On some 4-by-4 matrices with bi-elliptical numerical ranges

Titas Geryba<sup>a</sup> and Ilya M. Spitkovsky<sup>b</sup>

<sup>a</sup>Mathematisch Instituut, Universiteit Leiden, Leiden, Netherlands; <sup>b</sup>Division of Science and Mathematics, New York University Abu Dhabi (NYUAD), Abu Dhabi, United Arab Emirates

**ABSTRACT**

A complete description of 4-by-4 matrices $\begin{bmatrix} \alpha & C \\ D & \beta \end{bmatrix}$, with scalar 2-by-2 diagonal blocks, for which the numerical range is the convex hull of two non-concentric ellipses is given. This result is obtained by reduction to the leading special case in which $C - D^*$ also is a scalar multiple of the identity. In particular cases when in addition $\alpha - \beta$ is real or pure imaginary, the results take an especially simple form. An application to reciprocal matrices is provided.

**1. Introduction**

The numerical range of a matrix $A \in \mathbb{C}^{n \times n}$ is defined as

$$W(A) = \{ \langle Ax, x \rangle : x \in \mathbb{C}^n, \|x\| = 1 \}.$$ 

We are using the standard notation $\langle ., . \rangle$ for the inner product on the $n$-dimensional space $\mathbb{C}^n$ and $\| . \|$ for the norm associated with it: $\|x\|^2 = \langle x, x \rangle$.

It is well known that $W(A)$ is a convex compact subset of $\mathbb{C}$ containing the spectrum $\sigma(A)$ of $A$, and thus its convex hull $\text{conv} \ \sigma(A)$. For normal matrices $A$, the equality $W(A) = \text{conv} \ \sigma(A)$ holds. On the other hand, for non-normal $A \in \mathbb{C}^{2 \times 2}$ the numerical range is an elliptical disk, with the foci at the eigenvalues. These and other by now classical properties of $W(A)$ can be found, e.g. in [1, Chapter 1] or more recent [2, Chapter 6].

The shape of $W(A)$ for $A \in \mathbb{C}^{3 \times 3}$ is also known, see [3] (or its translation [4] into English) for the classification and [5] for the pertinent tests. However, for $n \geq 4$ many questions remain open. In particular, it is of interest when the boundary $\partial W(A)$ of the numerical range contains an elliptical arc [6].

A useful tool in studying properties of $W(A)$ is the so called *Numerical range (NR) generating curve* $C(A)$, also introduced in [3] – the *Kippenhahn curve* in the terminology of [2, Chapter 13], where a very lucid and detailed description of $C(A)$ is given. This curve is defined uniquely by having exactly $n$ tangent lines in each direction $e^{i\theta}$, intercepting the family of orthogonal lines at the eigenvalues $\lambda_j(\theta)$ of $\text{Im}(e^{-i\theta}A)$. As it happens,
$W(A) = \text{conv } C(A)$, and so $\partial W(A)$ consists of some arcs of $C(A)$ possibly connected by line segments.

We will denote the characteristic polynomial of $\text{Im}(e^{-i\theta}A)$ by $P_A(\lambda, \theta)$ and call it the \textit{NR generating polynomial} of $A$. From the description of $C(A)$ it follows in particular that it contains an ellipse if and only if $P_A(\lambda, \theta)$ is divisible by a polynomial quadratic in $\lambda$. More specifically (as can be established by direct computations similar to those carried out say in [7]), an ellipse centred at $p + i q$ ($p, q \in \mathbb{R}$) corresponds to a factor of the form

$$(\lambda + p \sin \theta - q \cos \theta)^2 + x \cos 2\theta + y \sin 2\theta - z$$

(1)

with some $x, y, z \in \mathbb{R}$ satisfying $z \geq \sqrt{x^2 + y^2}$. Note that if $z = \sqrt{x^2 + y^2}$, then the quadratic (1) factors further into two linear functions in $\lambda$, and the ellipse in question degenerates into the doubleton of its foci.

If $n = 4$ and $P_A(\lambda, \theta)$ is divisible by (1), then the quotient is of the same type. Consequently, for $A \in \mathbb{C}^{4 \times 4}$ the boundary of $W(A)$ contains an elliptical arc if and only if $C(A)$ consists of two ellipses, one of which is possibly degenerate. So, $\partial W(A)$ contains an elliptical arc if and only if $W(A)$ is an elliptical disk, the convex hull of two ellipses, or the convex hull of an ellipse and one or two points (the latter being an option only if $A$ is unitarily reducible). The respective criteria were established in [8]. However, these criteria are not stated in terms of $A$ directly and therefore not easy to verify. More can be done if $A$ enjoys some additional structure.

This paper is devoted to 4-by-4 matrices of the form

$$A = \begin{bmatrix} \alpha I & C \\ D & \beta I \end{bmatrix}. \quad (2)$$

In [9, Section 4], we have provided necessary and sufficient conditions for such matrices to have $W(A)$ in the shape of an elliptical disk or the convex hull of two concentric ellipses. Here we treat the remaining case, when $W(A)$ is the convex hull of two ellipses with distinct centres. For convenience of reference, this shape is called \textit{bi-elliptical} in what follows.

The special case of matrices (2) with $C - D^*$ being a scalar multiple of the identity is considered in Section 3, after some preliminary technicalities disposed of in Section 2. Under additional restrictions on $\alpha - \beta$, this result is further simplified in Section 4 which also contains several numerical examples. As it happens, the general case can be reduced to the one tackled in Section 3; this reduction is carried over in Section 5. The final Section 6 contains an application to so called reciprocal matrices.

2. Preliminary results

Passing from $A$ to $A - \frac{\alpha + \beta}{2} I$, we may (and will, in what follows) without loss of generality suppose that in (2) $\beta = -\alpha$. As in [9], we will also use the notation

$$H = C^* C + DD^*, \quad Z = DC. \quad (3)$$

Lemma 2.1: Let $A$ be given by (2), with $\beta = -\alpha$. Then:

(i) The eigenvalues of $A$ are $\pm \sigma_1, \pm \sigma_2$, where $\sigma_j = \sqrt{z_j + \alpha^2}$ and $z_1, z_2$ are the eigenvalues of $Z$,
(ii) The eigenvalues of \( \text{Im}(e^{-i\theta} A) \) are \( \pm \lambda_j(\theta) \), where

\[
\lambda_j(\theta) = \sqrt{\text{Im}(e^{-i\theta} \alpha)^2 + \mu_j(\theta)/4} \quad (j = 1, 2),
\]

and \( \mu_j(\theta) \) are the eigenvalues of

\[
(e^{i\theta} C^* - e^{-i\theta} D) (e^{-i\theta} C - e^{i\theta} D^*) = H - 2 \text{Re}(e^{-2i\theta} Z);
\]

(iii) The NR generating polynomial \( P_A \) of \( A \) is

\[
\Xi_1(\theta) = \frac{1}{4} \text{Tr} H + |\alpha|^2 - \left( \frac{1}{2} \text{Re Tr} Z + \text{Re} \alpha^2 \right) \cos 2\theta
\]

\[
- \left( \frac{1}{2} \text{Im Tr} Z + \text{Im} \alpha^2 \right) \sin 2\theta,
\]

\[
16\Xi_2(\theta) = 6|\alpha|^4 + 2|\alpha|^2 \text{Tr} H + 2 \text{Re}(\alpha^2 \text{Tr} Z) + \frac{1}{2} \text{Tr}(H^2) - \frac{1}{2} \text{Tr}(Z^2)
\]

\[
- \text{Tr}(ZZ^*) - \text{Re} \zeta_1 \cos 2\theta - \text{Im} \zeta_1 \sin 2\theta + \text{Re} \zeta_2 \cos 4\theta + \text{Im} \zeta_2 \sin 4\theta,
\]

with \( \zeta_1, \zeta_2 \) given by

\[
\zeta_1 = 8|\alpha|^2 \alpha^2 + 2\alpha^2 \text{Tr} H + 4|\alpha|^2 \text{Tr} Z + 2 \text{Tr} Z \text{Tr} H - 2 \text{Tr} ZH,
\]

\[
\zeta_2 = 2 \det Z + 2\alpha^2 \text{Tr} Z + 2\alpha^4.
\]

**Proof:** For (i) and (ii), the result follows by using Schur complement formula when computing the respective characteristic polynomials; the pertinent computation for (ii) is actually contained in the proof of [9, Lemma 2.1]. Expanding \( P_A(\lambda, \theta) = (\lambda^2 - \lambda_1^2(\theta))(\lambda^2 - \lambda_2^2(\theta)) \), we derive (iii) from (ii). \( \blacksquare \)

Observe that \( P_A \) is an even function of \( \lambda \). This agrees with the result of [9] for matrices (2) with arbitrary block sizes, implying that \( C(A) \) is symmetric with respect to the origin. From here we immediately obtain

**Proposition 2.1:** Suppose \( A \in \mathbb{C}^{4 \times 4} \) of the form (2) is such that \( C(A) \) consists of two non-concentric ellipses. Then, for an appropriate choice of signs of \( \sigma_j \), \( \partial W(A) \) contains a pair of parallel line segments coinciding in length and direction with \( \sigma_1 + \sigma_2 \).

**Proof:** Let \( C(A) \) consist of the ellipses \( E_1, E_2 \). Due to the central symmetry of \( C(A) \), either both \( E_1, E_2 \) also are symmetric with respect to the origin, or \( E_1(:= E) = -E_2 \neq -E \). The former case is excluded, because otherwise \( E_1, E_2 \) would be concentric. Furthermore, the foci of \( \pm E \) are the eigenvalues of \( A \), and so (relabelling \( z_1, z_2 \) if needed, and choosing the square roots signs appropriately) we may suppose that \( \pm E \) are the foci of \( E \).

Consider now the composition \( S \) of two symmetries, one with respect to the origin and the other with respect to the centre of \( E \). By its construction, \( S \) is a shift and, since
\( S(-(\sigma_1 + \sigma_2)/2) = (\sigma_1 + \sigma_2)/2 \), it is the shift by \( \sigma_1 + \sigma_2 \). So,
\[
E = S(-E) = -E + (\sigma_1 + \sigma_2)
\]

The flat portions on the boundary of \( W(A) \) are therefore the common tangents of \( E \) and \(-E\), and the endpoints of each differ by \( \sigma_1 + \sigma_2 \).

\[ \square \]

3. Leading special case

Let
\[
A = \begin{bmatrix}
\alpha I & B^* + I \\
B - I & -\alpha I
\end{bmatrix}.
\]

(7)

This is a particular case of (2) in which \( C = B^* + I \) and \( D = B - I \). Respectively, (3) takes the form
\[
H = 2(BB^* + I), \quad Z = BB^* - I + 2i \text{Im} B.
\]

(8)

Let us denote the eigenvalues of \( \text{Im} B \) by \( \beta_1, \beta_2 \) while keeping the notation \( \pm \sigma_j \) for the eigenvalues of \( A \). Let us also write \( \alpha \) as \( u + iv \) \((u, v \in \mathbb{R})\).

**Theorem 3.1:** For \( A \) as in (7), \( W(A) \) is bi-elliptical if and only if \( B \) is not normal and
\[
(1 + v^2)(\sigma_1 + \sigma_2)^2 = (\beta_1 - \beta_2)^2.
\]

(9)

Both in the proof of Theorem 3.1 and its further application, putting block \( B \) in an upper triangular form
\[
B = \begin{bmatrix}
b_1 & b \\
0 & b_2
\end{bmatrix}
\]

(10)

and rewording conditions in terms of its entries proves to be useful. This can be done by a block diagonal unitary similarity of \( A \), preserving its structure (7). Moreover, it can be arranged that \( b \geq 0 \). So, without loss of generality
\[
A = \begin{bmatrix}
\alpha & 0 & \overline{b_1} + 1 & 0 \\
0 & \alpha & b & \overline{b_2} + 1 \\
b_1 - 1 & b & -\alpha & 0 \\
0 & b_2 - 1 & 0 & -\alpha
\end{bmatrix}.
\]

(11)

Let us also denote \( \text{Re} b_j = \xi_j \), \( \text{Im} b_j = \eta_j \), \( j = 1, 2 \).
Plugging in the values of \( \sigma_j \) from Lemma 2.1, condition (9) can be rewritten as

\[
\Tr Z + 2\alpha^2 + 2\sqrt{\det Z} + \alpha^2 \Tr Z + \alpha^4 = 4p^2,
\]

where

\[
\sqrt{(\eta_1 - \eta_2)^2 + b^2} \quad \frac{1}{1 + v^2} := 2p.
\]

Rewriting (12) as

\[
2\sqrt{\det Z} + \alpha^2 \Tr Z + \alpha^4 = 4p^2 - \Tr Z - 2\alpha^2
\]

and taking the square, we find that (9) is equivalent to \( T = 0 \), where

\[
T := 16p^4 - (8\Tr Z + 16\alpha^2)p^2 + (\Tr Z)^2 - 4\det Z.
\]

This condition is in turn equivalent to both \( \Re T \) and \( \Im T \) being equal to zero. For future use observe therefore that

\[
\Re T = (4p^2 - (b^2 + |b_1|^2 + |b_2|^2))^2 - 16u^2p^2 - 4|b_1|^2|b_2|^2,
\]

\[
\Im T = 16v(\nu(\eta_1 + \eta_2) - 2u)p^2 + 4(\xi_1^2 - \xi_2^2)(\eta_1 - \eta_2).
\]

**Proof of Theorem 3.1:** Necessity. If \( B \) is normal then it follows from (8) that so is \( Z \). Moreover, \( Z \) and \( H \) commute. This situation falls under the setting of [9, Theorem 4.1], according to which \( W(A) \) is the convex hull of two ellipses, but these ellipses are concentric. So, \( B \) is not normal.

After yet another unitary (this time, permutational) similarity, the matrix (11) becomes

\[
A_0 = \begin{bmatrix}
\alpha & b_1 + 1 & 0 & 0 \\
\overline{b_1 - 1} & -\alpha & b & 0 \\
0 & b & \alpha & \overline{b_2 + 1} \\
0 & 0 & \overline{b_2 - 1} & -\alpha
\end{bmatrix}.
\]

This matrix is tridiagonal. Moreover, \( b \neq 0 \) since \( B \) is not normal. In terminology of [10] it implies that \( A_0 \) is proper tridiagonal (meaning that entries in the positions \((j-1,j)\) and \((j,j-1)\) cannot simultaneously equal zero). Invoking [10, Theorem 10], we conclude that the only flat portions of \( \partial W(A) \) are horizontal. Their \( y \)-coordinates are determined by the eigenvalues of \( \Im A_0 \), which is the direct sum of two copies of \( \begin{bmatrix} \nu & -i \\ i & \nu \end{bmatrix} \), and therefore equal \( \pm \sqrt{1 + v^2} \). The lengths of these portions are equal to the spread of the compression \( H_0 \) of \( \Re A_0 \) onto the (2-dimensional) subspace generated by the eigenvectors of \( \Im A_0 \) corresponding to either of its eigenvalues. Denoting \( f = \sqrt{1 + v^2} - \nu \) and \( k = 1/\sqrt{1 + f^2} \), an orthonormal basis in one of these subspaces can be chosen as

\[
k[-i,f,0,0]^T \quad \text{and} \quad k[0,0,-i,f]^T.
\]

The matrix of \( H_0 \) in this basis is

\[
\frac{1}{2\sqrt{1 + v^2}} \begin{bmatrix}
u(f^{-1} - f) + 2\eta_1 & -ib \\
ib & \nu(f^{-1} - f) + 2\eta_2
\end{bmatrix},
\]

and the spread of this matrix equals \( 2p \) defined by (13). It remains to observe that \( (\beta_1 - \beta_2)^2 = (\eta_1 - \eta_2)^2 + b^2 \) and to make use of Proposition 2.1.
Sufficiency. Suppose that (9) holds. We will now show that under this condition the NR generating polynomial of $A$ factors as

$$P_A(\lambda, \theta) = ((\lambda + p \sin \theta)^2 + \Omega(\theta)) (\lambda - p \sin \theta)^2 + \Omega(\theta)),$$

where $\Omega(\theta) = x \cos 2\theta + y \sin 2\theta - z$ for some appropriate choice of real parameters $x, y, z$.

According to Lemma 2.1(iii), (17) is equivalent to

$$2 \left( p^2 \sin^2 \theta - \Omega(\theta) \right) = \Xi_1(\theta), \quad \left( p^2 \sin^2 \theta + \Omega(\theta) \right)^2 = \Xi_2(\theta),$$

where $\Xi_1$ and $\Xi_2$ are given by (5) and (6), respectively. The first equality in (18) defines $x, y, z$ uniquely as

$$x = \frac{1}{4} \Re \text{Tr} Z + \frac{1}{2} \Re(\alpha^2) - \frac{p^2}{2},$$
$$y = \frac{1}{4} \Im \text{Tr} Z + \frac{1}{2} \Im(\alpha^2),$$
$$z = \frac{1}{8} \text{Tr} H + \frac{|\alpha|^2}{2} - \frac{p^2}{2},$$

or, in terms of $B$ explicitly:

$$x = \frac{1}{4} (b^2 + |b_1|^2 + |b_2|^2 - 2) + \frac{1}{2} \Re(\alpha^2) - \frac{b^2 + (\eta_1 - \eta_2)^2}{1 + v^2},$$
$$y = \frac{1}{2} (\eta_1 + \eta_2) + \frac{1}{2} \Im(\alpha^2),$$
$$z = \frac{1}{4} (b^2 + |b_1|^2 + |b_2|^2 + 2) + \frac{1}{2} |\alpha|^2 - \frac{b^2 + (\eta_1 - \eta_2)^2}{1 + v^2}.$$

Incidentally, with this choice of $x, y, z$, the second equality in (18) also holds. Here is the pertinent chain of computations. First, form (19):

$$4 \left[ p^2 \sin^2(\theta) + \Omega(\theta) \right] = -\frac{1}{2} \text{Tr} H - 2|\alpha|^2 + 4p^2$$
$$+ \left[ \Re \text{Tr} Z + 2 \Re(\alpha^2) - 4p^2 \right] \cos 2\theta$$
$$+ \left[ \Im \text{Tr} Z + 2 \Im(\alpha^2) \right] \sin 2\theta.$$

Squaring we obtain:

$$16 \left[ p^2 \sin^2(\theta) + \Omega(\theta) \right]^2$$
$$= 24p^4 - 4p^2 \left( \Re(\text{Tr} Z + 2\alpha^2) + \text{Tr} H + 4|\alpha|^2 \right)$$
$$+ \frac{1}{2} |\text{Tr} Z + 2\alpha^2|^2 + \left( \frac{1}{2} \text{Tr} H + 2|\alpha|^2 \right)^2$$
$$- 8 \Re \left[ \left( 2p^2 - \frac{1}{2} \text{Tr} Z - \alpha^2 \right) \left( 2p^2 - \frac{1}{4} \text{Tr} H - |\alpha|^2 \right) \right] \cos 2\theta.$$
\[
-8 \text{Im} \left[ \left( 2p^2 - \frac{1}{2} \text{Tr } Z - \alpha^2 \right) \left( 2p^2 - \frac{1}{4} \text{Tr } H - |\alpha|^2 \right) \right] \sin 2\theta \\
+ 2 \text{Re} \left[ \left( 2p^2 - \frac{1}{2} \text{Tr } Z - \alpha^2 \right)^2 \right] \cos 4\theta \\
+ 2 \text{Im} \left[ \left( 2p^2 - \frac{1}{2} \text{Tr } Z - \alpha^2 \right)^2 \right] \sin 4\theta.
\]

Plugging in \( \Xi_2 \) from (6):

\[
32 \left[ (p^2 \sin^2 \theta + \Omega(\theta))^2 - \Xi_2(\theta) \right] \\
= 3 \left[ 16p^4 - \frac{8}{3} \left( \text{Re}(\text{Tr } Z + 2\alpha^2) + \text{Tr } H + 4|\alpha|^2 \right) p^2 \\
+ \frac{2}{3} \text{Tr}(Z^*Z) - \frac{1}{3} |\text{Tr } Z|^2 + \frac{1}{3} \text{Tr}(H^2) - \frac{1}{6} \text{Tr}(H)^2 \right] \\
- 4 \text{Re} \left[ 16p^4 - 2(\text{Tr } H + 2 \text{Tr } Z + 4\alpha^2 + 4|\alpha|^2)p^2 + \text{Tr } ZH - \frac{1}{2} \text{Tr } Z \text{Tr } H \right] \cos 2\theta \\
- 2 \text{Im} \left[ -8(\text{Tr } Z + 2\alpha^2)p^2 + 2 \text{Tr } ZH - \text{Tr } Z \text{Tr } H \right] \sin 2\theta \\
+ \text{Re} \left[ 16p^4 - 8(\text{Tr } Z + 2\alpha^2)p^2 + (\text{Tr } Z)^2 - 4 \text{det}(Z) \right] \cos 4\theta \\
+ \text{Im} \left[ -8(\text{Tr } Z + 2\alpha^2)p^2 + (\text{Tr } Z)^2 - 4 \text{det}(Z) \right] \sin 4\theta.
\]

Rewriting the right-hand side of (21) in terms of the entries of \( B \) we obtain, with the use of (15)–(16):

\[
(p^2 \sin^2 \theta + \Omega(\theta))^2 - \Xi_2(\theta) \\
= \frac{1}{32} \left[ 3 \text{Re } T - 4 \text{Re } T \cos 2\theta - 2 \text{Im } T \sin 2\theta + \text{Re } T \cos 4\theta + \text{Im } T \sin 4\theta \right].
\]

So, condition (9), equivalent to \( T = 0 \), indeed implies the second equality in (18).

Finally, from (20): 

\[
z^2 - x^2 - y^2 = (1 + v^2)^2 + 2(1 + v^2)x - y^2 \\
= \frac{(1 + v^2)(\xi_1^2 + \xi_2^2)}{2} + \frac{(1 + 2v^2)b^2}{4} \\
+ \left[ u - \frac{(\eta_1 + \eta_2)v}{2} \right]^2 + \frac{v^2(\eta_1 - \eta_2)^2}{4} > 0
\]

since \( b \neq 0 \). Factorization (17) therefore generates \( C(A) \) consisting of two non-degenerate ellipses.

\[\blacksquare\]

4. \textbf{Follow up observations and examples}

Suppose that in (7) the parameter \( \alpha \) is real or pure imaginary. Criterion established in Section 3 can then be recast explicitly in terms of \( B \). This is done in two theorems below,
stated for $B$ as in (10). Note however that the results can be easily reworded without putting $B$ in a triangular form. Indeed, $\{b_1, b_2\}$ is the spectrum of $B$, while $b^2 = ||B||_F - |b_1|^2 - |b_2|^2$, with $||.||_F$ denoting the Frobenius norm.

**Theorem 4.1:** Let $\alpha = u \in \mathbb{R}$. Then the numerical range of matrix (11) is bi-elliptical if and only if $b \neq 0$ and either

(i) $\eta_1 = \eta_2$ and $4b^2u^2 = (\xi_1^2 - \xi_2^2)^2$, or (ii) $u = \xi_1 = \xi_2 = 0$.

**Proof:** Since $v = 0$, according to (16) $\text{Im} T = 0$ if and only if $\eta_1 = \eta_2$ or $\xi_1^2 = \xi_2^2$. On the other hand, $4p^2 = (\eta_1 - \eta_2)^2 + b^2$ due to (13), and so (15) takes the form

$$((\eta_1 - \eta_2)^2 - (|b_1|^2 + |b_2|^2))^2 - 4u^2 ((\eta_1 - \eta_2)^2 + b^2) - 4 |b_1|^2 |b_2|^2. \quad (22)$$

If $\eta_1 = \eta_2$, then (22) vanishing amounts to $(|b_1|^2 - |b_2|^2)^2 = 4u^2 b^2$, which is exactly case (i). If $\eta_1 \neq \eta_2$, denote the coinciding values of $\xi_1^2$ and $\xi_2^2$ by $\xi^2$. Expression (22) then simplifies to

$$-4\xi^2(\eta_1 - \eta_2)^2 - 4u^2 ((\eta_1 - \eta_2)^2 + b^2),$$

and so it vanishes if and only if $u = \xi = 0$. This is case (ii). \hfill \Box

**Corollary 4.1:** Let $A$ be of the form (11) with zero main diagonal. Then $W(A)$ is bi-elliptical if and only if $b \neq 0$ and either

(i) $\text{Im} b_1 = \text{Im} b_2$ and $|b_1| = |b_2|$, or (ii) $\text{Re} b_1 = \text{Re} b_2 = 0$.

Note that in [11, Theorem 3.14] the result of Corollary 4.1 was established under additional restrictions $b = \sqrt{(1 - |b_1|)^2(1 - |b_2|)^2}, |b_1|, |b_2| < 1$.

According to Corollary 4.1, for $\alpha = 0$ the value of $b$ is irrelevant (as long as it is different from zero, of course). The situation is quite the opposite for non-zero values of $\alpha$.

**Corollary 4.2:** Let $A$ be of the form (11) with $\alpha \neq 0$. Then there is at most one value of $b(> 0)$ for which $W(A)$ is bi-elliptical.

**Proof:** For $\alpha \in \mathbb{R}$ the result follows from Theorem 4.1, with an explicit expression for $b$. For non-real values of $\alpha$, observe that condition $\text{Re} T = 0$ with the use of (13) and (15) yields a quadratic equation for $b^2$ with one non-positive root, and thus a unique positive root. \hfill \Box

It is possible to derive the explicit value of $b$ for non-real $\alpha$ as well, but the expression is somewhat cumbersome. We will restrict our attention to another special case, when $\alpha$ is pure imaginary.

**Theorem 4.2:** Let $\alpha = iv \neq 0$. Then the numerical range of matrix (11) is bi-elliptical if and only if $b \neq 0$,

$$|b_1| = |b_2|, \quad \text{and} \quad v^2 b^2 = (\eta_1 - \eta_2)^2. \quad (23)$$
Proof: Setting $u = 0$ in (15) and (16), we see that conditions $\text{Re } T = 0$ and $\text{Im } T = 0$ simplify, respectively, to

$$4p^2 - b^2 = (|b_1| \pm |b_2|)^2$$

(24)

and

$$4v^2 p^2 (\eta_1 + \eta_2) + (\xi_1^2 - \xi_2^2)(\eta_1 - \eta_2) = 0.$$  

(25)

Moreover, due to (13) and (24):

$$4v^2 p^2 = (\eta_1 - \eta_2)^2 - (4p^2 - b^2) = (\eta_1 - \eta_2)^2 - (|b_1| \pm |b_2|)^2.$$  

Plugging this expression for $4v^2 p^2$ into (25), after some additional simplifications we arrive at

$$(|b_1| \pm |b_2|) (|b_1| \pm |b_2|) = 0,$$  

(26)

with the signs matching that in (24). Note that the system of conditions (24),(25) is equivalent to (24),(26), with $p$ given by (13).

The sufficiency of (23) for (9) to hold is now immediate. To prove the necessity, observe that from (13), since $b \neq 0$:

$$4p^2 - b^2 = \frac{(\eta_1 - \eta_2)^2 - v^2 b^2}{1 + v^2} < (\eta_1 - \eta_2)^2 \leq (|\eta_1| + |\eta_2|)^2 \leq (|b_1| + |b_2|)^2,$$

and therefore (24) holds with the lower sign. Then so does (26).

Suppose now that $|b_1| \neq |b_2|$. Condition (26) then boils down to $\eta_1 |b_2| = \eta_2 |b_1|$, which in turn implies that $\eta_1, \eta_2$ are of the same sign, and $\arg b_1 = \arg b_2$ or $\arg b_1 = \pi - \arg b_2$. The above inequality for $4p^2 - b^2$ can therefore be strengthened as follows:

$$4p^2 - b^2 = \frac{(\eta_1 - \eta_2)^2 - v^2 b^2}{1 + v^2} < (\eta_1 - \eta_2)^2 = (|\eta_1| - |\eta_2|)^2 \leq (|b_1| - |b_2|)^2.$$  

This is a contradiction with (24).

So, in fact $|b_1| = |b_2|$, and (24) (with the correct choice of the sign) implies $4p^2 - b^2 = 0$, thus leading to (23) (See Figure 1).
5. General case

We now return to arbitrary matrices (2), not supposing a priori the special form (7).

**Theorem 5.1:** Let $A$ be of the form (2), with the usual convention $\beta = -\alpha$. Then $W(A)$ is bi-elliptical if and only if

(i) there exists $\theta$ for which $e^{-i\theta} C - e^{i\theta} D^*$ is a scalar multiple of a unitary matrix while $(e^{-i\theta} C - e^{i\theta} D^*)D$ is not normal, and

(ii) for this value of $\theta$,

\[ 4\sqrt{\det(\text{Im}(e^{-i\theta} A))(\sigma_1(\theta) + \sigma_2(\theta))^2} = (s_1(\theta) - s_2(\theta))^2. \quad (27) \]

Here $\pm \sigma_1(\theta), \pm \sigma_2(\theta)$ are the eigenvalues of $e^{-i\theta} A$ (so that $\sigma_j(0) = \sigma_j, j = 1, 2$) and $s_1(\theta), s_2(\theta)$ are the non-repeating eigenvalues of $\text{Im}((e^{-i\theta} A)^2)$.

It will become clear from the proof of Theorem 5.1 that $\theta$ satisfying condition (i) is unique mod $\pi$. Also, if (i) holds then the eigenvalues $\pm \lambda_1(\theta), \pm \lambda_2(\theta)$ of $\text{Im}(e^{-i\theta} A)$ actually satisfy $\lambda_1(\theta) = \lambda_2(\theta)$ and, if (ii) also holds, then $\theta = \arg(\sigma_1 + \sigma_2) \mod \pi$. So, with an appropriate choice of the signs condition (27) can be rewritten as

\[ (\lambda_1(\theta) + \lambda_2(\theta)) (\sigma_1(\theta) + \sigma_2(\theta)) = s_1(\theta) - s_2(\theta). \]

Finally, condition $s_1(\theta) \neq s_2(\theta)$ follows from $(e^{-i\theta} C - e^{i\theta} D^*)D$ not being normal.

**Proof:** To establish necessity of (i), choose $e^{i\theta}$ as the direction of the line segments on $\partial W(A)$, as guaranteed by Proposition 2.1. The matrix $\text{Im}(e^{-i\theta} A)$ then must have repeated eigenvalues. According to Lemma 2.1(ii), this happens if and only if the 2-by-2 hermitian matrix defined by (4) has coinciding eigenvalues, and is therefore a scalar multiple of the identity. This, in turn, is equivalent to $e^{-i\theta} C - e^{i\theta} D^*$ being a scalar multiple of a unitary matrix.

This scalar multiple cannot be zero, since otherwise $Z = e^{2i\theta} C^* C$ while $H = 2C^* C$, and so $Z$ is a normal matrix commuting with $H$. As was already mentioned in the proof of Theorem 3.1, this situation corresponds to $W(A)$ being the convex hull of two concentric ellipses, and thus leads to a contradiction.

Both condition (27) and the shape of $W(A)$ persist under scaling of $A$. So, we may for the rest of the proof suppose that $\theta = 0$ and $C^* - D = 2W$ where $W$ is unitary. Furthermore, a unitary similarity of $A$ via $\text{diag}[I, W]$ leaves condition (27) invariant, while replacing $C$ by $C_1 = CW$ and $D$ by $D_1 = W^* D$. Since $C_1^* - D_1 = W^* (C^* - D) = 2I$, it suffices to consider the case $W = I$.

This brings us into the setting of Theorem 3.1, with $B = D_1 + I(= C_1^* - I)$. Since this $B$ is (or is not) normal simultaneously with $(e^{-i\theta} C - e^{i\theta} D^*)D$, condition (i) is indeed necessary.
Supposing that (i) holds, it only remains to show that (27) with \( \theta = 0 \) is (9) in disguise. But indeed, for \( A \) as in (7):

\[
1 + v^2 = \sqrt{\det(\text{Im}A)},
\]

while \( \text{Im}(A^2) \) is the direct sum of two copies of \( \text{Im}(\alpha^2)I + 2\text{Im}B \), and so \( s_1(0) - s_2(0) = 2(\beta_1 - \beta_2) \).

Example 5.1: Let

\[
A = \begin{bmatrix}
4 & 0 & 4 - 8i & 0 \\
0 & 4 & 5 - 5i & 4 - 12i \\
-2 + 6i & 5 - 5i & -4 & 0 \\
0 & 2 + 6i & 0 & -4
\end{bmatrix}.
\]

The matrix \( e^{-i\theta}C - e^{i\theta}D^* \) is upper triangular, and so it can be normal only if its lower left entry is also zero. This requirement is equivalent to \( \theta = -\pi/4 \mod \pi \). On the other hand, with this choice of \( \theta \) indeed \( e^{-i\theta}C - e^{i\theta}D^* \) is a scalar multiple of a unitary matrix (in this case, even the identity). Moreover, since \( D \) is not normal, the product \( (e^{-i\theta}C - e^{i\theta}D^*)D \) is not normal either. So, condition (i) of Theorem 5.1 holds.

Furthermore, since

\[
e^{i\pi/4}A = \sqrt{2} \begin{bmatrix}
2 + 2i & 0 & 6 - 2i & 0 \\
0 & 2 + 2i & 5 & 8 - 4i \\
-4 + 2i & 5 & -2 - 2i & 0 \\
0 & -2 + 4i & 0 & -2 - 2i
\end{bmatrix},
\]

somewhat lengthy but direct computations yield: \( \sigma_1(-\pi/4) + \sigma_2(-\pi/4) = 5\sqrt{2} \), \( s_1(-\pi/4) - s_2(-\pi/4) = 20\sqrt{29} \), and \( \sqrt{\det(\text{Im}(e^{i\pi/4}A))} = 58 \). So, condition (ii) of Theorem 5.1 also holds, and \( W(A) \) is bi-elliptical (See Figure 2). The bi-elliptical shape of \( W(A) \) in the setting of Theorem 5.1 means that in fact \( W(A) = W(A_1 \oplus A_2) \), where \( A_1 = -A_2 \in \mathbb{C}^{2 \times 2} \), \( \sigma(A_1) = \{\sigma_1, \sigma_2\} \), and \( \|A_1\|^2 = \|A\|^2/2 \). The following observation is therefore non-trivial and thus of some interest.

![Figure 2](image-url). The numerical range of \( A \) from Example 5.1.
**Proposition 5.2:** Any matrix (2) satisfying condition (i) of Theorem 5.1 is unitarily irreducible.

In what follows, denote by \( \mathcal{L}_{\pm} \) the span of \( e_1, e_2 \) (resp., \( e_3, e_4 \)) – the first/last two standard basis vectors of \( \mathbb{C}^4 \). The following auxiliary statement will be used repeatedly.

**Lemma 5.1:** Let a reducing subspace \( \mathcal{L} \) of the matrix (7) have a non-trivial intersection with \( \mathcal{L}_+ \) or \( \mathcal{L}_- \). Then \( \mathcal{L} = \mathbb{C}^4 \).

**Proof:** For any \( x \in \mathbb{C}^2 \)

\[
A \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha x \\ Bx - x \end{bmatrix}, \quad A^* \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} \overline{\alpha} x \\ Bx + x \end{bmatrix}
\]

and

\[
A \begin{bmatrix} 0 \\ x \end{bmatrix} = \begin{bmatrix} B^* x + x \\ -\alpha x \end{bmatrix}, \quad A^* \begin{bmatrix} 0 \\ x \end{bmatrix} = \begin{bmatrix} B^* x - x \\ -\overline{\alpha} x \end{bmatrix}.
\]

So, if for some \( x \in \mathbb{C}^2 \) the subspace \( \mathcal{L} \) contains one of the vectors \( \begin{bmatrix} 0 \\ x \end{bmatrix} \) or \( \begin{bmatrix} x \\ 0 \end{bmatrix} \), it also contains the other one, as well as \( \begin{bmatrix} B^* x \\ 0 \end{bmatrix} \) and \( \begin{bmatrix} 0 \\ Bx \end{bmatrix} \). Applying this observation to \( Bx \) and \( B^* x \) in place of \( x \), we conclude that \( \begin{bmatrix} 0 \\ B^* x \end{bmatrix}, \begin{bmatrix} Bx \\ 0 \end{bmatrix} \) also lie in \( \mathcal{L} \), and thus

\[
\mathcal{L} \supset \text{Span} \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix}, \begin{bmatrix} Bx \\ 0 \end{bmatrix}, \begin{bmatrix} B^* x \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ x \end{bmatrix}, \begin{bmatrix} 0 \\ Bx \end{bmatrix}, \begin{bmatrix} 0 \\ B^* x \end{bmatrix} \right\}.
\]

Since \( B \) is not normal, if such a non-zero \( x \) exists, it cannot be an eigenvector of both \( B \) and \( B^* \). Consequently, \( \text{Rank}\{x, Bx, B^* x\} = 2 \), and \( \mathcal{L} = \mathbb{C}^{4 \times 4} \).

**Proof of Proposition 5.2:** As it was shown in the proof of Theorem 5.1, condition (i) implies that by scaling, rotating, and unitary similarities the matrix \( A \) can be put in form (7) with a non-normal \( B \). Since these transformations preserve unitary (ir)reducibility, we only need to consider this special case. In particular, Lemma 5.1 is applicable.

Observe that any reducing subspace \( \mathcal{L} \) of \( A \) has to be invariant under

\[
A^2 = \alpha^2 I + \text{diag}[CD, DC] = (\alpha^2 - 1)I + \text{diag}[B^*B + B - B^*, BB^* + B - B^*],
\]

and therefore under \( \text{Im} A^2 \) and \( \text{Re} A^2 \). This implies the invariance under \( \text{diag}[B^*B, BB^*] \) and \( \text{diag}[\text{Im} B, \text{Im} B] \). Equivalently, \( \mathcal{L} \) is invariant under \( H_1 = \text{diag}[P_1, P_2] \) and \( H_2 = \text{diag}[P, P] \), where \( P_1, P_2 \) and \( P \) are spectral projections of \( B^*B, BB^* \) and \( \text{Im} B \), respectively. Note that these projections all have rank one, and \( P_1 \) does not commute with \( P_2 \), due to non-normality of \( B \).

With an appropriate choice of a unitary matrix \( W \in \mathbb{C}^{2 \times 2} \), a unitary similarity via \( \text{diag}[W, W] \) can be used to put \( P \) in the form \( \text{diag}[1, 0] \), without changing the structure (7) of \( A \). We will use the notation \( \begin{bmatrix} t_j & \overline{m} \\ \omega j & \overline{1-t_j} \end{bmatrix} \) for the resulting form of \( P_j, j = 1, 2 \). Here \( 0 \leq t_j \leq 1 \), with \( t_j \) not equal zero (or one) simultaneously, and \( |\omega|_1^2 = t_j(1 - t_j) \).

Since \( H_2 = \text{diag}[1, 0, 1, 0] \), a non-zero \( \mathcal{L} \) has to contain a non-zero vector \( x \) of the form \( [m, 0, n, 0]^T \) or \( [0, m, 0, n]^T \). If only one of \( m, n \) is non-zero, Lemma 5.1 implies that \( \mathcal{L} = \mathbb{C}^4 \).
So, we need now only to consider $m, n \neq 0$. The two options for the location of the non-zero entries can be treated similarly, and for the sake of definiteness we will assume that $x = [m, 0, n, 0]^T$.

Applying $H_1$ and then $H_2$ we see that

$$[t_1 m, \omega_1 m, t_2 n, \omega_2 n]^T, \quad [t_1 m, 0, t_2 n, 0]^T \in \mathcal{L}. \quad (28)$$

If $t_1 \neq t_2$, comparing the second vector from (28) with $x$ we conclude that $e_1$ or $e_3$ lies in $\mathcal{L}$. We can thus invoke Lemma 5.1 again.

It remains to consider the case $t_1 = t_2$. The second inclusion in (28) is then redundant while the first can be simplified to $y = [0, \omega_1 m, 0, \omega_2 n]^T \in \mathcal{L}$. (28)

Since $\mathcal{L}$ is also invariant under $\text{Im} A = \begin{bmatrix} vi & -v \omega_1 \\ i & -v \\ v & i \\ i & -v_2 \end{bmatrix}$, along with $x, y$ it will contain $\tilde{x} = [vm - in, 0, im - vn, 0]^T$ and $\tilde{y} = [0, \omega_1 m - i \omega_2 n, 0, i \omega_1 m - \omega_2 n]^T$. So, $\dim \mathcal{L} \geq \text{Rank}\{x, \tilde{x}, y, \tilde{y}\} = \text{Rank} \Delta_1 + \text{Rank} \Delta_2$, where

$$\Delta_1 = \begin{bmatrix} m & vm - in \\ n & im - vn \end{bmatrix}, \quad \Delta_2 = \begin{bmatrix} \omega_1 m & \omega_1 vm - i \omega_2 n \\ \omega_2 n & i \omega_1 m - \omega_2 vn \end{bmatrix}.$$

But $\det \Delta_1 = i(m^2 + n^2) - 2vmn$ and $\det \Delta_2 = i(\omega_1^2 m^2 + \omega_2^2 n^2) - 2\omega_1 \omega_2 vmn$ cannot both equal zero unless $\omega_1 = \omega_2$. This, however, is precluded by non-commutativity of $P_1$ with $P_2$. So, $\mathcal{L}$ is at least 3-dimensional, it therefore has a non-trivial intersection with 2-dimensional $\mathcal{L}_{\pm}$, and yet another application of Lemma 5.1 completes the proof. ■

6. Reciprocal matrices

To illustrate the applicability of Theorem 5.1, let us consider a so called reciprocal 4-by-4 matrix. By definition this is a tridiagonal matrix with the constant main diagonal and off-diagonal pairs of mutually inverse entries:

$$A = \begin{bmatrix} c & a_1 & 0 & 0 \\ a_1^{-1} & c & a_2 & 0 \\ 0 & a_2^{-1} & c & a_3 \\ 0 & 0 & a_3^{-1} & c \end{bmatrix}. \quad (29)$$

A diagonal unitary similarity can be used to change the arguments of $a_j$ independently, without any effect on $W(A)$. So, without loss of generality it suffices to consider matrices (29) with $a_j > 0$ for all $j = 1, 2, 3$.

Formally speaking, reciprocal matrices are not of the type considered in this paper. However, by a transpositional similarity (29) can be put in the form (2) with $\alpha = \beta = c$ and

$$C = \begin{bmatrix} a_1 \\ a_2^{-1} \\ a_3 \end{bmatrix}, \quad D = \begin{bmatrix} a_1^{-1} & a_2 \\ a_1 & a_2^{-1} \end{bmatrix}. \quad (30)$$

The criterion for the matrix (29) to have an elliptical numerical range was obtained in [12], along with some other results on reciprocal matrices of higher sizes. Stated in terms of

$$A_j = \frac{a_j^2 + a_j^{-2}}{2} (\geq 1), \quad (31)$$
this criterion looks as follows:

\[ A_2 = \phi A_1 - \phi^{-1} A_3 \text{ or } A_2 = \phi A_3 - \phi^{-1} A_1, \]

where \( \phi = \frac{1 + \sqrt{5}}{2} \) is the golden ratio, and at least one of the inequalities in (31) is strict.

Here is what follows for reciprocal matrices (29) by applying Theorem 5.1 to \( A \) given by (2),(30) with \( \alpha = \beta \).

**Theorem 6.1:** Let \( A \) be as in (29). Then \( W(A) \) is bi-elliptical if and only if \( A_1 = A_3 > 1 \) (equivalently: \( a_1 = a_3 \neq 1 \) or \( a_1 = a_3^{-1} \neq 1 \)) and \( A_2 = 1 \) (equivalently: \( a_2 = 1 \)).

**Proof:** With \( C \) and \( D \) given by (30), the matrix \( Z \) from (3) is simply

\[
Z = \begin{bmatrix} 2 & a_2 a_3 \\ a_2^{-1} a_3^{-1} & 1 \end{bmatrix}.
\]

So, the eigenvalues \( z_1, z_2 \) of \( Z \) are \( (3 \pm \sqrt{5})/2 \), and by Lemma 2.1(i) with \( \alpha = 0 \) the spectrum of \( A \) consists of the four point \( \pm (1 \pm \sqrt{5})/2 \).

The only candidates for \( \theta \) in the statement of Theorem 5.1 in our setting are therefore integer multiples of \( \pi \).

With this choice of \( \theta \), the matrix \( e^{-i\theta} C - e^{i\theta} D^* \) up to the sign equals

\[
C - D^* = \begin{bmatrix} a_1 - a_1^{-1} & 0 \\ a_2^{-1} - a_2 & a_3 - a_3^{-1} \end{bmatrix}.
\]

It is a scalar multiple of a unitary if and only if \( a_2^{-1} - a_2 = 0 \) and \( a_1 - a_1^{-1} = \pm (a_3 - a_3^{-1}) \). These are exactly the conditions \( A_2 = 1, A_1 = A_3 \).

Finally, with \( a_2 = 1 \) the matrix

\[
(C - D^*)D = \begin{bmatrix} 1 - a_1^{-2} & a_1 - a_1^{-1} \\ 0 & 1 - a_3^{-2} \end{bmatrix}
\]

is normal if and only if \( a_1 = 1 \) (equivalently: \( A_1 = 1 \)). This concludes the proof of necessity.

To prove sufficiency, we just need to show that (27) holds with \( \theta = 0 \). To this end, observe that

\[
\det(\Im A) = \frac{1}{16} |\det(C - D^*)|^2 = \frac{1}{16} (a_1 - a_1^{-1})^2 (a_3 - a_3^{-1})^2 = \frac{1}{4} (A_1 - 1)^2,
\]

and \( \Im(A^2) \) is unitarily similar to \( \Im Z \oplus \Im Z \) so that \( s_{1,2} = \pm (a_3 - a_3^{-1})/2 \). Consequently,

\[
(s_1 - s_2)^2 = (a_3 - a_3^{-1})^2 = 2A_3 - 2 = 2A_1 - 2.
\]

We see that (27) indeed holds if we pick \( \sigma_{1,2} \in \sigma(A) \) as say \( (1 \pm \sqrt{5})/2 \) (See Figure 3).
Figure 3. An example illustrating Theorem 6.1 with \( A_2 = 1 \) and \( A_1 = A_3 = 3 \).

**Acknowledgements**

There is an alternative approach to characterizing the shape of \( W(A) \) for 4-by-4 matrices, based on the singularities analysis of the curve dual to \( C(A) \), in the spirit of [13]. We thank the anonymous referee for pointing this out.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

The second author [IMS] was supported in part by Faculty Research funding from the Division of Science and Mathematics, New York University Abu Dhabi.

**References**

[1] Horn RA, Johnson CR. Topics in matrix analysis. Cambridge: Cambridge University Press; 1994. Corrected reprint of the 1991 original.
[2] Daepp U, Gorkin P, Shaffer A, et al. Finding ellipses. Providence (RI): MAA Press’ 2018. (Carus Mathematical Monographs. Vol. 34). What Blaschke products, Poncelet’s theorem, and the numerical range know about each other.
[3] Kippenhahn R. Über den Wertevorrat einer Matrix. Math Nachr. 1951;6:193–228.
[4] Kippenhahn R. On the numerical range of a matrix. Linear Multilinear Algebra. 2008;56(1–2):185–225. Translated from the German by Paul F. Zachlin and Michiel E. Hochstenbach.
[5] Keeler D, Rodman L, Spitkovsky I. The numerical range of 3 \( \times \) 3 matrices. Linear Algebra Appl. 1997;252:115–139.
[6] Gau H-L, Wu PY. Line segments and elliptic arcs on the boundary of a numerical range. Linear Multilinear Algebra. 2008;56(1–2):131–142.
[7] Chien MT, Hung K-C. Elliptical numerical ranges of bordered matrices. Taiwanese J Math. 2012;16(3):1007–1016.
[8] Gau H-L. Elliptic numerical ranges of 4 \( \times \) 4 matrices. Taiwanese J Math. 2006;10(1):117–128.
[9] Geryba T, Spitkovsky I. On the numerical range of some block matrices with scalar diagonal blocks. Linear Multilinear Algebra. 2020.
[10] Brown E, Spitkovsky I. On flat portions on the boundary of the numerical range. Linear Algebra Appl.. 2004;390:75–109.
[11] Yeh YT. Numerical range of 2-by-2 block matrices [Master’s thesis]. Taiwan: National Central University; 2011.

[12] Bebiano N, Providência J, Spitkovsky IM. Kippenhahn curves of some tridiagonal matrices. arXiv:2011.00849v1 [math.FA]. 2020. p. 1–20.

[13] Chien M-T, Nakazato H. Singular points of the ternary polynomials associated with 4-by-4 matrices. Electron J Linear Algebra. 2012;23:755–769.