RICE UNIVERSITY

Ritz Values and Arnoldi Convergence for Non-Hermitian Matrices

by

Russell Carden

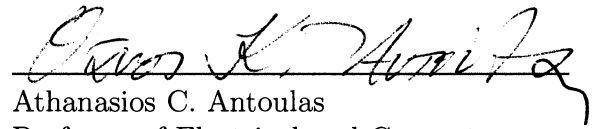
A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Approved, Thesis Committee:

Mark Embree, Chair
Professor of Computational and Applied Mathematics

Danny C. Sorensen
Noah G. Harding Professor of
Computational and Applied Mathematics



Athanasios C. Antoulas
Professer of Electrical and Computer Engineering

Houston, Texas

August, 2011

ABSTRACT

Ritz Values and Arnoldi Convergence for Non-Hermitian Matrices

by

Russell Carden

This thesis develops ways of localizing the Ritz values of non-Hermitian matrices. The restarted Arnoldi method with exact shifts, useful for determining a few desired eigenvalues of a matrix, employs Ritz values to refine eigenvalue estimates. In the Hermitian case, using selected Ritz values produces convergence due to interlacing. No generalization of interlacing exists for non-Hermitian matrices, and as a consequence no satisfactory general convergence theory exists. To study Ritz values, I propose the inverse field of values problem for k Ritz values, which asks if a set of k complex numbers can be Ritz values of a matrix. This problem is always solvable for $k = 1$ for any complex number in the field of values; I provide an improved algorithm for finding a Ritz vector in this case. I show that majorization can be used to characterize, as well as localize, Ritz values. To illustrate the difficulties of characterizing Ritz values, this work provides a complete analysis of the Ritz values of two 3×3 matrices: a Jordan block and a normal matrix. By constructing conditions for localizing the Ritz values of a matrix with one simple, normal, sought-after eigenvalue, this work develops sufficient conditions that guarantee convergence of the restarted Arnoldi method with exact shifts. For general matrices, the conditions provide insight into the subspace dimensions that ensure that shifts do not cluster near the wanted eigenvalue. As Ritz values form the basis for many iterative methods for determining

eigenvalues and solving linear systems, an understanding of Ritz value behavior for non-Hermitian matrices has the potential to inform a broad range of analysis.

Acknowledgments

First, I thank my family, for supporting me in all my endeavors. I would like to thank the members of my committee for their encouragement. I thank my adviser Dr. Embree whose skills as an applied mathematician, professor and writer are inspiring. I thank Josef Sifuentes for helping me throughout my graduate studies. I thank my officemate Nabor Reyna for keeping me on task. I thank all my friends from AGEP for helping me let of steam. I am grateful to the members of the eigenvalue clinic for listening to me talk about my research: Josef Sifuentes, Jeffery Hokanson, Derek Hansen, Cosmin Ionita and Sanda Lefteriu. I thank Dr. Richard Tapia and Dr. Pablo Tarazaga, without whom I would not have considered applying to Rice. I thank Dr. Stephen Sedory for sparking my interest in linear algebra.

Contents

Illustrations

Chapter 1

Introduction

This thesis analyzes the behavior of Ritz values and convergence of the restarted Arnoldi method for non-Hermitian matrices. Understanding the behavior of Ritz values is essential to establishing convergence of methods for approximating eigenvalues. In the Hermitian case, for which much is known, convergence of the restarted Arnoldi method with exact shifts follows from the interlacing property of Ritz values. In the non-Hermitian case, the restarted Arnoldi method with exact shifts performs well in practice. Researchers have observed typical patterns of behavior for Ritz values, but these are not sufficiently understood to establish criteria for convergence. In this thesis I show that Ritz values for non-Hermitian matrices can be localized. This information leads to sufficient criteria for convergence of the restarted Arnoldi method with exact shifts for these non-Hermitian matrices. Specifically, the criteria will address constraints on the starting vector and the spectral properties of the original matrix. I believe these to be the first such convergence results to be obtained for the restarted Arnoldi method with exact shifts for a nontrivial non-Hermitian matrix.

1.1 Eigenvalues

The restarted Arnoldi method is an indispensable tool for determining a few selected eigenvalues of a matrix, for example, those with the largest real part. Eigenvalues provide insight into the behavior of dynamical systems. They indicate how modeled

features will grow, decay or oscillate. An important use of eigenvalues is to determine the stability of systems modeled by $x'(t) = Ax(t)$. Real world systems cannot sustain the constant growth of a modeled feature that corresponds to eigenvalues with positive real part. Similarly, undampened oscillations associated with purely imaginary eigenvalues, when coupled with resonant driving forces, can lead to unsustainable behavior. The determination of system stability typically involves the computation of only a few of the eigenvalues. This is fortunate, as it limits the work needed to analyze real-word problems; moreover, many of the eigenvalues of the matrix can be spurious approximations to the true eigenvalues of an underlying infinite dimensional operator.

Determining eigenvalues of $A \in \mathbb{C}^{n\times n}$ involves finding x and λ such that

$$Ax = x\lambda,$$

where x, the *eigenvector*, is a nonzero vector of dimension n and λ, the *eigenvalue*, is a complex scalar. The vector x and scalar λ are said to be an eigenpair, and are denoted (x, λ). The set of all eigenvalues of a matrix is called the *spectrum* of the matrix, denoted by $\sigma(A)$. A space spanned by eigenvectors is an *eigenspace.*

Eigenvalues can be characterized as roots of the characteristic polynomial,

$$p(\lambda) = \det(\lambda I - A).$$

The roots of the characteristic polynomial correspond to the values λ for which $\lambda I - A$ is singular, meaning that $\lambda I - A$ has a nontrivial null space. For each unique root there exists at least one eigenvector. As the roots of polynomials of order greater than five cannot generally be determined in a finite number of steps, eigenvalues must be determined iteratively.

If $A \in \mathbb{C}^{n\times n}$ is *Hermitian*, $A = A^*$, then all its eigenvalues are real. If A is *normal*, $AA^* = A^*A$, then A is unitarily diagonalizable, i.e., the eigenvectors of A can be selected to form a complete orthonormal basis. As a consequence, $A = X\Lambda X^*$ where Λ is a diagonal matrix with the eigenvalues of A along the diagonal and $X \in \mathbb{C}^{n\times n}$ is *unitary*, $X^*X = I$, with eigenvectors of A in its columns. Hermitian matrices are normal; however, unlike the Hermitian case, the eigenvalues of a generic normal matrix can be complex. Any matrix with a set of eigenvectors that spans $\mathbb{C}^n$ is called *diagonalizable*, and may be decomposed as $A = X\Lambda X^{-1}$, where Λ is a diagonal matrix with eigenvalues along the diagonal and X is a nonsingular –not necessarily unitary– matrix whose columns are eigenvectors of A. A *nonnormal* matrix, $AA^* \neq A^*A$, need not be diagonalizable: the eigenvectors of A need not that span all of $\mathbb{C}^n$. Nonormality can give rise to much difficulty in the computation of eigenvalues.

1.2 Ritz Values

Knowing the action of a matrix upon various subspaces allows the approximation of a few of its eigenvalues. Eigenvalue approximations from a subspace are known as Ritz values. The set of all possible Ritz values of a matrix from a one dimensional subspace is known as the *field of values* or *numerical range*. The field of values (FOV) of $A \in \mathbb{C}^{n\times n}$ is

$$W(A) := \{v^*Av : v \in \mathbb{C}^n, v^*v = 1\}.$$

Hence the field of values is the set of all possible Rayleigh quotients of a matrix. Methods for sketching the field of values [32] make use of the following property: the Hermitian part of A, $\mathrm{H}(A) := \frac{1}{2}(A + A^*)$, satisfies $W(\mathrm{H}(A)) = \mathrm{Re}\, W(A) = \{\mathrm{Re}\, z : z \in W(A)\} = [\mu_1, \mu_n]$, where μ_1 and μ_n are the largest and smallest eigenvalues of $\mathrm{H}(A)$. Hence the boundary of the field of values of A can be determined by computing

$W(\mathrm{H}(e^{i\theta}A))$ for various values of θ.

Ritz values have many interesting properties. The Ritz value associated with unit vector v is the scalar θ such that $Av - v\theta$ is orthogonal to the space spanned by v, i.e., $v^*(Av - v\theta) = 0$. A Ritz value is optimal in the sense that it minimizes the norm of $Av - v\theta$. The Ritz value represents the action of A restricted to the subspace spanned by v. For a Ritz value to be a good approximation to an eigenpair, the residual $Av - v\theta$ should be small. For Ritz values generated from a subspace spanned by the columns of a matrix V, the residual is required to be orthogonal not only to the Ritz vector, but also to any vector in the subspace, i.e.,

$$V^*(Av - v\theta) = 0.$$

Since any $v \in \mathrm{Ran}(V)$ can be written as Vy,

$$V^*(AVy - Vy\theta) = 0,$$

which becomes

$$Hy = V^*Vy\theta,$$

where $H = V^*AV$. The problem of determining Ritz values from a subspace is then equivalent to solving the generalized eigenvalue problem $Hy = V^*Vy\theta$. If the columns of V are orthonormal, then $V^*V = I$, and the problem of determining Ritz values from a subspace reduces to an ordinary eigenvalue problem.

Eigenvalues can be thought of as are Ritz values for which $\|Av - vv^*Av\| = 0$. As not all Ritz values are good approximations to eigenvalues, the question then is: what exactly is the relationship between Ritz values and eigenvalues? How may Ritz values be used to localize and approximate eigenvalues? The restarted Arnoldi method with exact shifts uses Ritz values to approximate eigenvalues and often works

well in practice; however, our current knowledge of Ritz values does not explain why it works so well. As Ritz values are simply another means of observing the action of a matrix, such questions are naturally related to inclusion regions for eigenvalues.

1.3 Inclusion Regions for Eigenvalues

Different characterizations of eigenvalues give rise to different inclusion regions. Such inclusions are helpful for several reasons. If they are easy to compute, then they may in some cases quickly answer questions related to eigenvalues, such as system stability or rate of decay. Also, they may provide insight into the conditioning of the eigenvalues, i.e., sensitivity to perturbation, as well as methods for computing eigenvalues.

1.3.1 Min-max and majorization characterization

The eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ of a Hermitian matrix satisfy the min-max theorem of Courant and Fischer.

Theorem 1.1 (Courant-Fisher) If $A \in \mathbb{C}^{n\times n}$ is Hermitian with eigenvalues λ_i such that $\lambda_1 \leq \cdots \leq \lambda_n$, then

$$\min_{\substack{\dim(\mathcal{U})=k \\ \mathcal{U}\subset\mathbb{C}^n}} \max_{x\in\mathcal{U},\, x\neq 0} \frac{x^*Ax}{x^*x} = \lambda_k,$$

$$\max_{\substack{\dim(\mathcal{U})=n-k+1 \\ \mathcal{U}\subset\mathbb{C}^n}} \min_{x\in\mathcal{U},\, x\neq 0} \frac{x^*Ax}{x^*x} = \lambda_k.$$

Replacing the subspace $\mathcal{U}$ by a particular subspace spanned by, say, the columns of a matrix $V \in \mathbb{C}^{n\times m}$ with $V^*V = I_m$, $m \geq k$, gives an upper bound

$$\lambda_k \leq \min_{\substack{\dim(\mathcal{U})=k, \\ \mathcal{U}\subset \mathrm{Ran}(V)}} \max_{x\neq 0, x\in\mathcal{U}} \frac{x^*Ax}{x^*x} = \theta_k,$$

where θ_k denotes the kth eigenvalue of the matrix V^*AV, ordered by increasing real part (i.e., θ_k is a Ritz value for A from $\mathrm{Ran}(V)$). Similarly, for $k \geq n - m$ we have a lower bound

$$\theta_{m-n+k} \leq \lambda_k(A).$$

Combining these two results gives the Cauchy interlacing theorem.

Theorem 1.2 (Cauchy interlacing) If $A \in \mathbb{C}^{n\times n}$ is Hermitian with eigenvalues λ_i, and $V^*AV \in \mathbb{C}^{m\times m}$ with $V \in \mathbb{C}^{n\times m}$, $V^*V = I$ with eigenvalues θ_i, then

$$\lambda_k \leq \theta_k \leq \lambda_{n+m-k}$$

for $k = 1, 2, \ldots, m$.

Combining Cauchy interlacing with knowledge of the Ritz values of A from a subspace, we may determine inclusion regions for some of the interior eigenvalues, and exclusion regions for some of the exterior eigenvalues [44]. If we are interested in the largest or extreme eigenvalues, interlacing tells us that as we increase the size of our subspace $\mathrm{Ran}(V)$, the extreme Ritz values must march towards the extreme eigenvalues. Interlacing is precisely why the restarted Arnoldi method works for Hermitian matrices.

A much weaker notion than interlacing is majorization. A real vector $x \in \mathbb{R}^n$ is said to *majorize* a vector $y \in \mathbb{R}^n$, denoted $y \prec x$, provided that

$$\sum_{i=1}^{k} y_{(i)} \leq \sum_{i=1}^{k} x_{(i)} \quad \text{for } k = 1, 2, \ldots, n, \tag{1.1}$$

where $x_{(i)}$, denotes the ordering from smallest to largest of the entries of x, i.e., $x_{(i)} \leq x_{(i+1)}$ for $i = 1 \ldots, n-1$. The majorization is said to be *strong* if we have equality in (1.1) for $k = n$, and *weak* otherwise [38]. Care must be taken in reading the

literature about majorization, as some authors define majorization with the inequality reversed, or sum the entries from largest to smallest. Unless otherwise noted, we will be referring to strong majorization and our sums will be from smallest to largest.

A 1923 theorem of Schur uses majorization to relate the diagonal entries of a Hermitian matrix to its eigenvalues [38].

Theorem 1.3 If $A \in \mathbb{C}^{n\times n}$ *is Hermitian with diagonal elements* $a_{11}, \ldots, a_{nn}$ *and eigenvalues* $\lambda_1, \ldots, \lambda_n$, *then* $[\lambda_i] \prec [a_{ii}]$.

The proof of this theorem uses induction on the size of the matrix, along with Cauchy interlacing.

Amir-Moez and Horn [2] considered the implication of this theorem for non-Hermitian matrices.

Theorem 1.4 Let $A \in \mathbb{C}^n$ *have eigenvalues* λ_i *such that* $\operatorname{Re}\lambda_1 \leq \cdots \leq \operatorname{Re}\lambda_n$ *and eigenvalues of* $\frac{1}{2}(A^* + A)$, μ_i *such that* $\mu_1 \leq \cdots \leq \mu_n$. *Then* $[\mu_i] \prec [\operatorname{Re}\lambda_i]$.

This theorem follows from Theorem 1.3 by considering a Schur decomposition, $A = QTQ^*$, where Q is unitary and T is upper triangular, for then the diagonal of the Hermitian part of T gives the real part of the eigenvalues of A. This theorem tells us that if we know some of the eigenvalues of the Hermitian part of a A, then we may determine regions in the complex plane where the eigenvalues of A must lie. Unlike interlacing, which only uses the action of the matrix on a subspace to localize eigenvalues, this result requires knowledge of the of eigenvalues of the Hermitian part of A, a property of the full matrix.

1.3.2 Singular values and eigenvalues

The singular values of $A \in \mathbb{C}^{n\times n}$ denoted $\sigma_i(A)$ with $\sigma_i(A) \geq \sigma_{i+1}(A)$ are the square roots of the eigenvalues of A^*A, and thus they also have min-max/max-min characterizations, as well as interlacing properties [31].

Lemma 1.1 Let $A \in \mathbb{C}^{m\times n}$, $V \in \mathbb{C}^{m\times k}$ and $W \in \mathbb{C}^{n\times k}$ be given, where $k \leq \min\{m,n\}$ and V and W have orthonormal columns. Then

$$\begin{aligned} (a) \quad & \sigma_i(V^*AW) \leq \sigma_i(A),\ i = 1,\ldots,k, \text{ and} \\ (b) \quad & |\det V^*AW| \leq \sigma_1(A)\cdots\sigma_k(A). \end{aligned}$$

From the second inequality, Weyl derived an inequality relating the magnitudes of the eigenvalues to the singular values [31].

Theorem 1.5 (Weyl's Theorem) Let $A \in \mathbb{C}^{n\times n}$ have singular values $\sigma_1(A),\ldots,\sigma_n(A) \geq 0$ and eigenvalues $\{\lambda_1(A),\ldots,\lambda_n(A)\} \subset \mathbb{C}$ ordered so that $|\lambda_1(A)| \geq \cdots \geq |\lambda_n(A)|$. Then

$$|\lambda_1(A)\cdots\lambda_k(A)| \leq \sigma_1(A)\cdots\sigma_k(A) \text{ for } k = 1,\ldots,n, \text{ with equality for } k = n. \quad (1.2)$$

This result says that the magnitudes of the eigenvalues are *log-majorized* by the singular values,

$$[\log|\lambda_i|] \prec [\log\sigma_i].$$

As the exponential function is convex, this log majorization of the magnitudes of the eigenvalues gives rise to the following weak majorization.

$$[|\lambda_i(A)|] \prec [\sigma_i(A)].$$

Thus, knowledge of some of the singular values of the full matrix gives disks in the complex plane in which a prescribed number of eigenvalues must lie.

1.3.3 Lehmann bounds

These singular value based bounds are not unlike Lehmann bounds, which are also inclusion regions for eigenvalues. Lehmann bounds were originally developed for Hermitian matrices [4], but have since been generalized to non-Hermitian matrices [5]. Determined from the action of a matrix upon a subspace, they determine disks in the complex plane that contain a prescribed number of eigenvalues. For normal matrices the bounds are optimal, in the sense that there exist matrices with eigenvalues for which the bounds are sharp.

For diagonalizable matrices we have the following result from Beattie and Ipsen [5].

Theorem 1.6 Let $A \in \mathbb{C}^{n\times n}$ be diagonalizable with an eigenvalue decomposition $A = X\Lambda X^{-1}$. For a given complex number ρ and an $n \times m$ matrix S having orthonormal columns, denote by $\tau_{-1} \leq \tau_{-2} \leq \cdots \leq \tau_{-m}$, the singular values of the $n \times m$ matrix $(A - \rho I)S$. Then for each $i \in \{1, \ldots, m\}$, the disk

$$D_i := \{z : |z - \rho| \leq \kappa(X)\tau_{-i}\},$$

contains at least i eigenvalues of A.

Beattie and Ipsen also have a result for nondiagonalizable matrices [5].

Theorem 1.7 Let A be an $n \times n$ matrix. For a given complex number ρ and an $n \times m$ matrix S with orthonormal columns, denote by $\tau_{-1} \leq \tau_{-2} \leq \cdots \leq \tau_{-m}$ the singular values of the $n \times m$ matrix $(A - \rho I)S$. Then each for each $i \in \{1, \ldots, m\}$, the disk

$$D_i := \{z : z - \rho| \leq \tau_{-i} + \delta(A)\},$$

contains at least i eigenvalues of A, where $\delta(A)$ is Henrici's departure from normality [26]: $\delta(A) = \min \|N\|$ where N is the strictly upper triangle portion of a Schur

decomposition of A, $A = Q(\Lambda + N)Q^*$, *and the* min *is over all possible Schur decompositions of* A.

In the normal case, $\kappa(X) = 1$ and $\delta(A) = 0$, these bounds can be computed without having either an eigendecomposition or a Schur factorization. All that is required are the singular values of the rectangular matrix $(A - \rho I)S$. Both Lehmann bounds leave something to be desired in the nonnormal case. Both generalizations involve quantities that are not easily computed for nonnormal matrices; however, we must note that for a nonnormal matrix, without any additional assumptions or knowledge of the full matrix other than its action on a subspace, it is impossible to localize the eigenvalues. In the normal case, the action of a matrix on a subspace $\mathrm{Ran}(V)$ also provides some information about the action of the matrix on the orthogonal subspace $\mathrm{Ran}(V)^{\perp}$. In the nonnormal case, nothing can be said about the action on the orthogonal subspace, and as a consequence, the eigenvalues of the full matrix may occur anywhere throughout the complex plane. For example, let $A \in \mathbb{C}^{2\times 2}$ and suppose $V = e_1$. Then

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

and the first column of A is known simply from knowing the action of A upon V. With the first column of A fixed, the eigenvalues of A can be placed anywhere in the complex plane by the choice of a_{21} and a_{22}.

In this light, for bounds such as those in Theorems 1.6 and 1.7 for nonnormal matrices, we must ask: How much more are we willing to assume or compute in order to localize some of the eigenvalues? For example, supposing that we knew some of the largest singular values of the matrix, then Weyl's condition, equation (1.2), would give inclusion regions for all the eigenvalues, $n - 1$ of the eigenvalues, and so on. It

is likely that such bounds will be sharp for normal matrices and not necessarily so for nonnormal matrices. These results have implications for Ritz values, which I will explore in Chapter 3.

1.3.4 Pseudospectra

Another type of inclusion region for eigenvalues is the pseudospectrum [57]. The ε-pseudospectrum of a matrix A for a given $\varepsilon > 0$ is the set of λ in the complex plane for which there exists some $E \in \mathbb{C}^{n\times n}$ with $\|E\| < \epsilon$ such that λ is an eigenvalue of $A + E$, i.e.,

$$\sigma_\varepsilon(A) = \{z \in \mathbb{C} : \exists\ E \text{ with } \|E\| < \varepsilon \text{ and } z \in \sigma(A + E)\}.$$

This definition may also be stated in terms of the matrix valued function known as the *resolvent*, $R(z) = (zI - A)^{-1}$:

$$\sigma_\varepsilon(A) = \{z \in \mathbb{C} : \|(zI - A)^{-1}\| > \varepsilon^{-1}\}.$$

A Ritz pair (θ, v) from $H = V^*AV$, $V \in \mathbb{C}^{n\times k}$ and $V^*V = I$, is in the ε-pseudospectrum of A for all $\varepsilon > \|Av - v\theta\|$. To see this, take $E = -rx^*$, where $r = Av - v\theta$. The bound on how accurate a Ritz pair can be for a given $\|r\|$ is reflected in how large the ε-pseudospectrum of A is, relative to the spectrum of A. For a normal matrix, the ε-pseudospectrum for a given ε consists of the union of open disks in the complex plane of radius ε centered about the eigenvalues. Hence, small perturbations to a normal matrix produce small changes to the eigenvalues. In this case, the eigenvalues are said to be well conditioned. For a nonnormal matrix, the pseudospectra can differ significantly from the spectrum. The condition of the eigenvalues may vary; some may be more sensitive to perturbations than others. Typical applications require ap-

proximating well-conditioned eigenvalues, which can be complicated by nonnormality associated with the remaining eigenvalues.

One of the primary uses of pseudospectra in the analysis of matrices is for the study of functions of matrices. There are several different ways of defining functions of matrices. Perhaps the most elegant makes use of the resolvent. Evaluation of a function of a matrix using the resolvent is a generalization of the Cauchy integral formula:

$$f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - A)^{-1} dz, \tag{1.3}$$

where Γ is any simple closed rectifiable curve containing the spectrum of A in its interior. Since the ε-pseudospectrum of A always contains the spectrum, equation (1.3) can be evaluated using the boundary of the ε-pseudospectrum for Γ. This does not facilitate the computation of $f(A)$; rather it simplifies the analysis of $\|f(A)\|$:

$$\begin{aligned} \|f(A)\| &\leq \frac{1}{2\pi} \int_{\partial\sigma_\varepsilon(A)} |f(z)| \|(zI - A)^{-1}\| |dz| \\ &= \frac{1}{2\pi\varepsilon} \int_{\partial\sigma_\varepsilon(A)} |f(z)| |dz| \\ &\leq \frac{L_\varepsilon}{2\pi\varepsilon} \max_{\partial\sigma_\varepsilon(A)} |f(z)|, \end{aligned}$$

where L_ε is the length of the boundary of the ε-pseudospectrum. Thus pseudospectra allow us to bound the norm of a matrix function by knowing the length of the boundary of the pseudospectrum, as well as the maximum of the function on the boundary of the pseudospectrum or simply on the pseudospectrum itself, by the maximum principle. This can be particularly useful, as it turns a matrix problem into a scalar problem.

1.4 Krylov Subspaces

The Arnoldi method approximates eigenvalues by orthogonally projecting a matrix onto a Krylov subspace. The performance and analysis of many iterative methods for eigenvalues and systems of equations rely on Krylov subspaces. A Krylov subspace is spanned by the iterates of the power method,

$$\mathcal{K}_k(A, v) = \text{span}\{v, Av, A^2v, \ldots, A^{k-1}v\},$$

where v is the starting vector. The power method itself is an eigenvalue method that approximates the eigenvector associated with a distinct, largest-magnitude eigenvalue, if such an eigenvalue exists. The power method approximates the vector that A^kv approaches (in angle), the desired eigenvector, as k becomes large. A Krylov subspace contains not only all power method iterates, but it contains all shifted power method iterates $(A - \sigma I)^i v$ for any complex σ, $i = 1, \ldots, k-1$.

Associated with a Krylov subspace is a particular polynomial optimization, the Arnoldi minimization problem [23], that is relevant for the computation of eigenvalues:

$$\min_{x \in \mathcal{K}_k(A,v)} \|A^kv - x\| = \min_{p \in \mathbb{P}^k} \|p(A)v\|,$$

where $\mathbb{P}^k$ denotes the set of monic polynomials of degree less than or equal to k. If the Krylov subspace spans an eigenspace, the minimum for this problem will be zero and the roots of p will be eigenvalues of A. In this case, p is said to be the annihilating or minimal polynomial for (A, v). If the smallest k such that p annihilates v is n, then v is said to be a cyclic vector. If one had a means of determining p, and if the resulting $\|p(A)v\|$ was small, then one might be tempted to use the roots of p as approximations to eigenvalues. This is the idea behind using the Arnoldi method to approximate eigenvalues.

The Arnoldi minimization problem is essentially a matrix function problem. Utilizing diagonalizability or pseudospectra, one may bound the Arnoldi polynomial by a scalar problem. Assuming A is diagonalizable, $A = X\Lambda X^{-1}$,

$$\|p(A)v\| = \|Xp(\Lambda)X^{-1}v\| \leq \|X\|\|X^{-1}\| \max_{\lambda\in\sigma(A)} |p(\lambda)|\|v\|.$$

Assuming $\|v\| = 1$, this gives rise to the min-max problem

$$\min_{p\in\mathbb{P}^k} \|p(A)v\| \leq \kappa(X) \min_{p\in\mathbb{P}^k} \max_{\lambda\in\sigma(A)} |p(\lambda)|,$$

where $\kappa(X) = \|X\|\|X^{-1}\|$, the condition number of X. Similarly, a pseudospectral approach gives

$$\min_{p\in\mathbb{P}^k} \|p(A)v\| \leq \frac{L_\varepsilon}{2\pi\varepsilon} \min_{p\in\mathbb{P}^k} \max_{\lambda\in\sigma_\varepsilon(A)} |p(\lambda)|.$$

As a Krylov subspace is larger than the span of any single power method iterate, $(A - \sigma I)^i v$, provided that the starting vector v is not an eigenvector, the Krylov subspace should offer a better approximation to desired eigenspaces than any single vector method. Subspaces are particularly useful for eigenvalue estimation when the size of the matrix prohibits dense eigenvalue methods, and also when the number of eigenvalues desired makes single vector iteration/deflation techniques impractical.

1.5 The Arnoldi Method

The Arnoldi method determines Ritz values by generating an orthonormal basis for the Krylov subspace. The Arnoldi method was introduced in the early 1950s by W. E. Arnoldi [3] as a generalization of the Lanczos method for Hermitian matrices [34]. Though both Lanczos and Arnoldi recognized that their methods had some iterative potential, at the time they were proposed, both methods were seen primarily as ways of reducing a matrix to tridiagonal or upper Hessenberg form by a unitary similarity

transformation [44]. In the Hermitian case, for which the eigenvalues all fall on the real line, in their most basic form Sturm sequence methods determine eigenvalues by repeatedly evaluating the characteristic polynomial, locating eigenvalues by noting sign changes. The idea of the Arnoldi and Lanczos methods was that the evaluation of the characteristic polynomial is much simpler for such structured matrices. However, numerical instabilities in the methods lead to inaccuracies in the reduced matrices that limit the ability to accurately determine eigenvalues. As a result, until the 1970s neither method received much attention as anything other than a procedure for reducing a matrix to tridiagonal or upper Hessenberg form (see, e.g., Wilkinson [60]).

The utility of the Arnoldi and Lanczos methods comes from their ability to generate accurate eigenvalue approximations from a partial, rather than full, upper Hessenberg factorization of a matrix. Both the Arnoldi and Lanczos methods work by generating an orthonormal basis for the Krylov subspace. The Lanczos method can be viewed as a special case of the Arnoldi method for $A = A^*$. Each step of the Arnoldi method generates a new basis vector v_i such that $\|v_i\| = 1$ and $v_i \perp \mathcal{K}_j(A, v)$ for $j < i$ and $v_i \in \mathcal{K}_i(A, v)$. At the kth step, the columns of the matrix $V_k = [v_1, \ldots, v_k]$ span the kth Krylov subspace. Based on these properties, the v_i must satisfy

$$Av_i = \sum_{j=1}^{i} h_{ji} v_j + h_{i+1,i} v_{i+1},$$

where $h_{ji} = v_j^* A v_i$. This recurrence is a consequence of the nesting of Krylov subspaces, $\mathcal{K}_k(A, v) \subseteq \mathcal{K}_{k+1}(A, v)$. Combining these equations into a matrix equation, A and V_k must satisfy

$$AV_k = V_k H_k + h_{k+1,k} v_{k+1} e_k^*, \tag{1.4}$$

where $[H_k]_{ij} = h_{ij}$ is an upper Hessenberg matrix. Premultiply equation (1.4) by V_k^*

to get

$$V_k^* A V_k = H_k,$$

so H_k is the orthogonal restriction of A onto the kth Krylov subspace. Similar to the optimality of the Ritz values, H_k is optimal in that it minimizes the norm of $AV_k - V_k H$ over all H [54, p. 254]. H_k is also optimal in that the characteristic polynomial of H minimizes $\|p(A)v\|$, i.e., p solves the Arnoldi minimization problem for $\mathcal{K}_k(A, v)$. If $h_{k+1,k} = 0$, the columns of V_k span an eigenspace of A. If $h_{k+1,k}$ is small relative to $\|A\|$, then the entire kth Krylov subspace accurately approximates an eigenspace of A. Even if $h_{k+1,k}$ is large, there may be Ritz pairs that are good approximations to eigenvalues. To see this, note that the residual for a Ritz pair, $x = V_k y$ and θ, obeys

$$\begin{aligned} Ax - x\theta &= AV_k y - V_k y\theta \\ &= V_k H_k y + h_{k+1,k} v_{k+1} e_k^* y - V_k y\theta \\ &= V_k y\theta + h_{k+1,k} v_{k+1} e_k^* y - V_k y\theta \\ &= h_{k+1,k} v_{k+1} e_k^* y, \end{aligned}$$

and hence $\|Ax - x\theta\| = |h_{k+1,k}||e_k^* y|$. Thus, if together the product of $|h_{k+1,k}|$ and the kth component of y are small, then (θ, x) is likely to be a good approximation to an eigenpair. The matter of how small $|h_{k+1,k}|$ must be to ensure that the Ritz values are good approximations to eigenvalues depends on the sensitivity of the spectrum to perturbations. The sensitivity to perturbations can be measured using the pseudospectra of the matrix, as described in Section 1.3.4.

1.6 The Restarted Arnoldi Method

The advantage of both the Arnoldi and Lanczos methods is that the Hessenberg factorization can be updated incrementally with the size of the Krylov subspace. The primary difficulty with these methods is maintaining the orthogonal basis. The cost of doing so increases steadily as the dimension of the subspace increases. The costs of maintaining orthogonality and of storing the basis vectors are the primary reasons why these methods must be restarted. Loss of orthogonality in the Hermitian case is particularly acute, as the Lanczos method works on the assumption that basis vectors satisfy a three-term recurrence, a huge computational advantage, as the method only stores three basis vectors at any one time. However, the three-term recurrence only holds in exact arithmetic. In floating point arithmetic, more must be done to maintain orthogonality. The problem also worsens when the subspace develops a good approximation of a particular eigenvalue. Without modification, the Lanczos method can lead to dubious eigenvalue estimates. In the 1970s, Paige and Parlett determined the necessary modifications to the Lanczos method for addressing the loss of orthogonality due to floating point arithmetic [39, 44]. The knowledge and insight developed for Lanczos set the stage for Saad to introduce the restarted Arnoldi method as a means of calculating a few eigenvalues of a non-Hermitian matrix in 1980 [46].

The impact of Saad's paper on the restarted Arnoldi method was threefold. First, Saad provided an alternative eigenvalue algorithm. At the time, subspace iteration (then also known as simultaneous iteration) was the prominent iterative method for determining several eigenvalues of a non-Hermitian matrix. Subspace iteration is a generalization of the power method to subspaces. Unlike the original Arnoldi method, which requires ever larger subspaces and hence more memory and computation to

improve eigenvalue approximations, the restarted Arnoldi method (like subspace iteration) works to refine an approximate eigenspace.

Second, Saad provided the first *a priori* convergence bounds for the approximation of eigenvectors of non-Hermitian matrices from a Krylov subspace. Assuming that the eigenvalues of A are simple and that the starting vector represented in a normalized eigenvector basis $\{x_j\}$ is $v = \sum_{j=1}^{N} \alpha_j x_j$, then provided that $\alpha_i \neq 0$, the norm of the residual of the projection of the eigenvector x_i onto the Krylov subspace is bounded as

$$\|(I - V_k V_k^*)x_i\| \leq \delta_i \min_{\substack{p \in \mathbb{P}_k \\ p(\lambda_i)=1}} \max_{\substack{j=1,2,\ldots,N \\ j \neq i}} |p(\lambda_j)|, \tag{1.5}$$

where $\delta_i = \sum_{j=1, j\neq i}^{N} |\alpha_j|/|\alpha_i|$. Though presented in terms of a projection, the bound gives a measure of the angle between an eigenvector and a Krylov subspace. The bound involves a min-max problem for determining a polynomial that is small on the unwanted eigenvalues. The δ_j factors capture the nonnormality of the matrix, as evident from the conditioning of the eigenvector basis. Assuming v is normalized, the α_j can be large if some of the eigenvectors x_j are nearly linearly dependent.

Last, Saad motivated the need for restarting the Arnoldi method and proposed a technique to do so. To restart the Arnoldi method, one must pick a new starting vector, v^+, for the Krylov subspace. Saad suggested that this new starting vector be a weighted linear combination of the Ritz vectors, with the weights chosen based on how well the Ritz vectors approximate eigenvectors. As every vector in a Krylov subspace can be represented as a polynomial in the matrix times the starting vector, Saad's approach to restarting is equivalent to selecting the roots, μ_i, of a polynomial,

$$\psi(z) = \prod_{i=1}^{p} (z - \mu_i).$$

Hence the new starting vector is

$$v^+ = \psi(A)v.$$

In the years following his introduction of the restarted Arnoldi method, Saad used the polynomial representation of vectors in a Krylov subspace to suggest that the restart polynomial, ψ, should be small on the unwanted portion of the spectrum. Such a choice will amplify the components of the starting vector in the direction of desired eigenvectors. In the case that the spectrum is real, if one can determine an interval containing only the unwanted eigenvalues, then a Chebyshev polynomial can be constructed to be uniformly small on the interval and large everywhere else [48]. Chebyshev polynomials are optimal for intervals; i.e., any other polynomial on the same interval would not be as uniformly small on the interval. For complex spectra, if there exists an ellipse containing the unwanted eigenvalues but not the desired ones, then a Chebyshev polynomial can be constructed to be small on the ellipse. In this case the Chebyshev polynomial is asymptotically optimal for the ellipse: as the polynomial degree increases, the Chebyshev polynomial will improve asymptotically at the same rate as the optimal polynomials for the ellipse. In practice there may not exist an ellipse that contains only the unwanted eigenvalues. In this case, the problem of constructing a polynomial that is small on a region containing the unwanted eigenvalues becomes difficult; one could use another method such as conformal mapping to construct the restart polynomial [27, 53]. The main difficulty in this approach to constructing restart polynomials is estimating the location of the unwanted eigenvalues. The goal of the Arnoldi method is to compute solely the wanted eigenvalues, but in following the recipe above, one has to determine estimates of the unwanted eigenvalues as well. This problem remains relevant in analysis and application of the Arnoldi method. My work is partly focused on clarifying how reliably Ritz values can

be used to approximate the unwanted portion of the spectrum.

1.7 Implicitly Restarted Arnoldi with Exact Shifts

A decade after Saad introduced the Arnoldi method, Sorensen [52] formulated the implicitly restarted Arnoldi method. Explicit restarting of the method involves either directly applying the restart polynomial to the starting vector to generate the starting vector for the next iteration, or building v^+ as a linear combination of Ritz vectors. The new starting vector is then used to generate a basis for the new Krylov subspace, as well as the projection of the matrix onto that subspace. In floating point arithmetic, explicit restarting is numerically unstable. The direct application of the matrix polynomial to a vector can lead to rounding errors. By interpreting the Arnoldi method as a truncated version of the QR eigenvalue iteration, Sorensen developed a numerically stable method of implicitly applying the restart polynomial using the tools and concepts from the QR eigenvalue iteration, including implicit shifting and deflation. In addition to putting the restarted Arnoldi method into a reliable numerical form, Sorensen proposed a method for picking the roots of the restart polynomial, and showed that this approach, under mild assumptions, gives a convergent algorithm for Hermitian matrices.

To determine a restart polynomial, one must have some knowledge of the location of wanted and unwanted eigenvalues. Unless one has some prior knowledge of the system, this information has to be determined adaptively as the algorithm proceeds. With prior knowledge one can construct a fixed restart polynomial. In this case, the Arnoldi method is similar to applying the power method with a fixed polynomial. Not surprisingly, Sorensen showed that the convergence criteria for Arnoldi with a fixed restart polynomial is similar to the convergence criteria for the power method.

Fixed restart polynomials are rarely used in practice, but they are useful in theory for establishing convergence bounds. For a more potent practical algorithm, Sorensen proposed using some of the Ritz values as the roots of the restart polynomial. Depending upon the type of eigenvalue desired (largest/smallest magnitude, real part or imaginary part), the Ritz values are sorted, and a fixed number of the least desirable Ritz values at each iteration are used as roots for the restart polynomial. These Ritz values are referred to as *exact shifts*. Using Cauchy interlacing, Sorensen showed that for Hermitian matrices the implicitly restarted Arnoldi method with exact shifts would converge. As Hermitian matrices are encountered in numerous applications, Sorensen's proof validated the utility of the implicitly restarted Arnoldi method for determining eigenvalues of Hermitian matrices.

1.8 Convergence of Restarted Arnoldi for Non-Hermitian Matrices

Non-Hermitian eigenvalue problems arise frequently, and the use of the implicitly restarted Arnoldi method with exact shifts to solve them is common practice. Though the behavior of Ritz values for non-Hermitian problems is poorly understood, exact shifts perform well in practice. This thesis will identify and characterize the Ritz values of particular non-Hermitian matrices for which the restarted Arnoldi method will converge. This is a difficult problem because, for one, there are matrices and starting vectors for which the method will fail to converge in exact arithmetic; the restart polynomial annihilates the components of the starting vector in the direction of desired eigenspaces [20]. Matrices that allow for this type of failure are characterized by having wanted eigenvalues that lie in the field of values of the portion of the

matrix associated with the unwanted eigenvalues. Understanding Ritz value behavior alone is not sufficient to establish convergence, as the starting vector must be properly oriented to guarantee convergence. Even for the case of normal non-Hermitian matrices with perfectly conditioned eigenvalues, little is known about what is necessary for convergence.

Various lines of research have developed for understanding convergence of the restarted Arnoldi method. The most direct attacks on the problem focus on bounding convergence of the method using optimal shifts for some region of the complex plane containing the unwanted eigenvalues but not the desired eigenvalues [6, 7]. The convergence of the restarted Arnoldi method is best measured by calculating the *containment gap*, which is the sine of the largest canonical angle between the current subspace and the desired subspace:

$$\delta(\mathcal{W}, \mathcal{V}) = \max_{w \in \mathcal{W}} \min_{v \in \mathcal{V}} \frac{\|v - w\|}{\|w\|},$$

where typically $\mathcal{W}$ will be an invariant subspace and $\mathcal{V}$ some approximating subspace (possibly of unequal dimensions). Often the analysis of the Arnoldi method is concerned with approximating a single eigenvector w, in which case the containment gap reduces to

$$\delta(w, \mathcal{V}) = \frac{\|(I - P_{\mathcal{V}})w\|}{\|w\|},$$

where $P_{\mathcal{V}}$ is the orthogonal projector onto the subspace $\mathcal{V}$. In this light, Saad's original work for Arnoldi bounds precisely the containment gap; see equation (1.5). These approaches involve seeking polynomials that are small not only on the spectrum but also on larger sets that contain the spectrum, such as a pseudospectrum.

The main result [7] on the convergence of restarted Arnoldi bounds the containment gap between a desired invariant subspace $\mathcal{U}_g$ and the restarted Krylov subspace

as

$$\delta(\mathcal{U}_g, K_k(A, \Psi(A)v)) \leq C_1(A, v)C_2(A, \Omega_b) \max_{z \in \Omega_b} |1 - \Psi(z)\alpha_g(z)|.$$

The first term $C_1(A, v)$ accounts for the starting vector v. The second term $C_2(A, \Omega_b)$ accounts for the nonnormality of A associated with the unwanted eigenvalues contained in the complex set Ω_b. The last term, where Ψ is product of all the restart polynomials and α_g is a polynomial with roots at the wanted eigenvalues, captures the convergence behavior associated with constructing restart polynomials that are small on a set containing the unwanted portion of the spectrum and yet large on the wanted portion of the spectrum. Using potential theory, the last term can be shown to decay like ρ^k for some $\rho \in (0, 1)$. The disadvantage of this approach is that, though it does bound the rate of convergence and essentially identifies what would be ideal behavior for exact shifts, it does not provide insight into what is necessary for such ideal behavior to occur with exact shifts in practice. By localizing Ritz values, this thesis provides insight into why exact shifts should exhibit such behavior.

1.9 The Inverse Field of Value Problem

In 2008 Frank Uhlig [59] offered an interesting approach to studying Ritz values. Uhlig posed the inverse field of values problem (iFOV): for $A \in \mathbb{C}^{n \times n}$ and $z \in W(A)$ construct a unit vector $v \in \mathbb{C}^n$ such that $z = v^* A v$. Due to the convexity of the field of values, the problem is always solvable, and Uhlig offered a complicated algorithm for solving it based upon concepts from geometric computing. In this work we provide a simpler algorithm for solving the inverse field of values problem, and then propose a generalization that is much more relevant to understanding Ritz values.

So little is known about the behavior of Ritz values for non-Hermitian matrices that much effort must be invested simply in characterizing those $\{\theta_1, \ldots, \theta_k\} \subset W(A)$

that may be observed via Arnoldi eigenvalue estimates or any orthogonal restriction of a matrix. Such questions have the flavor of Uhlig's iFOV, and hence lead to what we call the inverse field of values problem for k Ritz values (iFOV-k): for $A \in \mathbb{C}^{n\times n}$ and $\{z_1, \ldots, z_k\} \subset W(A)$, is it possible to construct a $V \in \mathbb{C}^{n\times k}$ with $V^*V = I$ such that $\sigma(V^*AV) = \{z_1, \ldots, z_k\}$? iFOV-$k$ is much more difficult than Uhlig's problem: simply characterizing when the problem is solvable is nontrivial. Much of the work in this thesis consists of solving iFOV-k for particular matrices, and illustrating criteria that sets of complex values must satisfy to be a set of Ritz values.

Other authors have made different strides in understanding Ritz values. Inspired by results from Lie algebra, Shomron, Parlett and Strang [42, 50] consider the set of matrices having the same Ritz values, where by Ritz values they mean the eigenvalues of all *principal submatrices*, the n square matrices formed from columns $1, \ldots, i$ for $i = 1, \ldots, n$. Parlett and Strang's Ritz values consist of $n(n+1)/2$ complex numbers. One feature of their work is a parametrization of all, not necessarily Hessenberg, matrices that share the same Ritz values. They discussed Lie algebra methods for moving between these different matrices; however, as shown by Smit [51], if we are only interested in Arnoldi and Hessenberg matrices, then diagonal similarity is all that is needed to generate all Hessenberg matrices that share the same Ritz values. Tebbens and Meurant [19] noted that, as a consequence of these results, the convergence of the Arnoldi method for eigenvalues can be arbitrary: there exist $A \in \mathbb{C}^{n\times n}$ for which the Ritz values, eigenvalues of H_k, bear no resemblance to the eigenvalues of A up until the nth step of the Arnoldi method. This is similar to a result of Greenbaum, Pták and Strakoš [22], which says that any residual curve for GMRES is possible wfor a matrix with any specified spectrum. The primary use of such results is theoretical; they tell us about what is possible for all matrices, rather than for any given A that

we wish to study. The problem of prescribing Ritz values is much harder when the Ritz values must come from a particular matrix, as in iFOV-k. As these results are all constructive, they can be useful for determining necessary conditions for restarted Arnoldi convergence, through the construction of specific examples.

In this thesis I will solve completely iFOV-2 for a 3×3 Jordan block and for all 3×3 non-Hermitian normal matrices. I will show that even for the most nonnormal matrix, a Jordan block, the Ritz values can be "localized", in the sense that repeated Ritz values cannot occur throughout the entire field of values. I will show how this localization of Ritz values can be generalized to all matrices. For the 3×3 non-Hermitian normal matrix, I show that using fundamental matrix properties one can significantly restrict Ritz values for general normal matrices.

1.10 Ritz Values and Restarted Arnoldi Convergence

This thesis will identify criteria that give rise to localization of Ritz values, particularly Arnoldi Ritz values, for non-Hermitian matrices. In the Hermitian case, the classical interlacing result of Cauchy [44, §10.1] restricts the location of Ritz values with respect to the spectrum. The optimality of Arnoldi Ritz values gives rise to the separation of Ritz values by eigenvalues. These properties restrict the set of possible Arnoldi factorizations. Toward establishing similar properties for the non-Hermitian case, this thesis utilizes criteria for the solvability of iFOV-k to characterize the Ritz values.

The results of this thesis will allow for further analysis of methods that rely upon Ritz values for eigenvalue approximation. A better understanding of the behavior of Ritz values for non-Hermitian matrices can potentially aid in the analysis of deflated and augmented Krylov techniques, such as Morgan's GMRES-DR algorithm [14, 41]. These methods use information about eigenspaces derived from Ritz pairs to improve

the rate of convergence. In the standard restarted GMRES algorithm, at each restart information associated with certain eigenvalues (such as those closest to the origin) must be rediscovered before the algorithm can continue to make progress [49]. By supplementing the method with Ritz value information from previous steps, the time spent rediscovering the troublesome eigenvalues can be minimized. The question is, How well can one expect the Ritz values to approximate these eigenvalues?

This thesis will also establish criteria that are sufficient for convergence of the restarted Arnoldi method for certain scenarios in which the wanted eigenvalues are not in the field of values of the portion of the matrix associated with the unwanted eigenvalues. For this class of matrices, the type of failure demonstrated by Embree cannot occur [20]. The results of this thesis will also provide a different criteria for ruling out this type of failure.

Without loss of generality, the possible Ritz values will be characterized using the Schur decomposition, an invaluable tool in understanding many different eigenvalue problems. In 2001, Stewart generalized the notion of an Arnoldi decomposition, introducing what he called *Krylov decompositions* [55]. With this generalization and the Schur decomposition, he introduced a Krylov–Schur algorithm for determining eigenvalues. The algorithm is equivalent to the Arnoldi method, but the resulting factorization is upper triangular rather than upper Hessenberg, and allows for a simpler application of the exact-shift restart polynomial.

To the set of possible Krylov–Schur factorizations for a particular matrix there corresponds the set of equivalent matrices that can generate the same factorizations. Understanding the behavior of Ritz values for Arnoldi is largely about understanding the possible Krylov–Schur factorizations for a given matrix. However, Tebbens and Meurant [19] have shown that any sequence of Ritz values can be realized by

Arnoldi for some matrix. Thus in this thesis care is taken in utilizing this approach by respecting aspects of the original matrix: the field of values, pseudospectra or normality.

1.11 Summary of Contributions

The contributions of this thesis are:

1. an improved algorithm [12] for the inverse field of values problem (Chapter 2);
2. proof that the covering number of a point $z \in W(A)$ is always n;
3. bounds for the Ritz values of general matrices (Section 3.1);
4. a complete description of the Ritz values for a 3×3 Jordan block (Section 3.2);
5. a complete description of the Ritz values for a 3×3 normal non-Hermitian matrix (Section 3.3);
6. sufficient conditions for the convergence of the restarted Arnoldi method based on the bounds for Ritz values derived in Section 3.1 (Chapter 4).

Chapter 2

The Inverse Field of Values Problem

Given $z \in W(A)$, a unit vector x such that $z = x^*Ax$ is a *Ritz vector* for z. The point z and a corresponding Ritz vector x are referred to as a Ritz value/Ritz vector pair. A preliminary step toward understanding how Ritz values must distribute themselves throughout $W(A)$ involves the inverse field of values problem (iFOV) recently proposed by Uhlig [59]: given a point $z \in W(A)$, determine a unit vector x such that $z = x^*Ax$. Due to the convexity of the field of values, the inverse field of values problem is always solvable. To solve such problems, Uhlig proposed an algorithm that begins by generating points in the field of values from random unit vectors, in the hope that these Ritz values surround the desired z. Uhlig's algorithm then iteratively refines these preliminary Ritz vectors to better approximate a Ritz vector for z. Due to the difficulty of encircling the desired point using Ritz values generated by random vectors, the algorithm does not always converge to the desired accuracy, as shown in Figure 2.5. In posing the inverse field of values problem, Uhlig hoped to derive some insight into the structure of the field of values using the notion of the *covering number*, the number of linearly independent Ritz vectors for a particular point in the field of values. Uhlig conjectured that the covering number of any point in the interior of the field of values for $A \in \mathbb{C}^{n\times n}$ must be n.

In this chapter I present a simple algorithm for solving the inverse field of values problem, work that appeared in [12]. My algorithm is inspired by the standard algorithm used to determine the boundary of the field of values of a matrix [32]. I

prove that indeed the covering number is n for any point in the interior of the field of values. Finally, I propose a more general inverse field of values problem that involves the construction of subspaces of dimension greater than one that yield Ritz values corresponding to several points in $W(A)$. This problem provides a context for the analysis in the next chapter.

2.1 Main Ideas

In this section I state two results that are essential for the algorithm that follows. In two scenarios the inverse field of values problem can be solved immediately: (1) two dimensional matrices; (2) the point z is in the convex hull of three Ritz values, z_i, for which the corresponding Ritz vectors, x_i, are known. In both scenarios a Ritz vector x for z can be computed exactly. Both results follow directly from the argument used to prove that the field of values is convex [31, Ch.1].

The solution for two dimensional matrices follows from the proof of Lemma 1.3.1 in [31]. Without loss of generality, one may assume that the trace of A is zero. (If the trace were nonzero one would need only subtract $(\mathrm{tr}(A)/2)I$ from A and $\mathrm{tr}(A)/2$ from z.) Furthermore, one may assume that eigenvalues of A are real. (To accomplish this, one need only multiply both A and z by $e^{-i\psi}$ where the eigenvalues of A are $\pm|\lambda|e^{i\psi}$.) With this assumption, the eigenvalues sum to zero, and may be denoted as $+\lambda$ and $-\lambda$. As in [31, §1.3], one may determine a unitary matrix Q such that

$$Q^*AQ = \widehat{A} = \begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix}, \tag{2.1}$$

where α and β are both real and nonnegative and $\alpha \geq \beta$. Note that z is unchanged, as the field of values is invariant under unitary similarity transformations. As any

2×2 matrix can be shifted, scaled and unitarily transformed into this form, if we can solve the inverse field of values problem for $\widehat{A}$, then we are done.

Without loss of generality, represent the unit vector $x \in \mathbb{C}^2$ as $x = [\cos(\theta); e^{i\phi}\sin(\theta)]$, giving

$$x^* \widehat{A} x = \frac{\sin(2\theta)}{2}((\alpha+\beta)\cos(\phi) + (\alpha-\beta)i\sin(\phi)),$$

from which one can see that the field of values of $\widehat{A}$ is an ellipse in the complex plane centered at the origin with major axis $\alpha+\beta$ and minor axis $\alpha-\beta$, see Figure 2.1. Given $z \in W(\widehat{A})$, the θ and ϕ that specify the Ritz vector x must satisfy

$$\tan(\phi) = \frac{(\alpha+\beta)\mathrm{Im}(z)}{(\alpha-\beta)\mathrm{Re}(z)}, \qquad \sin(2\theta) = \frac{2|z|}{|(\alpha+\beta)\cos(\phi) + (\alpha-\beta)\sin(\phi)|}.$$

In terms of the coordinates for A, the Ritz vector is Qx. Thus the inverse field of values problem is solved for two-dimensional matrices. Note that the solution need not be unique, i.e., for z in the interior of $W(A)$, there exist linearly independent Ritz vectors, derived from θ and $\pi/2 - \theta$ for $\theta \neq \pi/4$. (When $\theta = \pi/4$, z is a point on the boundary, and its Ritz vector is unique provided that the ellipse is nondegenerate. If $\alpha = \beta$, then the field of values of A is a line segment connecting the eigenvalues, and the solution is unique only for the end-points, which must be the eigenvalues. If $\alpha = \beta = 0$, then A is the zero matrix, so the field of values is just a point and any unit vector $x \in \mathbb{C}^2$ will do.)

The procedure above offers a straightforward solution to solving iFOV for a 2×2 matrix. The approach also offers insight as to the uniqueness of such solutions. Other authors [16, 40] have usedd a similar method, also see [31, p. 25, Problem 10]. The approach avoids the reduction to the form (2.1), which would be an expensive reduction were we not dealing with 2×2 matrices. The idea is as follows, using the reasoning above for A and $\widehat{A}$. Let $A \in \mathbb{C}^{n\times n}$ with zero contained in the interior

of $W(A)$. Suppose that we have $x, y \in \mathbb{C}^n$ and $a, b \in \mathbb{R}$ such that $x^*Ax = a$ and $y^*Ay = b$, $a < 0 < b$. Suppose that we wish to construct a Ritz vector for 0; such a vector is known in the literature as an *isotropic vector*. Consider the vector valued function $v(t, \phi) = x + te^{i\phi}y$ for $t \in \mathbb{R}$ and $\phi \in [0, 2\pi]$. Then

$$\begin{aligned} v(t,\phi)^*Av(t,\phi) &= a + te^{-i\phi}x^*Ay + te^{i\phi}y^*Ax + t^2b \\ &= a + t(e^{-i\phi}x^*Ay + e^{i\phi}y^*A^*x) + te^{i\phi}(y^*Ax - y^*A^*x) + t^2b, \end{aligned}$$

where the coefficient of t is real if $\phi = -\arg(y^*Ax - y^*A^*x)$. For this choice of ϕ, $v(t, \phi)^*Av(t, \phi)$ is a parabola in t with real coefficents and unique roots, which may be used to determine a Ritz vector for 0 via $v(t, \phi)/\|v(t, \phi)\|$. Note that though this approach is formulated for a matrix of arbitrary size, it only makes use of knowledge of the action of the matrix upon the two dimensional subspace spanned by x and y.

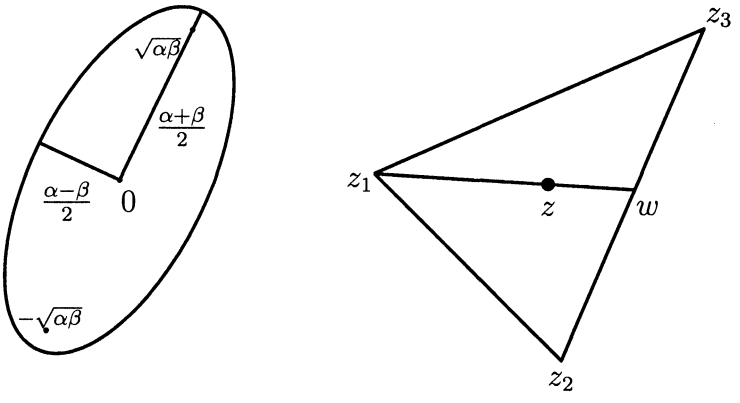


Figure 2.1 : On the left, an ellipse corresponding to the field of values of a general 2×2 matrix. On the right, an illustration of the procedure for determining a Ritz vector for a point located in the convex hull of three points in $W(A)$.

For the second scenario, we first recall why convexity of the field of values for 2×2 matrices implies the same for all larger matrices. Given any two unique points in the

field of values, z_1 and z_2, and their corresponding Ritz vectors, x_1 and x_2, one can construct a matrix $V \in \mathbb{C}^{n\times 2}$ with orthonormal columns such that $x_1, x_2 \in \text{Range}(V)$. Then the field of values of the restriction of A to Range(V), $H = V^*AV$, must be contained in the field of values of A. By construction, both $x_1, x_2 \subset \text{Range}(V)$, hence the field of values of H must contain both z_1 and z_2. Since the field of values of a 2×2 matrix is an ellipse and hence convex, $W(H)$ as well as $W(A)$ must contain all convex combinations of z_1 and z_2. Since z_1 and z_2 were arbitrary, $W(A)$ must be convex. Thus the solution to our first scenario indicates that if we are given two values z_1 and z_2 in the field of values of A, then we can determine vectors that generate any point that is a convex combination of z_1 and z_2.

Now suppose z is in the convex hull of three points, $z_i \in W(A)$, for which we have the corresponding Ritz vectors x_i. The convex hull of the three points is a (possibly degenerate) triangle. To apply the result for the 2×2 case, we would need z to fall on a segment with end-points for which we have — or can compute — corresponding Ritz vectors. Without loss of generality, take z_1 as one of the end-points for this proposed line segment. Let the other end-point of the line segment, w, be the intersection of the line through z_2 and z_3 with the line through z_1 and z. As w is a convex combination of z_2 and z_3, we can determine a corresponding Ritz vector for w. Now since z is a convex combination of w and z_1, we can determine a Ritz vector for z, as illustrated in Figure 2.1.

Given the results above, all that remains is to specify a way to generate a triangle in $W(A)$ that contains the desired value z. For this we consider Johnson's algorithm for the numerical determination of the field of values of a general complex matrix [31, 32]. Johnson's method constructs an approximation to the field of values by determining the largest and smallest eigenvalues of $H_\theta = (e^{i\theta}A + e^{-i\theta}A^*)/2$, the Hermitian part of

$e^{i\theta}A$, for various values of θ. The largest and smallest eigenvalues of H_θ provide sharp bounds on the extent of the field of values in the direction θ in the complex plane. These sharp bounds together form an *outer approximation* to the field of values, i.e., a polygon that contains $W(A)$ and whose sides intersect points on the boundary of $W(A)$. This outer approximation can be used to identify points that fall outside the field of values. The corresponding eigenvectors of H_θ generate points on the boundary of $W(A)$. These points, taken together, yield an *inner approximation* to $W(A)$; see Figure 2.2. By systematically selecting a series of angles θ, one can construct polygons that, to a desired accuracy, contain the desired value z. This is the substance of the algorithm we describe in Section 2.3.

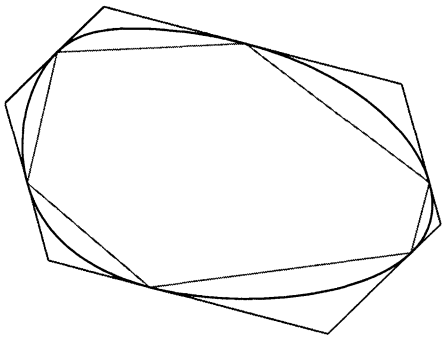


Figure 2.2 : Illustration of Johnson's algorithm for drawing the boundary of the fields of values. The boundary of the field of values (black line) and an inner and outer approximation (grey) of a 4×4 matrix from H_θ for $\theta = \{0, \pi/3, 2\pi/3\}$.

2.2 Covering number

Before detailing the algorithm, consider a conjecture posed by Uhlig [59].

Definition. For a matrix $A \in \mathbb{C}^{n\times n}$ *and a point* $z \in W(A) \subset \mathbb{C}$, *the maximal dimension of a linearly independent set of complex unit vectors* x *with* $x^*Ax = z$ *is the* **covering number** *of* (A, z).

Having shown that the Ritz vector that generates a point in the interior of $W(A)$ for $A \in \mathbb{C}^{2\times 2}$ is not unique, I have demonstrated that in the 2×2 case, the covering number of any point in the interior of $W(A)$ is two.

Uhlig introduced the covering number as a tool that might provide insight into the structure of the field of values, particularly for nonnormal matrices; however, based on numerical observations he conjectured that the covering number is n for all points in the interior of the field of values, which would imply that this initially appealing idea yields no further insight into the structure of the field of values. My algorithm can be used to prove that his conjecture indeed holds.

Theorem 2.1 Let $A \in \mathbb{C}^{n\times n}$ *and suppose* z *is in the interior of* $W(A)$. *(Note that if* $W(A)$ *is a line segment, then by interior we mean that* z *is not an end-point of the line segment.) Then the covering number of* (A, z) *is* n.

Proof. Consider any point z in the interior of $W(A)$. By definition, all points in $W(A)$ have at least one Ritz vector. Suppose that we have a set of $k < n$ linearly independent Ritz vectors for z. We will show that we can always find another linearly independent Ritz vector. As $k < n$, there exists a unit vector u that is orthogonal to the set of k linearly independent Ritz vectors for z. Assume that the point $a \in W(A)$ generated by u is not equal to z, for otherwise u would be a Ritz vector for z. Since z is in the interior of $W(A)$ and $W(A)$ is convex, there always exists some $b \in W(A)$, $b \neq z, a$, for which z is a convex combination of a and b. Since $b \in W(A)$, it must have some Ritz vector, $\widehat{u}$. Now z must be in the interior of the field of values of the

2×2 matrix corresponding to the restriction of A onto the span$\{u, \widehat{u}\}$. Thus from our argument in Section 2.1, we find that there exist two linearly independent Ritz vectors for z such that u is in their span. At least one of these vectors must have a component in u, thus giving an additional Ritz vector for z linearly independent from the rest. Hence we have at least $k+1$ linearly independent Ritz vectors for z. As $k < n$ and z were arbitrary, the covering number for any point in the interior of the field of values must be n. ■

Note that Uhlig's conjecture differs from the problem of determining how many distinct Ritz vectors generate a given $z \in W(A)$. There are generally at least $2n-4$ degrees of freedom available in specifying a Ritz vector for a given z. This may be seen by breaking up the Rayleigh quotient x^*Ax in terms of the real and imaginary parts of both A and x, and then requiring that x be a unit vector with its first nonzero entry being real and nonnegative. Hence there are in general infinitely many Ritz vectors for a given $z \in W(A)$. Theorem 2.1 shows that this set of infinitely many Ritz vectors always spans $\mathbb{C}^n$. Note however that the set of Ritz vectors for a point $z \in W(A)$ is not a subspace, i.e. a nontrivial linear combination of two Ritz vectors for z will in general not give another Ritz vector. Take for example the 2×2 case for a point in the interior. If the Ritz vectors for a point $z \in W(A)$ formed a subspace that would imply that the field of values consists of only z and $A = zI$.

2.3 Algorithm

Using the results above we can specify an algorithm for solving the inverse field of values problem. Given a matrix $A \in \mathbb{C}^{n \times n}$ and some $z \in \mathbb{C}$, let $\varepsilon > 0$ denote some tolerance (for example $\varepsilon = 10^{-16}\|A\|$ for a double precision computation).

I. Determine the leftmost and rightmost eigenvalues of H_θ for $\theta \in \{0, \pi/2\}$. The

resulting outer approximation will be a rectangle with sides parallel to the real and imaginary axes. If z is inside the outer approximation, then continue; otherwise conclude $z \notin W(A)$.

II. If the height or width of the outer approximation is less than ε, then $W(A)$ is approximately either Hermitian or skew-Hermitian (or a complex shift of one of these). If both the height and width of the outer approximation are less than ε, then the field of values is approximately a point. In either of these cases, one can easily determine a Ritz vector for z; otherwise, continue.

III. Using eigenvectors for the leftmost and rightmost eigenvalues of H_θ for $\theta = \{0, \pi/2\}$, determine an inner approximation to $W(A)$. The inner approximation will be a quadrilateral with vertices on the boundary of $W(A)$.

IV. Check if z lies in the inner approximation. If so, determine a vector that exactly generates z as described in Section 2.1, and successfully terminate. Otherwise, determine the side of the inner approximation to which z is closest.

V. Compute $\widehat{z}$, the closest point to z from the inner approximation. If $|\widehat{z} - z| < \varepsilon$, then compute the corresponding Ritz vector for $\widehat{z}$, accept this vector as a suitable Ritz vector for z, and successfully terminate.

VI. Update the inner and outer approximations by determining the largest eigenvalue and corresponding eigenvector of H_θ, where θ is the direction perpendicular to the side of the inner approximation to which z is closest. If this does not produce a point on the boundary of $W(A)$ that is not contained in the previous inner approximation, then $z \notin W(A)$.

VII. Check if z is in the new outer approximation. If so, go to step IV; otherwise, $z \notin W(A)$.

Figure 2.3 gives an example of this algorithm converging in one iteration. Figure 2.4 gives a similar example requiring four iterations.

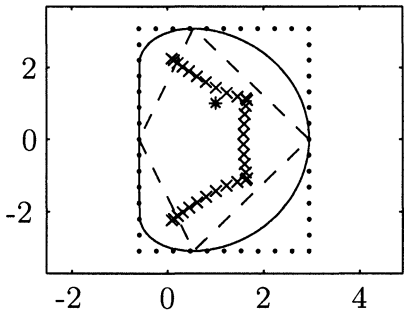
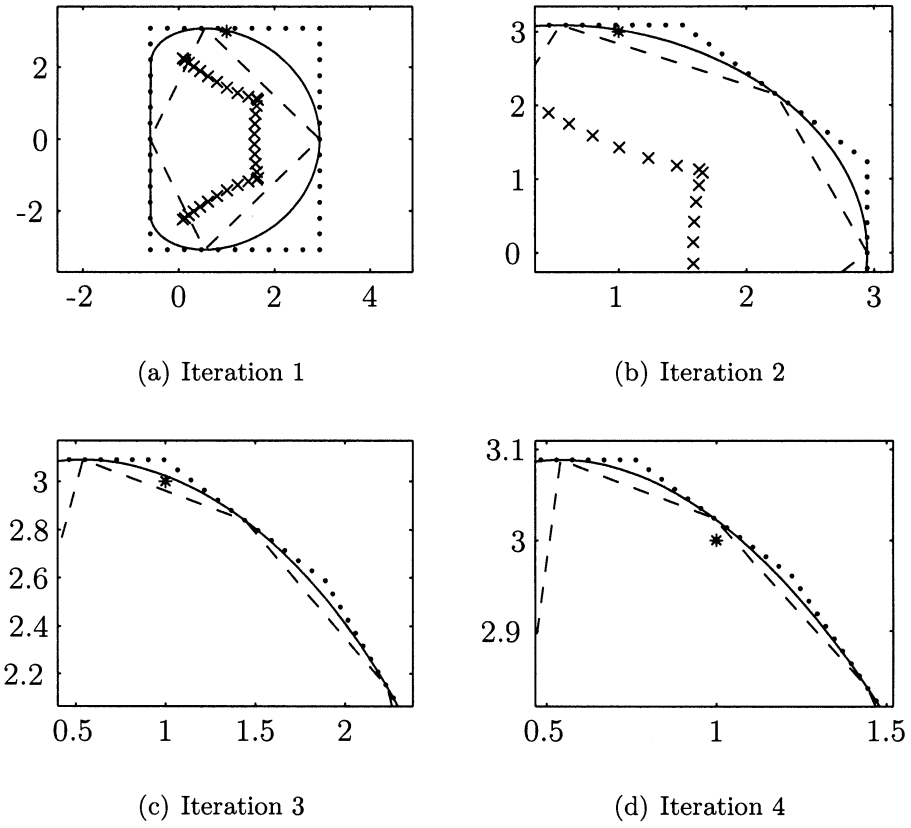


Figure 2.3 : Eigenvalues ($\times$), field of values boundary (solid line) and first inner and outer polygonal approximations (dashed, dotted) for Matlab's Grcar matrix (`gallery('grcar',32)`). In this case z (denoted by the $*$) falls inside the inner approximation, and an exact Ritz vector can be determined immediately.

Note that steps IV–VII are repeated until the inner approximation is within ε of z, or until the outer approximation does not contain z. This algorithm only updates the approximation to the field of values in the direction of z. To determine if z lies inside a particular polygonal approximation, there exist many efficient algorithms from computational geometry; see, e.g., [17]. In most cases the algorithm will terminate with an exact Ritz vector in step IV after only a few eigenvalue computations.

The algorithm above does not utilize the fact that the field of values of a 2×2 matrix is an ellipse. This property suggests that with the points and the corresponding Ritz vectors for the inner approximation, one may be able to exactly determine Ritz vectors for some points outside the inner approximation. This fact can be utilized with a modification of step V.



(a) Iteration 1

(b) Iteration 2

(c) Iteration 3

(d) Iteration 4

Figure 2.4 : Eigenvalues ($\times$), field of values boundary (solid line) and first inner and outer polygonal approximations (dashed, dotted) for four iterations of the algorithm for Matlab's Grcar matrix (`gallery('grcar',32)`) for $z = 1 + 3i$. At the fourth iteration, z (denoted by $*$) is in the inner approximation and a Ritz vector that exactly gives z can be determined.

V(a) Check if z is in the ellipse corresponding to the field of values of the restriction of A to the subspace spanned by the two Ritz vectors associated with the end points of the side of the inner approximation closest to z. If so, determine a vector that exactly generates z. Otherwise, compute $\widehat{z}$, the closest point to z from the ellipse. If $|\widehat{z} - z| < \varepsilon$, then compute the corresponding Ritz vector for $\widehat{z}$, accept it as a suitable Ritz vector for z, and successfully terminate.

Determining the closest point from an ellipse, a problem first solved by Greek geometer Apollonius of Perga, involves solving a quartic polynomial [25]. Depending on the nature of the ellipse, this modification could in some case reduce the number of eigenvalues computations required by one.

2.4 Discussion

The proposed algorithm differs from Uhlig's in several ways. Our algorithm is nearly deterministic; the only lack of determinism occurs if the extreme eigenvalues of H_θ have multiplicity greater than one, which is only possible when the boundary of the field of values contains a line segment. Otherwise, for a given z the algorithm will always return the same Ritz vector. A method for finding other Ritz vectors is suggested in the proof of Theorem 2.1. At the cost of determinism, Uhlig attempts to minimize the number of eigenvalue computations required by using random vectors to generate points in the field of values. If the randomly generated points sufficiently surround the desired point, then Uhlig's algorithm refines a quadrilateral that contains the desired point. Using great unit circles in $\mathbb{C}^n$ that contain pairwise Ritz vectors associated with the vertices of the quadrilateral, he refines his quadrilateral using the images of these great circles under the map $x \mapsto x^*Ax$, which happen to be ellipses. Points on these ellipses are determined numerically using equally spaced vectors on the great circle. I have demonstrated how one could use the Ritz vectors for the quadrilateral to exactly solve the problem, rather than iterate as in Uhlig's approach. We utilize ellipses associated with subspaces spanned by pairs of Ritz vectors. For these ellipses we can exactly generate any point in the interior as well as on the boundary; thus we are able to determine Ritz vectors for any point in the quadrilateral. Were Uhlig's algorithm to take full advantage of these ellipses, it could

potentially be faster and more accurate, particularly for points not near the boundary of the field of values.

As the Rayleigh quotients associated with random vectors tend to lie in the interior of $W(A)$, Uhlig's algorithm has difficulty finding Ritz values that surround z when z is near the boundary of $W(A)$. Also, as Uhlig observed, this problem becomes more acute for higher dimensions. If the randomly generated points do not sufficiently surround the desired point, Uhlig's algorithm utilizes the eigenvectors of H_θ for no more than six distinct θ. The θ are chosen in much the same way that our algorithm determines a polygon that contains the desired point. Uhlig's concern in limiting the number of eigenvalue computations that must be performed is that such calculations require $O(n^3)$ operations, whereas the refinement process only requires matrix vector operations of complexity $O(n^2)$. Our algorithm performs a minimum of two eigenvalue computation and could potentially perform more than six, depending upon the nature of the boundary of $W(A)$. To address this concern, in our implementation of the algorithm we allow for the use of Matlab's sparse eigensolver, `eigs` (which uses the ARPACK software [35]). In this case, for $\theta \in \{0, \pi/2\}$, a total of four calls must be made to `eigs` to determine the extreme eigenvalues of the H_θ. For additional angles utilized in step VI, only one call to `eigs` is necessary, as only the largest real eigenvalue is required. The advantage of using random vectors versus eigenvalue computations to generate a quadrilateral that contains the desired point will vary with the size of the matrix and the distance of z from the boundary of $W(A)$. Even with the concern for the number of eigenvalue computations required, our algorithm, using Matlab's dense eigensolver `eig`, often performed much faster for points throughout the majority of $W(A)$ for A of a variety of dimensions. Implementations of Uhlig's method and our algorithm are publicly available [10, 58].

Since the algorithm in Section 2.3 was published in 2010 [12], several other authors have refined this approach to yield slight improvements in certain situations. All these other algorithms follow from my observation that the inverse field of value problem can be solved exactly in certain scenarios. As mentioned in Section 2.1, other authors [16, 40] have made some savings in computation by not resorting to the form in equation (2.1). The savings are only significant when working with small matrices. These same authors have also attempted to minimize the number of eigenvalue computations performed. Depending on the matrix and the desired Ritz value, when successful, such attempts can solve the problem as accurately and as quickly as my algorithm. However there are always situations, particularly with Ritz values near the boundary, in which several eigenvalue computations cannot be avoided.

A comparison of run-times and errors for a Jordan block and a circulant shift,

$$\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 1 & & & 0 \end{pmatrix},$$

are shown in Figure 2.5. These matrices were chosen to illustrate extremes: a matrix with a large departure from normality, and a normal matrix for which the boundary of $W(A)$ contains line segments. For the Jordan block, one can see that with the exception of points on the boundary, our algorithm is nearly an order of magnitude faster for points throughout the field of values. For the circulant matrix, our algorithm clearly performs faster for all points not near the boundary. The error plots in Figure 2.5(e) shows that for Uhlig's algorithm, the error (the difference between the Rayleigh quotient from the computed Ritz vector and the desired z) varies over ten

orders of magnitude throughout the field of values, with the majority of the errors of the order 10^{-12}. The errors for my algorithm as shown in Figure 2.5(f) are much more uniform, varying only an order of magnitude about 10^{-14}. (These computations set the tolerance of our algorithm to 10^{-8} and use a built-in tolerance for Uhlig's of 10^{-12}.) These results are typical of a range of test cases we have investigated. The patterns seen for my algorithm are as one would expect: run-time depends on the number of eigenvalue computations performed, with points near the boundary taking longer; errors are more or less uniform with some loss of accuracy due to small angles encountered for points near the boundary of the inner approximation. The patterns for Uhlig's algorithm are completely unexpected, and as of yet, unexplained. In Figure 2.5(a), there are regions for which it takes longer to compute a Ritz vector; points in the lower right quadrant inexplicably require more time. There are similar features, though not as pronounced, in Figures 2.5(c) and 2.5(e).

2.5 A More General Inverse Problem

The problem we have just solved suggest a much more challenging and interesting generalization. I propose an inverse field of values problem, iFOV-k, involving multiple Ritz values:

> Given $z_1, \ldots, z_k \in W(A)$, does there exist $V \in \mathbb{C}^{n\times k}$, $V^*V = I$, such that the spectrum of V^*AV is $\{z_1, \ldots, z_k\}$, and if so, find such a V.

Note that the z_i need not be distinct, e.g, one can take $z_1 = z_2$. The algorithm in this chapter is concerned with iFOV-1. For iFOV-n, the only possible set $\{z_i\}$ is the spectrum of A. Even when $k = 2$, there exist matrices A and pairs $\{z_1, z_2\} \subset W(A)$ so that no such V exists (e.g., as can be deduced from the Courant–Fischer min-max

theorem [30]). The solution to this problem for general k would allow one to study how Ritz values distribute themselves throughout the field of values. Such knowledge would prove useful for analyzing eigenvalue algorithms such as the restarted Arnoldi method, and could suggest a kind of generalization of the Cauchy interlacing theorem for non-Hermitian matrices.

The solution of iFOV-1 might lead one to believe that iFOV-k can be solved using a combination of deflation and the solution of k iFOV-1 problems for each z_i. To demonstrate that this is not the case, consider the $n = 4$ circulant shift matrix

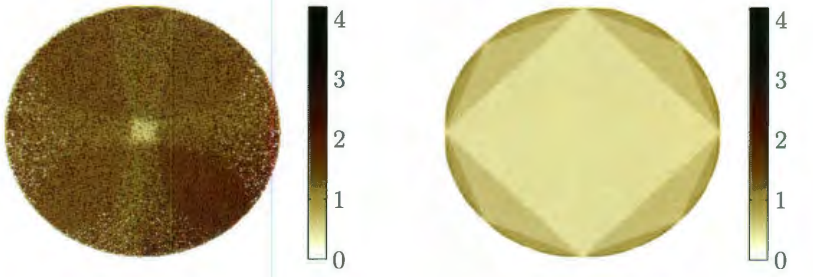
$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{2.2}$$

Suppose we wish to solve iFOV-2 for $z_1 = -3. + .3i$ and $z_2 = .2 + .3i$. To do so, we propose to first determine a Ritz vector, v_1, for z_1. Thoughts of deflation would tempt one to find v_2 for z_2 – if such a vector exists – in the subspace orthogonal to v_1, however this would not guarantee that z_1 is an eigenvalue of the restriction, H, of A onto Range($[v_1, v_2]$). Rather, it would only guarantee that $z_1, z \in W(H)$. Recall that for z_1 to be an eigenvalue of H we must require that $r = Av_1 - \theta v_1$ be orthogonal to v_1 as well as v_2, see Chapter 1 Section 1.2. Thus, to determine a Ritz vector v_2 for z_2 we consider the subspace orthogonal to v_1 and Av_1. Forcing v_2 to lie in this subspace amounts to solving iFOV-1 for A restricted to Null($[v_1\ Av_1]^*$). If such a v_2 exists, then restriction of A to the subspace spanned by v_1 and v_2 will have as its Ritz values z_1 and z_2. Figure 2.6 illustrate a success and a failure of this approach to the problem. The main issue is that for a given $z \in W(A)$, there exist numerous Ritz vectors v; however, the field of values of the restriction of A to the subspace

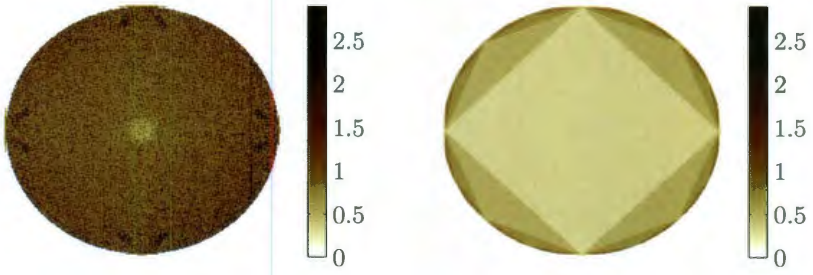
orthogonal to $[v\ Av]$ can vary significantly, depending on the choice of v.

A similar generalization of iFOV for joint numerical ranges is also not made any easier for having a solution for iFOV-1. Given $A_i \in \mathbb{C}^{n\times n}$ and $z_i \in W(A_i)$ for $i = 1,\ldots,k$ determine $x \in \mathbb{C}^n$ such that $z_i = x^*A_ix$. Determining when the restarted Generalized minimum Residual(GMRES) method stagnates is a problem of this sort [40]: If there exists an x such that $x^*A^ix = 0$ for $i = 1,\ldots k-1$, then GMRES restarted after k steps can stagnate. Through the use of compound matrices the iFOV-k problem may be written as a joint numerical range problem. There are two difficulties with solving these sort of problems. First, the x must simultaneously solve k iFOV-1 problems. Second, the sets $\{\{x^*A_ix\} : x \in \mathbb{C}^n\ \|x\| = 1\}$ are in general not convex. These issues make solving these problems extremely challenging.

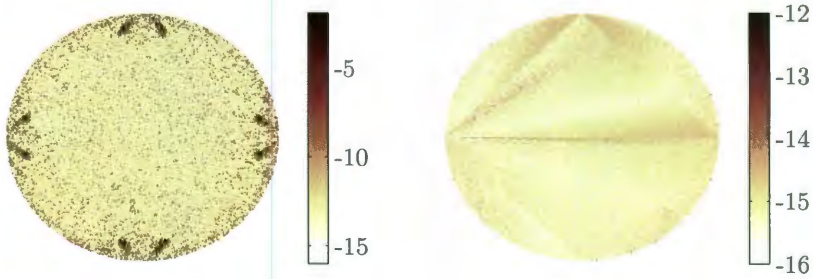
In the next chapter I will present general bounds on Ritz values that provide necessary conditions for iFOV-k to be solvable, then solve iFOV-k for two different $n = 3$ matrices, a Jordan block and a normal matrix. From these solutions we will determine criteria that $z_1,\ldots,z_k$ must satisfy in general in order for iFOV-k to be solvable for a given A.



(a) Jordan block-Uhlig's Algorithm (b) Jordan block-Section 2.3 Algorithm



(c) circulant shift-Uhlig's Algorithm (d) circulant shift-Section 2.3 Algorithm



(e) error for Uhlig's Algorithm (f) error for Section 2.3 Algorithm

Figure 2.5 : Comparison of Uhlig's algorithm versus our method for two examples: a Jordan block and a circulant shift matrix, both of dimension $n = 240$. For those points, z on a 256×256 uniform grid for $[-1, 1] \times [-1, 1]$ that fall in $W(A)$, plots (a) &(b) and (c) & (d) show run times in seconds. Plots (e)&(f) show $\log_{10}$ of the error in the calculated Ritz vector for z, $x^*Ax - z$, for the circulant shift matrix. From [12].

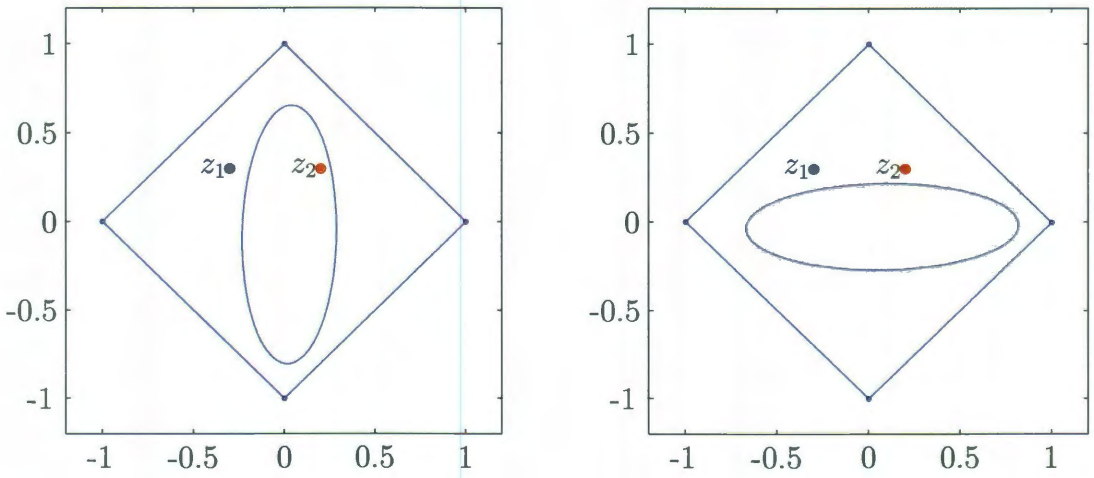


Figure 2.6 : Illustration of why iFOV-k cannot be solved using iFOV-1. For a $n = 4$ circulant shift matrix, equation (2.2), with $z_1 = -.3. + .3i$ and $z_2 = .2 + .3i$. On the left, the Ritz vector for z_1 determines a subspace for which iFOV-2 can be solved. On the right, z_2 does not lie in the field of values of the restriction of the subspace orthogonal to v and Av where v is a Ritz vector for z_1.

Chapter 3

Ritz Values for Non-Hermitian matrices

For Hermitian matrices, the Cauchy interlacing Theorem gives necessary and sufficient criteria for a set of points to be Ritz values from a k-dimensional subspace. Cauchy interlacing also ensures that Ritz values cannot cluster near the extreme eigenvalues. When using the restarted Arnoldi method to compute eigenvalues, having the shifts, Ritz values or other choices, well separated from these eigenvalues allows thee gap between the Krylov subspace and invariant subspaces associated with the extreme eigenvalues to go to zero with successive restarts. This is understood by constructing optimal polynomials on the remaining portion of the spectrum [6, 7, 8]. For restarted Arnoldi with exact shifts, Cauchy interlacing is also sufficient to guarantee the convergence of the Ritz values to the extreme eigenvalues [52]. The absence of interlacing for non-Hermitian matrices makes the development of a convergence theory for the restarted Arnoldi method difficult. Moreover, the presence of multiple Ritz values beyond the rightmost eigenvalue is essential to examples of extreme eigenvalue failure [20].

In this chapter, I demonstrate bounds for localizing Ritz values of non-Hermitian matrices. These bounds provide criteria for the solvability of iFOV-k . The results were motivated by efforts to understand iFOV-2 for two particular 3×3 matrices: a Jordan block and a normal non-Hermitian matrix. For such matrices I will completely characterize iFOV-2. For both types of matrices, Ritz values can be localized, i.e., the Ritz values cannot cluster throughout the entire field of values. These bounds address

the possibility failure associated with field of values of the unwanted eigenvalues containing the wanted eigenvalues [20].

3.1 A General Bound

Literature on the Ritz values of non-Hermitian matrices states little beyond the elementary fact that all Ritz values must lie in the field of values. Perhaps the scarcity of results for non-Hermitian matrices stems, in part, from the lack of a natural ordering for complex numbers. In the Hermitian case, all the Ritz values are real and there is a natural way of ordering them; this is not so for general matrices. Also, without interlacing, it is difficult to say which eigenvalue a Ritz value is converging towards. Consider the the Ritz values from an Arnoldi decomposition, see equation (1.4). The Ritz values are eigenvalues of $A + E$ with $E = -h_{k+1,k} v_{k+1} v_k^*$ [54, p. 255]. For the Grcar matrix, Figure 3.1 shows the curves in the complex plane traced out by the eigenvalues of $A + tE$ for $t \in [0, 1]$. In the figure, the paths of the eigenvalues of $A + tE$ nearly cross, or start out near one eigenvalue and yet converge to another.

In spite of such results, applications often motivate a particular ordering, either by real part, imaginary part, or magnitude. Even for such basic orderings, little has been said. Here we begin by ordering Ritz values by their real part, a natural ordering if one wishes to compute the rightmost eigenvalue. To illustrate the consequences of this ordering, take A to be the 3×3 Jordan block

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3.1}$$

The field of values of A is the closed disk centered at the origin of radius $\sqrt{2}/2$ in the complex plane [24]. Now consider the following simple numerical experiment:

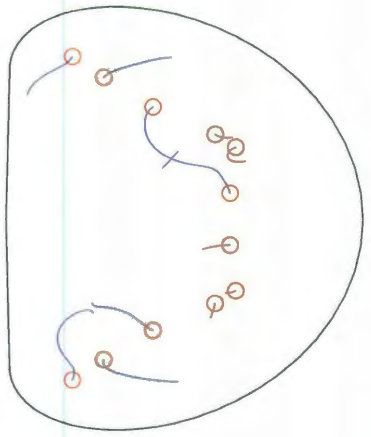


Figure 3.1 : Eigenvalues of $A + tE$ for $t \in [0, 1]$ and $E = h_3 v_3 v_2^*$ from having run two steps of Arnoldi. The circles and the black line indicate the eigenvalues and the boundary of the field of values of A, where A is Matlab's Grcar matrix (`gallery('grcar',12)`). The eigenvalues of $A + E$ take paths to the eigenvalues of A that make it difficult to associate the eigenvalues of one to the other.

generate random two-dimensional (complex) subspaces, compute the Ritz values, and sort them by their real parts. Figure 3.2 illustrates the results. The leftmost Ritz values appear to cover only a portion of the field of values. In none of these 10,000 experiments does the leftmost Ritz value fall near the rightmost extent of the field of values. Said another way, it does not appear possible to have two Ritz values in the far right portion of the field of values.

Exactly how far to the right the leftmost Ritz value can lie can be bounded using a simple trace argument. Consider the matrix $V \in \mathbb{C}^{3\times 2}$ with $V^*V = I$, and let $v \in \mathbb{C}^3$ denote the unit vector such that $U = [V\, v]$ is unitary. Let z_1 and z_2, with $\operatorname{Re} z_1 \leq \operatorname{Re} z_2$, denote the Ritz values from V, i.e., the eigenvalues of $T = V^*AV$. Also, let $z_3 = v^*Av$. Since the trace of a matrix is invariant under similarity transformations,

$$\operatorname{tr}(T) + z_3 = \operatorname{tr}(U^*AU) = \operatorname{tr}(A) = 0,$$

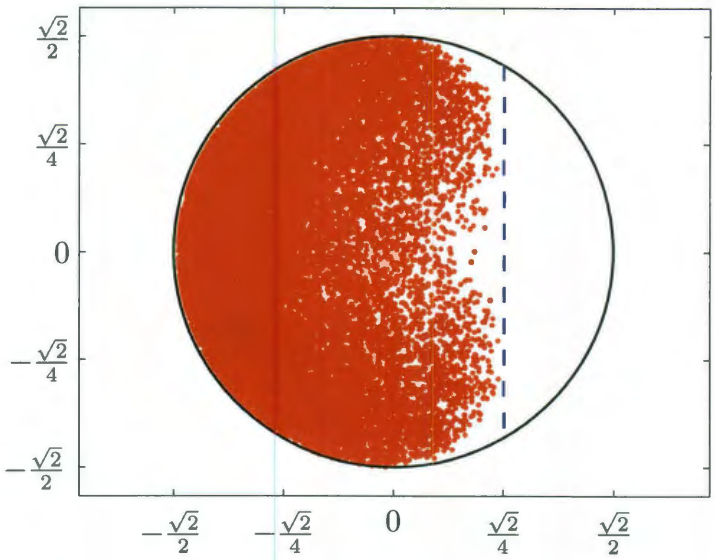


Figure 3.2 : Points in the region where left Ritz values of A may lie as determined by 10000 random subspaces. The dashed vertical line at $\mathrm{Re}\, z = \sqrt{2}/4$ indicates the bound for $\mathrm{Re}\, z_1$ based on a trace argument.

as the trace of a matrix is equal to the sum of the diagonal entries, and the diagonal entries of A are all zero. Taking the real part and noting that z_3, as a point in the field of values of A, must have real part no smaller than $-\sqrt{2}/2$, we conclude that

$$\mathrm{Re}\,(\mathrm{tr}(T)) = -\mathrm{Re}\, z_3 \leq \frac{\sqrt{2}}{2}.$$

Hence the real part of the trace of T can be no greater than $\sqrt{2}/2$. From the ordering of the Ritz values, $2\mathrm{Re}\, z_1 \leq \mathrm{Re}\, z_1 + \mathrm{Re}\, z_2 = \mathrm{Re}\,(\mathrm{tr}(T))$; thus, the real part of the leftmost Ritz value must satisfy

$$\mathrm{Re}\, z_1 \leq \frac{\sqrt{2}}{4}. \tag{3.2}$$

This result might at first seem quite jarring; however it is a natural consequence of some very fundamental matrix properties. Indeed, the simple bound (3.2) is actually just a special case of a far more general analysis based on eigenvalue majorization.

First recall that any matrix $A \in \mathbb{C}^{n\times n}$ can be decomposed as the sum of its Hermitian and skew-Hermitian parts; some call this the *Cartesian decomposition* [45]: $A = H + iS$ where $H = (A + A^*)/2$ and $S = (A - A^*)/2$. For now, we wish to study the Ritz values of A drawn from the subspace $\mathrm{Ran}(V)$, where $V \in \mathbb{C}^{n\times(n-1)}$ has orthonormal columns. Without loss of generality, we can assume that $R = V^*AV$ is upper triangular, and hence the Ritz values are on the diagonal of R. Label the Ritz values by increasing real part: $\mathrm{Re}\, z_1 \le \cdots \le \mathrm{Re}\, z_{n-1}$. Let z_n denote the Rayleigh quotient of a unit vector v orthogonal to $\mathrm{Ran}(V)$. Then the set $\{z_1, \ldots, z_n\}$ comprises the diagonal entries of the matrix $[V\ v]^*A[V\ v]$, while the real parts $\mathrm{Re}\, z_1, \ldots, \mathrm{Re}\, z_n$ are the diagonal entries of $[V\ v]^*R[V\ v]$.

Let $\mu_1 \le \mu_2 \le \cdots \le \mu_n$ denote the eigenvalues of H, and let $z_{(i)}$ denote the ordering of the z_i such that $\mathrm{Re}\, z_{(1)} \le \cdots \le \mathrm{Re}\, z_{(n)}$. (This relabeling is necessary because the value $z_n = v^*Av$ may be left of the rightmost Ritz value z_{n-1}.) By Theorem 1.3, the vector $[\mathrm{Re}\, z_{(j)}]$ of diagonal entries of $[V\ v]^*R[V\ v]$ majorizes the vector $[\mu_j]$ of eigenvalues, i.e.,

$$\sum_{j=1}^{k} \mu_j \le \sum_{j=1}^{k} \mathrm{Re}\, z_{(j)}, \quad k = 1, \ldots, n,$$

with equality for $k = n$. Since $\mathrm{Re}\, z_{(j)} \le \mathrm{Re}\, z_j$, we have

$$\sum_{j=1}^{k} \mu_j \le \sum_{j=1}^{k} \mathrm{Re}\, z_j, \quad k = 1, \ldots, n-1, \tag{3.3}$$

which means that the vector $[\mathrm{Re}\, z_j]_{j=1}^{n-1}$ weakly majorizes the vector $[\mu_j]_{j=1}^{n-1}$.

Theorem 3.1 For $A \in \mathbb{C}^{n\times n}$ and $V \in \mathbb{C}^{n\times k}$ with orthonormal columns, the real parts of the eigenvalues of V^*AV weakly majorize the k smallest eigenvalues of the Hermitian part of A.

From this majorization one may then derive bounds that localize where the Ritz values z_j of A must fall in the complex plane. For example, the weak majorization (3.3) with $k = 2$ implies

$$\mu_1 + \mu_2 \leq \mathrm{Re}\, z_1 + \mathrm{Re}\, z_2 \leq 2\mathrm{Re}\, z_2,$$

and so

$$\frac{\mu_1 + \mu_2}{2} \leq \mathrm{Re}\, z_2,$$

thus restricting the leftmost extent of the second Ritz value of A. This analysis clearly generalizes to more Ritz values, and can be adapted to limit the rightmost extent of the Ritz values by replacing A with $-A$.

Theorem 3.2 Let $z_1, \ldots, z_{n-1}$ denote the Ritz values of $A \in \mathbb{C}^{n\times n}$ drawn from an $n-1$-dimensional subspace, labeled by increasing real part: $\mathrm{Re}\, z_1 \leq \cdots \leq \mathrm{Re}\, z_{n-1}$. *Then for $j = 1, \ldots, n-1$,*

$$\frac{\mu_1 + \cdots + \mu_j}{j} \leq \mathrm{Re}\, z_j \leq \frac{\mu_{j+1} + \cdots + \mu_n}{n-j}, \tag{3.4}$$

where $\mu_1 \leq \cdots \leq \mu_n$ denote the eigenvalues of $(A + A^*)/2$, the Hermitian part of A.

Notice that for $j = 1$ and $j = n-1$, this bound yields the trivial statement

$$\mu_1 \leq \mathrm{Re}\ z_1 \leq \mathrm{Re}\ z_{n-1} \leq \mu_n,$$

which more directly follows from the fact that Ritz values must fall inside the field of values, and $\mathrm{Re}\,(W(A)) = [\mu_1, \mu_n]$, see Section 1.2. For $k \in \{2, \ldots, n-2\}$, the theorem provides considerable insight into the interior structure of the field of values.

The restriction to an $n-1$-dimensional subspace was arbitrary, and a similar result could be stated for a k-dimensional subspace where $k \in \{1, 2, \ldots, n\}$. The bound on the left would not change; however, the bound on the right would change

to $(\mu_{n-k+j+1}+\cdots+\mu_n)/(k-j+1)$. Note that these bounds are even applicable when $k=n$, in which case the z_i are the eigenvalues of A, see [15].

When A is an n-dimensional Jordan block (ones on the first superdiagonal, zeros elsewhere), Theorem 3.2 takes a very concrete form. In this case,

$$H = \frac{1}{2}\begin{bmatrix} 0 & 1 & & \\ 1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 0 \end{bmatrix}.$$

This is a very special matrix that arises from finite difference discretization of the second derivative, as well as orthogonal polynomials.

The eigenvalues of H are well-known, and can be computed by a small calculation. From this it follows that the eigenvalues of H are $\lambda_i = \cos(i\pi/(n+1))$ for $i = 1,\ldots,n$, with the largest eigenvalue $\rho(H) = \cos(\pi/(n+1))$.

Using these eigenvalues, Figure 3.3 shows the bounds from Theorem 3.2 for Ritz values for a Jordan block of dimension $n = 8$.

From $\mathrm{tr}(H) = 0$ it follows that $\mu_1 + \cdots + \mu_{n-1} = -\mu_n$, and so

$$\mathrm{Re}\, z_{n-1} \;\geq\; -\frac{1}{n-1}\mu_n = -\frac{1}{n-1}\cos\left(\frac{\pi}{n+1}\right).$$

When A is an n-dimensional Jordan block, $W(A)$ is the disk of radius $\cos(\pi/(n+1))$; Theorem 3.2 establishes a containment region for the rightmost Ritz value z_{n-1} that tends toward the right half of $W(A)$. It might initially seem a weakness that this bound does not require the rightmost Ritz value from an $n-1$ dimensional subspace to fall further to the right. However, if we take for V the first $n-1$ columns of the n-by-n identity matrix, then V^*AV is the $(n-1)$-by-$(n-1)$ upper-left corner of A. The corresponding Ritz values are $z_1 = z_2 = \cdots = z_{n-1} = 0$: *any* bound on $\mathrm{Re} z_{n-1}$

must thus contain the interval $[0, \cos(\pi/(n+1))]$. To put this another way, our Ritz value containment regions all include $\mathrm{Re}\,\mathrm{tr}(A)/n$.

Is Theorem 3.2 sharp? If A is Hermitian, then $H = A$, and $\mu_1, \ldots, \mu_n$ are the eigenvalues of A, allowing a comparison with the Cauchy interlacing theorem,

$$\mu_k \leq z_k \leq \mu_{k+1}.$$

This bound, which is sharp, will be considerably tighter than the bound of Theorem 3.2 when the eigenvalues of H are well-separated.

On a similar note, consider instead the sharpness of the majorization in Theo-

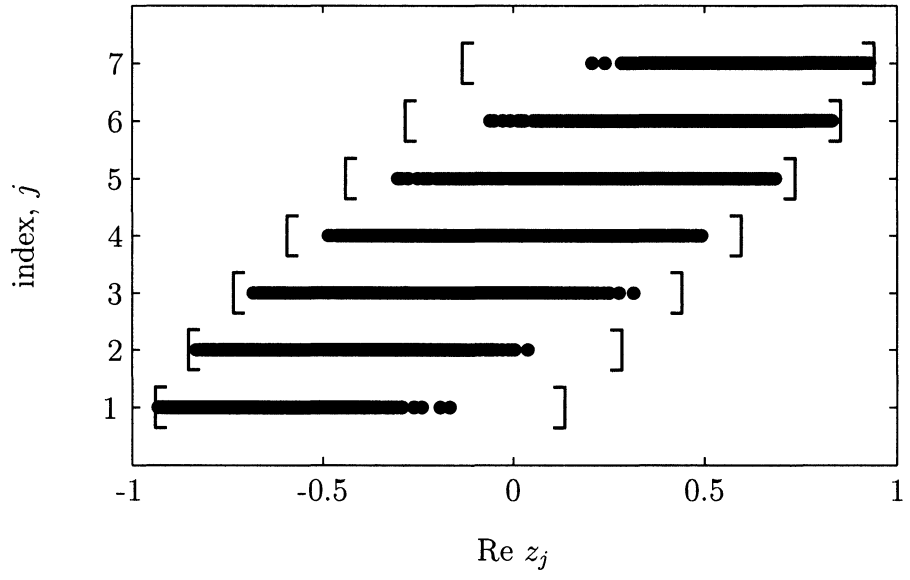


Figure 3.3 : Theorem 3.2 illustrated for a Jordan block of dimension $n = 8$. For each of $j = 1, \ldots, n-1$, the bound from Theorem 3.2 is shown as a bracket containing the real parts of the Ritz values z_j drawn from 2000 random real $n-1$ dimensional subspaces (solid black dots).

rem 3.1. Due to interlacing, the majorization can always be sharp. By selecting subspaces that coincide with eigenspaces of H, for any $k = 1, \ldots, n-1$, inequaltiy (3.3) can be sharp. Moreover, if the eigenvalues of the Hermitian part are distinct, then the corresponding subspaces are unique. For the $n = 3$ Jordan block, this is precisely why there appear to be only two complex conjugate points that attain the bound in the plots in Figure 3.2. These two points correspond to restricting to the eigenspace associated with the two rightmost eigenvalues of the Hermitian part of A.

A useful perspective to use when considering these results is that since the eigenvalues of the Hermitian part of a matrix and its restrictions must interlace, the boundary of the field of values also interlace as well, i.e., the projection property of the field of values, $W(V^*AV) \subset W(A)$, is due to interlacing. This is illustrated in Figure 3.4. Interlacing tells us that the real parts of the leftmost and right most points of the field of values of a $k \times k$ orthogonal restriction must fall in $[\mu_1, \mu_{n-k+1}]$ and $[\mu_k, \mu_n]$. This is evident in the figure. This observations suggests that for a given matrix, via interlacing, one can construct orthogonal restrictions having a prescribed Hermitian part.

Note that all the results in this section bound the real part of the Ritz values. By a simple complex scalar rotation the Ritz values can be bounded along any direction in the complex plane. The next section addresses fundamentally different orderings.

3.2 Other Orderings, Other Bounds

In the previous section we derived bounds for Ritz values ordered by real part. Here we consider related bounds that result when ordering Ritz values by their magnitude, $|z_i|$, or their phase, $\arg(z_i)$.

3.2.1 Ordering by magnitude

Results regarding the magnitude of Ritz values follow readily from properties of singular values. Alfred Horn in 1954 [28] showed necessary and sufficient conditions that must be satisfied by the eigenvalues of a matrix with prescribed singular values. The conditions are due to Weyl, see Theorem 1.5. Weyl's Theorem is a consequence of the interlacing of the singular values, see Lemma 1.1. Together, Weyl's Theorem and the lemma tell us two things that, when combined, give insight into Ritz values. First, the singular values of a matrix *log-majorize* the magnitudes of the eigenvalues (Theorem 1.5):

$$sum_i = 1^k \log |\lambda_i| \leq sum_i = 1^k \log \sigma_i,$$

for $k = 1, \ldots, n$ with equality for $k = n$, where λ_i and σ_i denote the eigenvalues and singular values ordered from largest to smallest magnitude. The singular values

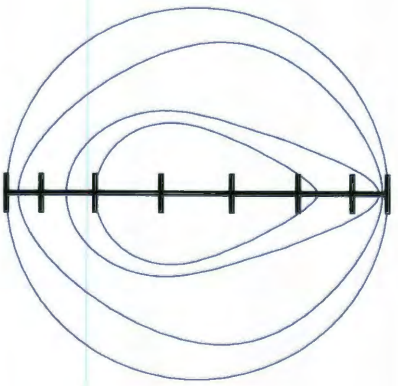


Figure 3.4 : Illustration of interlacing of the boundary of the field of values for submatrices of a $n = 8$ dimensional Jordan block. The tick marks indicate the eigenvalues of the Hermitian part of the Jordan block. The curves indicate the boundary of the field of values of submatrices of $V^* J_8 V$, where V is a random unitary matrix. Each matrix consists of columns and rows one through k of $V^* J_8 V$ for $k = 5, \ldots, 8$.

of a k-dimensional orthogonal restriction of a matrix are bounded by the k largest singular values of the full matrix (Lemma 1.1):

$$\sigma_i(V^*AV) \leq \sigma_i,$$

for $i = 1, \ldots, k$, where $\sigma_i(V^*AV)$ denotes the ith singular value of the restriction V^*AV for $V \in \mathbb{C}^{n\times k}$ having orthonormal columns. Since Ritz values are the eigenvalues of an orthogonal restriction V^*AV, the magnitudes of the Ritz values $|z_1| \geq \cdots \geq |z_k|$ are weakly log-majorized by the singular values,

$$\prod_{i=1}^{j} |z_i| \leq \prod_{i=1}^{j} \sigma_i,$$

for $i = 1, \ldots, k$. Thus if the singular values are known, we may bound the magnitudes of the Ritz values.

Theorem 3.3 Let $z_1, \ldots, z_k$ denote the Ritz values of $A \in \mathbb{C}^{n\times n}$ drawn from an k-dimensional subspace, labeled by decreasing magnitude: $|z_1| \geq \cdots \geq |z_k|$. Then for $j = 1, \ldots, k$,

$$|z_j| \leq (\sigma_1 \cdots \sigma_j)^{\frac{1}{j}}, \tag{3.5}$$

where $\sigma_1 \geq \cdots \geq \sigma_n$ denote the singular values of A.

Recall that Ritz values change accordingly for complex shifts, scalings and rotations of A: if $\{z\}$ are Ritz values of A, the corresponding Ritz values of $\alpha(A - \sigma I)$ for $\alpha, \sigma \in \mathbb{C}$ are simply $\{\alpha(z_i - \sigma)\}$. This is not the case when we concern ourselves with magnitudes and phases. Bounds obtained with singular values of A are not linearly related to bounds obtained for $\alpha(A - \sigma I)$ for σ nonzero. Depending on where we would like to localize the Ritz values, we should choose σ accordingly to get the best bounds possible.

Since singular values have some of the same min-max/max-min characteristics as the eigenvalues of Hermitian matrices, Theorem 3.5 is not all that surprising. Of course, the bounds can be sharp when A is normal. It is perhaps more surprising that a majorization bound can also be constructed when the Ritz values are sorted by their phase.

3.2.2 Ordering by phase

There are few known results regarding the phases of eigenvalues, let alone Ritz values. In this section I provide a motivating example for ordering by phases, then review what is known. To properly characterize the phases of Ritz values, I use a unitary matrix whose eigenvalues have interlacing properties, derive a majorization result for the phases of the Ritz values, then show how this result relates to the majorization result for the real parts of Ritz values.

As in Section 3.1, consider the 3×3 matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

This normal matrix, a *circulant shift*, has eigenvalues at the cubic roots of unity; the boundary of its field of values is the triangle formed by connecting its eigenvalues. Consider the Ritz values z_1 and z_2 from thousands of random two dimensional subspaces, but this time sort the Ritz values by their phases. We could measure the phases of the Ritz values about the origin; however, instead we will shift our matrix by $e^{-\pi i/3}$, one of the eigenvalues. The field of values of $A - e^{-i\pi/3}I$ lies in the upper half plane, and hence the phases of the Ritz values will be $\arg(z_j - e^{-\pi i/3}) \in [0, \pi]$. Again, identify a left and a right Ritz value. From the sorting by the real part, we

know that there should be regions where the Ritz values cannot cluster. The question is: What do these regions look like, subject to ordering by phase? Figure 3.5 shows the results. The Ritz values z_1 and z_2 appear to always lie on opposite sides of the angular bisector of $\angle\lambda_1\lambda_2\lambda_3$.

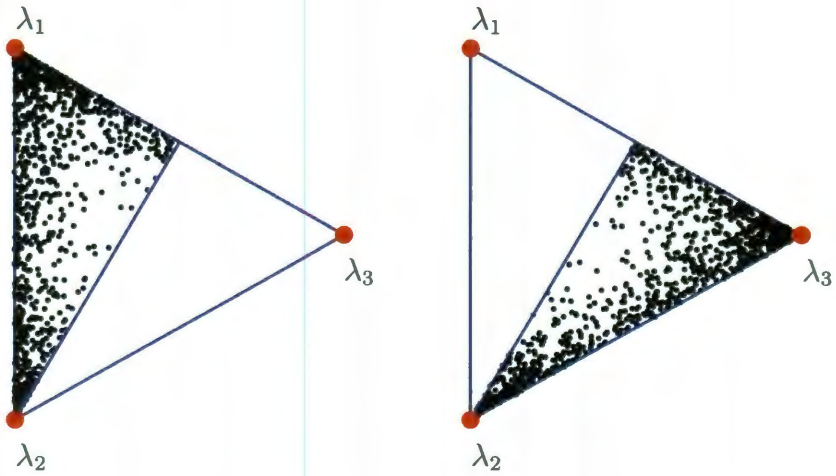


Figure 3.5 : Left and right Ritz values of A from a 2-dimensional subspace, ordered by phase about the eigenvalue at $\lambda_2 = e^{-\pi i/3}$. The left and right Ritz values are separated by the angular bisector of $\angle\lambda_1\lambda_2\lambda_3$ (as will be explained in Section 3.4).

The idea of ordering by the magnitudes and phases of the Ritz values is motivated by the polar form of a complex number, $z = re^{i\phi}$ for $r \in [0, \infty)$ and $\phi \in [0, 2\pi]$. The generalization to matrices is known as the polar decomposition. A polar decomposition of $A \in \mathbb{C}^{n\times n}$ consists of matrices P_i and U_i, such that $A = P_1U_1 = U_2P_2$, where the P_i are positive semidefinite matrices such that $P_1^2 = AA^*$, $P_2^2 = A^*A$, and the U_i are unitary. The U_i, like $e^{i\theta}$ in $z = re^{i\phi}$, account for rotation of a vector, while the P_i, like r, account for changes in length. If A is nonsingular, then $U_1 = U_2$. The P_i and A have the same singular values, hence the inequalities in Theorem 1.5 give insight into what the eigenvalues of the P_i, the *polar part* of a matrix, indicate about the eigenvalues of the matrix. In the same spirit, Alfred Horn and Robert Steinberg [29]

considered what the eigenvalues of the U_i, the *unitary part*, of a matrix indicate about its eigenvalues.

Unitary part versus unitary factor

The polar decomposition and the unitary parts U_i of a matrix are not compatible with orthogonal restrictions. Other than the results for the singular values just stated, there is not a clear way of relating the polar decomposition of an orthogonal restriction of A to the polar decomposition of the full matrix. To overcome this problem, define the *unitary factor* of a matrix A such that $0 \notin W(A)$.

Definition. A unitary matrix $U \in \mathbb{C}^{n\times n}$ is called a **unitary factor** *for a matrix $A \in \mathbb{C}^{n\times n}$ with $0 \notin W(A)$, if there exists a nonsingular $C \in \mathbb{C}^{n\times n}$ such that C^*AC.*

A unitary factor of A is a matrix to which A is congruent. (Recall that two matrices $A, B \in \mathbb{C}^{n\times n}$ are congruent if there exists a nonsingular $C \in \mathbb{C}^{n\times n}$ such that $A = C^*BC$. Congruence transformations preserve what is called the *angular field of values*, the smallest angular sector radiating from the origin in the complex plane that contains the field of values.) Note, the unitary factor of a matrix is not unique: if U is a unitary factor of A then so is Q^*UQ for any unitary $Q \in \mathbb{C}^{n\times n}$. Much as the phase of 0 is ill-defined, the unitary factor of A is only defined if $0 \notin W(A)$. Let us assume that we have used a complex scaling and a shift, so that $W(A)$ is contained in the open upper half of the complex plane. In this case, Horn and Steinberg [29] showed that A is always congruent to a unitary matrix, and this matrix is a *unitary factor* of A. For $A = H + iS$, where H and iS are Hermitian and skew-Hermitian parts of A, if $W(A)$ is in the strict upper half plane, then S is a Hermitian positive definite matrix, which may be factored as $S = S^{1/2}S^{1/2}$, with $S^{1/2}$ Hermitian positive definite. Then, $S^{-1/2}AS^{-1/2}$ is congruent to A with Hermitian part $\widehat{H} := S^{-1/2}HS^{-1/2}$ and

skew-Hermitian part iI. As $\widehat{H}$ is Hermitian, let X denote a unitary matrix that diagonalizes $\widehat{H}$. Then

$$S^{-1/2}AS^{-1/2} = X(\Lambda + iI)X^*, \tag{3.6}$$

where Λ is diagonal with $\sigma(\Lambda) = \sigma(\widehat{H})$. The matrix $C = S^{-1/2}X(\Lambda^2 + I)^{-1/4}$ is such that $U = C^*AC$ is a diagonal unitary matrix having eigenvalues in the strict upper half plane with phases equal to $\arg(\Lambda_{ii} + iI)$. Much like the signature of a Hermitian matrix, the eigenvalues of a unitary factor are unique

Lemma 3.1 If$A \in \mathbb{C}^{n\times n}$ is such that $W(A)$ lies in the strict upper half plane, then the eigenvalues of any unitary factor of a A are unique.

Proof. Suppose there exist two unitary factors of A: $A = C_1^*U_1C_1 = C_2^*U_2C_2$, where the C_i are nonsingular and the U_i are unitary. Without loss of generality assume that U_1 and U_2 are diagonal, and that the diagonal entries of U_2 are arranged so that all equal eigenvalues are grouped consecutively. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of U_2, labeled in the order they appear along the diagonal of U_2. As C_1 and C_2 are both nonsingular, U_1 is congruent to U_2: $U_1 = WU_2W^*$, where $W = C_1^{-*}C_2^*$. From the normality of U_1, we know $U_1^*U_1 = U_1U_1^*$, so

$$WU_2W^*WU_2^*W^* = WU_2^*W^*WU_2W^*,$$

which, after factoring off a W on the left and a W^* on the right, gives

$$U_2W^*WU_2^* = U_2^*W^*WU_2. \tag{3.7}$$

Partition the columns of W so that $W = [W_1, \ldots, W_m]$, where W_i comprises all columns associated with λ_i, the ith distinct eigenvalue of U_2. Since U_2 is diagonal, the (j,k) block of equation (3.7) is $\lambda_j\bar{\lambda}_kW_j^*W_k = \bar{\lambda}_j\lambda_kW_j^*W_k$. Note that $(\lambda_j\bar{\lambda}_k - \bar{\lambda}_j\lambda_k) =$

$2i|\lambda_j\bar{\lambda}_k|\sin\phi$, where $\phi = \arg(\lambda_j\bar{\lambda}_k)$. Since all the eigenvalues of U_2 are in the strict upper half plane, if $\lambda_j \neq \lambda_k$, then $\sin\phi \neq 0$. Hence $0 = (\lambda_j\bar{\lambda}_k - \bar{\lambda}_j\lambda_k)W_j^*W_k$ implies $W_j^*W_k = 0$, so the columns of W corresponding to distinct eigenvalues are orthogonal. Hence the matrix W^*W is block diagonal of the form

$$W^*W = \begin{bmatrix} W_1^*W_1 & & \\ & \ddots & \\ & & W_m^*W_m \end{bmatrix}.$$

Let $D = \mathrm{diag}(D_1, \cdots, D_m)$, where $D_i = (W_i^*W_i)^{-1/2}$. Then D is positive definite block diagonal, and $\widehat{W} := WD$ is unitary. In terms of $\widehat{W}$ and D, since U_2 is diagonal,

$$U_1 = WU_2W^* = \widehat{W}D^{-1}U_2D^{-1}\widehat{W}^* = \sum_{i=1}^{m} \lambda_i \widehat{W}_i D_i^{-2} \widehat{W}_i^*.$$

Performing a similarity transformation on U_1 with $\widehat{W}$ gives

$$\widehat{W}^*U_1\widehat{W} = \begin{bmatrix} \lambda_1 D_1^{-2} & & \\ & \ddots & \\ & & \lambda_m D^{-2} \end{bmatrix}.$$

This implies that $\sigma(U_1) = \bigcup_{i=1}^{m} \sigma(\lambda_i D_i^2)$. As the eigenvalues of U_1 all have magnitude one and the D_i are positive definite, each D_i must equal the identity. Thus, the matrix W is in fact unitary and $\sigma(U_1) = \sigma(U_2)$. Hence all unitary factors of a matrix have the same eigenvalues with the same multiplicities. ■

Contrast the flexibility in the unitary factor with the situation for the unitary part a polar decomposition $A = PU$, for which the matrix U, not just its eigenvalues, must be unique.

Horn and Steinberg [29] showed that, when a unitary factor exits, the phases of its eigenvalues satisfy the following min-max, max-min relations:

$$\alpha_i = \min_{\dim(\mathcal{U})=i} \max_{x\in\mathcal{U}} \arg(x^*Ax) = \max_{\dim(\mathcal{U})=n-i+1} \min_{x\in\mathcal{U}} \arg(x^*Ax); \tag{3.8}$$

the α_i are the phases of the eigenvalues of the unitary part of A, ordered such that $\alpha_1 \leq \cdots \leq \alpha_n$. This result can be seen by expressing x^*Ax as $xC^{-*}UC^{-1}x = \hat{x}^*U\hat{x} = \sum_{i=1}^{n} |\hat{x}_i|^2 e^{i\alpha_i}$, where $\hat{x} = C^{-1}x$. Then proof of these min-max, max-min relations follows as for the eigenvalues of a Hermitian matrix: pick a subspace for which the value is attained, and show that for any other subspace the quantity is either larger or smaller, as appropriate.

For matrices having a unitary factor, Horn and Steinberg [29] related the eigenvalues of a matrix to the eigenvalues of its unitary factor.

Theorem 3.4 Let λ_i denote the eigenvalues of a nonsingular matrix A and α_i the phases of the eigenvalues of a unitary factor of A, such that $0 \leq \arg\alpha_1 \leq \cdots \leq \arg\alpha_n \leq \pi$. *Then* $[\alpha_i] \prec [\arg\lambda_i]$.

This theorem follows in much the same manner as majorization of the diagonal entries by the eigenvalues in the Hermitian case: induction on the size of the matrix, combined with interlacing. This theorem is for a unitary factor, but a similar result holds for the unitary part. Let $A \in \mathbb{C}^{n\times n}$ have a polar decomposition, $A = PU$, with eigenvalues of the unitary part U satisfying the requirement of Theorem 3.4. Then the matrix

$$\widehat{A} = P^{-1/2}AP^{1/2} = P^{1/2}UP^{1/2} \tag{3.9}$$

is similar to A and has as a unitary factor the matrix U. As A and $\widehat{A}$ have the same eigenvalues, by the Theorem 3.4 the phases of the eigenvalues of U are majorized by the phases of the eigenvalues of A.

Horn and Steinberg also proved that the entries of any two length n vectors $\{z_i\}$ and $\{\hat{z}_i\}$ with nonzero entries satisfying $\arg\prod_{i=1}^{n} z_i = \arg\prod_{i=1}^{n} \hat{z}_i$, can be arranged such that $\max\arg\hat{z}_i - \min\arg\hat{z}_i \leq 2\pi$ and $\{\arg z_i\} \prec \{\arg\hat{z}_i\}$ [29]. This suggests

the eigenvalues of A and its unitary part U can always be arranged so that phases of eigenvalues of U are majorized by the phases of the eigenvalues of A. Furthermore, one could use the technique in equation (3.9) to relate the unitary part U of A from its polar decomposition to a matrix $\widehat{A}$ that is congruent to U. Why bother with a unitary factor? What sets a unitary factor apart in these scenarios is that the eigenvalues of a unitary factor satisfy min-max/max-min relations (3.8), whereas the eigenvalues of a unitary part of a matrix do not. Moreover, as mentioned before, a unitary factor is compatible with an orthogonal restriction, while a unitary part of a matrix is not, i.e., an orthogonal restriction of a full matrix is also a restriction of the unitary factor, whereas there is no similar relationship for the unitary part from the polar decomposition. This is important for our setting in Chapter 4, as the Arnoldi method involves taking orthogonal restrictions of the matrix.

Phase majorization for Ritz values

If A has a unitary factor, then one knows two things: the phases of the eigenvalues of the unitary part have max-min/min-max properties (3.8), and the phases of the eigenvalues of the unitary part are also majorized by the phases of the eigenvalues of A (Theorem 3.4). From the max-min/min-max property, for any restriction $T = V^*AV$ with $V \in \mathbb{C}^{n\times k}$ having orthonormal columns, the phases of the eigenvalues of the unitary part of T must interlace the phases of the eigenvalues of the unitary part of A:

$$\alpha_i \leq \hat{\alpha}_i \leq \alpha_{n+k-i},$$

for $i = 1, \ldots, k$, where α_i and $\hat{\alpha}_i$ denote the phases of the eigenvalues of the unitary parts of A and T. Combining these results, the phases of the Ritz values weakly

majorize the phases of the eigenvalues of the unitary part of A:

$$\sum_{i=1}^{j} \alpha_i \leq \sum_{i=1}^{j} \arg z_i.$$

Thus, if the phases of the α_i are known, then we may bound the phases of the Ritz values.

Theorem 3.5 Let $A \in \mathbb{C}^{n\times n}$ be a nonsingular matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ such that $W(A)$ is contained in the upper half plane. Let $z_1, \ldots, z_k$ denote the Ritz values of A drawn from a k-dimensional subspace, labeled by increasing phase: $\arg z_1 \leq \cdots \leq \arg z_k$. *Then for $j = 1, \ldots, k$,*

$$\frac{\alpha_1 + \cdots + \alpha_j}{j} \leq \arg z_j \leq \frac{\alpha_{n-k+j} + \cdots + \alpha_n}{k - j + 1}, \tag{3.10}$$

where $\alpha_1 \geq \cdots \geq \alpha_n$ denote the phases of the eigenvalues of the unitary part of A.

Though this result is for a matrix whose field of values is in the upper half of the complex plane, we can always shift and scale so that $W(\alpha(A - \sigma I))$ is in a half plane. Thus, for the more common case in applications, where $W(A)$ is in the strict left or right half plane, the above theorem holds for $-iA$ or iA.

These phase majorization results can also be related to the majorization result involving the real parts of the eigenvalues and the eigenvalues of the Hermitian part. For $W(A)$ in the upper half plane, let α_i denote phases of the eigenvalues of a unitary factor of A ordered such that $\alpha_1 \leq \cdots \leq \alpha_n$. The α_i satisfy

$$\cot \alpha_i = \xi_i,$$

where ξ_i denote the eigenvalues of $S^{-\frac{1}{2}} H S^{-\frac{1}{2}}$ with H and S being the Hermitian and skew-Hermitian parts of A. This can be seen from equation (3.6); the eigenvalues of the unitary part of A have the same phases as the eigenvalues of $S^{-\frac{1}{2}} H S^{-\frac{1}{2}} + iI$.

Since the α_i are increasing order, the ξ_i will be in decreasing order. In terms of the ξ_i, the majorization from Theorem 3.4 takes the form

$$\sum_{i=1}^{j} \cot^{-1} \xi_i \leq \sum_{i=1}^{j} \arg \lambda_i,$$

for $j = 1, \ldots, n$, where λ_i denotes the eigenvalues of A ordered by increasing phase. Now consider the matrix $B(y) = A + iyI$ and the corresponding $\xi_i(y)$, $\lambda_i(y)$, and $\alpha_i(y)$. Note, the phases of the eigenvalues of $B(y)$ and its unitary factor correspond to measuring the phases of the eigenvalues of A about the point $-yi$. For large y, to first order in $\xi_i(y)$, $\cot^{-1} \xi_i(y) \approx \pi/2 - \xi_i(y)$, and similarly $\arg \lambda_i(y) \approx \pi - \cot \arg \lambda(y)$. This is a consequence of the angles approaching $\pi/2$ as y becomes large. Multiplying the $\xi_i(y)$ and computing their limit as y goes to infinity,

$$\begin{aligned} \lim_{y \to \infty} y\xi_i &= \lim_{y \to \infty} \lambda_i (y(S + yI)^{-\frac{1}{2}} H (S + yI)^{-\frac{1}{2}}) && (3.11) \\ &= \lim_{y \to \infty} \lambda_i \left(\left(\frac{1}{y} S + I \right)^{-\frac{1}{2}} H \left(\frac{1}{y} S + I \right)^{-\frac{1}{2}} \right), && (3.12) \end{aligned}$$

where $\lambda_i(y(S+yI)^{-\frac{1}{2}} H (S+yI)^{-\frac{1}{2}})$ denotes the $i$the eigenvalue of $y(S+yI)^{-\frac{1}{2}} H (S+yI)^{-\frac{1}{2}}$, sorted from largest to smallest. Thus $\lim_{y \to \infty} y\xi_i(y) = \mu_i$, where μ_i denotes the ith eigenvalue of H sorted from largest to smallest. Similarly, $\lim_{y \to \infty} y \cot \arg \lambda_i(y) = \operatorname{Re} \lambda_i$, where λ_i denotes the ith eigenvalue of $B(0) = A$ sorted from largest to smallest real part. Hence the majorization of the phases for large y gives

$$\sum_{i=1}^{j} (\pi - \xi_i) \leq \sum_{i=1}^{j} (\pi - \cot \arg \lambda_i),$$

for $j = 1, \ldots, n$. Removing the π's and then scaling by y gives

$$\sum_{i=1}^{j} y \cot \arg \lambda_i(y) \leq \sum_{i=1}^{j} y\xi_i(y).$$

In the limit as y goes to infinity, this gives

$$\sum_{i=1}^{j} \operatorname{Re} \lambda_i \leq \sum_{i=1}^{j} \mu_i.$$

Hence in the limit as y approaches infinity majorization of the phases of the eigenvalues of $B(y)$ by phases of the eigenvalues of a unitary factor of $B(y)$ gives majorization of the real part of the eigenvalues A by the eigenvalues of the Hermitian part of A as implied by Theorem 3.1.

3.3 iFOV-2 for a 3×3 Jordan Block

Having bounded Ritz values in the previous sections, in this section and the next we seek to precisely describe the Ritz values of two extreme matrices: a defective matrix and a normal matrix. We wish to show what becomes of interlacing, and also the difficulty of characterizing Ritz values of non-Hermitian matrices. In this section I will solve iFOV-2 for a 3×3 Jordan block,

$$J_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

In doing so, I will characterize the Ritz value pairs $\{z_1, z_2\}$ for which iFOV-2 is solvable. These results motivated some of the bounds in the previous sections. To our knowledge, there do not exist any similar detailed analyses of the Ritz values of a nonnormal matrix.

Characterizing when iFOV-2 is solvable is equivalent to analyzing the eigenvalues of $T = V^* J_3 V$ for all $V \in \mathbb{C}^{3\times 2}$ such that $V^*V = I$. As T is a 2×2 matrix, it will have two eigenvalues, z_1 and z_2, which without loss of generality, I refer to as left and right eigenvalues, in the sense that $\mathrm{Re}(z_1) \leq \mathrm{Re}(z_2)$. With this ordering in mind, some questions arise: With z_1 fixed, where in $\mathbb{C}$ may z_2 lie? Where in $\mathbb{C}$ may $z_1 = z_2$? Recall that all Ritz values must lie in the field of values, and the leftmost Ritz value cannot fall to the right of the line $\mathrm{Re}\, z = \frac{\sqrt{2}}{4}$ in the complex plane: see Figure 3.2 and

Theorem 3.2. I will determine the curve for the boundary of the regions where the left and right Ritz values may lie. The matrix is sufficiently small that the solution can be specified completely, though this process hints at the underlying complexity of the problem for larger matrices. This result will be contrasted with the much simpler analogue for a normal 3×3 matrix in the following section. In the language of algebraic geometry, I will start from a parametric representation of all possible solutions to iFOV-2, then construct a more implicit representation. The analysis becomes quite technical, but ultimately allows us to compute the boundary of the regions where the leftmost and rightmost Ritz values can fall.

A detailed understanding of the Ritz values requires an expression for the Ritz values for all possible two-dimensional subspaces. For an $n-1$-dimensional subspace of $\mathbb{C}^n$, the task of parameterizing all subspaces is simplified, because every $n-1$-dimensional subspace of $\mathbb{C}^n$, represented by $V \in \mathbb{C}^{n\times(n-1)}$, $V^*V = I$, can be characterized by any nonzero unit vector v orthogonal to the subspace, $V^*v = 0$. The vector v uniquely determines the range of V. Any orthonormal basis for the range of V would determine the same Ritz values. We use these facts by way of the matrix adjugate. Recall that the adjugate of a matrix, sometimes referred to as the classical adjoint [1],[30, p. 21], satisfies

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

where A_{ji} is the matrix formed by deleting row j and column i from the matrix. (The second equality is true provided A is invertible). For unitary U, the adjugate satisfies $\mathrm{adj}(U^*AU) = U^*\mathrm{adj}(A)U$. If we form the unitary matrix $U = [V\ v] \in \mathbb{C}^{n\times n}$, then

$$\det(V^*AV) = \det((U^*AU)_{nn}) = [\mathrm{adj}(U^*AU)]_{nn} = [U^*\mathrm{adj}(A)U]_{nn} = v^*\mathrm{adj}(A)v.$$

Hence the characteristic polynomial of the restriction of a matrix A to the subspace

orthognonal to a vector v can be determined by computing the Rayliegh quotient of $\mathrm{adj}(\lambda I - A)$ with v. (Note that by using compound matrices, a similar approach may be taken for subspaces of size $k < n-1$; however there are issues with *decomposability*, the mapping of subspaces to vectors [1, 36].)

For J_n, an $n \times n$ Jordan block, we can compute

$$\mathrm{adj}(\lambda I - J_n) = \sum_{j=0}^{n-1} \lambda^{n-1-j} J_n^j.$$

Thus the coefficients c_j of the characteristic polynomial $p(\lambda) = \lambda^{n-1} \sum_j = 0^{n-2} c_j \lambda^j$ of $V^* J_n V$ are given by $c_j = v^* J_n^{n-1-j} v$. These coefficients are symmetric polynomials in the eigenvalues of $V^* J_n V$. For $n = 3$,

$$\begin{aligned} c_1 = v^* J_3 v = -(z_1 + z_2) &= -\mathrm{tr}(V^* J_3 V) \\ c_0 = v^* J_3^2 v = z_1 z_2 &= \det(V^* J_3 V), \end{aligned}$$

where z_1 and z_2 are the eigenvalues of $V^* J_3 V$. Without loss of generality, write the unit vector v in the form

$$v = \begin{bmatrix} \cos\phi_1 \\ -\sin\phi_1 \cos\phi_2 e^{i\phi_3} \\ \sin\phi_1 \sin\phi_2 e^{i\phi_4} \end{bmatrix} \tag{3.13}$$

for independent real parameters $\phi_1, \phi_2, \phi_3, \phi_4 \in [0, 2\pi)$, thus giving

$$z_1 + z_2 = \cos\phi_1 \sin\phi_1 \cos\phi_2 e^{i\phi_3} + \sin^2\phi_1 \cos\phi_2 \sin\phi_2 e^{i(\phi_4 - \phi_3)} \tag{3.14}$$

$$z_1 z_2 = \cos\phi_1 \sin\phi_1 \sin\phi_2 e^{i\phi_4}. \tag{3.15}$$

The Ritz values are completely parametrized by these two formulas. Hence, for there to be a two-dimensional subspace $\mathcal{V}$ that yields z_1 and z_2 as Ritz values, there must exist real $\phi_1, \ldots, \phi_4$ that satisfy (3.14)–(3.15). Without loss of generality, let

$\arg(z_1 z_2) = \phi_4$. (For given ϕ_1, ϕ_2, ϕ_3 and ϕ_4 for which $\arg(z_1 z_2) \neq \phi_4$, this may always be achieved through an appropriate change to ϕ_3 and either ϕ_1 or ϕ_2: set $\phi_4 \to \arg z_1 z_2$, $\phi_3 \to \phi_3 + \pi$, and either $\phi_1 \to \pi - \phi_1$ or $\phi_2 \to \phi_2 + \pi$.) Given this parametric representation of the set of possible Ritz value pairs, our goal is to make implicit expressions relating z_1 and z_2 [18]. We will also note the number of distinct subspaces that are possible for particular Ritz value combinations, and, where possible, state v in terms of z_1 and z_2.

A Ritz value at zero

We wish to use equation (3.15) to eliminate ϕ_4 from equation (3.14). To perform this elimination, $\cos\phi_1 \sin\phi_1 \sin\phi_2$ must be nonzero. First consider the special case where $\cos\phi_1 \sin\phi_1 \sin\phi_2 = 0$, which implies, by (3.15), that at least one of the Ritz values is zero. Without loss of generality, say $z_1 = 0$. Three scenarios are possible from (3.14):

- $\sin\phi_1 = 0$, in which case $z_2 = 0$;
- $\cos\phi_1 = 0$, in which case $z_2 = \cos\phi_2 \sin\phi_2 e^{i(\phi_4 - \phi_3)}$, allowing z_2 to take any value in the disk $\{z \in \mathbb{C} : |z| \leq 1/2\}$;
- $\sin\phi_2 = 0$, in which case $z_2 = \cos\phi_1 \sin\phi_1 e^{i\phi_3}$, again allowing z_2 to take any value in the disk $\{z \in \mathbb{C} : |z| \leq 1/2\}$.

Hence, any pair $\{0, z\}$ is a valid set of Ritz values for $|z| \leq 1/2$: $z = 0$ only corresponds to the subspaces defined by $v = e_1, e_2$ or e_3, the canonical basis vectors; each nozero $|z| \leq 1/2$ corresponds to four possible subspaces determined by the vectors

$$v = \begin{bmatrix} 0 \\ -\sqrt{\frac{1 \mp \sqrt{1-4|z|^2}}{2}} \\ \frac{z}{|z|}\sqrt{\frac{1 \pm \sqrt{1-4|z|^2}}{2}} \end{bmatrix}, \qquad v = \begin{bmatrix} -\sqrt{\frac{1 \mp \sqrt{1-4|z|^2}}{2}} \\ \frac{z}{|z|}\sqrt{\frac{1 \pm \sqrt{1-4|z|^2}}{2}} \\ 0 \end{bmatrix}.$$

Note that the corresponding subspaces $\mathcal{V} = v^{\perp}$ contain either the left or the right eigenvector of J_3. As we will show in the following section, for a normal matrix, the magnitudes of the entries of v are uniquely determined by polynomial expressions in z_i [13]. This nonnormal case already involves square roots of the Ritz values. (Notice that we can already conclude that it is impossible to have one Ritz value at zero and the other Ritz value near the boundary of $W(J_3)$, i.e., in the region $\{z \in \mathbb{C} : 1/2 < |z| \leq \sqrt{2}/2\}$.)

Trace equals zero

With all $z_1 z_2 = 0$ cases understood, now assume $z_1 z_2 \neq 0$. Using (3.15), substitute

$$e^{i\phi_4} = \frac{z_1 z_2}{\cos\phi_1 \sin\phi_1 \sin\phi_2}, \tag{3.16}$$

into (3.14) to eliminate ϕ_4:

$$(\cos^2\phi_1 \sin\phi_1 e^{i\phi_3} + z_1 z_2 \sin\phi_1 e^{-i\phi_3}) \cos\phi_2 \;=\; (z_1 + z_2)\cos\phi_1. \tag{3.17}$$

If the coefficient of $\cos\phi_2$ is zero, then the expression on the right of equation (3.17) must also be zero. This implies that the trace $z_1 + z_2 = 0$ (since $\cos\phi_1 \neq 0$, for otherwise the determinant would be zero), hence $\cos^2\phi_1 e^{2i\phi_3} = z_1^2$, i.e., $z_1 = -z_2 = \pm\cos\phi_1 e^{i\phi_3}$. Since $z_1, z_2 \in W(J_3)$, we must have $|\cos\phi_1| \leq \sqrt{2}/2$, i.e., $\phi_1 \in [\pi/4, 3\pi/4] \cup [5\pi/4, 7\pi/4]$. In this case, there are only two possible solutions, corresponding to the vectors

$$v = \begin{bmatrix} \bar{z}_1 \\ \pm\sqrt{1 - 2|z_1|^2} \\ -z_1 \end{bmatrix}. \tag{3.18}$$

This scenario happens to correspond to all possible normal $V^* J_3 V$. We will note this in the following section when we consider Ritz values of normal matrices.

Some conjugate pairs

Now assume that the trace is nonzero: $z_1 + z_2 \neq 0$. Then from equation (3.17),

$$\cos(\phi_2) = \frac{(z_1 + z_2)\cos(\phi_1)}{\cos^2\phi_1 \sin\phi_1 e^{i\phi_3} + z_1 z_2 \sin\phi_1 e^{-i\phi_3}}, \tag{3.19}$$

thus determining ϕ_2 in terms of ϕ_1, ϕ_3, and the Ritz values.

To simplify the coefficients, write $d := z_1 z_2$ and $t := z_1 + z_2$ for the determinant and trace of $V^* J_3 V$:

$$\cos\phi_2 = \frac{t\cos\phi_1}{\cos^2\phi_1 \sin\phi_1 e^{i\phi_3} + d\sin\phi_1 e^{-i\phi_3}}. \tag{3.20}$$

Requiring the imaginary part of (3.20) to be zero yields

$$(\mathrm{Im}\,(t)\cos^2\phi_1 - \mathrm{Im}\,(d\bar{t}))\cos\phi_3 = (\mathrm{Re}\,(t)\cos^2\phi_1 - \mathrm{Re}\,(d\bar{t}))\sin\phi_3. \tag{3.21}$$

If the coefficients of $\cos\phi_3$ and $\sin\phi_3$ are both zero, then $t\cos^2\phi_1 = d\bar{t}$. This expression is invariant to rotations of the Ritz values,

$$\begin{aligned} t\cos^2\phi &= (de^{2i\psi})e^{-2i\psi}\bar{t} \\ (te^{i\psi})\cos^2\phi &= (de^{2i\psi})\overline{e^{i\psi}t}, \end{aligned} \tag{3.22}$$

as rotating the Ritz values by ψ corresponds to multiplying the determinant by $e^{2i\psi}$ and the trace by $e^{i\psi}$. Hence for this case, it suffices to assume the trace is real and positive. This implies that the determinant, which we have assumed is nonzero, is also real and positive. Thus if equation (3.22) is satisfied, then Ritz values are equivalent, via rotation, to a complex conjugate pair located in the right half plane. These are not the only Ritz values that can be associated with complex conjugate pairs, they are just the ones that satisfy (3.22), i.e., the complex conjugate pairs that can be generated by complex subspaces, $\mathrm{Im}\, e^{i\phi_3} \neq 0$.

With $d = \cos^2 \phi_1$ and $t > 0$, we eliminate ϕ_1 and ϕ_2 from equation (3.17):

$$8d^2 - 4d + t^2 \sec^2 \phi_3 = 0. \tag{3.23}$$

If we require that d and t be such that $\sec^2 \phi_3 \geq 1$, then we find that the Ritz values must satisfy

$$\left(\left| z + \frac{\sqrt{2}}{4} i \right|^2 - \frac{1}{8} \right) \left(\left| z - \frac{\sqrt{2}}{4} i \right|^2 - \frac{1}{8} \right) \leq 0. \tag{3.24}$$

This expression is negative for all z in the union of the disks of radius $\sqrt{\frac{1}{8}}$ centered at $\pm i\sqrt{2}/4$ in the complex plane. Thus all the Ritz values for this scenario must come in complex conjugate pairs and lie in these disks.

If we consider the Ritz values corresponding to $d, t > 0$ and rotate them by $e^{i\phi}$, there are two possible v:

$$v = \begin{bmatrix} \sqrt{d} e^{-i\phi} \\ -\frac{t}{2\sqrt{d}} \pm i\sqrt{1 - 2d - \frac{t^2}{4d}} \\ \sqrt{d} e^{i\phi} \end{bmatrix}, \tag{3.25}$$

where if $t = 0$, this expression reduces to equation (3.18). If $1 - 2d - t^2/4d = 0$, there is only one possible vector. This corresponds to the Ritz values lying on the boundary of the discs mentioned above. Also if $t = 0$, this vector reduces to equation (3.18).

General case

Now assuming that $t \cos^2 \phi_1 \neq d\bar{t}$, equation (3.21) gives

$$e^{i\phi_3} = \frac{t \cos^2 \phi_1 - d\bar{t}}{\left| t \cos^2 \phi_1 - d\bar{t} \right|}. \tag{3.26}$$

Using the expressions for ϕ_2, ϕ_3 and ϕ_4, equations (3.20), (3.26), and (3.16) determine an expression for v in terms of d, t, and $\cos\phi_1$:

$$v = \begin{bmatrix} \cos\phi_1 \\ -\dfrac{\cos\phi_1(t\cos^2\phi_1 - d\bar{t})}{\cos^4\phi_1 - |d|^2} \\ \dfrac{d}{\cos\phi_1} \end{bmatrix}. \tag{3.27}$$

For this vector to have norm one, $\cos\phi_1$ must satisfy

$$(\cos^{12}\phi_1 + |d|^6) - (\cos^{10}\phi_1 + |d|^4\cos^2\phi_1) + (|t|^2 - |d|^2)(\cos^8\phi_1 + |d|^2\cos^4\phi_1) + (2|d|^2 - d\bar{t}^2 - \bar{d}t^2)\cos^6\phi_1 = 0. \tag{3.28}$$

This equation is a polynomial in $\cos\phi$ that involves only even powers, consistent with $\pm v$ generating the same subspace $\mathrm{Ran}(V)$. The terms are arranged to emphasize that the polynomial is $|d|$-self reciprocal, i.e., if $\cos^2\phi_1$ is a solution then $d/\cos^2\phi_1$ is also a solution. Making use of this property, we make the substitution $\cos^2\phi \to |d|e^y$, which reduces (3.28) to

$$4d^2\cosh^3 y - 2d\cosh^2 y + (-4d^2 + t^2)\cosh y - t^2\cos\psi + 2d = 0, \tag{3.29}$$

where $\psi := \arg(d\bar{t}^2)$. This equation is cubic in $\cosh y$, which means one can write out the solution exactly in terms of d, t, and $\cos\psi$; however the complexity of the expressions limits the amount of insight that can be gained. From numerics, at most two of the solutions to this equation correspond to actual Ritz value pairs.

3.3.1 Where can z_2 lie?

If z_1 is specified, the general case does not readily indicate where z_2 lie. For bounding the region containing the leftmost Ritz value, the Ritz values that satisfy $d = \cos^2\phi_1$

for $d, t > 0$ only give a sharp bounds for the right upper and right lower portions where the leftmost Ritz value can lie; recall Figure 3.2. To address both these issues, we must better understand the general case.

Understanding where z_2 can lie once z_1 is specified amounts to solving the following optimization problems: Given $z_1 \in \mathbb{R}$, and $\operatorname{Im} z_2 \in [-\sqrt{2}/2, \sqrt{2}/2]$, determine $\max \operatorname{Re} z_2$ and $\min \operatorname{Re} z_2$, such that $z_1, z_2 \in W(J_3)$ and there exists $\cos^2 \phi_1 \in [0, 1]$ that satisfies equation (3.28). Note we have assumed z_1 is real; to handle complex z_1, the result of this optimization can be rotated, given the radiual symmetry of $W(J_3)$. Understanding this problem for $\operatorname{Im} z_2 \in [-\sqrt{2}/2, \sqrt{2}/2]$ would answer the question of when iFOV-2 is solvable.

Now consider the similar problem: Given $\operatorname{Im} z_1 \in [-\sqrt{2}/2, \sqrt{2}/2]$, determine $\max \operatorname{Re} z_1$ and $\min \operatorname{Re} z_1$ such that $z_1, z_2 \in W(J_3)$ and there exists a $\cos^2 \phi_1 \in [0, 1]$ that satisfies equation (3.28) and $\operatorname{Re} z_1 \leq \operatorname{Re} z_2$. The solution to this problem for $\operatorname{Im} z_2 \in [-\sqrt{2}/2, \sqrt{2}/2]$ would determine the region where the leftmost Ritz value may lie.

Both these optimization problems have a polynomial constraint in equation (3.28). Let $g(\cos^2 \phi_1)$ denote the polynomial in $\cos^2 \phi_1$ in equation (3.28). Then using a Lagrange multiplier to solve the problems above requires the discriminant of g to be zero. Recall the computation of the discriminant of a polynomial requires taking the resultant of the polynomial and its derivative. The resultant of two polynomials can be computed by taking the determinant of the Sylvester matrix [18], a matrix whose rank determines the degree of their greatest common divisor, i.e., how many roots two polynomials have in common. The discriminant of g gives a multivariate polynomial in $\operatorname{Re} z_1, \operatorname{Im} z_1, \operatorname{Re} z_2$ and $\operatorname{Im} z_2$. This polynomial determines, given z_1, the boundary of the region where z_2 must lie for iFOV-2 to be solvable; see Figure 3.6. As $g(\cos^2 \phi_1)$

has a repeated root for z_2 on this boundary, there are fewer subspaces that generate the Ritz value pair z_1, z_2.

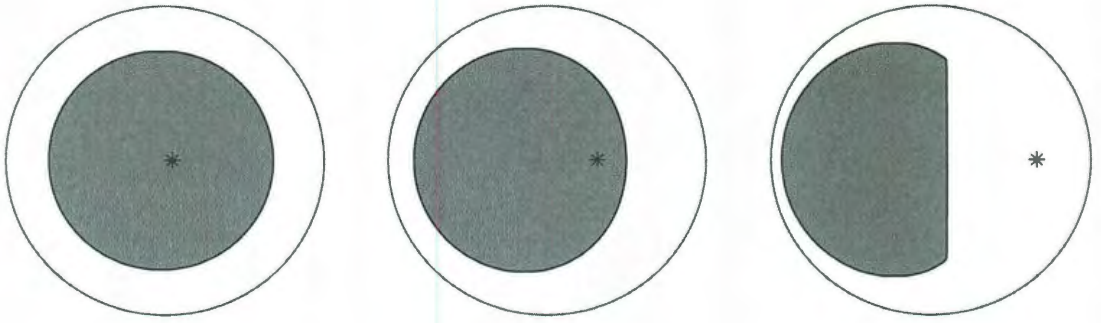


Figure 3.6 : The region in $W(J_3)$ where z_2 may lie once z_1 $(*)$ has been specified, shown in gray, for three choices of z_1. (The solid line denotes the boundary of $W(J_3)$.)

The discriminant alone is insufficient to determine where the leftmost Ritz value may lie. For in this case we are maximizing over $\operatorname{Re} z_1$, $\operatorname{Re} z_2$, $\operatorname{Im} z_2$ and $\cos^2 \phi_1$. From symmetry, for $\operatorname{Im} z_1$ the constraint $\operatorname{Re} z_1 \leq \operatorname{Re} z_2$ must be active. This may be seen by considering z_1, z_2 such that $|z_1| = R$ and $\operatorname{Re} z_1 < \operatorname{Re} z_2$. All such pairs can be rotated to determine $\hat{z}_1, \hat{z}_2$ such that either $\operatorname{Re} z_1 = \operatorname{Re} z_2$ or $z_1 \in \mathbb{R}$ and $z_1 < \operatorname{Re} z_2$. The second scenario is only relevant for the case $\operatorname{Im} z_1 = 0$; hence in general we may assume that along the right half of the boundary of the region containing the leftmost Ritz value, $\operatorname{Re} z_1 = \operatorname{Re} z_2$. With this constraint and the discriminant, to determine the region where the leftmost Ritz value can lie for a given $\operatorname{Im} z_1$, we need only maximize $\operatorname{Re} z_1$ over $z_1, z_2 \in W(J_3)$, for which the discriminant of g is zero. This requires taking the discriminant of the discriminant of g, but this time with respect to $\operatorname{Im} z_2$. Doing so for the boundary of the region containing the leftmost Ritz value, the upper and lower portions of the right half of the region are determined by equation (3.24), and

the middle portion of the region is determined by

$$\operatorname{Im}\theta_1^2(1 - 4\operatorname{Re}\theta_1^2 + 4\operatorname{Re}\theta_1^4) + (\operatorname{Re}\theta_1^2 - 12\operatorname{Re}\theta_1^4 + 4\operatorname{Re}\theta_1^6) = 0. \tag{3.30}$$

This gives $\operatorname{Im}\theta_1$ as the square root of a rational function in $\operatorname{Re}\theta_1$. The curves are shown in Figure 3.7.

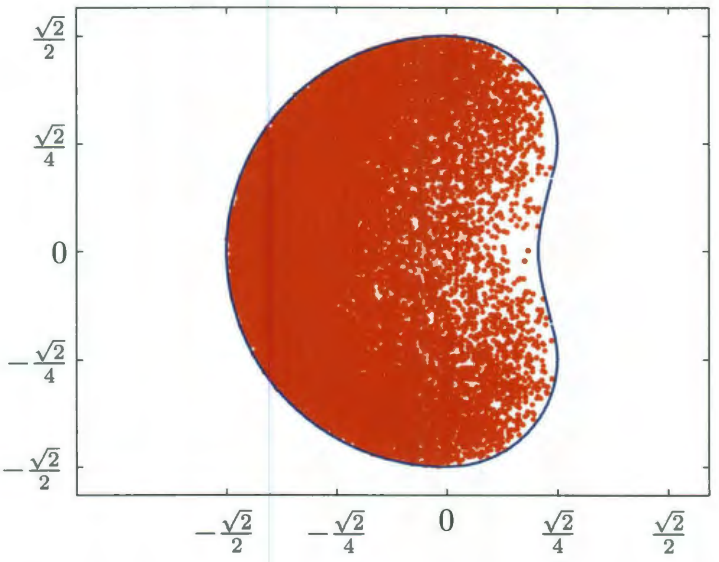


Figure 3.7 : Boundary of the region where leftmost Ritz value z_1 of J_3 may lie for all possible choices of z_2. Points show z_1 from 10000 randomly generated complex 2-dimensional subspaces.

In this section we were able to work out expressions that the Ritz values of a Jordan block must satisfy. Determining if a particular pair $\{z_1, z_2\}$ is a solution to iFOV-2 involves checking if they satisfy any of the four cases. In the most general case, determining whether $\{z_1, z_2\}$ is possible involves computing the roots of a cubic polynomial. This suggests that in general solving iFOV-$(n-1)$ requires computing the roots of a polynomial, possibly of degree at most $2n$. For the 3×3 Jordan block, symmetry allowed us to reduce the polynomial to degree n; such a reduction may

not be possible for general matrices. Interlacing may not generalize for nonnormal matrices, but certain max-min/min-max problems can still be posed, and though such problems may not be readily associated with eigenvalue problem as in the Hermitian case, they provide useful information for localizing Ritz values.

3.4 iFOV-2 for a 3×3 Normal Matrix

In the previous section we considered a 3×3 matrix having only one eigenvector. Here we consider a much nicer matrix, a normal matrix which has complete set of orthonormal eigenvectors. I will solve iFOV-2 for a 3×3 normal matrix, showing that in most cases, once z_1 has been specified, there is a unique choice for z_2. Contrast this with the Jordan block, see Figure 3.6, where specifying one of the Ritz values determined regions in $W(J_3)$ where the other Ritz value can fall. The results for normal matrices motivated the results involving marjorization of the phases of Ritz values. Some remarks will be made about solving iFOV-k for larger normal matrices.

General case, $k = n - 1$

Begin by assuming that A is a normal matrix with distinct eigenvalues. As normal matrices are unitarily diagonalizable, for normal A the field of values of $W(A)$ is the convex hull of $\sigma(A)$. Remember that Hermitian matrices are normal matrices that have all real eigenvalues, in which case $W(A)$ is a line. For a generic normal matrix having complex eigenvalues, $W(A)$ is some polygon in the complex plane [31].

Recall from Section 3.3 that a $n-1$ -dimensional subspace $\mathcal{V}$ of $\mathbb{C}^n$ is uniquely determined by any unit vector $v \in \mathbb{C}^n$ orthogonal to $\mathcal{V}$, and through the use of the adjugate, the characteristic polynomial of the restriction of A to $\mathcal{V}$ is

$$p(\lambda) = v^* \text{adj}(\lambda I - A) v.$$

Both Malamud [37] and Thompson [56] used these facts to characterize the eigenvalues of the $(n-1)$-dimensional principal submatrices of a normal matrix. We can assume, without loss of generality, that A is diagonal, $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then

$$p(\lambda) = \sum_{j=1}^{n} |v_j|^2 \prod_{k \neq j} (\lambda - \lambda_k), \tag{3.31}$$

hence the magnitudes of the entries of v determine the Ritz values from the orthogonal restriction of A to $\mathcal{V}$. As the eigenvalues are distinct, the $|v_j|^2$ may be thought of as the coefficients of $p(\lambda)$ in a Lagrange-like basis.

Evaluating $p(\lambda)$ at the eigenvalues reveals that for $z_1, \ldots, z_{n-1}$ to be the Ritz values of A, i.e., the roots of $p(\lambda)$, it is necessary and sufficient that the vector $x \in \mathbb{C}^n$, determined componentwise by

$$x_j = \frac{\prod_{\ell=1}^{n-1} (\lambda_j - \theta_\ell)}{\prod_{k \neq j} (\lambda_j - \lambda_k)}, \quad j = 1, \ldots, n, \tag{3.32}$$

have nonnegative entries. By construction, the x_j's determine a monic polynomial with $\sum x_j = 1$. If the entries of x are nonnegative, then $|v_i|^2 = x_i$, i.e., the entries of x_i are equal to the magnitude of the entries of v, squared. The vector v is only unique up to the magnitude of its entries, hence there are numerous subspaces that generate the same Ritz values; however they can all be related by diagonal unitary similarity transformation. Note that equation (3.32) breaks down if the eigenvalues of A are not distinct; however, a similar formulation using the distinct eigenvalues is possible.

An expansion of equation (3.31) for the coefficients of the λ in $p(\lambda)$ reveals that certain products and sums of the Ritz values must lie in the convex hull of the corresponding products and sums of the eigenvalues. This is ultimately how Malamud characterized the Ritz values of a normal matrix [37].

If a Ritz value coincides with an eigenvalue, then an eigenvector for that eigenvalue must be in $\mathcal{V}$, regardless of whether the eigenvalue is on the exterior or interior of $W(A)$, where an eigenvalue is *exterior* if it cannot be represented as a nontrivial convex combination of other eigenvalues and *interior* otherwise. For example, in the Hermitian case all points between the largest and smallest eigenvalues would be considered to be in the interior of $W(A)$, while the largest and smallest eigenvalues would be exterior. In the generic normal case, all eigenvalues at the corners of the boundary of $W(A)$ are exterior. So if a Ritz value coincides with an exterior eigenvalue, then an eigenvector for that eigenvalue must be in $\mathcal{V}$, as the only Ritz vector for an exterior eigenvalue is its eigenvector. For an eigenvalue in the interior of $W(A)$, there are numerous such vectors, most of them not eigenvectors. Thus, it is surprising in the normal case that having a Ritz value coincide with an eigenvalue requires that the corresponding eigenvector be in $\mathcal{V}$. Malamud's result for Ritz values of normal matrices provides an algebraic means of characterizing Ritz values. For $n = 3$, we provide a geometric interpretation of Malamud's result, which accounts for the Ritz values seen in Figure 3.5.

Ceva's theorem and iFOV-2

The requirement that equation (3.32) be real and positive is essentially a constraint on the phases of the Ritz values. This characterization of Ritz values of a normal matrix lends itself naturally to Ceva's theorem, a connection made by my colleague Derek Hansen and described in [13], from which this section is adapted. Let $A \in \mathbb{C}^{3\times 3}$ be normal, with eigenvalues λ_1, λ_2, and λ_3 that are distinct and not collinear, and let z_1 and z_2 be in the interior of $W(A)$. The x_j from (3.32) are real and nonnegative if

and only if $\arg(x_j) = 0$. Thus,

$$\begin{aligned}\arg(x_1) &= \arg\left(\frac{(\lambda_1 - z_1)(\lambda_1 - z_2)}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)}\right) \\ &= \arg(z_1 - \lambda_1) - \arg(\lambda_2 - \lambda_1) + \arg(z_2 - \lambda_1) - \arg(\lambda_3 - \lambda_1) \\ &= \angle(\lambda_2, \lambda_1, z_1) + \angle(\lambda_3, \lambda_1, z_2), \end{aligned} \tag{3.33}$$

and similarly for x_2 and x_3, where $\angle(\lambda_2, \lambda_1, z_1)$ denotes the signed angle from λ_2 to z_1 as centered about λ_1, being positive if counterclockwise and negative otherwise, such that $|\angle(\lambda_2, \lambda_1, z_1)| < \pi$, and similarly for $\angle(\lambda_3, \lambda_1, z_2)$. In classical geometry, a *Cevian* is a line segment joining the vertex of a triangle with a point on the opposite side [33]. We can interpret (3.33) to mean that $\arg(x_j) = 0$ if and only if z_1 and z_2 each lie on opposite Cevians that are reflections about the angle bisector through the vertex at λ_j of triangle $W(A)$; that is, z_1 and z_2 must lie on what are called *isogonal Cevians*; see Figure 3.8.

We now show that given any z_1 in the interior of $W(A)$, there exists one and only one other choice of $z_2 \in W(A)$ such that this geometric relationship is satisfied for all three angle bisectors. To this end, regard z_1 as a *Cevian point*, i.e., the point of intersection of three Cevians. Reflect each of these Cevians across the associated angle bisector to get the isogonal Cevians. According to Ceva's theorem [33], these three isogonal Cevians intersect at the same point 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 3.9. This equality must hold, since the original three Cevians are concurrent (at z_1). The point z_2 is known as the *isogonal conjugate* of the Cevian point z_1 [33]. Our observations are summarized in the following theorem.

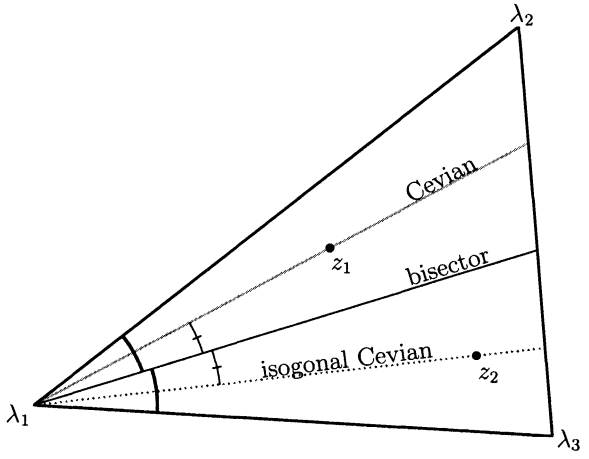


Figure 3.8 : The Cevian from λ_1 through z_2 must be the isogonal Cevian of the Cevian from λ_1 through z_1. From [13].

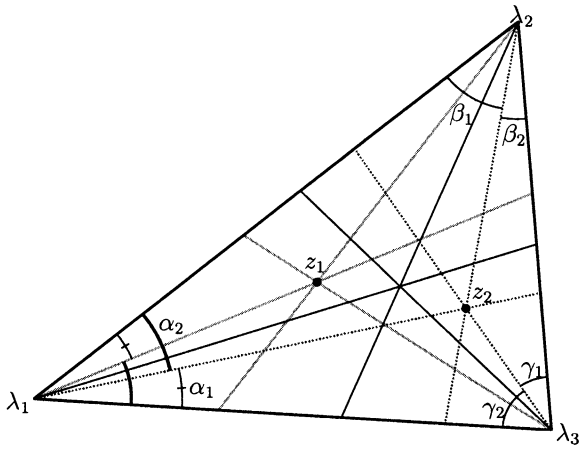


Figure 3.9 : The Cevians through z_2 are the isogonals of the Cevians through z_1. The Ritz value z_2 is the isogonal conjugate of z_1. From [13].

Theorem 3.6 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. For every* $z_1 \in W(A) \setminus \sigma(A)$, *there exists a unique* $z_2 \in W(A)$ *such that* $\sigma(V^*AV) = \{z_1, z_2\}$ *for some subunitary* $V \in \mathbb{C}^{3\times 2}$.

This theorem implies that any Ritz value $z_1 \in W(A) \setminus \sigma(A)$ *uniquely* determines the second Ritz value z_2. Note, if $z_1 \notin \sigma(A)$ is on an edge of the triangle $W(A)$, then z_2 must be the eigenvalue (vertex) opposite the edge. Likewise, if $z_1 \in \sigma(A)$, then z_2 can be *any* point on the opposite side. We also lose uniqueness when $W(A)$ is a degenerate triangle, that is, when A is a shifted and scaled Hermitian matrix, what some call *rotationaly Hermitian* [21]. Suppose λ_1, λ_2 and λ_3 are collinear with, say, λ_2 lying between λ_1 and λ_3. Divide $W(A)$ into two closed line segments: one joining λ_1 to λ_2, the other joining λ_2 to λ_3. The condition that

$$\angle(\lambda_2, \lambda_1, z_1) + \angle(\lambda_3, \lambda_1, z_2) = 0$$

(and similarly for λ_2 and λ_3) is equivalent to the statement that z_2 must lie on the opposite segment of z_1, but it may lie *anywhere* on this segment. In other words, z_1 and z_2 can be any pair that interlace the eigenvalues of A. Though perhaps at first surprising, this nonuniqueness of the choice of z_2 when $W(A)$ is degenerate is consistent with Theorem 3.6 as a limiting case; moreover, it exactly recovers the Cauchy interlacing theorem.

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

- Specify $z_1 \in W(A) \setminus \sigma(A)$;
- Draw the Cevians;
- Draw the Bisectors;
- Reflect Cevians across the angular bisectors to get the isogonal Cevians;

- z_2 lies at the intersection of the isogonal Cevians.

The result of this section accounts for Figure 3.5, where the eigenvalues of A are the cubic roots of unity. We are observing the phases of the Ritz values as measured about the eigenvalue $e^{-\pi i/3}$. Sorting the Ritz values by their phase, $\arg(z_i - e^{-\pi i/3})$, the result of this section gives that the Ritz values lie on opposite sides of the angular bisector at $e^{-\pi i/3}$.

iFOV-$(n-1)$ for larger normal matrices

Building on some of the results of the previous section, we will discuss some numerical approaches to studying $n-1$ Ritz values for an $n > 3$ dimensional normal matrix. The key feature that sets Ritz values of normal matrices apart from Ritz values of general matrices is evident from equation (3.31). The possible $n-1$ complex Ritz values can be thought of as $2n-2$ real numbers, the real and imaginary parts of the z_i. From equation (3.32) the possible $n-1$ Ritz values of a normal matrix can be parametrized using n 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 $\|v\| = 1$ and that only $n-1$ real numbers are required. For $n = 3$ and the eigenvalues not collinear, this resulted in one Ritz value almost always uniquely determining the other. For $n > 3$, one would expect that specifying roughly half of the Ritz values should determine the other half. Similar in spirit to the previous section, we provide recipes to determine where the remaining Ritz values may lie provided we have specified at most half of them.

From the properties of the field of values, for any $z \in W(A)$ one can construct a vector v such that $v^*Av = z$. For a normal matrix, assuming A is diagonal, such a v

must satisfy

$$\begin{bmatrix} \operatorname{Re}\lambda_1 & \cdots & \operatorname{Re}\lambda_n \\ \operatorname{Im}\lambda_1 & \cdots & \operatorname{Im}\lambda_n \\ 1 & \cdots & 1 \end{bmatrix} x = \begin{bmatrix} \operatorname{Re} z \\ \operatorname{Im} z \\ 1 \end{bmatrix}, \tag{3.34}$$

where $x_i = |v_i|^2$. Since the matrix has dimension $3 \times n$, the set of all possible x that satisfy equation (3.34) 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 (3.34) are linearly dependent, and the affine 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 then affine subspace for x is $n-1$ dimensional. The intersection of this affine subspace with the positive orthant determines all v that would generate z.

Suppose we wish to have a Ritz value at z. To determine where the remaining Ritz values must lie, we pick a v from the affine subspace corresponding to z. We would like to determine a $\widehat{V}$ such that $V = [v\,\widehat{V}] \in \mathbb{C}^{n\times(n-1)}$ has orthonormal columns and so that

$$[v\,\widehat{V}]^* A [v\,\widehat{V}] = \begin{bmatrix} z & v^*A\widehat{V} \\ 0 & \widehat{V}^*A\widehat{V} \end{bmatrix}, \tag{3.35}$$

in which case z is a Ritz value of V^*AV and the remaining Ritz values of V^*AV are the eigenvalues of $\widehat{V}^*A\widehat{V}$. For this to happen, the range of $(A - zI)v$ must be orthogonal to V, which is equivalent to having $\operatorname{Ran}(\widehat{V})$ be the kernel of $[v\,Av]^*$. Hence the remaining Ritz values are entirely dependent upon the particular v. To determine where the Ritz values may lie given only z, this process must be repeated for additional v; for an example, see Figure 3.10. The union of the field of values of the $\widehat{V}^*A\widehat{V}$ for all such v would indicate where the remaining Ritz values must lie for any size restriction having one prescribed Ritz value at z.

One thing should be clear from Figure 3.10: 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 (perhaps we already have part of a valid combination), then the above procedure is unsuitable. Starting with an affine space corresponding to one of the Ritz values, we would then have to search through all possible subspaces to determine those that had the second Ritz value. Using the adjugate approach, we may do something similar in spirit to the procedure above. Suppose that we have some $z_1, \ldots, z_p$ for $p \leq \lfloor (n-1)/2 \rfloor$, and we wish to determine all possible V such that $z_1, \ldots, z_p$ are the Ritz values from some $V^*AV \in \mathbb{C}^{(n-1)\times(n-1)}$. Then equation (3.31) gives the form of the characteristic polynomial $p(\lambda)$ of V^*AV in terms of the vector v orthogonal to the range of V. If the $z_1, \ldots, z_p$ are to be roots of $p(\lambda)$, then v must satisfy

$$\begin{bmatrix} \operatorname{Re}\ell_1(z_1) & \cdots & \operatorname{Re}\ell_n(z_1) \\ \operatorname{Im}\ell_1(z_1) & \cdots & \operatorname{Im}\ell_n(z_1) \\ \vdots & & \vdots \\ \operatorname{Re}\ell_1(z_p) & \cdots & \operatorname{Re}\ell_n(z_p) \\ \operatorname{Im}\ell_1(z_p) & \cdots & \operatorname{Im}\ell_n(z_p) \\ 1 & \cdots & 1 \end{bmatrix} x = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{3.36}$$

where again $x_i = |v_i|^2$ and $\ell_i(\lambda) = \prod_{j\neq i}(\lambda - \lambda_j)$. Equation (3.36) determines an affine subspace of $\mathbb{R}^n$ for x of at least dimension $n - 2p - 1$. If this subspace intersects the positive orthant, then there exists V such that $z_1, \ldots, z_p$ are Ritz values of V^*AV. To determine where the remaining Ritz values lie for all possible V, one would explore the affine subspace for v.

One must contrast these results with those for the Jordan block. In the normal case, due to A being unitarily diagonalizable, there exist affine subspaces of $\mathbb{R}^n$ associ-

ated with $n-1$ dimensional subspaces having prescribed Ritz values. Also, specifying some of the Ritz values can significantly reduce where the remaining Ritz values may lie. In the nonnormal case, as evident from the Jordan block, there is a nonlinear relationship between the Ritz values and the subspaces that generate them, and specifying one of the Ritz values may only marginally restrict where the remaining Ritz values may lie.

These examples illustrate the difficulties in analyzing Ritz values of non-Hermitian matrices. Chapter 4 uses the bounds developed in this chapter to analyze the restarted Arnoldi method with exact shifts, giving sufficient conditions for the convergence. One of the restarted Arnoldi examples will rely on the nice features of the normal case.

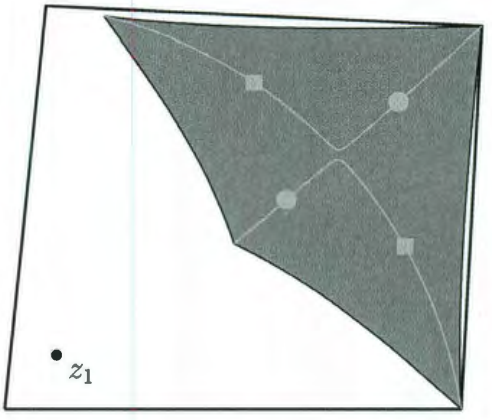


Figure 3.10 : For restrictions onto 2-dimensional subspaces z_2 can fall anywhere in the gray regions. For restrictions onto 3-dimensional subspaces, the light gray curves inside the gray region indicate where z_2 and z_3 have to lie. The correspondence between the points on the two curves is one-to-one: if z_2 is at the the square along the top curve, then z_3 must lie at the square on the bottom curve.

Chapter 4

Convergence of the Restarted Arnoldi Method

Having established in the last chapter various majorization-based bounds for the Ritz values of general nonsymmetric matrices, I exploit these bounds to develop some sufficient conditions for convergence of the restarted Arnoldi method with exact shifts.

The problem of determining a few eigenvalues of a non-Hermitian matrix using an iterative method such as restarted Arnoldi is complicated by the nonnormality of the eigenvalues: both those that are desired, which restarted Arnoldi seeks to compute, and those undesired, which restarted Arnoldi suppresses via the restart polynomial. The possibility of failure or stagnation adds further challenges. The nonnormality of eigenvalues reflects how sensitive the eigenvalues are to perturbations in the matrix. The possibility of failure is dependent upon whether the starting vector leads to either "lucky breakdown," in which case a desired eigenspace has been found, or misconvergence to undesired eigenvalues. In applications, additional issues arise due to the finite precision of floating point arithmetic and the cost of performing real versus complex arithmetic. Such concerns necessitate modifications to the algorithm, such as reorthogonalization to counteract the loss of orthogonality due to finite precision, and double shifts to avoid complex arithmetic.

Addressing all the factors above would be a rather daunting task; in this chapter I address some of these issues. First, I present examples demonstrating two different types of failure. These examples represent worse case scenarios that limit what can be said when dealing with general matrices. The first example demonstrates the possibil-

ity of stagnation: the Ritz values converge, but not to eigenvalues. This type of failure is dependent on the starting vector. The second example comes from Embree [20] and involves extreme breakdown: the restart polynomial annihilates the desired eigenvector from the starting vector, thereby precluding the possibility of convergence to the desired eigenvalue. This type of failure is due to the wanted eigenvalue being in the field of values of the matrix restricted to the unwanted invariant subspace. With my knowledge of Ritz values, to eliminate this type of failure I can make restrictions on the properties of the matrix and the size of the subspace used for the restarted Arnoldi method. The projection property of the field of values is used to improve upon the bounds of the previous chapter. Throughout I assume exact arithmetic, in which case the implicitly restarted Arnoldi, explicitly restarted Arnoldi and restarted Krylov–Schur methods are all mathematically equivalent.

Since in practical applications the desired eigenvalues tend to be well-conditioned, I will consider matrices that have a *simple normal eigenvalue.* An eigenvalue is *simple* if it has multiplicity one. An eigenvalue is *normal* if its eigenvector is orthogonal to the complementary invariant subspace associated with the other eigenvalues. Hence, the class of matrices I consider are all unitarily similar to a block diagonal matrix with diagonal entries λ and D:

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & D \end{pmatrix}, \tag{4.1}$$

where $\sigma(D)$ contains all the unwanted eigenvalues and we wish to compute the eigenvalue λ. Without loss of generality, assume λ is real and nonnegative. (Later this will be generalized to allow for more wanted eigenvalues, as well as a nonnormal coupling between the wanted eigenvalue and the block associated with the unwanted eigenvalues.) Ultimately, we will assume that $\|D\| < \lambda$, so λ is both the rightmost and largest

magnitude eigenvalue of A. Note that since $\lambda > \|D\|$, our analysis also applies to the computation of the largest magnitude eigenvalue of A.

The development of a convergence theory for the matrices I consider will proceed in the following manner. I will establish that there is a Ritz value near the wanted eigenvalue, then I will show that the other Ritz values cannot be arbitrarily close to the wanted eigenvalue. These results lead to conditions on the spectrum and on the starting vector that together ensure convergence. While these conditions may seem quite strong for restarted Arnoldi, it is likely that they are reasonable for shift invert Arnoldi. To test my results, I will consider a case where D is skew symmetric, $D^* = -D$. In this case A is normal, and the behavior of restarted Arnoldi with exact shifts can be worked out explicitly.

4.1 Examples

In this section, two examples will be considered; one demonstrates extreme breakdown, and the other demonstrates stagnation. All these involve computing the eigenvalue with largest real part. In each example, the wanted eigenvalue is simple and normal, and thus the matrices in question could each be presented in the block diagonal form (4.1).

4.1.1 Stagnation

In this section I will present a matrix and starting vector for which the restarted Arnoldi method stagnates. Consider

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

a circulant matrix whose largest real eigenvalue $\lambda = 1$ has an eigenvector with equal components in each entry. Use the restarted Arnoldi method with one exact shift to compute the largest eigenvalue with the starting vector

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

This gives $\mathcal{K}_2(A, v_1) = \text{span}\{v_1, Av_1\} = \text{span}\{v_1, v_2\}$, where

$$v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The upper Hessenberg matrix H_2, the restriction of A onto $K_2(A, v_1)$ using $V_2 = [v_1\ v_2]$, then is

$$H_2 = V_2^* A V_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Clearly H_2 has but one eigenvalue, thus $z_1 = z_2 = 0$. Using an exact shift of zero to generate the new starting vector,

$$v_1^{(2)} = v^+ = (A - 0I)v_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

where the superscript denotes that $v_1^{(2)}$ is the starting vector for the second iteration of the restarted Arnoldi method.

For the second iteration, the Arnoldi basis vectors are

$$v_1^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad v_2^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

As in the previous iteration, the restriction of A to the current Krylov subspace is

$$H_2^{(2)} = (V_2^{(2)})^* A V_2^{(2)} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

As before, both Ritz values are zero. Proceeding with further restarted Arnoldi cycles produces the successive starting vectors

$$v_1^{(3)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_1^{(4)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Thus at the fourth cycle of restarted Arnoldi, the starting vector $v_1^{(4)} = v_1^{(1)}$: the new starting vector is equal to the first starting vector. Hence, for this example the restarted Arnoldi method stagnates, and the Ritz value never converges to an eigenvalue, wanted or unwanted. All Ritz values are zero.

This example is particularly striking because A is a normal matrix with a unique rightmost eigenvalue $\lambda = 1$, and all the other eigenvalues are in the left half plane. If put into the form (4.1), then $\lambda \notin W(D)$. The starting vector v_1 has a significant component in the desired eigenvector direction; in fact, the problem arises because v_1 is equally weighted in each of the eigenvectors. Moreover, this example readily generalizes to n-dimensional circulant shift matrices with Krylov subspaces of dimension k for $2 \leq k < n$. This matrix is also related to a well known example of stagnation for GMRES; see [9].

In this example and its generalization to higher dimensions, the Ritz values are as far as possible from the eigenvalues. They represent the worst that one could do in terms of gaining information about the eigenvalues from the Ritz values. Any bounds for Ritz values will have to work around such benign-looking examples.

If one were to alter the starting vector slightly, making it closer to the desired eigenvector, then the restarted Arnoldi method would converge. The stagnation demonstrated is not a stable fixed point. This example suggests that for some matrices there exist criteria for *local convergence*. In other words, if the starting vector is sufficiently rich in the desired eigenvector, then the restarted Arnoldi process will converge. Later in this chapter, I will consider a class of matrices for which local convergence as well as stagnation can occur.

4.1.2 Extreme failure

This example is taken from Embree [20] and demonstrates extreme breakdown.

Consider the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 6 & -2 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.2}$$

of the form (4.1) with largest eigenvalue $\lambda = 1$ and corresponding eigenvector e_1. Using the restarted Arnoldi algorithm with one exact shift to compute the largest eigenvalue with a starting vector that has equal components in each entry leads to the following Arnoldi basis for $\mathcal{K}_2(A, v_1)$:

$$v_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \frac{1}{2\sqrt{35}}\begin{pmatrix} -3 \\ 9 \\ 1 \\ -7 \end{pmatrix}.$$

Restricting the matrix A to $\mathcal{K}_2(A, v_1)$ gives

$$H_2 = V_2^* A V_2 = \begin{pmatrix} 7/4 & 3/(4\sqrt{35}) \\ \sqrt{35}/4 & 5/4 \end{pmatrix}.$$

The characteristic polynomial of H_2 is

$$p(z) = \det(zI - H_2) = z^2 - 3z + 2 = (z-1)(z-2).$$

Thus the eigenvalues of H_2 are $z_1 = 1$, $z_2 = 2$. The strategy for computing the rightmost eigenvalue would use z_1 as the exact shift. Since $z_1 = \lambda$, this particular shift results in the new starting vector

$$v^+ = (A - z_1 I)v_1 = \begin{pmatrix} 0 \\ 3 \\ 1 \\ -1 \end{pmatrix},$$

which does not have a component in e_1, the eigenvector associated with the rightmost eigenvalue. Due to the structure of A, all further starting vectors of the restarted Arnoldi method will be orthogonal to e_1. Hence convergence to e_1 for this particular starting vector, v_1, is impossible. This failure is not unique to just this particular starting vector. Failure can also occur for any vector of the form

$$v_1 = \frac{1}{\sqrt{\alpha^2 + 3}} \begin{pmatrix} \alpha \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

where α is any scalar. This form shows that the starting vector can be arbitrarily rich in the desired eigenvector and yet restarted Arnoldi can still fail to converge

to the desired eigenvalue. Similarly, in Section 4.3, we show a normal matrix for which the starting vector can be arbitrarily rich and the desired eigenvector, and yet initially the restarted Arnoldi method will diverge. Such examples are troubling for the convergence theory of the restarted Arnoldi method for general matrices. Unlike the previous example involving stagnation, local convergence is not possible for this matrix.

Embree went on to generalize this example, allowing for more desired eigenvalues and more shifts. In all his examples, this type of failure occurs where the wanted eigenvalues are in the field of values of the portion of the matrix associated with the unwanted eigenvalues. All of his examples involved block diagonal matrices, $A = \text{diag}(T, D)$: a block T with wanted eigenvalues and a block D such that $\sigma(T) \subset W(D)$. Note that in the notation of (4.1), $\lambda \in W(D)$ for the matrix (4.2). In Section 4.2.1, we provide a criterion to help rule out this type of failure. Note that eigenvalues of the Hermitian part of A in (4.2) are -3.5616, $.5616$, 1 and 3. For subspaces up to size three, equation (3.4) indicates that the leftmost Ritz value could still fall on the eigenvalue $\lambda = 1$. For this matrix, using a larger subspace may not reduce the possibility of failure.

4.2 The General Case

Having shown two types of failure for the restarted Arnoldi method, in this section I develop a convergence theory for a class of matrices that addresses the more serious type of failure. Throughout this section, assume that $\lambda \in \mathbb{R}$ is not in the field of values of D, and that $\|D\| < \lambda$. For simplicity, I always assume we are computing a single rightmost eigenvalue, and hence will use all but one Ritz value as exact shifts.

The development of the convergence theory rests upon the localization of the Ritz

values. I show there must be a Ritz value within a certain distance of the wanted eigenvalue, and that the rest of the Ritz values are bounded away from the desired eigenvalue. Sufficient criteria for convergence are then based upon these localization results. Throughout this section I assume A has the form (4.1), and the starting vector v is represented as

$$v = \begin{pmatrix} c \\ r \end{pmatrix},$$

where $c \in \mathbb{C}$ is a nonzero scalar that the gives component of the starting vector in the direction of the desired eigenvector, e_1, and $r \in \mathbb{C}^{n-1}$ is the rest of the starting vector.

4.2.1 Ritz Value Localization

In this section I prove two lemmas that localize the Ritz values. (For theorems that localize Ritz values in more general settings, see Chapter 3.) The first lemma shows that not all the Ritz values can be arbitrarily far away from the desired eigenvalue.

Lemma 4.1 For a Krylov subspace $\mathcal{K}_k(A, v)$, there must exist at least one Ritz value, z_1, that is within η of the desired eigenvalue, $|z_1 - \lambda| \leq \eta$, where

$$\eta := (\|D\| + \lambda)\left(\frac{\|r\|}{|c|}\right)^{\frac{1}{k}}. \tag{4.3}$$

Proof. Ritz values from a Krylov subspace are optimal, in the sense that they are the roots of the monic polynomial that minimizes

$$\left\| \prod_{i=1}^{k} (A - z_i I) v \right\| = \min_{p \in \mathbb{P}_k} \|p(A)v\|,$$

where $\mathbb{P}_k$ is the set of all monic polynomials of degree k [47]. Define $\hat{\eta} := \min_{i\in\{1,\ldots,k\}} |\lambda - z_i|$. Due to the block diagonal structure of A,

$$\begin{aligned}\left|\prod_{i=1}^{k}(\lambda - z_i)c\right|^2 &\leq \left|\prod_{i=1}^{k}(\lambda - z_i)c\right|^2 + \left\|\prod_{i=1}^{k}(D - z_i I)r\right\|^2 \\ &= \min_{p\in\mathbb{P}_k} \|p(A)v\|^2.\end{aligned}$$

Thus

$$\hat{\eta}^k|c| \leq \min_{p\in\mathbb{P}_k} \|p(A)v\|.$$

Since the Ritz values are optimal, no other monic polynomial $\hat{p}$ with different roots can produce a smaller norm, so taking $\hat{p}(z) = (z - \lambda)^k$, one obtains

$$\min_{p\in\mathbb{P}_k} \|p(A)v\| \leq \|(D - \lambda)^k r\|;$$

this comes from the fact that this particular $\hat{p}$ annihilates the first component of the starting vector. Applying the definition of the operator norm and the triangle inequality, the term on the right gives

$$\min_{p\in\mathbb{P}_k} \|p(A)v\| \leq (\|D\| + \lambda)^k \|r\|.$$

Combining the bounds from above and below for $\min_{p\in\mathbb{P}_k} \|p(A)v\|$ yields

$$\hat{\eta}^k|c| \leq (\|D\| + \lambda)^k \|r\|.$$

This implies that $\hat{\eta} \leq (\|D\| + \lambda)(\|r\|/|c|)^{1/k}$, indicating that at least one Ritz value must be less than a distance of $(\|D\| + \lambda)(\|r\|/|c|)^{1/k}$ from λ. Denoting the closest Ritz value to λ as z_1, we see $|\lambda - z_1| \leq \eta$. ■

This lemma suggests that as the size of the subspace increases, there need not be a Ritz value any closer than $\|D\| + \lambda$ from λ. Such behavior should be expected; recall the example in Section 4.1.1, where all the Ritz values were as far as possible from the eigenvalues.

The next lemma localizes the exact shifts, i.e., the Ritz values $z_2, \ldots, z_k$. The weak majorization result, equation (3.3), is used.

Lemma 4.2 Let $\vartheta_1, \vartheta_2, \ldots, \vartheta_{n-1}$ denote the eigenvalues of the Hermitian part of A from equation (4.1) with $\vartheta_1 \geq \cdots \geq \vartheta_{n-1}$, and η be as in Lemma 4.1. If the Ritz values are ordered such that $\operatorname{Re} z_1 \geq \cdots \geq \operatorname{Re} z_k$, then for each z_j, $j = 2, \ldots, k$,

$$\operatorname{Re} z_j \leq f(j,\eta) \;:=\; \eta + \sum_{i=2}^{j-1} (\vartheta_i - \operatorname{Re} z_i) + \vartheta_j. \tag{4.4}$$

Furthermore

$$|z_j| \leq \rho(j,\eta) := \sqrt{f(j,\eta)^2 + \mu(D)^2}, \tag{4.5}$$

where $\mu(D) := \max_{z \in W(D)} |z|$ is the numerical radius of D.

Proof. From the weak majorization of the real parts of the Ritz values by eigenvalues of the Hermitian part of A, equation (3.3), we have

$$\sum_{i=1}^{j} \operatorname{Re} z_i \;\leq\; \sum_{i=1}^{j} \vartheta_i,$$

that is,

$$\operatorname{Re} z_1 + \sum_{i=2}^{j-1} \operatorname{Re} z_i + \operatorname{Re} z_j \;\leq\; \vartheta_1 + \sum_{i=2}^{j-1} \vartheta_i + \vartheta_j.$$

For A of the form (4.1), we have $\lambda = \vartheta_1$, so

$$\operatorname{Re} z_j \;\leq\; \lambda - \operatorname{Re} z_1 + \sum_{i=2}^{j-1} (\vartheta_i - \operatorname{Re} z_i) + \vartheta_j, \tag{4.6}$$

for $j = 1, \ldots, k$. From Lemma 4.1, $\lambda - \operatorname{Re} z_1 \leq \eta$, and equation (4.4) follows. Bounding $\operatorname{Re} z_j$ with equation (4.4) and $\operatorname{Im} z_j$ with $\mu(D)$ gives equation (4.5). ∎

As z_1 gets closer to λ, the bound on the other Ritz values lies closer to $W(D)$. There are several simpler bounds that follow from Lemma 4.2. Rather than derive them as needed, we state them all here.

Corollary 4.1 The quantity $f(j,\eta)$ from Lemma 4.2 for $j = 2,\ldots,k$ satisfies

$$f(j,\eta) \leq \frac{\eta + \vartheta_2 + \cdots + \vartheta_j}{j-1} \tag{4.7}$$

$$\leq \frac{\lambda + \vartheta_2}{2}. \tag{4.8}$$

Proof. Both bounds follow from the ordering of the Ritz values. From rearranging (4.6), for $j = 2,\ldots,k$, we have

$$\sum_{i=2}^{j} \operatorname{Re} z_i \leq \lambda - \operatorname{Re} z_1 + \sum_{i=2}^{j} \vartheta_i,$$

and hence from the ordering of the $\operatorname{Re} z_j$

$$(j-1)\operatorname{Re} z_j \leq \lambda - \operatorname{Re} z_1 + \sum_{i=2}^{j} \vartheta_i,$$

from which equation (4.7) follows. Equation (4.8) follows similarly. ∎

These bounds allow us to assume that we know nothing about the shifts $z_2,\ldots,z_k$. The quantity $\rho(j,\eta)$ in Lemma 4.2, gives a bound for the magnitude of the shifts such that $\rho(j,\eta) > \mu(D)$ for all η. A more reasonable bound having the property that $\rho(j,\eta) \to \|D\|$ as $\eta \to 0$ can derived using the weak majorization of the magnitudes of the Ritz value by the singular values of the matrix.

Lemma 4.3 Let $\lambda > \|D\|$ and η be as in Lemma 4.1, and let $\sigma_1 \geq \cdots \geq \sigma_n$ denote the singular values of A. If the Ritz values are ordered such that $\operatorname{Re} z_1 \geq \cdots \geq \operatorname{Re} z_k$, then for each z_j, $j = 2,\ldots,k$,

$$|z_j| \leq \eta + \mu(D),$$

where $\mu(D)$ is the numerical radius of D.

Proof. Let $|\hat{z}_1| \geq \cdots \geq |\hat{z}_k|$ denote the Ritz valuess ordered by magnitude. Recall the weak log-majorization of the magnitudes of the Ritz values by the singular values

of the matrix, see equation (3.5):

$$\prod_{i=1}^{j} |\hat{z}_i| \leq \prod_{i=1}^{j} \sigma_i,$$

for $j = 1, \ldots, k$. The two largest singular values of A are $\sigma_1 = \lambda$ and $\sigma_2 = \|D\|$. As weak log-majorization implies weak majorization,

$$\sum_{i=1}^{j} |\hat{z}_i| \leq \sum_{i=1}^{j} \sigma_i,$$

for $j = 1, \ldots, k$. Subject to the two different orderings, either $\hat{z}_1 = z_1$ or $\hat{z}_1 \neq z_1$. In the first case,

$$|\hat{z}_1| + |\hat{z}_2| \leq \lambda + \|D\|.$$

From this we may be bound the magnitude of the second largest shift:

$$\begin{aligned} |\hat{z}_2| &\leq \lambda - |\hat{z}_1| + \|D\| \\ &\leq \lambda - |z_1| + \|D\|. \end{aligned}$$

Using the triangle inequality for $|z_1|$ and recalling from Lemma 4.1 that $|\lambda - z_1| < \eta$ gives that

$$\begin{aligned} |\hat{z}_2| &\leq \lambda - \operatorname{Re} z_1 + \|D\| \\ &\leq \eta + \|D\|. \end{aligned}$$

Now suppose that $\hat{z}_1 \neq z_1$. Then

$$\begin{aligned} |\hat{z}_1| + |z_1| &\leq |\hat{z}_1| + |\hat{z}_2| \\ &\leq \lambda + \|D\|. \end{aligned}$$

Arguing as in the first case then gives the desired result:

$$|\hat{z}_1| \leq \eta + \|D\|.$$

Thus the magnitudes of the shifts are such that

$$|z_j| \leq \eta + \|D\|$$

for $j = 2, \ldots, k$. ■

With our bounds for the wanted and unwanted Ritz values, we may now construct a convergence theory for the restarted Arnoldi method.

Implications for Arnoldi Convergence

Building upon the lemmas above, in this section I demonstrate a condition sufficient for convergence of the restarted Arnoldi method with exact shifts. The result gives conditions on the starting vector and the spectrum of the matrix which, if satisfied, guarantee that the restarted Arnoldi method with exact shifts will converge.

To ensure convergence, I determine conditions on A and starting vector v that ensure the tangent of the angle Φ between the desired eigenvector and the Krylov subspace decreases at each restart. For the model problem (4.1), the desired eigenspace is spanned by the first canonical vector e_1. Again, write the starting vector as

$$v = \begin{pmatrix} c \\ r \end{pmatrix},$$

where $c \in \mathbb{C}$ and $r \in \mathbb{C}^{n-1}$ are such that $\|v\| = 1$. Then the tangent of the angle between v and e_1 is simply

$$\tan \Phi = \frac{\|r\|}{|c|}.$$

Thus convergence occurs when the ratio of the norm of r to the absolute value of c is driven to zero by successive restarts. The relationship between the starting vector from one cycle to the next involves a restart polynomial, ψ, whose roots are the $k-1$

leftmost Ritz values of A restricted to the Krylov subspace $\mathcal{K}_k(A, v)$. The starting vector, at the next iteration is

$$v^+ = \frac{\psi(A)v}{\|\psi(A)v\|}.$$

Due to the structure of A,

$$\psi(A)v = \begin{pmatrix} c\psi(\lambda) \\ \psi(D)r \end{pmatrix}.$$

Using $p = k - 1$ exact shifts, Lemma 4.2 indicates that each shift z_j has real part less than or equal to $f(j, \eta)$ and magnitude bounded by both $\rho_j = \sqrt{f(j,\eta)^2 + \mu(D)^2}$ and $\eta + \|D\|$. Each of these bounds depends on the starting vector through the quantity η. Having the tangent of the angle between the starting vector and the desired eigenvector decrease at a prescribed rate $\gamma \in (0, 1)$ at each step is equivalent to having

$$\frac{\|\psi(D)r\|}{|c\psi(\lambda)|} \leq \gamma < \frac{\|r\|}{|c|},$$

for any valid exact shift polynomial ψ. Thus for convergence, it is sufficient to show that there exists some $\gamma \in (0, 1)$ such that

$$\frac{\|\psi(D)r\|}{\|r\||\psi(\lambda)|} \leq \gamma < 1 \tag{4.9}$$

for all exact shift restart polynomials ψ. With this notation in place, the following theorem employs the bounds on the shifts z_j for $j = 2, \ldots, k$ to determine conditions that ensure that if $\tan\Phi$ is sufficiently small at the first step, then the restarted Arnoldi method with exact shifts will converge.

Theorem 4.1 Let $\gamma \in (0, 1)$ and η be as in Lemma 4.1. Suppose $\lambda > 2\|D\|\gamma^{-1/p} + \vartheta_2$. If $\eta < (\gamma^{1/p}(\lambda - \vartheta_2) - 2\|D\|)/(1 + \gamma^{1/p})$, then the tangent of the angle between the starting vector and the desired eigenvector will decrease by a factor of at least γ at each restart.

Proof. The bound (4.9) is implied by the more stringent criterion

$$\frac{\|\psi(D)\|}{|\psi(\lambda)|} \leq \gamma < 1. \tag{4.10}$$

This quantity is too difficult to work with; hence to generate an even stronger criterion, recall that $\psi(z) = \prod_{i=2}^{k}(z - z_i)$, where the z_i are the exact shifts, the unwanted Ritz values. If we have a upper bound $\hat{z}$ for the real part of all the exact shifts, then the worst possible scenario for the denominator of (4.10) would be that all the shifts occur at $\hat{z}$, for this would minimize the distance between the shifts and λ. If we have a upper bound $\hat{\rho}$ for the magnitude of all the shifts, then the numerator of (4.10) can also be bounded. Each term in $\psi(D) = \prod_{i=2}^{k}(D - z_i I)$ satisfies

$$\|D - z_i I\| \leq \|D\| + |z_i| \leq \|D\| + \hat{\rho},$$

hence

$$\frac{\|\psi(D)\|}{|\psi(\lambda)|} \leq \frac{(\|D\| + \hat{\rho})^p}{(\lambda - \hat{z})^p}.$$

These bounds give an even stronger criterion that implies (4.10): If

$$\frac{(\|D\| + \hat{\rho})^p}{(\lambda - \hat{z})^p} < \gamma, \tag{4.11}$$

then

$$\frac{\|\psi(D)\|}{|\psi(\lambda)|} \leq \gamma.$$

Taking the pth root of both sides of (4.11) gives

$$\frac{\|D\| + \hat{\rho}}{\lambda - \hat{z}} < \gamma^{\frac{1}{p}}. \tag{4.12}$$

Rearranging equation (4.12) leads to the inequality

$$\hat{z}\gamma^{\frac{1}{p}} + \hat{\rho} < \lambda\gamma^{\frac{1}{p}} - \|D\|.$$

From Lemma 4.2, the real parts of all the shifts are bounded by $f(2,\eta) := \eta + \vartheta_2$, which can be used as the bound $\hat{z}$. From Lemma 4.3, the magnitude of all the shifts can be bounded by $\eta + \|D\|$, which can be used for $\hat{\rho}$. Thus, progress is ensured if

$$(\eta + \vartheta_2)\gamma^{\frac{1}{p}} + \eta + \|D\| < \lambda\gamma^{\frac{1}{p}} - \|D\|. \tag{4.13}$$

Rearranging the expression gives a criterion for η:

$$\eta < \frac{\gamma^{\frac{1}{p}}(\lambda - \vartheta_2) - 2\|D\|}{1 + \gamma^{\frac{1}{p}}}. \tag{4.14}$$

This criterion implies that if

$$\lambda > \vartheta_2 + 2\|D\|\gamma^{-\frac{1}{p}}$$

and if the starting vector v is sufficiently rich in the desired eigenvector, then the tangent of the angle between the new starting vector v^+ and e_1 will be at least a factor γ less than the tangent of the angle between v and e_1. This criterion is sufficient for convergence of the restarted Arnoldi method using p exact shifts at a rate no worse than γ. Since (4.14) and (4.14) ensures that η decrease when restarting, these criteria will thus be satisfied at the next cycle of restarted Arnoldi and thus the method will converge.

Exactly how rich the starting vector must be in the desired eigenvector can be determined from the expression for η:

$$\eta = (\|D\| + \lambda)\left(\frac{\|r\|}{|c|}\right)^{\frac{1}{k}} \leq \frac{\gamma^{\frac{1}{p}}(\lambda - \vartheta_2) - 2\|D\|}{1 + \gamma^{\frac{1}{p}}},$$

which implies that

$$\frac{\|r\|}{|c|} \leq \left(\frac{\gamma^{\frac{1}{p}}(\lambda - \vartheta_2) - 2\|D\|}{(1 + \gamma^{\frac{1}{p}})(\|D\| + \lambda)}\right)^{k}.$$

The quantity γ is a bound for the initial rate of convergence. The bound for the rate of convergence improves with each cycle of restarted Arnoldi. In the limit as $\eta \to 0$, (4.11) combined with the bounds $\hat{z}$ and $\hat{\rho}$ determine a bound for the asymptotic rate of convergence

$$\frac{2\|D\|}{\lambda - \vartheta_2}. \quad \blacksquare$$

The criterion in Theorem 4.1 does not seem particularly sharp, that is, the wanted eigenvalue must be well separated from the unwanted eigenvalues. The proof involved bounding $\|\psi(D)r\|/\|r\|$ with $\|\psi(D)\|$. Requiring $\|\psi(D)r\|/\|r\|$ to be small, depending on r, may necessitate only that ψ be small on some of the unwanted eigenvalues of D, whereas requiring $\|\psi(D)\|$ to be small means that ψ must be small on all the unwanted eigenvalues. In bounding $\|\psi(D)\|/|\psi(\lambda)|$, each of the shifts was treated independently, twice; different bounds were used for the numerator and in the denominator of equation (4.10). The bounds limited the magnitude of the shifts as well as the proximity of the shifts to λ. Such bounds are unlikely to be attained simultaneously for all the shifts. A sharper bound would require bounding the expression as a whole, and treating the shifts as an ensemble rather than independently. The latter could be done using $f(j, \eta)$ for each of the Ritz values; however this would significantly increase the complexity of the resulting expressions. Nonetheless, the analysis above (begun in [11]) do indeed give criteria that ensure convergence of the restarted Arnoldi algorithm with exact shifts, the first such results of which I am are aware.

Improvements on the bound

Several things can be done to extend the result of Theorem 4.1. Here we consider keeping more shifts, computing more normal eigenvalues, and using interlacing of the boundary of the field of values to improve the bounds for the exact shifts.

Suppose that instead of keeping just one shift at each restart and using the rest as exact shifts, that we instead use the $p \leq k-2$ right most Ritz values as exact shifts. Then without assuming any further knowledge about the $k-p$ rightmost Ritz values that are not being used as shifts, we have from Corollary 4.1 that real part of each of the exact shifts z_j will be bounded by

$$f(j,\eta) := \frac{\eta + \vartheta_2 + \cdots + \vartheta_j}{j-1}.$$

As we are keeping $k-p$ of the Ritz values, a bound for the real parts of the exact shifts $z_{k-p+1}, \ldots, z_k$ would be $f(k-p+1,\eta)$. Using the same arguments as in Theorem 4.1, assuming that the starting vector is sufficiently rich in the desired eigenvector, the restarted Arnoldi method would converge if

$$\frac{\vartheta_2 + \cdots + \vartheta_{k-p+1}}{k-p} + 2\|D\|\gamma^{-\frac{1}{p}} < \lambda. \tag{4.15}$$

From equation (4.15), we can see that as we keep more of the shifts, the criteria for convergence is not as strong, i.e., λ need not lie so far from the unwanted eigenvalues. Note that in the limit, the quantity on the left of (4.15) is the real part of the trace of D. This suggests that all the criteria for convergence could be made tighter by an appropriate shift of the whole matrix. There is an interesting implication of this result and Corollary 4.1: if we are concerned about having shifts near a wanted eigenvalue and we know the eigenvalues of the Hermitian part of the matrix, then we can determine the size of the subspace to use and the number of Ritz values to keep,

to guarantee that the shifts will be well separated from the wanted eigenvalue. It is unclear if this approach is an artifact of the analysis, or would be genuinely better in some situations than keeping all $k-1$ shifts.

The analysis above assumed there was one rightmost normal eigenvalue that we wished to compute. Suppose instead that there are m rightmost simple real normal eigenvalues, $\lambda_1, \ldots, \lambda_m$ that we wish to compute, and that their field of values is disjoint from the field of values of D. Our matrix would then be of the form

$$A = \begin{pmatrix} \Lambda & 0 \\ 0 & D \end{pmatrix}, \tag{4.16}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$. The analysis would not change significantly. The quantity η, instead of bounding the distance of the nearest Ritz value to just one of the eigenvalues, could be defined to bound the distance of m Ritz values to all the eigenvalues of Λ. The angle between the starting vector and the desired subspace would have to be sufficiently small so as to guarantee that each desired eigenvalue has its own Ritz value near it. The bound for the magnitude of the Ritz values could still use the weak majorization of the magnitudes of the Ritz values by the singular values. A bound for $\|\psi(\Lambda)\|$ could be stated in terms of the distance of the shifts to λ_m, the smallest eigenvalue of Λ. Thus, similar to subspace iteration, the convergence to the desired eigenspace can be understood in terms of the eigenvalue that is closest to the unwanted eigenvalues.

For a different bound for the Ritz values, recall from Section 3.1 that the boundary of the field of values satisfies certain interlacing inequalities. Let $T = V^*AV$ be the restriction of A onto a subspace. The bounds for the real part of $W(T)$ are determined by the eigenvalues of the Hermitian part of T: $(T+T^*)/2 = V^*(A+A^*)V/2$. The Hermitian part of T is simply a restriction of the Hermitian part of A. Hence the

eigenvalues of the Hermitian part of T must interlace the eigenvalues of the Hermitian part of A. Suppose that we are restricting onto a 2-dimensional subspace. Then in the notation used for Theorem 4.1, the two largest eigenvalues of the Hermitian part of A are $\vartheta_1 = \lambda$ and ϑ_2. Denote with $\hat{\vartheta}_1$ and $\hat{\vartheta}_2$ ($\hat{\vartheta}_1 \geq \hat{\vartheta}_2$) the eigenvalues of the Hermitian part of T. From interlacing, we must have that

$$\hat{\vartheta}_2 \leq \vartheta_2, \tag{4.17}$$

thus the field of values of T must always contain points with real part less than or equal to ϑ_2. This limits how nonnormal T can be; see Figure 4.1. In the figure, the problem of determining how far the eigenvalues z_1 and z_2 of H_2 can be to the right of ϑ_2 is equivalent to determining the largest circle contained in $W(A)$ that is tangent and to the right of the vertical line $\operatorname{Re} z = \vartheta_2$ and contained in $W(A)$. This would correspond to T being similar to a Jordan block. Let h denote a bound for the size of the imaginary part of $W(A)$ along $\operatorname{Re} z = \vartheta_2$. As in the figure, θ denotes half the angle made by $W(D)$ at λ. (Note that the angle θ is associated with the phases of the eigenvalues of the unitary factor of $D - \lambda$; see Section 3.2.2.) Let r denote the radius of the circle mentioned above. Then from the geometry, we have

$$r = (\lambda - \vartheta_2 - r)\sin\theta.$$

We also have that $\tan\theta = h/(\lambda - \vartheta_2)$, thus

$$r = \left(\frac{h}{\tan\theta} - r\right)\sin\theta = \frac{h\cos\theta}{1+\sin\theta}.$$

Thus we can never have two Ritz values to the right of $\vartheta_2 + h\cos\theta/(1+\sin\theta)$. To determine when this bound is sharper than the weak majorization-based bound of

$(\lambda + \vartheta_2)/2$, consider

$$\vartheta_2 + \frac{h\cos\theta}{1+\sin\theta} \;<\; \frac{\lambda+\vartheta_2}{2},$$

so,

$$\frac{2h}{\sec\theta+\tan\theta} \;<\; \lambda - \vartheta_2,$$

and hence

$$2h - \tan\theta(\lambda-\vartheta_2) \;<\; (\lambda-\vartheta_2)\sec\theta.$$

Eliminating θ gives

$$2h - \frac{h}{\lambda-\vartheta_2}(\lambda-\vartheta_2) \;<\; (\lambda-\vartheta_2)\sqrt{1+\left(\frac{h}{\lambda-\vartheta_2}\right)^2},$$

so,

$$h \;<\; \sqrt{(\lambda-\vartheta_2)^2+h^2}.$$

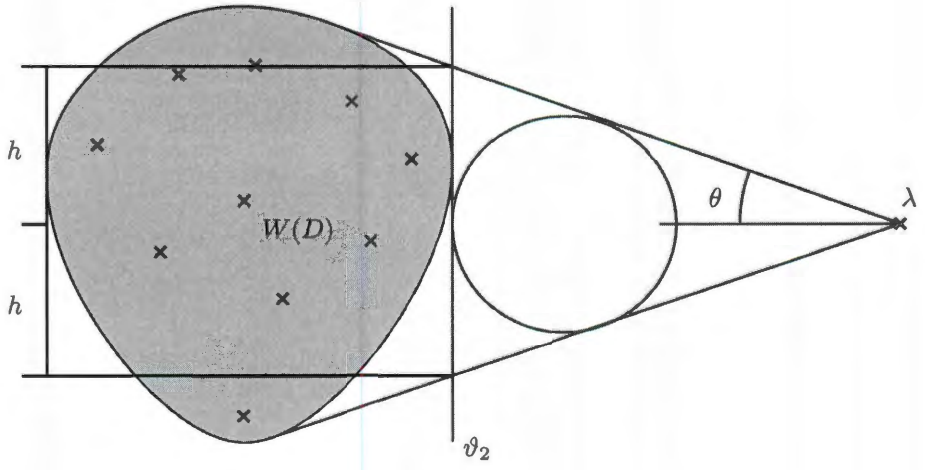


Figure 4.1 : Illustration of the bound based on interlacing of the eigenvalues of the Hermitian part of V^*AV and A.

Since this inequality always holds, the field of value interlacing bound for the shifts is always sharper than the majorization-based bound. This field of value interlacing based bound, unlike the quantity $f(j,\eta)$, does not incorporate the richness of the starting vector in the desired subspace. This offers an instance in which the weak majorization-based bounds are not sharp. To further illustrate, if we were to move λ in Figure 4.1 to the right, keeping D, fixed, then the weak majorization-based bound $(\vartheta_2+\lambda)/2$ would increase at the rate of one half, regardless of the size of λ. However, the field of values interlacing based bound can be no greater than $\vartheta_2+\mu(D)$.

4.2.2 Some Nonnormality

Up to now we have only been considering a normal eigenvalue. In this section, we relax this requirement and show that the results above for convergence need only be modified slightly. Let $E = e_1[0\ g^*]$. Then for A in (4.1), we have

$$A+E = \begin{pmatrix} \lambda & g^* \\ 0 & D \end{pmatrix}. \tag{4.18}$$

Bellalij, Saad and Sadok [8] analyzed the same type of matrix for the unrestarted Arnoldi method. We will draw upon several observations they collected: a form for polynomials of $A+E$ and an expression for the left eigenvector of λ. The matrix $\psi(A+E)$ has the form

$$\psi(A+E) = \begin{pmatrix} \psi(\lambda) & g^*(\lambda I - D)^{-1}(\psi(\lambda)I - \psi(D)) \\ 0 & \psi(D) \end{pmatrix}.$$

This is a result that Parlett noted in [43]. The left eigenvector of $A+E$ associated with λ is

$$t = \begin{pmatrix} 1 \\ (\lambda I - D)^{-*}g \end{pmatrix}. \tag{4.19}$$

Now if we apply $\psi(A+E)$ to our starting vector v,

$$\psi(A+E)v = \begin{pmatrix} \psi(\lambda)c + g^*(\lambda I - D)^{-1}(\psi(\lambda)I - \psi(D))r \\ \psi(D)r \end{pmatrix} \tag{4.20}$$

$$= \begin{pmatrix} \psi(\lambda)t^*v - g^*(\lambda I - D)^{-1}\psi(D)r \\ \psi(D)r \end{pmatrix}. \tag{4.21}$$

As before, in order for the restarted Arnoldi method to make progress, we must have

$$\frac{\|\psi(D)r\|}{|\psi(\lambda)t^*v - g^*(\lambda I - D)^{-1}\psi(D)r|} < \frac{\|r\|}{|c|}\gamma, \tag{4.22}$$

where $\gamma \in (0,1)$. Rearranging the expression,

$$\frac{\|\psi(D)r\|}{\|r\|} < \frac{|\psi(\lambda)t^*v - g^*(\lambda I - D)^{-1}\psi(D)r|}{|c|}\gamma. \tag{4.23}$$

The strategy from Theorem 4.1 involving the magnitudes and real parts of the shifts applied here amounts to bounding the expression on the left from above and the expression on the right from below,

$$\frac{\|\psi(D)r\|}{\|r\|} \leq \|\psi(D)\|$$

and

$$\frac{|\psi(\lambda)t^*v| - \|g^*(\lambda I - D)^{-1}\|\|\psi(D)\|\|r\|}{|c|} \leq \frac{|\psi(\lambda)t^*v - g^*(\lambda I - D)^{-1}\psi(D)r|}{|c|}.$$

Combining these two bounds gives a stronger criteria for convergence that implies 4.23:

$$\|\psi(D)\| < \gamma\frac{|\psi(\lambda)t^*v| - \|g^*(\lambda I - D)^{-1}\|\|\psi(D)\|\|r\|}{|c|}.$$

Isolating terms involving the starting vector and the left eigenvector t gives

$$h(v,\gamma) := \gamma\frac{|t^*v|}{|c| + \gamma\|g^*(\lambda I - D)^{-1}\|\|r\|} > \frac{\|\psi(D)\|}{|\psi(\lambda)|}. \tag{4.24}$$

Note as $v \to e_1$, $h(v) \to \gamma$. Also, if $g = 0$, then $h(v) = \gamma$. To bound the expression on the right, we will use the argument in Theorem 4.1. However, we must adjust our bounds for the shifts to take into account the effect of E. The field of values, and hence the Ritz values, of A can be perturbed by E by no more than $\|E\|$. Hence the bounds in Theorem 4.1 for the magnitude and real part of the shifts can change by no more than $\|E\|$. Thus a yet more stringent requirement for convergence is

$$\frac{\|D\| + \hat{\rho} + \|E\|}{\lambda - \hat{z} - \|E\|} = \frac{\|D\| + \eta + \|D\| + \|E\|}{\lambda - \eta - \vartheta_2 - \|E\|} < h(v,\gamma)^{\frac{1}{p}}, \tag{4.25}$$

where $\hat{\rho}$ and $\hat{\theta}$ are as in Theorem 4.1. This gives a criterion for η:

$$\eta < \frac{(h(v,\gamma))^{\frac{1}{p}}(\lambda - \vartheta_2) - 2\|D\|}{h(v,\gamma)^{\frac{1}{p}} + 1} - \|E\|. \tag{4.26}$$

Thus, if $\lambda > 2\|D\|h(v,\gamma)^{-1/p} + \vartheta_2 + \|E\|((h(v,\gamma))^{-1/p} + 1)$, then for v sufficiently rich in the desired eigenvector, the restarted Arnoldi method for $A+E$ will converge. This criterion reduces to that of Theorem 4.1 in the case that $\|E\| \to 0$, as would be expected.

It is likely that a similar result could be derived for computing multiple eigenvalues or keeping several shifts, by noting that for polynomials of block triangular matrices,

$$\psi\left(\begin{pmatrix} R_1 & G \\ 0 & R_2 \end{pmatrix}\right) = \begin{pmatrix} \psi(R_1) & \psi(R_1)T - T\psi(R_2) \\ 0 & \psi(R_2) \end{pmatrix}, \tag{4.27}$$

where T is such that $G = R_1T - TR_2$ and the rows of $[I\ T]$ form a basis for the left eigenspace corresponding to the eigenvalues of R_1. For T to be well defined, the spectra of R_1 and R_2 must be disjoint, i.e., R_1 and R_2 do not share any eigenvalues. An analysis of the restarted Arnoldi method for matrices of the form (4.27) would have to account for both conditioning of the wanted eigenvalues, as well as the coupling between the eigenvalues of R_1 and R_2 due to G, and hence could provide much insight into the restarted Arnoldi method for non-Hermitian matrices.

Having derived some convergence criteria for the restarted Arnoldi method, in the next section I analyze convergence for a normal matrix.

4.3 Normal Example

Here I demonstrate some of the notions developed in the previous section for a small normal matrix for which Arnoldi convergence can be completely analyzed. Though simple, this example illustrates in a clean manner a variety of convergence behavior.

Given a matrix with $D = -D^*$ of the form (4.1),

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & D \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \alpha i & 0 \\ 0 & 0 & -\alpha i \end{pmatrix},$$

with $\lambda, \alpha > 0$, I will show the restarted Arnoldi method can be understood completely in terms of the Ritz values. We are interested in computing the rightmost eigenvalue, λ, using restarted Arnoldi with exact shifts drawn from two dimensional subspaces. As the matrix is complex, I will be dealing with complex shifts. We will have two Ritz values, z_1 and z_2, and we shall assume $\operatorname{Re} z_2 \leq \operatorname{Re} z_1$; hence, z_2 will be the shift. I will address the following questions concerning restarted Arnoldi with Ritz values of 2×2 restrictions of A.

- What are the possible Ritz values, z_1 and z_2?
- How do Ritz values from one cycle relate to Ritz values from the next cycle?
- Does restarted Arnoldi always converge?
- At what rate does the method converge/diverge?

- Are there any restrictions that must be placed on the starting vector to ensure that z_1 converges to λ?

For restarted Arnoldi, starting with some $V \in \mathbb{C}^{3\times 2}$ having orthonormal columns and spanning a Krylov subspace, compute the Ritz values z_1 and z_2 from V^*AV. Use the leftmost Ritz value z_2 to compute $\widehat{V}$:

$$\text{Range}(\widehat{V}) = \text{Range}\left((A - z_2 I)V\right).$$

Update $V \longrightarrow \widehat{V}$ and repeat this process until (hopefully) z_1 converges to the eigenvalue λ. We wish to understand the convergence of this process.

The field of values of $W(A)$ is a triangle, and from Section 3.4, we know that if we specify one of the Ritz values, then the other is determined uniquely. Also, the Ritz values from V uniquely determine the magnitudes of the entries of a unit vector v orthogonal to the range of V, see equation (3.32):

$$|v|^2 = \begin{bmatrix} \frac{(\lambda - z_1)(\lambda - z_2)}{(\lambda - \alpha i)(\lambda + \alpha i)} \\ \frac{(\alpha i - z_1)(\alpha i - z_2)}{(\alpha i - \lambda)(\alpha i + \alpha i)} \\ \frac{(-\alpha i - z_1)(-\alpha i - z_2)}{(-\alpha i - \lambda)(-\alpha i - \alpha i)} \end{bmatrix}, \tag{4.28}$$

where $|v|^2$ denotes the vector consisting of the magnitudes entries of v squared. For z_1 and z_2 to be Ritz values of A, each entry in the vector in (4.28) must be nonzero. This constraint yields an equation relating z_1 to z_2:

$$z_1 = \lambda + \frac{(\alpha^2 + \lambda^2)\text{Re}\, z_2}{\lambda(\frac{1}{4}(\lambda + \frac{\alpha^2}{\lambda})^2 - |z_2 - \frac{1}{2}(\lambda - \frac{\alpha^2}{\lambda})|^2)}(\bar{z}_2 - \lambda). \tag{4.29}$$

In the language of Section 3.4, this expression shows that z_1 lies on the isogonal of the Cevian from λ to z_2, i.e., the phase of $\bar{z}_2 - \lambda$ gives the angle of isogonal Cevian to the Cevian from λ to z_2. A similar expression could be derived emphasizing the Cevians through the eigenvalues $i\alpha$ and $-i\alpha$. The denominator for the second term in

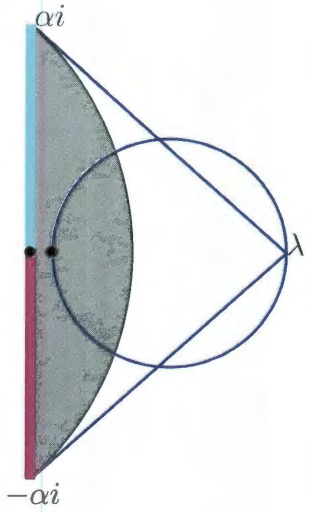


Figure 4.2 : Shifts, fixed points, limit cycles and progress for $\lambda = 1.1\alpha$. The gray region indicates where all the shifts, the z_2, can lie. If a shift falls in the circle then the previous iteration of Arnoldi did not improve the eigenvector estimate. The black dots indicate the fixed points, with the one on the left being attractive and the one on the right being repulsive. The line segments on the left indicate the period two limit cycles to which the shift converges.

(4.29) is the equation for the circle passing through the eigenvalues. This expression breaks down for $z_2 \in \sigma(A)$: there no longer is a one to one correspondence between z_2 and z_1. If $\operatorname{Re} z_2 \leq \operatorname{Re} z_1$, then taking the real part of (4.29) and substituting the right hand side into $\operatorname{Re} z_2 \leq \operatorname{Re} z_1$, one can see that z_2 must lie inside the circle of radius $\alpha\sqrt{1+\alpha^2/\lambda^2}$ centered at $-\alpha^2/\lambda$. To understand restarted Arnoldi behavior, we must know how z_2 moves about this circle. For z_1 to converge to λ, $\operatorname{Re} z_2$ must go to zero. For illustrations of these circles containing the shifts for different λ and α see Figures 4.2, 4.3, and 4.4.

Equations (4.28) and (4.29) allow us to represent the unit vector v orthogonal to the range of V solely in terms of z_2. The vector orthogonal to $\widehat{V}$, denoted $\hat{v}$, is

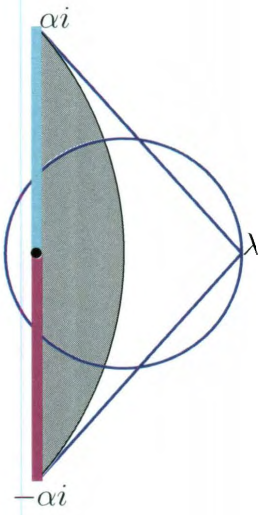


Figure 4.3 : Shifts, fixed points, limit cycles and progress for $\lambda = 7\alpha/8$. The gray region indicates where all the shifts, the z_2, can lie. If a shift falls in the circle then the previous iteration of Arnoldi did not improve the eigenvector estimate. There is only one fixed point, shown as a black dot at the origin; it is repulsive The line segments on the left indicate the period two limit cycles to which the shift converges. Note that the portion of the segments contained in the circle correspond to repulsive limit cycles.

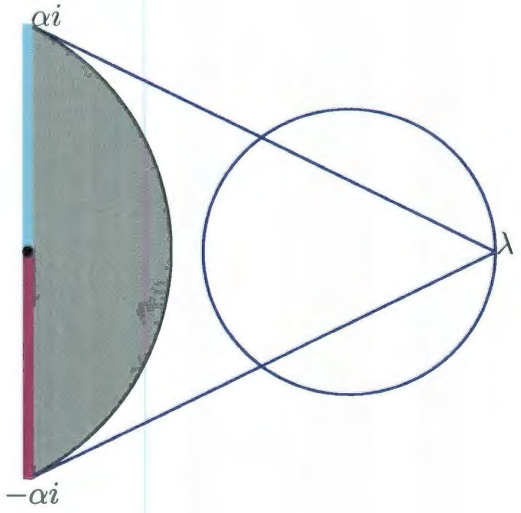


Figure 4.4 : Shifts, fixed points, limit cycles and progress for $\lambda = 2\alpha$. The gray region indicates where all the shifts, the z_2, can lie. If a shift falls in the circle then the previous iteration of Arnoldi did not improve the eigenvector estimate. There is only one fixed point, shown as a black dot at the origin; it is attractive. The line segments on the left indicate the period two limit cycles to which the shift converges. In this case, restarted Arnoldi with exact shifts converges monotonically.

proportional to $(A - z_2 I)^{-*} v$, hence we may also represent $\hat{v}$ in terms of z_2. Thus, for this normal matrix, we may study restarted Arnoldi solely through the magnitudes of the entries of the vectors v and $\hat{v}$. Note that the only vectors v such that the range of V is not a Krylov subspace correspond to V containing only one eigenvector of A. Expressing v and $\hat{v}$ in terms of z_1,

$$|v_1|^2 = \frac{|\hat{v}_1|^2 |z_2 - \lambda|^2}{\frac{1}{4}(\lambda + \frac{\alpha^2}{\lambda})^2 - |z_2 - \frac{1}{2}(\lambda - \frac{\alpha^2}{\lambda})|^2}, \quad |\hat{v}_1|^2 = \frac{\operatorname{Re} z_2}{\lambda}, \tag{4.30}$$

$$|v_2|^2 = \frac{|\hat{v}_2|^2 |z_2 - \alpha i|^2}{\frac{1}{4}(\lambda + \frac{\alpha^2}{\lambda})^2 - |z_2 - \frac{1}{2}(\lambda - \frac{\alpha^2}{\lambda})|^2}, \quad |\hat{v}_2|^2 = \frac{\operatorname{Im}((\lambda - \alpha i)(z_2 + \alpha i))}{2\lambda\alpha}, \tag{4.31}$$

where we have omitted the expressions for v_3 and $\hat{v}_3$, as both v and $\hat{v}$ are unit vectors. Using these same equations for $\hat{v}$ and $\hat{z}_2$, where $\hat{z}_2$ is the leftmost eigenvalue of $\widehat{V}^* A \widehat{V}$, by eliminating the $|v_i|^2$ and the $|\hat{v}_i|^2$, we can relate z_2 and $\hat{z}_2$:

$$\operatorname{Re} z_2 = \frac{\operatorname{Re} \hat{z}_2 |\hat{z}_2 - \lambda|^2}{\frac{1}{4}(\lambda + \frac{\alpha^2}{\lambda})^2 - |\hat{z}_2 - \frac{1}{2}(\lambda - \frac{\alpha^2}{\lambda})|^2} \tag{4.32}$$

$$\operatorname{Im} z_2 = -\frac{\operatorname{Im} \hat{z}_2 (\alpha^2 (1 + \frac{\alpha^2}{\lambda^2}) - |\hat{z}_2 + \frac{\alpha^2}{\lambda}|^2)}{\frac{1}{4}(\lambda + \frac{\alpha^2}{\lambda})^2 - |\hat{z}_2 - \frac{1}{2}(\lambda - \frac{\alpha^2}{\lambda})|^2}. \tag{4.33}$$

These expressions are interesting, in that they describe a discrete dynamical system in z_2 that corresponds to running restarted Arnoldi in reverse. Though not immediately evident from these expressions because equations (4.32) and (4.33) allow for a subspace to specified by either of its two Ritz values, the pairing of the subspaces from running Arnoldi forward or backward is unique. Equation (4.33) for $\operatorname{Im} z_2$ involves the ratio of the distance of $\hat{z}_2$ to two circles; in the denominator, the circle passing through the eigenvalues, and in the numerator, the circle containing all possible $\hat{z}_2$. As the shift 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 z_2, if nonzero, alternates in sign, and its magnitude is increasing with each iteration of

restarted Arnoldi. An example of this alternating of the sign of the shifts can be seen in Figure 4.5.

In order for the restarted Arnoldi method to converge, the range of V must, in the limit, contain $[1,0,0]^T$, the eigenvector associated with λ. This corresponds to $|v_1|$ going to zero. From (4.30) expressions above, this is equivalent to having $\mathrm{Re}\, z_2$ go to zero. For progress to be made during one iteration, $\mathrm{Re}\, \hat{z}_2/\mathrm{Re}\, z_2 < 1$. From (4.32), for $\mathrm{Re}\, \hat{z}_2$ to be less than $\mathrm{Re}\, z_2$, $\hat{z}_2$ must lie inside the circle of radius $(\lambda^2+\alpha^2)/4\lambda$ centered at $(3\lambda - \alpha^2/\lambda)/4$. Note that this circle indicates when progress has been made at the last iteration, rather then when progress will be made at the next iteration. From this circle and the circle containing z_2, for our matrix we can completely understand the behavior of the restarted Arnoldi method. This is illustrated in Figure 4.5.

There are several special regions to note. For $\alpha < \lambda \leq \sqrt{3}\alpha$, there is a fixed point for z_2 located at $(\lambda^2 - \alpha^2)/2\lambda$. The latter corresponds to stagnation, in which case z_1 never converges to the eigenvalue at λ. This is a generalization of the stagnation observed for the circulant shift in Section 4.1.1. This fixed point is repulsive, hence for nearly all starting subspaces, the restarted Arnoldi method does not stagnate. For $\lambda > \alpha$, 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 part of z_2 alternates in sign, while z_1 remains fixed at λ. These are the limit cycles for restarted Arnoldi. There are also sources, corresponding to subspaces that could only be generated by having used z_1, the rightmost Ritz value, to generate $\widehat{V}$. In terms of z_2, the sources lie near the top and bottom portions of the boundary of the circle containing z_2. Illustrations of these different scenarios for α and λ can be seen in Figures 4.2, 4.3, and 4.4.

From above, we know that for nearly all starting V, the restarted Arnoldi method

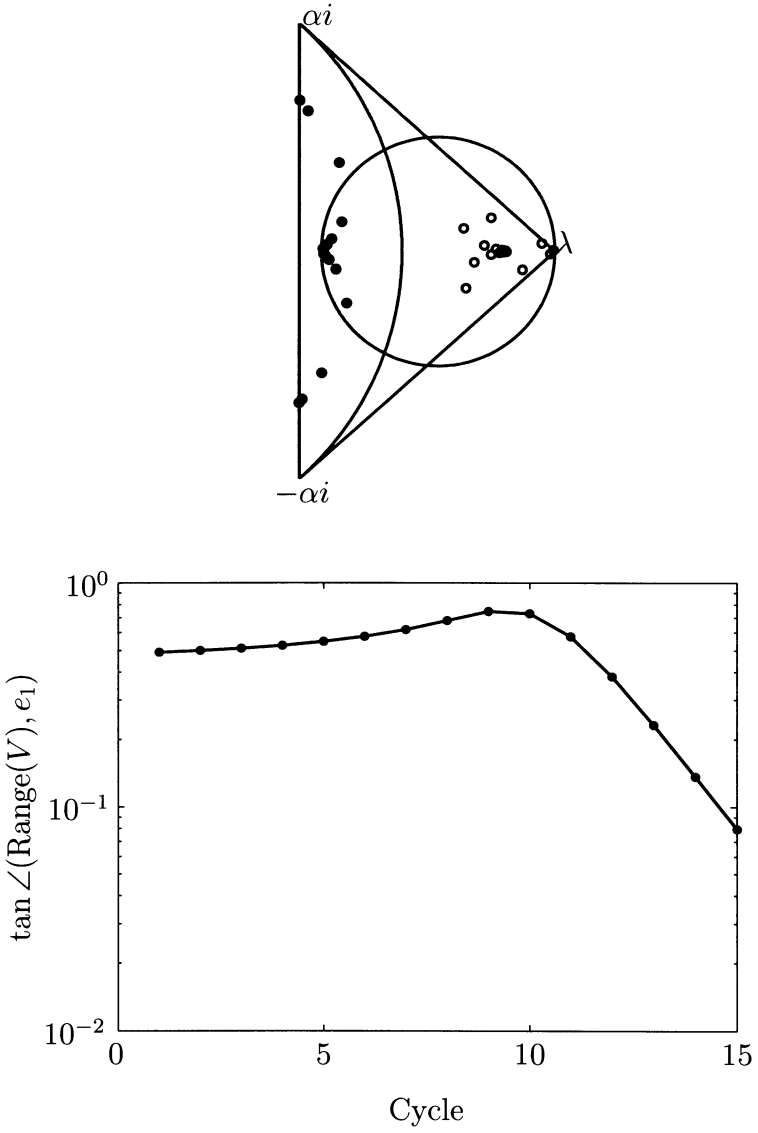


Figure 4.5 : Convergence example for $\lambda = 1.1\alpha$. The top plot shows the shift z_2, the black dots, from several cycles of restarted Arnoldi with exact shifts for an initial shift lying in the circle corresponding to progress not having been made in the previous iteration. The circles correspond to z_1, the desired Ritz value at each cycle. The curves plotted are as in Figure 4.2. The bottom plot shows the tangent of the angle between the Krylov subspace and the desired eigenvector at each cycle of restarted Arnoldi with exact shifts. After a period of divergence, the Krylov subspaces starts converges to a subspace that contains the wanted eigenvector.

will converge, though it may initially diverge. The rate at which it converges or diverges can be determined from equation (4.32). Maximizing the ratio of $\mathrm{Re}\,\hat{z}_2$ to $\mathrm{Re}\,z_2$ gives a rate of $\alpha^2/\lambda^2 + \alpha\sqrt{\alpha^2+\lambda^2}/\lambda^2$. 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 convergence in the sense that if we initially have a good approximation to the desired eigenvector for λ; then the approximations can only improve. For $\lambda < \alpha$, z_1 can be arbitrarily close to λ, and z_1 can initially diverge from λ. With the exception of the stagnation point, all z_2 must eventually converge to a period two cycle along the imaginary axis. The asymptotic rate at which z_2 converges to such a cycle is $(\alpha^2 - y^2)/(\lambda^2 + y^2)$, where y is the limit of the imaginary part of z_2 as restarted Arnoldi converges. This rate is bounded above by α^2/λ^2, which is the rate you would expect from the power method provided, $\alpha < \lambda$.

We now relate the results of this section to those of the previous. Here, in order for the method to converge for a starting vector that is sufficiently rich in the desired eigenvector e_1, we require $\lambda > \alpha$. Theorem 4.1 has that for convergence at rate γ, $\lambda > 2\|D\|/\gamma + \vartheta_1$. For this normal matrix, $\|D\| = \alpha$ and $\vartheta_1 = 0$, hence the theorem requires $\lambda > 2\alpha/\gamma$. Thus, for this normal matrix, the theorem is not sharp. This is due to our independently bounding the magnitude of the shifts and the proximity of the shifts to λ. Theorem 4.1 only gives a criterion for monotone convergence. A theorem that accounts for non-monotone convergence would be much more complicated. It would be interesting to know if there are any D for which Theorem 4.1 is sharp.

4.4 Discussion

In this chapter I developed sufficient conditions for the convergence of the restarted Arnoldi algorithm for a matrix with one simple normal wanted eigenvalue, then gen-

eralized these conditions to allow for nonnormality. The conditions are stronger than having the wanted eigenvalue lie outside of the the field of values associated with the unwanted eigenvalues. If one is interested solely in preventing extreme breakdown, the requirement on the field of values need not be as stringent, and with knowledge of the eigenvalues of the Hermitian part of the matrix, one can determine the size of the subspace necessary to keep the shifts away from the wanted eigenvalue. The criteria are rather weak, in that they ask that the wanted eigenvalue be well separated from the unwanted eigenvalues. However, these criteria may be more realistic for shift-invert Arnoldi. In this setting the standard Arnoldi algorithm is applied to $(A - \mu I)^{-1}$. If μ is close to the desired eigenvalue, then the wanted desired eigenvalue of the shifted and inverted matrix can be well separated from the rest of the spectrum. Developing less stringent criteria will require accounting not just for how the Ritz values may cluster about the wanted eigenvalue, but also for how the Ritz values must distribute themselves throughout the rest of the field of values.

For a 3×3 normal matrix with a skew symmetric block and a real wanted eigenvalue, I showed that the restarted Arnoldi method with an exact complex shift nearly always converges. Depending on the size of the wanted real eigenvalue, there may exist starting vectors for which the method stagnates. The approach for this normal matrix took advantage of the properties of $n-1$ Ritz values for an n-dimensional normal matrix discussed in Section 3.4. Building on this approach for normal matrices, it may be possible to nicely characterize restarted Arnoldi behavior for normal matrices for $n-1$ dimensional subspaces with $(n-1)/2$ exact shifts. Such a result could provide a great deal of insight on how convergence depends on subspace dimension.

Chapter 5

Conclusion

This thesis has shown that the Ritz values of nonsymmetric matrices can be localized and that the localization of the Ritz values can be used to determine sufficient conditions for convergence of the restarted Arnoldi method with exact shifts.

The analysis of Ritz values in this thesis has been posed in terms of the inverse field of values problem with k Ritz values (iFOV-k). For just one Ritz value, iFOV-1, the problem can be solved easily and exactly due to the convexity of the field of values, as shown in Chapter 2. For more than one Ritz value, iFOV-k with $k > 1$, determining if the problem is solvable is nontrivial. In Chapter 3, I showed majorization results for the Ritz values ordered by their real part, magnitude and phase by the eigenvalues of different matrices: the Hermitian part, the polar part and the unitary factor. These same results are also useful for localizing the eigenvalues of the matrix. Hence, these results suggest that for nonsymmetric matrices, these bounds localize the Ritz values no better than they can localize the eigenvalues.

It remains to be discovered if there is anything sharper than majorization for localizing Ritz values for general. I showed one way of improving upon these bounds using the projection property of the field of values. By limiting the size of the field of values of a restriction, the projection property limits how many Ritz values can cluster near the boundary of the field of values. There are generalizations of the field of values that take into account multiple Ritz values; however, they lack convexity [36]. An analysis of iFOV-2 for a Jordan block indicates the difficulty of precisely

understanding Ritz values in general. Understanding iFOV-$(n-1)$ for normal matrices is something that may be tractable, as our analysis for $n = 3$ indicates.

With bounds for the Ritz values, I determined sufficient conditions for the convergence of the restarted Arnoldi method with exact shifts. In Chapter 4, for matrices with one simple normal eigenvalue, the conditions determine when convergence is possible, i.e., if the starting vector well approximates the desired eigenvector and the desired eigenvalue is well-separated from the rest of the spectrum, then the restarted Arnoldi method will converge. I provided a criterion that guarantees that exact shifts do not fall near wanted eigenvalues. This criterion differs from requiring that wanted eigenvalues not lie in the field of values associated with the unwanted eigenvalues. With knowledge of the eigenvalues of the Hermitian part of the matrix, one can get a sense of how large a subspace one should use and how many Ritz values to keep to prevent shifts from falling near wanted eigenvalues. Majorization limits the regions where large numbers of Ritz values can cluster. The density of the Ritz values can be greatest around $\mathrm{tr}(A)/n$, and smallest near the boundary of the field of values. This provides some insight into why the restarted Arnoldi method works so well, but does not address how exact shifts can, in the limit, behave optimally.

Future work could include determining the properties of Ritz values that give rise to convergence of the restarted Arnoldi method with exact shifts. Just as eigenvalues for non-Hermitian matrices do not completely explain convergence of restarted Arnoldi, to what extent do the characteristics of Ritz values from k-dimensional subspaces determine convergence of the restarted Arnoldi method? The normal 3×3 example in Chapter 4 included examples of stagnation and non-monotonic convergence, as well as limit cycles of the restarted Arnoldi method. Working out a similar example for a nonnormal matrix could provide much insight. In addition to the pos-

sibility of chaotic behavior, the 3×3 Jordan block showed that the Ritz values do not uniquely determine a subspace. If the subspace determined by Ritz values is not unique, can we understand Arnoldi behavior solely in terms of Ritz values? The answers to these questions would improve our understanding of other iterative methods that depend on Ritz values.

Bibliography

[1] A. Aitken, *Determinants and Matrices*, Oliver and Boyd, Edinburgh, 1939.

[2] A. R. Amir-Moez and A. Horn, *Singular values of a matrix*, Amer. Math. Monthly, 65 (1958), pp. 742–748.

[3] W. E. Arnoldi, *The principle of minimized iteration in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 17–29.

[4] C. Beattie, *Harmonic Ritz and Lehmann bounds*, Electron. Trans. Numer. Anal., 7 (1998), pp. 18–39.

[5] C. Beattie and I. C. F. Ipsen, *Inclusion regions for matrix eigenvalues*, Linear Algebra Appl., 358 (2003), pp. 281–291.

[6] C. A. Beattie, M. Embree, and J. Rossi, *Convergence of restarted Krylov subspaces to invariant subspaces*, SIAM J. Matrix Anal. Appl., 25 (2004), pp. 1074–1109.

[7] C. A. Beattie, M. Embree, and D. C. Sorensen, *Convergence of polynomial restart Krylov methods for eigenvalue computations*, SIAM Rev., 47 (2005), pp. 492–515.

[8] M. Bellalij, Y. Saad, and H. Sadok, *Further analysis of the Arnoldi process for eigenvalue problems*, SIAM J. Numer. Anal., 48 (2010), pp. 393–407.

[9] P. N. BROWN, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 58–78.

[10] R. CARDEN, MATLAB *m-file* `inversefov.m`. `http://www.caam.rice.edu/tech_reports/2009/inversefov.m`, 2009.

[11] R. CARDEN, *Ritz values and Arnoldi convergence for nonsymmetric matrices*, Master's thesis, Rice University, 2009.

[12] R. CARDEN, *A simple algorithm for the inverse field of values problem*, Inverse Problems, 25 (2009), p. 115019.

[13] R. CARDEN AND D. HANSEN, *Ritz values of normal matrices and Ceva's theorem*. In preparation.

[14] A. CHAPMAN AND Y. SAAD, *Deflated and augmented Krylov subspace techniques*, Numer. Linear Algebra Appl., 4 (1997), pp. 43–66.

[15] C.-M. CHENG, R. A. HORN, AND C.-K. LI, *Inequalities and equalities for the Cartesian decomposition of complex matrices*, Linear Algebra Appl., 341 (2002), pp. 219 – 237.

[16] C. CHORIANOPOULOS, P. PSARRAKOS, AND F. UHLIG, *A method for the inverse numerical range problem*, Electron. J. Linear Algebra, 20 (2010), pp. 198–206.

[17] T. H. CORMEN, C. STEIN, R. L. RIVEST, AND C. E. LEISERSON, *Introduction to Algorithms*, MIT Press, Cambridge, 1990.

[18] D. COX, J. LITTLE, AND D. O'SHEA, *Ideals, Varieties, and Algorithms*, Springer, New York, third ed., 2007.

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

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

[21] S. Furtado and C. R. Johnson, *Unitary similarity classes within the cospectral-congruence class of a matrix*, Linear Algebra Appl., 394 (2005), pp. 291 – 307.

[22] A. Greenbaum, V. Pták, and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469.

[23] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.

[24] K. E. Gustafson and D. K. M. Rao, *Numerical range*, Springer-Verlag, New York, 1997.

[25] F. Hartmann and R. Jantzen, *Apollonius's ellipse and evolute revisited.* `http://mathdl.maa.org/mathDL/46/?pa=content&sa=viewDocument&nodeId=2554`, June 2009.

[26] P. Henrici, *Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices*, Numer. Math., 4 (1962), pp. 24–40. 10.1007/BF01386294.

[27] V. Heuveline and M. Sadkane, *Arnoldi–Faber method for large non-Hermitian eigenvalue problems*, Electron. Trans. Numer. Anal., 5 (1997), pp. 62–76.

[28] A. Horn, *On the eigenvalues of a matrix with prescribed singular values*, Proc. Amer. Math. Soc., 5 (1954), pp. 4–7.

[29] A. Horn and R. Steinberg, *Eigenvalues of the unitary part of a matrix*, Pacific J. Math., 9 (1959), pp. 541–550.

[30] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985. Corrected reprint of the 1985 original.

[31] ——, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1994.

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

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

[34] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Research Nat. Bur. Standards, 45 (1950), pp. 255–282.

[35] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK Users' Guide*, SIAM, Philadelphia, 1998.

[36] C.-K. Li, *Polynomials and numerical ranges*, Proc. Amer. Math. Soc., 104 (1988), pp. 369–373.

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

[38] A. W. MARSHALL AND I. OLKIN, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.

[39] G. MEURANT, *The Lanczos and Conjugate Gradient Algorithms: From Theory to Finite Precision Computations*, SIAM, Philadelphia, 2006.

[40] ——, *The computation of isotropic vectors*, 2011. Manuscript.

[41] R. B. MORGAN, *GMRES with deflated restarting*, SIAM J. Sci. Comput., 24 (2002), pp. 20–37.

[42] B. PARLETT AND G. STRANG, *Matrices with prescribed Ritz values*, Linear Algebra Appl., 428 (2008), pp. 1725 – 1739.

[43] B. N. PARLETT, *A recurrence among the elements of functions of triangular matrices*, Linear Algebra Appl., 14 (1976), pp. 117 – 121.

[44] B. N. PARLETT, *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, 1998.

[45] J. F. QUEIRÓ AND A. L. DUARTE, *On the Cartesian decomposition of a matrix*, Linear Multilinear Algebra, 18 (1985), pp. 77–85.

[46] Y. SAAD, *Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices*, Linear Algebra Appl., 34 (1980), pp. 269–295.

[47] Y. SAAD, *Projection methods for solving large sparse eigenvalue problems*, in Matrix Pencils, B. K. gström and A. Ruhe, eds., vol. 973 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1983, pp. 121–144.

[48] Y. SAAD, *Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems*, Math. Comp., 42 (1984), pp. 567–588.

[49] Y. SAAD AND M. H. SCHULTZ, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[50] N. SHOMRON AND B. N. PARLETT, *Linear algebra meets Lie algebra: The Kostant-Wallach theory*, Linear Algebra Appl., 431 (2009), pp. 1745 – 1767.

[51] P. SMIT, *Generating identical Ritz values*, Research Memorandum 696, Tilburg University, Faculty of Economics and Business Administration, May 1995.

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

[53] G. STARKE AND R. S. VARGA, *A hybrid Arnoldi–Faber iterative method for nonsymmetric systems of linear equations*, Numer. Math., 64 (1993), pp. 213–240.

[54] G. W. STEWART, *Matrix Algorithms. Vol. II: Eigensystems*, SIAM, Philadelphia, 2001.

[55] ——, *A Krylov–Schur algorithm for large eigenproblems*, SIAM J. Matrix Anal. Appl., 23 (2001/02), pp. 601–614.

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

[57] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton,

2005.

[58] F. UHLIG, MATLAB *m-file* `wberpoint.m`. `http://www.auburn.edu/~uhligfd/m_files/wberpoint.m`, 2008.

[59] F. UHLIG, *An inverse field of values problem*, Inverse Problems, 24 (2008), p. 055019 (19pp).

[60] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Oxford University Press, New York, 1963.