X\|' = \|X\|$ (i.e., the usual operator norm).

(b) $\|X\|' = (\|X_{1\cdot}\|, \ldots, \|X_{n_i\cdot}\|)^T$ (i.e., a column vector whose elements are the operator norms of the rows of X in the sense of Sec. 2.5, i.e., those rows considered as elements of S_j^D).

(c) $\|X\|' = \|(\|X_{1\cdot}\|, \ldots, \|X_{n_i\cdot}\|)^T\|''$, where $\|\cdot\|''$ denotes some *monotonic* vector norm (in the sense of [2]; examples are the Hölder norms) depending on i. As a special case we may take $\|\cdot\| = \|\cdot\|_p$ (the Hölder p-norm) and $\|\cdot\|'' = \|\cdot\|_q$ (the Hölder q-norm) with $1/p + 1/q = 1$; then *rows* are measured in the Hölder q-norm (cf. Sec. 2.5), and hence $\|X\|'$ is the Hölder q-norm of the column vector having as coordinates all the elements of X in some order. Taking $q = 1, 2, \infty$, we get (where x_{rs} denotes an *element* of X):

(c1) $\|X\|' = \sum_{r,s} |x_{rs}|$;

(c2) $\|X\|' = \|X\|_F = \sqrt{\sum_{r,s} |x_{rs}|^2}$ (the Frobenius norm);

(c3) $\|X\|' = \max_{r,s} |x_{rs}|$.

The notation $\|X\|' \leqslant \|Y\|'$ has an obvious meaning not only in cases (a) and (c) but also in case (b), provided that $X \in L_{ij}$ and $Y \in L_{ik}$ for the same value of i.

3.2

We note that in all cases the following properties hold—*and those are the properties that we shall use later on*—even if for different values of i we make different choices from Sec. 3.1, (a), (b) and (c):

(p) For any i, j, $\|\cdot\|'$ is a mapping of L_{ij} into the set of non-negative elements of a partially ordered real linear vector space P_i (as defined in [9, Definition 11.1], for example).

(q) $\|X+Y\|' \leqslant \|X\|' + \|Y\|'$ if X and Y belong to the same L_{ij}.

(r) $\|X\|' = 0_{P_i} \Leftrightarrow X = 0_{L_{ij}}$ if $X \in L_{ij}$ (0 denoting the respective zero elements).

(s) $\|XM\|' \leqslant \|M\| \, \|X\|'$ if $X \in L_{ij}$ and $M \in L_{jk}$. This is obvious for (a); for (b) and (c) it follows from the fact that for any row vector u the product uM may be considered as the result of applying the adjoint M^* of M to the linear functional u, and that the norm of M^* on the dual space equals $\|M\|$ (cf. [6, Chapter I, (4.8)]), whence $\|uM\| \leqslant \|u\| \, \|M\|$.

(t) If $X \in L_{ii}$, and a and b are non-negative elements of P_i such that $\|X\|' \leqslant a+b$, then there exist Y and $Z \in L_{ii}$ such that $X = Y + Z$, $\|Y\|' \leqslant a$, $\|Z\|' \leqslant b$.

As a consequence of (s) we have:

(u) $\|tX\|' \leqslant t\|X\|'$ for any scalar t with $0 \leqslant t \leqslant 1$.

3.3

Properties (p), (q), (r) and (u) are very similar to the ones in [4, (1.1), (1.2), (1.3)], which define a *vectorial norm*. Likewise, property (t) is very similar to the one in [4, Proposition 1, statement (iii)], which says that a vectorial norm is *regular*.

Hence, $\|\cdot\|'$ is very similar to a regular vectorial norm, which, on account of (s), might be said to be *compatible* with the norms on the spaces S_i.

Actually, *all mappings* $\|\cdot\|'$ *considered in Sec.* 3.1 *are indeed regular vectorial norms*, as may be checked right away. More generally we have:

THEOREM 3.1. *Let all P_i be finite-dimensional cartesian spaces in which $a \geqslant 0$ means that all coordinates of a are non-negative. Let a mapping $\|\cdot\|$ satisfy* (p), (q), (r) *and* (s). *Then:*

(a) $\|\cdot\|$ *is a vectorial norm on all L_{ij}.*

(b) *This vectorial norm is regular on L_{ii} if and only if* (t) *is satisfied.*

Proof. (a) is trivial, whereas (b) follows from the following lemma, which may have some interest on its own. ■

N.B. In the following lemma we will not make the assumption about the P_i that was made in Theorem 3.1.

LEMMA 3.2. *Let* (t′) *denote the property which arises from* (t) *if the* $\leqslant$ *signs are replaced by* $=$ *signs. Then:*

(a) *The properties* (p), (q), (t) *imply* (t′).

(b) *The properties* (p), (t′) *imply* (t) *if the spaces* P_i *are Riesz spaces* (*cf.* [9, *Definition* 11.1]; *finite-dimensional cartesian spaces ordered as in Theorem* 3.1 *are special cases*).

Proof.

(a) If $\|X\|' = a+b$ with $a,b \geqslant 0$, then $X = Y+Z$ with $\|Y\|' \leqslant a$, $\|Z\|' \leqslant b$. Then $\|Y'\| \geqslant \|X\|' - \|Z\|' \geqslant a+b-b=a$. Hence $\|Y\|' = a$. Likewise $\|Z\|' = b$.

(b) If $\|X\|' \leqslant a+b$ with $a,b \geqslant 0$, then there exist a' and b' such that $0 \leqslant a' \leqslant a$, $0 \leqslant b' \leqslant b$, $\|X\|' = a'+b'$ (cf. [9, Corollary 15.6]). Hence $X = Y+Z$ with $\|Y\|' = a' \leqslant a$, $\|Z\|' = b' \leqslant b$. ∎

4. RESULTS

Let $\|\cdot\|'$ be as defined in Sec. 3.1, (a), (b) or (c), or more generally, let $\|\cdot\|'$ satisfy Sec. 3.2, (p), (q), (r), (s), (t). Then we have the following theorem (which is our main theorem):

THEOREM 4.1. *Let* A *be a given matrix, partitioned as above. Let* G_i *denote the set of all eigenvalues of the matrices* $A_{ii} + X$ *for all* $X \in L_{ii}$ *with* $\|X\|' \leqslant \sum_{j,\neq i} \|A_{ij}\|'$. *Then*

(a) *All eigenvalues of* A *are contained in* $\bigcup G_i$.

(b) *If, for some integer* m, $G' =_{\mathrm{def}} G_1 \cup G_2 \cup \cdots \cup G_m$ *is isolated* (*in the sense of having a positive distance*) *from* $G'' =_{\mathrm{def}} G_{m+1} \cup \cdots \cup G_N$, *then* G' *contains exactly* $n_1 + \cdots + n_m$ *eigenvalues of* A, *each one counted according to its multiplicity.*

(c) *If* $\bigcap G_i \neq \varnothing$ *and* μ *is any number contained in* $\bigcap G_i$, *then there exist matrices* $\tilde{A}_{ij} \in L_{ij}$, $\|\tilde{A}_{ij}\|' \leqslant \|A_{ij}\|'$, *such that* μ *is an eigenvalue of the matrix* $\tilde{A}$ *arising from* A *by substituting* $\tilde{A}_{ij}$ *for* A_{ij} *for all* $i \neq j$.

Proof.

(a) Suppose $Av=\lambda v$ for a certain $v\neq 0$. Let $\|v_i\|=\max_j\|v_j\|$ (v_j denoting the jth element of the partitioning of v according to Sec. 2.1). Now define $M_{ji}=v_jv_i^D$, whence $\|M_{ji}\|\leqslant 1$ (cf. Sec. 2.6), and take $X=\Sigma_{j,\neq i}A_{ij}M_{ji}$. Then $\|X\|'\leqslant \Sigma_{j,\neq i}\|A_{ij}\|'$ and $(A_{ii}+X)v_i=\Sigma_jA_{ij}v_j=\lambda v_i$. Hence $\lambda\in G_i$.

(b) Consider the matrices $A(t)$ arising from A by multiplying all off-diagonal blocks A_{ij} by t, $0\leqslant t\leqslant 1$. Then $\|X\|'\leqslant\Sigma_{j,\neq i}\|tA_{ij}\|'$ implies $\|X\|'\leqslant t\Sigma_{j,\neq i}\|A_{ij}\|'\leqslant\Sigma_{j,\neq i}\|A_{ij}\|'$, and hence $G'(t)\subset G'$ and $G''(t)\subset G''$ for the corresponding G' and G''. Since the eigenvalues of $A(t)$ depend continuously on t, the number of eigenvalues in $G'(t)$ is independent of t. Because of Sec. 3.2, (r), $G_i(0)$ consists exactly of the eigenvalues of A_{ii}. Hence $G'(0)$ contains exactly $n_1+\cdots+n_m$ eigenvalues.

(c) For all i there are now $X_i\in L_{ii}$ with $\|X_i\|'\leqslant\Sigma_{j,\neq i}\|A_{ij}\|'$ and $v_i\in S_i$ with $\|v_i\|=1$ such that $(A_{ii}+X_i)v_i=\mu v_i$. Write X_i as $\Sigma_{j,\neq i}X_{ij}$ with $X_{ij}\in L_{ii}$ and $\|X_{ij}\|'\leqslant\|A_{ij}\|'$ [cf. Sec. 3.2, (t)]. Define $M_{ij}=v_iv_j^D$. Hence $\|M_{ij}\|=1$ and $M_{ij}v_j=v_i$. Define $\tilde{A}_{ij}=X_{ij}M_{ij}$. Then $\|\tilde{A}_{ij}\|'\leqslant\|A_{ij}\|'$ and $\Sigma_{j,\neq i}\tilde{A}_{ij}v_j=X_iv_i$. Hence $A_{ii}v_i+\Sigma_{j,\neq i}\tilde{A}_{ij}v_j=A_{ii}v_i+X_iv_i=\mu v_i$. ■

Remark 4.2. The sets G_i do not alter if in their definition we restrict X to be a matrix of rank $\leqslant 1$. Indeed, if $(A_{ii}+X)v_i=\lambda v_i$, then $(A_{ii}+Xv_iv_i^D)v_i=\lambda v_i$ and $\|Xv_iv_i^D\|'\leqslant\|X\|'$. Likewise, the matrices $\tilde{A}_{ij}$ in part (c) of the theorem may be restricted to rank $\leqslant 1$, as appears from the proof.

Remark 4.3. Regarding the *conditions* of Theorem 4.1 we note:

(a) We may replace Sec. 3.2, (s), by the weaker requirement that $\|XM\|'\leqslant\|M\|\,\|X\|'$ for matrices M of rank $\leqslant 1$. But then Sec. 3.2, (u), should be required explicitly, unless we are not interested in part (b) of the theorem.

(b) We may omit Sec. 3.2, (t), if we are not interested in part (c) of the theorem.

(c) We may omit Sec. 3.2, (t), also in the important case $N=2$, as is apparent from the proof.

Remark 4.4. Regarding the *proof* of Theorem 4.1, we note that the somewhat abstract-looking apparatus in Secs. 2 and 3 was developed only to make possible a unified treatment of the various cases in Sec. 3.1. For any concrete choice of norms we could virtually omit Secs. 2 and 3 without much effect on the proof of the theorem.

5. COMMENTS ON THEOREM 4.1

In this section we give a number of comments on and consequences of Theorem 4.1.

REMARK 5.1. Regarding Theorem 4.1(a), we note:

(a) It is identical with the well-known first Gershgorin circle theorem (cf. [11, Chapter 2, Sec. 13, Theorem 3]) for the maximal partitioning (cf. Sec. 2.1).

(b) It is equivalent to Theorem 1.1 (and thus to [5, Theorem 2]) if $\|\cdot\|'$ has been chosen according to Sec. 3.1, (a), as a consequence of Theorem 1.3.

REMARK 5.2. Regarding Theorem 4.1(b), we note:

(a) It is identical with the well-known second Gershgorin circle theorem (cf. [11, Chapter 2, Sec. 13, Theorem 4]) for the maximal partitioning.

(b) It is equivalent to [5, Theorem 4] if $\|\cdot\|'$ has been chosen according to Sec. 3.1, (a).

(c) If $\|\cdot\|'$ has been chosen according to Sec. 3.1, (a), (b) or (c), the requirement that G' be isolated from G'' is satisfied as soon as $G'\cap G''=\varnothing$, since the sets G' and G'' are now compact.

REMARK 5.3. Regarding Theorem 4.1(c), we note:

(a) It doesn't seem to be similar to anything in the literature [possibly because of (b) below].

(b) It becomes trivial for the maximal partitioning.

THEOREM 5.4. *If a block diagonal matrix D, in which the eigenvalues of the block D_{11} are distinct from the other ones, is perturbed by εB, where ε is a sufficiently small parameter, then the n_1 eigenvalues of the matrix $D+\varepsilon B$ which are nearest to the eigenvalues of D_{11} are eigenvalues of $D_{11}+\varepsilon B_{11}+O(\varepsilon^2)$.*

Proof. Choose norms on the S_i and take the corresponding operator norms for the blocks. Multiply the first block-row by ε/k and the first block-column by k/ε for some suitable constant k (this leaves the eigenvalues unaltered), and apply Theorem 4.1(b). ■

REMARK 5.5. Regarding this theorem and its proof, we note:

(a) The technique used in this proof is the same as used in [11, Chapter 2, Sec. 15] to prove the perturbation properties of a simple eigenvalue of a diagonalizable matrix.

(b) The result in Theorem 5.4 comprises, however, the cases 2, 3 and 4 in [11, Chapter 2, Secs. 17, 19, 21] as well.

(c) We did not have to use the artifact used in deriving [11, Chapter 2, (20.5)]—i.e., apply a rather skew similarity transform to a Jordan block matrix in order to reduce the ones on the codiagonal to smaller quantities.

In view of Remark 5.1(b), it is interesting to know whether Theorem 4.1(a) may give sharper results if $\|\cdot\|'$ has been chosen according to Sec. 3.1, (b) or (c), than with Sec. 3.1, (a).

DEFINITION 5.6. A choice (1) for $\|\cdot\|'$ will be called *stronger* for Theorem 4.1(a) than a choice (2) if the corresponding Gershgorin domains $G^{(1)}$ and $G^{(2)}$ satisfy $G^{(1)} \subset G^{(2)}$ for all matrices A and $G^{(1)} \neq G^{(2)}$ for at least one A. The choices will be called *equivalent* if $G^{(1)} = G^{(2)}$ for all A. The choices will be called *disparate* if they are not equivalent and if neither one is stronger than the other.

REMARK 5.7. Regarding the notion of disparateness we note:

(a) The choices (1) and (2) are disparate if and only if there exists a matrix A for which $G^{(1)} \setminus G^{(2)}$ is nonempty and there exists a matrix A for which $G^{(2)} \setminus G^{(1)}$ is nonempty.

(b) If a pair of choices is disparate, this does not mean that it doesn't matter which choice one uses. Indeed, it may happen that $G^{(1)}$ is a small subset of $G^{(2)}$ for one matrix A and that $G^{(2)}$ is a small subset of $G^{(1)}$ for another. It may also happen that $G^{(1)}$, say, extends at most a little beyond $G^{(2)}$ for all matrices A, whereas $G^{(2)}$ extends a great deal beyond $G^{(1)}$ for some matrices A, which means that for practical purposes one might consider choice (1) to be stronger than choice (2).

(c) See Remark 5.11.

THEOREM 5.8. *Table 1 displays, for quite a few choices of* $\|\cdot\|'$, *whether or not one is stronger than the other for Theorem* 4.1(a). *In cases 5.8.1 and 5.8.2 the partitioning is assumed to be submaximal; in 5.8.3, 5.8.4 and 5.8.5 it is assumed to be strongly submaximal (cf. Sec. 2.1).*

TABLE 1

Case	Choices for $\|\cdot\|'$[a]	Notes
5.8.1	(1): (a), with $\|\cdot\|^{(1)}$ on the S_i (2): (a), with $\|\cdot\|^{(2)}$ on the S_i	Equivalent or disparate for any norms $\|\cdot\|^{(1)}$ and $\|\cdot\|^{(2)}$
5.8.2	(1): (a), with $\|\cdot\|_\infty$ on the S_i (2): (b), with $\|\cdot\|_\infty$ on the S_i	(2) stronger than (1)
5.8.3	(1): (a), with $\|\cdot\|_p$ on the S_i (2): (b), with $\|\cdot\|_q$ on the S_i	Disparate for all cases distinct from 5.8.2
5.8.4	(1): (a), with $\|\cdot\|_p$ on the S_i (2): (c), with $\|\cdot\|_p$ on the S_i and $\|\cdot\|'' = \|\cdot\|_p$	(1) stronger than (2) for $p \neq \infty$, equivalent if $p = \infty$.
5.8.5	(1): (a), with $\|\cdot\|_p$ on the S_i (2): (c1) or (c3)	Disparate for all p

[a]The indications (a), (b), (c), (c1), (c3) refer to Sec. 3.1.

The proof is deferred to Sec. 6.

This theorem leads us to the following remarks:

Remark 5.9. Unfortunately it is very common for a pair of choices to be disparate.

Remark 5.10. The pair of choices in case 5.8.2 is a very important one in view of the easy accessibility of the ∞-norm. Moreover, it will become clear from the proof that for some matrices $G^{(2)}$ is indeed a very small subset of $G^{(1)}$.

Remark 5.11. If a pair of choices for $\|\cdot\|'$ is disparate for Theorem 4.1(a), then nevertheless one of these choices may be stronger than the other in practical applications, in view of the information which is available about the matrix A. If, for example, only the largest modulus α_{ij} of the elements of A_{ij} is known for all $i \neq j$ and we look at case 5.8.5 for $p = 1$, then we could do hardly better than to estimate $\|A_{ij}\|_1 \leqslant n_i \alpha_{ij}$, and then have to use all X with $\|X\|_1 \leqslant \sum_{j, \neq i} n_i \alpha_{ij}$ for choice (1), leading to a superset $\tilde{G}^{(1)}$ of $G^{(1)}$. On the other hand, using Sec. 3.1, (c3), for choice (2), we may say that $\|X\|' \leqslant \sum_{j, \neq i} \|A_{ij}\|'$ implies $\|X\|_1 \leqslant n_i \|X\|' \leqslant \sum_{j, \neq i} n_i \alpha_{ij}$. Hence $G^{(2)} \subset \tilde{G}^{(1)}$ for all A, and on account of the disparateness there certainly exists an A such that $G^{(2)} \neq \tilde{G}^{(1)}$. The same applies if we take case 5.8.5 with $p = \infty$ provided that all n_i are equal.

6. PROOF OF THEOREM 5.8

The proof uses a few lemmata, of which 6.1 and 6.6 may have some interest on their own account, whereas 6.3 and 6.4 are trivial but just serve to facilitate the formulation of the proof later on.

LEMMA 6.1. *If $\|v\|^{(2)}/\|v\|^{(1)}$ does not have the same value for all $v\in S_1\cup S_2$, then there exist $P\in L_{12}$ and $Q\in L_{21}$ such that either $\|P\|^{(1)}<\|P\|^{(2)}$ and $\|Q\|^{(1)}>\|Q\|^{(2)}$ or the other way around.*

Proof. There exist $v_1\in S_1$ and $v_2\in S_2$ such that

$$\frac{\|v_1\|^{(2)}}{\|v_1\|^{(1)}}\neq\frac{\|v_2\|^{(2)}}{\|v_2\|^{(1)}}.$$

Suppose

$$\frac{\|v_1\|^{(2)}}{\|v_1\|^{(1)}}>\frac{\|v_2\|^{(2)}}{\|v_2\|^{(1)}}.$$

Then

$$\frac{\|v_1\|^{(1)}}{\|v_2\|^{(1)}}<\frac{\|v_1\|^{(2)}}{\|v_2\|^{(2)}}.$$

Now take $P=v_1v_2^{D_1}$, $Q=v_2v_1^{D_2}$, where D_1 and D_2 refer to the dual with respect to $\|\cdot\|^{(1)}$ and $\|\cdot\|^{(2)}$, respectively. Then

$$\|P\|^{(1)}=\frac{\|Pv_2\|^{(1)}}{\|v_2\|^{(1)}}=\frac{\|v_1\|^{(1)}}{\|v_2\|^{(1)}}<\frac{\|v_1\|^{(2)}}{\|v_2\|^{(2)}}=\frac{\|Pv_2\|^{(2)}}{\|v_2\|^{(2)}}\leqslant\|P\|^{(2)},$$

and likewise $\|Q\|^{(1)}>\|Q\|^{(2)}$. ■

REMARK 6.2. We note about this lemma:

(a) It allows $\|\cdot\|^{(1)}$ and $\|\cdot\|^{(2)}$ to be proportional on S_1 and also to be proportional on S_2 provided that the proportionality factors are different.

(b) It applies to infinite-dimensional normed spaces as well.

(c) It contains as a special case the result: If on a space S there are two nonproportional norms $\|\cdot\|^{(1)}$ and $\|\cdot\|^{(2)}$, then it is impossible that

$\|P\|^{(1)} \leqslant \|P\|^{(2)}$ holds for all linear mappings P from S into itself. This result was probably first proved by Bonsall (cf. [3, Theorem 8]). It was also obtained by Schneider and Strang as a special case of a more general result (cf. [10, Theorem 1]), and by Ljublič (cf. [8]).

LEMMA 6.3. *In order that a choice* (1) *for* $\|\cdot\|'$ *shall be stronger for Theorem* 4.1(*a*), *than a choice* (2), *the combined conditions* (*a*) *and* (*b*) *are necessary and sufficient*:

(a) $G_i^{(1)} \subset G_i^{(2)}$ *for all matrices A and all i;*
(b) *There exists a matrix A such that* $G_i^{(1)} \neq G_i^{(2)}$ *for some i.*

LEMMA 6.4. *In order that the choices* (1) *and* (2) *for* $\|\cdot\|'$ *shall be disparate for Theorem* 4.1(a), *the combined conditions* (a) *and* (b) *are necessary and sufficient*:

(a) *There exists a matrix A such that* $G_i^{(1)} \not\supset G_i^{(2)}$ *for some i.*
(b) *There exists a matrix A such that* $G_i^{(1)} \not\subset G_i^{(2)}$ *for some i.*

Proofs. For any $j \neq i$ replace A_{jj} by λI, λ an eigenvalue of A_{ii}, and replace A_{jk} by 0 for $k \neq j$, $j \neq i$. ■

DEFINITION 6.5. A matrix A is of type $\alpha_{ij}(S,T)$ if $A_{ii}=S$, $A_{ij}=T$, $A_{ik}=0$ for $k \neq i,j$.

LEMMA 6.6. *For any i and j let T denote an arbitrary* $n_i \times n_j$ *matrix, and define* $n_i \times n_j$ *matrices*

$$E=\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}, \qquad V=\begin{bmatrix} \varepsilon & 0 & \cdots & 0 \\ & & & \\ & & 0 & \end{bmatrix},$$

$$W=\begin{bmatrix} \varepsilon & \cdots & \varepsilon \\ \vdots & & \vdots \\ \varepsilon & \cdots & \varepsilon \end{bmatrix}, \qquad Z=\begin{bmatrix} 1 & 0 & \cdots & 0 \\ & & & \\ & & 0 & \end{bmatrix}$$

(where in E the entire main diagonal is filled with ones), and $n_i \times n_i$ *matrices*

$$U=\begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}, \qquad Y=\begin{bmatrix} 0 & & & 1 \\ & & 1 & \\ & ⋰ & & \\ 1 & & & 0 \end{bmatrix}.$$

Then the results in Table 2 hold for G_i of any matrix A of the given type.

In this table and later on p' and r' are defined by

$$\frac{1}{p'}+\frac{1}{p}=1, \qquad \frac{1}{r'}+\frac{1}{r}=1.$$

Furthermore, $B(0,R)$ denotes a disc in $\mathbb{C}$ with center 0 and radius R. The notation $G_i \approx B(0,R)$ for $\varepsilon \to 0$ indicates that $B(0,R_1) \subset G_i \subset B(0,R_2)$ with $R_1/R \to 1$ and $R_2/R \to 1$ for $\varepsilon \to 0$.

TABLE 2

Case	Type	$\|\cdot\|'$ according to[a]	G_i satisfies
6.6.1	$\alpha_{ij}(0,T)$	(a)	$G_i = B(0,R)$ with $R = \|T\|$
6.6.2	$\alpha_{ij}(0,T)$	(b), with $\|\cdot\| = \|\cdot\|_\infty$	$G_i = B(0,R)$ with $R = \|T\|_\infty$
6.6.3	$\alpha_{ij}(0,E)$	(b), with $\|\cdot\| = \|\cdot\|_p$	$G_i = B(0,R)$ with $R = n_i^{1/p}$ if $n_i \leqslant n_j$
6.6.4	$\alpha_{ij}(0,E)$	(c), with $\|\cdot\| = \|\cdot\|_r$ on all S_j, $\|\cdot\|'' = \|\cdot\|_p$	$G_i = B(0,R)$ with $R = \max(n_i^{1/p}, n_i^{1/r})$ if $n_i \leqslant n_j$
6.6.5	$\alpha_{ij}(U,V)$	(a)	$G_i \approx B(0,R)$ with $R = \lvert c\varepsilon\rvert^{1/n_i}$ for $\varepsilon \to 0$, c depending on $\|\cdot\|$
6.6.6	$\alpha_{ij}(U,V)$	(b), with $\|\cdot\| = \|\cdot\|_p$	$G_i = B(0,R)$ with $R = \lvert\varepsilon\rvert$
6.6.7	$\alpha_{ij}(U,W)$	(a), with $\|\cdot\| = \|\cdot\|_p$	$G_i \approx B(0,R)$ with $R = \lvert n_{ij}\varepsilon\rvert^{1/n_i}$ for $\varepsilon \to 0$, $n_{ij} = n_i^{1/p} n_j^{1/p'}$
6.6.8	$\alpha_{ij}(U,W)$	(c3)	$G_i \approx B(0,R)$ with $R = \lvert\varepsilon\rvert^{1/n_i}$ for $\varepsilon \to 0$
6.6.9	$\alpha_{ij}(Y,Z)$	(a), with $\|\cdot\| = \|\cdot\|_p$	G_i contains -2 if $n_i \geqslant 2$
6.6.10	$\alpha_{ij}(Y,Z)$	(c1)	G_i does not contain -2 if $n_i \geqslant 2$

[a]The indications (a), (b), (c), (c1), (c3) refer to Sec. 3.1.

Proof. In this proof ω will denote any scalar with $|\omega| \leqslant 1$; ρ will denote the spectral radius.

Case 6.6.1. $\|X\| \leqslant \|T\| \Rightarrow \rho(X) \leqslant \|T\|$. Taking $X = \omega\|T\|I$ yields all λ with $|\lambda| \leqslant \|T\|$.

Case 6.6.2. $\|X\|' \leqslant \|T\|' \Rightarrow \rho(X) \leqslant \|T\|_\infty$. Let the kth row of T have the largest sum of moduli $(=\|T\|_\infty)$. Then take the (k,k) element of X equal to $\omega\|T\|_\infty$, and all other elements of X equal to zero.

Case 6.6.3. $\|X\|' \leqslant \|E\|' \Rightarrow \|X\|_p \leqslant n_i^{1/p}$, implying $\rho(X) \leqslant n_i^{1/p}$. Taking $X = \omega v y^T$ with $v = (1,\dots,1)^T$, $y^T = v^D/\|v^D\|_{p'}$, we have $\|X\|' \leqslant \|E\|'$ and $Xv = \omega v v^D v/\|v^D\| = \omega v\|v\| = \omega n_i^{1/p} v$.

Case 6.6.4. $\|E\|' = n_i^{1/p}$. We need only consider X of rank $\leqslant 1$ (cf. Remark 4.2); hence $X = vy^T$, $\|X\|' = \|v\|_p \|y\|_{r'}$. The only possible nonzero eigenvalue of X is $y^T v$ on account of $Xv = vy^T v$. Since G_i obviously is a circle whose radius is the largest modulus the eigenvalues of any X may have, we want to know the quantity $z = \max\{|y^T v| \mid \|v\|_p \|y\|_{r'} = n_i^{1/p}\}$. Since $|y^T v| \leqslant \|v\|_p \|y\|_{p'}$ and since for any given y there is a v with a given value of $\|v\|_p$ such that $|y^T v| = \|v\|_p \|y\|_{p'}$, we have $z = \max\{\|v\|_p \|y\|_{p'} \mid \|v\|_p \|y\|_{r'} = n_i^{1/p}\} = n_i^{1/p} \max\{\|y\|_{p'} \mid \|y\|_{r'} = 1\}$. The latter max equals 1 if $p' \geqslant r'$ and equals $n_i^{1/p' - 1/r'} = n_i^{1/r - 1/p}$ if $p' \leqslant r'$.

Case 6.6.5. If X has $c\varepsilon$ in the bottom left-hand corner and $O(\varepsilon)$ anywhere else, then $U + X$ has eigenvalues $(c\varepsilon)^{1/n_i}[1 + o(1)]$. If $\tilde{X}$ denotes a matrix with $\|\tilde{X}\| = \|V\|$ whose bottom left-hand element has the largest possible modulus, then this modulus equals $|c\varepsilon|$ for some constant c which is independent of ε. Now consider all $X = \omega\tilde{X}$.

Case 6.6.6 is trivial.

Case 6.6.7. $W = \varepsilon v y^T$, where v and y are vectors with coordinates 1 only. Consequently, we have $\|W\|_p = |\varepsilon| \|v\|_p \|y\|_{p'} = |\varepsilon| n_i^{1/p} n_j^{1/p'} = |\varepsilon| n_{ij}$. Thus, all X with $\|X\|_p \leqslant |\varepsilon| n_{ij}$ are to be considered. Hence, all elements of X are at most $|\varepsilon| n_{ij}$ in modulus. Now take the matrix X with $\omega\varepsilon n_{ij}$ in the bottom left-hand corner and zeros elsewhere.

Case 6.6.8. All elements of X are at most $|\varepsilon|$ in modulus. Now see the proof of case 6.6.7.

Case 6.6.9. $\|Y\|_p = \|Z\|_p = 1$. Take $X = Y$; then -2 is an eigenvalue of $Y + X$.

Case 6.6.10. Suppose that -2 is an eigenvalue of $Y + X$, $X = (x_{rs})$ with $\Sigma|x_{rs}| < 1$, and let (ξ_s) be a corresponding eigenvector. Assume $|\xi_r| = \max_s |\xi_s| = 1$, and let for the moment $r \neq (n_i + 1)/2$. Then $-2\xi_r = \xi_{n+1-r} + \theta_1$, $-2\xi_{n+1-r} = \xi_r + \theta_2$ and $|\theta_1| + |\theta_2| \leqslant 1$. This implies $4\xi_r + 2\theta_1 = \xi_r + \theta_2$, and hence $3\xi_r = \theta_2 - 2\theta_1$, which is impossible, since $|\theta_2 - 2\theta_1| \leqslant 2$. If $r = (n_i + 1)/2$, then $-2\xi_r = \xi_r + \theta$ with $|\theta| \leqslant 1$, which is impossible again. ∎

We now come to the proof of Theorem 5.8.

Proof of case 5.8.1. If $\|v\|^{(2)}/\|v\|^{(1)}$ has the same value for all $v \in S_1 \cup S_2 \cup \cdots \cup S_N$, then $\|B\|^{(1)} = \|B\|^{(2)}$ for all matrices $B \in \cup L_{ij}$, and hence $G^{(1)} = G^{(2)}$ for all A.

If $\|v\|^{(2)}/\|v\|^{(1)}$ assumes different values on $S_1 \cup S_2 \cup \cdots \cup S_N$, we may as well suppose that it assumes different values on $S_1 \cup S_2$. Let P and Q be as in Lemma 6.1. On account of case 6.6.1, Lemma 6.4 is applicable with a matrix with first block-row $(0 \;\; P \;\; 0 \;\; \cdots \;\; 0)$ in (a) and a matrix with second block-row $(Q \;\; 0 \;\; \cdots \;\; 0)$ in (b) or the other way around. ■

Proof of case 5.8.2. From the obvious implication

$$\|X\|'^{(2)} \leqslant \sum_{j, \neq i} \|A_{ij}\|'^{(2)} \quad \Rightarrow \quad \|X\|_\infty \leqslant \sum_{j, \neq i} \|A_{ij}\|_\infty$$

we have $G_i^{(2)} \subset G_i^{(1)}$ for all i. In view of cases 6.6.5 and 6.6.6, Lemma 6.3 is applicable for ε small enough, and then indeed $G^{(2)}$ may be a very small subset of $G^{(1)}$ (as stated in Remark 5.10). ■

Proof of case 5.8.3. If $q = \infty$, we should have $p \neq \infty$, and on account of case 6.6.2 we may apply Lemma 6.4 in the same way as we did in the proof of case 5.8.1.

Now consider the situation $q \neq \infty$. Cases 6.6.5 and 6.6.6 give $G_i^{(1)} \not\subset G_i^{(2)}$ for ε small and $n_i > 1$. Case 6.6.1 with $T = E$ (note that $\|E\|_p = 1$) and case 6.6.3 give $G_i^{(1)} \not\supset G_i^{(2)}$ if $n_i > 1$ and $n_i \leqslant n_j$ for some $j \neq i$ (which can always be arranged, since the partitioning is strongly submaximal). Now apply Lemma 6.4. ■

Proof of case 5.8.4. Restricting ourselves to matrices X of rank 1 (cf. Remark 4.2), we have $\|X\|'^{(2)} = \|X\|^{(1)}$, whereas $\|A_{ij}\|'^{(2)} \geqslant \|A_{ij}\|^{(1)}$. Hence $G_i^{(1)} \subset G_i^{(2)}$ for all i. Case 6.6.1 with $T = E$ and case 6.6.4 with $r = p$ then give $G_i^{(1)} \neq G_i^{(2)}$ if $p \neq \infty$ and $n_i > 1$ and $n_i \leqslant n_j$ for some $j \neq i$. Now apply Lemma 6.3. ■

Proof of case 5.8.5. Case 6.6.4 gives $G_i^{(2)} = \{\lambda \mid |\lambda| \leqslant n_i\}$. Case 6.6.1 with $T = E$ then gives $G_i^{(1)} \not\supset G_i^{(2)}$ if $n_i > 1$ and $n_i \leqslant n_j$ for some $j \neq i$.

If choice (2) is according to Sec. 3, (c1), then cases 6.6.9 and 6.6.10 give $G_i^{(1)} \not\subset G_i^{(2)}$. If choice (2) is according to Sec. 3, (c3), then the same is obtained from cases 6.6.7 and 6.6.8 if $n_i, n_j > 1$. ■

The author recalls with pleasure the discussions about Sec. 3 with his student B. van Putten, who also found Lemma 3.2(a).

Professor H. Schneider (Madison) kindly mentioned to the author the references [3] and [8], given in Remark 6.2(c).

REFERENCES

1 F. L. Bauer, On the field of values subordinate to a norm, *Numer. Math.* 4:103–113 (1962).
2 F. L. Bauer, J. Stoer, and C. Witzgall, Absolute and monotonic norms, *Numer. Math.* 3:257–264 (1961).
3 F. F. Bonsall, A minimal property of the norm in some Banach algebras, *J. London Math. Soc.* 29:156–164 (1954).
4 E. Deutsch, On vectorial norms and pseudonorms, *Proc. Amer. Math. Soc.* 28:18–24 (1971).
5 D. G. Feingold and R. S. Varga, Block diagonally dominant matrices and generalizations of the Gershgorin circle theorem, *Pacific J. Math.* 12:1241–1250 (1962).
6 T. Kato, *Perturbation Theory for Linear Operators*, Springer, Berlin, 1966.
7 Z. V. Kovarik and D. D. Olesky, Sharpness of generalized Gershgorin disks, *Linear Algebra and Appl.* 8:477–482 (1974).
8 Ju. I. Ljubič, On operator norms of matrices, *Uspehi Mat. Nauk* 18(4):161–164 (1963).
9 W. A. J. Luxemburg and A. C. Zaanen, *Riesz Spaces*, North-Holland, Amsterdam, 1971.
10 H. Schneider and W. G. Strang, Comparison theorems for supremum norms, *Numer. Math.* 4:15–20 (1962).
11 J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965.

Received 11 April 1977; revised 23 October 1978