The Convex Algebraic Geometry of Higher-Rank Numerical Ranges

Jonathan Andrés Niño-Cortés

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:
Cynthia Vinzant, Chair
Rekha R. Thomas
Jarod Alper

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

The Convex Algebraic Geometry of Higher-Rank Numerical Ranges

Jonathan Andrés Niño-Cortés

Chair of the Supervisory Committee:
Cynthia Vinzant
Mathematics

The higher-rank numerical range is a convex compact set generalizing the classical numerical range of a square complex matrix, first appearing in the study of quantum error correction. In this thesis, we will discuss some of the real algebraic and convex geometry of these sets, including a generalization of Kippenhahn's theorem, and describe an algorithm to explicitly calculate the higher-rank numerical range of a given matrix. We will also discuss the inverse field of values problem, an inverse problem on the numerical range. We focus on the geometric properties of the set of solutions. Finally, we consider an analogous problem for higher-rank numerical ranges and show how to solve it using the ideas behind the proof of convexity for these sets.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank my advisor Cynthia Vinzant for all her support and guidance through my PhD. Also, for introducing me to the double life of numerical ranges back in the course on Convex Algebraic Geometry on Winter 2022. I would also like to thank all the people that have been a part of Vallenato Panamericano. This is a Latinamerican band that a couple of friends and I started in 2022 and that, throughout the years, has grown to a full-size orchestra with over 18 musicians and dancers. They have become my family overseas and practicing and performing with them has given me some of the most endearing memories of my life. Finally, to my gramma Guma who was my first math teacher and to my parents and sisters whom, despite the distance, were always very close to me and supported me in the moments when I needed it the most.

DEDICATION

A Henry, Marla, Cris y Sofi

Chapter 1

INTRODUCTION

The main objects of study of this thesis are numerical ranges of matrices, also known as field of values, and a generalization thereof. Given a complex square matrix $A \in \mathbb{C}^{n\times n}$ we define the numerical range as

$$\mathcal{W}(A) = \{x^*Ax : x \in \mathbb{C}^n \text{ and } x^*x = 1\}. \tag{1.1}$$

This is a set in the plane of convex numbers. The fact that these sets are nonempty, compact and convex is a classical result in functional analysis known as the Toeplitz-Hausdorff theorem. The introduction of these sets was attributed to Toeplitz in 1918 [36] while their convexity was established by Hausdorff in 1919 [16].

It is also worth mentioning that the *spectrum* of A, i.e. the set of eigenvalues of A, is contained in $\mathcal{W}(A)$. This set has been studied extensively in the numerical analysis community and it plays a key role in determining the convergence rate for iterative algorithms that rely on powers and exponentials of A [11].

But numerical ranges have also captured the attention of computational algebraic geometers. A classical result in this field is known as the Kippenhahn theorem [23]. For a complex matrix A, let

$$\Re A = \frac{A + A^*}{2} \quad \text{and} \quad \Im A = \frac{A - A^*}{2\mathbb{i}}.$$

Observe that $\Re A$ and $\Im A$ are Hermitian matrices and also $A = \Re A + \mathbb{i}\Im A$. Kippenhanh showed that $\mathcal{W}(A)$ is the convex hull of a real algebraic curve. This curve is dual to the algebraic curve defined by the vanishing of the *Kippenhahn polynomial* of A, namely

$$f_A = \det\left(tI_n + x\Re(A) + y\Im(A)\right)$$

More details on projective duality and how to compute a generator for this dual curve are given in Section 2.5.

Given a matrix $A \in \mathbb{C}^{n\times n}$ and a positive integer $1 \le k \le n$, the *rank-k* numerical range of A is defined to be

$$\Lambda_k(A) = \{\mu \in \mathbb{C} : \exists \text{ a rank-}k \text{ orthogonal projection } P \text{ s.t. } PAP = \mu P\}. \tag{1.2}$$

In the context of quantum error correction, elements $\mu \in \Lambda_k(A)$ are known as compression values [9]. An equivalent description of this set is

$$\Lambda_k(A) = \left\{\mu \in \mathbb{C} : \exists X \in \mathbb{C}^{n\times k} \text{ such that } X^*X = I_k \text{ and } X^*AX = \mu I_k\right\}. \tag{1.3}$$

From this description, one can see that the higher-rank numerical ranges are nested and that the rank-one numerical range coincides with $\mathcal{W}(A)$. That is, $\mathcal{W}(A) = \Lambda_1(A)$ and $\Lambda_k(A) \supseteq \Lambda_{k+1}(A)$ for all k. Higher-rank numerical ranges have special behavior when f_A is a power of a single linear form, i.e. $f_A = (t + ax + by)^n$. In this case $\Lambda_k(A) = \{a + \mathbb{i}b\}$ for all $k = 1, \ldots, n$ [8, Prop. 2.2].

Choi et. al. conjectured that higher-rank numerical ranges are always convex, which was proved by Woerdeman in 2008 [37]. Independently, Li and Sze also proved that $\Lambda_k(A)$ is convex [26] by proving a description as an intersection of half-planes, namely

$$\Lambda_k(A) = \bigcap_{\theta\in[0,2\pi)} \left\{\mu \in \mathbb{C} : \mathrm{Re}(e^{-i\theta}\mu) \le \lambda_k(\Re(e^{-i\theta}A))\right\} \tag{1.4}$$

where $\lambda_1(M) \ge \ldots \ge \lambda_n(M)$ denote the real eigenvalues of a Hermitian matrix M.

Example 1.1. The 4×4 weighted cyclic shift matrix

$$A = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 6 \\ 8 & 0 & 0 & 0 \end{pmatrix} \quad \text{gives} \quad f_A = \det \begin{pmatrix} t & x - \mathbb{i}y & 0 & 4x + 4\mathbb{i}y \\ x + \mathbb{i}y & t & 2x - 2\mathbb{i}y & 0 \\ 0 & 2x + 2\mathbb{i}y & t & 3x - 3\mathbb{i}y \\ 4x - 4\mathbb{i}y & 0 & 3x + 3\mathbb{i}y & t \end{pmatrix}$$

$$= 25(x^4 + y^4) + 434x^2y^2 - 30(x^2 + y^2)^2 + t^4.$$

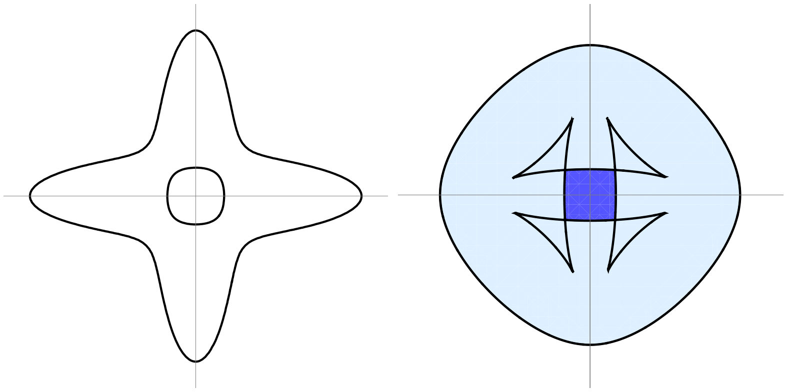

Figure 1.1: The curve $f_A = 0$ and numerical ranges $\Lambda_1(A)$ and $\Lambda_2(A)$ of the 4×4 matrix A from Example 1.1. The zero set of the polynomial g_A vanishes on the boundaries of $\Lambda_1(A)$ and $\Lambda_2(A)$.

Figure 1.1 shows the zero set of $f_A(1, x, y)$ together with rank-one and rank-two numerical ranges of A. The dual curve of f_A is defined by the polynomial

$$\begin{aligned} g_A = {} & 15625a^{12} + 273750a^{10}b^2 + 90375a^8b^4 + 549236a^6b^6 + 90375a^4b^8 + 273750a^2b^{10} + 15625b^{12} \\ & - 1368750a^{10} - 17139750a^8b^2 + 44934900a^6b^4 + 44934900a^4b^6 - 17139750a^2b^8 - 1368750b^{10} \\ & + 47610625a^8 + 429249700a^6b^2 - 1058169786a^4b^4 + 429249700a^2b^6 + 47610625b^8 \\ & - 838188000a^6 - 5975989920a^4b^2 - 5975989920a^2b^4 - 838188000b^6 + 7621461600a^4 \\ & + 39076977600a^2b^2 + 7621461600b^4 - 30526848000a^2 - 30526848000b^2 + 21083040000. \end{aligned}$$

By Kippenhahn's theorem, $\mathcal{W}(A)$ is the convex hull of $\{a + \mathbb{i}b \in \mathbb{C} : g_A(a, b) = 0\}$.

Higher-rank numerical ranges arose in the study of quantum error-correcting codes. In this context, errors are represented as sets of operators that are applied to vectors in the Hilbert space of possible quantum states. The goal is to find a projection that makes the action of these operators on the data negligible, usually at the expense of restricting to a subspace, thereby lowering the capacity of the channel. Each projection is associated with a compression value, and the set of all possible compression values for rank-k projections corresponds to the rank-k numerical range. See [8], [9] for more details. Gau and Wu [13] showed that the Kippenhahn polynomial f_A of a matrix A completely determines all of its higher-rank numerical ranges and, conversely, any two matrices all of whose higher-rank numerical ranges coincide have the same Kippenhahn polynomial. Bebiano, Providência, and Spitkovsky [5] give a description of the rank-k numerical range for certain families of matrices.

The contents of this thesis are organized as follows. In Chapter 2, we give an overview of some of the basic concepts and theorems used throughout this work. We cover basic results in linear algebra, algebraic geometry, and convex geometry. We finish this section by giving a proof of the classical Toeptlitz-Hausdorff theorem using the concepts and theorems introduced in the preliminaries.

In Chapter 3, we study how to represent higher-rank numerical ranges and give algorithms to compute such representations. Emphasis is placed on how to compute higher-rank numerical ranges when they are not full-dimensional.

In Chapter 4, we look at the *inverse Field of Values Problem* (iFVP). This is well studied problem in numerical analysis (see [4] and [6]). We focus on the set of solutions of the iFVP and try to determine its algebraic structure and properties. One of our main tools is the *-congruence canonical form of a complex squared matrix which is introduced at the beginning of the section.

Chapter 5 extends the iFVP to higher-rank numerical ranges. In this chapter, we formulate this problem and give a symbolic method to obtain a solution. This solution is obtained by studying a proof for convexity of the higher-rank numerical range in depth.

Chapter 6 covers a handful of results that do not fit the general themes of the other chapters but are nonetheless valuable in their own right. We end with a recompilation of open questions in the field that might be of use to the interested reader.

Chapter 2

PRELIMINARIES

2.1 Notation

This is some of the notation that we will be using throughout this thesis.

- $\mathbb{R}$ and $\mathbb{C}$ denote the fields of real and complex numbers, respectively. $\mathbb{Z}$ denotes the ring of integers and $\mathbb{N}$ the set of natural numbers (including 0).
- $\mathbb{R}_{\geq 0}$ denotes the set of nonnegative real numbers.
- $[n]$ denotes the set of natural numbers from 1 to n where $n \in \mathbb{N}$.
- $\mathbb{i}$ denotes the complex square root of -1 .
- For a field $\mathbb{K}$, we use $\mathbb{P}^n(\mathbb{K})$ to denote the projective space of dimension n over $\mathbb{K}$. Primarily, we would be working with $\mathbb{P}^n(\mathbb{R})$ and $\mathbb{P}^n(\mathbb{C})$.
- The identity matrix in $\mathbb{C}^{n\times n}$ ($\mathbb{R}^{n\times n}$) is denoted by I_n. The zero matrix in $\mathbb{C}^{n\times n}$ ($\mathbb{R}^{n\times n}$) is denoted by 0_n.
- For a Hermitian matrix M, we use $\lambda_1(M) \geq \ldots \geq \lambda_n(M)$ to denote the eigenvalues of M. We often fix a matrix A and use the notation

  $$\lambda_k(\theta) = \lambda_k(\Re(e^{-i\theta}A)) = \lambda_k(\cos(\theta)\Re A + \sin(\theta)\Im A) \quad \text{and} \quad \lambda_k(x,y) = \lambda_k(x\Re A + y\Im A),$$

  where $\theta \in [0, 2\pi)$ and $(x, y) \in \mathbb{R}$. Any $(x, y) \in \mathbb{R}^2$ can be written as $(x, y) = r(\cos(\theta), \sin(\theta))$ for some $r \geq 0$ and $\theta \in [0, 2\pi)$, in which case $\lambda_k(x, y) = r\lambda_k(\theta)$.

2.2 Real structure of a complex vector space

The vector space $\mathbb{C}^n$ also has an underlying structure as a real vector space of dimension $2n$. If $\{e_1, \ldots, e_n\}$ is the canonical basis for $\mathbb{C}^n$, then a basis for the underlying real vector space consists of the vectors $\{e_1, \ldots, e_n, \mathbb{i}e_1, \ldots, \mathbb{i}e_n\}$.

Let $M = A + \mathbb{i}B$ be a complex square matrix representing a linear automorphism of $\mathbb{C}^n$, with $A, B \in \mathbb{R}^{n\times n}$. This map corresponds to an automorphism of $\mathbb{R}^{2n}$ with its respective matrix in $\mathbb{R}^{2n\times 2n}$ which we denote by $\widehat{M}$. Given our choice of basis, this matrix would be

$$\widehat{M} = \begin{bmatrix} A & -B \\ B & A \end{bmatrix}. \tag{2.1}$$

For a Hermitian matrix H, consider the map $f_H : V \to \mathbb{R}$ given by $w_H(v) = v^*Hv$. This is a quadratic form over $\mathbb{R}^{2n}$. The next lemma shows that the associated symmetric matrix to this quadratic form is in fact $\widehat{H}$.

Lemma 2.1. *Let H be a Hermitian matrix in $\mathbb{C}^{n\times n}$ and let $z = (x_1 + \mathbb{i}y_1, \ldots, x_n + \mathbb{i}y_n) \in \mathbb{C}^n$ and $\mathbf{z} = (x_1, \ldots, x_n, y_1, \ldots, y_n) \in \mathbb{R}^{2n}$. Then $z^*Hz = \mathbf{z}^\top \widehat{H}\mathbf{z}$. Moreover,* $\operatorname{rank}(\widehat{H}) = 2\operatorname{rank} H$.

Proof. Let $H = A + \mathbb{i}B$ with A, $B \in \mathbb{R}^{n\times n}$. Since H is Hermitian, we must have that A is symmetric and B is antisymmetric. Then

$$\begin{aligned} z^*Hz &= (x - \mathbb{i}y)^\top (A + \mathbb{i}B)(x + \mathbb{i}y) \\ &= x^\top Ax + \underline{\mathbb{i}x^\top Ay - \mathbb{i}y^\top Ax} + y^\top Ay + \underline{\mathbb{i}x^\top Bx} - x^\top By + y^\top Bx + \underline{\mathbb{i}y^\top By} \\ &= x^\top Ax + y^\top Ay - x^\top By + y^\top Bx. \end{aligned}$$

The first two underlined terms above cancel by the symmetry of A since

$$x^\top Ay = y^\top Ax.$$

The last two underlined terms also cancel by the antisymmetry of B since for any vector $v \in \mathbb{R}^n$ and any antisymmetric matrix $B \in \mathbb{R}^{n\times n}$ we have

$$v^\top Bv = -v^\top Bv = 0.$$

On the other hand,

$$\mathbf{z}^\top \widehat{H}\mathbf{z} = \begin{pmatrix} x^\top & y^\top \end{pmatrix} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^\top Ax + y^\top Ay - x^\top By + y^\top Bx.$$

To check the rank of $\widehat{H}$, set $S = \begin{pmatrix} -\mathbb{i}I_n & \mathbb{i}I_n \\ I_n & I_n \end{pmatrix}$. Then we can check that

$$S^{-1} \begin{pmatrix} A & -B \\ B & A \end{pmatrix} S = \begin{pmatrix} A - \mathbb{i}B & 0 \\ 0 & A + \mathbb{i}B \end{pmatrix} = \overline{H} \oplus H.$$

Which clearly has rank equal to $2\,\mathrm{rank}(H)$. Since S is invertible and rank is preserved by conjugation with S, this is also the rank of $\widehat{H}$. □

To complement our discussion, we give a counterpart to the process of obtaining $\mathbb{R}^{2n}$ from $\mathbb{C}^n$. Starting with the vector space $\mathbb{R}^{2n}$ we can give it the structure of an n-dimensional complex vector space via a construction called the *linear complex structure* [24, Chapter IX, Section 1]. This involves finding a matrix $J \in \mathbb{R}^{2n\times 2n}$ such that $J^2 = -I$. Then the scalar multiplication in $\mathbb{C}$ is defined as $(x + iy)v = xv + yJ(v)$. Commonly, J is taken to be the matrix

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

Observe that the matrices $\widehat{M}$ we defined in Eq. (2.1) satisfy the property that $J\widehat{M} = \widehat{M}J$. This characterizes all matrices that correspond to linear maps of the linear complex structure of $\mathbb{R}^{2n}$.

2.3 Convex geometry

A set $S \subset \mathbb{R}^n$ is *convex* if for any points $\mathbf{x}, \mathbf{y} \in S$, the line segment between them, $\{\lambda\mathbf{x} + (1-\lambda)\mathbf{y} : 0 \le \lambda \le 1\}$, is contained in S. We call S a *convex cone* if it is also closed under nonnegative scaling, or equivalently if for any two points $\mathbf{x}, \mathbf{y} \in S$, the set of conic combinations $\{\lambda\mathbf{x} + \mu\mathbf{y} : \lambda, \mu \in \mathbb{R}_{\ge 0}\}$ is contained in S. We call $F \subseteq S$ a *face* of a convex set if for every $\mathbf{p} \in F$, if $\mathbf{p} = \lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ for some $\lambda \in (0,1)$ where $\mathbf{x}, \mathbf{y} \in S$, then $\mathbf{x}, \mathbf{y} \in F$. That is, any way to express a point from F as a convex combination of points from S involves only points from F.

The *convex hull* of a set $S \subset \mathbb{R}^n$ is the smallest convex set containing S and its *conic hull* is the smallest convex cone containing S. We can write these as

$$\operatorname{conv}(S) = \left\{ \sum_{i=1}^{k} \lambda_i p_i : \ k \in \mathbb{N},\ p_i \in S,\ \lambda_i \geq 0,\ \sum_{i=1}^{k} \lambda_i = 1 \right\}, \text{ and}$$

$$\operatorname{conicalHull}(S) = \left\{ \sum_{i=1}^{k} \lambda_i p_i : \ k \in \mathbb{N},\ p_i \in S,\ \lambda_i \geq 0 \right\}.$$

It is often more convenient to work with convex cones than convex sets. Following the notation of [34], for any set $S \subseteq \mathbb{R}^n$, we define $\widehat{S}$ to be the conical hull of S embedded into $\mathbb{R}^{n+1}$ at height one:

$$\widehat{S} = \operatorname{conicalHull}(\{1\} \times S) = \left\{ \sum_{i=1}^{k} \lambda_i (1, p_i) : \ k \in \mathbb{N},\ p_i \in S,\ \lambda_i \geq 0 \right\} \subset \mathbb{R}^{n+1}.$$

For any convex cone C, the set of linear functions that take only nonnegative values on C also forms a convex cone, known as the *dual cone* of C, denoted C^*. For $C \subseteq \mathbb{R}^n$ this is

$$C^* = \{w \in \mathbb{R}^n : \langle w, v \rangle \geq 0 \text{ for all } v \in C\}.$$

This is a closed convex cone in $\mathbb{R}^n$. It is a classical theorem in convex geometry that the dual of C^* coincides with the closure of C in the Euclidean topology. See Theorem 1.2, the subsequent Problem 2, and Lemma 1.4 of [2, Section IV.1]. This is known as the biduality theorem:

Theorem 2.2 (Biduality). *For any nonempty convex cone C, $(C^*)^* = \overline{C}$.*

Note that when $C = \emptyset$, $C^* = \mathbb{R}^n$ and $(C^*)^* = \{(0, \ldots, 0)\}$. From an inequality description of a cone C, we can therefore understand its dual.

Corollary 2.3. *Let $K = \{\mathbf{a} \in \mathbb{R}^n : \langle \mathbf{x}(\theta), \mathbf{a} \rangle \geq 0$ for all $\theta \in \Theta\}$ where $\mathbf{x}(\theta) \in \mathbb{R}^n$. The dual cone of K is*

$$K^* = \overline{\operatorname{conicalHull}\{\mathbf{x}(\theta) : \theta \in \Theta\}}.$$

Proof. The condition that $\langle \mathbf{x}(\theta), \mathbf{a} \rangle \geq 0$ for all $\theta \in \Theta$ is equivalent to the condition that $\langle \mathbf{x}, \mathbf{a} \rangle \geq 0$ for all $\mathbf{x}$ in the conical hull of $\{\mathbf{x}(\theta) : \theta \in \Theta\}$. This shows that K is the dual cone of $C = \text{conicalHull}\{\mathbf{x}(\theta) : \theta \in \Theta\}$. The result then follows from the biduality theorem. $\square$

A set in $\mathbb{R}^n$ is called an *affine space* if it can be written as $W + x$ where $x \in \mathbb{R}^n$ and W is a linear subspace of $\mathbb{R}^n$. The dimension of $x + W$ is defined to be the dimension of W as a linear subspace. The *affine span* (also known as *affine hull*) of a set $S \in \mathbb{R}^n$ is the intersection of all affine subspaces that contain S.

$$\text{aff}(S) = \left\{ \sum_{i=1}^{k} \lambda_i p_i : k \in \mathbb{N},\ p_i \in S,\ \lambda_i \in \mathbb{R},\ \sum_{i=1}^{k} \lambda_i = 1 \right\}$$

The dimension of a convex set S is defined to be the dimension of its affine span.

The *relative interior* of a set S is its interior relative to the affine span of S.

$$\text{ri}(S) = \{x \in S : \exists \epsilon > 0 \text{ such that } B(x, \epsilon) \cap \text{aff}(S) \subseteq S\},$$

where $B(x, \epsilon)$ denotes the open ball of radius epsilon around x.

$$B(x, \epsilon) = \{y \in \mathbb{R}^n : \|x - y\|_2 < \epsilon\}.$$

The following property of relative interiors would be useful for our purposes.

Theorem 2.4 (Theorem 6.6 in [32])**.** *Let K be convex set in $\mathbb{R}^n$ and $\psi : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Then $\psi(\text{ri}(K)) = \text{ri}(\psi(K))$.*

2.4 Properties of the PSD cone of Hermitian matrices

This section follows Barvinok's discussion about the cone of positive semidefinite matrices [2, Chapter II, Section 12]. Here, as in most references, only the PSD cone for real symmetric matrices is studied in detail. All analogous statements on the PSD cone of Hermitian matrices are left as exercises to the reader. For completeness, we give solutions to some of these exercises.

For any complex square matrix A we denote its conjugate transpose by A^*. A is *Hermitian* if $A = A^*$. We denote the set of Hermitian matrices of size n by Herm_n.

Theorem 2.5. *The set of $n \times n$ Hermitian matrices of size n is a real vector space of dimension n^2.*

Proof. Let A, B be two Hermitian matrices and $b \in \mathbb{R}$. Observe that $(A + bB)^* = A^* + (bB)^* = A+bB$. This shows that Hermitian matrices are closed under addition and real scalar multiplication. Denote the elementary matrix with 1 in the ij entry and zero everywhere else by E_{ij}. To check that the dimension is n^2 observe that a basis of $\mathrm{Herm_n}$ is given by the matrices E_{ii} for $1 \leq i \leq n$, and $E_{ij} + E_{ji}$, $\mathbb{i}E_{ij} - \mathbb{i}E_{ji}$, $1 \leq i < j \leq n$. The basis has size $n + 2\binom{n}{2} = n^2$. $\square$

This space comes equipped with an inner product

$$\langle A, B\rangle = \mathrm{tr}(AB) = \sum_{1\leq i,j\leq n} \overline{a_{ij}} b_{ij}.$$

A Hermitian matrix $A \subseteq \mathbb{R}^{n\times n}$ is *positive semidefinite* (PSD) if all of its eigenvalues are non-negative. If all of its eigenvalues are strictly positive then we say that the matrix is *positive definite* (PD).

Theorem 2.6. *The following conditions are equivalent.*

1. *A is* PSD.

2. $x^* Ax \geq 0$ *for all* $x \in \mathbb{C}^n$.

3. *All principal minors of A are nonnegative.*

4. $A = B^* B$ *for some matrix* $B \in \mathbb{C}^{n\times n}$.

The PD property has very similar equivalences.

Theorem 2.7. *The following conditions are equivalent.*

1. *A is* PD.

2. *$x^*Ax > 0$ for all $x \in \mathbb{C}^n$.*

3. *All leading principal minors of A are strictly positive.*

4. *$A = B^*B$ for some full-rank matrix $B \in \mathbb{C}^{n\times n}$.*

To check if a matrix is PD using (3) it is only necessary to check the leading principal minors rather than all of them. Computationally, this is a big difference, since there are n leading principal minors and $2^n - 1$ principal minors in total.

We use the PSD property to introduce a partial ordering in Herm_n. We say that $A \succeq B$ if $A - B$ is PSD. In particular, A is PSD if and only if $A \succeq 0$. We denote the set of all Hermitian PSD matrices of size n as $\mathrm{PSD}_n(\mathbb{C})$. Building on our previous section, we will show that this is a convex cone inside the real vector space of Hermitian matrices.

Theorem 2.8. *The set of positive semidefinite matrices is a closed pointed convex cone in* Herm_n.

Proof. To show that PSD_n is a cone we can use the second characterization in Theorem 2.8. For any two PSD matrices A and B and for any two scalars $a, b \in \mathbb{R}_{\geq 0}$ we have

$$x^*\left(aA + bB\right)x = ax^*Ax + bx^*Bx \geq 0.$$

Observe that for a fixed nonzero $x \in \mathbb{C}^n$, $\mathrm{PSD}_n(\mathbb{C})$ satisfies the closed constraint $x^*Hx \geq 0$. Moreover,

$$x^*Ax = \mathrm{tr}(xx^*A) = \langle xx^*, A\rangle.$$

Therefore, $\mathrm{PSD}_n(\mathbb{C})$ is equal to the intersection of the half-spaces $\langle xx^*, A\rangle \geq 0$ for all $x \in \mathbb{C}^n \backslash \mathbf{0}$. In particular, $\mathrm{PSD}_n(\mathbb{C})$ is closed.

Finally, we can check that $\mathrm{PSD}_n(\mathbb{C})$ is a pointed cone by observing that in $\mathrm{PSD}_n(\mathbb{C})$ there is no straight line that passes through the origin. This is not possible because for any nonzero PSD matrix A, $-A$ has at least one strictly negative eigenvalue and therefore $-A$ is not in $\mathrm{PSD}_n(\mathbb{C})$. $\square$

Theorem 2.9. $\mathrm{PSD}_n(\mathbb{C})$ *is generated by the set of rank-1* $n \times n$ *Hermitian matrices.*

Proof. Observe that all rank-1 Hermitian matrices can be written as vv^* up to sign for some vector v. Let H be a Hermitian matrix and let $\{\lambda_1, \ldots, \lambda_n\}$ be its eigenvalues. Then there exists a unitary matrix U such that $UDU^* = H$ with $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Let u_i denote the i-th column of U. Then observe that u_i is an eigenvector of H.

$$Hu_i = UDU^*u_i = UDe_i = U\lambda_i e_i = \lambda u_i.$$

Then we claim that

$$H = \sum_{j=1}^{n} \lambda_j u_j u_j^*. \tag{2.2}$$

Since the u_i form a basis of $\mathbb{C}^n$ we can show this by checking that

$$\left(\sum_{j=1}^{n} \lambda_i u_j u_j^* \right) u_i = \lambda_i u_i,$$

for all $i \in [n]$. But this is immediate from the fact that u_i is a unitary basis, so $u_j^* u_i = 0$ for all $j \neq i$ and the only surviving term above is $\lambda_i u_i u_i^* u_i = \lambda_i u_i$. □

The expression on the right-hand side of Eq. (2.2) is known as the rank-1 decomposition of H. If we assume that H is PSD then its rank-1 decomposition is just a conical combination of rank-1 PSD matrices.

Theorem 2.10. *The faces of* PSD_n *are indexed by the subspaces of* $\mathbb{C}^n$. *They are of the form*

$$F_V = \{A \in \mathrm{PSD}_n : \ker A \supset V\}.$$

where V *is a subspace of* $\mathbb{C}^n$.

Proof. A proof for the symmetric case is given in [2, Proposition 12.3]. The proof for Hermitian matrices is analogous. □

Theorem 2.11. *A Hermitian matrix* A *is in the interior of* $\mathrm{PSD}_n(\mathbb{C})$ *if and only if* A *is positive definite.*

Proof. We can use the characterization that A is PD if $\det(A_k) > 0$ for all leading principal submatrices A_k of A. The determinant is a continuous function from Herm_k to $\mathbb{R}$. So PD_n is the finite intersection of preimages of open sets and therefore it is open. The other direction follows from Theorem 2.10 because if a Hermitian matrix A is PSD and it is not full-rank then its kernel is non trivial and therefore it must be contained in the proper face of PSD_n indexed by $\ker A$. $\square$

Corollary 2.12. $\mathrm{PSD}_n(\mathbb{C})$ *is a full-dimensional cone.*

Proof. This follows from the previous lemma because the interior of $\mathrm{PSD}_n(\mathbb{C})$ is nonempty. In particular I_n is an interior point of $\mathrm{PSD}_n(\mathbb{C})$. $\square$

A *spectahedron* is the intersection of PSD_n with an affine linear subspace of Herm_n. For example, for any $n \times n$ Hermitian matrices X and Y, the set $\{I_n + aX + bY \succeq 0 : (a, b) \in \mathbb{R}^2\}$ is a spectahedron. Other examples are sets of the form $\{H \in \mathrm{PSD}_n(\mathbb{C}) : \langle X, H\rangle = b\}$ for $X \in \mathrm{Herm}_n$ and $b \in \mathbb{R}$. Spectahedrons are convex because they are the intersection of two convex sets ($\mathrm{PSD}_n(\mathbb{C})$ and an affine linear subspace).

On the other hand a linear tranformation of a spectahedron is called a *spectahedral shadow*. For example, for X and Y as above

$$\{(\langle X, H\rangle, \langle Y, H\rangle) : H \in \mathrm{PSD}_n \text{ and } \langle H, I_n\rangle = 1\}$$

is a spectahedral shadow. Spectahedral shadows are also convex since convexity is preserved under affine linear maps.

2.5 Duality of plane curves

Let $f \in \mathbb{R}[t, x, y]$ be a homogeneous polynomial of degree n. That is, $f = \sum_{i+j \le n} c_{ij} t^i x^j y^{n-i-j}$ for some constants $c_{ij} \in \mathbb{R}$. We use $\mathcal{V}(f)$ and $\mathcal{V}_{\mathbb{R}}(f)$ to denote the variety of f in $\mathbb{P}^2(\mathbb{C})$ and $\mathbb{P}^2(\mathbb{R})$, respectively. Up to scaling, the polynomial f has a unique factorization into irreducible polynomials $f = \prod_{i=1}^{s} f_i^{m_i}$ where $\gcd(f_i, f_j) = 1$ for $i \neq j$. We use f^{red} to denote the square-free product $\prod_{i=1}^{s} f_i$. Then $\mathcal{V}(f) = \mathcal{V}(f^{\mathrm{red}})$.

A point $p \in \mathcal{V}(f)$ is called a *smooth point* of $\mathcal{V}(f)$ if the gradient $\nabla(f^{\text{red}})$ is nonzero at p, in which case $\frac{\partial f^{\text{red}}}{\partial t}(p)t + \frac{\partial f^{\text{red}}}{\partial x}(p)x + \frac{\partial f^{\text{red}}}{\partial y}(p)y = 0$ defines the tangent line of $\mathcal{V}(f)$ at p. Otherwise, we call p a *singular point* of $\mathcal{V}(f)$. We use $\operatorname{Sing}(f)$ and $\operatorname{Sing}_{\mathbb{R}}(f)$ to denote the set of singular points of $\mathcal{V}(f)$ and $\mathcal{V}_{\mathbb{R}}(f)$, respectively.

For $\mathbb{P}^n(\mathbb{C})$, the set of hyperplanes $H \subseteq$ has the structure of a projective space. We call this set the *dual projective space* of dimension n and denoted by $\mathbb{P}^n(\mathbb{C})^*$. This set is, in fact, isomorphic to $\mathbb{P}^n(\mathbb{C})$ via the following map. For any hyperplane H, map H to its normal vector.

The *dual variety* of $\mathcal{V}(f)$, denoted $\mathcal{V}(f)^*$, is defined as the closure of the image of $\mathcal{V}(f) \setminus \operatorname{Sing}(f)$ under the map $p \mapsto [\nabla f^{\text{red}}(p)]$. That is,

$$\mathcal{V}(f)^* = \left\{[c : a : b] \in \mathbb{P}^2(\mathbb{C}) : ct + ax + by = 0 \text{ is tangent to } \mathcal{V}(f) \text{ at some point } p\right\}.$$

When f is irreducible and has degree ≥ 2, $\mathcal{V}(f)^*$ is an irreducible plane curve. When $f = ct + ax + by$, the dual variety $\mathcal{V}(f)^*$ is a single point $[c : a : b]$. In general, if $f = \prod_{i=1}^{s} f_i^{m_i}$ is an irreducible factorization of f, the dual variety of f is the union of the dual varieties of its irreducible factors, $\mathcal{V}(f)^* = \cup_{i=1}^{s} \mathcal{V}(f_i)^*$.

2.6 Saturation and variable elimination

This is an overview of the algorithms an operations on ideals described in [10]. Consider two ideals I and J in $\mathbb{R}[x_1, \ldots, x_n]$. The saturation of J over I, denoted by $(I : J^\infty)$ is defined as

$$(I : J^\infty) = \bigcup_{k=1}^{\infty} (I : J^k).$$

By the ascending chain condition, the chain of ideals $(I : J^k)$ stabilizes and so $(I : J^\infty)$ is equal to $(I : J^k)$ for big enough k. The saturation ideal has the property that $\mathcal{V}(I : J^\infty) = \overline{\mathcal{V}(I) \setminus \mathcal{V}(J)}$. Thus, saturation is useful when we want to remove irreducible components from a variety.

Theorem 2.13 (Elimination of variables)**.** *Let I be a homogeneous ideal in* $\mathbb{C}[x_0, \ldots, x_m, y_0, \ldots, y_m]$ *and consider the variety* $\mathcal{V}(I) = \mathcal{P} \in \mathbb{P}^n(\mathbb{C}) \times \mathbb{P}^m(\mathbb{C})$*. Then*

$$\pi_1(\mathcal{P}) = \mathcal{V}((I : (y_0, \ldots, y_m)^\infty \cap \mathbb{C}[x_1, \ldots, x_n]) ,$$

where π_1 *denotes the projection of* $\mathbb{P}^n(\mathbb{C}) \times \mathbb{P}^m(\mathbb{C})$ *onto* $\mathbb{P}^n(\mathbb{C})$*.*

Proof. See [10, Chapter 8, Section 5, Proposition 5]. □

Given $f \in \mathbb{R}[t, x, y]$ we would like to find an ideal $I \subseteq \mathbb{R}[a, b, c]$ such that $\mathcal{V}(I) = \mathcal{V}(f)^*$. For this we introduce a new variety in $\mathbb{P}^n \times (\mathbb{P}^n)^*$ called the *conormal variety.*

$$\mathrm{CN}(f) = \{(x, H) : x \in \mathcal{V}(f) \backslash \mathrm{Sing}(f) \text{ and } H \supseteq T_x(\mathcal{V}(f))\}.$$

This consists of pairs of points p in the variety $\mathcal{V}(f)$ and hyperplanes that are tangent at $\mathcal{V}(f)$ on p. The projection of $\mathrm{CN}(X)$ onto $(\mathbb{P}^n)^*$ is the dual curve.

Consider the ideal I,

$$I = \langle f,\ ct + ax + by\rangle + \left\langle 2 \times 2 \text{ minors} \begin{pmatrix} c & a & b \\ \frac{\partial f}{\partial t} & \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \right\rangle \tag{2.3}$$

By saturating I with the ideal $J = \mathcal{I}(\mathrm{Sing}(f)) = \langle \frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle$, we obtain the vanishing ideal for $\mathrm{CN}(f)$.

$$\mathcal{I}(\mathrm{CN}(f)) = (I : J^\infty).$$

We can get the ideal of a projection through variable elimination. In particular, if we eliminate the variables t, x, y we obtain the ideal for $\mathcal{V}(f)^*$. If $\mathcal{V}(f)$ is not a line, then the above variety has dimension 1 and its defining ideal is principal.

2.7 Coordinate changes

Numerical ranges behave nicely under affine-linear maps on $\mathbb{C} \cong \mathbb{R}^2$. Consider an affine-linear map $L : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$L(a, b) = (u_{01} + u_{11}a + u_{21}b, u_{02} + u_{12}a + u_{22}b), \tag{2.4}$$

where $u_{ij} \in \mathbb{R}$. We can extend this to an action of the affine-linear group on $n \times n$ matrices as follows. Given a matrix $A \in \mathbb{C}^{n\times n}$, define

$$L \cdot A = (u_{01} + \mathbb{i}u_{02})I + (u_{11} + \mathbb{i}u_{12})\Re(A) + (u_{21} + \mathbb{i}u_{22})\Im(A).$$

One can check that $L(\Lambda_k(A)) = \Lambda_k(L \cdot A)$. Note that this also corresponds to a linear change of coordinates on the Kippenhahn polynomial:

$$f_{L\cdot A}(t, x, y) = f_A\left(t + u_{01}x + u_{02}y, u_{11}x + u_{12}y, u_{21}x + u_{22}y\right).$$

For an $n \times n$ matrix M, we defined the corank of M to be $n - \operatorname{rank}(M)$.

Lemma 2.14. *For $A \in \mathbb{C}^{n\times n}$ and $p = (p_0, p_1, p_2) \in \mathbb{R}^3$ with $(p_1, p_2) \neq (0, 0)$, the following quantities coincide:*

(i) the corank of the matrix $p_0I + p_1\Re(A) + p_2\Im(A)$,

(ii) the maximum d such that r^d divides $f_A(rq + sp) \in \mathbb{R}[r, s]$ for every $q \in \mathbb{R}^3$, and

(iii) the minimum degree of a monomial of $f_{L\cdot A}(t, x, 1)$ where $L : \mathbb{R}^2 \to \mathbb{R}^2$ is any invertible affine linear transformation as in (2.4) *with $(u_{02}, u_{12}, u_{22}) = p$.*

This number is called the *multiplicity* of f_A at p. The multiplicity of f_A is ≥ 1 if and only if $f_A(p) = 0$ and ≥ 2 if and only if both $f_A(p) = 0$ and $\nabla f_A(p) = 0$.

Proof. For $(t, x, y) \in \mathbb{R}^3$, let $M_A(t, x, y)$ denote the linear matrix pencil

$$M_A(t, x, y) = tI + x\Re(A) + y\Im(A).$$

(i)=(iii) For an invertible affine linear transformation $L : \mathbb{R}^2 \to \mathbb{R}^2$ as in (2.4) with $(u_{02}, u_{12}, u_{22}) = p$, $\Im(L \cdot A) = M_A(p)$. Therefore

$$f_{L\cdot A}(t, x, y) = \det(tI + x\Re(L \cdot A) + yM_A(p)).$$

By the Laplace expansion of the determinant, we see that the degree of $f_{L\cdot A}$ in the variable y, denoted $\deg_y(f_{L\cdot A})$, is $\leq \operatorname{rank}(M_A(p))$. Conversely, the restriction $f_{L\cdot A}(t,0,y)$ factors as $t^d\prod_i(t+\lambda_i y)$ where the product is taken over the nonzero eigenvalues of $M_A(p)$. This shows that $\deg_y(f_{L\cdot A}) \geq \operatorname{rank}(M_A(p))$. Since $f_{L\cdot A}(t,x,y)$ is homogeneous of degree n, $n-\deg_y(f_{L\cdot A})$ is the quantity in (iii). Similarly $n-\operatorname{rank}(M_A(p))$ is the quantity in (i).

(i)=(ii) Similarly, for $q \in \mathbb{R}^3$, $f_A(rq+sp) = \det(rM_A(q)+sM_A(p))$. By the Laplace expansion of the determinant, we see that the degree of s in this polynomial is $\leq \operatorname{rank}(M_A(p))$. Since it is homogeneous of degree n in r, s, it is therefore divisible by r^d where $d = n - \operatorname{rank}(M_A(p))$. Conversely, for $q = (1,0,0)$, this restriction factors as $r^d\prod_i(r+\lambda_i s)$ where the product is taken over the nonzero eigenvalues of $M_A(p)$. $\square$

2.8 Convexity of the numerical range

Recall that the numerical range of a complex squared matrix A is given by

$$\mathcal{W}(A) = \{x^*Ax : x \in \mathbb{C}^n \text{ and } x^*x = 1\}.$$

Throughout our work, we have come across numerous proofs of the convexity of the numerical range. In Barvinok's book, a proof is given that relies on properties of the PSD cone [2, Theorem 14.2]. The crux of the proof is showing that under suitable conditions a Hermitian matrix that satisfies a linear equation can be replaced with another Hermitian matrix of relatively low rank [2, Propostion 13.4]. This follows from the topological fact that there are no continuous injective maps between $\mathbb{S}^n$ and $\mathbb{P}_n(\mathbb{R})$ [2, Topological Fact 13.5].

Another proof using numerical analysis is given in Horn's book [20, Section 1.3]. The key step is to reduce the problem to the case of 2 by 2 matrices. For such matrices, it is explicitly shown that their numerical ranges are ellipses (possibly degenerate) and, therefore, convex.

Here we give a unified proof that combines ideas from the two references cited above. Specifically, we use the reduction to the 2 by 2 case but we show that convexity in this case follows easily from the results shown in the preliminary section on PSD cones (no topological fact 13.5 needed).

Lemma 2.15. *Let* $A \in \mathbb{C}^{n\times n}$ *and* $X \in \mathbb{C}^{n\times k}$ *such that* $X^*X = I_k$. *Then* $\mathcal{W}(X^*AX) \subseteq \mathcal{W}(A)$.

Proof. Suppose that $\mu \in \mathcal{W}(X^*AX)$. Then there exists some $u \in \mathbb{C}^k$ with $\|u\|_2 = 1$ such that

$$u^*X^*AXu = \mu.$$

It follows that $\mu \in \mathcal{W}(A)$, since for $v = Xu$ we have $v^*v = u * X^*Xu = u^*u = 1$ and $v^*Av = \mu$. $\square$

Lemma 2.16. *For any* $A \in \mathbb{C}^{2\times 2}$, $\mathcal{W}(A)$ *is convex.*

Proof. Let $\Re A$ and $\Im A$ be the real and imaginary part of A respectively. Observe that $x^*Ax = x^*\Re Ax + \mathbb{i}x^*\Im Ax$. Moreover,

$$x^*\Re Ax = \operatorname{tr}(x^*\Re Ax) = \operatorname{tr}(xx^*\Re A) = \langle x^*x, \Re A\rangle.$$

For the condition $\|x\|_2 = 1$ we have the following equivalences:

$$\|x\|_2^2 = x^*x = x^*I_nx = \langle xx^*, I_n\rangle.$$

Finally, for all rank-1 PSD Hermitian matrices H there exists a vector x such that $x^*x = H$. Putting all of these assertions together we get the following representation of the numerical range.

$$\mathcal{W}(A) = \{(\langle X, \Re A\rangle + \mathbb{i}\langle X, \Im A\rangle) : X \in \mathrm{PSD}_2(\mathbb{C}), \langle X, I_2\rangle = 1 \text{ and } X \text{ is rank-1}\} \tag{2.5}$$

In particular,

$$\mathcal{W}(A) \subseteq \{(\langle X, \Re A\rangle + \mathbb{i}\langle X, \Im A\rangle) : X \in \mathrm{PSD}_2(\mathbb{C}) \text{ and } \langle X, I_2\rangle = 1\}. \tag{2.6}$$

The set on the right-hand side of Eq. (2.6) is convex since it is a spectahedral shadow. It is obtained by intersecting the PSD cone with a hyperplane and then projecting this plane onto $\mathbb{R}^2$. To finish, we must show that Eq. (2.6) is actually an equality. Suppose that $a + \mathbb{i}b$

belongs to the right-hand side of the previous containment. In particular, there exists an $X \in \mathrm{PSD}_2(\mathbb{C})$ such that

$$\langle X, \Re A\rangle = a \qquad \langle X, \Im A\rangle = b \quad \text{and} \quad \langle X, I_2\rangle = 1. \tag{2.7}$$

If we think about the set of Hermitian matrices that satisfy the previous equations, they form an affine subspace $\mathfrak{U}$ of codimension at most 3 that intersects with $\mathrm{PSD}_2(\mathbb{C})$. Since $\mathrm{PSD}_2(\mathbb{C})$ is a closed pointed cone it cannot contain any line in Herm_n (see Theorem 2.8). Therefore, $\mathfrak{U}$ must intersect the boundary of $\mathrm{PSD}_2(\mathbb{C})$. But given that we are in the 2-by-2 case, this boundary corresponds to the rank-1 Hermitian matrices and the zero matrix by Theorem 2.10. Since $\langle 0_2, I_2\rangle = 0$, the zero matrix is not contained in the line $\mathfrak{U}$ and thus we conclude that a matrix in the intersection must correspond to a rank-1 matrix. $\square$

In the case where $\Re A$, $\Im A$ and I_2 are linearly independent then $\mathfrak{U}$ corresponds to a line. In this case there are at most 2 rank-1 Hermitian matrices that satisfy Eq. (2.7). With this we are ready to prove the main theorem.

Theorem 2.17. *$\mathcal{W}(A)$ is a convex set.*

Proof. Consider two values μ_1 and μ_2 and suppose there exist v_1 and $v_2 \in \mathbb{C}^n$ such that $v_i^* v_i = 1$ and $v_i^* A v_i = \mu_i$. We aim to show that $\mathcal{W}(A)$ contains the convex hull of μ_1 and μ_2. We can assume without loss of generality that v_1 and v_2 are linearly independent. Otherwise, v_1 and v_2 would be proportional and then it would follow that $\mu_1 = \mu_2$ in which case there is nothing to prove. Take X to be an n by 2 matrix such that $X^*X = I_2$ and the span of the columns of X equals the span of v_1 and v_2. This can be achieved, for example, by applying Gram-Schmidt to the matrix $[v_1\, v_2]$.

Then we can consider the matrix X^*AX. Because v_1 and v_2 are in the column span of X, there exists some vector u_1 and u_2 such that $Xu_i = v_i$.

Moreover, $u_i^* u_i = u_i^* X^* X u_i = v_i^* v_i = 1$. It follows that μ_1 and $\mu_2 \in \mathcal{W}(X^*AX)$. By Lemma 2.16 it follows that $\Lambda_1(X^*AX)$ is convex and therefore the line segment between μ_1

and μ_2 is contained in $\mathcal{W}(X^*AX)$. Then by Lemma 2.15 it follows that this line segment is also contained in $\mathcal{W}(A)$ from which the claim follows. $\square$

Corollary 2.18. *For any complex matrix* $A \in \mathbb{C}^{n\times n}$

$$\mathcal{W}(A) = \{(\langle X, \Re A\rangle + \mathbb{i}\langle X, \Im A\rangle) : X \in \mathrm{PSD}_n(\mathbb{C}) \text{ and } \langle X, I_n\rangle = 1\}. \tag{2.8}$$

Proof. We follow the same argument as in [18, Lemma 1]. By the same reasoning as in Lemma 2.16 we can conclude the same equality as in Eq. (2.5). That is,

$$\mathcal{W}(A) = \{(\langle X, \Re A\rangle + \mathbb{i}\langle X, \Im A\rangle) : X \in \mathrm{PSD}_n(\mathbb{C}), \langle X, I_n\rangle = 1 \text{ and } X \text{ is rank-1}\} \tag{2.9}$$

By the characterization of the basis of $\mathrm{PSD}_n(\mathbb{C})$ in Theorem 2.9, the convex hull of the right hand side of Eq. (2.9) is precisely the right hand side of Eq. (2.8). But, since we established that $\mathcal{W}(A)$ is convex, the convex hull of the right-hand side of Eq. (2.9) is equal to itself. $\square$

The following lemma is useful when working with block diagonal matrices.

Lemma 2.19. *Suppose that* $A = A_1 \oplus A_2$ *with* $A_1 \in \mathbb{C}^{k\times k}$ *and* $A_2 \in \mathbb{C}^{(n-k)\times(n-k)}$*. Then* $\mathcal{W}(A) = \mathrm{conv}(\mathcal{W}(A_1) \cup \mathcal{W}(A_2))$.

Proof. ($\supseteq$): Observe that $A_1 = X^*AX$ where $X = \begin{bmatrix} I_k \\ 0 \end{bmatrix} \in \mathbb{C}^{n\times k}$. It is clear that $X^*X = I_k$. Therefore, by Lemma 2.15, $\mathcal{W}(A_1) \subseteq \mathcal{W}(A)$. A similar argument shows that $\mathcal{W}(A_2) \subseteq \mathcal{W}(A)$. From these two containments, we conclude that $\mathcal{W}(A) \supseteq \mathrm{conv}(\mathcal{W}(A_1) \cup \mathcal{W}(A_2))$.

($\subseteq$): If $z_1 = 0$, then $\|z_2\|_2 = \|z\|_2 = 1$ and $z^*Az = z_2^*Az_2 \in \mathcal{W}(A_2)$. Similarly, if $z_2 = 0$ we can show that $z^Az \in \mathcal{W}(A_1)$.

If $z = (z_1, z_2)$ with $z_1 \in \mathbb{C}^k\backslash\{0\}$ and $z_2 \in \mathbb{C}^{n-k}\backslash\{0\}$ then

$$z^*Az = z_1^*A_1z_1 + z_2^*A_2z_2 = \|z_1\|_2^2 \left(\frac{z_1}{\|z_1\|_2}\right)^* A_1 \left(\frac{z_1}{\|z_1\|_2}\right) + \|z_2\|_2^2 \left(\frac{z_2}{\|z_2\|_2}\right)^* A_2 \left(\frac{z_2}{\|z_2\|_2}\right).$$

The condition $\|z\|_2 = 1$ implies that $\|z_1\|_2^2 + \|z_2\|_2^2 = 1$. This shows that any element in $\mathcal{W}(A)$ can be written as a convex combination of elements in $\mathcal{W}(A_1)$ and $\mathcal{W}(A_2)$. $\square$

Chapter 3

COMPUTATION OF HIGHER-RANK NUMERICAL RANGES

The content of this chapter can be found in [28] which was written jointly with Cynthia Vinzant.

3.1 The dual convex cone of the rank-k numerical range

Li and Sze [26] gave a description (1.4) of $\Lambda_k(A)$ as the intersection of half-planes. Taking $\mu = a + \mathbb{i}b$, we rewrite this description as

$$\Lambda_k(A) = \{a + \mathbb{i}b \in \mathbb{C} : a\cos(\theta) + b\sin(\theta) \leq \lambda_k(\theta) \text{ for all } \theta \in [0, 2\pi)\} \tag{3.1}$$

where $\lambda_k(\theta)$ is short-hand notation for $\lambda_k(\cos(\theta)\Re A + \sin(\theta)\Im A)$. Using this description, we use convex duality to study $\Lambda_k(A)$. For convenience, we lift $\Lambda_k(A)$ to a convex cone in $\mathbb{R}^3$. Namely, let

$$\widehat{\Lambda_k(A)} = \{(\lambda, \lambda a, \lambda b) : a + \mathbb{i}b \in \Lambda_k(A), \lambda \in \mathbb{R}_{\geq 0}\}$$

when $\Lambda_k(A)$ is nonempty. When $\Lambda_k(A)$ is empty, we use the convention $\widehat{\Lambda_k(A)} = \{(0, 0, 0)\}$.

To understand the dual cone, consider the closed curve in $\mathbb{R}^3$ parametrized by

$$O_k(A) = \{(\lambda_k(\theta), -\cos\theta, -\sin\theta) : \theta \in [0, 2\pi)\}. \tag{3.2}$$

This curve is contained in the set of points $(t, x, y) \in \mathbb{R}^3$ with $f_A(t, x, y) = 0$ and $x^2 + y^2 = 1$.

Theorem 3.1. *The convex cone $\widehat{\Lambda_k(A)}$ has the inequality description*

$$\widehat{\Lambda_k(A)} = \{(c, a, b) : c \geq 0 \text{ and } \cos(\theta)a + \sin(\theta)b \leq \lambda_k(\theta)c \text{ for all } \theta \in [0, 2\pi)\}$$

and its dual cone is given by

$$(\widehat{\Lambda_k(A)})^* = \overline{\text{conicalHull}(\{(1, 0, 0)\} \cup O_k(A))}.$$

Proof. From Corollary 2.3, it suffices to prove the inequality description for $\widehat{\Lambda_k(A)}$ and the description of $(\widehat{\Lambda_k(A)})^*$ follows. The containment $\subseteq$ follows directly from the description of (1.4) of Li and Sze.

For the reverse containment, suppose that (c, a, b) satisfies $c \geq 0$ and $\cos(\theta)a + \sin(\theta)b \leq \lambda_k(\theta)c$ for all $\theta \in [0, 2\pi)$. If $c > 0$, then $\frac{1}{c}(a, b)$ satisfies all the inequalities that define $\Lambda_k(A)$ and so $(c, a, b) = c(1, a/c, b/c)$ belongs to $\widehat{\Lambda_k(A)}$. If $c = 0$, we claim that $(a, b) = (0, 0)$. Taking $\theta \in \{0, \pi/2, \pi, 3\pi/2\}$, we see that (a, b) must satisfy the inequalities $-a \geq 0$, $-b \geq 0$, $a \geq 0$ and $b \geq 0$, which imply $(a, b) = (0, 0)$. Therefore $(c, a, b) = (0, 0, 0) \in \widehat{\Lambda_k(A)}$. □

Remark 3.2. Since $\Lambda_k(A)$ is compact, $\widehat{\Lambda_k(A)}$ is a closed, pointed cone. That is, it is closed and does not contain any line through the origin.

Corollary 3.3. *A point $a + \mathbb{i}b$ belongs to $\Lambda_k(A)$ if and only if the linear function $t + ax + by$ is nonnegative on $O_k(A)$.*

Proof. This follows directly from the description (1.4) of Li and Sze. □

The following example shows why the inclusion of $(1, 0, 0)$ in Theorem 3.1 is necessary:

Example 3.4. Consider the 3×3 matrix $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$. The Kippenhahn polynomial is $f_A(t, x, y) = t(t^2 - x^2 - y^2)$. One can check that $\lambda_2(\theta) = 0$ for all $\theta \in [0, 2\pi)$ and therefore $\Lambda_2(A) = \{0\}$. The conical hull of $O_2(A)$ is therefore the plane $\{c = 0\}$ and so the dual cone of this conical hull is the whole real line spanned by $(1, 0, 0)$. This strictly contains $\widehat{\Lambda_2(A)}$, which is the ray $\{(c, 0, 0) : c \geq 0\}$. The conical hull of $\{(1, 0, 0)\} \cup O_2(A)$ is the half-space $\{(t, x, y) : t \geq 0\}$, whose dual cone is $\widehat{\Lambda_2(A)}$, as promised by Theorem 3.1.

In Example 3.4, the curve $O_k(A)$ was contained in a plane. As we will see below, this is the only way for $(\widehat{\Lambda_k(A)})^*$ to differ from the closure of the conical hull of $O_k(A)$. To see this, we prove some very simple but useful inequalities.

Lemma 3.5. *For any* $\theta \in [0, 2\pi)$,

$$\lambda_k(\theta) + \lambda_k(\theta + \pi) \geq 0 \text{ for } k \leq (n+1)/2, \text{ and}$$
$$\lambda_k(\theta) + \lambda_k(\theta + \pi) \leq 0 \text{ for } k \geq (n+1)/2.$$

Proof. Note that $k \leq (n+1)/2$ if and only if $k \leq n - k + 1$. By definition, we have that $\lambda_k(\theta) \geq \lambda_{n-k+1}(\theta)$. Also we have $\lambda_k(\theta + \pi) = -\lambda_{n-k+1}(\theta)$. Together this gives

$$\lambda_k(\theta) + \lambda_k(\theta + \pi) = \lambda_k(\theta) - \lambda_{n-k+1}(\theta) \geq 0.$$

The second statement follows analogously by taking the inequality $\lambda_k(\theta) \leq \lambda_{n-k+1}(\theta)$. $\square$

Corollary 3.6. *Suppose that* $O_k(A)$ *is not contained in a plane through the origin in* $\mathbb{R}^3$. *If* $k \leq (n+1)/2$, *then*

$$(\widehat{\Lambda_k(A)})^* = \overline{\text{conicalHull}(O_k(A))}$$

and if $k \geq (n+1)/2$, *then* $(\widehat{\Lambda_k(A)})^* = \mathbb{R}^3$, $\widehat{\Lambda_k(A)} = \{(0,0,0)\}$, *and* $\Lambda_k(A) = \emptyset$.

Proof. We show that when $k \leq (n+1)/2$, $(1,0,0)$ belongs to the closed convex cone $C = \overline{\text{conicalHull}(O_k(A))}$. Suppose, for the sake of contradiction, that it does not. Then there exists a hyperplane separating these sets. That is, there exists $(c, a, b) \in \mathbb{R}^3$ with $ct + ax + by \geq 0$ for all $(t, x, y) \in C$ and $c < 0$.

Since $O_k(A)$ is not contained in the plane $\{(t, x, y) : ct + ax + by = 0\}$, there must be some point of $O_k(A)$ not on this plane. That is, there is some $\theta \in [0, 2\pi)$ for which

$$\lambda_k(\theta)c - \cos(\theta)a - \sin(\theta)b > 0.$$

Since $ct + ax + by$ is nonnegative on $O_k(A)$, we also have that

$$\lambda_k(\theta + \pi)c + \cos(\theta)a + \sin(\theta)b = \lambda_k(\theta + \pi)c - \cos(\theta + \pi)a - \sin(\theta + \pi)b \geq 0.$$

Adding these two inequalities gives that $c(\lambda_k(\theta) + \lambda_k(\theta + \pi)) > 0$. Since $c < 0$, this contradicts the first inequality in Lemma 3.5 when $k \leq (n+1)/2$. Therefore $(1,0,0) \in \overline{\text{conicalHull}(O_k(A))}$. The claim then follows from Theorem 3.1.

For $k \geq (n+1)/2$, we use that $O_k(A) = -O_{n-k+1}(A)$ and $n-k+1 \leq (n+1)/2$. By the arguments above, the point $(-1,0,0)$ belongs to C. The cone $\overline{\text{conicalHull}(\{(1,0,0)\} \cup O_k(A))}$ therefore contains the whole real line $\{(t,0,0) : t \in \mathbb{R}\}$. By Theorem 3.1, this cone coincides with $(\widehat{\Lambda_k(A)})^*$. From this and the parametrization of $O_k(A)$, we see that it also contains the set $\{(0, -\cos(\theta), -\sin(\theta)) : \theta \in [0, 2\pi)\}$. It follows that $(\widehat{\Lambda_k(A)})^*$ is all of $\mathbb{R}^3$. By the biduality theorem, $\widehat{\Lambda_k(A)} = \{(0,0,0)\}$. It follows that $\Lambda_k(A) = \emptyset$. $\square$

Example 3.7. The 4×4 symmetric matrix

$$A = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 2 & 0 & \mathbb{i} & 0 \\ 0 & \mathbb{i} & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{gives} \quad f_A = \det \begin{pmatrix} t & 2x & 0 & 0 \\ 2x & t & y & 0 \\ 0 & y & t & x \\ 0 & 0 & x & t \end{pmatrix} = t^4 - 5t^2x^2 + 4x^4 - t^2y^2.$$

Figure 3.1 shows the curves $O_k(A)$ for $k = 1,2,3,4$ as well as the convex hull of $O_2(A)$. Every point (t,x,y) in the curve $O_1(A)$ satisfies $t \geq 1$. Its convex hull does not contain the origin and its conical hull is a pointed convex cone given by

$$(\widehat{\Lambda_1(A)})^* = \{(t,x,y) \in \mathbb{R}^3 : f_A(t,x,y) \geq 0, 2t^2 - 5x^2 - y^2 \geq 0, t \geq 0\}.$$

On the other hand, the convex hull of $O_2(A)$ has a family of edges whose tangent planes approach $\{c = 0\}$. The plane $\{c = 0\}$ meets this convex body in an edge between the two points $(0, \pm 1, 0)$. The *conical hull* of $O_2(A)$ is therefore $\{(t,x,y) : t > 0\} \cup \{(0,x,0) : x \in \mathbb{R}\}$. The closure is the half-plane $(\widehat{\Lambda_2(A)})^* = \{(t,x,y) : t \geq 0\}$. The dual cone is the ray $\widehat{\Lambda_2(A)} = \{(c,0,0) : c \geq 0\}$, giving $\Lambda_2(A) = \{0\} \subset \mathbb{C}$.

We end with the following corollary, which is useful for studying rank-k numerical ranges when $k \geq (n+1)/2$.

Corollary 3.8. *Let $k \geq (n+1)/2$ and suppose that $\Lambda_k(A) = \{a + \mathbb{i}b\}$. Then $\Lambda_j(A) = \{a + b\mathbb{i}\}$ for all $n-k+1 \leq j \leq k$.*

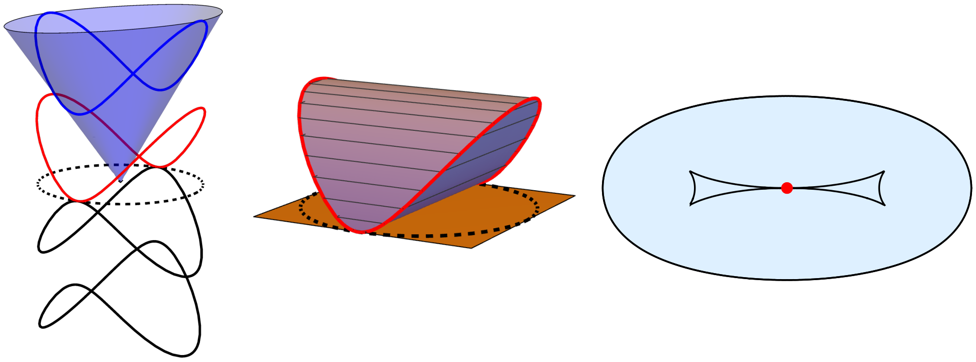

Figure 3.1: The curves $O_k(A)$ from Example 3.7 with the cone $(\widehat{\Lambda_1(A)})^*$, the convex hull of $O_2(A)$, and higher-rank numerical ranges $\Lambda_1(A)$ and $\Lambda_2(A)$.

Proof. By Corollary 3.6, if $\Lambda_k(A)$ is nonempty then $O_k(A)$ has to be contained in a plane. By duality, the functional $t+ax+by$ must be nonnegative in this plane. This is only possible if the plane in question is $t+ax+by=0$. In particular, for any $\theta \in [0,2\pi)$ $\lambda_k(\theta) = a\cos\theta + b\sin\theta$. Now consider,

$$\lambda_{n-k+1}(\theta) = -\lambda_k(\theta+\pi) = -a\cos(\theta+\pi) - b\sin(\theta+\pi) = \lambda_k(\theta).$$

It then follows that $\lambda_{n-k+1}(\theta) = \lambda_{n-k+1}(\theta) = \ldots = \lambda_{k-1}(\theta) = \lambda_k(\theta) = a\cos\theta + b\sin\theta$. Therefore, all $O_j(A)$ are the same for $n-k+1 \leq j \leq k$ and the statement follows.

$\square$

3.2 A membership test for the rank-k numerical range

In this section we describe an algorithm for testing membership of a given point $a+\mathrm{i}b$ in $\Lambda_k(A)$. This relies on the description of the dual cone from Theorem 3.1. Note that a point $(1,a,b)$ belongs to $\widehat{\Lambda_k(A)}$ if and only if the functional $\ell(t,x,y) = t+ax+by$ is nonnegative on $O_k(A)$. This results in many different equivalent conditions for membership in $\Lambda_k(A)$.

Theorem 3.9 (Membership test)**.** *Let $(a,b)\in\mathbb{R}^2$ and $\ell = t+ax+by$. The following are equivalent:*

(a) $a + \mathbb{i}b \in \Lambda_k(A)$*,*

(b) $tI_n + x\Re A + y\Im A$ *has at most* $n-k$ *strictly positive eigenvalues for all* $(t, x, y) \in \mathcal{V}_{\mathbb{R}}(\ell)$*,*

(c) $tI_n + x\Re A + y\Im A$ *has at most* $n-k$ *strictly negative eigenvalues for all* $(t, x, y) \in \mathcal{V}_{\mathbb{R}}(\ell)$*,*

(d) $\lambda_{n-k+1}(x\Re A + y\Im A) \leq ax + by$ *for all* $(x, y) \in \mathbb{R}^2$*,*

(e) $\lambda_k(x\Re A + y\Im A) \geq ax + by$ *for all* $(x, y) \in \mathbb{R}^2$*, and*

(f) $\lambda_k(\Re A + y\Im A) \geq a + by$ *and* $\lambda_{n-k+1}(\Re A + y\Im A) \leq a + by$ *for all* $y \in \mathbb{R}$*.*

Proof. (a)⇔(e) The ⇒ direction follows directly from the description (3.1) of $\Lambda_k(A)$. For the converse, suppose $(x, y) \in \mathbb{R}^2$. We can write $(x, y) = r(\cos\theta, \sin\theta)$ for some $\theta \in [0, 2\pi)$ and $r \in \mathbb{R}_{\geq 0}$. By (3.1), $\lambda_k(\theta) \geq a\cos(\theta) + b\sin(\theta)$. Rescaling both sides by r gives the claim.

(d)⇔(e) For any Hermitian matrix M, $\lambda_k(-M) = -\lambda_{n-k+1}(M)$. Since $\mathbb{R}^2$ is closed under negation and $(-x)\Re A + (-y)\Im A = -(x\Re A + y\Im A)$, the claim follows.

(d,e)⇔(f) Specializing to $x = 1$ gives the forward direction. Conversely, suppose that (f) holds and take $(x, y) \in \mathbb{R}^2$. If $x > 0$, then

$$\lambda_k(x\Re A + y\Im A) = x\lambda_k(\Re A + (y/x)\Im A) \geq x(a + b(y/x)) = ax + by.$$

The inequality follows from (f). Similarly, if $x < 0$, then

$$\begin{aligned} \lambda_k(x\Re A + y\Im A) &= -x\lambda_k(-\Re A - (y/x)\Im A) \\ &= x\lambda_{n-k+1}(\Re A + (y/x)\Im A) \geq x(a + b(y/x)) = ax + by. \end{aligned}$$

Finally, we note that $\lambda_k(x\Re A + y\Im A)$ is a continuous function of (x, y). Since it is $\geq ax + by$ for all (x, y) with $x \neq 0$, it is $\geq ax + by$ for any (x, y). This shows that condition (e) holds, which implies (d) as above.

(c)⇔(e) Condition (c) is equivalent to the condition that $\lambda_k(tI_n + x\Re A + y\Im A) \geq 0$ for all $(t, x, y) \in \mathcal{V}_{\mathbb{R}}(\ell)$. Note that $(t, x, y) \in \mathcal{V}_{\mathbb{R}}(\ell)$ if and only if $t = -ax - by$. We see that

$$\lambda_k((-ax - by)I_n + x\Re A + y\Im A) = -ax - by + \lambda_k(x\Re A + y\Im A).$$

This is ≥ 0 if and only if $\lambda_k(x\Re A + y\Im A) \geq ax + by$.

(b)⇔(c) Since $\mathcal{V}_{\mathbb{R}}(\ell)$ is closed under scaling and $\lambda_k(-M) = -\lambda_{n-k+1}(M)$. □

These equivalences reduce the problem of deciding whether an element $a + \mathbb{i}b$ belongs to $\Lambda_k(A)$ to checking the signs of the eigenvalues in the matrix pencil $-(a+by)I_n + \Re A + y\Im A$. An eigenvalue can only change signs at the roots of the determinant $\det(-(a+by)I_n + \Re A + y\Im A) = f_A(-a-by, 1, y)$. It therefore suffices to check the signature of this matrix pencil at a single point between consecutive roots of this univariate polynomial.

One subtlety is that the restriction of f_A to the line $t + ax + by = 0$ could be identically zero, in which case $\ell = t + ax + by$ is a factor of f_A. To deal with this, we first factor out all powers of ℓ from f_A. This gives the following algorithm:

Algorithm 1 Membership test

Input: $A \in \mathbb{C}^{n\times n}$, $k \in \mathbb{Z}$ with $1 \leq k \leq n$, $a + \mathbb{i}b \in \mathbb{C}$

Output: "True" if $a + \mathbb{i}b \in \Lambda_k(A)$, "False" otherwise.

```
Take $f = \det(tI_n + x\Re A + y\Im A)$ and $\ell = t + ax + by$.
Let $d = \max\{e \in \mathbb{N} : \ell^e \text{ divides } f\}$ and define $\tilde{f} := f/\ell^d$.
Let $h(y) = \tilde{f}(-a-by, 1, y) \in \mathbb{R}[y]$ and compute the distinct real roots $r_1 < \ldots < r_m$ of $h$.
if $h(y)$ has no real roots then
    define $s_0 = 0$
else
    define $s_0 = r_1 - 1$ and $s_m = r_m + 1$.
    for $i = 1, \ldots, m-1$ do
        define $s_i = (r_i + r_{i+1})/2$.
for $i = 0, \ldots, m$ do
    if $\lambda_k(\Re A + s_i\Im A) < a + bs_i$ or $\lambda_{n-k+1}(\Re A + s_i\Im A) > a + bs_i$ then
        return "False" and stop
return "True"
```

Proposition 3.10. *Algorithm 1 correctly determines membership in $\Lambda_k(A)$.*

Proof. If the algorithm returns "False", then, for some i, $\lambda_k(\Re A + s_i \Im A) < a + bs_i$ or $\lambda_{n-k+1}(\Re A + s_i \Im A) > a + bs_i$. By Theorem 3.9, $a + \mathbb{i}b$ does not belong to $\Lambda_k(A)$.

Conversely, if $a+\mathbb{i}b \notin \Lambda_k(A)$, then $\lambda_k(\Re A + s\Im A) < a+bs$ or $\lambda_{n-k+1}(\Re A + s\Im A) > a+bs$ for some $s \in \mathbb{R}$. Since these inequalities hold for an open set of $s \in \mathbb{R}$, we can assume that s is not a root of h. It follows that s belongs to some connected component I of $\mathbb{R}\backslash\{r_1, \ldots, r_m\}$, which has some representative $s_i \in I$. We claim that, for any j, the value of

$$\lambda_j((-a - by)I_n + \Re A + y\Im A) = -a - by + \lambda_j(\Re A + y\Im A)$$

is either identically zero or has constant sign on the open interval I. Note that by definition

$$f(t - a - by, 1, y) = t^d \tilde{f}(t - a - by, 1, y) = \det((t - a - by)I_n + \Re A + y\Im A).$$

For given y, the roots of this polynomial (in t) are the eigenvalues of the matrix $(a + by)I_n - \Re A - y\Im A$. The values of y for which $t = 0$ is a root of multiplicity $> d$ are exactly the roots of $h(y) = \tilde{f}(-a - by, 1, y)$. Therefore for $y \in I$, the roots of $\tilde{f}(-a - by, 1, y)$ are nonzero and have constant sign. It follows that for $y \in I$, $\lambda_j((-a - by)I_n + \Re A + y\Im A)$ is either identically zero or has constant sign. In particular, since I contains s and s_i, $\lambda_k(\Re A + s_i\Im A) < a + bs_i$ or $\lambda_{n-k+1}(\Re A + s_i\Im A) > a + bs_i$ and the algorithm returns "False". $\square$

We finish this section by illustrating the membership test on a small example.

Example 3.11. Consider the 3×3 matrix $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1/2 \\ 0 & 0 & 0 \end{pmatrix}$, $k = 2$, and $a+\mathbb{i}b = 1+0\mathbb{i}$. The Kippenhahn polynomial is $f(t, x, y) = \frac{1}{16}(t+x)(16t^2-x^2-y^2)$. We take $\ell = t+ax+by = t+x$. Note that ℓ is the highest power of ℓ that divides f, so $\tilde{f} = f/\ell = \frac{1}{16}(16t^2 - x^2 - y^2)$. Then

$$h(y) = \tilde{f}(-1, 1, y) = \frac{1}{16}(16(-1)^2 - 1^2 - y^2) = \frac{1}{16}(15 - y^2)$$

which has $m = 2$ real roots, $r_1 = -\sqrt{15}$ and $r_2 = \sqrt{15}$. From this, we define three test points $s_0 = -\sqrt{15} - 1$, $s_1 = 0$, and $s_2 = \sqrt{15} + 1$. For any $y \in \mathbb{R}$, the matrix

$$(-a - by)I_n + \Re A + y\Im A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & \frac{1-iy}{4} \\ 0 & \frac{1+iy}{4} & -1 \end{pmatrix}$$

has eigenvalues $0, -1 \pm \frac{1}{4}\sqrt{1+y^2}$. In particular, for $y = s_1 = 0$, it has two negative eigenvalues. By Theorem 3.9, we conclude that $1 \notin \Lambda_2(A)$. The curve $\mathcal{V}(f_A)$ is shown in Figure 3.2 in the chart $\{t = 1\}$. Connected components of the complement are labelled with the signs of the eigenvalues of the matrix $I_n + x\Re A + y\Im A$.

3.3 *The algebraic boundary of* $\Lambda_k(A)$

In this section we define a polynomial $g_A \in \mathbb{R}[a, b]$ that vanishes on the boundary of $\Lambda_k(A)$ for any k. This polynomial will be well-defined whenever f_A is not a power of a linear form. We define g_A as the minimal polynomial vanishing on the dual curve of $\mathcal{V}(f_A)$ and all lines coming from singularities of f_A. Formally, we compute g_A as follows. Let $f = f_A^{\mathrm{red}}$ denote the square-free part of the factorization of f_A, i.e. the minimal polynomial vanishing on $\mathcal{V}(f_A)$. Consider the ideal $I \subset \mathbb{R}[t, x, y, a, b]$ given by

$$\begin{aligned} I &= \langle f,\ t + ax + by\rangle + \left\langle 2 \times 2 \text{ minors} \begin{pmatrix} 1 & a & b \\ \frac{\partial f}{\partial t} & \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \right\rangle \\ &= \left\langle f,\ t + ax + by, \frac{\partial f}{\partial x} - a\frac{\partial f}{\partial t}, \frac{\partial f}{\partial y} - b\frac{\partial f}{\partial t}, a\frac{\partial f}{\partial y} - b\frac{\partial f}{\partial x} \right\rangle \end{aligned}$$

and the ideal J obtained from I by saturating by the irrelevant ideal $\langle t, x, y\rangle$ and then eliminating the variables (t, x, y):

$$J = \text{radical}\left(\text{eliminate}\left(I : \langle t, x, y\rangle^\infty, \{t, x, y\}\right)\right).$$

See [10] for more details on the operations of taking radicals, saturation, and elimination of variables. We claim that this is a principal ideal whose variety consists of the dual curve of $\mathcal{V}(f_A)$ and lines corresponding to singularities of $\mathcal{V}(f_A)$.

Lemma 3.12. *If f_A is not a power of a linear form, then the ideal J above is principal and*

$$\mathcal{V}(J) = \mathcal{V}(f_A)^* \cup \bigcup_{[p_0:p_1:p_2]\in \mathrm{Sing}(f_A)} \mathcal{V}(p_0 + p_1 a + p_2 b). \tag{3.3}$$

Proof. We first show that the variety of J is as claimed. As in the construction of J, let $f = f_A^{\mathrm{red}}$ denote the squarefree part of the factorization of f_A. By construction, the variety of J (over $\mathbb{C}$) consists of the set of $(a, b) \in \mathbb{C}^2$ for which there exists $p = (p_0, p_1, p_2) \in \mathbb{C}^3 \backslash \{(0, 0, 0)\}$ and $\lambda \in \mathbb{C}$ with $f(p) = 0$, $p_0 + ap_1 + bp_2 = 0$, and $\lambda(1, a, b) = \nabla f(p)$. Under the additional restriction that $\nabla f(p)$ is nonzero, this is exactly the dual curve of $\mathcal{V}(f)$. We see that for any singular point p, any point $(a, b) \in \mathbb{C}^2$ satisfying $p_0 + p_1 a + p_2 b = 0$ belongs to $\mathcal{V}(J)$, since $0 \cdot (1, a, b) = \nabla f(p)$.

To show that $\mathcal{V}(J)$ is principal, it suffices to show that its variety is a union of plane curves. Suppose $f = \ell_1 \cdots \ell_r \cdot f_1 \cdots f_s$ where $\ell_1, \dots, \ell_r$ are linear forms and $f_1, \dots, f_s$ have degree ≥ 2. The dual curve of f is the union of the dual curves of its factors, i.e. $\mathcal{V}(f)^* = (\cup_{i=1}^r \mathcal{V}(\ell_i)^*) \cup \left(\cup_{j=1}^s \mathcal{V}(f_j)^*\right)$. For each i and j, $\mathcal{V}(\ell_i)^*$ is a point and $\mathcal{V}(f_j)^*$ is a curve. It suffices to show that each point $\mathcal{V}(\ell_i)^*$ is contained in $\mathcal{V}(p_0 + p_1 a + p_2 b)$ for some singular point p of $\mathcal{V}(f)$. To do this, consider the factorization $f = \ell_i \cdot h$. By assumption, f_A is not a power of a linear form and so its reduced polynomial f is not linear, implying that $\deg(h) \geq 1$. It follows that the intersection of $\mathcal{V}(\ell_i)$ and $\mathcal{V}(h)$ in $\mathbb{P}^2(\mathbb{C})$ is nonempty. Let p be a point in this intersection. Then p will be a singular point of $\mathcal{V}(f)$. Moreover, since $\ell_i(p) = 0$, the point $\mathcal{V}(\ell_i)^*$ is contained in the line $\mathcal{V}(p_0 + p_1 a + p_2 b)$. Therefore the variety of J is

$$\mathcal{V}(J) = \left(\bigcup_{j=1}^{s} \mathcal{V}(f_j)^*\right) \cup \left(\bigcup_{[p_0:p_1:p_2]\in \mathrm{Sing}(f)} \mathcal{V}(p_0 + p_1 a + p_2 b)\right),$$

which is a union of plane curves. It follows that J is a principal ideal. $\square$

Definition 3.13 (The polynomial g_A)**.** Let $A \in \mathbb{C}^{n\times n}$ and suppose that f_A is not a power of a linear form. Define the polynomial $g_A \in \mathbb{R}[a, b]$ to be the generator of the ideal J above. When $\mathcal{V}(f_A)$ is singular, the variety of g_A includes both the dual curve of $\mathcal{V}(f_A)$ and lines corresponding to its singular points.

Lemma 3.14. *Let* $f = f_A^{\mathrm{red}}$*. If* ∇f *is nonzero at* $p = (\lambda_k(x^*, y^*), -x^*, -y^*)$ *for some* $(x^*, y^*) \in \mathbb{R}^2$*, then the function* $\lambda_k(x, y)$ *is differentiable at* (x^*, y^*) *with*

$$\frac{\partial \lambda_k}{\partial x}(x^*, y^*) = \frac{\frac{\partial f}{\partial x}(p)}{\frac{\partial f}{\partial t}(p)} \quad \textit{and} \quad \frac{\partial \lambda_k}{\partial y}(x^*, y^*) = \frac{\frac{\partial f}{\partial y}(p)}{\frac{\partial f}{\partial t}(p)}.$$

Proof. Suppose that $\nabla f(p)$ is nonzero. By [30, Lemma 2.4], $\frac{\partial f}{\partial t}$ is nonzero at p. Note that for any $(x, y) \in \mathbb{R}^2$, the point $(\lambda_k(x, y), -x, -y)$ belongs to the variety of f. The result then follows from the implicit function theorem. $\square$

Theorem 3.15. *If* $a + \mathbb{i}b$ *belongs to* $\partial\Lambda_k(A)$*, then* $g_A(a, b) = 0$*.*

Proof. Since $a + \mathbb{i}b \in \Lambda_k(A)$, the linear function $\ell = t + ax + by$ is nonnegative on the curve $O_k(A) = \{(\lambda_k(\theta), -\cos(\theta), -\sin(\theta)) : \theta \in [0, 2\pi)\} \subset \mathcal{V}(f_A)$. Moreover, since it belongs to the boundary of $\Lambda_k(A)$, ℓ must be zero at some point $p = (p_0, p_1, p_2) \in O_k(A)$ (otherwise ℓ is positive on $O_k(A)$ and any small enough perturbation of ℓ would remain positive).

If p is a singular point of $\mathcal{V}(f_A)$, then the linear form $p_0 + p_1 a + p_2 b$ is a factor of the polynomial g_A. Since $0 = \ell(p) = p_0 + ap_1 + bp_2$, (a, b) belongs to the variety of g_A.

Now, suppose that $\nabla f_A^{\mathrm{red}}(p)$ is nonzero and consider $F_{(a,b)} : \mathbb{R}^2 \to \mathbb{R}$ given by function

$$F_{(a,b)}(x, y) = \ell(\lambda_k(x, y), -x, -y) = \lambda_k(x, y) - ax - by.$$

Since $\lambda_k(\mu x, \mu y) = \mu\lambda_k(x, y)$ for any $\mu \in \mathbb{R}_{>0}$ and ℓ is nonnegative on $O_k(A)$, $F_{(a,b)}$ is nonnegative on all of $\mathbb{R}^2$. Moreover, it is zero at (p_1, p_2). By Lemma 3.14, $F_{(a,b)}$ is also differentiable at (p_1, p_2) with

$$\frac{\partial F_{(a,b)}}{\partial x}(p_1, p_2) = \frac{\frac{\partial f_A^{\mathrm{red}}}{\partial x}(p)}{\frac{\partial f_A^{\mathrm{red}}}{\partial t}(p)} - a \quad \text{and} \quad \frac{\partial F_{(a,b)}}{\partial y}(p_1, p_2) = \frac{\frac{\partial f_A^{\mathrm{red}}}{\partial y}(p)}{\frac{\partial f_A^{\mathrm{red}}}{\partial t}(p)} - b.$$

Since $F_{(a,b)}$ achieves its minimum value at (p_1, p_2), these derivatives must be zero, from which we conclude that $(1, a, b)$ is a scalar multiple of $\nabla f_A^{\mathrm{red}}(p)$. Since $p \neq 0$ belongs to $\mathcal{V}(f_A)$, it follows by definition that $(1, a, b)$ belongs to $\mathcal{V}(f_A)^*$. $\square$

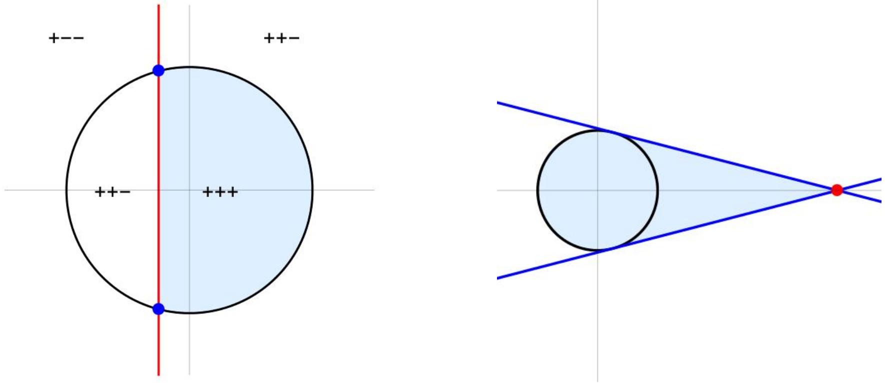

Figure 3.2: The curves $\mathcal{V}(f_A)$ and $\mathcal{V}(g_A)$ from Example 3.11 and Example 3.16.

Example 3.16 (Example 3.11 continued)**.** Consider the 3×3 matrix A from Example 3.11 with $f_A(t,x,y) = \frac{1}{16}(t+x)(16t^2 - x^2 - y^2)$. The dual curve of $\mathcal{V}(f_A)$ is the union of the duals of its two components $\mathcal{V}(f_A)^* = \{[1:1:0]\} \cup \mathcal{V}(1 - 16a^2 - 16b^2)$. The curve $\mathcal{V}(f_A)$ has two singular points, at which the two factors are both zero, namely $[1 : -1 : \pm\sqrt{15}]$. These both contribute a linear factor to g_A, giving

$$g_A(a,b) = (1 - 16a^2 - 16b^2)(1 - a + \sqrt{15}b)(1 - a - \sqrt{15}b).$$

3.4 *Lower dimensional* $\Lambda_k(A)$

In this section, we study the geometric and algebraic conditions under which $\Lambda_k(A)$ is nonempty but has dimension ≤ 1. We do this by considering the convex hull, $\operatorname{conv}(O_k(A))$, of the curve $O_k(A) \subset \mathbb{R}^3$ defined in (3.2). We will see that when $\Lambda_k(A)$ has dimension 0 or 1, the origin $(0,0,0)$ lies on a face F of $\operatorname{conv}(O_k(A))$ with $1 \leq \dim(F) \leq 2$. We then differentiate when F contains *antipodal points* of $O_k(A)$ and when it does not. Here, we say that $(\lambda_k(\theta), -\cos(\theta), -\sin(\theta))$ and $(\lambda_k(\theta+\pi), -\cos(\theta+\pi), -\sin(\theta+\pi))$ are *antipodal points* of $O_k(A)$ when $\lambda_k(\theta) = \lambda_{n-k+1}(\theta)$, in which case,

$$\begin{aligned}(\lambda_k(\theta+\pi), -\cos(\theta+\pi), -\sin(\theta+\pi)) &= (-\lambda_{n-k+1}(\theta), \cos(\theta), \sin(\theta)) \\ &= (-\lambda_k(\theta), \cos(\theta), \sin(\theta)),\end{aligned}$$

giving antipodal points $\pm(\lambda_k(\theta), -\cos(\theta), -\sin(\theta))$. This occurs only when the curve $\mathcal{V}(f_A)$ has a real singularity with a prescribed signature. First we consider the dimension of F.

Proposition 3.17. *Suppose* $O_k(A)$ *is not contained in a plane through the origin and* $k \leq (n+1)/2$. *Then* $\dim(\Lambda_k(A)) \in \{0,1\}$ *if and only if* $(0,0,0)$ *belongs to the relative interior of a face* F *of* $\operatorname{conv}(O_k(A))$ *with* $1 \leq \dim(F) \leq 2$.

- *(i) If* $\dim(F) = 1$, *then it is the convex hull of two antipodal points on* $O_k(A)$.
- *(ii) If* $\dim(F) = 2$ *and* $F \subseteq \{(t,x,y) : t + ax + by = 0\}$, *then* $\Lambda_k(A) = \{a + \mathbb{i}b\}$.

Proof. Recall that the cone over $\Lambda_k(A)$, $\widehat{\Lambda_k(A)} \subseteq \mathbb{R}^3$, has dimension one more than that of $\Lambda_k(A)$. By Corollary 3.6, $\widehat{\Lambda_k(A)}$ is the convex dual cone of the *conical hull* of the curve $O_k(A)$. The cone $\widehat{\Lambda_k(A)}$ is full dimensional if and only if its dual cone, $\overline{\text{conicalHull}(O_k(A))}$ is pointed, which occurs if and only if $(0,0,0)$ does not belong to $\operatorname{conv}(O_k(A))$. We can therefore assume that $(0,0,0)$ belongs to $\operatorname{conv}(O_k(A))$ and let F be the unique nonempty face of $\operatorname{conv}(O_k(A))$ containing $(0,0,0)$ in its relative interior.

Note that $\Lambda_k(A) = \emptyset$ is equivalent to $\widehat{\Lambda_k(A)} = \{(0,0,0)\}$, which happens if and only if its dual cone, $\overline{\text{conicalHull}(O_k(A))}$, is all of $\mathbb{R}^3$. This occurs if and only if $\operatorname{conv}(O_k(A))$ is three-dimensional and $(0,0,0)$ belongs to its interior. That is, if $\dim(F) = 3$.

It follows that $\dim(\Lambda_k(A)) \in \{0,1\}$ if and only if $(0,0,0)$ belongs to the boundary of $\operatorname{conv}(O_k(A))$. The 0-dimensional faces of this convex body are precisely the points on the curve $O_k(A)$. Since $(0,0,0)$ does not lie on this curve, it lies on the boundary of $\operatorname{conv}(O_k(A))$ if and only if it lies in the relative interior of some face F with $1 \leq \dim(F) \leq 2$.

($\dim(F) = 1$) If $\dim(F) = 1$, it is the convex hull of two points $(\lambda_k(\theta), -\cos(\theta), -\sin(\theta))$ and $(\lambda_k(\theta'), -\cos(\theta'), -\sin(\theta'))$. Since $(0,0,0)$ is a convex combination of these points and the last two coordinates lie on the unit circle, these points must be negatives of each other. It follows that $\theta' \equiv \theta + \pi \mod 2\pi$ and $\lambda_k(\theta + \pi) = -\lambda_k(\theta)$.

($\dim(F) = 2$) Suppose $\dim(F) = 2$ and $F \subseteq \{(t,x,y) : t + ax + by = 0\}$. If $a' + \mathbb{i}b' \in \Lambda_k(A)$, then by (3.1), the linear function $t + a'x + b'y$ is nonnegative on $O_k(A)$, in which case it must vanish identically on F. Since F has dimension two, there is a unique linear function vanishing on F, up to scaling, showing that $(a', b') = (a, b)$. $\square$

We first explore the case when the face F has no antipodal points. In this case, $\Lambda_k(A)$ will consist of a single point that corresponds to a tritangent line of the curve $\mathcal{V}(f_A)$ or linear factor of f_A.

Definition 3.18. Let $\ell = t + ax + by$, $f \in \mathbb{C}[t, x, y]$ be homogeneous of degree one and d, respectively. Let f^{red} denote the square free part of f. We say that ℓ is *tritangent* to $\mathcal{V}(f)$ if the restriction of f to the line $\ell = 0$ has ≥ 3 double points. That is, $f^{\text{red}}(-ax - by, x, y) = (\ell_1 \cdot \ell_2 \cdot \ell_3)^2 \cdot h$ for some $\ell_1, \ell_2, \ell_3, h \in \mathbb{C}[x, y]$ with $\deg(\ell_i) = 1$ for $i = 1, 2, 3$.

Note that each square factor ℓ_i^2 gives a point v_i at which the line $\mathcal{V}(\ell)$ intersects $\mathcal{V}(f)$ with multiplicity ≥ 2. That is, either the the curve $\mathcal{V}(f)$ is smooth at v_i and $\mathcal{V}(\ell)$ is tangent to $\mathcal{V}(f)$ at this point or the curve $\mathcal{V}(f)$ has multiplicity ≥ 2 at v_i and v_i is contained in $\mathcal{V}(\ell)$. When the linear factors ℓ_1, ℓ_2, ℓ_3 are distinct, the line $\mathcal{V}(\ell)$ is tangent to $\mathcal{V}(f)$ at three distinct points, justifying the name "tritangent".

Theorem 3.19. *Suppose* $\dim(\Lambda_k(A)) \in \{0, 1\}$ *and that* $\lambda_k(\theta) \neq \lambda_{n-k+1}(\theta)$ *for all* θ. *Then* $\Lambda_k(A) = \{a + \mathbb{i}b\}$ *where* $t + ax + by$ *is either tritangent to* $\mathcal{V}(f_A)$ *or divides* f_A.

When the curve $O_k(A)$ is smooth, this follows from a characterization of faces for convex hulls of smooth algebraic curves in $\mathbb{R}^3$ [31]. In general, this curve may have singularities.

Proof. Note that the assumption $\lambda_k(\theta) \neq \lambda_{n-k+1}(\theta)$ guarantees that $O_k(A)$ is not contained in a plane through the origin. Proposition 3.17 then implies that $(0, 0, 0)$ is contained in the relative interior of a face F of $\operatorname{conv}(O_k(A))$ with $\dim(F) = 2$ and that $\Lambda_k(A) = \{a + \mathbb{i}b\}$, where $t + ax + by = 0$ defines the unique plane through the origin containing F.

Let $\ell = t + ax + by$ and suppose that ℓ does not divide f_A. It follows that there are only finitely many points of the curve $O_k(A)$ with $\ell = 0$. In particular, the vertices of the face F belong to this finite intersection. Since $\dim(F) = 2$, it has three vertices v_1, v_2, v_3 (possibly among others), which necessarily belong to $O_k(A)$. As the curve $O_k(A)$ has no antipodal points, the images, $[v_1], [v_2], [v_3]$, of these points in $\mathbb{P}^2(\mathbb{R})$ are all distinct.

We claim that the restriction of f_A^{red} to the line $\ell = 0$ has roots of multiplicity ≥ 2 at each of these points. Suppose to the contrary that $[v_i]$ is a simple root of the restriction of f_A^{red} to $\ell = 0$. It follows that $[v_i]$ is not a singular point of $\mathcal{V}(f_A^{\text{red}})$ and that $\mathcal{V}(f_A^{\text{red}})$ and $\mathcal{V}(\ell)$ meet transversely at $[v_i]$. From this, we see that v_i is a smooth point of the curve $O_k(A)$ and that $O_k(A)$ and the plane $\mathcal{V}(\ell)$ meet transversely at v_i. This implies that $\ell(\lambda_k(\theta), -\cos(\theta), -\sin(\theta))$ changes sign at v_i, contradicting the fact that ℓ is nonnegative on $O_k(A)$. Therefore, the restriction of f_A^{red} to the line $\ell = 0$ has ≥ 3 roots of multiplicity ≥ 2 and ℓ is tritangent to $\mathcal{V}(f_A)$. □

Up to rescaling by arbitrary real scalars, antipodal points $\pm(\lambda_k(\theta), -\cos(\theta), -\sin(\theta))$ are arbitrary points $\pm(-\lambda_k(p_1, p_2), p_1, p_2)$ in the variety of f_A with $\lambda_k(p_1, p_2) = \lambda_{n-k+1}(p_1, p_2)$.

Lemma 3.20. *Suppose that $O_k(A)$ contains a pair of antipodal points. That is, suppose $\lambda_k(p_1, p_2) = \lambda_{n-k+1}(p_1, p_2)$ for some $(p_1, p_2) \in \mathbb{R}^2 \backslash \{(0,0)\}$. Then*

(i) $\Lambda_k(A) \subseteq \{a + \mathbb{i}b : \lambda_k(p_1, p_2) = ap_1 + bp_2\}$, and

(ii) if $k < (n+1)/2$, then $p = [-\lambda_k(p_1, p_2) : p_1 : p_2]$ satisfies $f_A(p) = 0$ and $\nabla f_A(p) = 0$.

Proof. (i) After rescaling, we can take $p = (\lambda_k(\theta), -\cos(\theta), -\sin(\theta))$. Suppose that $a + \mathbb{i}b \in \Lambda_k(A)$. Then the function $\ell(t, x, y) = t + ax + by$ is nonnegative on $O_k(A)$. In particular, $\ell(p) \geq 0$ and $\ell(-p) = -\ell(p) \geq 0$. It follows that

$$0 = \ell(p) = \lambda_k(\theta) - a\cos(\theta) - b\sin(\theta).$$

(ii) When $k < (n+1)/2$, $k < n - k + 1$. By definition for any $k \leq j \leq n - k + 1$, $\lambda_{n-k+1}(\theta) \leq \lambda_j(\theta) \leq \lambda_k(\theta)$. Since $\lambda_{n-k+1}(\theta) = \lambda_k(\theta)$ we see that these are equalities. It follows that the matrix $I_n\lambda_k(\theta) - \cos(\theta)\Re A - \sin(\theta)\Im A$ has corank $\geq n - 2k + 2 \geq 2$. By Lemma 2.14, f_A has multiplicity ≥ 2 at $[\lambda_k(\theta) : -\cos(\theta) : -\sin(\theta)]$. □

Corollary 3.21. *Let* $S = \{[p_0 : p_1 : p_2] \in \mathcal{V}(f_A) : -p_0 = \lambda_k(p_1, p_2) = \lambda_{n-k+1}(p_1, p_2)\}$ *and consider the projective linear space* $V \subset \mathbb{P}^2(\mathbb{R})$ *spanned by* S*. Then*

$$\dim(\Lambda_k(A)) \leq 1 - \dim(V).$$

In particular, if $\dim(V) = 2$ *then* $\Lambda_k(A)$ *is empty and if* $\dim(V) = 1$ *then* $\Lambda_k(A)$ *is either empty or the single point* $\{a + \mathbb{i}b : [1 : a : b] \in V^{\perp}\}$*.*

Note that $[1 : a : b] \in V^{\perp}$ if and only if $t + ax + by = 0$ for all $[t : x : y] \in V$.

Remark 3.22. The condition $\lambda_k(p_1, p_2) = \lambda_{n-k+1}(p_1, p_2)$ is well-defined on points in $\mathbb{P}^2(\mathbb{R})$. It is clearly invariant under nonnegative scaling. Moreover, by definition $\lambda_k(-p_1, -p_2) = -\lambda_{n-k+1}(p_1, p_2)$. Similarly, $\lambda_{n-k+1}(-p_1, -p_2) = -\lambda_k(p_1, p_2)$.

When $\Lambda_k(A)$ is one-dimensional, we see from Proposition 3.17 and Lemma 3.20 that the curve $\mathcal{V}(f_A)$ has a point p of multiplicity ≥ 2. In this case, we would like to determine the end points of the line segment $\Lambda_k(A)$. These will correspond to lines passing through p with additional tangency to $\mathcal{V}(f_A)$.

Definition 3.23. Let $p \in \mathcal{V}(f_A)$. Let f_A^{red} denote the square-free part of the factorization of f_A into irreducible polynomials. Then p is a point of $\mathcal{V}(f_A^{\text{red}})$ of some multiplicity $d \geq 1$. Recall from Lemma 2.14 that for any $q \in \mathbb{C}^3$ the restriction $f_A^{\text{red}}(rq + sp)$ has a factor of r^d. We say that the line $\{[rq + sp] : [r : s] \in \mathbb{P}^1(\mathbb{C})\}$ is *p-tangent* to $\mathcal{V}(f_A)$ if one of the following holds:

(I) $f_A^{\text{red}}(rq + sp) = r^d \cdot \ell^2 \cdot h$ for some $\ell \in \mathbb{C}[r, s]_1$ and $h \in \mathbb{C}[r, s]$, or

(II) $f_A^{\text{red}}(rq + sp) = r^{d+1} \cdot h$ for some $h \in \mathbb{C}[r, s]$.

Observe that part (I) in this definition allows for $\ell = r$, which satisfies (II). Similarly, both conditions are satisfied when the restriction $f_A^{\text{red}}(rq + sp)$ is identically zero, in which case the linear form defining the line is a factor of f_A^{red} and hence f_A.

Theorem 3.24. *If* $k < (n+1)/2$ *and* $p = [p_0 : p_1 : p_2] \in \mathcal{V}(f_A)$ *with* $-p_0 = \lambda_k(p_1, p_2) = \lambda_{n-k+1}(p_1, p_2)$ *then* $\Lambda_k(A)$ *is the convex hull of its elements* $a + \mathbb{i}b$ *for which* $t + ax + by$ *is* p*-tangent. That is,*

$$\Lambda_k(A) = \operatorname{conv}(\{a + \mathbb{i}b \in \Lambda_k(A) : t + ax + by = 0 \textit{ is } p\textit{-tangent to } \mathcal{V}(f_A)\}).$$

We will often perform a change of coordinates to take p to the point $[0 : 0 : 1]$. The following will be a useful way of discovering p-tangent lines.

Lemma 3.25. *Suppose that* $[0 : 0 : 1]$ *is a point of* $\mathcal{V}(f_A)$ *of multiplicity* $d \geq 1$*. Then* $t + cx$ *is* $[0 : 0 : 1]$*-tangent to* $\mathcal{V}(f_A)$ *of type (II) if and only if* $t + cx$ *divides the lowest degree part of* $f_A^{\text{red}}(t, x, 1)$*. Moreover, this condition is satisfied when* $\lim_{x \to 0^+} \lambda_k(x, 1)/x = c$ *or* $\lim_{x \to 0^-} \lambda_k(x, 1)/x = c$ *for some* k*.*

Proof. Recall the notation $\lambda_k(x, 1) := \lambda_k(x\Re A + \Im A)$. For ease of notation let f denote f_A^{red}. By assumption $[0 : 0 : 1]$ is a point of $\mathcal{V}(f)$ of some multiplicity $d \geq 1$. Consider the expansion $f(t, x, y) = y^{m-d} h_d + y^{m-d-1} h_{d+1} + \ldots + h_m$, where $h_j \in \mathbb{R}[t, x]$ is homogeneous of degree j and $m = \deg(f)$.

To see the first equivalence, note that the line $t + cx = 0$ is spanned by $q = (-c, 1, 0)$ and $p = (0, 0, 1)$. We expand $f(rq + sp) = \sum_{j=d}^{m} s^{m-j} r^j h_j(-c, 1)$ to see that $h_d(-c, 1) = 0$ if and only if r^{d+1} divides $f(rq + sp)$. Since $h_d(t, x)$ is homogeneous and bivariate, $t + cx$ divides $h_d(t, x)$ if and only if $h_d(-c, 1) = 0$.

For the further claim, we find the limit

$$\lim_{x \to 0^\pm} \frac{f(-\lambda_k(x, 1), x, 1)}{x^d} = 0$$

by noting that its numerator is identically zero. We can then rewrite this limit as

$$\begin{aligned} 0 &= \lim_{x \to 0^\pm} \frac{h_d(-\lambda_k(x, 1), x)}{x^d} + \frac{h_{d+1}(-\lambda_k(x, 1), x)}{x^d} + \cdots + \frac{h_m(-\lambda_k(x, 1), x)}{x^d} \\ &= \lim_{x \to 0^\pm} h_d\left(-\frac{\lambda_k(x, 1)}{x}, 1\right) + x h_{d+1}\left(-\frac{\lambda_k(x, 1)}{x}, 1\right) + \cdots + x^{m-d} h_m\left(-\frac{\lambda_k(x, 1)}{x}, 1\right) \\ &= h_d(-c, 1). \end{aligned}$$

In the last limit, note that the $\lim_{x\to 0^\pm} h_j(-\frac{\lambda_k(x,1)}{x},1) = h_j(-c,1)$ for all j. This shows that $t + cx$ must be a linear factor of $h_d(t,x)$. □

Lemma 3.26. *Let* $\varphi : \mathbb{R} \to \mathbb{R}$ *be a continuous semialgebraic function. If* $\inf\{\varphi(y) : y \in \mathbb{R}\}$ *equals zero, then either* $\varphi(y^*) = 0$ *for some* $y^* \in \mathbb{R}$*,* $\lim_{y\to\infty} \varphi(y) = 0$*, or* $\lim_{y\to-\infty} \varphi(y) = 0$*.*

Proof. By definition, there is some polynomial $F(y,z) \in \mathbb{R}[y,z]$ for which $F(y,\varphi(y)) = 0$ for all $y \in \mathbb{R}$. Then φ is differentiable at all but finitely many points and has finitely many critical points, at which φ' is zero or undefined. See, for example, [3, Thm 5.56]. Let $[a,b]$ be a closed interval containing those critical points. The minimum of $\varphi(y)$ over $y \in [a,b]$ is attained. If this minimum is zero, then there is a point $y^* \in [a,b]$ with $\varphi(y^*) = 0$. If not, then the infimum of φ over either $(-\infty,a)$ or (b,∞) must be zero. On $(-\infty,a)$ and (b,∞) the function φ is monotone. It follows that either $\lim_{y\to\infty}\varphi(y)$ or $\lim_{y\to-\infty}\varphi(y)$ equals zero. □

Proof of Theorem 3.24. After a change of coordinates, we may assume that $p = [0:0:1]$. From Lemma 3.20, it follows that $\Lambda_k(A) \subset \mathbb{R}$. If $\Lambda_k(A)$ is empty then there is nothing to prove. Therefore, we assume that $\Lambda_k(A)$ is non-empty. By applying another affine linear transformation we can assume that $\Lambda_k(A) = [-1,0]$ or $\Lambda_k(A) = \{0\}$. In either case $0 \in \Lambda_k(A)$ and $\varepsilon \notin \Lambda_k(A)$ for all $\varepsilon \in \mathbb{R}_{>0}$. It suffices to show that $t = 0$ is a p-tangent line of $\mathcal{V}(f_A)$. By definition, if t is a factor of f_A, then, since it passes through p, it is p-tangent. Therefore we can assume that t is not a factor of f_A.

Recall the notation $\lambda_j(x,y) = \lambda_j(x\Re(A)+y\Im(A))$. Since $0 \in \Lambda_k(A)$, by Theorem 3.9(a)(f), we have that for all $y \in \mathbb{R}$,

$$\lambda_{n-k+1}(1,y) \leq 0 \leq \lambda_k(1,y). \tag{3.4}$$

Similarly, since $\varepsilon \notin \Lambda_k(A)$, there exists $y_\varepsilon \in \mathbb{R}$ for which $\lambda_k(1,y_\varepsilon) < \varepsilon$ or $\lambda_{n-k+1}(1,y_\varepsilon) > \varepsilon$. Since $\lambda_{n-k+1}(1,y) \leq 0$ for all $y \in \mathbb{R}$, it must be that $\lambda_k(1,y_\varepsilon) < \varepsilon$. It follows that

$$\inf_{y\in\mathbb{R}} \lambda_k(1,y) = 0. \tag{3.5}$$

Because the function $y \mapsto \lambda_k(1,y)$ is continuous and semialgebraic, it follows from

Lemma 3.26 that either $\lambda_k(1, y^*) = 0$ for some $y^* \in \mathbb{R}$, $\lim_{y\to\infty} \lambda_k(1, y) = 0$, or $\lim_{y\to-\infty} \lambda_k(1, y) = 0$.

For $x > 0$, $\lambda_k(1, 1/x) = \lambda_k(x, 1)/x$. Then $\lim_{y\to\infty} \lambda_k(1, y) = \lim_{x\to 0^+} \lambda_k(x, 1)/x$. Similarly, for $x < 0$, $\lambda_k(1, 1/x) = \lambda_{n-k+1}(x, 1)/x$. So $\lim_{y\to-\infty} \lambda_k(1, y) = \lim_{x\to 0^-} \lambda_{n-k+1}(x, 1)/x$. If either of these limits equals zero, Lemma 3.25 gives that $t = 0$ is a p-tangent line of $\mathcal{V}(f_A)$.

Finally, suppose $\lambda_k(1, y^*) = 0$ for some $y^* \in \mathbb{R}$. If the point $(0, 1, y^*)$ is a singularity of $\mathcal{V}(f_A^{\text{red}})$, then the restriction $f_A^{\text{red}}(0, 1, y)$ has a zero of multiplicity ≥ 2 at y^*, by Lemma 2.14.

In the case that this point is smooth, we can consider the function $\lambda_k(x, y)$. This function is locally minimized at $(1, y^*)$. Therefore, it must be the case that its gradient is zero. Hence $\nabla f^{\text{red}}(0, 1, y^*)$ is proportional to $(1, 0, 0)$. See Lemma 3.14. From here we can conclude that the line $t = 0$ is tangent to $\mathcal{V}(f_A^{\text{red}})$ at $(0, 1, y^*)$. In particular, the restriction $f_A^{\text{red}}(0, 1, y)$ is either identically zero or has a root of multiplicity ≥ 2. In either case, $t = 0$ is a p-tangent of $\mathcal{V}(f_A)$. $\square$

Example 3.27. Consider the Kippenhahn polynomial

$$f_A = t^4 - \frac{13}{36}t^2x^2 + \frac{1}{36}x^4 - \frac{1}{4}t^2y^2 - \frac{2}{25}txy^2 + \frac{1}{100}x^2y^2.$$

The curve $\mathcal{V}(f_A)$ is singular at the point $p = [0 : 0 : 1]$. The lowest degree part of $f_A(t, x, 1)$ is $h_2(t, x) = \frac{1}{4}t^2 - \frac{2}{25}tx + \frac{1}{100}x^2 = (t + \frac{1}{25}(4 + \sqrt{41})x)(t + \frac{1}{25}(4 - \sqrt{41})x)$. Both factors give p-tangent lines of f_A. One can check that $a = \frac{1}{25}(4 - \sqrt{41})$ belongs to $\Lambda_2(A)$ and will form one of its end points. The line $t + \frac{1}{3}x = 0$ passes through $[0 : 0 : 1]$ and is additionally tangent to $\mathcal{V}(f_A)$ at the point $[1 : -3 : 0]$, showing that it is also p-tangent to $\mathcal{V}(f_A)$. In this case, $\Lambda_2(A)$ is the line segment $[\frac{1}{25}(4 - \sqrt{41}), 1/3]$.

The curve $O_2(A)$ along with a portion of the boundary of its convex hull is shown in Figure 3.3. The plane $t + \frac{1}{3}x = 0$ defines a two-dimensional face of this convex hull, whereas $t + \frac{1}{25}(4 - \sqrt{41})x = 0$ is a limit of supporting hyperplanes of edges.

Lastly, we consider the case when $k \geq (n + 1)/2$. It is known that if $\Lambda_k(A) = \{a + \mathbb{i}b\}$ is nonempty then $a + \mathbb{i}b$ must be an eigenvalue of A with geometric multiplicity $\geq 2k - n$ [8, Proposition 2.2]. Here we give an algebraic analogue of this result.

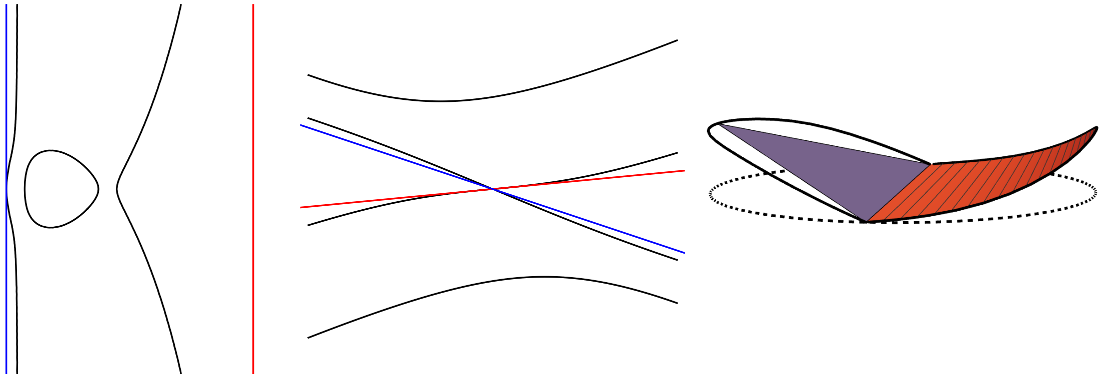

Figure 3.3: The curve $\mathcal{V}(f_A)$ and two $[0:0:1]$-tangent lines from Example 3.27 in the $\{t=1\}$ and $\{y=1\}$ affine charts and curve $O_2(A)$.

Lemma 3.28. *If* $k \geq (n+1)/2$ *and* $a+\mathbb{i}b \in \Lambda_k(A)$ *then* $(t+ax+by)^{2k-n}$ *divides the Kippenhahn polynomial* $\det(tI_n + x\Re A + y\Im A)$.

Proof. Let $\ell = t+ax+by$. By the Membership test Theorem 3.9, for all real points in the line $\mathcal{V}_{\mathbb{R}}(\ell)$, the matrix $M = tI_n + x\Re A + y\Im A$ has at most $n-k$ strictly positive eigenvalues and at most $n-k$ strictly negative eigenvalues. It follows that the rank of M at this point is at most $2(n-k)$, meaning that its $d \times d$ minors are all zero for $d = 2(n-k)+1$. Since this holds for every point $(t,x,y) \in \mathcal{V}_{\mathbb{R}}(\ell)$ and ℓ is irreducible, ℓ divides the $d \times d$ minors of M. By Lemma 3.29 below, it follows that $\ell^{n-d+1} = \ell^{2k-n}$ divides $f_A = \det(M)$. □

The converse of this lemma is not true even for the case $n=3$ and $k=2$. For example, one could take A to be a 3×3 diagonal matrix whose diagonal entries are not co-linear in $\mathbb{C}$. One can check that in this case f_A has three linear factors but that $\Lambda_2(A)$ is empty.

Lemma 3.29. *Let* M *be an* $n \times n$ *matrix with entries in a unique factorization domain* R. *Let* $\ell \in R$ *be irreducible and suppose that* ℓ *divides all* $d \times d$ *minors of* M, *where* $d \leq n$. *Then* ℓ^{n-d+1} *divides* $\det(M)$.

Proof. We induct on $n-d$ and follow the proof of [30, Lemma 4.7]. For $n-d=0$, this holds by assumption. Suppose that $n-d \geq 1$ and that ℓ divides all $d \times d$ minors of M. This also holds for any of the $d \times d$ minors of any $(n-1) \times (n-1)$ submatrix of M, so, by induction ℓ^{n-d} divides the $(n-1) \times (n-1)$ minors of M. We use the identity $M^{\text{adj}}M = \det(M)I_n$, where M^{adj} denotes the adjugate matrix of M, whose entries are signed $(n-1) \times (n-1)$ minors of M. Taking the determinant of boths sides, we find that $\det(M^{\text{adj}}) = \det(M)^{n-1}$.

Suppose that $\det(M) = \ell^m h$ where ℓ does not divide h. Then $\det(M^{\text{adj}}) = \ell^{m(n-1)}h^{n-1}$. By assumption, ℓ^{n-d} divides all entries of M^{adj} and so $\ell^{n(n-d)}$ divides its determinant. Since ℓ is irreducible, it follows that $\ell^{n(n-d)}$ must divide $\ell^{m(n-1)}$, giving $n(n-d) \leq m(n-1)$. If $m \leq n-d$, then $m(n-1) \leq (n-d)(n-1) < (n-d)n$. Since this contradicts the inequality above, we conclude that $n-d+1 \leq m$, as desired. □

3.5 Algorithms

In this section we discuss an algorithm to compute the rank-k numerical range of an $n \times n$ matrix. This builds off Algorithm 1, which tests membership of a given point in $\Lambda_k(A)$, as well as the characterizations of the boundary of $\Lambda_k(A)$ in Section 3.3 and of algebraic conditions satisfied when $\Lambda_k(A)$ is lower dimensional in Section 3.4. To build up to the full algorithm, we first establish some subroutines to compute p-tangents and tritangent lines of $\mathcal{V}(f_A)$, as defined in Section 3.4.

Many of these computations involve the discriminant. Formally, given a polynomial $p(t) = \sum_{k=0}^{d} a_k t^k$, the discriminant is a polynomial $\text{Disc}_t(p)$ in $\mathbb{Q}[a_0, \ldots, a_d]$. It is the unique minimal polynomial (up to scaling) satisfying $\text{Disc}_t(p) = 0$ whenever $p(t)$ that has a double root. For a multivariate polynomial $f \in \mathbb{C}[x_1, \ldots, x_n]$, we use $\text{Disc}_{x_i}(f)$ to denote the discriminant of f interpreted as a univariate polynomial in x_i. This is a polynomial in $\mathbb{C}[x_j : j \neq i]$. For a detailed discussion on discriminants, see [3, Section 4.1].

Lemma 3.30. *The output of Algorithm 2 is the set of* $(a, b) \in \mathbb{R}^2$ *for which* $t + ax + by$ *is* p*-tangent to* $\mathcal{V}(f_A)$.

Proof. Note that $t+ax+by$ is a p-tangent line of f_A if and only if $t+cx$ is a $[0:0:1]$-tangent line of $f_{L \cdot A}$, where $(c, 0) = L(a, b)$. Therefore, it suffices to show that the computed set $\mathcal{S}$ is

Algorithm 2 Singularity Tangents

```
Input: $A \in \mathbb{C}^{n\times n}$ and $p \in \mathcal{V}(f_A)$
Output: $\mathcal{T} = \{(a,b) \in \mathbb{R}^2 : t + ax + by$ is $p$-tangent to $\mathcal{V}(f_A)\}$.
  Take an invertible $L : \mathbb{R}^2 \to \mathbb{R}^2$ as in (2.4) with $[u_{02} : u_{12} : u_{22}] = p$.
  Let $f = f_{L\cdot A}$ and let $f^{\mathrm{red}}$ be the square-free part of $f$.
  Let $h_d(t,x)$ to be the smallest degree part of $f^{\mathrm{red}}(t,x,1)$.
  Take $\mathcal{S} = \{c \in \mathbb{R} : h_d(-c,1) = 0\} \cup \{c \in \mathbb{R} : \mathrm{Disc}_y(f^{\mathrm{red}})(-c,1) = 0\}$
  return $\mathcal{T} = \{L^{-1}(c,0) : c \in \mathcal{S}\}$.
```

the set of $c \in \mathbb{R}$ for which $t + cx$ is $[0:0:1]$-tangent to $\mathcal{V}(f_{L\cdot A})$. By Lemma 3.25, $t + cx$ is a $[0:0:1]$-tangent line of the form (II) if and only if $h_d(-c,1) = 0$.

We claim that $t + cx$ is a $[0:0:1]$-tangent line of the form (I) with $\ell \neq r$ if and only if $\mathrm{Disc}_y(f^{\mathrm{red}})(-c,1) = 0$. As above, the line $t + cx = 0$ is parametrized by the points $(-c,1,0)$ and $(0,0,1)$. In particular, $f^{\mathrm{red}}(r(-c,1,0) + s(0,0,1)) = f^{\mathrm{red}}(-cr,r,s)$ has a square factor $\ell^2 \neq r^2$ in $\mathbb{C}[r,s]$ if and only if its restriction to $r = 1$ has a repeated root. This happens if and only if its discriminant, $\mathrm{Disc}_y(f^{\mathrm{red}})(-c,1)$, is zero. □

As discussed in Corollary 3.21, the presence of antipodal points gives a restriction on higher-rank numerical ranges and guaranties that its dimension is at most 1. Given our definition of antipodal points it is possible to have infinitely many of them, in particular, when the Kippenhahn curve has irreducible components of multiplicity ≥ 2. However, we show that it is only necessary to consider antipodal points that are also singularities of $\mathcal{V}(f_A)$.

Lemma 3.31. *Suppose $\lambda_k(p_1,p_2) = \lambda_{n-k+1}(p_1,p_2)$ for some $(p_1,p_2) \in \mathbb{R}^2\backslash\{(0,0)\}$ and $k < (n+1)/2$. Then one of the following holds:*

(i) $[-\lambda_k(p_1,p_2) : p_1 : p_2]$ is a singularity of $\mathcal{V}(f_A^{\mathrm{red}})$,

(ii) $\Lambda_k(A) = \{a + \mathbb{i}b\}$ for some $a,b \in \mathbb{R}$ where $(t + ax + by)^2$ divides f_A, or

(iii) $\Lambda_k(A)$ *is empty.*

Proof. Let $\ell = t+ax+by$. Suppose that (i) and (iii) are false. That is, there exists some point $a + \mathbb{i}b \in \Lambda_k(A)$ and also $[-\lambda_k(p_1, p_2) : p_1 : p_2]$ is a smooth point of $\mathcal{V}(f_A^{\text{red}})$. In particular, $\frac{\partial f_A^{\text{red}}}{\partial t} \neq 0$ by [30, Lemma 2.4]. By the implicit function theorem there exists a neighborhood $(p_1, p_2) \in U \subseteq \mathbb{R}^2$ and a unique continuously differentiable function $\varphi(x, y)$ such that

$$f_A^{\text{red}}(-\varphi(x, y), x, y) = 0 \text{ for all } (x, y) \in U.$$

Since $\lambda_k, \ldots, \lambda_{n-k}$ are locally continuously differentiable functions that satisfy the previous statement they must all agree on U. Thus, for every (q_1, q_2) in U, the point $[-\lambda_k(q_1, q_2) : q_1 : q_2]$ is a smooth point of $\mathcal{V}(f_A^{\text{red}})$ and $\lambda_k(q_1, q_2) = \lambda_{n-k+1}(q_1, q_2)$. Following the arguments of Lemma 3.20 we can conclude that $-\lambda_k(q_1, q_2) + aq_1 + bq_2 = 0$. Since there are infinitely many points (t, x, y) in $\mathcal{V}(\ell)$ at which the matrix $tI_n + x\Re A + y\Im A$ has corank ≥ 2, it follows that ℓ must divide all of the $(n-1) \times (n-1)$ minors by Bezout's theorem. Then, by Lemma 3.29, $\ell^{n-(n+1)+1} = \ell^2$ divides f_A. $\square$

As shown in Lemma 3.28, when $k \geq (n+1)/2$, we only need to check points corresponding to linear factors with a multiplicity of $2k - n$. For the sake of completeness we present a simple algorithm that computes $\Lambda_k(A)$ under this assumption in Algorithm 3.

Algorithm 3 Computing the rank-k numerical range of a matrix for $(n+1)/2 \leq k \leq n$

Input: $A \in \mathbb{C}^{n \times n}$, $k \in \mathbb{N}$ with $(n+1)/2 \leq k \leq n$

Output: $d \in \{-1, 0\}$: dimension of $\Lambda_k(A)$

If $d = -1$, return the empty set.

If $d = 0$, returns $(a, b) \in \mathbb{R}^2$ such that $\Lambda_k(A) = \{a + bi\}$.

```
  $f_A = \det(tI_n + x\Re A + y\Im A)$.
  Compute all distinct linear factors of $f_A$ (or all factors appearing with power $\geq 2k - n$)
  and perform membership test on each.
```

After this, we present the main algorithm for computing rank-k numerical ranges, Algorithm 4. Roughly speaking, this algorithm splits the problem in different cases given by the presence of antipodal points and the possible dimension of $\Lambda_k(A)$. Then, for each of these cases the problem is reduced to applying the membership test to a finite list of points to determine which of those points belong to $\Lambda_k(A)$.

Algorithm 4 Computing the rank-k numerical range of a matrix for $k < (n+1)/2$

Input: $A \in \mathbb{C}^{n\times n}$, $k \in \mathbb{Z}$ with $1 \leq k < (n+1)/2$

Output: $\dim \in \{-1,0,1,2\}$: dimension of $\Lambda_k(A)$

If $\dim = 0$, returns $a + \mathbb{i}b \in \mathbb{C}$ such that $\Lambda_k(A) = \{a + \mathbb{i}b\}$.

If $\dim = 1$, returns $a + \mathbb{i}b, c + \mathbb{i}d \in \mathbb{C}$ such that $\Lambda_k(A) = \operatorname{conv}\{a + \mathbb{i}b, c + \mathbb{i}d\}$.

If $\dim = 2$, returns a polynomial g_A vanishing on the boundary of $\Lambda_k(A)$ and a representative for each connected component of $\mathbb{C}\backslash\{a + \mathbb{i}b : g_A(a,b) = 0\}$ that is contained in $\Lambda_k(A)$.

```
Take $f_A = \det(tI_n + x\Re A + y\Im A)$ and let $f_A^{\mathrm{red}}$ be the square-free part of $f_A$
Compute $S = \{[p_0 : p_1 : p_2] \in \mathrm{Sing}_{\mathbb{R}}(f_A^{\mathrm{red}}) : -p_0 = \lambda_k(p_1,p_2) = \lambda_{n-k+1}(p_1,p_2)\}$
Compute $V = \operatorname{span} S \subseteq \mathbb{P}^2$ and $\dim V \in \{-1,0,1,2\}$
if $\dim V = 2$ then return $\dim = -1$
if $\dim V = 1$ then compute $\{[1 : a : b]\} = V^{\perp}$
    Use Algorithm 1 to test $a + \mathbb{i}b \in \Lambda_k(A)$.
    if $a + \mathbb{i}b \in \Lambda_k(A)$ then return $\dim = 0$, $\{a + \mathbb{i}b\}$
    else return $\dim = -1$
if $\dim V = 0$ then $S = \{p\}$
    Use Algorithm 2 to compute $\mathcal{T} = \{a + \mathbb{i}b : t + ax + by$ is $p$-tangent to $\mathcal{V}(f_A)\}$
    Use Algorithm 1 to compute $\mathcal{T} \cap \Lambda_k(A)$.
    if $|\mathcal{T} \cap \Lambda_k(A)| \geq 2$ then compute end points $\{a + \mathbb{i}b, c + \mathbb{i}d\}$ of $\operatorname{conv}(\mathcal{T} \cap \Lambda_k(A))$
        return $\dim = 1$, $\{a + \mathbb{i}b, c + \mathbb{i}d\}$
    if $\mathcal{T} \cap \Lambda_k(A) = \{a + \mathbb{i}b\}$ then return $\dim = 0$, $\{a + \mathbb{i}b\}$
```

```
        if $\mathcal{T} \cap \Lambda_k(A) = \emptyset$ then return dim $= -1$
    if $\dim V = -1$ then compute $g_A \in \mathbb{R}[a,b]$ as in Definition 3.13
        Compute representatives $\mathcal{R}$ for each connected component of $\mathbb{C} \backslash \{a + \mathbb{i}b : g_A(a,b) = 0\}$
        Use Algorithm 1 to compute $\mathcal{R} \cap \Lambda_k(A)$.
        if $\mathcal{R} \cap \Lambda_k(A) \neq \emptyset$ then return dim $= 2$, $g_A$, $\mathcal{R} \cap \Lambda_k(A)$
        else Compute $\mathcal{C} = \{a + \mathbb{i}b : t + ax + by$ divides $f_A$ or is tritangent to $f_A^{\mathrm{red}}\}$ (Remark 3.33).
            Use Algorithm 1 to compute $\mathcal{C} \cap \Lambda_k(A)$.
            if $\mathcal{C} \cap \Lambda_k(A) = \{a + \mathbb{i}b\}$ then return dim $= 0$, $\{a + \mathbb{i}b\}$
            else return dim $= -1$
```

Theorem 3.32. *Let $A \in \mathbb{C}^{n \times n}$ for which f_A is not a power of a linear form. For any $k < (n+1)/2$, Algorithm 4 computes the dimension and stated description of $\Lambda_k(A)$.*

Proof. Let S and V be as computed in the algorithm. Note that V is a projective linear subspace of $\mathbb{P}^2$, so $\dim(V) \in \{-1, 0, 1, 2\}$. Let

$$S' = \{[p_0 : p_1 : p_2] \in \mathbb{P}^2(\mathbb{R}) : -p_0 = \lambda_k(p_1, p_2) = \lambda_{n-k+1}(p_1, p_2)\} \text{ and } V' = \mathrm{span}(S') \subseteq \mathbb{P}^2.$$

The set S is contained in S' and so $\dim(V') \geq \dim(V)$. By Corollary 3.21, if $\dim(V) = 2$, then $\Lambda_k(A)$ is empty and if $\dim(V) = 1$, then $\Lambda_k(A)$ is either empty or the single point $a + \mathbb{i}b$ where $[1 : a : b] = V^{\perp}$, which can be determined by the membership test. If $\dim(V) = 0$, then S is a single point p. By Lemma 3.20(i), $\Lambda_k(A)$ is contained in the line $\{a + \mathbb{i}b : p_0 + ap_1 + bp_2 = 0\}$. By Theorem 3.24, $\Lambda_k(A)$ is the convex hull of the set of $a + \mathbb{i}b \in \mathcal{T} \cap \Lambda_k(A)$.

It remains to examine the case $S = \emptyset$. If S' is nonempty, then by Lemma 3.31, $\Lambda_k(A)$ is either empty or equal to $\{a + \mathbb{i}b\}$ where $t + ax + by$ is a linear factor of f_A, in which case $a + \mathbb{i}b \in \mathcal{C}$. Otherwise, S' is empty. By Theorem 3.19, $\Lambda_k(A)$ is then either two-dimensional, a singleton $\{a + \mathbb{i}b\}$ with $a + \mathbb{i}b \in \mathcal{C}$, or empty.

We claim that $\Lambda_k(A)$ is two-dimensional if and only if $\mathcal{R} \cap \Lambda_k(A)$ is nonempty. By Theorem 3.15, the boundary of $\Lambda_k(A)$ is contained in $\mathcal{V}_{\mathbb{R}}(g_A)$. In particular, a point in $\mathcal{R}$

belongs to $\Lambda_k(A)$ if and only if it belongs to the interior of $\Lambda_k(A)$. If there is any such point, then $\Lambda_k(A)$ is two-dimensional. Conversely, if $\Lambda_k(A)$ is two-dimensional, then it will contain some non-empty bounded connected component of $\mathbb{C}\backslash\{a + \mathbb{i}b : g(a,b) = 0\}$. The representative of this connected component gives a point in $\mathcal{R}\cap\Lambda_k(A)$. Moreover, by applying the membership test to $\mathcal{R}$, a set of representatives of each bounded connected component of the complement of g_A, we obtain at least one point for each connected component of $\Lambda_k(A)\backslash\{a + \mathbb{i}b : g_A(a,b) = 0\}$. If $\Lambda_k(A) = \{a + \mathbb{i}b\}$, then, by Theorem 3.19, we conclude that $t + ax + by$ is either a tritangent or a linear factor of f_A. In particular, $a + \mathbb{i}b \in \mathcal{C}\cap\Lambda_k(A)$. Finally, if $\Lambda_k(A)$ is empty, then no points in $\mathcal{T}$, $\mathcal{R}$ or $\mathcal{C}$ pass the membership test in Algorithm 1 and we correctly conclude that $\dim(\Lambda_k(A)) = -1$. □

We finish the discussion of the previous algorithm by describing how to compute the sets $\mathcal{R}$ and $\mathcal{C}$. We give an algorithm for computing $\mathcal{R}$ based on the theory of cylindrical algebraic decompositions for semialgebraic sets. For a thorough presentation see [3, Section 5.1].

Algorithm 5 Bounded connected components of $\mathbb{R}^2\backslash\mathcal{V}(g)$

Input: $g(a,b)$ a square-free polynomial in $\mathbb{R}[a,b]$

Output: A set of representatives for each bounded connected component of $\mathbb{R}^2\backslash\mathcal{V}(g)$.

```
Set $h = \mathrm{Disc}_a(g)$.
Compute the distinct real roots $b_1 < \ldots < b_m$ of $h$.
for $i = 1, \ldots, m-1$ do
    define $b'_i = (b_i + b_{i+1})/2$.
for $i = 1, \ldots, m-1$ do
    Find the distinct real roots $a_{i1} < \ldots < a_{in_i}$ of $g(a, b'_i)$.
    for $j = 1, \ldots, n_i$ do
        define $a'_{ij} = (a_{ij} + a_{i(j+1)})/2$.
Set $\mathcal{R} = \bigcup_{i=1}^{m-1}\{a'_{ij} + \mathbb{i}b'_i, \ldots, a'_{i(n_i-1)} + \mathbb{i}b'_i\}$.
return $\mathcal{R}$.
```

Remark 3.33 (Computing tritangent lines and linear factors)**.** Let $f \in \mathbb{C}[t,x,y]$ be a homogeneous polynomial and let f^{red} be the square-free part of f, which is homogeneous of some degree d. We can compute the set (a,b) for which the line $t+ax+by$ is tritangent to $\mathcal{V}(f)$ or divides f via elimination theory. We note that when $t+ax+by$ divides f, the restriction of f to this line is identically zero, which satisfies our definition for being tritangent, by taking $h=0$. Note that if $d<6$, then the only lines tritangent to f are its linear factors. The set of (a,b) for which $t+ax+by$ divides f^{red} is the variety of the ideal generated by the $d+1$ coefficients of $f^{\text{red}}(-ax-by,x,y)$ as a polynomial in x and y.

For $d \geq 6$, consider the ideal $I \subset \mathbb{C}[a,b,\ell_{00},\ell_{01},\ell_{10},\ell_{11},\ell_{20},\ell_{21},h_0,\ldots,h_{d-6}]$ generated by the $d+1$ coefficients of

$$f^{\text{red}}(-ax-by,x,y) - (\ell_{00}x+\ell_{01}y)^2 \cdot (\ell_{10}x+\ell_{11}y)^2 \cdot (\ell_{11}x+\ell_{12}y)^2 \cdot \left(\sum_{j=0}^{d-6} h_j x^j y^{d-6}\right) \quad (3.6)$$

in x and y. The line $t+ax+by$ is tritangent to $\mathcal{V}(f)$ if and only if there exists values of ℓ_{ij} and h_j for which the polynomial in (3.6) is identically zero. The set of (a,b) for which this holds is therefore equal to the variety of the elimination ideal

$$J = \text{radical}\left(\text{eliminate}\left(I, \{\ell_{00},\ell_{01},\ell_{10},\ell_{11},\ell_{20},\ell_{21},h_0,\ldots,h_{d-6}\}\right)\right).$$

Remark 3.34. For the above algorithms we focus on correctness rather than computational efficiency. For an implementation of the above algorithm some modifications can be made in order to make it more computationally efficient. It is also worth noting that most of the sets of points that we need can be reused when computing $\Lambda_k(A)$ for different k's. The polynomial g_A is also independent of k and therefore it needs to be computed only once.

3.6 Gallery of Examples

In this section we illustrate the behavior of the given algorithms of Section 3.5 and highlight the many possible subtle behaviors of higher-rank numerical ranges.

Example 3.35. In this example, we construct a matrix $\tilde{A}$ with 0-dimensional $\Lambda_k(\tilde{A})$ coming from a tritangent line not passing through any singularities of $\mathcal{V}(f_{\tilde{A}})$.

To start, let A denote the matrix from Example 1.1. The polynomial f_A defines a quartic plane curve. We define a parametric family of 6×6 matrices by

$$\tilde{A}(u) = \begin{pmatrix} B + u(1+\mathbb{i})I_2 & 0 \\ 0 & A \end{pmatrix} \quad \text{where} \quad B = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}$$

and I_2 denotes the 2×2 identity matrix. The boundary of the numerical range of B is the circle $\{a + \mathbb{i}b : 4a^2 + 4b^2 = 9\}$. The boundary of the numerical range of $B + u(1+\mathbb{i})I_2$ is this circle translated by $u(1+\mathbb{i})$. Fig. 3.4 shows the Kippenhahn curves and numerical ranges of $\tilde{A}(u)$ for $u = -1, \hat{u}$ and $-9/4$ where

$$\hat{u} = \frac{-4\sqrt{540 + 330\sqrt{3}} - 3\sqrt{1374 + 792\sqrt{3}}}{72 + 44\sqrt{3}}.$$

We see that for $u = -1$, $\Lambda_3(\tilde{A})$ is full dimensional, corresponding to the two-dimensional family of lines not intersecting the curve $\mathcal{V}_{\mathbb{R}}(f_{\tilde{A}})$. On the other hand, for $u = -9/4$, $\Lambda_3(\tilde{A})$ is empty, as every line intersects regions of $\mathbb{R}^3 \backslash \mathcal{V}_{\mathbb{R}}(f_{\tilde{A}})$ at which the matrix $tI + x\Re(\tilde{A}) + y\Im(\tilde{A})$ has four positive or negative eigenvalues. The value of $\hat{u}$ was computed so that for $u = \hat{u}$, $\Lambda_3(\tilde{A})$ is a single point. This corresponds to the unique line on which the matrix pencil has at most three positive and at most three negative eigenvalues. This line is tritangent to the curve $\mathcal{V}_{\mathbb{R}}(f_{\tilde{A}})$, as promised by Theorem 3.19.

In Fig. 3.4, the sets $\Lambda_2(\tilde{A})$ are shown in light blue. As promised by Theorem 3.15, the boundary of these regions is contained in the zero locus of $g_{\tilde{A}}$.

Example 3.36. Here we give a 6×6 matrix $\tilde{A}$ for which the curve $\mathcal{V}(f_{\tilde{A}})$ is irreducible and smooth but the numerical range $\Lambda_3(\tilde{A})$ is empty.

Let A denote the diagonal matrix $\operatorname{diag}(1, 1, -1+\mathbb{i}, -1+\mathbb{i}, -1-\mathbb{i}, -1-\mathbb{i})$. By the formula of $\Lambda_k(A)$ for normal matrices we know that $\Lambda_3(A)$ is empty. To construct a smooth curve

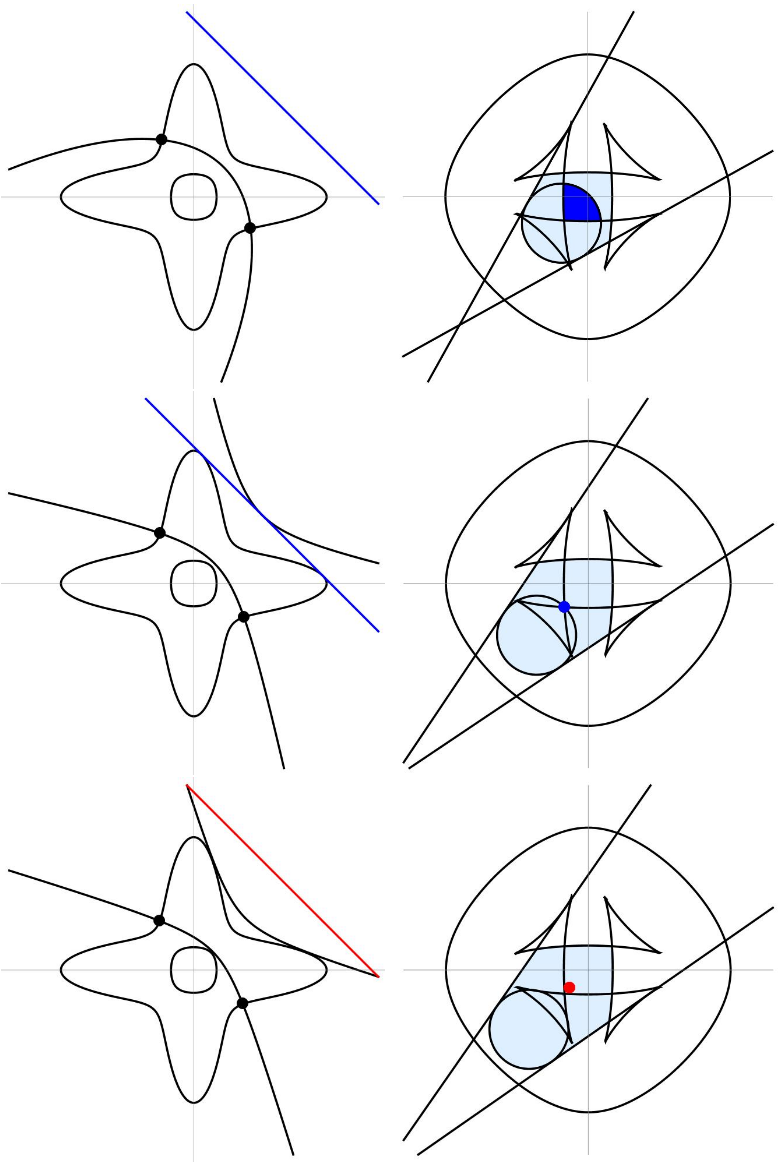

Figure 3.4: Kippenhahn curve and dual curve for Example 3.35. Here we study a family of curves obtained by allowing the circle on the right to be shifted along the line $a = b$. The regions shaded by light and dark blue correspond to $\Lambda_2(\tilde{A})$ and $\Lambda_3(\tilde{A})$, respectively. As we translate the circle, the rank-3 numerical ranges passes from being 2-dimensional to 0-dimensional to empty.

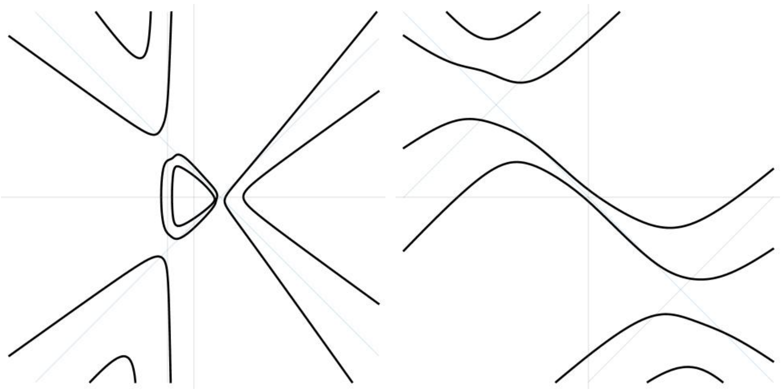

Figure 3.5: Kippenhahn curve for Example 3.36 in the charts $t = 1$ and $y = 1$.

with the same property we do a small perturbation. Specifically, define

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

The matrix $\tilde{A} = A + (1/5)B$ has the desired properties. The curve $\mathcal{V}(f_{\tilde{A}})$ is shown in the affine charts $t = 1$ and $y = 1$ in Fig. 3.5.

Example 3.37. As we saw in Section 3.4, zero-dimensional rank-k numerical ranges can also come from antipodal points on the curve $O_k(A)$, which give rise to singularities in the Kippenhahn curve $\mathcal{V}_{\mathbb{R}}(f_A)$. Consider the 4×4 matrix A from Example 3.7 with Kippenhahn polynomial $f_A = t^4 - 5t^2x^2 + 4x^4 - t^2y^2$. The point $0 + 0\mathbb{i}$, corresponding to the line $t = 0$, is the only one that passes the membership test for $\Lambda_2(A)$. The curve $\mathcal{V}_{\mathbb{R}}(f_A)$ and line $t = 0$ in the affine chart $y = 1$ are shown in Fig. 3.6.

If $a + \mathbb{i}b$ is the endpoint of a one-dimensional $\Lambda_k(A)$, the line $t + ax + by$ is either tangent to $\mathcal{V}_{\mathbb{R}}(f_A)$ at multiple points or has a zero of multiplicity $d+1$ at a singularity of multiplicity

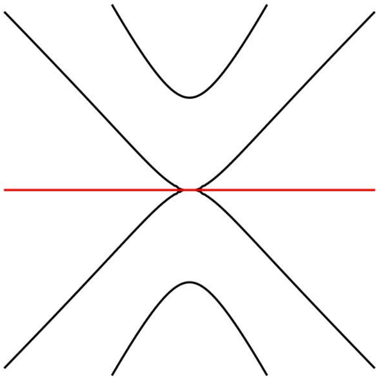

Figure 3.6: The Kippenhahn curve $\mathcal{V}_{\mathbb{R}}(f_A)$ from Example 3.37.

d. One might think that the algebraic condition for $\Lambda_k(A)$ to be a single point should be a degeneration of the algebraic condition for two endpoints of a one-dimensional $\Lambda_k(A)$. This would suggest that the corresponding line is either tangent to the curve $\mathcal{V}_{\mathbb{R}}(f_A)$ at multiple points or has a zero of multiplicity $d+2$ at a singularity of multiplicity d. The next example shows that this is not the case.

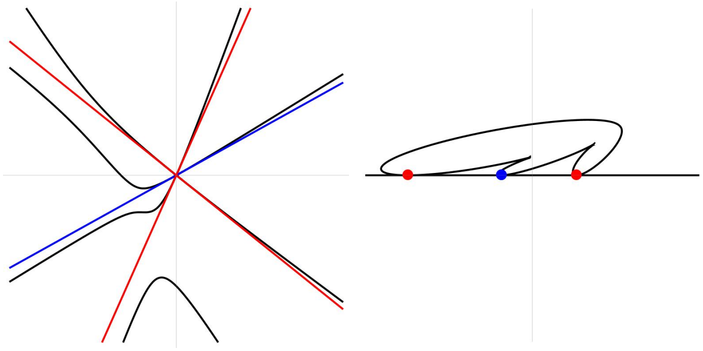

Figure 3.7: The curve $\mathcal{V}(f_A)$ from Example 3.38 along with its three $[0:0:1]$-tangent lines. On the right is the curve $\mathcal{V}_{\mathbb{R}}(g_A)$ along with the points corresponding to these lines. The unique point in $\Lambda_2(A)$ and corresponding line are shown in blue.

Example 3.38. Consider the matrix

$$A = \begin{pmatrix} 1+\mathbb{i} & 1 & 1 & 0 \\ 1 & -1 & -1 & 0 \\ 1 & -1 & -1 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

The Kippenhahn polynomial is

$$f_A(t,x,y) = t^4 - t^3x + t^3y - 5t^2x^2 - 2t^2xy - tx^2y + 2x^4 + x^3y$$

and its dual curve is given by

$$g_A = 49a^6 - 644a^5b + 196a^5 + 3824a^4b^2 - 2212a^4b + 98a^4 - 12172a^3b^3 + 8942a^3b^2 + 56a^3b - 294a^3 + 21248a^2b^4 - 16084a^2b^3 - 3420a^2b^2 + 2324a^2b - 147a^2 - 18836ab^5 + 11860ab^4 + 9244ab^3 - 4754ab^2 + 56ab + 98a + 7260b^6 - 5132b^5 - 1739b^4 - 1600b^3 + 2402b^2 - 672b + 49.$$

The curve $\mathcal{V}_{\mathbb{R}}(f_A)$ has a singularity at $p = [0:0:1]$ and $\lambda_2(0,1) = \lambda_3(0,1) = \lambda_4(0,1) = 0$. By Theorem 3.24, it follows that $\Lambda_2(A)$ is contained in the line $b = 0$. The curve has three p-tangent lines. The lines and the corresponding points of $\mathcal{V}_{\mathbb{R}}(g_A)$ are shown in Fig. 3.7. Of these, only one point passes the membership test for $\Lambda_2(A)$ and we conclude that $\Lambda_2(A)$ is this single point.

The curve $\mathcal{V}(f_A)$ has multiplicity three at p. The restriction of f_A to each p-tangent line has a single root of multiplicity four. In particular, this shows that we are unable to use algebraic methods to distinguish between the point of a zero-dimensional set $\Lambda_k(A)$ and the endpoint of a one-dimensional set $\Lambda_k(A)$.

The curve $\mathcal{V}_{\mathbb{R}}(f_A)$ in the affine chart $y = 1$ is the union of the graphs of the function $t = -\lambda_k(x\Re(A) + \Im(A))$ for $k = 1, 2, 3, 4$. In many other examples, the branches of $\mathcal{V}_{\mathbb{R}}(f_A)$ meeting at a singularity with a common tangent line correspond to the λ_k and λ_{n-k} eigenfunctions. Here the branches λ_2 and λ_4 meet along common tangent lines at p.

Chapter 4

THE FIBER OF SOLUTIONS OF THE INVERSE FIELD OF VALUES PROBLEM

The inverse Field of Values Problem (iFVP) is stated as follows. For $A \in \mathbb{C}^{n\times n}$ and $a+\mathbb{i}b \in \mathcal{W}(A)$ find a vector $z \in \mathbb{S}^{n-1}(\mathbb{C})$ such that $z^*Az = a+\mathbb{i}b$. We extend this question further by asking what is the structure of the set of all unit vectors such that $z^*Az = a+\mathbb{i}b$.

*4.1 A review of the *-congruence canonical form*

To solve the general field of values problem for the general case we use the *-congruence canonical decomposition. In order to actually compute U we need to explicitly find S such that S^*AS is in *-congruence canonical form.

Consider the space of complex square matrices in $\mathbb{C}^{n\times n}$. We say that two matrices A, B in $\mathbb{C}^{n\times n}$ are *-congruent if there exists an invertible matrix S such that $S^*AS = B$. Observe that this is a generalization of unitary equivalence, since any unitary matrix is invertible. Also, this should not be confused with the notion of *similarity.* Two matrices A and B are similar if there exists an invertible matrix S such that $S^{-1}AS = B$.

The *-congruence canonical form of A is a representative of the equivalence class of A under the relation of *-congruence. It plays the same role as the Jordan canonical form for the relation of similarity.

Theorem 4.1. *Any matrix A is *-congruent to a matrix B which is a direct summand of blocks of the following form.*

i) $J_m(0)$: upper triangular Jordan block of size m with eigenvalue 0.

ii) $H_{2m} = \begin{pmatrix} 0_m & I_m \\ J_m(\mu) & 0_m \end{pmatrix}$*: where* $|\mu| > 1$.

iii) $e^{\mathrm{i}\theta} K_m$*, where* $[K_m]_{ij} = \begin{cases} 1 & \textit{if } i+j = m+1 \\ \mathrm{i} & \textit{if } i+j = m+2 \end{cases}$, $\theta \in [0, 2\pi)$.

Moreover, B is unique up to permutation of its blocks.

Proof. This is proven in [21, Theorem 1.b] □

The *-congruence canonical form was used by Li and Sze [26] to prove that the higher-rank numerical ranges are convex. However, since our main motivating questions are concerned with computing higher-rank numerical ranges and their associated fibers we need to address the following questions: Given the matrix A, how can we compute its *-congruence canonical form? How to find an invertible S such that S^*AS is in *-congruence canonical form? The answers we have found work for symbolical computations. Unfortunately, the methods discussed, even just for finding the *-congruence canonical form, are known to be highly unstable numerically. That is, they are very susceptible to small perturbations in the entries of M.

An answer to the first question is given explicitly in [21, Section 4]. The algorithm can be divided into two parts.

1. Reduce the matrix A to the following form

   $$B \oplus J_{r_1}(0) \oplus J_{r_2}(0) \oplus \ldots \oplus J_{r_p}(0) \tag{4.1}$$

   where B is invertible.

2. Compute $B^{-*}B$, the *cosquare of B and compute is Jordan Canonical form. Group the blocks of the Jordan canonical form into pairs of the form $J_m(\mu) \oplus J_m(\overline{\mu}^{-1})$ for $|\mu| > 1$. The remaining blocks would be of the form $J_m(e^{\mathrm{i}\theta})$. Then for each pair of

Jordan blocks in the grouping above there is one block of the form $H_{2m}(\mu)$ and for each of the remaining Jordan blocks there is a corresponding block of the form $\xi e^{\mathrm{i}\theta/2}K_n$ with $\xi \in \{-1, 1\}$.

The first part is achieved via the *regularization algorithm* which is studied in detail in [22]. It is an iterative process that involves using *-congruence relations to write a singular matrix A in the form

$$\begin{bmatrix} A_{(1)} & B & 0 \\ 0 & C & D \\ 0 & 0 & 0_{m_1} \end{bmatrix}.$$

Then, if $A_{(i)}$ is singular, the same type of reductions are applied to it to obtain a new matrix $A_{(i+1)}$. This process is repeated until we find a matrix $A_{(\tau)}$ that is invertible.

The *-congruence relations used in the stage can be performed using unitary matrices [22, Theorem 2]. At this point, rather than giving an explicit S such that S^*AS has the form Eq. (4.1) they show that by applying successive *-congruence reductions they can get that decomposition. Therefore, a strategy to find S is to keep track of all the *-congruence reductions used.

For example, in [22, Lemma 4], for a full-rank complex rectangular matrix $M \in \mathbb{C}^{n\times m}$ with $m \geq n$, they find an invertible matrix V such that $MV = \begin{bmatrix} I & 0 \end{bmatrix}$. Such a matrix V can be obtained, for example, by reducing M to reduced column-echelon form and keeping track of all the column operations performed.

Once the regularization of A is computed the next step is to determine the *-congruence form of B. There is some nuance in determining the correct signs ξ for each of the blocks of the form $\xi e^{\mathrm{i}\theta/2}K_n$.

However, once they are found we could compute S in the following way.

1. Construct a block diagonal matrix C as explained in Part 2 of the algorithm above.

2. Use Jordan decomposition to compute an invertible matrix V such that $V^{-1}B^{-*}BV = J$ and J is in Jordan Canonical Form.

3. Use Jordan decomposition to compute an invertible matrix T such that $T^{-1}C^{-*}CT = J$. This is possible since both *cosquares are similar.

4. Take M to be $[VT^{-1}]^*B[VT^{-1}]C^{-1}$. M has the property that $MB = BM^*$.

5. Find matrix P such that $P^2 = M$ and $PB = BP^*$. Such a matrix can be found, via the theory of primary functions.

6. Then $P^{-*}[VT^{-1}]^*B[VT^{-1}]P^{-1} = C$, so $S = [VT^{-1}]P^{-1}$.

4.2 The fiber of solutions of the inverse field of values problem

Consider the map $w_A : \mathbb{C}^n \to \mathbb{C}$ given by

$$w_A(z) = z^*Az = z^*\Re Az + \mathbb{i}z^*\Im Az. \tag{4.2}$$

This maps the unit sphere to the numerical range $\mathcal{W}(A)$. The map above is not algebraic over $\mathbb{C}^n$ since we are taking conjugates of z. However, if we consider the map as a function of the real and imaginary parts of z then it becomes algebraic over $\mathbb{R}^{2n}$. To make this more evident, we give the following notation. For any Hermitian matrix $H \in \mathrm{Herm}_n$, let $q_H(x, y) := (x - \mathbb{i}y)^\top H(x + \mathbb{i}y)$, where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Observe that q_H is a homogeneous quadratic polynomial in $\mathbb{R}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ by Lemma 2.1. Then we can express the map w_A as

$$w_A(x + \mathbb{i}y) = q_{\Re A}(x, y) + \mathbb{i}q_{\Im A}(x, y).$$

Later we will discuss how to find $q_{\Re A}$ and $q_{\Im A}$ explicitly.

In this section, we are interested in studying the geometric and algebraic properties of the fibers $w_A^{-1}(a + \mathbb{i}b)$. These fibers are real affine algebraic varieties in $\mathbb{R}^{2n}$ defined by the following equations:

$$q_{\Re A}(x,y) = a, \quad q_{\Im A}(x,y) = b \quad \text{and} \quad q_{I_n}(x,y) = 1. \tag{4.3}$$

We start by studying the case where A is normal. We begin by proving some simple statements about polytopes, i.e., convex hulls of finite subsets of $\mathbb{R}^m$. Let K be the convex hull of $\{p_1, \ldots, p_n\} \in \mathbb{R}^m$. Consider the linear map $\psi : \mathbb{R}^n \to \mathbb{R}^m$ given by

$$(t_1, \ldots, t_n) \mapsto t_1 p_1 + \cdots + t_n p_n.$$

Denote by $\{e_1, \ldots e_n\}$ the canonical basis of $\mathbb{R}^n$. Let $\Delta_{[n]}$ be the simplex with n vertices $\operatorname{conv}\{e_i : i \in [n]\}$. By the definition of convex hull, we have $\psi(\Delta_{[n]}) = K$. The faces of $\Delta_{[n]}$ correspond to subsets of $[n]$. We denote by Δ_I the face of $\Delta_{[n]}$ corresponding to $I \subseteq [n]$, $\Delta_I = \operatorname{conv}\{e_i : i \in I\}$.

Lemma 4.2. *Let $p \in K$ and suppose that $p \in \operatorname{ri} F$ where F is a face of K. Let $I = \{i \in [n] : p_i \in F\}$. Then $\psi^{-1}(p) \cap \Delta_{[n]} \subseteq \Delta_I$. Moreover, $\dim(\psi^{-1}(p) \cap \Delta_{[n]}) = |I| - 1 - \dim F$.*

Proof. Consider a supporting hyperplane of F given by some linear equation $\ell(\cdot) = d$. Then $\ell(p) = d$, but also $\ell(p_i) = d$ for all $i \in I$ and $\ell(p_j) > d$ for all $j \notin I$. Consider $t = (t_1, \ldots, t_n) \in \Delta_{[n]}$ such that $\psi(t_1, \ldots, t_n) = p$. Then

$$\ell(p) = \ell \circ \psi(t) = \sum_{i=1}^{n} t_i \ell(p_i) \geq \sum_{i=1}^{n} t_i d = d.$$

Observe that the above is an equality if and only if t_j is equal to 0 for all $j \notin I$. In other words, $\psi(t) \in F$ if and only if $t \in \Delta_I$.

For the statement about the dimension we use the notion of affine hull. We denote the affine hull of a set S by $\operatorname{aff}(S)$. Recall that for a convex set S, $\dim(S) = \dim(\operatorname{aff}(S))$. The linear map ψ gives an affine surjective map between $\operatorname{aff}(\Delta_I)$ and $\operatorname{aff}(F)$. In particular,

$$|I| - 1 = \dim \Delta_I = \dim F + \dim(\psi^{-1}(p) \cap \operatorname{aff} \Delta_{[n]}).$$

Therefore, $\psi^{-1}(p) \cap \operatorname{aff} \Delta_{[n]}$ is an affine subspace of $\mathbb{R}^n$ with dimension $|I| - 1 - \dim F$. On the other hand, since ψ is a linear map, we also have $\psi(\operatorname{ri} \Delta_I) = \operatorname{ri} F$ by Theorem 2.4. In particular, $\psi^{-1}(p) \cap \Delta_{[n]}$ must contain points in the relative interior of Δ_I. $\square$

The previous lemma allow us to prove the following theorem about the fibers of w_A. Recall that for normal matrices the numerical range $\mathcal{W}(A)$ is a polygon in $\mathbb{C} \simeq \mathbb{R}^2$.

Theorem 4.3. *Suppose that A is a normal $n \times n$ matrix and let $\{p_1, \ldots, p_n\} \subseteq \mathbb{C}$ be its eigenvalues. Let $a + \mathbb{i}b \in \mathcal{W}(A)$ and F be a minimal face of $\mathcal{W}(A)$ containing $a + \mathbb{i}b$. Then $w_A^{-1}(a + \mathbb{i}b)$ is a real affine algebraic variety of dimension $2|I| - \dim F - 1$, where $I = \{i \in [n] : p_i \in F\}$.*

Proof. We may assume without loss of generality that A is diagonal. Then

$$w_A(x + \mathbb{i}y) = (x - \mathbb{i}y)^\top A(x + \mathbb{i}y) = \sum_{i=1}^{n} p_i(x_i^2 + y_i^2).$$

We may decompose the map $w_A : \mathbb{R}^{2n} \to \mathbb{C}$ as follows. Consider the map $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^n$ given by

$$(x_1, \ldots x_n, y_1, \ldots, y_n) \mapsto (x_1^2 + y_1^2, \ldots, x_n^2 + y_n^2).$$

This map sends $\mathbb{S}^{2n-1}$ to $\Delta_{[n]}$. Moreover, consider the map $\psi : \mathbb{R}^n \to \mathbb{C}$ given by

$$t \mapsto \sum_{i=1}^{n} p_i t_i.$$

Then $w_A(x + \mathbb{i}y) = \psi \circ \varphi(x, y)$. By Lemma 4.2, $\psi^{-1}(a + \mathbb{i}b) \cap \Delta_{[n]}$ is contained in Δ_I. The minimality of F guarantees that $a + \mathbb{i}b$ is contained in the relative interior of F. Then $\psi^{-1}(a + \mathbb{i}b) \cap \Delta_{[n]}$ is a section of Δ_I of dimension $|I| - 1 - \dim F$.

Now, consider any point $q = (q_1, \ldots, q_n) \in \mathbb{R}^n_{\geq 0}$ in the interior of $\psi^{-1}(a + \mathbb{i}b)$. The elements of the fiber $\varphi^{-1}(q)$ are of the form

$$(q_1 \cos\theta_1, \ldots, q_n \cos\theta_n, q_1 \sin\theta_1, \ldots, q_n \sin\theta_n), \qquad (\theta_1, \ldots, \theta_n) \in [0, 2\pi)^n.$$

Thus, the fiber $\varphi^{-1}(q)$ is a product of circles. The dimension of this fiber is the number of circles in the product, which corresponds to the number of nonzero coordinates in q. Because q is in the relative interior of Δ_I we know that this number is equal to $|I|$. Therefore,

$$\dim w_A^{-1}(a + \mathbb{i}b) = \dim \psi^{-1}(a + b\mathbb{i}) + \dim \varphi^{-1}(q) = 2|I| - \dim F - 1.$$

□

Corollary 4.4. *If A is normal and μ is in the interior of $\mathcal{W}(A)$ then $w_A^{-1}(\mu)$ has dimension $2n-3$.*

Corollary 4.5. *If $\mathcal{W}(A)$ has dimension 1 and μ is in the relative interior of $\mathcal{W}(A)$ then $w_A^{-1}(\mu)$ has dimension $2n-2$.*

Proof. If $\dim \mathcal{W}(A) = 1$ then all higher-rank numerical ranges are contained in a line segment. We conclude that A is normal by [13, Corollary 6]. □

Next, we study the case where A is a general matrix and μ lies on the boundary of $\mathcal{W}(A)$. It turns out that the behavior is very similar to that of the normal case. Note that fibers are invariant under translations and rotations.

Lemma 4.6. *Suppose that $0 \in \mathcal{W}(A)$. Also, suppose that $a \geq 0$ for any $a + b\mathbb{i} \in \mathcal{W}(A)$. Then $\Re A$ is positive semidefinite and at least one of its eigenvalues is zero.*

Proof. If we consider the cone $\widehat{\mathcal{W}(A)}$ then the linear functional $\langle(0,1,0),(c,a,b)\rangle = a$ is nonnegative over $\widehat{\mathcal{W}(A)}$. Therefore, (0,1,0) belongs to $\widehat{\mathcal{W}(A)}^*$. But

$$\widehat{\mathcal{W}(A)}^* = \{(t,x,y) \in \mathbb{R}^3 : tI_n + x\Re A + y\Im A \succeq 0\},$$

so $\Re A \succeq 0$.

On the other hand, since $0 \in \mathcal{W}(A)$, then $(1,0,0) \in \widehat{\mathcal{W}(A)}$. By duality, the linear functional $\langle(t,x,y),(1,0,0)\rangle = t$ is non-negative on $\widehat{\mathcal{W}(A)}^*$. Moreover, (0,1,0), satisfies the equation $t = 0$. The plane $t = 0$ is therefore a supporting hyperplane for a proper face of $\widehat{\mathcal{W}(A)}^*$ containing (1,0,0) which shows that this point is on the boundary of $\widehat{\mathcal{W}(A)}^*$.

Finally, points in the boundary of $\widehat{\mathcal{W}(A)}^*$ must satisfy that $\det(tI_n + x\Re A + y\Im A) = 0$. Therefore, the kernel of $\Re A$ must be nontrivial. □

After applying a unitary change of basis, we can assume that $\Re A$ is diagonal and that the first m elements in the diagonal are zero, and the rest of the elements are strictly greater than zero.

Theorem 4.7. *Suppose that A satisfies the assumptions in Lemma 4.6 and let $m = \operatorname{corank}(\Re A)$. Furthermore, suppose that $\Re A$ is diagonal and the first m entries in its diagonal are zero. Then $w_A^{-1}(0) = w_B^{-1}(0)$ where B is the m-by-m leading principal submatrix of $\Im A$. Moreover, the dimension of $w_B^{-1}(0)$ is equal to $2m-2$ if $\mathcal{W}(B)$ has dimension 1 and the point 0 is in the relative of $\mathcal{W}(B)$. Otherwise, the dimension is $2l-1$ where $l = \operatorname{corank}(B)$.*

Proof. Recall that $w_A^{-1}(0)$ is defined by the following equations:

$$q_{\Re A}(x,y) = 0, \qquad q_{\Im A}(x,y) = 0 \quad \text{and} \quad q_{I_n}(x,y) = 1. \tag{4.4}$$

Due to our assumption about $\Re A$, we can write the first equation as

$$\sum_{i=m+1}^{n} d_i(x^2 + y^2) = 0,$$

with $d_i > 0$. Because this is a sum of squares, its vanishing set is the linear subspace $V \subseteq \mathbb{R}^{2n}$ defined by the equations $x_i = y_i = 0$ for all $m+1 \leq i \leq n$. Thus, $w_A^{-1}(0)$ can be obtained by taking the last two equations and restricting them to V. These equations are the same as those that define the fiber $w_B^{-1}(0)$.

The last two statements in the proof follow by Theorem 4.3 and the fact that B is Hermitian if and only if $\mathcal{W}(B)$ is contained in the real line. ☐

We end this section by considering the case where A is a general matrix and $a + \mathbb{i}b$ is in the interior of A. First, we introduce a projective variety version of the fiber $w_A^{-1}(a + \mathbb{i}b)$.

We start by observing that vectors that satisfy the third equation $q_{I_n} = 1$ in Eq. (4.3), must also satisfy $aq_{I_n} = a$. Then we can rewrite the first two equations for $w_A^{-1}(a + \mathbb{i}b)$ in the following way.

$$q_{\Re A} - aq_{I_n} = 0, \qquad q_{\Im A} - bq_{I_n} = 0 \quad \text{and} \quad q_{I_n} = 1.$$

Notice that the first two equations are now homogeneous of degree 2. Then consider the projective variety

$$\mathcal{P}_{A,a+\mathbb{i}b} := \mathcal{V}(q_{\Re A} - aq_{I_n}, q_{\Im A} - bq_{I_n}) \subseteq \mathbb{P}^{2n-1}(\mathbb{C}),$$

and its real projective counterpart

$$\mathcal{P}_{A,a+\mathbb{i}b}(\mathbb{R}) := \mathcal{V}_{\mathbb{R}}(q_{\Re A} - aq_{I_n}, q_{\Im A} - bq_{I_n}) \subseteq \mathbb{P}^{2n-1}(\mathbb{R}).$$

Observe that there is a 2-to-1 correspondence between $w_A^{-1}(a+\mathbb{i}b)$ and $\mathcal{P}_{a+\mathbb{i}b}(\mathbb{R})$. In particular, the dimension of these varieties is the same. Also, observe that we can obtain the same system by taking the translation $A-(a+\mathbb{i}b)I_n$ and obtaining the equations in Eq. (4.3) for the fiber at zero. We use this new approach to prove the converse of Corollary 4.5.

Lemma 4.8. *If $w_A^{-1}(a+\mathbb{i}b)$ has an irreducible component $\mathcal{X}$ of dimension $2n-2$ then $\mathcal{W}(A)$ has dimension 1 and μ is in its relative interior. In this case $w_A^{-1}(a+\mathbb{i}b) = \mathcal{X} \cup (-\mathcal{X})$.*

Proof. This irreducible component corresponds to an irreducible component $\mathcal{Y}$ of dimension $2n-2$ of $\mathcal{P}_{a+\mathbb{i}b}$. Due to its dimension, there exists a homogeneous polynomial h such that $\mathcal{V}_{\mathbb{C}}(p) = \mathcal{Y}$. Then h must be a common divisor of $q_{\Re A} - aq_{I_n}$ and $q_{\Im A} - bq_{I_n}$. In particular, its degree must be either 2 or 1.

In the first case, we have that $q_{\Re A} - aq_{I_n}$ and $q_{\Im A} - bq_{I_n}$ are scalar multiples of h. This implies that $\Re A$ is a scalar multiple of $\Im A$. From here we conclude that A is normal and $\dim \mathcal{W}(A) \leq 1$. But, by Theorem 4.3 in order for the fiber to have dimension $2n-2$, $\dim \mathcal{W}(A) = 1$ and $a+\mathbb{i}b$ must be in the relative interior of $\mathcal{W}(A)$. Otherwise, if $a+\mathbb{i}b$ was a vertex of $\mathcal{W}(A)$ then its fiber would have dimension $2l-1$, where l is the algebraic multiplicity of the eigenvalue $a+\mathbb{i}b$ in A. This dimension is odd and, therefore, can never equal $2n-2$.

The second case, $\deg(h) = 1$ is actually impossible. Suppose by contradiction that $q_{\Re A} - aq_{I_n}$ and $q_{\Im A} - bq_{I_n}$ have h as a common linear factor. By Lemma 2.1,

$$q_{\Re A} - aq_{I_n} = \mathbf{z}^\top \widehat{\Re A - aI_n} \mathbf{z},$$

where $\mathbf{z} = (x, y)$.

Then $q_{\Re A} - aq_{I_n}$ is reducible if and only if $\operatorname{rank} \widehat{\Re A - aI_n} \leq 2$ [15, Exercise 5.12]. This, in turn, happens if and only if $\operatorname{rank}(\Re A - aI_n) \leq 1$. In particular $\Re A - aI_n$ is either positive or negative semidefinite. As we saw in Theorem 4.7 this restricts $a + \mathbb{i}b$ to be in the boundary of $\mathcal{W}(A)$. But the dimension of $w_A^{-1}(a + \mathbb{i}b)$ can only be $2n - 2$ if $\Re A - aI_n = 0_n$.

By applying the same argument, we conclude that $\Im A - bI_n = 0_n$. But then $w_A^{-1}(a+\mathbb{i}b) = \mathbb{S}^{2n-1}$ which in particular has no irreducible components of dimension $2n - 2$.

The last statement follows by observing that $\mathcal{P}_{A,a+\mathbb{i}b} = \mathcal{V}(h) = \mathcal{Y}$. □

Next we use *-congruence, which is defined in Section 4.1, to reduce the general case to the case when A is in *-congruence canonical form. Through the remaining of this section we assume without loss of generality that $a + \mathbb{i}b = 0$.

To simplify notation we denote $\mathcal{P}_{A,0}$ as $\mathcal{P}_A$.

Lemma 4.9. *For any invertible $S \in \mathbb{C}^{n\times n}$ and $A \in \mathbb{C}^{n\times n}$, the projective varieties $\mathcal{P}_A$ and $\mathcal{P}_{S^*AS}$ are isomorphic.*

Proof. The linear automorphism given by $(x + \mathbb{i}y) \mapsto S(x + \mathbb{i}y)$ in $\mathbb{C}^n$ extends to a linear automorphism in $\mathbb{R}^{2n}$ which in turn gives rise to a linear automorphism ϕ in $\mathbb{P}^{2n-1}$. The proof follows by observing that $\phi(\mathcal{P}_{S^*AS}) = \mathcal{P}_A$.

Take (x, y) such that $(x + \mathbb{i}y)^*S^*AS(x + \mathbb{i}y) = 0$ and set

$$(x', y') := \phi(x, y) = (\operatorname{Re} S(x + \mathbb{i}y), \operatorname{Im} S(x + \mathbb{i}y)).$$

Observe that

$$(x' + \mathbb{i}y')^*A(x' + \mathbb{i}y') = (S(x + \mathbb{i}y))^*A(S(x + \mathbb{i}y)) = (x + \mathbb{i}y)^*S^*AS(x + \mathbb{i}y) = 0.$$

□

Remark 4.10. In general, the isomorphism above does not give an isomorphism of affine varieties between $w_A^{-1}(0)$ and $w_{S^*AS}^{-1}(0)$. However, we do obtain the map

$$\varphi : w_{S^*AS}^{-1}(0) \to w_{S^*AS}^{-1}(0), \qquad z \mapsto \frac{Sz}{\|Sz\|_2}.$$

This map in bijective and semialgebraic.

Lemma 4.11. *If* 0 *is in the relative interior of* $\mathcal{W}(A)$ *then it is also in the relative interior of* $\mathcal{W}(S^*AS)$ *for any invertible matrix* S*.*

Proof. Recall that $\mathcal{W}(A)$ is a spectahedral shadow as is shown in Corollary 2.18. It is the image of $Q = \{X \in \mathrm{PSD}_n : \langle X, I_n \rangle = 1\}$ under the projection map

$$\Phi_A(X) = \langle \Re A, X \rangle + \mathbb{i}\langle \Im A, X \rangle.$$

Observe that Q is convex and Φ_A is a linear map between convex sets. Therefore, $\Phi(\operatorname{ri} Q) = \operatorname{ri} \mathcal{W}(A)$ and in particular this implies that there exists $X \in \mathrm{PD}_n$ such that $\langle X, I_n \rangle = 1$ and $\Phi(X) = 0$. Now, we show that $\Phi_{S^*AS}(S^{-1}XS^{-*}) = 0$. Indeed,

$$\langle S^*(\Re A)S, S^{-1}XS^{-*} \rangle = \operatorname{tr}(S^{-1}XS^{-*}S^*(\Re A)S) = \operatorname{tr}(X\Re A) = 0$$

Similarly, we can show that $\langle S^*(\Re A)S, S^{-1}XS^{-*} \rangle = 0$. Observe that $S^{-1}XS^{-*} \in \mathrm{PD}_n$ since we are multiplying invertible matrices and positive semidefiniteness is preserved under *-congruence. By scaling $S^{-1}XS^{-*}$ so that it has trace 1 we have found a Hermitian matrix in the interior of Q, such that $\Phi_{S^*AS}(0)$. By the same argument about the image of relative interiors we conclude that 0 is in the relative interior of $\mathcal{W}(S^*AS)$. □

The previous lemmas allow us to study the fiber of w_A by taking the *-congruence canonical form of A and finding the fiber of that element instead.

Theorem 4.12. *Let* B *be the *-congruence canonical form of* A *as defined in Theorem 4.1. Then 0 is in the interior of* $\mathcal{W}(A)$ *if and only* B *can be written as* $B_1 \oplus B_2$ *where* B_1 *is one of the following:*

i) $J_m(0)$ *with* $m \geq 2$.

ii) $H_{2m}(\mu)$ *with* $m \geq 1$ *and* $|\mu| > 1$.

iii) $e^{\mathbb{i}\theta}K_m$ *with* $m \geq 3$.

iv) $e^{\mathbb{i}\theta}K_2 \oplus [e^{\mathbb{i}(\theta+\theta')}]$ *with* $0 \leq \theta < 2\pi$ *and* $\pi/2 < \theta' < 3\pi/2$.

v) $e^{\mathbb{i}\theta_1}K_2 \oplus e^{\mathbb{i}\theta_2}K_2$ *with* $0 \leq \theta_1 < \theta_2 < 2\pi$*.*

vi) $[e^{\mathbb{i}\theta_1}] \oplus [e^{\mathbb{i}\theta_2}] \oplus [e^{\mathbb{i}\theta_3}]$ *such that 0 is in the interior of* $\operatorname{conv}\{e^{\mathbb{i}\theta_1}, e^{\mathbb{i}\theta_2}, e^{\mathbb{i}\theta_3}\}$*.*

Proof. ($\Leftarrow$): It is enough to assume that $B = B_1$, in other words, there are no other blocks appearing in the *-congruence canonical form of A. By Lemma 2.19 if 0 is in the interior of $\mathcal{W}(B_1)$ it would also be in the interior of $\mathcal{W}(B)$. Denote by $\{e_1, \dots e_n\}$ the canonical basis of $\mathbb{C}^n$.

Type i) Suppose that $B = J_n(0)$. To show that 0 is in the interior, observe that if we take $X = [e_1\, e_2]$ then $X^*BX = J_2(0)$ and

$$\mathcal{W}(J_2(0)) = \{x + \mathbb{i}y \in \mathbb{C} : x^2 + y^2 \leq 1/4\}.$$

This is a circle of 1/2 radius centered at the origin. Therefore, $0 \in \operatorname{int} \mathcal{W}(X^*BX) \subseteq \operatorname{int} \mathcal{W}(B)$.

Type ii) After a rotation and a unitary change of basis we can assume that $B = H_{2m}(|\mu|)$. Take $X = [e_1\, e_{m+1}]$. Then $X^*BX = H_2(|\mu|)$ and

$$\mathcal{W}(H_2(|\mu|)) = \left\{ x + \mathbb{i}y : (x, y) \in \mathbb{R}^2 \text{ and } \left(\frac{x}{|\mu|+1}\right)^2 + \left(\frac{y}{|\mu|-1}\right)^2 \leq \frac{1}{4} \right\}.$$

This is a nondegenerate ellipse centered at the origin and therefore 0 is in its interior.

Type iii) Without loss of generality we can assume that $\theta = 0$. To check that 0 is in the interior of $\mathcal{W}(B)$ consider the following vectors: $x_1 = (e_1 + \mathbb{i}e_2)/\sqrt{2}$ and $x_2 = e_n$ and take $X = [e_1\, e_2]$. We have two cases.

For $n = 3$, we have

$$X^*BX = \begin{bmatrix} \frac{1}{2} & 0 \\ \sqrt{2} & 0 \end{bmatrix}.$$

The numerical range of this 2 by 2 matrix is an ellipse containing 0 in its origin.

For $n \geq 4$, we have

$$X^*BX = \begin{bmatrix} 0 & 0 \\ \sqrt{2} & 0 \end{bmatrix}.$$

The numerical range in this case is the circular disk centered at 0 with radius $1/\sqrt{2}$.

Type iv) Without loss of generality assume that $\theta = 0$. Observe that

$$\mathcal{W}(B) = \operatorname{conv}(\mathcal{W}(K_2) \cup e^{i\theta'}).$$

Note that $\mathcal{W}(K_2)$ is the ellipse

$$\{a + \mathbb{i}b : a^2 + \frac{(b - 1/2)^2}{4} \leq 1\}.$$

The point 0 is at the boundary of the ellipse and the only supporting hyperplane is $b = 0$. In particular, any other line passing through zero contains points of $\mathcal{W}(K_2)$ in both half-planes defined by the line. Therefore, if we assume by contradiction that 0 is in the boundary of $\mathcal{W}(B)$ then $b \geq 0$ for all points in $\mathcal{W}(B)$ but this is not true for $e^{\mathbb{i}\theta'}$ since $\sin(\theta') < 0$.

Type v) Assume that $\theta = 0$. Then $\mathcal{W}(B) = \operatorname{conv}(\mathcal{W}(K_2) \cup \mathcal{W}(e^{i\theta'}K_2))$.

As in the previous case, if 0 was in the boundary of $\mathcal{W}(B)$ then $b \geq 0$ for all $a + \mathbb{i}b \in \mathcal{W}(B)$. But for any $\theta' \neq 0$, $\mathcal{W}(e^{i\theta'}K_2))$ has points such that $b < 0$. In fact, the only supporting line for 0 in $\mathcal{W}(e^{i\theta}K_2)$ is $-a\sin\theta + b\cos\theta = 0$.

Type vi) This is true by assumption.

($\Rightarrow$): Observe that if B has no blocks of the types above then its decomposition after a rotation must be of the form

$$[0]^{\oplus m_1} \oplus K_2^{\oplus m_2} \oplus \bigoplus_{i=1}^{m_3} [e^{\mathbb{i}\theta_i}], \qquad 0 \leq \theta_i \leq \pi,$$

Then it follows that $b \geq 0$ for all $a + \mathbb{i}b \in \mathcal{W}(B)$. This is enough to show that either $0 \notin \mathcal{W}(B)$ or it is in the boundary of $\mathcal{W}(B)$. □

A consequence of the previous lemma is that for 0 to be in the boundary of $\mathcal{W}(A)$ then the *-congruence canonical form of A can only consist of blocks of size 1 or 2×2 blocks of the form $e^{\mathbb{i}\theta}K_2$. We can then determine the dimension of the fiber by exploiting this structure and using the following general fact from real algebraic geometry.

Lemma 4.13. *Let* q_1 *and* q_2 *be two quadratic homogeneous polynomials in* $\mathbb{R}[x_1, \ldots, x_{n+1}]$*. Suppose that* q_1 *and* q_2 *have no linear factors in common. Let* $X = \mathcal{V}(q_1, q_2) \subset \mathbb{P}^n(\mathbb{C})$ *and* $X(\mathbb{R}) = \mathcal{V}_{\mathbb{R}}(q_1, q_2) \subset \mathbb{P}^n(\mathbb{R})$*. Furthermore, suppose that there exists a point* $p \in X(\mathbb{R})$ *such that the Jacobian*

$$\begin{bmatrix} \nabla q_1(p) \\ \nabla q_2(p) \end{bmatrix}$$

has rank 2. Then p *is smooth in* X*. Moreover,* $\dim X(\mathbb{R}) = \dim X = n - 2$*.*

Proof. Let Y be an irreducible component of X. Because $X = \mathcal{V}(q_1, q_2)$ it follows that $\operatorname{codim} Y \geq 2$. The codimension is 0 if and only if q_1 and q_2 are both equal to 0 in which case the Jacobian above would be 0 as well. Suppose by contradiction that there exists some irreducible component Y such that $\operatorname{codim}(X) = 1$. Then there exists an irreducible polynomial h such that $\mathcal{V}(h) = Y$ [15, Proposition 1.13]. The polynomial h must be a common factor of q_1 and q_2. In particular, the degree of h is at most two. But since q_1 and q_2 have no common linear factors, it follows that the degree of h is 2 and q_1 and q_2 are scalar multiples of q_2. Then $X = Y$ and the Jacobian has rank at most 1 for any $p \in X$, contradicting our assumption.

Let q_1 and q_2 be part of a set of generators for X. Since the Jacobian above has rank 2 it follows that p is a smooth point of X.

Let Y be an irreducible component of X that contains p. By [35, Theorem 5.1] or [27, Theorem 12.6.1], $Y(\mathbb{R})$ is dense in Y. In particular, $Y(\mathbb{R})$ has the same dimension as Y, which is $2n - 2$. □

We end the section by proving the following theorem on the dimension of fibers for points in the interior of $\mathcal{W}(A)$.

Theorem 4.14. *If* 0 *is in the interior of* $\mathcal{W}(A)$ *then* $w_A^{-1}(0)$ *has dimension* $2n - 3$*.*

Proof. Since $\dim w_A^{-1}(0) = \dim \mathcal{P}_A(\mathbb{R})$, we focus on checking that this last projective variety has dimension $2n - 3$. By Lemma 4.11, we can substitute A with its *-congruence canonical form. Let B be the *-congruence canonical form of A. By Lemma 4.11, we can assume that

$B = B_1 \oplus B_2$ where B_1 is of the form i)-vi). By Lemma 4.13, it is enough to find a real point $\mathbf{z} \in \mathcal{P}_B(\mathbb{R}) \subseteq \mathbb{P}^{2n-1}(\mathbb{R})$ such that its Jacobian has rank 2. By Lemma 2.1 this Jacobian is of the form

$$\begin{pmatrix} \nabla q_{\Re B}(\mathbf{z}) \\ \nabla q_{\Im B}(\mathbf{z}) \end{pmatrix} = \begin{pmatrix} \nabla \mathbf{z}^\top \widehat{\Re B} \mathbf{z} \\ \nabla \mathbf{z}^\top \widehat{\Im B} \mathbf{z} \end{pmatrix} = \begin{pmatrix} 2\widehat{\Re B} \mathbf{z} \\ 2\widehat{\Im B} \mathbf{z} \end{pmatrix}.$$

We can further reduce this to the case where $B_2 = 0$. Suppose that there exists $\mathbf{z_1} \in \mathcal{P}_{B_1}$ such that $2\Re B_1 \mathbf{z_1}$ and $2\Im B_1 \mathbf{z_1}$ are linearly independent. Observe that $q_{\Re B} = q_{\Re B_1} + q_{\Re B_2}$ and that these two polynomials do not have any variables in common (similarly for $q_{\Im B}$). Then $\mathbf{z} = \mathbf{z_1} \oplus 0 \in \mathcal{P}_B$ would be the point that we are looking for on the account that

$$\begin{pmatrix} \nabla q_{\Re B}(\mathbf{z}) \\ \nabla q_{\Im B}(\mathbf{z}) \end{pmatrix} = \begin{pmatrix} 2\widehat{\Re B_1} \mathbf{z_1} & 0 \\ 2\widehat{\Im B_1} \mathbf{z_1} & 0 \end{pmatrix}$$

and this Jacobian is also full-rank. Let $\{e_1^x, \dots, e_n^x, e_1^y, \dots, e_n^y\}$ denote the canonical base of $\mathbb{R}^{2n}$.

Type i) Consider $\mathbf{z} = e_1^x$.

$$2\Re B = (J_n(0) + J_n(0)^\top) \quad \text{and} \quad 2\Im B = \mathbb{i}(J_n(0)^\top - J_n(0)).$$

Then $2\widehat{\Re B} e_1^x = e_2^x$ and $2\widehat{\Im B} e_1^x = e_2^y$. Clearly, these two vectors are linearly independent.

Type ii) Under our assumptions, $n = 2m$. Consider $\mathbf{z} = e_1^x$. Let $\mu = a + \mathbb{i}b$.

$$2\Re B = \begin{pmatrix} 0_m & I_m + J_m(\mu^*)^\top \\ I_m + J_m(\mu) & 0_m \end{pmatrix} \qquad 2\Im B = -\mathbb{i} \begin{pmatrix} 0_m & I_m - J_m(\mu^*)^\top \\ I_m - J(\mu) & 0_m \end{pmatrix}$$

Then $2\widehat{\Re B} e_1^x = (1+a)e_{m+1}^x + b e_{m+1}^y$ and $2\widehat{\Im B} e_1^x = b e_{m+1}^x + (a-1)e_{m+1}^y$. The only non zero elements in these vector form the following minor of the Jacobian:

$$\begin{vmatrix} 1+a & b \\ b & 1-a \end{vmatrix} = 1 - a^2 - b^2.$$

Since $\mu = a + \mathbb{i}b$ has norm bigger than 1 we conclude that the previous minor is nonnegative and that these two vectors are linearly independent.

Type iii) Without loss of generality assume that $B = K_n$, i.e., $\theta = 0$. Take $\mathbf{z} = e_n^x$. Observe that $(e_n^x)^* K_n e_n^x = (e_n^x)^*(e_1^x + \mathbb{i} e_2^x) = 0$.

$$\Re K_n = \begin{bmatrix} & & & & 1 \\ & & & 1 & \\ & & \iddots & & \\ & 1 & & & \\ 1 & & & & \end{bmatrix} \qquad \Im K_n \begin{bmatrix} & & & & 0 \\ & & & 0 & 1 \\ & & \iddots & & \\ & 0 & & & \\ 0 & 1 & & & \end{bmatrix}$$

$\widehat{\Re A} e_n^x = e_1^x$ while $\widehat{\Im A} e_n^x = e_2^x$. Again, this are linearly independent so we are done.

Type iv) Without loss of generality we may assume that $B = K_2 \oplus [e^{\mathbb{i}\theta}] K_2$ with $\pi < \theta < 2\pi$.

Then

$$\Re A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & \cos\theta \end{bmatrix} \qquad \Im A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sin\theta \end{bmatrix}$$

Then $q_{\Re B} = 2x_1x_2 + 2y_1y_2 + \cos\theta(x_3^2 + y_3^2)$ and $q_{\Im B} = x_2^2 + y_2^2 + \sin\theta(x_3^2 + y_3^2)$. Consider the point $\mathbf{z} = \left[\frac{\cos\theta}{2\sin\theta} : 1 : \sqrt{\frac{2}{-\sin\theta}} : \frac{\cos\theta}{2\sin\theta} : 1 : 0\right]$ which belongs to $\mathcal{P}_B$.

Then

$$\widehat{\Re B}\mathbf{z} = \left[1 : \frac{\cos\theta}{2\sin\theta} : \cos\theta\sqrt{\frac{2}{-\sin\theta}} : 1 : \frac{\cos\theta}{2\sin\theta} : 0\right] \text{ and } \widehat{\Im A}\mathbf{z} = \left[0 : 1 : \cos\theta\sqrt{\frac{2}{-\sin\theta}} : 0 : 1 : 0\right].$$

By checking the first two entries in each of these vectors we conclude that they are linearly independent and therefore $\mathbf{z}$ is a smooth point.

Type v) Without loss of generality we may assume that $B = K_2 \oplus e^{\mathbb{i}\theta} K_2$ with $0 < \theta < 2\pi$.

Then

$$\Re A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta \\ 0 & 0 & \cos\theta & -\sin\theta \end{bmatrix} \qquad \Im A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{bmatrix}$$

Then consider $\mathbf{z} = [1+\cos\theta : \sin\theta : -1-\cos\theta : \sin\theta : \sin\theta : 0 : \sin\theta : 0]$.

Excluding the case $\theta = \pi$, this is a nonzero vector that belongs in $\mathcal{P}_B$. We have

$$\widehat{\Re B}\mathbf{z} = [\sin\theta : 1+\cos\theta : \cos\theta\sin\theta : -1-\cos\theta : \sin\theta : 0 : \cos\theta\sin\theta : -\sin^2\theta]$$

$$\widehat{\Im B}\mathbf{z} = [0 : 1 : \sin\theta : -1 : 0 : 1 : \sin\theta : \cos\theta]$$

The minor of the Jacobian consisting of the first two columns is

$$\begin{vmatrix} \sin\theta & 1+\cos\theta \\ 0 & 1 \end{vmatrix} = \sin\theta$$

which is not zero given our assumptions on θ.

For the case $\theta = \pi$, consider $\mathbf{z} = e_2^x + e_4^x$. It can be checked that $\mathbf{z} \in \mathcal{P}_B$. On the other hand $\widehat{\Re B}\mathbf{z} = e_1^x - e_3^x$ and $\widehat{\Im B}\mathbf{z} = e_2^x - e_4^x$ which are linearly independent so we are done.

Type vi) Without loss of generality, we assume that $\theta_1 = 0$. Then,

$$\Re B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_2 & 0 \\ 0 & 0 & \cos\theta_3 \end{bmatrix} \quad \text{and} \quad \Im B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sin\theta_2 & 0 \\ 0 & 0 & \sin\theta_3 \end{bmatrix}$$

B is a normal matrix. By Theorem 4.3 we know what the fiber $w_B^{-1}(0)$ looks like. This fiber consists of points of the form

$$\mathbf{z} = (r_1\cos\phi_1, r_2\cos\phi_2, r_3\cos\phi_3, r_1\sin\phi_1, r_2\sin\phi_2, r_3\sin\phi_3), \quad \phi_i \in [0, 2\pi)$$

where $r_i \in \mathbb{R}_{\geq 0}$ and $r_1^2 + r_2^2 e^{\mathrm{i}\theta_2} + r_3^2 e^{\mathrm{i}\theta_3} = 0$. From the assumption that 0 is in the interior of $\mathcal{W}(B)$ we conclude that (1) $r_i \neq 0$ and (2) $\sin(\theta_2) \neq 0$.

Now consider the point $\mathbf{z} = [r_1 : r_2 : r_3 : 0 : 0 : 0] \in \mathcal{P}_B$. Then

$$\widehat{\Re B}\mathbf{z} = [r_1 : r_2\cos\theta_2 : r_3\cos\theta_3 : 0 : 0 : 0] \quad \text{and} \quad \widehat{\Im A}\mathbf{z} = [0 : r_2\sin\theta_2 : r_3\sin\theta_3 : 0 : 0 : 0]$$

By (1) and (2), it follows that these two vectors are linearly independent. $\square$

Chapter 5

THE RANK-K FIELD OF VALUES PROBLEM

As was discussed in the introduction, one of the initial motivations in studying higher-rank numerical ranges was their connection to quantum error correcting codes. Choi et. al, discussed ways of finding all projections P for which $PTP = \lambda P$ [8, Section 4]. However, they restrict themselves to the case where T is a Hermitian matrix. In this section, we give a way of computing one such projection for any type of matrix. Since this is similar in nature to the inverse field of values problem, we call this the rank-k inverse field of values problem (rank-k iFVP). The problem is as follows: Given a matrix $A \in \mathbb{C}^{n\times n}$ and $\mu \in \Lambda_k(A)$, find a matrix $X \in \mathbb{C}^{n\times k}$ such that $X^*X = I_k$ and $X^*AX = \mu I_k$. The projection matrix can be obtained from X by taking $P = XX^*$.

An equivalent way to formulate this problem is to find a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that U^*AU has μI_k as a principal submatrix. We call such a matrix U a solution for the rank-k iFVP of A at μ. If $\mu = 0$ we say that U is a solution for the rank-k iFVP of A. The proof of convexity of $\Lambda_k(A)$ by Li and Sze establishes the existence of such a matrix U [26]. In this section, we revisit their arguments with the aim of pointing out how to construct U explicitly.

5.1 The normal case

We first consider the case where the matrix A is normal. Without loss of generality, we may assume that A is diagonal and that $\mu = 0$. Up to permutation of the entries in the diagonal, we can write A in the following way:

$$A = 0_j \oplus \bigoplus_{i=1}^{p} B_{\theta_i} \oplus C, \tag{5.1}$$

where 0_j denotes the $j \times j$ zero matrix. B_{θ_i} denotes a 2-by-2 diagonal matrix such that its two entries have *opposite arguments* θ_i and $\theta_i + \pi$, i.e, it has the following form:

$$B_{\theta_i} = \begin{bmatrix} \alpha_i e^{\mathrm{i}\theta_i} & 0 \\ 0 & \beta_i e^{\mathrm{i}(\theta_i+\pi)} \end{bmatrix}, \quad \theta_i \in [0, 2\pi) \text{ and } \alpha_i, \beta_i > 0. \tag{5.2}$$

Finally, C is a diagonal matrix such that it is invertible and no pair of its eigenvalues have opposite arguments.

We proceed to show that we can solve the rank-k iFVP by solving a ranked iFVP for each of the blocks in Eq. (5.1) individually. Let M be a complex $n \times n$ matrix. By [26, Theorem 2.2], $0 \in \Lambda_t(M)$ if and only if the matrices

$$M(\xi) = \Re M \cos\xi + \Im M \sin\xi$$

have at least t nonnegative eigenvalues for all $\xi \in [0, 2\pi)$. We begin by checking that we can relax this condition so that it is satisfied by all but finitely many $\xi \in [0, 2\pi)$.

Lemma 5.1. *Suppose that for a complex $n \times n$ matrix M, $M(\xi)$ has at least t nonnegative eigenvalues for all but finitely many $\xi \in [0, 2\pi]$. Then the statement holds for all $\xi \in [0, 2\pi]$.*

Proof. The above statement is equivalent to $\lambda_t(M(\xi)) \geq 0$. The function $\lambda_t(M(\xi))$ gives a continuous map from $[0, 2\pi]$ to the real numbers. Therefore, the set of $\xi \in [0, 2\pi]$ such that $\lambda_t(M(\xi)) \geq 0$ must be a closed subset of $[0, 2\pi]$. □

Theorem 5.2. *Let A be as in Eq. (5.1). Suppose that $0 \in \Lambda_k(A)$. Then the following statements are true:*

1. $0 \in \Lambda_j(0_j)$
2. $0 \in \Lambda_1(B_{\theta_i})$ *for all* $1 \leq i \leq p$
3. *If* $k > p + j$ *then* $0 \in \Lambda_{k-j-p}(C)$

In particular, $U = I_j \oplus \bigoplus_{i=1}^{p} U_i \oplus U_C$ *is a solution for the rank-k iFVP of A, where* U_i *is the solution for the rank-1 iFVP of* B_{θ_i} *and* U_C *is a solution for the rank-*$(k-p-j)$ *iFVP of* C*.*

Proof. Claim (1) follows easily from the fact that $\Lambda_k(0) = \{0\}$ for all $1 \leq k \leq j$.

For claim (2) we observe that $\Lambda_1(B_{\theta_i}) = \mathcal{W}(B_{\theta_i}) = \text{conv}\{\alpha e^{\mathrm{i}\theta}, -\beta e^{\mathrm{i}\theta}\}$. Therefore, $\mathcal{W}(B_{\theta_i})$ is a line segment that contains the point 0.

For claim (3) we will show that, for any $\xi \in [0, 2\pi)$, $C(\xi)$ has at least $k-p-j$ nonnegative eigenvalues. By our assumptions, $A(\xi)$ has at least k nonnegative eigenvalues. The matrix $A(\xi)$ has the same block decomposition as A.

$$A(\xi) = 0_j \oplus \bigoplus_{i=1}^{p} B_{\theta_i}(\xi) \oplus C(\xi).$$

We make the following observations:

a) The 0_j matrix has j nonnegative (actually zero) eigenvalues.

b) If $\theta_i - \xi \notin \{-\pi, 0, \pi\}$ then $B_{\theta_i}(\xi)$ has exactly one nonnegative eigenvalue. This is because

$$B_{\theta_i}(\xi) = \begin{bmatrix} \alpha_i \cos(\theta_i - \xi) & 0 \\ 0 & -\beta_i \cos(\theta_i - \xi) \end{bmatrix}$$

and by our assumptions $\cos(\theta_i - \xi) \neq 0$.

By a) and b) we know that for all but finitely many $\xi \in [0, 2\pi)$, the first two types of blocks in the decomposition of $A(\xi)$ account for exactly $j + p$ nonnegative eigenvalues of $A(\xi)$. Therefore, $C(\xi)$ must have at least $k - j - p$ nonnegative eigenvalues. By Lemma 5.1 it follows that $0 \in \Lambda_{k-j-p}(C)$. □

We now proceed to solve the rank-k iFVP problem block by block. Notice that for the blocks of type B_{θ_i}, we only need to solve the classical iFVP. In Section 4.2 we give the set of all iFVP solutions for any normal matrix. For the sake of completeness, we give one of these solutions explicitly.

Lemma 5.3. *There exists a unitary matrix U_i such that the top leftmost entry of $U_i^* B_{\theta_i} U_i$ is zero.*

Proof. Consider the following orthogonal matrix

$$U_i = \frac{1}{\sqrt{\alpha_i + \beta_i}} \begin{pmatrix} \sqrt{\beta_i} & -\sqrt{\alpha_i} \\ \sqrt{\alpha_i} & \sqrt{\beta_i} \end{pmatrix}. \tag{5.3}$$

The top-leftmost entry of $U_i^* B_{\theta_i} U_i$ corresponds to $u^* B_{\theta_i} u$ where $u = (\sqrt{\beta_i}, \sqrt{\alpha_i})/\sqrt{\alpha_i + \beta_i}$. Therefore, $u^* B_{\theta_i} u = \frac{1}{\alpha_i + \beta_i}(\sqrt{\beta_i}, \sqrt{\alpha_i})^\top (e^{\mathrm{i}\theta_i}\alpha_i\sqrt{\beta_i}, -e^{\mathrm{i}\theta_i}\beta_i\sqrt{\alpha_i}) = 0$ ☐

We now proceed to solve the iFVP of C. In order to do this, we need the following lemma.

Lemma 5.4. *Suppose there exists an invertible matrix $S \in \mathbb{C}^{n\times n}$ such that S^*AS has 0_t as a leading principal submatrix. Then there exists a unitary matrix U such that 0_t is a leading principal submatrix of U^*AU.*

Proof. As it is shown in [26, Lemma 2.5], we can use QR-decomposition to find a unitary matrix U and an upper triangular matrix R such that $UR = S$. Since S is invertible so is R. In particular $U = SR^{-1}$ and

$$U^*AU = R^{-*}S^*ASR^{-1} = R^{-*} \begin{bmatrix} 0_t & * \\ * & * \end{bmatrix} R^{-1}.$$

The inverse of an upper triangular matrix is still upper triangular. Its conjugate on the other hand is lower triangular. Multiplying by R^{-1} on the right is equivalent to applying columns operations to S^*AS. Moreover, the i-th column is only affected by operations involving columns 1 through i. On the other hand R^{-*} is now lower triangular and multiplying by it on the left is equivalent to doing row operations. Again the i-th row is only affected by rows 1 through i. Both of these types of operations preserve the submatrix 0_t. ☐

To simplify notation, throughout the rest of the section we assume that $j = p = 0$. Let $k' := \max\{t \in [n] : 0 \in \Lambda_t(C)\}$. The next step in our construction is to find a line in $\mathbb{C}$ passing through the origin such that k of the eigenvalues are in one side and $n - k$ are in the other. However, this line only exists when $k = k'$. Luckily for us, $\Lambda_{k'}(C) \subseteq \Lambda_k(C)$. Furthermore, a solution for the rank-k' iFVP of C would also be a solution for the rank-k iFVP of C. Therefore, without loss of generality, we may assume that $k = k'$, i.e., $0 \notin \Lambda_{k+1}(C)$.

Lemma 5.5. *Suppose that* $0 \in \Lambda_k(C)\backslash\Lambda_{k+1}(C)$. *Then there exists* $\xi \in [0, 2\pi)$ *such that* k *eigenvalues of* C *are strictly on one side of the line* $\{a\cos\xi + b\sin\xi = 0\}$ *and* $n-k$ *are strictly in the other.*

Proof. By the membership test (Theorem 3.9), we know that for all $\xi \in [0, 2\pi)$, $\lambda_k(C(\xi)) \geq 0$. Moreover, by Lemma 5.1 there must be infinite ξ such that $\lambda_{k+1}(C(\xi)) < 0$. We claim that under our assumptions of C, there are only finitely many ξ such that $\lambda_k(C(\xi)) = 0$.

If $\lambda_k(C(\xi)) = 0$ then $C(\xi)$ would have to be singular. But such $C(\xi)$ would correspond to an intersection point of $\mathcal{V}(f_C)$, the Kippenhahn curve of C, with the circle $x^2 + y^2 = 1$. Since C is normal, $\mathcal{V}(f_C)$ is the union of n lines. By Bezout's theorem, there can only be finitely many of these intersection points. Therefore, we can select ξ such that

$$\lambda_k(C(\xi)) > 0 > \lambda_{k+1}(C(\xi)).$$

In particular, k of the eigenvalues of $C(\xi)$ are strictly positive and $n-k$ of the eigenvalues of $C(\xi)$ are strictly negative.

Let $\{\lambda_1, \ldots, \lambda_n\} \subseteq \mathbb{C}$ be the eigenvalues of C. Because C is normal, the eigenvalues of $C(\xi)$ are of the form $\operatorname{Re}\lambda_i \cos\xi + \operatorname{Im}\lambda_i \sin\xi$. Therefore, the condition above is equivalent to $\operatorname{Re}\lambda_i \cos\xi + \operatorname{Im}\lambda_i \sin\xi > 0$ for k of the eigenvalues λ_i of C and $\operatorname{Re}\lambda_i \cos\xi + \operatorname{Im}\lambda_i \sin\xi < 0$ for the rest of them. ☐

Such ξ can be found explicitly using the procedure described in Algorithm 1 taking C, $k+1$ and 0 as inputs. The algorithm returns False at one of the points $(1, s_i)$ used to test membership. These points are chosen so that $[1 : s_i] \notin \mathcal{V}(f_A)$. Finally, by normalizing the point, we can write it as $[r\cos\xi : r\sin\xi]$ with $r > 0$. Then ξ would be the value we are looking for.

We then obtain a matrix $C_{(1)} = e^{\mathrm{i}\xi}C$ with the property that the separating line in Lemma 5.5 is $b = 0$. Here we note that if U_C is a solution for the k-rank iFVP of $e^{\mathrm{i}\xi}C$ then it is also a solution for the k-rank iFVP of C. Therefore, we continue our discussion with $C_{(1)}$. Also, after an appropriate permutation of the entries in C we may assume that

$$C_{(1)} = \operatorname{diag}\{c_1 + \mathbb{i}d_1, \ldots, c_k + \mathbb{i}d_k, -c_{k+1} + \mathbb{i}d_{k+1}, \ldots - c_n + \mathbb{i}d_n\}$$

where $c_i > 0$ for all $i \in [n]$.

Then consider the matrix $T = \operatorname{diag}\{\sqrt{c_1}^{-1}, \ldots, \sqrt{c_n}^{-1}\}$. By conjugating $C_{(1)}$ by T and permuting the entries of C as necessary, we obtain a matrix with the following form.

$$C_{(2)} = T^* C_{(1)} T = \begin{pmatrix} I_k + \mathbb{i}D & 0 \\ 0 & -I_{n-k} + \mathbb{i}G \end{pmatrix}, \tag{5.4}$$

with $D = \operatorname{diag}(g_1, \ldots, g_k)$, $G = \operatorname{diag}(g_{k+1}, \ldots, g_n)$, $g_1 \geq g_2 \geq \ldots \geq g_k$ and $g_{k+1} \geq \ldots \geq g_n$. It turns out that under our assumption on C, we must have $k < n - k + 1$.

Lemma 5.6. *Let C be a normal invertible matrix such that no pair of eigenvalues have opposite arguments. Then if $0 \in \Lambda_k(C)$ the size of C must be at least $2k + 1$.*

Proof. Suppose there exists one eigenvalue λ of C such that no other eigenvalue has the opposite argument and additionally one of the open half-planes defined by the line that passes between 0 and λ has $k-1$ eigenvalues of C on it. Then if we consider a small perturbation on the angle of the line we would have a closed half-plane that contains only $k-1$-eigenvalues contradicting our hypothesis of C. Thus, there must be at least k eigenvalues on each side of the line, meaning that the size of C has to be at least $2k + 1$. □

In fact, if $2k \geq n$ and $0 \in \Lambda_k(C)$ then for every eigenvalue λ of C there must exist another eigenvalue μ with the opposite argument. Next, we introduce the notion of imbeddable matrices.

Definition 5.7. Let $A \in \mathrm{Herm}_m$ and $B \in \mathrm{Herm}_n$ with $m \geq n$. Then B is *imbeddable* in A if there exists a unitary $m \times m$ matrix V such that V^*AV has B as a leading principal submatrix.

The following theorem is very useful to establish if one matrix is imbeddable in another.

Theorem 5.8 (Theorem 1 in [12]). *Let* $\{a_1, \dots a_m\}$ *and* $b_1, \dots, b_n$ *be the eigenvalues of* A *and* B *respectively. Suppose additionally that* $a_1 \geq \dots \geq a_m$ *and* $b_1 \geq \dots \geq b_n$. *Then* B *is imbeddable in* A *if and only if the following inequalities hold:*

$$a_i \geq b_i \geq a_{i+m-n}, \textit{ for all } 1 \leq i \leq n$$

We refer to the sequences $\{a_i\}$ and $\{b_i\}$ that satisfy the inequalities in the previous theorem as having the *interlacing property.* In Section 5.2 we explore the interlacing property in more depth and give an explicit construction of V. For now, we use V to make further manipulations on $C_{(2)}$.

Proposition 5.9. *Suppose that* $C_{(2)}$ *is as described in Eq.* (5.4) *and* $0 \in \Lambda_k(C)$. *Then* D *is imbeddable in* $-G$.

Proof. The following inequalities are established in the proof of [26, Lemma 2.7]:

$$-g_n > g_1 > -g_{2k}, \quad -g_{n-1} > g_2 > -g_{2k-1}, \quad \dots \quad -g_{n-k+1} > g_k > -g_{k+1}.$$

The proposition then follows by Theorem 5.8. □

By the previous proposition, we conclude that there exists a unitary matrix V such that $-V^*GV$ has D as a leading principal submatrix. We are ready to transform $C_{(2)}$ even further.

$$C_{(3)} = (I_k \oplus V^*)C_{(2)}(I_k \oplus V) = \begin{bmatrix} I_k + \mathbb{i}D & 0 & 0 \\ 0 & -I_k - \mathbb{i}D & * \\ 0 & * & * \end{bmatrix} \tag{5.5}$$

A final step is necessary to transform C into the desired form. Consider the matrix

$$W = \frac{1}{\sqrt{2}} \begin{pmatrix} I_k & I_k \\ I_k & -I_k \end{pmatrix}.$$

Observe that W is a $2k \times 2k$ matrix that is both symmetric and unitary. Then take

$$C_{(4)} = (W \oplus I_{n-2k})C_{(3)}(W \oplus I_{n-2k}) = \begin{bmatrix} 0_k & I_k + \mathbb{i}D & 0_{n-2k} \\ I_k + \mathbb{i}D & 0_k & * \\ 0_{n-2k} & * & * \end{bmatrix}. \tag{5.6}$$

The matrix $C_{(4)}$ has the desired 0_k submatrix that we are looking for. We summarize how to obtain $C_{(4)}$ from C and the desired invertible matrix S in the following theorem.

Theorem 5.10. *There exists a unitary matrix U_c such that $U_c^* C U_c$ has 0_k as a leading principal submatrix.*

Proof. Let ξ be as described in Lemma 5.5 and consider the invertible matrix $S = T(I_k \oplus V)(W \oplus I_{n-2k})$ where T, V and W are as defined previously in the section. Then $S^*CS = e^{-\mathbb{i}\xi}C_{(4)}$, which has the desired leading submatrix 0_k. Using QR-decomposition, we can find a unitary matrix U_c and an upper triangular matrix R such that $S = U_cR$. Then by Lemma 5.4, U_c has the desired property. □

5.2 *Eigenvalue interlacing property*

In this section, we dive deeper into the work of Fan and Pall [12] in order to come up with a construction of a unitary matrix V that imbeds a Hermitian matrix B onto another Hermitian matrix A. We begin by checking some of the details of their proof of Theorem 5.8.

Proof of Theorem 5.8. The proof is done for the case $n = m - 1$. For the general case, note that it is possible to construct a series of interlacing sequences between $\{a_i\}$ and $\{b_i\}$ each with one fewer member. Then for each of these intermediate sequences the proof is applied.

The solution for $n = m-1$ is very constructive. It amounts to solving an m-by-m system of linear equations. To work out the system, we observe that if such a V existed then V^*AV would be unitarily equivalent to a matrix of the following form:

$$C = \begin{pmatrix} b_1 & & & & \bar{z}_1 \\ & b_2 & & & \bar{z}_2 \\ & & \ddots & & \vdots \\ & & & b_{m-1} & \bar{z}_{m-1} \\ z_1 & z_2 & \cdots & z_{m-1} & \gamma \end{pmatrix}, \tag{5.7}$$

with $z_i \in \mathbb{C}$ and $\gamma \in \mathbb{R}$.

We can assume that

$$a_1 < b_1 < a_2 < \ldots < a_{m-1} < b_{m-1} < a_m.$$

If this was not the case, say $b_i = a_i$, then we can take $z_i = 0$. We may then proceed by removing that corresponding row and column from the matrix C. This is equivalent to removing a_i and b_i from their sequences, which preserves the interlacing property. In the case where $b_i = a_{i+1}$ we would also set $z_i = 0$ and proceed by removing these elements from the set of interlacing eigenvalues.

From Eq. (5.7) we can deduce that the characteristic polynomial of C is equal to

$$\left(\lambda - \gamma - \sum_{j=1}^{m-1} \frac{|z_j|^2}{\lambda - b_j}\right) \prod_{i=1}^{m-1} (\lambda - b_i). \tag{5.8}$$

For C to be unitarily similar to A, the a_i should be roots of this polynomial. This happens if and only if

$$\gamma + \sum_{j=1}^{m-1} \frac{|z_j|^2}{a_i - b_j} = a_i, \text{ for all } 1 \leq i \leq m-1. \tag{5.9}$$

This gives a linear system of equations over the variables γ and $x_i = |z_i|^2$. In [12] it is shown that this system has a unique solution and that all of the values of x_i are positive. □

If we want to construct V explicitly, we could do so by choosing values for the intermediate interlacing sequences between $\{a_i\}$ and $\{b_i\}$. Then we must iteratively solve Eq. (5.9) and find $V_{(i)}$ that diagonalizes the matrix Eq. (5.7).

Example 5.11. To illustrate this, we show how to find V when $m = 4$ and $n = 2$. First of all we need to begin by finding values $a_1^{(2)}$, $a_2^{(2)}$ and $a_3^{(2)}$ such that the following inequalities are met:

$$\begin{array}{ccccccccccccc}
a_1 & & & & a_2 & & & & a_3 & & & & a_4 \\
& \rotatebox{-45}{$\geq$} & & \rotatebox{45}{$\geq$} & & \rotatebox{-45}{$\geq$} & & \rotatebox{45}{$\geq$} & & \rotatebox{-45}{$\geq$} & & \rotatebox{45}{$\geq$} & \\
& & a_1^{(2)} & & & & a_2^{(2)} & & & & a_3^{(2)} & & \\
& & & \rotatebox{-45}{$\geq$} & & \rotatebox{45}{$\geq$} & & \rotatebox{-45}{$\geq$} & & \rotatebox{45}{$\geq$} & & & \\
& & & & b_1 & & & & b_2 & & & &
\end{array} \tag{5.10}$$

We extend a_i to be defined by all $i \in \mathbb{Z}$ by setting $a_i = \infty$ if $i < 1$ and $a_i = -\infty$ if $i > m$. Similarly, $b_i = \infty$ if $i < 1$ and $b_i = -\infty$ if $i > n$. Then $a_i^{(2)}$ must satisfy the following inequalities:

$$\min\{a_i, b_{i-1}\} \geq a_i^{(2)} \geq \max\{b_i, a_{i+1}\} \text{ for all } i \in [3]. \tag{5.11}$$

The interlacing property between $\{a_i\}$ and $\{b_i\}$ guarantees $\min\{a_i, b_{i-1}\} \geq \max\{b_i, a_{i+1}\}$. It should also be clear by Eq. (5.10) that $\{a_i^{(2)}\}$ interlaces $\{a_i\}$ and $\{b_i\}$ interlaces $\{a_i^{(2)}\}$. Once the sequence $\{a_i^{(2)}\}$ is set we construct the following matrix:

$$C_1 = \begin{bmatrix} a_1^{(2)} & 0 & 0 & \bar{z}_1 \\ 0 & a_2^{(2)} & 0 & \bar{z}_2 \\ 0 & 0 & a_3^{(2)} & \bar{z}_3 \\ z_1 & z_2 & z_3 & \gamma \end{bmatrix}$$

We then proceed to find values for $|z_i|^2$ and γ by setting $z_i = 0$, whenever $a_i = a_i^{(2)}$ or $a_{+1} = a_i^{(2)}$. We then solve for the linear system of equations in Eq. (5.9) to obtain a value for γ and the rest of the $x_i = |z_i|^2$. The value $|z_i|^2$ is invariant under multiplication of z_i by any root of unity. We may choose $z_i = \sqrt{x_i}$. Once all the values in C_1 have been found, we proceed to find a unitary matrix V_1 that diagonalizes this matrix. We could do so by finding a unitary basis of eigenvectors of C_1 and setting these vectors as the columns of V_1. Then $V_1^* C_1 V_1 = \operatorname{diag}\{a_1, \ldots, a_4\} = A$ and $V_1 A V_1^*$ has $A_2 = \operatorname{diag}\{a_1^{(2)}, \ldots, a_3^{(2)}\}$ as a leading principal submatrix.

Repeating the same procedure as before, we can find a unitary 3×3 matrix V_2 such that $V_2 A_{(2)} V_2^*$ has $B = \mathrm{diag}\{b_1, b_2\}$ as a leading principal submatrix. Then the matrix $V = V_1^*(V_2^* \oplus [1])$ is the unitary matrix that we are looking for.

5.3 Solution of the general case

We now focus on the case where A is any complex $n \times n$ matrix. Suppose that $0 \in \Lambda_k(A)$. Recall that by [26, Lemma 2.7] it is enough to find an invertible matrix T such that T^*AT has 0_k as a principal submatrix. We turn our attention once again to the *-congruent canonical form of A.

Let S_A be an invertible matrix such that $B = S_A^* A S_A$ is in *-congruence canonical form. By Theorem 4.1, we know that

$$B = \bigoplus_{i=1}^{p} J_{k_i}(0) \oplus \bigoplus_{i=1}^{q} J_{2l_i}(\mu_j) \oplus \bigoplus_{i=1}^{r} e^{\mathrm{i}\xi_i} K_{2m_i} \oplus \bigoplus_{i=1}^{s} e^{\mathrm{i}\theta_i} K_{2n_i+1} \tag{5.12}$$

Following the arguments in [26, Lemma 2.8] for the blocks of type $e^{\mathrm{i}\theta} K_n$ we distinguish between those of even and odd size. Excluding the last type of blocks in the decomposition above, each of these blocks B_i has the property that $B_i(\xi)$ has a minimum number t of nonnegative eigenvalues for all $\xi \in [0, 2\pi)$ and that the block B_i has 0_t as a principal submatrix.

For the blocks of type $J_{k_i}(0)$, $t = \lfloor \frac{k_i+1}{2} \rfloor$ and we can obtain 0_t as a principal submatrix by taking the submatrix consisting of entries at odd rows and columns of $J_n(0)$. For $H_{2l_i}(\mu_i)$ and $e^{\mathrm{i}\xi_i} K_{2m_i}$, $t = m_i$ or $t = l_i$, respectively, and 0_t is the leading principal submatrix.

Finally, for the blocks of type $e^{\mathrm{i}\theta_i} K_{2n_i+1}$ it is argued that $e^{\mathrm{i}\theta} K_{2n_i+1}(\xi)$ is congruent to $D(\xi) \oplus [e^{\mathrm{i}\theta+\xi} + e^{-\mathrm{i}\theta-\xi}]$ where $D(\xi)$ contributes n_i nonnegative eigenvalues. Meanwhile, the leading principal submatrix of size $n_i + 1$ $e^{\mathrm{i}\theta_i} K_{2n_i+1}$ is $0_{n_i} \oplus [e^{\mathrm{i}\theta_i}]$. Therefore, we can find a permutation matrix P such that P^*BP has a principal leading submatrix $0_u \oplus N$, where $N = \mathrm{diag}(e^{i\theta_1}, \dots, e^{i\theta_s})$ and $u = \sum_{i=1}^{p} \lfloor \frac{k_i+1}{2} \rfloor + \sum_{i=1}^{q} l_i + \sum_{i=1}^{r} m_i + \sum_{i=1}^{s} n_i$

By a pigeonhole theorem argument, it is shown that $N(\xi) = e^{-\mathrm{i}\xi} N$ must have at least $k - u$ nonnegative eigenvalues for all $\xi \in [0, 2\pi)$. Since N is normal, we can use the procedure discussed in Section 5.1 to find an invertible matrix S_N such that $S_N^* N S_N$ has 0_{k-u} as a leading principal submatrix. Then the invertible matrix $T = S_A P(I_u \oplus S_N \oplus I_{n-s-u})$ satisfies that T^*AT has 0_k as a leading principal submatrix. The solution of the rank-k iFVP is the unitary matrix U obtained in the QR decomposition of T by Lemma 5.4.

Chapter 6

ADDITIONAL TOPICS ON HIGHER-RANK NUMERICAL RANGES

6.1 The emptiness of the higher rank numerical range is an open condition

In Corollary 3.6, we learned that $\Lambda_k(A)$ is empty if and only if $(-1,0,0)$ is contained in the dual cone $\widehat{\Lambda_k(A)}^*$. By Carathéodory's theorem [2, Theorem 2.3], for some $c>0$, $(-c,0,0)$ can be represented as a convex combination of three points $p_i := (-\lambda_k(\theta_i), \cos\theta_i, \sin\theta_i) \in O_k(\theta)$, $i \in [3]$. That is,

$$\mu_1 p_1 + \mu_2 p_2 + \mu_3 p_3 = (-c,0,0), \quad 0 \le \mu_i \le 1 \text{ and } \sum_{i=1}^{3} \mu_i = 1$$

We call such a convex combination a *certificate of emptiness* for $\Lambda_k(A)$. Then in particular, we have

$$\mu_1 \lambda_k(A(\theta_1)) + \mu_2 \lambda_k(A(\theta_2)) + \mu_3 \lambda_k(A(\theta_3)) = -c.$$

We use the following fact from matrix analysis.

Lemma 6.1 (Hoffman-Wiedland Theorem)**.** *For two Hermitian matrices A and B,*

$$\sum_{k=1}^{n} |\lambda_k(A) - \lambda_k(B)|^2 \le \|A-B\|_F^2.$$

In particular each $\lambda_k(\cdot)$ is Lipschitz continuous.

Proof. See [19, Theorem 1] □

Theorem 6.2. *Under the topology induced by the Frobenius norm, if $\Lambda_k(A) = \emptyset$ then there exists an open neighborhood U of A such that $\Lambda_k(B)$ is empty for all matrices B in U.*

Proof. Let θ_1, θ_2 and θ_3 be the angles corresponding to the points that appear on a certificate that $\Lambda_k(A)$ is empty. Consider the neighborhood of matrices B such that $\|B - A\|_F < \epsilon$ for some $\epsilon > 0$. First, let us observe that for any $\theta \in [0, 2\pi]$,

$$|\lambda_k(B(\theta)) - \lambda_k(A(\theta))| \leq \|B(\theta) - A(\theta)\|_F = \|(B - A)(\theta)\|_F \leq 2\|B - A\|_F.$$

Then we have that

$$\lambda_k(B(\theta)) \leq \lambda_k(A(\theta)) + |\lambda_k(B(\theta)) - \lambda_k(A(\theta))| < \lambda_k(A(\theta)) + 2\epsilon$$

Applying this inequality to the three terms in a certificate for the emptiness of $\Lambda_k(B)$ with the same angles, we conclude that

$$\mu_1\lambda_k(B(\theta_1)) + \mu_2\lambda_k(B(\theta_2)) + \mu_3\lambda_k(B(\theta_3)) = -c + 2\epsilon.$$

Given that ϵ is arbitrary, we can choose it such that the above sum remains negative. $\square$

6.2 *Non-emptiness of a generalization of higher rank numerical matrices*

During our discussion about higher-rank numerical ranges, the following question arose. Given $A \in \mathbb{C}^{n\times n}$ and $1 \leq k \leq n$, what conditions guarantee that $\Lambda_k(A)$ is nonempty? Li, Poon and Sze showed $\Lambda_k(A)$ is nonempty if $k < 1 + n/3$ [25].

We propose the following generalization. Suppose you take d Hermitian matrices $\{A_1, \ldots, A_d\}$ of size n and a unit vector u in $\mathbb{R}^d$. Then let $A(u) = \langle u, (A_i, \ldots, A_d)\rangle$. Next, following the characterization of higher-rank numerical range given in [26], for each u take the affine half-space

$$H_k(u) = \{x \in \mathbb{R}^d : \langle u, x\rangle \leq \lambda_k(A(u))\}$$

where $\lambda_k(A)$ is the k-th largest eigenvalue of the matrix A. Then we define the generalized higher-rank numerical range as

$$\Lambda_k(A_1, \ldots, A_d) = \bigcap_{u\in\mathbb{S}^{d-1}} H_k(u).$$

By the same proof as in [25], we can show the following.

Theorem 6.3. *For $k \leq (n+d)/(d+1) < n/(d+1)+1$, $\Lambda_k(A_1, \ldots, A_d)$ is nonempty.*

Proof. By Helly's theorem [2, Theorem 4.2], it suffices to show that for any $d+1$ points $\{u_1, \ldots u_{d+1}\} \subseteq \mathbb{S}^{d-1}$ the set $\bigcap_{i=1}^{d+1} H_k(u_i)$ is nonempty. For each u_i consider the following vector space.

$$V_k(u_i) = \operatorname{span}\{w_k(u_i), \ldots, w_n(u_i)\}$$

where $w_i(u)$ is an eigenvector corresponding to the i-th eigenvalue of $A(u)$. Observe that $\operatorname{codim} V_k(u_i) = k-1$. Then $\operatorname{codim} \bigcap_{i=1}^{d+1} V_k(u_i) \leq (d+1)(k-1)$. Our assumption guarantees $(d+1)(k-1) < n$. Observe that

$$\begin{aligned} (d+1)(k-1) &< n \\ k-1 &< n/(d+1) \\ k &< n/(d+1)+1 \end{aligned}$$

Therefore, there exists a unit vector $\omega \in \bigcap_{i=1}^{d+1} V_k(u_i)$. It turns out that for any $u \in \mathbb{S}^{d-1}$ and any $x \in \mathbb{S}^{d-1} \cap V_k(u)$ the vector $\mu = (x^* A_1 x, \ldots x^* A_d x)$ belongs to $H_k(u)$. By this fact, the vector $\mu = (\omega^* A_1 \omega, \ldots \omega^* A_d \omega)$ belongs to $\bigcap_{i=1}^{d+1} H_k(u_i)$.

To show this last fact, observe that $\langle u, \mu \rangle = x^* A(u) x$, by linearity. Next, given that $\{w_1(u), \ldots w_n(u)\}$ are the eigenvectors of A(u) we can write this product as

$$x^* A(u) x = \sum_{i=1}^{n} \lambda_i(u) x^* w_i(u) w_i(u)^* x = \sum_{i=1}^{n} \lambda_i(u) \langle x, w_i(u) \rangle^2$$

Because $x \in V_k(u)$ we know that $\langle x, w_i(u) \rangle = 0$ for the first $k-1$ terms. Therefore,

$$\sum_{i=1}^{n} \lambda_i(u) \langle x, w_i(u) \rangle^2 = \sum_{i=k}^{n} \lambda_i(u) \langle x, w_i(u) \rangle^2 \leq \lambda_k(u) \sum_{i=k}^{n} \langle x, w_i(u) \rangle^2$$

By the Pythagorean theorem, this last sum is equal to the norm square of x which is one. So in conclusion $\langle u, \mu \rangle \leq \lambda_k(A(u))$ which shows that $\mu \in H_k(u)$. □

6.3 Relation between the rank-k numerical ranges and Kippenhanh's polynomial

This was a question proposed by one of the reviewers for [28]. Consider the following theorem.

Theorem 6.4 (Theorem 1 in [13]). *Any two complex square matrices A, B have their rank-k numerical range equal to each other for $1 \leq k \leq \lfloor \frac{n}{2} \rfloor + 1$ if and only if the Kippenhahn polynomials f_A and f_B are the same.*

In [13], Gau and Wu showed that their bound on k for the above theorem cannot be relaxed further via counterexamples. They showed matrices A and B where $\Lambda_k(A) = \lambda_k(B)$ for $k < k'$ and $\Lambda_{k'}(A) \neq \Lambda_{k'}(B)$ for $k' = \lfloor \frac{n}{2} \rfloor + 1$. However, in all of their examples n is odd. It is not clear if the same holds when n is even. In fact, we believe the following conjecture to be true.

Conjecture 6.5. *Let $n = 2m$. Let A and $B \in \mathbb{C}^{n \times n}$ such that $\Lambda_k(A) = \lambda_k(B)$ for $1 \leq k \leq m$. Then $\Lambda_{m+1}(A) = \Lambda_{m+1}(B)$*

This would imply that Theorem 6.4 can be refined so that equality of $\Lambda_k(A)$ and $\Lambda_k(B)$ is sufficient for $1 \leq k \leq \lfloor \frac{n+1}{2} \rfloor$ to conclude f_A and f_B are the same. Let us suppose that there exist some matrices A and B such that $\Lambda_k(A) = \Lambda_k(B)$ for $1 \leq k \leq m$ and $\Lambda_k(A) \neq \Lambda_k(B)$. Now, $\Lambda_{m+1}(A)$ is either a singleton or the empty set. By Corollary 3.8, if $\Lambda_{m+1}(A)$ is equal to $\{\mu\}$ then so would be $\Lambda_m(A)$. So we can exclude the case where $\Lambda_{m+1}(A)$ and $\Lambda_{m+1}(B)$ are both nonempty. Therefore, if we were to disprove our conjecture, we would need to find two matrices A and B whose rank-k numerical ranges agree for $k \leq m$ and so that $\Lambda_m(A) = \Lambda_m(B) = \Lambda_{m+1}(A) = \{\mu\}$ and $\Lambda_{m+1}(B) = \emptyset$.

6.4 Dimension of the fiber for the rank-k inverse field of values problem

In Chapter 5 we showed a way to explicitly compute a solution to the rank-k iFVP. Following the discussion in Chapter 4 we could study the structure of the set of all solutions to this problem.

For now, we focus on the case where A is normal. If we follow the construction of the solution given in Section 5.1 we see that there is one degree of freedom in the choice of the half-plane that separates k eigenvalues of A from the other $n - k$ (see Lemma 2.1).

Some extra degrees of freedom are also found in the choices of intermediate interlacing sequences. For given n and k there are $(n-k)(n-k-1)/2-(k+1)k/2$ of these parameters as Example 5.11 suggests.

Moreover, each of the variables z_i in Eq. (5.9) contributes one degree of freedom because we can multiply this value by a root of unity and still get a solution. Considering all systems of equations that must be solved, there is a total of $(n-k)(n-k-1)/2-(k)(k-1)/2$.

This suggests that the dimension of this fiber should be $-2kn+k+n^2-n+1$. However, a more thorough analysis must be conducted to establish that this is indeed the case.

BIBLIOGRAPHY

[1] Hamid Reza Afshin, Mohammad Mehrjoofard, and Akbar Chavoshi. Pauli group: Classification and joint higher rank numerical range. *Linear Algebra and its Applications*, 549, 03 2018.

[2] Alexander Barvinok. *Course in Convexity.* Graduate studies in mathematics. American Mathematical Society, 2015.

[3] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry (Algorithms and Computation in Mathematics).* Springer-Verlag, Berlin, Heidelberg, 2006.

[4] Natália Bebiano, João da Providência, Ana Nata, and João P. da Providência. Revisiting the inverse field of values problem. *Electron. Trans. Numer. Anal.*, 42:1–12, 2014.

[5] Natália Bebiano, João da Providência, and Ilya M. Spitkovsky. On Kippenhahn curves and higher-rank numerical ranges of some matrices. *Linear Algebra and its Applications*, 629:246–257, 2021.

[6] Russell Carden. A simple algorithm for the inverse field of values problem. *Inverse Problems*, 25(11):115019, nov 2009.

[7] Man-Duen Choi, Michael Giesinger, John A. Holbrook, and David W. Kribs. Geometry of higher-rank numerical ranges. *Linear and Multilinear Algebra*, 56(1-2):53–64, 2008.

[8] Man-Duen Choi, David W. Kribs, and Karol Życzkowski. Higher-rank numerical ranges and compression problems. *Linear Algebra and its Applications*, 418(2):828–839, 2006.

[9] Man-Duen Choi, David W. Kribs, and Karol Życzkowski. Quantum error correcting codes from the compression formalism. *Reports on Mathematical Physics*, 58(1):77–91, 2006.

[10] David A. Cox, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3/e (Undergraduate Texts in Mathematics).* Springer-Verlag, Berlin, Heidelberg, 2007.

[11] Michel Crouzeix. Numerical range and functional calculus in Hilbert space. *Journal of Functional Analysis*, 244(2):668–690, 2007.

[12] Ky Fan and Gordon Pall. Imbedding conditions for hermitian and normal matrices. *Canadian Journal of Mathematics*, 9:298–304, 1957.

[13] Hwa-Long Gau and Pei Yuan Wu. Higher-rank numerical ranges and Kippenhahn polynomials. *Linear Algebra and its Applications*, 438(7):3054–3061, April 2013. arXiv:1208.1333 [math].

[14] Eugene Gutkin, Edmond A. Jonckheere, and Michael Karow. Convexity of the joint numerical range: topological and differential geometric viewpoints. *Linear Algebra and its Applications*, 376:143–171, 2004.

[15] Robin Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Mathematics*. Springer, 1977.

[16] Felix Hausdorff. Der wertvorrat einer bilinearform. *Mathematische Zeitschrift*, 3:314–316, 1919.

[17] J. William Helton and Victor Vinnikov. Linear matrix inequality representation of sets. *Communications on Pure and Applied Mathematics*, 60(5):654–674, 2007.

[18] Didier Henrion. Semidefinite geometry of the numerical range. *ELA. The Electronic Journal of Linear Algebra [electronic only]*, 20:322–332, 2010.

[19] Alan J. Hoffman and Helmut W. Wielandt. The variation of the spectrum of a normal matrix. *Duke Mathematical Journal*, 20(1):37 – 39, 1953.

[20] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge; New York, 1994.

[21] Roger A. Horn and Vladimir V. Sergeichuk. Canonical forms for complex matrix congruence and congruence. *Linear Algebra and its Applications*, 416(2):1010–1032, 2006.

[22] Roger A. Horn and Vladimir V. Sergeichuk. A regularization algorithm for matrices of bilinear and sesquilinear forms. *Linear Algebra and its Applications*, 412(2):380–395, 2006.

[23] Rudolf Kippenhahn. Über den wertevorrat einer matrix. *Mathematische Nachrichten*, 6(3-4):193–228, 1951.

[24] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry. Vol. II.* Interscience Tracts in Pure and Applied Mathematics, No. 15 Vol. II. Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney, 1969.

[25] Chi-Kwong Li, Yiu-Tung Poon, and Nung-Sing Sze. Condition for the higher rank numerical range to be non-empty. *Linear and Multilinear Algebra*, 57(4):365–368, 2009.

[26] Chi-Kwong Li and Nung-Sing Sze. Canonical forms, higher rank numerical ranges, totally isotropic subspaces, and matrix equations. *Proceedings of the American Mathematical Society*, 136(9):3013–3023, 2008.

[27] Murray Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2008.

[28] Jonathan Niño-Cortés and Cynthia Vinzant. The convex algebraic geometry of higher-rank numerical ranges. *Journal of Symbolic Computation*, page 102457, 2025.

[29] Daniel Plaumann, Rainer Sinn, and Stephan Weis. Kippenhahn's theorem for joint numerical ranges and quantum states. *SIAM Journal on Applied Algebra and Geometry*, 5(1):86–113, 2021.

[30] Daniel Plaumann and Cynthia Vinzant. Determinantal representations of hyperbolic plane curves: An elementary approach. *Journal of Symbolic Computation*, 57:48–60, 2013. Computational algebraic geometry.

[31] Kristian Ranestad and Bernd Sturmfels. On the convex hull of a space curve. *Advances in Geometry*, 12(1):157–178, 2012.

[32] R. Tyrrell Rockafellar. *Convex analysis.* Princeton Mathematical Series. Princeton University Press, Princeton, N. J., 1970.

[33] Joschka Roffe. Quantum error correction: an introductory guide. *Contemporary Physics*, 60(3):226–245, July 2019.

[34] Rainer Sinn. *Algebraic Boundaries of Convex Semi-Algebraic Sets.* PhD thesis, Universität Konstanz, Konstanz, 2014.

[35] Frank Sottile. Real algebraic geometry for geometric constraints. In *Handbook of Geometric Constraint Systems Principles*, Discrete Math. Appl. (Boca Raton), pages 273–285. CRC Press, Boca Raton, FL, 2019.

[36] Otto Toeplitz. Das algebraische analogon zu einem satze von fejér. *Mathematische Zeitschrift*, 2:187–197, 1918.

[37] Hugo J. Woerdeman. The higher rank numerical range is convex. *Linear and Multilinear Algebra*, 56(1-2):65–67, 2008.