Matrix perturbations: bounding and computing eigenvalues

Reis da Silva, R.J.

[Link to publication](#)

Chapter 3

Normal matrices

When A and E are not Hermitian and even when the analysis is restricted to normal matrices the study of perturbations is difficult. One important reason is that for normal matrices many fundamental questions remain unanswered. For instance, given A and E normal matrices, it is still not fully understood under which conditions is $A + E$ also normal.

In this chapter we deal with normality preserving perturbations and augmentations of normal matrices and their consequences to the eigenvalues. We revisit the normality preserving augmentation of normal matrices studied by Ikramov and Elsner [45] in 1998 and complement their results by showing how the eigenvalues of the original matrix are perturbed by the augmentation. Moreover, we construct all augmentations that result in normal matrices with eigenvalues on a quadratic curve in the complex plane, using the stratification of normal matrices presented by Huhtanen [42] in 2001. To make this construction feasible, but also for its own sake, we study normality preserving normal perturbations of normal matrices. For 2×2 and for rank-one matrices, the analysis is complete. For higher rank, all essentially Hermitian normality perturbations are described. In all cases, the effect of the perturbation on the eigenvalues of the original matrix is given.

3.1   Introduction

A complex number $z \in \mathbb{C}$ is often split up as $z = \Re(z) + \Im(z)$, where $2\Re(z) = z + \overline{z}$ and $2\Im(z) = z - \overline{z}$. This interprets the complex plane as a two-dimensional real vector space with as basis the numbers 1 and i. In this chapter, it will be convenient to decompose z differently. For this, we introduce a family of decompositions parametrized in $\theta \subset \mathbb{T} \subset \mathbb{C}$, the circle group of unimodular numbers. For a given $\theta \in \mathbb{T}$ we let

$$z = \Theta(z) + \Theta^\perp(z), \quad \text{where} \quad \Theta(z) = \Re(\overline{\theta}z)\theta \quad \text{and} \quad \Theta^\perp(z) = \Im(\overline{\theta}z)\theta. \quad (3.1)$$

Moreover, apart from the standard Toeplitz or Cartesian decomposition of a square matrix, A, into its Hermitian and skew-Hermitian parts,

$$A = \mathcal{H}(A) + \mathcal{S}(A), \quad \text{where} \quad \mathcal{H}(A) = \frac{1}{2}(A + A^*) \quad \text{and} \quad \mathcal{S}(A) = \frac{1}{2}(A - A^*),$$

in accordance with (3.1) we consider the family of matrix decompositions

$$A = \Theta(A) + \Theta^\perp(A), \quad \text{where} \quad \Theta(A) = \mathcal{H}(\overline{\theta}A)\theta \quad \text{and} \quad \Theta^\perp(A) = \mathcal{S}(\overline{\theta}A)\theta. \quad (3.2)$$

We call this decomposition the θ-*Toeplitz* decomposition of A. The matrix $\Theta(A)$ is then the θ-*Hermitian part* of A and $\Theta^\perp(A)$ its θ-*skew-Hermitian part* and subsequently, A is θ-*Hermitian* if $A = \Theta(A)$ and θ-*skew-Hermitian* if $A = \Theta^\perp(A)$. Note that θ-skew-Hermitian matrices are $i\theta$-Hermitian. We now generalize a well-known result to the θ-Toeplitz decomposition of normal matrices.

Lemma 3.1.1. *Let $\theta \in \mathbb{T}$ be arbitrarily given. For any normal matrix A we have that*

$$Av = \lambda v \quad \Leftrightarrow \quad \Theta(A)v = \Theta(\lambda)v \quad and \quad \Theta^\perp(A)v = \Theta^\perp(\lambda)v,$$

where Θ and $\Theta^\perp$ and their relation to θ are defined in (3.2).

Proof. Let $Au = \lambda u$ for some $u \neq \vec{0}$. Since A is normal, there exists a unitary matrix U with u as first column and $U^*AU = \Lambda$ diagonal. But then $U^*A^*U = \Lambda^*$, showing that $A^*u = \overline{\lambda}u$. Since the argument can be repeated with A^* instead of A, that yields that A and A^* have the same eigenvectors and that the corresponding

eigenvalues are each other's complex conjugates. Therefore,

$$2\Theta(A)v = (\bar{\theta}A + \theta A^*)\theta v = (\bar{\theta}\lambda + \theta\bar{\lambda})\theta v = 2\Theta(\lambda)v$$

and, similarly, also $\Theta^\perp(A)v = \Theta^\perp(\lambda)v$. The reverse implication is trivial. $\square$

Corollary 3.1.2. *Let λ_1 be an eigenvalue of a normal matrix A. For given $\theta \in \mathbb{T}$, consider the line $\ell \subset \mathbb{C}$ through λ_1 defined by*

$$\ell : \{\lambda_1 + \rho\theta \mid \rho \in \mathbb{R}\}.$$

Assume that $\lambda_1, \ldots, \lambda_p$ are all eigenvalues of A that lie on ℓ. Then the eigenspace $\mathcal{U}$ of the eigenvalue $\Theta^\perp(\lambda_1)$ of $\Theta^\perp(A)$ equals the invariant subspace of A spanned by $u_1, \ldots, u_p$. Restricted to $\mathcal{U}$, the matrix $A - \lambda_1 I_p$ is θ-Hermitian.

Proof. Let λ_a and λ_b be eigenvalues of A, then we have that

$$\Theta^\perp(\lambda_a) = \Theta^\perp(\lambda_b) \iff \Im(\bar{\theta}\lambda_a) = \Im(\bar{\theta}\lambda_b) \iff \bar{\theta}(\lambda_a - \lambda_b) = \rho \in \mathbb{R} \iff \lambda_a - \lambda_b = \theta\rho.$$

Thus, $\Theta^\perp(A)u_j = \Theta^\perp(\lambda_1)u_j$ for all $j \in \{1, \ldots, p\}$, and conversely, if we have $\Theta^\perp(A)u = \Theta^\perp(\lambda_1)u$ then u is a linear combination of $u_1, \ldots, u_p$. Writing U_p for the matrix with columns $u_1, \ldots, u_p$ we moreover find that

$$AU_p = U_p\Lambda_p \quad \text{with} \quad \Lambda_p = \theta R + \lambda_1 I_p$$

where Λ_p is the $p \times p$ diagonal matrix whose eigenvalues are $\lambda_1, \ldots, \lambda_p$ and R is real diagonal. Thus, $\Theta^\perp(U_p^* AU_p - \lambda_1 I_p) = O_p$ proving the last statement. $\square$

3.1.1 Eigenvalues on polynomial curves

If A is θ-Hermitian then A is normal. Moreover, by the spectral theorem for Hermitian matrices, all eigenvalues of A lie on the line $\ell : \{\rho\theta \mid \rho \in \mathbb{R}\}$. In the literature, for instance [7, 23, 27], the matrix A is called *essentially Hermitian* if there exists an $\alpha \in \mathbb{C}$ such that $A - \alpha I$ is θ-Hermitian for some $\theta \in \mathbb{T}$. Clearly, the spectrum of an essentially Hermitian matrix lies on an affine line shifted over

$\alpha \in \mathbb{C}$. Conversely, if a normal matrix has all its eigenvalues on a line $\ell \subset \mathbb{C}$, it is essentially Hermitian. This includes *all* normal 2×2 matrices and *all* normal rank-one perturbations of αI for $\alpha \in \mathbb{C}$. Larger and higher rank normal matrices have their eigenvalues on a polynomial curve $\mathcal{C} \subset \mathbb{C}$ of higher degree.

Polynomial curves of degree $k \geq 2$

Each matrix $A \in \mathbb{C}^{n \times n}$ has its eigenvalues on a polynomial curve $\mathcal{C} \subset \mathbb{C}$ of degree $k \leq n - 1$. This can be explained as follows (see also [42]). First, fix $\theta \in \mathbb{T}$ such that for each pair λ_p, λ_q of eigenvalues of A

$$\lambda_p \neq \lambda_q \Rightarrow \Theta(\lambda_p) \neq \Theta(\lambda_q). \tag{3.3}$$

Note that there exists at most $n(n-1)$ values of θ for which this cannot be realized. These values correspond to the at most $\frac{1}{2}n(n-1)$ lines going through each pair of distinct eigenvalues of A. Once 3.3 is satisfied, the points

$$\left(\bar{\theta}\Theta(\lambda_1) \, , \, \bar{i\theta}\Theta^{\perp}(\lambda_1) \right), \ldots, \left(\bar{\theta}\Theta(\lambda_n) \, , \, \bar{i\theta}\Theta^{\perp}(\lambda_n) \right) \in \mathbb{R} \times \mathbb{R}$$

form a feasible set of points in $\mathbb{R} \times \mathbb{R}$ through which a Lagrange interpolation polynomial $\wp \in \mathcal{P}^{n-1}(\mathbb{R})$ can be constructed that satisfies

$$\Theta^{\perp}(\lambda_j) = i \cdot \theta \wp \left(\bar{\theta}\Theta(\lambda_j) \right) \quad \text{for all } j \in \{1, \ldots, n\}.$$

If A is normal, however, we may draw additional consequences. We summarize these in the following lemma.

Lemma 3.1.3. *Let A be normal and let θ be such that (3.3) is satisfied. Then there is a $\wp \in \mathcal{P}^{n-1}(\mathbb{R})$ such that $\Theta^{\perp}(A) = i \cdot \theta \wp \left(\bar{\theta}\Theta(A) \right)$, and thus the θ-Toeplitz decomposition of A can be written as*

$$A = \Theta(A) + i \cdot \theta \wp \left(\bar{\theta}\Theta(A) \right),$$

or, in terms of $\bar{\theta}A$ and its classical Toeplitz decomposition,

$$\bar{\theta}A = \mathcal{H}(\bar{\theta}A) + i\wp\left(\mathcal{H}(\bar{\theta}A)\right).$$

Moreover, the eigenvalues of A lie on the image $\mathcal{C}$ of the function

$$c : \mathbb{R} \to \mathbb{C} : \quad \rho \mapsto \theta\rho + i \cdot \theta\wp(\rho).$$

Proof. The statement is proved by applying the spectral theorem for normal and Hermitian matrices. $\qquad\square$

Remark 3.1.4. *Note that A is essentially Hermitian if and only if there exists a $\theta \in \mathbb{T}$ such that the interpolating polynomial $\wp \in \mathcal{P}^1(\mathbb{R})$. In fact, $\wp$ may even be in $\mathcal{P}^0(\mathbb{R})$.*

Remark 3.1.5. *If for some normal matrix A the degree of the interpolation polynomial equals, say, two for some value of θ, it may well be of degree $n-1$ for almost all other values of θ, since in that case the eigenvalues lie on a rotated parabola (see Figure 3.1).*

Remark 3.1.6. *There does not seem to be an easy way to determine θ for which the polynomial degree is minimal, although for given θ, the polynomial can be computed in a finite number of arithmetic operations without knowing the eigenvalues. This is explained in the following section.*

Computing the polynomial $\wp$ for given $\theta \in \mathbb{T}$

For almost any fixed value of θ, the interpolation polynomial $\wp$ belonging to a normal matrix A can be computed, without knowing the eigenvalues of A, in a finite number of arithmetic operations. This provides us with a curve $\mathcal{C} \subset \mathbb{C}$ on which all eigenvalues of A lie. Indeed, if the degree of $\wp$ equals k then $A - \Theta(A)$ is a linear combination of

$$i\theta I , \; i\theta\Theta(A) , \; i\left(\theta\Theta(A)\right)^2 , \; \ldots, \; i\left(\theta\Theta(A)\right)^k .$$
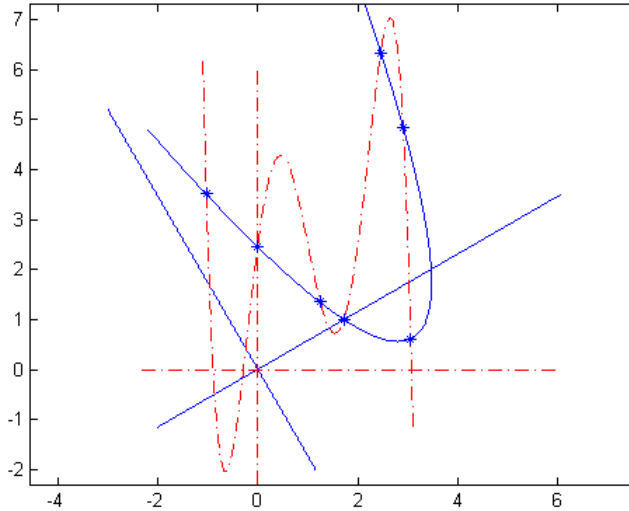
Figure 3.1: Interpolating polynomials of different degree for different values of θ. The seven asterisks represent the eigenvalues.

Making the combination explicit is equivalent to finding the coefficients of $\wp$. To obtain $\wp$ in practice, notice that for any $v \in \mathbb{C}^n$,

$$\wp\left(\overline{\theta}\Theta(A)\right)v \in \mathcal{K}^k\left(\overline{\theta}\Theta(A), v\right) = \mathrm{span}\{v, \, \overline{\theta}\Theta(A)v, \ldots, \, \left(\overline{\theta}\Theta(A)\right)^k v\},$$

the Krylov subspace of dimension k generated by the matrix $\overline{\theta}\Theta(A)$ and the vector v. Since $\overline{\theta}\Theta(A)$ is Hermitian, an orthonormal basis for $\mathcal{K}^k\left(\overline{\theta}\Theta(A), v\right)$ can be constructed using a three-term recursion, and solving the linear system can be done cheaply. These and other considerations led Huhtanen to the development of efficient structure preserving eigensolvers [42] and linear systems [43] for problems involving normal matrices.

3.1.2   Commuting normal matrices

We begin by recalling a well-known result [39, §1.3] on commuting normal matrices to which we shall refer often. Moreover, because we will also need to draw conclu-

sions about commuting normal matrices with distinct or with multiple eigenvalues, we opt to give its complete proof.

Lemma 3.1.7. *Normal matrices $A \in \mathbb{C}^{n \times n}$ and $E \in \mathbb{C}^{n \times n}$ commute if and only if they are simultaneously unitarily diagonalizable.*

Proof. If both $\Lambda = W^* A W$ and $\Delta = W^* E W$ are diagonal for a unitary matrix W, then clearly $[A, E] = W[\Lambda, \Delta]W^* = O_n$. Conversely, assume that $[A, E] = O_n$. Let U be a unitary matrix such that $\Delta = U^* E U$ is diagonal with multiple eigenvalues being neighbors on the diagonal of Δ. Thus

$$\Delta = \begin{bmatrix} \delta_1 I_{m_1} & & \\ & \ddots & \\ & & \delta_\ell I_{m_\ell} \end{bmatrix},$$

where m_j denotes the multiplicity of δ_j. Let $S = U^* A U$ and write s_{pq} for its entries. Then equating the entries of $S\Delta$ and ΔS in view of the relation

$$[S, \Delta] = U^*[A, E]U = O_n,$$

shows that $s_{pq} = 0$ whenever $\delta_p \neq \delta_q$. Thus, S is block diagonal with respective blocks $S_1, \ldots, S_\ell$ of sizes $m_1, \ldots, m_\ell$. For each $j \in \{1, \ldots, \ell\}$, S_j is normal. Let S_j be such that $S_j = Q_j \Lambda_j Q_j^*$ for some unitary Q_j and diagonal Λ_j. Now, with

$$Q = \begin{bmatrix} Q_1 & & \\ & \ddots & \\ & & Q_\ell \end{bmatrix} \quad \text{and} \quad \Lambda = \begin{bmatrix} \Lambda_1 & & \\ & \ddots & \\ & & \Lambda_\ell \end{bmatrix} \quad \text{and} \quad W = UQ$$

we find that $\Delta = Q^* \Delta Q = W^* E W$ and $\Lambda = Q^* S Q = W^* A W$. $\qquad\square$

Remark 3.1.8. *In case all eigenvalues of E are distinct, then the matrix S in the proof is itself diagonal and the proof is complete. In case E has an eigenvalue, say δ_1, of multiplicity $m_1 > 1$, then there is freedom in the choice of the first m_1 columns $u_1, \ldots, u_{m_1}$ of U that correspond to δ_1. Even though each choice diagonalizes E, not each choice diagonalizes A as well. This is expressed as S*

having a diagonal block S_1 of size m_1. Writing U_1 for the matrix with columns $u_1, \ldots, u_{m_1}$ we have that

$$AU_1 = U_1 S_1,$$

hence the column span of U_1 is an invariant subspace $\mathcal{V}$ of A. The matrix Q_1 determines, through the transformation $W_1 = U_1 Q_1$, an orthonormal basis for $\mathcal{V}$ of eigenvectors of A. If S_1 has multiple eigenvalues, again there may be much freedom in the choice of Q_1.

3.2   Normality preserving augmentation

In this section we revisit, from an alternative point of view, a problem studied by Ikramov and Elsner in [45]. It concerns the augmentation by a number m of rows and columns of a normal matrix in such a way that normality is preserved. Our analysis differs from the one in [45], and we add details on the eigendata of A_+ in terms of those of A.

Normality preserving augmentation.   *Let $A \in \mathbb{C}^{n \times n}$ be normal. Characterize all $n \times m$ matrices $V, W \in \mathbb{C}^{n \times m}$, and all $\Gamma \in \mathbb{C}^{m \times m}$ such that*

$$A_+ = \begin{bmatrix} A & V \\ W^* & \Gamma \end{bmatrix}, \tag{3.4}$$

is normal, too. In other words, characterize all normality preserving augmentations of A.

Remark 3.2.1. *Note that Hermicity, θ-Hermicity and essentially Hermicity preserving augmentation problems are all trivial, because each of these properties is inherited by principal submatrices. For unitary matrices this does not hold. However, if A and A_+ are both unitary, their rows and columns all have length one. Thus $V = W = O$ and Γ is unitary. This solves the unitarity preserving augmentation problem.*

3.2.1 Normality preserving augmentation for $m = 1$

First consider the case $m = 1$, and write v, w and γ instead of V, W and Γ. It is easily verified that $[A_+, A_+^*] = O_n$ if and only if

$$ww^* = vv^*, \qquad w^*w = v^*v \qquad \text{and} \qquad A^*v + \gamma w = Aw + \overline{\gamma}v. \qquad (3.5)$$

The two leftmost relations hold if and only if $v = \phi w$ for some $\phi \in \mathbb{T}$. The right most relation may add further restrictions on v, w and ϕ. Before studying these, however, note that ϕ is the square of a unique $\theta \in \mathbb{T}_U \subset \mathbb{T}$, where

$$\mathbb{T}_U = \{\tau \in \mathbb{T} \mid \arg(\tau) \in [0, \pi)\}.$$

This yields a reformulation of $v = \phi w$ that better reveals the underlying structure,

$$v = \phi w \quad \Leftrightarrow \quad v = \theta^2 w \quad \Leftrightarrow \quad \overline{\theta}v = \theta w = u \quad \Leftrightarrow \quad v = \theta u \text{ and } w = \overline{\theta}u$$

for some $u \in \mathbb{C}^n$. Further restrictions on the vector u and the scalar $\theta \in \mathbb{T}_U$ follow from substituting $v = \theta u$ and $w = \overline{\theta}u$ into the rightmost equation in (3.5). After some rearrangements we obtain the condition

$$\Theta^\perp(A)u = \Theta^\perp(\gamma)u.$$

Because the eigenpairs of $\Theta^\perp(A)$ were already characterized in Corollary 3.1.2, we have solved the augmentation problem for $m = 1$. The theorem below summarizes this solution, constructively, in terms of the eigendata of A. Note that this result was proved already in [45], though in a different manner.

Theorem 3.2.2. *Let A be normal. Let $\gamma \in \mathbb{C}$, and let $\ell : \{\gamma + \rho\theta \mid \rho \in \mathbb{R}\}$ be a line in $\mathbb{C}$ through γ with slope $\theta \in \mathbb{T}_U$. Moreover, let $\lambda_1, \ldots, \lambda_p$ be the eigenvalues of A that lie on ℓ. Then, the matrix*

$$A_+ = \begin{bmatrix} A & v \\ w^* & \gamma \end{bmatrix}$$

is normal if $v = \theta u$ and $w = \bar{\theta} u$, where u is a linear combination of eigenvectors corresponding to $\lambda_1, \ldots, \lambda_p$, with the convention that $u = \vec{0}$ if $p = 0$. Conversely, if A_+ is normal then $v = \theta u$ and $w = \bar{\theta} u$, for some θ and the vector u is a linear combination of eigenvectors of A whose corresponding eigenvalues all lie on a line $\ell : \{ \gamma + \rho \theta \mid \rho \in \mathbb{R} \}$ for some fixed $\gamma \in \mathbb{C}$.

Proof. Corollary 3.1.2 shows the relation between the eigendata of $\Theta^\perp(A)$ and A, and together with the derivation in this Section this proves the statement.   □

3.2.2   Eigenvalues of the augmented matrix

We now augment the analysis in [45] with a study of the eigenvalues of A_+ in relation to those of A. Let $\Lambda_p \in \mathbb{C}^{p \times p}$ be the diagonal matrix whose eigenvalues are the p eigenvalues of A that lie on $\ell : \mathbb{R} \to \mathbb{C} : \quad \gamma + \rho \theta$. Then, as already mentioned in the proof of Corollary 3.1.2,

$$\Lambda_p = \theta R + \gamma I_p$$

for some real diagonal matrix R. Let $U \in \mathbb{C}^{n \times n}$ be any unitary matrix whose last p columns are eigenvectors of A belonging to $\lambda_1, \ldots, \lambda_p$ and let U_p contain those last p columns of U. Then, assuming that A and A_+ are normal, Theorem 3.2.2 shows, with $r = U_p^* u$, that

$$\begin{bmatrix} U & \\ & 1 \end{bmatrix}^* \begin{bmatrix} A & v \\ w^* & \gamma \end{bmatrix} \begin{bmatrix} U & \\ & 1 \end{bmatrix} = \begin{bmatrix} B & & \\ & \Lambda_p & \theta r \\ & \theta r^* & \gamma \end{bmatrix}. \tag{3.6}$$

Moreover,

$$\begin{bmatrix} \Lambda_p & \theta r \\ \theta r^* & \gamma \end{bmatrix} = \theta R_+ + \gamma I_{p+1}, \quad \text{where} \quad R_+ = \begin{bmatrix} R & r \\ r^* & 0 \end{bmatrix}. \tag{3.7}$$

The above observations reveal some additional features of the solution of the augmentation problem, that we formulate as a new theorem.

Theorem 3.2.3. *The only normality preserving 1-augmentations of A are the ones that, on an orthonormal basis of eigenvectors of A, augment a $p \times p$ essentially Hermitian submatrix of A. Hence, $n - p$ eigenvalues of A are also eigenvalues of A_+. The remaining $p + 1$ eigenvalues of A_+ lie on the same line as, and are interlaced by, the remaining p eigenvalues of A.*

Proof. The block form in Equation(3.6) shows that the eigenvalues of B are eigenvalues of both A and A_+, whereas (3.7) shows that to locate the remaining $p + 1$ eigenvalues of A_+, one only needs to observe that, by Cauchy Interlace Theorem (Theorem 1.1.10) the (real) eigenvalues of R interlace those of R_+. □

Remark 3.2.4. *Note that the case $p = 0$, covered by Theorem 3.2.2, is also included in the above analysis if one is willing to interpret on the same line as the remaining p eigenvalues of A as any line in $\mathbb{C}$. This just reflects that the additional eigenvalue, $\gamma \in \mathbb{C}$, of A_+ can lie anywhere.*

For an illustration of the constructions described in Theorems 3.2.2 and 3.2.3 we refer the reader to Section 3.5.1. There, we augment a given 3×3 matrix A in two different ways and compute the eigenvalues of the augmented matrix, A_+.

3.2.3   Normal matrices with normal principal submatrices

By applying the procedure for $m = 1$ several times consecutively, we may also construct m-augmentations with $m > 1$. In particular, *all* normal matrices having the property that all their leading principal submatrices are normal can be constructed. Since, generally, normal matrices do not have normal principal submatrices, this shows that the m-augmentation for $m > 1$ has not yet been completely solved. In Section 3.4 we investigate the principal submatrices of normal matrices from the point of view of Section 3.1.1. This study will also give more insight into the m-augmentation for $m = 2$. In [45], this case already proved to be quite complicated. In particular, we give a procedure to augment A that does not reduce to m-fold application of the 1-augmentation. Before that, we investigate normality preserving normal perturbations. Apart from being of interest on its own, we shall make use of some of the conclusions obtained here in Section 3.4.

3.3   Normality preserving normal perturbations

In this section we consider a question related to the augmentation problem, and we study it using the same techniques as the ones from the previous sections. In particular, we investigate normality preserving θ-Hermitian perturbations. These type of perturbations play a role also in the augmentation problem of Section 3.4.1.

Normality preserving normal perturbation.   *Let $A \in \mathbb{C}^{n \times n}$ be normal. Characterize all normal E such that $A^{+} = A + E$ is normal. In other words, characterize the normality preserving normal perturbations E of A.*

Remark 3.3.1. *Recall that any matrix A may be written as the sum of two normal matrices, for instance, as $A = \Theta(A) + \Theta^{\perp}(A)$, where both $\Theta(A)$ and $\Theta^{\perp}(A)$ are normal. This illustrates why the problem above is non-trivial: the sum of normal matrices can be, literally,* any *matrix.*

We begin by formulating a multi-functional lemma summarizing the technicalities of writing out commutators of linear combinations of matrices.

Lemma 3.3.2. *Let $A, E \in \mathbb{C}^{n \times n}$ and $\gamma, \mu \in \mathbb{C}$. Then, with $\theta = \overline{\gamma}\mu/|\gamma\mu|$,*

$$[\gamma A + \mu E, (\gamma A + \mu E)^{*}] = |\gamma|^{2} [A, A^{*}] + 2\gamma\overline{\mu}\,\Theta([A, E]) + |\mu|^{2} [E, E^{*}], \quad (3.8)$$

and

$$\Theta([A, E^{*}]) = [\Theta(A), \Theta^{\perp}(E^{*})] + [\Theta^{\perp}(A), \Theta(E^{*})]. \quad (3.9)$$

Therefore, if A and E are normal, then $\gamma A + \mu E$ is normal if and only if

$$\Theta([A, E^{*}]) = O_{n}, \quad (3.10)$$

or, in other words, if and only if $[A, E^{*}]$ is $\overline{\gamma}\mu$-skew-Hermitian.

Proof. The statements are obtained from straight-forward manipulations with the commutator. $\square$

Corollary 3.3.3. *Let $A, E \in \mathbb{C}^{n \times n}$ be normal. Then $A + E$ is normal if and only if $\gamma A + \mu E + \alpha I$ is normal for all $\gamma, \mu, \alpha \in \mathbb{C}$ with the restriction that $\gamma\overline{\mu} \in \mathbb{R}$.*

Corollary 3.3.4. *If $A, E \in \mathbb{C}^{n \times n}$ are normal then $[A, E^*] = O_n \Leftrightarrow [A, E] = O_n$. Thus, if either term vanishes, both $A + E^*$ and $A + E$ are normal.*

Proof. As was shown in the proof of Lemma 3.1.1, E and E^* have the same eigenvectors. Thus, A and E^* are simultaneously unitarily diagonalizable if and only if A and E are. Lemma 3.1.7 now proves that $[A, E^*] = O_n \Leftrightarrow [A, E] = O_n$, and Lemma 3.3.2 proves the conclusion. $\qquad\square$

Corollary 3.3.5. *The matrix $\gamma A + \mu E$ with $\gamma, \mu \in \mathbb{C}$ and A and E Hermitian is normal if and only if $\overline{\gamma}\mu \in \mathbb{R}$ or $[A, E] = O_n$.*

Proof. The commutator of Hermitian matrices is always skew-Hermitian. Thus, for $[A, E]$ to be θ-Hermitian in Equation (3.10), $\overline{\gamma}\mu$ must be real, or $[A, E]$ should vanish. $\qquad\square$

3.3.1 Normality preserving normal rank-one perturbations

This section aims to show the similarities between the normality preserving normal rank-one perturbation problem and the m-augmentation problem of Section 3.2.1. Indeed, let $E = vw^*$ with $v, w \in \mathbb{C}^n$. Then E is normal if and only if

$$\|w\|_2^2 vv^* = \|v\|_2^2 ww^*,$$

and thus if and only if $v = zw$ for some $z \in \mathbb{C}$. Write $z = \theta\rho$ with $\theta \in \mathbb{T}$ and $0 \le \rho \in \mathbb{R}$. This shows that a rank-one matrix E is normal if and only if E is θ-Hermitian,

$$E = \theta uu^*, \qquad \theta \in \mathbb{T}.$$

With $A \in \mathbb{C}^{n \times n}$ normal, we look for the conditions on $u \in \mathbb{C}^n$ and $\theta \in \mathbb{T}$ such that $A + \theta uu^*$ is normal. Since θuu^* is θ-Hermitian, from Equations (3.8) and (3.9) in Lemma 3.3.2 we obtain

$$[(A + \theta uu^*), (A + \theta uu^*)^*] = 2\overline{\theta}\left[\Theta^{\perp}(A), uu^*\right].$$

Therefore, $A + \theta uu^*$ is normal if and only if $\Theta^{\perp}(A)$ and uu^* commute. According to Lemma 3.1.7, this is true if and only if they are simultaneously unitarily

diagonalizable. For this, it is necessary and sufficient that u be an eigenvector of $\Theta^\perp(A)$. As in Theorem 3.2.2 we formulate this result constructively in terms of the eigendata of A.

Theorem 3.3.6. *Let A be normal and let $\gamma \in \mathbb{C}$ and $\ell : \{\gamma + \rho\theta \mid \rho \in \mathbb{R}\}$ be a line in $\mathbb{C}$ through γ with slope $\theta \in \mathbb{T}$. Let $\lambda_1, \ldots, \lambda_p$ be the eigenvalues of A that lie on ℓ. Then, the matrix*

$$A^+ = A + \theta uu^*$$

is normal if and only if u is a linear combination of eigenvectors corresponding to $\lambda_1, \ldots, \lambda_p$, with the convention that $u = \vec{0}$ if $p = 0$.

Proof. Corollary 3.1.2 shows the relation between the eigendata of $\Theta^\perp(A)$ and A, and together with the derivation above this proves the statement. $\square$

Remark 3.3.7. *An interesting consequence of adding the normality preserving normal rank-one perturbation $E = \theta uu^*$ is that*

$$\Theta^\perp(A + \theta uu^*) = \Theta^\perp(A) + \Theta^\perp(\theta uu^*) = \Theta^\perp(A),$$

because θuu^* is θ-Hermitian. Thus, the conditions under which adding another θ-Hermitian normal rank-one perturbation $F = \theta ww^*$ to $A + E$ lead to a normal $A + E + F$ are identical to the conditions just described for E. We shall get back to this observation in Section 3.3.3.

3.3.2   Eigenvalues of the perturbed matrix

To study the eigenvalues of A^+ in relation to those of A, let $\Lambda_p \in \mathbb{C}^{p \times p}$ be the diagonal matrix whose eigenvalues are the p eigenvalues of A that lie on the line $\ell : \{\gamma + \rho\theta \mid \rho \in \mathbb{R}\}$. Then

$$\Lambda_p = \theta R + \gamma I_p$$

for some real diagonal matrix R. Let $U \in \mathbb{C}^{n \times n}$ be any unitary matrix whose first p columns are eigenvectors of A belonging to $\lambda_1, \ldots, \lambda_p$. Then, assuming that A

and A^+ are normal, Theorem 3.3.6 shows that

$$U^* A^+ U = U^* (A + \theta u u^*) U = \begin{bmatrix} \Lambda_p & \\ & B \end{bmatrix} + \begin{bmatrix} \theta r r^* & \\ & O_{n-p} \end{bmatrix}$$

because u is a linear combination of the first p columns of U. This leads to the following theorem, in which we summarize the above analysis.

Theorem 3.3.8. *The only normality preserving normal rank-one perturbations of A are the ones that, on an orthonormal basis of eigenvectors of A, are θ-Hermitian rank-one perturbations of a $p \times p$, θ-Hermitian submatrix. Hence, $n-p$ eigenvalues of A are also eigenvalues of A^+. The remaining p eigenvalues of A^+ are the eigenvalues of*

$$\Lambda_p + \theta r r^* = \theta (R + r r^*) + \gamma I_p. \tag{3.11}$$

These interlace the p eigenvalues of A on ℓ with the additional $(p+1)$st point $+\infty \theta$.

Proof. The eigenvalues of A^+ are the eigenvalues of B together with the eigenvalues of the matrix in Equation (3.11). Obviously, all eigenvalues of B are eigenvalues of A as well. Since $r r^*$ is a positive semi-definite rank-one perturbation of R, the eigenvalues $\rho_1 \leq \ldots \leq \rho_p$ of $R + r r^*$ and the eigenvalues $r_1 \leq \ldots \leq r_p$ of R satisfy

$$r_1 \leq \rho_1 \leq r_2 \leq, \ldots, \leq \rho_{p-1} \leq r_p \leq \rho_p$$

as a result of Weyl's Theorem (Theorem 1.1.7). Multiplying by θ and shifting over γ yields the proof. $\qquad \square$

Remark 3.3.9. *Note that if $p = 1$, only one eigenvalue is perturbed, and we have $\rho_1 = r_1 + \|r\|_2^2$. In terms of the original perturbation $E = \theta u u^*$ this becomes $\tilde{\lambda} = \lambda + \|u\|_2^2$, where λ is the eigenvalue of A belonging to the eigenvector u (see also Corollary 1.1.8).*

As a consequence of the following theorem, it is possible to indicate where the eigenvalues of the family of matrices $A + tE$ are located. This can only be done for *normal* normality preserving perturbations.

Theorem 3.3.10. *Let $A, B \in \mathbb{C}^{n \times n}$ be normal. Consider the line ℓ through A and B,*

$$\ell : \mathbb{R} \to \mathbb{C}^{n \times n} : \quad t \mapsto tA + (1 - t)B.$$

If $E = B - A$ is normal, all matrices on ℓ are normal; if E is not normal, A and B are the only normal matrices on ℓ.

Proof. Observe that $\ell(t) = A + (1 - t)E$. If E is normal, then Corollary 3.3.3 shows that all matrices $\gamma A + \mu E$ with $\gamma \bar{\mu} \in \mathbb{R}$ are normal, which includes the line ℓ. Assume now that E is not normal. Because A and $B = A + E$ both are normal, Equation (3.3.2) in Lemma 3.8 gives that

$$(1 - t) \left(2 \mathcal{H} \left([A, E^*] \right) + (1 - t)[E, E^*] \right) = O_n \quad \text{with} \quad [E, E^*] \neq O_n.$$

The solution $t = 1$ confirms the normality of A, and the linear matrix equation

$$2 \mathcal{H} \left([A, E^*] \right) + (1 - t)[E, E^*] = O_n \quad \text{with} \quad [E, E^*] \neq O_n$$

allows at most one solution in t which, by assumption, is $t = 0$. $\qquad\square$

Thus, any line in $\mathbb{C}^{n \times n}$ parametrized by a real variable that does not lie entirely in the set of normal matrices, contains at most two normal matrices.

Remark 3.3.11. *Let A be normal. Lemma 3.1.1 shows that if E is such that $A + E$ is normal, then*

$$\sigma(A + E) \subset \sigma \left(\Theta(A) + \Theta(E) \right) \times \sigma \left(\Theta^\perp(A) + \Theta^\perp(E) \right).$$

and perturbation theory for θ-Hermitian matrices can be used to derive statements about the eigenvalues of $A + E$. According to Theorem 3.3.10, if E itself is normal too, this relation is valid continuously in t along the line $A + tE$:

$$\sigma(A + tE) \subset \sigma \left(\Theta(A) + t\Theta(E) \right) \times \sigma \left(\Theta^\perp(A) + t\Theta^\perp(E) \right).$$

For non-normal E this is, generally, not true, as is illustrated in Section 3.5.2.

Corollary 3.3.12. *Let A be normal. As a result of Theorem 3.3.10 and Remark 3.3.11, the perturbed eigenvalues of*

$$t \mapsto A + t\theta uu^*, \quad 0 \le t \le 1$$

seen as functions of t, form line segments that all lie on the same line with slope θ.

3.3.3 θ-Hermitian rank-k perturbations of normal matrices

Consider for given k with $1 \le k \le n$ the θ-Hermitian rank-k matrix

$$E = \theta H, \quad \text{where} \quad \theta \in \mathbb{T}_U \quad \text{and} \quad H = H^*. \tag{3.12}$$

Let A be normal. Then, since $\Theta^\perp(E^*) = O_n$, Lemma 3.3.2 shows that $A + E$ is normal if and only if

$$[\Theta^\perp(A), H] = O_n. \tag{3.13}$$

By Lemma 3.1.7 this is equivalent to $\Theta^\perp(A)$ and H being simultaneously diagonalizable by a unitary transformation U. Thus, H needs to be of the form

$$H = U\Delta U^* \tag{3.14}$$

where $\Delta \in \mathbb{R}^{k \times k}$ is diagonal with diagonal entries $\delta_1, \ldots, \delta_k$ and the columns $u_1, \ldots, u_k$ of U are orthonormal eigenvectors of $\Theta^\perp(A)$. But then, writing

$$E = E_1 + \ldots + E_k, \quad \text{where for all} \quad j \in \{1, \ldots, k\}, \quad E_j = \theta \delta_j u_j u_j^*,$$

the observation in Remark 3.3.7 reveals that perturbing A by E is equivalent to perturbing A consecutively by the rank-one matrices $E_1, \ldots, E_k$. The above analysis is summarized in the following theorem. An illustration of this theorem is provided in Section 3.5.2.

Theorem 3.3.13. *Let $E = \theta H$ be a θ-Hermitian rank-k perturbation of a normal*

matrix A. Then E is normality preserving if and only if E can be decomposed as

$$E = E_1 + \ldots + E_k,$$

where $E_1, \ldots, E_k$ are all normality preserving θ-Hermitian rank-one perturbations of A. In fact, for each permutation σ of $\{1, \ldots, k\}$ and each $m \in \{1, \ldots, k\}$, the partial sum

$$A + \sum_{j=1}^{m} E_{\sigma(j)}$$

is normal, too.

Remark 3.3.14. *In accordance with Remark 3.1.8, if the matrix Δ in Equation (3.14) has multiple eigenvalues, there exist non-diagonal unitary matrices Q such that $\Delta = Q\Delta Q^*$. As a result, H can be written as $Z\Delta Z^*$ where the orthonormal columns of Z span an invariant subspace of $\Theta^{\perp}(A)$. This implicitly writes the perturbation θH as the sum of rank-one normal perturbations that do not necessarily preserve normality. This aspect is also illustrated in Section 3.5.2.*

Theorem 3.3.15. *The only normality preserving θ-Hermitian rank-k perturbations of A are the ones that, on an orthonormal basis of eigenvectors of A, are θ-Hermitian perturbations of θ-Hermitian submatrices of size $s_1 \times s_1, \ldots, s_m \times s_m$ of ranks $k_1, \ldots, k_m$, where $k_1 + \ldots + k_m = k$. As a result of this perturbation, at most $s_1 + \ldots + s_m$ eigenvalues of A are perturbed, which are located on at most m distinct parallel lines $\ell_1, \ldots, \ell_m$, defined by $\theta \in \mathbb{T}$ and $\gamma_1, \ldots, \gamma_m \in \mathbb{C}$ as*

$$\ell_j : \{\gamma_j + \theta\rho \mid \rho \in \mathbb{R}\}.$$

Moreover, the eigenvalues of $A + tE$ with $t \in [0, 1]$ connect the eigenvalues of A with those of A^+ by line segments that lie on $\ell_1, \ldots, \ell_m$.

Proof. Write $\theta H = \theta(\delta_1 u_1 u_1^* + \ldots + \delta_k u_k u_k^*)$, where $u_1, \ldots, u_k$ are eigenvectors $\Theta^{\perp}(A)$, and repeatedly apply Theorem 3.3.8. The statement about the eigenvalues of $A + tE$ follows from Corollary 3.3.12. $\qquad\square$

For a qualitative illustration of the effect on the eigenvalues due to a rank-one θ-Hermitian perturbation and a rank-k θ-Hermitian perturbation of a normal

matrix, see Figure 3.2. The asterisks in the pictures are eigenvalues of A, the circles represent different choices for γ, and the boxes are the perturbed eigenvalues.
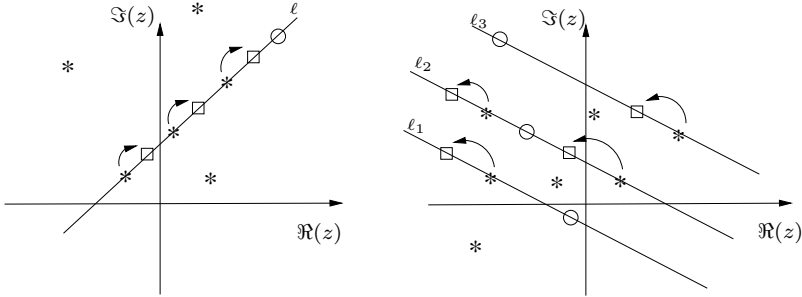

Figure 3.2: Eigenvalue perturbation by a rank-one matrix (left) and a rank-k matrix (right).

Remark 3.3.16. *If H in (3.12) is semi-definite, then the eigenvalues of $A+tE$ all move in the same direction over the parallel lines $\ell_1, \ldots, \ell_m$ from Theorem 3.3.15.*

3.3.4 Normality preserving normal perturbations

The above analysis of θ-Hermitian normality preserving perturbations also gives sufficient conditions for when normal perturbations E of the form $E = \theta_1 H_1 + \theta_2 H_2$ with H_1, H_2 Hermitian and $\theta_1, \theta_2 \in \mathbb{T}$ are normality preserving in case $\theta_2 \neq \pm\theta_1$. Notice that in order for E to be normal itself, $[H_1, H_2] = O_n$ by Corollary 3.3.5. Obviously, E is, in general, not θ-Hermitian for some value of θ. Nevertheless, the following holds.

Theorem 3.3.17. *Let A be normal, H_1, H_2 Hermitian with $[H_1, H_2] = O_n$, and $\theta_1, \theta_2 \in \mathbb{T}$ with $\theta_1\overline{\theta_2} \notin \mathbb{R}$. Then $E = E_1 + E_2 = \theta_1 H_1 + \theta_2 H_2$ is a normality preserving normal perturbation of A if E_1 and E_2 both are normality preserving perturbations of A.*

Proof. Corollary 3.3.5 covers the normality of E. Furthermore, assuming that $A + E_1$ is normal, Equation (3.13) gives that $A + E_1 + E_2$ is normal if and only if $[\Theta^\perp(A + E1), H_2] = O_n$, where $\Theta^\perp$ refers to the θ_2-skew-Hermitian part. But

then

$$\left[\Theta^{\perp}(A + E_1), H_2\right] = \left[\Theta^{\perp}(A), H_2\right] + \left[\Theta^{\perp}(\theta_1 H_1), H_2\right] = \left[\Theta^{\perp}(A), H_2\right],$$

because $[H_1, H_2] = O_n$, proving the statement. $\qquad\qquad\qquad\qquad\qquad\square$

Remark 3.3.18. *A similar result holds for normal perturbations $E_1 + \ldots + E_k$ where $E_j = \theta_j H_j$ with H_j Hermitian and $\theta_j \in \mathbb{T}$ for each $j \in \{1, \ldots, k\}$. Moreover, each normal perturbation E can be written in this form in several different ways.*

We are now ready to return to the augmentation problem of Section 3.3 and to study augmentations of normal matrices A that result in an augmented matrix A_+ whose eigenvalues all lie on the graph of a quadratic polynomial.

3.4   Further augmentations

We now return to the m-augmentation problem of Section 3.3 and concentrate on the case $m > 1$. If all eigenvalues of A_+ lie on a line, then A_+ is essentially Hermitian, a property that is inherited by principal submatrices. In that case it is clear which matrices A can be augmented into A_+. The next simplest case is the case where all eigenvalues of A_+ lie on a curve $\mathcal{C}$ that is the image of a *quadratic* function in a rotated complex plane.

3.4.1   Matrices A_+ with all eigenvalues on a quadratic curve

Assume that for $\theta \in \mathbb{T}$ there exists a $\wp \in \mathcal{P}^2(\mathbb{R})$ such that

$$A_+ = \Theta(A_+) + i \cdot \theta \wp \left(\bar{\theta} \Theta(A_+)\right), \quad \text{where} \quad A_+ = \begin{bmatrix} A & V \\ W^* & \Gamma \end{bmatrix}. \qquad (3.15)$$

Then, A_+ is normal with all its eigenvalues on the rotated parabola $\mathcal{C} \subset \mathbb{C}$ defined as the image of q, where

$$q : \mathbb{R} \to \mathbb{C} : \quad \rho \mapsto \theta \rho + i \cdot \theta \wp(\rho)$$

and

$$\wp(x) = r_0 + r_1 x + r_2 x^2 \quad \text{with} \quad r_0, r_1, r_2 \in \mathbb{R}.$$

Remark 3.4.1. *Throughout this section we assume, without loss of generality, that $r_2 > 0$. The case $r_2 = 0$, as argued above, is trivial and concerns essentially Hermitian matrices, whereas the case $r_2 < 0$ can be avoided by replacing θ by $-\theta$, which is nothing else than a trivial change of coordinates that transforms the polynomial $\wp$ into $-\wp$.*

The curve $\mathcal{C}$ now divides the complex plane $\mathbb{C}$ in three disjoint parts

$$\mathbb{C} = \mathcal{C}_+ \cup \mathcal{C} \cup \mathcal{C}_- \tag{3.16}$$

where $\mathcal{C}_+$ is the open part of $\mathbb{C}$ that lies on the one side of $\mathcal{C}$ that is convex.

The principal submatrices of A_+ and their eigenvalues

For convenience, write

$$X = \bar{\theta}\Theta(A), \quad M = \bar{\theta}\Theta(\Gamma) \quad \text{and} \quad 2Z = \bar{\theta}V + \theta W,$$

then,

$$\bar{\theta}\Theta(A_+) = \begin{bmatrix} X & Z \\ Z^* & M \end{bmatrix} \text{ and } \left(\bar{\theta}\Theta(A_+)\right)^2 = \begin{bmatrix} X^2 + ZZ^* & XZ + ZM \\ Z^*X + MZ^* & M^2 + Z^*Z \end{bmatrix} \tag{3.17}$$

Thus, explicitly evaluating $\wp$ at A_+ using the block forms in (3.17), and comparing the result with the block form of A_+ displayed in (3.15) yields

$$A = \Theta(A) + i \cdot \theta\wp\left(\bar{\theta}\Theta(A)\right) + i \cdot \theta r_2 ZZ^* \tag{3.18}$$

and

$$\Gamma = \Theta(\Gamma) + i \cdot \theta\wp\left(\bar{\theta}\Theta(\Gamma)\right) + i \cdot \theta r_2 Z^* Z.$$

The results that follow will sometimes be stated for A only, even though similar statements obviously hold for Γ. The first proposition simply translates Equation

(3.18) in words.

Proposition 3.4.2. *The $n \times n$ principal submatrix A of A_+ is a θ-skew-Hermitian rank-k (with $k \leq \min(m,n)$) perturbation of a normal matrix that has all the eigenvalues on $\mathcal{C}$.*

Lemma 3.4.3. *If A in Equation (3.18) is normal then $\sigma(A) \subset \mathcal{C} \cup \mathcal{C}_+$.*

Proof. If A is normal, then $i \cdot \theta r_2 Z Z^*$ is a *normality preserving* θ-skew-Hermitian matrix perturbation of the normal matrix $\Theta(A) + i \cdot \theta \wp \left(\bar{\theta} \Theta(A) \right)$ that has all its eigenvalues on $\mathcal{C}$. By Theorem 3.3.15 each perturbed eigenvalue $\lambda \in \mathcal{C}$ of A moves along a line $\ell : \{ \lambda + i \cdot \theta \rho \mid \rho \in \mathbb{R} \}$. Note that ℓ is vertical in the θ-rotated complex plane. By Remark 3.3.16 and because $Z Z^*$ is positive semi-definite, the direction is the same for each perturbed eigenvalue and is determined by the sign of r_2. In Remark 3.4.1 we assumed that $r_2 > 0$, and thus the direction is directed into $\mathcal{C}_+$ defined in (3.16). $\qquad\square$

Remark 3.4.4. *Note that a* multiple *eigenvalue λ of $\Theta(A) + i \cdot \theta \wp \left(\bar{\theta} \Theta(A) \right)$, located on $\mathcal{C}$, may be perturbed by $i \cdot \theta r_2 Z Z^*$ into several distinct eigenvalues of A. Those will all be located on $\ell : \mathbb{R} \to \mathbb{C} : \quad \rho \mapsto \lambda + i \cdot \theta \rho$ with $\rho > 0$.*

Corollary 3.4.5. *Assume that A in (3.18) is normal. Then $\sigma(A) \subset \mathcal{C}$ if and only if A_+ is block diagonal with blocks A and Γ.*

Proof. If $Z \neq O$ then $\text{trace}(Z Z^*) \neq 0$ and at least one eigenvalue is perturbed. Lemma 3.4.3 shows that a perturbed eigenvalue cannot stay on $\mathcal{C}$ and necessarily moves from $\mathcal{C}$ into $\mathcal{C}_+$. $\qquad\square$

Augmentations with eigenvalues on a quadratic curve

We may now reverse the previous observations. Given A, we choose a parabolic curve $\mathcal{C}$ and construct $Z \in \mathbb{C}^{n \times m}$ such that $i \cdot \theta r_2 Z Z^*$ perturbs the eigenvalues of A onto $\mathcal{C}$. We then use $\mathcal{C}$ to define the corresponding m-augmentation A_+ of A.

Corollary 3.4.6. *Necessary for a normal matrix A to be m-augmentable into a normal matrix A_+ with all eigenvalues on a quadratic curve $\mathcal{C}$ is that $\sigma(A) \subset \mathcal{C} \cup \mathcal{C}_+$.*

Proof. This is just another corollary of Lemma 3.4.3. $\qquad\square$

Clearly, for any given finite set of points in $\mathbb{C}$, there are infinitely many candidates for such quadratic curves $\mathcal{C}$. It is the purpose of this section to show that each of this candidates can be used, and to construct essentially all possible corresponding augmentations A_+.

Theorem 3.4.7. *Let $A \in \mathbb{C}^{n \times n}$ be normal, and let $\theta \in \mathbb{T}$ and $\wp \in \mathcal{P}^2(\mathbb{R})$ be such that $\sigma(A) \subset \mathcal{C} \cup \mathcal{C}_+$ where $\mathcal{C}$ is the graph of*

$$q : \mathbb{R} \to \mathbb{C} : \quad \rho \mapsto \theta\rho + i \cdot \theta\wp(\rho).$$

Then there exist p-augmentations A_+ of A such that

$$\sigma(A_+) \subset \mathcal{C},$$

where p is the number of eigenvalues of A in $\mathcal{C}_+$.

Proof. Write $\Lambda_p \in \mathbb{C}^{p \times p}$ for the diagonal matrix with precisely the eigenvalues $\lambda_1, \ldots, \lambda_p$ of A that do not lie on $\mathcal{C}$ and let $U_p \in \mathbb{C}^{n \times p}$ have corresponding orthonormal eigenvectors $u_1, \ldots, u_p$ as columns. Since $\sigma(\Lambda_p) \subset \mathcal{C}_+$, for each $j \in \{1, \ldots, p\}$ there exists a positive real number ψ_j such that

$$(\lambda_j - i \cdot \theta\psi_j) \in \mathcal{C}.$$

Write $\Psi \in \mathbb{C}^{p \times p}$ for the diagonal matrix with $\sigma(\Psi) = \{\sqrt{\psi_1}, \,\ldots, \, \sqrt{\psi_p}\}$ and set

$$Z = U_p\Psi.$$

By Lemma 3.1.1 the columns of U_p are also eigenvectors of $\Theta(A)$ and thus $i \cdot \theta ZZ^*$ is a θ-skew-Hermitian normality preserving perturbation of A. Moreover,

$$u_j^*(A - i \cdot \theta ZZ^*)u_j = (\lambda - i \cdot \theta\psi) \in \mathcal{C}.$$

Because the matrix $A - i \cdot \theta ZZ^*$ is normal with all eigenvalues in $\mathcal{C}$, the equality

$\Theta(A - i \cdot \theta Z Z^*) = \Theta(A)$ leads to

$$A - i \cdot \theta Z Z^* = \overline{\theta}\Theta(A) + i \cdot \theta \wp\left(\overline{\theta}\Theta(A)\right).$$

The assumption $r_2 > 0$, justified in Remark 3.4.1, now gives that

$$A = \overline{\theta}\Theta(A) + i \cdot \theta \wp\left(\overline{\theta}\Theta(A)\right) + i \cdot \theta r_2 \left(\frac{Z}{\sqrt{r_2}}\right)\left(\frac{Z}{\sqrt{r_2}}\right)^*,$$

and according to Equation (3.18) this is precisely the $n \times n$ leading principal submatrix of A_+, where A_+ is defined as $A_+ = \overline{\theta} H + i \cdot \theta \wp(H)$, where

$$H = \begin{bmatrix} \overline{\theta}\Theta(A) & \hat{Z} \\ \hat{Z}^* & M \end{bmatrix}, \quad \text{with} \quad \hat{Z} = \frac{ZQ}{\sqrt{r_2}} \tag{3.19}$$

and $M, Q \in \mathbb{C}^{m \times m}$ are arbitrary Hermitian and unitary matrices.   □

For a given $n \times n$ normal matrix A, the typical situation is that after selecting θ suitably, at least $p = n - 3$ of its eigenvalues do not lie on a quadratic curve, and a matrix Z of rank p is needed to push those outliers onto $\mathcal{C}$. This is illustrated in Figure 3.3.

Remark 3.4.8. *It is, of course, possible to move each eigenvalue of A from $\mathcal{C}_+$ onto $\mathcal{C}$ as a result of an arbitrary amount of rank-one perturbations. This would increase the number of columns of Z, and give m-augmentations of A with $m > p$. However, this would not increase the rank of Z, and ZZ^* would remain the same. Together with the analysis of Section 3.3.3, that shows which θ-Hermitian pertur- bations are normality preserving, this shows that in essence, each p-augmentation A_+ with Z of full rank, is of the form (3.19). In Section 3.5.3 we give an explicit example of the construction in the proof of Theorem 3.4.7.*

Augmentations without computing eigenvalues

So far, explicit knowledge about the eigenvalues and eigenvectors of A was used to construct augmentations A_+. There are, however, cases in which it is sufficient to know the polynomial curve $\mathcal{E}$ on which the eigenvalues of A lie. To see this,
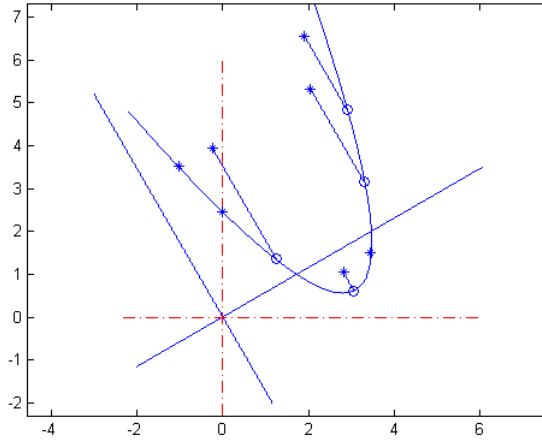
Figure 3.3: Illustration of the construction in the proof of Theorem 3.4.7. Three of the seven eigenvalues of A, indicated by the circles, are already on the quadratic curve $\mathcal{C}$, and a rank-4 matrix Z is needed to push the remaining four eigenvalues onto $\mathcal{C}$, after which the augmented matrix can be formed.

assume that A is normal and

$$A = \Theta(A) + i \cdot \theta \wp \left(\overline{\theta} \Theta(A) \right), \qquad \wp \in \mathcal{P}^{2k}(\mathbb{R}) \quad \text{and} \quad \wp \notin \mathcal{P}^{2k-1}(\mathbb{R})$$

for some integer $k \geq 1$. Since $\wp$ has even degree, there exist polynomials $p \in \mathcal{P}^2(\mathbb{R})$ such that

$$\wp(x) - p(x) \geq 0 \quad \text{for all} \quad x \in \mathbb{R}.$$

This implies that the matrix $(\wp - p) \left(\overline{\theta} \Theta(A) \right)$, is positive semi-definite and, hence, it can be factorized as

$$(\wp - p) \left(\overline{\theta} \Theta(A) \right) = ZZ^*, \tag{3.20}$$

after which we have that

$$A = \Theta(A) + i \cdot \theta p \left(\overline{\theta} \Theta(A) \right) + i \cdot \theta ZZ^*.$$

It is trivial that the matrix $i \cdot \theta ZZ^*$ is a normality preserving perturbation of $\Theta(A) + i \cdot \theta p \left(\overline{\theta} \Theta(A) \right)$ and by choosing between θ and $-\theta$, as explained in Remark 3.4.1, this leads to an m-augmentation of A, with generally $m = n - 1$. Section 3.1.1 explained how $\wp$ can be computed in a finite number of arithmetic operations, and the same is valid for the factorization (3.20). Of course, the problem of finding a minorizing polynomial $p \in \mathcal{P}(\mathbb{R})$ may prove to be difficult in specific situations.

3.4.2 Polynomial curves of higher degree

If one tries to generalize the approach of Section 3.4.1 to polynomial curves $\mathcal{C}$ of higher degree, the situation rapidly becomes more difficult. As an illustration, consider the cubic case. The third power of the matrix $\overline{\theta} \Theta(A_+)$ in (3.17) equals

$$\begin{bmatrix} X^3 + XZZ^* + ZZ^*X + ZMZ^* & X^2Z + XZM + ZM^2 + ZZ^*Z \\ Z^*X^2 + Z^*ZZ + MZ^*X + M^2Z & M^3 + Z^*ZM + MZ^*Z + Z^*XZ \end{bmatrix},$$

and thus, comparing the leading principal submatrices,

$$A = \Theta(A) + i \cdot \theta\theta\wp(\overline{\theta}A) + i \cdot \theta r_2 ZZ^* + i \cdot \theta r_3 (XZZ^* + ZZ^*X + ZMZ^*),$$

where r_3 is the coefficient of x^3 of $\wp$. Thus, A is a rank-k (with $k \leq 2m$), θ-skew-Hermitian perturbation of the normal matrix $\Theta(A) + i \cdot \theta\wp(\overline{\theta}A)$. Of course, if ZZ^* commutes with $\Theta(A)$, then it commutes with X, and this may help the analysis. However, it becomes much harder to control the perturbation in such a way, that A will be augmented into a matrix A_+ with $\sigma(A_+) \subset \mathcal{C}$. Therefore, we will not pursue this idea any further.

3.5 Illustrations

In this section we present some illustrations of the main constructions and theorem of this chapter. By making them explicit, we hope to create more insight in their structure.

3.5.1 Illustrations belonging to Section 3.2

This example, illustrates Theorem 3.2.2 and Theorem 3.2.3. Let A be the matrix

$$A = \begin{bmatrix} 2i & & \\ & 2+i & \\ & & -3 \end{bmatrix}.$$

Take $\gamma = 1$ and choose ℓ the line through $\gamma = 1$ and the eigenvalue $2 + i$ of A,

$$\ell : \mathbb{R} \to \mathbb{C} : \quad \rho \mapsto 1 + \rho\theta, \quad \text{with} \quad \theta = e^{i\frac{\pi}{4}} = \frac{1+i}{\sqrt{2}}.$$

Thus, in order for A_+ to be normal, u must be a multiple of e_2, $v = \theta u$ and $w^* = \theta u^*$, which yields that for all $\mu \in \mathbb{C}$,

$$A_+ = \left[\begin{array}{ccc|c} 2i & & & \\ & 2+i & & \theta\mu \\ & & -3 & \\ \hline & \theta\overline{\mu} & & 1 \end{array} \right]$$

is a normal augmentation of A. Moreover, for the choice $\gamma = 1$ and ℓ, these are all the normal augmentations of A. The eigenvalues of A_+ that are not eigenvalues of A are the eigenvalues of

$$\begin{bmatrix} 2+i & \theta\mu \\ \theta\overline{\mu} & 1 \end{bmatrix} = \begin{bmatrix} 1+\theta\sqrt{2} & \theta\mu \\ \theta\overline{\mu} & 1 \end{bmatrix} = \theta \begin{bmatrix} \sqrt{2} & \mu \\ \overline{\mu} & 0 \end{bmatrix} + I,$$

and thus equal to

$$\lambda = \theta \left(\frac{\sqrt{2}}{2} \pm \sqrt{2\mu^2 + 1} \right) + 1,$$

which lie on ℓ and have the eigenvalue $2 + i$ (that was perturbed) as average, as is depicted in the left in Figure 3.4. Here, the stars represent eigenvalues of A, the circles depict different choices for γ and the squares indicate the perturbed eigenvalues.

A second option is to choose $\gamma = -2 + 3i$ instead of $\gamma = 1$. This gives other possibilities to construct normality preserving augmentations. The first one is to choose $\gamma = -2 + 3i$ instead of $\gamma = 1$. This gives other possibilities to construct normality preserving augmentations. The first one is to choose ℓ through γ and -3, which shows that u must be a multiple of e_3 and $v = \theta u$ and $w^* = \theta u^*$ where $\theta = (1 + 3i)/\sqrt{10}$, which is a similar situation as for $\gamma = 1$. The second non-trivial option is to choose ℓ through γ and both $2i$ and $2 + i$. Then with $\theta = (-2 + i)/\sqrt{5}$, we may take u as a linear combination of e_1 and e_2, showing that

$$
A_+ = \left[
\begin{array}{ccc|c}
2i & & & \theta\alpha \\
 & 2+i & & \theta\mu \\
 & & -3 & 0 \\
\hline
\theta\overline{\alpha} & \theta\overline{\mu} & 0 & -2+3i
\end{array}
\right].
$$

is normal for all $\alpha, \mu \in \mathbb{C}$. For the given value of γ those two options are the only possible normality preserving augmentations. The perturbed eigenvalues are depicted in the right picture in Figure 3.4.
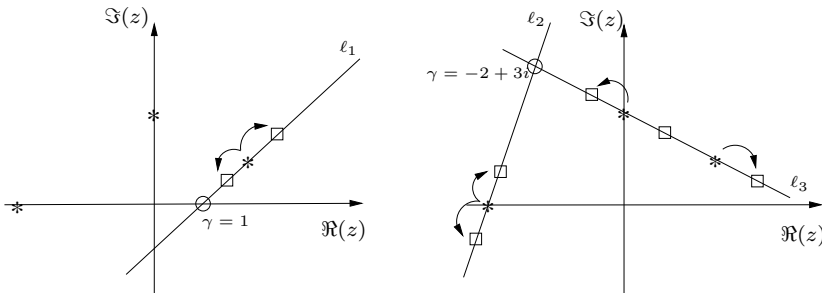

Figure 3.4: Normality preserving augmentations of a 3×3 matrix for two different values of γ.

3.5.2 Illustrations belonging to Section 3.3

Normality preserving non-normal perturbations

First we illustrate Theorem 3.3.10 by presenting an example of a line through two normal matrices that contains only two normal matrices. For this, let

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \text{and} \quad E = B - A = \begin{bmatrix} 0 & 0 \\ -2 & 0 \end{bmatrix}.$$

Thus, E is a *non-normal normality preserving perturbation* of A. Hence, apart from A and B, no other matrix of the form $A + tE$ with $t \in \mathbb{R}$ is normal. Indeed,

$$[A + tE, (A + tE)^*] = \begin{bmatrix} 1 - (1 - 2t)^2 & 0 \\ 0 & (1 - 2t)^2 - 1 \end{bmatrix},$$

and this matrix is only zero for $t = 0$ and $t = 1$. Moreover, the eigenvalues of $A + tE$ are

$$\sqrt{1 - 2t} \quad \text{and} \quad -\sqrt{1 - 2t}$$

whereas the corresponding sums of the eigenvalues of $\mathcal{H}(A + tE)$ and $\mathcal{S}(A + tE)$ equal

$$(1 - t) + it \quad \text{and} \quad -(1 - t) - it.$$

Thus, for instance at $t = \frac{1}{2}$, the eigenvalue zero of $A + tE$ is not the sum of the eigenvalues of the the Hermitian and skew-Hermitian parts of $A + tE$.

Normality preserving normal perturbations

Now we illustrate Theorem 3.3.13. This concerns *normality preserving normal perturbations*. For this, we take $\theta = 1$ and consider the matrix $A + E$, where

$$A = \begin{bmatrix} 1 & & \\ & i & \\ & & 1 + i \end{bmatrix} \quad \text{and} \quad E = u_1 u_1^* + 2u_2 u_2^*$$

with $u_1, u_2 \in \mathbb{C}^3$ mutually orthonormal. Then $A + E$ is normal if and only of both u_1 and u_2 are eigenvectors of

$$2\Theta^\perp(A) = 2S_A = \begin{bmatrix} 0 & & \\ & i & \\ & & i \end{bmatrix}.$$

However, $A + E$ where $E = u_1 u_1^* + u_2 u_2^*$ with $u_1, u_2 \in \mathbb{C}^3$ mutually orthonormal, is normal if and only if both u_1 and u_2 are *linear combinations of the same two eigenvectors* v_1 and v_2 of $2\Theta^\perp(A)$. Thus, with

$$u_1 = \frac{1}{4} \begin{bmatrix} \sqrt{2} \\ 1 \\ 1 \end{bmatrix}, \quad u_2 = \frac{1}{4} \begin{bmatrix} -\sqrt{2} \\ 1 \\ 1 \end{bmatrix}, \quad \text{where } v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } v_2 = \begin{bmatrix} 0 \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix},$$

we have that

$$E_1 + E_2 = \frac{1}{16} \begin{bmatrix} 2 & \sqrt{2} & \sqrt{2} \\ \sqrt{2} & 1 & 1 \\ \sqrt{2} & 1 & 1 \end{bmatrix} + \frac{1}{16} \begin{bmatrix} 2 & -\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & 1 & 1 \\ -\sqrt{2} & 1 & 1 \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 4 & & \\ & 2 & 2 \\ & 2 & 2 \end{bmatrix}$$

is a normality preserving rank-two perturbation of A, written as the sum of two rank-one perturbations that individually do not preserve normality. However, we also have that

$$\frac{1}{4} v_1 v_1^* + \frac{1}{2} v_2 v_2^* = \frac{1}{16} \begin{bmatrix} 4 & & \\ & 0 & \\ & & 0 \end{bmatrix} + \frac{1}{16} \begin{bmatrix} 0 & & \\ & 2 & 2 \\ & 2 & 2 \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 4 & & \\ & 2 & 2 \\ & 2 & 2 \end{bmatrix}$$

and this expresses the perturbation as a sum of two normality preserving normal rank-one perturbations. The eigenvalues of $A + tE$ are $1 + \frac{1}{4}t$ due to the term $\frac{1}{4} v_1 v_1^*$, together with the eigenvalues of

$$\begin{bmatrix} i & \\ & 1+i \end{bmatrix} + \frac{1}{8}t \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \text{which are} \quad i + \frac{1}{2}\left(1 + \frac{t}{4} \pm \sqrt{1 + \frac{t^2}{16}}\right),$$

which are due to the term $\frac{1}{4}v_2v_2^*$. As stated in Theorem 3.3.15, the rank-2 pertur-
bation moves the eigenvalues of A in the horizontal direction. Since the eigenvalues
i and $1+i$ of A are on the same horizontal line, they can be simultaneously per-
turbed by a rank-one perturbation. For $t \in [0,4]$ those eigenvalues are plotted
by circles in Figure 3.5. We also computed the eigenvalues of $A + tE_1$ and of
$A + 4E_1 + tE_2$ for $t \in [0,4]$ and indicated them by asterisks and boxes, respec-
tively. As is visible in Figure 3.5, the eigenvalues leave the straight line before
returning to the eigenvalues of the normal matrix $A + 4E_1 + 4E_2 = A + 4E$.
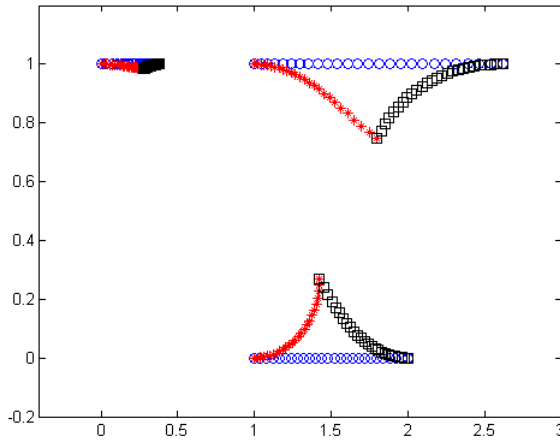

Figure 3.5: Eigenvalue trajectories of a normality preserving perturbation, and of the
same perturbation written as the sum of non-normality preserving normal perturbations.

3.5.3   Illustrations belonging to Section 3.4

We now illustrate Theorem 3.4.7. The starting point is a 3×3 matrix A, chosen
so that the conditions of the theorem are easy to satisfy. Let

$$A = \begin{bmatrix} 5i & & \\ & 1 & \\ & & 2+2i \end{bmatrix} \quad \text{and thus,} \quad \mathcal{H}(A) = X = \begin{bmatrix} 0 & & \\ & 1 & \\ & & 2 \end{bmatrix}.$$

With $\theta = 1$, the eigenvalues already lie on a parabolic curve, and thus also with the trivial choice $Z = O_n$ augmentations A_+ can be constructed having eigenvalues on the same curve. More interesting is to choose a $Z \neq O_n$ such that $A - iZZ^*$ is normal. Since $\mathcal{H}(A)$ has distinct eigenvalues, Z needs to have eigenvectors of $\mathcal{H}(A)$ as columns. Take for example

$$Z = \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{hence} \quad ZZ^* = \begin{bmatrix} 5 & & \\ & 0 & \\ & & 1 \end{bmatrix} \quad \text{and} \quad A - iZZ^* = \begin{bmatrix} i & & \\ & 1 & \\ & & 2 + i \end{bmatrix}.$$

The eigenvalues of $A - iZZ^*$ lie on the curve $\mathcal{C}$ that is the image of

$$q : \mathbb{R} \to \mathbb{C} : \quad \rho \mapsto \rho + i(1 - \rho)^2.$$

Augmentation A_+ can now be constructed by choosing an arbitrary Hermitian 2×2 matrix M and an arbitrary unitary matrix Q, for instance

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad Q = \frac{1}{2}\sqrt{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

and then to form the Hermitian part of A_+ as

$$\mathcal{H}(A_+) = \begin{bmatrix} Z & ZQ \\ Q^*Z^* & M \end{bmatrix} = \begin{bmatrix} 0 & & & \sqrt{2} & \sqrt{2} \\ & 1 & & 0 & 0 \\ & & 2 & \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ \sqrt{2} & 0 & \frac{1}{2}\sqrt{2} & 1 & 1 \\ \sqrt{2} & 0 & -\frac{1}{2}\sqrt{2} & 1 & 2 \end{bmatrix}$$

Finally, A_+ itself can be formed as $A_+ = \mathcal{H}(A) + iq(\mathcal{H}(A)) = \mathcal{H}(A) + i(I - \mathcal{H}(A))^2$, resulting in

$$A_+ = \begin{bmatrix} 5i & & & \sqrt{2} & \sqrt{2}+\sqrt{2}i \\ & 1 & & 0 & 0 \\ & & 2+2i & \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2}-\frac{1}{2}\sqrt{2}i \\ \sqrt{2} & 0 & \frac{1}{2}\sqrt{2} & 1+3\frac{1}{2}i & 1+2\frac{1}{2}i \\ \sqrt{2}+\sqrt{2}i & 0 & -\frac{1}{2}\sqrt{2}-\frac{1}{2}\sqrt{2}i & 1+2\frac{1}{2}i & 2+4\frac{1}{2}i \end{bmatrix}$$

Indeed, A_+ is a 2-augmentation of A. From this example we observe that if Z, M, Q are chosen real, then $\mathcal{H}(A_+)$ is real symmetric and A_+ complex symmetric, being the sum of a real symmetric matrix and i times a polynomial of this real symmetric matrix. Note that not all complex symmetric matrices are normal. In fact, the leading 4×4 principal submatrix of A_+ in the above example is not normal, nor is the trailing 2×2 principal submatrix. Thus A_+ could not have been constructed using the procedure for $m = 1$ twice.