SIAM J. Matrix Anal. Appl.
Vol. 38, No. 1, pp. 96–117
© 2017 Society for Industrial and Applied Mathematics

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THE COEFFICIENTS OF THE FOM AND GMRES RESIDUAL POLYNOMIALS*

GÉRARD MEURANT†

Abstract. In this paper we derive closed-form expressions for the coefficients of the residual and characteristic polynomials in the full orthogonalization method and GMRES iterative Krylov method for solving linear systems with diagonalizable matrices. The coefficients are given as functions of the eigenvalues and eigenvectors of the matrix A and of the right-hand side b. These results yield the residual vectors and the explicit solution of the optimization problem $\min_{p\in\pi_k} \|p(A)b\|$, where π_k is the set of polynomials of degree k with a value 1 at the origin. In addition, the Ritz values and harmonic Ritz values can be written explicitly for the first four iterations of the Arnoldi algorithm. Moreover, from the coefficients of the characteristic polynomials, we obtain lower bounds for the distances of the eigenvalues of A to the Ritz and harmonic Ritz values.

Key words. FOM and GMRES algorithms, Ritz values, harmonic Ritz values, residual polynomials, normal matrices

AMS subject classifications. 65F10, 65F15, 15A18

DOI. 10.1137/15M1053827

1. Introduction. This paper is mainly concerned with the polynomials generated by the FOM (full orthogonalization method) and GMRES (generalized minimum residual method) iterative Krylov methods for solving nonsymmetric linear systems $Ax = b$, where A is a nonsingular real or complex matrix of order n and b is a given vector; see [27, 28, 29, 30]. These two methods use the Arnoldi process [3] to compute an orthonormal basis of the Krylov subspace. We are also interested in the convergence of the Ritz values which are the approximations of the eigenvalues of A generated by the Arnoldi process; see [31]. We will assume for simplicity that the Arnoldi process started with b as the initial vector does not stop before step n and that $\|b\| = 1$. Given A and b, the Krylov matrix is defined as

$$K = \begin{pmatrix} b & Ab & A^2b & \cdots & A^{n-1}b \end{pmatrix}.$$

The columns of K span the Krylov space $\mathcal{K}_n(A, b)$. Similarly we define the Krylov subspaces $\mathcal{K}_j(A, b)$ which are spanned by $b, Ab, \ldots, A^{j-1}b$ for $j < n$. Starting from $v_1 = b$ the Arnoldi process computes a unitary matrix V with columns v_i and $v_1 = Ve_1 = b$ (e_j being the jth column of the identity matrix) and an upper Hessenberg matrix H with positive real subdiagonal entries $h_{j+1,j}$, $j = 1, \ldots, n-1$, such that

$$AV = VH.$$

This corresponds to the choice of a zero starting vector $x_0 = 0$ for FOM and GMRES. Let $V_{n,k}$ be the matrix of the k first columns of V, $k < n$. The relation satisfied by $V_{n,k}$ is

$$AV_{n,k} = V_{n,k}H_k + h_{k+1,k}v_{k+1}e_k^T, \tag{1.1}$$

where H_k is the principal matrix of order k of H and e_k is the last column of the identity matrix of order k. Note that $H_k = V_{n,k}^* AV_{n,k}$. Relation (1.1) shows how

*Received by the editors December 21, 2015; accepted for publication (in revised form) by V. Simoncini November 23, 2016; published electronically January 24, 2017.
http://www.siam.org/journals/simax/38-1/M105382.html
†30 rue du sergent Bauchat, 75012 Paris, France (gerard.meurant@gmail.com).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

to compute the next column v_{k+1} of the matrix V and the kth column of H by orthogonalizing v_{k+1} against the previous columns of V. We define the iterates x_k as a linear combination of the k first basis vectors,

$$x_k = V_{n,k} y_k, \quad k = 1, 2, \ldots.$$

since $x_0 = 0$. The residual vector r_k is defined as usual as $r_k = b - Ax_k$. FOM is defined (provided that H_k is nonsingular) by computing y_k as the solution of the linear system

$$H_k y_k = e_1.$$

In GMRES the vector y_k is computed as the solution of the minimization problem

$$\min_y \|e_1 - \underline{H}_k y\|,$$

where $\underline{H}_k$ is the matrix H_k appended with the k first entries of the $k+1$st row of H. This minimizes the norm of the residual r_k at iteration k.

It is easy to see that the residual vectors r_k^F in FOM and r_k^G in GMRES can be expressed as polynomials of degree k in A applied to the initial residual b,

$$r_k^F = p_k^F(A)b, \quad r_k^G = p_k^G(A)b.$$

These polynomials are so-called residual polynomials and they have a value 1 at 0. The GMRES residual polynomial $p_k^G(A)$ is the solution of a minimization problem

$$\|r_k^G\| = \min_{p \in \pi_k} \|p(A)b\|, \tag{1.2}$$

where π_k is the set of polynomials of degree k with a value 1 at the origin.

Most studies on GMRES convergence use the upper bound

$$\|r_k^G\| \leq \left(\min_{p \in \pi_k} \|p(A)\| \right) \|b\|$$

that decouples the matrix polynomial from the right-hand side b. The polynomial giving the minimum in the right-hand side of the inequality is known as the *ideal* GMRES polynomial; see [11, 8]. It is different from the residual polynomial.

However, a closed-form expression of the norm of the GMRES residual was provided in [23] as a function of the eigenvalues and eigenvectors of A and the right-hand side b when A is diagonalizable. The main contribution of the present paper is to obtain closed-form expressions for the coefficients of the polynomial which yields the solution of the minimization problem in relation (1.2). This solution gives an exact expression for the GMRES residual vectors as functions of the eigenvalues, the eigenvectors, and the right-hand side.

The Ritz values $\theta_j^{(k)}$ at iteration k, which are the roots of the polynomial p_k^F, are the eigenvalues of the upper Hessenberg matrix H_k and therefore the roots of its characteristic polynomial. Hence, the FOM residual polynomial is proportional to the characteristic polynomial of H_k which is a monic polynomial. To simplify the notation in the following we will drop the upper index (k) since we will only consider iteration k.

The goal of introducing the harmonic Ritz values $\zeta_j^{(k)} = \zeta_j$ at iteration k is to provide better approximations of the interior eigenvalues. A parameter is used to shift the eigenvalues of interest to the exterior of the spectrum. Here, for simplicity

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

of notation, we will use a zero shift. Then, the harmonic Ritz values are defined as the reciprocals of the standard Ritz values obtained from the Arnoldi process applied to the matrix A^{-1} and the starting vector Ab; see [25] and the earlier paper [26] for symmetric matrices. They can also be seen as the eigenvalues of the matrix

$$\hat{H}_k = H_k + h_{k+1,k}^2 H_k^{-*} e_k e_k^T$$

when H_k is nonsingular. It was proven in [9] that they are the roots of the GMRES residual polynomial p_k^G which is proportional to the characteristic polynomial of $\hat{H}_k$.

In the past, very ingenious methods had been proposed for computing the coefficients of characteristic polynomials; see [13]. The goal was to compute eigenvalues of matrices as roots of their characteristic polynomials. Of course, today we do not any longer compute eigenvalues in this way, the approximations of eigenvalues of nonsymmetric matrices are computed as the eigenvalues of H_k or $\hat{H}_k$. Here we are interested in a different problem. We will obtain closed-form expressions for the coefficients of the characteristic polynomials of H_k and $\hat{H}_k$ as functions of the eigenvalues, the eigenvectors of A, and the right-hand side b. We will consider the case where the matrix A is diagonalizable and give particular attention to the case where A is normal. As a by-product of our study, the closed-form expressions of the coefficients of the residual polynomials will give the locations of the Ritz and harmonic Ritz values for the first four iterations of the Arnoldi algorithm, for which we can explicitly write down the roots from the coefficients. Moreover, we will also obtain lower bounds for the distances of Ritz values to eigenvalues by using bounds which are available in the literature for roots of polynomials as functions of their coefficients.

The contents of the paper are as follows. Section 2 is devoted to the FOM polynomials. We compute the coefficients of a new polynomial whose roots are the Ritz values. This polynomial is not monic nor with a value 1 at the origin but the characteristic polynomial of H_k is obtained dividing by the coefficients of the monomial of highest power and the residual polynomial is obtained dividing by the value at the origin. Moreover, the characterization of the residual vectors leads to a simplified way of constructing linear systems with a prescribed convergence curve for FOM or GMRES, prescribed eigenvalues and Ritz values. We also study some relations of the coefficients with the entries of the inverses of the matrices M_k and M_{k+1}, where $M_k = K_{n,k}^* K_{n,k}$, $K_{n,k}$ is the $n \times k$ Krylov matrix (the first k columns of K). In section 3 we do the same for the GMRES polynomials. In section 4, as an illustration of what can be obtained from the polynomial coefficients, we consider the Ritz and harmonic Ritz values for $k = 2$. Explicit formulas for the Ritz and harmonic Ritz values can also be written for $k = 3, 4$ but, of course, the formulas are more involved. Section 5 considers a small example with a normal matrix. In section 6 we explain how to use the previous results to obtain lower bounds for the distances of an eigenvalue of A to Ritz or harmonic Ritz values. This is illustrated by a small numerical example with a normal matrix. Finally, we give some conclusions.

2. The FOM residual polynomial. The matrix H can be factored as $H = UCU^{-1}$, where U is upper triangular with real positive entries on the diagonal and C is the companion matrix of the eigenvalues of A; see, for instance, [21]. Note that $h_{j+1,j} = \frac{u_{j+1,j+1}}{u_{j,j}}$, $j = 1, \ldots, n-1$, $u_{i,j}$ denoting the entries of U. The matrix U is linked to the Krylov matrix K and the matrix of the basis vectors V by $K = VU$. From [7] we have that

$$H_k = U_k C^{(k)} U_k^{-1},$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

where U_k is the principal submatrix of order k of U and $C^{(k)}$ is the companion matrix of the Ritz values at iteration k. Let us denote $M_k = U_k^* U_k = K_{n,k}^* K_{n,k}$.

Let A be diagonalizable with $A = X\Lambda X^{-1}$. The matrix $K_{n,k}$ can be factorized as

$$K_{n,k} = \begin{pmatrix} b, & Ab, & \ldots, & A^{k-1}b \end{pmatrix} = X \begin{pmatrix} c, & \Lambda c, & \ldots, & \Lambda^{k-1}c \end{pmatrix} = X D_c \mathcal{V}_{n,k},$$

where D_c is a diagonal matrix with the components of $c = X^{-1}b$ on the diagonal and

$$\mathcal{V}_{n,k} = \begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{k-1} \end{pmatrix}$$

is an $n \times k$ Vandermonde matrix.

2.1. Normal matrices. Let us assume that A is normal with $A = X\Lambda X^*$. Then $c = X^*b$. The eigenvalues θ_j of H_k are the roots of its characteristic polynomial. Therefore we have,

$$\begin{aligned} 0 = \det(H_k - \theta I) &= \det(V_{n,k}^* A V_{n,k} - \theta I), \\ &= \det(U_k^{-*} K_{n,k}^* A K_{n,k} U_k^{-1} - \theta I), \\ &= \det(U_k^{-*}[K_{n,k}^* A K_{n,k} - \theta U_k^* U_k] U_k^{-1}), \\ &= \det(M_k^{-1}) \det(K_{n,k}^* X\Lambda X^* K_{n,k} - \theta M_k), \\ &= \frac{1}{\det(M_k)} \det(\mathcal{V}_{n,k}^* D_c^* \Lambda D_c \mathcal{V}_{n,k} - \theta \mathcal{V}_{n,k}^* D_c^* D_c \mathcal{V}_{n,k}). \end{aligned}$$

Clearly $p_k(\theta) = \det(\mathcal{V}_{n,k}^* D_c^* \Lambda D_c \mathcal{V}_{n,k} - \theta \mathcal{V}_{n,k}^* D_c^* D_c \mathcal{V}_{n,k})$ is a polynomial of degree k in θ. We see that the two polynomials on the left- and right-hand sides are the same up to a constant factor. We just have to know when M_k is singular. This can only happen when at least one diagonal entry of U_k is zero, but this corresponds to H_k being reducible.

The following theorem gives a closed-form expression of the polynomial p_k.

THEOREM 2.1. *Let A be a normal matrix with the spectral factorization $A = X\Lambda X^*$. Then, the polynomial p_k whose roots are the Ritz values at iteration k is given by*

$$p_k(\theta) = \sum_{\mathcal{I}_k} (\lambda_{i_1} - \theta) \cdots (\lambda_{i_k} - \theta) \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2, \tag{2.1}$$

where the summation is over all sets $\mathcal{I}_k$ of k indices $\begin{pmatrix} i_1, & i_2, & \ldots, & i_k \end{pmatrix}$ such that $1 \le i_1 < \cdots < i_k \le n$ and $c = X^*b$. In the product the indices i_p and i_q belong to the set of indices $\mathcal{I}_k$.

The characteristic polynomial $\tilde{p}_k^F$ of H_k (which is monic) is obtained by dividing p_k by

$$(-1)^k \sum_{\mathcal{I}_k} |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2 = (-1)^k \det(M_k).$$

The residual polynomial p_k^F for FOM is obtained by dividing p_k by

$$\sum_{\mathcal{I}_k} \lambda_{i_1} \cdots \lambda_{i_k} \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

if this quantity is different from zero. If it is zero, then $p_k(0) = 0$, the matrix H_k is singular, and the FOM iterate is not defined.

Proof. To simplify the notation, let $B = D_c \mathcal{V}_{n,k}$. Then, the Ritz values are solutions of the equation

$$\det(B^* \Lambda B - \theta B^* B) = \det(B^*(\Lambda - \theta I)B) = 0.$$

Let $G = (\Lambda - \theta I)B$; we have $\det(B^* G) = 0$. We need to compute the determinant of the product of two rectangular matrices which are, respectively, $k \times n$ and $n \times k$. We can use the Cauchy–Binet formula (see [10]) to get

$$\det(B^* G) = \sum_{\mathcal{I}_k} \overline{\det(B_{\mathcal{I}_k,:})} \det(G_{\mathcal{I}_k,:}),$$

where $\mathcal{I}_k$ is defined above and $B_{\mathcal{I}_k,:}$ (resp., $G_{\mathcal{I}_k,:}$) is the submatrix of B (resp., G) constructed with the rows in $\mathcal{I}_k$ and all the k columns. We have

$$\det(B_{\mathcal{I}_k,:}) = c_{i_1} \cdots c_{i_k} \prod_{i_1 \le i_p < i_q \le i_k} (\lambda_{i_q} - \lambda_{i_p}),$$

and, since $\Lambda - \theta I$ is diagonal,

$$\det(G_{\mathcal{I}_k,:}) = (\lambda_{i_1} - \theta) \cdots (\lambda_{i_k} - \theta)\, c_{i_1} \cdots c_{i_k} \prod_{i_1 \le i_p < i_q \le i_k} (\lambda_{i_q} - \lambda_{i_p}). \tag{2.2}$$

Therefore, we have the new polynomial

$$\det(B^* G) = \sum_{\mathcal{I}_k} (\lambda_{i_1} - \theta) \cdots (\lambda_{i_k} - \theta)\, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2,$$

whose roots are the Ritz values at iteration k. We remark that the product $(\lambda_{i_1} - \theta) \cdots (\lambda_{i_k} - \theta)$ is a polynomial of degree k in θ whose roots are k of the eigenvalues of A. Hence, p_k is a sum of polynomials whose roots are eigenvalues whose indices are given by all possible ordered subsets of k elements of $\{1, \dots, n\}$. We have

$$(\lambda_{i_1} - \theta) \cdots (\lambda_{i_k} - \theta) = \sum_{j=0}^{k} (-1)^{k-j} e_{(j)}(\lambda_{i_1}, \dots, \lambda_{i_k}) \theta^{k-j},$$

where $e_{(j)}$ is a symmetric elementary polynomial that is

$$e_{(j)}(\mu_1, \dots, \mu_k) = \sum_{1 \le j_1 < j_2 < \cdots < j_j \le k} \mu_{j_1} \cdots \mu_{j_j}, \quad e_{(0)} \equiv 1.$$

The new polynomial is not monic nor with a value 1 at 0, but it is a constant multiple of the characteristic polynomial of H_k. If we want to obtain a monic polynomial (the characteristic polynomial of H_k) we have to divide by the coefficient of θ^k which is

$$(-1)^k \sum_{\mathcal{I}_k} |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2 = (-1)^k \det(M_k).$$

If we want a value 1 at the origin (the FOM residual polynomial p_k^F) we have to divide by the constant coefficient which is

$$\sum_{\mathcal{I}_k} \lambda_{i_1} \cdots \lambda_{i_k}\, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

provided that it is nonzero. The other coefficients involve the symmetric elementary polynomials; the coefficient of θ^{k-j}, $j = 0, \dots, k$, is

$$(-1)^{k-j} \sum_{\mathcal{I}_k} e_{(j)}(\lambda_{i_1}, \dots, \lambda_{i_k}) \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2.$$

□

According to Theorem 2.1 the FOM residual vector at iteration k is

$$\begin{aligned} r_k^F &= p_k^F(A)b, \\ &= X p_k^F(\Lambda) X^* b, \\ &= \sum_{j=1}^n c_j p_k^F(\lambda_j) x^{(j)}, \end{aligned}$$

where $x^{(j)}$ is the column of X corresponding to the eigenvalue λ_j and $c = X^*b$ is the vector whose components are the projections of b on the eigenvectors. We also have $(x^{(j)})^T r_k^F = p_k^F(\lambda_j) c_j$. Let

$$\mu_0 = \sum_{\mathcal{I}_k} \lambda_{i_1} \cdots \lambda_{i_k} \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2 \neq 0.$$

The value of the polynomial p_k^F at the eigenvalue λ_j is

$$p_k^F(\lambda_j) = \frac{1}{\mu_0} \sum_{\hat{\mathcal{I}}_k} (\lambda_{i_1} - \lambda_j) \cdots (\lambda_{i_k} - \lambda_j) \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2,$$

where the sum is over all ordered set of indices $\begin{pmatrix} i_1, & i_2, & \dots, & i_k \end{pmatrix}$ with $i_\ell \neq j$, $\ell = 1, \dots, k$. Hence, $p_k^F(\lambda_j)$ depends on the distances of λ_j to all the other eigenvalues and also on the pairwise distances of these other eigenvalues.

Note that the knowledge of the coefficients of the polynomial p_k^F gives the decomposition of the residual vector r_k^F on the natural basis of the Krylov subspace $\begin{pmatrix} b & Ab & \cdots & A^k b \end{pmatrix}$. In fact, the coefficients of the residual polynomial allow us to characterize the nonzero entries of the inverse of the matrix U. It its well known that the orthonormal basis vectors are linked to the FOM residual vectors by

$$r_k^F = -h_{k+1,k} (y_k)_k v_{k+1}.$$

Hence, the residual vectors are proportional to the basis vectors. Let us assume that we have real data and that H_k is nonsingular for all $k = 1, \dots, n$. Since $K = VU$, we have

$$V = KU^{-1} \Rightarrow v_{k+1} = KU^{-1} e_{k+1} = \pm \frac{1}{\|r_k^F\|} p_k^F(A) b.$$

Let $p_k^F(A)b = \xi_k^{(k)} A^k b + \cdots + \xi_1^{(k)} Ab + b$. By identification of the coefficients, we obtain

$$U^{-1} e_{k+1} = \pm \frac{1}{\|r_k^F\|} \begin{pmatrix} 1 \\ \xi_1^{(k)} \\ \vdots \\ \xi_k^{(k)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

In particular, we see that the entry $(U^{-1})_{1,k+1}$ on the first row is equal to the inverse of the FOM residual norm at iteration k. This fact was already proved differently in [7, Theorem 3.2] for any Q-OR method. The previous results are also valid when A is diagonalizable and not normal.

We can also write $U^{-1} = \hat{U}^{-1} D_r^{-1}$ with

$$D_r = \begin{pmatrix} 1 & & & \\ & \|r_1^F\| & & \\ & & \ddots & \\ & & & \|r_{n-1}^F\| \end{pmatrix} S, \quad \hat{U}^{-1} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ & \xi_1^{(1)} & \cdots & \xi_1^{(n-1)} \\ & & \ddots & \vdots \\ & & & \xi_{n-1}^{(n-1)} \end{pmatrix},$$

where S is a diagonal matrix with ± 1 on the diagonal. Using this result we can solve the problem of constructing a matrix and a right-hand side with prescribed eigenvalues, a prescribed FOM convergence curve, and prescribed Ritz values (provided they are different from zero). This problem was considered and solved in [6]; see also [2] for an early paper describing how to prescribe the GMRES convergence curve. But, the following construction is a little bit simpler. Prescribing the Ritz values, we can compute the coefficients $\xi_i^{(k)}$ of the polynomials p_k^F for $k = 1, \ldots, n-1$ and construct the matrix $\hat{U}^{-1}$. Prescribing the FOM residual norms (or, equivalently, the GMRES residual norms provided GMRES does not stagnate) we can construct the matrix D_r. The prescribed eigenvalues yield their companion matrix C. Now, let V be any unitary matrix and

$$A = V D_r \hat{U} C \hat{U}^{-1} D_r^{-1} V^*, \quad b = V e_1.$$

Then, A has the prescribed eigenvalues and using FOM with (A, b) gives the prescribed residual norms and the prescribed Ritz values. Note that the right-hand side can also be chosen as we wish since we can freely choose the first column of V but, of course, this yields a different matrix. The case of some matrices H_k being singular is treated in [6].

2.2. Diagonalizable matrices. Let us see if we can generalize the results of the previous subsection to the case A diagonalizable with $A = X\Lambda X^{-1}$. In fact, the only difference is that we have to use the Cauchy–Binet formula twice, but, the formulas are much more intricate.

THEOREM 2.2. *Let A be a diagonalizable matrix with the spectral factorization $A = X\Lambda X^{-1}$. Then, the polynomial p_k whose roots are the Ritz values at iteration k is given by*

$$\begin{aligned} p_k(\theta) = \sum_{\mathcal{I}_k} (\lambda_{i_1} - \theta) \cdots (\lambda_{i_k} - \theta)\, c_{i_1} \cdots c_{i_k} \prod_{i_1 \le i_p < i_q \le i_k} (\lambda_{i_q} - \lambda_{i_p}) \\ \times \left\{ \sum_{\mathcal{J}_k} \overline{\det((X^*X)_{\mathcal{I}_k,\mathcal{J}_k})}\, \overline{c_{j_1}} \cdots \overline{c_{j_k}} \prod_{j_1 \le j_p < j_q \le j_k} (\overline{\lambda_{j_q}} - \overline{\lambda_{j_p}}) \right\}, \end{aligned}$$

where the summations are over all sets $\mathcal{I}_k$ and $\mathcal{J}_k$ of k indices defined as in Theorem 2.1 and $c = X^{-1}b$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

The characteristic polynomial $\tilde{p}_k^F$ of H_k (which is monic) is obtained by dividing p_k by

$$(-1)^k \quad \sum_{\mathcal{I}_k} c_{i_1} \cdots c_{i_k} \prod_{i_1 \le i_p < i_q \le i_k} (\lambda_{i_q} - \lambda_{i_p}) \times \left\{ \sum_{\mathcal{J}_k} \overline{\det((X^*X)_{\mathcal{I}_k,\mathcal{J}_k})} \overline{c_{j_1}} \cdots \overline{c_{j_k}} \prod_{j_1 \le j_p < j_q \le j_k} (\overline{\lambda_{j_q}} - \overline{\lambda_{j_p}}) \right\}.$$

The residual polynomial p_k^F for FOM is obtained by dividing p_k by

$$\sum_{\mathcal{I}_k} c_{i_1} \cdots c_{i_k} \lambda_{i_1} \cdots \lambda_{i_k} \prod_{i_1 \le i_p < i_q \le i_k} (\lambda_{i_q} - \lambda_{i_p}) \times \left\{ \sum_{\mathcal{J}_k} \overline{\det((X^*X)_{\mathcal{I}_k,\mathcal{J}_k})} \overline{c_{j_1}} \cdots \overline{c_{j_k}} \prod_{j_1 \le j_p < j_q \le j_k} (\overline{\lambda_{j_q}} - \overline{\lambda_{j_p}}) \right\},$$

if this quantity is different from zero.

Proof. The proof is essentially the same as for Theorem 2.1; we just outline the differences. Now, the equation for the Ritz values is

$$\begin{aligned} 0 = \det(H_k - \theta I) &= \det(M_k^{-1}) \det(K_{n,k}^* X \Lambda X^{-1} K_{n,k} - \theta M_k), \\ &= \frac{1}{\det(M_k)} \det(\mathcal{V}_{n,k}^* D_c^* (X^*X) \Lambda D_c \mathcal{V}_{n,k} - \theta \mathcal{V}_{n,k}^* D_c^* D_c \mathcal{V}_{n,k}). \end{aligned}$$

Using the same notation as above the equation reads

$$\det(B^*(X^*X)\Lambda B - \theta B^*(X^*X)B) = \det(B^*(X^*X)(\Lambda - \theta I)B) = 0.$$

Let us define $G = (\Lambda - \theta I)B$. Then we have $\det(B^*(X^*X)G) = 0$. Once again we can use the Cauchy–Binet formula to obtain

$$\det(B^*(X^*X)G) = \sum_{\mathcal{I}_k} \overline{\det((X^*XB)_{\mathcal{I}_k,:})} \det(G_{\mathcal{I}_k,:}).$$

As before, $\det(G_{\mathcal{I}_k,:})$ is given by (2.2). For $(X^*XB)_{\mathcal{I}_k,:}$ we again use the Cauchy–Binet formula. It yields

$$\det((X^*XB)_{\mathcal{I}_k,:}) = \sum_{\mathcal{J}_k} \det((X^*X)_{\mathcal{I}_k,\mathcal{J}_k}) c_{j_1} \cdots c_{j_k} \prod_{j_1 \le j_p < j_q \le j_k} (\lambda_{j_q} - \lambda_{j_p}).$$ ∎

Let $\mu_0 = p_k(0) \neq 0$. Then, the residual vector is $r_k^F = \sum_{j=1}^n c_j p_k^F(\lambda_j) x^{(j)}$ with $c = X^{-1}b$ and

$$p_k^F(\lambda_j) = \frac{1}{\mu_0} \sum_{\hat{\mathcal{I}}_k} (\lambda_{i_1} - \lambda_j) \cdots (\lambda_{i_k} - \lambda_j)\, c_{i_1} \cdots c_{i_k} \prod_{i_1 \le i_p < i_q \le i_k} (\lambda_{i_q} - \lambda_{i_p}) \times \left\{ \sum_{\mathcal{J}_k} \overline{\det((X^*X)_{\mathcal{I}_k,\mathcal{J}_k})} \overline{c_{j_1}} \cdots \overline{c_{j_k}} \prod_{j_1 \le j_p < j_q \le j_k} (\overline{\lambda_{j_q}} - \overline{\lambda_{j_p}}) \right\}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

2.3. Relations between the coefficients of the FOM polynomials and the matrix M_k. There exist interesting relationships between the coefficients of the characteristic polynomial of H_k as well as those of the FOM residual polynomial and the entries of the inverse of the matrix $M_k = K_{n,k}^* K_{n,k}$. The vector of the coefficients of the characteristic polynomial of H_k (in ascending order of the powers) denoted as $\begin{pmatrix}\alpha_0 & \dots & \alpha_{k-1}\end{pmatrix}^T$ is the solution of the linear system,

$$M_k \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_{k-1} \end{pmatrix} = -M_{[1:k],k+1}. \tag{2.3}$$

For a proof, see [22].

Theorem 2.3. *The coefficients of the characteristic polynomial $\tilde{p}_k^F$ of H_k (in ascending order of powers) are given by*

$$\alpha_{j-1} = \frac{(M_{k+1}^{-1})_{j,1} - (M_k^{-1})_{j,1}}{(M_{k+1}^{-1})_{k+1,1}}, \quad j = 1, \dots, k. \tag{2.4}$$

Moreover, we have

$$\frac{\alpha_{j-1}}{\alpha_0} = \frac{(M_{k+1}^{-1})_{j,1} - (M_k^{-1})_{j,1}}{(M_{k+1}^{-1})_{1,1} - (M_k^{-1})_{1,1}}, \ j = 1, \dots, k. \tag{2.5}$$

Proof. Let us consider the first column of the inverse of M_{k+1}, given by the solution of the linear system

$$M_{k+1} x = \begin{pmatrix} M_k & M_{[1:k],k+1} \\ (M_{[1:k],k+1})^* & m_{k+1,k+1} \end{pmatrix} \begin{pmatrix} \tilde{x} \\ x_{k+1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

It yields the two equations

$$\begin{aligned} M_k \tilde{x} + x_{k+1} M_{[1:k],k+1} &= e_1, \\ (M_{[1:k],k+1})^* \tilde{x} + x_{k+1} m_{k+1,k+1} &= 0. \end{aligned}$$

Solving the first equation for $\tilde{x}$ we have

$$\tilde{x} = M_k^{-1}(e_1 - x_{k+1} M_{[1:k],k+1})$$

and multiplying by e_j^T, $j = 1, \dots, k$, we obtain

$$\begin{aligned} e_j^T \tilde{x} = (M_{k+1}^{-1})_{j,1} &= (M_k^{-1})_{j,1} - (M_{k+1}^{-1})_{k+1,1} e_j^T M_k^{-1} M_{[1:k],k+1}, \\ &= (M_k^{-1})_{j,1} + (M_{k+1}^{-1})_{k+1,1} \alpha_{j-1}, \end{aligned}$$

since $x_{k+1} = (M_{k+1}^{-1})_{k+1,1}$. The last relation proves (2.4). Now we can use the second equation which, after elimination of $\tilde{x}$, yields

$$x_{k+1}[m_{k+1,k+1} - (M_{[1:k],k+1})^* M_k^{-1} M_{[1:k],k+1}] = -(M_{[1:k],k+1})^* M_k^{-1} e_1.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The term within brackets is a Schur complement equal to $1/(M_{k+1}^{-1})_{k+1,k+1}$ and the right-hand side is $\bar{\alpha}_0$. Therefore $(M_{k+1}^{-1})_{k+1,1} = \bar{\alpha}_0 (M_{k+1}^{-1})_{k+1,k+1}$. Plugging this result in the expression for α_0 in (2.4) yields

$$|\alpha_0|^2 = \frac{(M_{k+1}^{-1})_{1,1} - (M_k^{-1})_{1,1}}{(M_{k+1}^{-1})_{k+1,k+1}}.$$

Therefore we obtain the following simple formula,

$$\frac{\alpha_{j-1}}{\alpha_0} = \frac{(M_{k+1}^{-1})_{j,1} - (M_k^{-1})_{j,1}}{|\alpha_0|^2 (M_{k+1}^{-1})_{k+1,k+1}} = \frac{(M_{k+1}^{-1})_{j,1} - (M_k^{-1})_{j,1}}{(M_{k+1}^{-1})_{1,1} - (M_k^{-1})_{1,1}}, \; j = 1, \dots, k.$$

This proves (2.5). ∎

Relation (2.5) gives the coefficients of the FOM residual polynomial. We should mention that one can obtain closed-form expressions for the entries of the inverse of M_k. Due to lack of space we cannot explain this in this paper. But, this gives another way to obtain the coefficients of the FOM polynomials. However, the formulas that are obtained are even more intricate that the ones we derived in this paper.

3. The GMRES residual polynomial. From the definition of the harmonic Ritz values and vectors one can see that they are also given by solving

$$\underline{H}_k^* \underline{H}_k x = \zeta H_k^* x.$$

We have $H_k = V_{n,k}^* A V_{n,k}$ and $\underline{H}_k = V_{n,k+1}^* A V_{n,k}$. It yields

$$V_{n,k}^* A^* (V_{n,k+1} V_{n,k+1}^*) A V_{n,k} x = \zeta V_{n,k}^* A^* V_{n,k} x.$$

But, $V_{n,k+1} V_{n,k+1}^*$ is the orthogonal projector on the Krylov subspace $\mathcal{K}_{k+1}(A, v)$. Since $AV_{n,k}x$ is in this subspace, we have $(V_{n,k+1} V_{n,k+1}^*) A V_{n,k} x = A V_{n,k} x$ and

$$V_{n,k}^* A^* A V_{n,k} x = \zeta V_{n,k}^* A^* V_{n,k} x.$$

3.1. Normal matrices. When A is normal this yields the following result.

THEOREM 3.1. *Let A be a normal matrix with the spectral factorization $A = X \Lambda X^*$. Then, the polynomial q_k whose roots are the harmonic Ritz values at iteration k is given by*

$$\sum_{\mathcal{I}_k} (|\lambda_{i_1}|^2 - \zeta \overline{\lambda_{i_1}}) \cdots (|\lambda_{i_k}|^2 - \zeta \overline{\lambda_{i_k}}) \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2, \tag{3.1}$$

where the summation is over all sets $\mathcal{I}_k$ of k indices $(i_1, \; i_2, \; \dots, \; i_k)$ such that $1 \le i_1 < \cdots < i_k \le n$ and $c = X^* b$. In the product the indices i_p and i_q belong to the set of indices $\mathcal{I}_k$.

The characteristic polynomial $\tilde{p}_k^G$ of $\hat{H}_k$ (which is monic) is obtained by dividing q_k by

$$(-1)^k \sum_{\mathcal{I}_k} \overline{\lambda_{i_1}} \cdots \overline{\lambda_{i_k}} \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2.$$

The residual polynomial p_k^G for GMRES is obtained by dividing q_k by

$$\sum_{\mathcal{I}_k} |\lambda_{i_1}|^2 \cdots |\lambda_{i_k}|^2 \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. With the same notation as in the previous section the equation for the harmonic Ritz values is

$$\det(B^*\Lambda^*\Lambda B - \zeta B^*\Lambda^* B) = \det(B^*(\Lambda^*\Lambda - \zeta\Lambda^*)B) = 0.$$

Let $G = (\Lambda^*\Lambda - \zeta\Lambda^*)B$; we have $\det(B^*G) = 0$. Therefore, the new polynomial is

$$\det(B^*G) = \sum_{\mathcal{I}_k}(|\lambda_{i_1}|^2 - \zeta\overline{\lambda_{i_1}})\cdots(|\lambda_{i_k}|^2 - \zeta\overline{\lambda_{i_k}})\,|c_{i_1}|^2\cdots|c_{i_k}|^2 \prod_{i_1\le i_p<i_q\le i_k}|\lambda_{i_q}-\lambda_{i_p}|^2.$$

We have

$$(|\lambda_{i_1}|^2 - \zeta\overline{\lambda_{i_1}})\cdots(|\lambda_{i_k}|^2 - \zeta\overline{\lambda_{i_k}}) = \sum_{j=0}^{k}(-1)^{k-j}e_{(j)}(\lambda_{i_1},\ldots,\lambda_{i_k})\overline{\lambda_{i_1}}\cdots\overline{\lambda_{i_k}}\zeta^{k-j}.$$

The coefficient of ζ^{k-j}, $j=0,\ldots,k$, in the polynomial q_k whose roots are the harmonic Ritz values is

$$(-1)^{k-j}\sum_{\mathcal{I}_k}e_{(j)}(\lambda_{i_1},\ldots,\lambda_{i_k})\,\overline{\lambda_{i_1}}\cdots\overline{\lambda_{i_k}}\,|c_{i_1}|^2\cdots|c_{i_k}|^2\prod_{i_1\le i_p<i_q\le i_k}|\lambda_{i_q}-\lambda_{i_p}|^2.$$

From this we can obtain the characteristic polynomial of $\hat{H}_k$ and the GMRES residual polynomial dividing by the appropriate coefficient. ∎

The residual vector at iteration k is $r_k^G = \sum_{j=1}^n c_j p_k^G(\lambda_j)x^{(j)}$ with $c = X^*b$ and

$$p_k^G(\lambda_j) = \frac{1}{\gamma_0}\sum_{\hat{\mathcal{I}}_k}(|\lambda_{i_1}|^2 - \lambda_j\overline{\lambda_{i_1}})\cdots(|\lambda_{i_k}|^2 - \lambda_j\overline{\lambda_{i_k}})\,|c_{i_1}|^2\cdots|c_{i_k}|^2\prod_{i_1\le i_p<i_q\le i_k}|\lambda_{i_q}-\lambda_{i_p}|^2$$

with

$$\gamma_0 = \sum_{\mathcal{I}_k}|\lambda_{i_1}|^2\cdots|\lambda_{i_k}|^2\,|c_{i_1}|^2\cdots|c_{i_k}|^2\prod_{i_1\le i_p<i_q\le i_k}|\lambda_{i_q}-\lambda_{i_p}|^2.$$

3.2. Diagonalizable matrices. Let us now assume that A is diagonalizable. Then we have the following result. We just give a sketch of the proof since it is essentially the same as what we did for the FOM residual polynomial.

THEOREM 3.2. *Let A be a diagonalizable matrix with the spectral factorization $A = X\Lambda X^{-1}$. Then, the polynomial q_k whose roots are the harmonic Ritz values at iteration k is given by*

$$\begin{aligned} q_k(\zeta) = &\sum_{\mathcal{I}_k}(\lambda_{i_1}-\zeta)\cdots(\lambda_{i_k}-\zeta)\,\overline{\lambda_{i_1}}\cdots\overline{\lambda_{i_k}}c_{i_1}\cdots c_{i_k}\prod_{i_1\le i_p<i_q\le i_k}(\lambda_{i_q}-\lambda_{i_p})\\ &\times\left\{\sum_{\mathcal{J}_k}\overline{\det((X^*X)_{\mathcal{I}_k,\mathcal{J}_k})}\overline{c_{j_1}}\cdots\overline{c_{j_k}}\prod_{j_1\le j_p<j_q\le j_k}(\overline{\lambda_{j_q}}-\overline{\lambda_{j_p}})\right\},\end{aligned}$$

where the summations are over all sets $\mathcal{I}_k$ and $\mathcal{J}_k$ of k indices defined as in Theorem 2.1 and $c = X^{-1}b$.

The characteristic polynomial $\tilde{p}_k^G$ of $\hat{H}_k = H_k + h_{k+1,k}^2 H_k^{-*}e_k e_k^T$ (which is monic) is obtained by dividing q_k by

$$\begin{aligned}(-1)^k\quad&\sum_{\mathcal{I}_k}\overline{\lambda_{i_1}}\cdots\overline{\lambda_{i_k}}c_{i_1}\cdots c_{i_k}\prod_{i_1\le i_p<i_q\le i_k}(\lambda_{i_q}-\lambda_{i_p})\\ &\times\left\{\sum_{\mathcal{J}_k}\overline{\det((X^*X)_{\mathcal{I}_k,\mathcal{J}_k})}\overline{c_{j_1}}\cdots\overline{c_{j_k}}\prod_{j_1\le j_p<j_q\le j_k}(\overline{\lambda_{j_q}}-\overline{\lambda_{j_p}})\right\}.\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

The residual polynomial p_k^G for GMRES is obtained by dividing q_k by

$$\sum_{\mathcal{I}_k} |\lambda_{i_1}|^2 \cdots |\lambda_{i_k}|^2 c_{i_1} \cdots c_{i_k} \prod_{i_1 \le i_p < i_q \le i_k} (\lambda_{i_q} - \lambda_{i_p}) \times \left\{ \sum_{\mathcal{J}_k} \overline{\det((X^*X)_{\mathcal{I}_k,\mathcal{J}_k})} \overline{c_{j_1}} \cdots \overline{c_{j_k}} \prod_{j_1 \le j_p < j_q \le j_k} (\overline{\lambda_{j_q}} - \overline{\lambda_{j_p}}) \right\}.$$

Proof. The equation for the harmonic Ritz values is

$$\det(B^*(X^*X)\Lambda^*\Lambda B - \zeta B^*(X^*X)\Lambda^* B) = \det(B^*(X^*X)(\Lambda^*\Lambda - \zeta\Lambda^*)B) = 0.$$

If we choose $G = (\Lambda^*\Lambda - \zeta\Lambda^*)B$, we have $\det(B^*(X^*X)G) = 0$. Using the Cauchy–Binet formula twice, we obtain,

$$\det(B^*(X^*X)G) = \sum_{\mathcal{I}_k} (\lambda_{i_1} - \zeta) \cdots (\lambda_{i_k} - \zeta)\, \overline{\lambda_{i_1}} \cdots \overline{\lambda_{i_k}} c_{i_1} \cdots c_{i_k} \prod_{i_1 \le i_p < i_q \le i_k} (\lambda_{i_q} - \lambda_{i_p}) \times \left\{ \sum_{\mathcal{J}_k} \overline{\det((X^*X)_{I_k,J_k})} \overline{c_{j_1}} \cdots \overline{c_{j_k}} \prod_{j_1 \le j_p < j_q \le j_k} (\overline{\lambda_{j_q}} - \overline{\lambda_{j_p}}) \right\}.$$

∎

Again the residual vector at iteration k is $r_k^G = \sum_{j=1}^n c_j p_k^G(\lambda_j) x^{(j)}$ but with $c = X^{-1}b$. The formula for $p_k^G(\lambda_j)$ is similar to what we have for normal matrices except for the additional sum over $\mathcal{J}_k$.

3.3. Relations between the coefficients of the GMRES polynomials and the matrix M_k. In this subsection we first look for a factorization of $\hat{H}_k$ similar to the one we have for H_k.

THEOREM 3.3. *For $k < n$ assume that H_k is nonsingular. Then the matrix $\hat{H}_k$ can be written as $\hat{H}_k = U_k \hat{C}^{(k)} U_k^{-1}$, U_k being upper triangular and*

$$\hat{C}^{(k)} = C^{(k)} - \frac{u_{k+1,k+1}^2}{\bar{\alpha}_0} U_k^{-1} U_k^{-*} e_1 e_k^T,$$

being a companion matrix, where $C^{(k)}$ is the companion matrix in $H_k = U_k C^{(k)} U_k^{-1}$. Let $\begin{pmatrix} \beta_0 & \dots & \beta_{k-1} \end{pmatrix}^T$ be the coefficients of the characteristic polynomial of $\hat{H}_k$. We have

$$M_k \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_{k-1} \end{pmatrix} = -M_{[1:k],k+1} + \gamma_{k+1} e_1, \tag{3.2}$$

where $M_k = K_{n,k}^* K_{n,k} = U_k^* U_k$ and

$$\gamma_{k+1} = \frac{u_{k+1,k+1}^2}{\bar{\alpha}_0} = \frac{1}{(M_{k+1}^{-1})_{k+1,1}}.$$

Proof. Let us first consider $H_k^{-*} e_k$. We have $H_k^{-*} = U_k^{-*} [C^{(k)}]^{-*} U_k^*$. Since U_k is upper triangular we have $U_k^* e_k = u_{k,k} e_k$ with $u_{k,k}$ real and positive. The companion matrix $C^{(k)}$ is

$$C^{(k)} = \begin{pmatrix} 0 & -\alpha_0 \\ I_{k-1} & -\hat{\alpha} \end{pmatrix}$$

with $\hat{\alpha} = \begin{pmatrix} \alpha_1 & \cdots & \alpha_{n-1} \end{pmatrix}^T$, the α_j's being the coefficients of the characteristic

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

polynomial of H_k. We have $\alpha_0 \neq 0$ since H_k is nonsingular. The inverse of the companion matrix is

$$[C^{(k)}]^{-1} = \begin{pmatrix} -\hat{\alpha}/\alpha_0 & I_{k-1} \\ -1/\alpha_0 & 0 \end{pmatrix}.$$

Therefore if we (Hermitian) transpose and take the last column of the result, we have

$$[C^{(k)}]^{-*} e_k = -\frac{1}{\bar{\alpha}_0} e_1.$$

Finally we obtain

$$H_k^{-*} e_k = -\frac{u_{k,k}}{\bar{\alpha}_0} U_k^{-*} e_1.$$

On the other hand, we have $h_{k+1,k} = u_{k+1,k+1}/u_{k,k}$. Since $u_{1,1} = 1$ (because $\|b\| = 1$) and $h_{k+1,k}$ is real and positive, this implies that all the diagonal entries of U are real and positive as we said before. Then

$$\hat{H}_k = U_k C^{(k)} U_k^{-1} - \frac{u_{k+1,k+1}^2}{\bar{\alpha}_0 u_{k,k}} U_k^{-*} e_1 e_k^T.$$

Let us factor U_k on the left and U_k^{-1} on the right. We obtain

$$\hat{H}_k = U_k \left[C^{(k)} - \frac{u_{k+1,k+1}^2}{\bar{\alpha}_0 u_{k,k}} U_k^{-1} U_k^{-*} e_1 e_k^T U_k \right] U_k^{-1}.$$

We remark that $e_k^T U_k = u_{k,k} e_k^T$. Hence $\hat{H}_k$ is similar to the matrix

$$\hat{C}^{(k)} = C^{(k)} - \frac{u_{k+1,k+1}^2}{\bar{\alpha}_0} U_k^{-1} U_k^{-*} e_1 e_k^T.$$

The second term on the right-hand side modifies only the last column. Therefore, $\hat{C}^{(k)}$ is a companion matrix and the coefficients of the characteristic polynomial of $\hat{H}_k$ are given by the negative of the last column of $\hat{C}^{(k)}$. We remark that

$$u_{k+1,k+1}^2 = \frac{1}{(M_{k+1}^{-1})_{k+1,k+1}}.$$

From Theorem 2.3 we have

$$\begin{pmatrix} \beta_0 \\ \vdots \\ \beta_{k-1} \end{pmatrix} = -U_k^{-1} U_{[1:k],k+1} + \gamma_{k+1} U_k^{-1} U_k^{-*} e_1$$

with

$$\gamma_{k+1} = \frac{u_{k+1,k+1}^2}{\bar{\alpha}_0} = \frac{1}{(M_{k+1}^{-1})_{k+1,1}}.$$

Multiplying both sides by $M_k = U_k^* U_k$ we obtain

$$M_k \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_{k-1} \end{pmatrix} = -M_{[1:k],k+1} + \gamma_{k+1} e_1.$$

Note that if A is real, the coefficient α_0 and γ_{k+1} are real. We see that, compared to the linear system (2.3) for the characteristic polynomial of H_k, only the first entry of the right-hand side is modified. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

The previous theorem shows that we have a simple relation between the coefficients of the characteristic polynomials for the harmonic Ritz values and the standard Ritz values,

$$\begin{pmatrix} \beta_0 \\ \vdots \\ \beta_{k-1} \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_{k-1} \end{pmatrix} + \gamma_{k+1} M_k^{-1} e_1.$$

From the results in Theorem 2.3 we have simple expressions for the coefficients of the characteristic polynomial of $\hat{H}_k$ as functions of the entries of the inverse of M_{k+1}.

THEOREM 3.4. *The coefficients of the characteristic polynomial* $\tilde{p}_k^G$ *of* $\hat{H}_k$ *(in ascending order of powers) are given by*

$$\beta_{j-1} = \frac{(M_{k+1}^{-1})_{j,1}}{(M_{k+1}^{-1})_{k+1,1}}, \quad j = 1, \ldots, k. \tag{3.3}$$

The coefficients of the GMRES residual polynomial p_k^G *(in ascending order of powers) at iteration* k *are*

$$1, \quad \frac{\beta_{j-1}}{\beta_0} = \frac{(M_{k+1}^{-1})_{j,1}}{(M_{k+1}^{-1})_{1,1}}, \ j = 2, \ldots, k, \quad \frac{1}{\beta_0} = \frac{(M_{k+1}^{-1})_{k+1,1}}{(M_{k+1}^{-1})_{1,1}}.$$

Proof. We have

$$\begin{aligned} \beta_{j-1} &= \alpha_{j-1} + \gamma_{k+1} (M_k^{-1})_{j,1} \\ &= \frac{(M_{k+1}^{-1})_{j,1} - (M_k^{-1})_{j,1}}{(M_{k+1}^{-1})_{k+1,1}} + \frac{(M_k^{-1})_{j,1}}{(M_{k+1}^{-1})_{k+1,1}} = \frac{(M_{k+1}^{-1})_{j,1}}{(M_{k+1}^{-1})_{k+1,1}}. \end{aligned}$$ ∎

We note that we have a relation between the coefficients of the FOM and GMRES residual polynomials,

$$\frac{1}{\alpha_0} = \frac{1}{\beta_0} \left[\frac{1}{1 - \frac{(M_k^{-1})_{1,1}}{(M_{k+1}^{-1})_{1,1}}} \right], \quad \frac{\alpha_{j-1}}{\alpha_0} = \frac{\beta_{j-1}}{\beta_0} \left[\frac{1 - \frac{(M_k^{-1})_{j,1}}{(M_{k+1}^{-1})_{j,1}}}{1 - \frac{(M_k^{-1})_{1,1}}{(M_{k+1}^{-1})_{1,1}}} \right].$$

4. The Ritz and harmonic Ritz values for $k = 2$. Since we know the coefficients of the polynomials associated with FOM and GMRES, we can write the Ritz and harmonic Ritz values for $k \leq 4$. Let us assume for simplicity that A is normal and $k = 2$. We consider the polynomial p_2 written as

$$p_2(\theta) = \delta_2 \theta^2 + \delta_1 \theta + \delta_0$$

with

$$\begin{aligned} \delta_2 &= \sum_{\mathcal{I}_2} |c_{i_1}|^2 |c_{i_2}|^2 |\lambda_{i_2} - \lambda_{i_1}|^2, \\ \delta_1 &= \sum_{\mathcal{I}_2} (\lambda_{i_1} + \lambda_{i_2}) \, |c_{i_1}|^2 |c_{i_2}|^2 |\lambda_{i_2} - \lambda_{i_1}|^2, \\ \delta_0 &= \sum_{\mathcal{I}_2} \lambda_{i_1} \lambda_{i_2} \, |c_{i_1}|^2 |c_{i_2}|^2 |\lambda_{i_2} - \lambda_{i_1}|^2, \end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where the summations are over all the ordered pairs of indices (i_1, i_2) such that $1 \leq i_1 < i_2 \leq n$. The Ritz values at the second iteration are given by the roots of p_2,

$$\theta_{1,2} = \frac{-\delta_1 \pm \sqrt{\delta_1^2 - 4\delta_2\delta_0}}{2\delta_2}.$$

The harmonic Ritz values are the roots of the polynomial

$$\begin{aligned} q_2(\zeta) &= \nu_2\zeta^2 + \nu_1\zeta + \nu_0, \\ \nu_2 &= \sum_{\mathcal{I}_2} \overline{\lambda_{i_1}}\,\overline{\lambda_{i_2}}\,|c_{i_1}|^2|c_{i_2}|^2|\lambda_{i_2} - \lambda_{i_1}|^2, \\ \nu_1 &= \sum_{\mathcal{I}_2} (\lambda_{i_1} + \lambda_{i_2})\,\overline{\lambda_{i_1}}\,\overline{\lambda_{i_2}}\,|c_{i_1}|^2|c_{i_2}|^2|\lambda_{i_2} - \lambda_{i_1}|^2, \\ \nu_0 &= \sum_{\mathcal{I}_2} |\lambda_{i_1}|^2|\lambda_{i_2}|^2\,|c_{i_1}|^2|c_{i_2}|^2|\lambda_{i_2} - \lambda_{i_1}|^2. \end{aligned}$$

Then,

$$\zeta_{1,2} = \frac{-\nu_1 \pm \sqrt{\nu_1^2 - 4\nu_2\nu_0}}{2\nu_2}.$$

Note that when A is real all the coefficients δ_j, ν_j are real. These formulas can be used, for instance, to study what happens to the two Ritz or harmonic Ritz values when we move around one or several eigenvalues or when we change the starting vector in the Arnoldi process. This is what we do in the next example.

We can also write down closed-form expressions for the Ritz and harmonic Ritz values for $k = 3, 4$ but the expressions are much more involved and we do not consider them in this paper.

5. A small example. Let us consider a small example from [22]. The matrix A is real normal of order 4. Its eigenvalues are $\lambda_1 = -0.432565 + 1.66558\imath$, $\lambda_2 = -0.432565 - 1.66558\imath$, $\lambda_3 = 0.187377$, and $\lambda_4 = -1.20852$. We have two real eigenvalues and a pair of complex conjugate eigenvalues. As one can see in Figure 1 the field of values of A which contains the Ritz values is a quadrilateral. Let $\omega_i = |c_i|^2$. The two values corresponding to λ_1 and λ_2 are equal, $\omega_1 = \omega_2$, and, therefore, since $\|c\| = 1$, $2\omega_1 + \omega_3 + \omega_4 = 1$ with $0 < \omega_i < 1$. Therefore we can compute the Ritz values at $k = 2$ for given values of ω_3 and ω_4 using the formulas given above. This gives all the possible locations of the Ritz values when we change the starting vector. This is what was done for Figure 1 where we used 50 uniformly distributed values in $(0, 1)$ for ω_3 and ω_4. One can check that the result agrees with what was computed in [22] by using another technique. The Ritz values do not cover the whole surface of the field of values. As already noticed in [22], this is due to the fact that we are considering a problem with real data. The same phenomenon does not happen with a complex matrix. On these issues, see also [4]. For our problem we see that there exist starting vectors (that is, values of ω_3 and ω_4) for which the pairs of complex conjugate Ritz values are close to the pair of complex eigenvalues of A or very far from it. The boundary of the region containing the pairs of complex Ritz values is given by $\omega_3 = 0$ or $\omega_4 = 0$. Another way of computing this boundary is given in [22].

Of course, we can do the same thing for the harmonic Ritz values. Figure 2 displays the inverses of the harmonic Ritz values in the field of values of A^{-1}. We see that the inverses of the complex harmonic Ritz values are also clustered in a small region of the field of values. The inverses of the real harmonic Ritz values do not

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

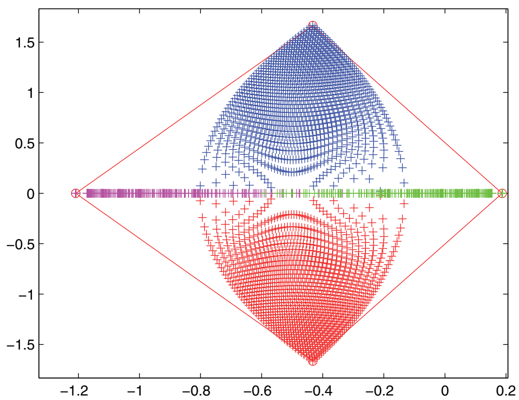

FIG. 1. *Example, $n = 4, k = 2$, A normal real, location of the Ritz values in the field of values of A.*

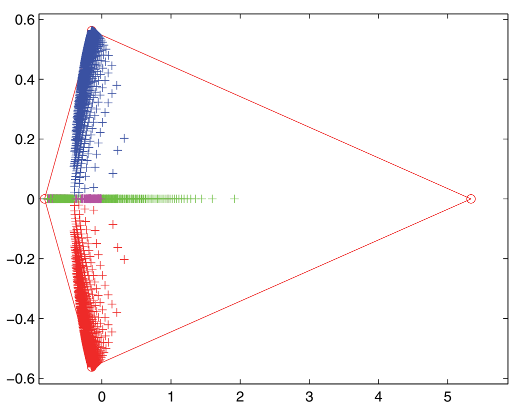

FIG. 2. *Example, $n = 4, k = 2$, A normal real, location of the inverses of the harmonic Ritz values in the field of values of A^{-1}.*

completely fill the intersection of the field of values with the real axis. It would be interesting to explain this phenomenon.

Let us now consider the third iteration of GMRES applied to this matrix. The sum with $\mathcal{I}_3$ is over the set of indices

$$(1,2,3),\ (1,2,4)\ (1,3,4),\ (2,3,4).$$

When considering $p_3^G(\lambda_1)$ only one term remains in the sum (corresponding to the indices (2, 3, 4)) and we have

$$p_3^G(\lambda_1) = \frac{1}{\gamma_0}[\overline{\lambda}_2\lambda_3\lambda_4\,(\lambda_2-\lambda_1)(\lambda_3-\lambda_1)(\lambda_4-\lambda_1) \\ \times |c_2|^2\,|c_3|^2\,|c_4|^2\,|\lambda_2-\lambda_3|^2\,|\lambda_2-\lambda_4|^2\,\lambda_3-\lambda_4|^2].$$

In this small example none of the pairwise distances between eigenvalues is small.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Hence, generally, $p_3^G(\lambda_1)$ can only be small if at least one of the $|c_j|^2$, $j = 2, 3, 4$, is small. Let us first consider a right-hand side such that $|c_i| = 1$, $j = 1, \ldots, 4$. We obtain the following residual norms,

$$2.0000, 1.8763, 1.4799, 1.2169, 1.2009\ 10^{-15}.$$

The dot products of the first eigenvector and the residual vectors are

$$1, 0.98564, 0.30887, 0.053046, 2.7433\ 10^{-16}.$$

As predicted, they are not small before iteration 4. Let us now take a right-hand side for which $|c_1| = 0.9$, $|c_2| = 0.1$, $|c_3| = 0.01$, $|c_4| = 1.0113\ 10^{-4}$. The right-hand side is large in the direction of the first eigenvector. Then, the residual norms are

$$0.91104, 0.21835, 0.10648, 8.4935\ 10^{-4}, 2.3488\ 10^{-16}.$$

The dot products of the first eigenvector and the residual vectors are

$$0.9, 0.021404, 0.6861\ 10^{-4}, 2.8713\ 10^{-8}, 1.4048\ 10^{-16}.$$

The component of the residual on the first eigenvector becomes quite small at iteration 3 because $p_3^G(\lambda_1)$ is small. This is an example of what can be obtained with the formulas derived in this paper.

6. Distances of Ritz values to eigenvalues.

For $k > 4$ we cannot obtain closed-form expressions for the Ritz values. Therefore, we are interested in finding bounds for the distance of an eigenvalue of A to the Ritz values at a given iteration of the Arnoldi algorithm. We have seen above that we know closed-form expressions for the coefficients of the characteristic polynomial of H_k in terms of the eigenvalues and eigenvectors of A. Therefore the problem is reduced to finding bounds for the roots of a given monic polynomial as a function of its coefficients.

6.1. Bounds for roots of polynomials.

There exist many lower and upper bounds for the moduli of the roots of a given polynomial as a function of its coefficients in the literature. Some of them use Gershgorin disks for the companion matrix associated with the polynomial coefficients or their generalizations; see [19, 20]. There are hundreds of papers providing upper bounds. But one can obtain lower bounds by applying their results to the polynomial

$$\tilde{p}(z) = \frac{z^k}{\alpha_0} p\left(\frac{1}{z}\right),$$

whose roots are the inverses of the roots of p. A few of these papers are (in chronological order) [14, 17, 24, 16, 15, 1, 5]. Note that [5] does not use the companion matrix but the Fiedler matrices.

For the moment we consider a monic polynomial

$$p(z) = z^k + \alpha_{k-1} z^{k-1} + \cdots + \alpha_1 z + \alpha_0,$$

and its roots θ_j. Some of the bounds that one can find in the literature are the following (see [12, 18, 5]):

- Cauchy's bounds,

$$\frac{|\alpha_0|}{\max_{i=1,\ldots,k-1}(1, |\alpha_0| + |\alpha_i|)} \leq |\theta_j| \leq \max_{i=1,\ldots,k-1} (|\alpha_0|, 1 + |\alpha_i|);$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

- Montel's bounds,

$$\frac{|\alpha_0|}{\max(|\alpha_0|, 1+|\alpha_1|+\cdots+|\alpha_{k-1}|)} \le |\theta_j| \le \max(1, |\alpha_0|+\cdots+|\alpha_{k-1}|);$$

- Carmichael and Mason's bounds,

$$\frac{|\alpha_0|}{(1+|\alpha_0|^2+\cdots+|\alpha_{k-1}|^2)^{\frac{1}{2}}} \le |\theta_j| \le (1+|\alpha_0|^2+\cdots+|\alpha_{k-1}|^2)^{\frac{1}{2}};$$

- De Terán, Dopico, and Pérez's bounds,

$$\frac{|\alpha_0|}{\max_{i=2,\dots,k-1}(1+|\alpha_1|, |\alpha_0|(1+|\alpha_i|))} \le |\theta_j| \le \max_{i=1,\dots,k-2}\left(1+\frac{|\alpha_i|}{|\alpha_0|}, |\alpha_0|+|\alpha_{k-1}|\right).$$

6.2. Application to the Arnoldi algorithm. At iteration k of the Arnoldi algorithm we have the characteristic polynomial of H_k,

$$\tilde{p}_k^F(z) = z^k + \alpha_{k-1}z^{k-1} + \cdots + \alpha_1 z + \alpha_0.$$

When A is normal the coefficients are given by Theorem 2.1.

The results that we have seen above give us bounds for $|\theta_j|$ depending on the coefficients of the polynomial. However, this is not particularly interesting for us. What we would like to know is if some Ritz values are close to some eigenvalues of A. Hence, we are interested in the distances of the Ritz values to a given eigenvalue of A, say λ_i. A way to obtain this is to do a change of variable in the polynomial,

$$z \to z + \lambda_i.$$

The roots of the new polynomial $\hat{p}_k(z)$ will be $\theta_j - \lambda_i$ and the lower bounds will give us lower bounds for the distances of λ_i to the Ritz values. As long as they are not small, there won't be convergence of Ritz values to λ_i. The upper bounds are useless since they are always larger than 1 and they are upper bounds of the distance from λ_i to any Ritz value which means that they are usually large. Note that to do this change of variable we have to know the eigenvalue λ_i (or at least a target that has to be reached). Therefore these bounds are only of theoretical interest and cannot be used to stop the Arnoldi algorithm. Of course, we have to compute the coefficients of the modified polynomials $\hat{p}_k$. Let us denote

$$\hat{p}_k(z) = z^k + \mu_{k-1}z^{k-1} + \cdots + \mu_1 z + \mu_0.$$

There could have been an index i for the coefficients, but we are only considering one eigenvalue at a time, so there is no need for that. We have

$$(z+\lambda_i)^j = \sum_{\ell=0}^{j} \binom{j}{\ell} \lambda_i^{j-\ell} z^\ell.$$

Then we have to collect the coefficients of z^ℓ for $j = 1, \dots, k$. Hence,

$$\mu_\ell = \sum_{j=1}^{k} \alpha_j \binom{j}{\ell} \lambda_i^{j-\ell}, \quad \ell = 1, \dots, k-1$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with $\binom{j}{\ell} = 0$ if $\ell > j$ and $\alpha_k = 1$. The constant coefficient is

$$\mu_0 = \alpha_0 + \sum_{j=1}^{k} \alpha_j \lambda_i^j.$$

Therefore, in the normal case the coefficients are

$$\mu_\ell = \frac{1}{d_k} \sum_{j=1}^{k} (-1)^j \binom{j}{\ell} \lambda_i^{j-\ell} \sum_{\mathcal{I}_k} e_{(k-j)}(\lambda_{i_1}, \dots, \lambda_{i_k}) \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2$$

with

$$d_k = \sum_{\mathcal{I}_k} |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2$$

and

$$\mu_0 = \alpha_0 + \frac{1}{d_k} \sum_{j=1}^{k} (-1)^j \lambda_i^j \sum_{\mathcal{I}_k} e_{(k-j)}(\lambda_{i_1}, \dots, \lambda_{i_k}) \, |c_{i_1}|^2 \cdots |c_{i_k}|^2 \prod_{i_1 \le i_p < i_q \le i_k} |\lambda_{i_q} - \lambda_{i_p}|^2.$$

For instance the Cauchy lower bound when A is normal is

$$\frac{|\mu_0|}{\max_{\ell=1,\dots,k-1}(1, |\mu_0| + |\mu_\ell|)} \le |\theta_j - \lambda_i|,$$

and we have similar relations for the other bounds described above. The behavior of these lower bounds requires further study.

6.3. A numerical experiment for the Ritz values. Let us give a small example to illustrate the behavior of the lower bounds defined above from the coefficients of the polynomials. We consider a random normal real matrix of order 10. Figure 3 (resp., Figure 4) displays the eigenvalues of A (as circles) which are well separated and the Ritz values (as crosses) at iterations 4 and 6 (resp., 8 and 9). At iteration 6 four of the Ritz values are close to four outer complex eigenvalues. The smallest real eigenvalue is almost reached at iteration 8. Figure 5 shows the smallest distance of a Ritz value to $\lambda_1 = -0.43256 + \imath 1.6656$ (the eigenvalue with the largest imaginary part) which is an outlier and the lower bounds as functions of the iteration number. After iteration 6 the best result is given by the De Terán, Dopico, and Pérez bound. The

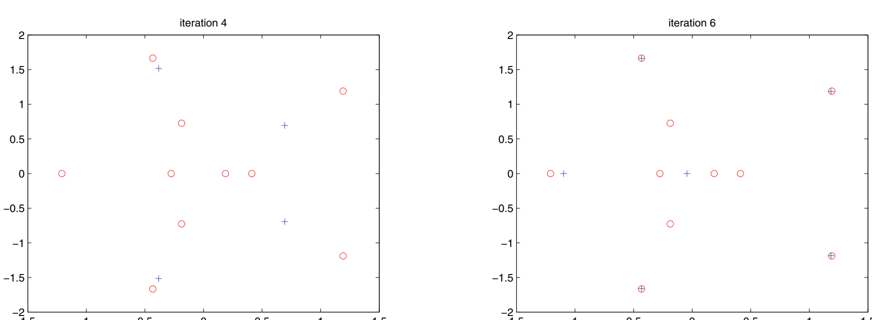

FIG. 3. *Example, $n = 10$, eigenvalues (circles) and Ritz values (+), iterations 4 and 6.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

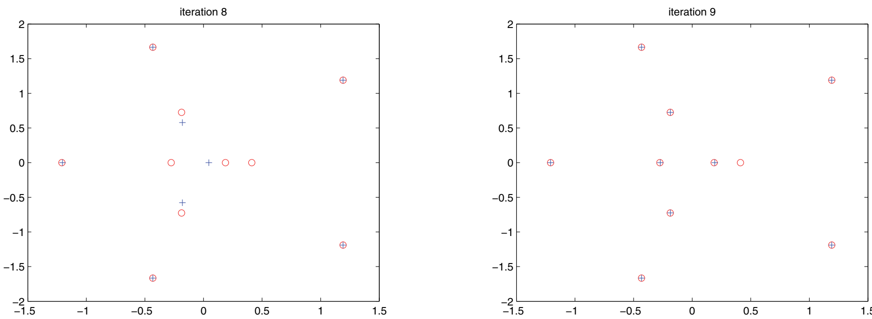

FIG. 4. *Example, $n = 10$, eigenvalues (circles) and Ritz values (+), iterations 8 and 9.*

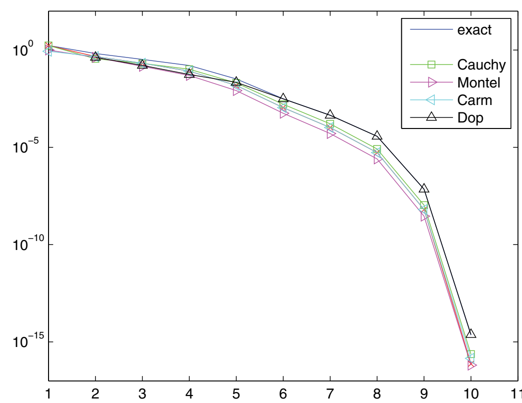

FIG. 5. *Example, $n = 10$, minimum distance to λ_1 and lower bounds.*

next best is Cauchy's bound. The worst result is given by Montel's bound. Figure 6 displays the distance and the bounds for a real eigenvalue $\lambda_8 = 0.41257$ which is the largest real eigenvalue. There is no decrease of the distance before iteration 9. On this example the lower bounds are quite sharp (at least on a log-scale) and they well describe the decrease of the distance between the given eigenvalues and Ritz values. Other numerical experiments with different problems show that the relative merits of the different bounds are not always the same.

Numerical experiments not reported here show that the lower bounds describe quite well the distances of Ritz values to an eigenvalue when A is diagonalizable but not normal.

7. Conclusion. We have derived closed-form expressions for the coefficients of the FOM and GMRES residual polynomials as functions of the eigenvalues, the eigenvectors, and the right-hand side for diagonalizable matrices. It yields the solution of the minimization problem (1.2). These expressions can be used to write down the Ritz and harmonic Ritz values for the first four iterations of the Arnoldi algorithm. They can also yield lower bounds for the minimum distances of eigenvalues to Ritz values. Unfortunately, these expressions are intricate functions of the eigenvalues and eigenvectors, even though they simplify when A is a normal matrix. It remains to

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

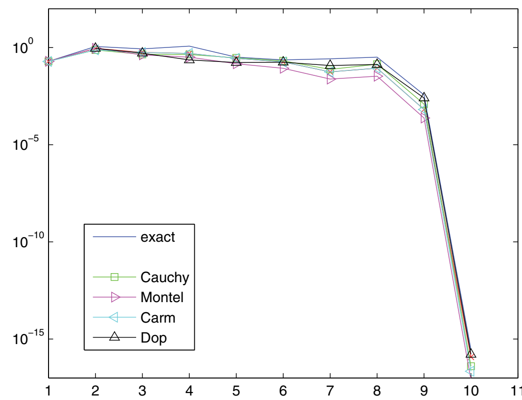

FIG. 6. *Example, $n = 10$, minimum distance to λ_8 and lower bounds.*

study more carefully the coefficients and the lower bounds to obtain more information about the convergence of the Ritz values. It would also be interesting to find good upper bounds for the minimum distances, but this requires other techniques than the ones developed here.

Acknowledgment. The author thanks the referees for interesting comments and suggestions.

REFERENCES

[1] A. Abu-Omar and F. Kittaneh, *Estimates for the numerical radius and the spectral radius of the Frobenius companion matrix and bounds for the zeros of polynomials*, Ann. Funct. Anal., 5 (2014), pp. 56–62.

[2] M. Arioli, V. Pták, and Z. Strakoš, *Krylov sequences of maximal length and convergence of GMRES*, BIT, 38 (1998), pp. 636–643.

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

[4] Z. Bujanović, *On the permissible arrangements of Ritz values for normal matrices in the complex plane*, Linear Algebra Appl., 438 (2013), pp. 4606–4624.

[5] F. De Terán, F. M. Dopico, and J. Pérez, *New bounds for roots of polynomials based on Fiedler companion matrices*, Linear Algebra Appl., 451 (2014), pp. 197–230.

[6] J. Duintjer Tebbens and G. Meurant, *Any Ritz value behavior is possible for Arnoldi and for GMRES*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 958–978.

[7] J. Duintjer Tebbens and G. Meurant, *On the convergence of Q-OR and Q-MR Krylov methods for solving linear systems*, BIT, 56 (2016), pp. 77–97.

[8] V. Faber, J. Liesen, and P. Tichý, *Properties of worst-case GMRES*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1500–1519.

[9] S. Goossens and D. Roose, *Ritz and harmonic Ritz values and the convergence of FOM and GMRES*, Numer. Linear Algebra Appl., 6 (1999), pp. 281–293.

[10] F.R. Gantmacher, *The Theory of Matrices*, Vol. 1, Chelsea, New York, 1959.

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

[12] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1987.

[13] A.S. Householder and F.L. Bauer, *On certain methods for expanding the characteristic polynomials*, Numer. Math., 1 (1959), pp. 29–37.

[14] F. Kittaneh, *Singular values of companion matrices and bounds on zeros of polynomials*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 333–340.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

[15] F. Kittaneh, *Bounds and a majorization for the real parts of the zeros of polynomials*, Proc. Amer. Math. Soc., 135 (2007), pp. 659–664.
[16] F. Kittaneh and K. Shebrawi, *Some decomposition results for companion matrices*, J. Math. Anal. Appl., 318 (2006), pp. 626–633.
[17] H. Linden, *Bounds for the zeros of polynomials from eigenvalues and singular values of some companion matrices*, Linear Algebra Appl., 271 (1998), pp. 41–82.
[18] M. Marden, *Geometry of Polynomials*, 2nd ed., Amer. Math. Soc. Math. Surveys 3, AMS, Providence, RI, 1966.
[19] A. Melman, *Generalizations of Gershgorin disks and polynomial zeros*, Proc. Amer. Math. Soc., 138 (2010), pp. 2349–2364.
[20] A. Melman, *Modified Gershgorin disks for companion matrices*, SIAM Rev., 54 (2012), pp. 355–373.
[21] G. Meurant, *GMRES and the Arioli, Pták and Strakoš parametrization*, BIT, 52 (2012), pp. 687–702.
[22] G. Meurant, *On the location of the Ritz values in the Arnoldi process*, Electron. Trans. Numer. Anal., 43 (2015), pp. 188–212.
[23] G. Meurant and J. Duintjer Tebbens, *The role eigenvalues play in forming GMRES residual norms with non-normal matrices*, Numer. Algorithms, 68 (2015), pp. 143–165.
[24] G.V. Milovanović and T.M. Rassias, *Inequalities for polynomial zeros*, in Survey on Classical Inequalities, Springer, Netherlands, 2000, pp. 165–202.
[25] R.B. Morgan and M. Zeng, *Harmonic projection methods for large non-symmetric eigenvalue problems*, Numer. Linear Algebra Appl., 5 (1998), pp. 33–55.
[26] C.C. Paige, B.N. Parlett, and H. Van der Vorst, *Approximate solutions and eigenvalue bounds from Krylov subspaces*, Numer. Linear Algebra Appl., 2 (1995), pp. 115–133.
[27] Y. Saad, *Krylov subspace methods for solving large nonsymmetric linear systems*, Math. Comp., 37 (1981), pp. 105–126.
[28] Y. Saad, *The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems*, SIAM J. Numer. Anal., 19 (1982), pp. 485–506.
[29] Y. Saad, *Practical use of some Krylov subspace methods for solving indefinite and nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 5 (1984), pp. 203–228.
[30] Y. Saad and M.H. Schultz, *GMRES: A generalized minimum residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.
[31] G.W. Stewart, *Matrix Algorithms, Volume* II: *Eigensystems*, SIAM, Philadelphia, 2001.

Downloaded 01/24/17 to 128.138.73.68. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.