A NOTE ON THE FIELD OF VALUES OF NON-NORMAL MATRICES

ILSE C. F. IPSEN*

Abstract. It is shown that the distance from zero of the field of values of a matrix A depends on how large the departure from normality is compared to the distance from the zero of the field of values of the normal part of A. A connection is made to the convergence of Krylov methods for the solution of linear systems.

Key words. 15A60, 65F10, 15A18, 65F15, 15A42

AMS subject classification. field of values, numerical range, numerical radius, departure from normality, eigenvalues, Krylov methods

1. Introduction. The *field of values* (or *numerical range*) of a square matrix A is the set of all complex numbers[1]

$$\mathrm{F}(A) \equiv \{\frac{x^*Ax}{x^*x},\ x \neq 0 \text{ is a vector}\}.$$

The field of values is used in the convergence analysis of iterative methods for the solution of systems of linear equations, such as, for instance, asymptotically stationary k-step methods [2], ADI methods [15], and the Krylov methods GMRES and FOM [16], and Orthomin [5, §2.2]. Take Orthomin(1), for instance. Applied to the linear system $Ax = b$ it produces approximate solutions x_k whose residuals $r_k \equiv b - Ax_k$ have strictly decreasing two-norms if and only if $0 \notin \mathrm{F}(A^*)$ [5, Theorem 2.2.1]. Moreover, let

$$\mathrm{dist}(\mathrm{F}(A), 0) \equiv \min\{|z|,\ z \in \mathrm{F}(A)\}$$

be the *distance from zero* of the field of values of A. Then [5, Theorem 2.2.2] the residual norms produced by Orthomin(1) decrease fast if the distance from zero of $\mathrm{F}(A^*)$ is large,

$$\|r_k\| \leq \sqrt{1 - \frac{\mathrm{dist}(\mathrm{F}(A^*), 0)^2}{\|A\|^2}}\ \|r_{k-1}\|.$$

Although existing work on Krylov methods deals with non-normal matrices, no explicit use appears to be made of the departure from normality. Here we express the numerical radius and the distance from zero of the field of in terms of the departure from normality. In particular we show that the distance from zero of $\mathrm{F}(A)$ depends on how large the departure of A from normality is compared to the distance from zero of the field of values of the normal part of A.

The only other work that connects field of values and departure from normality seems to Henrici's [7]. Henrici used the departure from normality to bound the distance from the convex hull of the eigenvalues to the boundary of the field of values boundary [7, §5].

* Center for Research In Scientific Computation, Department of Mathematics, North Carolina State University, P. O. Box 8205, Raleigh, NC 27695-8205, USA (ipsen@math.ncsu.edu, http://www4.ncsu.edu/~ipsen/info.html). This research was supported in part by NSF grant CCR-9400921.

[1] The superscript * represents the congugate transpose.

1.1. Notation. The norm $\|\cdot\|$ is the Euclidean two-norm, or spectral norm. The identity matrix of order n is $I = (\, e_1 \quad \ldots \quad e_n \,)$ with columns e_i. The transpose of a matrix A is A^T, and the conjugate transpose is A^*. The eigenvalues of a square matrix A are $\lambda_i(A)$ and the numerical radius is

$$\mathrm{r}(A) \equiv \max_{x \neq 0} \frac{|x^* A x|}{|x^* x|}.$$

2. A Single Jordan Block. The numerical radius and distance from zero of the field of values of a single Jordan block are determined.

Let

$$A = \begin{pmatrix} \lambda & \eta & & \\ & \lambda & \ddots & \\ & & \ddots & \eta \\ & & & \lambda \end{pmatrix} \tag{2.1}$$

be a matrix of order n. The two-norm departure of A from normality [7, §1.2] is $|\eta|$. When $\eta = 0$ then $A = \lambda I$ is normal. When $\eta \neq 0$ then A is diagonally similar to a Jordan block, $A = XJX^{-1}$, where

$$J = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}, \qquad X = \begin{pmatrix} 1 & & & \\ & \eta & & \\ & & \ddots & \\ & & & \eta^{n-1} \end{pmatrix},$$

Hence the single eigenvalue λ is maximally defective for $\eta \neq 0$.

Below it is shown that the numerical radius of A and the distance from zero of the field of values of A depend on the departure from normality.

THEOREM 2.1. *Let A be as in (2.1). Then $F(A)$ is a disk centered at λ with*

$$\mathrm{r}(A) = |\eta| \, \cos \frac{\pi}{n+1}.$$

If $|\lambda| > |\eta| \, \cos \frac{\pi}{n+1}$ then

$$\mathrm{dist}(F(A), 0) = |\lambda| - \cos \frac{\pi}{n+1} \, |\eta|,$$

otherwise $0 \in \mathrm{F}(A)$.

Proof. Write $A = \lambda I + \eta Z$, where

$$Z \equiv \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$$

is a shift. $\mathrm{F}(Z)$ is a disk centered at the origin with radius

$$\mathrm{r}(Z) = \cos \frac{\pi}{n+1}$$

[11, Lemma 3], [6, Example in §1.3].

The scalar multiplication property [8, §1.2.4] of the field of values implies that $\mathrm{F}(\eta Z) = \eta\,\mathrm{F}(Z)$. Hence $\mathrm{F}(\eta Z)$ is a disk centered at zero with radius $|\eta|\ \mathrm{r}(Z)$. The translation property [8, §1.2.3] implies that $\mathrm{F}(A)$ is a disk centered at λ with the same radius, i.e.

$$\mathrm{r}(A) = |\eta|\,\mathrm{r}(Z).$$

The expression for the radius also follows from [12, p 107] and [2, Corollary 4].

To determine $\mathrm{dist}(\mathrm{F}(A), 0)$ first derive the lower bound

$$\begin{aligned}\mathrm{dist}(F(A),0) &= \min_{x\neq 0} \frac{|x^*Ax|}{|x^*x|} = \min_{x\neq 0} \left|\lambda + \eta\,\frac{x^*Zx}{x^*x}\right| \geq |\lambda| - |\eta|\ \max_{x\neq 0}\left|\frac{x^*Zx}{x^*x}\right| \\ &= |\lambda| - |\eta|\ \mathrm{r}(Z) = |\lambda| - |\eta|\ \cos\frac{\pi}{n+1}.\end{aligned}$$

Now show that equality is attained if $|\lambda| > |\eta|\ \mathrm{r}(Z)$. Write $\lambda = e^{\sqrt{-1}\,\theta}\,|\lambda|$ for some $0 \leq \theta \leq 2\pi$. Since $F(Z)$ is a disk there is a vector v such that

$$\frac{v^*Zv}{v^*v} = -e^{\sqrt{-1}\,\theta}\,\mathrm{r}(Z).$$

Hence

$$\mathrm{dist}(\mathrm{F}(A),0) \leq \left|e^{\sqrt{-1}\,\theta}\,(|\lambda| - |\eta|\ \mathrm{r}(Z))\right| = |\lambda| - |\eta|\ \mathrm{r}(Z).$$

Otherwise, if $|\lambda| \leq |\eta|\ \mathrm{r}(Z)$, select a vector w such that

$$\frac{w^*Zw}{w^*w} = e^{\sqrt{-1}\,\theta}\,\mathrm{r}(Z).$$

Hence both

$$\frac{w^*Aw}{w^*w} = e^{\sqrt{-1}\,\theta}\,(|\lambda| + |\eta|\ \mathrm{r}(Z)) \qquad \text{and} \qquad \frac{v^*Av}{v^*v} = e^{\sqrt{-1}\,\theta}\,(|\lambda| - |\eta|\ \mathrm{r}(Z))$$

are in $\mathrm{F}(A)$, where

$$|\lambda| - |\eta|\ \mathrm{r}(Z) \leq 0 \qquad \text{and} \qquad |\lambda| + |\eta|\ \mathrm{r}(Z) \geq 0.$$

Convexity of the field of values [8, §1.2.2] implies that also $e^{\sqrt{-1}\,\theta}\,0 = 0 \in \mathrm{F}(A)$. □

Theorem 2.1 says that the numerical radius of the matrix in (2.1) is proportional to its departure from normality, and that the distance from zero of the field of values depends on how large the departure from normality is compared to the magnitude of the eigenvalue. In particular, the field of values contains 0 if the deaprture from normality is too large compared to the magnitude of the eigenvalue. A large departure from normality can make things only worse: it can only move the field of values closer to zero but it cannot push it further away. For fixed λ and η, $\cos\frac{\pi}{n+1} \to 1$ as $n \to \infty$. Hence the field of values increases with growing matrix size.

3. Connection to Krylov Methods. It is shown that GMRES applied to a linear system with coefficient matrix (2.1) converges fast if the field of values of A is far away from zero.

In [9, §3] we consider the linear system

$$
(3.1)\quad Ax = b \qquad \text{where } A = \begin{pmatrix} \lambda & \eta & & \\ & \lambda & \ddots & \\ & & \ddots & \eta \\ & & & \lambda \end{pmatrix}, \quad \lambda \neq 0, \qquad b = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},
$$

which has the coefficient matrix from (2.1). If this linear system is solved by the Krylov method GMRES, the relative residual norms depend on how large the departure from normality is compared to the magnitude of the eigenvalue. In particular, the relative residual norm in iteration i of GMRES can be estimated as follows [9, Corollary 3.4]. If $|\eta| \geq |\lambda|$ then

$$
\min_{z \in \mathcal{K}_i} \frac{\|b - Az\|}{\|b\|} \geq \frac{1}{\sqrt{i+1}},
$$

and if $|\eta| \ll |\lambda|$ then

$$
\min_{z \in \mathcal{K}_i} \frac{\|b - Az\|}{\|b\|} \approx \left|\frac{\eta}{\lambda}\right|^i.
$$

This means the minimal residual norm decreases fast if the departure from normality is much smaller than the magnitude of the eigenvalue. From Theorem 2.1 follows that GMRES applied to the linear system (3.1) converges fast if the field of values of A is far away from zero.

In [9, §3] the matrix A in (3.1) is called *weakly non-normal in the context of Krylov methods* if the magnitude of the eigenvalue is larger than the departure from normality, $|\lambda| \gg |\eta|$. Hence the field of values of A in is far away from zero if A is weakly non-normal.

4. Jordan Blocks of Different Sizes. Theorem 2.1 is extended to a matrix consisting of Jordan blocks of different sizes.

Let

$$
(4.1)\qquad A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}, \qquad \text{where} \quad J_i \equiv \begin{pmatrix} \lambda & \eta & & \\ & \lambda & \ddots & \\ & & \ddots & \eta \\ & & & \lambda \end{pmatrix}
$$

is a Jordan block of order $n_i \geq 1$. As before, $|\eta|$ is the departure of A from normality. Let $n_{max} \equiv \max_{1 \leq i \leq k} n_i$ be the order of the largest Jordan block, i.e. the index of λ. Since A is a direct sum of Jordan blocks, the numerical radius and distance to zero of the field of values now depend on the index.

THEOREM 4.1. *Let A be as in (4.1). Then $F(A)$ is a disk centered at λ with*

$$
\mathrm{r}(A) = |\eta| \, \cos \frac{\pi}{n_{max}+1}.
$$

If $|\lambda| > |\eta| \, \cos \frac{\pi}{n_{max}+1}$ then

$$
\mathrm{dist}(F(A), 0) = |\lambda| - \cos \frac{\pi}{n_{max}+1} \, |\eta|,
$$

otherwise $0 \in \mathrm{F}(A)$.

Proof. The expression for $\mathrm{r}(A)$ follows from [11, Theorem], [6, Theorem 5.3-1] and [17, Example 1]. The rest of the proof is analogous to that of Theorem 2.1. $\square$

The remarks immediately after Theorem 2.1 also apply here. In addition, the field of values of A in (4.1) is further away from zero if A consists of only small Jordan blocks than if A contains a huge Jordan block.

5. A More General Jordan Block. The results in the previous sections are extended to Jordan blocks with different off-diagonal elements.

Let

$$A = \begin{pmatrix} \lambda & \eta_1 & & \\ & \lambda & \ddots & \\ & & \ddots & \eta_{n-1} \\ & & & \lambda \end{pmatrix} \tag{5.1}$$

be a matrix of order n. The departure of A from normality is $\max_{1\leq i\leq n-1}|\eta_i|$. In contrast to Theorems 2.1 and 4.1, we are only able to derive bounds rather than equalities.

THEOREM 5.1. *Let A be as in (5.1). Then*

$$\mathrm{r}(A) \leq \min\left\{\max_{1\leq i\leq n-1}\{|\eta_i|\}\cos\frac{\pi}{n+1},\quad \frac{1}{2}\max_{1\leq i\leq n}\{|\eta_{i-1}+\eta_i|\}\right\},$$

where $\eta_0 = \eta_n \equiv 0$, and

$$|\lambda| - \min\left\{\max_{1\leq i\leq n-1}\{|\eta_i|\}\cos\frac{\pi}{n+1},\quad \frac{1}{2}\max_{1\leq i\leq n}\{|\eta_{i-1}+\eta_i|\}\right\} \leq \mathrm{dist}(\mathrm{F}(A), 0) \leq |\lambda|.$$

Proof. As in the proof of Theorem 2.1 write $A = \lambda I + Z$, where

$$Z \equiv \begin{pmatrix} 0 & \eta_1 & & \\ & 0 & \ddots & \\ & & \ddots & \eta_{n-1} \\ & & & 0 \end{pmatrix}$$

is a weighted shift. Since $\mathrm{F}(Z)$ is a disk centered at the origin with radius bounded by [1, Theorem 3],

$$\mathrm{r}(Z) \leq \min\left\{\max_{1\leq i\leq n-1}\{|\eta_i|\}\cos\frac{\pi}{n+1},\quad \frac{1}{2}\max_{1\leq i\leq n}\{|\eta_{i-1}+\eta_i|\}\right\},$$

where $\eta_0 = \eta_n \equiv 0$, $\mathrm{F}(A)$ is a disk of the same radius centered at λ.

The lower bound on $\mathrm{dist}(\mathrm{F}(A), 0)$ follows from the fact that $\mathrm{dist}(\mathrm{F}(A), 0) \geq |\lambda| - \mathrm{r}(Z)$, and for the upper bound use the fact that $x^*Zx = 0$ when x is the first column of the identity. $\square$

The first term in the bound on $\mathrm{r}(A)$ is as in Theorems 2.1 and 4.1. This means the numerical radius of A is bounded above by the departure from normality. When all off-diagonal elements are the same, i.e. $\eta_i \equiv \eta$, then Theorem 5.1 reduces to Theorem 2.1. However it is not clear how tight the bound in Theorem 5.1 is.

EXAMPLE 1. *Let A in (5.1) have*

$$\eta_1 \equiv \sqrt{2}, \qquad \eta_2 = \cdots = \eta_{n-1} = 1.$$

Since A has non-negative elements, its numerical radius is the spectral radius of the real part of A [3, Theorem 2.1],

$$\mathrm{r}(A) = \frac{1}{2} \max_{1 \le i \le n} |\lambda_i(A + A^T)|,$$

and the spectral radius of $A + A^T$ is [13, §2.6.2], [17, page 3585]

$$\max_{1 \le i \le n} |\lambda_i(A + A^T)| = \cos \frac{\pi}{2(n+1)}.$$

Thus the numerical radius of A equals

$$\mathrm{r}(A) = \frac{1}{2} \cos \frac{\pi}{2(n+1)}.$$

However Theorem 5.1 gives the bound

$$\mathrm{r}(A) \le \min \left\{ \sqrt{2} \cos \frac{\pi}{n+1}, \frac{1}{2}(\sqrt{2}+1) \right\} = \frac{1}{2}(\sqrt{2}+1) \qquad \text{for} \quad n \ge 5.$$

For large n, $\mathrm{r}(A) \approx \frac{1}{2}$, while the bound in Theorem 5.1 is $\frac{1}{2} + \frac{1}{\sqrt{2}}$.

6. Jordanblocks of Order One and Two. The results from §2 are extended to a matrix that is similar to a Jordan matrix consisting of Jordanblocks of order one or two,

$$A = X \begin{pmatrix} \ddots & & \\ & J_i & \\ & & \ddots \end{pmatrix} X^{-1}, \qquad \text{where} \quad J_i = \begin{pmatrix} \lambda & 1 \\ & \lambda \end{pmatrix} \text{ or } J_i = \lambda.$$

Suppose there are k Jordanblocks of order 2 and $n-2k$ Jordanblocks of order 1. Then A has a Schur decomposition [4, §10], [10], [14]

$$A = QTQ^*, \qquad \text{where} \quad T \equiv \begin{pmatrix} \lambda I_{n-k} & N \\ & \lambda I_k \end{pmatrix}, \tag{6.1}$$

N is $(n-k) \times k$, and Q is unitary.

THEOREM 6.1. *Let A be as in (6.1). Then $\mathrm{F}(A)$ is a disk centered at λ with*

$$\mathrm{r}(A) = \frac{1}{2}\|N\|.$$

If $|\lambda| > \frac{1}{2}\|N\|$ then

$$\mathrm{dist}(F(A), 0) = |\lambda| - \frac{1}{2}\|N\|,$$

otherwise $0 \in \mathrm{F}(A)$.

Proof. Since Q is unitary $\mathrm{F}(A)=\mathrm{F}(T)$ [8, §1.2.8]. Write

$$T=\lambda I+N_0, \qquad \text{where} \quad N_0 \equiv \begin{pmatrix} 0 & N \\ 0 & 0 \end{pmatrix}.$$

Since $\mathrm{F}(N_0)$ is a disk centered at the origin with radius [8, Problem 1.2.25(a)],

$$\mathrm{r}(N_0)=\frac{1}{2}\|N\|,$$

$\mathrm{F}(T)$ is a disk of the same radius centered at λ.

The expression for $\mathrm{dist}(F(A),0)$ is derived as in the proof of Theorem 2.1. □

The remarks made after Theorem 2.1 also apply here. Theorems 2.1 and 6.1 are identical in the case $n=2$, where A has a Schur decomposition

$$A=Q\begin{pmatrix} \lambda & \eta \\ & \lambda \end{pmatrix}Q^*.$$

7. General Matrices. The previous results for matrices with a single eigenvalue are extended to general matrices. Unfortunately the conclusions are not as strong.

Let $A=Q(\Lambda+N)Q^*$ be a Schur decomposition of A, where the diagonal matrix Λ contains the eigenvalues of A, and N is strictly upper triangular. As in Theorem 5.1 we were only able to derive bounds.

We show that the distance from zero of the field of values of A is bounded below by the distance from zero of the field of values of the normal part minus the numerical radius of the non-normal part.

THEOREM 7.1.

$$\mathrm{dist}(\mathrm{F}(\Lambda),0)-\mathrm{r}(N)\leq \mathrm{dist}(\mathrm{F}(A),0)\leq \min_i|\lambda_i|.$$

Proof. Since Q is unitary, $\mathrm{F}(A)=\mathrm{F}(\Lambda+N)$, and

$$\begin{aligned}\mathrm{dist}(\mathrm{F}(A),0) &= \min_{x\neq 0}\left|\frac{x^*\Lambda x}{x^*x}+\frac{x^*Nx}{x^*x}\right| \geq \min_{x\neq 0}\left|\frac{x^*\Lambda x}{x^*x}\right|-\max_{x\neq 0}\left|\frac{x^*Nx}{x^*x}\right| \\ &= \mathrm{dist}(F(\Lambda),0)-\mathrm{r}(N).\end{aligned}$$

Regarding the upper bound, if x is the ith column of the identity matrix then $x^*Nx=0$ since N is strictly upper triangular. Hence $\mathrm{dist}(\mathrm{F}(A),0)\leq|\lambda_i|$. □

When A is as in (2.1), (4.1) or (6.1) then the lower bound for $\mathrm{dist}(\mathrm{F}(A),0)$ in Theorem 7.1 is the same as the expressions in Theorems 2.1, 4.1 and 6.1, respectively.

To illustrate the harmful effect of the departure from normality we consider a situation where the field of values of the normal part of A does not contain zero. This is the case, for instance, when all eigenvalues of A lie on a ray in the complex plane emanating from the origin (i.e. all eigenvalues are positive, or all eigenvalues lie on the upper half of the imaginary axis). Then it can be shown that the departure from normality can make things only worse, i.e. move the field of values closer to zero but not push it further away.

COROLLARY 7.2. *If* $\lambda_i=e^{\sqrt{-1}\,\theta}\,|\lambda_i|$ *for some* $0\leq\theta\leq 2\pi$ *then*

$$\min_i|\lambda_i|-\mathrm{r}(N)\leq \mathrm{dist}(\mathrm{F}(A),0)\leq \min_i|\lambda_i|.$$

Proof. Since $\Lambda = e^{\sqrt{-1}\,\theta}\,|\Lambda|$, where $|\Lambda|$ is Hermitian positive semi-definite,

$$\mathrm{dist}(\mathrm{F}(\Lambda), 0) = \min_{x\neq 0}\left|\frac{x^*\Lambda x}{x^*x}\right| = \min_{x\neq 0}\frac{x^*|\Lambda|x}{x^*x} = \min_i |\lambda_i|.$$

□

Since the numerical radius is close to the two-norm [3, Theorem 1.2], [6, Theorem 1.3-1]

$$\frac{1}{2}\|N\| \leq \mathrm{r}(N) \leq \|N\|,$$

one does not lose very much by replacing $\mathrm{r}(N)$ by $\|N\|$.

COROLLARY 7.3.

$$\mathrm{dist}(\mathrm{F}(\Lambda), 0) - \|N\| \leq \mathrm{dist}(\mathrm{F}(A), 0) \leq \min_i |\lambda_i|.$$

This means, if the departure from normality is small then the field of values of A is not much farther away from zero than the field of values of its normal part.

REFERENCES

[1] M. CHIEN, *On the numerical range of tridiagonal operators*, Linear Algebra Appl., 246 (1996), pp. 203–14.

[2] M. EIERMANN, *Fields of values and iterative methods*, Linear Algebra Appl., 180 (1993), pp. 167–97.

[3] M. GOLDBERG AND E. TADMOR, *On the numerical radius and its applications*, Linear Algebra Appl., 42 (1982), pp. 263–84.

[4] G. GOLUB AND J. WILKINSON, *Ill-conditioned eigensystems and the computation of the Jordan canonical form*, SIAM Review, 18 (1976), pp. 578–619.

[5] A. GREENBAUM, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.

[6] K. GUSTAFSON AND D. RAO, *Numerical Range*, Springer Verlag, New York, 1997.

[7] P. HENRICI, *Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices*, Numer. Math., 4 (1962), pp. 24–40.

[8] R. HORN AND C. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[9] I. IPSEN, *A different approach to bounding the minimal residual norm in Krylov methods*, Tech. Rep. CRSC-TR98-19, Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, 1998. Submitted to SIAM J. Numer. Anal.

[10] B. KÅGSTRÖM AND A. RUHE, *An algorithm for numerical computation of the Jordan normal form of a complex matrix*, ACM Trans. Math. Softw., 6 (1980), pp. 398–419.

[11] M. MARCUS AND B. SHURE, *The numerical range of certain* $0,1$*-matrices*, Linear And Multilinear Algebra, 7 (1979), pp. 111–20.

[12] W. RIDGE, *Numerical range of a weighted shift with periodic weights*, Proc. Amer. Math. Soc., 55 (1976), pp. 107–10.

[13] P. ROEBUCK AND S. BARNETT, *A survey of Toeplitz and related matrices*, Int. J. Systems Sci., 9 (1978), pp. 921–34.

[14] A. RUHE, *An algorithm for numerical determination of the structure of a general matrix*, BIT, 10 (1970), pp. 196–216.

[15] G. STARKE, *Fields of values and the ADI method for non-normal matrices*, Linear Algebra Appl., 180 (1993), pp. 199–218.

[16] ———, *Fields-of-values analysis of preconditioned iterative methods for nonsymmetric elliptical problems*, Numer. Math., 78 (1997), pp. 103–17.

[17] T. TAM, *On a conjecture of Ridge*, Proc. Amer. Math. Soc., 125 (1997), pp. 3581–92.