6-2013

On Almost Normal Matrices

Tyler J. Moran
College of William and Mary

On Almost Normal Matrices

A thesis submitted in partial fulllment of the requirement
for the degree of Bachelor of Science in Mathematics from
The College of William and Mary

by

Tyler J. Moran

Accepted for ______________________________

____
Ilya Spitkovsky, Director

____
Charles Johnson

____
Ryan Vinroot

____
Donald Campbell

Williamsburg, VA
April 3, 2013

Abstract

An n-by-n matrix A is called almost normal if the maximal cardinality of a set of orthogonal eigenvectors is at least $n-1$. We give several basic properties of almost normal matrices, in addition to studying their numerical ranges and Aluthge transforms. First, a criterion for these matrices to be unitarily irreducible is established, in addition to a criterion for A^* to be almost normal and a formula for the rank of the self commutator of A. We then show that unitarily irreducible almost normal matrices cannot have flat portions on the boundary of their numerical ranges and that the Aluthge transform of A is never normal when $n > 2$ and A is unitarily irreducible and invertible.

Contents

1. Introduction

First, let us fix the notation. Throughout this paper, $\mathbb{C}^n$ (resp., $\mathbb{C}^{n\times n}$) will represent the vector space of all n-vectors (resp., the algebra of all n-by-n matrices) with complex entries. We also let $e_1, \dots, e_n$ denote the standard basis of $\mathbb{C}^n$.

A matrix $A \in \mathbb{C}^{n\times n}$ is *normal* if it commutes with A^* (the conjugate transpose of A). For a non-normal A, several measures of non-normality can be used to characterize its deviation from being normal. They include the distance from A to the set of all normal matrices, various norms of the *self-commutator* $[A] := A^*A - AA^*$, etc. (see, e.g., [**?**]). More recently, the following notion was introduced in [**?**]: $A \in \mathbb{C}^{n\times n}$ is *almost normal* if it has (at least) $n-1$ pairwise orthogonal eigenvectors.

As with any matrix, an almost normal matrix will have Schur decomposition. The Schur decomposition of a matrix A will satisfy $A = QUQ^{-1}$ where Q is unitary (meaning that the conjugate transpose of Q is the inverse of Q and U is upper triangular. U is then called the "Schur form" of A. The Schur form of A will be unitarily similar to A (by definition) and therefore have the same multiset of eigenvalues. For an almost normal matrix, this form is

$$\begin{bmatrix} \lambda_1 & 0 & \dots & 0 & \beta_1 \\ 0 & \lambda_2 & \dots & 0 & \beta_2 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & \lambda_{n-1} & \beta_{n-1} \\ 0 & 0 & \dots & 0 & \mu \end{bmatrix}. \tag{1.1}$$

Using an additional diagonal unitary similarity, it is possible to arrange that $\beta_j \geq 0$, $j = 1, \dots, n-1$ because the β_i are defined by A only up to their modulus (where the modulus of a complex number $z = a + bi$ is defined as $|z| = \sqrt{a^2 + b^2}$. On the other hand, the eigenvalues $\lambda_1, \dots, \lambda_{n-1}$, (up to the permutation) and μ are defined by A uniquely.

In this paper, we study aspects of almost normality not covered by [6]. In Section 2, we look at several well known properties of normal matrices and explore the extent to which almost normal matrices differ in such regards. These properties include the criterion for an almost normal matrix to be unitarily irreducible (that is, not unitarily similar to a block diagonal matrix), the conditions under which A^* to be almost normal when A is, as well as the formula for the rank of the self commutator of A

Section 5 is devoted to the numerical range of an almost normal matrix. The numerical range of a matrix is a convex subset of the complex plane defined by $W(A) = \{\langle Ax, x\rangle : x \in \mathbb{C}^n,\ \|x\| = 1\}$.

Lastly, Section 6 looks at the relationship between the Aluthge transform and almost normal matrices (looking at both the Aluthge transform of an almost normal matrix as well as the possibility of an almost normal matrix lying in the range of the aluthge transform). The Aluthge transform of a matrix A is defined as $\Delta(A) = R^{1/2}UR^{1/2}$ where U and R arise from the polar decomposition of A. That is, $A = UR$ where U is unitary and R is a positive semidefinite Hermitian matrix.

A paper based on several of these results has been accepted for publication in Textos de Matemática for 2013.

2. Basic Properties

Under a unitary similarity transformation, it's possible that a matrix will become block diagonal- that is, equal to direct sum of 2 or more matrices. We call such matrices unitarily irreducible. For such matrices, it is also possible that they could be represented as the direct sum of varying numbers of blocks. For example, a 4-by-4 matrix could be unitarily similar to the direct sum of two 2-by-2 matrices, while also being unitarily similar to the direct sum of four singletons (that is, it would be diagonalizable). There will, of course, be a maximal amount of blocks that a matrix can be represented as a direct sum of.

Using a unitarily similarity transformation to represent A as a direct sum of the maximal possible number of blocks, we see that

$$A_n \oplus A_a, \tag{2.1}$$

where the block A_n is normal while A_a is almost normal and unitarily irreducible.The blocks A_n and A_a in (**??**) are each defined up to unitary similarity.

If A itself is normal (which is not excluded by the formal definition of almost normality), the block A_a disappears from (**??**). In the other extreme, that is, if $A = A_a$, we will say that A is *pure almost normal*. We will refer to the size of the block A_a in (**??**) as the *PAN-rank* of A. That is to say, the PAN rank of almost normal matrix A is defined as the rank of its unitarily irreducible component. Thus, an almost normal n-by-n matrix is normal (resp., pure almost normal) if its PAN-rank is 0 (resp, n). Note also that the PAN-rank cannot equal one, because a 1-by-1 block (a singleton) will necessarily commute with itself and therefore be normal. To see this, let

$$U^{-1}AU = B = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{bmatrix}$$

such that k is maximized. Then, since the eigenspace is rotated under unitary similarity, at most one B_j will have a non orthogonal eigenvector. Thus, a matrix constructed of the remaining blocks will be normal (A_n), and therefore diagonalizable. Since k is maximized, each block B_i that does not correspond to a non othogonal eigenvector will be of dimension 1 by 1. The remaining block, B_j, will be almost normal if A has a nonorthogonal eigenvector and will not exist if no such eigenvector exists. If B_j does exist, then it must be unitarily irreducible- otherwise, it would be unitarily similar to the direct sum of two matrices, and k would not have been maximized in the above representation.

In terms of representation (**??**), pure almost normal matrices can be characterized as follows.

Theorem 2.1. *An almost normal matrix A is unitarily irreducible if and only if in its canonical form* (**??**) *all β_j are different from zero and all λ_j are distinct:*

$$\beta_j \neq 0, \quad \lambda_i \neq \lambda_j \quad (i, j = 1, \dots, n-1;\ i \neq j). \tag{2.2}$$

Proof. *Necessity.* If $\beta_m = 0$ for some m, then A under a permutational similarity corresponding to the $(1, m)$ transposition turns into the direct sum of a one dimensional block (λ_m) with an almost normal matrix from $\mathbb{C}^{(n-1)\times(n-1)}$ and thus is unitarily reducible. On the other hand, if $\lambda_i = \lambda_j$, then under an appropriate unitary similarity affecting only i, j-th rows and columns we obtain from (**??**) a matrix with the same first $n-1$ columns and the zero (i, n)-entry. Thus, the already proven part of the statement applies.

Sufficiency. Direct computations show that, under conditions (**??**), $\overline{\mu}$ as an eigenvalue of A^* has geometric multiplicity one, with the corresponding eigenvector equal e_n.

Consider now a reducing subspace L of A, that is, suppose that both L and its orthogonal complement $L^\perp$ are invariant under A, and thus under A^* as well. The simple eigenvector e_n of the latter then must lie either in L or in $L^\perp$; switching the notation if necessary, without loss of generality we have $e_n \in Ł$. Since L is invariant under A, we conclude from here that

$$x := (A-\mu I)e_n = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{n-1} \\ 0 \end{bmatrix} \in L.$$

Moreover,

$$A^k x = \begin{bmatrix} \lambda_1^{k+1} \\ \lambda_2^{k+1} \\ \vdots \\ \lambda_{n-1}^{k+1} \\ 0 \end{bmatrix} \in L, \quad k = 1, 2, \dots$$

From the Vandermonde determinant formula it follows that the matrix

$$[x \ Ax \ \dots \ A^{n-2}x] = \begin{bmatrix} \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{n-1} \\ \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{n-1} \\ \vdots & \vdots & \dots & \vdots \\ \lambda_{n-1} & \lambda_{n-1}^2 & \dots & \lambda_{n-1}^{n-1} \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

has full rank, due to the second part of condition (**??**). Thus, the span of the vectors $x, \dots A^{n-2}x$ is $(n-1)$-dimensional and, lying in the span of $e_1, \dots, e_{n-1}$, actually coincides with it. Consequently, $e_1, \dots, e_{n-1} \in L$. Since we already know that $e_n \in L$ as well, in fact we must have $L = \mathbb{C}^n$. In other words, A has no non-trivial reducing subspace, and is therefore unitarily irreducible. □

For a pure almost normal matrix A all its eigenvalues have geometric multiplicity one, even if μ in the representation (**??**) coincides with one of the λ_j. Consequently, the choice of a basis in which A takes form (**??**) is unique, up to (trivial) multiplications by unimodular scalars and permutations of the first $n-1$ vectors. The canonical form (**??**) therefore also is unique, up to the permutational similarities involving the first $n-1$ rows and columns.

2.1. **Almost normality of A^*.** The definition of normality immediately implies that a matrix A is normal if and only simultaneously with A^*. We brought up this trivial observation to emphasize that for almost normality the situation changes.

Theorem 2.2. *Let A be an almost normal matrix. Then A^* is almost normal if and only if its PAN-rank is at most 2.*

Proof. Using the decomposition (**??**) for A, we observe that A^* is unitarily similar to $A_n^* \oplus A_a^*$. Since A_n^* is normal along with A_n, A^* will be almost normal only simultaneously with A_a^*.

If the PAN-rank of A is 2, then A_a^* is almost normal, as any 2-by-2 matrix. If the PAN-rank is zero, the situation is even simpler: A^* is normal, since A is. This proves the sufficiency.

Suppose now that the PAN-rank of A is at least 3. Using the canonical form of A_a, we see that A_a^* is unitarily similar to

$$\begin{bmatrix} \overline{\lambda_1} & 0 & \dots & 0 & 0 \\ 0 & \overline{\lambda_2} & \dots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & \overline{\lambda_{n-1}} & 0 \\ \beta_1 & \beta_2 & \dots & \beta_{n-1} & \overline{\mu} \end{bmatrix}. \tag{2.1}$$

The eigenvalues $\overline{\lambda_j}$ of the matrix (**??**) have geometric multiplicities one, and the respective eigenvectors, up to scalar multiples, are

$$x_j = [0 \ \dots \ 0 \ \overline{\lambda_j - \mu} \ 0 \ \dots 0 \ \beta_j]^T \tag{2.2}$$

($\overline{\lambda_j - \mu}$ in the jth position), $j = 1, \ldots, n-1$. If μ is different from all λ_j, there is also an eigenvector $x_n = e_n$ corresponding to the eigenvalue $\overline{\mu}$ of (**??**); otherwise the set $\{x_1, \ldots, x_{n-1}\}$ of the eigenvectors of (**??**) is complete.

Either way, no pair of the eigenvectors x_i, x_j is orthogonal. Thus, starting with $n = 3$, the matrix (**??**) is not almost normal. This proves the necessity. □

2.2. **Self-commutator.** Yet another way to define normality is as follows: a matrix A is normal if and only if its self-commutator $[A]$ has rank zero. Since $[A]$ is traceless for any A, non-normal matrices have self-commutators of rank at least two. On the other hand, it was shown in [**?**, Section 4] that for almost normal matrices the rank of self-commutators does not exceed three. Here is an ϵ-improvement of this result.

Theorem 2.3. *Let A be an almost normal matrix. Then*

$$\operatorname{rank}[A] = \min\{3, k\},$$

where k is the PAN-rank of A.

Proof. Since $[A] = 0 \oplus [A_a]$, it suffices to consider pure almost normal matrices only. They are not normal, so the rank of their self-commutators is between 2 and n, which proves the statement for $n = 2$. It remains to show that for a pure almost normal A with $n \geq 3$ the rank of the self-commutator equals 3.

A direct computation reveals that the self-commutator of (**??**) equals

$$-\left[\begin{array}{ccc|c} & & & \overline{(\mu - \lambda_1)}\beta_1 \\ & (\beta_i\beta_j)_{i,j=1}^{n-1} & & \vdots \\ & & & \overline{(\mu - \lambda_{n-1})}\beta_{n-1} \\ \hline (\mu - \lambda_1)\beta_1 & \ldots & (\mu - \lambda_{n-1})\beta_{n-1} & -(\beta_1^2 + \cdots + \beta_{n-1}^2) \end{array}\right].$$

With $\beta_1 \neq 0$, multiplication of the latter matrix on the left by the triangular invertible matrix

$$\frac{1}{\beta_1}[0 \ \beta_2 \ \ldots \ \beta_{n-1} \ \mu - \lambda_1]^T[1 \ 0 \ldots 0] - I$$

yields

$$\begin{bmatrix} \beta_1^2 & \beta_1\beta_2 & \ldots & \beta_1\beta_{n-1} & \overline{(\mu - \lambda_1)}\beta_1 \\ 0 & 0 & \ldots & 0 & \overline{(\lambda_1 - \lambda_2)}\beta_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \overline{(\lambda_1 - \lambda_{n-1})}\beta_{n-1} \\ 0 & (\lambda_1 - \lambda_2)\beta_2 & \ldots & (\lambda_1 - \lambda_{n-1})\beta_{n-1} & * \end{bmatrix}.$$

Under conditions (**??**) and with $n \geq 3$, the first, last, and any one additional row form a basis of the row space of $[A]$. Thus, $\operatorname{rank}[A] = 3$. □

3. Numerical range

The *numerical range* (also known as the *field of values*) of $A \in \mathbb{C}^{n\times n}$ is defined as

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

(here of course $\langle .,.\rangle$ stand for the usual scalar product on $\mathbb{C}^n$, and $\|.\|$ for the norm associated with it).

This is an extensively studied object, see e.g. [**?**]. In particular, it is known that $W(A)$ is a compact convex subset of $\mathbb{C}$ containing the spectrum of A, and thus the convex hull of the spectrum as well:

$$W(A) \supseteq \operatorname{conv} \sigma(A). \tag{3.1}$$

Moreover, the numerical range is invariant under unitary similarities. If A is unitarily similar say to $A_1 \oplus A_2$, then $W(A) = \operatorname{conv}\{W(A_1), W(A_2)\}$. This implies in particular that, for normal matrices, the equality holds in (**??**), so in this case $W(A)$ is a polygon. For almost normal A, in turn,

$$W(A) = \operatorname{conv}\{W(A_n), W(A_a)\} = \operatorname{conv}\{\sigma(A_n), W(A_a)\},$$

according to the decomposition (**??**). Consequently, only the case of pure almost normal matrices is of interest.

3.1. Roundness of $W(A)$.

Being unitarily irreducible, pure almost normal matrices cannot have sharp points on the boundary of their numerical range, since every such point is a normal eigenvalue. Flat portions on the boundary, however, are potentially possible for unitarily irreducible matrices, starting with $n = 3$, see e.g. [**?**]. Therefore it is not trivial to observe that they do not materialize in the case of pure almost normal matrices of any size.

Theorem 3.1. *Let A be a pure almost normal matrix. Then the boundary of $W(A)$ is a smooth curve with no flat portions on its boundary*

Proof. According to [**?**], it suffices to show that for any $\theta \in \mathbb{R}$ the maximal eigenvalue of $\operatorname{Re}(e^{i\theta}A)$ is simple (meaning that it is not repeated). Since the matrices $e^{i\theta}A$ are pure almost normal along with A, we need only to prove the claim for $\theta = 0$. But, for A given by (**??**),

$$\operatorname{Re} A = \begin{bmatrix} \xi_1 & 0 & \dots & 0 & \beta_1/2 \\ 0 & \xi_2 & \dots & 0 & \beta_2/2 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & \xi_{n-1} & \beta_{n-1}/2 \\ \beta_1/2 & \beta_2/2 & \dots & \beta_{n-1}/2 & \nu \end{bmatrix} \tag{3.2}$$

(here $\xi_j = \operatorname{Re} \lambda_j$, $j = 1, \dots, n-1$; $\nu = \operatorname{Re} \mu$).

Due to the interlacing eigenvalues theorem, multiple eigenvalues of (**??**), if any exist, must coincide with some ξ_j. On the other hand, ξ_j is not an eigenvalue (and thus not an endpoint of the numerical range) of the 2-by-2 block of (**??**) located in its j, n row and column, since $\beta_j \neq 0$. Consequently, for all $j = 1, \dots, n-1$, ξ_j are not endpoints of $W(\operatorname{Re} A)$, and therefore not the extremal eigenvalues of $\operatorname{Re} A$. □

3.2. The 3-by-3 case.

According to Kippenhahn's classification ([**?**], see also [**?**]), for $n = 3$ there are three possible shapes of $W(A)$ for unitarily irreducible A, and one of these types has a flat portion on the boundary. This leaves two other options available for pure almost normal matrices: the elliptical shape and the so called ovular shape. For our purposes, an "ovular" region is one defined by a curve of order 6. The next theorem allows us to distinguish easily between the two.

Theorem 3.2. *Let A be a pure almost normal 3-by-3 matrix, with a canonical form* (**??**). *Then $W(A)$ is an ellipse if*

$$\mu = \frac{\lambda_1\beta_2^2 + \lambda_2\beta_1^2}{\beta_1^2 + \beta_2^2}, \tag{3.3}$$

and has an ovular shape otherwise. Under condition (**??**) *the ellipse $W(A)$ has its foci at the eigenvalues λ_1, λ_2 and the minor axis of the length $\sqrt{\beta_1^2 + \beta_2^2}$.*

Proof. According to the ellipticity criterion from [**?**], a unitarily irreducible 3-by-3 matrix

$$\begin{bmatrix} \lambda_1 & x & y \\ 0 & \lambda_2 & z \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

has an elliptical numerical range if and only if the number

$$\frac{\lambda_1 |z|^2 + \lambda_2 |y|^2 + \lambda_3 |x|^2 - x\overline{y}z}{|x|^2 + |y|^2 + |z|^2} \tag{3.4}$$

coincides with one of the eigenvalues λ_j. In our setting (and in our notation) (**??**) simplifies to the right hand side of (**??**). Being a convex combination of λ_1 and λ_2 with positive coefficients, it cannot coincide with either of them (recall that $\lambda_1 \neq \lambda_2$ due to (**??**)). Since μ is the only remaining eigenvalue, the ellipticity criterion boils down to (**??**). The description of $W(A)$, provided that (**??**) holds, also follows from [**?**, Theorem 2.4].

Pure almost normal A will not had an elliptical numerical range when this fails, since we showed the condition is necessary. On the other hand, Theorem **??** shows that flat portions cannot occur for pure almost normal A , so the ovular shape is the only option left for $W(A)$ if (**??**) fails. □

Corollary 3.3. *A pure almost normal 3-by-3 matrix cannot have a circular disk as its numerical range.*

A circle is an ellipse in which the foci coincide. For a pure almost normal matrix with an elliptical numerical range, the foci of are λ_1 and λ_2. Since $\lambda_1 \neq \lambda_2$ for pure almost normal A according to Theorem 2.1, a circular numerical range cannot occur.

4. Aluthge transform

The *Aluthge transform* of a square matrix A is defined as

$$\Delta(A) = R^{1/2}UR^{1/2},$$

where $A = UR$ is the (right) polar representation of A. Recall that $R = (A^*A)^{1/2}$, and is therefore invertible for invertible A. Thus, for such A

$$\Delta(A) = R^{1/2}AR^{-1/2}. \tag{4.1}$$

For singular A the choice of U is not unique; $\Delta(A)$ is nevertheless still defined by A uniquely. The transformation $A \mapsto \Delta(A)$ was introduced by Aluthge [**?**], and has since been studied extensively.

4.1. **Unitary reducibility of $\boldsymbol{\Delta(A)}$.** If A is unitarily reducible, then the Aluthge transform acts on each block independently, so unitary reducibility is preserved under Δ. On the other hand, $\Delta(A)$ may be unitarily reducible even when A is not. In particular, the kernel of a singular matrix A is a reducing subspace for $\Delta(A)$. The criterion for $\Delta(A)$ to be normal (and thus unitarily reducible) in case of invertible A was obtained in [**?**, Theorems 8 and 9], stated in terms of either the polar representation of A or its unitary similarity to a certain canonical form. In the case of almost normal matrices, we propose here an alternative treatment which yields a more constructive result.

Noting that a normal eigenvector is a matrix A is an eigenvector of both A and A^*

Lemma 4.1. *Let $A \in \mathbb{C}^{n\times n}$ be an invertible pure almost normal matrix, $n \geq 3$. Then $\Delta(A)$ has no normal eigenvectors.*

Proof. Suppose that $\Delta(A)$ and $\Delta(A)^*$ have a common eigenvector x:

$$\Delta(A)x = \zeta x, \quad \Delta(A)^*x = \overline{\zeta}x.$$

Using (**??**), these equalities can be equivalently rewritten in terms of A and R:

$$R^{1/2}AR^{-1/2}x = \zeta x, \quad R^{-1/2}A^*R^{1/2}x = \overline{\zeta}x.$$

Denoting $R^{-1/2}x = y$ and $R^{1/2}x = z$, we conclude from here that $Ay = \zeta y$, $A^*z = \overline{\zeta}z$. In other words, y is an eigenvector of A and z is an eigenvector of A^* corresponding to complex conjugate eigenvalues. It is crucial for the rest of the proof that

$$z = R^{1/2}x = R^{1/2}(R^{1/2}y) = Ry. \tag{4.2}$$

We now consider separately two situations, depending on which eigenvalue of A plays the role of ζ.

Case 1. $\zeta = \lambda_j$ for some j. Without loss of generality then y is e_j, and z is a scalar multiple of the vector (**??**). Due to (**??**), z actually is the jth column of R. Since R is positive definite, its diagonal entries are positive, so μ must be different from λ_j. Moreover,

$$z = c[0\ \dots\ 0\ 1\ 0\ \dots\ 0\ \beta_j/\overline{\lambda_j - \mu}]^T$$

for some $c > 0$.

Consequently, in the notation $R = (r_{ik})_{i,k=1}^n$ we have

$$r_{ij} = r_{ji} = 0 \text{ for } i \neq j, n.$$

Thus, for $i \neq j$ the (i,j)-entry of R^2 equals

$$\textstyle\sum_{k=1}^n r_{ik}r_{kj} = r_{in}r_{nj}.$$

But

$$R^2 = A^*A = \left[\begin{array}{ccc|c} |\lambda_1|^2 & & & \beta_1\overline{\lambda_1} \\ & \ddots & & \vdots \\ & & |\lambda_{n-1}|^2 & \beta_{n-1}\overline{\lambda_{n-1}} \\ \hline \beta_1\lambda_1 & \dots & \beta_{n-1}\lambda_{n-1} & |\mu|^2 + \sum_{j=1}^{n-1}\beta_j^2 \end{array}\right] \tag{4.3}$$

has zero (i,j)-entries for all $i \neq j, n$. Since $r_{nj} \neq 0$, we conclude from here that $r_{in} = 0$, $i \neq j, n$.

Choosing any such i (which is possible, starting with $n = 3$), we observe that the (i,n)-entry of R^2 is

$$\textstyle\sum_{k=1}^n r_{ik}r_{kn} = r_{ij}r_{jn} + r_{in}r_{nn} = 0,$$

which is in contradiction with (**??**).

Case 2. $\zeta = \mu$ and is different from all λ_j, $j = 1, \ldots, n-1$. Without loss of generality, $z = e_n$, while

$$y = c \begin{bmatrix} \dfrac{\beta_1}{\mu - \lambda_1} & \ldots & \dfrac{\beta_{n-1}}{\mu - \lambda_{n-1}} & 1 \end{bmatrix}^T \tag{4.4}$$

for some scalar multiple c. According to (**??**), (**??**) is nothing but the last column of R^{-1}. Moreover, due to the positive definiteness of R (and then of R^{-1} as well), the constant c must be positive.

From (**??**) and the trivial equality $R = R^2 R^{-1}$ we conclude that for $j = 1, \ldots, n-1$:

$$r_{jj} = |\lambda_j|^2 (R^{-1})_{jj} + c\beta_j^2 \overline{\eta_j}, \text{ where } \eta_j = \frac{\lambda_j}{\mu - \lambda_j}.$$

(Here $(R^{-1})_{jj}$ is the j-th diagonal entry of R^{-1} while, as in Case 1, we denote the entries of R by r_{ij}.) Since $r_{jj}, (R^{-1})_{jj} > 0$, all η_j must be real. Equivalently, $\lambda_j/\mu \in \mathbb{R}$, $j = 1, \ldots, n-1$. Considering $e^{-i \arg \mu} A$ in place of A if needed, we may without loss of generality suppose that $\mu, \lambda_j \in \mathbb{R}$, $j = 1, \ldots, n-1$.

With this simplification in mind, and using subsequently (**??**) and (**??**), we compute the last column of R as

$$w := Re_n = R^2 R^{-1} e_n = R^2 y = cR^2 \begin{bmatrix} \frac{\beta_1}{\mu - \lambda_1} \\ \vdots \\ \frac{\beta_{n-1}}{\mu - \lambda_{n-1}} \\ 1 \end{bmatrix} = c\mu \begin{bmatrix} \beta_1 \lambda_1/(\mu - \lambda_1) \\ \vdots \\ \beta_{n-1}\lambda_{n-1}/(\mu - \lambda_{n-1}) \\ \mu + \sum_{j=1}^{n-1} \beta_j^2/(\mu - \lambda_j) \end{bmatrix}.$$

Since $w^T w$ is the (n,n)-entry of R^2, from here and (**??**) we conclude:

$$\mu^2 + \sum_{j=1}^{n-1} \beta_j^2 = c^2\mu^2 \left(\sum_{j=1}^{n-1} \frac{\beta_j^2 \lambda_j^2}{(\mu - \lambda_j)^2} + \left(\mu + \sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu - \lambda_j} \right)^2 \right). \tag{4.5}$$

On the other hand,

$$1 = \langle e_n, e_n \rangle = \langle Re_n, R^{-1} e_n \rangle = \langle w, y \rangle = c^2 \mu^2 \left(1 + \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu - \lambda_j)^2} \right). \tag{4.6}$$

Plugging in the value of $c^2\mu^2$ from (**??**) into (**??**), we arrive at

$$\left(\mu^2 + \sum_{j=1}^{n-1} \beta_j^2 \right) \left(1 + \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu - \lambda_j)^2} \right) = \sum_{j=1}^{n-1} \frac{\beta_j^2 \lambda_j^2}{(\mu - \lambda_j)^2} + \left(\mu + \sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu - \lambda_j} \right)^2.$$

We see that the left side is equal to

$$\left(\mu^2 + \sum_{j=1}^{n-1} \beta_j^2\right)\left(1 + \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu-\lambda_j)^2}\right) = \mu^2 + \sum_{j=1}^{n-1} \beta_j^2 + \mu^2 \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu-\lambda_j)^2} + \sum_{j=1}^{n-1} \beta_j^2 \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu-\lambda_j)^2}$$

,

while the right side is equal to

$$\sum_{j=1}^{n-1} \frac{\beta_j^2\lambda_j^2}{(\mu-\lambda_j)^2} + \left(\mu + \sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu-\lambda_j}\right)^2 = \sum_{j=1}^{n-1} \frac{\beta_j^2\lambda_j^2}{(\mu-\lambda_j)^2} + \mu^2 + \left(\sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu-\lambda_j}\right)^2 + 2\mu \sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu-\lambda_j}$$

.

So

$$\mu^2 + \sum_{j=1}^{n-1} \beta_j^2 + \mu^2 \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu-\lambda_j)^2} + \left(\sum_{j=1}^{n-1} \beta_j^2\right) \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu-\lambda_j)^2} = \sum_{j=1}^{n-1} \frac{\beta_j^2\lambda_j^2}{(\mu-\lambda_j)^2} + \mu^2 + \left(\sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu-\lambda_j}\right)^2 + 2\mu \sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu-\lambda_j}$$

, or, canceling the μ^2 on both sides,

$$\sum_{j=1}^{n-1} \beta_j^2 + \mu^2 \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu-\lambda_j)^2} + \left(\sum_{j=1}^{n-1} \beta_j^2\right) \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu-\lambda_j)^2} = \sum_{j=1}^{n-1} \frac{\beta_j^2\lambda_j^2}{(\mu-\lambda_j)^2} + \left(\sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu-\lambda_j}\right)^2 + 2\mu \sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu-\lambda_j}$$

We show that

$$\sum_{j=1}^{n-1} \beta_j^2 + \mu^2 \sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu-\lambda_j)^2} = \sum_{j=1}^{n-1} \frac{\beta_j^2\lambda_j^2}{(\mu-\lambda_j)^2} + 2\mu \sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu-\lambda_j}$$

,

which follows from

$$\begin{aligned} \beta_j^2 + \frac{\mu^2\beta_j^2}{(\mu-\lambda_j)^2} &= \frac{\beta_j^2\lambda_j^2}{(\mu-\lambda_j)^2} + 2\mu\frac{\beta_j^2}{\mu-\lambda_j} \\ \to \beta_j^2(\mu-\lambda_j)^2 + \mu^2\beta_j^2 &= \beta_j^2\lambda_j^2 + 2\mu\beta_j^2(\mu-\lambda_j) \\ \to (\mu-\lambda_j)^2 + \mu^2 &= \lambda_j^2 + 2\mu(\mu-\lambda_j) \\ \to 2\mu^2 - 2\mu\lambda + \lambda_j^2 &= 2\mu^2 - 2\mu\lambda + \lambda^2 \end{aligned}$$

and is therefore correct. This leaves us with

$$\left(\sum_{j=1}^{n-1} \beta_j^2\right)\left(\sum_{j=1}^{n-1} \frac{\beta_j^2}{(\mu-\lambda_j)^2}\right) = \left(\sum_{j=1}^{n-1} \frac{\beta_j^2}{\mu-\lambda_j}\right)^2. \tag{4.7}$$

By the Cauchy-Schwarz inequality, (**??**) holds if and only if the vectors $[\beta_1 \ \dots \ \beta_{n-1}]$ and $[\beta_1/(\mu - \lambda_1) \ \dots \ \beta_{n-1}/(\mu - \lambda_{n-1})]$ are collinear. But, starting with $n = 3$, this is not the case — a contradiction. $\square$

With Lemma **??** at our disposal, several significant results follow.

Theorem 4.2. *Let A be an invertible 3-by-3 pure almost normal matrix. Then $\Delta(A)$ is unitarily irreducible.*

PROOF. $\Delta(A)$ has no normal eigenvalues according to Lemma **??**. If a 3-by-3 matrix was unitarily reducible, it would be unitarily similar to a direct sum of either 3 singletons or a 2-by-2 matrix and a singleton. In either case, the eigenvector corresponding to the singleton would not change when the conjugate transpose is taken. Thus, an eigenvector corresponding to a singleton would be normal. This would contradict Lemma **??**, so we cannot have $\Delta(A)$ be unitarily reducible for 3-by-3 A. $\square$

We also establish a criterion for $\Delta(A)$ to be normal when A is almost normal and invertible. $\Delta(A)$ may be normal for 2-by-2 non-normal (and thus pure almost normal) matrices A. This shows that the restriction on n in Lemma **??** is essential.

Theorem 4.3. *Let A be an invertible almost normal matrix. Then $\Delta(A)$ is normal if and only if either A is normal itself or its PAN-rank equals 2 and the eigenvalues μ_1, μ_2 of the block A_a of its representation* (**??**) *satisfy $\overline{\mu_1}\mu_2 \leq 0$.*

PROOF. The Aluthge transform of A is unitarily similar to $\Delta(A_n) \oplus \Delta(A_a)$. Since $\Delta(A_n) = A_n$ is normal, $\Delta(A)$ is normal if and only if A_a is either absent (that is, A itself is normal) or $\Delta(A_a)$ is normal. According to Lemma **??**, this cannot happen if the size of A_a is bigger than 2. On the other hand, for a 2-by-2 matrix A_a the normality of its Aluthge transform is given by [**?**, Corollary 1] and amounts exactly to the condition $\overline{\mu_1}\mu_2 \leq 0$ on its eigenvalues μ_1, μ_2. $\square$

4.2. **Almost normal matrices in the range of $\boldsymbol{\Delta}$.** The Aluthge transform is a non-linear mapping of $\mathbb{C}^{n\times n}$ into itself. We now address the question of which pure almost normal matrices lie in its range. To formulate the result, an additional notion has to be introduced. We partition the set $\{\lambda_1, \dots, \lambda_{n-1}\}$ from the representation (**??**) into equivalence classes according to the equivalence relation

$$\lambda_i \equiv \lambda_j \text{ if and only if } \arg\lambda_i = \arg\lambda_j \mod \pi,$$

and call them *clusters*. In other words, each cluster is formed by all λ_i lying on the same line passing through the origin. Let $J_1, \dots, J_k$ be the respective partition of the index set $\{1, \dots, n-1\}$. This is the case because if arg λ_i = arg λ_j mod π, then arg λ_i = arg λ_j or arg λ_i = arg $-\lambda_j$, both of which are equivalent to the two points lying on a line that goes through the origin.

Theorem 4.4. *A pure almost normal matrix A lies in the range of the Aluthge transform if and only if it is invertible,*

$$c_i := \sum_{j\in J_i} \frac{\beta_j^2}{\lambda_j} \neq 0 \tag{4.8}$$

for each cluster consisting of more then one element, and all the eigenvalues of $A^{-1}A^*$ have absolute value one.

PROOF. Suppose $A = \Delta(B)$ for some B. If A is not invertible then neither is B, since the spectrum is invariant under Aluthge transform. This would imply unitary reducibility of A, a contradiction. Thus, invertibility of A is necessary, and will be imposed in the rest of the proof.

According to [**?**, Corollary 3], an invertible matrix A lies in the range of Δ if and only if $A^{-1}A^*$ is similar to a unitary matrix, that is, it is diagonalizable with all the eigenvalues having absolute value one. So, it remains to show that conditions (**??**) are necessary and sufficient for $A^{-1}A^*$ to be diagonalizable.

For technical reasons, it is more convenient to consider A^*A^{-1} in place of $A^{-1}A^*$; the two matrices are similar, so this switch is allowed. Since A is unitarily similar to (**??**), in its turn A^*A^{-1} is (unitarily) similar to

$$
\left[\begin{array}{ccc|c} \overline{\lambda_1} & & & 0 \\ & \ddots & & \vdots \\ & & \overline{\lambda_{n-1}} & 0 \\ \hline \beta_1 & \dots & \beta_{n-1} & \overline{\mu} \end{array}\right]
\left[\begin{array}{ccc|c} {\lambda_1}^{-1} & & & -\frac{\beta_1}{\mu\lambda_1} \\ & \ddots & & \vdots \\ & & {\lambda_{n-1}}^{-1} & -\frac{\beta_{n-1}}{\mu\lambda_{n-1}} \\ \hline 0 & \dots & 0 & \mu^{-1} \end{array}\right]
\tag{4.9}
$$

$$
= \left[\begin{array}{ccc|c} \overline{\lambda_1}/\lambda_1 & & & -\frac{\beta_1\overline{\lambda_1}}{\mu\lambda_1} \\ & \ddots & & \vdots \\ & & \overline{\lambda_{n-1}}/\lambda_{n-1} & -\frac{\beta_{n-1}\overline{\lambda_{n-1}}}{\mu\lambda_{n-1}} \\ \hline \beta_1/\lambda_1 & \dots & \beta_{n-1}/\lambda_{n-1} & \frac{\overline{\mu}}{\mu} - \frac{1}{\mu}\sum_{j=1}^{n-1}\frac{\beta_j^2}{\lambda_j} \end{array}\right].
$$

A multiple eigenvalue ω of (**??**), if it exists, has to also be an eigenvalue of its left upper $(n-1)$-by-$(n-1)$ submatrix, and thus coincide with one of $\overline{\lambda_j}/\lambda_j$. By inspection it is easy to see that ω is an eigenvalue of (**??**) with geometric multiplicity $m-1$, where m is the size of the cluster containing λ_j.

On the other hand, (**??**) is an arrow-head matrix, and it remains such after subtracting a scalar multiple of the identity. Thus, for the characteristic polynomial we obtain:

$$
\det(A^{-1}A^* - \xi I) = \det(A^*A^{-1} - \xi I) \tag{4.10}
$$

$$
= \left(\frac{\overline{\mu}}{\mu} - \xi\right)\prod_{j=1}^{n-1}\left(\frac{\overline{\lambda_j}}{\lambda_j} - \xi\right) + \frac{\xi}{\mu}\sum_{j=1}^{n-1}\left(\frac{\beta_j^2}{\lambda_j}\prod_{k\neq j}\left(\frac{\overline{\lambda_k}}{\lambda_k} - \xi\right)\right).
$$

The algebraic multiplicity of ω is therefore equal to $m-1$ if condition (**??**) holds for the cluster containing λ_j, and is not smaller than m otherwise. Consequently, (**??**) is necessary and sufficient for geometric and algebraic multiplicities of all the eigenvalues of $A^{-1}A^*$ to coincide, that is, for the matrix to be diagonalizable. □

Note that for any invertible A the characteristic polynomial of $A^{-1}A^*$ is *self-inversive*, that is, satisfies the identity

$$
f(z) = \kappa z^{\deg f}\overline{f(1/\overline{z})}
$$

for some unimodular constant κ. Indeed,

$$
\overline{\det(A^{-1}A^* - \overline{z}^{-1}I)} = \det(A^{-1}A^* - \overline{z}^{-1}I)^*
$$

$$
= \det(AA^{*-1} - z^{-1}I) = z^{-n}\det A(\overline{\det A})^{-1}\det(A^{-1}A^* - zI).
$$

There are various tests known for such polynomials to have all roots on the unit circle, see e.g. [**?**, **?**, **?**]. We, however, demonstrate the applicability of Theorem **??** in a particular situation when the root location for (**??**) can be handled rather elementary.

Corollary 4.5. *Let A be an invertible pure almost normal matrix with real eigenvalues $\lambda_1, \dots, \lambda_{n-1}$ corresponding to its orthogonal eigenvectors. Then A lies in the range of the Aluthge transform if and only if*

$$
c := \sum_{j=1}^{n-1}\frac{\beta_j^2}{\lambda_j} \neq 0 \tag{4.11}
$$

and

$$
|c - 2\operatorname{Re}\mu| \leq 2|\mu|. \tag{4.12}
$$

Proof. If all λ_j are real, they all lie on the same line passing through the origin (namely, the real axis), there is only one cluster, and thus (**??**) takes the form (**??**). On the other hand, (**??**) simplifies to

$$
\frac{1}{\mu}(1-\xi)^{n-2}(\mu\xi^2 + (c - 2\operatorname{Re}\mu)\xi + \overline{\mu}). \tag{4.13}
$$

It remains to observe that the roots of the quadratic factor in (**??**) lie on the unit circle if and only if its discriminant is non-positive, which is equivalent to (**??**). This holds because the discriminant of the quadratic term is $c^2 - 4c\mathrm{Re}\mu + 4\mathrm{Re}\mu^2 - 4\mu\bar{\mu} = c^2 - 4c\mathrm{Re}\mu + 4\mathrm{Re}\mu^2 - 4|\mu|^2$ so we see that the following are equivalent

$$\begin{aligned} c^2 - 4c\mathrm{Re}\mu + 4\mathrm{Re}\mu^2 - 4|\mu|^2 &\leq 0 \\ c^2 - 4c\mathrm{Re}\mu + 4\mathrm{Re}\mu^2 &\leq 4|\mu|^2 \\ |c - 2\mathrm{Re}\mu| &\leq 2|\mu| \end{aligned}$$

□

If μ is also real, then (**??**) simplifies further to

$$0 \leq \frac{c}{\mu} \leq 4.$$

4.3. **The Aluthge Transform of an Almost Normal Matrix.** The conditions under which $\Delta(A)$ will be almost normal for almost normal A has not yet been completely determined at the time of this writing. We present some preliminary results for particular forms of matrices. First, note that all 2 by 2 matrices are almost normal. Thus, the Althuge will always be almost normal when $n = 2$. However the situation changes with $n \neq 3$.

Theorem 4.6. *Let A be an invertible pure almost normal matrix. Then $\Delta(A)$ cannot be almost normal with the same multiset of eigenvalues corresponding to pairwise orthogonal eigenvectors as A.*

Proof. Let

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & \beta_1 \\ 0 & \lambda_2 & 0 & \dots & \beta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{n-1} & \beta_{n-1} \\ 0 & 0 & \dots & 0 & \mu \end{pmatrix}$$

be an invertible pure almost normal matrix and let $A = UR$ be the polar decomposition of A. Then $\Delta(A) = R^{1/2}UR^{1/2}$. Let $e_1, e_2, ..., e_{n-1}$ by the orthogonal eigenvectors of A Since A is invertible, so is R so $\Delta(A) = R^{1/2}URR^{-1/2} = R^{1/2}AR^{-1/2}$. Because this is now a similarity transformation on A, the corresponding eigenvectors of $\Delta(A)$, $v_i = R^{1/2}e_i$. Let the first $n-1$ eigenvectors of $\Delta(A)$ be orthogonal. Then $(R^{1/2}e_i, R^{1/2}e_j) = 0$ for $i \neq j$, which is equivalent to $(Re_i, e_j) = 0$ for $i \neq j$. e_i, e_j are orthogonal, so this is equivalent to R being an arrowhead matrix. However, $R = (A^*A)^{1/2}$ since $A = UR$ is the polar decomposition of A. Then $R^2 = A^*A$ and

$$A^*A = \begin{pmatrix} \lambda_1\bar{\lambda}_1 & 0 & 0 & \dots & \bar{\lambda}_1\beta_1 \\ 0 & \lambda_2\bar{\lambda}_2 & 0 & \dots & \bar{\lambda}_2\beta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{n-1}\bar{\lambda}_{n-1} & \bar{\lambda}_{n-1}\beta_{n-1} \\ \lambda_1\beta_1 & \lambda_2\beta_2 & \dots & \lambda_{n-1}\beta_{n-1} & \beta_1^2 + \beta_2^2 + \dots + \beta_{n-1}^2 + \mu\bar{\mu} \end{pmatrix}$$

However, if

$$R = \begin{pmatrix} a_1 & 0 & 0 & \dots & b_1 \\ 0 & a_2 & 0 & \dots & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & a_{n-1} & b_{n-1} \\ \bar{b}_1 & \bar{b}_2 & \dots & \bar{b}_{n-1} & d \end{pmatrix}$$

then

$$R^2 = \begin{pmatrix} a_1^2 + b_1\bar{b}_1 & b_1\bar{b}_2 & b_1\bar{b}_3 & \dots & b_1(a_1 + d) \\ \bar{b}_1 b_2 & a_2 + b_2\bar{b}_2 & b_2\bar{b}_3 & \dots & b_2(a_2 + d) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \bar{b}_1 b_{n-1} & \bar{b}_2 b_{n-1} & \dots & a_{n-1} + b_{n-1}\bar{b}_{n-1} & b_{n-1}(a_{n-1} + d) \\ \bar{b}_1(a_1 + d) & \bar{b}_2(a_2 + d) & \dots & \bar{b}_{n-1}(a_{n-1} + d) & d^2 + \sum(b_i\bar{b}_i) \end{pmatrix}$$

For these two to be equal, then at least one $b_i, \bar{b}_j$ must equal 0 for each off diagonal entry of R^2. Note that these entries do not appear when $n = 2$, this does not apply in the 2 by 2 case. If all $c, b = 0$, then R^2 is a diagonal matrix. If $R^2 = A^*A$, then A^*A would also have to be diagonal- however, this can't occur since $\beta_i \neq 0$ and $\lambda_i \neq \lambda_j$ $(i \neq j)$. Thus, not all $b_i, \bar{b}_j = 0$. Furthermore, consider $R_{s,t}$ and $R_{t.s}$ for $s \neq t$. These entries must be 0, so at least 1 of $b_s, \bar{b}_t$ and $b_t, \bar{b}_s$ must equal 0. Assume, without loss of generality. let $b_t = 0$ and $\bar{b}_t = 0$. Then $R^2_{t,n} = R^2_{s,n} = 0$. For A^*A this means that $\lambda_s \beta_s = 0$ and $\bar{\lambda}_t \beta_t = 0$. This cannot occur as shown above. Thus, R^2 can't be an arrowhead matrix if R and so R can't be an arrowhead matrix, so $\Delta(A)$ isn't almost normal with the same set of orthogonal eigenvectors. □

However, we find that it is possible for $\Delta(A)$ to be almost normal with different eigenvalues corresponding to orthogonal eigenvectors, which is only possible when $\mu \neq \lambda_j$ for all j. We consider $n = 3$, letting

$$A = \begin{pmatrix} \lambda_1 & 0 & \beta_1 \\ 0 & \lambda_2 & \beta_2 \\ 0 & 0 & \mu \end{pmatrix}$$

then the eigenvector corresponding to μ is determined, up to scalar multiplication, as

$$x = \begin{pmatrix} \frac{\beta_1}{\mu - \bar{\lambda}_1} \\ \frac{\beta_2}{\mu - \bar{\lambda}_2} \\ 1 \end{pmatrix}$$

.

We show, without loss of generality, that it is possible for the eigenvector corresponding to μ to become orthogonal to that of λ_1 (denoted as e_1) under the Aluthge transform. Moreover, it is possible for this to occur when all elements of A and R are real. R and R^2 are still defined as they were in the previous proof. Because of the equality of R^2 and A^*A, we can determine the terms of A given the terms of R. Letting all $R_{i,j}$ be real (making R symmetric) , the following equalities arise:

$$\lambda_1^2 = r_{11}^2 + r_{12}^2 + r_{13}^2 \tag{4.14}$$

$$\lambda_1^2 = r_{12}^2 + r_{22}^2 + r_{23}^2 \tag{4.15}$$

$$(r_{11} + r_{22})r_{12} + r_{12}r_{23} = 0 \tag{4.16}$$

$$(r_{11} + r_{33})r_{13} + r_{12}r_{23} = \beta_1\lambda_1 \tag{4.17}$$

$$(r_{22} + r_{33})r_{23} + r_{12}r_{13} = \beta_2\lambda_2 \tag{4.18}$$

$$\mu = \frac{\det(\mathrm{R})}{\lambda_1\lambda_2} \tag{4.19}$$

For e_1 to become orthogonal to x under the Aluthge transform, it must be that e_1 is orthogonal to Rx. This means

$$\frac{r_{11}\beta_1}{\mu - \lambda_1} + \frac{r_{12}\beta_2}{\mu - \lambda_2} + r_{13} = 0$$

Fixing all r_{ij} aside from r_{33} we that β_1, β_2, μ will be the only parameters in the equation which vary and all are linear functions of r_{33}. The equation as a whole will be a quadratic function of r_{33}. An additional condition must be imposed, however, as R arises from the polar decomposition of A and must therefore be positive definite (since A is invertible, R may not be positive semi-definite). We show that such can occur via example.

Our calculations begin by fixing, somewhat arbitrarily, $r_{11} = r_{22} = r_{23} = 1$ and $r_{13} = .5$ Then, because of (6.16), $r_{12} = -.25$. We then treat $r_{33} = x$ as our variable. Because $A^*A = R^2$, this also allows us to uniquely determine the values of the parameters of A.

$$\lambda_1 = (r_{11}^2 + r_{12}^2 + r_{13}^2)^{1/2} = 1.1456$$
$$\lambda_2 = (r_{12}^2 + r_{22}^2 + r_{23}^2)^{1/2} = 1.4361$$
$$\beta_1 = \frac{r_{11}r_{13} + r_{12}r_{23} + r_{13}x}{\lambda_1} = \frac{8 \cdot 21^{1/2}(x + .5)}{21}$$
$$\beta_2 = \frac{r_{12}r_{13} + r_{22}r_{23} + r_{23}x}{\lambda_2} = \frac{4 \cdot 33^{1/2}(x + \frac{7}{8})}{33}$$
$$\mu = -\frac{(xr_{12}^2 - 2r_{12}r_{13}r_{23} + r_{22}r_{13}^2 + r_{11}r_{23}^2 - r_{11}r_{22}x)}{(r_{11}^2 + r_{12}^2 + r_{13}^2)^{1/2}(r_{12}^2 + r_{22}^2 + r_{23}^2)^{1/2}} = \frac{16 * 77^{1/2}(\frac{15}{16}x - 1.5)}{231}$$

Then the orthogonality condition will be satisfied whenever

$$\frac{12196929975433925x^2 - 73691793049977520x + 86682728142853184}{48 * (264569985433600 * x^2 - 2045394144563019x + 3936050964178994)} = 0$$

which is equivalent to

$$12196929975433925x^2 - 73691793049977520x + 86682728142853184 = 0$$

so long as the denominator remains nonzero. This is satisfied at two points- $\frac{8}{5}$, which results in $\mu = 0$, violating our assumption that A is invertible, and approximately 4.4418.

Thus,

$$A \doteq \begin{pmatrix} 1.1456 & 0 & 2.1568 \\ 0 & 1.4361 & 3.7022 \\ 0 & 0 & 1.6193 \end{pmatrix}$$

and

$$R \doteq \begin{pmatrix} 1 & -.25 & .5 \\ -.25 & 1 & 1 \\ .5 & 1 & 4.4418 \end{pmatrix}$$

We see that the eigenvalues of R are approximately .04561, 1.2276, and 4.7581, so we have positive definiteness. Moreover,

$$\Delta(A) = R^{1/2}AR^{-1/2} \doteq \begin{pmatrix} .9169 & -.4066 & .8433 \\ -.3806 & .7512 & 1.6578 \\ -.3423 & -.04693 & 2.5330 \end{pmatrix}$$

which has the approximate eigenvector

$$\begin{pmatrix} -.0603 \\ -.8785 \\ -.4739 \end{pmatrix}$$

correspond to eigenvalue of $1.6193 \doteq \mu$ and

$$\begin{pmatrix} 0.9691 \\ 0.2484 \\ -0.1832 \end{pmatrix}$$

corresponding to the eigenvalue of $1.1456 \doteq \lambda_1$.

For confirmation, we calculated the dot product of these two vectors and found it to be on the order of 10^{-15}, quite close to machine ϵ. However, the dot product of the eigenvectors corresponding to 1.4361 and 1.6193 are not orthogonal, so we do not have a normal matrix.

5. Appendix

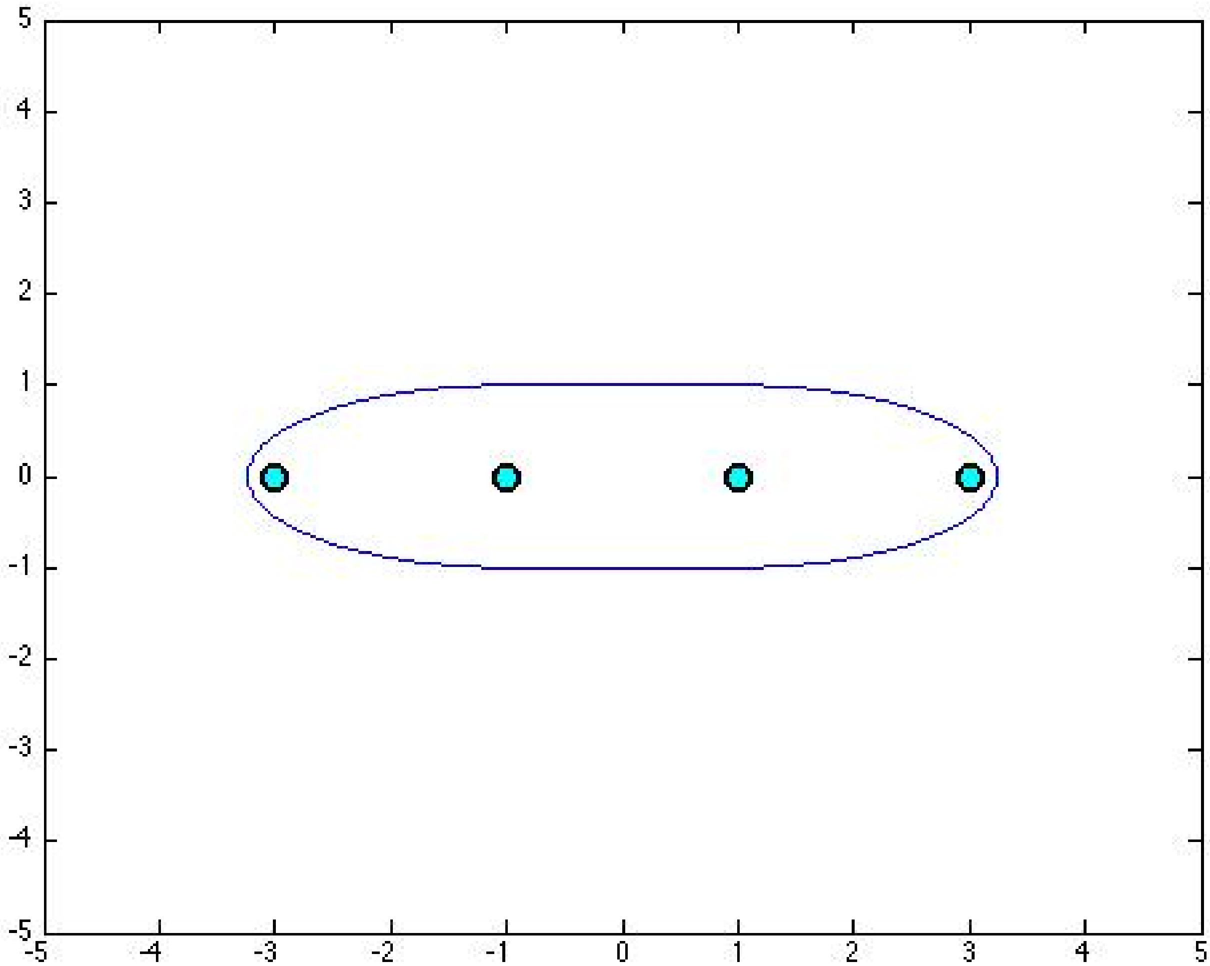
The following includes several examples of matrix K and the boundary of $\Delta(K)$ in the complex plane (with the imaginary part being denoted on the vertical axis and real part being denoted on the horizontal axis). The circles within the boundary represent the eigenvalues of that matrix.

The first represents a matrix with two horizontal portions on the boundary of its numerical range, and the second represents a 90 degree rotation of said matrix (so its numerical range now has vertical flat portions). The third shows a an almost normal matrix with an elliptical numerical range, while the fourth gives an example of a matrix with an ovular numerical range. Lastly, a normal matrix with a triangular numerical range is shown.

All images were generated in Matlab.

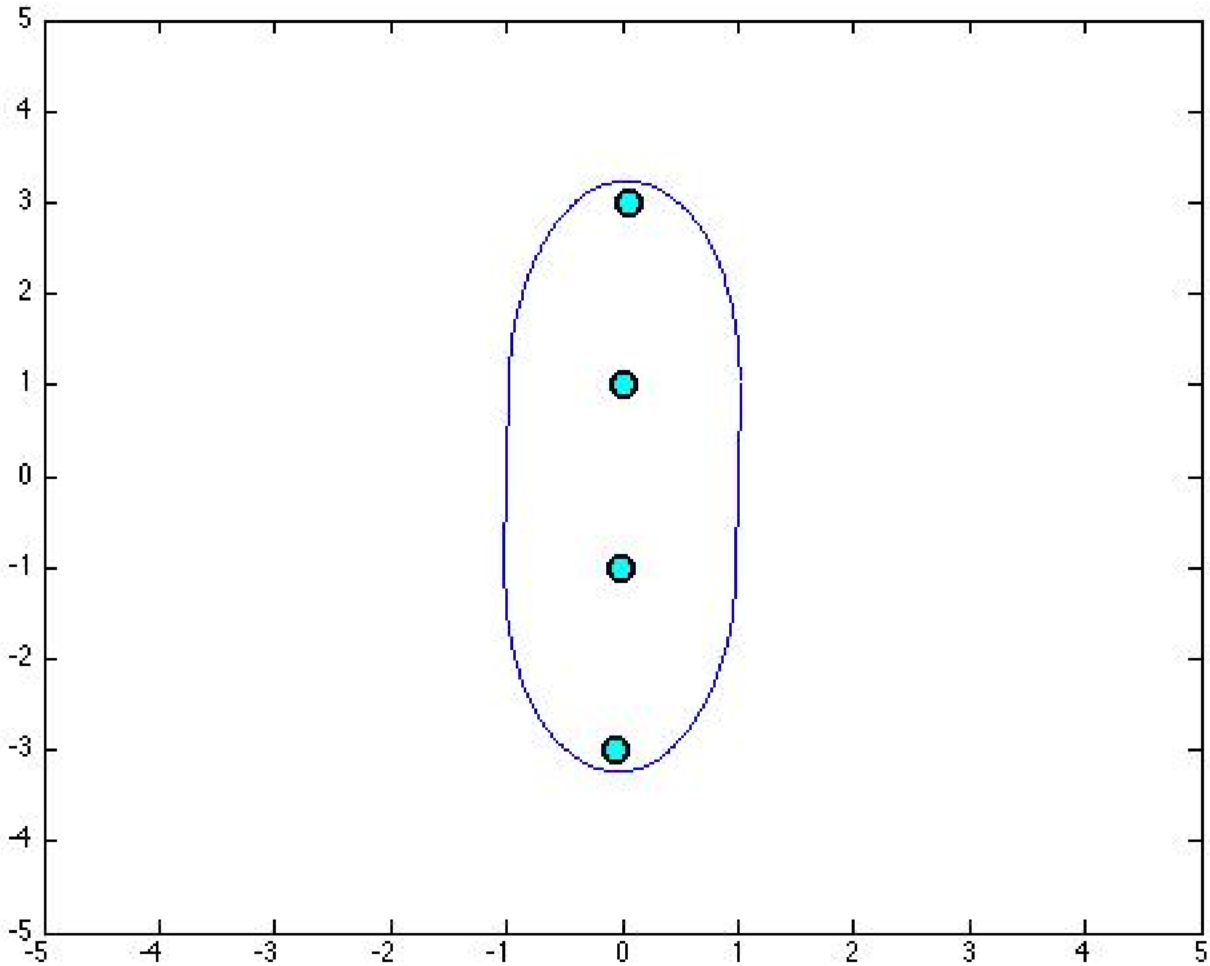
This matrix has flat portions on the boundary of its numerical range and, therefore, cannot be pure almost normal.

$$K = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 3 & 0 & 2 & 0 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$



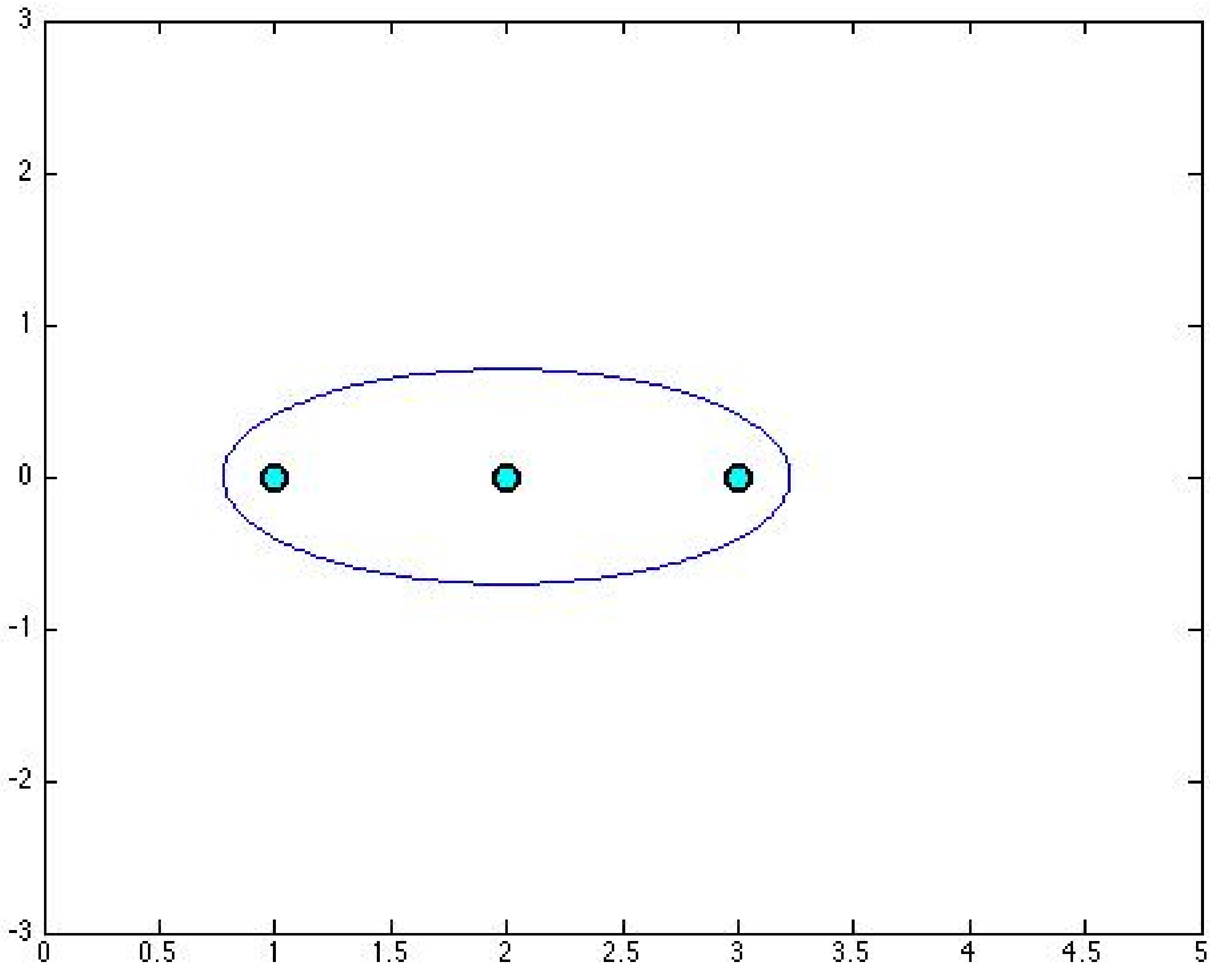
Via a rotation, the matrix from the preceding page now has two vertical flat portions on the boundary of its numerical range.

$$e^{i\frac{pi}{2}}K = \begin{pmatrix} 0 & 0+1i & 0 & 0 \\ 0+3i & 0 & 0+2i & 0 \\ 0 & 0+2i & 0 & 0+3i \\ 0 & 0 & 0+1i & 0 \end{pmatrix}$$

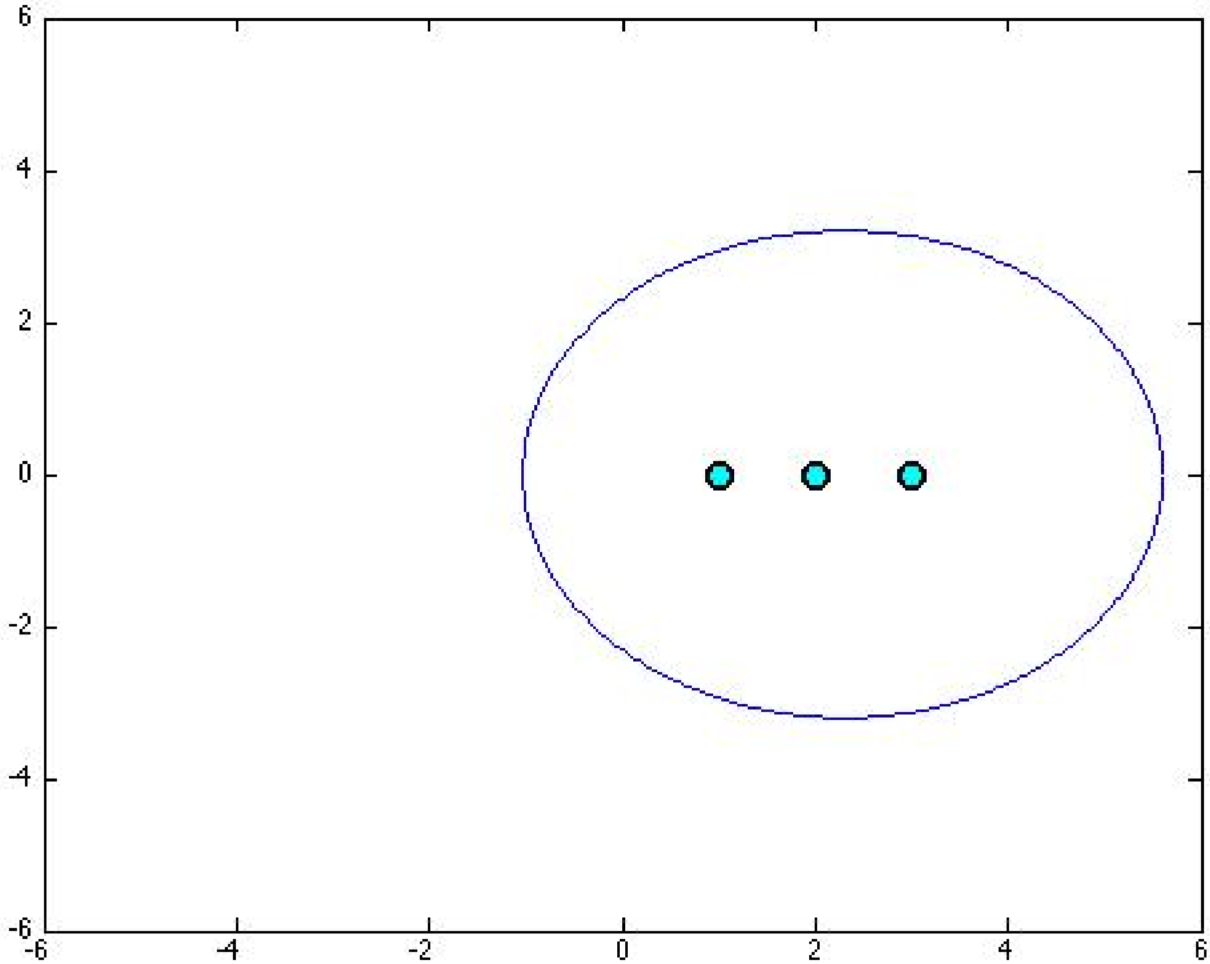


This is an almost normal matrix that satisfies (5.3) since $\frac{1^2 1 + 1^2 3}{1^2 + 1^2} = \frac{4}{2} = 2$

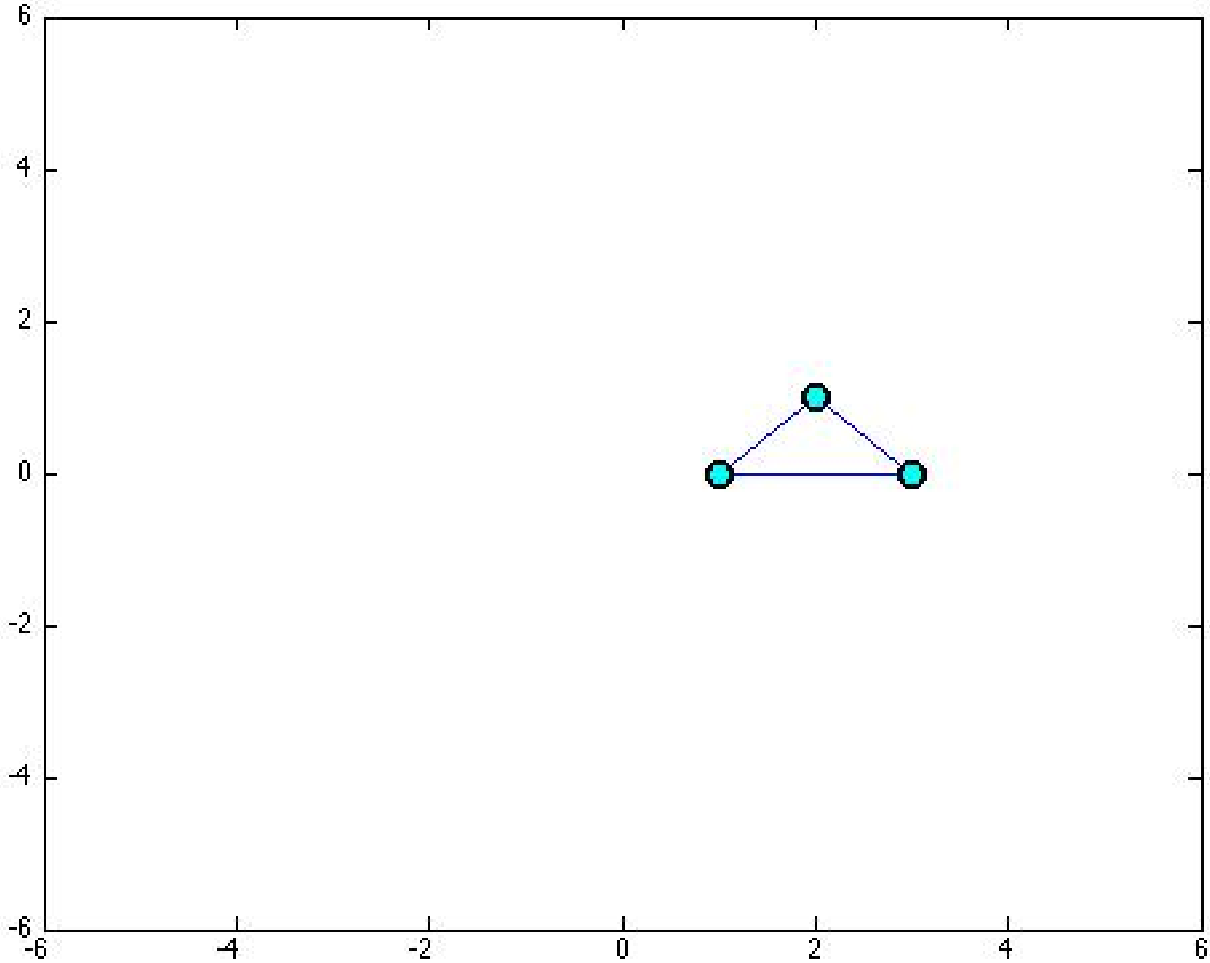
$$K = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$



$$K = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & 5 \\ 0 & 0 & 3 \end{pmatrix}$$



$$K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2+i & 0 \\ 0 & 0 & 3 \end{pmatrix}$$



References

[1] A. Aluthge, *On p-hyponormal operators for* $0 < p < 1$, Integral Equations and Operator Theory **13** (1990), no. 3, 307–315.
[2] F. F. Bonsall and M. Marden, *Zeros of self-inversive polynomials*, Proc. Amer. Math. Soc. **3** (1952), 471–475.
[3] L. Elsner and M. H. C. Paardekooper, *On measures of nonnormality of matrices*, Linear Algebra Appl. **92** (1987), 107–123.
[4] M. Fiedler, *Numerical range of matrices and Levinger's theorem*, Linear Algebra Appl. **220** (1995), 171–180.
[5] K. E. Gustafson and D. K. M. Rao, *Numerical range. The field of values of linear operators and matrices*, Springer, New York, 1997.
[6] Kh. D. Ikramov, *On almost normal matrices*, Vestnik Moskov. Univ. Ser. XV Vychisl. Mat. Kibernet. (2011), no. 1, 5–9, 56.
[7] D. Keeler, L. Rodman, and I. Spitkovsky, *The numerical range of* 3×3 *matrices*, Linear Algebra Appl. **252** (1997), 115–139.
[8] R. Kippenhahn, *Über den Wertevorrat einer Matrix*, Math. Nachr. **6** (1951), 193–228.
[9] ———, *On the numerical range of a matrix*, Linear Multilinear Algebra **56** (2008), no. 1-2, 185–225, Translated from the German by Paul F. Zachlin and Michiel E. Hochstenbach.
[10] K. L. Petersen and C. D. Sinclair, *Conjugate reciprocal polynomials with all roots on the unit circle*, Canad. J. Math. **60** (2008), no. 5, 1149–1167.
[11] D. E. V. Rose and I. M. Spitkovsky, *On the stabilization of the Aluthge sequence*, Int. J. Inf. Syst. Sci. **4** (2008), no. 1, 178–189.
[12] C. D. Sinclair and J. D. Vaaler, *Self-inversive polynomials with all zeros on the unit circle*, Number theory and polynomials, London Math. Soc. Lecture Note Ser., vol. 352, Cambridge Univ. Press, Cambridge, 2008, pp. 312–321.

Department of Mathematics, College of William and Mary, Williamsburg, VA 23187, USA
E-mail address: tjmoran@email.wm.edu (TM)