TAIWANESE JOURNAL OF MATHEMATICS
Vol. 10, No. 1, pp. 117-128, January 2006
This paper is available online at http://www.math.nthu.edu.tw/tjm/

ELLIPTIC NUMERICAL RANGES OF 4×4 MATRICES

Hwa-Long Gau

Abstract. Let A be an $n \times n$ (complex) matrix. Recall that the *numerical range* $W(A)$ of A is the set $\{\langle Ax, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\}$ in the plane, where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $\mathbb{C}^n$. In this paper a series of tests is given, allowing one to determine when the numerical range of a 4×4 matrix A is an elliptic disc.

1. Introduction

Let A be an n-by-n (complex) matrix. Recall that the *numerical range* $W(A)$ of A is the set $\{\langle Ax, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\}$ in the plane, where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $\mathbb{C}^n$. It is well known that $W(A)$ is a convex compact subset of $\mathbb{C}$, which contains all the eigenvalues of A. For properties of numerical ranges, a good reference is [6, Chapter 1].

For 2×2 matrices A a complete description of the numerical range $W(A)$ is well known. Namely, $W(A)$ is the (closed) elliptic disc with foci the eigenvalues λ_1 and λ_2 of A and the minor axis of length $(\mathrm{tr}\,(A^*A) - |\lambda_1|^2 - |\lambda_2|^2)^{1/2}$ [10]. Here, for a $n \times n$ matrix B, $\mathrm{tr}\, B$ denotes its trace.

In [8] R. Kippenhahn studied the numerical range of 3×3 matrices. He showed that there are four classes of shapes which the numerical range of a 3×3 matrix A can assume. His classification is based on the factorability of the associated polynomial $P_A(x, y, z) \equiv \det(x\mathrm{Re}\, A + y\mathrm{Im}\, A + zI)$, where Re $A = (A + A^*)/2$ and Im $A = (A - A^*)/(2i)$ are the real and imaginary parts of A, respectively, and I_n denotes the n-by-n identity matrix. This was improved in [7] by expressing the conditions in terms of the eigenvalues and entries of A, which are easier to apply.

For general n, the following Kippenhahn's result is useful: For any n-by-n matrix A, consider the homogeneous degree-n polynomial $P_A(x, y, z) = \det(x\mathrm{Re}\, A +$

Received February 27, 2005.
Communicated by Ngai-Ching Wong.
2000 *Mathematics Subject Classification*: 15A18, 15A60.
Key words and phrases: Numerical range, Kippenhahn curve.
Research partially supported by the National Science Council of the Republic of China.

$y\mathrm{Im}\,A+zI_n)$ and the algebraic curve $C(A)$ which is dual to the algebraic curve determined by $P_A(x,y,z)=0$ in the complex projective plane $\mathbb{CP}^2$, that is, $C(A)$ consists of all points $[u,v,w]$ in $\mathbb{CP}^2$ such that $ux+vy+wz=0$ is a tangent line to $P_A(x,y,z)=0$. As usual, we identify the point (x,y) in $\mathbb{C}^2$ with $[x,y,1]$ in $\mathbb{CP}^2$ and identify any point $[x,y,z]$ in $\mathbb{CP}^2$ such that $z\neq 0$ with $(x/z,y/z)$ in $\mathbb{C}^2$. Thus, in particular, the plane $\mathbb{R}^2$ (identified with $\mathbb{C}$) sits in $\mathbb{CP}^2$ by way of the identification of the point (a,b) of $\mathbb{R}^2$ (or $a+bi$ of $\mathbb{C}$) with $[a,b,1]$ in $\mathbb{CP}^2$. The algebraic curve $p(x,y,z)=0$ in $\mathbb{CP}^2$, where p is a homogeneous polynomial, can be dehomogenized to yield the curve $p(x,y,1)=0$ in $\mathbb{C}^2$ and, conversely, an algebraic curve $q(x,y)=0$ in $\mathbb{C}^2$ can be homogenized to a curve in $\mathbb{CP}^2$ with equation obtained by simplifying $q(x/z,y/z)=0$. A result of Kippenhahn says that the numerical range $W(A)$ is the convex hull of the real points of $C(A)$ (cf. [8, p. 199]). The real part of the curve $C(A)$ in the complex plane, namely, the set $\{a+bi\in\mathbb{C}\,;\,a,b\in\mathbb{R}$ and $ax+by+z=0$ is tangent to $P_A(x,y,z)=0\}$, will be denoted by $C_R(A)$ and is called the Kippenhahn curve of A. Note that, as proved in [3, Theorem 1.3]3, if $x_0u+y_0v+z_0w=0$ is a supporting line of $W(A)$, then det $(x_0\mathrm{Re}\ A+y_0\mathrm{Im}\ A+z_0I_n)=0$. Since the dual of $C(A)$ is the original curve $P_A(x,y,z)=0$, we infer, in particular, that every supporting line of $W(A)$ is tangent to $C(A)$.

There have been some attempts to classify the numerical range of 4×4 matrices using an analogous strategy as [7]. A complete solution seems rather difficult. The aim of this paper is to offer a series of tests, in terms of a 4×4 matrix A itself or its canonical unitarily equivalent forms, to determine when the numerical range of A is an elliptic disc. We will also express the conditions in terms of the eigenvalues and entries of A. These characterizations will be useful to construct a 4×4 matrix with an elliptic numerical range.

2. The Main Result

In this section, we want to formulate a necessary and sufficient condition for a 4×4 matrix A to have an elliptic disc as its numerical range.

Let A be a 4×4 matrix. We have known that if $W(A)$ is an elliptic disc, then $C(A)$ has a factor of order 2. By duality, it follows that the homogeneous polynomial P_A also has a factor of degree 2. Note that P_A is of degree 4. Therefore, if $W(A)$ is an elliptic disc, then P_A can be decomposed either by two factors of degree 2 or by one factor of degree 2 and two factors of degree 1. Therefore, we will discuss these two cases of $C_R(A)$, respectively. Now, let

$$A=\begin{bmatrix}\lambda_1 & a & d & f\\ 0 & \lambda_2 & b & e\\ 0 & 0 & \lambda_3 & c\\ 0 & 0 & 0 & \lambda_4\end{bmatrix},\qquad (2.1)$$

and $\lambda_j = \alpha_j + i\beta_j$, where α_j and β_j are real for $j = 1, 2, 3, 4$. Then

$$\begin{aligned}
P_A(x,y,z) &\equiv \det(x\mathrm{Re}\,A + y\mathrm{Im}\,A + zI_4) \\
&= \det\begin{bmatrix} \alpha_1 x+\beta_1 y+z & \frac{a}{2}(x-iy) & \frac{d}{2}(x-iy) & \frac{f}{2}(x-iy) \\ \frac{\bar a}{2}(x+iy) & \alpha_2 x+\beta_2 y+z & \frac{b}{2}(x-iy) & \frac{e}{2}(x-iy) \\ \frac{\bar d}{2}(x+iy) & \frac{\bar b}{2}(x+iy) & \alpha_3 x+\beta_3 y+z & \frac{c}{2}(x-iy) \\ \frac{\bar f}{2}(x+iy) & \frac{\bar e}{2}(x+iy) & \frac{\bar c}{2}(x+iy) & \alpha_4 x+\beta_4 y+z \end{bmatrix} \\
&= (\alpha_1 x+\beta_1 y+z)(\alpha_2 x+\beta_2 y+z)(\alpha_3 x+\beta_3 y+z)(\alpha_4 x+\beta_4 y+z) \\
&\quad -\frac{x^2+y^2}{4}Q(x,y,z),
\end{aligned}$$

where

$$\begin{aligned}
Q(x,y,z) \equiv\ & |a|^2(\alpha_3 x+\beta_3 y+z)(\alpha_4 x+\beta_4 y+z) \\
&+|b|^2(\alpha_1 x+\beta_1 y+z)(\alpha_4 x+\beta_4 y+z) \\
&+|c|^2(\alpha_1 x+\beta_1 y+z)(\alpha_2 x+\beta_2 y+z) \\
&+|d|^2(\alpha_2 x+\beta_2 y+z)(\alpha_4 x+\beta_4 y+z) \\
&+|e|^2(\alpha_1 x+\beta_1 y+z)(\alpha_3 x+\beta_3 y+z) \\
&+|f|^2(\alpha_2 x+\beta_2 y+z)(\alpha_3 x+\beta_3 y+z) \\
&+\frac{\mathrm{Re}\,(abc\bar f)}{2}(x^2-y^2) + \mathrm{Im}\,(abc\bar f)xy \\
&-(\alpha_1 x+\beta_1 y+z)(\mathrm{Re}\,(bc\bar e)x + \mathrm{Im}\,(bc\bar e)y) \\
&-(\alpha_2 x+\beta_2 y+z)(\mathrm{Re}\,(cd\bar f)x + \mathrm{Im}\,(cd\bar f)y) \\
&-(\alpha_3 x+\beta_3 y+z)(\mathrm{Re}\,(ae\bar f)x + \mathrm{Im}\,(ae\bar f)y) \\
&-(\alpha_4 x+\beta_4 y+z)(\mathrm{Re}\,(ab\bar d)x + \mathrm{Im}\,(ab\bar d)y) \\
&-\frac{x^2+y^2}{4}(|a|^2|c|^2 + |d|^2|e|^2 + |b|^2|f|^2 - 2\mathrm{Re}\,(a\bar c\bar d e) - 2\mathrm{Re}\,(b\bar d\bar e f)).
\end{aligned}$$

Let the polynomial

$$(*)\qquad \begin{aligned} P_A(x,y,z) = &(\alpha_1 x+\beta_1 y+z)(\alpha_2 x+\beta_2 y+z)(\alpha_3 x+\beta_3 y+z) \\ &\cdot(\alpha_4 x+\beta_4 y+z) - \frac{x^2+y^2}{4}Q(x,y,z) \end{aligned}$$

be denoted by $(*)$.

We now state and prove our main result. Firstly, we prove some lemmas which will be needed.

Lemma 1. *Let A be a 4×4 matrix with eigenvalues λ_1, λ_2, λ_3 and λ_4. Then the Kippenhahn curve $C_R(A)$ consists of two points and one ellipse if and only if*

$$\begin{aligned} P_A(x,y,z) &= (\alpha_1 x + \beta_1 y + z)(\alpha_2 x + \beta_2 y + z) \\ &\quad \cdot [(\alpha_3 x + \beta_3 y + z)(\alpha_4 x + \beta_4 y + z) - \frac{r^2}{4}(x^2 + y^2)], \end{aligned}$$

where $\lambda_j = \alpha_j + i\beta_j$ for all j and the α_j's and β_j's are real. In this case, the Kippenhahn curve $C_R(A)$ is the union of these two points λ_1, λ_2 and the ellipse with foci λ_3, λ_4 and the minor axis of length r.

Proof. Let

$$B = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & r \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix}.$$

Since $C_R(A) = C_R(B)$, the polynomials P_A and P_B have to be the same. Hence

$$\begin{aligned} P_A(x,y,z) &= (\alpha_1 x + \beta_1 y + z)(\alpha_2 x + \beta_2 y + z) \\ &\quad \cdot [(\alpha_3 x + \beta_3 y + z)(\alpha_4 x + \beta_4 y + z) - \frac{r^2}{4}(x^2 + y^2)]. \end{aligned}$$

The converse is clear. ∎

Lemma 2. *Let A be a 4×4 matrix. Then the Kippenhahn curve $C_R(A)$ consists of two ellipses, one with foci λ_1, λ_2 and minor axis of length s, and the other with foci λ_3, λ_4 and minor axis of length r if and only if*

$$\begin{aligned} P_A(x,y,z) &= [(\alpha_1 x + \beta_1 y + z)(\alpha_2 x + \beta_2 y + z) - \frac{s^2}{4}(x^2 + y^2)] \\ &\quad \cdot [(\alpha_3 x + \beta_3 y + z)(\alpha_4 x + \beta_4 y + z) - \frac{r^2}{4}(x^2 + y^2)], \end{aligned}$$

where $\lambda_j = \alpha_j + i\beta_j, j = 1,2,3,4,$ and the α_j's and β_j's are real.

Proof. The proof is similar to Lemma 1. Let

$$B = \begin{bmatrix} \lambda_1 & s & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & r \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix}.$$

Since $C_R(A) = C_R(B)$, the polynomials P_A and P_B have to be the same. Hence

$$\begin{aligned} P_A(x,y,z) \; = \; & [(\alpha_1 x + \beta_1 y + z)(\alpha_2 x + \beta_2 y + z) - \frac{s^2}{4}(x^2+y^2)] \\ & \cdot [(\alpha_3 x + \beta_3 y + z)(\alpha_4 x + \beta_4 y + z) - \frac{r^2}{4}(x^2+y^2)]. \end{aligned}$$

The converse is clear ■

With the above lemmas, we have the following theorems.

Theorem 3. *Let A be in upper-triangular form* (2.1). *Then $C_R(A)$ consists of two points and one ellipse if and only if*

(a) $r^2 = |a|^2 + |b|^2 + |c|^2 + |d|^2 + |e|^2 + |f|^2$,

(b) $r^2\lambda_i\lambda_j = |a|^2\lambda_3\lambda_4 + |b|^2\lambda_1\lambda_4 + |c|^2\lambda_1\lambda_2 + |d|^2\lambda_2\lambda_4 + |e|^2\lambda_1\lambda_3 + |f|^2\lambda_2\lambda_3 - (\lambda_1 bc\bar{e} + \lambda_2 cd\bar{f} + \lambda_3 ae\bar{f} + \lambda_4 ab\bar{d}) + abc\bar{f}$,

(c) $r^2(\lambda_i + \lambda_j) = (|b|^2 + |c|^2 + |e|^2)\lambda_1 + (|c|^2 + |d|^2 + |f|^2)\lambda_2 + (|a|^2 + |e|^2 + |f|^2)\lambda_3 + (|a|^2 + |b|^2 + |d|^2)\lambda_4 - (bc\bar{e} + cd\bar{f} + ae\bar{f} + ab\bar{d})$, *and*

(d) $r^2\alpha_i\alpha_j = |a|^2\alpha_3\alpha_4 + |b|^2\alpha_1\alpha_4 + |c|^2\alpha_1\alpha_2 + |d|^2\alpha_2\alpha_4 + |e|^2\alpha_1\alpha_3 + |f|^2\alpha_2\alpha_3 - [\alpha_1 \mathrm{Re}\,(bc\bar{e}) + \alpha_2 \mathrm{Re}\,(cd\bar{f}) + \alpha_3 \mathrm{Re}\,(ae\bar{f}) + \alpha_4 \mathrm{Re}\,(ab\bar{d})] - \frac{1}{4}(|a|^2|c|^2 + |d|^2|e|^2 + |b|^2|f|^2 - 2\mathrm{Re}\,(a\bar{c}\bar{d}e) - 2\mathrm{Re}\,(b\bar{d}\bar{e}f) - 2\mathrm{Re}\,(abc\bar{f}))$.

If these conditions are satisfied, then $C_R(A)$ is the union of two points λ_i,λ_j with the ellipse having its foci at two other eigenvalues of A and minor axis of length r.

Proof. By Lemma 1,

$$\begin{aligned} P_A(x,y,z) \; = \; & (\alpha_i x + \beta_i y + z)(\alpha_j x + \beta_j y + z) \\ & \cdot [(\alpha_k x + \beta_k y + z)(\alpha_l x + \beta_l y + z) - \frac{r^2}{4}(x^2+y^2)]. \end{aligned}$$

Comparing this with polynomial $(*)$, we have

$$Q(x,y,z) = r^2(\alpha_i x + \beta_i y + z)(\alpha_j x + \beta_j y + z)$$

and then obtain the following equalities by computing the coefficients of x^2, y^2, z^2, xy, xz and yz, respectively. Therefore,

(1) $r^2\alpha_i\alpha_j = |a|^2\alpha_3\alpha_4 + |b|^2\alpha_1\alpha_4 + |c|^2\alpha_1\alpha_2 + |d|^2\alpha_2\alpha_4 + |e|^2\alpha_1\alpha_3 + |f|^2\alpha_2\alpha_3 - [\alpha_1 \mathrm{Re}\,(bc\bar{e}) + \alpha_2 \mathrm{Re}\,(cd\bar{f}) + \alpha_3 \mathrm{Re}\,(ae\bar{f}) + \alpha_4 \mathrm{Re}\,(ab\bar{d})] - \frac{1}{4}(|a|^2|c|^2 + |d|^2|e|^2 + |b|^2|f|^2 - 2\mathrm{Re}\,(a\bar{c}\bar{d}e) - 2\mathrm{Re}\,(b\bar{d}\bar{e}f) - 2\mathrm{Re}\,(abc\bar{f}))$,

(2) $r^2\beta_i\beta_j = |a|^2\beta_3\beta_4 + |b|^2\beta_1\beta_4 + |c|^2\beta_1\beta_2 + |d|^2\beta_2\beta_4 + |e|^2\beta_1\beta_3 + |f|^2\beta_2\beta_3 - [\beta_1 \mathrm{Im}\,(bc\bar{e}) + \beta_2 \mathrm{Im}\,(cd\bar{f}) + \beta_3 \mathrm{Im}\,(ae\bar{f}) + \beta_4 \mathrm{Im}\,(ab\bar{d})] - \frac{1}{4}(|a|^2|c|^2 + |d|^2|e|^2 + |b|^2|f|^2 - 2\mathrm{Re}\,(a\bar{c}\bar{d}e) - 2\mathrm{Re}\,(b\bar{d}\bar{e}f) + 2\mathrm{Re}\,(abc\bar{f}))$,

(3) $r^2 = |a|^2 + |b|^2 + |c|^2 + |d|^2 + |e|^2 + |f|^2$,

(4) $r^2(\alpha_i\beta_j + \alpha_j\beta_i) = |a|^2(\alpha_3\beta_4 + \alpha_4\beta_3) + |b|^2(\alpha_1\beta_4 + \alpha_4\beta_1) + |c|^2(\alpha_1\beta_2 + \alpha_2\beta_1) + |d|^2(\alpha_2\beta_4 + \alpha_4\beta_2) + |e|^2(\alpha_1\beta_3 + \alpha_3\beta_1) + |f|^2(\alpha_2\beta_3 + \alpha_3\beta_2) - (\alpha_1 \mathrm{Im}\,(bc\bar{e}) + \beta_1 \mathrm{Re}\,(bc\bar{e})) - [\alpha_2 \mathrm{Im}\,(cd\bar{f}) + \beta_2 \mathrm{Re}\,(cd\bar{f})] - [\alpha_3 \mathrm{Im}\,(ae\bar{f}) + \beta_3 \mathrm{Re}\,(ae\bar{f})] - [\alpha_4 \mathrm{Im}\,(ab\bar{d}) + \beta_4 \mathrm{Re}\,(ab\bar{d})] + \mathrm{Im}\,(abc\bar{f})$,

(5) $r^2(\alpha_i + \alpha_j) = (|b|^2 + |c|^2 + |e|^2)\alpha_1 + (|c|^2 + |d|^2 + |f|^2)\alpha_2 + (|a|^2 + |e|^2 + |f|^2)\alpha_3 + (|a|^2 + |b|^2 + |d|^2)\alpha_4 - [\mathrm{Re}\,(bc\bar{e}) + \mathrm{Re}\,(cd\bar{f}) + \mathrm{Re}\,(ae\bar{f}) + \mathrm{Re}\,(ab\bar{d})]$, *and*

(6) $r^2(\beta_i + \beta_j) = (|b|^2 + |c|^2 + |e|^2)\beta_1 + (|c|^2 + |d|^2 + |f|^2)\beta_2 + (|a|^2 + |e|^2 + |f|^2)\beta_3 + (|a|^2 + |b|^2 + |d|^2)\beta_4 - [\mathrm{Im}\,(bc\bar{e}) + \mathrm{Im}\,(cd\bar{f}) + \mathrm{Im}\,(ae\bar{f}) + \mathrm{Im}\,(ab\bar{d})]$.

Note that the combination of (1), (2) and (4) is equivalent to the one of (b) and (d) since $(1) - (2) + i(4)$ yields (b). Moreover, the combination of (5) and (6) is equivalent to (c) since $(5) + i(6)$ yields (c). This completes the proof. ∎

A similar argument shows the following theorem.

Theorem 4. *Let A be in upper-triangular form* (2.1). *Then $C_R(A)$ consists of two ellipses, one with foci λ_k, λ_l and minor axis of length r, the other with foci λ_i, λ_j and minor axis of length s if and only if*

(a) $r^2 + s^2 = |a|^2 + |b|^2 + |c|^2 + |d|^2 + |e|^2 + |f|^2$,

(b) $r^2\lambda_i\lambda_j + s^2\lambda_k\lambda_l = |a|^2\lambda_3\lambda_4 + |b|^2\lambda_1\lambda_4 + |c|^2\lambda_1\lambda_2 + |d|^2\lambda_2\lambda_4 + |e|^2\lambda_1\lambda_3 + |f|^2\lambda_2\lambda_3 - (\lambda_1 bc\bar{e} + \lambda_2 cd\bar{f} + \lambda_3 ae\bar{f} + \lambda_4 ab\bar{d}) + abc\bar{f}$,

(c) $r^2(\lambda_i + \lambda_j) + s^2(\lambda_k + \lambda_l) = (|b|^2 + |c|^2 + |e|^2)\lambda_1 + (|c|^2 + |d|^2 + |f|^2)\lambda_2 + (|a|^2 + |e|^2 + |f|^2)\lambda_3 + (|a|^2 + |b|^2 + |d|^2)\lambda_4 - (bc\bar{e} + cd\bar{f} + ae\bar{f} + ab\bar{d})$, *and*

(d) $r^2\alpha_i\alpha_j + s^2\alpha_k\alpha_l - \frac{1}{4}r^2s^2 = |a|^2\alpha_3\alpha_4 + |b|^2\alpha_1\alpha_4 + |c|^2\alpha_1\alpha_2 + |d|^2\alpha_2\alpha_4 + |e|^2\alpha_1\alpha_3 + |f|^2\alpha_2\alpha_3 - [\alpha_1 \mathrm{Re}\,(bc\bar{e}) + \alpha_2 \mathrm{Re}\,(cd\bar{f}) + \alpha_3 \mathrm{Re}\,(ae\bar{f}) + \alpha_4 \mathrm{Re}\,(ab\bar{d})] - \frac{1}{4}(|a|^2|c|^2 + |d|^2|e|^2 + |b|^2|f|^2 - 2\mathrm{Re}\,(a\bar{c}\bar{d}e) - 2\mathrm{Re}\,(b\bar{d}\bar{e}f) - 2\mathrm{Re}\,(abc\bar{f}))$.

Proof. By Lemma 2,

$$\begin{aligned} P_A(x,y,z) \;=\; & [(\alpha_i x + \beta_i y + z)(\alpha_j x + \beta_j y + z) - \frac{s^2}{4}(x^2 + y^2)] \\ & \cdot[(\alpha_k x + \beta_k y + z)(\alpha_l x + \beta_l y + z) - \frac{r^2}{4}(x^2 + y^2)]. \end{aligned}$$

Comparing this with polynomial $(*)$, we have

$$Q(x,y,z) = r^2(\alpha_i x + \beta_i y + z)(\alpha_j x + \beta_j y + z) + s^2(\alpha_k x + \beta_k y + z)(\alpha_l x + \beta_l y + z) - \frac{r^2 s^2}{4}(x^2 + y^2)$$

and then obtain the following equalities by computing the coefficients of x^2, y^2, z^2, xy, xz, and yz, respectively. Therefore,

(1) $r^2\alpha_i\alpha_j + s^2\alpha_k\alpha_l - \frac{r^2s^2}{4} = |a|^2\alpha_3\alpha_4 + |b|^2\alpha_1\alpha_4 + |c|^2\alpha_1\alpha_2 + |d|^2\alpha_2\alpha_4 + |e|^2\alpha_1\alpha_3 + |f|^2\alpha_2\alpha_3 - [\alpha_1\mathrm{Re}\,(bc\bar{e}) + \alpha_2\mathrm{Re}\,(cd\bar{f}) + \alpha_3\mathrm{Re}\,(ae\bar{f}) + \alpha_4\mathrm{Re}\,(ab\bar{d})] - \frac{1}{4}(|a|^2|c|^2 + |d|^2|e|^2 + |b|^2|f|^2 - 2\mathrm{Re}\,(a\bar{c}\bar{d}e) - 2\mathrm{Re}\,(b\bar{d}\bar{e}f) - 2\mathrm{Re}\,(abc\bar{f}))$,

(2) $r^2\beta_i\beta_j + s^2\beta_k\beta_l - \frac{r^2s^2}{4} = |a|^2\beta_3\beta_4 + |b|^2\beta_1\beta_4 + |c|^2\beta_1\beta_2 + |d|^2\beta_2\beta_4 + |e|^2\beta_1\beta_3 + |f|^2\beta_2\beta_3 - [\beta_1\mathrm{Im}\,(bc\bar{e}) + \beta_2\mathrm{Im}\,(cd\bar{f}) + \beta_3\mathrm{Im}\,(ae\bar{f}) + \beta_4\mathrm{Im}\,(ab\bar{d})] - \frac{1}{4}(|a|^2|c|^2 + |d|^2|e|^2 + |b|^2|f|^2 - 2\mathrm{Re}\,(a\bar{c}\bar{d}e) - 2\mathrm{Re}\,(b\bar{d}\bar{e}f) + 2\mathrm{Re}\,(abc\bar{f}))$,

(3) $r^2 + s^2 = |a|^2 + |b|^2 + |c|^2 + |d|^2 + |e|^2 + |f|^2$,

(4) $r^2(\alpha_i\beta_j + \alpha_j\beta_i) + s^2(\alpha_k\beta_l + \alpha_l\beta_k) = |a|^2(\alpha_3\beta_4 + \alpha_4\beta_3) + |b|^2(\alpha_1\beta_4 + \alpha_4\beta_1) + |c|^2(\alpha_1\beta_2 + \alpha_2\beta_1) + |d|^2(\alpha_2\beta_4 + \alpha_4\beta_2) + |e|^2(\alpha_1\beta_3 + \alpha_3\beta_1) + |f|^2(\alpha_2\beta_3 + \alpha_3\beta_2) - (\alpha_1\mathrm{Im}\,(bc\bar{e}) + \beta_1\mathrm{Re}\,(bc\bar{e})) - [\alpha_2\mathrm{Im}\,(cd\bar{f}) + \beta_2\mathrm{Re}\,(cd\bar{f})] - [\alpha_3\mathrm{Im}\,(ae\bar{f}) + \beta_3\mathrm{Re}\,(ae\bar{f})] - [\alpha_4\mathrm{Im}\,(ab\bar{d}) + \beta_4\mathrm{Re}\,(ab\bar{d})] + \mathrm{Im}\,(abc\bar{f})$,

(5) $r^2(\alpha_i + \alpha_j) + s^2(\alpha_k + \alpha_l) = (|b|^2 + |c|^2 + |e|^2)\alpha_1 + (|c|^2 + |d|^2 + |f|^2)\alpha_2 + (|a|^2 + |e|^2 + |f|^2)\alpha_3 + (|a|^2 + |b|^2 + |d|^2)\alpha_4 - [\mathrm{Re}\,(bc\bar{e}) + \mathrm{Re}\,(cd\bar{f}) + \mathrm{Re}\,(ae\bar{f}) + \mathrm{Re}\,(ab\bar{d})]$, and

(6) $r^2(\beta_i + \beta_j) + s^2(\beta_k + \beta_l) = (|b|^2 + |c|^2 + |e|^2)\beta_1 + (|c|^2 + |d|^2 + |f|^2)\beta_2 + (|a|^2 + |e|^2 + |f|^2)\beta_3 + (|a|^2 + |b|^2 + |d|^2)\beta_4 - [\mathrm{Im}\,(bc\bar{e}) + \mathrm{Im}\,(cd\bar{f}) + \mathrm{Im}\,(ae\bar{f}) + \mathrm{Im}\,(ab\bar{d})]$.

Note that the combination of (1), (2) and (4) is equivalent to the one of (b) and (d) since $(1) - (2) + i(4)$ yields (b). Moreover, the combination of (5) and (6) is equivalent to (c) since $(5) + i(6)$ yields (c). This completes the proof. ∎

Although every matrix is unitarily equivalent to an upper-triangular matrix, it is not easy to obtain the upper-triangular form of a matrix. For generality, we obtain the unitary invariant forms of Theorems 3 and 4.

Corollary 5. *Let A be a 4×4 matrix with eigenvalues λ_1, λ_2, λ_3 and λ_4. Then $C_R(A)$ consists of two points λ_i, λ_j and one ellipse having its foci at two other eigenvalues of A and minor axis of length r if and only if*

(a) $r^2 = \mathrm{tr}\,(A^*A) - \sum_{i=1}^4 |\lambda_i|^2$,

(b) $r^2\lambda_i\lambda_j = \sum_{1\le i<j\le 4}(r^2+|\lambda_i|^2+|\lambda_j|^2)\lambda_i\lambda_j + \mathrm{tr}\,(A^*A^3) - \mathrm{tr}\,(A)\mathrm{tr}\,(A^*A^2)$,

(c) $r^2(\lambda_i+\lambda_j) = r^2\mathrm{tr}\,(A) - \mathrm{tr}\,(A^*A^2) + \sum_{i=1}^4 |\lambda_i|^2\lambda_i$, *and*

(d) $r^2\alpha_i\alpha_j = 4\alpha_1\alpha_2\alpha_3\alpha_4 - 4\det(\mathrm{Re}\,A)$.

Proof. Let B be in upper-triangular form (2.1) which is unitarily equivalent to A. After a little computation, we obtain

$$\begin{aligned}
\mathrm{tr}\,(B^*B) &= \sum_{i=1}^4 |\lambda_i|^2 + |a|^2 + |b|^2 + |c|^2 + |d|^2 + |e|^2 + |f|^2,\\
\mathrm{tr}\,(B^*B^2) &= \sum_{i=1}^4 |\lambda_i|^2\lambda_i + (|a|^2+|d|^2+|f|^2)\lambda_1 + (|a|^2+|b|^2+|e|^2)\lambda_2\\
&\quad +(|b|^2+|c|^2+|d|^2)\lambda_3 + (|c|^2+|e|^2+|f|^2)\lambda_4\\
&\quad +(ab\bar{d} + ae\bar{f} + cd\bar{f} + bc\bar{e}),
\end{aligned}$$

$$\begin{aligned}
\mathrm{tr}\,(B^*B^3) &= \sum_{i=1}^4 |\lambda_i|^2{\lambda_i}^2 + (|a|^2+|d|^2+|f|^2){\lambda_1}^2 + (|a|^2+|b|^2+|e|^2){\lambda_2}^2\\
&\quad +(|b|^2+|c|^2+|d|^2){\lambda_3}^2 + (|c|^2+|e|^2+|f|^2){\lambda_4}^2 + |a|^2\lambda_1\lambda_2 + |b|^2\lambda_2\lambda_3\\
&\quad +|c|^2\lambda_3\lambda_4 + |d|^2\lambda_1\lambda_3 + |e|^2\lambda_2\lambda_4 + |f|^2\lambda_1\lambda_4 + ab\bar{d}(\lambda_1+\lambda_2+\lambda_3)\\
&\quad +ae\bar{f}(\lambda_1+\lambda_2+\lambda_4) + cd\bar{f}(\lambda_1+\lambda_3+\lambda_4) + bc\bar{e}(\lambda_2+\lambda_3+\lambda_4) + abc\bar{f},
\end{aligned}$$

and

$$\det(\mathrm{Re}\,B) = \alpha_1\alpha_2\alpha_3\alpha_4 - \frac{1}{4}Q(1,0,0).$$

By the condition (a) in Theorem 3, we have

$$\mathrm{tr}\,(B^*B) - \sum_{i=1}^4 |\lambda_i|^2 = |a|^2 + |b|^2 + |c|^2 + |d|^2 + |e|^2 + |f|^2 = r^2.$$

By the condition (b) in Theorem 3, we have

$$\begin{aligned}
&\sum_{1\le i<j\le 4}(r^2+|\lambda_i|^2+|\lambda_j|^2)\lambda_i\lambda_j + \mathrm{tr}\,(B^*B^3) - \mathrm{tr}\,(B)\mathrm{tr}\,(B^*B^2)\\
&= |a|^2\lambda_3\lambda_4 + |b|^2\lambda_1\lambda_4 + |c|^2\lambda_1\lambda_2 + |d|^2\lambda_2\lambda_4 + |e|^2\lambda_1\lambda_3 + |f|^2\lambda_2\lambda_3\\
&\quad -(\lambda_1 bc\bar{e} + \lambda_2 cd\bar{f} + \lambda_3 ae\bar{f} + \lambda_4 ab\bar{d}) + abc\bar{f}.
\end{aligned}$$

By the condition (c) in Theorem 3, we have

$$\begin{aligned} & r^2 \mathrm{tr}\,(B) - \mathrm{tr}\,(B^* B^2) + \sum_{i=1}^{4} |\lambda_i|^2 \lambda_i \\ = & \ (|b|^2 + |c|^2 + |e|^2)\lambda_1 + (|c|^2 + |d|^2 + |f|^2)\lambda_2 + (|a|^2 + |e|^2 + |f|^2)\lambda_3 \\ & +(|a|^2 + |b|^2 + |d|^2)\lambda_4 - (bc\bar{e} + cd\bar{f} + ae\bar{f} + ab\bar{d}). \end{aligned}$$

By the condition (d) in Theorem 3, we have

$$\begin{aligned} 4\alpha_1\alpha_2\alpha_3\alpha_4 - 4\det(\mathrm{Re}\,B) \ &= \ Q(1,0,0) \\ &= \ |a|^2\alpha_3\alpha_4 + |b|^2\alpha_1\alpha_4 + |c|^2\alpha_1\alpha_2 \\ & \quad +|d|^2\alpha_2\alpha_4 + |e|^2\alpha_1\alpha_3 + |f|^2\alpha_2\alpha_3 \\ & \quad -[\alpha_1 \mathrm{Re}\,(bc\bar{e}) + \alpha_2 \mathrm{Re}\,(cd\bar{f}) \\ & \quad +\alpha_3 \mathrm{Re}\,(ae\bar{f}) + \alpha_4 \mathrm{Re}\,(ab\bar{d})] \\ & \quad -\frac{1}{4}(|a|^2|c|^2 + |d|^2|e|^2 + |b|^2|f|^2 \\ & \quad -2\mathrm{Re}\,(a\bar{c}\bar{d}e) - 2\mathrm{Re}\,(b\bar{d}\bar{e}f) - 2\mathrm{Re}\,(abc\bar{f})). \end{aligned}$$

Since trace and determinant are unitary invariant, completing the proof. ∎

Corollary 6. *Let A be a 4×4 matrix with eigenvalues λ_1, λ_2, λ_3 and λ_4. Then $C_R(A)$ consists of two ellipses, one with foci λ_k, λ_l and minor axis of length r, the other with foci λ_i, λ_j and minor axis of length s if and only if*

(a) $r^2 + s^2 = \mathrm{tr}\,(A^*A) - \sum_{i=1}^{4} |\lambda_i|^2 \equiv \gamma^2$,

(b) $r^2\lambda_i\lambda_j + s^2\lambda_k\lambda_l = \sum_{1\le i<j\le 4}(\gamma^2 + |\lambda_i|^2 + |\lambda_j|^2)\lambda_i\lambda_j + \mathrm{tr}\,(A^*A^3) - \mathrm{tr}\,(A)\mathrm{tr}\,(A^*A^2)$,

(c) $r^2(\lambda_i + \lambda_j) + s^2(\lambda_k + \lambda_l) = \gamma^2 \mathrm{tr}\,(A) - \mathrm{tr}\,(A^*A^2) + \sum_{i=1}^{4} |\lambda_i|^2\lambda_i$, *and*

(d) $r^2\alpha_i\alpha_j + s^2\alpha_k\alpha_l - \frac{1}{4}r^2s^2 = 4\alpha_1\alpha_2\alpha_3\alpha_4 - 4\det(\mathrm{Re}\,A)$.

Proof. Let B be in upper-triangular form (2.1) which is unitarily equivalent to A. A direct computation yields that

$$\mathrm{tr}\,(B^*B) \ = \ \sum_{i=1}^{4} |\lambda_i|^2 + |a|^2 + |b|^2 + |c|^2 + |d|^2 + |e|^2 + |f|^2,$$

$$\mathrm{tr}\,(B^*B^2) \ = \ \sum_{i=1}^{4} |\lambda_i|^2\lambda_i + (|a|^2 + |d|^2 + |f|^2)\lambda_1 + (|a|^2 + |b|^2 + |e|^2)\lambda_2$$

$$
\begin{aligned}
&+(|b|^2+|c|^2+|d|^2)\lambda_3+(|c|^2+|e|^2+|f|^2)\lambda_4\\
&+(ab\bar{d}+ae\bar{f}+cd\bar{f}+bc\bar{e}),\\
\mathrm{tr}\,(B^*B^3) \;=\; &\sum_{i=1}^{4}|\lambda_i|^2{\lambda_i}^2+(|a|^2+|d|^2+|f|^2){\lambda_1}^2+(|a|^2+|b|^2+|e|^2){\lambda_2}^2\\
&+(|b|^2+|c|^2+|d|^2){\lambda_3}^2+(|c|^2+|e|^2+|f|^2){\lambda_4}^2+|a|^2\lambda_1\lambda_2+|b|^2\lambda_2\lambda_3\\
&+|c|^2\lambda_3\lambda_4+|d|^2\lambda_1\lambda_3+|e|^2\lambda_2\lambda_4+|f|^2\lambda_1\lambda_4+ab\bar{d}(\lambda_1+\lambda_2+\lambda_3)\\
&+ae\bar{f}(\lambda_1+\lambda_2+\lambda_4)+cd\bar{f}(\lambda_1+\lambda_3+\lambda_4)+bc\bar{e}(\lambda_2+\lambda_3+\lambda_4)+abc\bar{f},
\end{aligned}
$$

and

$$\det(\mathrm{Re}\,B) \;=\; \alpha_1\alpha_2\alpha_3\alpha_4-\frac{1}{4}Q(1,0,0).$$

By the condition (a) in Theorem 4, we have

$$\mathrm{tr}\,(B^*B)-\sum_{i=1}^{4}|\lambda_i|^2=|a|^2+|b|^2+|c|^2+|d|^2+|e|^2+|f|^2=r^2.$$

By the condition (b) in Theorem 4, we have

$$
\begin{aligned}
&\sum_{1\le i<j\le 4}(r^2+|\lambda_i|^2+|\lambda_j|^2)\lambda_i\lambda_j+\mathrm{tr}\,(B^*B^3)-\mathrm{tr}\,(B)\mathrm{tr}\,(B^*B^2)\\
=\;&|a|^2\lambda_3\lambda_4+|b|^2\lambda_1\lambda_4+|c|^2\lambda_1\lambda_2+|d|^2\lambda_2\lambda_4+|e|^2\lambda_1\lambda_3+|f|^2\lambda_2\lambda_3\\
&-(\lambda_1bc\bar{e}+\lambda_2cd\bar{f}+\lambda_3ae\bar{f}+\lambda_4ab\bar{d})+abc\bar{f}.
\end{aligned}
$$

By the condition (c) in Theorem 4, we have

$$
\begin{aligned}
&r^2\mathrm{tr}\,(B)-\mathrm{tr}\,(B^*B^2)+\sum_{i=1}^{4}|\lambda_i|^2\lambda_i\\
=\;&(|b|^2+|c|^2+|e|^2)\lambda_1+(|c|^2+|d|^2+|f|^2)\lambda_2+(|a|^2+|e|^2+|f|^2)\lambda_3\\
&+(|a|^2+|b|^2+|d|^2)\lambda_4-(bc\bar{e}+cd\bar{f}+ae\bar{f}+ab\bar{d}).
\end{aligned}
$$

By the condition (d) in Theorem 4, we have

$$
\begin{aligned}
4\alpha_1\alpha_2\alpha_3\alpha_4-4\det(\mathrm{Re}\,B) \;&=\; Q(1,0,0)\\
&=\;|a|^2\alpha_3\alpha_4+|b|^2\alpha_1\alpha_4+|c|^2\alpha_1\alpha_2\\
&\quad+|d|^2\alpha_2\alpha_4+|e|^2\alpha_1\alpha_3+|f|^2\alpha_2\alpha_3
\end{aligned}
$$

$$-[\alpha_1 \mathrm{Re}\,(bc\bar{e}) + \alpha_2 \mathrm{Re}\,(cd\bar{f})$$
$$+\alpha_3 \mathrm{Re}\,(ae\bar{f}) + \alpha_4 \mathrm{Re}\,(ab\bar{d})]$$
$$-\frac{1}{4}(|a|^2|c|^2 + |d|^2|e|^2 + |b|^2|f|^2$$
$$-2\mathrm{Re}\,(a\bar{c}\bar{d}e) - 2\mathrm{Re}\,(b\bar{d}\bar{e}f) - 2\mathrm{Re}\,(abc\bar{f})).$$

Since trace and determinant are unitary invariant, the results follow obviously. ■

Now we are ready to formulate a sufficient condition for a 4×4 matrix A to have an elliptic disc as its numerical range.

Corollary 7. *Let A be a 4×4 matrix with eigenvalues λ_1, λ_2, λ_3 and λ_4 which satisfies the following conditions:*

(a) $r^2 = \mathrm{tr}\,(A^*A) - \sum_{i=1}^4 |\lambda_i|^2$,

(b) $r^2\lambda_i\lambda_j = \sum_{1 \le i < j \le 4}(r^2 + |\lambda_i|^2 + |\lambda_j|^2)\lambda_i\lambda_j + \mathrm{tr}\,(A^*A^3) - \mathrm{tr}\,(A)\mathrm{tr}\,(A^*A^2)$,

(c) $r^2(\lambda_i + \lambda_j) = \gamma^2\mathrm{tr}\,(A) - \mathrm{tr}\,(A^*A^2) + \sum_{i=1}^4 |\lambda_i|^2\lambda_i$,

(d) $r^2\alpha_i\alpha_j = 4\alpha_1\alpha_2\alpha_3\alpha_4 - 4\det(\mathrm{Re}\,A)$, *and*

(e) $(|\lambda - \lambda_k| + |\lambda - \lambda_l|)^2 - |\lambda_k - \lambda_l|^2 \le r^2$, *where* $\lambda = \lambda_i, \lambda_j$ *and* λ_k, λ_l *are other two eigenvalues of* A.

Then $W(A)$ is an elliptic disc with foci λ_k, λ_l and the minor axis of length r.

Proof. By Corollary 5, $C_R(A)$ consists of two points λ_i, λ_j and one ellipse whose foci are λ_k, λ_l and whose minor axis has length r. Moreover, condition (e) means that these two points λ_i, λ_j lie inside the ellipse. Hence $W(A)$ is an elliptic with foci λ_k, λ_l and the minor axis of length r. ■

Corollary 8. *Let A be a 4×4 matrix with eigenvalues λ_1, λ_2, λ_3 and λ_4. If conditions* (a)-(d) *of* Corollary 6 *hold and, in addition,*

(e) $\sqrt{|\lambda_k - \lambda_l|^2 + r^2} + |\lambda_k - \lambda_i| + |\lambda_l - \lambda_j| \le \sqrt{|\lambda_i - \lambda_j|^2 + s^2}$.

Then $W(A)$ is an elliptic disc with foci λ_i, λ_j and the minor axis of length s.

Proof. By Corollary 6, $C_R(A)$ consists of two ellipses, one with foci λ_k, λ_l and the minor axis of length r and the other with foci λ_i, λ_j and the minor axis of length s. Moreover, for λ in $\mathbb{C}$ such that

$$|\lambda - \lambda_k| + |\lambda - \lambda_l| \le \sqrt{|\lambda_k - \lambda_l| + r^2},$$

we have

$$\begin{aligned} |\lambda-\lambda_i|+|\lambda-\lambda_j| &\le |\lambda-\lambda_k|+|\lambda-\lambda_l|+|\lambda_k-\lambda_i|+|\lambda_l-\lambda_j| \\ &\le \sqrt{|\lambda_k-\lambda_l|+r^2}+|\lambda_k-\lambda_i|+|\lambda_l-\lambda_j| \\ &\le \sqrt{|\lambda_i-\lambda_j|^2+s^2} \end{aligned}$$

by condition (e). Thus we conclude that $W(A)$ is an elliptic disc with foci λ_i, λ_j and the minor axis of length s. ∎

REFERENCES

1. E. Brieskorn and H. Knörrer, *Plane Algebraic Curves*, Birkhäuser Verlag, Basel, 1986.
2. J. L. Coolidge, *A Treatise on Algebraic Plane Curves*, Dover, New York, 1959.
3. M. Fiedler, Geometry of the numerical range of matrices, *Linear Algebra Appl*. **37** (1981), 81-96.
4. H.-L. Gau and Y.-H. Lu, *Elliptical numerical ranges of* 4×4 *matrices*, Master Thesis, National Central University, 2003.
5. K. E. Gustafson and D. K. M. Rao, *Numerical Range, the Field of Values of Linear Operators and Matrices*, Springer, New York, 1997.
6. R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1991.
7. D. S. Keeler, L. Rodman and I. M. Spitkovsky, The Numerical Range of 3×3 Matrices. *Linear Algebra Appl.*, **252** (1997), 115-139.
8. R. Kippenhahn, Über den Wertevorrat einer Matrix, *Math. Nachr.*, **6** (1951), 193-228.
9. F. Kirwan, *Complex Algebraic Curves*, Cambridge Univ. Press, Cambridge, 1992.
10. C.-K. Li, A simple proof of the elliptical range theorem, *Proc. Amer. Math. Soc*., **124** (1996), 1985-1986.
11. F. D. Murnaghan, On the field of values of a square matrix, *Proc. Nat. Acad. Sci. U.S.A.*, **18** (1932), 246-248.

Hwa-Long Gau
Department of Mathematics,
National Central University,
Chung-Li 32001, Taiwan.
E-mail: hlgau@math.ncu.edu.tw