Ritz Values of Normal Matrices and Ceva's Theorem

Russell Carden[☆], Derek J. Hansen,[☆☆]

Rice University, Department of Computational and Applied Mathematics,
6100 Main St. MS-134, Houston, TX 77005-1892

Abstract

The Cauchy interlacing theorem for Hermitian matrices provides an indispensable tool for understanding eigenvalue estimates and various numerical algorithms that rely on the Ritz values of a matrix. No generalization of interlacing is known for non-Hermitian matrices, and as a consequence, many useful algorithms for such matrices are not fully understood. Toward filling this gap, we consider the behavior of Ritz values of normal matrices. We apply Ceva's theorem, a classical geometric result, to understand two Ritz values of a 3×3 normal matrix and analyze the implications for larger matrices. Unlike the Hermitian case, specifying at most half of the Ritz values significantly restricts where the remaining Ritz values may fall. We use our results to analyze the restarted Arnoldi method with exact shifts applied to a 3×3 normal, non-Hermitian matrix.

Keywords: eigenvalues, Ritz values, Ceva's theorem, normal matrices, interlacing, restarted Arnoldi method
2000 MSC: 65F35,15A29,15A60

1. Introduction

For $A \in \mathbb{C}^{n\times n}$ and $V \in \mathbb{C}^{n\times k}$ with $k \leq n$ orthonormal columns, the eigenvalues of $V^*AV \in \mathbb{C}^{k\times k}$ are known as the *Ritz values* of the pair (A, V). Though they are well understood for Hermitian matrices, surprisingly little is known about the relative locations of Ritz values for general matrices, despite their widespread use in applications and computations. In this paper, we broaden this understanding through a geometric description of Ritz values of normal matrices that reveals a deeper structure and hints at the challenge of understanding Ritz values for nonnormal matrices.

[☆]Supported by National Science Foundation grant DMS-CAREER-0449973.
[☆☆]Supported by National Science Foundation grants DMS-0240058 (VIGRE), and DMS-0739420 (VIGRE).

Email addresses: `rlc2@rice.edu` (Russell Carden[☆]), `derek.j.hansen@gmail.com` (Derek J. Hansen,[☆☆])

When we use the term *eigenvalue* without qualification, we refer to the eigenvalues of A, denoted by $\lambda_1, \ldots, \lambda_n$. The *spectrum* of A, denoted $\sigma(A)$, is the set of all eigenvalues. Ritz values will typically be denoted by $\theta_1, \ldots, \theta_k$. The *field of values* of A, also known as the *numerical range*, is the set

$$W(A) := \{x^* A x : x \in \mathbb{C}^n, \|x\| = 1\},$$

where $\|x\| = \sqrt{x^* x}$ is the Euclidean norm. For each fixed $k < n$, the set of all Ritz values, taken over all $V \in \mathbb{C}^{n \times k}$ with orthonormal columns, is simply $W(A)$. Thus, every Ritz value is a *Rayleigh quotient*, i.e., a number of the form $v^* A v / v^* v$, and every Rayleigh quotient is a Ritz value.

Given a particular $\theta \in W(A)$, the problem of finding a $v \in \mathbb{C}^n$ such that $\theta = v^* A v / v^* v$ is known as the *inverse field of values problem* (iFOV) [1]. A solution can be constructed using Johnson's algorithm for drawing the boundary of the field of values [2] and an explicit solution for 2×2 matrices [3, 4]. The inverse field of values problem with k Ritz values (iFOV-k) is more general, but less understood: Given k values $\theta_1, \ldots, \theta_k$ in $W(A)$, does there exist a $V \in \mathbb{C}^{n \times k}$ with orthonormal columns for which $\sigma(V^* A V) = \{\theta_1, \ldots, \theta_k\}$? Understanding this problem would give further insight into the behavior of Ritz values and the role they play, for example, in the convergence of iterative schemes for solving large-scale eigenvalue problems and linear systems.

If A is Hermitian, i.e., $A = A^*$, then all the eigenvalues are real. For V having orthonormal columns, $V^* A V \in \mathbb{C}^{k \times k}$ is also Hermitian, so the Ritz values are also real. Labeled in increasing order, $\lambda_1 \le \cdots \le \lambda_n$ and $\theta_1 \le \cdots \le \theta_k$, these values must obey the Cauchy interlacing theorem (see, e.g., [5, Theorem 10.1.1]):

$$\lambda_i \le \theta_i \le \lambda_{n-k+i}, \quad i = 1, \ldots, k. \tag{1}$$

Moreover, this interlacing property is sharp in the following sense: given A and k real values θ_i interlacing the eigenvalues as in (1), there exists a V with orthonormal columns for which the θ_i are the associated Ritz values [5, p. 205].

If A is *normal*, i.e., $AA^* = A^* A$, then it is unitarily diagonalizable: $A = U \Lambda U^*$ for some unitary U and diagonal Λ. Consequently, when A is normal, $W(A)$ is the convex hull of the spectrum. In this case, it would be too much to hope for a characterization of the possible locations of the Ritz values as simple as the Cauchy interlacing theorem. One complicating factor is that, whereas $V^* A V$ must be Hermitian if A is Hermitian, $V^* A V$ need not be normal when A is normal. However, if $V^* A V$ does happen to be normal, some results are known. It is a trivial consequence of the Cauchy interlacing theorem that the real and imaginary parts of the Ritz values must interlace the eigenvalues of the Hermitian and skew-Hermitian parts of A, respectively. More interesting is the classic result of Fan and Pall [6] that addresses the case of $k = n - 1$.

Theorem 1. *Let $A \in \mathbb{C}^{n \times n}$ be normal and $V \in \mathbb{C}^{n \times (n-1)}$ have orthonormal columns, and suppose $V^* A V \in \mathbb{C}^{(n-1) \times (n-1)}$ is normal, and no eigenvalue of $V^* A V$ coincides with an eigenvalue of A. Then the eigenvalues and Ritz values are collinear, with the Ritz values interlacing the eigenvalues.*

This theorem is sharp in the trivial case, where A is a shifted and scaled Hermitian matrix.

For $k < n-1$, while no sharp geometric characterization of the Ritz values has been found, Carlson and Marques de Sá [7] have proved the following.

Theorem 2. *Let $A \in \mathbb{C}^{n\times n}$ be normal with $0 \notin W(A)$ (so that the field of values is entirely contained in a closed half-plane). Order the eigenvalues by increasing argument. Let $V \in \mathbb{C}^{n\times k}$ have orthonormal columns, and suppose $V^*AV \in \mathbb{C}^{k\times k}$ is normal. If the Ritz values are ordered by increasing argument, then their arguments interlace those of the eigenvalues:*

$$\arg \lambda_i \le \arg \theta_i \le \arg \lambda_{n-k+i}, \quad i = 1, \dots, k.$$

Queiró and Duarte [8] later provided a similar result.

For A normal and V^*AV not necessarily normal, Thompson (1966) and Malamud (2005) independently determined the criteria for a set of complex numbers to be Ritz values of a normal matrix [9, 10]. Their criteria take the form of intricate algebraic restrictions (see equation (5) and (7) here) and hence lack a natural geometric interpretation, in the sense that one may delineate regions of $W(A)$ where sets of Ritz values must lie.

We show that Ritz values of normal matrices satisfy much stronger constraints than interlacing. For a normal non-Hermitian matrix with $n = 3$ and $k = 2$, specifying one of the Ritz values to lie in the interior of $W(A)$ uniquely determines the other, as a consequence of Ceva's theorem. We use this result to study the restarted Arnoldi algorithm for a particular 3×3 normal matrix, and explore similar results for $n > 3$, and establish a connection to recent results of Bujanović for Ritz values from Krylov subspaces [11].

2. Algebraic criteria for the solvability of iFOV-$(n-1)$ for normal matrices

In this section we recall the relevant result of Malamud [9] and Thompson [10]. Let $A \in \mathbb{C}^{n\times n}$ be normal with $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$. Our goal is to determine those sets $\{\theta_1, \dots, \theta_{n-1}\} \subset W(A)$ for which it is possible to construct $V \in \mathbb{C}^{n\times(n-1)}$ with orthonormal columns such that V^*AV has eigenvalues $\theta_1, \dots, \theta_{n-1}$. If such a V exists, we say that iFOV-$(n-1)$ has a solution for $\{\theta_1, \dots, \theta_{n-1}\}$. By unitarily diagonalizing A, it is clear that given $\{\theta_1, \dots, \theta_{n-1}\}$, iFOV-$(n-1)$ is solvable for A if and only if it is solvable for $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. Through the remainder of this section, we will therefore assume A is diagonal.

If such a V exists, then any matrix $\widehat{V}$ with orthonormal columns and the same column space will work just as well. Since this column space, denoted $\mathcal{R}(V)$, has dimension $n-1$, it is determined by any single vector $v \in \mathbb{C}^n$ orthogonal to it. For this reason, the case of $k = n-1$ is more amenable to study than the general case.

Recall the *adjugate* (sometimes called the *classical adjoint*) of a square matrix:

$$[\mathrm{adj}(M)]_{ij} :=(-1)^{i+j} \det M(j,i),$$

where $M(j,i)$ is the matrix formed by deleting row j and column i from the matrix. If M is invertible, then

$$\mathrm{adj}(M) = (\det M)M^{-1}.$$

For unitary U, the adjugate satisfies

$$\mathrm{adj}(U^*MU) = U^* \,\mathrm{adj}(M)U.$$

Given a $V \in \mathbb{C}^{n\times(n-1)}$ with orthonormal columns and a unit vector $v \perp \mathcal{R}(V)$, we may form the unitary matrix $U = [V \ v]$ and find

$$\det(V^*MV) = [\mathrm{adj}(U^*MU)]_{nn} = [U^* \,\mathrm{adj}(M)U]_{nn} = v^* \,\mathrm{adj}(M)v.$$

Hence the characteristic polynomial p of $V^*\Lambda V$, whose roots are the Ritz values, satisfies

$$\begin{aligned} p(\lambda) &= \prod_{k=1}^{n-1} (\lambda - \theta_k) && (2) \\ &= v^* \,\mathrm{adj}(\lambda I - \Lambda)v = \sum_{j=1}^{n} |v_j|^2 \prod_{\substack{=1\\ k\neq j}}^{n} (\lambda - \lambda_k). && (3) \end{aligned}$$

Suppose the eigenvalues are distinct. Evaluating p at one of the eigenvalues λ_ℓ gives that

$$\begin{aligned} p(\lambda_\ell) = \prod_{k=1}^{n-1} (\lambda_\ell - \theta_k) &= \sum_{j=1}^{n} |v_j|^2 \prod_{\substack{k=1\\ k\neq j}}^{n} (\lambda_\ell - \lambda_k) \\ &= |v_\ell|^2 \prod_{\substack{k=1\\ k\neq \ell}}^{n} (\lambda_\ell - \lambda_k). && (4) \end{aligned}$$

Thus if we define x_ℓ such that

$$x_\ell := \frac{\prod_{k=1}^{n-1} (\lambda_\ell - \theta_k)}{\prod_{k\neq j} (\lambda_\ell - \lambda_k)} \tag{5}$$

for $\ell = 1, \ldots, n$, then for any $V \in \mathbb{C}^{n\times(n-1)}$ having orthonormal columns and corresponding unit vector $v \perp \mathcal{R}(V)$, the corresponding Ritz values of (Λ, V) must be such that the x_ℓ are nonnegative as $x_\ell = |v_\ell|^2$. Conversely, for any choice of $\theta_1, \ldots, \theta_{n-1}$ that makes all the x_ℓ nonnegative, there exists a $V \in$

$\mathbb{C}^{n\times(n-1)}$ with orthonormal columns such that $\{\theta_1,\ldots,\theta_{n-1}\}$ are the Ritz values of (Λ, V). This can be seen by first writing the polynomial $\prod_{k=1}^{n-1}(\lambda-\theta_k)$ as a Lagrange-like interpolant at the eigenvalues, and then noting that the x_j are the unique coefficients:

$$\prod_{k=1}^{n-1}(\lambda-\theta_k)=\sum_{j=1}^{n}x_j\prod_{\substack{k=1\\k\neq j}}^{n}(\lambda-\lambda_k). \tag{6}$$

If all the x_ℓ are nonnegative, then one may then choose any unit vector v for which $|v_\ell|^2=x_\ell$, and subsequently choose any $V\in\mathbb{C}^{n\times(n-1)}$ having orthonormal columns such that $\mathcal{R}(V)$ is orthogonal to v. Note that we always have $\sum x_j=1$: since the left hand side of (6) has leading term λ^{n-1}, so too must the right hand side.

The above argument can be adapted to the general case when the eigenvalues are not necessarily distinct. Suppose $\mu_1,\ldots,\mu_h$ are the distinct eigenvalues of A, with μ_j having multiplicity m_j. Then a necessary condition for $\theta_1,\ldots,\theta_{n-1}$ to be Ritz values is that for each j, $1\leq j\leq h$, at least m_j-1 of the Ritz values must equal μ_j. Order the Ritz values $\theta_1,\ldots,\theta_{n-1}$ so that those that do not coincide with eigenvalues (of which there are at most $h-1$) come first. In order for there to exist a $V\in\mathbb{C}^{n\times(n-1)}$ with orthonormal columns such that the θ_j are the Ritz values of (A,V), it is necessary and sufficient that each of the h rational expressions

$$y_j=\frac{\prod\limits_{k=1}^{h-1}(\mu_j-\theta_k)}{\prod\limits_{k\neq j}(\mu_j-\mu_k)},\quad j=1,\ldots,h, \tag{7}$$

be nonnegative. If each y_j is nonnegative, one may chose any $v\in\mathbb{C}^n$ for which

$$\sum_{i=1}^{m_j}|v_{j,i}|^2=y_j,\quad j=1,\ldots,h,$$

where the components of v are labeled and ordered in the natural way. Using less elementary means, Malamud arrived previously at an equivalent result [9, Proposition 3.1].

Note that for Ritz values from a $n-1$ dimensional subspace, (5) implies that if one of the $\theta_1,\ldots,\theta_{n-1}$ coincides with an eigenvalue, then there must be an eigenvector for that eigenvalue in $\mathcal{R}(V)$, regardless of whether the eigenvalue is a *sharp point* of $W(A)$. A sharp point is a corner on the boundary of $W(A)$, as opposed to a degenerate vertex that is a convex combination of two other eigenvalues. For example, in the Hermitian case, only the smallest and largest eigenvalues would be sharp points of $W(A)$. In the generic normal case, only eigenvalues at corners of the boundary of $W(A)$ are sharp. So if a Ritz value coincides with such an eigenvalue, then an eigenvector for that eigenvalue must be in $\mathcal{R}(V)$, as any vector whose Rayleigh quotient equals a sharp point is an eigenvector [12, Theorem 1.6.3]. An eigenvalue that does not lie at a sharp point, has numerous such generating vectors, most of which are not eigenvectors. Thus, it may at first seem surprising that in the normal case, having a Ritz value coincide with an eigenvalue requires that a corresponding eigenvector be in $\mathcal{R}(V)$. However, if one notes that the adjugate of $zI - A$, when evaluated at an eigenvalue λ_i, has rank equal to the geometric multiplicity of λ_i and range equal to the eigenspace corresponding to λ_i, then the result follows immediately. In fact, by the same reasoning this result holds in general, i.e., if $A \in \mathbb{C}^{n\times n}$ and $V \in \mathbb{C}^{n\times(n-1)}$ has orthonormal columns such that an eigenvalue λ of A is also an eigenvalue of V^*AV, then $\mathcal{R}(V)$ contains an eigenvector of A for λ. For more details, see [13].

Malamud and Thomspon's results for Ritz values of normal matrices provide (even for Hermitian matrices) an algebraic means of characterizing Ritz values. For $n = 3$, we provide a geometric interpretation of their results.

3. Geometric solution of iFOV-2 for 3×3 normal matrices: Ceva's theorem

We can use the results of the previous section to provide geometric criteria for the solvability of iFOV-2 for 3×3 normal matrices. In the case of a generic Hermitian 3×3 matrix, we know that given any θ_1 between the minimum and maximum eigenvalues, there are infinitely many choices for θ_2 for which $\{\theta_1, \theta_2\}$ are the Ritz values for some 3×2 matrix V with orthonormal columns; $\{\theta_1, \theta_2\}$ need only interlace the eigenvalues. However, as we shall see, it is generally the case that for a normal 3×3 matrix, given $\theta_1 \in W(A)$, there is *only one* θ_2 for which $\{\theta_1, \theta_2\}$ can be the Ritz values.

Assume $A \in \mathbb{C}^{3\times 3}$ is normal with distinct and noncollinear eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\}$, and assume θ_1 and θ_2 lie in the interior of $W(A)$. In this case, $W(A)$ is a triangle in the complex plane with vertices λ_1, λ_2, and λ_3, having θ_1 and θ_2 in its interior. The x_j from (5) are real and nonnegative if and only if $\arg(x_j) = 0$. Let $\angle(z_1, z_2, z_3) \in (-\pi, \pi]$ denote the signed angle from z_1 to z_3 about the vertex z_2. Observe that

$$\begin{aligned}
0 = \arg(x_1) &= \arg\left(\frac{(\lambda_1 - \theta_1)(\lambda_1 - \theta_2)}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)}\right) \\
&= \arg(\theta_1 - \lambda_1) - \arg(\lambda_2 - \lambda_1) + \arg(\theta_2 - \lambda_1) - \arg(\lambda_3 - \lambda_1) \\
&= \angle(\lambda_2, \lambda_1, \theta_1) + \angle(\lambda_3, \lambda_1, \theta_2), \qquad (8)
\end{aligned}$$

and similarly for x_2 and x_3. In classical geometry, a *Cevian* is a line segment joining the vertex of a triangle with a point on the opposite side [14]. We deduce from (8) that $\arg(x_j) = 0$ if and only if θ_1 and θ_2 each lie on separate Cevians, with each Cevian a reflection of the other about the angle bisector through the vertex at λ_j of the triangle formed by λ_1, λ_2 and λ_3. That is, θ_1 and θ_2 must lie on what are called *isogonal* Cevians (see Figure 1). Figure 2 gives one coarse consequence of this observation.

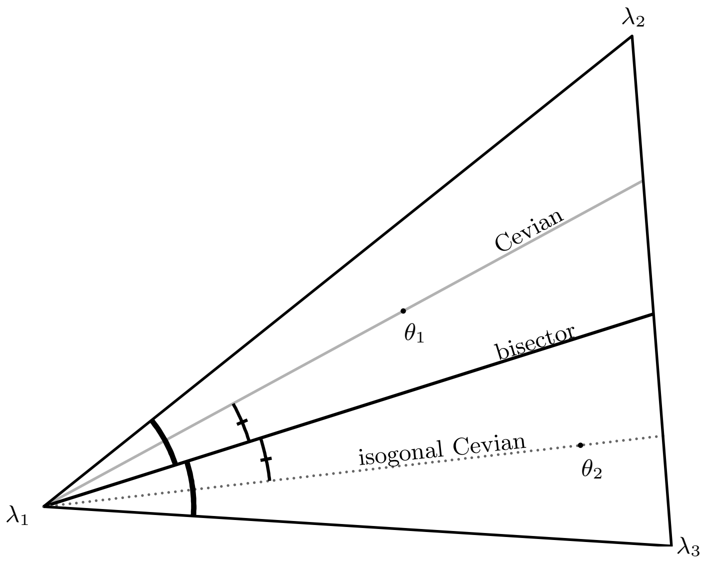

Figure 1: The Cevian from λ_1 through θ_2 is isogonal to the Cevian from λ_1 through θ_1.

We now show that given any $\theta_1 \in W(A) \setminus \sigma(A)$, there exists one and only one other choice of $\theta_2 \in W(A)$ such that this geometric relationship is satisfied for all three angle bisectors. To this end, regard θ_1 as a *Cevian point*, that is, as the point of intersection of three Cevians, each emanating from its own vertex. Reflect each of these Cevians across the associated angle bisector to get the isogonal Cevians. According to Ceva's theorem [14], these three isogonal Cevians intersect at θ_2 if and only if

$$\frac{\sin \alpha_1}{\sin \alpha_2} \frac{\sin \beta_1}{\sin \beta_2} \frac{\sin \gamma_1}{\sin \gamma_2} = 1,$$

where the angles are as labeled in Figure 3. But this equality must hold, since

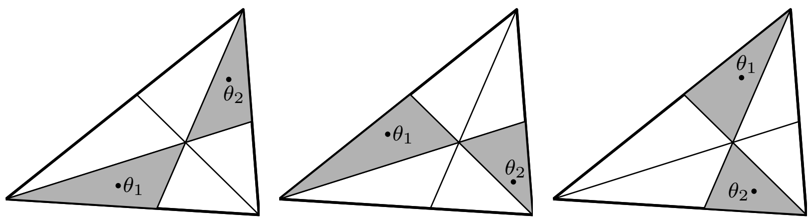

Figure 2: The regions in which θ_1 and θ_2 can lie are determined by the angular bisectors. On the left, if θ_1 lies in the lower shaded region, then θ_2 must lie in the upper shaded region. Similarly, for the other plots, if θ_1 lies in one of the shaded regions, then θ_2 must lie in the other. Given θ_1, the choice for θ_2 is unique; see Theorem 3.

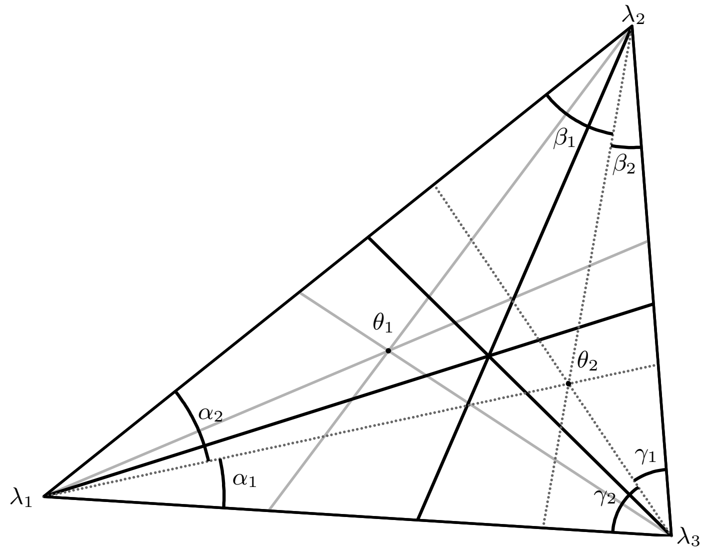

Figure 3: The Cevians through θ_2 are the isogonals of the Cevians through θ_1. The point θ_2 is the isogonal conjugate of θ_1. It is known from classical geometry that the angle bisectors (solid black lines) are concurrent at the *incenter* of the triangle, i.e., the center of the inscribed circle.

the original three Cevians are concurrent (at θ_1). The point θ_2 is known as the *isogonal conjugate* of the Cevian point θ_1 [14]. These observations are summarized in the following theorem.

Theorem 3. *Let $A \in \mathbb{C}^{3\times 3}$ be normal with spectrum $\sigma(A) = \{\lambda_1, \lambda_2, \lambda_3\}$, assume the eigenvalues of A are not collinear, and let $\{\theta_1, \theta_2\} = \sigma(V^*AV)$ for $V \in \mathbb{C}^{3\times 2}$ with orthonormal columns. If $\theta_1 \in W(A) \setminus \sigma(A)$, then θ_2 must be the isogonal conjugate of θ_1 in the triangle formed by λ_1, λ_2, and λ_3.*

This theorem implies that any one Ritz value $\theta_1 \in W(A) \setminus \sigma(A)$ *uniquely* determines the second Ritz value θ_2. If $\theta_1 \notin \sigma(A)$ is on an edge of the boundary of $W(A)$, then θ_2 must be the eigenvalue (vertex) opposite the edge. Moreover, the converse holds: if $\theta_1 \in \sigma(A)$, then θ_2 can be *any* point on the opposite side. A similar loss of uniqueness occurs when $W(A)$ is a degenerate triangle, that is, when A is a shifted and scaled Hermitian matrix. Suppose λ_1, λ_2 and λ_3 are collinear with, say, λ_2 lying between λ_1 and λ_3. The condition that

$$\angle(\lambda_1, \lambda_2, \theta_1) + \angle(\lambda_3, \lambda_2, \theta_2) = 0 \pmod{2\pi}$$

is equivalent to the statement that θ_1 and θ_2 interlace the eigenvalues of A. Though perhaps at first surprising, this nonuniqueness of the choice of θ_2 (given θ_1) when $W(A)$ is degenerate is consistent with Theorem 3 as a limiting case; see Figure 4. It exactly recovers the Cauchy interlacing theorem.

In summary, the steps for constructing a valid Ritz pair are as follows:

- Specify any $\theta_1 \in W(A)$.

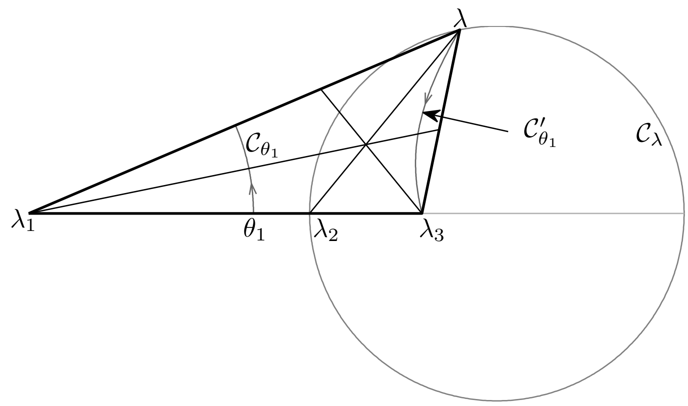

Figure 4: Suppose λ_1, λ_2 and λ_3 are collinear. Let $x = |\lambda_2 - \lambda_1|$, $y = |\lambda_3 - \lambda_2|$, and assume $x > y$. The locus of points λ for which the angle bisector of $\angle(\lambda_1, \lambda, \lambda_3)$ meets λ_2 is the circle of radius $r = xy/(x - y)$ that passes through λ_2 and is centered on the ray from λ_2 through λ_3. Let $\mathcal{C}_\lambda$ denote this circle. Extend a circular arc with center λ_1 from θ_1 to the segment connecting λ_1 and λ. Let $\mathcal{C}_{\theta_1}$ denote this arc. Each point on $\mathcal{C}_{\theta_1}$ is mapped to its isogonal conjugate on the curve denoted by $\mathcal{C}'_{\theta_1}$ connecting λ and λ_3; the arrows indicate the direction of the mapping. As $\lambda \to \lambda_2$ along $\mathcal{C}_\lambda$, the arc $\mathcal{C}_{\theta_1}$ shrinks to the point θ_1, while the curve $\mathcal{C}'_{\theta_1}$ tends to the closed line segment connecting λ_2 and λ_3. It is therefore clear that given any θ_2 on this closed line segment, one may define a function θ from $\mathcal{C}_\lambda$ to $\mathcal{C}_{\theta_1}$ such that as $\lambda \to \lambda_2$, $\theta(\lambda) \to \theta_1$, while the isogonal conjugate of $\theta(\lambda)$ approaches θ_2.

- If θ_1 is on the boundary of $W(A)$:
  - if it coincides with a vertex of the triangle, i.e., $\theta_1 \in \sigma(A)$, then θ_2 can be any point on the edge opposite the vertex (including the two other eigenvalues);
  - if it lies on an open edge, i.e., $\theta_1 \in \partial W(A) \setminus \sigma(A)$ where $\partial W(A)$ denotes the boundary of $W(A)$, then θ_2 must coincide with the vertex opposite the edge.
- Otherwise, θ_1 lies in the interior of $W(A)$:
  - Draw the three Cevians through θ_1.
  - Draw the angle bisectors emanating from each vertex.
  - Reflect the Cevians across the bisectors to get the isogonal Cevians.
  - These three isogonal Cevians intersect at the isogonal conjugate of θ_1. This point is the unique location of θ_2.
- Construct a v whose entries satisfy $|v_\ell|^2 = x_\ell$, for x_ℓ as in (5). Any V with orthonormal columns whose range is orthogonal to v will generate θ_1 and θ_2: $\sigma(V^*AV) = \{\theta_1, \theta_2\}$.

4. Geometric approach for iFOV-$(n-1)$ when $n > 3$

Building on the results of the previous section, we will discuss some numerical approaches to studying $n-1$ Ritz values for a $n > 3$ dimensional normal matrix. The key feature of such Ritz values is evident from (3). The possible $n-1$ Ritz values can be thought of as $2n-2$ real numbers, the real and imaginary parts of the θ_i. From (3), the possible $n-1$ Ritz values of a normal matrix can be parametrized using $n-1$ real numbers: the magnitudes of the entries of the unit vector v orthogonal to the subspace that generates the Ritz values. Note that, as v is a unit vector, only $n-1$ real numbers are required to specify the magnitudes of its entries. For $n = 3$ and noncollinear eigenvalues, we showed that this implies that one Ritz value often determines the other uniquely. For $n > 3$, one might then expect, roughly speaking, that specifying half of the Ritz values should determine the other half. In this spirit, we will provide constructions to determine where the remaining Ritz values can lie, provided we have specified no more than half of them.

From the properties of the field of values, for any $\theta \in W(A)$ one can construct a unit vector v such that $v^*Av = \theta$. For a normal matrix, assuming A is diagonal, such a v must satisfy, for $x_j = |v_j|^2$,

$$\begin{bmatrix} \operatorname{Re}\lambda_1 & \cdots & \operatorname{Re}\lambda_n \\ \operatorname{Im}\lambda_1 & \cdots & \operatorname{Im}\lambda_n \\ 1 & \cdots & 1 \end{bmatrix} x = \begin{bmatrix} \operatorname{Re}\theta \\ \operatorname{Im}\theta \\ 1 \end{bmatrix}. \tag{9}$$

Since the matrix in (9) has dimension $3 \times n$, the set of all possible x that satisfy (9) is an affine subspace of $\mathbb{R}^n$ of dimension at least $n-3$. If the eigenvalues are collinear, then $\operatorname{Im}\lambda_i = a\operatorname{Re}\lambda_i + b$ for some $a, b \in \mathbb{R}$, in which case the rows of the matrix in (9) are linearly dependent, and the subspace for x is $n-2$ dimensional. If A is a multiple of the identity, then all the rows of the matrix are linearly dependent, and the subspace for x is $n-1$ dimensional. The intersection of this subspace with the positive orthant determines all v such that $v^*Av = \theta$.

Suppose we wish to know where the eigenvalues of V^*AV can be located for all $V \in \mathbb{C}^{n\times(n-1)}$ such that $\theta \in \sigma(V^*AV)$. To do so, we pick a v from the vectors that satisfy (9). We would like to determine a $\widehat{V}$ such that $V = \begin{bmatrix} v \ \widehat{V} \end{bmatrix} \in \mathbb{C}^{n\times(n-1)}$ has orthonormal columns and

$$V^*AV = \begin{bmatrix} v^* \\ \widehat{V}^* \end{bmatrix} A \begin{bmatrix} v & \widehat{V} \end{bmatrix} = \begin{bmatrix} \theta & v^*A\widehat{V} \\ 0 & \widehat{V}^*A\widehat{V} \end{bmatrix}, \tag{10}$$

so θ is a Ritz value of (A, V). Equation (10) gives that $V^*(A-\theta I)v = 0$. As the columns of V are orthonormal, this implies that $\widehat{V}^*Av = 0$, i.e., $(Av)^*\widehat{V} = 0$. So $\mathcal{R}(\widehat{V})$ must be orthogonal to both v and Av. Since $\dim(\mathcal{R}(\widehat{V})) = n-2$, in general $\mathcal{R}(\widehat{V})$ must be the kernel of $[v \ Av]^*$. The remaining Ritz values of V^*AV are the eigenvalues of $\widehat{V}^*A\widehat{V}$. Thus, the remaining Ritz values are entirely dependent upon the particular v that satisfied equation (9). To determine where the Ritz

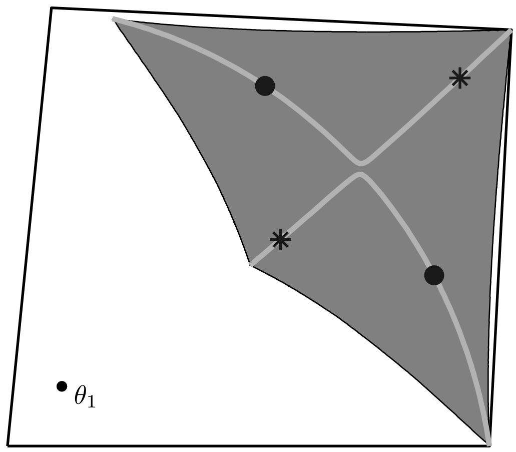

Figure 5: An example when $n = 4$. The eigenvalues lie at the vertices of the quadrilateral; θ_1 has been specified. If $k = 2$, the shaded region represents all allowable locations for the other Ritz value. Note that this region lies in the convex hull of two of the eigenvalues and two other points, both of which are isogonal conjugates of θ_1 from triangles formed from the eigenvalues. If $k = 3$, the remaining two Ritz values must lie on the two curves shown within the shaded region: θ_2 may be chosen anywhere on these curves. Once θ_2 is chosen, θ_3 is uniquely determined: it must lie at a unique point on the other curve. In the figure, the two circles show the locations of one possible choice for the pair $\{\theta_2, \theta_3\}$; the two stars show another.

values may lie given only θ, one must repeat this process for additional v; for an example, see Figure 5. The union of the field of values of $\widehat{V}^* A \widehat{V}$ for all such v would indicate where the remaining Ritz values must lie for any size restriction having one prescribed Ritz value at θ.

One thing should be clear from Figure 5: if we specify one Ritz value, it is unlikely that we will be able to place a second Ritz value wherever we want in the field of values. But if we wish to attempt to specify more than one Ritz value $\theta_1, \theta_2, \ldots, \theta_k$ (perhaps we already have part of a valid combination) then the above procedure is unsuitable. Starting with the affine space corresponding to θ_1, we would then have to search through all possible $\widehat{V}$ such that θ_2 is an eigenvalue of $\widehat{V}^* A \widehat{V}$, to determine those subspaces that give the second Ritz value, and so on. From the adjugate-based approach, we may do something similar in spirit to the procedure for specifying just one of the Ritz values above. Suppose we have some $\theta_1, \ldots, \theta_k$ for $k \le \lfloor (n-1)/2 \rfloor$, and we wish to determine all possible V such that $\theta_1, \ldots, \theta_k$ are the Ritz values from some $V^* A V \in \mathbb{C}^{(n-1)\times(n-1)}$. Equation (3) gives the form of the characteristic polynomial p of $V^* A V$ in terms of the vector v orthogonal to the range of V. If $\theta_1, \ldots, \theta_k$ are to be roots of p, then v must satisfy

$$\begin{bmatrix} \operatorname{Re}\ell_1(\theta_1) & \cdots & \operatorname{Re}\ell_n(\theta_1) \\ \operatorname{Im}\ell_1(\theta_1) & \cdots & \operatorname{Im}\ell_n(\theta_1) \\ \vdots & & \vdots \\ \operatorname{Re}\ell_1(\theta_k) & \cdots & \operatorname{Re}\ell_n(\theta_k) \\ \operatorname{Im}\ell_1(\theta_k) & \cdots & \operatorname{Im}\ell_n(\theta_k) \\ 1 & \cdots & 1 \end{bmatrix} x = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{11}$$

where again $x_j = |v_j|^2$ and $\ell_i(\lambda) = \prod_{j \neq i}(\lambda - \lambda_j)$. Equation (11) determines an affine subspace for x of dimension at least $n-2k-1$. If this subspace intersects the positive orthant, then there exist V such that $\theta_1, \ldots, \theta_p$ are Ritz values of (A, V). To determine where the remaining Ritz values lie for all possible V, one would have to consider all v that satisfy (11).

5. Nonnormal Example

To contrast the results for normal matrices, consider the nonnormal matrix

$$A = \begin{bmatrix} \gamma & & \\ & i\alpha & \beta \\ & & -i\alpha \end{bmatrix},$$

for $\alpha, \beta, \gamma > 0$. The field of values $W(A)$ is the convex hull of the origin and an ellipse with foci at $i\alpha$ and $-i\alpha$ and semimajor axis of length $\sqrt{\alpha^2 + \beta^2/4}$. The adjugate approach for $p(\lambda) = \det(V^*(\lambda I - A)V)$ still applies, and reveals that

$$p(\gamma) = |v_1|^2(\gamma^2 + \alpha^2) = (\gamma - \theta_1)(\gamma - \theta_2),$$

hence guaranteeing that the Ritz values lie on "isogonal Cevians" emanating from γ and reflected across the real axis. Given θ_1, this restricts where θ_2 can lie. Evaluating $p(i\alpha)$ reveals

$$p(i\alpha) = (i\alpha - \gamma)(2i\alpha|v_2|^2 + \beta v_3 \overline{v_2}) = (i\alpha - \theta_1)(i\alpha - \theta_2). \tag{12}$$

If β were zero, and hence A normal, then given θ_1 this condition would uniquely determine θ_2. However, for β not equal to zero, while (12) does further restrict the location of θ_2, it does not uniquely determine it; see Figure 6. Moreover, the subspaces determined by the Ritz values are no longer simply related; i.e., in the normal case, all possible v have the same magnitude entries. Now, the Ritz values may be parameterized by three rather than two real variables, v_1, $|v_2|$ and $\arg v_2$, with the assumption that v_1 and v_3 are both real. In general, having normal eigenvalues reduces the number of variables needed to parametrize the set of possible Ritz values; however the Ritz values must then satisfy constraints involving their phases.

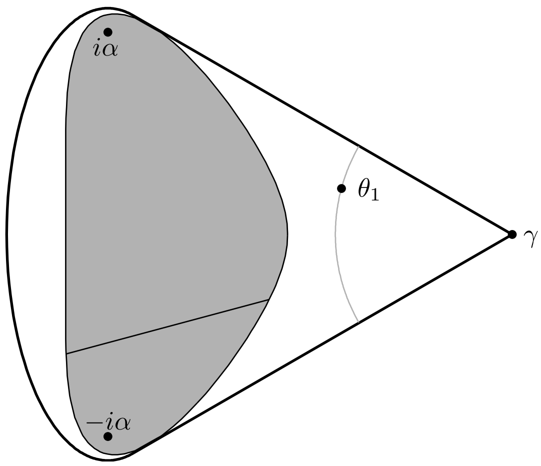

Figure 6: For $\alpha = 1/2$, $\beta = 1/2$ and $|\theta_1 - \gamma| = 7/16$. The line segment indicates where θ_2 can lie for the plotted θ_1. The gray region indicates the union of all such line segments for θ_1 lying in the field of values and on the arc of the circle centered at γ of radius 7/16.

6. Application: restarted Arnoldi convergence

In this section we use the uniqueness of the pairing of the Ritz values for $n = 3$ to analyze the restarted Arnoldi method applied to $A = \operatorname{diag}(\lambda, i\alpha, -i\alpha)$, with $\lambda, \alpha > 0$. For this case one can derive analytic expressions that show that the restarted Arnoldi method must converge. In general for non-Hermitian matrices, such analytic expressions are not possible, nor is there any straightforward qualitative analysis. Even for this small normal matrix, the restarted Arnoldi method has not been fully studied, although we note that some convergence criteria have been established in [15].

The restarted Arnoldi method is useful for computing a few eigenvalues of a matrix. Here we recall relevant features of the method; for more details, see, for example, [16]. The method works by utilizing Ritz values of (A, V) to compute a $\widehat{V}$ with orthonormal columns for which $\mathcal{R}(\widehat{V})$ contains better approximations to some desired eigenspace (i.e., a subspace spanned by eigenvectors). For our matrix, we start with some $V \in \mathbb{C}^{3\times 2}$ having orthonormal columns and compute the Ritz values θ_1 and θ_2 of (A, V). Label the Ritz values so that $\operatorname{Re}\theta_1 \leq \operatorname{Re}\theta_2$. Use the leftmost Ritz value θ_1 to compute $\widehat{V}$ such that

$$\mathcal{R}(\widehat{V}) = \mathcal{R}\left((A - \theta_1 I)V\right).$$

Assign $\widehat{V} \to V$ and repeat this process until (hopefully) θ_2 converges to the rightmost eigenvalue λ. This approach to restarting is known as "exact shifts" [16]. We wish to understand the convergence of this process.

By Theorem 3, θ_2 is uniquely determined by a given $\theta_1 \in W(A) \setminus \sigma(A)$.

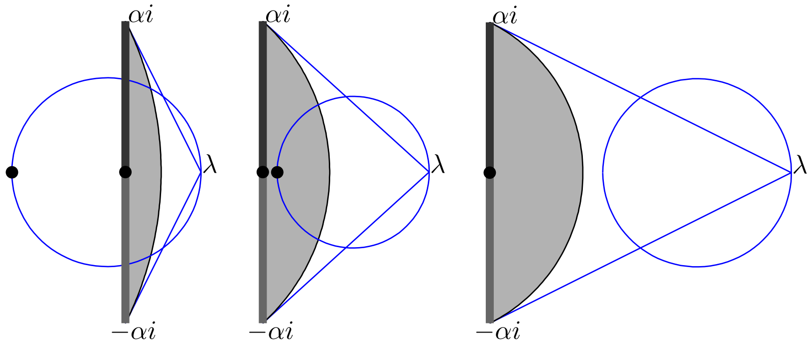

Figure 7: Shifts, fixed points, limit cycles and convergence basins for $\lambda = \alpha/2$, 1.1α and 2α. The gray region indicates where all the shifts, the θ_1, can be located. If a shift falls inside the circle, then the previous iteration of Arnoldi did not improve the eigenvector estimate. The black dots indicate the fixed points. For $\lambda = 1.1\alpha$ there are two fixed points, with the one on the left being attractive and the one on the right being repulsive. The line segments connecting the eigenvalues at $i\alpha$ and $-i\alpha$ indicate the period two limit cycles to which the shifts converge. Note that the portion of the segments contained in the circle correspond to repulsive limit cycles. For $\lambda = 2\alpha$, the circle does not intersect the gray region, hence the restarted Arnoldi method with exact shifts converges monotonically.

Explicitly, we have

$$\theta_2 = \lambda + \frac{(\alpha^2 + \lambda^2)\operatorname{Re}\theta_1}{\lambda\left(\frac{1}{4}\left(\lambda + \frac{\alpha^2}{\lambda}\right)^2 - \left|\theta_1 - \frac{1}{2}\left(\lambda - \frac{\alpha^2}{\lambda}\right)\right|^2\right)}(\overline{\theta_1} - \lambda). \tag{13}$$

This expression shows that θ_2 lies on the isogonal of the Cevian from λ through θ_1. Similar expressions for θ_2 as the isogonal of θ_1 could be derived using the eigenvalues $i\alpha$ and $-i\alpha$. The denominator of the second term in (13) gives the signed distance of θ_1 from the circle that passes through all the eigenvalues. Hence, this expression breaks down for $\theta_1 \in \sigma(A)$: there no longer is a one-to-one correspondence between θ_1 and θ_2. If $\operatorname{Re}\theta_1 \leq \operatorname{Re}\theta_2$, then taking the real part of (13) and substituting the right hand side into $\operatorname{Re}\theta_2 \leq \operatorname{Re}\theta_1$, one can see that θ_1 must lie inside the circle of radius $\alpha\sqrt{1 + \alpha^2/\lambda^2}$ centered at $-\alpha^2/\lambda$. To understand the restarted Arnoldi method, we must know how θ_1 moves about this circle. For θ_2 to converge to λ, $\operatorname{Re}\theta_1$ must go to zero. For illustrations of these circles containing the shifts for different λ and α, see Figure 7.

Equations (5) and (13) allow us to represent the unit vector v orthogonal to $\mathcal{R}(V)$ solely in terms of θ_1. The vector orthogonal to $\widehat{V}$, denoted $\widehat{v}$, is proportional to $(A - \theta_1 I)^{-*}v$, hence we may also represent $\widehat{v}$ in terms of θ_1. Thus, for this normal matrix, we may study the restarted Arnoldi method solely through the mangitudes of the entries of the vectors v and $\widehat{v}$. Expressing v and

$\widehat{v}$ in terms of θ_2:

$$|v_1|^2 = \frac{|\widehat{v}_1|^2|\theta_1-\lambda|^2}{\frac{1}{4}\big(\lambda+\frac{\alpha^2}{\lambda}\big)^2-\big|\theta_1-\frac{1}{2}\big(\lambda-\frac{\alpha^2}{\lambda}\big)\big|^2}, \quad |\widehat{v}_1|^2 = \frac{\operatorname{Re}\theta_1}{\lambda}, \tag{14}$$

$$|v_2|^2 = \frac{|\widehat{v}_2|^2|\theta_1-\alpha i|^2}{\frac{1}{4}\big(\lambda+\frac{\alpha^2}{\lambda}\big)^2-\big|\theta_1-\frac{1}{2}\big(\lambda-\frac{\alpha^2}{\lambda}\big)\big|^2}, \quad |\widehat{v}_2|^2 = \frac{\operatorname{Im}\big((\lambda-\alpha i)(\theta_1+\alpha i)\big)}{2\lambda\alpha}, \tag{15}$$

where we have omitted the expressions for v_3 and $\widehat{v}_3$, as both v and $\widehat{v}$ are unit vectors. Using these same equations, but for the next iteration, we can relate θ_1 and $\widehat{\theta}_1$:

$$\operatorname{Re}\theta_1 = \frac{|\widehat{\theta}_1-\lambda|^2}{\frac{1}{4}\big(\lambda+\frac{\alpha^2}{\lambda}\big)^2-\big|\widehat{\theta}_1-\frac{1}{2}\big(\lambda-\frac{\alpha^2}{\lambda}\big)\big|^2}\operatorname{Re}\widehat{\theta}_1 \tag{16}$$

$$\operatorname{Im}\theta_1 = -\frac{\alpha^2\big(1+\frac{\alpha^2}{\lambda^2}\big)-\big|\widehat{\theta}_1+\frac{\alpha^2}{\lambda}\big|^2}{\frac{1}{4}\big(\lambda+\frac{\alpha^2}{\lambda}\big)^2-\big|\widehat{\theta}_1-\frac{1}{2}\big(\lambda-\frac{\alpha^2}{\lambda}\big)\big|^2}\operatorname{Im}\widehat{\theta}_1. \tag{17}$$

These expressions are interesting in that they describe a discrete dynamical system in θ_1 that corresponds to running restarted Arnoldi in reverse. Though not immediately evident from these expressions, the pairing of the subspaces from running Arnoldi forward or backward is unique. This is not evident from (16) and (17), as they allow for a subspace to specified by either of its two Ritz values. Equation (17) for $\operatorname{Im}\theta_1$ involves the ratio of the distance of $\widehat{\theta}_1$ to two circles; in the denominator the circle passes through the eigenvalues and in the numerator the circle contains all possible θ_1. As θ_1 always lies in both these circles, this ratio is nonnegative. Moreover this ratio is always less than or equal to one. Hence the imaginary part of θ_1, if nonzero, alternates in sign and its magnitude is increasing with each iteration of the restarted Arnoldi method. An example of this alternating of the sign of the shifts can be seen in Figure 8.

In order for the restarted Arnoldi method to converge, the range of V must in the limit contain $e_1 = [1,0,0]^T$, the eigenvector associated with λ. This corresponds to $|v_1|$ going to zero. From (14), this is equivalent to having $\operatorname{Re}\theta_1$ go to zero. For progress to be made during one iteration, $\operatorname{Re}\widehat{\theta}_1$ must be less than $\operatorname{Re}\theta_1$. From (16), for $\operatorname{Re}\widehat{\theta}_1$ to be less than $\operatorname{Re}\theta_1$, $\widehat{\theta}_1$ must lie inside the circle of radius $(\lambda^2+\alpha^2)/4\lambda$ centered at $(3\lambda-\alpha^2/\lambda)/4$. From this circle and the circle containing θ_1, for our matrix we can completely understand the behavior of the restarted Arnoldi method. This is illustrated in Figure 8.

There are several special regions to note. For $0 \le \lambda \le \sqrt{3}\alpha$, there is a fixed point for θ_1 located at $(\lambda^2-\alpha^2)/2\lambda$. This fixed point corresponds to stagnation, in which case θ_2 never converges to the eigenvalue at λ. This fixed point is repulsive, hence for nearly all starting subspaces the restarted Arnoldi method does not stagnate. For $\lambda > \sqrt{3}\alpha$ there is a fixed point located at 0, and it is attractive. All the points connecting the eigenvalues $i\alpha$ and $-i\alpha$ are periodic points of order two. They correspond to cycles for which the imaginary

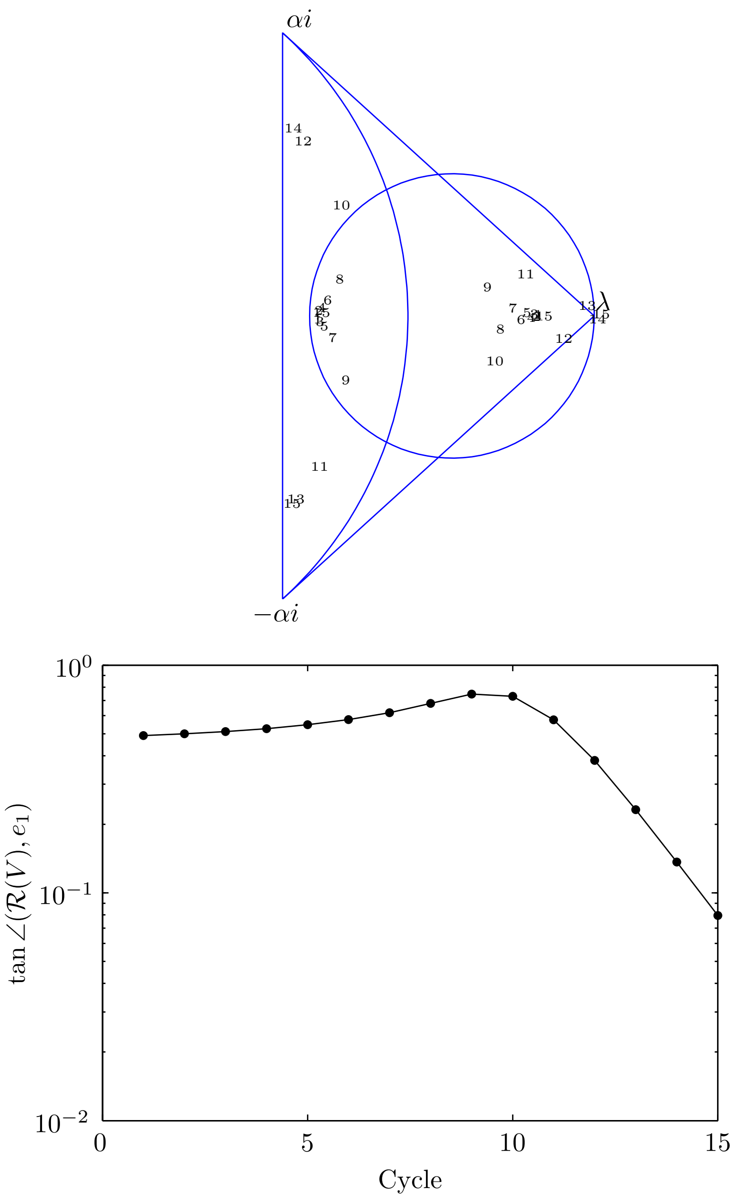

Figure 8: Convergence example for $\lambda = 1.1\alpha$. The top plot shows the Ritz values from several cycles of restarted Arnoldi for an initial shift lying in the circle corresponding to progress not having been made in the previous iteration. For each cycle the Ritz values are plotted using the cycle number. The curves plotted are as in the lefmost plot of Figure 7. The bottom plot shows the tangent of the angle between the $\mathcal{R}(V)$ and the desired eigenvector at each cycle of the restarted Arnoldi method. After a period of divergence, $\mathcal{R}(V)$ starts to converge to a subspace that contains the wanted eigenvector.

part of θ_1 alternates in sign while θ_2 remains fixed at λ. These are the limit cycles for the restarted Arnoldi method. There are also sources, corresponding to subspaces that could only be generated by having used θ_2, the rightmost Ritz value, to generate $\widehat{V}$. In terms of θ_1, the sources lie near the top and bottom portions of the boundary of the circle containing θ_1. Illustrations of these different scenarios for α and λ were shown in Figure 7.

The type of failure exhibited for $0 \leq \lambda \leq \sqrt{3}\alpha$ must be contrasted with the construction given by Embree [17] for constructing examples in which a shift falls on a wanted eigenvalue, resulting in the annihilation, in exact arithmetic, of the desired eigenvector from the $\mathcal{R}(V)$ for all successive iterations. Restarted Arnoldi may still converge, but not to the wanted eigenvalue. For this stagnation point restarted Arnoldi never converges. Embree's construction relies upon the field of values associated with the unwanted eigenvalues containing the desired eigenvalue. This is not possible here due to the matrix being normal.

From above we know that for nearly all starting V, the restarted Arnoldi method will converge, though it may initially diverge. The rate at which it does so can be determined from (16). Maximizing the ratio of $\operatorname{Re}\widehat{\theta}_1$ to $\operatorname{Re}\theta_1$ gives a rate of $(\alpha^2 + \alpha\sqrt{\alpha^2 + \lambda^2})/\lambda^2$. This quantity is less than one for $\lambda > \sqrt{3}\alpha$, in which case we have unconditional convergence. If $\alpha < \lambda < \sqrt{3}\alpha$, we have local convergence, in the sense that if the initial approximation is close to the desired eigenvector for λ, then the approximations can only improve. For $\lambda < \alpha$, θ_2 can be arbitrarily close to λ and θ_2 can still initially diverge from λ. With the exception of the stagnation point, all θ_1 must eventually converge to a period two cycle along the imaginary axis. The asymptotic rate at which θ_1 converges to such a cycle is $(\alpha^2 - y^2)/(\lambda^2 + y^2)$, where y is the limit of the imaginary part of θ_1. This rate is bounded above by α^2/λ^2, which is the rate you would expect from the power method, provided that $\alpha < \lambda$.

7. Future work

The results of this work noted the constraints that normality places upon $n-1$ Ritz values for a n-dimensional normal matrix. Future work could include developing a better understanding of $k < n-1$ Ritz values for a normal matrix. Majorization results for the real and imaginary parts, as well as the phases of the Ritz values, are known to hold for general matrices [18]. An open question is whether for normal matrices there is anything sharper than majorization.

While preparing this paper, we became aware that Bujanović [11] had shown for normal matrices that Ritz values $\theta_1, \ldots, \theta_k$ that are distinct from the eigenvalues and arising when $\mathcal{R}(V)$ is a *Krylov subspace* (a subspace spanned by vectors of the form $A^i b$ for $i = 0, 1, \ldots, k-1$ for some starting vector b) must satisfy a particular system of equations. After seeing Bujanović's result, we realized this system of equations is the very same system of equations that would arise were one to attempt to prescribe k Ritz values coming from a $n-1$ dimensional subspace; see (11). Thus any set of Ritz values coming from a Krylov subspace can be extended to a set of Ritz values coming from an $n-1$ dimensional subspace. A vector orthogonal to this $n-1$ dimensional subspace is simply $p(A)b$, where b is the starting vector of the Krylov subspace and p is a polynomial whose roots are the k Ritz values. Bujanović used the optimality of Ritz values from a Krylov subspace to derive his result. The optimality property of Ritz values from Krylov subspaces holds for nonnormal matrices as well. Moreover, an expression similar to (11), though now nonlinear in v, can be derived for nonnormal matrices as well, and it can be shown to be the very same expression that Ritz values from a Krylov subspace must satisfy due to optimality. Thus if one wishes to study Ritz values arising from Krylov subspaces, it suffices to study the Ritz values from $n-1$ dimensional subspaces.

A characterization of $n-1$ Ritz values from a normal matrix could provide much insight into the restarted Arnoldi method for eigenvalues. Duintjer Tebbens and Meraunt [19] have already shown that any sequence of Ritz values may be realized by the Arnoldi method, i.e., the Ritz values from a Krylov subspace can be arbitrarily far away from the eigenvalues up until the last iteration. However, Duintjer Tebbens and Meurant's result indicates what types of Ritz value behavior are possible for all matrices where one can adjust the normality to obtain the desired behavior. In practice, one is only concerned with the Ritz values coming from a particular matrix with fixed nonnormality. A better understanding of the inverse field of values problem for k Ritz values could also offer insight into the behavior of Ritz values for particular matrices. Forthcoming work address Ritz value localization for nonnormal matrices [20].

References

[1] F. Uhlig, An inverse field of values problem, Inverse Problems 24 (2008) 055019.

[2] C. R. Johnson, Numerical determination of the field of values of a general complex matrix, SIAM J. Numer. Anal. 15 (1978) 595–602.

[3] R. Carden, A simple algorithm for the inverse field of values problem, Inverse Problems 25 (2009) 115019.

[4] C. Chorianopoulos, P. Psarrakos, F. Uhlig, A method for the inverse numerical range problem, Electron. J. Linear Algebra 20 (2010) 198–206.

[5] B. N. Parlett, The Symmetric Eigenvalue Problem, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.

[6] K. Fan, G. Pall, Imbedding conditions for Hermitian and normal matrices, Canad. J. Math. 9 (1957) 298–304.

[7] D. Carlson, E. Marques de Sá, Generalized minimax and interlacing theorems, Linear and Multilinear Algebra 15 (1984) 77–103.

[8] J. F. Queiró, A. L. Duarte, Imbedding conditions for normal matrices, Linear Algebra Appl. 430 (2009) 1806–1811.

[9] S. M. Malamud, Inverse spectral problem for normal matrices and the Gauss-Lucas theorem, Trans. Amer. Math. Soc. 357 (2005) 4043–4064.

[10] R. C. Thompson, Principal submatrices of normal and Hermitian matrices, Illinois J. Math. 10 (1966) 296–308.

[11] Z. Bujanović, Krylov Type Methods for Large Scale Eigenvalue Computations, Ph.D. thesis, University of Zagreb, 2011.

[12] R. A. Horn, C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1994.

[13] J.-P. M. Zemke, Hessenberg eigenvalueeigenmatrix relations, Linear Algebra Appl. 414 (2006) 589–606.

[14] R. A. Johnson, Advanced Euclidean Geometry: An Elementary Treatise on the Geometry of the Triangle and the Circle, Dover, New York, 1960.

[15] R. Carden, Ritz Values and Arnoldi Convergence for Nonsymmetric Matrices, Master's thesis, Rice University, 2009.

[16] D. C. Sorensen, Implicit application of polynomial filters in a k-step Arnoldi method, SIAM J. Matrix Anal. Appl. 13 (1992) 357–385.

[17] M. Embree, The Arnoldi eigenvalue iteration with exact shifts can fail, SIAM J. Matrix Anal. Appl. 31 (2009) 1–10.

[18] R. Carden, Ritz Values and Arnoldi Convergence for Non-Hermitian Matrices, Ph.D. thesis, Rice University, 2011.

[19] J. Duintjer Tebbens, G. Meurant, Any Ritz value convergence behavior is possible for the Arnoldi method and a consequence for the GMRES method, 2011. Submitted.

[20] R. Carden, M. Embree, Ritz value localization for non-Hermitain matrices (2011). In preparation.