A proof of the elliptical range theorem via Kippenhahn's theorem

Pietro Paparella[1], Luis J. Ramirez[2], and Yen-Fen Wang[2]

[1]Division of Engineering and Mathematics, University of Washington Bothell, Bothell, WA 98011-8246, USA (pietrop@uw.edu)
[2]University of Washington Bothell, Bothell, WA 98011-8246, USA

September 27, 2018

Abstract

The elliptical range theorem asserts that the field of values (or numerical range) of a two-by-two matrix with complex entries is an elliptical disk, the foci of which are the eigenvalues of the given matrix. Many proofs of this result are available in the literature, but most, with one exception, are computational and quite involved. In this note, it is shown that the elliptical range theorem follows from the properties of plane algebraic curves and a straightforward application of a well-known result due to Kippenhahn.

1 Introduction

If A is an an n-by-n matrix, then the *field (of values)* or the *numerical range of A*, denoted by $F(A)$, is defined by $F(A) = \{x^*Ax : x^*x = 1\} \subseteq \mathbb{C}$. The field possesses many desirable properties and, as such, is useful in operator theory and other subjects including quantum computing (e.g., see Horn and Johnson [3] and references therein). In particular, the celebrated Toeplitz-Hausdorff theorem asserts that the field is convex.

Most proofs of this result begin by observing that the problem reduces to the two-dimensional case [1, 2, 3, 6]. In this particular case we obtain the following fundamental result, which is known as the *elliptical range theorem*.

Theorem 1.1 *[Elliptical range theorem] If A is a two-by-two matrix with complex entries and eigenvalues λ_1 and λ_2, then $F(A)$ is an elliptical disk centered at $(1/2)\operatorname{tr}(A)$, has foci λ_1 and λ_2, and has minor axis length equal to*

$$\sqrt{\operatorname{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}.$$

There are many proofs of this result [2, 3, 4, 7, 8] and most, with one exception [7], are computational and quite involved. In the sequel, it will be shown that the elliptical range theorem follows from properties of plane algebraic curves and a straightforward application of a well-known result due to Kippenhahn.

2 Background

In this section we recall relevant background.

2.1 Algebraic Curves

If $p \in \mathbb{C}[x, y]$ is a polynomial of degree n, then the *plane algebraic curve with respect to p*, denoted by $\gamma = \gamma_p$, is defined by

$$\gamma = \left\{ (x, y) \in \mathbb{C}^2 \mid p(x, y) = 0 \right\}.$$

If $P \in \mathbb{C}[x, y, z]$ is a homogeneous polynomial of degree n, then the *plane projective curve with respect to P,* denoted by $\kappa = \kappa_P$, is defined by

$$\kappa = \left\{ (x, y, z) \in \mathbb{CP}^2 \mid P(x, y, z) = 0 \right\}.$$

The *degree of γ* (*degree of κ*), denoted by $\deg \gamma$ (respectively, $\deg \kappa$), is defined by $\deg \gamma = \deg p$ (respectively, $\deg \kappa = \deg P$). The *real part* of a plane algebraic curve γ (plane projective curve κ) is defined by $\mathrm{Re}(\gamma) = \left\{ (x, y) \in \mathbb{R}^2 \mid p(x, y) = 0 \right\}$ (respectively, $\mathrm{Re}(\kappa) = \left\{ (x, y, z) \in \mathbb{RP}^2 \mid P(x, y, z) = 0 \right\}$).

If $p \in \mathbb{C}[x, y]$, then $H[p](x, y, z) := z^{\deg p} p(x/z, y/z)$ is a homogeneous polynomial. If $P \in \mathbb{C}[x, y, z]$ is a homogeneous polynomial, then $B[P](x, y) := P(x, y, 1)$ is a bivariate polynomial. Thus, every plane algebraic curve can be identified with a plane projective curve and vice-versa.

If κ_P is a plane projective curve of degree n, then the *dual of κ_P,* denoted by $(\kappa_P)^\delta = \kappa_{P^\delta}$, is the unique plane projective curve of degree m such that

$$P^\delta(u, v, w) = 0 \tag{1}$$

if and only if the line

$$ux + vy + wz = 0 \tag{2}$$

is tangent to κ_P. It is well-known that $((\kappa_P)^\delta)^\delta = \kappa_P$.

To find the *point-equation* $P(x, y, z) = 0$ given the *tangential equation* $P^\delta(u, v, w) = 0$, one can eliminate the variables u, v, w, λ from (1), (2), and

$$\frac{\partial P^\delta}{\partial u} + \lambda x = 0, \quad \frac{\partial P^\delta}{\partial v} + \lambda y = 0, \quad \frac{\partial P^\delta}{\partial w} + \lambda z = 0 \tag{3}$$

(see, e.g., Salmon [9, p. 76]).

2.2 The Field of Values

Let A be an n-by-n matrix with complex entries. The following properties are well-known and otherwise easy to establish:

P1 $F(\alpha A + \beta I) = \alpha F(A) + \beta$ [3, Properties 1.2.3 & 1.2.4].

P2 If A is *normal,* i.e., if $A^* A = A A^*$, then $F(A) = \mathrm{conv}\,(\sigma(A))$, in which $\sigma(A)$ denotes the spectrum of A [3, Property 1.2.9].

P3 If U is a *unitary matrix,* i.e., if $U^* U = U U^* = I_n$, then $F(U^* A U) = F(A)$.

The following result is due to Kippenhahn [5].

Theorem 2.1 ([5, Theorem 10]) *If A is an n-by-n matrix with complex entries, then there is a plane algebraic curve κ_P of class n such that*

$$F(A) = \mathrm{conv}\left(\mathrm{Re}(\gamma_{P(x,y,1)})\right).$$

Furthermore, if $H_1 := (A + A^*)/2$ and $H_2 := (A - A^*)/(2i)$, then

$$P^\delta = |H_1 u + H_2 v + I_n w|. \tag{4}$$

3 The Proof

Proof of Theorem 1.1. Let A be a two-by-two matrix with complex entries and eigenvalues λ_1 and λ_2. We follow the cases established by [7].

If A is normal, then, as a consequence of **P2**, $F(A) = \mathrm{conv}\,(\lambda_1, \lambda_2)$, which can be viewed as an ellipse with foci $\{\lambda_1, \lambda_2\}$ and minor axis length of zero.

Suppose A is not normal. Without loss of generality, we may assume that $\mathrm{tr}\,A = 0$ (othwerwise, we can replace A with $A - (\mathrm{tr}\,A/2)I_2$ in view of **P1**). By the Schur decomposition theorem, the matrix A is unitarily similar to the matrix

$$B = \begin{bmatrix} \lambda & b \\ 0 & -\lambda \end{bmatrix}, \ b \neq 0.$$

Consider the following cases:

(i) $\lambda = 0$. Without loss of generality, we only consider the case when $b = 1$ (if $b \neq 1$, one may consider the matrix B/b in view of **P1**). A straightforward computation shows that

$$P^\delta = |H_1 u + H_2 v + I_n w| = -\frac{u^2}{4} - \frac{v^2}{4} + w^2.$$

Eliminating the variables u, v, w, and λ from (1), (2), and (3) yields the plane projective curve

$$\kappa_P = \left\{ (x, y, z) \in \mathbb{CP}^2 \mid \frac{x^2}{\frac{1}{4}} + \frac{y^2}{\frac{1}{4}} - z^2 = 0 \right\}.$$

The real plane algebraic curve

$$\mathrm{Re}(\gamma_{P(x,y,1)}) = \left\{ (x, y) \in \mathbb{R}^2 \mid \frac{x^2}{\frac{1}{4}} + \frac{y^2}{\frac{1}{4}} - 1 = 0 \right\}$$

is a circle centered at the origin with diameter

$$2\left(\frac{1}{2}\right) = 1 = \sqrt{\mathrm{tr}\,B^*B}.$$

(ii) $\lambda \neq 0$. Without loss of generality, we only consider the case when $\lambda = 1$ (if $\lambda \neq 1$, one may consider the matrix B/λ in view of **P1**). A straightforward computation shows that

$$P^\delta = |H_1 u + H_2 v + I_n w| = -\left(1 + \frac{b\bar{b}}{4}\right) u^2 - \frac{b\bar{b}}{4} v^2 + w^2.$$

Eliminating the variables u, v, w, and λ from (1), (2), and (3) yields the plane projective curve

$$\kappa_P = \left\{ (x, y, z) \in \mathbb{CP}^2 \mid \frac{x^2}{1 + \frac{b\bar{b}}{4}} + \frac{y^2}{\frac{b\bar{b}}{4}} - z^2 = 0 \right\}.$$

The plane algebraic curve

$$\mathrm{Re}(\gamma_{P(x,y,1)}) = \left\{ (x, y) \in \mathbb{R}^2 \mid \frac{x^2}{1 + \frac{b\bar{b}}{4}} + \frac{y^2}{\frac{b\bar{b}}{4}} - 1 = 0 \right\}$$

is an ellipse centered at the origin with minor axis length equal to

$$2\left(\frac{\sqrt{b\bar{b}}}{2}\right) = b\bar{b} = \sqrt{\mathrm{tr}\,B^*B - 1^2 - (-1)^2}$$

and foci

$$\pm c = \sqrt{\left(1 + \frac{b\bar{b}}{4}\right) - \frac{b\bar{b}}{4}} = \pm 1.$$

3

References

[1] C. Davis. The Toeplitz-Hausdorff theorem explained. *Canad. Math. Bull.*, 14:245–246, 1971.

[2] W. F. Donoghue, Jr. On the numerical range of a bounded operator. *Michigan Math. J.*, 4:261–263, 1957.

[3] R. A. Horn and C. R. Johnson. *Topics in matrix analysis.* Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[4] C. R. Johnson. Computation of the field of values of a 2×2 matrix. *J. Res. Nat. Bur. Standards Sect. B*, 78B:105–107, 1974.

[5] R. Kippenhahn. On the numerical range of a matrix. *Linear Multilinear Algebra*, 56(1-2):185–225, 2008. Translated from the German by Paul F. Zachlin and Michiel E. Hochstenbach [MR0059242].

[6] C.-K. Li. *C*-numerical ranges and *C*-numerical radii. *Linear and Multilinear Algebra*, 37(1-3):51–82, 1994. Special Issue: The numerical range and numerical radius.

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

[8] F. D. Murnaghan. On the field of values of a square matrix. *Proceedings of the National Academy of Sciences of the United States of America*, 18(3):246–248, 1932.

[9] G. Salmon. *A treatise on the higher plane curves: intended as a sequel to "A treatise on conic sections".* 3rd ed. Chelsea Publishing Co., New York, 1960.