Electronic Transactions on Numerical Analysis.
Volume 33, pp. 207-220, 2009.

ISSN 1068-9613.

ETNA
Kent State University
http://etna.math.kent.edu

ALGEBRAIC PROPERTIES OF THE BLOCK GMRES AND BLOCK ARNOLDI METHODS*

L. ELBOUYAHYAOUI†, A. MESSAOUDI‡, AND H. SADOK§

Abstract. The solution of linear systems of equations with several right-hand sides is considered. Approximate solutions are conveniently computed by block GMRES methods. We describe and study three variants of block GMRES. These methods are based on three implementations of the block Arnoldi method, which differ in their choice of inner product.

Key words. block method, GMRES method, Arnoldi method, matrix polynomial, multiple right-hand sides, block Krylov subspace, Schur complement, characteristic polynomial.

AMS subject classifications. 65F10.

1. Introduction. Many problems in science and engineering require the solution of large linear systems of equations with multiple right-hand sides

$$AX = B, \quad A \in \mathbb{C}^{n\times n}, \quad B \in \mathbb{C}^{n\times s}, \quad X \in \mathbb{C}^{n\times s}, \quad 1 \le s \ll n. \tag{1.1}$$

Instead of applying a standard iterative method to the solution of each one of the linear systems of equations

$$Ax^{(i)} = b^{(i)} \quad \text{for} \quad i = 1, \ldots, s \tag{1.2}$$

independently, it is often more efficient to apply a block method to (1.1). The first block method, a block conjugate gradient method, was introduced by O'Leary [14] for the solution of a linear system of equations with multiple right-hand sides (1.1) and a symmetric positive definite matrix. For systems (1.1) with a nonsymmetric matrix, a block version of GMRES was introduced in [29] and studied in [26, 27]. This method is based on a block version of the standard Arnoldi process [1]; see, for example, [17, 28].

The purpose of the present paper is to compare three variants of GMRES for multiple right-hand sides, including the block GMRES method considered in [26, 27, 29]. These schemes are based on block Arnoldi-type methods and differ in the choice of inner product. We provide a unified description of the methods discussed, and derive new expressions and bounds for the residual errors.

The paper is organized as follows. Section 2 defines block GMRES iterates with the aid of Schur complements, and presents a connection with matrix-valued polynomials. We use these polynomials to derive some new relations in Section 3. A few examples that illustrate the theory are provided in Section 4.

We conclude this section by introducing notation used in the remainder of this paper. We first recall the definition of the Schur complement. [24].

DEFINITION 1.1. *Let M be a matrix partitioned into four blocks*

$$M = \begin{bmatrix} C & D \\ E & F \end{bmatrix},$$

*Received April 27, 2009. Accepted August 22, 2009. Published online on December 16, 2009. Recommended by Lothar Reichel.

†Département de Mathématiques, Faculté des Sciences et Techniques, Université de Mohammadia, Mohammedia, Morocco. lakhdar.elbouyahyaoui@gmail.com.

‡Département de Mathématiques, Ecole Normale Supérieure, Rabat, Morocco. abderrahim.messaoudi@gmail.com.

§L.M.P.A, Université du Littoral, 50 rue F. Buisson BP699, F-62228 Calais Cedex, France. sadok@lmpa.univ-littoral.fr.

ETNA
Kent State University
http://etna.math.kent.edu

where the submatrix F is assumed to be square and nonsingular. The Schur complement of F in M, denoted by (M/F), is defined by

$$(M/F) = C - DF^{-1}E.$$

Throughout this paper, I and I_s denote identity matrices and e_k their kth column. For two matrices Y and Z in $\mathbb{C}^{n\times s}$, we define the inner product $\langle Y, Z\rangle_F = \mathrm{trace}(Y^H Z)$, (where Y^H denotes the conjugate transpose of Y). The associated norm is the Frobenius norm $\|\cdot\|_F$. The 2-norm of a matrix $X \in \mathbb{C}^{n\times s}$ is denoted by $\|X\|_2$. The Kronecker product of the matrices $C = [c_{i,j}]$ and D is given by $C \otimes D = [c_{i,j}D]$. If X is an $n \times s$ matrix, $x = \mathrm{vec}(X)$ is the ns vector obtained by stacking the s columns of the matrix X.

Finally, the roots of a matrix-valued polynomial $\mathbb{P}$, which is a square matrix whose entries are ordinary polynomials, are defined to be the roots of the ordinary polynomial $\det(\mathbb{P}(t))$.

2. Block minimal residual-type methods. The nonsingular linear system with multiple right-hand sides (1.1) can be solved by Krylov subspace methods in two distinct ways. The first approach is to apply classical GMRES [16] for linear systems of equations with single-vector right-hand sides to the s linear systems separately. The second approach is to treat all the right-hand sides simultaneously.

Before investigating the two approaches, we need some notation. Let the initial approximation of the solution of (1.1) be $X_0 = [x_0^{(1)}, x_0^{(2)}, \ldots, x_0^{(s)}]$, and let $R_0 = [r_0^{(1)}, \ldots, r_0^{(s)}] = B - AX_0$ be the corresponding residual, with $r_0^{(i)} = b^{(i)} - A\, x_0^{(i)}$ and $B = [b^{(1)}, \ldots, b^{(s)}]$. In what follows, we let $\mathcal{K}_k$ denote the block Krylov matrix

$$\mathcal{K}_k = [R_0, AR_0, \ldots, A^{k-1}R_0],$$

and $K_{i,k}$ the Krylov matrix

$$K_{i,k} = [r_0^{(i)}, A\, r_0^{(i)} \ldots, A^{k-1}\, r_0^{(i)}]$$

for $i = 1, \ldots, s$. We also introduce the matrix

$$\mathcal{W}_k = A\mathcal{K}_k.$$

2.1. Standard GMRES applied to systems with multiple right-hand sides. In this section we apply the standard GMRES method to each one of the s linear systems of equations (1.2). Define the ith classical Krylov subspace $\mathrm{K}_k(A, r_0^{(i)})$ by

$$\mathrm{K}_k(A, r_0^{(i)}) = \mathrm{span}\{r_0^{(i)}, A\, r_0^{(i)}, \ldots A^{k-1} r_0^{(i)}\} \subset \mathbb{C}^n. \tag{2.1}$$

It is well-known that the kth approximation $x_{k,S}^{(i)}$ of GMRES applied to the ith linear system (1.2) satisfies

$$x_{k,S}^{(i)} - x_0^{(i)} \in \mathrm{K}_k(A, r_0^{(i)}) \quad \text{and} \quad (A^j r_0^{(i)})^H r_{k,S}^{(i)} = 0 \quad \text{for } j = 1, \ldots k, \tag{2.2}$$

where $r_{k,S}^{(i)} = b^{(i)} - A\, x_{k,S}^{(i)}$. It follows that the residual vector $r_{k,S}^{(i)}$ can be written as a linear combination of the vectors $A^j r_0^{(i)}$, $j = 0, 1, \ldots, k$, i.e.,

$$r_{k,S}^{(i)} = p_{k,S}^{(i)}(A) r_0^{(i)},$$

ETNA
Kent State University
http://etna.math.kent.edu

where

$$p_{k,S}^{(i)}(t) = \frac{\det\left(\begin{bmatrix} 1 & t & \dots & t^k \\ K_{i,k}^H A^H r_0^{(i)} & & K_{i,k}^H A^H A K_{i,k} & \end{bmatrix}\right)}{\det(K_{i,k}^H A^H A K_{i,k})}.$$

The residual error for each one of the s linear systems satisfies

$$\|r_{k,S}^{(i)}\|_2^2 = \frac{1}{e_1^T (K_{i,k+1}^H \, K_{i,k+1})^{-1} e_1}, \quad i = 1, \dots, s; \tag{2.3}$$

see [21, 22]. Therefore, the Frobenius norm of the residual

$$R_{k,S} = [r_{k,S}^{(1)}, \dots, r_{k,S}^{(2)}] = [p_{k,S}^{(1)}(A) r_0^{(1)}, \dots, p_{k,S}^{(s)}(A) r_0^{(s)}] \tag{2.4}$$

can be written as

$$\|R_{k,S}\|_F^2 = \sum_{i=1}^{s} \frac{1}{e_1^T (K_{i,k+1}^H \, K_{i,k+1})^{-1} e_1}. \tag{2.5}$$

Similarly to the situation for standard GMRES, the residual $R_{k,S}$ can be expressed in terms of a polynomial in A. We deduce from (2.4) that

$$\mathrm{vec}(R_{k,S}) = \mathbb{P}_{k,S}^G(A) \, \mathrm{vec}(R_0), \quad \text{where} \quad \mathbb{P}_{k,S}^G(t) = \mathrm{diag}(p_{k,S}^{(1)}(t), \dots, p_{k,S}^{(s)}(t)). \tag{2.6}$$

2.2. The global GMRES method. Instead of using standard GMRES to solve each linear system (1.2) separately, we may apply GMRES to a block diagonal matrix. The s linear systems (1.2) can be rewritten in a compact form as $(A \otimes I_s)x = \mathrm{vec}(B)$, with $x = \mathrm{vec}(X)$. This gives the following linear system with a single right-hand side

$$\begin{bmatrix} A & & \\ & \ddots & \\ & & A \end{bmatrix} x = \begin{bmatrix} b^{(1)} \\ \vdots \\ b^{(s)} \end{bmatrix}. \tag{2.7}$$

Application of standard GMRES to (2.7) yields the global GMRES method, which also can be defined as follows. Let

$$\mathbf{K}_k^G(A, U) = \mathrm{span}\{U, A\,U, \dots, A^{k-1}\,U\} \subset \mathbb{C}^{n \times s}$$

denote the matrix Krylov subspace spanned by the matrices $U, AU, \dots, A^{k-1}U$, where U is an $n \times s$ matrix. Note that $Z \in \mathbf{K}_k^G(A, U)$ implies that

$$Z = \sum_{j=1}^{k} \alpha_j A^{j-1} U, \qquad \alpha_j \in \mathbb{C}, \qquad j = 1, \dots, k.$$

At step k, the global GMRES method constructs the approximation $X_{k,G}$, which satisfies the relations

$$X_{k,G} - X_0 \in \mathbf{K}_k^G(A, R_0) \quad \text{and} \quad \langle A^j R_0, R_{k,G} \rangle_F = 0, \qquad j = 1, \dots, k.$$

The residual $R_{k,G} = B - AX_{k,G}$ satisfies the minimization property

$$\|R_{k,G}\|_F = \min_{Z \in \mathbf{K}_k^G(A, R_0)} \|R_0 - AZ\|_F. \tag{2.8}$$

ETNA
Kent State University
http://etna.math.kent.edu

The problem (2.8) is solved by applying the global Arnoldi process [8].

Global GMRES is a generalization of the global MR method proposed by Saad for approximating the inverse of a matrix [18, p. 300]. The global method also is effective, compared to block Krylov subspace methods, when applied to the solution of large and sparse Lyapunov and Sylvester matrix equations with right-hand sides of low rank; see [9, 10, 23]. Applications of the global Arnoldi method in control theory, model reduction, and quadratic matrix equations are given in [4–6, 30].

It is convenient to introduce the matrix product $\diamond$. Let $Y = [Y_1, Y_2, \ldots, Y_p]$ and $Z = [Z_1, Z_2, \ldots, Z_l]$ be matrices of dimension $n \times ps$ and $n \times ls$, respectively, where Y_i and Z_j ($i = 1, \ldots, p$; $j = 1, \ldots, l$) are $n \times s$ matrices. Then $\diamond$ is defined by

$$Y^H \diamond Z = \begin{bmatrix} \langle Y_1, Z_1\rangle_F & \langle Y_1, Z_2\rangle_F & \ldots & \langle Y_1, Z_l\rangle_F \\ \langle Y_2, Z_1\rangle_F & \langle Y_2, Z_2\rangle_F & \ldots & \langle Y_2, Z_l\rangle_F \\ \vdots & \vdots & \ddots & \vdots \\ \langle Y_p, Z_1\rangle_F & \langle Y_p, Z_2\rangle_F & \ldots & \langle Y_p, Z_l\rangle_F \end{bmatrix} \subset \mathbb{C}^{p\times l}.$$

THEOREM 2.1 ([3]). *Let the matrix $(A\mathcal{K}_k)^H \diamond (A\mathcal{K}_k)$ be nonsingular. Then*

$$\begin{aligned} R_{k,G} &= R_0 - A\mathcal{K}_k \left((A\mathcal{K}_k)^H \diamond (A\mathcal{K}_k)\right)^{-1} \left((A\mathcal{K}_k)^H \diamond R_0\right) \\ &= \mathbb{P}^G_{k,G}(A) R_0, \end{aligned}$$

where

$$\mathbb{P}^G_{k,G}(t) = \frac{\det\left(\begin{bmatrix} 1 & t & \ldots & t^k \\ (A\mathcal{K}_k)^H \diamond R_0 & & (A\mathcal{K}_k)^H \diamond (A\mathcal{K}_k) & \end{bmatrix}\right)}{\det((A\mathcal{K}_k)^H \diamond (A\mathcal{K}_k))}. \tag{2.9}$$

Moreover,

$$\|R_{k,G}\|_F^2 = \frac{1}{e_1^T\,(\mathcal{K}^H_{k+1} \diamond \mathcal{K}_{k+1})^{-1} e_1} = \frac{1}{e_1^T\,(\sum_{i=1}^s K^H_{i,k+1}\, K_{i,k+1})^{-1} e_1}.$$

We also can write $R_{k,G} = \left[\mathbb{P}^G_{k,G}(A) r_0^{(1)}, \ldots, \mathbb{P}^G_{k,G}(A) r_0^{(s)}\right]$ or equivalently

$$\mathrm{vec}(R_{k,G}) = \mathrm{diag}\left(\mathbb{P}^G_{k,G}(A), \ldots, \mathbb{P}^G_{k,G}(A)\right) \mathrm{vec}(R_0).$$

The matrix-valued polynomial involved in the preceding two studied methods are both diagonal. In the next section, we consider a general matrix-valued polynomial.

2.3. The block GMRES method. Another approach to solving (1.1) is to consider all the s right-hand side vectors $b^{(i)}$, $i = 1, \ldots, s$, as an $n \times s$ matrix. This leads to the block GMRES method (BGMRES). This method determines at step k an approximate solution $X_{k,B}$ of (1.1) from the requirements

$$X_{k,B} - X_0 \in \mathbf{K}^B_k(A, R_0), \quad \text{and} \quad R_{k,B} = B - AX_{k,B} \perp \mathbf{K}^B_k(A, AR_0), \tag{2.10}$$

where

$$\mathbf{K}^B_k(A, U) = \text{block span}\{U, AU, \ldots, A^{k-1}U\},$$

and “block span” is defined by

$$\mathbf{K}^B_k(A, U) = \left\{X \in \mathbb{C}^{n\times s}\,|X = \sum_{i=0}^{k-1} A^i U\,\Omega_i;\ \Omega_i \in \mathbb{C}^{s\times s} \text{ for } i = 0, \ldots, k-1\right\} \subset \mathbb{C}^{n\times s}.$$

Alternatively, BGMRES can be defined by considering the approximate solution of the ith system (1.2), which is determined by

$$x_{k,B}^{(i)} - x_0^{(i)} \in \mathbb{K}_k(A, R_0) \text{ and } (A^j R_0)^H\, r_{k,B}^{(i)} = 0, \quad j = 1, \ldots, k;\ i = 1, \ldots, s, \tag{2.11}$$

where

$$\mathbb{K}_k(A, R_0) = \mathrm{Range}([R_0, AR_0, \ldots, A^{k-1} R_0]) \subset \mathbb{C}^n. \tag{2.12}$$

Note that the Krylov subspace $\mathbb{K}_k(A, R_0)$ is a sum of s classical Krylov subspaces

$$\mathbb{K}_k(A, R_0) = \sum_{i=1}^{s} K_k(A, r_0^{(i)}).$$

Each column of the residual matrix $R_{k,B}$ is obtained by projecting orthogonally the corresponding column of R_0 onto the block Krylov subspace $A\,\mathbb{K}_k(A, R_0)$. Therefore, BGMRES is a minimization method

$$\|R_{k,B}\|_F = \min_{Z \in \mathbf{K}_k^B(A, R_0)} \|R_0 - AZ\|_F.$$

The following result will be used in the sequel.

THEOREM 2.2. *([2]) Let the matrix $\mathcal{W}_k = A\mathcal{K}_k$ be of full rank. Then*

$$\begin{aligned} R_{k,B} &= R_0 - A\mathcal{K}_k \left(\mathcal{W}_k^H \mathcal{W}_k\right)^{-1} \mathcal{W}_k^H R_0 \\ &= \left(\begin{bmatrix} R_0 & \mathcal{W}_k \\ \mathcal{W}_k^H R_0 & \mathcal{W}_k^H \mathcal{W}_k \end{bmatrix} / \mathcal{W}_k^H\, \mathcal{W}_k \right). \end{aligned}$$

Following Vital [29], we introduce the operator

$$\mathbb{P}_{k,B}^G(A) \circ R_0 = \sum_{i=0}^{k} A^i R_0 \Omega_i,$$

where $\Omega_0 = I_s$, $[\Omega_1, \ldots, \Omega_k] = -\left(\mathcal{W}_k^H \mathcal{W}_k\right)^{-1} \mathcal{W}_k^H R_0$, and $\mathbb{P}_{k,B}^G$ is the matrix-valued polynomial defined by

$$\mathbb{P}_{k,B}^G(t) = \sum_{i=1}^{k} t^i \Omega_i = \left(\begin{bmatrix} I_s & tI_s & \ldots & t^k I_s \\ \mathcal{W}_k^H R_0 & & \mathcal{W}_k^H \mathcal{W}_k & \end{bmatrix} / \mathcal{W}_k^H \mathcal{W}_k \right). \tag{2.13}$$

Then the residual $R_{k,B}$ can be expressed as

$$R_{k,B} = \mathbb{P}_{k,B}^G(A) \circ R_0.$$

Theorem 2.2 helps us compare the residuals of standard GMRES applied to (1.2) and of BGMRES applied to (1.1). The relation (2.3) is the key to developing convergence results for GMRES [21, 22]. We have the following expression for the norm of the residuals determined by BGMRES.

THEOREM 2.3. *Assume that the matrix $\mathcal{W}_k$ is of full rank. Then*

$$\|r_{k,B}^{(i)}\|^2 = \frac{1}{e_1^T \begin{bmatrix} {r_0^{(i)}}^H r_0^{(i)} & {r_0^{(i)}}^H \mathcal{W}_k \\ \mathcal{W}_k^H r_0^{(i)} & \mathcal{W}_k^H\, \mathcal{W}_k \end{bmatrix}^{-1} e_1} \quad \textit{for } i = 1, \ldots, s. \tag{2.14}$$

ETNA
Kent State University
http://etna.math.kent.edu

Proof. From the first expression in Theorem 2.2, we deduce that

$$r_{k,B}^{(i)} = r_0^{(i)} - \mathcal{W}_k \left(\mathcal{W}_k^H \mathcal{W}_k\right)^{-1} \mathcal{W}_k^H r_0^{(i)} = (I - \mathcal{W}_k \left(\mathcal{W}_k^H \mathcal{W}_k\right)^{-1} \mathcal{W}_k^H) r_0^{(i)}. \tag{2.15}$$

Consequently,

$$\|r_{k,B}^{(i)}\|^2 = r_0^{(i)\,H} r_{k,B}^{(i)} = (r_0^{(i)})^H r_0^{(i)} - (r_0^{(i)})^H \mathcal{W}_k \left(\mathcal{W}_k^H \mathcal{W}_k\right)^{-1} \mathcal{W}_k^H r_0^{(i)},$$

and so we obtain

$$\|r_{k,B}^{(i)}\|^2 = \left(\begin{bmatrix} (r_0^{(i)})^H (r_0^{(i)}) & (r_0^{(i)})^H \mathcal{W}_k \\ \mathcal{W}_k^H r_0^{(i)} & \mathcal{W}_k^H \mathcal{W}_k \end{bmatrix} / \, \mathcal{W}_k^H \, \mathcal{W}_k \right). \tag{2.16}$$

Hence, $\|r_{k,B}^{(i)}\|^2$ is the Schur complement of $\mathcal{W}_k^H \mathcal{W}_k$ in the matrix

$$\begin{bmatrix} (r_0^{(i)})^H r_0^{(i)} & (r_0^{(i)})^H \mathcal{W}_k \\ \mathcal{W}_k^H r_0^{(i)} & \mathcal{W}_k^H \mathcal{W}_k \end{bmatrix},$$

which can be factored into a product of a block upper and a block lower triangular matrix (UL factorization)

$$\begin{bmatrix} (r_0^{(i)})^H r_0^{(i)} & (r_0^{(i)})^H \mathcal{W}_k \\ \mathcal{W}_k^H r_0^{(i)} & \mathcal{W}_k^H \mathcal{W}_k \end{bmatrix} = \begin{bmatrix} 1 & (r_0^{(i)})^H \mathcal{W}_k \left(\mathcal{W}_k^H \mathcal{W}_k\right)^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \|r_{k,B}^{(i)}\|^2 & 0 \\ \mathcal{W}_k^H r_0^{(i)} & \mathcal{W}_k^H \mathcal{W}_k \end{bmatrix}.$$

This factorization yields

$$e_1^T \begin{bmatrix} r_0^{(i)\,H} \, r_0^{(i)} & r_0^{(i)\,H} \, \mathcal{W}_k \\ \mathcal{W}_k^H r_0^{(i)} & \mathcal{W}_k^H \, \mathcal{W}_k \end{bmatrix}^{-1} e_1 = \frac{1}{\|r_{k,B}^{(i)}\|^2}, \tag{2.17}$$

which proves the theorem. □

The above theorem allows us to improve the well-known result

$$\min_{Z \in \mathbf{K}_k^B(A,R_0)} \|R_0 - AZ\|_\psi \leq \max_{i=1\ldots s} \; \min_{z_i \in K_k(A, r_0^{(i)})} \|r_0^{(i)} - z_i\|,$$

which was stated in [26, 27] with $\|Z\|_\psi = \max_{i=1,\ldots,s} (\|z_i\|)$, and was shown by Vital in her thesis [29]. It shows that the residual obtained by BGMRES is bounded by the maximum of the norm of the residuals obtained by applying standard GMRES to each one of the s systems (1.2).

THEOREM 2.4. *Let the matrix $\mathcal{W}_k$ be of full rank. Then*

$$\|r_{k,B}^{(i)}\| \leq \|r_{k,S}^{(i)}\| \quad \textit{for} \quad i = 1, \ldots, s$$

and

$$\|R_{k,B}\|_F \leq \|R_{k,S}\|_F \leq \|R_{k,G}\|_F.$$

Proof. It suffices to show the first part of the theorem for $i = 1$. Let us first remark that there exists a permutation matrix P, such that $\mathcal{W}_k = A[K_{1,k}, \ldots, K_{s,k}]\, P$. Therefore, we can rewrite the expression of $\|R_{k,B}^{(1)}\|^2$ as

$$\|r_{k,B}^{(1)}\|^2 = \frac{1}{e_1^T F_k^{-1} e_1},$$

where

$$F_k = \begin{bmatrix} {r_0^{(1)}}^H r_0^{(1)} & {r_0^{(1)}}^H AK_{1,k} & \ldots & {r_0^{(1)}}^H AK_{s,k} \\ (AK_{1,k})^H r_0^{(1)} & (AK_{1,k})^H AK_{1,k} & \ldots & (AK_{1,k})^H AK_{s,k} \\ \vdots & \vdots & \ldots & \vdots \\ (AK_{s,k})^H r_0^{(1)} & (AK_{s,k})^H AK_{1,k} & \ldots & (AK_{s,k})^H AK_{s,k} \end{bmatrix}.$$

By noticing that the $(k+1) \times (k+1)$ principal submatrix of F_k is $K_{1,k+1}^H K_{1,k+1}$, using Theorem 6.2 of [31, p. 177] and (2.3), we deduce that

$$\frac{1}{\|r_{k,B}^{(1)}\|^2} = e_1^T F_k^{-1} e_1 \geq e_1^T (K_{1,k+1}^H K_{1,k+1})^{-1} e_1 = \frac{1}{\|r_{k,S}^{(1)}\|^2}.$$

To prove the last inequality, we apply (2.5) and Theorem 7.2 of [15], and obtain

$$\|R_{k,S}\|_F = \sum_{i=1}^{s} \frac{1}{e_1^T (K_{i,k+1}^H K_{i,k+1})^{-1} e_1} \leq \frac{1}{e_1^T \left(\sum_{i=1}^{s} K_{i,k+1}^H K_{i,k+1}\right)^{-1} e_1} = \|R_{k,G}\|_F^2,$$

which completes the proof. □

We now examine the zeros of the matrix-valued polynomial $\mathbb{P}_{k,B}^G$.

THEOREM 2.5. *Let the matrix $\mathcal{W}_k$ be of full rank. Then*

$$\det(\mathbb{P}_{k,B}^G(t)) = \prod_{i=1}^{ks} \frac{(\alpha_i^{(k)} - t)}{\alpha_i^{(k)}},$$

where the $\alpha_i^{(k)}$, for $i = 1, \ldots, ks$, are the generalized eigenvalues of the matrix pair $\{\mathcal{W}_k^H \mathcal{W}_k, \mathcal{W}_k^H \mathcal{K}_k\}$.

Proof. Let α be a root of $\det(\mathbb{P}_{k,B}^G)$. Then from Theorem 2.2, we deduce that

$$\det\left(\begin{bmatrix} I_s & \alpha I_s & \ldots & \alpha^k I_s \\ \mathcal{W}_k^H R_0 & & \mathcal{W}_k^H \mathcal{W}_k & \end{bmatrix}\right) = 0. \tag{2.18}$$

Let us denote the ith block column of this determinant by C_i. Then by replacing the block column C_i by $C_i - \alpha C_{i-1}$ for $i = 2, \ldots, k$, we obtain

$$\det(\mathcal{W}_k^H \mathcal{W}_k - \alpha \mathcal{W}_k^H \mathcal{K}_k) = 0, \tag{2.19}$$

which shows that α is a generalized eigenvalue of the matrix pair $\{\mathcal{W}_k^H \mathcal{W}_k, \mathcal{W}_k^H \mathcal{K}_k\}$. The proof is completed by noticing that $\det(\mathbb{P}_{k,B}^G(0)) = 1$. □

ETNA
Kent State University
http://etna.math.kent.edu

3. Block Arnoldi-type algorithms for multiple starting vectors. This section provides the framework for block Arnoldi-type algorithms. These algorithms are used for determining multiple or clustered eigenvalues. They also are applied in implementations of block and global GMRES methods.

We give a unified presentation of Arnoldi-type algorithms, which include the standard Arnoldi algorithm applied to each column of the starting block vector, the global Arnoldi method, and the block Arnoldi method. Let U be an $n \times s$ matrix. The Arnoldi-type algorithms construct a basis $\{V_1^\bullet, \ldots, V_k^\bullet\}$ of a subspace of $\mathbf{K}_k^B(A, U)$. The basis satisfies an orthogonality property and $\mathcal{H}_k^\bullet = (\mathcal{V}_k^\bullet)^H A\mathcal{V}_k^\bullet$ is upper block Hessenberg.

We examine three possibly choices of orthogonality. Let $\Phi^\bullet : \mathbb{C}^{n\times s} \times \mathbb{C}^{n\times s} \to \mathbb{C}^{s\times s}$ be defined for $\bullet \in \{B, S, G\}$ by

$$\begin{cases} \Phi^B(X,Y) &=& X^H Y, \\ \Phi^S(X,Y) &=& \text{the diagonal of the matrix } X^H Y, \\ \Phi^G(X,Y) &=& \operatorname{trace}(X^H Y) I_s = \langle X, Y\rangle_F I_s, \end{cases}$$

for all $X \in \mathbb{C}^{n\times s}$ and for all $Y\mathbb{C}^{n\times s}$.

If $\Phi^B(X, Y) = X^H Y = 0$, then the block-vectors X, Y are said to be block-orthogonal; see Gutknecht [11]. Moreover, X is said to be block-normalized if $X^H X = I_s$. Of course, the vector space of block vectors is a finite-dimensional inner product space with inner product $\langle X, Y\rangle_F = \operatorname{trace}(X^H Y)$. If $\langle X, Y\rangle_F = 0$, then X and Y are said to be F-orthogonal. If $\Phi^S(X, Y) = \operatorname{diag}(X^H Y) = 0$, then we say that X and Y are diagonally orthogonal.

Using the map $\Phi^\bullet$, we will show how the matrices $\mathcal{V}_k^\bullet$ and $\mathcal{H}_k^\bullet$ are computed.

Block Arnoldi-type algorithms

1. Let U be an $n \times s$ matrix.
2. Compute $V_1^\bullet \in \mathbb{C}^{n\times s}$ by determining the factorization of U: $U = V_1^\bullet H_{1,0}^\bullet$, $H_{1,0}^\bullet \in \mathbb{C}^{s\times s}$, such that $H_{1,0}^\bullet = \Phi^\bullet(V_1^\bullet, U)$ and $\Phi^\bullet(V_1^\bullet, V_1^\bullet) = I_s$.
3. for $i = 1, \ldots, k$ do
   - Compute $W = AV_i^\bullet$.
   - for $j = 1, \ldots, i$ do
     (a) $H_{j,i}^\bullet = \Phi^\bullet(V_j^\bullet, W)$
     (b) $W = W - V_j^\bullet H_{j,i}^\bullet$
   - End
   - Compute $H_{i+1,i}^\bullet$ by determining the decomposition of W : $W = V_{i+1}^\bullet H_{i+1,i}^\bullet$, such that $H_{i+1,i}^\bullet = \Phi^\bullet(V_{i+1}^\bullet, W)$ and $\Phi^\bullet(V_{i+1}^\bullet, V_{i+1}^\bullet) = I_s$.
4. End

We now consider two particular choices.

3.1. The block Arnoldi algorithm. For $\Phi^\bullet(X,Y) = \Phi^B(X,Y) = X^H Y$, the preceding algorithm reduces to block Arnoldi algorithm [11, 18–20, 25–27, 29], which builds an orthonormal basis $\{V_1^B, \ldots, V_k^B\}$ such that the block matrix $\mathcal{V}_k^B = [V_1^B, \ldots, V_k^B]$ satisfies $(\mathcal{V}_k^B)^H \mathcal{V}_k^B = I_{ks}$. It is well known that

$$A\mathcal{V}_k^B = \mathcal{V}_k^B \mathcal{H}_k^B + V_{k+1}^B H_{k+1,k}^B E_k^T, \tag{3.1}$$

where $E_k^T = [0_s, \ldots, 0_s, I_s] \in \mathbb{R}^{s\times ms}$. Multiplying equation (3.1) by E_k, we deduce that

$$V_{k+1}^B H_{k+1,k}^B = AV_k^B - \mathcal{V}_k^B \mathcal{V}_k^{B^H} AV_k^B.$$

ETNA
Kent State University
http://etna.math.kent.edu

We also have $AV_1^B H_{1,0} = AU$ and $V_2^B H_{2,1}^B = AV_1^B - \mathcal{V}_1^B \mathcal{V}_1^{B^H} AV_1^B$, which imply that $V_2^B H_{2,1}^B H_{1,0} = AU - \mathcal{V}_1^B \mathcal{V}_1^{B^H} AU$. Thus, by induction, we deduce that

$$V_{k+1}^B H_{k+1,k}^B H_{k,k-1}^B \cdots H_{1,0}^B = A^k U - \mathcal{V}_k^B \mathcal{V}_k^{B^H} A^k U.$$

Furthermore, if $\mathcal{K}_k = \mathcal{V}_k^B \mathcal{R}_k^B$ is the QR decomposition of the full-rank matrix $\mathcal{K}_k$, then $\mathcal{V}_k \mathcal{V}_k{}^B = \mathcal{K}_k (\mathcal{K}_k^H \mathcal{K}_k)^{-1} \mathcal{K}_k^H$. Hence, we have

$$V_{k+1}^B H_{k+1,k}^B H_{k,k-1}^B \cdots H_{1,0}^B = A^k U - \mathcal{K}_k (\mathcal{K}_k^H \mathcal{K}_k)^{-1} \mathcal{K}_k^H A^k U. \tag{3.2}$$

Consider the representation $V_{k+1}^B = \mathbb{P}_{k,B}^A(A) \circ U$. Since it is not easy to express $\mathbb{P}_{k,B}^A$ in terms of Krylov matrices, we consider a monic matrix-valued polynomial, which, apart from a multiplicative matrix, is the polynomial $\mathbb{P}_{k,B}^A$. Thus, let $\widetilde{\mathbb{P}}_{k,B}^A$ denote the matrix-valued polynomial

$$\widetilde{\mathbb{P}}_{k,B}^A(t) = \mathbb{P}_{k,B}^A(t) H_{k+1,k}^B \cdots H_{1,0}^B,$$

and let $\{Z_k\}$ be the block vectors defined by $Z_1 = U$ and

$$Z_{k+1} = V_{k+1}^B H_{k+1,k}^B \cdots H_{1,0}^B.$$

Then

$$Z_{k+1} = \left(I - \mathcal{K}_k \, (\mathcal{K}_k^H \mathcal{K}_k)^{-1} \mathcal{K}_k^H\right) \, A^k U = \widetilde{P}_{k,B}^A(A) \circ U \quad \text{for} \quad k \geq 1. \tag{3.3}$$

The matrix-valued polynomial $\widetilde{\mathbb{P}}_{k,B}^A$ can be expressed as

$$\widetilde{\mathbb{P}}_{k,B}^A(t) = \left(\begin{bmatrix} t^k I_s & I_s & \dots & t^{k-1} I_s \\ & & & \\ \mathcal{K}_k^H A^k U & & \mathcal{K}_k^H \mathcal{K}_k & \end{bmatrix} / \mathcal{K}_k^H \mathcal{K}_k \right). \tag{3.4}$$

Applying the determinant function to this Schur complement, we obtain

$$\det(\widetilde{\mathbb{P}}_{k,B}^A(t)) = \frac{\det \left(\begin{bmatrix} t^k I_s & I_s & \dots & t^{k-1} I_s \\ & & & \\ \mathcal{K}_k^H A^k U & & \mathcal{K}_k^H \mathcal{K}_k & \end{bmatrix} \right)}{\det(\mathcal{K}_k^H \mathcal{K}_k)}. \tag{3.5}$$

The following result examines the zeros of $\widetilde{\mathbb{P}}_{k,B}^A$.

THEOREM 3.1. *Let the matrix $\mathcal{K}_k$ be of full rank. Then*

$$\det(\widetilde{\mathbb{P}}_{k,B}^A(t)) = \prod_{i=1}^{ks} (t - \theta_i^{(k)}),$$

where $\theta_i^{(k)}$, $i = 1, ..., ks$, are the eigenvalues of the matrix $(\mathcal{K}_k^H \mathcal{K}_k)^{-1} (\mathcal{K}_k^H A \mathcal{K}_k)$.

Proof. Let θ be a root of $\det(\widetilde{\mathbb{P}}_{k,B}^A(t))$. It follows from (3.5) that

$$\det \left(\begin{bmatrix} I_s & \dots & \theta^{k-1} I_s & \theta^k I_s \\ & & & \\ & \mathcal{K}_k^H \mathcal{K}_k & & \mathcal{K}_k^H A^k R_0 \end{bmatrix} \right) = 0.$$

ETNA
Kent State University
http://etna.math.kent.edu

Let C_i denote the i^{th} block column of this determinant. Then replacing C_i by $C_i - \theta C_{i-1}$ for $i = 2, \ldots, k$, we obtain

$$\det(\mathcal{K}_k^H A \mathcal{K}_k - \theta \mathcal{K}_k^H \mathcal{K}_k) = 0.$$

Since the matrix $\mathcal{K}_k^H \mathcal{K}_k$ is nonsingular, θ is an eigenvalue of $(\mathcal{K}_k^H \mathcal{K}_k)^{-1} \mathcal{K}_k^H A \mathcal{K}_k$. This result, for the special case $s = 1$, is shown in [21]. □

Using the QR decomposition of the full-rank matrix $\mathcal{K}_k = \mathcal{V}_k^B \mathcal{R}_k^B$, we deduce that

$$(\mathcal{K}_k^H \mathcal{K}_k)^{-1} \mathcal{K}_k^H A \mathcal{K}_k = \left(\mathcal{R}_k^B\right)^{-1} \mathcal{H}_k^B \, \mathcal{R}_k^B.$$

Consequently the roots of $\mathbb{P}_{k,B}^A(t)$ are the eigenvalues of $\mathcal{H}_k^B$.

3.2. The global Arnoldi algorithm. We have

$$\Phi^\bullet(X, Y) = \Phi^G(X, Y) = \langle X, Y \rangle_F I_s.$$

Hence, the global Arnoldi process builds an F-orthonormal basis $\{V_1^G, \ldots, V_k^G\}$ of $\mathbf{K}_k^B(A, U)$, such that the matrix $\mathcal{V}_k^G = [V_1^G, \ldots, V_k^G]$ satisfies

$$A \mathcal{V}_k^G = \mathcal{V}_k^G \mathcal{H}_k^G + V_{k+1}^G H_{k+1,k}^G E_k^T,$$

where $\mathcal{H}_k^G = H_k^G \otimes I_s$ and the matrix H_k^G is a $k \times k$ Hessenberg matrix whose nonzero entries $(h_{i,j}^G)$ are defined by the following algorithm.

Global Arnoldi algorithm

1. Let U be an $n \times s$ matrix.
2. Compute $V_1^G \in \mathbb{C}^{n \times s}$ by $V_1^G = U / \|U\|_F$,
3. for $i = 1, \ldots, k$ do
   - Compute $W = A V_i^G$.
   - for $j = 1, \ldots, i$ do
     (a) $h_{j,i}^G = \langle V_j^G, W \rangle_F$
     (b) $W = W - h_{j,i}^G V_j^G$
   - End
   - Compute $h_{i+1,i}^G = \|W\|_F$ and set $V_{i+1}^G = W / h_{i+1,i}$.
4. End

It is easy to see that $H_{k+1,k}^G = h_{k+1,k}^G I_s$ and that $V_{k+1}^G = \mathbb{P}_{k,G}^A(A) U$. Moreover, if we set

$$\widetilde{\mathbb{P}}_{k,G}^A(t) = h_{k+1,k}^G \cdots h_{2,1}^G \|U\|_F \mathbb{P}_{k,G}^A(t)$$

and use the explicit form of $\widetilde{\mathbb{P}}_{k,G}^A$,

$$\widetilde{\mathbb{P}}_{k,G}^A(t) = \frac{\det\left(\begin{bmatrix} t^k & 1 & \ldots & t^{k-1} \\ \mathcal{K}_k^H \diamond (A^k U) & & (\mathcal{K}_k^H \diamond \mathcal{K}_k) & \end{bmatrix}\right)}{\det(\mathcal{K}_k^H \diamond \mathcal{K}_k)}, \tag{3.6}$$

we can characterize the roots.

THEOREM 3.2. *Let the matrix $(\mathcal{K}_k^H \diamond \mathcal{K}_k)$ be nonsingular. Then*

$$\widetilde{\mathbb{P}}_{k,G}^A(t)) = \prod_{i=1}^{s} (t - \widetilde{\theta}_i^{(k)}),$$

ETNA
Kent State University
http://etna.math.kent.edu

where $\widetilde{\theta}_i^{(k)}$, *for* $i = 1, ..., s$, *are the eigenvalues of the matrix* $(\mathcal{K}_k^H \diamond \mathcal{K}_k)^{-1}(\mathcal{K}_k^H \diamond (A\mathcal{K}_k))$.

The eigenvalues $\widetilde{\theta}_i^{(k)}$ can be called the F-Ritz values, since they also are the eigenvalues of the Hessenberg matrix $H_k^G = (\mathcal{V}_k^{G\,H} \diamond (A\mathcal{V}_k^G))$.

When we apply the global or the block Arnoldi processes with $s = 1$ and with the ith columns of U, we obtain the standard Arnoldi process. Hence, the standard Arnoldi vectors obtained with the ith columns of U can be written as $v_k^{(i)} = p_{k,S}^{(i)}(A)Ue_i$. Let $V_{k,S}^A$ be the vector whose columns are $v_k^{(1)}, \ldots, v_k^{(s)}$. We have

$$V_{k,S}^A = [p_{k,S}^{(1)}(A)Ue_1, \ldots, p_{k,S}^{(s)}(A)Ue_s].$$

Consequently,

$$\operatorname{vec}(V_{k,S}^A) = \mathbb{P}_{k,S}^A(A)\operatorname{vec}(U), \tag{3.7}$$

where $\mathbb{P}_{k,S}^A(t) = diag(p_{k,S}^{(1)}(t), \ldots, p_{k,S}^{(s)}(t))$.

4. Examples. We illustrate the theory developed in this paper with two examples. The first one involves a diagonalizable matrix; the matrix of the second example is defective. In these examples, we set $X_0 = 0$ and $U = B$.

EXAMPLE 4.1. Consider the matrix and right-hand sides

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

Results obtained by the block Arnoldi and BGMRES methods are reported in Table 4.1. Moreover, we have for block Arnoldi, $\det(\mathbb{P}_{1,B}^A(t)) = \frac{\sqrt{2}}{18}(2t+1)(t+2)$ and $\det(\mathbb{P}_{2,B}^A(t)) = \frac{\sqrt{2}}{9}(t^2-1)(t^2-4)$. Hence, the eigenvalues of the matrix A are the roots of $\mathbb{P}_{2,B}^A$. We also remark that the roots of $\mathbb{P}_{1,B}^A$ are $-\frac{1}{2}$ and -2.

k	$\mathbb{P}_{k,B}^A(t)$	$\mathbb{P}_{k,B}^G(t)$
1	$\frac{\sqrt{3}}{3}\begin{bmatrix} \frac{\sqrt{2}}{3}(t+1) & \frac{1}{3}(t-2) \\ -\frac{\sqrt{2}}{6} & t+\frac{4}{3} \end{bmatrix}$	$\begin{bmatrix} \frac{3}{26}t+1 & \frac{5}{26}t \\ -\frac{1}{13}t & \frac{7}{13}t+1 \end{bmatrix}$
2	$\frac{\sqrt{3}}{3}\begin{bmatrix} \frac{\sqrt{2}}{3}(t^2+\frac{2}{3}t-2) & \frac{1}{3}(t^2-\frac{10}{3}t+4) \\ \frac{\sqrt{2}}{3}(\frac{t}{3}+1) & t^2-\frac{5}{9}t-\frac{8}{3} \end{bmatrix}$	$\begin{bmatrix} (-\frac{3}{4}t^2-\frac{1}{6}t+1 & -\frac{1}{2}t^2+\frac{1}{3}t \\ -\frac{1}{4}t^2-\frac{1}{12}t & -\frac{1}{2}t^2-\frac{1}{6}t+1 \end{bmatrix}$

TABLE 4.1
Polynomials obtained by the block Arnoldi and BGMRES methods.

On the other hand, the upper block Hessenberg matrix $\mathcal{H}_2^B$ determined by block Arnoldi algorithm is

$$\mathcal{H}_2^B = \begin{bmatrix} -1 & \frac{\sqrt{2}}{2} & \frac{5\sqrt{2}}{18} & \frac{1}{9} \\ \frac{\sqrt{2}}{2} & \frac{-3}{2} & \frac{5}{18} & \frac{\sqrt{2}}{18} \\ \frac{3\sqrt{2}}{2} & -\frac{1}{2} & \frac{11}{18} & -\frac{5\sqrt{2}}{18} \\ 0 & \frac{\sqrt{2}}{2} & -5\frac{\sqrt{2}}{18} & \frac{17}{9} \end{bmatrix},$$

k	$P_{k,G}^{A}(t)$	$P_{k,G}^{G}(t)$
1	$\frac{1}{\sqrt{23}}(3t+4)$	$(\frac{4}{13}t+1)$
2	$\frac{1}{\sqrt{15042}}(69t^2+64t-91)$	$(\frac{-21}{113}t^2+\frac{88}{1467}t+1)$
3	$\frac{1}{3037830}(981t^3+1482t^2-2066t-2182)$	$(\frac{-2182}{11809}t^3-\frac{4589}{11809}t^2+\frac{716}{1687}t+1)$
4	$3\sqrt{\frac{327}{9290}}(t^2-1)(t^2-4)$	$\frac{1}{4}(t^2-1)(t^2-4)$

TABLE 4.2
Polynomials obtained by the global Arnoldi and global GMRES methods.

k	$P_{k,S}^{A}(t)$
1	$\frac{\sqrt{5}}{5}\begin{bmatrix} t+1 & 0 \\ 0 & 2t+3 \end{bmatrix}$
2	$\begin{bmatrix} \frac{\sqrt{6}}{6}(3t^2+2t-5) & 0 \\ 0 & \frac{\sqrt{870}}{870}(15t^2+22t-7) \end{bmatrix}$
3	$\begin{bmatrix} \frac{\sqrt{30}}{120}(15t^3+15t^2-28t-8) & 0 \\ 0 & \frac{\sqrt{174}}{870}(87t^3+108t^2-323t-352) \end{bmatrix}$
4	$(t^2-1)(t^2-4)\begin{bmatrix} \frac{\sqrt{30}}{8} & 0 \\ 0 & \frac{\sqrt{174}}{10} \end{bmatrix}$

TABLE 4.3
Polynomials obtained by the standard Arnoldi method.

k	$P_{k,S}^{G}(t)$
1	$\begin{bmatrix} \frac{1}{6}t+1 & 0 \\ 0 & \frac{3}{7}t+1 \end{bmatrix}$
2	$\begin{bmatrix} (\frac{-75}{161}t^2-\frac{44}{161}t+1) & 0 \\ 0 & (\frac{-3}{71}t^2+\frac{178}{497}t+1) \end{bmatrix}$
3	$\begin{bmatrix} (\frac{-1}{22}t^3-\frac{1}{2}t^2-\frac{2}{11}t+1) & 0 \\ 0 & (\frac{-352}{1567}t^3-\frac{443}{1567}t^2+\frac{1358}{1567}t+1) \end{bmatrix}$
4	$\frac{1}{4}(t^2-1)(t^2-4)I_2$

TABLE 4.4
Polynomials obtained using the Standard GMRES method.

with the characteristic polynomial

$$P_{\mathcal{H}_2}(t) = \det(tI_4 - \mathcal{H}_2^B) = (t^2-1)(t^2-4).$$

The polynomials determined by global Arnoldi and global GMRES are displayed in Table 4.2. The standard Arnoldi methods yields the polynomials of Table 4.3, and the polynomial determined by standard GMRES are shown in Table 4.4.

ETNA
Kent State University
http://etna.math.kent.edu

k	$P^A_{k,B}(t)$	$P^G_{k,B}(t)$
1	$\frac{\sqrt{10}}{5}\begin{bmatrix}\frac{1}{4}(4t-9) & \frac{3}{5}t-2\\ \frac{3}{5} & \frac{t}{4}-1\end{bmatrix}$	$\begin{bmatrix}-\frac{7}{18}t+1 & \frac{t}{18}\\ -\frac{13}{54}t & -\frac{29}{54}t+1\end{bmatrix}$
2	$\frac{t-1}{5\sqrt{10}}\begin{bmatrix}4t-7 & \frac{3}{4}(t-3)\\ -\frac{1}{4} & \frac{5t-7}{4}\end{bmatrix}$	$\frac{t-1}{8}\begin{bmatrix}5t-8 & -3t\\ -t & 7t-8\end{bmatrix}$

TABLE 4.5
Polynomials obtained by the block Arnoldi and BGMRES methods.

EXAMPLE 4.2. Define the defective matrix

$$A = \begin{bmatrix}1 & 2 & 1 & 0\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 2\end{bmatrix}$$

and let

$$B = U = \begin{bmatrix}1 & 2\\ 1 & 0\\ 0 & 1\\ 0 & 1\end{bmatrix}.$$

Table 4.5 shows the results obtained by the block algorithms. The upper block Hessenberg matrix determined by the block Arnoldi algorithm is given by

$$\mathcal{H}_2^B = \frac{1}{20}\begin{bmatrix}45 & 5 & -3\sqrt{10} & 4\sqrt{10}\\ -15 & 25 & 5\sqrt{10} & 0\\ 5\sqrt{10} & -3\sqrt{10} & -2 & -4\\ 0 & 4\sqrt{10} & 16 & 32\end{bmatrix}.$$

It has the characteristic polynomial

$$P_{\mathcal{H}_2}(t) = (t-1)^3(t-2).$$

We also have

$$\begin{aligned}
P^A_{3,G}(t) &= \frac{2\sqrt{19}}{\sqrt{67}}(t-1)^2(t-2),\\
P^G_{3,G}(t) &= -\frac{1}{2}(t-1)^2(t-2),\\
P^A_{3,S}(t) &= (t-1)^2(t-2)\begin{bmatrix}\sqrt{38} & 0\\ 0 & \frac{\sqrt{66}}{9}\end{bmatrix},\\
P^G_{3,S}(t) &= -\frac{1}{2}(t-1)^2(t-2).
\end{aligned}$$

We remark that for all iterations except for the last one, the roots of the BGMRES polynomials are, in general, different from those of the corresponding Arnoldi polynomials. Moreover, apart from a multiplicative scalar, the determinant of the Arnoldi polynomial is the characteristic polynomial of the Hessenberg matrix obtained from the Arnoldi-type algorithms.

REFERENCES

[1] W. E. ARNOLDI, *The principle of minimized iterations in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 17–29.
[2] R. BOUYOULI, K. JBILOU, A. MESSAOUDI, AND H. SADOK, *On block minimal residual methods*, Appl. Math. Lett., 20 (2007), pp. 284–289.
[3] R. BOUYOULI, K. JBILOU, R. SADAKA, AND H. SADOK, *Convergence properties of some block Krylov subspace methods for multiple linear systems*, J. Comput. Appl. Math., 196 (2006), pp. 498–511.
[4] C. C. CHU, M. H. LAI, AND W.S. FENG, *MIMO interconnects order reductions by using the multiple point adaptive-order rational global Arnoldi algorithm*, IECE Trans. Elect., E89-C (2006), pp 792–808.
[5] C.C. CHU, M.H. LAI, AND W.S. FENG, *The multiple point global Lanczos method for multiple-inputs multiple-outputs interconnect order reduction*, IECE Trans. Elect., E89-A (2006), pp. 2706–2716.
[6] M. HEYOUNI AND K. JBILOU, *Matrix Krylov subspace methods for large scale model reduction problems*, Appl. Math. Comput., 181 (2006), pp. 1215–1228.
[7] K. JBILOU AND A. MESSAOUDI, *Matrix recursive interpolation algorithm for block linear systems. Direct methods*, Linear Algebra Appl. 294 (1999), pp. 137–154 .
[8] K. JBILOU, A. MESSAOUDI, AND H.SADOK, *Global FOM and GMRES algorithms for matrix equations*, Appl. Numer. Math., 31 (1999), pp. 49–43.
[9] K. JBILOU AND A.J. RIQUET, *Projection methods for large Lyapunov matrix equations*, Linear Algebra Appl., 31 (2006), pp. 49–43.
[10] Y. Q. LIN, *Implicitly restarted global FOM and GMRES for nonsymmetric matrix equations and Sylvester equations*, Appl. Math. Comput., 167 (2005), pp. 1004–1025.
[11] M. H. GUTKNECHT, *Block Krylov space methods for linear systems with multiple right-hand sides: an introduction*, in Modern Mathematical Models, Methods and Algorithms for Real World Systems, A. H. Siddiqi, I. S. Duff, and O. Christensen, eds., Anamaya Publishers, New Delhi, 2005, pp. 420–447.
[12] A. MESSAOUDI, *Recursive interpolation algorithm: a formalism for linear equations I. Direct methods*, J. Comput. Appl. Math., 76 (1996), pp. 13–30
[13] ———, *Recursive interpolation algorithm: a formalism for linear equations II. Iterative methods*, J. Comput. Appl. Math., 76 (1996), pp. 31–53.
[14] D. O'LEARY, *The block conjugate gradient algorithm and related methods*, Linear Algebra Appl., 29 (1980), pp. 293–322.
[15] C. R. RAO, *Statistical proofs of some matrix inequalities*, Linear Algebra Appl., 321 (2000), pp. 307–320.
[16] Y. SAAD AND M. SCHULTZ, *GMRES: A Generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[17] Y. SAAD, *Numerical Methods for Large Eigenvalue Problems*, Halstead, New York, 1992.
[18] ———, *Iterative Methods for Sparse Linear Systems*, PWS, Boston, 1996.
[19] M. SADKANE, *Block Arnoldi and Davidson methods for unsymmetric large eigenvalue problems*, Numer. Math., 64 (1993), pp. 687–706.
[20] M. SADKANE AND M. ROBBÉ, *Exact and inexact breakdowns in the block GMRES method*, Linear Algebra Appl., 419 (2006), pp. 265–285.
[21] H. SADOK, *Méthodes de projections pour les systèmes linéaires et non linéaires*. Habilitation thesis, University of Lille 1, Lille, France, 1994.
[22] ———, *Analysis of the convergence of the minimal and the orthogonal residual methods*, Numer. Algorithms, 40 (2005), pp. 101–115.
[23] D. K. SALKUYEH AND F. TOUTOUNIAN, *New approaches for solving large Sylvester equations*, Appl. Math. Comput., 173 (2006), pp. 9–18
[24] I. SCHUR, *Potenzreihen im Innern des Einheitskreises*, J. Reine Angew. Math., 147 (1917), pp. 205–232.
[25] V. SIMONCINI, *Ritz and Pseudo-Ritz values using matrix polynomials*, Linear Algebra Appl., 241-243 (1996), pp. 787–801.
[26] V. SIMONCINI AND E. GALLOPOULOS, *Convergence properties of block GMRES and matrix polynomials*, Linear Algebra Appl., 247 (1996), pp. 97–119.
[27] ———, *An iterative method for nonsymmetric systems with multiple right-hand sides*, SIAM J. Sci. Comput., 16 (1995), pp. 917–933.
[28] G. W. STEWART, *Matrix Algorithms II: Eigensystems*, SIAM, Philadelphia, 2001.
[29] B. VITAL, *Etude de quelques méthodes de résolution de problèmes linéaires de grande taille sur multiprocesseur*, Ph.D. thesis, Université de Rennes, Rennes, France, 1990.
[30] N. WAGNER, *A Krylov based method for quadratic matrix equations*, in Angewandte und Experimentelle Mechanik – Ein Querschnitt, K. Willner and M. Hanss, eds., Der Andere Verlag, Tönning, 2006, pp. 283–303 .
[31] F. ZHANG, *Matrix Theory*, Springer, New York, 1999.