The Numerical Radius and Spectral Matrices

MOSHE GOLDBERG, EITAN TADMOR and GIDEON ZWAS

Department of Mathematical Sciences, Tel-Aviv University, Israel

(Received April 11, 1974)

In this paper we investigate spectral matrices, i.e., matrices with equal spectral and numerical radii. Various characterizations and properties of these matrices are given.

1. INTRODUCTION

Let $A$ be an $n$-square complex matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, and let

$$\rho(A) = \max_{1 \leq j \leq n} |\lambda_j|$$

(1.1)

be the spectral radius of $A$. Let

$$r(A) = \max_{|x|=1} |(Ax, x)|$$

(1.2)

be the numerical radius of $A$, and

$$\|A\| = \max_{|x|=1} |Ax|$$

(1.3)

the spectral norm of $A$. Here $(x, y)$ is the unitary inner product of the vectors $x$ and $y$, and $|x| = (x, x)^{\frac{1}{2}}$.

It is well known that

$$\rho(A) \leq r(A) \leq \|A\| \leq 2r(A).$$

(1.4)

In this paper we investigate matrices for which

$$\rho(A) = r(A).$$

(1.5)

Following Halmos ([3] p. 115), we call matrices which satisfy (1.5), spectral matrices. Our main purpose is to characterize the spectral matrices and find some of their properties.

Before turning to some new results, we recall a few known results which we shall use later on. It is known that
M. GOLDBERG, E. TADMOR AND G. ZWAS

(1.6) $r(A) = 0$ if and only if $A = 0$.

(1.7) $r(\alpha A) = |\alpha| r(A)$, for every scalar $\alpha$.

(1.8) $r(A + B) \leq r(A) + r(B)$.

However, the numerical radius is not a matrix-norm, since in general it is not true that $r(AB) \leq r(A) \cdot r(B)$ even if $A$ and $B$ are powers of the same matrix [8]. On the other hand, we always have the Halmos inequality

(1.9) $r(A^k) \leq r^k(A)$, $k = 1, 2, 3, \ldots$

This power inequality was conjectured by Halmos and proved by Berger. The proof was simplified by Pearcy [8]. Generalizations of the power inequality were given by Kato [4], and by Berger and Stampfli [1]. It is also known that

(1.10) $r(A_1 \oplus \cdots \oplus A_m) = \max_{1 \leq j \leq m} r(A_j)$.

Another concept associated with the numerical radius of a matrix is the numerical range $F(A)$, defined by

(1.11) $F(A) = \{(Ax, x), |x| = 1\}$.

The numerical range, known also as the field of values of $A$, is a convex set in the complex plane. If $U$ is a unitary transformation, then

(1.12) $F(U^*AU) = F(A), \quad r(U^*AU) = r(A)$.

If $M$ is any principle sub-matrix of $A$, then

(1.13) $F(M) \subseteq F(A), \quad r(M) \leq r(A)$.

For a $2 \times 2$ matrix it is known that $F(A)$ is an ellipse whose foci are the eigenvalues $\lambda_1$ and $\lambda_2$ of $A$. In particular, if $A$ is of the form

(1.14) $A = \begin{pmatrix} \lambda_1 & 0 \\ \sigma & \lambda_2 \end{pmatrix}$,

then $|\sigma|/2$ is the semi-minor axis of the ellipse $F(A)$. We shall refer to this result as the Elliptic Range Theorem (see for example [6]).

Most of the above mentioned results can be found in [3]. A survey of properties of the numerical range and the numerical radius, some of which were proven by Parker, is given in [7].

The investigation of spectral matrices is motivated by stability problems related to finite difference schemes, where the uniform boundedness of $\|A^k\|, k = 1, 2, 3, \ldots$ plays a central role. In general we have

(1.15) $\rho(A) \leq 1$ is a necessary condition for uniform boundedness of the powers of $A$. However, if $A$ is spectral, this condition is sufficient as well, and implies that $\|A^k\| \leq 2$ for all $k$. Such an idea was applied first by Lax and Wendroff [5].
2. STRUCTURE-CHARACTERIZATION OF SPECTRAL MATRICES

Before we start characterizing the class of spectral matrices we note that this class is wider than the class of normal matrices. For, if $A$ is normal, then it is unitarily similar to a diagonal matrix, and by (1.10) and (1.12)

$$r(A) = \max |\lambda_j| = \rho(A),$$  

(2.1)

so $A$ is spectral. However, not every spectral matrix is normal. To see that, take the non-normal matrix

$$A = I \oplus B, \quad I = I_{n-2}, \quad B = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}, \quad n \geq 3.$$  

(2.2)

We have $\rho(A) = 1$, where by (1.10) and the Elliptic Range Theorem $r(A) = 1$ too. Thus for $n \geq 3$ the class of normal matrices is a proper subclass of the class of spectral matrices.

The example just given shows that the spectrality of a direct sum does not imply the spectrality of each of the summands. On the other hand, it is clear that a direct sum of spectral matrices is spectral.

Let us now order the eigenvalues of an arbitrary $n$-square matrix $A$ such that

$$\rho(A) = |\lambda_1| = \cdots = |\lambda_s| > |\lambda_{s+1}| \geq \cdots \geq |\lambda_n|,$$  

(2.3)

where $s = s(A)$ is the number of eigenvalues of $A$ on the spectral circle $|z| = \rho(A)$.

**Theorem 1** The matrix $A$ is spectral if and only if $A$ is unitarily similar to a triangular matrix of the form

$$\begin{pmatrix} \Lambda & 0 \\ 0 & B \end{pmatrix},$$  

(2.4a)

where

$$\Lambda = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_s \end{pmatrix}, \quad B = \begin{pmatrix} \lambda_{s+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ (B_{ij}) & \cdots & \lambda_n \end{pmatrix},$$  

(2.4b)

and where

$$r(B) \leq \rho(A).$$  

(2.5)

**Proof** It is known that $A$ is unitarily similar to a triangular matrix $T$, where the eigenvalues $\lambda_i$, are ordered along its diagonal as in (2.3). Since $\rho(A) = \rho(T)$ and $r(A) = r(T)$, we may assume that $A$ is already triangular.

Suppose that $A$ is spectral and take $j$ and $k$ with $1 \leq j \leq s$ and $j \leq k \leq n$. By (1.13)

$$\rho(A) = r(A) \geq r\begin{pmatrix} \lambda_j & 0 \\ \lambda_k & \lambda_j \end{pmatrix} \geq r(\lambda_j) = |\lambda_j| = \rho(A),$$  

(2.6)
and therefore
\[ r\begin{pmatrix} \lambda_j & 0 \\ a_{ij} & \lambda_i \end{pmatrix} = |\lambda_j|. \tag{2.7} \]

Using the Elliptic Range Theorem, it is clear that (2.7) is satisfied if and only if \( a_{ij} = 0 \). Thus \( A = \Lambda + B \) as in (2.4). Now by (1.10)
\[ \rho(A) = r(A) = \max\{r(A), r(B)\}, \tag{2.8} \]
hence (2.5) holds.

If (2.4) and (2.5) are satisfied, then by (1.10)
\[ r(A) = \max\{r(\Lambda), r(B)\}. \tag{2.9} \]

Since \( \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \) we have \( \rho(\Lambda) = r(\Lambda) \), and by (2.5) \( \rho(A) = r(A) \). Thus \( A \) is spectral.

**COROLLARY 1** If \( s = s(A) \geq n - 1 \), then \( A \) is spectral if and only if \( A \) is normal.

**Proof** Normality implies spectrality. If \( A \) is spectral and \( s \geq n - 1 \), then by Theorem 1 it is unitarily similar to a diagonal matrix and \( A \) is normal.

In particular we obtain the following:

**COROLLARY 2** If \( n = 2 \), then \( A \) is a spectral if and only if it is normal.

Corollary 2 follows also directly from the Elliptic Range Theorem. For if \( A \) is not normal then, without restriction, it is of the form (1.14) with \( \sigma \neq 0 \). Therefore, the ellipse \( F(A) \) includes points \( z \) with \( r(A) \geq |z| \max\{|\lambda_1|, |\lambda_2|\} = \rho(A) \), and \( A \) is not spectral.

For \( A = [a_{ij}] \), denote \( A^* = [\bar{a}_{ij}] \). By the definition of the numerical radius we find that
\[ r(A^*) \geq r(A). \tag{2.10} \]

Therefore, Theorem 1 yields the following:

**COROLLARY 3** If \( s(A) = n - 2 \), then a sufficient condition for \( A \) to be spectral is that \( A \) is unitarily similar to a matrix of the form (2.4), where
\[ B = \begin{pmatrix} \lambda_{n-1} & 0 \\ \beta & \lambda_n \end{pmatrix}, \tag{2.11} \]
and
\[ |\beta| \leq 2(\rho(A) - |\lambda_{n-1}|)^g[\rho(A) - |\lambda_n|]^g. \tag{2.12} \]

**Proof** In order to satisfy (2.5), it is sufficient, by (2.10), to require that
\[ r\begin{pmatrix} |\lambda_{n-1}| & 0 \\ |\beta| & |\lambda_n| \end{pmatrix} \leq r(A). \tag{2.13} \]

By the Elliptic Range Theorem, (2.13) means that the circle \( |z| = \rho(A) \) contains the ellipse with the non-negative foci \( |\lambda_{n-1}|, |\lambda_n| \), and minor axis \( |\beta| \). This clearly holds if and only if (2.12) is satisfied.
In the case $s(A) = n - 2$ we remark that if $\arg(\lambda_{n-1}) = \arg(\lambda_n)$, then (2.12) is also necessary for the spectrality of $A$. However, for general $\lambda_{n-1}$ and $\lambda_n$, finding a condition on the size of $|\beta|$ in (2.11) which is necessary as well as sufficient for the spectrality of $A$, involves the solution of a general quadric equation.

3. CRITICAL POWER CHARACTERIZATION

We start with the following theorem.

**Theorem 2** The matrix $A$ is spectral if and only if

$$r(A^k) = r^k(A), \quad k = 1, 2, 3, \ldots$$

(3.1)

**Proof** If $A$ is spectral, then by (1.9)

$$r(A^k) \leq r(A^k) \leq r^k(A) = \rho^k(A) = \rho(A^k), \quad k = 1, 2, 3, \ldots,$$

and (3.1) holds. Conversely, we know that

$$\|A^k\|^{1/k} \to \rho(A).$$

(3.3)

Therefore, if (3.1) is satisfied, then using (1.4) we have

$$\rho(A) = \rho^{1/k}(A^k) \leq r^{1/k}(A^k) = r(A) \leq \|A^k\|^{1/k} \to \rho(A),$$

(3.4)

and the theorem follows.

Theorem 2 leads to the following conclusion.

**Corollary 4** If $A$ is spectral, then any power of $A$ is spectral.

**Proof** Consider $A^n$. By Theorem 2 we have, for all $k$,

$$r((A^n)^k) = r(A^{nk}) = r^{nk}(A) = (r^n(A))^k = r^k(A^n).$$

(3.5)

Hence, using Theorem 2 once again, $A^n$ is spectral.

Note that if a power of $A$ is spectral, then $A$ is not necessarily spectral. To see this, take

$$A = I_{n-2} \oplus B, \quad B = \begin{pmatrix} 0 & 0 \\ \alpha & 0 \end{pmatrix}, \quad |\alpha| > 2, \quad n \geq 3.$$  

(3.6)

By (1.10) and the Elliptic Range Theorem, $A$ is not spectral. On the other hand, all the powers, $A^m = I_{n-2} \oplus O_{2 \times 2}, \quad m \geq 2$, are normal and hence spectral.

Equation (3.1) of Theorem 2 provides infinitely many conditions, whose simultaneous satisfaction is equivalent to spectrality. However, the finite nature of a matrix leads us to conjecture the existence of a finite integer $k_0 = k_0(A)$ such that the validity of (3.1) for $k = k_0$ only, is sufficient as well as necessary for $A$ to be spectral. The remainder of this section deals with this question.
Let \( m \) be a positive integer and let \( \omega_j = e^{2\pi i/jm}, \) \( 1 \leq j \leq m, \) be the \( m \)th roots of unity. The following polynomial identities are well known:

\[
1 - z^m = \prod_{k=1}^{m} (1 - \omega_k z); \tag{3.7}
\]

\[
m = \sum_{j=1}^{m} \prod_{k=1 \atop k \neq j}^{m} (1 - \omega_k z). \tag{3.8}
\]

Using these identities, which hold also when \( z \) is replaced by any square matrix \( B, \) Pearcy [8] proved the following lemma.

**Lemma (Pearcy)** Let \( B \) be a square matrix, \( m \) a positive integer, and \( x \) a unit vector. Then

\[
1 - (B^m x, x) = \frac{1}{m} \left[ \sum_{j=1}^{m} |x_j|^2 - \sum_{j=1}^{m} \omega_j (B x_j, x_j) \right], \tag{3.9}
\]

where the vectors \( x_j \) are defined by

\[
x_j = \left[ \prod_{k=1 \atop k \neq j}^{m} (1 - \omega_k) \right] x, \quad 1 \leq j \leq m. \tag{3.10}
\]

By the known identity

\[
\prod_{k=1 \atop k \neq j}^{m} (1 - \omega_k z) = \sum_{k=0}^{m-1} \omega_j^k z^k, \tag{3.11}
\]

the vectors \( x_j \) in Pearcy's Lemma, may be rewritten in the form

\[
x_j = \sum_{k=0}^{m-1} \omega_j^k B^k x, \quad 1 \leq j \leq m. \tag{3.12}
\]

From (3.7) and (3.11) we obtain \((1 - B^m)x = (1 - \omega_jB)x_j; \) thus

\[
\omega_j B x_j = x_j + B^m x - x, \quad 1 \leq j \leq m. \tag{3.13}
\]

Now let \( A \neq 0 \) be an \( n \)-square matrix, \( m \) a positive integer, and \( x = x(m) \) a unit vector such that

\[
|(A^m x, x)| = r(A^m) \tag{3.14}
\]

Define the matrix

\[
B = \frac{e^{i\theta}}{r(A)} A, \tag{3.15a}
\]

where

\[
\theta = \theta(m) = -\frac{1}{m} \arg(A^m x, x). \tag{3.15b}
\]

Note that \( r(B) = 1; \) moreover \( B = B(m) \) is spectral if and only if \( A \) is spectral. We are now in a position to prove the following lemma.
SPECTRAL MATRICES

Lemma 1 Let \(A \neq 0\) be a square matrix, \(x = x(m)\) a unit vector satisfying (3.14), \(B\) as defined by (3.15), and \(x_1, \ldots, x_m\) the vectors in (3.12). Then

\[ r(A^m) = r^m(A) \tag{3.16} \]

if and only if

\[ (B^m x - x, x_j) = 0, \quad 1 \leq j \leq m. \tag{3.17} \]

Proof By (3.14)

\[ (B^m x, x) = \frac{e^{-m \theta}}{r^m(A)}(A^m x, x) = \frac{|(A^m x, x)|}{r^m(A)} = \frac{r(A^m)}{r^m(A)}. \tag{3.18} \]

Therefore, Pearcy's Lemma implies that

\[ \frac{1}{m} \left[ \sum_{j=1}^{m} |x_j|^2 - \sum_{j=1}^{m} \omega_j(Bx_j, x_j) \right] = 1 - \frac{r(A^m)}{r^m(A)}. \tag{3.19} \]

Now, if \(r(A^m) = r^m(A)\), then by (3.19)

\[ \sum_{j=1}^{m} \omega_j(Bx_j, x_j) = \sum_{j=1}^{m} |x_j|^2. \tag{3.20} \]

Hence the left hand side of (3.20) is real and non-negative. Since for all \(x_j\)

\[ |(Bx_j, x_j)| \leq r(B)|x_j|^2 = |x_j|^2, \tag{3.21} \]

we find that

\[ \sum_{j=1}^{m} |x_j|^2 = \sum_{j=1}^{m} \omega_j(Bx_j, x_j) \leq \sum_{j=1}^{m} |(Bx_j, x_j)| \leq \sum_{j=1}^{m} |x_j|^2. \tag{3.22} \]

That is,

\[ \sum_{j=1}^{m} \omega_j(Bx_j, x_j) = \sum_{j=1}^{m} |(Bx_j, x_j)| = \sum_{j=1}^{m} |x_j|^2. \tag{3.23} \]

From the left equality in (3.23) we have \(\omega_j(Bx_j, x_j) \geq 0\); from the right equality and (3.21), \(|(Bx_j, x_j)| = |x_j|^2\). Therefore

\[ \omega_j(Bx_j, x_j) = |x_j|^2, \quad 1 \leq j \leq m. \tag{3.24} \]

Now, substituting \(\omega_j Bx_j\) from (3.13) into (3.24), we find that

\[ |x_j|^2 = (x_j + B^m x - x, x_j) = |x_j|^2 + (B^m x - x, x_j), \quad 1 \leq j \leq m, \tag{3.25} \]

and (3.17) follows.

Conversely, if (3.17) holds, then by (3.13),

\[ (\omega_j Bx_j, x_j) = (x_j + B^m x - x, x_j) = |x_j|^2 + (B^m x - x, x_j) = |x_j|^2, \quad 1 \leq j \leq m. \tag{3.26} \]

Hence (3.20) is satisfied, and by (3.19) \(r(A^m) = r^m(A)\).

Lemma 1 enables us to prove the following theorem.
Theorem 3. Let A be an n-square matrix with minimal polynomial of degree p, and m an integer with \( m \geq p \). Then A is spectral if and only if \( r(A^m) = r^m(A) \).

Proof. By Theorem 2, the spectrality of A implies \( r(A^m) = r^m(A) \). If \( A = 0 \), then A is obviously spectral.

Assume \( A \neq 0 \) and let \( x = x(m) \), \( B = B(m) \), and \( x_1, \ldots, x_m \) be as in Lemma 1. By (3.12) and (3.17) we have

\[
\sum_{k=0}^{m-1} \bar{a}_j(B^kx - x, B^kx) = (B^m x - x, \sum_{k=0}^{m-1} \bar{a}_j B^kx) = (B^m x - x, x_j) = 0. \tag{3.27}
\]

We conclude that the polynomial \( P(x) = \sum_{k=0}^{m-1} (B^kx - x, B^kx)x^k \) has degree \( m - 1 \) at most, has \( m \) roots, \( \bar{a}_1, \ldots, \bar{a}_m \). Hence all its coefficients vanish, i.e.,

\[
(B^m x - x, B^k x) = 0, \quad k = 0, \ldots, m - 1. \tag{3.28}
\]

Clearly, the minimal polynomials of A and B are of the same degree, since \( m \geq p \), there exists scalars \( \alpha_j \), \( 0 \leq j \leq m - 1 \), such that

\[
B^m = \sum_{j=0}^{m-1} \alpha_j B^j. \tag{3.29}
\]

Therefore by (3.28) and (3.29)

\[
(B^m x - x, B^m x) = \sum_{j=0}^{m-1} \bar{a}_j (B^m x - x, B^j x) = 0. \tag{3.30}
\]

By (3.30) and by (3.28) with \( k = 0 \), we obtain

\[
(x, B^m x) = 1, \quad (x, x) = (B^m x, B^m x) = 1, \tag{3.31}
\]

i.e., the inner product of the unit vectors \( x \) and \( B^m x \) is 1. This is true if and only if

\[
B^m x = x. \tag{3.32}
\]

Hence \( \mu = 1 \) is an eigenvalue of \( B^m \), and we have \( \rho(B^m) \geq 1 \). Since \( r(B) = 1 \),

\[
1 \leq \rho(B^m) = \rho^m(B) \leq r^m(B) = 1. \tag{3.33}
\]

Consequently B is spectral and the spectrality of A follows.

Since the degree of the minimal polynomial of an n-square matrix does not exceed \( n \), we may conclude the following.

Corollary 4. An n-square matrix is spectral if and only if \( r(A^m) = r^m(A) \).

At this point it seems natural to ask whether in general an equality of the form \( r(A^m) = r^m(A) \), for some \( m < n \), implies spectrality. In general, the answer is negative even for the case \( m = n - 1 \), as can be seen from the example

\[
A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{3.34}
\]

Clearly \( \rho(A) = 0 \) and it can be verified that \( r(A^2) = r^2(A) = \frac{1}{2} \).
Another result which follows immediately from Theorems 2 and 3 is given below.

**Corollary 5** If \( r(A^m) = r^m(A) \) for some \( m \) with \( m \geq p \), where \( p \) is the degree of the minimal polynomial of \( A \), then \( r(A^k) = r^k(A) \) for all \( k \).

An additional result can be derived from Corollaries 2 and 4.

**Corollary 6** For \( n = 2 \), \( A \) is normal if and only if \( r(A^2) = r^2(A) \).

We remark that Corollary 6 can be obtained directly by geometrical reasoning, using results of A. Brown [2].

Using Theorem 2 and Corollary 4 we prove our next theorem.

**Theorem 4** Let \( A \), with eigenvalues \( \lambda_1, \ldots, \lambda_n \), be unitarily similar to a matrix of the form
\[
Q = \text{diag}(\lambda_1, \ldots, \lambda_n) \oplus C, \quad (C = C_{n-1 \times n-1}),
\]
so that at least one of the eigenvalues \( \lambda_1, \ldots, \lambda_n \) is on the spectral circle \( |z| = \rho(A) \). Let \( m \) be an integer such that \( m > n - 1 \). Then \( A \) is spectral if and only if \( r(A^m) = r^m(A) \).

**Proof** If \( A \) is spectral, then (3.36) holds by Theorem 2. Conversely, suppose \( A \) is unitarily similar to a matrix \( Q \) of the form (3.35), and that (3.36) holds. Since \( Q = U^*AU \), we have \( Q^m = U^*A^mU \), and by (1.12) and (3.36)
\[
r(Q^m) = r^m(Q).
\]
In addition
\[
Q^m = \text{diag}(\lambda_1^m, \ldots, \lambda_n^m) + C^m.
\]
Thus by (3.37) and (1.10)
\[
\max\{\rho^m(Q), r(C^m)\} = r(Q^m) = r^m(Q) = \max\{\rho^m(Q), r^m(C)\}.
\]
If
\[
r^m(Q) = r^m(C) > \rho^m(Q),
\]
then by (3.39) we also have
\[
r(Q^m) = r(C^m) > \rho^m(Q),
\]
from which
\[
r(C^m) = r^m(C).
\]
Since \( C \) is \((n-1)\)-square and \( m > n - 1 \), it follows from Theorem 3 that \( C \) is spectral; hence \( r(C) = \rho(C) \leq \rho(Q) \). This contradicts (3.40), so we must have
\[
r^m(Q) = \rho^m(Q) \geq r^m(C).
\]
This leads to
\[
r(C^k) \leq r^k(C) \leq \rho^k(Q), \quad k = 1, 2, 3, \ldots.
\]
and for \( k = n \) we obtain
\[
r(Q') = \max\{r(Q), r(C')\} = \max\{\rho(Q), r(C)\} = r(Q).
\] (3.45)

By Corollary 4, \( Q \) is spectral. Hence \( A \) is spectral and the theorem follows.

Combining Theorems 1, 2 and 4, we may derive yet another final result.

**Corollary 7** Let \( A \) have eigenvalues \( \lambda_1, \ldots, \lambda_n \), ordered as in (2.3), and let \( m \) be an integer such that \( m > n - s \). Then \( A \) is spectral if and only if \( A \) is unitarily similar to a triangular matrix \( T = \Lambda \oplus B \) of the form (2.4) and \( r(A^m) = r^n(A) \).

The above results tend to show that the numerical radius can be useful in various applications. It is hoped that further research will actually bear this out.

**References**


