SIAM J. Matrix Anal. Appl.
Vol. 31, No. 2, pp. 853–863

ON BEST APPROXIMATIONS OF POLYNOMIALS IN MATRICES IN THE MATRIX 2-NORM*

JÖRG LIESEN† AND PETR TICHÝ‡

Abstract. We show that certain matrix approximation problems in the matrix 2-norm have uniquely defined solutions, despite the lack of strict convexity of the matrix 2-norm. The problems we consider are generalizations of the ideal Arnoldi and ideal GMRES approximation problems introduced by Greenbaum and Trefethen [*SIAM J. Sci. Comput.*, 15 (1994), pp. 359–368]. We also discuss general characterizations of best approximation in the matrix 2-norm and provide an example showing that a known sufficient condition for uniqueness in these characterizations is not necessary.

Key words. matrix approximation problems, polynomials in matrices, matrix functions, matrix 2-norm, GMRES, Arnoldi's method

AMS subject classifications. 41A52, 15A60, 65F35

DOI. 10.1137/080728299

1. Introduction.

Much of the work in approximation theory concerns the approximation of a given function f on some (compact) set Ω in the complex plane by polynomials. Classical results in this area deal with the *best approximation problem*

$$\min_{p\in\mathcal{P}_m} \|f-p\|_\Omega\,, \tag{1.1}$$

where $\|g\|_\Omega \equiv \max_{z\in\Omega}|g(z)|$ and $\mathcal{P}_m$ denotes the set of polynomials of degree at most m. (Note that since in (1.1) we seek an approximation from a finite-dimensional subspace, the minimum is indeed attained by some polynomial $p_* \in \mathcal{P}_m$.)

Scalar approximation problems of the form (1.1) have been studied since the mid 1850s. Accordingly, numerous results on the existence and uniqueness of the solution as well as estimates for the value of (1.1) are known. Here we consider a problem that at first sight looks similar, but apparently is much less understood: Let f be a function that is analytic in an open neighborhood of the spectrum of a given matrix $A \in \mathbb{C}^{n\times n}$, so that $f(A)$ is well defined, and let $|\cdot|$ be a given matrix norm. Consider the *matrix* approximation problem

$$\min_{p\in\mathcal{P}_m} |f(A)-p(A)|\,. \tag{1.2}$$

Does this problem have a unique solution?

An answer to this question of course depends on the norm used in (1.2). A norm $|\cdot|$ on a vector space $\mathcal{V}$ is called *strictly convex* when for all vectors $v_1, v_2 \in \mathcal{V}$ the equation $|v_1| = |v_2| = \frac{1}{2}\,|v_1+v_2|$ implies that $v_1 = v_2$. A geometric interpretation of strict convexity is that the unit sphere in $\mathcal{V}$ with respect to the norm $|\cdot|$ does not

*Received by the editors June 24, 2008; accepted for publication (in revised form) by M. Embree May 21, 2009; published electronically July 30, 2009.
http://www.siam.org/journals/simax/31-2/72829.html

†Institute of Mathematics, Technical University of Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany (liesen@math.tu-berlin.de). The work of this author was supported by the Heisenberg Program of the Deutsche Forschungsgemeinschaft.

‡Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, 18207 Prague, Czech Republic (tichy@cs.cas.cz). The work of this author was supported by GAAS grant IAA100300802.

contain any line segments. If $\mathcal{S} \subseteq \mathcal{V}$ is a finite-dimensional subspace, then for any given $v \in \mathcal{V}$ there exists a *unique* $s_* \in \mathcal{S}$ so that

$$|v - s_*| \;=\; \min_{s \in \mathcal{S}} |v - s| \,.$$

A proof of this classical result can be found in most books on approximation theory; see, e.g., [3, Chapter 1]. In particular, if the norm is strictly convex, then (1.2) is guaranteed to have a unique solution as long as the value of (1.2) is positive.

A useful matrix norm that is met in many applications is the matrix 2-norm (or spectral norm), which for a given matrix A is equal to the largest singular value of A. We denote the 2-norm of A by $\|A\|$. This norm is *not* strictly convex, as can be seen from the following simple example: Suppose that we have two matrices $A_1, A_2 \in \mathbb{C}^{n\times n}$ of the form

$$A_1 \;=\; \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}, \qquad A_2 \;=\; \begin{bmatrix} B & 0 \\ 0 & D \end{bmatrix},$$

with $\|A_1\| = \|A_2\| = \|B\| \geq \frac{1}{2}\|C + D\|$. Then $\frac{1}{2}\|A_1 + A_2\| = \|B\|$, but whenever $C \neq D$, we have $A_1 \neq A_2$. Consequently, in the case of the matrix 2-norm, the classical uniqueness result mentioned above does not apply, and our question about the uniqueness of the solution of the matrix approximation problem (1.2) is nontrivial.

It is well known that when the function f is analytic in an open neighborhood of the spectrum of the matrix $A \in \mathbb{C}^{n\times n}$, then $f(A)$ is a well-defined complex $n \times n$ matrix. In fact, $f(A) = p_f(A)$, where p_f is a polynomial that depends on the values and possibly the derivatives of f on the spectrum of A. The recent book of Higham [5] gives an extensive overview of definitions, applications, and computational techniques for matrix functions. Our above question now naturally leads to the following mathematical problem: *Let a polynomial b and a nonnegative integer $m < \deg b$ be given. Determine conditions so that the best approximation problem*

$$\min_{p \in \mathcal{P}_m} \|b(A) - p(A)\| \tag{1.3}$$

has a unique solution, where $\|\cdot\|$ is the matrix 2-norm and $\mathcal{P}_m$ denotes the set of polynomials of degree at most m.

When searching the literature we found a number of results on general characterizations of best approximations in normed linear spaces of matrices, e.g., in [7, 9, 15, 16], but just a few papers related to our specific problem. In particular, Greenbaum and Trefethen consider in [4] the two approximation problems

$$\min_{p \in \mathcal{P}_m} \left\|A^{m+1} - p(A)\right\|, \tag{1.4}$$

$$\min_{p \in \mathcal{P}_m} \|I - Ap(A)\|. \tag{1.5}$$

They state that both (1.4) and (1.5) (for nonsingular A) have a unique minimizer.[1] The problem (1.4) is equal to (1.3) with $b(A) = A^{m+1}$. Because of its relation to the convergence of the Arnoldi method [1] for approximating eigenvalues of A, the uniquely defined monic polynomial $z^{m+1} - p_*$ that solves (1.4) is called the $(m+1)st$

[1] The statement of uniqueness is true, but the proof given in [4], which was later repeated in [14, Chapter 29], contains a small error at the very end. After the error was spotted by Michael Eiermann, it was fixed by Anne Greenbaum in 2005, but the correction has not been published.

ideal Arnoldi polynomial of A. In a paper that is mostly concerned with algorithmic and computational results, Toh and Trefethen [13] call this polynomial the $(m+1)$*st Chebyshev polynomial of* A. The reason for this terminology is the following: When the matrix A is *normal*, i.e., unitarily diagonalizable, problem (1.4) becomes a scalar approximation problem of the form (1.1) with $f(z) = z^{m+1}$ and Ω being the spectrum of A. The resulting monic polynomial is the $(m+1)$st Chebyshev polynomial on this (discrete) set Ω, i.e., the unique monic polynomial of degree $m+1$ with minimal maximum norm on Ω. In this sense, the matrix approximation problem (1.3) we study here can be considered a generalization of the classical scalar approximation problem (1.1). Some further results on Chebyshev polynomials of matrices are given in [11] and [14, Chapter 29].

The quantity (1.5) can be used for bounding the relative residual norm in the GMRES method [8]; for details see, e.g., [10, 12]. Therefore, the uniquely defined polynomial $1 - z\,p_*$ that solves (1.5) is called the $(m+1)$*st ideal GMRES polynomial of* A.

In this paper we show that, despite the lack of strict convexity of the matrix 2-norm, the approximation problem (1.3) as well as a certain related problem that generalizes (1.5) have a unique minimizer. Furthermore, we discuss some of the above-mentioned general characterizations of best approximations with respect to the 2-norm in linear spaces of matrices. On the example of a Jordan block, we show that a sufficient condition for the uniqueness of the best approximation obtained by Ziętak [15] does not hold. We are not aware that such an example for a nonnormal matrix has been given before.

2. Uniqueness results. Let $\ell \geq 0$ and $m \geq 0$ be given integers, and consider a given polynomial b of the form

$$b \;=\; \sum_{j=0}^{\ell+m+1} \beta_j z^j \;\in\; \mathcal{P}_{\ell+m+1}\,.$$

Let us rewrite the approximation problem (1.3) in a more convenient equivalent form:

$$\begin{aligned}\min_{p\in\mathcal{P}_m} \|b(A) - p(A)\| &= \min_{p\in\mathcal{P}_m} \left\| b(A) - \left(p(A) + \sum_{j=0}^{m} \beta_j A^j \right) \right\| \\ &= \min_{p\in\mathcal{P}_m} \left\| \sum_{j=m+1}^{\ell+m+1} \beta_j A^j \;-\; p(A) \right\| \\ &= \min_{p\in\mathcal{P}_m} \left\| A^{m+1} \sum_{j=0}^{\ell} \beta_{j+m+1} A^j \;-\; p(A) \right\|. \end{aligned} \tag{2.1}$$

The polynomials in (2.1) are of the form $z^{m+1}g + h$, where the polynomial $g \in \mathcal{P}_\ell$ is *given* and $h \in \mathcal{P}_m$ is *sought*. Hence (1.3) is equivalent to the problem

$$\min_{h\in\mathcal{P}_m} \left\| A^{m+1} g(A) + h(A) \right\|, \tag{2.2}$$

where $g \in \mathcal{P}_\ell$ is a given polynomial or

$$\min_{p\in\mathcal{G}^{(g)}_{\ell,m}} \|p(A)\|, \;\text{ where }\; \mathcal{G}^{(g)}_{\ell,m} \equiv \left\{ z^{m+1}g + h \,:\, g \in \mathcal{P}_\ell \text{ is given}, h \in \mathcal{P}_m \right\}. \tag{2.3}$$

With $\ell = 0$ and $g = 1$, (2.3) reduces to (1.4).

Similarly, we may consider the approximation problem

$$\min_{p \in \mathcal{H}_{\ell,m}^{(h)}} \|p(A)\|, \quad \text{where } \mathcal{H}_{\ell,m}^{(h)} \equiv \left\{ z^{m+1}g + h \,:\, h \in \mathcal{P}_m \text{ is given}, \; g \in \mathcal{P}_\ell \right\}. \tag{2.4}$$

Setting $m = 0$ and $h = 1$ in (2.4), we retrieve a problem of the form (1.5).

The problems (2.3) and (2.4) are trivial for $g = 0$ and $h = 0$, respectively. Both cases are unconstrained minimizations problems, and it is easily seen that the resulting minimum value is zero. In the following we will therefore exclude the cases $g = 0$ in (2.3) and $h = 0$ in (2.4). Under this assumption, both $\mathcal{G}_{\ell,m}^{(g)}$ and $\mathcal{H}_{\ell,m}^{(h)}$ are subsets of $\mathcal{P}_{\ell+m+1}$, where certain coefficients are *fixed*. In the case of $\mathcal{G}_{\ell,m}^{(g)}$, these are the coefficients at the $\ell + 1$ largest powers of z, namely, $z^{m+1}, \dots, z^{\ell+m+1}$. For $\mathcal{H}_{\ell,m}^{(h)}$ these are the coefficients at the $m + 1$ smallest powers of z, namely, $1, \dots, z^m$.

We start with conditions so that the values of (2.3) and (2.4) are positive for all given nonzero polynomials $g \in \mathcal{P}_\ell$ and $h \in \mathcal{P}_m$, respectively.

Lemma 2.1. *Consider the approximation problems* (2.3) *and* (2.4), *where* $\ell \geq 0$ *and* $m \geq 0$ *are given integers. Denote by* $d(A)$ *the degree of the minimal polynomial of the given matrix* $A \in \mathbb{C}^{n\times n}$. *Then the following two assertions are equivalent:*

(1) $\min_{p\in\mathcal{G}_{\ell,m}^{(g)}} \|p(A)\| > 0$ *for all nonzero polynomials* $g \in \mathcal{P}_\ell$.

(2) $m + \ell + 1 < d(A)$.

If A *is nonsingular, the two assertions are equivalent with*

(3) $\min_{p\in\mathcal{H}_{\ell,m}^{(h)}} \|p(A)\| > 0$ *for all nonzero polynomials* $h \in \mathcal{P}_m$.

Proof. (1) $\Rightarrow$ (2): We suppose that $m + \ell + 1 \geq d(A)$ and show that (1) fails to hold. Denote the minimal polynomial of A by Ψ_A. If $m + 1 \leq d(A) \leq \ell + m + 1$, then there exist uniquely determined polynomials $\widehat{g} \in \mathcal{P}_\ell$, $\widehat{g} \neq 0$, and $\widehat{h} \in \mathcal{P}_m$, so that $z^{m+1} \cdot \widehat{g} + \widehat{h} = \Psi_A$. Hence $\min_{p\in\mathcal{G}_{\ell,m}^{(g)}} \|p(A)\| = 0$ for $g = \widehat{g}$. If $0 \leq d(A) \leq m$, let $\widehat{g}$ be any nonzero polynomial of degree at most ℓ. By the division theorem for polynomials,[2] there exist uniquely defined polynomials $q \in \mathcal{P}_{m+\ell+1-d(A)}$ and $h \in \mathcal{P}_{m-1}$, so that $z^{m+1}\cdot\widehat{g} = q\cdot\Psi_A - h$, or, equivalently, $z^{m+1}\cdot\widehat{g} + h = q\cdot\Psi_A$. Hence $A^{m+1}\widehat{g}(A) + h(A) = 0$, which means that $\min_{p\in\mathcal{G}_{\ell,m}^{(g)}} \|p(A)\| = 0$ for the nonzero polynomial $g = \widehat{g} \in \mathcal{P}_\ell$.

(2) $\Rightarrow$ (1): If $m+\ell+1 < d(A)$, then $\mathcal{G}_{\ell,m}^{(g)} \subset \mathcal{P}_{m+\ell+1}$ implies $\min_{p\in\mathcal{G}_{\ell,m}^{(g)}} \|p(A)\| > 0$ for every nonzero polynomial $g \in \mathcal{P}_\ell$.

(2) $\Rightarrow$ (3): If $m+\ell+1 < d(A)$, then $\mathcal{H}_{\ell,m}^{(h)} \subset \mathcal{P}_{m+\ell+1}$ implies $\min_{p\in\mathcal{H}_{\ell,m}^{(h)}} \|p(A)\| > 0$ for every nonzero polynomial $h \in \mathcal{P}_m$.

(3) $\Rightarrow$ (2): For this implication we use that A is nonsingular. Suppose that (2) does not hold, i.e., that $0 \leq d(A) \leq m + \ell + 1$. Then there exist uniquely defined polynomials $\widehat{g} \in \mathcal{P}_\ell$ and $\widehat{h} \in \mathcal{P}_m$ such that $z^{m+1} \cdot \widehat{g} + \widehat{h} = \Psi_A$. Since A is assumed to be nonsingular, we must have $\widehat{h} \neq 0$. Consequently, $\min_{p\in\mathcal{H}_{\ell,m}^{(h)}} \|p(A)\| = 0$ for the nonzero polynomial $h = \widehat{h} \in \mathcal{P}_m$. $\square$

In the following Theorem 2.2, we show that the problem (2.3) has a uniquely defined minimizer when the value of this problem is positive (and not zero). In the previous lemma we have shown that $m + \ell + 1 < d(A)$ is necessary and sufficient so that the value of (2.3) is positive *for all* nonzero polynomials $g \in \mathcal{P}_\ell$. However, it is possible that *for some* nonzero polynomial $g \in \mathcal{P}_\ell$ the value of (2.3) is positive even when $m+1 \leq d(A) \leq m+\ell+1$. It is possible to further analyze this special case, but

[2]If f and $g \neq 0$ are polynomials over a field $\mathbb{F}$, then there exist uniquely defined polynomials s and r over $\mathbb{F}$ such that (i) $f = g \cdot s + r$, and (ii) either $r = 0$ or $\deg r < \deg g$. If $\deg f \geq \deg g$, then $\deg f = \deg g + \deg s$. For a proof of this standard result, see, e.g., [6, Chapter 4].

for the ease of the presentation we simply assume that the value of (2.3) is positive. The same assumption is made in Theorem 2.3 below, where we prove the uniqueness of the minimizer of (2.4) (under the additional assumption that A is nonsingular).

We point out that Lemma 2.1 implies that the approximation problems (1.4) and (1.5) for nonsingular A have positive values if and only if $m+1 < d(A)$. Of course, if $m+1 = d(A)$, then the value of both problems is zero. In this case, the $(m+1)$st ideal Arnoldi polynomial that solves (1.4) is equal to the minimal polynomial of A, and the $(m+1)$st ideal GMRES polynomial that solves (1.5) is a scalar multiple of that polynomial.

THEOREM 2.2. *Let $A \in \mathbb{C}^{n\times n}$ be a given matrix, $\ell \geq 0$ and $m \geq 0$ be given integers, and $g \in \mathcal{P}_\ell$ be a given nonzero polynomial. If the value of* (2.3) *is positive, then this problem has a uniquely defined minimizer.*

Proof. The general strategy in the following is similar to the construction in [4, section 5]. We suppose that $q_1 = z^{m+1}g + h_1 \in \mathcal{G}^{(g)}_{\ell,m}$ and $q_2 = z^{m+1}g + h_2 \in \mathcal{G}^{(g)}_{\ell,m}$ are two distinct solutions to (2.3) and derive a contradiction. Suppose that the minimal norm attained by the two polynomials is

$$C = \|q_1(A)\| = \|q_2(A)\|.$$

By assumption, $C > 0$. Define $q \equiv \frac{1}{2}(q_1 + q_2) \in \mathcal{G}^{(g)}_{\ell,m}$, then

$$\|q(A)\| \;\leq\; \frac{1}{2}(\|q_1(A)\| + \|q_2(A)\|) = C.$$

Since C is assumed to be the minimal value of (2.3), we must have $\|q(A)\| = C$. Denote the singular value decomposition of $q(A)$ by

$$q(A) \;=\; V\,\mathrm{diag}(\sigma_1,\ldots,\sigma_n)\,W^*\,. \tag{2.5}$$

Suppose that the maximal singular value $\sigma_1 = C$ of $q(A)$ is J-fold, with left and right singular vectors given by $v_1,\ldots,v_J$ and $w_1,\ldots,w_J$, respectively.

It is well known that the 2-norm for vectors $v \in \mathbb{C}^n$, $\|v\| \equiv (v^*v)^{1/2}$, is strictly convex. For each w_j, $1 \leq j \leq J$, we have

$$C = \|q(A)w_j\| \leq \frac{1}{2}\left(\|q_1(A)w_j\| + \|q_2(A)w_j\|\right) \leq C\,,$$

which implies

$$\|q_1(A)w_j\| = \|q_2(A)w_j\| = C, \qquad 1 \leq j \leq J\,.$$

By the strict convexity of the vector 2-norm,

$$q_1(A)w_j = q_2(A)w_j\,, \qquad 1 \leq j \leq J\,.$$

Similarly, one can show that

$$q_1(A)^*v_j = q_2(A)^*v_j, \qquad 1 \leq j \leq J.$$

Thus,

$$(q_2(A) - q_1(A))w_j = 0, \qquad (q_2(A) - q_1(A))^*v_j = 0, \qquad 1 \leq j \leq J. \tag{2.6}$$

By assumption, $q_2 - q_1 \in \mathcal{P}_m$ is a nonzero polynomial. By the division theorem for polynomials (see footnote 2), there exist uniquely defined polynomials s and r, with $\deg s \leq \ell + m + 1$ and $\deg r < \deg (q_2 - q_1) \leq m$ (or $r = 0$), so that

$$z^{m+1} g \;=\; (q_2 - q_1) \cdot s + r\,.$$

Hence we have shown that for the given polynomials $q_2 - q_1$ and g there exist polynomials s and r such that

$$\widetilde{q} \;\equiv\; (q_2 - q_1) \cdot s \;=\; z^{m+1} g - r \;\;\in\;\; \mathcal{G}^{(g)}_{\ell,m}.$$

Since $g \neq 0$, we must have $\widetilde{q} \neq 0$. For a fixed $\epsilon \in (0, 1)$, consider the polynomial

$$q_\epsilon \;=\; (1 - \epsilon) q + \epsilon \widetilde{q} \;\in\; \mathcal{G}^{(g)}_{\ell,m}\,.$$

By (2.6),

$$\widetilde{q}(A) w_j = 0, \qquad \widetilde{q}(A)^* v_j = 0, \qquad 1 \leq j \leq J,$$

and thus

$$\begin{aligned} q_\epsilon(A)^* q_\epsilon(A) w_j &= (1 - \epsilon) q_\epsilon(A)^* q(A) w_j = (1 - \epsilon) C q_\epsilon(A)^* v_j \\ &= (1 - \epsilon)^2 C q(A)^* v_j = (1 - \epsilon)^2 C^2 w_j\,, \end{aligned}$$

which shows that $w_1, \ldots, w_J$ are right singular vectors of $q_\epsilon(A)$ corresponding to the singular value $(1 - \epsilon)C$. Note that $(1 - \epsilon)C < C$ since $C > 0$.

Now there are two cases: Either $\|q_\epsilon(A)\| = (1 - \epsilon)C$, or $(1 - \epsilon)C$ is not the largest singular value of $q_\epsilon(A)$. In the first case we have a contradiction to the fact that C is the minimal value of (2.3). Therefore, the second case must hold. In that case, none of the vectors $w_1, \ldots, w_J$ correspond to the largest singular value of $q_\epsilon(A)$. Using this fact and the singular value decomposition (2.5), we get

$$\begin{aligned} \|q_\epsilon(A)\| &= \|q_\epsilon(A) W\| \\ &= \|q_\epsilon(A) [w_{J+1}, \ldots, w_n]\| \\ &= \|(1 - \epsilon) q(A) [w_{J+1}, \ldots, w_n] \;+\; \epsilon \widetilde{q}(A) [w_{J+1}, \ldots, w_n]\| \\ &\leq (1 - \epsilon)\, \|[v_{J+1}, \ldots, v_n] \operatorname{diag}(\sigma_{J+1}, \ldots, \sigma_n)\| \;+\; \epsilon \|\widetilde{q}(A) [w_{J+1}, \ldots, w_n]\| \\ &\leq (1 - \epsilon) \sigma_{J+1} \;+\; \epsilon \|\widetilde{q}(A) [w_{J+1}, \ldots, w_n]\|\,. \end{aligned} \tag{2.7}$$

Note that the norm $\|\widetilde{q}(A)[w_{J+1}, \ldots, w_n]\|$ in (2.7) does not depend on the choice of ϵ and that (2.7) goes to σ_{J+1} as ϵ goes to zero. Since $\sigma_J > \sigma_{J+1}$, one can find a positive $\epsilon_* \in (0, 1)$ such that (2.7) is less than σ_J for all $\epsilon \in (0, \epsilon_*)$. Any of the corresponding polynomials q_ϵ gives a matrix $q_\epsilon(A)$ whose norm is less than σ_J. This contradiction finishes the proof. ∎

In the following theorem we prove that the problem (2.4), and hence in particular the problem (1.5), has a uniquely defined minimizer.

THEOREM 2.3. *Let $A \in \mathbb{C}^{n \times n}$ be a given nonsingular matrix, $\ell \geq 0$ and $m \geq 0$ be given integers, and $h \in \mathcal{P}_m$ be a given nonzero polynomial. If the value of* (2.4) *is positive, then this problem has a uniquely defined minimizer.*

Proof. Most parts of the following proof are analogous to the proof of Theorem 2.2 and are stated only briefly. However, the construction of the polynomial q_ϵ used to derive the contradiction is different.

We suppose that $q_1 = z^{m+1}g_1+h \in \mathcal{H}^{(h)}_{\ell,m}$ and $q_2 = z^{m+1}g_2+h \in \mathcal{H}^{(h)}_{\ell,m}$ are two distinct solutions to (2.4) and that the minimal norm attained by them is $C = \|q_1(A)\| = \|q_2(A)\|$. By assumption, $C > 0$. Define $q \equiv \frac{1}{2}(q_1 + q_2) \in \mathcal{H}^{(h)}_{\ell,m}$; then $\|q(A)\| = C$. Denote the singular value decomposition of $q(A)$ by $q(A) = V\mathrm{diag}(\sigma_1, \ldots, \sigma_n)W^*$, and suppose that the maximal singular value $\sigma_1 = C$ of $q(A)$ is J-fold, with left and right singular vectors given by $v_1, \ldots, v_J$ and $w_1, \ldots, w_J$, respectively. As previously, we can show that

$$(q_2(A) - q_1(A))w_j = 0, \qquad (q_2(A) - q_1(A))^* v_j = 0, \qquad 1 \le j \le J.$$

Since A is nonsingular and $q_2 - q_1 = z^{m+1}(g_2 - g_1)$, these relations imply that

$$(g_2(A) - g_1(A))w_j = 0, \qquad (g_2(A) - g_1(A))^* v_j = 0, \qquad 1 \le j \le J. \tag{2.8}$$

By assumption, $0 \ne g_2 - g_1 \in \mathcal{P}_\ell$. Hence there exists an integer d, $0 \le d \le \ell$, so that

$$g_2 - g_1 \;=\; \sum_{i=d}^{\ell} \gamma_i z^i\,, \;\text{ with }\; \gamma_d \ne 0\,.$$

Now define

$$\widetilde{g} \;\equiv\; z^{-d}(g_2 - g_1) \;\in\; \mathcal{P}_{\ell-d}\,.$$

By construction, $\widetilde{g}$ is a polynomial with a nonzero constant term. Furthermore, define

$$\widehat{h} \;\equiv\; z^{-m-1-\ell+d}\,h \qquad \text{and} \qquad \widehat{g} \;\equiv\; z^{-\ell+d}\,\widetilde{g}\,.$$

After a formal change of variables $z^{-1} \mapsto y$, we obtain

$$\widehat{h}(y) \;\in\; \mathcal{P}_{m+1+\ell-d} \qquad \text{and} \qquad \widehat{g}(y) \;\in\; \mathcal{P}_{\ell-d} \setminus \mathcal{P}_{\ell-d-1}\,.$$

(Here $\mathcal{P}_{-1} \equiv \emptyset$ in case $d = \ell$.) By the division theorem for polynomials (see footnote 2), there exist uniquely defined polynomials $s(y)$ and $r(y)$ with $\deg s \le m + 1$ (since $\widehat{g} \ne 0$ is of exact degree $\ell - d$) and $\deg r < \ell - d$ (or $r = 0$) such that

$$\widehat{h}(y) \;=\; \widehat{g}(y) \cdot s(y) - r(y)\,.$$

We now multiply the preceding equation by $y^{-m-1-\ell+d}$, which gives

$$y^{-m-1-\ell+d}\,\widehat{h}(y) \;=\; \left(y^{-\ell+d}\widehat{g}(y)\right) \cdot \left(y^{-m-1}s(y)\right) - y^{-m-1}\,\left(y^{-\ell+d}\,r(y)\right)\,.$$

Since $y^{-1} = z$, this equation is equivalent to

$$h \;=\; \widetilde{g} \cdot \widetilde{s} - z^{m+1}\,\widetilde{r},$$

where $\widetilde{s} \in \mathcal{P}_{m+1}$ and $\widetilde{r} \in \mathcal{P}_{\ell-d-1}$. Hence we have shown that for the given polynomials h and $\widetilde{g}$ there exist polynomials $\widetilde{s} \in \mathcal{P}_{m+1}$ and $\widetilde{r} \in \mathcal{P}_{\ell-d-1}$ such that

$$\widetilde{q} \;\equiv\; \widetilde{g} \cdot \widetilde{s} \;=\; z^{m+1}\,\widetilde{r} + h \;\;\in\;\; \mathcal{H}^{(h)}_{\ell,m}.$$

For a fixed $\epsilon \in (0, 1)$, consider

$$q_\epsilon \;=\; (1 - \epsilon)q + \epsilon\widetilde{q} \;\in\; \mathcal{H}^{(h)}_{\ell,m}\,.$$

Since $\widetilde{q} = \widetilde{s} z^{-d}(g_2 - g_1)$, (2.8) implies that

$$\widetilde{q}(A) w_j = 0, \qquad \widetilde{q}(A)^* v_j = 0, \qquad 1 \le j \le J,$$

which can be used to show that

$$q_\epsilon(A)^* q_\epsilon(A) w_j = (1-\epsilon)^2 C^2 w_j, \qquad 1 \le j \le J.$$

Now the same argument as in the proof of Theorem 2.2 gives a contradiction to the original assumption that $q_2 \neq q_1$. □

Remark 2.4. Similarly as in Lemma 2.1, the assumption of nonsingularity in the previous theorem is in general necessary. In other words, when A is singular the approximation problem (2.4) might have more than one solution even when the value of (2.4) is positive. The following example demonstrating this fact was pointed out to us by Ziętak: Consider a normal matrix $A = U \Lambda U^*$, where $U^* U = I$ and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Suppose that A is singular with n distinct eigenvalues, and $\lambda_1 = 0$. Furthermore, suppose that $h \in \mathcal{P}_m$ is any given polynomial that satisfies $h(0) \neq 0$ and $|h(0)| > |h(\lambda_j)|$ for $j = 2, \ldots, n$. Then for *any* integer $\ell \ge 0$,

$$\min_{p \in \mathcal{H}_{\ell,m}^{(h)}} \|p(A)\| = \min_{g \in \mathcal{P}_\ell} \max_j \left| \lambda_j^{m+1} g(\lambda_j) + h(\lambda_j) \right| \; = \; |h(0)| \; > \; 0\,.$$

One solution of this problem is given by the polynomial $g = 0$. Moreover, the minimum value is attained for any polynomial $g \in \mathcal{P}_\ell$ that satisfies

$$\min_{g \in \mathcal{P}_\ell} \max_{2 \le j \le n} \left| \lambda_j^{m+1} g(\lambda_j) + h(\lambda_j) \right| \le |h(0)|\,,$$

i.e., for any polynomial $g \in \mathcal{P}_\ell$ that is close enough to the zero polynomial.

3. Characterization of best approximation with respect to the matrix 2-norm.

In this section we discuss general characterizations of best approximation in linear spaces of matrices with respect to the matrix 2-norm obtained by Ziętak [15, 16], and we give an example from our specific problem. To state Ziętak's results, we need some notation. Suppose that we are given m matrices $A_1, \ldots, A_m \in \mathbb{C}^{n \times n}$ that are linearly independent in $\mathbb{C}^{n \times n}$. We assume that $1 \le m < n^2$ to avoid trivialities. Denote $\mathbb{A} \equiv \operatorname{span}\{A_1, \ldots, A_m\}$, which is an m-dimensional subspace of $\mathbb{C}^{n \times n}$. As above, let $\|\cdot\|$ denote the matrix 2-norm. For a given matrix $B \in \mathbb{C}^{n \times n} \backslash \mathbb{A}$, we consider the best approximation (or matrix nearness) problem

$$\min_{M \in \mathbb{A}} \|B - M\|\,. \tag{3.1}$$

A matrix $A_* \in \mathbb{A}$ for which this minimum is achieved (such a matrix exists, since $\mathbb{A}$ is finite dimensional) is called a *spectral approximation of* B *from the subspace* $\mathbb{A}$. The corresponding matrix $R(A_*) = B - A_*$ is called a *residual matrix*.

The approximation problems (2.3) and (2.4) studied in the previous section are both special cases of (3.1). In the case of (2.3),

$$B = A^{m+1} g(A), \;\text{ where } g \in \mathcal{P}_\ell \text{ is given and } \;\; \mathbb{A} = \{I, A, \ldots, A^m\}\,,$$

while in case of (2.4),

$$B = h(A), \;\text{ where } h \in \mathcal{P}_m \text{ is given and } \;\; \mathbb{A} = \left\{A^{m+1}, \ldots, A^{\ell+m+1}\right\}.$$

We have shown that when the values of these approximation problems are positive (which is true if $\ell + m + 1 < d(A)$), for both these problems there exists a uniquely defined spectral approximation A_* of B from the subspace $\mathbb{A}$ (in the case of (2.4), we have assumed that A is nonsingular). Another approximation problem that fits into the template (3.1) arises in the convergence theory for Arnoldi eigenvalue iterations in [2], where the authors study the problem of minimizing $\|I - h(A)p(A)\|$ over polynomials $p \in \mathcal{P}_{\ell-2m}$, $\ell \geq 2m \geq 2$, and $h \in \mathcal{P}_m$ is a given polynomial.

In general, the spectral approximation of a matrix $B \in \mathbb{C}^{n\times n}$ from a subspace $\mathbb{A} \subset \mathbb{C}^{n\times n}$ is not unique. Ziętak [15] studies the problem (3.1) and gives a general characterization of spectral approximations based on the singular value decomposition of the residual matrices. In particular, combining results of [16] with [15, Theorem 4.3] yields the following sufficient condition for uniqueness of the spectral approximation.

LEMMA 3.1. *In the notation established above, let A_* be a spectral approximation of B from the subspace $\mathbb{A}$. If the residual matrix $R(A_*) = B - A_*$ has an n-fold singular value, then the spectral approximation A_* of B from the subspace $\mathbb{A}$ is unique.*

It is quite obvious that the sufficient condition in Lemma 3.1 is, in general, not necessary. To construct a nontrivial counterexample, we recall that the dual norm to the matrix 2-norm is the trace norm (also called energy norm or c_1-norm)

$$|||\, M \,||| \equiv \sum_{j=1}^{r} \sigma_j(M)\,, \tag{3.2}$$

where $\sigma_1(M), \ldots, \sigma_r(M)$ denote the singular values of the matrix $M \in \mathbb{C}^{n\times n}$ with $\operatorname{rank}(M) = r$. For $X \in \mathbb{C}^{n\times n}$ and $Y \in \mathbb{C}^{n\times n}$ we define the inner product $\langle X, Y\rangle \equiv \operatorname{trace}(Y^* X)$. Using this notation, we can state the following result, which is given in [16, p. 173].

LEMMA 3.2. *The matrix $A_* \in \mathbb{A}$ is a spectral approximation of B from the subspace $\mathbb{A}$ if and only if there exists a matrix $Z \in \mathbb{C}^{n\times n}$, with $|||\, Z \,||| = 1$ such that*

$$\langle Z, X\rangle = 0 \ \textit{ for all } \ X \in \mathbb{A} \qquad \textit{and} \qquad \operatorname{Re}\,\langle Z, B - A_*\rangle = \|B - A_*\|\,. \tag{3.3}$$

Remark 3.3. Lemmas 3.1 and 3.2 are both stated here for square complex matrices. Originally, Lemma 3.1 is formulated in [15] for real rectangular matrices and Lemma 3.2 given in [16] for square complex matrices. A further generalization to rectangular complex matrices seems possible, but it is out of our focus here.

Based on Lemma 3.2 we can prove the following result.

THEOREM 3.4. *For $\lambda \in \mathbb{C}$, consider the $n \times n$ Jordan block*

$$J_\lambda \ \equiv \ \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}.$$

Then for any nonnegative integer m with $m+1 \leq n$, the solution to the approximation problem (1.4) with $A = J_\lambda$, i.e., the $(m+1)$st ideal Arnoldi (or Chebyshev) polynomial of J_λ, is uniquely defined and given by $(z-\lambda)^{m+1}$.

Proof. With $A = J_\lambda$, the approximation problem (1.4) reads

$$\min_{p \in \mathcal{P}_m} \left\| J_\lambda^{m+1} - p(J_\lambda) \right\|. \tag{3.4}$$

In the notation established in this section, we seek a spectral approximation A_* of $B = J_\lambda^{m+1}$ from the subspace $\mathbb{A} = \operatorname{span}\{I, J_\lambda, \dots, J_\lambda^m\}$. We claim that the uniquely defined solution is given by the matrix $A_* = J_\lambda^{m+1} - (J_\lambda - \lambda I)^{m+1}$. For this matrix A_* we get

$$B - A_* = J_\lambda^{m+1} - A_* = (J_\lambda - \lambda I)^{m+1} = J_0^{m+1}.$$

For $m+1 = n$, $A_* = J_\lambda^n - (J_\lambda - \lambda I)^n = J_\lambda^n$ yields $B - A_* = J_0^n = 0$. The corresponding ideal Arnoldi polynomial of J_λ is uniquely defined and equal to $(z-\lambda)^n$, the minimal polynomial of J_λ.

For $m+1 < n$, the value of (3.4) is positive, and hence Theorem 2.2 ensures that the spectral approximation of J_λ^{m+1} from the subspace $\mathbb{A}$ is uniquely defined. We prove our claim using Lemma 3.2. Define $Z \equiv e_1 e_{m+2}^T$ then $|||Z||| = 1$,

$$\left\langle Z, J_\lambda^j \right\rangle = 0 \quad \text{for } j = 0, \dots, m,$$

and $\|B - A_*\| = \|J_0^{m+1}\| = 1$, so that

$$\langle Z, B - A_* \rangle = \left\langle Z, J_0^{m+1} \right\rangle = 1 = \|B - A_*\|,$$

which shows (3.3) and completes the proof. ∎

The proof of this theorem shows that the residual matrix of the spectral approximation A_* of $B = J_\lambda^{m+1}$ from the subspace $\mathbb{A} = \operatorname{span}\{I, J_\lambda, \dots, J_\lambda^m\}$ is given by $R(A_*) = J_0^{m+1}$. This matrix $R(A_*)$ has $m+1$ singular values equal to zero and $n-m-1$ singular values equal to one. Hence, for $m+1 < n$, the maximal singular value of the residual matrix is not n-fold, and the sufficient condition of Lemma 3.1 does not hold. Nevertheless, the spectral approximation of B from the subspace $\mathbb{A}$ is unique whenever $m+1 < n$.

As shown above, for $m = 0, 1, \dots, n-1$, the polynomial $(z-\lambda)^{m+1}$ solves the ideal Arnoldi approximation problem (1.4) for $A = J_\lambda$. For $\lambda \neq 0$, we can write

$$(z-\lambda)^{m+1} = (-\lambda)^{m+1} \cdot \left(1 - \lambda^{-1} z\right)^{m+1}.$$

Note that the rightmost factor is a polynomial that has value one at the origin. Hence it is a candidate for the solution of the ideal GMRES approximation problem (1.5) for $A = J_\lambda$. More generally, it is tempting to assume that the $(m+1)$st ideal GMRES polynomial for a given matrix A is equal to a scaled version of its $(m+1)$st ideal Arnoldi (or Chebyshev) polynomial. However, this assumption is false, as we can already see in case $A = J_\lambda$. As shown in [10], the determination of the ideal GMRES polynomials for a Jordan block is an intriguing problem, since these polynomials can become quite complicated. They are of the simple form $(1 - \lambda^{-1} z)^{m+1}$ if and only if $0 \le m+1 < n/2$ and $|\lambda| \ge \varrho_{m+1,n-m-1}^{-1}$; cf. [10, Theorem 3.2]. Here $\varrho_{k,n}$ denotes the radius of the polynomial numerical hull of degree k of an $n \times n$ Jordan block (this radius is independent of the eigenvalue λ).

Now let n be even, and consider $m+1 = n/2$. If $|\lambda| \le 2^{-2/n}$, the ideal GMRES polynomial of degree $n/2$ of J_λ is equal to the constant polynomial 1. If $|\lambda| \ge 2^{-2/n}$, the ideal GMRES polynomial of degree $n/2$ of J_λ is equal to

$$\frac{2}{4\lambda^n + 1} + \frac{4\lambda^n - 1}{4\lambda^n + 1} \left(1 - \lambda^{-1} z\right)^{n/2}; \tag{3.5}$$

cf. [10, p. 465]. Obviously, neither the polynomial 1 nor the polynomial (3.5) are scalar multiples of $(z-\lambda)^{n/2}$, the ideal Arnoldi polynomial of degree $n/2$ of J_λ.

Acknowledgments. We thank Krystyna Ziętak for many discussions and suggestions that helped to improve the content of this paper. We also thank Mark Embree, Shmuel Friedland, Anne Greenbaum, Nick Trefethen, and two anonymous referees for their helpful comments.

REFERENCES

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

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

[3] E. W. Cheney, *Introduction to Approximation Theory*, McGraw-Hill, New York, 1966.

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

[5] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.

[6] K. Hoffman and R. Kunze, *Linear Algebra*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1971.

[7] K. K. Lau and W. O. J. Riha, *Characterization of best approximations in normed linear spaces of matrices by elements of finite-dimensional linear subspaces*, Linear Algebra Appl., 35 (1981), pp. 109–120.

[8] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[9] I. Singer, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, Springer-Verlag, Berlin, 1970.

[10] P. Tichý, J. Liesen, and V. Faber, *On worst-case GMRES, ideal GMRES, and the polynomial numerical hull of a Jordan block*, Electron. Trans. Numer. Anal., 26 (2007), pp. 453–473 (electronic).

[11] K.-C. Toh, *Matrix Approximation Problems and Nonsymmetric Iterative Methods*, Ph.D. thesis, Cornell University, Ithaca, NY, 1996.

[12] K.-C. Toh, *GMRES vs. ideal GMRES*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 30–36.

[13] K.-C. Toh and L. N. Trefethen, *The Chebyshev polynomials of a matrix*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 400–419.

[14] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra*, Princeton University Press, Princeton, NJ, 2005.

[15] K. Ziętak, *Properties of linear approximations of matrices in the spectral norm*, Linear Algebra Appl., 183 (1993), pp. 41–60.

[16] K. Ziętak, *On approximation problems with zero-trace matrices*, Linear Algebra Appl., 247 (1996), pp. 169–183.