PRESCRIBING THE BEHAVIOR OF EARLY TERMINATING GMRES AND ARNOLDI ITERATIONS

JURJEN DUINTJER TEBBENS* AND GÉRARD MEURANT†

May 21, 2012

Abstract. We generalize the results of Arioli, Pták and Strakoš [BIT v 38 (1998)] and Duintjer Tebbens and Meurant [accepted in SIMAX (2012)] which describe the class of matrices with right-hand sides generating prescribed GMRES residual norm convergence curves as well as prescribed Ritz values when solving linear systems. These results assumed that the underlying Arnoldi orthogonalization processes are breakdown-free. We extend the results with parametrizations of classes of matrices with right-hand sides allowing the early termination case and also give analogues for the early termination case of other results related to the theory in Arioli, Pták and Strakoš [BIT v 38 (1998)].

1. Introduction. We consider solving linear systems

$$Ax = b, \tag{1.1}$$

where A is a nonsingular matrix of order n and b a given nonzero n-dimensional vector with the GMRES algorithm; see [10]. Assuming that the matrix A is non-derogatory and that GMRES terminates at iteration n, the results of a series of papers by Arioli, Greenbaum, Pták and Strakoš [4, 3, 1] show that for an arbitrary prescribed decreasing residual norm history there exists a class of matrices and right-hand sides that gives these residual norms. Moreover, the eigenvalues of those matrices can also be chosen freely. The last paper [1] of the series shows explicitly how to construct matrices and right-hand sides with prescribed residual norms and eigenvalues, see Theorem 2.1 and Corollary 2.4 in that paper. These results show that GMRES convergence for general matrices does not depend on the eigenvalues of A.

The GMRES algorithm is based on the Arnoldi process that generates upper Hessenberg matrices whose eigenvalues are known as Ritz values. In the Arnoldi method (see e.g. [8]), these values are used as approximations of the eigenvalues of A. Based on the results of [1], Duintjer Tebbens and Meurant [2] have shown that one can construct a class of matrices and right-hand sides with a prescribed residual norm convergence curve and prescribed Ritz values in *every* GMRES iteration, i.e. from the first until the nth iteration. This shows that there exists a class of matrices and right-hand sides for which the Ritz values generated in the iterations of the Arnoldi method (or, equivalently, of the GMRES method) can be arbitrary and fully independent of the spectrum and that they need not have any influence on the n residual norms generated in the GMRES method.

For practical problems one rarely computes all n iterations of an Arnoldi method or of GMRES if n is large. Often one will stop at a low iteration number with the value of the last subdiagonal entry of the Hessenberg matrix being below a small tolerance. Depending on the tolerance, this might be considered to correspond to early termination of the Arnoldi orthogonalization process in exact arithmetic. In this paper we would like to extend the above mentioned results to the early termination case when GMRES or Arnoldi terminates before iteration n. Some results for this

*Institute of Computer Science, Academy of Sciences of the Czech Republic. Pod Vodárenskou věží 2, 18 207 Praha 8 - Libeň, Czech Republic (duintjertebbens@cs.cas.cz). The work of J. Duintjer Tebbens is a part of the Institutional Research Plan AV0Z10300504 and it was supported by the project IAA100300802 of the Grant Agency of the ASCR.

†30 rue du sergent Bauchat, 75012 Paris, France (gerard.meurant@gmail.com). The work on this paper was started in 2011 during a visit of G. Meurant to the Institute of Computer Science of the Czech Republic Academy of Sciences

problem were described in the Ph.D. thesis of Liesen [5]; see also [6]. It was shown, for example, that any non-increasing GMRES convergence curve terminating before iteration n is possible with any spectrum. The conclusion of the paper [1] mentions that it is desirable to formulate the parametrizations of matrices and right-hand sides of that paper also for the early termination case. Some aspects of the early termination case are pointed out in the next to last section of that paper, but a parametrization is not given. In this paper we will give complete parametrizations of the matrices and right-hand sides giving a prescribed non-increasing GMRES convergence curve terminating before or at iteration n and, in addition, giving prescribed Ritz values in all iterations. We also prove some additional properties for the case with early termination similar to those proven in [7] for the case with termination at iteration n.

The contents of the paper are as follows. Section 2 first gives a new parametrization of the class of matrices and right-hand sides with a prescribed convergence curve and prescribed Ritz values with termination at iteration n. This result is then used to handle the case of early termination. It also shows how to practically construct these matrices. Section 3 generalizes the parametrization given in [1] to the case of early termination. In Section 4 we prove some properties of the matrices involved in the parametrizations and also give an expression for the GMRES iterates as well as the error vectors.

Throughout the paper we use the same notation as in [1] and [2] and e_i denotes the ith column of the identity matrix of appropriate dimension. The entry on position i, j of a matrix M is denoted as $m_{i,j}$. ¡in this paper we assume exact arithmetic; hence, early termination corresponds to a zero residual vector.

2. Prescribed Ritz values and GMRES residual norms with early termination. The Arnoldi orthogonalization process applied to an input matrix $A \in \mathbb{C}^{n\times n}$ with an initial nonzero vector $b \in \mathbb{C}^n$ yields, if it does not terminate before the nth iteration, the so-called Arnoldi decomposition

$$AV = VH, \qquad Ve_1 = b/\|b\|, \quad V^*V = I_n,$$

where H is an unreduced upper Hessenberg matrix containing the coefficients of the Arnoldi recursion and V is the orthormal matrix whose columns are basis vectors of the Krylov subspace $\mathcal{K}_n(A, b)$. If the orthogonalization process *does* break down at an iteration number $k, k < n$, this means that $h_{k+1,k} = 0$ and we obtain an Arnoldi decomposition which we will write as

$$AV_{n,k} = V_{n,k}H_k, \qquad V_{n,k}e_1 = b/\|b\|, \quad V_{n,k}^*V_{n,k} = I_k, \tag{2.1}$$

where $V_{n,k} \in \mathbb{C}^{n\times k}$ and $H_k \in \mathbb{C}^{k\times k}$ is an unreduced upper Hessenberg matrix of order k.

Since GMRES residual norms are invariant under unitary transformation of the linear system, the convergence curve generated by A and b is identical with the convergence curve generated by $H = V^*AV$ and $\|b\|e_1 = V^*b$. Thus essentially all information about Ritz values and residual norms is contained in H. Moreover, all information on the Ritz values and residual norms generated during the first k iterations must be contained in H_k. In order to characterize the Hessenberg matrices H_k that generate prescribed GMRES residual norms and prescribed Ritz values with early termination at the kth iteration, we therefore use a characterization of the Hessenberg matrices H with these properties when the Arnoldi orthogonalization process does *not* break down. Such a description was given in [2, Corollary 3.7], but here we introduce a simpler characterization. The proof of the fact that the Hessenberg matrices in this characterization generate the prescribed GMRES residual norms is new, but the proof of the fact that they generate the desired Ritz values is only a slight modification of the proof of [2, Proposition 2.1]. Nevertheless, we give this proof below for completeness. The next proposition can be seen as a complement of [2].

PROPOSITION 2.1. *Let the set*

$$\begin{aligned}\mathcal{R} = \{\, & \rho_1^{(1)}, \\ & (\rho_1^{(2)},\, \rho_2^{(2)}), \\ & \vdots \\ & (\rho_1^{(n-1)}, \dots, \rho_{n-1}^{(n-1)}), \\ & (\lambda_1, \dots\dots\dots, \lambda_n)\, \},\end{aligned}$$

represent any choice of $n(n+1)/2$ complex Ritz values. An unreduced upper Hessenberg matrix H has the spectrum $\lambda_1, \dots, \lambda_n$ and its kth leading principal submatrix has eigenvalues $\rho_1^{(k)}, \dots, \rho_k^{(k)}$ for all $k = 1, \dots, n-1$ if and only if it has the form

$$H = \begin{bmatrix} & g^T & \\ 0 & & T \end{bmatrix}^{-1} C^{(n)} \begin{bmatrix} & g^T & \\ 0 & & T \end{bmatrix}, \tag{2.2}$$

where $C^{(n)}$ is the companion matrix of the polynomial with roots $\lambda_1, \dots, \lambda_n$

$$C^{(n)} = \begin{bmatrix} 0 & -\alpha_0 \\ I_{n-1} & \vdots \\ & -\alpha_{n-1} \end{bmatrix},$$

the first entry g_1 of the vector g is nonzero, T is nonsingular upper triangular of order $n-1$ and if

$$q_k(\lambda) = [1, \lambda, \dots, \lambda^k] \begin{bmatrix} & g^T & \\ 0 & & T \end{bmatrix} e_{k+1},$$

then $q_k(\lambda)$ is a polynomial with roots $\rho_1^{(k)}, \dots, \rho_k^{(k)}$ for $k = 1, \dots, n-1$.

Proof. Clearly, H is unreduced upper Hessenberg and its spectrum is $\lambda_1, \dots, \lambda_n$ by the definition of $C^{(n)}$. We will show that the spectrum of the $k \times k$ leading principal submatrix of H is $\rho_1^{(k)}, \dots, \rho_k^{(k)}$. Let U be the nonsingular upper triangular matrix

$$\begin{bmatrix} & g^T & \\ 0 & & T \end{bmatrix}$$

and let U_k denote the $k \times k$ leading principal submatrix of U. Also, for $j > k$, let $\tilde{u}_j$ denote the vector of the first k entries of the jth column of U^{-1}. The spectrum of the $k \times k$ leading principal submatrix of H is the spectrum of

$$[I_k, 0]\, U^{-1} C^{(n)} U \begin{bmatrix} I_k \\ 0 \end{bmatrix} = [U_k^{-1}, \tilde{u}_{k+1}, \dots, \tilde{u}_n] \begin{bmatrix} 0 \\ U_k \\ 0 \end{bmatrix} = [U_k^{-1}, \tilde{u}_{k+1}] \begin{bmatrix} 0 \\ U_k \end{bmatrix}.$$

It is also the spectrum of the matrix

$$U_k [U_k^{-1}, \tilde{u}_{k+1}] \begin{bmatrix} 0 \\ U_k \end{bmatrix} U_k^{-1} = [I_k, U_k \tilde{u}_{k+1}] \begin{bmatrix} 0 \\ I_k \end{bmatrix},$$

which is a companion matrix with last column $U_k\tilde{u}_{k+1}$. From

$$e_{k+1} = U_{k+1}U_{k+1}^{-1}e_{k+1} = \begin{bmatrix} U_k & \begin{matrix} g_{k+1} \\ t_{1,k} \\ \vdots \end{matrix} \\ 0 & t_{k,k} \end{bmatrix} \begin{bmatrix} U_k^{-1} & \tilde{u}_{k+1} \\ 0 & 1/t_{k,k} \end{bmatrix} e_{k+1} = \begin{bmatrix} U_k\tilde{u}_{k+1} + \begin{bmatrix} g_{k+1}/t_{k,k} \\ t_{1,k}/t_{k,k} \\ \vdots \\ t_{k-1,k}/t_{k,k} \end{bmatrix} \\ 1 \end{bmatrix},$$

we obtain that the coefficients corresponding to λ^0 till λ^{k-1} of the polynomial with roots $\rho_1^{(k)}, \dots, \rho_k^{(k)}$ are the entries of $U_k\tilde{u}_{k+1}$. $\square$

The vector g in (2.2) is arbitrary (except for that its first entry must be nonzero). In the next theorem we show that g can be chosen such that it forces prescribed GMRES residual norms when GMRES is applied to H with right hand side $\|b\|e_1$. This immediately gives a parametrization of the class of matrices and right-hand sides such that GMRES generates residual norms and prescribed Ritz values in all iterations.

THEOREM 2.2. *Consider a set of tuples of complex numbers*

$$\begin{aligned} \mathcal{R} = \{&\rho_1^{(1)}, \\ &(\rho_1^{(2)}, \rho_2^{(2)}), \\ &\vdots \\ &(\rho_1^{(n-1)}, \dots, \rho_{n-1}^{(n-1)}), \\ &(\lambda_1, \dots\dots\dots, \lambda_n)\}, \end{aligned}$$

such that $(\lambda_1, \dots, \lambda_n)$ contains no zero number and n positive numbers

$$f(0) \ge f(1) \ge \cdots \ge f(n-1) > 0,$$

such that $f(k-1) = f(k)$ if and only if the k-tuple $(\rho_1^{(k)}, \dots, \rho_k^{(k)})$ contains a zero number. If A is a matrix of order n and b a nonzero n-dimensional vector, then the following assertions are equivalent:

1. *The GMRES method applied to A and right-hand side b with zero initial guess yields residuals $r^{(k)}$, $k = 0, \dots, n-1$ such that*

   $$\|r^{(k)}\| = f(k), \quad k = 0, \dots, n-1,$$

   A has eigenvalues $\lambda_1, \dots, \lambda_n$ and $\rho_1^{(k)}, \dots, \rho_k^{(k)}$ are the eigenvalues of the kth leading principal submatrix of the generated Hessenberg matrix for all $k = 1, \dots, n-1$.
2. *The matrix A and the right-hand side b are of the form*

   $$A = V \begin{bmatrix} & g^T \\ 0 & T \end{bmatrix}^{-1} C^{(n)} \begin{bmatrix} & g^T \\ 0 & T \end{bmatrix} V^*, \qquad b = f(0)Ve_1, \tag{2.3}$$

   where V is any unitary matrix, $C^{(n)}$ is the companion matrix of the polynomial with roots $\lambda_1, \dots, \lambda_n$,

   $$g_1 = \frac{1}{f(0)}, \qquad g_k = \frac{\sqrt{f(k-2)^2 - f(k-1)^2}}{f(k-2)f(k-1)}, \qquad k = 2, \dots, n$$

and T is nonsingular upper triangular of order $n-1$ such, that if

$$q_k(\lambda) = [1, \lambda, \ldots, \lambda^k] \left[\begin{array}{cc} \multicolumn{2}{c}{g^T} \\ 0 & T \end{array} \right] e_{k+1},$$

then $q_k(\lambda)$ is a polynomial with roots $\rho_1^{(k)}, \ldots, \rho_k^{(k)}$ for $k = 1, \ldots, n-1$.

Proof. We have to show three claims, namely that we have the parametrization (2.3) if and only if we have the prescribed spectrum, the prescribed Ritz values and the prescribed GMRES residual norms of the first assertion. The first claim follows immediately from the definition of the companion matrix $C^{(n)}$. The second claim follows from Proposition 2.1, because the Hessenberg matrix generated by GMRES applied to A and b in (2.3) is

$$H = \left[\begin{array}{cc} \multicolumn{2}{c}{g^T} \\ 0 & T \end{array} \right]^{-1} C^{(n)} \left[\begin{array}{cc} \multicolumn{2}{c}{g^T} \\ 0 & T \end{array} \right].$$

Let us consider the last claim on GMRES residual norms. If the QR decomposition $H = QR$ of H is computed with Givens rotations that zero out the subsequent subdiagonal entries of H, the individual rotation parameters give the residual norms. More precisely, if the kth subdiagonal entry was eliminated with Givens cosine c_{k-1} and sine s_{k-1}, then

$$\|r^{(k)}\| = \|b\| \prod_{j=1}^{k} |s_j| = f(0) \prod_{j=1}^{k} |s_j|, \tag{2.4}$$

see, e.g., [9, Section 6.5.5, p. 166]. The Q factor of the QR decomposition $H = QR$ of H is the same as the Q factor of the QR decomposition of

$$\left[\begin{array}{cc} \multicolumn{2}{c}{g^T} \\ 0 & T \end{array} \right]^{-1} C^{(n)}$$

because $\left[\begin{array}{cc} \multicolumn{2}{c}{g^T} \\ 0 & T \end{array} \right]$ is nonsingular upper triangular. If we define $\hat{g} = [g_2, \ldots, g_n]^T$ and write

$$\left[\begin{array}{cc} \multicolumn{2}{c}{g^T} \\ 0 & T \end{array} \right]^{-1} C^{(n)} = \left[\begin{array}{cc} 1 & \hat{g}^T/g_1 \\ 0 & I \end{array} \right]^{-1} \left[\begin{array}{cc} g_1 & 0 \\ 0 & T \end{array} \right]^{-1} C^{(n)} = \left[\begin{array}{cc} 1 & \hat{g}^T/g_1 \\ 0 & I \end{array} \right]^{-1} C^{(n)} \hat{R},$$

then it can be easily checked that

$$\hat{R} = (C^{(n)})^{-1} \left[\begin{array}{cc} 1/g_1 & 0 \\ 0 & T^{-1} \end{array} \right] C^{(n)}$$

is upper triangular. Hence in fact, the Q factor of the QR decomposition $H = QR$ of H is the same as the Q factor of the QR decomposition of

$$\hat{H} \equiv \left[\begin{array}{cc} 1 & \hat{g}^T/g_1 \\ 0 & I \end{array} \right]^{-1} C^{(n)} = \left[\begin{array}{cc} 1 & -\hat{g}^T/g_1 \\ 0 & I \end{array} \right] C^{(n)}.$$

Let us zero out the first subdiagonal entry of the upper Hessenberg matrix $\hat{H}$. With $\hat{h}_{1,1} = -g_2/g_1$ and $\hat{h}_{2,1} = 1$ we obtain the Givens cosine and sine satisfying

$$|c_1| = \frac{|g_2/g_1|}{\sqrt{1 + (g_2/g_1)^2}}, \qquad |s_1| = \frac{1}{\sqrt{1 + (g_2/g_1)^2}}.$$

Thus

$$|s_1| = \frac{1}{\sqrt{1+\frac{f(0)^2-f(1)^2}{f(1)^2}}} = \frac{f(1)}{f(0)}$$

and with (2.4) we have $\|r^{(1)}\| = f(1)$ as desired. Now assume $|s_j| = \frac{f(j)}{f(j-1)}$ for $j = 1, \dots, k$. Then the application of all previous k Givens rotations to the $(k+1)$st column of $\hat{H}$, that is to the vector $[-g_{k+2}/g_1, 0, \dots, 0, 1, 0, \dots, 0]^T$, yields a vector whose $(k+1)$st entry is $-\prod_{j=1}^{k}(-s_j)g_{k+2}/g_1$ and its $(k+2)$nd entry is 1. Then we obtain the Givens cosine and sine

$$|c_{k+1}| = \frac{\prod_{j=1}^{k}|s_j||g_{k+2}/g_1|}{\sqrt{1+\prod_{j=1}^{k}s_j^2(g_{k+2}/g_1)^2}}, \qquad |s_{k+1}| = \frac{1}{\sqrt{1+\prod_{j=1}^{k}s_j^2(g_{k+2}/g_1)^2}}.$$

Thus

$$|s_{k+1}| = \left(1+(g_{k+2}/g_1)^2\prod_{j=1}^{k}s_j^2\right)^{-\frac{1}{2}} = \left(1+g_{k+2}^2 f(k)^2\right)^{-\frac{1}{2}} = \left(1+\frac{f(k)^2-f(k+1)^2}{f(k+1)^2}\right)^{-\frac{1}{2}} = \frac{f(k+1)}{f(k)}$$

and with (2.4) we have $\|r^{(k+1)}\| = f(k+1)$ as desired. $\square$

Theorem 2.2 shows how to construct a Hessenberg matrix H giving n prescribed GMRES residual norms and giving the prescribed Ritz values of all n iterations. Now, it suffices to consider the first k columns of H to prescribe the behavior of GMRES terminating at iteration number k. The remaining columns are fully discarded in this case, see (2.1), and can be chosen arbitrarily. This observation results in the next theorem.

Theorem 2.3. *Consider a set of tuples of complex numbers*

$$\begin{aligned}\mathcal{R} = \{&\rho_1^{(1)},\\ &(\rho_1^{(2)}, \rho_2^{(2)}),\\ &\vdots\\ &(\rho_1^{(k-1)}, \dots, \rho_{k-1}^{(k-1)}),\\ &(\lambda_1, \dots\dots\dots, \lambda_k)\},\end{aligned}$$

such that $(\lambda_1, \dots, \lambda_k)$ contains no zero number and k positive numbers

$$f(0) \geq f(1) \geq \cdots \geq f(k-1) > 0,$$

such that $f(j-1) = f(j)$ if and only if the j-tuple $(\rho_1^{(j)}, \dots, \rho_j^{(j)})$ contains a zero number. If A is a matrix of order n and b a nonzero n-dimensional vector, then the following assertions are equivalent:

1. *The GMRES method applied to A and right-hand side b with zero initial guess yields residuals $r^{(j)}$ such, that*

$$\|r^{(j)}\| = f(j), \quad j = 0, \dots, k-1, \qquad \|r^{(j)}\| = 0, \quad j = k, \dots, n,$$

the spectrum of A contains the eigenvalues $\lambda_1, \dots, \lambda_k$ and $\rho_1^{(j)}, \dots, \rho_j^{(j)}$ are the eigenvalues of the jth leading principal submatrix of the generated Hessenberg matrix for all $j = 1, \dots, k-1$.

2. *The matrix A and the right-hand side b are of the form*

$$A = V \begin{bmatrix} H_k & B \\ 0 & D \end{bmatrix} V^*, \qquad b = f(0) V e_1, \tag{2.5}$$

where V is any unitary matrix and $B \in \mathbb{C}^{k\times(n-k)}$, $D \in \mathbb{C}^{(n-k)\times(n-k)}$ are submatrices with arbitrary entries. The unreduced upper Hessenberg matrix H_k has the form

$$H_k = \begin{bmatrix} & \tilde{g}^T \\ 0 & T_{k-1} \end{bmatrix}^{-1} C^{(k)} \begin{bmatrix} & \tilde{g}^T \\ 0 & T_{k-1} \end{bmatrix},$$

with $C^{(k)}$ being the companion matrix of the polynomial with roots $\lambda_1, \ldots, \lambda_k$, with the k-dimensional vector $\tilde{g}$ being defined as

$$\tilde{g}_1 = \frac{1}{f(0)}, \qquad \tilde{g}_j = \frac{\sqrt{f(j-2)^2 - f(j-1)^2}}{f(j-2)f(j-1)}, \qquad j = 2, \ldots, k$$

and with a nonsingular upper triangular matrix T_{k-1} of order $k-1$ such, that if

$$q_j(\lambda) = [1, \lambda, \ldots, \lambda^j] \begin{bmatrix} & \tilde{g}^T \\ 0 & T_{k-1} \end{bmatrix} e_{j+1},$$

then $q_j(\lambda)$ is a polynomial with roots $\rho_1^{(j)}, \ldots, \rho_j^{(j)}$ for $j = 1, \ldots, k-1$.

Theorem 2.3 gives a complete parametrization of the matrices with right hand sides generating a prescribed GMRES residual norm history with prescribed Ritz values and allowing the early termination case. Of course, it holds for $k = n$, too. Note that the system matrix A in (2.5) is allowed to be singular, because B, D are fully arbitrary. For example, B and D can both be zero matrices.

3. Early termination and the parametrization of [1]. In this section we address the issues and the open question on early termination formulated in [1]. The authors were concerned with prescribing GMRES residual norms only, not Ritz values. The central question is whether the following parametrization, which is the main result of [1], can be extended to the early termination case.

THEOREM 3.1 (see [1]). *Assume we are given $n+1$ nonnegative numbers*

$$f(0) \geq f(1) \geq \cdots \geq f(n-1) > 0, \quad f(n) = 0$$

and n complex numbers $\lambda_1, \ldots, \lambda_n$ all different from 0. The following assertions are equivalent:

1. *The spectrum of A is $\{\lambda_1, \ldots, \lambda_n\}$ and GMRES applied to A and b with zero initial guess yields residuals r_k, $k = 0, \ldots, n-1$ such that*

$$\|r_k\| = f(k), \quad k = 0, \ldots, n.$$

2. *The matrix A is of the form*

$$A = WYC^{(n)}Y^{-1}W^*$$

and $b = Wh$, where W is any unitary matrix, the matrix Y is given by

$$Y = \begin{bmatrix} h & \begin{matrix} R \\ 0 \end{matrix} \end{bmatrix}, \tag{3.1}$$

R being any nonsingular upper triangular matrix of order $n-1$, h a vector describing the convergence curve such that

$$h = [\eta_1, \dots, \eta_n]^T, \qquad \eta_k = (f(k-1)^2 - f(k)^2)^{1/2}, \quad k < n, \quad \eta_n = f(n-1) \tag{3.2}$$

and $C^{(n)}$ is the companion matrix corresponding to the polynomial $q(\lambda)$ defined as

$$q(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n) = \lambda^n + \sum_{j=0}^{n-1} \alpha_j \lambda^j.$$

In [1, Section 3] some aspects involved in generalizing this theorem were described. In particular, properties of the components of b in the Jordan canonical vector basis of A and the relation with the minimal polynomial of A with respect to b were investigated for the early termination case. Note that Theorem 2.3 gives the minimal polynomial of A with respect to b in the early termination case; it is the polynomial with roots $\lambda_1, \dots, \lambda_k$ which takes the value one at the origin.

In the following theorem we give a direct, brute force generalization of Theorem 3.1 to the early termination case. It may look rather technical but we have chosen this formulation, close to that in Theorem 3.1, to emphasize the instances where Theorem 3.1 is modified. It reveals not only the minimal polynomial of A with respect to b but also the minimal polynomial of A itself.

THEOREM 3.2. *Assume we are given k positive numbers*

$$f(0) \geq f(1) \geq \cdots \geq f(k-1) > 0$$

and m distinct complex numbers $\lambda_1, \dots, \lambda_m$, all different from 0. The following assertions are equivalent for a matrix A of order n having m distinct eigenvalues and $n \geq k$, $n \geq m$,:

1. *$\lambda_1, \dots, \lambda_m$ are eigenvalues of A and GMRES applied to A and b with a zero initial guess yields residuals r_j, $j = 0, \dots, n$ such that*

$$\|r_j\| = f(j), \quad j = 0, \dots, k-1, \quad \|r_j\| = 0, \quad j = k, \dots, n.$$

2. *The right-hand side b is of the form $b = W_{n,k} h_k$ where $W_{n,k} \in \mathbb{C}^{n \times k}$ has orthonormal columns and*

$$\begin{aligned} h_k &= [\eta_1, \dots, \eta_k]^T, \\ \eta_j &= (f(j-1)^2 - f(j)^2)^{1/2}, \quad 1 \leq j < k, \qquad \eta_k = f(k-1). \end{aligned} \tag{3.3}$$

The matrix A satisfies the equation

$$A W_{n,k} Y_{k,\ell} = W_{n,k} Y_{k,\ell} C^{(\ell)}, \tag{3.4}$$

where $C^{(\ell)} \in \mathbb{C}^{\ell \times \ell}$ is the companion matrix corresponding to a polynomial

$$q(\lambda) = \Pi_{j=1}^m (\lambda - \lambda_j)^{\ell_j} = \lambda^\ell + \sum_{j=0}^{\ell-1} \alpha_j \lambda^j$$

with integers $\ell_j > 0$ such that $\sum_{j=1}^m \ell_j = \ell \geq k$. The matrix $Y_{k,\ell} \in \mathbb{C}^{k \times \ell}$ is given by

$$Y_{k,\ell} = \begin{bmatrix} \eta_1 & & \\ \vdots & R_{k-1} & \hat{R} \\ \eta_k & 0 & 0 \end{bmatrix},$$

with R_{k-1} being a nonsingular upper triangular matrix of order $k-1$ and $\hat{R} \in \mathbb{C}^{k\times(\ell-k)}$ being the matrix whose columns are given recursively through the relations

$$\hat{R}e_i = Y_{k,\ell}\,[e_i,\ldots,e_{i+k-1}]\begin{bmatrix} -\beta_0 \\ \vdots \\ -\beta_{k-1} \end{bmatrix}, \quad i = 1,\ldots,\ell-k, \tag{3.5}$$

for coefficients $\beta_0,\ldots,\beta_{k-1}$ of a polynomial $p(\lambda)$ of degree k of the form

$$p(\lambda) = \Pi_{j=1}^{m}(\lambda-\lambda_j)^{\tilde{\ell}_j} = \lambda^k + \sum_{j=0}^{k-1}\beta_j\lambda^j, \tag{3.6}$$

with $0 \le \tilde{\ell}_j \le \ell_j$.

Proof. The proof follows closely the proof of Theorem 2.1 and Proposition 2.4 of [1].

Let us first prove that $1 \to 2$. Let $q(\lambda)$ be the minimal polynomial of A, $q(\lambda) = \Pi_{j=1}^{m}(\lambda-\lambda_j)^{\ell_j}$ with integers $\ell_j > 0$ such that $\sum_{j=1}^{m}\ell_j = \ell \ge k$, let

$$z = [\zeta_1,\ldots,\zeta_\ell]^T, \quad \text{where} \quad \frac{q(\lambda)}{(-1)^\ell\prod_{j=1}^{m}\lambda_j^{\ell_j}} = 1-(\zeta_1\lambda+\ldots+\zeta_\ell\lambda^\ell)$$

and define $K_{n,\ell} \equiv [b, Ab, \ldots, A^{\ell-1}b]$ and $B_{n,\ell} = [Ab, A^2b, \ldots, A^\ell b]$. From $q(A) = 0$ and $q(A)b = 0$ we get

$$AK_{n,\ell} = K_{n,\ell}C^{(\ell)}, \qquad AB_{n,\ell} = B_{n,\ell}C^{(\ell)}, \qquad b = B_{n,\ell}z. \tag{3.7}$$

Consider a QR decomposition of $B_{n,\ell}$, $B_{n,\ell} = \tilde{W}\tilde{R}_{n,\ell}$ with $\tilde{W} \in \mathbb{C}^{n\times n}$ and $\tilde{R}_{n,\ell} \in \mathbb{C}^{n\times\ell}$. Because GMRES terminates at the kth iteration, $A^{k+i}b$ is linearly dependent on $Ab,\ldots,A^kb$ for all $i > 0$ and the rows $k+1$ until n of $\tilde{R}_{n,\ell}$ must be zero. The prescribed residual norms imply that

$$b = \tilde{W}\Gamma h, \quad h = \begin{pmatrix} h_k \\ 0 \end{pmatrix}$$

where Γ is a diagonal unitary matrix and h_k is defined by (3.3), see [3, p. 466]. Define $W \equiv \tilde{W}\Gamma$ and $R_{n,\ell} \equiv \Gamma^*\tilde{R}_{n,\ell}$. Then $b = Wh$ as desired. Furthermore we have

$$AWR_{n,\ell} = AB_{n,\ell} = B_{n,\ell}C^{(\ell)} = WR_{n,\ell}C^{(\ell)}$$

and

$$WR_{n,\ell}z = B_{n,\ell}z = b = Wh, \quad \text{i.e.} \quad R_{n,\ell}z = h. \tag{3.8}$$

Then from (3.7) we have $B_{n,\ell} = AK_{n,\ell} = K_{n,\ell}C^{(\ell)}$, i.e. $K_{n,\ell} = B_{n,\ell}[C^{(\ell)}]^{-1}$. With $B_{n,\ell} = WR_{n,\ell}$ it follows that $K_{n,\ell} = WR_{n,\ell}[C^{(\ell)}]^{-1}$ and with (3.7) that

$$AWR_{n,\ell}[C^{(\ell)}]^{-1} = (WR_{n,\ell}[C^{(\ell)}]^{-1})C^{(\ell)}.$$

Define the matrix $Y_{n,\ell} \in \mathbb{C}^{n\times\ell}$ as $Y_{n,\ell} \equiv R_{n,\ell}[C^{(\ell)}]^{-1}$ and note that $[C^{(\ell)}]^{-1} = \begin{bmatrix} z & \begin{matrix} I_{\ell-1} \\ 0 \end{matrix} \end{bmatrix}$. Because of $R_{n,\ell}z = h$, $Y_{n,\ell}$ has the form $Y_{n,\ell} = [h, R_{n,\ell}e_1, \ldots, R_{n,\ell}e_{\ell-1}]$. However, the rows $k+1$ to n of $Y_{n,\ell}$ must be zero (so are the corresponding rows of $R_{n,\ell}$) and thus $Y_{n,\ell}$ has the form

$$Y_{n,\ell} = \begin{bmatrix} \eta_1 & & \\ \vdots & R_{k-1} & \hat{R} \\ \eta_k & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where R_{k-1} denotes the leading principal submatrix of order $k-1$ of $R_{n,\ell}$ and $\hat{R} \in \mathbb{C}^{k\times(\ell-k)}$ the remaining nonzero part of $R_{n,\ell}$. Denoting the first k columns of W by $W_{n,k}$ and the first k rows of $Y_{n,\ell}$ by $Y_{k,\ell}$, we have $W_{n,k}Y_{k,\ell} = WY_{n,\ell}$. Then we obtain equation (3.4) from

$$AW_{n,k}Y_{k,\ell} = AWY_{n,\ell} = AWR_{n,\ell}[C^{(\ell)}]^{-1} = (WR_{n,\ell}[C^{(\ell)}]^{-1})C^{(\ell)} = WY_{n,\ell}C^{(\ell)} = W_{n,k}Y_{k,\ell}C^{(\ell)}.$$

Finally, because GMRES terminates at the kth iteration, the minimal polynomial of A with respect to b is a polynomial $p(\lambda)$ of degree k with $p(A)b = 0$ which is a divisor of $q(\lambda)$. If we write $p(\lambda)$ as $p(\lambda) = \lambda^k + \sum_{j=0}^{k-1} \beta_j \lambda^j$ and since h_k denotes the first k entries of h, then with $[b, Ab, \dots, A^{k-1}b] = W_{n,k}\begin{bmatrix} h_k & \begin{matrix} R_{k-1} \\ 0 \end{matrix} \end{bmatrix}$ we have

$$A^k b = -\sum_{j=0}^{k-1} \beta_j A^j b = W_{n,k}\begin{bmatrix} h_k & \begin{matrix} R_{k-1} \\ 0 \end{matrix} \end{bmatrix}\begin{bmatrix} -\beta_0 \\ \vdots \\ -\beta_{k-1} \end{bmatrix} = W_{n,k}Y_{k,\ell}[e_1, \dots, e_k]\begin{bmatrix} -\beta_0 \\ \vdots \\ -\beta_{k-1} \end{bmatrix}.$$

Because $A^k b = B_{n,\ell}e_k = W_{n,k}\hat{R}e_1$, this shows the first condition in (3.5) for $i = 1$. Recursively we obtain

$$\begin{aligned} A^{k+i-1}b = A^{i-1}(A^k b) &= -\sum_{j=0}^{k-1} \beta_j A^{j+i-1} b \\ &= [b, Ab, \dots, A^{k-1}b]\begin{bmatrix} 0 \\ \vdots \\ 0 \\ -\beta_0 \\ \vdots \\ -\beta_{k-i-1} \end{bmatrix} - W_{n,k}\left(\beta_{k-i}\hat{R}e_1 + \dots + \beta_{k-1}\hat{R}e_i\right), \end{aligned}$$

for $i = 2, \dots, \ell-k$. Using $A^{k+i-1}b = B_{n,\ell}e_{k+i-1} = W_{n,k}\hat{R}e_i$ one obtains the remaining conditions in (3.5).

Now, let us consider the implication $2 \to 1$. Denote the eigenpairs of $C^{(\ell)}$ by $\{\lambda_i, y_i\}$ for $i = 1, \dots, m$. Then

$$AW_{n,k}Y_{k,\ell}y_i = W_{n,k}Y_{k,\ell}C^{(\ell)}y_i = \lambda_i W_{n,k}Y_{k,\ell}y_i,$$

hence λ_i is an eigenvalue of A for $i = 1, \dots, m$ and these m distinct eigenvalues are the only distinct eigenvalues by the assumptions of the theorem. To show that GMRES generates the nonzero residual norms $f(0), \dots, f(k-1)$ it suffices to show that $W_{n,k}$ is a unitary basis of $A\mathcal{K}_k(A, b)$, see [3, p. 466]. First we introduce the notation $C_p^{(k)}$ for the companion matrix of the polynomial $p(\lambda)$ in (3.6) and Y_k for the first k columns of $Y_{k,\ell}$. Then if we equate the first k columns in (3.4) and $k < \ell$, we obtain

$$AW_{n,k}Y_k = W_{n,k}Y_{k,\ell}[e_2, \dots, e_{k+1}] = W_{n,k}Y_kC_p^{(k)} \tag{3.9}$$

because of (3.5). In case $k = \ell$, the polynomials $q(\lambda)$ and $p(\lambda)$ are identical and we also obtain

$$AW_{n,k}Y_k = W_{n,k}Y_{k,\ell}C^{(\ell)}[e_1, \dots, e_k] = W_{n,k}Y_{k,\ell}C^{(k)} = W_{n,k}Y_kC_p^{(k)}.$$

We will prove that $W_{n,k}$ is a unitary basis of $A\mathcal{K}_k(A, b)$ by induction. We have

$$Ab = AWh = AW_{n,k}h_k = AW_{n,k}Y_ke_1 = W_{n,k}Y_kC_p^{(k)}e_1 = W_{n,k}Y_ke_2 = r_{1,1}W_{n,k}e_1 = r_{1,1}w_1,$$

where w_1 is the first column of W and the entries of $R_{n,l}$ are denoted as $r_{i,j}$. Because R_{k-1} is nonsingular, $r_{1,1} \neq 0$. Now let $A^{j-1}b = W_{n,k}Y_k e_j$ be the assumption of the induction. Then

$$A^j b = AW_{n,k}Y_k e_j = W_{n,k}Y_k C_p^{(k)} e_j = W_{n,k}Y_k e_{j+1} = W_{n,k}R_{k-1}e_j = W_{n,k}[r_{1,j}, \dots, r_{j,j}, 0, \dots, 0]^T,$$

where $r_{j,j} \neq 0$ for $j \leq k-1$. For $j = k$ we have

$$A^k b = AW_{n,k}Y_k e_k = W_{n,k}Y_k C_p^{(k)} e_k = -W_{n,k}Y_k[\beta_0, \dots, \beta_{k-1}]^T,$$

and

$$-(e_k)^T Y_k[\beta_0, \dots, \beta_{k-1}]^T = -\beta_0 \eta_k = (-1)^{k+1}\eta_k \prod_{j=1}^{m} \lambda_j^{\tilde{\ell}_j} \neq 0.$$

It proves that $W_{n,k}$ is a basis of $A\mathcal{K}_k(A,b)$. □

Theorem 3.2 does not describe how to construct the matrices A generating a prescribed convergence curve terminating at the kth iteration. It only gives the condition (3.4) that such a matrix A must satisfied. The next result shows how to construct A.

THEOREM 3.3. *Under the assumptions and with the notation of Theorem 3.2, the assertions 1- and 2- are equivalent to*

3- The matrix A is of the form

$$A = W \begin{bmatrix} Y_k C_p^{(k)} Y_k^{-1} & H_{1,2} \\ 0 & H_{2,2} \end{bmatrix} W^* \tag{3.10}$$

where W is unitary, $C_p^{(k)}$ is the companion matrix for the polynomial $p(\lambda)$ from (3.6),

$$C_p^{(k)} = \begin{bmatrix} 0 & -\beta_0 \\ I_{k-1} & \vdots \\ & -\beta_{k-1} \end{bmatrix}, \tag{3.11}$$

Y_k is the principal submatrix of order k of $Y_{k,\ell}$ that is,

$$Y_k = \begin{bmatrix} h_k & \begin{matrix} R_{k-1} \\ 0 \end{matrix} \end{bmatrix}$$

and the union of the spectra of $C_p^{(k)}$ and $H_{2,2}$ is $\{\lambda_1, \dots, \lambda_m\}$. The right-hand side b is of the form $b = W_{n,k}h_k$ where $W_{n,k} \in \mathbb{C}^{n \times k}$ contains the first k columns of W.

Proof. Let us prove that $2 \to 3$.

From the second assertion of Theorem 3.2 we have that A satisfies $AW_{n,k}Y_k = W_{n,k}Y_k C_p^{(k)}$, see (3.9). Then for a matrix $\tilde{W} \in \mathbb{C}^{n \times n-k}$ such that $[W_{n,k}, \tilde{W}]$ is unitary, A also satisfies

$$A[W_{n,k}, \tilde{W}] \begin{bmatrix} Y_k & 0 \\ 0 & I_{n-k} \end{bmatrix} = [W_{n,k}, \tilde{W}] \begin{bmatrix} Y_k C_p^{(k)} & W_{n,k}^* A\tilde{W} \\ 0 & \tilde{W}^* A\tilde{W} \end{bmatrix}.$$

With the notation $W \equiv [W_{n,k}, \tilde{W}]$, $H_{1,2} \equiv W_{n,k}^* A\tilde{W}$ and $H_{2,2} \equiv \tilde{W}^* A\tilde{W}$ this immediately gives (3.10). At the beginning of the proof of the implication $2 \to 1$ in Theorem 3.2 is was shown that the distinct eigenvalues of A are $\lambda_1, \ldots, \lambda_m$. Therefore the union of the spectra of $C_p^{(k)}$ and $H_{2,2}$ is $\{\lambda_1, \ldots, \lambda_m\}$.

To prove $3 \to 1$, we first note that by assumption the union of the spectra of $C_p^{(k)}$ and $H_{2,2}$ is $\{\lambda_1, \ldots, \lambda_m\}$ and therefore A has distinct eigenvalues $\{\lambda_1, \ldots, \lambda_m\}$. Now it suffices to show that $W_{n,k}$ is a unitary basis of $A\mathcal{K}_k(A, b)$, see [3, p. 466]. We will prove this again by induction. We have, using (3.10),

$$Ab = W \begin{bmatrix} Y_k C_p^{(k)} Y_k^{-1} & H_{1,2} \\ 0 & H_{2,2} \end{bmatrix} W^* b = W \begin{bmatrix} Y_k C_p^{(k)} Y_k^{-1} h_k \\ 0 \end{bmatrix} = W \begin{bmatrix} Y_k e_2 \\ 0 \end{bmatrix} = r_{1,1} w_1.$$

Now let $A^{j-1}b = W_{n,k} Y_k e_j$ be the induction assumption. Then if $j < k$,

$$A^j b = W \begin{bmatrix} Y_k C_p^{(k)} Y_k^{-1} & H_{1,2} \\ 0 & H_{2,2} \end{bmatrix} W^* W_{n,k} Y_k e_j = W \begin{bmatrix} Y_k C_p^{(k)} Y_k^{-1} Y_k e_j \\ 0 \end{bmatrix} = W \begin{bmatrix} Y_k e_{j+1} \\ 0 \end{bmatrix}$$

and we have

$$W \begin{bmatrix} Y_k e_{j+1} \\ 0 \end{bmatrix} = W_{n,k} \begin{bmatrix} R_{k-1} e_j \\ 0 \end{bmatrix}$$

with $r_{j,j} \neq 0$. If $j = k$,

$$A^k b = W \begin{bmatrix} Y_k C_p^{(k)} Y_k^{-1} & H_{1,2} \\ 0 & H_{2,2} \end{bmatrix} W^* W_{n,k} Y_k e_k = W \begin{bmatrix} Y_k C_p^{(k)} Y_k^{-1} Y_k e_k \\ 0 \end{bmatrix} = W \begin{bmatrix} Y_k C_p^{(k)} e_k \\ 0 \end{bmatrix}$$

and with the definition (3.6) of the polynomial p, we have

$$W \begin{bmatrix} Y_k C_p^{(k)} e_k \\ 0 \end{bmatrix} = W_{n,k} Y_k \begin{bmatrix} -\beta_0 \\ \vdots \\ -\beta_{k-1} \end{bmatrix},$$

where the last entry of $Y_k \left[\beta_0, \ldots, \beta_{k-1}\right]^T$ is $\beta_0 \eta_k \neq 0$, see (3.3). □

Let us summarize the degrees of freedom in constructing b and A with prescribed distinct eigenvalues and with termination in step k according to (3.10) in Theorem 3.3. The unitary matrix W is chosen arbitrarily. The non-singular upper triangular matrix R_{k-1} contained in Y_k is arbitrary. The companion matrix $C_p^{(k)}$ is constructed from an arbitrary polynomial $p(\lambda)$ of degree k whose roots belong to the prescribed distinct eigenvalues. The matrix $H_{1,2}$ is fully arbitrary and $H_{2,2}$ is arbitrary except that its spectrum must guarantee that the union of the spectrum with the roots of $p(\lambda)$ add up to the complete set of prescribed distinct eigenvalues.

Should we compare to the V parametrization?

4. Some additional properties. In this section we generalize some relations and properties satisfied by the matrices in the parametrization of [1] that were proved in [7] for termination at iteration n, as well as their relations with the parametrization of Section 2. First, in the next two theorems, we prove some relations similar to those in Theorem 3.1 in [7].

THEOREM 4.1. *The Krylov matrix $K_{n,k} = [b, Ab, \ldots, A^{k-1}b]$ can be factorized as*

$$K_{n,k} = V_{n,k} \hat{U}_k, \tag{4.1}$$

where $V_{n,k}$ is the matrix whose columns are the orthonormal basis vectors of the Krylov subspace $\mathcal{K}_k(A,b)$ and $\hat{U}_k$ is an upper triangular matrix with a real positive diagonal. Moreover,

$$\hat{U}_k = \|b\| \begin{bmatrix} e_1 & H_k e_1 & \cdots & H_k^{k-1} \end{bmatrix} \tag{4.2}$$

and

$$\hat{U}_k^{-1} = U_k = \begin{bmatrix} & \tilde{g}^T \\ 0 & T_{k-1} \end{bmatrix},$$

the entries of the last matrix being defined in Theorem 2.3 describing the parametrization of Section 2.

Proof. We have $b = \|b\| V_{n,k} e_1$. Let us prove that $A^j V_{n,k} = V_{n,k} H_k^j$, $j = 1, \ldots, k-1$. This is true for $j = 1$ since we have $AV_{n,k} = V_{n,k} H_k$. Let us assume that $A^{j-1} V_{n,k} = V_{n,k} H_k^{j-1}$. Then,

$$A^j V_{n,k} = A(A^{j-1} V_{n,k}) = AV_{n,k} H_k^{j-1} = V_{n,k} H^j.$$

Therefore,

$$K_{n,k} = \begin{bmatrix} b & Ab & \cdots & A^{k-1}b \end{bmatrix} = \|b\| V_{n,k} \begin{bmatrix} e_1 & H_k e_1 & \cdots & H_k^{k-1} e - 1 \end{bmatrix}.$$

The matrix H_k being upper Hessenberg, one can prove easily that the matrix $\hat{U}_k$ is upper triangular. Moreover, since H_k has a positive first subdiagonal, the diagonal entries of $\hat{U}_k$ are positive. From $AK_{n,k} = K_{n,k} C_p^{(k)}$, we obtain that $H_k \hat{U}_k = \hat{U}_k C_p^{(k)}$. Therefore, $\hat{U}_k$ is the inverse of the upper triangular matrix involved in the factorization of H_k in Theorem 2.3. □

THEOREM 4.2. *Using the notation of Theorem 3.2, the matrix $AK_{n,k}$ can be factorized as*

$$AK_{n,k} = W_{n,k} \tilde{\mathcal{R}}_k,$$

where the upper triangular matrix $\tilde{\mathcal{R}}_k$ is equal to $Y_k C_p^{(k)}$. The first $k-1$ columns of $\tilde{\mathcal{R}}_k$ are

$$\begin{bmatrix} & R_{k-1} & \\ 0 & \cdots & 0 \end{bmatrix},$$

the matrix R_{k-1} being defined in Theorem 3.2. The upper Hessenberg matrix H_k can be factorized as $H_k = Q_k \mathcal{R}_k$ where

$$Q_k = V_{n,k}^* W_{n,k} = \hat{U}_k Y_k^{-1}$$

is upper Hessenberg and such that its first row is $h_k^T / \|h_k\|$. The matrix $\mathcal{R}_k$ is linked to $\tilde{\mathcal{R}}_k$ by

$$\tilde{\mathcal{R}}_k = \mathcal{R}_k \hat{U}_k.$$

The upper triangular matrix $\hat{U}_k$ is defined in Theorem 4.1.

Proof. One can prove that $K_{n,k} = W_{n,k} Y_k$ and

$$B_{n,k} = A \begin{bmatrix} b & Ab & \cdots & A^{k-1}b \end{bmatrix} = AK_{n,k} = WR_{n,k}.$$

Therefore,

$$\begin{aligned} AK_{n,k} &= WR_{n,k}, \\ &= W_{n,k} \tilde{\mathcal{R}}_k, \\ &= AW_{n,k} Y_k, \\ &= W_{n,k} Y_k C_p^{(k)}, \end{aligned}$$

from (3.9). This yields $\tilde{\mathcal{R}}_k = Y_k C_p^{(k)}$ and, from the structure of $C_p^{(k)}$, the first $k-1$ columns of $\tilde{\mathcal{R}}_k$ are

$$\left[\begin{array}{c} R_{k-1} \\ 0 \quad \cdots \quad 0 \end{array} \right].$$

We have

$$H_k = \hat{U}_k C_p^{(k)} \hat{U}_k^{-1}, \quad \hat{U}_k = V_{n,k}^* W_{n,k} Y_k,$$

from $K_{n,k} = V_{n,k}\hat{U}_k = W_{n,k}Y_k$. Let $Q_k = V_{n,k}^* W_{n,k}$ and $\mathcal{R}_k = Y_k C_p^{(k)} \hat{U}_k^{-1}$ which is an upper triangular matrix. Then,

$$H_k = V_{n,k}^* W_{n,k} Y_k C_p^{(k)} U_k^{-1} = Q_k \mathcal{R}_k,$$

and $\mathcal{R}_k \hat{U}_k = \tilde{\mathcal{R}}_k$. Moreover,

$$Q_k Y_k = V_{n,k}^* W_{n,k} Y_k = \hat{U}_k.$$

Instead of considering the first row of Q_k, let us look at the first column of $Q_k^* = W_{n,k}^* V_{n,k}$,

$$Q_k^* e_1 = W_{n,k}^* V_{n,k} e_1 = W_{n,k}^* \frac{b}{\|b\|} = \frac{h_k}{\|h_k\|},$$

since $b = W_{n,k} h_k$ and $\|b\| = \|h_k\|$. Therefore, the first row of Q_k is real positive and describes the convergence of GMRES. □

By using Theorem 4.2 we can obtain a relation between the η_js defined in Theorem 3.2 and the components of the vector $\tilde{g}$ in Theorem 2.3. Considering the first row of Q_k we have

$$e_1^T Q_k = e_1^T V_{n,k}^* W_{n,k} = e_1^T \hat{U}_k Y_k^{-1}.$$

By computing $\hat{U}_k Y_k^{-1}$ we obtain the two relations

$$\left[\begin{array}{ccc} \eta_1 & \cdots & \eta_{k-1} \end{array} \right] R_{k-1} = -\|b\| \left[\begin{array}{ccc} \tilde{g}_2 & \cdots & \tilde{g}_k \end{array} \right] T_{k-1}^{-1},$$

$$\eta_k^2 = \|b\| \left(1 + \left[\begin{array}{ccc} \tilde{g}_2 & \cdots & \tilde{g}_k \end{array} \right] T_{k-1}^{-1} R_{k-1}^{-1} \left[\begin{array}{c} \eta_1 \\ \vdots \\ \eta_{k-1} \end{array} \right] \right).$$

The two matrices R_{k-1} and T_{k-1} involved in the two parametrizations are linked through the entries of Q_k. Let $\hat{Q}_{k-1}$ be the upper triangular submatrix of rows 2 to k and columns 1 to $k-1$ of Q_k. Then, by identification, we have $\hat{Q}_{k-1} = T_{k-1}^{-1} R_{k-1}^{-1}$.

The following theorem is the analog of Theorem 3.2 in [7].

THEOREM 4.3. *The GMRES residual norm convergence curve described by h_k is characterized by the following relation,*

$$[b^* A b,\, b^* A^2 b,\, \ldots, b^* A^{k-1} b] = h_{k-1}^T R_{k-1}, \tag{4.3}$$

where h_{k-1} is the vector of the first $k-1$ components of h_k defined in Theorem 3.2 ($h_{k-1}^T = \begin{bmatrix} \eta_1 & \cdots & \eta_{k-1} \end{bmatrix}$) and the upper triangular matrix R_{k-1} is such that

$$R_{k-1}^* R_{k-1} = \begin{bmatrix} b^* A^* \\ b^*(A^2)^* \\ \vdots \\ b^T(A^{k-1})^* \end{bmatrix} \begin{bmatrix} Ab & A^2 b & \dots & A^{k-1} b \end{bmatrix}. \tag{4.4}$$

Proof. The result is obtained by identification using the relation $K_{n,k}^* K_{n,k} = Y_k^* Y_k$. □

The matrix on the right-hand side of (4.4) is a Gram (moment) matrix. This result fully (implicitly) describes GMRES convergence using the factor of the Gram matrix. Let us now consider the GMRES iterates. They can be expressed using the matrices in the parametrization of Section 3.

THEOREM 4.4. *The GMRES iterates x_j, $j < k$ are given by*

$$x_j = W_{n,k} Y_k \begin{bmatrix} R_j^{-1} h_j \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad h_j = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_j \end{bmatrix},$$

where R_j is the principal submatrix of order j of R_{k-1}.

Proof. The residual vector r_j at iteration $j < k$ can be written as $r_j = b - u$ where $u \in A\mathcal{K}_j$ yields the minimum of

$$\|r_j\| = \min_{u \in A\mathcal{K}_j} \|b - u\|.$$

The solution is given by the orthogonal projection of b on $A\mathcal{K}_j$. But, we have an orthogonal basis of the subspace $A\mathcal{K}_j$ given by the columns of $W_{n,j}$ and the solution can be written as

$$u = W_{n,k} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_j \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Since $b = W_{n,k} h_k$ we obtain that the residual vector is

$$r_j = W_{n,k} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \eta_{j+1} \\ \vdots \\ \eta_k \end{bmatrix}.$$

The corresponding iterate is given by

$$x_j = A^{-1}(b - r_j) = A^{-1}W_{n,k}\begin{bmatrix} \eta_1 \\ \vdots \\ \eta_j \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

From (3.9) we have $A^{-1}W_{n,k} = W_{n,k}Y_k(C_p^{(k)})^{-1}Y_k^{-1}$ and

$$x_j = W_{n,k}Y_k(C_p^{(k)})^{-1}Y_k^{-1}\begin{bmatrix} \eta_1 \\ \vdots \\ \eta_j \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The inverse of the matrix Y_k is

$$Y_k^{-1} = \begin{bmatrix} 0 & \cdots & 0 & 1/\eta_k \\ & R_{k-1}^{-1} & & -R_{k-1}^{-1}h_{k-1}/\eta_k \end{bmatrix}.$$

Finally, we obtain

$$x_j = W_{n,k}Y_k(C_p^{(k)})^{-1}\begin{bmatrix} 0 \\ R_j^{-1}h_j \\ 0 \\ \vdots \\ 0 \end{bmatrix} = W_{n,k}Y_k\begin{bmatrix} R_j^{-1}h_j \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

using the structure of the inverse of the companion matrix. □

Note that $K_{n,k} = W_{n,k}Y_k = V_{n,k}\hat{U}_k$. Hence Theorem 4.4 explains what are the coordinates of the iterates x_j in the three bases given by $K_{n,k}, V_{n,k}, W_{n,k}$. It also shows that the GMRES iterates do not depend on the eigenvalues of the matrix A in the sense that, in the parametrization of Section 3, one can change the coefficients of the last column of the companion matrix without changing the iterates. By looking at the exact solution of the linear system $Ax = b$, we can obtain an expression of the error vector. This is a generalization of Theorem 5.1 in [7].

THEOREM 4.5. *The GMRES error vector ε_j can be written as*

$$\varepsilon_j = W_{n,k}Y_k\left((C_p^{(k)})^{-1}e_1 - \begin{bmatrix} R_j^{-1}h_j \\ 0 \end{bmatrix}\right).$$

Proof. We have the relation

$$W_{n,k}Y_k(C_p^{(k)})^{-1} = A^{-1}W_{n,k}Y_k.$$

Looking at the first columns, it yields what is the solution vector of the linear system,

$$W_{n,k}Y_k(C_p^{(k)})^{-1}e_1 = A^{-1}W_{n,k}h_k = A^{-1}b = x.$$

Subtracting the iterate x_j from Theorem 4.4 gives the result for ε_j. $\square$

Contrary to the iterates, the error vectors do depend on the eigenvalues of A through the exact solution x.

5. Conclusion. In this paper we have generalized the results proved in [1] and [2] to the case of early termination of the Arnoldi process. We gave two characterizations of the class of matrices having the same GMRES residual norm convergence curve and early termination. Moreover, we also showed how to construct matrices having a prescribed residual norm convergence curve as well as prescribed Ritz values at all the iterations.

The conclusion must be expanded.

REFERENCES

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

[2] J. Duintjer Tebbens and G. Meurant, *Any Ritz value behavior is possible for Arnoldi and for GMRES with any nonincreasing convergence curve*, accepted in SIMAX, 2012.

[3] A. Greenbaum, V. Pták, and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469.

[4] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent advances in iterative methods, vol. 60 of IMA Vol. Math. Appl., Springer, New York, 1994, pp. 95–118.

[5] J. Liesen, *Construction and analysis of polynomial iterative methods for non-Hermitian systems of linear equations*, PhD thesis, University of Bielefeld, Germany, 1998.

[6] ———, *Computable convergence bounds for GMRES*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 882–903.

[7] G. Meurant, *GMRES and the Arioli, Pták and Strakoš factorization*, Accepted in BIT Numerical Mathematics, (2012).

[8] Y. Saad, *Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices*, Linear Algebra Appl., 34 (1980), pp. 269–295.

[9] Y. Saad, *Iterative methods for sparse linear systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, second ed., 2000.

[10] 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.