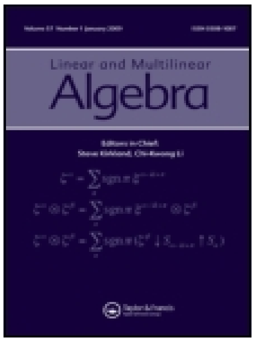

Linear and Multilinear Algebra

ISSN: 0308-1087 (Print) 1563-5139 (Online) Journal homepage: http://www.tandfonline.com/loi/glma20

On Matrices with Elliptical Numerical Ranges

Ethan S. Brown & Ilya M. Spitkovsky

To cite this article: Ethan S. Brown & Ilya M. Spitkovsky (2004) On Matrices with Elliptical Numerical Ranges, Linear and Multilinear Algebra, 52:3-4, 177-193, DOI: 10.1080/03081080310001112589

To link to this article: http://dx.doi.org/10.1080/03081080310001112589

Published online: 31 Aug 2006.

Linear and Multilinear Algebra, 2004, Vol. 52, Nos. 3–4, pp. 177–193

Taylor & Francis
Taylor & Francis Group

On Matrices with Elliptical Numerical Ranges

ETHAN S. BROWN[a,*] and ILYA M. SPITKOVSKY[b,†]

[a]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA;
[b]Department of Mathematics, College of William and Mary, Williamsburg, VA 23187, USA

(Received 11 December 2002)

It is known that for several classes of matrices, including quadratic and certain tridiagonal matrices, the numerical range is an ellipse. We prove this result for a larger class of matrices, encompassing these published results as well as providing other sufficient conditions for ellipticity. The proof gives an explanation for this phenomenon and, in certain cases, provides an explicit description of the numerical range.

Keywords: Numerical range; Tridiagonal matrices

1991 Mathematics Subject Classifications: Primary 15A60; Secondary 47A12; 47B36

1. INTRODUCTION

Let $\mathbb{C}^{n\times n}$ denote the set of all $n \times n$ matrices with entries from the field $\mathbb{C}$ of complex numbers. The *numerical range* of a matrix $A \in \mathbb{C}^{n\times n}$, introduced by Toeplitz early in the 20th century [13], is defined as the set

$$W(A) = \{x^*Ax\colon\ x \in \mathbb{C}^n,\ x^*x = 1\}.$$

It is well known that $W(A)$ is a convex compact subset of $\mathbb{C}$. The set is unitarily invariant; that is, $W(A) = W(B)$ whenever A and B are unitarily equivalent ($B = U^*AU$ for some unitary matrix U). It is also well behaved under direct sums: if $A = A_1 \oplus \cdots \oplus A_N$, then

$$W(A) = \operatorname{conv}\{W(A_1) \cup \cdots \cup W(A_N)\},$$

denoting by $\operatorname{conv}\{S\}$ the convex hull of a set S. For these properties of the numerical range, and for a history of the subject, see [6,8].

The numerical range of a 2×2 matrix is an ellipse. More precisely, the following result holds [8] (see also [11] for a short proof).

THEOREM 1.1 *Let A be a 2×2 matrix with eigenvalues λ_1 and λ_2. Then $W(A)$ is an ellipse with the foci at λ_1 and λ_2, and minor axis of length* $(\operatorname{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2)^{1/2}$. *In particular, the center of the ellipse $W(A)$ is at* $\operatorname{tr}A/2$.

*E-mail: esbrown@mit.edu
†Corresponding author. E-mail: ilya@math.wm.edu

ISSN 0308-1087 print: ISSN 0000-000X online/01/010000-00 © 2004 Taylor & Francis Ltd
DOI: 10.1080/03081080310001122589

In fact, for any n, the numerical range of an $n \times n$ matrix A is a convex hull of a certain algebraic curve, sometimes referred to as the *associated curve* of A [10] (see also [12]). The *class* of this curve (that is, the degree of its dual) equals n, but its degree may be substantially higher than n for $n > 2$. Nevertheless, for several types of matrices recently considered in the literature, it was shown that, independently of their size, the numerical range is an ellipse. This is the case for all *quadratic matrices* (that is, those with minimal polynomial of degree 2) [14]. It is also true for tridiagonal matrices of the form

$$\begin{pmatrix} a_1 & 1 & 0 & \cdots & 0 \\ -1 & a_2 & 1 & \ddots & \vdots \\ 0 & -1 & a_1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & -1 & \end{pmatrix} \tag{1.1}$$

with a 2-periodic main diagonal [3], and of the form

$$\begin{pmatrix} 0 & b_1 & 0 & \cdots & 0 \\ c_1 & 0 & b_2 & \ddots & \vdots \\ 0 & c_2 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b_{n-1} \\ 0 & \cdots & 0 & c_{n-1} & 0 \end{pmatrix} \tag{1.2}$$

with zero main diagonal and the off-diagonal entries such that at least one of b_j and c_j is equal to zero for all $j = 1, \dots, n-1$. In the latter case the numerical range is even a disk [2].

In this article, we present a unifying result explaining what we believe is an intrinsic reason for this phenomenon. Along the way, new classes of matrices with elliptical numerical ranges are found and described.

The article is organized as follows. Section 2 describes a class of matrices whose numerical range is the convex hull of a small number of ellipses, and a subclass for which this number drops down to one. In Section 3, these results are applied to tridiagonal matrices with 2-periodic main diagonal. For 3×3 tridiagonal matrices, an alternative approach is available, based on results of [9]. This approach is presented in Section 4. Finally, in Section 5 we explain why some seemingly natural generalizations of the results of Section 2 are not possible.

2. GENERAL RESULT

THEOREM 2.1 *Let the matrix A be unitarily equivalent to a matrix of the form*

$$\begin{pmatrix} a_1 I_{n_1} & X \\ Y^* & a_2 I_{n_2} \end{pmatrix}, \tag{2.1}$$

where XY^* and Y^*X are normal matrices. Then the numerical range of A is the convex hull of at most $\min\{n_1, n_2\}$ *ellipses, all of which centered at* $(a_1 + a_2)/2$.

Proof Due to unitary invariance of $W(A)$, we may suppose without loss of generality that A itself is of the form (2.1). Moreover, we can further apply a unitary transformation $A \mapsto U^*AU$ with

$$U = \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}$$

and blocks U_j of the size $n_j \times n_j$. Under this transformation the scalar diagonal blocks of (2.1) are preserved and the off-diagonal blocks are changed as $X \mapsto U_1^* X U_2$, $Y \mapsto U_1^* Y U_2$. According to singular value decomposition, with an appropriate choice of unitary U_1 and U_2, we have that $U_1^* X U_2$ is a diagonal matrix Σ with the diagonal entries $\Sigma_{jj} = \sigma_j (\geq 0)$, $j = 1, \ldots, \min\{n_1, n_2\}$, being the *singular values* of the matrix X. As it happens, the normality of XY^* and Y^*X is a necessary and sufficient condition under which it is possible to simultaneously make the matrix $\Delta = U_1^* Y U_2$ diagonal as well (see, for example, [7, p. 426]). Observe that the diagonal entries δ_j of Δ are not necessarily nonnegative, but their absolute values $|\delta_j|$ up to their order coincide with the singular values of Y.

Thus, the numerical range of the matrix A is the same as that of the matrix

$$U^*AU = \begin{pmatrix} a_1 I_{n_1} & \Sigma \\ \Delta^* & a_2 I_{n_2} \end{pmatrix}.$$

Let now P be a permutation matrix corresponding to the permutation

$$1, n_1 + 1, 2, n_1 + 2, \ldots, n_1, 2n_1, 2n_1 + 1, \ldots, n_1 + n_2$$

if $n_1 \leq n_2$ and

$$1, n_1 + 1, 2, n_1 + 2, \ldots, n_2, n_1 + n_2, n_2 + 1, \ldots, n_1$$

if $n_1 \geq n_2$. Then $(UP)^* A (UP)$ is the direct sum of the matrix

$$B = \bigoplus_{i=1}^{\min\{n_1, n_2\}} B_i, \quad \text{where } B_i = \begin{pmatrix} a_1 & \sigma_i \\ \overline{\delta_i} & a_2 \end{pmatrix},$$

with $n_1 - n_2$ one-dimensional blocks (a_1) if $n_1 > n_2$ or $n_2 - n_1$ one-dimensional blocks (a_2) if $n_1 < n_2$. Since both a_1 and a_2 lie in the numerical range of B, in either case $W(A)$ is the same as $W(B)$, that is, it coincides with the convex hull of the numerical ranges of its 2×2 diagonal blocks B_i. It remains to apply Theorem 1.1. ∎

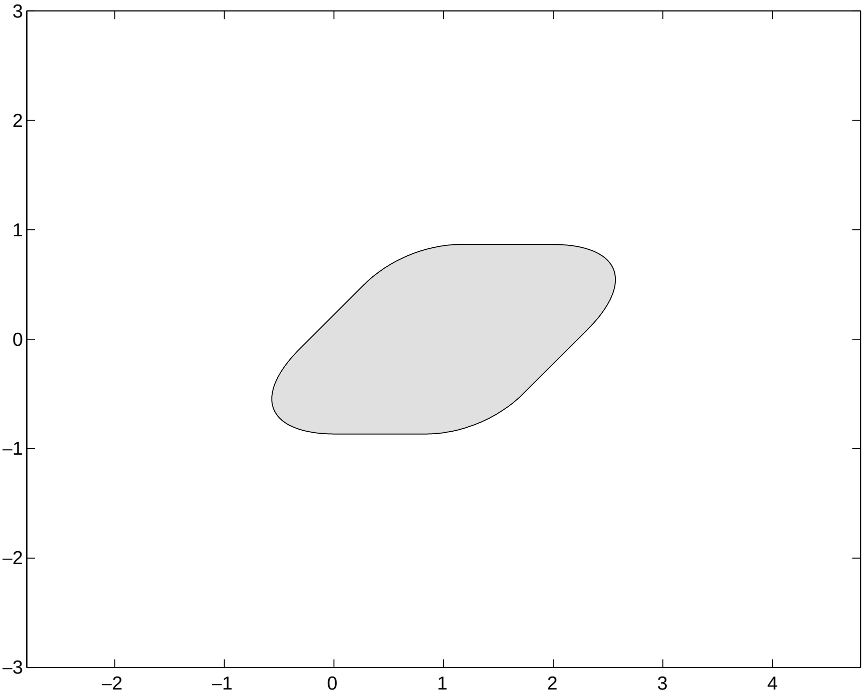

FIGURE 1 The numerical range of a 4×4 matrix A, unitarily reducible to two 2×2 block matrices, that is the convex hull of two non-nested ellipses.

Example 2.2 Let A be the 4×4 matrix

$$\begin{pmatrix} 2 & 0 & -\frac{1}{2}+\frac{1}{2}i & \frac{1}{2}+\frac{3}{2}i \\ 0 & 2 & \frac{1}{2}-\frac{1}{2}i & -\frac{1}{2}+\frac{1}{2}i \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

This matrix is of the form (2.1). In addition, the matrices

$$XY^* = \begin{pmatrix} \frac{1}{2}+\frac{3}{2}i & \frac{1}{2}-\frac{1}{2}i \\ -\frac{1}{2}+\frac{1}{2}i & -\frac{1}{2}+\frac{1}{2}i \end{pmatrix} \quad \text{and} \quad Y^*X = \begin{pmatrix} -\frac{1}{2}+\frac{1}{2}i & \frac{1}{2}-\frac{1}{2}i \\ -\frac{1}{2}+\frac{1}{2}i & \frac{1}{2}+\frac{3}{2}i \end{pmatrix}$$

are normal. One can check that A is unitarily reducible to

$$\begin{pmatrix} 2 & \sqrt{2+\sqrt{2}} \\ 1 & 0 \end{pmatrix} \oplus \begin{pmatrix} 2 & \sqrt{2-\sqrt{2}} \\ 1 & 0 \end{pmatrix}.$$

The numerical range of A is shown[1] in Fig. 1.

[1]All figures in this article were produced using the program given in [4].

Let us now consider the case when Y is a scalar multiple of X. It is clear then that XY^* and Y^*X are scalar multiples of the Hermitian matrices XX^* and X^*X, and are therefore normal. Thus, Theorem 2.1 applies. However, more can be said in this particular situation.

COROLLARY 2.3 *Let A be unitarily equivalent to a matrix*

$$\begin{pmatrix} a_1 I_{n_1} & X \\ kX^* & a_2 I_{n_2} \end{pmatrix}. \tag{2.2}$$

Then the numerical range of A is the ellipse with the foci

$$\lambda_{1,2} = \frac{a_1 + a_2 \pm \sqrt{(a_1 - a_2)^2 + 4k\sigma^2}}{2},$$

where $\sigma = \|X\|$, major axis of length

$$\sqrt{\frac{1}{2}|a_1 - a_2|^2 + \sigma^2(|k|^2 + 1) + \left|\frac{1}{2}(a_1 - a_2)^2 + 2\sigma^2 k\right|},$$

and minor axis of length

$$\sqrt{\frac{1}{2}|a_1 - a_2|^2 + \sigma^2(|k|^2 + 1) - \left|\frac{1}{2}(a_1 - a_2)^2 + 2\sigma^2 k\right|}.$$

Proof Following the proof of Theorem 2.1, even without a recourse to the criterion of simultaneous diagonalizability from [7], it is obvious that for the matrix (2.2), $\Delta = \overline{k}\Sigma$ is diagonal together with Σ, and that $\delta_j = \overline{k}\sigma_j$. Thus, $W(A)$ is the convex hull of the numerical ranges of the matrices A_{σ_j}, where

$$A_s = \begin{pmatrix} a_1 & s \\ ks & a_2 \end{pmatrix},$$

and $\sigma_j (\geq 0)$ is the jth singular value of X. We claim that, moreover, $W(A_s)$ is an increasing function of the nonnegative parameter s; that is,

$$W(A_{s_1}) \subset W(A_{s_2}) \quad \text{for any } (0 \leq) s_1 < s_2. \tag{2.3}$$

As soon as this fact is established, it follows that $W(A)$ coincides with $W(A_{\sigma_1})$, where σ_1 is $\|X\|$, the largest singular value of X. The rest follows directly from Theorem 1.1.

It remains to show (2.3). To this end, observe that directly from the definition of numerical range,

$$W(A_s) = \left\{ a_1|x_1|^2 + a_2|x_2|^2 + s(\overline{x_1}x_2 + kx_1\overline{x_2}) : \ |x_1|^2 + |x_2|^2 = 1 \right\}.$$

Now let $|x_1| = \cos\phi$, $|x_2| = \sin\phi$, and $\arg(\overline{x_1}x_2) = \theta$ for some $\phi \in [0, \pi/2]$ and $\theta \in [0, 2\pi)$. Then

$$W(A_s) = \bigcup_{\phi}\bigcup_{\theta}\{a_1\cos^2\phi + a_2\sin^2\phi + s\sin\phi\cos\phi[(k+1)\cos\theta + i(1-k)\sin\theta]\}. \tag{2.4}$$

Denote by $\mathcal{E}$ the curve

$$\{(k+1)\cos\theta + i(1-k)\sin\theta\colon\ \theta \in [0, 2\pi)\}.$$

Depending on the value of k, $\mathcal{E}$ is either the boundary of an ellipse or a line segment, in both cases centered at the origin. From (2.4),

$$W(A_s) = \bigcup_{\phi}\{a_1\cos^2\phi + a_2\sin^2\phi + (s\sin\phi\cos\phi)\mathcal{E}\}.$$

Any $z \in W(A_{s_1})$ therefore lies on the curve $\Gamma_1 = a_1\cos^2\phi_1 + a_2\sin^2\phi_1 + (s_1\sin\phi_1\cos\phi_1)\mathcal{E}$ for a certain choice of ϕ_1. If $s_2 > s_1$, then Γ_1 lies in the domain bounded by $\Gamma_2 = a_1\cos^2\phi_1 + a_2\sin^2\phi_1 + (s_2\sin\phi_1\cos\phi_1)\mathcal{E}$. Since $\Gamma_2 \subset W(A_{s_2})$ and the latter set is convex, it follows that $(z \in)\Gamma_1 \subset W(A_{s_2})$. Thus, (2.3) holds. ■

If A is a quadratic matrix, then A is unitarily equivalent to a matrix

$$\begin{pmatrix} \lambda_1 I & X \\ 0 & \lambda_2 I \end{pmatrix},$$

where $\{\lambda_1, \lambda_2\}$ is the spectrum of A. Thus, conditions of Corollary 2.3 are satisfied with $a_j = \lambda_j$, $k = 0$, and the result of [14, Theorem 2.1] follows.

3. TRIDIAGONAL MATRICES

Recall that a *tridiagonal matrix* is a matrix $T \in \mathbb{C}^{n\times n}$ such that $t_{ij} = 0$ whenever $|i-j| > 1$. We will usually label the nonzero entries of an arbitrary tridiagonal matrix in the following way:

$$\begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 \\ c_1 & a_2 & b_2 & \ddots & \vdots \\ 0 & c_2 & a_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b_{n-1} \\ 0 & \cdots & 0 & c_{n-1} & a_n \end{pmatrix}. \tag{3.1}$$

As it happens, the numerical range of such matrices is invariant under the switching of any two *corresponding* off-diagonal entries b_j and c_j.

LEMMA 3.1 *The numerical range of an $n \times n$ tridiagonal matrix is invariant under interchange of the $i, i+1$ and $i+1, i$ entries for any $i = 1, \dots, n-1$.*

Proof Let

$$A = \begin{pmatrix} a_1 & b_1 & 0 & 0 & 0 & 0 \\ c_1 & \ddots & \ddots & 0 & 0 & 0 \\ 0 & \ddots & a_j & b_j & 0 & 0 \\ 0 & 0 & c_j & a_{j+1} & \ddots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & b_{n-1} \\ 0 & 0 & 0 & 0 & c_{n-1} & a_n \end{pmatrix}$$

and

$$\widehat{A} = \begin{pmatrix} a_1 & b_1 & 0 & 0 & 0 & 0 \\ c_1 & \ddots & \ddots & 0 & 0 & 0 \\ 0 & \ddots & a_j & c_j & 0 & 0 \\ 0 & 0 & b_j & a_{j+1} & \ddots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & b_{n-1} \\ 0 & 0 & 0 & 0 & c_{n-1} & a_n \end{pmatrix}$$

be $n \times n$ tridiagonal matrices that differ only by interchanging b_j and c_j. Consider an arbitrary point $z = x^*Ax \in W(A)$. Let $x = (x_1, \dots, x_n)^T \in \mathbb{C}^n$. Then x^*Ax can be represented as

$$\begin{aligned} &a_1|x_1|^2 + \cdots + a_{j-1}|x_{j-1}|^2 + a_j|x_j|^2 + a_{j+1}|x_{j+1}|^2 + \cdots + a_n|x_n|^2 \\ &\quad + b_1\overline{x_1}x_2 + \cdots + b_{j-1}\overline{x_{j-1}}x_j + b_j\overline{x_j}x_{j+1} + b_{j+1}\overline{x_{j+1}}x_{j+2} + \cdots + b_{n-1}\overline{x_{n-1}}x_n \\ &\quad + c_1x_1\overline{x_2} + \cdots + c_{j-1}x_{j-1}\overline{x_j} + c_jx_j\overline{x_{j+1}} + c_{j+1}x_{j+1}\overline{x_{j+2}} + \cdots + c_{n-1}x_{n-1}\overline{x_n}. \end{aligned}$$

We claim that there exists $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n)^T \in \mathbb{C}^n$ such that $|x_i| = |\hat{x}_i|$ for all $i = 1, \dots, n$ and $x^*Ax = \hat{x}^*\widehat{A}\hat{x}$. For the latter equality to hold, it suffices that

$$x_j\overline{x_{j+1}} = \overline{\hat{x}_j}\hat{x}_{j+1} \quad \text{and} \quad x_i\overline{x_{i+1}} = \hat{x}_i\overline{\hat{x}_{i+1}} \quad \text{for all } i \neq j;\ i = 1, \dots, n-1. \tag{3.2}$$

If $x_j = 0$, we can simply put $\hat{x} = x$. Otherwise, let $\hat{x}_j = \overline{x_j}$ and $\hat{x}_{j+1} = \overline{x_{j+1}}$. Also, define p as the greatest positive integer less than j such that either $x_{p-1} = 0$ or $b_p = 0$ and $c_p = 0$. If no such integer exists, let $p = 0$. Similarly, let k be the smallest integer greater than j such that either $x_{k+1} = 0$ or $b_k = 0$ and $c_k = 0$, and if no such integer exists, let $k = n+1$. Now choose the arguments of $\hat{x}_i$ for $i \neq j, j+1$ according to the formulas

$$\arg(\hat{x}_i) = \arg(x_i) - 2\arg(x_j) \quad \text{for } p+1 \leq i \leq j-1$$

and

$$\arg(\hat{x}_i) = \arg(x_i) - 2\arg(x_{j+1}) \quad \text{for } j+2 \le i \le k-1.$$

For all $i \le p$ or $i \ge k$ we need not change the arguments of $\hat{x}_i$ and can assign $\hat{x}_i = x_i$ for such i. Then (3.2) holds, so that $x^*Ax = \hat{x}^*\widehat{A}\hat{x}$. Thus, $z \in W(\widehat{A})$. We just proved that $W(A) \subset W(\widehat{A})$; the reverse inclusion follows by switching the roles of A and $\widehat{A}$. ■

Observe that the operation of interchanging the corresponding off-diagonal entries in general cannot be reduced to unitary equivalence.

Example 3.2 For all $a \in \mathbb{C}$, the matrix

$$A = \begin{pmatrix} a & 1 & 0 & 0 \\ 0 & -a & 0 & 0 \\ 0 & 1 & a & 1 \\ 0 & 0 & 0 & -a \end{pmatrix}$$

is unitarily reducible, but the matrix

$$B = \begin{pmatrix} a & 1 & 0 & 0 \\ 0 & -a & 1 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & 0 & -a \end{pmatrix},$$

obtained from one interchange of corresponding off-diagonal entries of A, is unitarily irreducible.

Our next result describes a class of tridiagonal matrices with elliptical numerical ranges. It consists of the matrices (3.1) with 2-*periodic* main diagonal, that is, satisfying the condition

$$a_j = \begin{cases} a_1 & \text{if } j \text{ odd}, \\ a_2 & \text{if } j \text{ even}, \end{cases} \tag{3.3}$$

and off-diagonal entries b_j, c_j such that the following condition holds:

$$\text{For each } j = 1, \ldots, n-1, \text{ either } k\overline{b_j} = c_j \text{ or } k\overline{c_j} = b_j \text{ for } k \in \mathbb{C} \text{ constant.} \tag{3.4}$$

We emphasize that the first equality in (3.4) holds for j from some subset J_1 of $\{1, \ldots, n-1\}$; the second equality holds for $j \in J_2 = \{1, \ldots, n-1\} \setminus J_1$.

To formulate the result, let $d_j = b_j$ for $j \in J_1$ and $d_j = c_j$ for $j \in J_2$. Denote also by X the $\lceil n/2 \rceil \times \lfloor n/2 \rfloor$ bidiagonal matrix with the diagonal entries $d_1, d_3, \ldots$ and the entries

$d_2, d_4, \ldots$ directly below them:

$$X = \begin{pmatrix} d_1 & & & \\ d_2 & d_3 & & \\ & d_4 & d_5 & \\ & & \ddots & \ddots \end{pmatrix}. \tag{3.5}$$

(Here we use the standard notation $\lceil x \rceil$ and $\lfloor x \rfloor$ for the least integer greater than or equal to $x \in \mathbb{R}$ and the greater integer less than or equal to x, respectively.) Finally, let $\rho = \|X\|$.

THEOREM 3.3 *Let A be a tridiagonal matrix* (3.1) *satisfying conditions* (3.3) *and* (3.4). *Then the numerical range of A is the ellipse with center* $(a_1 + a_2)/2$, *foci*

$$\frac{a_1 + a_2 \pm \sqrt{(a_1 - a_2)^2 + 4\rho^2 k}}{2},$$

major axis length

$$\sqrt{\frac{1}{2}|a_1 - a_2|^2 + \rho^2(|k|^2 + 1) + \left|\frac{1}{2}(a_1 - a_2)^2 + 2\rho^2 k\right|},$$

and minor axis length

$$\sqrt{\frac{1}{2}|a_1 - a_2|^2 + \rho^2(|k|^2 + 1) - \left|\frac{1}{2}(a_1 - a_2)^2 + 2\rho^2 k\right|}.$$

Proof Due to Lemma 3.1, the numerical ranges of all matrices A satisfying the conditions of the theorem are the same, independent of the partition of the set $\{1, \ldots, n-1\}$ into subsets J_1 and J_2. So, it suffices to provide the proof for one partition of our choice. As it happens, the convenient situation is when $J_1 \cup J_2$ is the partition into the subsets of odd and even numbers.

Let P be the permutation matrix corresponding to the permutation of $1, \ldots, n$ into the sequence of all odd and then all even terms. Then P^*AP is a matrix of the form (2.1), with $n_1 = \lceil n/2 \rceil$ and $n_2 = \lfloor n/2 \rfloor$. Its right upper block X is indeed given by (3.5), due to our choice of J_1 and J_2. Condition (3.4) implies then that $Y = \overline{k}X$; that is, the matrix P^*AP is in fact of the form (2.2). It remains to apply Corollary 2.3, observing that $\rho = \sigma$. ■

Note that condition (3.4) is satisfied in case of constant super and sub diagonals: $b_1 = \cdots = b_{n-1}(= b)$ and $c_1 = \cdots = c_{n-1}(= c)$. In this case, the matrix X is simply

$$b\begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ & 1 & 1 & \\ & & \ddots & \ddots \end{pmatrix},$$

and its norm can be computed explicitly:

$$\rho = 2|b| \cos\left(\frac{\pi}{n+1}\right).$$

When applied to this setting, Theorem 3.3 covers, in particular, the result on the ellipticity of numerical ranges for Toeplitz tridiagonal matrices [5, Corollary 4]. On the other hand, the case of matrices (1.1) is covered as well, with $k = -1$ and $b = 1$. This implies the result of [3, Theorem 2].

From the expressions for the foci of the ellipse, we get the following corollary to Theorem 3.3 describing when these ellipses are circular discs.

Corollary 3.4 *Let A be a tridiagonal matrix* (3.1) *satisfying conditions* (3.3) *and* (3.4). *Then $W(A)$ is a circular disc if and only if $(a_1 - a_2)/(2\rho)$ is the square root of $-k$.*

Proof Observe first of all that $\rho = 0$ if and only if the matrix A is in fact diagonal, and its numerical range is then the line segment joining a_1 and a_2. When $\rho \neq 0$, the condition

$$\left(\frac{a_1 - a_2}{2\rho}\right)^2 + k = 0$$

is necessary and sufficient for the foci of the ellipse $W(A)$ to coincide. ■

Conditions of Corollary 3.4 are satisfied, in particular, if $a_1 = a_2$ (that is, the main diagonal of A is constant) and $k = 0$. This covers matrices (1.2), so that [2, Theorem 3] follows. A little more generally:

Corollary 3.5 *Let A be a tridiagonal matrix* (3.1) *with a* 2*-periodic main diagonal and such that $b_i c_i = 0$ for all $i = 1, \ldots, n-1$. Then $W(A)$ is an ellipse. If, in addition, the main diagonal of A equals some constant $a \in \mathbb{C}$, then $W(A)$ is a circular disc centered at a.*

4. MATRICES OF SMALL SIZE

Let us consider a 3×3 tridiagonal matrix A with 2-periodic main diagonal:

$$A = \begin{pmatrix} a_1 & b_1 & 0 \\ c_1 & a_2 & b_2 \\ 0 & c_2 & a_1 \end{pmatrix}. \tag{4.1}$$

Theorem 3.3 implies that $W(A)$ is an ellipse provided that $b_1 c_2 = b_2 c_1$ or $b_1 b_2 = c_1 c_2$. However, for arbitrary 3×3 matrices there is an independent test for the ellipticity of the numerical range, see [9, Theorems 2.3 and 2.4]:

Theorem 4.1 *Let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of $A \in \mathbb{C}^{3\times 3}$. The numerical range of A is a non-degenerate ellipse if and only if*

$$d = \operatorname{tr}(A^*A) - \sum_{j=1}^{3} |\lambda_j|^2 > 0, \tag{4.2}$$

the number

$$\lambda = \operatorname{tr}(A) + \frac{1}{d}\left(\sum_{j=1}^{3} |\lambda_j|^2 \lambda_j - \operatorname{tr}(A^* A^2)\right) \tag{4.3}$$

coincides with at least one of the eigenvalues λ_j *of* A (*say*, λ_3), *and*

$$(|\lambda_1 - \lambda_3| + |\lambda_2 - \lambda_3|)^2 - |\lambda_1 - \lambda_2|^2 \leq d. \tag{4.4}$$

When applied to matrices of the form (4.1), it leads to the following.[2]

THEOREM 4.2 *The numerical range of any* 3×3 *tridiagonal matrix with a* 2-*periodic main diagonal is an ellipse.*

In other words, no additional conditions on the off-diagonal entries are required for ellipticity of the numerical range in this low dimensional setting.

The results of [9] also allow us to describe $W(A)$ explicitly. Namely, for the matrix (4.1) its numerical range has foci at

$$\frac{a_1 + a_2 \pm \sqrt{(a_2 - a_1)^2 + 4(b_1 c_1 + b_2 c_2)}}{2},$$

and a minor axis of length

$$\sqrt{\frac{|a_2 - a_1|^2 + 2|b_1|^2 + 2|b_2|^2 + 2|c_1|^2 + 2|c_2|^2 - \left|(a_2 - a_1)^2 + 4(b_1 c_1 + b_2 c_2)\right|}{2}}.$$

Theorem 4.2 for the case of zero (and thus any constant) main diagonal was established in [2, Theorem 4]. It shows that, even for $n = 3$, Theorem 3.3 does not provide the most general conditions on the off-diagonal entries of a tridiagonal matrix with 2-periodic (or even constant) main diagonal under which the numerical range is elliptical. On the other hand, for $n > 3$ some restrictions on the off-diagonal entries are in order for the numerical range to be of an elliptical shape. The next example illustrates this point for $n = 4$.

Example 4.3 Let $A =$

$$\begin{pmatrix} 1+i & 2+2i & 0 & 0 \\ -i & 0 & i & 0 \\ 0 & -1+i & 1+i & -3 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Then A is a 4×4 tridiagonal matrix with 2-periodic main diagonal. The numerical range of A is not an ellipse and is shown in Fig. 2.

[2]After the article was submitted, the authors noticed that Theorem 4.2 was proved earlier in [1, Lemma 2.1], and its proof also used the results of [9]. Thus, we omit the proof of Theorem 4.2.

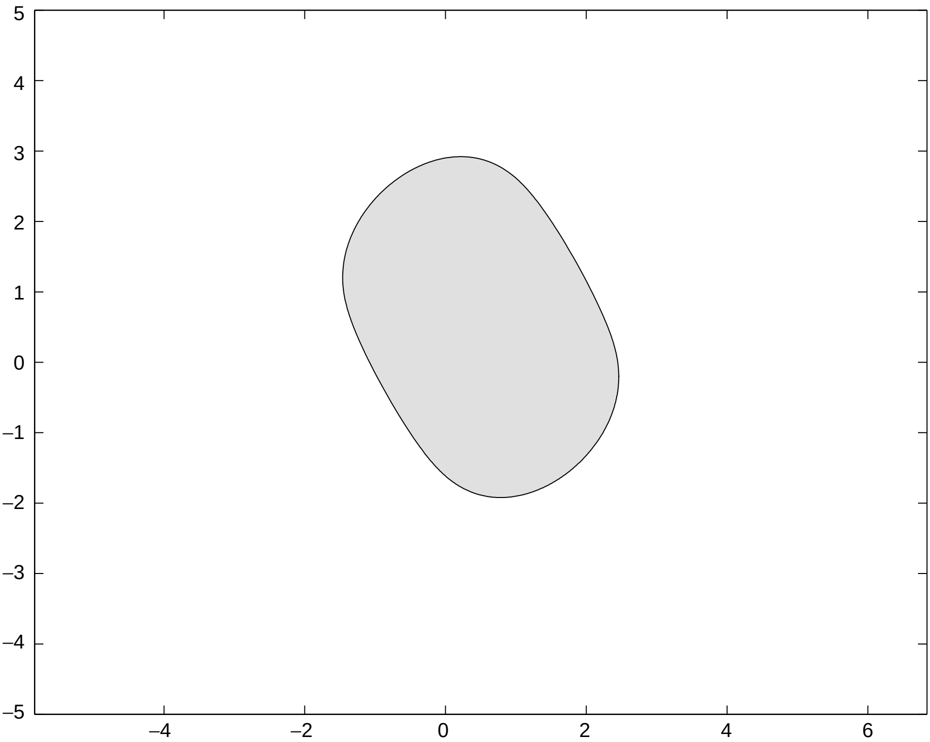

FIGURE 2 The nonelliptical numerical range of a 4×4 matrix with 2-periodic main diagonal.

Yet another case of a tridiagonal 3×3 matrix with an elliptical numerical range is delivered by the following theorem. Note that for the matrices considered there the main diagonal is not 2-periodic unless it is constant.

THEOREM 4.4 *Let*

$$A = \begin{pmatrix} a_1 & b & 0 \\ c & a_2 & b \\ 0 & c & a_3 \end{pmatrix}.$$

If $a_2 = (a_1 + a_3)/2$, *then* $W(A)$ *is an ellipse.*

Proof Without loss of generality, translate the matrix to

$$A = \begin{pmatrix} w & b & 0 \\ c & 0 & b \\ 0 & c & -w \end{pmatrix},$$

where $w = (a_1 - a_3)/2$. The spectrum of this matrix is set $\sigma(A) = \{0, \pm\sqrt{w^2 + 2bc}\}$. The parameter d defined by (4.2) is zero if and only if the matrix A is normal, in which case its numerical range is the line segment with the endpoints $\pm\sqrt{w^2 + 2bc}$, that is, a degenerate ellipse. Otherwise, $d > 0$, and we recourse to Theorem 4.1.

Computing λ by the formula (4.3):

$$\begin{aligned}\lambda &= \mathrm{tr}(A) + \frac{1}{d}\left(\sum_{j=1}^{3} |\lambda_j|^2\lambda_j - \mathrm{tr}(A^*A^2)\right)\\ &= 0 + (1/d)\Big(0 + \big|w^2 + 2bc\big|\Big(\sqrt{w^2+2bc} + (-\sqrt{w^2+2bc})\Big) - 0\Big)\\ &= 0 \in \sigma(A).\end{aligned}$$

Now observe that

$$(|\lambda_1| + |\lambda_2|)^2 - |\lambda_1 - \lambda_2|^2 = (|\lambda_1| + |-\lambda_1|)^2 - \big|\lambda_1 - (-\lambda_1)\big|^2 = 0.$$

Thus, condition (4.4) is satisfied as well. ■

5. FINAL REMARKS

Recall the main idea of the proof of Theorem 3.3. At its first step, a tridiagonal matrix (3.1) with a 2-periodic main diagonal was put in the form (2.1) via a permutation equivalence. Condition (3.4) combined with Lemma 3.1 was then used to make the off-diagonal blocks X, Y^* of the latter matrix satisfy condition $Y = \overline{k}X$ of Corollary 2.3. So, it seems natural to try finding conditions on the off-diagonal entries of the matrix (3.1), less restrictive than (3.4), which would still guarantee the normality of XY^*, Y^*X and thus the applicability of Theorem 2.1. The purpose of this section is to show that such a generalization is not possible.

Since the matrices X, Y arising in consideration of (3.1) are bidiagonal (with only i, i and $i+1, i$ entries possibly different from zero), the products XY^* and Y^*X are tridiagonal. Thus, the following normality criterion is useful.

Lemma 5.1 *Suppose T is an $n \times n$ tridiagonal matrix of the form* (3.1). *Then T is normal if and only if*

$$\big|b_j\big| = \big|c_j\big| \quad \text{for all } j = 1, \ldots, n-1, \tag{5.1}$$

$$\arg b_j + \arg c_j \ \ \textit{does not depend on } j \textit{ for all consecutive } j \textit{ such that } b_j \neq 0, \tag{5.2}$$

and

$$2\arg(a_{j+1} - a_j) = \arg b_j + \arg c_j \quad \textit{whenever } a_j \neq a_{j+1}. \tag{5.3}$$

Proof Condition (5.1) is necessary and sufficient for the respective diagonal entries of TT^* and T^*T to coincide. Comparing the off-diagonal entries we see that T is normal if and only if, on top of (5.1),

$$b_j(\overline{a_{j+1}} - \overline{a_j}) = \overline{c_j}(a_{j+1} - a_j) \quad \text{and} \quad b_j\overline{c_{j+1}} = \overline{c_j}b_{j+1}, \quad j = 1, \ldots, n-1. \tag{5.4}$$

Observe that the absolute values of the left- and right-hand sides in (5.4) are the same due to (5.1). Equating their arguments (when absolute values are different from zero) yields (5.2) and (5.3). ■

THEOREM 5.2 *Let*

$$X = \begin{pmatrix} x_1 & & & & \\ \xi_1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \xi_{j-1} & x_j & \\ & & & \ddots & \ddots \end{pmatrix}, \quad Y = \begin{pmatrix} \overline{y_1} & & & & \\ \overline{\eta_1} & \ddots & & & \\ & \ddots & \ddots & & \\ & & \overline{\eta_{j-1}} & \overline{y_j} & \\ & & & \ddots & \ddots \end{pmatrix}$$

be lower triangular bidiagonal matrices of the same size such that their respective diagonal entries x_j and y_j, as well as the respective subdiagonal entries ξ_j and η_j, do not simultaneously equal zero. Then XY^* and Y^*X are normal if and only if one of the matrices X, Y is a scalar multiple of the other.

Note that for our purposes, the condition of non-simultaneous vanishing of the respective elements of X, Y is quite natural: if it does not hold, the tridiagonal matrix A from which X and Y arise, as in the proof of Theorem 3.3, becomes a direct sum of tridiagonal matrices of smaller sizes.

Proof The comments preceding Corollary 2.3 show sufficiency; let us prove that necessity holds as well. For the sake of definiteness, consider the case of square (say, $n \times n$) matrices X, Y. The consideration of rectangular matrices differs only by minor details.

Direct computation shows that

$$XY^* = \begin{pmatrix} x_1y_1 & x_1\eta_1 & & \\ y_1\xi_1 & x_2y_2 + \xi_1\eta_1 & \ddots & \\ & \ddots & \ddots & x_{n-1}\eta_{n-1} \\ & & y_{n-1}\xi_{n-1} & x_ny_n + \xi_{n-1}\eta_{n-1} \end{pmatrix}$$

and

$$Y^*X = \begin{pmatrix} x_1y_1 + \xi_1\eta_1 & x_2\eta_1 & & \\ y_2\xi_1 & \ddots & \ddots & \\ & \ddots & x_{n-1}y_{n-1} + \xi_{n-1}\eta_{n-1} & x_n\eta_{n-1} \\ & & y_n\xi_{n-1} & x_ny_n \end{pmatrix}.$$

Being normal, the matrices XY^* and Y^*X satisfy, in particular, condition (5.1) of Lemma 5.1. Thus,

$$|x_j\eta_j| = |y_j\xi_j| \quad \text{and} \quad |x_{j+1}\eta_j| = |y_{j+1}\xi_j|, \quad j = 1, \ldots, n-1. \tag{5.5}$$

Since ξ_j and η_j cannot equal zero simultaneously, (5.5) implies that $\xi_j = 0$ if and only if $x_j = 0$ if and only if $x_{j+1} = 0$. But then either all the entries x_j, ξ_j are different from zero, or $X = 0$. Since in the latter case the statement of the theorem holds trivially, it suffices to consider the former case only.

Similarly, without loss of generality all the entries y_j, η_j are different from zero. Then (5.5) can be rewritten as

$$\left|\frac{y_{j+1}}{x_{j+1}}\right| = \left|\frac{y_j}{x_j}\right| = \left|\frac{\eta_j}{\xi_j}\right|, \quad j = 1, \dots, n-1.$$

Consequently, there exists a positive constant c such that

$$|y_j| = c|x_j| \quad (j = 1, \dots, n), \qquad |\eta_j| = c|\xi_j| \quad (j = 1, \dots, n-1). \tag{5.6}$$

In its turn, applying condition (5.2) to matrices XY^*, Y^*X we obtain:

$$\phi_j + \psi_j = \alpha \quad \text{and} \quad \phi_{j+1} + \psi_j = \beta, \quad j = 1, \dots, n-1, \tag{5.7}$$

where

$$\phi_j = \arg x_j + \arg y_j, \qquad \psi_j = \arg \xi_j + \arg \eta_j,$$

and the constants α, β do not depend on j.

On the other hand, recall that normality of XY^* and Y^*X means the existence of diagonal matrices Λ, Σ and unitary matrices U and V such that $X = V\Lambda U$, $Y = V\Sigma U$. But then $X^*X = U^*|\Lambda|^2U$ commutes with $Y^*Y = U^*|\Sigma|^2U$. In other words, the matrix $X^*X + iY^*Y$ is normal. It is also tridiagonal, because

$$X^*X = \begin{pmatrix} |x_1|^2+|\xi_1|^2 & x_2\overline{\xi_1} & & & \\ \overline{x_2}\xi_1 & \ddots & \ddots & & \\ & \ddots & |x_{n-1}|^2+|\xi_{n-1}|^2 & x_n\overline{\xi_{n-1}} \\ & & \overline{x_n}\xi_{n-1} & |x_n|^2 \end{pmatrix}$$

and

$$Y^*Y = \begin{pmatrix} |y_1|^2+|\eta_1|^2 & \overline{y_2}\eta_1 & & \\ y_2\overline{\eta_1} & \ddots & \ddots & \\ & \ddots & |y_{n-1}|^2+|\eta_{n-1}|^2 & \overline{y_n}\eta_{n-1} \\ & & y_n\overline{\eta_{n-1}} & |y_n|^2 \end{pmatrix}.$$

Condition (5.1) applied to the matrix $X^*X + iY^*Y$ means that

$$\left|x_{j+1}\overline{\xi_j} + i\,\overline{y_{j+1}}\eta_j\right| = \left|x_{j+1}\overline{\xi_j} - i\,\overline{y_{j+1}}\eta_j\right|.$$

But this is possible only if the complex numbers $x_{j+1}\overline{\xi_j}$ and $\overline{y_{j+1}}\eta_j$ lie on the same line with zero. In terms of ϕ_j, ψ_j:

$$\phi_{j+1} = \psi_j \ (\mathrm{mod}\pi), \quad j = 1, \dots, n-1. \tag{5.8}$$

Similarly, from the normality of $XX^* + iYY^*$ it follows that

$$\phi_j = \psi_j \ (\mathrm{mod}\pi), \quad j = 1, \dots, n-1. \tag{5.9}$$

Comparing equalities (5.7), (5.8) and (5.9), we see that either (i) $\beta = \alpha$ and

$$\phi_1 = \cdots = \phi_n (= \phi), \quad \psi_1 = \cdots = \psi_{n-1} (= \psi), \tag{5.10}$$

or (ii) $\beta = \alpha + \pi$ and $\{\phi_j\}$, $\{\psi_j\}$ are arithmetic progressions with the differences $\pm\pi$. In situations (i) and (ii), all the numbers $x_j y_j$ and $\xi_j \eta_j$ lie on the same line passing through zero. But then all the diagonal entries of the matrices XY^*, Y^*X lie on this same line, and so do their pairwise differences.

Keeping this in mind, invoke now part (5.3) of Lemma 5.1. When applied to normal matrices XY^*, Y^*X, it yields

$$\phi_j + \psi_j = \phi_{j+1} + \psi_j = 2\phi_j \ (\mathrm{mod}2\pi).$$

So, only the case (i) is possible and, moreover, in (5.10) ϕ is the same as ψ.

Thus, not only (5.6) holds, but in fact

$$\overline{y_j} = ce^{-i\phi}x_j, \quad \overline{\eta_j} = ce^{-i\phi}\xi_j.$$

In other words, $Y = \mu X$ for $\mu = ce^{-i\phi}$.

Acknowledgements

This research was conducted during the summer of 2002 at the College of William and Mary's research Experience for Undergraduates program and was supported by NSF REU grant DMS-99-87803 and NSF grant DMS-99-88579. The second author was also supported by the Semester Research grant from the College of William and Mary. The authors thank the referee for the valuable comments. The first author would also like to thank Christopher Hillar of the University of California, Berkeley, for his useful discussions during the Research.

References

[1] M. Adam and J. Maroulas (2002). The joint numerical range of bordered and tridiagonal matrices, Linear operators and matrices. *Oper. Theory Adv. Appl.*, Vol. 130, pp. 29–41. Birkhäuser, Basel.

[2] Mao-Ting Chien (1996). On the numerical range of tridiagonal operators. *Linear Algebra Appl.*, **246**, 203–214. MR **97i**:47007

[3] Mao-Ting Chien and Jun-Ming Huang (2001). Numerical range of a continuant matrix. *Appl. Math. Lett.*, **14**(2), 213–216. MR **2001j**:15015

[4] C. Cowen and E. Harel (1995). An effective algorithm for computing the numerical range. http://www.math.purdue.edu/˜cowen/NumRange.html.

[5] M. Eiermann (1993). Fields of values and iterative methods. *Linear Algebra Appl.*, **180**, 167–197.

[6] K.E. Gustafson and D.K.M. Rao (1997). Numerical range. *The Field of Values of Linear Operators and Matrices*. Springer, New York.

[7] R.A. Horn and C.R. Johnson (1985). *Matrix Analysis*. Cambridge University Press, Cambridge.

[8] R.A. Horn and C.R. Johnson (1991). *Topics in Matrix Analysis*. Cambridge University Press, Cambridge.

[9] D.Keeler, L. Rodman and I. Spitkovsky (1997). The numerical range of 3×3 matrices. *Linear Algebra Appl.*, **252**, 115–139.

[10] R. Kippenhahn (1951). Über den Wertevorrat einer Matrix. *Math. Nachr.*, **6**, 193–228 (in German).

[11] Chi-Kwong Li (1996). A simple proof of the elliptical range theorem. *Proc. Amer. Math. Soc.*, **124**(7), 1985–1986.

[12] H. Shapiro (1982). A conjecture of Kippenhahn about the characteristic polynomial of a pencil generated by two Hermitian matrices. II. *Linear Algebra Appl.*, **45**, 97–108.

[13] O. Toeplitz (1918). Das algebraische Analogon zu einen Satz von Fejér. *Math. Z.*, **2**, 187–197.

[14] Shu-Hsien Tso and Pei Yuan Wu (1999). Matricial ranges of quadratic operators. *Rocky Mountain J. Math.*, **29**(3), 1139–1152.