

WHEN IS THE ADJOINT OF A MATRIX A LOW DEGREE RATIONAL FUNCTION IN THE MATRIX?

JÖRG LIESEN*

Abstract. We show that the adjoint A^+ of a matrix A with respect to a given inner product is a rational function in A , if and only if A is normal with respect to the inner product. We consider such matrices and analyze the McMillan degrees of the rational functions r such that $A^+ = r(A)$. We introduce the McMillan degree of A as the smallest among these degrees, characterize this degree in terms of the number and distribution of the eigenvalues of A , and compare the McMillan degree with the normal degree of A , which is defined as the smallest degree of a polynomial p for which $A^+ = p(A)$. We show that unless the eigenvalues of A lie on a single circle in the complex plane, the ratio of McMillan degree and normal degree of A is bounded by a small constant that depends neither on the number nor on the distribution of the eigenvalues of A . Our analysis is motivated by applications in the area of short recurrence Krylov subspace methods.

Key words. normal matrices, representation of matrix adjoints, rational interpolation, Krylov subspace methods, short recurrences

AMS subject classifications. 15A21, 30C15, 65F10.

1. Introduction. Consider a unitary matrix A with $n \geq 2$ distinct eigenvalues. Since A is normal, its adjoint A^* is a polynomial in A [9, condition 17],

$$(1.1) \quad A^* = p(A).$$

It has been observed in several publications, e.g. [4, pp. 774–775], that for a unitary matrix A , (1.1) does not hold for a polynomial p of “small” degree. In this paper we strengthen this observation by showing that the smallest degree of such polynomial is equal to $n - 1$. On the other hand, the *McMillan degree* of a rational function $r = p/q$, where p and q are relatively prime polynomials (i.e. their only common divisor is the constant polynomial 1), is defined as

$$(1.2) \quad \deg r = \max \{ \deg p, \deg q \}.$$

Hence

$$(1.3) \quad A^* = r(A),$$

where, since A is unitary, $r(z) = 1/z$ is a rational function of McMillan degree one. In summary, the adjoint of a unitary matrix A is a large degree polynomial and a small (McMillan) degree rational function in A . The observation that the adjoint of a normal matrix may be represented as a polynomial as well as a rational function in the matrix, and that the degrees of these representations may vastly differ, is more than a curiosity. In fact, it is of great importance for the construction of short recurrence Krylov subspace methods.

On the one hand, the fundamental theorem of Faber and Manteuffel [6] shows that if (1.1) holds for a matrix A and a polynomial of degree s , then orthogonal Krylov subspace bases for A can be generated by an $(s + 2)$ -term Arnoldi recurrence (this condition is not only sufficient but also necessary; see [6] or [14] for more details).

*Institute of Mathematics, Technical University of Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, email: liesen@math.tu-berlin.de. This work was supported by the Emmy Noether-Programm of the Deutsche Forschungsgemeinschaft.

For a unitary matrix A with n distinct eigenvalues, $A^* = p(A)$ with the smallest possible degree of p being $n - 1$. Thus, generating orthogonal Krylov subspace bases for unitary matrices via the Arnoldi process requires a full recurrence.

On the other hand, as shown by Barth and Manteuffel [3, 4], if (1.3) holds for a matrix A and a rational function $r = p/q$, where p and q are relatively prime polynomials of respective degrees ℓ and m , then orthogonal Krylov subspace bases for A can be generated by a recurrence containing $\ell + m + 2$ terms. The generic type of this recurrence is displayed in [4, equation (4.1)], from which it is easily seen that this recurrence is not of Arnoldi-type (partial necessary conditions for the existence of this recurrence are given in the unpublished report [5]). For stability reasons this (single) recurrence should be implemented in form of coupled, or multiple recurrences (see [4] for details), but the actual implementation is not important for us here. The important point here is that when ℓ and m are small, an orthogonal Krylov subspace basis can be generated by a short recurrence. For a unitary matrix A , $r(z) = 1/z$, hence $\ell = 0$ and $m = 1$, so that the length of this recurrence is three. In practical applications one uses coupled two-term recurrences instead of the numerically unstable three-term version. The resulting algorithm, which originally was discovered by Gragg in the context of orthogonal polynomials on the unit circle [8], is called the *isometric Arnoldi algorithm*. This algorithm has been used for solving unitary eigenvalue problems (see [17] for a survey) as well as for constructing a minimal residual method for solving linear systems with shifted unitary matrices [11].

For efficiency reasons we would like to use the shortest possible recurrence, and thus we would like to characterize, for a given matrix A , the smallest degrees of p and r (if any) such that (1.1) and (1.3), respectively, hold. While the smallest degree of the representation (1.1) has been characterized in the literature (we here give new proofs of some results), comparably little has been done to characterize (1.3). The only related work we are aware of is in the afore mentioned papers of Barth and Manteuffel. There, for a given Hermitian positive definite (HPD) matrix B , a matrix A is called *B-normal*(ℓ, m), if A is normal with respect to the inner product generated by B , and if its adjoint A^+ with respect to this inner product satisfies $A^+ = r(A)$, where $r = p/q$ for relatively prime polynomials of respective degrees ℓ and m , cf. [4, Definition 3.1]. For a given representation $A^+ = r(A)$ with known degrees ℓ and m , Barth and Manteuffel derive bounds on the maximal number of distinct eigenvalues of A in terms of ℓ and m , cf. [4, Theorem 3.1], or [3, Theorem 4.1]. However, they provide no characterization how small, or large ℓ and m may be for a given matrix A , which is the question of interest in this paper (see Remarks 2.4 and 3.7 below for further comments on the *B-normal*(ℓ, m) matrices).

To allow a rigorous characterization of (1.3), we here introduce the concept of the McMillan degree of A , which we define as the smallest McMillan degree of a rational function r such that $A^+ = r(A)$ (Section 2). In Section 3 we then completely answer the question raised in the title, which, as outlined above, has direct applications in the area of short recurrence Krylov subspace methods. Moreover, we show that unless the eigenvalues of A lie on a single circle in the complex plane, the ratio of the McMillan degree of A and the smallest degree of a polynomial representation of A^+ (called normal degree of A), is bounded from above by a small constant (less than five), that depends neither on the number nor on the distribution of the eigenvalues of A . In our derivations we apply results from rational interpolation theory, which apparently have not been used in this context before.

2. B -normal matrices. Suppose that A is a square matrix and B is an HPD matrix. Throughout the paper we will assume that these matrices are of the same size. The matrix B generates an inner product, $\langle x, y \rangle_B = y^* B x$, and the adjoint of A with respect to this inner product, or, shortly, the B -adjoint of A , is $A^+ = B^{-1} A^* B$. If A^+ is a polynomial in A , then A is said to be B -normal. This is a straightforward generalization of the common concept of normal matrices, which we here call I -normal. Of particular interest is the degree of the polynomial representation of the adjoint.

DEFINITION 2.1. *Let A be a square matrix and let B be an HPD matrix. If*

$$(2.1) \quad A^+ = p(A),$$

where p is a polynomial of the smallest possible degree s having this property, then A is called normal of degree s with respect to B , or, shortly, B -normal(s).

The property that A is B -normal(s) is completely characterized in the following result [14, Theorem 3.1].

THEOREM 2.2. *Let A be a square matrix and let B be an HPD matrix. Then the following two assertions are equivalent:*

- (1) A is B -normal(s).
- (2) a) A is diagonalizable with the eigendecomposition $A = W \Lambda W^{-1}$ (without loss of generality we consider the eigenvalues and eigenvectors of A ordered so that equal eigenvalues form a single diagonal block in Λ),

and

b) using the eigenvector matrix W of A , the matrix B^{-1} has the decomposition $B^{-1} = W D W^*$, where D is an HPD block diagonal matrix with block sizes corresponding to those of Λ ,

and

c) there exists a polynomial p of degree s such that $p(\Lambda) = \Lambda^*$, and s is the smallest degree of all polynomials with this property. The polynomial p is uniquely determined.

In [14] this result is stated only for nonsingular matrices A , which is due to the focus of the work in that paper. It is easy to see, however, that the assertion is true also for singular matrices A . Using Theorem 2.2, we can characterize all A and B for which $A^+ = r(A)$, where r is a rational function.

LEMMA 2.3. *Let A be a square matrix and let B be an HPD matrix. If there exists a rational function r such that $A^+ = r(A)$, then $r(\lambda) = \bar{\lambda}$ for all eigenvalues λ of A , and, moreover, A is B -normal(s), where s is the degree of the (uniquely determined) interpolation polynomial p of smallest degree that satisfies $p(\lambda) = \bar{\lambda}$ for all eigenvalues λ of A .*

Proof. We adopt the strategy of the proof of [6, Lemma 2]. Let (λ, x) be an eigenpair of A , $Ax = \lambda x$. Then $A^+x = r(A)x = r(\lambda)x$, so that

$$r(\lambda) \langle x, x \rangle_B = \langle r(\lambda)x, x \rangle_B = \langle A^+x, x \rangle_B = \langle x, Ax \rangle_B = \langle x, \lambda x \rangle_B = \bar{\lambda} \langle x, x \rangle_B,$$

from which we receive $r(\lambda) = \bar{\lambda}$.

Now suppose that there is a nontrivial Jordan block associated with λ . Then there exists a nonzero vector y such that $(A - \lambda I)y = x$. But then

$$\begin{aligned} \langle Ay, x \rangle_B &= \langle \lambda y + x, x \rangle_B = \lambda \langle y, x \rangle_B + \langle x, x \rangle_B, \\ \langle Ay, x \rangle_B &= \langle y, A^+x \rangle_B = \langle y, \bar{\lambda}x \rangle_B = \bar{\lambda} \langle y, x \rangle_B, \end{aligned}$$

which means that $\langle x, x \rangle_B = 0$. This contradiction shows that A is diagonalizable, i.e., that (2a) of Theorem 2.2 holds.

If (η, y) is another eigenpair of A with $\eta \neq \lambda$, then

$$\lambda \langle x, y \rangle_B = \langle \lambda x, y \rangle_B = \langle Ax, y \rangle_B = \langle x, A^+ y \rangle_B = \langle x, \bar{\eta} y \rangle_B = \eta \langle x, y \rangle_B.$$

Since $\lambda \neq \eta$ we must have $\langle x, y \rangle_B = 0$, which shows that the eigenvectors of A form a complete B -orthogonal set. In particular, when we consider the diagonalizable matrix A as in (2a) of Theorem 2.2, then $W^* B W = D$, where D is HPD and block diagonal, showing that B is as stated in (2b) of Theorem 2.2.

Finally, the polynomial p in (2c) of Theorem 2.2 is the uniquely determined interpolation polynomial of smallest degree that satisfies $p(\lambda) = \bar{\lambda}$ for all eigenvalues λ of A . \square

REMARK 2.4. According to Lemma 2.3, the existence of a representation of the form $A^+ = r(A)$, where r is a rational function, implies that A is B -normal(s). Therefore the assumption that A be B -normal in the definition of the B -normal(ℓ, m) matrices of Barth and Manteuffel, cf. [3, Definition 4.2] or [4, Definition 3.1], is redundant.

The converse of Lemma 2.3 is obviously true as well: If A is B -normal(s), then there exists a rational function, namely $r = p$ from (2c) in Theorem 2.2, such that $A^+ = r(A)$. We therefore have the following corollary.

COROLLARY 2.5. *For a square matrix A there exists an HPD matrix B such that A^+ is a rational function in A , if and only if A is diagonalizable.*

If A is diagonalizable, $A = W \Lambda W^{-1}$, then the HPD matrices B for which A^+ is a rational function in A are completely characterized in (2b) of Theorem 2.2. Moreover, if B is any such matrix, then $A^+ = r(A)$ holds for a rational function r , if and only if $r(\Lambda) = \Lambda^$.*

Proof. Only the necessity part in the last sentence remains to be shown. Let B be any matrix as characterized in (2b) of Theorem 2.2, i.e., $B = W^{-*} D W^{-1}$. Then

$$A^+ = B^{-1} A^* B = (W D^{-1} W^*) (W^{-*} \Lambda^* W^*) (W^{-*} D W^{-1}) = W \Lambda^* W^{-1} = r(A),$$

where in the last equation we have used that $r(\Lambda) = \Lambda^*$. \square

By Corollary 2.5, for a nondiagonalizable matrix A there exists no HPD matrix B such that the corresponding A^+ is a rational function in A . We therefore can restrict our attention to diagonalizable matrices. The last part of the corollary shows that if, for some HPD matrix B , A^+ is a rational function in A , $A^+ = r(A)$, then r is completely determined by the eigenvalues of A . We use the following concepts in our further development.

DEFINITION 2.6. *Let A be a diagonalizable matrix.*

- (1) *The (uniquely determined) smallest degree of a polynomial p that satisfies $p(\lambda) = \bar{\lambda}$ for all eigenvalues λ of A is called the normal degree of A , and is denoted by $d_p(A)$.*
- (2) *The (uniquely determined) smallest McMillan degree of a rational function r that satisfies $r(\lambda) = \bar{\lambda}$ for all eigenvalues λ of A is called the McMillan degree of A , and is denoted by $d_r(A)$.*

We immediately observe that $d_r(A) \leq d_p(A) \leq n - 1$, where n is the number of distinct eigenvalues of A .

Let us put the degrees $d_p(A)$ and $d_r(A)$ into the picture of short recurrence Krylov subspace methods that is described in the Introduction: On the one hand, if A is normal with respect to an HPD matrix B and $d_p(A) = s$, then B -orthogonal Krylov subspace bases for A can be generated with an $(s + 2)$ -term Arnoldi recurrence [6]. On the other hand, if for an HPD matrix B the B -adjoint of A satisfies $A^+ = r(A)$, where $r = p/q$ for relatively prime polynomials p and q of respective degrees ℓ and m , so that $d_r(A) = \deg r = \max\{\ell, m\}$, then B -orthogonal Krylov subspace bases for A can be generated using a non-standard recurrence containing $\ell + m + 2 \leq 2d_r(A) + 2$ terms [4]. If $d_r(A) \ll d_p(A)$, then the non-standard recurrence is significantly more efficient than the standard Arnoldi recurrence. It is therefore of great practical interest to characterize the (diagonalizable) matrices A for which $d_r(A) \ll d_p(A)$.

3. Characterization of the McMillan degree of A . We will study the McMillan degree of a diagonalizable matrix A using results from rational interpolation theory. The results we employ were originally developed by Antoulas and Anderson [2], and are summarized in Antoulas' book [1, Chapter 4.5].

Let $\lambda_1, \dots, \lambda_n$ be the *distinct* eigenvalues of A . We want to determine a rational function $r = p/q$, where p and q are relatively prime polynomials, such that $r(\lambda_j) = \bar{\lambda}_j$, $j = 1, \dots, n$. We assume $n \geq 2$, for otherwise the problem is trivial. If there exists such a rational function of McMillan degree m , then m is called an *admissible McMillan degree*. By definition, the smallest admissible McMillan degree is equal to $d_r(A)$.

Consider the array \mathbb{P} containing the interpolation points $(\lambda_j, \bar{\lambda}_j)$, $j = 1, \dots, n$,

$$(3.1) \quad \mathbb{P} = \{(\lambda_j, \bar{\lambda}_j) : j = 1, \dots, n\}.$$

We choose an integer n_1 , $1 \leq n_1 < n$, and partition \mathbb{P} into two disjoint subarrays \mathbb{J} and \mathbb{I} ,

$$\mathbb{J} = \{(\lambda_j, \bar{\lambda}_j) : j = 1, \dots, n_1\}, \quad \mathbb{I} = \{(\lambda_j, \bar{\lambda}_j) : j = n_1 + 1, \dots, n\}.$$

For notational convenience, we now write $\mu_j \equiv \lambda_{j+n_1}$ for $j = 1, \dots, n - n_1$. Then the *Löwner matrix* L corresponding to the arrays \mathbb{J} and \mathbb{I} is defined by

$$(3.2) \quad L = [l_{i,j}]_{i=1, \dots, n-n_1, j=1, \dots, n_1}, \quad \text{where } l_{i,j} = \frac{\bar{\mu}_i - \bar{\lambda}_j}{\mu_i - \lambda_j}.$$

Note that L is of size $(n - n_1) \times n_1$. Moreover, the *rank of the array* \mathbb{P} is defined as

$$(3.3) \quad \text{rank } \mathbb{P} = \max_L \{ \text{rank } L \},$$

where the maximum is taken over all possible Löwner matrices which can be formed from \mathbb{P} by partitioning into two subarrays as described above, cf. [1, Definition 4.51].

A similar construction can be made for any subarray of interpolation points. More precisely, we may take any $\mathbb{Q} \subset \mathbb{P}$ containing at least two points, partition \mathbb{Q} into two disjoint subarrays, and form the corresponding Löwner matrix according to (3.2). In this way we can form Löwner matrices from \mathbb{P} that are of size $k_1 \times k_2$ with $k_1 + k_2 < n$.

THEOREM 3.1. (cf. [1, Theorem 4.55 and Corollary 4.56]) *Suppose that the rank of the array \mathbb{P} in (3.1) is equal to m .*

- (1) If $2m < n$, and all $m \times m$ Löwner matrices that can be formed from \mathbb{P} are nonsingular, then there exists a uniquely determined rational function $r = p/q$ (with p and q relatively prime) of McMillan degree m with $r(\lambda_j) = \bar{\lambda}_j$, $j = 1, \dots, n$. Moreover, the admissible McMillan degrees are m and all integers greater than or equal to $n - m$.
- (2) Otherwise, the admissible McMillan degrees are all integers greater than or equal to $n - m$, and there exists no uniquely determined rational function $r = p/q$ (with p and q relatively prime) of McMillan degree $n - m$ with $r(\lambda_j) = \bar{\lambda}_j$, $j = 1, \dots, n$.

The following is a straightforward consequence.

COROLLARY 3.2. *For any diagonalizable matrix A with n distinct eigenvalues, $d_r(A) \leq \lceil n/2 \rceil$. In particular, if $n \in \{2, 3\}$, then $d_r(A) = 1$.*

Having characterized the cases $n = 2$ and $n = 3$, we will now focus on matrices with at least four distinct eigenvalues.

LEMMA 3.3. *Let $\lambda_1, \dots, \lambda_4$ be any four distinct complex numbers, and consider the corresponding array \mathbb{P} of the form (3.1). Then $\text{rank } \mathbb{P} = 1$, if and only if $\lambda_1, \dots, \lambda_4$ are either collinear (i.e. they lie on a single line in the complex plane) or concyclic (i.e. they lie on a single circle in the complex plane).*

Proof. We partition \mathbb{P} into two subarrays \mathbb{J} and \mathbb{I} containing two interpolation points each. The corresponding Löwner matrix is

$$L = \left[\begin{array}{cc} \bar{\mu}_i - \bar{\lambda}_j & \\ \mu_i - \lambda_j & \end{array} \right]_{i,j=1,2},$$

giving

$$\det L = \frac{\bar{\mu}_1 - \bar{\lambda}_1}{\mu_1 - \lambda_1} \frac{\bar{\mu}_2 - \bar{\lambda}_2}{\mu_2 - \lambda_2} - \frac{\bar{\mu}_2 - \bar{\lambda}_1}{\mu_2 - \lambda_1} \frac{\bar{\mu}_1 - \bar{\lambda}_2}{\mu_1 - \lambda_2}.$$

Hence $\det L = 0$, if and only if

$$(3.4) \quad \frac{(\mu_1 - \lambda_1)(\mu_2 - \lambda_2)}{(\mu_1 - \lambda_2)(\mu_2 - \lambda_1)} \in \mathbb{R}.$$

We denote by $\widehat{\mathbb{C}}$ the extended complex plane. Recall that a circle in $\widehat{\mathbb{C}}$ is either a true circle in the complex plane, or a line in the complex plane with the point at infinity adjoined. In (3.4) we replace μ_2 by a variable z , and consider the function

$$f(z) = \frac{(\mu_1 - \lambda_1)(z - \lambda_2)}{(\mu_1 - \lambda_2)(z - \lambda_1)}.$$

The function $f(z)$ is the unique Moebius transformation satisfying

$$f(\lambda_1) = \infty, \quad f(\lambda_2) = 0, \quad f(\mu_1) = 1.$$

Now realize that through the points $\lambda_1, \lambda_2, \mu_1$ passes one and only one circle \mathcal{C} in $\widehat{\mathbb{C}}$. Since the Moebius transformation f conformally maps circles in $\widehat{\mathbb{C}}$ onto circles in $\widehat{\mathbb{C}}$, we see that $f(\mathcal{C}) = \mathbb{R} \cup \{\infty\}$, and, in particular, $f(\mu_2) \in \mathbb{R}$, if and only if $\mu_2 \in \mathcal{C}$ (see, e.g., [15, Chapter 3] for more on Moebius transformations). Consequently, L is

singular, if and only if $\lambda_1, \lambda_2, \mu_1, \mu_2$ lie on the same circle in $\widehat{\mathbb{C}}$, i.e. if and only if these points in the complex plane are either collinear or concyclic, which completes the proof. \square

Using this lemma we can characterize the diagonalizable matrices of McMillan degree one. To do so, we recall that a matrix has rank k , if and only if it has a nonsingular $k \times k$ submatrix, and all its $(k + 1) \times (k + 1)$ submatrices are singular (see, e.g., [10, pp. 12–13]).

LEMMA 3.4. *Let A be a diagonalizable matrix with at least four distinct eigenvalues. Then $d_r(A) = 1$, if and only if the eigenvalues of A are either collinear or concyclic.*

Proof. Let L be any Löwner matrix with at least two rows and columns formed from the array \mathbb{P} corresponding to the $n \geq 4$ distinct eigenvalues of A . Clearly, $\text{rank } L \geq 1$.

If the eigenvalues of A are either collinear or concyclic, then Lemma 3.3 shows that every 2×2 submatrix of L is singular. Therefore, $\text{rank } L = 1$, which shows that $\text{rank } \mathbb{P} = 1$. Since $2 < n$, and every 1×1 Löwner matrix that can be formed from \mathbb{P} is nonsingular, case (1) in Theorem 3.1 applies, showing that $d_r(A) = 1$.

On the other hand, if the eigenvalues of A are neither collinear nor concyclic, then by Lemma 3.3 there exists a 2×2 submatrix of L that is nonsingular. Hence $\text{rank } L \geq 2$, which implies that $\text{rank } \mathbb{P} \geq 2$ and hence $d_r(A) \geq 2$. \square

If A is any diagonalizable matrix with $n \geq 2$ distinct eigenvalues that all are collinear, then Corollary 3.2 and Lemma 3.4 show that $d_r(A) = 1$. These matrices are also known to be B -normal(1) for some HPD matrix B (cf. [14, Theorem 3.3] and the references given there), and thus they satisfy $d_p(A) = 1$.

More interesting is the class of the diagonalizable matrices with $n \geq 3$ distinct eigenvalues that all are concyclic. For such matrices A , Corollary 3.2 and Lemma 3.4 show that $d_r(A) = 1$. Moreover, case (1) in Theorem 3.1 shows that the rational function r of McMillan degree one that satisfies $r(\lambda) = \bar{\lambda}$ for all eigenvalues λ is uniquely determined. In fact, this function can be easily computed. Suppose that the eigenvalues of A are given by

$$\lambda_j = \rho e^{i\varphi_j} + \zeta, \quad j = 1, \dots, n,$$

where $\rho \in \mathbb{R} \setminus \{0\}$ and $\zeta \in \mathbb{C}$ do not depend on j , while $\varphi_j \in [0, 2\pi)$. Then

$$r(z) = \frac{\bar{\zeta}z + (\rho^2 - |\zeta|^2)}{z - \zeta}$$

satisfies $r(\lambda_j) = \bar{\lambda}_j$, $j = 1, \dots, n$. Clearly, r is not a polynomial. Case (1) in Theorem 3.1 also shows that the next smallest admissible McMillan degree is $n - 1$. Apparently, a corresponding rational function is the uniquely determined (Lagrange) interpolation polynomial p that satisfies $p(\lambda) = \bar{\lambda}$ for all eigenvalues λ . This means that $d_p(A) = n - 1$.

It is easy to see that, for a diagonalizable matrix A , $d_r(A)$ is equal to the smallest possible McMillan degree of a rational function r such that the eigenvalues of A are zeros of the function $r(z) - \bar{z}$. When $n \geq 4$ and the eigenvalues of A are neither collinear nor concyclic, Lemma 3.4 implies that $d_r(A) \geq 2$. Hence in this case we

search for a rational function r of (smallest possible) $\deg r \geq 2$, such that the eigenvalues of A are zeros of $r(z) - \bar{z}$. The following result summarizes what is known about the zeros of such functions.

THEOREM 3.5.

- (1) A function of the form $p(z) - \bar{z}$, where p is a polynomial of degree $s \geq 2$, has at most $3s - 2$ zeros. For any $s \geq 2$ there exists a polynomial p of degree s , such that $p(z) - \bar{z}$ has $3s - 2$ zeros.
- (2) A function of the form $r(z) - \bar{z}$, where r is a rational function of McMillan degree $s \geq 2$, has at most $5s - 5$ zeros. For any $s \geq 2$ there exists a rational function r of McMillan degree s , such that $r(z) - \bar{z}$ has $5s - 5$ zeros.

The bounds in (1) and (2) have been shown in [13] and [12], respectively. The corresponding sharpness results have been shown in [7] and [16]. Using these bounds we can prove the following result.

THEOREM 3.6. *Let A be a diagonalizable matrix with $n \geq 4$ distinct eigenvalues.*

- (1) *If the eigenvalues are collinear, then $d_r(A) = d_p(A) = 1$.*
- (2) *If the eigenvalues are concyclic, then $d_r(A) = 1$ and $d_p(A) = n - 1$.*
- (3) *In all other cases, $d_r(A) \geq \lfloor n/5 + 1 \rfloor$, $d_p(A) \geq \lfloor (n + 2)/3 \rfloor$, and*

$$(3.5) \quad 1 \leq \frac{d_p(A)}{d_r(A)} \leq 5 \frac{n-1}{n+5} < 5.$$

Proof. Cases (1) and (2) were shown above, so only case (3) needs to be proven. Here the eigenvalues are neither collinear nor concyclic, and thus by Lemma 3.4 we must have $d_r(A) \geq 2$. From case (2) in Theorem 3.5 we know that any function of the form $r(z) - \bar{z}$, where r is a rational function of $\deg r \geq 2$, may have at most $5 \deg r - 5$ zeros. Since any function for which the McMillan degree of A is attained must have (at least) n distinct zeros, we must have $n \leq 5d_r(A) - 5$, and thus $d_r(A) \geq \lfloor n/5 + 1 \rfloor$. The lower bound on $d_p(A)$ follows in a similar way from case (1) in Theorem 3.5. Finally, the leftmost and rightmost inequalities in (3.5) are straightforward, while the middle inequality follows from the lower bound on $d_r(A)$ and from noting that $d_p(A) \leq n - 1$. \square

Using Theorem 2.2, Corollary 3.2, and Theorem 3.6 we can derive the following well known result: There exists an HPD matrix B with respect to which a matrix A with at least two distinct eigenvalues is normal of degree one, if and only if A is diagonalizable and has collinear eigenvalues (cf. [14, Theorem 3.3] and the references given there). Here we have given a new proof of this result using rational interpolation theory and conformal mappings.

A surprising fact shown by Theorem 3.6 is that the ratio $d_r(A)/d_p(A)$ is bounded from above by five, unless the eigenvalues of A are concyclic, in which case the ratio is equal to $n - 1$. In this sense, the diagonalizable matrices with concyclic eigenvalues form a very special class.

Theorem 3.6 also shows that if the eigenvalues of a diagonalizable matrix A are neither collinear nor concyclic, then $d_r(A)$ is small, if and only if A has only a small number (at most $5d_r(A) - 5$) of distinct eigenvalues.

REMARK 3.7. A related observation is made after the statement of [4, Theorem 3.1], but it is not fully justified from the theory presented there. According to Barth and Manteuffel, their result “says that if A is B -normal(ℓ, m) and either ℓ or m

greater than 1, then A has a relatively small number of distinct eigenvalues” [4, p. 775]. However, in terms of [4, Definition 3.1], any unitary matrix A with $n \geq 3$ distinct eigenvalues is I -normal(0,1) and I -normal($n - 1,0$). Hence, for $\ell = n - 1 > 1$ and $m = 0$, A is B -normal(ℓ,m), but A may have arbitrarily many distinct eigenvalues. The confusion is caused by the lack of uniqueness of the “smallest degrees” ℓ and m . In general, there exist no “simultaneously smallest” ℓ and m for which $A^+ = r(A)$ with $r = p/q$ for relatively prime polynomials of respective degrees ℓ and m .

We next show by examples that the two weak inequalities in (3.5) cannot be improved in general. First consider the lower bound on $d_p(A)/d_r(A)$. This bound is attained, if and only if a rational function r of smallest possible McMillan degree, that satisfies $r(\lambda) = \bar{\lambda}$ for all eigenvalues λ of A , is a polynomial (this always holds when the eigenvalues of A are collinear, cf. case (1) in Theorem 3.6, where $d_r(A) = d_p(A) = 1$). Consider a diagonalizable matrix A with $n = 4$ distinct eigenvalues given by

$$\lambda_1 = 1 + \sqrt{1/2}, \quad \lambda_2 = 1 - \sqrt{1/2}, \quad \lambda_3 = i\sqrt{1/2}, \quad \lambda_4 = -i\sqrt{1/2}.$$

The polynomial

$$p(z) = z^2 - z + \frac{1}{2}$$

is the unique polynomial of smallest possible degree that satisfies $p(\lambda_j) = \bar{\lambda}_j$, $j = 1, \dots, 4$, so that $d_p(A) = 2$. On the other hand, $d_r(A) \leq 2$ by Corollary 3.2, and since the four eigenvalues are neither collinear nor concyclic, Lemma 3.4 implies that $d_r(A) = 2$, showing that the lower bound in (3.5) is attained.

To give an example that the upper bound is attained, consider any diagonalizable matrix A with $n = 5$ distinct eigenvalues that are neither collinear nor concyclic. By Corollary 3.2, $d_r(A) \leq \lceil 5/2 \rceil = 2$, and by Lemma 3.4, $d_r(A) > 1$, showing that $d_r(A) = 2$. Suppose that the five eigenvalues are

$$\lambda_1 = 0, \quad \lambda_2 = 1, \quad \lambda_3 = 2, \quad \lambda_4 = i, \quad \lambda_5 = -i.$$

Obviously, these are neither collinear nor concyclic. An elementary computation (that may be performed by any computer algebra package) shows that the unique polynomial p of smallest possible degree that satisfies $p(\lambda_j) = \bar{\lambda}_j$, $j = 1, \dots, 5$, is given by

$$p(z) = \frac{3}{5}z - \frac{3}{5}z^2 + \frac{8}{5}z^3 - \frac{3}{5}z^4,$$

so that $d_p(A) = 4$. Therefore, $d_p(A)/d_r(A) = 2$, showing that the weak upper bound in (3.5) is attained.

Finally, we remark that it may be possible to extend our approach to give an alternative proof of the sharpness of the bound of [12] on the maximal number of zeros of $r(z) - \bar{z}$, where r is rational with $\deg r \geq 2$ (cf. case (2) in Theorem 3.5). For example, let five distinct complex numbers be given, such that any four of them are neither collinear nor concyclic. Then, by case (1) in Theorem 3.1, there exists a unique rational function r of $\deg r = 2$, so that the five complex numbers are zeros of $r(z) - \bar{z}$. This function r can be explicitly computed along the lines of [1, pp. 105–107], and it attains the bound of [12].

Acknowledgments. Thanks to Jurjen Duintjer Tebbens, Daniel Szyld and Petr Tichý for comments that helped to improve the presentation.

REFERENCES

- [1] A. C. ANTOULAS, *Approximation of large-scale dynamical systems*, vol. 6 of Advances in Design and Control, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005. With a foreword by Jan C. Willems.
- [2] A. C. ANTOULAS AND B. D. O. ANDERSON, *On the scalar rational interpolation problem*, IMA J. Math. Control Info., 3 (1986), pp. 61–88.
- [3] T. BARTH AND T. MANTEUFFEL, *Conjugate gradient algorithms using multiple recursions*, in Proceedings of the AMS-IMS-SIAM Summer Research Conference held at the University of Washington, Seattle, WA, July 9–13, 1995, L. Adams and J. L. Nazareth, eds., Philadelphia, 1996, SIAM, pp. 107–123.
- [4] ———, *Multiple recursion conjugate gradient algorithms. I. Sufficient conditions*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 768–796.
- [5] ———, *Multiple recursion conjugate gradient algorithms. II. Necessary conditions*, unpublished manuscript, (2000).
- [6] V. FABER AND T. MANTEUFFEL, *Necessary and sufficient conditions for the existence of a conjugate gradient method*, SIAM J. Numer. Anal., 21 (1984), pp. 352–362.
- [7] L. GEYER, *Sharp bounds for the valence of certain harmonic polynomials*, Tech. Report arXiv:math.CV/0510539, 2005.
- [8] W. B. GRAGG, *Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle*, J. Comput. Appl. Math., 46 (1993), pp. 183–198.
- [9] R. GRONE, C. R. JOHNSON, E. M. DE SÁ, AND H. WOLKOWICZ, *Normal matrices*, Linear Algebra and its Applications, 87 (1987), pp. 213–225.
- [10] R. A. HORN AND C. R. JOHNSON, *Matrix analysis*, Cambridge University Press, Cambridge, 1985.
- [11] C. F. JAGELS AND L. REICHEL, *A fast minimal residual algorithm for shifted unitary matrices*, Numer. Linear Algebra Appl., 1 (1994), pp. 555–570.
- [12] D. KHAVINSON AND G. NEUMANN, *On the number of zeros of certain rational harmonic functions*, Proc. Amer. Math. Soc., 134 (2006), pp. 1077–1085 (electronic).
- [13] D. KHAVINSON AND G. ŚWIĄTEK, *On the number of zeros of certain harmonic polynomials*, Proc. Amer. Math. Soc., 131 (2003), pp. 409–414.
- [14] J. LIESEN AND Z. STRAKOŠ, *On optimal short recurrences for generating orthogonal Krylov subspace bases*, submitted to SIAM Rev., (2006).
- [15] T. NEEDHAM, *Visual complex analysis*, The Clarendon Press Oxford University Press, New York, 1997.
- [16] S. H. RHIE, *n-point gravitational lenses with $5(n - 1)$ images*, Tech. Report arXiv:astro-ph/0305166, 2003.
- [17] D. S. WATKINS, *Some perspectives on the eigenvalue problem*, SIAM Rev., 35 (1993), pp. 430–471.